Is the relationship between the Credit Default Swap Spread and Distance to Default dependent on the state of the economy?

M.Sc. Finance & Investments

Master's Thesis

Supervisor: Søren Agergaard Andersen

Copenhagen Business School: 17-09-2018

Author:

Dennis Dyg Jørgensen Hannesbo

Student number:

32338

Number of characters (incl. spacing, excluding appendix and references): 129,062

Number of pages (excluding appendix and references): 80

1. Abstract

We examine the relationship between the credit default swap spread and the Merton distance to default model, which is based on the Merton bond pricing model (1974). We examine the relationship for the estimated default probability and for the estimated distance to default. For both variables we find, in a first difference regression setting, that we must lag them by one period for them to be significant predictors of the CDS spread. The results of our study indicate that distance to default explains slightly more of the variability in the CDS spread than the probability of default. We test if the sensitivity of the CDS spread to the Merton DD model changes when the economy goes into recession relative to normal times. We examine this by including interaction-terms. For the default probability the results indicate that the change in the relationship is insignificant. However, for the distance to default the results show that the first difference of CDS spread becomes less sensitive to changes in the first difference of distance to default lagged one period when the economy is in recession. Furthermore, we examine if the CDS spread can be explained by S&P500 index and the VIX index and if the relationships depend of the state of the economy. The S&P500 index is a significant predictor of the CDS spread, but as for the PD and DtD the first difference of the S&P500 index must be lagged by one period to be significant. The change in relationship between the S&P500 index and the CDS spread from normal times to crisis is insignificant. The conclusion drawn for the S&P500 index is the same for the VIX index in the univariate setting. However, in the multivariate setting the VIX index is insignificant whereas S&P500 remains a significant predictor of the CDS spread.

Contents

2. Background, introduction, problem statement and delimitation.

2.1 Background

Credit risk has throughout the last decades caught the attention of both researchers and practitioners. The financial crisis raved during the period of 2007-2009, a period where the importance of controlling risk was exposed. During the crisis, large and well-established firms experienced large drops in equity value, and banks suffered enormous losses on loans given to private firms. The significant losses during the crisis may indicate that the lenders had a difficult time predicting the default risk of the borrowing firm. However, the explanation can also be that the models that may work fine during normal times do a poor job when the economy is in crisis.

Credit risk modeling is complex and by nature quantitative, and there are numerous models out there that try to price and predict default risk. It is beyond the scope of the thesis to cover them all.

The thesis investigates how the Merton DD model explains a market measure of credit risk. The market measure of default could be the corporate bond yield spread, defined as the bond yield minus the risk-free rate. This is the premium that bondholders receive for taking on the risk the bond-seller default on its obligation. But is has been shown in studies that a part of the bond yield spread is not due to credit risk but is explained by liquidity factors (Chen, Lesmond, & Wei, 2007) and (Longstaff, Mithal, & Neis, 2005). Another market measure for credit risk is the credit default swap spread (CDS spread), which is considered to be a more pure measure of default risk (Zhang, Zhou, & Zhu, 2009).

The CDS spread can be considered the market's price for taking on the default risk for the entity firm, hence the first place to look for an explanation for the movement in the CDS spread will be in the movements in the default risk of the company.

2.2 Introduction

The structural approach has been used extensively in credit risk modeling, the approach originates out of the arbitrage pricing framework by Black & Scholes (1973); and (Merton, 1974).The structural approach will be the theoretical foundation of the thesis. A specific structural model that has won great attention both in practice and in academic research is the one developed by the KMV Corporation (Crosbie & Bohn, 2003), which is a practical application of the Merton model (1974). The thesis follows the study of Bharath & Shumway (2008) and refer to the KMV model as the Merton distance to default model or simply the Merton DD model.

We will investigate how well movements in the Merton DD model measure of credit risk explains the movements in the CDS spread, and if the potential relationship changes for an economy in crisis relative to normal times.

The Merton DD model applies the setting of Merton (1974), where the equity of a firm is viewed as a call option on the asset value of the specific firm, where the exercise price equals the face value of the firm's debt. The market value of the assets and the asset volatility are not directly observable. But the assumptions of the model make it possible to infer the values from the equity value using an iterative procedure. Having these values, the Merton DD model produces a measure for default risk, the Distance to Default (DtD).

Although the structural approach is widely used in credit risk modeling, several studies find that the structural models do a poor job in explaining the magnitude of credit spreads, a result often referred to as the *credit spread puzzle* (Augustin, Subrahmanyam, Tang, & Wang, 2014). Numerous studies examine the credit spread puzzle based on the structural models using bond spreads, for evidence of the *credit spread puzzle*, see for example Eom, Helwege, & Huang (2004b); and Huang & Huang (2012).

In line with the *credit spread puzzle* other studies test the structural models using CDS spreads, where Huang & Zhou (2008) conclude that the structural models fail to predict the CDS spreads and capture the time-series changes accurately. On the other hand, a study by Ericsson, Jacobs, & Oviedo (2010), concludes that covariates inspired by the structural models, such as leverage and volatility in a linear regression explain a great fraction of the variation in the CDS spread. In the study by Bharath & Shumway (2008): "Forecasting Default with the Merton Distance to Default Model" the default probability is calculated from the Merton DD model and they find that the measure is insufficient in predicting the CDS spreads.

Numerous studies have been carried out testing the Merton DD model's ability to predict default using CDS spreads and especially corporate bond yield spread. The results of these studies have pointed in different directions, however with the tendency to conclude that the Merton DD model is insufficient when it comes to predicting the CDS spread and/or the bond spread.

What have not been tested extensively, to our knowledge, is if the potential relationship between the CDS spread and the estimate from the Merton DD model changes as a function of the state of the economy. This thesis setoff to test this but will in addition tests if the S&P 500 index level and the volatility measure from the VIX index have explanatory power on the CDS spread, and if the relationship changes when the economy is in crisis.

2.3 Problem statement

This thesis addresses the following problem statement:

"Does the Merton DD model estimate explain the variation in the CDS spread in a linear regression framework, and is relationship between Merton DD model estimate and CDS spread dependent on the state of the economy? Does the explanatory power of the model increase by adding macro variables and does the relationship between the CDS spread and the macro variables depend on the state of the economy?"

This thesis answers the problem statement through the following questions:

1. How well does the Merton DD model estimate explain the variability in CDS spread in an OLS regression setup?
2. Does the potential relationship between the Merton DD model and the CDS spread depend on the state of the economy?
3. Is it possible to improve the regression model by including non-firm-specific variables such as the S&P 500 index and the volatility measure from the VIX index?
4. Is the potential relationship between the non-firm-specific variables and the CDS spread dependent on the state of the economy?

2.4 Delimitations

Assessing the performance of Merton DD model is important because many practitioners and academics apply the model. In addition, it is relevant to know if the model performance depends on the state of the economy. Questions 1-4 in the problem statement are answered by empirically testing how the risk measure from the Merton DD model, The S&P500 index and the VIX index predict the CDS spread. The analysis is done for a single American firm for a period of 6 years, from 2007 to 2013. The conclusions drawn from the analysis will only be valid for the single firm investigated.

