



UNIVERSITY OF GOTHENBURG

SCHOOL OF BUSINESS, ECONOMICS AND LAW

A study on the relation between VIX, S&P500 and the CDX-index

Authors: Alexander Winberg, Niklas Rugås

Supervisor: Alexander Herbertsson

NEG-300 H14 Project Paper with Discussant – Finance (15 ECTS)

Keywords: Implied volatility, Credit default swaps, Credit spreads, Stock index, Correlation, VIX

Acknowledgement

We would like to thank our supervisor Alexander Herbertsson for his expertise and support during this thesis.

Abstract

In this thesis we investigate the relationship between the VIX-index, CDX NA IG and S&P500. Our goal is to study how well the market volatility (traded in VIX) can be explained by stock prices(S&P500) and credit indices (CDX NA IG)

The VIX-index is a measure of implied volatility in the S&P500 and is often referred to as a fear index. CDX.NA.IG is a credit default swap-index consisting of 125 North American investment grade companies and the S&P500 is a stock index consisting of the 500 largest companies in USA.

We use ordinary least square (OLS) regression to study the relationship between our variables and find that the VIX, CDX NA IG and S&P500 have a high correlation.

Table of content

1. Introduction	6
2. Theory	8
2.1 Product	8
2.1.1 Volatility	8
2.1.2 S&P500	9
2.1.3 VIX-index	9
2.1.4 Computing VIX	11
2.1.5 Credit default swaps	13
2.1.6 Credit default swap-indices and the CDX-index	14
2.2 Econometric theory	17
2.2.1 Regression analysis	17
2.2.2 R-squared	19
2.2.3 Adjusted R-squared	19
2.2.4 F-tests	20
2.2.5 Homoscedasticity	20
2.2.6 Newey west standard errors	20
2.2.7 Logarithmic variables	21
2.2.8 Trend	22
2.2.9 Lag in the model	22
3. Methodology	23
3.1 Data	23
3.2 Econometrics	24
3.2.1 Regression analysis	24
4. Empirical findings	25
4.1 Simple regression model	25
4.2 Multiple regression model	30
4.3 Trend	34
4.4 Lag in the model	35
4.5 Predicting lagged movements	38
5. Conclusions	40
References	41

List of figures

1. A historical price chart over VIX	10
2. Structure of a CDS contract	13
3. Scatter plot over VIX and CDX	25
4. Historical price chart of VIX and CDX	26
5. Scatter plot with fitted values against residual (regression 1)	28
6. Historical price chart of VIX and S&P500	30
7. Scatter plot with fitted values against residual (regression 2)	32
8. Scatter plot with fitted values against residual (regression 3)	36
9. Cross-correlogram between VIX and CDX	38
10. Cross-correlogram between VIX and S&P500	39
11. Cross-correlogram between CDX and S&P500	39

List of table

1. Omitted variable description	19
2. Variables	23
3. Output regression 1	27
4. R-squared regression 1	27
5. Output regression of lagged error term (regression 1)	27
6. R-squared regression with lagged error term (regression 1)	28
7. Output Newey-West regression 1	29
8. Correlation-Matrix	30
9. Output regression 2	31
10. R-squared regression 2	31
11. Output regression of lagged error term (regression 2)	32
12. R-squared regression with lagged error term (regression 2)	32
13. Output Newey-West regression 2	33
14. Output regression 3	35
15. R-squared regression 3	35
16. Output regression of lagged error term (regression 3)	36
17. R-squared regression with lagged error term (regression 3)	36

List of major equations

1. Standard deviation	9
2. VIX formula	11
3. 30 days weighted average variance	12
4. Adjusted R-squared	19
5. Newey-West standard error	21
6. Simple regression	24
7. Multiple regression	24
8. Multiple regression with lag	24

1. Introduction

In this thesis we study the correlation between VIX, CDX.NA.IG and S&P500. Our goal is to find variables that can explain changes in the VIX-index. The two variables that we choose to examine as independent variables are the stock prices of S&P500 and the credit default swap-index CDX.NA.IG.

VIX is a measurement of implied volatility in the S&P500. It is constructed by put- and call options of S&P500. The Chicago Board Options Exchange compiles the VIX-index. The credit default swap-index (CDS index) that we choose to study contains of 125 investment grade companies that are included in the S&P500. By studying CDX (below we will for notational convenience denote the CDX.NA.IG by CDX or the CDX-index) we get a sense of the credit risk associated by the companies. In cases when companies perform badly and heading in to economic turmoil the risk increases that they will default on their corporate bonds and the credit spread increases. S&P500 is a stock index consisting of 500 of the largest companies in North America with respect to market capitalization. These companies are the same companies used when determining the VIX-index by finding the so-called implied volatility for S&P500. The 125 companies in CDX are also included in S&P500.

A previous study about the correlation between CDX, VIX and S&P500 is made by Che and Kapadia (2012). They are looking to answer why VIX and market return explain changes in credit spreads. Unlike us Che and Kapadia (2012) also investigates how effectively CDS can be hedged in the equity market. In their paper they found that VIX significantly explains CDS spreads. A credit default swap (CDS) is a credit derivative used to hedge the risk against a credit event such as a default on a bond issued by a company, it can be seen as insurance on credit loss and will be discussed further by us in Subsection 2.1.5.

The study by Che and Kapadia (2012) shows that there is a high correlation between VIX, S&P500 and Credit default swaps. It also demonstrates that the VIX and market returns predict both the root-mean-square error (RMSE) as well as the improvement in hedging effectiveness that occurs over time. This thesis takes a similar approach as Che and Kapadia (2012). Instead we will use regression analysis to study if VIX can be explained by CDS and S&P500.

So instead of explaining CDS by VIX and S&P500, we will try to explain VIX without putting any weight at hedging effectiveness.

We use ordinary least squares-regression (OLS) to test our hypothesis, that VIX, CDX and S&P500 are correlated. The idea is to test two different models, one simple regression model and one multiple model. The simple regression model has CDX as an independent variable and VIX as the dependent variable.

In the multiple regression model, S&P500 will be included as another independent variable. An additional multiple regression model will be considered where a lagged value of VIX is used as an independent variable. We also try to find a lag reaction between VIX, S&P500 and CDX.

The rest of this thesis is organized as follows. In Subsection 2.1 we describe the theory behind the products VIX, CDX and S&P500. Subsection 2.1 also explains the concept of volatility, which is used to compute the VIX-index. CDS is also covered to get a better understanding of the CDS-index, CDX. In Subsection 2.2 we will explain the theory behind the econometric methods we used in our study.

In Section 3 we introduce the methodology, Subsection 3.1 will cover our data behind the analysis and 3.2 show the models that are used. Then in Section 4 we provide the empirical findings and also comments on the results that are found. At last, in Section 5 the conclusion that is found in our study will briefly be discussed.

2. Theory

In this section we introduce the theory behind the products VIX, S&P500 and CDX as well as the methodology used in this thesis. In Subsection 2.1 the products are explained and then in 2.2 we cover the econometric.

2.1 Products

Subsection 2.1 covers the theory behind the different products that is used in the thesis.

2.1.1 Volatility

This subsection will define the concept of volatility and implied volatility. This is done to give a better understanding of the VIX-index that will be explained in Subsection 2.1.3.

Volatility is a measure of how much the asset prices is moving from their mean values and is defined as the standard deviation of the assets data (Herscovic et al. 2014, p.7). Volatility increases during uncertain times on the financial markets and during the financial crisis of 2007-2008 it increased dramatically. One characteristic observed with volatility is that it has a negative correlation with return, this is due to the higher risk associated with high volatility (Ibotsson, 2014). The price movements can be seen as a risk (Ibotsson, 2014). By studying empirical data one gets the historical variance. The markets expectation of the volatility for the underlying asset can be calculated as the implied volatility, by example from an implied Black-Scholes model (Canina and Figlewski, 1993, p.660-661). The Black-Scholes model is an option pricing formula that will value an option. The implied volatility is the volatility in the Black-Scholes model that will generate the current market value of the option (Bodie, Kane and Marcus, 2011, p.637).

To price an option one need to know the volatility from the present day until the option expires (Poon and Granger, 2003, 478). Therefore we need to estimate a forecasted standard deviation. In finance the historical standard deviation is often calculated for a volatility- and risk measurement (Poon and Granger, 2003, p. 480). The sample standard deviation s of the asset return R_t during the period $t=1, \dots, N$ is computed as

$$s = \sqrt{\left(\frac{1}{N-1}\right) \sum_{t=1}^N [R_t - \bar{R}]^2} \quad (1)$$

where \bar{R} is the mean return during the period t=1 until N and R_t is the return at time t.

2.1.2 S&P 500

The Standard & Poor's 500-index, abbreviated S&P500 is a weighted stock index combined of 500 companies from the North American stock market. S&P500 contains the 500 largest companies using market capitalizations, which means the value of outstanding stocks. It is designed to measure the performance of the American equity markets (Bloomberg).

2.1.3 VIX-index

In this subsection we will explain the background of VIX. Subsection 2.1.4 provides the calculation of VIX.

The VIX-index is an index that measures the implied volatility of out-of-the-money put- and call options on the S&P500. The implied volatility in this case is as mentioned in subsection 2.1.1 the volatility in the Black-Scholes model that will generate the corresponding market value of the option. VIX is often referred to as the fear-index since the index reaches higher values in times of economic uncertainty. The Chicago Board Options Exchange introduced it in 1993 and its purpose was to measure the expectation of 30-days volatility implied by at-the-money S&P 100 Index option prices (Chicago Board Options Exchange Technical Notes, 2009). VIX has become a benchmark for the volatility of the US stock market and has an inverse relationship with the stock market. When the market is performing well VIX is decreasing and when the market is performing bad VIX increases (Mitchell, 2014).

In Figure 1 we present historical data for the VIX-index from October 2006 until November 2014. The mean value of VIX was 21.68, max was 80.83 and the minimum value was 9.89.

