UNIVERSITY OF VAASA

FACULTY OF BUSINESS STUDIES

DEPARTMENT OF ACCOUNTING AND FINANCE

Sakari Pasanen

THE BEHAVIOUR OF VOLATILITY RATIO

Implied vs. Realized Volatility of the S&P 500 and DAX Stock Indices

Master's Degree
Programme in Finance

VAASA 2016

TABLE OF CONTENTS **page**

LIST OF FIGURES AND TABLES

FIGURES

TABLES

UNIVERSITY OF VAASA
Faculty of Business Studies
Author: Sakari Pasanen
Topic of the thesis: The Behaviour of Volatility Ratio: Implied vs. Realized Volatility of the S&P 500 and DAX Stock Indices
Name of the Supervisor: Sami Vähämaa
Degree: Master of Science in Economics and Business Administration
Department: Department of Accounting and Finance
Major Subject: Finance
Year of Entering the University: 2004
Year of Completing the Thesis: 2016 **Pages:** 61

ABSTRACT

The volatility ratio is defined as an implied volatility divided by a realized volatility. The purpose of this thesis is to investigate if some chosen macroeconomic variables make an impact on the volatility ratio. The variables are base index return, unemployment rate, risk free interest rate, economic uncertainty, consumer consumption, market mode, and inflation. It is also investigated whether the time of the years 2007 – 2008 financial crisis have had an effect on the volatility ratio level. The literature review of the thesis goes through the studies on option implied standard deviations, implied volatility, and relationship between the implied and realized volatilities. The second part of the thesis covers the basics on option pricing, volatility models and theory on correlation.

The measure for the realized volatility is the monthly sample standard deviation of the daily index return. The indices used in this thesis are the S&P 500 and DAX. Values of the volatility indices VIX and VDAX are used as implied volatilities. Multiple linear regression analysis is used as a method for the final analysis with stepwise variable selection. Data used for an independent variable (volatility ratio) in the thesis is daily values for the indices S&P 500, VIX, DAX, and VDAX, which have been obtained from Thomson Datastream –database. Data for dependent variables is monthly values for earlier mentioned macroeconomic variables. Data spread is from the start of the year 2002 through the year 2014.

The only macroeconomic variables of the ones investigated in this thesis that have statistically significant effect on the volatility ratios are index return and market mode. It is concluded that the financial crisis did not have an effect on the level of the DAX volatility ratio. However the level of the S&P 500 volatility ratio has been higher after the crisis.

KEYWORDS: volatility ratio, implied and realized volatility, volatility index, regression analysis

1. INTRODUCTION

The relationship between different technical variables derived from the stock markets and stock returns have been studied in the numerous studies after the introduction of the stock markets. The field of the study is important because of the always ongoing search for bigger and more predictable returns from the stock markets. Volatilities have been under constant scrutiny after Black and Scholes (1973), and Merton (1973) introduced their model for the option pricing. The model is embedded with a constant volatility but there have been numerous studies weakening that constraint. Different volatility models and introduction of the high-frequency daily price data for asset prices has brought the importance of the volatility modelling to the new level. The most basic volatility concepts in a current literature are the implied and realized volatilities.

The implied volatility is a volatility that can be computed from the prices of different financial derivatives, like options and futures, using pricing models. These pricing models usually have a foundation on the already mentioned Black-Scholes-Merton pricing model. The implied volatility can be described as a market's expectation of the future volatility. The realized volatility concept can be divided to the past or historical volatility, and to the future realized volatility. The realized volatility is computed from asset price fluctuations over different time ranges from high-frequency intraday-prices observed on few-minute-basis to the daily and monthly observations. The better the relationship between the implied and future realized volatilities is known, the more accurate predictions of the future volatility the market traders are able to do (and the harder it is to get a profit on the efficient markets).

1.1. Hypotheses

The literature review goes through several studies that forecast the future volatility with the implied and past realized volatilities. However there are only few studies on the volatility ratio. The ratio between the implied and past realized volatilities is taken under the inspection in this thesis. Thesis investigates the effect that several individual variables have on the ratio. These independent variables include return, unemployment rate, risk free interest rate, economic uncertainty, consumer consumption, market state, inflation, and the time of the years 2007 – 2008 financial crisis.

The behaviour of the volatility ratio is inspected with the multiple linear regression analysis. A dependent variable in the regression will be the volatility ratio that is modelled in the Chapter 4.3.1. There is total of nine independent variables investigated to be used in the regression analysis. Six of the variables are linear variables; economic policy uncertainty index, unemployment rate, consumer consumption rate of change, monthly index return, inflation rate, and risk free interest rate. Three of the variables are dummy variables; indicator for the bull/bear market, and indicators for splitting time series unto three phases before, during, and after the years 2007 – 2008 financial crisis.

Hypotheses to be studied can be expressed individually for each independent variable as follows:

H_1: The level of the economic policy uncertainty index has an impact on the volatility ratio.

H_2: The unemployment rate has an impact on the volatility ratio.

H_3: The change in the consumer consumption has an impact on the volatility ratio.

H_4: The monthly index return has an impact on the volatility ratio.

H_5: The risk free interest rate has an impact on the volatility ratio.

H_6: The inflation rate has an impact on the volatility ratio.

H_7: The market mode has an impact on the volatility ratio.

H_8: The volatility ratio level during the years 2007 – 2008 financial crisis is distinctive.

H_9: The volatility ratio level after the years 2007 – 2008 financial crisis is distinctive.

Hypotheses H_8 and H_9 indicate the third time sensitive hypothesis that the volatility ratio level before year 2007 is also distinctive. The confirmation of the hypothesis requires for both H_8 and H_9 to hold true in the analysis and it is not taken under the inspection as an individual hypothesis in this thesis. Knowledge of the possible presence and strength of the impacts described in hypotheses would be valuable when looking to get trading opportunities from the imbalance between the realized and implied volatilities.

1.2. Structure of the thesis

This study consists of a theoretical and an empirical part. The theoretical part will go through previous studies and basic concepts of the topic. In the empirical part the behaviour of the ratio between the implied and realized volatilities is inspected. Indices under the scrutiny are the S&P 500 index (SPX) and the CBOE Volatility Index (VIX), and DAX index with its volatility index VDAX. In the end of the empirical part the conclusions will be based on acquired results.

In the second chapter the study will introduce previous literature that has been made on the topic of volatility forecasting and the implied-realized volatility ratio. Chapter 3 will go through the theoretical background of the option pricing and volatility. The basics of the equation called the Black-Scholes-Merton differential equation are introduced. This equation is widely used when pricing financial derivatives. The importance of the equation was recognized when the developers received the Nobel Prize for economics in 1997. Chapter 3 also describes few facts about different volatility models. Chapter 3.2 gives a short introduction to the stochastic volatility models and GARCH models. In addition, there will be few words about the realized volatility. Chapter 3.3 presents the basics of the correlation. The chapter is about differences between the regression and correlation analyses with a short dip to the population correlation coefficient. Fourth chapter has an explanation about the data and the methodology used in the study. Results and conclusions for the study are in chapters five and six.

2. LITERATURE REVIEW

This chapter will go through some earlier studies that have been made concerning the future volatility forecasting abilities of the implied and past realized volatilities. Chapter 2.1 will introduce few early studies that examine an option implied standard deviation which is later called volatility. Chapter 2.2 presents studies where the implied volatility is concluded to be a poor predictor of the future volatility. Chapter 2.3 presents studies where the implied volatility is found to be a valuable asset in forecasting future return volatilities. More detailed introduction to the volatility forecasting literature, especially methodologies used and empirical findings obtained in different studies, can be found in Poon and Granger (2003). Valuable information can be found on handbooks written about the volatility forecasting like Knight and Satchell (2007), and Bauwens, Hafner and Laurent (2012).

2.1. Early studies on the option implied standard deviations

There have been quite many studies investigating the relationship of the realized and implied volatilities from different point of views, and with different data sets. The early study by Latané and Rendleman in 1976 is one of the first to use the Black-Scholes-Merton option pricing model, which is introduced in Chapter 3.1, in order to generate weighted average implied standard deviations from actual option prices. Latané et al. (1976) then compare these implied standard deviations with realized standard deviations of log prices for stocks. Data on the study consists of the weekly option and stock prices for 24 companies that had options traded on the CBOE for the time period of 38 weeks from October 1973 through to June 1974. The authors conclude that their implied standard deviations correlate strongly with the realized standard deviations. They conclude also that the implied standard deviation presented in the study is better estimator for the future deviation than the realized standard deviation. Chiras and Manaster (1978), and Beckers (1981) make some minor modifications to the work done by Latané et al. (1976), but in the end they agree with the result that implied standard deviation is a good predictor of the future deviations.

2.2. Underperforming implied volatility

Lamoureux and Lastrapes (1993), who examine the implied and realized variances, obtain different result than Latané et al. (1976), Chiras et al. (1978), or Beckers (1981), and conclude that the realized variance beats the implied variance in forecasting future variances. They state that forecasts based on realized variances are more accurate. Data in their study is option and stock prices for 10 individual stocks that have options traded in the CBOE. Data range is from 19th April, 1982, to 31st March, 1984. Lamoureux et al. (1993) have chosen the individual stocks on the basis that there are no cash dividends which would cause jumps on the prices observed.

The study made by Canina and Figlewski (1993) is interesting especially from the view of this thesis as they look into the predictive power of the volatility implied by index options, namely the S&P 100 index options. The data they use in the study consists of the daily closing prices of call options on the S&P 100 index for 48 months from 15th March, 1983, to 28th March, 1987. Modification to the data is that far outliers, the nearest and the farthest options to expiration time, and options far in- and out-of-the-money, are eliminated. As the older studies, Canina et al. (1993) use Black-Scholes-Merton model to calculate implied volatilities from the option prices. They conclude that the implied and realized volatilities are not correlated, and also that neither the implied nor realized volatility is any good when forecasting the volatility. This result is totally a contradiction of older studies.

2.3. Implied volatility outperforms realized volatility

Jorion (1995) looks for the forecasting power of the implied volatility of foreign currency futures. Data in the study consists of Chicago Mercantile Exchange's prices for currency futures and options on those futures. Data range is about seven years up to February 1992, with different start times for individual currencies used in the study. Jorion (1995) concludes that the implied volatility is an efficient estimator of the future volatility. The author comments on older studies (in example Lamoureux et al. 1993, Canina et. al 1993), which have found contradictory results, that on those either the test procedures are incorrect or there is an inefficiency on the options market. Andersen and Bollerslev (1998) have made another study that looks into the currency exchange rate volatility confirming Jorion's (1995) findings.

Kawaller, Koch and Peterson (1994) take a little different approach to the relationship between the option implied and historical volatility of an asset. The authors compare said volatilities on futures of the S&P 500 index, deutsche marks, Eurodollars, and live cattle. The type of the contracts is not of interest in the context of this thesis but the way the data is obtained is more so. Kawaller et al. (1994) obtain high-frequency minute-to-minute prices for all of the assets to form a time-series for the daily return. They then calculate a standard deviation for the time-series as a daily historical volatility. This standard deviation is calculated twice, once for a full set of the daily data, and once for the daily data divided to ten subsections for to better capture intraday movements. The option implied volatility is obtained as an average of all option implied volatilities on the respective time interval. The data for all assets is for the fourth quarter, 1988, with a little variation on actual starting and ending days between the assets. Kawaller et al. (1994) conclude that the causality between the implied and historical volatilities cannot be generalized as there are differences between markets and time intervals.

Christensen and Prabhala (1998) set out to check out the results presented by Canina et al. (1993). Time period for the S&P 100 index options price data used in the study overlaps with Canina et al. (1993) as the data is for 139 months period from November 1983 to May 1995. Options included in the data are at-the-money call options. They found that the implied volatility predicts the future volatility better than the realized volatility when forecasting. There are three reasons given for the different results when compared to the results presented by Canina et al. (1993). The first reason is that Christensen et al. have longer time period for the data to use in the study. The second reason is the data sampling as authors use a monthly data for the option and index prices, and options are those that are expiring just before the next sample date. The sampling made in this way gives the results some robustness against the autocorrelation in a daily returns. The third reason is that according to the authors the October 1987 stock market crash caused a shift in both implied and realized volatility levels. After the crash the explanatory value of the implied volatility for the future volatility is significantly better than before the crash. The result of Christensen et al. is supported by Gwilym and Buckle (1999) on one-month-forward forecasts.