If one wants to make more general statements about the relationship between the CDS spread and probability of default, distance to default, the S&P500 index, and the VIX index, then it would be necessary to increase the sample size significantly.

The empirical test carried out in the thesis can be relevant for analysts who does analysis on single firm level. In section: “9. Data collection”, we go into why an American firm has been chosen, the data frequency and the exact period of study.

3. Structure of thesis

In this section we present the structure of the thesis. It serves the purpose to prepare the reader for what to come and make it clear why each section is covered.

In section 4 we put forth theoretical foundation and its application. We present the Merton model (1974), that serves as the theoretical foundation of the thesis. Hereafter follows the practical application of the model, first with the KMV approach to default point. But the KMV approach still leaves us with the challenge of estimating the asset value and its volatility which we need to calculate the distance to default and the probability of default. In the subsection, “The VX-algorithm”, we put forth the method of Vassalou & Xing (2004) so we can estimate the asset value and asset volatility. We present their iterative procedure, which we call the VX-algorithm, it shows how we solve the two nonlinear equations for the two unknowns, asset value and asset volatility.

In section 5 we cover the credit market, where we show how the CDS spread can be calculated from default probabilities.

In section 6, the statistical method is put forth and we show how we make use of interaction-terms to test if the relationship between CDS spread and for example DtD changes when the economy is in crisis relative to normal times.

In section 7, we shortly present the frim Alcoa Inc. which is used to answer the problem statement

In section 8, we go over the data collection process. In section 9, we give a summary statistic of the variables, and do a plot analysis of all variables included in the study. In section 10, we run the regressions that allow us to answer the problem statement. In section 11 it is discussed how the model could have been improved and in section 12 we conclude.

4. Theoretical foundation and application

4.1 Merton model

First it is assumed that we operate within the settings of the standard Black-Scholes model (1973), hence we are in a market which trades continuous in time and the market is frictionless and competitive, that is

1. Agents are price takers, hence they do not have any effect on the prices on the assets traded in the market.
2. No transaction costs on traded assets.
3. No restrictions on short selling.
4. No indivisibilities of assets.
5. Lending and borrowing at the at risk-free rate are done at the same rate, and the risk-free rate is continuously compounded.

(Lando, 2004)

In the Merton model (1974), a firm's equity value is viewed as a call option on the firm's assets. To understand why, one must recognize that equity is a residual claim on the firm's assets after all other obligations have been met. The strike price of the call option is then the book value of the firm's liabilities.

It is assumed that the dynamics of the assets value, V, follows a Geometric Brownian Motion (GBM):

$$dVt = \mu V_t \, dt + \sigma V_t \, dW_t$$

Where μ is the expected continously compoinded return on the asset value, σ is its volatility, and W is a standard Brownian motion under the physical probability measure P.

It then follows that the asset value at time t can be written as

$$V_t = V_0 e^{(\mu - 0.5\sigma^2)t + \sigma W_t}$$

Where W_t is normally distributed with mean 0 and variance t, and the increments of W are independent, that is $W_{t(1)} - W_{t(0)}$ is independendetn of $W_{t(3)} - W_{t(2)}$ for $t_0 < t_1 < t_2 < t_3$ and so on.

Taking the logarithm of the underlying assets value

$$V_t = V_0 e^{(\mu - 0.5 * \sigma^2)t + \sigma W_t}$$

$$logV_t = \log\left[V_0 e^{(\mu - 0.5 * \sigma^2)t + \sigma W_t}\right]$$

$$logV_t - logV_0 = (\mu - 0.5 * \sigma^2)t + \sigma W_t$$

This means that the logarithm of the price increment is normal distributed with mean $(\mu - 0.5\sigma^2)t$ and variance $\sigma^2 t$.

It is also assumed that there exists a money market account with a constant risk-free rate, r, that is deterministic:

$$\beta_t = e^{r*t}$$

The price of a contingent claim paying $C(V_T)$ at time T is equal to

$$C_0 = E^Q(e^{-r*t} C_T)$$

Where Q is the equivalent martingale measure under which the dynamics of V are given as

$$V_t = V_0 e^{(r - 0.5*\sigma^2)t + \sigma W_t^Q}$$

And W_t^Q is a Brownian motion, and μ is replaced by r.

In general, the market value of a firm's assets are far from observable, and a critical assumption is that the asset value process is given and does not change by financing decisions made by the firm's owners.

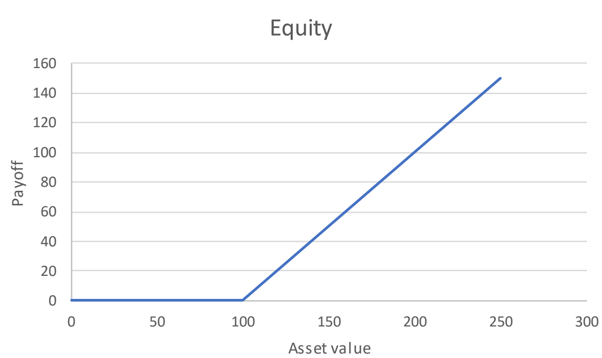
Now assume at time $t = 0$, that the firm issues two types of claims: equity and debt. Furthermore, assume that the debt has face value of D and is of the type zero coupon bond. Armed with these assumptions, the payoff at time T to equity holders, E_T, and debt holders, B_T are:

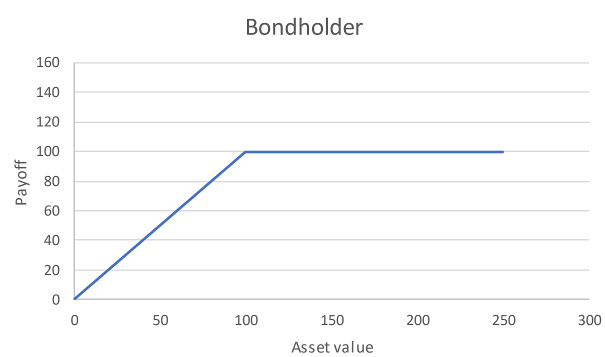
$$E_T = \mathrm{Max}(V_T - D, 0) = \begin{cases} V_T - D, & for\ V_T > D \\ 0, & for\ V_T < D \end{cases} \quad (1)$$

$$B_T = \min(V_T, D) = \mathrm{D} - \max(D - V_T, 0) = \begin{cases} D, & for\ V_T > D \\ V_T, & for\ V_T < D \end{cases} \quad (2)$$

(Lando, 2004)

The graph below illustrates the payoff to equity and bondholders for D=100 at maturity of the debt at time T as a function of asset value.





The firm is run by the equity owners, hence they control the assets. At maturity of the debt the equity owners pay the debt holders the face value of the bond, D, if the assets value exceeds the face value of the bond. The equity owners do that to keep the ownership over the assets. If the assets value is below the face value of the debt at maturity the equity owners do not want to pay D to keep the assets. Hence, when $V < D$ is the case at maturity then bond holders overtake the ownership of the assets, and recover V_T instead of the D.