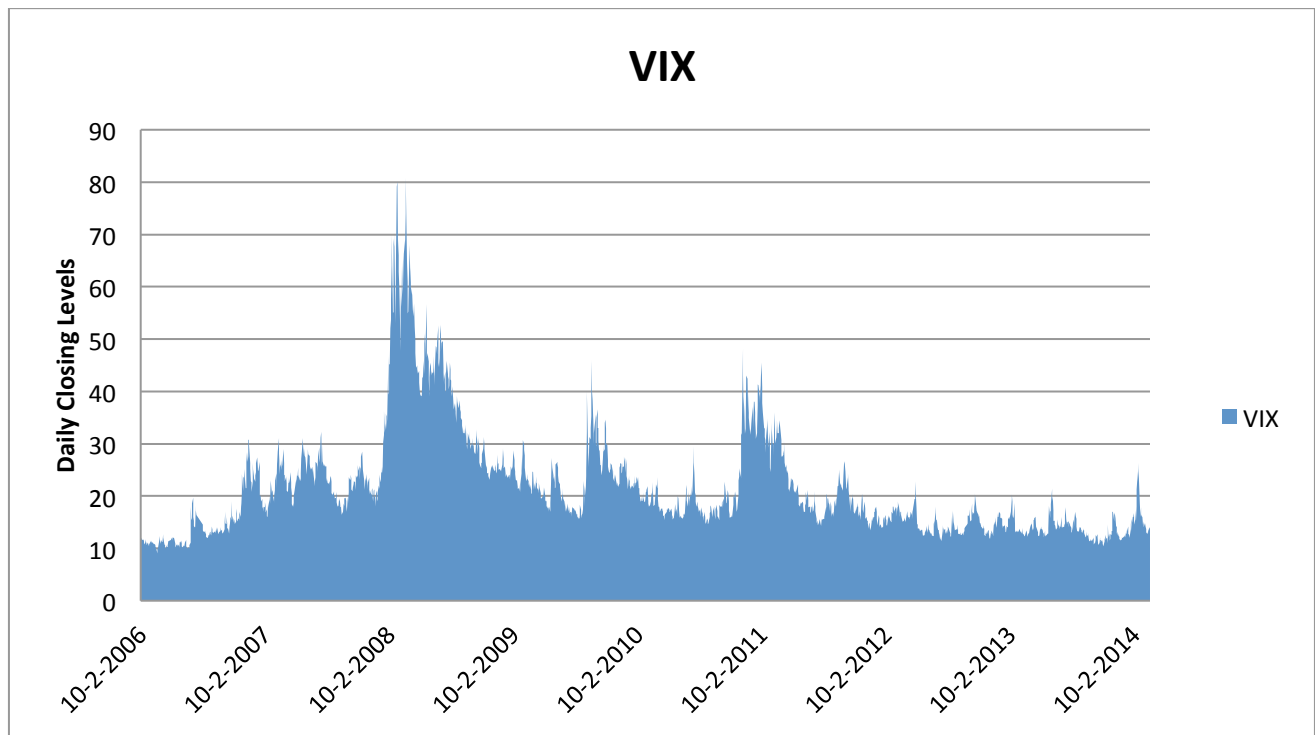


Figure 1. The VIX-index during the period October 2006 to November 2014

VIX is measured as the weighted 30-day implied standard deviation of annual changes in Standard & Poor's 500. For example if the value of VIX is 20, then S&P500 is expected to increase or decrease by 20% over the next year (Williams, 2013). This will be true in 68% of the cases, because standard deviation is in this case assumed to come from a normally distributed random variable, where 68% of the outcomes lies within one standard deviation from its mean, see in Narasimhan (1996). Values of VIX above 30 are often observed in distressed markets while values below 20 are associated with calm periods (Pepitone, 2011).

2.1.4 Computing VIX

Subsection 2.1.4 explains how VIX is computed using options on the S&P500-index. Below we closely follow the notation and outline presented in Chicago Board Options Exchange (Chicago Board Options Exchange Technical Notes, 2009) VIX-index Guide.

Until 2003 VIX was computed from a Black Schools model, since then the VIX-index is calculated by using an equation constructed from the VIX behavior, which we will be presented below. VIX is constructed by the variance of a near and a next-term options. The near options must at least have one week to expiration. If this does not hold, one constructs VIX with the next term (the second contract). Then one uses the second and third contract instead of the first and second, this to avoid using a contract shorter than one week, which violates the models assumptions (Chicago Board Options Exchange Technical Notes, 2009). The first step in calculating VIX is to select which options to use. It should be out-of-the-money puts and calls quoted with non-zero bid prices. The price is determined by the difference between call and put prices. The strike price that has the lowest difference between call and put price is the strike price used in the formula. The next step is to calculate the forward index level derived from index option price for both the near and next-term options. After this one choses the out-of-the-money call and put option. Then the mid-quote price for the put/call averages is calculated for the near and next-term options. We calculate the variance of the near-term and next-term option, using the VIX-formula. In order to compute the VIX value we first need to determine the variance σ^2 from S&P500 options. This is done by using the following formula

$$\sigma^2 = \frac{2}{T} \sum_i \frac{\Delta K_i}{K_i^2} e^{RT} Q(K_i) - \frac{1}{T} \left[\frac{F}{K_0} - 1 \right]^2 \quad (2)$$

where

T = time to expiration

$$T = \frac{M_{\text{current day}} + M_{\text{Settlement day}} + M_{\text{other days}}}{\text{Minutes in a year}}$$

$M_{\text{current day}}$ = minutes reaming until midnight of the current day

$M_{\text{Settlement day}} =$ minutes from midnight until 8.30AM on SPX settlement day

$M_{\text{other days}} =$ total minutes in the days between current day and settlement day

$F =$ Forward index level derived from index option price which is detrmed as Strike price + $e^{RT} * (\text{call price} - \text{Put price})$

$K_0 =$ First strike below the forward index level F

$K(i) =$ Strike price of the i^{th} out – of – the – money option;

a call if $K_i > K_0$ and a put if $K_i < K_0$

$\Delta K_i =$ Intervall between strike price – half the difference between the strike on either side of K_i :

$$\Delta K_i = \frac{K_{i+1} - K_{i-1}}{2}$$

$R =$ Risk – free interest rate to expiration

$Q(K_i) =$ The midpoint of the bid and ask spread for each option with strike K_i

After calculating the variance σ_1^2 and σ_2^2 of the near and next term-option using Equation (2), the next step is to calculate the 30-days weighed average. To do this one use Equation (3) and take the variances calculated by Equation (2).

$$VIX = 100 * \sqrt{T_1 \sigma_1^2 \left(\frac{N_{T_2} - N_{30}}{N_{T_2} - N_{T_1}} \right) + T_2 \sigma_2^2 \left(\frac{N_{30} - N_{T_1}}{N_{T_2} - N_{T_1}} \right) \frac{N_{365}}{N_{30}}} \quad (3)$$

The properties in the equation are as follows:

$\sigma_1^2 =$ Varaiance for near – term, less than 30 days left

$\sigma_2^2 =$ Varaince for next – term, more than 30 days left

$N_{T_1} =$ number of minutes to settlement of the near – term options

N_{T_2} = number of minutes to settlement of the next – term options

N_{30} = number of minutes in 30 days (43,200)

N_{365} = number of minutes in a 365 – day year (525,600)

T_1 = Time to expiration for near – term

T_2 = Time to expiration for next – term

2.1.5 Credit default swaps

This subsection introduces the so called credit default swaps (CDS). We provide subsection 2.1.5 to give a better understanding of the CDS-Index, CDX that will be studied in this thesis.

A credit default swap is a credit derivative used to hedge the risk against a credit event such as a default on a bond issued by a company. More specific, a CDS is a contract between two parties, one that buys a protection for credit loss and one that offers the protection for continuous coupon payments (Markit B, 2014).

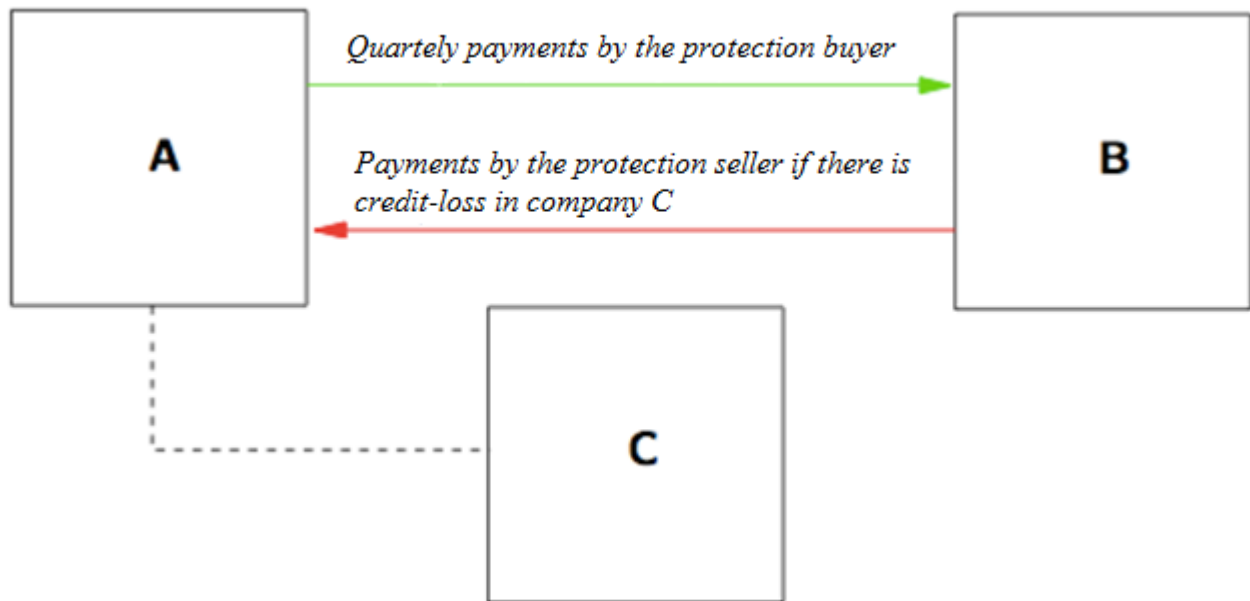


Figure 2 The structure of a credit default swap (CDS) Herbertsson (2010).