Fleming (1998) has also made a study to examine the future volatility forecasting ability of the S&P 100 implied volatility. Data on the study is the S&P 100 option prices and index level from October 1985 through April 1992. Excluded from the data are all observations during the October 1987 stock market crash. Fleming's conclusion is that the implied volatility is an efficient forecast of the realized volatility. He also comments that the past realized volatility rate cannot explain any components of the realized volatility left unexplained by the implied volatility.

Blair, Poon and Taylor (2001) compare the index volatility forecasting abilities of implied volatilities and intraday returns. The authors use a weighted implied volatility index calculated as in Fleming, Ostdiek, and Whaley (1995) to obtain daily volatilities. Blair et al. are first to obtain another volatility measure from high-frequency (5-min) stock index returns. The data in their study has 13 year span from the start of 1987 to the end of 1999. Blair et al. support the findings of several studies (i.e. Jorion 1995, Christensen et al. 1998, Fleming 1998) and states that the implied volatility index has more predictive power on the future volatility than the past realized volatility has. The new volatility measure on high-frequency index returns and a more traditional volatility on daily index returns both are outperformed by implied volatilities. Authors comment that there could be some increased forecasting efficiency gained if used a mixture of the implied and past realized volatilities on a 1-day-ahead forecasts, but for longer forecasts the volatility index has all relevant forecasting information.

Hansen (2001) makes an interesting comparison on the implied and past realized volatilities in an attempt to prove Christensen et al. (1998) results on an illiquid option market. The index which returns are from and options are on is the Danish KFX share index. The data is for 52 months from September 1995 to December 1999. The author has had to be satisfied with some missing values on the data due to the illiquidity of the market. Straightforward results on the study give poor results for the implied volatility but after dealing with the errors-in-variance problem, which originates from the poor liquidity, the results of Christensen et al. are confirmed. In the end implied volatility outperforms the past realized volatility as a predictor of the future volatility.

Christensen and Hansen (2002) extend the Christensen et al. (1998) by adding in the study a wider range of options and a trade weighted average of implied options. The authors divide the implied volatility to a call and put implied volatilities as a robustness check. Data on the study is one month to expiration call and put options on the S&P 100 index from April 1993 to February 1997. The result is that the call implied volatility is better predictor of the future volatility than the put implied volatility. Christensen et al. also make an analysis on a full set of the data which shows that additional valuable volatility information can be obtained from put option prices.

Giot (2005) makes an interesting study where he inspects the relationship between relative changes in the implied volatility and both current and future stock market returns. Data he uses is the stock indices S&P 100 and Nasdaq 100 and their implied volatility indices. Data range from August 1994 through January 2003 is divided to three subperiods which are a low-volatility bull market (1994-1997), a high-volatility bull market (1997-2000), and a high-volatility bear market (2000-2003). The author concludes that implied volatilities and stock index returns have a strong negative correlation with asymmetricalness in especially in the relationship between the S&P 100 and its' implied volatility. This means that there is a difference between the negative and positive return periods as negative return periods have much greater relative changes on the contemporaneous volatility.

Corrado and Miller (2005) take a wider look on the forecast abilities of the volatility indices as they examine implied volatilities for the S&P 100, S&P 500, and Nasdaq 100 stock indices. Volatility indices for the stock indices are the VXO, VIX, and VXN, respectively. Data for the study is index returns and implied volatilities from January 1988 for the VXO and S&P 100, from January 1990 for the VIX and S&P 500, and from January 1995 for the VXN and Nasdaq 100, through December 2003. The authors divide their data to pre-1995 and after-1995 periods. Conclusion of the study is that implied volatility indices clearly outperform past realized volatilities in the forecasting. Further result on the study is that the forecast error in the earlier data series is significant when it was found to be almost extinct in the latter data series. The calculation of the realized volatility on this study is very interesting from the view of this thesis as Corrado et al. calibrate the realized volatility rate in a way which results as a good comparability between the volatility indices and the realized volatility series.

Koopman, Jungbacker and Hol (2005) compare the forecasting power of the past (daily) realized volatility, implied volatility, and high-frequency intraday realized volatility. Data is for the S&P 100 stock index and its volatility index from January 6th, 1997, through November 14th, 2003. Result on the paper is that the high-frequency intraday realized volatility has more predictive power on the future volatilities.

Gosponidov, Gavala and Jiang (2006) look for forecasting results of different volatility models with the S&P 100 index volatility. Data in the study is daily measures for the S&P 100 and its implied volatility index VIX from June 1st, 1988, through May 17th, 2002. The authors create the daily realized volatility using 22 trading day rolling volatility to estimate monthly volatility on a daily basis as opposed to the non-overlapping monthly measure used on this thesis. This method can possibly incur an autocorrelation problem in the volatility measure but gives more data points to be used in the forecasting. Gosponidov et al. also test the implied volatility as a predictor for the realized volatility and find it to be an optimal and unbiased forecast for the future volatility. The authors conclude first that the implied volatility is valuable when forecasting volatility and second that the combination of the forecasts of different volatility models can improve the strength of the forecast.

Giot and Laurent (2007) make one more study between forecasting strength of the implied and historical volatilities against the future volatility. Data in this study is the S&P100 and S&P500 indices and their volatility indices VXO and VIX. Datasets are 20-minute intraday sample for base indices and the daily data for volatility indices. The time period for the data is January 2nd, 1990, through March 5th, 2003. Giot et al. extend their historical volatility measure by decomposing it to the daily, weekly and monthly time components. All additional regressors, time structure decomposition and GARCH volatility, included in the study confirm the fact that the implied volatility has most relevant information available for volatility forecasts.

Shephard and Sheppard (2010) study volatility models based on a high-frequency data. These models are called HEAVY models and they allow for both mean reversion and momentum. The authors use the daily returns and realized kernels from a database (Oxford-Man Institute’s realized library) which covers some traded currencies and indices as well as some computed indices. Datasets are tick by tick data with median duration between price updates ranging from 2 to 60 seconds. Data series start between 1996 and 2002 and run through March 2009. Shephard et. al conclude their analysis stating that the models are more robust to level breaks and display also a momentum which is not featured in GARCH models.

Another case of an intraday data is a study made by Byun, Rhee, and Kim (2011) where they focus on the value of implied volatilities from a stochastic volatility model and Black and Scholes model when forecasting the realized volatility. The data used in the study is minute by minute transaction prices of Korean KOSPI200 index options from July 1st, 2004, through June 29th, 2007. Conclusion in the study is that although both implied volatilities have information value over the realized volatility neither of them beats another in the forecasting.

Han and Park (2013) study forecasting abilities of the realized and option implied volatilities. The study is particularly interesting because the data used is the daily S&P 500 index returns and the VIX volatility index, which are also used in this thesis. The data range for the indices is from 3 January 1996 to 27 February 2009. The realized kernel is used in the study for a measure of the realized volatility. The authors test models which include information from either the realized kernel or VIX or both of them. The results of the study give a bit of a benefit for both versions of the volatility. The realized volatility beats the implied when tested with the in-sample data but the opposite is true when the information content is tested with the out-of-sample forecasting.

2.4. Relationship between implied and realized volatilities

There are several recent studies where the relationship between the implied and realized volatilities has been inspected. These studies concentrate on the effect that different parameters, uncertainties and shocks have on the implied/realized volatility ratio. This ratio has also been presented in a little different way as a difference between variances and called volatility spread or variance risk premium (or premia).

One of these studies was made by Garvey and Gallagher (2012). The authors study the relationship between the implied and realized volatilities and the value of the implied volatility in forecasting the future volatility. Equity option prices data used in the calculation of implied volatilities is for sixteen FTSE-100 companies from October 1st, 1997, through December 31st, 2003. The high-frequency data is converted to price series with 30 minute intervals. The authors find that the presence of a fractional integration between the implied and realized equity volatilities cannot be rejected. Another conclusion in the paper is that the implied volatility of the high-frequency option price data gives additional information when compared with the historical realized volatility of the underlying assets. This is in line with several other studies presented in earlier chapters.

3. THEORETICAL BACKGROUND

This chapter presents a review of the basic option pricing, volatility modelling and regression analysis. Chapter 3.1 deals with the Black-Scholes-Merton option pricing model which has long been accepted as a basic pricing model for options and has been used as a benchmark for other models. Pricing equations and model assumptions will be introduced. Chapter 3.2 introduces some different volatility models as follows: Sub-chapter 3.2.1 stochastic volatility models, sub-chapter 3.2.2 generalized autoregressive conditionally heteroskedastic models and sub-chapter 3.2.3 realized volatility. Chapter 3.3 introduces basics of the correlation analysis.

3.1. Black-Scholes-Merton Option Pricing Model

One big breakthrough in the stock market derivatives pricing was in the early 1970s when Black et al. (1973) and Merton (1973) introduced a model which grounded the foundation for the pricing of the options. Starting point for a model was an assumption of a riskless portfolio. The model was the first reliable tool for the pricing of the options. The option pricing theory has been further developed after that to soften tight constraints of the model. There was and is a need for a more efficient option pricing models because the derivative markets continue to develop and become more effective.

Black et al. (1973: 637) concludes that it should not be possible to get sure gain by doing portfolios from put and call options with corresponding shares if options have been priced correctly. The starting point of the Black-Scholes-Merton model is previous idea of the riskless portfolio that is made of the option and asset. When options price is derived that in mind, then the result is a differential equation that is usually called to the Black-Scholes-Merton differential equation (equation (1)). The expected return and the risk premium, which made previous models unpractical, have been secluded from this model. In its original form the Black-Scholes-Merton equation gives an answer to the pricing of the European options but not that of the American options.

Use of the model is possible only with some basic assumptions of the state of stock markets and also of the option itself. Assumptions work better on the short options than on long options or on options that have an asset that is tied to the interest rates (Jarrow 1999: 233–234). Black's et al. (1973) assumptions are that 1) the short interest rate is known and it is constant, 2) the price of the stock follows a random-walk and the variance of the return of the stock is constant, 3) dividend is not paid, 4) the option is European, 5) selling and buying of options and the stock is possible without costs, 6) loans are given at the riskless rate whenever needed and 7) short selling is possible without excess costs. When these assumptions are correct only the price of the stock and time and some assumed constants effect on the price of the option.

If we know the right limits for the derivatives, it's possible to solve prices from the Black-Scholes-Merton equation. Equation is as follows (see Hull 2000: 247; compare Black et al. 1973: 643; Broadie and Detemple 2004: 1146):

(1) $$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf,$$

where f = price of the derivative, t = time, r = risk-free rate of return, S = price of the stock at time t and σ = volatility of the stock price.

As mentioned the solution of the differential equation depends on the used limits. In example the limit for European call option is f = max(S – X, 0), when t = T. In limit X = strike price and T = strike time. Black-Scholes-Merton pricing equations for European call (c) and put (p) options are (Hull 2000: 250, compare Black et al. 1973: 644, equation 13)

(2) $$c = S_0 e^{-qT} N(d_1) - Xe^{-rT} N(d_2)$$

and

(3) $$p = Xe^{-rT} N(-d_2) - S_0 e^{-qT} N(-d_1),$$

where

(4) $$d_1 = \frac{\ln(S_0/X) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}$$

and

(5) $$d_2 = \frac{\ln(S_0/X) + (r - q - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}.$$

In equations (2)-(5) S_0 = price of an asset at time 0, N(x) = cumulative normal distribution (probability that a variable is smaller than x), T = validity time, X = strike price, r = risk-free rate of return and σ = volatility of the price of an asset. Variable q is the rate factor that will get different values depending of the type of the asset. When the asset is (i) the stock that doesn't have dividend payments, q = 0, (ii) the stock index, q = mean dividend of the index, (iii) the currency, q = foreign risk-free interest rate r_f, and (iv) the future, q = domestic risk-free interest rate r. (Hull 2000: 250, 275, 284, 294.)

It is possible to use market values as values for all variables, the volatility excluded, in the Black-Scholes-Merton pricing equations. The volatility of a stock is calculable from some popular option that is in the market (Hull 2000: 255). The implicit volatility is the volatility that is needed to get the market value and the Black-Scholes-Merton value of the stock price to be similar with each other. It is not possible to get the volatility implicitly from pricing equations so the only way is to iterate. The implicit volatility is used when someone calculates the prices of more rare securities. (Jarrow 1999.)

The practical uses of the Black-Scholes-Merton model are Greeks (different derivatives of the equation (1)) that are used when taking cover. *Delta* is the first derivate of the asset time against the asset price. If delta is zero then the price of an option will not change when the price of an asset changes. *Gamma* is the second derivate of the asset time against the asset price. The proportions of options and assets have to be changed if the target is the risk-free portfolio and gamma is different from zero. *Tau* is the first derivative of the price against time so it is like a value of portfolios risk as a function of time. *Rho* is the derivate of the price against the rate of return so it is the risk of the portfolio for changes in rate of return. With Greeks it is possible to study sensitivity of the price of an option. (Broadie et al. 2004: 1147–1148; see Hull 2000: 307–341.)