From the payoff structure it clear that equity can be viewed as a call option on the firm's assets and debt can be viewed as a riskless bond and short position in a put option on the firm's assets.

We now make use of the Black-Scholes pricing machinery for an European call option, given the current asset level V and it's volatility σ, the risk-free rate r, and strike price D and maturity T, the call price is

$$C(V,D,T,\sigma,r) = VN(d_1) - De^{-rT}N(d_2)$$

Where

$$d_1 = \frac{\log\left(\frac{V}{D}\right)+(r+0.5*\sigma^2)T}{\sigma\sqrt{T}},$$

$$d_2 = \frac{\log\left(\frac{V}{D}\right)+(r-0.5*\sigma^2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

The function $N(\cdot)$ is the cumulative distribution function of the normal distribution.

We now apply the put-call parity for European options on non-divided paying stock to derive price of the put option

$$C(V_T) - P(V_T) = V_T - De^{-rT}$$

$$P(V_T) = De^{-rT} - C(V_T) - V_T$$

Applying this, we obtain the Merton model for risky debt, the value of equity and debt at time t becomes

$$E_t = C(V,D,T-t,\sigma,r)$$

$$B_t = De^{-r(T-t)} - P(V,D,T-t,\sigma,r)$$

For an European call option $N(d_2)$ is defined as the risk neutral probability that the option finish in the money. To see why it is noted that the dynamics of V are given as

$$V_t = V_0 e^{(r-0.5*\sigma^2)t+\sigma W_t^Q}$$

where Q is the equivalent martingale measure, to calculate $P(D < V_T)$ is now straight forward:

$$P(V_T < D) = P(V_0 e^{(r-0.5*\sigma^2)T+\sigma W_T} < D)$$

$$P(V_T < D) = P\left((r-0.5*\sigma^2)T + \sigma W_T < \log\left(\frac{D}{V_0}\right)\right)$$

$$P(V_T < D) = P\left(\sigma W_T < \log\left(\frac{D}{V_0}\right) - (r-0.5*\sigma^2)T\right)$$

$$P(V_T < D) = P\left(\frac{\sigma W_T}{\sigma\sqrt{T}} < \frac{\log\left(\frac{D}{V_0}\right) - (r-0.5*\sigma^2)T)}{\sigma\sqrt{T}}\right)$$

$$P(V_T < D) = N\left(\frac{\log\left(\frac{D}{V_0}\right) - (r-0.5*\sigma^2)T)}{\sigma\sqrt{T}}\right)$$

$$P(D < V_T) = 1 - N\left(\frac{\log\left(\frac{D}{V_0}\right) - (r - 0.5 * \sigma^2)T)}{\sigma\sqrt{T}}\right)$$

$$P(D < V_T) = N\left(-\frac{\log\left(\frac{D}{V_0}\right) - (r - 0.5 * \sigma^2)T)}{\sigma\sqrt{T}}\right)$$

$$P(D < V_T) = N\left(\frac{\log\left(\frac{V_0}{D}\right) + (r - 0.5 * \sigma^2)T)}{\sigma\sqrt{T}}\right) = N(\mathrm{d}_2)$$

In Merton model setting this is equivalent to the risk neutral probability that the firm does not default. But from risk neutral probability that the firm does not default the probability of default is obtained as $1 - N(d_2) = N(-d_2)$, because the normal distribution is symmetric. Hence $P(V_T < D) = N(-d_2)$.

To calculate the risk neutral probability of default we need V_0 and σ, neither are directly observable in the market. But if the firm is publicly traded then E_t is observable, and we have the following equation:

$$E_t = C(V, D, T - t, \sigma, r)$$

Which provides one condition that must be satisfied by V_0 and σ.

From historical data or options, it is possible to estimate the volatility of the equity, σ_E . We now make use Itô's lemma to set up another equation that must be satisfied by V_0 and σ we have:

$$\sigma_E E_0 = \frac{\delta E}{\delta V}\sigma V_0$$

Or

$$\sigma_E E_0 = N(d_1)\sigma V_0$$

We now have two non-linear equations in two unknows, the non-linearity means that there is no closed form solution, but it is possible to determine V and σ by solving the equations numerically using an iterative procedure (Hull, 2012).

We now go over the properties of the Merton model:

The sum of equity and debt is equal to the asset value at any point in time, that is, we can write $B_t = V_t - C(V, D, T - t, \sigma, r)$, which makes it easier to see how debt value depends on movements in the parameters V, D, T, σ, and r.

B_t increase in V, it is seen from the fact that the put option decrease in value when V increase, all else equal.

B_t is increasing in D, increasing the face value of debt equals a lager payoff in the state $V_T > D$, but also seen from the fact that the call option in $B_t = V_t - C(V, D, T - t, \sigma, r)$ decrease when D increase.

B_t is decreasing r. It is easiest seen from rearranging $B_t = V_t - C(V, D, T - t, \sigma, r)$ to $B_t + C(V, D, T - t, \sigma, r) = V_t$ the call increase in r, the sum of equity and debt must remain the same. For this to be the case B_t must decrease.

B_t is decreasing in T. The increasing discounting of the risk-free bond has the dominant effect here (Lando, 2004).

B_t is decreasing in σ. It is easy to see why, both the put and the call options increase in value. The bond holder has a short position in the put, hence debt value decreases. And in the call option increase in value it is also easily seen from $B_t = V_t - C(V, D, T - t, \sigma, r)$. The fact, that both the put and the call increase simultaneously in volatility is the key to understand "asset substitution". Increasing asset volatility at time t=0, can be done by selling the assets and reinvest the sum in more risky assets, the value of the assets remains the same. It will, on the other hand, move the wealth from equity holders to bond holders. In practice the bond holders meet the risk of wealth transfer by adding covenants to the bonds, so that the bondholders exercise some control over the investment decisions. In the Merton model, the control is incorporated by the fact that the volatility is constant over the period studied (Lando, 2004).

Shortcomings of the Merton model

1. The model only takes one kind of debt, the zero-coupon bond into account. In the real world it is often observed that debt take on different characteristics.
2. The model assumes constant volatility.
3. The default time restriction, it is assumed that default only can occur if the firm does not manage to repay the principal of the one-year zero-coupon bond at maturity. In practice default can happen any time the firm fails to meet its obligations, it could for example be a short time liability to one of its suppliers.

It is rare that a model achieves everything one wants, and the Merton model is no exception. For this thesis, the Merton model works as the theoretical foundation and clarify our conceptual thinking.