Consider company C that has issued bonds, company C might default at a random time. The protection buyer A wants to buy protection on credit losses in C the coming T years from the protection seller B, where T is the time to maturity of the contract.

If there is no default on company C up to time T, protection buyer A pays a fee to the protection seller B for the entire period. This fee is paid on a quarterly basis.

In case of a default on company C up to time T, the protection seller B is obliged to compensate the protection buyer A with the value that has been lost due to the credit event (Markit B, 2014). Hence, the protection buyer A buys an insurance against credit loss on C up to time T and pays a risk premium to the protection seller.

The fee paid from A to B is often called the CDS-spread and is determined so that the expected discounted cash-flows between A and B are equal at the start of the CDS-contract.

2.1.6 Credit default swap-indices and the CDX-index

In this subsection we outline Credit default swap-indices (CDS indices) and also discuss one specific CDS-index, the CDX-index.

A CDS index gives the market an insight of the overall credit quality of the companies in the index and tells us if the companies are performing well or not. Credit indices have expanded dramatically in recent years, with an increase in volumes and decreasing costs of trading. The visibility of the contracts has also grown across the financial markets (Markit B, 2014).

Participants in credit markets have developed standardized indices to track portfolios of credit default swaps. In 2004 there were agreements between different producers of indices that led to some consolidation. (Hull, p524. 2008)

A CDS-index is a generalization of a single-name CDS where the protection buyer A buys insurance against all credit losses in an equally weighted portfolio of m obligors (typically $m \geq 100$), from a protection buyer B, up to time T.

At each default in the portfolio before time T , the protection buyer B pays the corresponding credit losses to A and the nominal size of the portfolio is written down the amount $[N(1/n)]$ where N is the amount to be protected at the start of the CDS-index contract. Just as in the CDS-contract, A pays B a fee up to time T , or until all m obligors in the portfolio have defaulted. The fee is set so that the expected discounted cash flows between A and B are equal at the start of the contract, and this fee is often denoted the CDS-index spread. In a model where all obligors are assumed to have the same characteristics (i.e. a homogeneous portfolio) it is possible to show that the CDS-index spread is equal to the individual CDS-spread (which is the same for all obligors since the portfolio is homogeneous), see e.g. Herbertsson, Jang and Schmidt (2011)

Two important standard portfolios used by index providers are:

1. CDX.NA.IG, a portfolio of 125 investment grade companies in North America
2. iTraxx Europe, a portfolio of 125 investment grade names in Europe (Hull, p524. 2008)

Investment grade refers to the company's credit rating. A company can be considered an investment grade company if it's rated BBB or higher by credit rating firms such as Standard and Poor or Moody's.

These indices roll every six months and then a new series of the index is created with updated constituents. The previous series continues trading, although liquidity is concentrated on the on-the-run series. (Markit B, 2014) When the liquidity list is calculated and other criteria is met a new series is created, it gets a specific number so that each time period easily can be tracked (Markit A, 2013).

The value of debt that is insured in the CDX index is the summation of the individual company's debt. If there is a default in the index, the new value that is insured is given by the remaining 124 companies' debt.

CDX indices are divided into two different groups:

1. The investment grade index IG that we are going to study for this thesis.
2. A high yield index constructed by 100 high yielding companies in the CDS market.

Both indices are owned and managed by Markit North America. (Markit A, 2013)

The CDX.NA.IG index is constructed by credit default swaps on the most liquid 125 investment grade companies', which means that credit risk is a key component of pricing the index.

CDX.NA.IG can be divided into sub-indices and is currently divided into: Sector- and High volatility sub-indices. Sector sub-indices: Consumer Cyclical, Energy, Financials, Industrial and Telecom, Media & Technology. The high volatility sub-index comprises the 30 entities in the IG Index with the widest 5-year Average CDS spreads, the last 90 days before the index is composed. (Markit A, 2013)

The two roll dates for the IG index are September 20 and March 20. Those that have a roll date in September will be issued with the maturity date at December 20 and those from March 20 will have the maturity date at June 20. The time to maturity is either 3, 5, 7 or 10 years.

There is a higher risk associated with a longer maturity because of the risk associated with uncertainty of the future. This makes the risk premiums to differ between different contracts. The contracts on 5 years are the most liquid and the ones that we will study for this thesis. (Markit B, 2014)

At the composition dates when the index is constructed there might be changes in which companies that are included in the index. (Markit A, 2013)

To qualify for inclusion in the IG index the companies have to live up to some general criteria listed below:

- The company cannot be a swap dealer in products related to the IG index.
- The company must have issued a certain amount of publicly traded debt securities.
- The company must have relevant rating, by Moody's or Standard and Poor.

The weighting of the entities in the index will be equally weighted. The weighting of each entity is thus given by: $\frac{1}{\text{number of companies in the index}}$ then rounded to the nearest one-thousandth of a percent (Markit A, 2013). So if the index contains for example 125 companies, then the weights are $\frac{1}{125} = 0,008 = 0,8\%$

The following example is taken from Hull (2008, p.525) and illustrates how a CDS-index works. Let's say a 5 year CDX.NA.IG is quoted by a market maker with bid at 65 basis points, and offer at 66 basis points. Roughly speaking this means that a trader can buy CDS protection on the 125 companies in the index for 66 basis points per company. If a trader wants \$800,000 of protection on each company, the total cost per year is: $[(0.0066 * 800,000) * 125] = \$660,000$

When a company defaults, the protection buyer receives the usual CDS payoff and the annual payments is reduced by $660,000/125 = \$5,280$

As mentioned on page 15 the CDS-index can be seen as the average of the CDS spreads for the 125 companies that constitute the CDX-index.

2.2 Econometric theory

Subsection 2.2 gives a brief introduction to the econometric theory that will be used in this thesis. Below all notations and concepts are taken from Wooldridge (2014) except for Subsection 2.2.8-2.2.9.

2.2.1 Regression analysis

In regression analysis one tries to create a linear model that best explains variation in a dependent variable by studying changes in independent variables.

Simple regression model

In a simple regression model the explanatory variable is explained by one independent variable X_1 . Thus the model is specified as: $y = \beta_0 + \beta_1 x_1 + u$, where β_0 is a constant and the intercept of the model, β_1 is the coefficient of variable x_1 , u is an error term that cannot be explained by any changes in x_1 .

Multiple regression model

In a multiple regression model one or more explanatory variables are added to help explain changes in the dependent variable. Hence the extended model is given by $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$, where β_2 is the coefficient of the second explanatory variable and x_2 is the value of the second variable.

Time series regression model

In this thesis we are observing time series data of VIX, S&P500 and CDX NA IG. This means that the observations are collected over time and given in chronological order. This means that the regression will be extended to include time t so that e.g. $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + u_t$

When one works with time series data in regression analysis there are some additional assumptions that have to be made in order to test for statistical significance in the model

Assumptions

The following assumptions are taken from Wooldridge (2014, p.279-285)

In order to obtain unbiasedness of OLS there are three assumptions that have to be made.

Assumption 1: Linear in parameters: The stochastic process $t = 1, 2, 3 \dots, n$ follows a linear model.

Assumption 2: no perfect collinearity: any of the explanatory variables are constant or perfectly collinear

Assumption 3: Zero conditional mean: The expected value of the error term is zero given any value of the explanatory variables, that is $E(u_t|x) = 0$, for $t = 1, 2, 3 \dots, n$. This is the most critical assumption. When one does a regression and excludes a variable that actually belongs in the model, one is suffering from omitted variable bias. The error term is then correlated with some of the variables. To fix this one can try to find that variable that is excluded and extend the regression model.

If the model is extended, the excluded variable is extracted from the error term and put in the model as an independent variable. But if the model is suffering from omitted variable bias, the model will be estimated wrong. To conclude in which direction the estimation is wrong we use Table 1 taken from Wooldridge (2014, P.78).

	$\text{Corr}(x_1, x_2) > 0$	$\text{Corr}(x_1, x_2) < 0$
$\beta_2 > 0$	Positive Bias	Negative Bias
$\beta_2 < 0$	Negative Bias	Positive Bias

Table 1. Omitted variable bias description

To make the model BLUE which means the best linear unbiased estimator we add two assumptions.

Assumption 4: Homoscedasticity: Conditional on all independent variables, x , the error term have the same variance for every t : $\text{Var}(u_t|x) = \text{Var}(u_t) = \sigma^2$

Assumption 5: No serial correlation: Conditional on x , the errors are uncorrelated with each other in two different time periods: $\text{Corr}(u_t, u_s|x) = 0$ for all $t \neq s$

2.2.2 R-squared

R-Squared is a goodness of fit test and it measure how much of the independent variables in the model help explain the dependent variable as a measure of percent. For example, if R-squared is 0.75 it means that 75% of the changes in the model can be explain by the independent variables.

2.2.3 Adjusted R-squared

In multiple regression models the adjusted R-squared penalizes additional variables that help explain the model. This is done so that we don't overestimate the effects of the additional variables. One calculates this by using the degrees of freedom adjustment when estimating the error variance. In a regression with time series the R-squared is often higher than a regression with cross-sessional data. This is because time series comes in aggregated form, which makes the model easier to explain, it tends to be higher. But a high R-squared can also be calculated because of a trend in the dependent variable. The formula for the adjusted R-squared is given by

$$\bar{R}^2 = 1 - \left[\frac{\hat{\sigma}_u^2}{\hat{\sigma}_y^2} \right] \quad (4)$$

where $\hat{\sigma}_u^2$ is the unbiased estimator of the error variance and $\hat{\sigma}_y^2$ is given by $\text{SST}/(n-1)$ for SST computed as $\text{SST} = \sum_{t=1}^n (y_t - \bar{y})^2$.