Later it was necessary to get rid of the limitations of the Black-Scholes-Merton model, notably the constant rate of return and the constant volatility of a price distribution of the stock. Already in 70s and 80s the model was improved with extensions where the volatility was not constant (in example Hull and White 1987; Scott 1987). Fluctuation of rate of returns increased in 80s when several researchers developed extensions to the Black-Scholes-Merton model without an assumption of a constant rate of return (see Jarrow 1999).

Broadie et al. (2004) have calculated implicit volatilities from the Black-Scholes-Merton model for the S&P 500 -index for the options of the stocks of the New York's stock exchanges 500 biggest companies at February 4th 1985 and September 16th 1999. Results can be seen in the Figure 1. It can be seen from the figure that the strike price and the maturity of an option affect the implicit volatility. In year 1985 volatilities fluctuated between 14–18 % but in 1999 the fluctuation was significantly bigger (from 16 to 53 %). As Figure 1 implies the Black-Scholes-Merton model's assumption of the constant volatility is quite far from the reality. One of the restraints in the Black-Scholes-Merton model is an assumption of the constant volatility of an asset. However the model implied volatilities for the options of an asset are different for different strike prices. This so called volatility smile has been explained by different stochastic models.

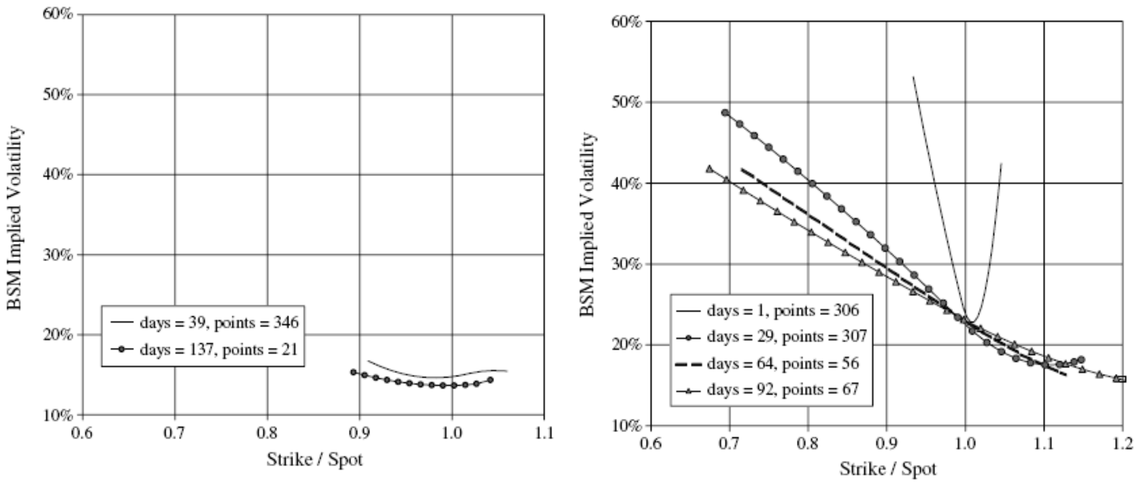

Figure 1. Volatilities for options of stocks in the S&P 500 index as implied by the Black-Scholes-Merton model. Left February 4th 1985, right September 16th 1999. Days means the maturity of an option and points means how many events have been used in drawing. (Broadie et al. 2004: 1160.)

3.2. Volatility models

There are two ongoing branches of research which try to soften the constant volatility constraint of the Black-Scholes-Merton model. The first one of these is the continuous time stochastic volatility models and another branch is the discrete time GARCH models of volatility. The abbreviation GARCH stands for Generalized Autoregressive Conditionally Heteroskedastic models. Main difference, simultaneously with the time scale, between models is a relationship between the volatility and the asset return. In stochastic volatility models the volatility and the asset return behave differently although they can be correlated. In GARCH models the volatility goes hand in hand with the asset return. There are some recent publications that make a good summary on volatility models like publication edited by Bauwens et al. (2012) and a study specifically on the stochastic volatility papers edited by Shephard (2005).

3.2.1. Stochastic volatility

Clark (1973) brought forward a stochastic process to model the price series. He stated in his article that the log-price of an asset can be modelled with process for the changes on the levels and another process for the volatility of the asset.

The stochastic volatility process as a model was first described by Taylor (1982). In his model he used discrete time series instead of continuous which has later become a describing quality on the stochastic volatility models. Taylor's stochastic volatility model as denoted by Bauwens et al. (2012: 25) is as follows:

(6) $$y_t - \mu_t = \varepsilon_t = \sigma_t z_t,$$

(7) $$\log \sigma_{t-1}^2 = \omega + \beta \log \sigma_t^2 + \sigma_u u_t,$$

where the innovations z_t and u_t are normally distributed and independent.

Hull et al. (1987) developed a model for the pricing of a European call option on an asset with the stochastic volatility. In their study they also examine the effect that the correlation between the stochastic volatility and asset price has on the modelled option price when compared to the Black-Scholes-Merton price of the said option. The result is that (i) for the uncorrelated volatility the Black-Scholes-Merton prices on the at-the-money options are overvalued and on the deep in- and out-of-the-money options are undervalued. (ii) The positive correlation means undervalued out-of-the-money option prices and overvalued in-the-money option prices by the Black-Scholes-Merton formula. (iii) With the negative correlation the effect is opposite from that with the positive correlation.

Heston (1993) derives a closed-form solution for the price of a European call option on an asset with the stochastic volatility. Heston comments that the correlation between the volatility and spot-asset returns is needed to explain a return skewness and strike-price biases in the Black-Scholes-Merton model. The effect of the correlation on a price difference between his model and the Black-Scholes-Merton model is similar to Hull et al. (1987). Lately there have been many studies on the multivariate stochastic volatility models for multivariate asset returns. These are given a good introduction to by Omori and Ishihara (2012).

3.2.2. GARCH

Engle (1982) was first to introduce an autoregressive conditionally heteroskedastic model in his study on the variance of United Kingdom inflation. He describes a ARCH process as a *"mean zero, serially uncorrelated process with nonconstant variances conditional on the past, but constant unconditional variances"*. Engle gives the ARCH model few characteristics that help in the forecasting of the economic process. The variance in the model can change with the time and it correlates with the past and the model also allows for an unexplained exogenous variable in the variance. Bollerslev (1986) generalized Engle's model to the so called GARCH model.

Bos (2012: 148) gives two reasons that favour the use of the GARCH models of the volatility against the use of the stochastic volatility models. First reason being the easiness, as the GARCH model estimates can be found in the most statistical packages. The stochastic volatility so far does not have any options for that. They state that the second reason is that there is only one estimation method for many different GARCH models against multiple methods for few stochastic volatility models.

Shephard (2005: 21) for his part concludes that the stochastic volatility models have come to compete with the GARCH models in research. He gives two reasons for the growth spurt in the stochastic volatility models popularity. The first reason is the high frequency price data that has become more and more available for different commodities. The second reason according to Shephard is that there are many papers giving examples on how to use a high frequency data. Caporin and McAleer (2012) discuss the model selection between the different stochastic volatility and GARCH models on time series analysis.

3.2.3. Realized volatility

In words of Andersen and Benzoni (2009) the realized volatility is *"a nonparametric ex-post estimate of the return variation"*. The realized volatility is calculated from the return of the asset and it has become increasingly popular in volatility studies. This is mostly due to the availability of the high frequency data for financial assets. (Andersen et al. 2009, Bauwens et al. 2012.)

The realized volatility calculated from the high frequency intraday returns closes on the underlying integrated volatility (e.g. Andersen, Bollerslev, Diebold and Labys 2001). Bauwens et al. (2012) define the latent integrated volatility as follows:

(8) $$\mathrm{IV_t} \equiv \int_{t-1}^{t} \sigma^2(s)ds \,.$$

Bauwens et al. (2012: 35) also comment positively for the use of the GARCH on estimation of the integrated volatility as the result is much less noisy and that the intraday realized volatility still estimates well on the level of integrated volatility.

3.3. Correlation

The dependence between two variables can be modelled with the equation of a straight line, Y=A+BX. In the equation A is Y intercept and B is the slope of the line. In the finance and economics a usual notation is β_0 for the intercept and β_1 for the slope. When the randomness is taken in to the account, we have to include an error term ε to the equation. The model with the error term is called a simple linear regression model:

(9) $$Y = \beta_0 + \beta_1 X + \varepsilon.$$

The equation (9) is very useful in the physics and mechanics, where the laws behind the phenomenon are generally well known. The equation can be successfully applied also in the economics and finance when we examine the dependence of two known factors. A simple example is the price of the apartments which can be modelled as a dependent variable Y with a surface area as an independent variable X. Estimates of constants β_0 and β_1 can be found using the method of least squares (Aczel & Sounderpandian 2006: 433).

In the method of least squares estimators b_0 and b_1 for constants β_0 and β_1 can be obtained with the error term e of the equation $y = b_0 + b_1x + e$. The best line of regression is certainly the one that has the smallest collective distance to the data points. This collective sum is the smallest when the sum of absolute values of error terms e is as small as possible. This situation is equal to the minimum of sum of squared error terms Σe^2.

Possible minimum is always where all first degree partial differentials of the equation are equal to zero simultaneously. This means that $\partial e^2/\partial b_0=0$ and $\partial e^2/\partial b_1=0$. The estimators, b_0 and $b_{1,}$ for constants β_0 and β_1 are obtained by solving these two equations. (Aczel et al. 2006: 434-435.)

3.3.1. The basics of correlation

The assumption in the regression analysis is always that the values of independent variable (here X) are fixed and not random (error term ε makes the Y random variable in the simple linear regression model). The situation in the economics and finance is usually that we want to interpret the relationship between two random variables. The study of relationship between two random variables is called correlation analysis. (Aczel et al.: 458) Aczel et al. defines the correlation as follows

> *"The correlation between two random variables X and Y is a measure of the degree of linear association between the two variables."*

Both variables are equal random variables in the correlation analysis. This is where the correlation analysis differs from the regression analysis where one of the variables is fixed and only one variable is investigated as a random variable. The correlation between samples of two variables is indicated by the sample correlation coefficient r (the population correlation coefficient is symbolized by ρ). Brealey, Myers and Allen (2006: 995) define the correlation coefficient as follows

> *"(A correlation is a) measure of the closeness of the relationship between two variables"*

The correlation coefficient can be anything from -1 to 1. This can be seen from the expression of the sample correlation coefficient $\Sigma xy/\sqrt{(\Sigma x^2)(\Sigma y^2)}$. Negative values of the correlation coefficient tell us that values of one variable are negatively linearly related to the values of the other variable. The linearity decreases from a perfect negative linear relationship at r (or ρ) equal to -1 to no correlation at all at r equal to zero. A situation with positive values of the correlation coefficient is analogous. The linearity increases from no correlation at r equal to zero to a perfect positive linear correlation at r equal to 1. (Snedecor & Cochran 1989: 177-180, Brealey et al. 2006: 458-459)

The correlation can be illustrated with the regression analysis. The equations $y = b_0 + b_1x$ and $x = c_0 + c_1y$, obtained by the method of least squares, give us identical declining lines (in XY-coordinates) when r is equal to -1 (Figure 2 (a)). When r is between -1 and zero, the lines are declining but they cut each other (Figure 2 (b)). An angle between the lines increases when the correlation coefficient approaches zero until r is equal to zero and lines are in the right angle (Figure 2 (c)). Similarly the angle between the lines decreases when the correlation coefficient increases from zero to 1. Only this time both lines are rising (Figure 2 (d)). At last the lines are equal rising lines when the correlation coefficient is equal to 1 (Figure 2 (e)).

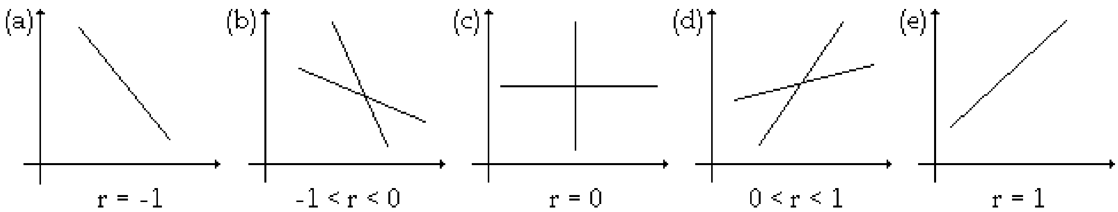

Figure 2. Five different stages of the correlation coefficient r. (a) Perfect negative correlation. (b) Negative correlation. (c) No correlation. (d) Positive correlation. (e) Perfect positive correlation. (Figure made by author).