4.2 KMV Approach

The KMV approach (Crosbie & Bohn, 2003) offers a practical use of the methodology for prediction of default. The approach assumes that the asset value evolves as

$$V_t = V_0 e^{(\mu - 0.5 * \sigma^2)t + \sigma W_t}$$

Which makes it possible to calculate the probability that the asset value falls below the face value of the debt at maturity under physical probability measure P:

$$P(V_T < D) = P(V_0 e^{(\mu - 0.5*\sigma^2)T + \sigma W_T} < D)$$

$$P(V_T < D) = P\left((\mu - 0.5*\sigma^2)T + \sigma W_T < \log\left(\frac{D}{V_0}\right)\right)$$

$$P(V_T < D) = P\left(\sigma W_T < \log\left(\frac{D}{V_0}\right) - (\mu - 0.5*\sigma^2)T\right)$$

At this point it is noted that if a random variable, X, is normally distributed with mean μ and variance σ^2 then $\frac{X-\mu}{\sigma}$ is a standard normal variable, i.e. with mean 0 and variance 1. We also remember that W_T is normally distributed with mean 0 and variance T.

$$P(V_T < D) = P\left(\sigma W_T < \log\left(\frac{D}{V_0}\right) - (\mu - 0.5*\sigma^2)T\right)$$

$$P(V_T < D) = P\left(\frac{\sigma W_T}{\sigma\sqrt{T}} < \frac{\log\left(\frac{D}{V_0}\right) - (\mu - 0.5*\sigma^2)T}{\sigma\sqrt{T}}\right)$$

$$P(V_T < D) = N\left(\frac{\log\left(\frac{D}{V_0}\right) - (\mu - 0.5*\sigma^2)T}{\sigma\sqrt{T}}\right)$$

$$P(V_T < D) = N\left(-\frac{\log\left(\frac{V_0}{D}\right) + (\mu - 0.5*\sigma^2)T}{\sigma\sqrt{T}}\right)$$

The function $N(\cdot)$ is the cumulative distribution function of a normal distribution. In the literature distance to default (DtD) is often used instead of the probability of default. DtD measures the number of standard deviations the expected asset value is away from default. DtD is given as:

$$DtD = \frac{E[\log(V_T)] - \log(D)}{\sigma\sqrt{T}} = \frac{\log(V_0) + (\mu - 0.5*\sigma^2)T - \log(D)}{\sigma\sqrt{T}} = \frac{\log(\frac{V_0}{D}) + (\mu - 0.5*\sigma^2)T}{\sigma\sqrt{T}}$$

In the Merton model, the DtD estimate can be negative, but for $V_T = D$ it is 0. The probability of default and distance to default contain the same information and the relationship is clear:

$$P(V_T < D) = 1 - N(DtD) = N(-DtD)$$

DtD is simply another way of stating the probability of default.

Where $N(\cdot)$ is the cumulative distribution function of the normal distribution.

At this point it seems straight forward to calculate both DtD and the probability of default, but it is not. For a typical firm the market value of assets, let along the asset volatility, are not directly observable (Lando, 2004). What is observable is the book value of the assets, but often it will make no sense to use it as a proxy for the market value because it can deviate from its significantly.

If we cannot observe the market value of asset then we cannot calculate distance to default, in addition we cannot use observed asset values to calculate an estimate of the asset volatility σ.

In the fowling section we solve the challenge of the unknow asset value and volatility by introducing an iterative procedure.

We also need to define the default point, D. In the KMV approach Crosbie (2003) the default point is defined as:

$$D = 0.5 \times longterm\ debt + shortterm\ debt$$

The idea is that the short-term debt requires a payment of the face value soon, whereas long-term debt only requires coupon payments in the short run.

4.3 The VX-algorithm

We know put forth a method to estimate the two unknown parameters: asset value, V, and asset volatility, σ. Several approaches to estimate V and σ from equity exist, but the iterative procedure of Vassalou and Xing (2004) has been shown to work very well (Lando, 2004). We follow Jessen & Lando (2015) and from now on the procedure is described as the VX-algorithm.

The iterative procedure uses the fact that equity is equal to the value of a call option on the firm's asset with strike equal to the face value of debt as described in the Merton model.

$$E_t = C(V, D, T - t, \sigma, r)$$

Assume that we observe a series of equity values, $E_{t0}, E_{t1}, \dots, E_{tN}$

By inverting the formula, we can solve for the asset value for a given σ. Let $V_{t_i}(\sigma)$ be the asset value obtained for a given σ. The then *n*th step brings us from an estimate of σ_V^n asset volatility to an improved estimate of asset volatility σ_V^{n+1}, the VX-algorithm follows the form:

Start with an initial guess on σ_V^n:

1. Calculate $V_{t_0}, V_{t_1}, \dots, V_{t_N}$ from the observed share prices $S_{t0}, S_{t1}, \dots, S_{tN}$ using the inverse of the Black-Scholes formula as a function of the underlying assets.
2. Estimate σ_V^{n+1} by thinking of $V_{t_0}(\sigma_V^n), V_{t_1}(\sigma_V^n), \dots, V_{t_N}(\sigma_V^n)$ as a GBM, i.e. let

$$\sigma_V^{n+1} = \sqrt{\frac{1}{N\Delta t}\sum_{i=1}^{N}\left(\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right) - \xi\right)^2}$$

Where

$$\xi = mean = \frac{1}{N}\sum_{i}^{N}\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right)$$

Now use the updated value and continue with σ_V^{n+1} in place of σ_V^n, run the VX-algorithm until $(\sigma_V^{n+1}-\sigma_V^n)$ converges to 0.

To truly understand what goes on in the VX-algorithm, we again note that the asset value follows a process of

$V_t = V_0 e^{(\mu - 0.5*\sigma^2)t + \sigma W_t}$

That can be written as

$V_t = V_{t-1} e^{(\mu - 0.5*\sigma^2)\Delta t + \sigma\sqrt{\Delta t}\epsilon}$

Where ϵ is standard normal distributed and Δt is the distance between the asset value of the firm. We now remember from the properties of Brownian motions that the increments are independent if the time intervals do not overlap.

It gives us

$$logV_t - logV_{t-1} = (\mu - 0.5*\sigma^2)\Delta t + \sigma\sqrt{\Delta t}\epsilon \sim N\big((\mu - 0.5*\sigma^2)\Delta t, \sigma^2\Delta t\big)$$

Where $N(mean, variance)$. That is, we have observations from the iid normal distribution.

From elementary statistics it is straight forward to find mean and variance in a normal distribution.

$$Mean = \xi = \frac{1}{N}\sum_{i}^{N}(\log V_{t_i} - \log V_{t_i}) \ = \frac{1}{N}\sum_{i}^{N}\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right)$$

$$Variance = \frac{1}{N}\sum_{i=1}^{N}\left(\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right) - \xi\right)^2$$

But we want to find μ and σ, hence we put them into the expression.

For the variance

$$Variance = \frac{1}{N}\sum_{i=1}^{N}\left(\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right) - \xi\right)^2 = \sigma^2\sqrt{\Delta t}$$

Moving around solving for σ

$$\hat{\sigma} = \sqrt{\frac{1}{N\Delta t}\sum_{i=1}^{N}\left(\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right) - \xi\right)^2}$$

For the mean we have

$$\hat{\xi} = \frac{1}{N}\sum_{i}^{N}\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right) = (\mu - 0.5 * \sigma^2)\Delta t$$

Moving around and we get:

$$\hat{\mu} = \frac{1}{N\Delta t}\sum_{i}^{N}\left[\log\left(\frac{V_{t_i}}{V_{t_{i-1}}}\right)\right] + 0.5 * \sigma^2$$

In the last equation, we put in the estimated value of σ (Dick-Nielsen, n.d.).