When one estimate y_t and it is trending and a time trend is not included in the model, $\hat{\sigma}_y^2$ will not be an unbiased or consistent estimator. With a trending dependent variable and without a time trend included, the estimation of the adjusted R-squared will be wrong according to Wooldridge (2014, p.300). When using time series data and the model is suffering from serial correlation in the error term, the R-squared and the adjusted R-squared will still be a good estimation. This holds as long as the data is stationary (not following a trend) .To see how trend is estimated in a model, see Subsection 2.2.8.

2.2.4 F-test

An F-test is made to conclude how good the variable is explaining the model. It calculates explained variance against the unexplained variance. The explained variance is the variance that can be derived to the variables and the unexplained variance is the variance that appears but not can be explained by the variables.

2.2.5 Homoscedasticity

Homoscedasticity is when the error term has the same variance given any value of the explanatory variables $Var(u|x_1, \dots, x_k) = \sigma^2$. If this does not hold the model is suffering from heteroskedasticity and this will make the calculation of the t-statistic wrong.

2.2.6 Newey-West standard errors

If the data in the given model is both heteroskedastic and has serial correlation we can treat this with Newey-West standard errors that corrects for misleading standard errors and makes the inference tests valid.

Consider the following OLS regression where $y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + u_t$ and where we want to do a valid t-test. Then we want to estimate the standard error for β_1 with consideration to serial correlation and heteroskedasticity, with Newey-west standard errors.

We let $se(\hat{\beta}_1)$ denote the standard error of parameter β_1 given by the OLS regression. Then we let $\hat{\sigma}$ be the standard error of the regression. We construct a linear function of x_{t1} , including the remaining independent variable and an error term $x_{t1} = \delta_0 + \beta_2 x_{t2} + r_t$. The error term r_t has a

zero mean and is uncorrelated with every x in the model. To estimate the true standard error with consideration to serial correlation and heteroskedasticity we use the equation

$$se(\hat{\beta}_1) = \left[\frac{("se(\hat{\beta}_1)")}{\hat{\sigma}} \sqrt{\hat{\nu}} \right]. \quad (5)$$

where “ $se(\hat{\beta}_1)$ ” is the standard error estimated by the OLS regression and $\hat{\nu}$ is an expression of how much serial correlation we allow in the equation and $\hat{\nu} \geq 0$, see e.g. Wooldridge (2014).

2.2.7 Logarithmic variables.

In a model where the variables are greater than zero, a logged version of the variables often generates values that in a regression, better satisfies the assumptions presented in subsection 2.2.1. When working with data containing only positive values, one often has heteroskedasticity, which sometimes can be treated by transforming the variable into logarithmic form. A rule of thumb is to use logged form when working with monetary units or large integer values. If the unit is expressed in percentage one can use both logged form and the original percentage form as long as the values are not between 0-1. By using log one could also fix the problem with large difference between values. For example if one variable has a mean of 2000 and the other has a mean of 20, a 2 unit change is large for the second variable but not the first one and that must be taken in to account.

2.2.8 Trend

If there is a trend in the dependent variable, the model will be affected by time passing. This means that as time is passing the variable is changing. To account for that, one must use time as a variable in the model (Wooldridge, 2014, p.293-294). To search for a trend in the variable you can use Dickey-Fuller test. It takes the true model, on the form $y_t = \beta_0 + y_{t-1} + u_t$ and then adds a variable in the model that will account for a trend, and hence the model will be $y_t = \beta_0 + \beta_1 y_{t-1} + \delta_t + u_t$, where δ_t is the trend variable. In the test the null hypothesis states that the variable is following a time-trend and the alternative hypothesis states that it is not following a time-trend. However this model can suffer from serial correlation, so the Dickey-Fuller test correct for this by adding lags and delta-change (Stata A,2013, p.2).

2.2.9 Lag in model

If previous values of the dependent variable are suspected to affect the current value, a lagged version of the dependent variable has to be added as independent variable. When doing this it corrects for omitted variable bias that will occur otherwise. The model can also suffer from serial correlation and by adding a lagged value it will treat for the serial correlation that might exist (Baum, 2013, p.69). The serial correlation does not affect the prediction of the model, it only affect the standard errors. So we could use another model to fix for serial correlation instead of adding a lagged variable (Achen, 2001, p.5). If the independent variable is trending with the dependent variable, the lagged independent variable will take effect from the dependent variable. It means that the lagged value will “steal effect” from the other variable and make them less important in the model and we might underestimate the effect of these variables (Achen, 2001, p.7). Including the lagged variable when they are heavily trending, makes us to underestimate the other variables (Achen, 2001,p.5). It could also violate the assumptions for the regression and make the dependent variable insignificant or even changing the sign of the coefficient (Achen, 2001, p.20).

To perform a test of lag in the model one can use the lag-order selection of via e.g. Akaike’s information criterion (AIC), final prediction error (FPE), Hannan & Quinn information criterion (HQIC) or the Bayesian information criterion (SBIC) (Stata B,2013, p.2-3).

3. Methodology

This section explains how our models are constructed. The computer software that is used to do the analysis and calculations is Stata. We also cover information of the data that we have used.

3.1 Data

In this subsection we describe the data used in our regressions. Our primary source is Bloomberg. Our indices are:

VIX-Index(VIX)

CDX.NA.IG(YCCI0674)

S&P500(SPX:IND)

Where the names in the parenthesis are the corresponding Bloomberg ticker ID:s.

All indices contain daily sampled data in the period 2006-10-02 until 2014-11-18. Thus VIX, CDX, SP500 has 2031 observations each. In Table 2 the names and a variable description is presented.

Name	Description
Dates	Traded dates between 2006-10-02 - 2014-11-18
VIX	Daily close price, in US dollar
CDX	Daily midprice, in US dollar
SP500	Daily close price, in US dollar
LVIX	Logged form of VIX
LCDX	Logged form of CDX
LSP500	Logged form of SP500
LLVIX	LVIX lagged one day, $LVIX_{-1}$
RES	Residual from regression one
RES2	Residual from regression two
RES3	Residual from regression three
LRES	Residual from regression one, lagged by one day
LRES2	Residual from regression two, lagged by one

	day
LRES3	Residual from regression three, lagged by one day
YHAT	Fitted values from regression one
YHAT2	Fitted values from regression two
YHAT3	Fitted values from regression three

Table 2 A list over the variable used in the regression and their meaning.

3.2 Econometrics

Here we will explain how we use the different econometric models in our work to reach the results in our study. All the variables are constructed as presented in Table 2.

3.2.1 Regression analysis

Simple regression model:

In our simple regression model we use the log of VIX as our dependent variable, which is explained by changes in the logged form CDX. The model is given by

$$LVIX_t = \beta_0 + \beta_1 * LCDX_t + u_t \quad (6)$$

Multiple regression model:

In this model we added S&P500 as another factor that we think might help explain changes in the VIX-index. S&P500 is also expressed in log form

$$LVIX_t = \beta_0 + \beta_1 * LCDX_t + \beta_2 * LS\&P500_t + u_t \quad (7)$$

We also add a lagged dependent variable as an independent variable which gives us

$$LVIX_t = \beta_0 + \beta_1 * LCDX_t + \beta_2 * LS\&P500_t + \beta_3 * LLVIX_t + u_t \quad (8)$$

4. Empirical findings

This section presents the empirical findings of our regressions and related results. We also provide a discussion over the result that is obtained. Subsection 4.1 contains the simple regression model, 4.2 multiple regression model, 4.3 investigates trend in the model, Subsection 4.4 study a lagged variable in the regression model and at last 4.5 contains predictions of lagged movements.

4.1 Simple regression model

In this subsection we will study a simple regression between VIX and CDX by using Equation (6). We will then test the different assumption from subsection 2.2.1.

Figure 3 presents a two-way scatter of VIX and CDX and an estimated linear relationship between them. As seen in Figure 3 there is a linear relationship between VIX and CDX and therefore we will predict a model, using linear regression.

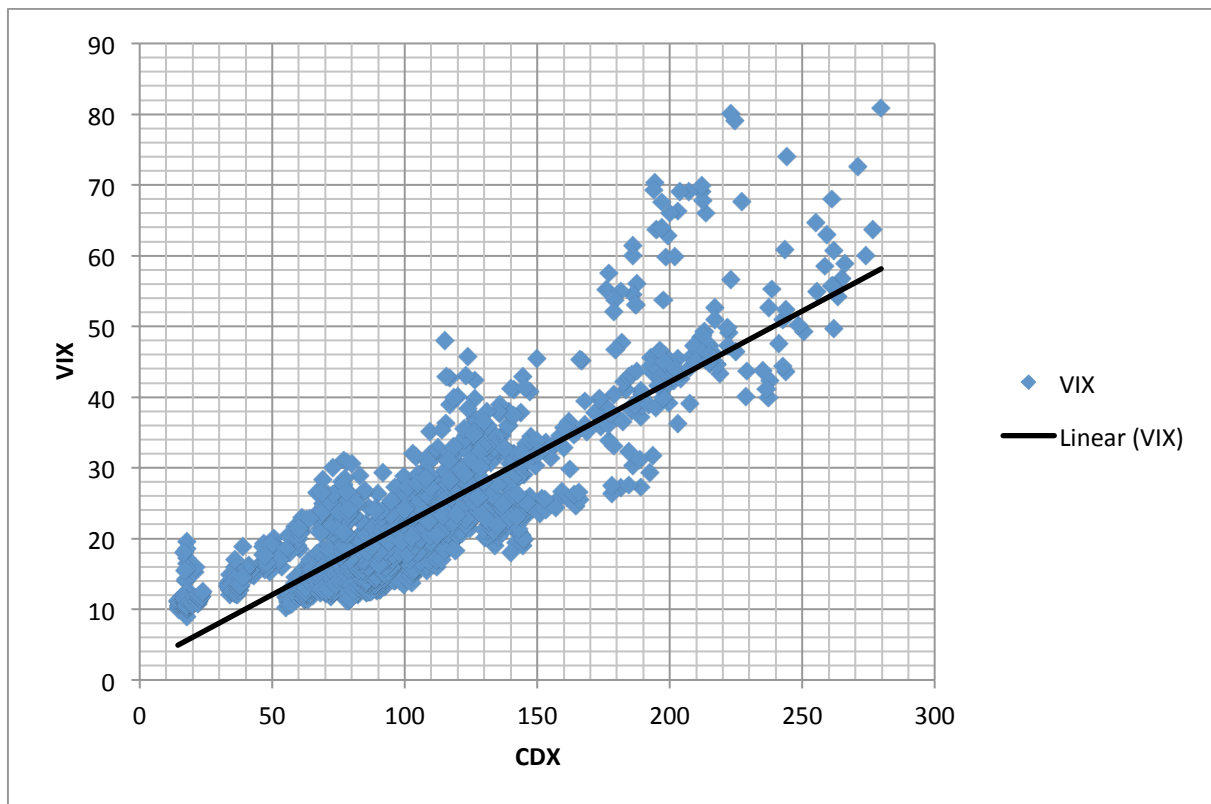


Figure 3. A two-way scatter and a fitted line that estimates a linear relationship between VIX and CDX.