3.3.2. Population correlation coefficient

The population correlation coefficient ρ measures how close the relationship between populations is. The equation for the population correlation coefficient is (Aczel et al. 2006: 450)

(10) $$\rho = \mathrm{Cov}(X,Y)/\sigma_X\sigma_Y \,,$$

where $\mathrm{Cov}(X,Y) = E[(X-\mu_x)(Y-\mu_y)]$, μ_x is mean of X and μ_y is mean of Y, σ_x and σ_y are standard deviations of X and Y, respectively. The assumption, as stated by Aczel et al. (2006), for variables X and Y in the correlation analysis is that they are *"normally distributed random variables with means μ_x and μ_y and standard deviations σ_x and σ_y"*.

It is possible to use the sample correlation coefficient r, which is known also as the Pearson product-moment correlation coefficient, as an estimator for population correlation coefficient. This is beneficial because ρ, as population parameters tend to, is not known. This estimator r we can calculate from known data by using an equation (Aczel et al. 2006: 450)

(11) $$r = SS_{XY}/\sqrt{SS_X SS_Y}\ ,$$

where SS_X and SS_Y are sums of squares for X's and Y's and SS_{XY} is sum of cross-products between X and Y.
With the sample correlation coefficient we can test whether the correlation between variables X and Y exists at all. The null hypothesis of this test is $\rho = 0$ and the alternative hypothesis is $\rho \neq 0$. The test statistics for the correlation test is (Snedecor et al. 1989: 187)

(12) $$t_{n-2} = r\sqrt{n-2}/\sqrt{1-r^2}.$$

4. DATA AND METHODOLOGY

Data used in this study consists of daily price indices from January 2002 through December 2014 for indices the S&P 500 Composite – price index (SPX) and the CBOE SPX Volatility VIX – price index (VIX). Second set of data consists of equally daily price indices from January 2002 through December 2014 for indices DAX 30 Performance – price index (DAX) and VDAX – New Volatility index – price index (VDAX). Daily returns for these four indices are derived to be used in the analysis. Producers of the used indices are as follows: SPX is produced by Standard & Poor's, VIX is produced by Chicago Board Options Exchange (CBOE), DAX and VDAX are both produced by Deutsche Börse.

Chapter 4.1 introduces index producers CBOE and Deutsche Börse and basics behind the indices used in the thesis. Chapter 4.2 describes data details and chapter 4.3 outlines calculations used in the study to obtain the time series for a dependent variable. Chapter 4.4 has a description for the process that is used when choosing independent variables for the regression analysis. Chapter 4.5 will define the regression analysis used to obtain results of the thesis. Chapter 4.6 introduces all independent variables.

4.1. Publishers and indices

Standard & Poor's originates from the publication *History of the Railroads and Canals of the United States*. That was published by Henry Varnum Poor at 1860 and is one of the first stock analyses in the world. Standard & Poor's has provided wide range of credit ratings, indices, investment research, risk evaluation and data for the investors. One of the most famous products of Standard & Poor's is index S&P 500. The S&P 500 index includes 500 leading companies in leading industries of the U.S. economy. The index covers about 75% of the U.S equities market and it is almost ideal proxy for the total market. (Standard & Poor's 2013.) The level of the index from early 2002 to late 2014 is plotted in a Figure 3.

Chicago Board Options Exchange (CBOE) was founded 1973 as first U.S. options exchange and it provided the platform for the trade of standardized, listed options. In 1993 CBOE introduced the CBOE Volatility Index (VIX). VIX is widely followed and considered to be a good barometer of investor sentiment and market volatility. It measures 30-day implied volatility conveyed by the S&P 500 stock index (SPX) option prices and is calculated as a weighted average of the prices of all out-of-the-money call and put options from two nearby expiration dates. In other words the VIX is annualized 30-day variance expressed in percentages.

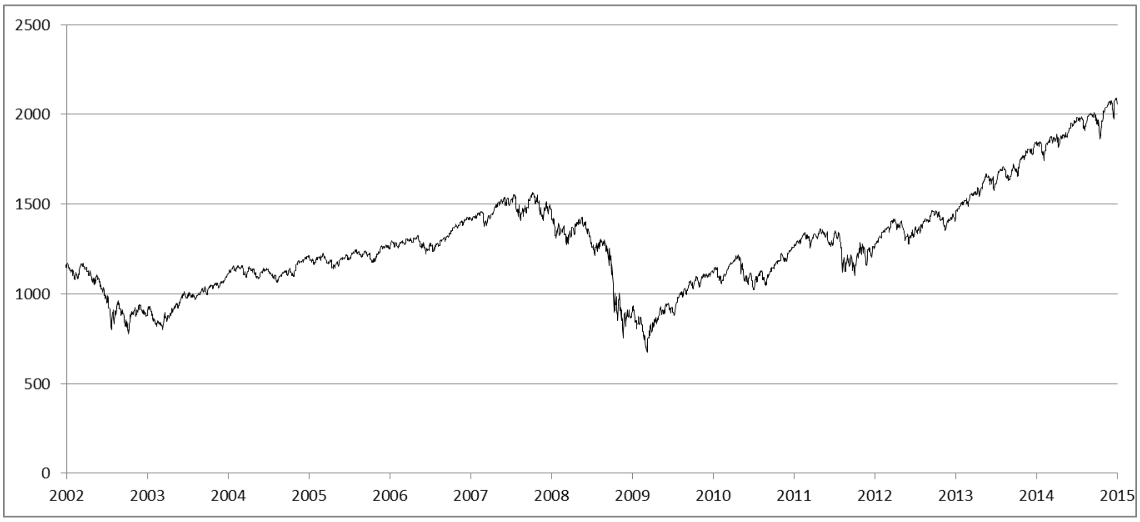

Figure 3. SPX level from the start of 2002 to the end of 2014 (Data by Datastream).

The level of the index from early 2002 to late 2014 is plotted in the Figure 4. Historically the VIX has got its highest levels during times of financial crises. So when markets have declined rapidly it has caused VIX levels to climb up and as markets recover, VIX levels tend to drop. This can be seen when comparing the Figure 3 and Figure 4 where the VIX and SPX are plotted from 2002 to 2014. Several drops in the SPX are accompanied by a rise in the VIX level. It has to be noted that historical performance does not indicate future results. (CBOE 2009, Corrado et al. 2005.)

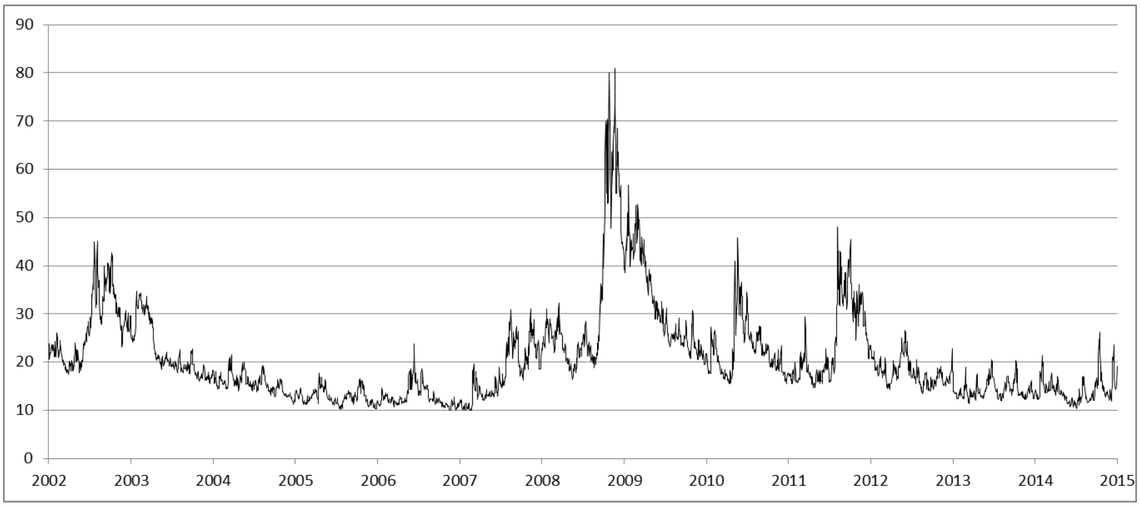

Figure 4. VIX level from the start of 2002 to the end of 2014 (Data by Datastream).

Deutsche Börse Group is a German company that operates Frankfurt Stock Exchange. Deutscher Aktien IndeX 30 (DAX) consists of the 30 largest German companies in terms of order book volume and market capitalization trading on the Frankfurt Stock Exchange. The VDAX index represents the implied volatility of the DAX calculated from the DAX option contracts. The VDAX indicates the volatility of the DAX to be expected in the next 30 days. (Deutsche Börse 2013.)

The levels of the DAX and VDAX from 2002 to 2014 are plotted in Figure 5 and Figure 6. From these two figures we can see the same kind of the behaviour as there is with the SPX and VIX. There are several large drops in the DAX level associated with simultaneous rises in the VDAX level.

Figure 5. DAX level from the start of 2002 to the end of 2014 (Data by Datastream).

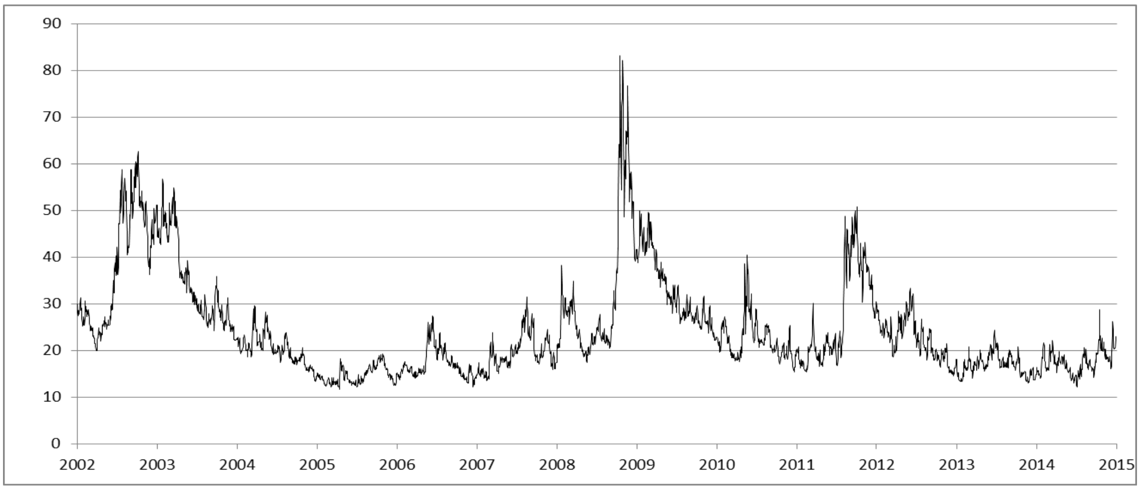

Figure 6. VDAX level from the start of 2002 to the end of 2014 (Data by Datastream).

4.2. Data description

The data used in this thesis was obtained from Thomson Datastream provided by University of Vaasa. Thomson Datastream is a well-known historical financial database which offers wide range of financial data from all-over the world. Datastream is acknowledged as a reliable source of scientifically accepted data.

All index data, SPX, VIX, DAX and VDAX, used in this thesis is from January 2002 to December 2014. The levels of all four indices are plotted in Figure 3 through Figure 6. The descriptive statistics for all data sets, including the mean, standard deviation, coefficient of skewness, and coefficient of kurtosis are provided in Table 1.

From the descriptive statistics it can be seen that the distributions of the VIX and VDAX have positive skewness which means that there is longer tail toward bigger values than toward small values. This is understandable as volatility cannot pass to negative but there is not any theoretical positive side boundary. Positive kurtosis on both volatility indices means that the distributions are more peaked than the normal distribution.

Table 1. Descriptive statistics for indices SPX, DAX, VIX, and VDAX from 2002 to 2014.