We have now obtained V, σ and μ, hence we can calculate distance to default and probability of default under physical probability measure P. Do we use the risk-free rate instead of the μ then we obtain distance to default and probability of default under the risk-neutral probability.

5. Credit market

Originally the credit market consisted of firms that issued debt to obtain capital to finance its activities. But there has been an ongoing development of credit products and credit derivatives throughout the last decades. The most popular single-name credit derivative is the credit default swap (CDS) (Hull, 2012). The payoff from this derivative depends on the credit risk of the reference firm. There are two sides to a CDS contract: a seller of protection and a buyer. There is a payout from the seller if the reference firm defaults on its obligations, whereas the buyer makes period payments to the seller until maturity of the contract or until a credit event occurs. The payments made from the buyer to the seller goes under the name the CDS spread and are often denoted in basis points (Hull, 2012).

5.1 CDS spread

Here we show how the CDS spread can be calculated, by presenting the simple numerical example that is put forth by Hull (2012):

Assume that we have calculated the risk neutral probability of default to be 2% for a year conditional upon no earlier defaults. The default probability of the first year is 2%, hence the survival probability is 98%. The probability of default in the second year becomes 2%*98%=1.96%, whereas the probability of survival until the end of the second year is 98%^2=96.04%. The probability of default during the third year becomes, 2%*96.04%=1.92% and so it continuous. We assume that defaults only happen halfway through a year and payments on the credit default swap are at the end of year. We also assume the risk-free rate is equal to 5% per annum, and it is continuous compounded. And we assume that the recovery rate is 40%. There are three steps to the calculations. The first step is to calculate the present value of the expected payments made by the protection buyer of the credit default swap. The second step is to calculate present value of the expected payments made by the protection seller of the credit default swap. As the third and final step we need to consider the accrual payments made in the event of default. For example, there is a 1.92% chance that there will be a final accrual payment halfway through the third year. The expected accrual payment half way through the third year becomes 1,92%*0.5*s.

We now calculate the expected payments made by the protection buyer,

Time	PD	Survival probability	E[payment]	Dis. Factor	PV(E[payment])
1	0.0200	0.9800	0.9800	0.9512	0.9322
2	0.0196	0.9604	0.9604	0.9048	0.8690
3	0.0192	0.9412	0.9412	0.8607	0.8101
4	0.0188	0.9224	0.9224	0.8187	0.7552
5	0.0184	0.9039	0.9039	0.7788	0.7040
Total					4.0704

Where we see that the total present value of the payments is 4.0704s

Next, we calculate the present value of the expected of the accrual payment in the event of default.

Time	PD	Survival probability	E[Accrual payment]	Dis. Factor	PV (E[Accrual payment])
0.5	0.0200	0.9800	0.0100	0.9753	0.0098
1.5	0.0196	0.9604	0.0098	0.9277	0.0091
2.5	0.0192	0.9412	0.0096	0.8825	0.0085
3.5	0.0188	0.9224	0.0094	0.8395	0.0079
4.5	0.0184	0.9039	0.0092	0.7985	0.0074
Total					0.0426

Where we see that the total present value of the payments is $0.0426s$

The total of the two present values is $4.1130s = (4.0704 + 0.0426)s$

The next step is to know the present value of the expected payments made by the protection seller, the following table shows the calculations

Time	PD	Survival probability	Recovery rate	E[payment]	Dis. Factor	PV(E[payment])
0.5	0.0200	0.9800	0.4	0.0120	0.9753	0.0117
1.5	0.0196	0.9604	0.4	0.0118	0.9277	0.0109
2.5	0.0192	0.9412	0.4	0.0115	0.8825	0.0102
3.5	0.0188	0.9224	0.4	0.0113	0.8395	0.0095
4.5	0.0184	0.9039	0.4	0.0111	0.7985	0.0088
Total						0.0511

The total present value of the expected payments of the accrual payments is 0.0511.

When the two parties agree on the CDS spread on the given example, they should agree up on a value that solves the following equation:

$$4.1130s = 0.0511$$

$$s = 0.0124$$

In basis points this will be 124 per year.

5.2 Corporate bond yield spread

A bond is a security sold by a government or a firm to raise capital today in exchange for a promised future payment to the bondholder.

A government bond is often assumed to be risk-free because it is assumed that a government will always be able to meet its obligations. Government bonds are often used as a proxy for the risk-free rate in a given currency. In contrast corporate bonds are not assumed to be default free. The difference in the risk exposure depends upon the firm and the government to which it is compared. To take on the additional risk the bondholder wants to be compensated. The difference in the yields is defined as the corporate bond yield spread (Berk & DeMarzo, 2013).

5.3 Why use the CDS spread over the corporate bond yield spread as a measure of credit risk?

In our empirical test the CDS spread has some advantages relative to the corporate bond yield spread, when used as a measure for credit risk.

The corporate bond yield spread account for more than just the risk of default, a significant part of the bond yield spread are determined by liquidity factors, these factors may not reflect the probability of default of the underlying firm (Chen et al., 2007) and (Longstaff et al., 2005).

It has also been shown that the CDS spread generally reacts faster to changes in credit risk of the reference firm than the corporate bond yield spread (Blanco, Brennan, & Marsh, 2005) and (Zhu, 2006). Furthermore the derivative nature of the CDS makes the CDS spread less sensitive to shocks in supply and demand compared to the bond spread (Cesare & Guazzarotti, 2010).

6. Regression method

Credit risk is by nature quantitative and it falls natural to use a quantitative method to answer the problem statement. The choice of method falls on the OLS regression.

The problem statement addresses the possibility that the relationship between the CDS spread and estimated credit risk measure from the Merton DD Model might depend on state of the economy. To capture the potential change, we make use of a binary variable to indicate the state of the economy and in addition we let the binary variable interact with the estimate for credit risk. We do so to allow for a different intercept and a different slope of the regression line during the financial crisis. In the following section we put forth the general theory of how we set up the regression models that allow to test for different intercept and different slope during the financial crisis.

6.1 Interactions between a continuous and a binary variable

The following section builds on Stock & Watson (2012) method to deal with interactions between a continuous and a binary variable. Their method allows the population regression line relating Y and the continuous variable X to depend on the binary variable D in three ways. The three ways are: different intercept, different intercept and slope, and different slope.

Different intercept:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D + \epsilon_i$$

It is a multiple regression model with a population regression function that is linear in X_i and D. D is a binary variable that can only take on the values 1 or 0.

When $D = 0$, then the estimated regression line is $\beta_0 + \beta_1 X_i$. When $D = 1$ the estimated regression line is $\beta_0 + \beta_1 X_i + \beta_2$, the slope, β_1, does not change, but the intercept does and becomes $(\beta_0 + \beta_2)$. Thus β_2 is the difference between the two intercepts of the two regression lines.