In Figure 4 we plot the time series of VIX and CDX during the period October 2006 to November 2014, which constitute our sample. From Figure 4 one clearly sees that VIX and CDX

are highly positive correlated. We also calculated the correlation between them and got a positive correlation of 0.8560 as can be seen in Table 8 in Subsection 4.2.

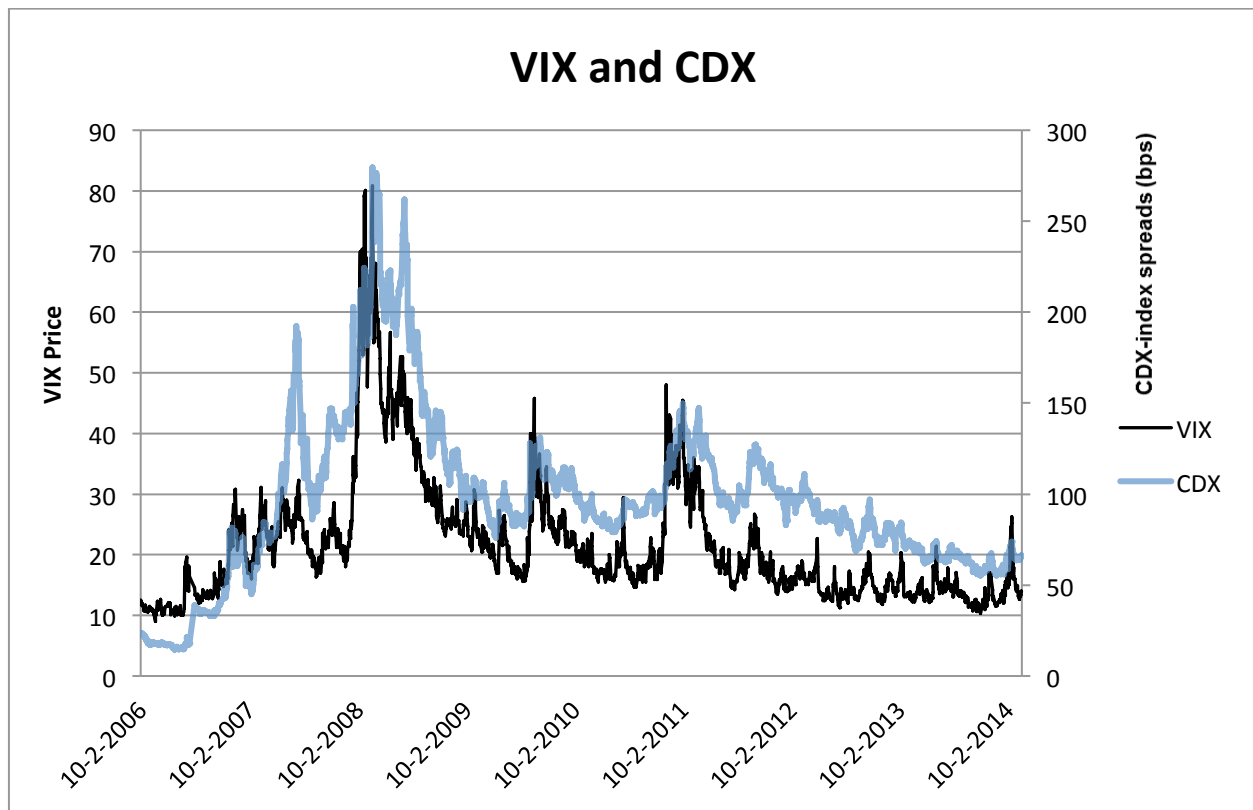


Figure 4. A historical price chart of VIX and CDX where CDX is in bps and VIX is expressed as traded price.

Recall that in this simple regression model we try to explain the implied volatility of the S&P500 by studying the credit spreads of 125 investment grade companies that also contains in S&P500.

The regression with $LVIX_t$ as a dependent variable and $LCDX_t$ as explanatory variable, is given by: $LVIX_t = 0.406 + 0.578 * LCDX_t + u_t$, where 0.406 is a constant and 0.578 is the beta coefficient of CDX. Table 3 shows the results from our regression.

Coefficient	Std.Error	P-value
-------------	-----------	---------

LCDX (β_1)	0,578238	0.0106791	0.000
Constant (β_0)	0.4060904	0.0480231	0.000

Table 3. Results from the simple regression given by Equation (6)

Both the constant and the independent variable are statistically significant at a 99% significance level, because our P-value is 0.000. The models R-squared is 0.5910 and is presented in Table 4.

R-squared	0.5910
Adj R-squared	0.5908

Table 4. R-squared from the simple regression given by Equation (6)

We do a regression of the error term, predicted by its own lag to determine for serial correlation.

The result of this regression is shown in Table 5. The predicted model is:

$RES_t = -0.00024 + 0.97 LRES_t + u_t$. The constant is highly statistically insignificant because a P-value of 0.862, and we would not want to use this because the high probability of estimating wrong. But our independent variable, the lagged residual is statistically significant at a 99% significant level, and we could accept this in our model and therefore say that the lagged residual affects the residual.

	Coefficient	Std.Error	P-value
LRES	0.968573	0.0055092	0.000
Constant	-0.0002424	0.001398	0.862

Table 5. Results from regression of error term with lagged error term.

By looking at Table 6 one sees that the lagged residual almost explains the whole model, with an R-squared of 0.9057. We can say that the lagged residual is affecting the residual and the model is suffering from serial correlation.

R-squared	0.9384
Adj R-squared	0.9384

Table 6. R-squared from regression of error term.

To determine if our predicted model is suffering from heteroskedasticity we plot the residuals with the fitted values, which are displayed in Figure 5. As seen in Figure 5 the spread is increasing and therefore one can say that our model is suffering from heteroskedasticity.

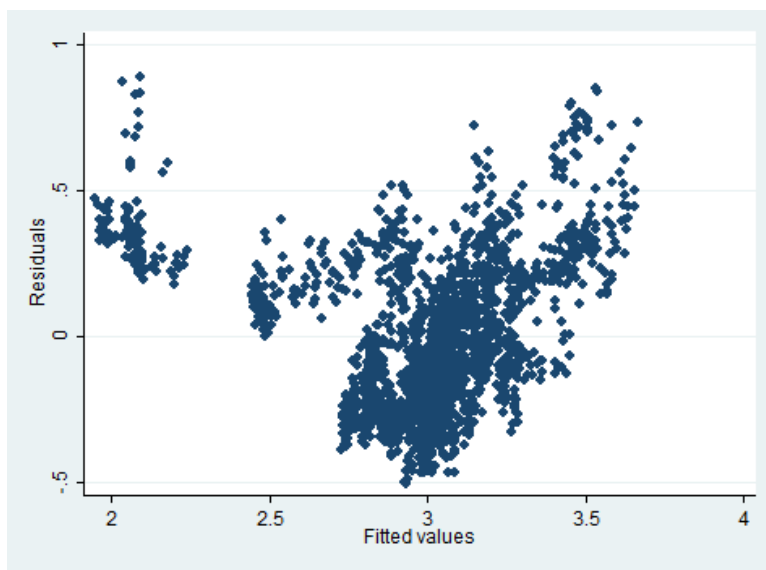


Figure 5. A graph of the fitted values against the residuals values from the simple regression

Since our model is suffering from serial correlation and heteroskedasticity we want to predict a model with this in consideration. Instead of an OLS regression we use a Newey-West test. Our model is still predicted as: $LVIX_t = 0.406 + 0578 LCDX_t + u_t$. The Newey-West test calculates new standard error for each variable with serial correlation and heteroskedasticity in to account. Then it performs new t-test on the variables to see if they are statically significant.

As seen in Table 7 our constant has a higher P-value than before and cannot be accepted, even at a 90% significant level. The independent variable, CDX, still has a P-value of 0.000 and can be accepted at a 99% significant level. The Newey-West test also reports an F-test with p-Value of 0.0000.

	Coefficient	Newey-West Std.Error	P-value
LCDX (β_1)	0.578238	0.1499279	0.000
Constant (β_0)	4060904	0.6833772	0.552

Table 7. Results from the Newey-West regression given by Equation (6)

In the simple regression model the changes in VIX is explained solely by studying changes in CDX. One can see that our model is so far not a good enough estimation. For example our constant is not statistically significant. One can see that CDX is statistically significant and as presented in Figure 4, CDX and VIX are highly correlated. We also think that CDX is economically significant, because of previous studies. A relationship is also seen between them, because they contain performance measurements of the same companies. Therefore we think that CDX belongs in the model and are both statistically and economically significant. The adjusted R-squared of 0, 5908 implies that 59, 08% of the changes in VIX can be explained by changes in CDX. But one should have in mind that R-squared often is high in time series models and we do not want to make all of our conclusions based on the R-squared. When performing a Newey-West test we also get an F-test with a p-value of 0.0000. This indicates that our variables explain the model well.