	SPX	DAX	VIX	VDAX
Mean	1 278,04	6 008,12	20,36	24,31
Standard Error	4,88	31,26	0,16	0,18
Standard Deviation	284,08	1 820,71	9,38	10,63
Kurtosis	0,34	-0,64	6,47	3,10
Skewness	0,74	0,22	2,14	1,71
Range	1 414,04	7 884,16	70,97	71,58
Minimum	676,53	2 202,96	9,89	11,65
Maximum	2 090,57	10 087,12	80,86	83,23
Count	3 392	3 392	3 392	3 392

4.3. Volatility measures

Implied volatilities will be denoted as and realized volatilities for indices will be calculated as proposed by Corrado et al. (2005). Measure for the realized volatility is the sample standard deviation of the index return for every month. This standard deviation is then level adjusted, from calendar month measure to 22 trading days measure per month, and annualized. The equation for a calculation is as follows (Corrado et al.: 342):

$$\text{(13)} \qquad \mathrm{VOL}_{\mathrm{m}}^{\mathrm{ind}} = \sqrt{\frac{30}{22} \times \frac{252}{\mathrm{n_m}-1} \sum_{\mathrm{d}=1}^{\mathrm{n_m}} \left(\mathrm{r_{d,m}} - \bar{\mathrm{r}}_{\mathrm{m}}\right)^2}\,.$$

In the equation (13), $r_{d,m}$ is an index return on day d in month m, ind is appropriate base index abbreviation, and n_m is the number of trading days in month m. This realized volatility measure is calculated for every calendar month on the data, so that nonoverlapping monthly volatility series is achieved. Sample size of the realized volatility is 156 points (156 months) for both SPX and DAX. Contemporaneous implied volatility will be denoted as VIX_m and $VDAX_m$, observed as the last volatility index value in month m. This will produce series of monthly volatilities with same length as the realized volatility series. The monthly implied volatility represents market expectation of the return volatility for next month so the study will be done on paired measures, where $\mathrm{VOL}_{\mathrm{m}}^{\mathrm{SPX}}$ is paired with VIX_{m-1}, and $\mathrm{VOL}_{\mathrm{m}}^{\mathrm{DAX}}$ is paired with $VDAX_{m-1}$. This means that the realized volatility calculated on a monthly-basis is aligned with the implied volatility observed on the prior month.

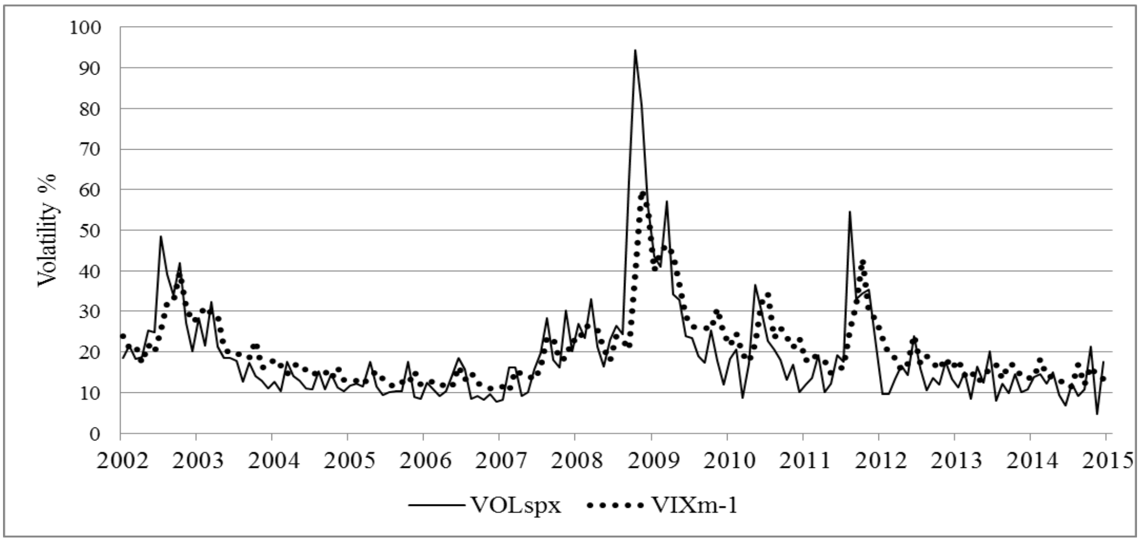

Figure 7. SPX realized volatility measure $\mathrm{VOL}^{\mathrm{SPX}}$ and monthly implied volatility $\mathrm{VIX}_{\mathrm{m-1}}$.

Figure 7 and Figure 8 present the time series of realized volatilities and implied volatilities for both measure pairs. Figure 7 plots the implied volatility VIX_{m-1} and the realized volatility VOL_m^{SPX} for SPX from 2002 through 2014. Figure 8 plots the implied volatility $VDAX_{m-1}$ and the realized volatility VOL_m^{DAX} for DAX from 2002 through 2014. Realized volatilities are plotted with solid lines and implied volatilities are plotted with dashed lines. Summary descriptive statistics, including the mean, standard deviation, coefficient of skewness, and coefficient of kurtosis, for monthly volatility measures are provided in Table 2.

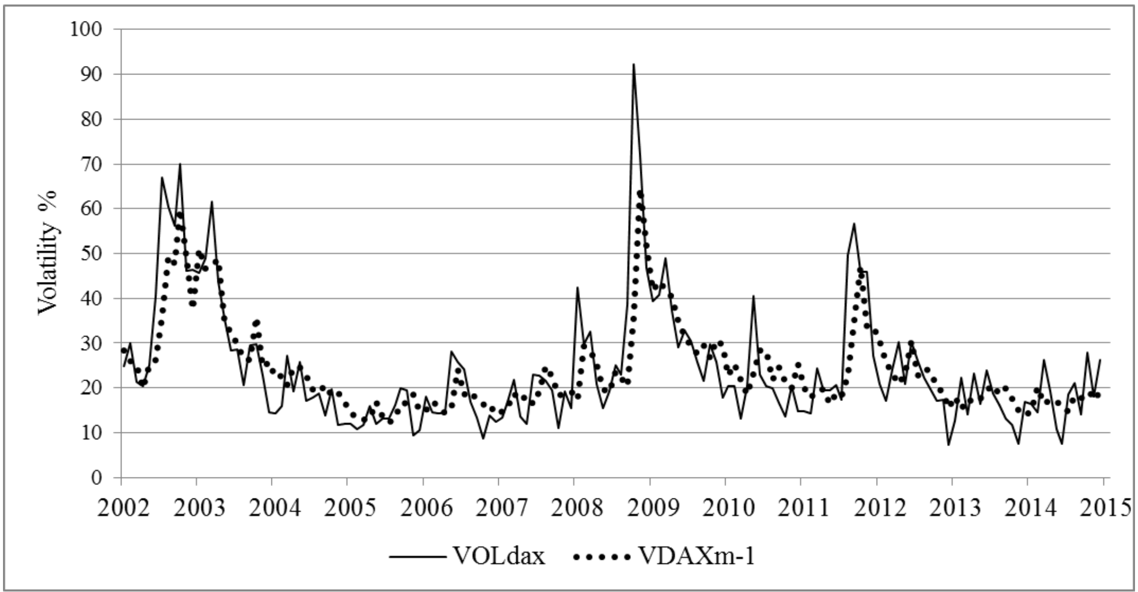

Figure 8. DAX realized volatility measure $\mathrm{VOL}^{\mathrm{DAX}}$ and monthly implied volatility $\mathrm{VDAX}_{\mathrm{m-1}}$.

Table 2. Descriptive statistics for monthly volatility measures VOL_m^{SPX}, VIX_{m-1}, VOL_m^{DAX}, and $VDAX_{m-1}$.

	VOL^{SPX}	VIX_{m-1}	VOL^{DAX}	$VDAX_{m-1}$
Mean	19,35	20,41	24,47	24,12
Standard Error	1,04	0,70	1,12	0,80
Standard Deviation	13,01	8,72	14,02	10,00
Kurtosis	10,34	4,07	4,50	2,76
Skewness	2,77	1,80	1,92	1,67
Range	89,70	49,47	85,00	52,36
Minimum	4,67	10,42	7,30	12,32
Maximum	94,37	59,89	92,30	64,68
Count	156	156	156	156

4.3.1. Volatility ratio – Dependent variable

The volatility ratio is calculated as the implied volatility divided by the realized volatility, VIX_{m-1} / $\mathrm{VOL}_{\mathrm{m}}^{\mathrm{SPX}}$ and $VDAX_{m-1}$ / $\mathrm{VOL}_{\mathrm{m}}^{\mathrm{DAX}}$. These volatility ratios for the SPX and DAX are plotted in Figure 9 and Figure 10, respectively. Graphical plot of the volatility ratio time series does not give much insight to the time behaviour of the ratio. Descriptive statistics for the volatility ratios are provided in Table 3. The most specific difference between the SPX and DAX volatility ratios is that in average the VIX is 19,9% higher that next months realized volatility when the same difference for the VDAX is only 9,2%. Range on both ratios is from a low 0,3 to a high value over 2.

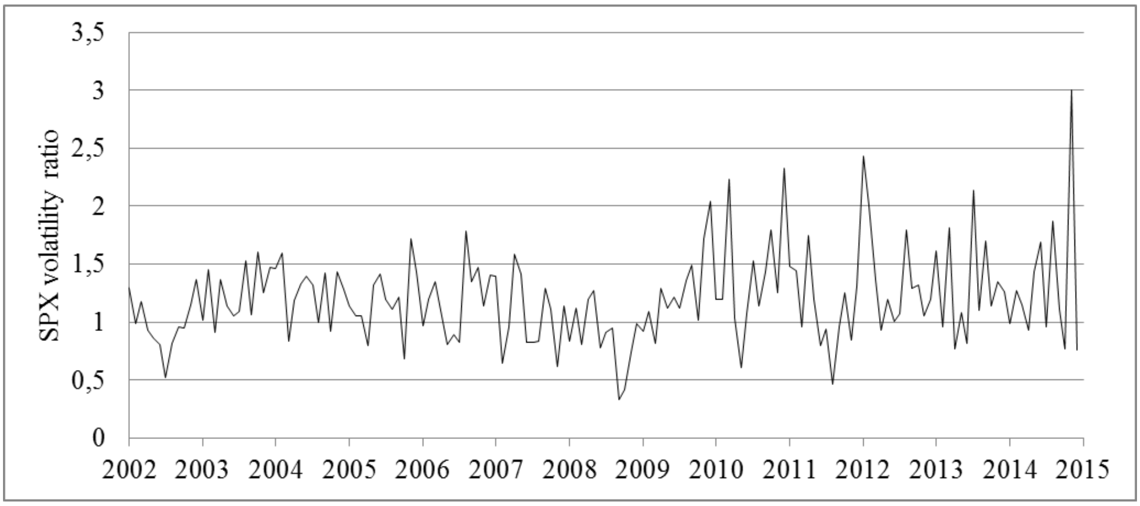

Figure 9. Level of the SPX volatility ratio from 2002 through 2014.

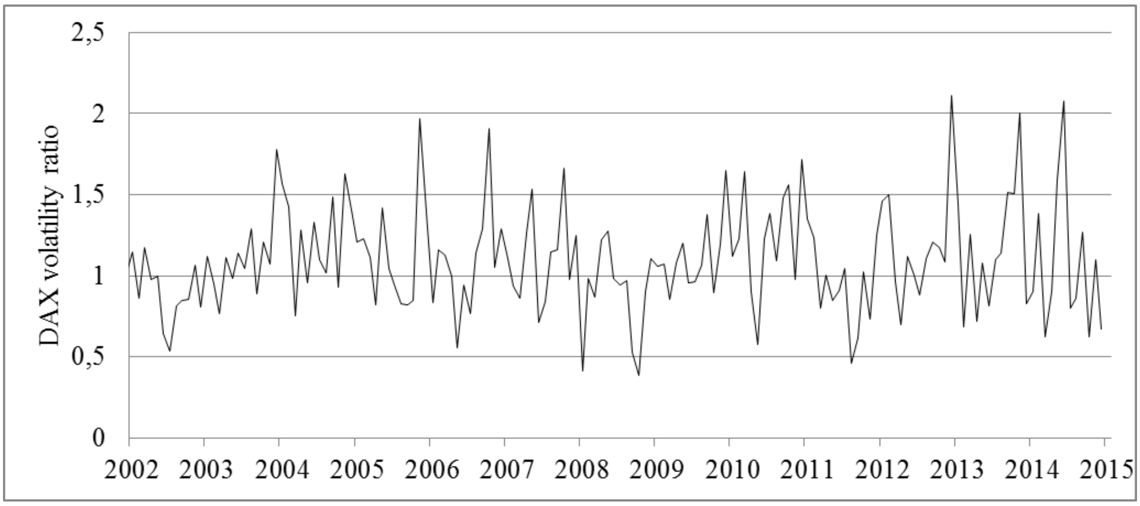

Figure 10. Level of the DAX volatility ratio from 2002 through 2014.