Different intercept and slope:

To allow for different slope we use the interaction between a X and D, which yields the following population regression line:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D + \beta_3 (X_i D) + \epsilon_i$$

The new variable $(X_i D)$ is the product of X_i and D, the term is called an interacted-regressor or interaction-term. It captures the potential change in the slope. If we let $D = 0$, then the population regression line is $\beta_0 + \beta_1 X_i$. But when $D = 1$ then population regression line is $\beta_0 + \beta_2 + (\beta_1 + \beta_3) X_i$. The difference between the intercepts in the two regression lines is β_2 and the difference between the slopes is β_3.

A third possibility is that the two lines have the same intercept but different slopes. The population regression line is:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 (X_i D) + \epsilon_i$$

$X_i D$ is still the interacted regressor and captures the potential change in the slope. Taking expected value of Y_i given $D = 0$ yields a regression line of $\beta_0 + \beta_1 X_i$. Letting $D = 1$, then the regression line becomes $\beta_0 + (\beta_1 + \beta_2) X_i$, the difference between the two slopes is β_2.

6.2 Applying the statistical method to our data

We now apply the statistical method of "Interactions between a continuous and a binary variable" to set up the statistical method to answer the problem statement. To ease the notation, we denote the CDS spread at time t as S_t.

The first question in the problem statement is: "how well does the Merton DD model estimate explain the variability in CDS spread in a linear regression setup?"

The question can be answered using either the DtD or PD estimated from the Merton DD model, to exemplify we use DtD.

To answer the first question, we can set up the following regression model:

$$S_t = \beta_0 + \beta_1 DtD_t + \epsilon_t$$

β_1 is the estimated slope of the regression line and capture the sensitivity of the CDS spread to the DtD. To test if the relationship is significant we set up the following hypothesis, H_0: $\beta_1 = 0$ versus the alternative H_1: $\beta_1 \neq 0$. Do we reject H_0, then we conclude that the estimate obtained from the Merton DD model is a significant covariate for the CDS spread. We report the R-square and the Adjusted R-square to answer how much of the variability in the CDS spread that the DtD estimate captures.

The second question in the problem statement is:

"Does the potential relationship between the Merton DD model and the CDS spread depends on the state of the economy?"

To answer the question, we make use of "*different intercept and different slope"* method and the regression model becomes:

$$S_t = \beta_0 + \beta_1 DtD_t + \beta_2 D + \beta_3 (DtD_t * D) + \epsilon_t$$

Where D is a binary variable, which is equal to 1 if financial crisis and else equal to zero.

We set up the hypothesis H_0: $\beta_3 = 0$ versus the alternative H_1: $\beta_3 \neq 0$. If we reject H_0, then we conclude that the estimated beta coefficient for the interaction-term is significant and we conclude that the relationship between the CDS spread and the estimated DtD does depend on the state of the economy. The sign of the estimated beta, β_3,shows if the CDS spread becomes more or less sensitive to the DtD when in crisis.
The regression also makes it possible to test if the CDS spread increase as a function of the economy being in crisis, all else equal. We set up the hypothesis: H_0: $\beta_2 = 0$ versus the alternative H_1: $\beta_2 \neq 0$, do we reject H_0, then we conclude that the CDS spread increase as function of being in crisis, all else equal.

Third and fourth question in the problem statement are: "Is it possible to improve the regression model by including non-firm-specific variables such as the S&P 500 index and the volatility measure from the VIX index?" and "Is the potential relationship between the non-firm-specific variables and the CDS spread dependent on the state of the economy?"
For each of the two variables, the S&P500 index and the VIX index, we run the same analysis as done for the CDS spread against the DtD.
If they are significant predictors of the CDS spread and their relationship do change as a function of the state of the economy, then the final regression model can have the form:

$$S_t = \beta_0 + \beta_1 DtD_t + \beta_2 D + \beta_3(DtD_t * D) + \beta_4 SP500_t + \beta_5(SPX_t * D) + \beta_6 VIX_t + +\beta_7(VIX_t * D) + \epsilon_t$$

The interpretation of the estimated beta coefficients is the same as just presented and the hypothesis to test for will follow the same procedure as just presented.
To answer if including the two macro variables improve the model we compare the Adjusted R-Square of the final model against the Adjusted R-square of the regression model: $S_t = \beta_0 + \beta_1 DtD_t + \beta_2 D + \beta_3(DtD_t * D) + \epsilon_t$

6.3 OLS regression model assumptions

We now present the assumptions the models rely on and what the consequences may be if the assumptions are violated. The 4 basic assumptions are: model form, independence of residuals, homoscedasticity, and normal distributed residuals, (Makridakis, Wheelwright, & Hyndman, 1998).

Model form the assumption is about the form of relationship between the dependent and the explanatory variable. For the OLS regression we assume that the relationship is linear. If the assumption is violated the forecasts may be inaccurate, and the F-test, t-test and confidence intervals are not strictly valid.

Independence of residuals is directly linked to the validity of F-test, t-tests, R-square and the confidence intervals. If the assumption is violated then the F-test, t-tests, and the confidence interval are not strictly valid, and the estimate of the coefficients might be unstable.

Homoscedasticity is a term used for constant variance of the residuals. The regression model assumes that the variance of the residuals is same throughout. Is the assumption violated then it affects the validity of the F-test, t-test, and confidence intervals.

Normal distributed residuals have no effect on the estimates of the coefficients nor the ability to predict the dependent variable. But if the assumption is violated then the F-test and t-test and confidence intervals are affected.
This assumption is the least serious of the 4 assumptions, because residuals are the result of many factors, of which many are unimportant factors acting together influencing the forecast variable. It is often reasonable to assume that the net effect of such influence is normal distributed. However, is the assumption seriously violated, then tests for significance are inappropriate (Makridakis et al., 1998).

Each of the assumptions can be examined by producing appropriate residual plots.
For model form, we plot the residuals produced by the regression model against the explanatory variable, if any curvature appears then it indicates that the relationship between the dependent and independent variable is non-linear, and the assumption about linearity has been violated.
For independence of residuals, we plot the residuals as a function of observation number to examine if they are independent. In addition, we calculate the autocorrelation of the residuals (serial correlation), if each residual is affected by the residual of the previous period, then the residuals show autocorrelation at lag 1 and we conclude that the assumption has be violated. It is noted that it is possible to have correlation between residuals that at more than one period apart.
Homoscedasticity, we plot the residuals against the fitted values from the regression model. If any pattern arises then the residuals do not have constant variance.
Normal distribution of residuals, we plot the residuals as a histogram to examine if the distribution is normal or approximately normal (Makridakis et al., 1998).

6.4 Stationarity

Although we are investigating the cross section between the CDS spread and the 4 potential covariates PD, DtD, the S&P500 index, and the VIX index, we must keep in mind that we are dealing with time series data since observations are obtained over multiple periods. When dealing with time series data, one must always worry about the time-series being stationary or non-stationary.

First, we need to define what a stationary series is, we follow Brooks (2008) and defined as a series with a constant mean, constant variance, and constant autocovariance structure as stationary, hence we are dealing with the concept of weak stationarity. Put more loosely, a stationary series means that the process does not "wander off" in upward or downward direction.