These results however would imply that CDX has a big influence on the VIX-index even if we are not completely satisfied with our model yet. We have confirmed that are highly positively correlated. It is logic to us that they have a high positive correlation. VIX measures the implied volatility of S&P500 and CDX measures the risk premium of 125 companies that is included in S&P500. So when the volatility of these companies increase the risk premium should also increase.

4.2 Multiple Regression model

Subsection 4.2 presents a multiple regression model. We extend our model with S&P500 as an independent variable and use Equation (7). Then the different assumptions from section 2.2.1 will be tested.

To get a better predicted model we add a variable that might be included in the error term in the previous model. The variable chosen is the index S&P500. As can be seen in Table 8 the correlation between all variables is high and we have a negative correlation between VIX and S&P500 of -0.7013. Therefore we think that S&P500 can help to predict our model better. One can also see in Figure 6 that S&P500 has a highly negative correlation with VIX.

	VIX	CDX	SP500
VIX	1.0000		
CDX	0.8560	1.0000	
SP500	-0.7009	-0.6847	1.0000

Table 8. Correlation-matrix between VIX, CDX and S&P50

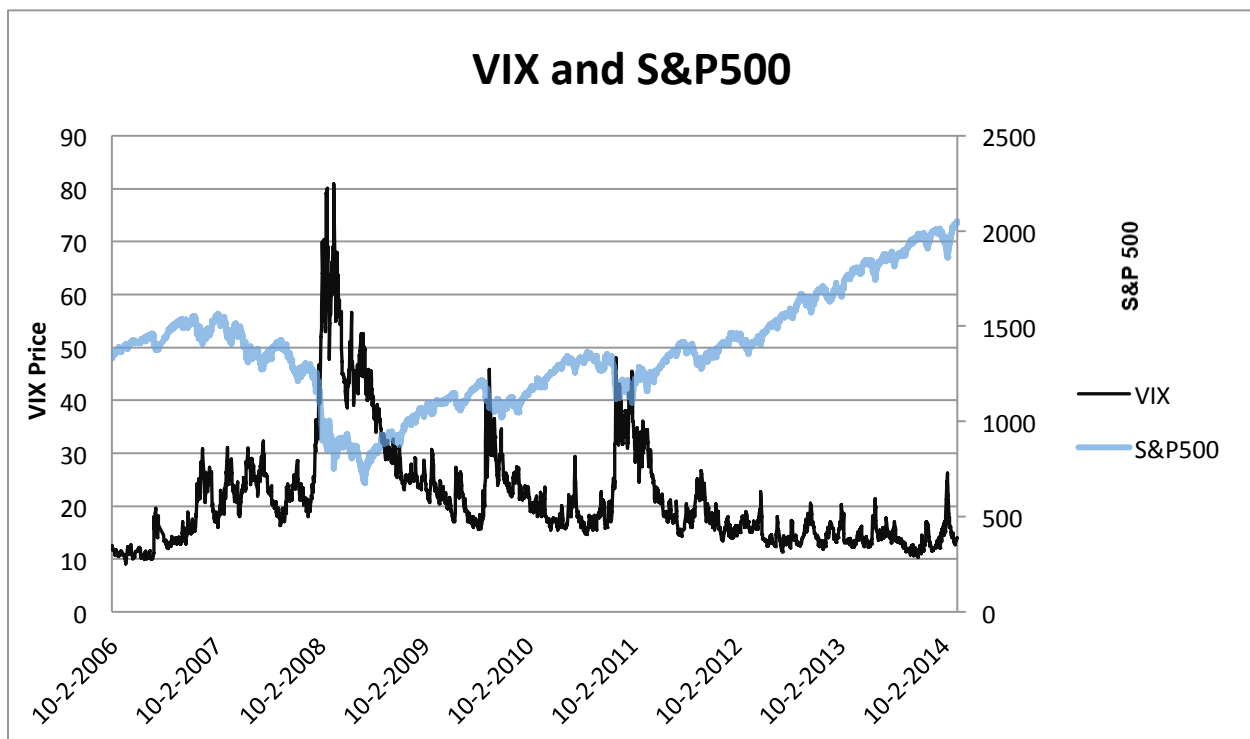


Figure 6. A historical price chart of VIX and S&P500.

Our predicted multiple regression model is: $LVIX_t = 8.151 + 0.353 LCDX_t - 0.936 LSP500_t + u_t$. In Table 9 one sees that all variable is statistically significant to a 99% significant level, because all P-value is 0.000.

	Coefficient	Std.Error	P-value
LCDX (β_1)	0.3528563	0.0098601	0.000
LSP500 (β_2)	-0.9363852	0.023741	0.000
Constant (β_0)	8.151272	0.1996671	0.000

Table 9. Results from our multiple regression given by Equation (7)

We look at the adjusted R-squared, because the R-squared is always increasing when you add another variable. In Table 10, one can see the adjusted R-squared has increased in comparison to the adjusted R-squared observed in the simple regression model, given in Table 4.

R-squared	0.7685
Adj R-squared	0.7683

Table 10. R-squared from the multiple regression given by Equation (7)

We want to examine if the predicted model is suffering from serial correlation and heteroskedasticity. We predict our residuals and then see if our residual can be predicted by its own lag. The model is model is: $RES2_t = -0.000065 + 0.96LRES2_t + u_t$. In Table 11 the output of this regression is shown. One can see, as in the simple regression that the constant is highly statistically insignificant. The independent variable, the lagged residual, on the other hand is highly statistically significant. Therefore one can say that the lagged residual is affecting the residual and we are suffering from serial correlation.

	Coefficient	Std.Error	P-value
LRES2	0.9547243	0.0066153	0.000
Constant	-0.000224	0.0012628	0.959

Table 11. Results from regression of error term with lagged error term.

In Table 12 the R-squared from the regression is presented. An R-squared of 0.9113 tells us that the lagged residual explains about 91.13% of the residual.

R-squared	0.9113
Adj R-squared	0.9112

Table 12. R-squared from the regression with error term.

To examine if our model is suffering from heteroskedasticity, the residual is plotted against the fitted values, predicted by Stata. The result is presented in Figure 7 and in this case it is not as clear as in the previous model to determine whether there is heteroskedasticity or not, we will perform a Newey-West test anyway.

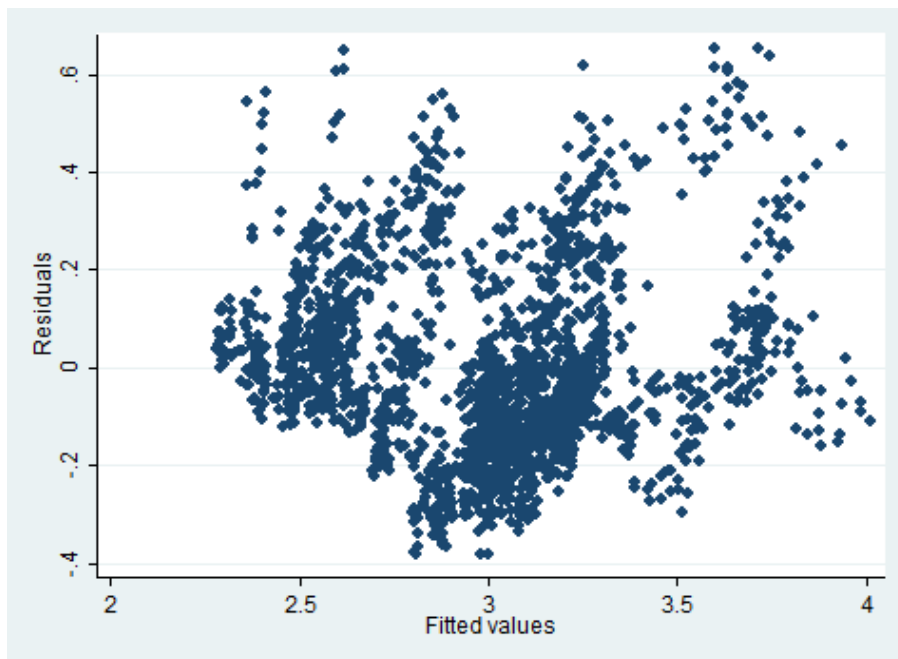


Figure 7. A graph of the fitted values against the residuals values from the multiple regression.

The output of the Newey-West test is given in Table 13. It predicts the new standard errors with consideration to serial correlation and heteroskedasticity. Then it performs a t-test with new standard errors. We can see that all variable is still statistic significant at a 99% significant level,

because the P-value is 0.000. The Newey-West test also report an F-test, with a p-Value of 0.0000.

	Coefficient	Newey-West Std.Error	P-value
LCDX (β_1)	0.3528563	0.0539551	0.000
LS&P500 (β_2)	-0.9363852	0.1438021	0.000
Constant (β_0)	9.151272	1.138365	0.000

Table 13. Results from the Newey-West regression given by Equation (7)

When adding S&P500 we got statistical significance in every variable in the model and an adjusted R-squared of 0.7683. The adjusted R-squared of 0, 7683 however would imply that 76,83% of the VIX is explained by the underlying stock prices and changes in the credit default swap-index. We still do not want to put too much of our analysis on the R-squared.

The F-test of our model is P-value is at 0.0000. This indicates that our variables have a strong explanatory effect on our model.

A difference from our simple regression and multiple regression is that, the constant is significant in the multiple regression model, with the Newey-West test. Therefore we can say that the model with S&P500 is a better estimation. Our variables should still be economically significant because of the theory in previous studies. It is intuitive that S&P500 will affect VIX because VIX measures the implied volatility of S&P500. The correlation is negative because the volatility is low when the market is performing well.