Table 3. Descriptive statistics for the SPX and DAX volatility ratios.

	SPX Vol.ratio	DAX Vol.ratio
Mean	1,199	1,092
Standard Error	0,031	0,026
Standard Deviation	0,389	0,323
Kurtosis	3,134	0,828
Skewness	1,160	0,670
Range	2,670	1,721
Minimum	0,333	0,388
Maximum	3,003	2,108
Count	156	156

4.4. Model selection

The model selection is an important factor that has to be taken in to the account when creating models for the regression analysis. It is not always a best practice to stack independent variables one after another in the regression equation. Some variables can have minimal impact on a dependent variable and some can be totally meaningless. Difficult question is how to choose the best model. Which variables are the most important predictors for the behaviour of the dependent variable?

Ratner (2010) and Draper & Smith (2014: 327-342) go through five widely used methods for the variable selection. These methods are *forward selection, backward elimination, stepwise, R-squared*, and *all-possible subsets*. Different test statistics are used in the methods. F statistic is used for forward selection, backward elimination, and stepwise methods. R-squared is used for R-squared method. With all-possible subsets method the choice is made between R-squared, adjusted R-squared, and Mallows C_p.

In a forward selection method the test statistic, which in this case is F statistic, is calculated individually for all independent variables. One that has the largest F statistic value is chosen to be included in the model. In the next step all other independent variables are again individually added to the model. The one with the largest partial F statistic value is included. The inclusion in all steps requires that chosen variable has the test statistic value more than a pre-set value which usually corresponds to p-values 0,05 or 0,10. The model is complete when none of the partial F statistics surpass a pre-set value.

Stepwise method is similar to forward selection method. The difference is that on each step the partial F statistics are calculated for all included independent variables. These values are then compared to the pre-set value to determine whether any variables should be excluded from the model. The model is completed when all variables included and none variables excluded have test statistics more than a pre-set value.

Backward elimination method uses the same test statistic as previous two methods. The difference is that a starting point is with all possible independent variables in the model. The one with the lowest test statistic is eliminated from the model in each step. The elimination continues until none of the variables in the model have test statistics less than a pre-set value.

R-squared method gives more leeway for a statistician in a model selection. The method generates a bunch of different size subsets of independent variables that have best predicting power on the dependent variable. A statistician makes a selection between given models using their own preferences.

All-possible subsets method is one where all subsets of independent variables are compered using chosen test statistics. Upside of the method is that all possible variable combinations are looked at. Downside is that when the count of independent variables is large the amount of possible models is huge. With 9 independent variables as is the case in this thesis there are $2^9 = 512$ different possible models including one with none of the variables.

The stepwise method is chosen to be used in this thesis. The method is fairly straightforward to execute but it still gives some consideration on the dependencies between all variables included in the model.

4.5. Regression analysis

The behaviour of the volatility ratio is inspected with the multiple linear regression analysis. A dependent variable in the regression will be the volatility ratio that is modelled in Chapter 4.3.1. Independent variables are linear variables; economic policy uncertainty index, unemployment rate, change in the consumer consumption, index return, risk free interest rate, and inflation rate and dummy variables; market indicator for the bull/bear market, and indicators for observations during and after the financial crisis on years 2007-2008.

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4 + \beta_5X_5 + \beta_6X_6 + \beta_7X_7 + \beta_8X_8 + \beta_9X_9 + \varepsilon, \quad (14)$$

where Y is a volatility ratio, ε is random error term and $\beta_0 - \beta_9$ are coefficients to be determined by the regression analysis. Independent variables $X_1 - X_9$ for volatility ratios corresponding to the DAX and SPX indices are listed in Table 4. Significant coefficients are depicted using the stepwise method which is introduced in chapter 4.4. Values of these significant coefficients are estimated with the regression analysis by ordinary least squares method.

Table 4. Independent variables in a multiple linear regression analysis for volatility ratios corresponding to the DAX and SPX indices.

	DAX	SPX
X_1	European economic policy uncertainty index	US economic policy uncertainty index
X_2	Germany unemployment rate	US unemployment rate
X_3	Germany consumer spending rate of change	US Real Personal Consumption Expenditures rate of change
X_4	Monthly base index return	
X_5	3-month Euribor annual rate	3-month Treasury bill secondary market rate
X_6	Euro area inflation rate	US inflation rate
X_7	Dummy indicator - Bull/Bear market	
X_8	Dummy indicator - observations during 2007-2008 financial crisis	
X_9	Dummy indicator - observations after 2007-2008 financial crisis	

4.6. Independent variables

Independent variables analysed for an effect to the volatility ratios are listed in Table 4. In total nine variables are tested. Three of the variables are based on different indices: economic policy uncertainty index, consumer spending, and base stock index return. Three of the variables are based on few macroeconomic rates: unemployment rate, estimate for the risk free interest rate, and inflation rate. The last three variables are dummies for the market mode and for the changes in the constant coefficient with time. All data has been acquired for years 2002 through 2014. Last data point from the year 2001 has been included where needed to get the calculated changes from the start of the year 2002.

Economic policy uncertainty indices (X_1) are work of Scott R. Baker, Steven J. Davis, and Nicholas Bloom. Methodology behind the indices can be found on the website www.policyuncertainty.com. In short the index is constructed from three components: component for the newspaper coverage, component for the federal tax code provisions (United States), and component for the forecasters' opinions. European economic policy uncertainty index is built similarly to the US index. More on the methodology behind the indices can be found in a working paper by Baker, Bloom, and Davis (2015). Policy uncertainty indices from year 2002 through 2014 are plotted in Figure 11. Descriptive statistics for US uncertainty are shown in Table 5 and for EU uncertainty in Table 6. Data was collected from the website 27.3.2016.

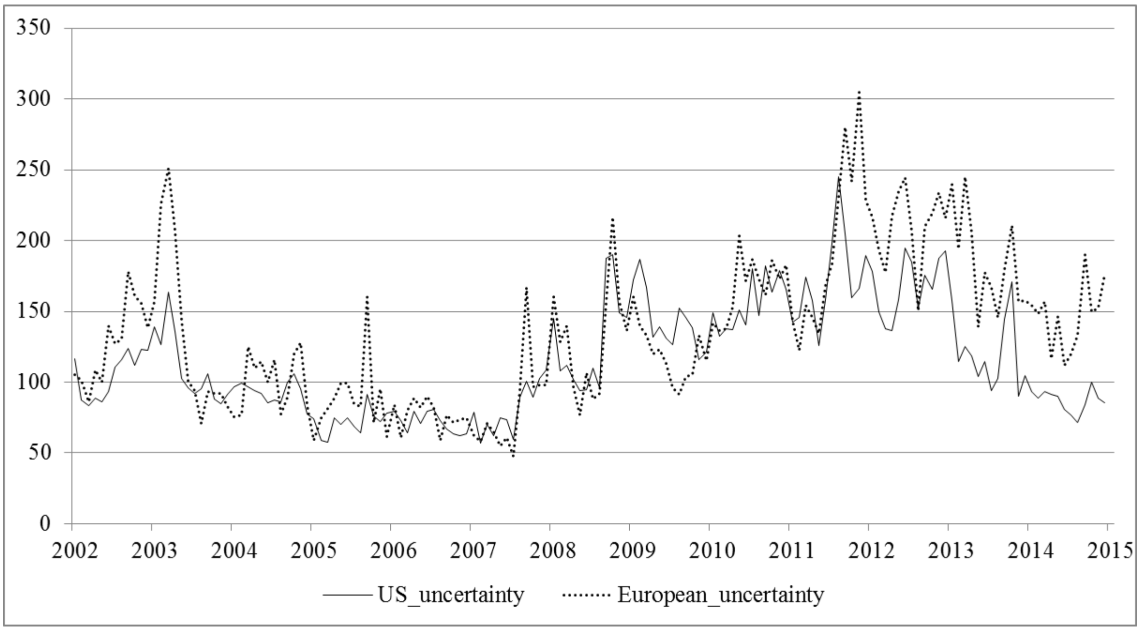

Figure 11. European economic policy uncertainty index (dashed line) and US economic policy uncertainty index (solid line) from year 2002 through year 2014.

Unemployment rates (X_2) used in the thesis are Germany unemployment rate and US unemployment rate. Data for Germany unemployment rate was obtained from European Central Bank Statistical Data Warehouse (http://sdw.ecb.europa.eu; "Germany - Standardised unemployment, Rate, Total (all ages), Total (male & female)”; 27.3.2016). Data for US unemployment rate was obtained from United States Department of Labor database (http://data.bls.gov/pdq/SurveyOutputServlet; Labor Force Statistics from the Current Population Survey, Series ID LNS14000000; 27.3.2016). Descriptive statistics for the US unemployment rate are shown in Table 5 and for Germany unemployment rate are shown in Table 6.

The measure for the consumer consumption (X_3) was hard to get and a little creativity was required for the final monthly time series. There is not ready to be used measure for Germany (or EU) personal consumption and the data used for the thesis is seasonally adjusted Germany consumer spending with constant prices. Data is from Thomson Datastream database via University of Vaasa. The problem with this measure is that it is obtained quarterly when the analysis in the thesis is done with the monthly time series. Quarterly rate of change in percentages is calculated and then compounded to get the annual rate. Thus obtained annual rate of change per quarter was copied for each month in the quarter. US data is real personal consumption expenditures obtained from Federal Reserve Bank of St. Louis database (http://research.stlouisfed.org/fred2; “Personal Consumption Expenditures: Chain-type Price Index, Index 2009=100, Quarterly, Seasonally Adjusted”; 27.3.2016). For the purposes of this thesis the monthly rate of change was calculated and then compounded to represent annual rate. Descriptive statistics for US consumer measures are shown in Table 5. Statistics for Germany measure are based on quarterly values without duplicates. Descriptive statistics for Germany consumer measures are shown in Table 6

Index returns (X_4) for the SPX and DAX indices were calculated on monthly basis. This was done by a simple calculation $r_m = (I_m - I_{m-1})/I_{m-1}$ (where m stands for a month). Descriptive statistics for these calculated returns are shown in Table 5 (SPX) and in Table 6 (DAX).

Inflation (X_6) used in the thesis is based on the OECD Consumer Price Index. Data has been obtained from the OECD database (https://data.oecd.org/price/inflation-cpi.htm) which has the inflation rate as monthly annual growth rates (percentage). Germany inflation rate is used as the DAX related inflation and United States inflation is used as the SPX related inflation. Descriptive statistics for inflation rates are shown in Table 5 and in Table 6.

The last independent linear variable is the risk free interest rate (X_5). This was fairly straightforward for US as there is a short term security issued by US government, 3-month Treasury bill. Rates for the Treasury bill were obtained from US Federal Reserve database (http://www.federalreserve.gov/releases/h15/data.htm; "3-month Treasury bill secondary market rate discount basis"; 27.3.2016). Euro area risk free interest rate proved to be a little more complicated. The reason is that though European Central Bank is printing the Euro currency it does not back any short term securities. In this thesis 3-month Euribor annual rate was chosen as an indicator for the short term risk free interest rate. Data for 3-month Euribor was obtained from European Central Bank Statistical Data Warehouse (http://sdw.ecb.europa.eu; "Euro area (changing composition) - Money Market - Euribor 3-month - Historical close, average of observations through period - Euro, provided by Reuters"; 27.3.2016). Descriptive statistics for interest rates are shown in Table 5 (Treasury bill) and in Table 6 (Euribor).

In addition to the six linear variables there are three dummy variables used in the thesis. The first one is market mode (X_7). Denotation for the market mode is a Bear market for negative monthly returns with value 1 for dummy variable and a Bull market for positive monthly returns with value 0 for dummy variable. Second and third dummy variables are tied together with implication of one more variable. Dummy variable for observations during the years 2007-2008 financial crisis (X_8) gets value 1 for observations during that time period and is otherwise 0. Next dummy has value 1 for observations after year 2008 (X_9) and is otherwise 0. These two time sensitive dummies implicate third time period, for observations before year 2007, which is effective when both time period dummy variables have a value 0.

Table 5. Descriptive statistics for the SPX volatility ratio related linear independent variables data from year 2002 through 2014. Columns from left to right: X_1, US economic policy uncertainty index; X_2, US unemployment rate; X_3, US real personal consumption expenditures annual rate of change; X_4, SPX monthly return; X_5, 3-month Treasury bill secondary market rate; X_6, US inflation rate.