The use of non-stationary data may lead to spurious regressions, and the regression model could have a high R-square even if the two variables are totally unrelated. If the assumption is not met then it may cause the hypothesis test and confidence intervals not to be valid (Stock & Watson, 2012).

Non-stationarity caused by trend

Trend is a movement of a variable over time, the variable then oscillates around its trend. There are two types of trends in time series data, stochastic and deterministic (Stock & Watson, 2012).

A stochastic trend varies over time and is random. For example, a stochastic trend in CDS spreads might display a long period of increase followed by a prolonged period of decrease. In contrast a deterministic trend is nonrandom in time, a deterministic trend may be a linear function of time. For example, if a stock index had a deterministic linear trend, so that it increases by 2.5 percentage per quarter, then the trend could be expressed as 2.5t, where t denotes number of quarters.

Economics is complex and it is difficult to accept the predictability implied by a deterministic trend for economic variables in the light of surprises and complications met month after month by governments, workers, and businesses. So many econometricians find it most appropriate to model economic time series as having a stochastic trend rather than deterministic (Stock & Watson, 2012). For this reason, we focus on stochastic trends in our dataset.

A very simple model with a stochastic trend is the random walk. A time series y_t follows a random walk if the change in y_t is i.i.d., that is: $y_t = y_{t-1} + u_t$ where u_i is i.i.d.

Often the term random walk is used more general to refer to a time series: $y_t = y_{t-1} + u_t$ where the conditional mean of the error term, u_t, is zero (Stock & Watson, 2012). The idea of a random walk is that the value of tomorrow is the value of today plus an unpredictable change. ´

A time series such as the value of S&P500 index has an obvious upward tendency, in such a case the best forecast must include an adjustment. It leads to an extension of the random walk model that includes the tendency to drift in one direction, hence a random walk with drift. We can model the relationship as

$y_t = \mu + y_{t-1} + u_t$

Where we have $E(u_t|y_{t-1}, y_{t-2}, \dots) = 0$, and μ is the drift in the random walk. The best forecast for tomorrow is today's value plus the drift. A random walk model with a drift has a unit root, the unit root is the drift, μ, (Stock & Watson, 2012).

If a series y_t follows a random walk, then the variance increases over time, hence the series is non-stationary. To see why, one must remember that u_t is uncorrelated with y_{t-1}, hence the variance of y_t is given by $Var(y_t) = Var(y_{t-1}) + Var(u_t)$. For y_t to be stationary $Var(y_t)$ must be independent of time and should be $Var(y_t) = Var(y_{t-1})$ but this can only hold if $Var(u_t) = 0$. Another way to see it is to set the y_0 equal to zero, then we have $y_1 = y_0 + u_1 = u_1$, $y_2 = y_1 + u_2 = u_1 + u_2$ and so it continuous so we have $y_t = u_1 + u_2 + \cdots + u_t$ and we have that u_t is serially uncorrelated then $Var(y_t) = \sigma_u^2 t$ and clearly dependent of time (Stock & Watson, 2012).

For an AR(p) to be stationary we need to look at its roots of the polynomial $1 - \beta_1 z - \beta_2 z^2 - \beta_3 z^3 - \cdots \beta_p z^p$, the roots are the values of z that makes $1 - \beta_1 z - \beta_2 z^2 - \beta_3 z^3 - \cdots \beta_p z^p = 0$. For an $AR(p)$ to be stationary the roots of the polynomial must all be greater than one 1 in absolute value (Stock & Watson, 2012). Does the AR(p) has a root that is equal 1 then the series has a unit root. If a series has a unit root then it has a stochastic trend and is non-stationary (Stock & Watson, 2012).

Problems that arise by stochastic trends.

If an explanatory variable has a unit root then the OLS estimator for the coefficient and its t-statistic may have non-normal distributions, even in large samples.

There are three specific aspects of the above problem, 1) the estimator of the autoregressive coefficient in an AR(1) is biased toward 0 when the true value is 1, 2) the t-statistic of the explanatory variable with a stochastic trend have nonstandard distribution in large samples, and 3) An example of the risk that stochastic trends pose is that two series that are independent will falsely appear to be related if they both have stochastic trends, known as spurious regression(Stock & Watson, 2012).

Detecting stochastic trends

Informal methods to detect trends involves examining the time series plot of data and computation of autocorrelation coefficients. The first autocorrelation coefficient will be close to 1, if the series has a stochastic trend, at least in large samples. If the autocorrelation coefficient at lag 1 is small in combination with a time series plot that has no apparent trend then it indicates that the series does not have a trend (Stock & Watson, 2012) .

A more formal test is the Dickey-Fuller test. It is not the only test for a stochastic trend, but it is the most commonly used in practice and among the most reliable (Stock & Watson, 2012).

The Dicky-fuller test in the AR(p) model

The augmented Dickey-Fuller test for a unit root tests the hypothesis $H_0\colon \phi = 0$ versus the one-sided alternative $H_1\colon \phi < 0$ in the regression model :

$$\Delta y_t = \mu + \phi y_{t-1} + \gamma_1 \Delta y_{t-1} + \gamma_2 \Delta y_{t-2} + \cdots + \gamma_p \Delta y_{t-p} + u_t$$

Under the null hypothesis the series, y_t, has a stochastic trend versus the alternative hypothesis where y_t is stationary. The augmented Dickey-Fuller statistic is the OLS t-stat testing $\phi = 0$ in the above equation.

In general, the lag length p is unknow, but it can be estimated using the information criteria AIC or BIC to decide on the lag length (Stock & Watson, 2012).

6.5 Omitted variable bias

The analysis only includes four explanatory variables, PD, DtD, S&P500 and the VIX index and plus one binary variable. The aim of the thesis it not to try to find the regression model that yields the highest possible R-square in a model with the CDS spread as the dependent variable. But rather to test if and how the four variables effect on the CDS spread depends on the state of the economy. However, only including four variables plus a dummy may cause omitted variable bias in the OLS estimators. The omitted variable bias can occur if either PD, DtD, S&P500 or the VIX index are correlated with the omitted variable, and the omitted variable is a determinant of the CDS spread (Stock & Watson, 2012).

7. Case study, Alcoa Inc

To answer the problem statement, we make a case study of Aluminum Company of America (Alcoa). Alcoa is a global industry leader in production of aluminum and has its headquarters in Pittsburgh, Pennsylvania, in the U.S (alcoa.com). Alcoa has been chosen more or less at random, but was picked because there was no missing data points in the period studied for the variables needed.

Alcoa dates back nearly 130 years to the ground-breaking discovery that made aluminum a vital and affordable part of production and modern life. Alcoa is a major producer of aluminum, such as primary aluminum, fabricated aluminum and alumina combined (alcoa.com).
In the sections to come we put forth information on equity value, debt, ratings and CDS spread for the period of study for Alcoa.