If we treat our variables as Che and Kapadia (2012) and use VIX as a good estimation of the market's volatility and S&P500 as a good estimation of the markets performance, we can say that when the market is performing well S&P500 is increasing and the market volatility will decrease, which means that VIX is decreeing. This is a fairly intuitively conclusion.

As can be seen if we compare our simple regression with our multiple regression, one can see that the coefficient of CDX has decreased, from 0.578 to 0.353. This is as expected. In Table 8

the correlation between CDX and S&P500 is negative and in Table 9 one sees that the coefficient of S&P500 is negative.

By using Table 1 we can make the conclusion that the simple regression will be positive bias and overestimate CDX: s coefficient when leaving S&P500 out of the model.

We want to conclude if our multiple regression is biased as well. Our independent variable S&P500, which roughly can be seen as an estimation of how well the market in USA is performing, will probably be correlated with some macro factors.

We could have tried to find more variables that affects VIX and is correlated with S&P500 and CDX. Because of time limitation and the purpose of our thesis we did not choose to expand our model further. We are aware of that our model could be biased. On the other hand evidence was found that our three variables have explaining power on each other, which were our purpose of our study.

4.3 Trend

A Dickey-Fuller test is performed to search for a trend in our model. The p-value is 0.035. It means that we can reject our Null hypothesis, at a 5% significant level. Therefore the alternative hypothesis that it is not following a time-trend is accepted. Therefore we will not add a time variable in our model. We can from these results also say that our Adjusted R-squared will not suffer from a trend in y_t and be biased as discussed in theory section 2.2.8.

4.4 Lag in the model

In this subsection will we extend our multiple regression with a lagged variable of LVIX as an independent variable. Equation (8) will be used to estimate the model. We will also test for the assumptions discussed in section 2.2.1.

We use the method varsoc in Stata and read the results from the SBIC-test. The best way to estimate our model, according to the test, is to add one lag of LVIX. Our model is:

$$LVIX_t = 0.636 + 0.026LCDX_t - 0.073LSP500_t + 0.923LLVIX_t + u_t$$

	Coefficient	Std.Error	P-value
LCDX (β_1)	0.0260467	0.0046478	0.000
LS&P500 (β_2)	-0.0734177	0.011609	0.000
LLVIX (β_3)	0.9249628	0.008155	0.000
Constant (β_0)	0.6363139	0.0991589	0.000

Table 14. Results from multiple regression given by Equation (8)

The regression output is presented in Table 14. All variables are statistically significant. Our models adjusted R-squared is shown in Table 15 and it is 0.9685.

R-squared	0.9685
Adjusted R-squared	0.9684

Table 15. Reported R-squared from multiple regression given by Equation (8).

We examine as before if our model is suffering from serial correlation and heteroskedasticity. We test for serial correlation by doing a regression of the residual and the lagged residual as a dependent variable. The result is shown in Table 16. The model is:

$$RES3_t = 0.0000131 - 0.082LRES3_t + u_t$$

, one can see that $LRES3$ is statistically significant but the constant is not.

	Coefficient	Std.Error	P-value
--	--------------------	------------------	----------------

LRES3	-0.0815565	0.0221365	0.000
Constant	0.000131	0.0015593	0.993

Table 16. Results from regression of error term with lagged error term.

By looking at the adjusted R-squared in Table 17, we see that it has decreased drastically comparing to the previous model. The adjusted R-squared is now 0.0062

R-squared	0.0067
Adjusted R-squared	0.0062

Table 17. Reported R-squared from regression of error term with lagged error term.

We examine for heteroskedasticity by plotting the residual against its fitted value. The results are shown in Figure 8 and there is no strong evidence of heteroskedasticity anymore. We will not perform a Newey-West on this model. As mentioned in theory subsection 2.2.9, adding a lagged dependent variable will treat for serial correlation and heteroskedasticity

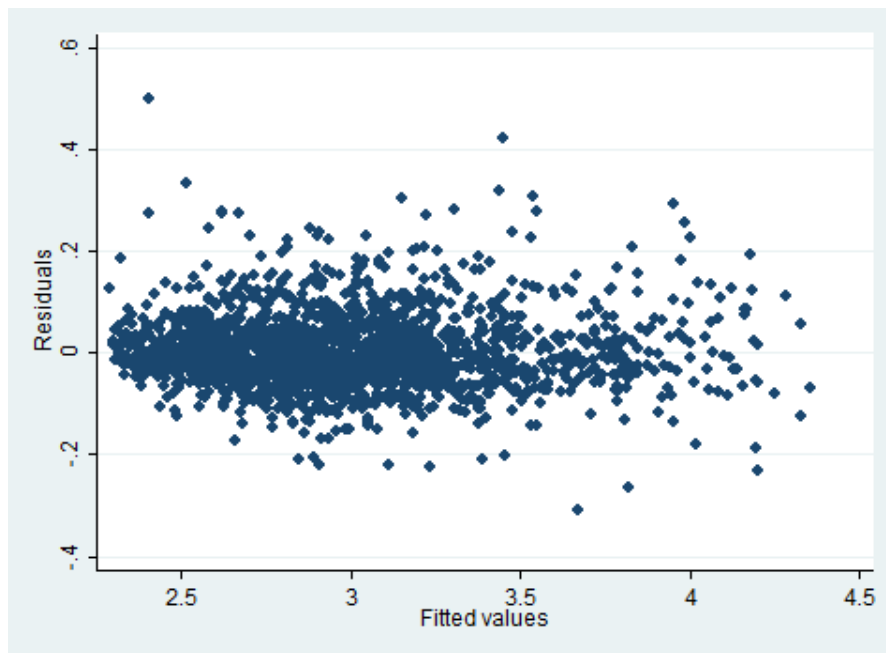


Figure 8. A graph of the fitted values against the residuals values from the multiple regression with a lagged independent variable.

We decided to add a lagged version of our dependent variable as an independent variable in our model, because we were afraid that our model would be biased otherwise. The rule is that if a previous value affects the current value, a lagged dependent variable should be added to avoid omitted variable bias. We might think that if VIX was 20 yesterday it will affect what VIX is today. Therefore a previous value affects the current value. But we could also suspect that if something drastic were to happen, yesterday's value will no longer be relevant. By this conclusion it is not clear if a lagged value of VIX should be in the model to explain VIX.

An advantage of adding a lagged value is that it is treating for the serial correlation and heteroskedasticity that might exist. Something important to say is that serial correlation does not affect the prediction of the model, it will only affect the standard errors. To perform a t-test and see if the variable is significant in the model one needs to estimate the standard errors without serial correlation, this can be done without a lagged variable. Adding a lagged variable to correct for serial correlation is therefore irrelevant and does not help us, instead we could use Newey-West standard errors.

By adding the lagged dependent variable we can see some benefits in our model, but it will also give us some disadvantages that can harm our model. If the independent variables (CDX and S&P500) are trending with the dependent variable (VIX), the lagged VIX will take effect from dependent variables (CDX and S&P500). It means that the lagged value will "steal effect" from the other variable and make us underestimate the effect of these variables so they seem less important. So when adding the lagged VIX it's a chance that we now underestimate CDX and S&P500:s effect in the model. If we instead leave the lagged VIX out and it should be in the model, it will suffer from omitted variable bias. As discussed before, it is not clear if the lagged VIX belongs in the model or not. We get a high adjusted R-squared, but this only says that it is easy to predict VIX with the lagged VIX. It does not tell us if the lagged VIX belongs in the model and explain VIX.

But as disused in theory section 2.2.9 a real problem appears when the added lagged variable is making the independent variable insignificant or change the sign of the coefficient. This does not appear in our model. If we study all three regressions one can say something important.

In all regressions the independent variable CDX and S&P500 is highly significant and has an effect on VIX. From this one can see that they have an effect on VIX and that CDX has a positive effect and S&P500 has a negative effect. Even if the exact coefficient is not estimated correctly we have found a good estimation of how VIX, CDX and S&P500 is correlating and affecting each other.

Since VIX is referred to as a fear-index, we can make the conclusion that the fear on the market today will affect the market-fear of tomorrow. This is because the lagged VIX-variable affects VIX with a high significance.

4.5 Predicting lagged movements

Subsection 4.5 will examine if the variables are reacting to each other immediately or if it is a lagged effect between them.

We have examined if it is a lagged effect between the variables by using a cross-correlogram. In Figure 9 VIX and CDX is shown. Here we can see that the highest value is at lag zero. This means that they respond to each other immediately at daily data. We can't say anything about minutes or hours by using this data. One thing that was detected is that every seventh day a high value was observed in the lags. We have not been able to figure out why we get these findings due to the limitations of time, we do however think that it is an interesting observation.

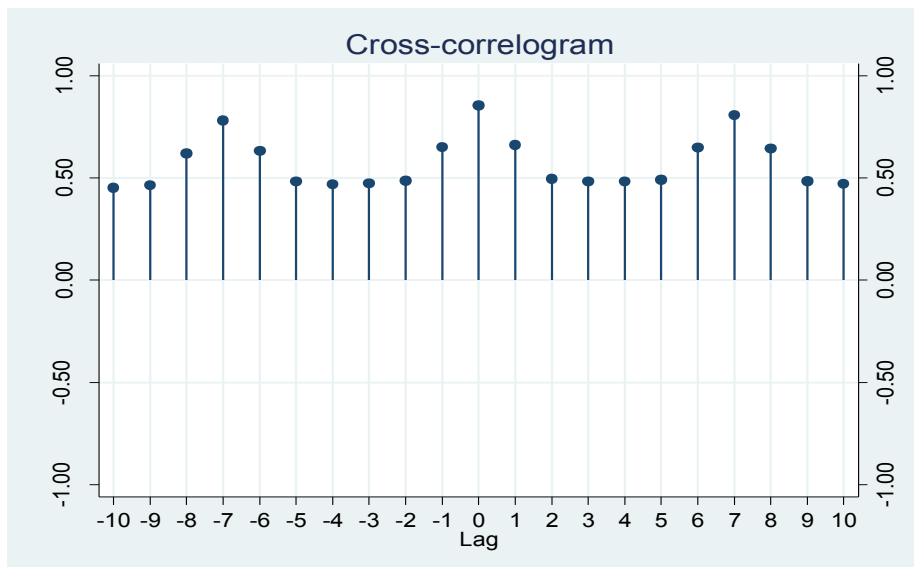


Figure 9. The lag relationship between VIX and CDX, in actual value.