	X1	X2	X3	X4	X5	X6
Mean	115,89548	6,68205	0,02149	0,00467	1,39186	2,26770
Standard Error	3,16622	0,13920	0,00324	0,00340	0,13173	0,10637
Standard Deviation	39,54607	1,73861	0,04041	0,04249	1,64535	1,32855
Kurtosis	-0,37819	-1,12725	0,63880	1,70900	-0,22902	1,06199
Skewness	0,63977	0,51010	-0,25739	-0,77053	1,06387	-0,42281
Range	187,92365	5,60000	0,24792	0,27715	5,02000	7,69728
Minimum	57,20263	4,40000	-0,11703	-0,16942	0,01000	-2,09716
Maximum	245,12628	10,00000	0,13089	0,10772	5,03000	5,60012
Count	156	156	156	156	156	156

Table 6. Descriptive statistics for the DAX volatility ratio related linear independent variables data from year 2002 through 2014. Columns from left to right: X_1, EU economic policy uncertainty index; X_2, Germany unemployment rate; X_3, Germany consumer spending annual rate of change (quarterly data); X_4, DAX monthly return; X_5, 3-month Euribor annual rate; X_6, Euro area inflation rate.

	X_1	X_2	X_3	X_4	X_5	X_6
Mean	135,55722	7,82744	0,00743	0,00605	2,02830	1,54561
Standard Error	4,28240	0,16097	0,00327	0,00491	0,11584	0,05624
Standard Deviation	53,48712	2,01057	0,02361	0,06128	1,44688	0,70245
Kurtosis	-0,17915	-1,33408	3,50839	3,06605	-0,83933	0,56688
Skewness	0,63028	0,04124	-0,63895	-0,66943	0,43590	0,08055
Range	256,90894	6,38000	0,14862	0,46800	5,03220	3,82546
Minimum	47,69431	4,86000	-0,07953	-0,25422	0,08090	-0,50251
Maximum	304,60324	11,24000	0,06909	0,21378	5,11310	3,32295
Count	156	156	52	156	156	156

5. RESULTS

Results on the model building and regression analysis for the SPX and DAX volatility ratios are introduced and analysed in this chapter. The model building is done with the stepwise method which was introduced in Chapter 4.4. The regression analysis on chosen variables is then done following denotations in the equation (14). All calculations are performed using Microsoft Excel spreadsheet with Data Analysis toolkit. Chapter 5.1 presents results from the model building and multiple linear regression analysis for the SPX volatility ratio. Chapter 5.2 presents same results for the DAX volatility ratio.

5.1. SPX volatility ratio regression model

Model for the SPX volatility ratio was obtained in three steps. Results of the stepwise method are detailed in Table 7. First step has all 9 independent variables (see Table 4) added individually to the regression with constant coefficient. Variable X_4 (index return) has the largest F statistic and is chosen as the first independent variable on the regression equation. In the second step the largest partial F statistic is for variable X_9 (observations after financial crisis). Partial F statistic for X_4 in the model with X_4 and X_9 is calculated to determine if X_4 is still significant. Result of this analysis is in Table 8. Significance level for both variables is less than 0,05 and variables are retained in the model. In the third step only significant added variable is X_7 (market mode). Variable X_7 is included in the model and all variables are tested for the significance (Table 8). The partial F statistics in the fourth step are all statistically insignificant. This concludes that none of the remaining variables are significant additions to the model and thus the model is complete after three steps.

Table 7. Partial F statistics and corresponding P-values for variables added to the SPX volatility ratio regression model on each step. On each step variable with largest F statistic is chosen to be included in to the model. Step 1 variable X_4, step 2 variable X_9, and step 3 variable X_7. Model is complete after step 3 as on step 4 none of the added variables have partial F statistic with P-value less than 0,05.

	Current variables	New variables	Partial F	P-value
STEP 1	-	X1	0,1636	0,6865
	-	X2	5,3006	0,0227
	-	X3	6,0423	0,0151
	-	**X4**	**59,1755**	**0,0000**
	-	X5	3,5617	0,0610
	-	X6	3,6645	0,0574
	-	X7	46,5600	0,0000
	-	X8	11,7213	0,0008
	-	X9	9,8189	0,0021
STEP 2	X4	X41	0,1079	0,7430
	X4	X42	3,2200	0,0747
	X4	X43	1,7790	0,1843
	X4	X45	3,4267	0,0661
	X4	X46	0,3981	0,5290
	X4	X47	3,3182	0,0705
	X4	X48	5,1351	0,0248
	X4	**X49**	**5,4685**	**0,0207**
STEP 3	X49	X491	2,0658	0,1527
	X49	X492	0,0577	0,8104
	X49	X493	2,2783	0,1333
	X49	X495	0,0378	0,8462
	X49	X496	0,1974	0,6574
	X49	**X497**	**4,0720**	**0,0454**
	X49	X498	2,2906	0,1322
STEP 4	X497	X4971	2,3353	0,1286
	X497	X4972	0,0093	0,9232
	X497	X4973	2,0512	0,1542
	X497	X4975	0,0775	0,7811
	X497	X4976	0,2985	0,5856
	X497	X4978	2,4359	0,1207

Table 8. Analysis for retaining previous independent variables in the SPX volatility ratio regression analysis on each step of model building with stepwise method. Part 1: Step 1 with constant coefficient and X_4. Part 2: Step 2 with constant coefficient, X_4, and X_9. Part 3: Step 3 with constant coefficient, X_4, X_9, and X_7. All added variables on every step have partial F statistic with P-value less than 0,05 and all are significant within 95% confidence level.

	Current variables	New variables	Partial F	P-value	
STEP 1		X4	59,1755	0,0000	Significant
STEP 2	X4	X49	5,4685	0,0207	Significant
	X9	X49	53,2131	0,0000	Significant
STEP 3	X49	X497	4,0720	0,0454	Significant
	X47	X497	6,2189	0,0137	Significant
	X79	X497	10,4790	0,0015	Significant

Full regression analysis is made for the model with independent variables X_4, X_9, and X_7. Results of the analysis are shown in Table 9. Value of R^2 for the regression is 0,32072. This means that about 32% of the volatility ratio variation is explained by the regression. Significance of the full model F statistic shown in ANOVA table (middle part, Table 9) is less than 0,001 and results are statistically significant within 99,9% confidence interval. The lower part of Table 9 shows that values for coefficients β_0, β_4, β_7, and β_9 are statistically significant within 99,9%, 99,5%, 95%, and 95% confidence intervals, respectively.

The model for the SPX volatility ratio with values from the regression may be written as

$$Y = 1{,}187 + 3{,}101 \times X_4 - 0{,}167 \times X_7 + 0{,}132 \times X_9 \tag{15}$$

More informative way to write out the same equation is as

$$\frac{\mathrm{VIX}_{m-1}}{\mathrm{VOL}_m^{\mathrm{SPX}}} = \begin{cases} \begin{matrix} 1{,}020+3{,}101\times r_{spx}, & \text{Bear market} \\ 1{,}187+3{,}101\times r_{spx}, & \text{Bull market} \end{matrix}, & \text{during } 2002-2008 \\ \begin{matrix} 1{,}152+3{,}101\times r_{spx}, & \text{Bear market} \\ 1{,}319+3{,}101\times r_{spx}, & \text{Bull market} \end{matrix}, & \text{during } 2009-2014 \end{cases}. \tag{16}$$

Equation (16) means that the ratio of the implied volatility VIX_{m-1} and the realized volatility $\mathrm{VOL}_m^{\mathrm{SPX}}$ is smaller during a declining market than during an ascending market. Difference on the volatility ratio between the time periods 2002-2008 and 2009-2014 is also statistically significant. The ratio during the latter time period is on average more than the ratio before and during the time of the financial crisis.

Table 9. Regression statistics (upper part), ANOVA table (middle part), and correlation coefficients (lower part) for SPX multiple linear regression equation $Y = \beta_0 + \beta_4 X_4 + \beta_7 X_7 + \beta_9 X_9 + \varepsilon$.

Regression Statistics	
Multiple R	0,56632
R Square	0,32072
Adjusted R Square	0,30731
Standard Error	0,32367
Observations	156

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	7,51819	2,50606	23,92182	0,000
Residual	152	15,92361	0,10476		
Total	155	23,44180			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	1,18653	0,04838	24,52389	0,00000
X_4	3,10143	0,95808	3,23712	0,00148
X_7	-0,16703	0,08277	-2,01791	0,04536
X_9	0,13184	0,05287	2,49377	0,01371

5.2. DAX volatility ratio regression model

Base structure of the results for the DAX volatility ratio model is similar to the SPX model in the previous chapter. Difference being that the model for the DAX volatility ratio required only two steps. Results of the stepwise method are detailed in Table 10. The first step has all 9 independent variables (see Table 4) added individually to the regression with a constant coefficient. Variable X_7 (market mode) has the largest F statistic and is chosen as the first independent variable on the regression equation. In the second step the largest partial F statistic is for variable X_4 (index return). Partial F statistic for X_7 in the model with X_7 and X_4 is calculated to determine if X_7 is still significant. Result of this analysis is in Table 11. Significance level for both variables is less than 0,05 and variables are retained in the model. The partial F statistics in the step 3 are all statistically insignificant. This concludes that none of the remaining variables are significant additions to the model and thus model is complete after two steps.

Table 10. Partial F statistics and corresponding P-values for variables added to the DAX volatility ratio regression model on each step. On each step variable with largest F statistic is chosen to be included in to the model. Step 1 variable X_7, and step 2 variable X_4. Model is complete after step 2 as on step 3 none of the added variables have partial F statistic with P-value less than 0,05.

Step	Current variables	New variables	Partial F	P-value
STEP 1	-	X1	3,0569	0,0824
	-	X2	0,0449	0,8325
	-	X3	0,0588	0,8088
	-	X4	45,8092	0,0000
	-	X5	4,8124	0,0298
	-	X6	1,2140	0,2723
	-	**X7**	**48,9616**	**0,0000**
	-	X8	2,2800	0,1331
	-	X9	1,0679	0,3030
STEP 2	X7	X71	2,1425	0,1453
	X7	X72	0,0573	0,8111
	X7	X73	0,0262	0,8716
	X7	**X74**	**6,5511**	**0,0115**
	X7	X75	2,7826	0,0973
	X7	X76	0,6837	0,4096
	X7	X78	2,0601	0,1532
	X7	X79	0,5160	0,4736
STEP 3	X74	X741	1,4457	0,2311
	X74	X742	0,0050	0,9435
	X74	X743	0,2365	0,6275
	X74	X745	1,7369	0,1895
	X74	X746	0,2568	0,6130
	X74	X748	1,2768	0,2603
	X74	X749	0,3706	0,5436

Table 11. Analysis for retaining previous independent variables in the DAX volatility ratio regression analysis on each step of model building with stepwise method. Part 1: Step 1 with constant coefficient and X_7. Part 2: Step 2 with constant coefficient, X_7, and X_4. All added variables on every step have partial F statistic with P-value less than 0,05 and all are significant within 95% confidence level.

Current variables	New variables	Partial F	P-value	
	X7	48,9616	0,0000	Significant
X7	X74	6,5511	0,0115	Significant
X4	X74	9,0684	0,0030	Significant

Full regression analysis is made for the model with independent variables X_7, and X_4. Results of the analysis are shown in Table 12. Value of R^2 for the regression is 0,27239. This means that about 27% of the volatility ratio variation is explained by the regression. Significance of the full model F statistic shown in ANOVA table (middle part, Table 12) is less than 0,001 and results are statistically significant within 99,9% confidence interval. The lower part of Table 12 shows that values for coefficients β_0, β_4, and β_7 are statistically significant within 99,9%, 95%, and 99,5% confidence intervals, respectively.

Model for the DAX volatility ratio with values from the regression may be written as

$$Y = 1{,}166 + 1{,}360 \times X_4 - 0{,}199 \times X_7. \tag{17}$$

More informative way to write out the same equation is as

$$\frac{\mathrm{VDAX}_{\mathrm{m-1}}}{\mathrm{VOL}_{\mathrm{m}}^{\mathrm{DAX}}} = \begin{cases} 0{,}967+1{,}360\times r_{dax}\,, & \text{on Bear market} \\ 1{,}166+1{,}360\times r_{dax}\,, & \text{on Bull market} \end{cases}. \tag{18}$$

The first part of the equation (18) means that the implied volatility VDAX_{m-1} during a declining market is less than the realized volatility $\mathrm{VOL}_{\mathrm{m}}^{\mathrm{DAX}}$. More negative return means smaller ratio as r_{dax} goes down. The second part of the equation means that on an ascending market the situation is the other way around and the realized volatility passes the implied volatility. More so when r_{dax} goes up.

Table 12. Regression statistics (upper part), ANOVA table (middle part), and correlation coefficients (lower part) for DAX multiple linear regression equation Y = $\beta_0 + \beta_4 X_4 + \beta_7 X_7 + \varepsilon$.

Regression Statistics	
Multiple R	0,52191
R Square	0,27239
Adjusted R Square	0,26288
Standard Error	0,27763
Observations	156

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	4,41476	2,20738	28,63878	0,000
Residual	153	11,79271	0,07708		
Total	155	16,20747			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	1,16567	0,03694	31,55790	0,00000
X_4	1,36024	0,53145	2,55951	0,01145
X_7	-0,19874	0,06600	-3,01137	0,00304

6. CONCLUSION

The aim of this thesis is to investigate the behaviour of the volatility ratio and impact that chosen variables have on the ratio. The volatility ratio is defined as a relation between the implied and realized volatility. Values of the volatility indices VIX and VDAX are taken as a measure for the implied volatility and the realized volatility is calculated for values of the SPX and DAX indices. The behaviour of the ratio is investigated with monthly data for period from 2002 to 2014. This period is divided to three time series; time before, time during, and time after the financial crisis on years 2007-2008. Macroeconomic variables investigated for an impact are unemployment rate, risk free interest rate, economic uncertainty, consumer consumption, market state, and inflation.

The significant variables for the SPX and DAX related volatility ratios are slightly different. The model building and investigation for significant variables shows that for the DAX volatility ratio the retained variables are the market mode and return for the DAX index (see equation (18)). The SPX volatility ratio indicates also a time sensitive behaviour in addition to the market mode and return (return for the SPX index). The ratio is greater after than before or during the financial crisis 2007-2008 (see equation (16)). This means that only hypotheses H_4 and H_7 hold true for both volatility ratios and in addition the hypothesis H_9 holds true for the SPX volatility ratio. The results of analyses regarding the hypotheses H_1-H_9 are shown in Table 13.

Table 13. The results of analyses on significant variables for the SPX and DAX volatility ratio models are shown in the table. Significant variables are indicated by “x” and insignificant variables are indicated by “o”.

Hypothesis		SPX	DAX
H_1	The level of the economic policy uncertainty index has an impact on the volatility ratio	o	o
H_2	The unemployment rate has an impact on the volatility ratio	o	o
H_3	The change in the consumer consumption has an impact on the volatility ratio	o	o
H_4	The monthly index return has an impact on the volatility ratio	x	x
H_5	The risk free interest rate has an impact on the volatility ratio	o	o
H_6	The inflation rate has an impact on the volatility ratio	o	o
H_7	The market mode has an impact on the volatility ratio	x	x
H_8	The volatility ratio level during the years 2007 – 2008 financial crisis is distinctive	o	o
H_9	The volatility ratio level after the years 2007 – 2008 financial crisis is distinctive	x	o

The conclusion of the thesis is that when the monthly index return and market mode are implemented in the volatility ratio models then the impact of the other analysed macroeconomic variables on the ratio is insignificant. Also there is no time sensitive behaviour on the DAX volatility ratio. The level of the SPX volatility ratio has been higher after the years 2007-2008 financial crisis when compared to the earlier years.

Different approaches on the problem could have been used. One would have been to analyse the model with all variables in and then determine which ones are significant within chosen confidence interval. Another way would have been to build models separately on years 2002-2006, 2007-2008, and 2009-2014 (differentiated by dummy variables X_8 and X_9 in the chosen approach) and then investigate if models are significantly different during the periods. These and any other approaches would probably have resulted on the slightly different values for the coefficients and maybe also different significant variables. Selection of independent variables is a qualitative process and there are many other variables that could have been included in to the analysis like liquidity and GDP growth. Different data partitions and independent variables could be used for the further research.

REFERENCES

Aczel, Amir D. & Jayavel Sounderpandian (2006). *Complete Business Statistics*. 6th edition. Singapore: McGraw-Hill. ISBN 007-124416-6.

Andersen, Torben G. & Luca Benzoni (2009). *Realized Volatility*. In Andersen, T.G., R.A.Davis, J.-P.Krei & T.Mikosch (Eds.), Handbook of Financial Time Series, 555-575. Berlin, Germany: Springer-Verlag. ISBN 978-3-540-71296-1.

Andersen, Torben G. & Tim Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39:4, 885–905.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold & Paul Labys (2001). The Distribution of Realized Exchange Rate Volatility. *Journal of the American Statistical Association* 96:453, 42-55.

Baker, Scott R., Nicholas Bloom & Steven J.Davis (2015). Measuring Economic Policy Uncertainty. *NBER Working Paper* No. 21633 [online] [cited 2016-03-20]. 75 p. Available from internet <URL:http://www.nber.org/papers/w21633>.

Bauwens, Luc, Christian Hafner & Sébastien Laurent (2012). *Handbook of Volatility Models and Their Applications.* 1st edition. Hoboken USA: John Wiley & Sons. 565 p. pISBN 978-0-470-87251-2.

Beckers, Stan (1981). Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Variability. *Journal of Banking and Finance* 5:3, 363-381.

Black, Fischer & Myron Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81:3, 637-654.

Blair, Bevan J., Ser-Huang Poon & Steven J. Taylor (2001). Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns. *Journal of Econometrics* 105, 5-26.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.

Bos, Charles S. (2012). Relating Stochastic Volatility Estimation Methods. In: *Handbook of Volatility Models and Their Applications*, 147-174. Ed. Bauwens, Luc, Christian Hafner & Sébastien Laurent. Hoboken, USA: John Wiley & Sons.

Brealey, Richard A., Stewart C. Myers & Franklin Allen (2006). *Corporate Finance*. 8th ed. Boston, MA, USA: McGraw-Hill/Irwin. ISBN 978-0-07-111551-3.

Broadie, Mark & Jerome B. Detemple (2004). Option Pricing: Valuation Models and Applications. *Management Science* 50:9, 1145-1177.

Byun, Suk Joon, Dong Woo Rhee & Sol Kim (2011). Intraday volatility forecasting from implied volatility. *International Journal of Managerial Finance* 7:1, 83-100.

Canina, Linda & Stephen Figlewski (1993). The Informational Content of Implied Volatility. *The Review of Financial Studies* 6:3, 659-681.

Caporin, Massimiliano & Michael McAleer (2012). Model selection and testing of conditional and stochastic volatility models. In: *Handbook of Volatility Models and Their Applications*, 199-221. Ed. Bauwens, Luc, Christian Hafner & Sébastien Laurent. Hoboken, USA: John Wiley & Sons.

CBOE (2009). The CBOE Volatility Index – VIX. [cited 17-December-2013]. Available from World Wide Web: http://www.cboe.com/micro/vix/vixwhite.pdf

Chiras, Donald P. & Steven Manaster (1978). The Information Content of Option Prices and a Test of Market Efficiency. *Journal of Financial Economics* 6, 213-234

Christensen, Bent Jesper & Charlotte Strunk Hansen (2002). New evidence on the implied-realized volatility relation. *The European Journal of Finance* 8:2, 187-205.

Christensen, B.J. & N.R. Prabhala (1998). The relation between implied and realized volatility. *Journal of Financial Economics* 50, 125-150.

Clark, Peter K. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* 41, 135-155.

Corrado, Charles J. & Thomas W. Miller, Jr. (2005). The Forecast Quality of CBOE Implied Volatility Indexes. *The Journal of Futures Markets* 25:4, 339-373.

Draper, Norman R. & Harry Smith (2014). *Applied Regression Analysis*. 3rd edition. New York etc.: John Wiley & Sons, Inc. 706 p. ISBN 0-471-17082-8.

Deutche Börse (2013). [cited 17-December-2013] Available from World Wide Web: http://deutsche-boerse.com /dbg/dispatch/en/kir/dbg_nav/about_us/10_Deutsche_Boerse_Group

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987-1007.

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

Fleming, Jeff, Barbara Ostdiek & Robert E. Whaley (1995). Predicting stock market volatility: A new measure. *The Journal of Futures Markets* 15:3, 265–302.

Garvey, John & Liam A. Gallagher (2012). The Realised-Implied Volatility Relationship: Recent Empirical Evidence from FTSE-100 Stocks. *Journal of Forecasting* 31, 639-660.

Giot, Pierre (2005). Relationship Between Implied Volatility Indexes and Stock Index Returns. *The Journal of Portfolio Management* Spring 2005, 92-100.

Giot, Pierre & Sébastian Laurent (2007). The Information Content of Implied Volatility in Light of the Jump/Continuous Decomposition of Realized Volatility. *The Journal of Futures Markets* 27:4, 337-359.

Gospodinov, Nikolay, Athanasia Gavala & Deming Jiang (2006). Forecasting Volatility. *Journal of Forecasting* 25, 381-400.

Gwilym, Owain A. & Mike Buckle (1999). Volatility forecasting in the framework of the option expiry cycle. *The European Journal of Finance* 5, 73–94.

Han, Heejoon & Myung D. Park (2013). Comparison of Realized Measure and Implied Volatility in Forecasting Volatility. *Journal of Forecasting* 32, 522-533.

Hansen, Charlotte Strunk (2001). The relation between implied and realised volatility in the Danish option and equity markets. *Accounting and Finance* 41, 197-228.

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

Hull, John & Alan White (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance* 42:2, 281-300.

Hull, John C. (2000). *Option, Futures, & Other Derivatives*. 4th edition. New Jersey: Prentice-Hall Inc. 698 p. ISBN 0-13-015822-4.

Jarrow, Robert A. (1999). In Honor of the Nobel Laureates Robert C. Merton and Myron S. Scholes: A Partial Differential Equation That Changed the World. *Journal of Economic Perspectives* 13:4, 229-248.

Jorion, Philippe (1995). Predicting Volatility in the Foreign Exchange Market. *The Journal of Finance* 50:2, 507-528.

Kawaller, Ira G., Paul D. Koch & John E. Peterson (1994). Assessing the Intraday Relationship Between Implied and Historical Volatility. *The Journal of Futures Markets* 14:3, 323-346.

Knight, John & Stephen Satchell (2007). *Forecasting Volatility in the Financial Markets*. 3. ed. The Netherlands: Elsevier Ltd. 428 p. ISBN 0-7506-6942-X.

Koopman, Siem Jan, Borus Jungbacker & Eugenie Hol (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance* 12, 445-475.

Lamoureux, Christopher G. & William D. Lastrapes (1993). Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities. *The Review of Financial Studies* 6:2, 293-326.

Latané, Henry A. & Richard J. Rendleman (1976). Standard Deviations of Stock Price Ratios Implied in Option Prices. *The Journal of Finance* 31:2, 369-381.

Merton, Robert (1973). Theory of Rational Option Pricing. *Bell Journal of Economics and Management Science* 4, 141-183.

Omori, Yasuhiro & Tsunehiro Ishihara (2012). Multivariate Stochastic Volatility Models. In: *Handbook of Volatility Models and Their Applications*, 175-195. Ed. Bauwens, Luc, Christian Hafner & Sébastien Laurent. Hoboken, USA: John Wiley & Sons.

Poon, Ser-Huang & Clive Granger (2003). Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature* 41, 478-539.

Ratner, Bruce (2010). Variable selection methods in regression: Ignorable problem, outing notable solution. *Journal of Targeting, Measurement and Analysis for Marketing* 18, 65-75.

Scott, Louis (1987). Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application. *Journal of Financial and Quantitative Analysis* 22:4, 419-438.

Shephard, Neil (2005). *Stochastic volatility: Selected readings*. Oxford, UK: Oxford University Press. 534 p.

Shephard, Neil & Kevin Sheppard (2010). Realising the Future: Forecasting with High-Frequency-Based Volatility (HEAVY) Models. *Journal of Applied Econometrics* 25, 197-231.

Snedecor, George W. & William G Cochran (1989). *Statistical Methods*. 8th edition. Ames, Iowa, USA: Iowa State University Press. 503 p. ISBN 0-8138-1561-4.

Standard & Poor’s (2013). [cited 17-December-2013] Available from World Wide Web: http://www.standardandpoors.com/about-sp/main/en/us

Taylor, Stephen J. (1982). Financial Returns Modelled by the Product of Two Stochastic Processes – A Study of Daily Sugar Prices, 1961-79. In: *Stochastic volatility: Selected readings*, 60-82. Ed. Shephard, Neil. Oxford, UK: Oxford University Press.