We do the same thing between VIX and S&P500 and get the same result that lag zero is the best prediction. This is seen in Figure 10. In Figure 11 we study the lag relationship between CDX and S&P500. It shows the same results that lag zero gives the strongest prediction.

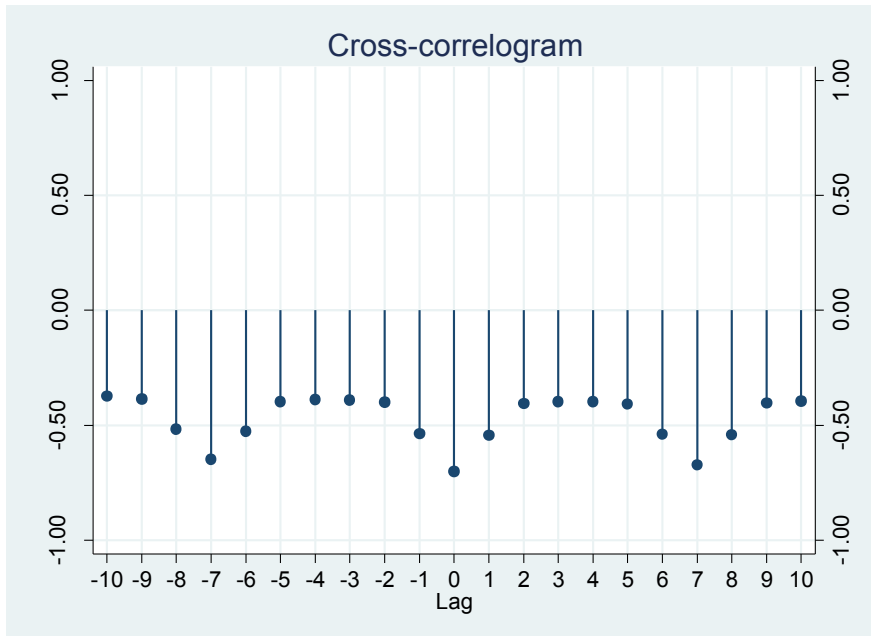


Figure 10. The lag relationship between VIX and S&P500, in actual value.

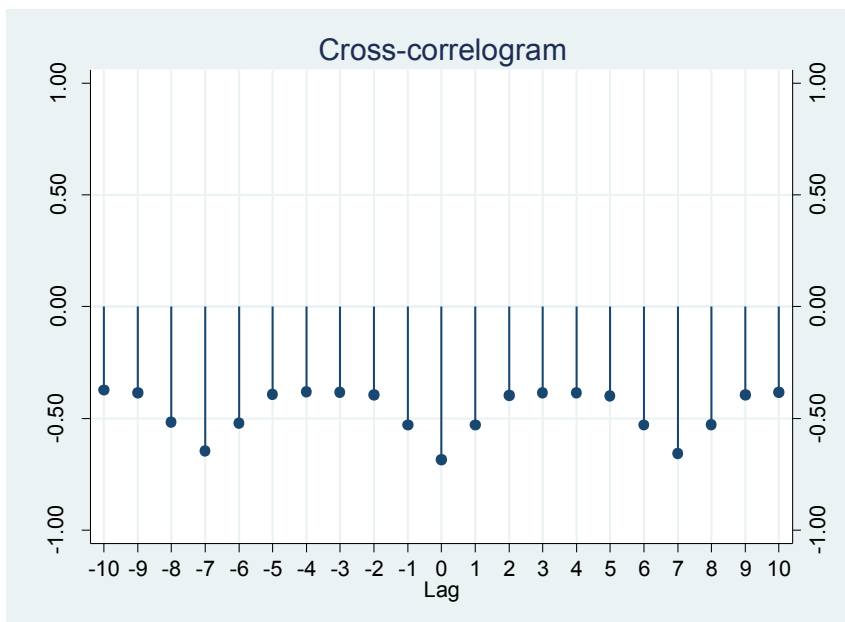


Figure 11. The lag relationship between S&P500 and CDX, in actual value.

As we can see in the Figures 9, 10 and 11 above, all the variables have the strongest correlation at lag 0. This means that the strongest reaction between the variables is immediate. The conclusion that is made from our tests is that we cannot study the current values of our dependent variables to predict the future values of VIX.

If we want to predict a future value of VIX from our model one have to estimate the future values of our independent variables because of the immediately reaction to each other. However our model explains previous values rather than predicting future values. If future values of CDX and S&P500 could be estimated we can predict the future value of VIX by using our model.

5. Conclusions

We have found that in a simple regression model CDX has a high correlation with VIX. The model has an insignificant constant when using a valid t-test, this model is however still a good estimation of how well CDX explains VIX.

When working with an extend model an adding S&P500, we found a strong correlation between VIX, CDX and S&P500. All the variables are statically significant and our model is a better estimation. By study the changes in the beta coefficient of CDX, we determined that the simple regression was suffering from positive bias.

The last regression with a lagged VIX variable as an independent variable gave us a high R-squared, and may be a good estimated model. But this model may be wrong estimated because of trending variables and we might overestimate the effect of the lagged variable.

Our study of a lag between VIX CDX and S&P500 shows that there is no lag and the largest reaction between them is immediately. In our cross-correlograms we noticed that there is a relatively high lag reaction on every seventh day. This study has shown that there is a strong positive correlation between VIX and CDX as well as a strong negative relationship between VIX and S&P500.

References

- Achen, C. (2001). *Why lagged dependent variables can suppress the explanatory power of independent variables*. Paper presented at the annual meeting of the political methodology section of the American political science association, UCLA, 20-22 July 2000.
- Baum, C. (2013). *Time series estimation and forecasting*. University of Mauritius. (Online): http://sites.uom.ac.mu/wtochair/attachments/article/3/MRUS4_BC29.slides.pdf (2014-12-02)
- Bloomberg. *S&P 500 Index*. (Online): <http://www.bloomberg.com/quote/SPX:IND> (2014-11-17)
- Bodie, Z., Kane, A., & Marcus, A. J. (2011). *Investments*. Ninth Edition (International). McGraw-Hill.
- Canina, L., & Figlewski, S. (1993). The Informational Content of Implied Volatility. *Review of Financial Studies*, 6(3), 659-681.
- Chicago Board Options Exchange Technical Notes. (2009). *The CBOE Volatility Index – VIX*. (Online): <http://www.cboe.com/micro/vix/vixwhite.pdf> (2014-11-05)
- Che, X., & Kapadia, N. (2012). *Understanding the Role of VIX in Explaining Movements in Credit Spreads*. Working paper. University of Massachusetts.
- Herbertsson, A. (2010). *Lecture notes on credit derivatives*. University of Gothenburg, School of Business, Economic and Law.
- Herbertsson, A., Jang, J., & Schmidt, T. (2011). Pricing basket default swaps in a tractable shot-noise model. *Statistics & Probability Letters*, 81(8), 1196-1207.
- Herscovic, B., Kelly, B., Lustig, H., & Van Nieuwerburgh, S. (2014). *The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications*. Working paper. NYU, Chicago Booth and UCLA.
- Hull, J. (2009). *Options, Futures and other Derivatives*. Seventh Edition. Prentice Hall.
- Ibottson, R. *Why does volatility Matters*. Yale Insight. (Online): <http://insights.som.yale.edu/insights/why-does-market-volatility-matter> (2014-11-26)
- Markit A. (2013). *Markit CDX High Yield & Markit CDX Investment Grade Index Rules*. (Online): <http://www.markit.com/assets/en/docs/products/data/indices/credit-index-annexes/Markit%20CDX%20HY%20and%20IG%20Rules%20Mar%202013.pdf> (2014-11-14)
- Markit B. (2014). *Markit Credit Indices A Primer*. (Online): <http://www.markit.com/assets/en/docs/products/data/indices/credit-index-annexes/Markit%20Credit%20Indices%20Primer.pdf> (2014-11-09)

Mitchell, C. (2014). *Understanding Volatility and How to Trade the VIX*. TraderHQ. (Online): <http://traderhq.com/understanding-volatility-how-to-trade-vix/>

Narasimhan, B. (1996). *The Normal Distribution*. Stanford edu. (Online): <http://www-stat.stanford.edu/~naras/jsm/NormalDensity/NormalDensity.html> (2014-12-22)

Pepitone, J. (2011). *Runnin' scared: VIX fear gauge spikes 35%*. CNN. (Online): http://money.cnn.com/2011/08/18/markets/VIX_fear_index/ (2015-01-04)

Poon, S., & Granger, C. (2003). Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature*, 41(2), 478-539.

Stata A.(2013). *Dfuller-Augmented Dickey–Fuller unit-root test*. (Online): <http://www.stata.com/manuals13/tsdfuller.pdf> (2014-12-10)

Stata B.(2013). *Varsoc-Obtain lag-order selection statistics for VARs and VECMs*. (Online): <http://www.stata.com/manuals13/tsvarsoc.pdf> (2014-11-29)

Williams, M. (2013). *VIX: What Is It, What Does It Mean, And How To Use It*. Seeking Alpha. (Online): <http://seekingalpha.com/article/1233251-vix-what-is-it-what-does-it-mean-and-how-to-use-it> (2014-11-08)

Wooldridge, J. (2014). *Introductory Econometrics: EMEA adaptation*. First Edition. Cengage Learning.