8. Data collection

To answer the research questions, we collect observations for Alcoa. In the data collection process certain questions arise, and the process is not straight forward. The questions are: what is the exact period of study, from which country should the firm be, and what data frequency is optimal.
We want to investigate if the relationship between the CDS spread and the risk estimate from the Merton DD model depends on the state of the economy. A way to do so is to include the financial crisis in the period of study. There is no official start date of the global financial crisis, however the U.S economy experience a recession period of 19 months from December 2007 to June 2009, according to the U.S National Bureau of Economic Research (2010). We choose to define this period as the financial crisis. The period defined as non-crisis then becomes July 2009 forward, then end date of the study is picked arbitrary to be November 2013.
A vast amount of research focusing on credit spread has been carried out on U.S data. We do the same and chose an American firm making comparability between the results of the thesis and existing literature possible.
At which frequency should the data points be observed, in general, when it comes to data, the more the merrier, indicating that daily observations may be the optimal choice. Using daily observations, however cause two main problems: first it can cause the data to by biased by market microstructure noise, which in turn could bias the results (Cohen & Frazzini, 2008). The second problem using daily observations may be that the variability in the variables are not great, making it harder to capture the relationship relative to monthly data. Hence, we answer the research questions using monthly data.
To perform the study, the following data has been collected:
1) Shares outstanding, 2) Share price, 3) Short-term debt, 4) Long-term debt, 5) One-year swap rate, 6) CDS spread, 7) S&P 500 index, and 8) VIX index
1)-6) were obtained from Wharton Research Data Service, whereas 7)-8) was obtained from Yahoo Finance.
To estimate asset value and asset volatility we need the equity value. Equity value is calculated as shares outstanding multiplied by the price per share, data was available at monthly frequency. Book values of short and long-term debt are only published quarterly. To by-pass the missing data points between publication

dates, we use the method of (Gündüz & Uhrig-Homburg, 2014) and (Eom, Helwege, & Huang, 2004a). The quarterly data points for short and long-term debt are kept constant during quarter up until new quarterly data is published.

The risk-free rate is a theoretical concept and no investment is truly risk-free. As a proxy for the risk-free rate we use the one year swap rate, the choice is motivated by the study of Feldhütter & Lando (2008) where they conclude that the swap rate is a better proxy for the risk-free rate than the Treasury rate is for all maturities. They argue that the Treasury yields are affected by convenience yield and if used as a proxy will underestimate the risk-free rate.

The CDS spreads are obtained from CDS contracts with a maturity of 5 years. The maturity of 5 years is motivated by the fact that market is concentrated on maturities of this length (Ericsson et al., 2010).

S&P 500 index, short for Standard and Poor's 500 index, constitutes 500 of the largest U.S stocks, value-weighted. Size however is only one selection criteria, the S&P try to have a representation of different sectors of the economy and picks firms that are industry leaders to represent the broader economy. Many investors consider the S&P 500 an adequate measure of overall US stock market performance (Berk & DeMarzo, 2013).

The VIX is the ticker name of the Chicago Board Options Exchange Volatility index. It tracks the one-month implied volatility of options written on the S&P 500 index. The volatility measure is quoted in percent per annum. The index is one of the most-cited measures of market volatility (Berk & DeMarzo, 2013).

9. Empirical analysis

9.1 Data and summary statistics

To calculate asset value and the asset volatility, we need the equity and debt value of Alcoa, and the risk-free rate. Table 1 in the appendix contains start-of-month value for the three variables over the period December 2007 to November 2013, yielding 73 monthly observations for each variable. In addition, the table contains the 5-year CDS spread (in basis points), the rating and recovery rate of Alcoa.

Table 2 in the appendix contains the estimated asset value, and the calculated distance to default and the probability of default for all 73 points in time. Table 3 in the appendix shows the index values of the S&P 500 index and the VIX index for 73 points in time.

Summary statistics for equity, debt, the risk-free rate, CDS spread, Recovery, Asset value, Distance to default, default probability, S&P500 and the VIX index are presented in the following table:

Summary statistics							
Panel: Means, standard deviations, and quantiles							
			Quantiles				
Variable	*Mean*	Std. dev.	*Min*	*0.25*	*median*	*0.75*	*Max*
Equity	13520	6616	4995	9126	10842	15691	33084
Debt	5193	450	4509	4814	5246	5487	6324
rf(%)	0.95%	0.96%	0.28%	0.39%	0.52%	0.95%	4.29%
CDS spread	294	163	47	192	275	348	883
Recovery	0.40	0.00	0.40	0.40	0.40	0.40	0.40
Asset	18661	6631	10535	14388	16192	20538	38181
DtD	3.37	0.87	1.64	2.80	3.09	3.82	5.66
PD (%)	0.2716%	0.6808%	0.0000%	0.0068%	0.1009%	0.2585%	5.0421%
S&P 500 index	1272	243	735	1099	1286	1408	1848
VIX index	24	10	13	17	22	26	60

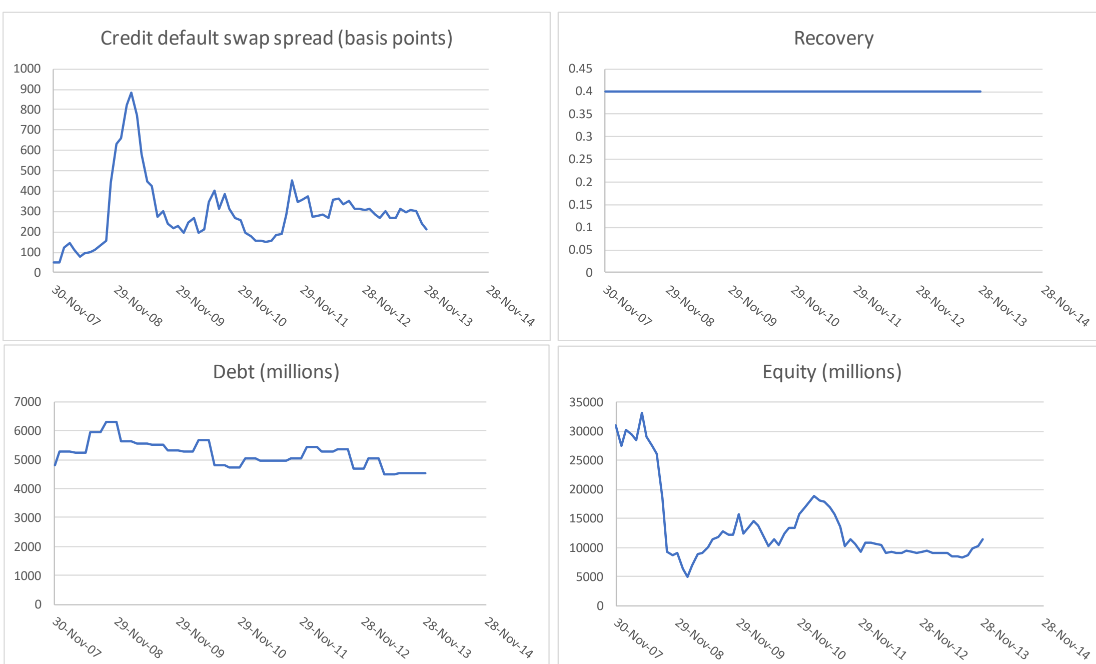
And from the VX-algorithm we have estimated μ to -0.0733, and σ to 0.3503.

9.1.1 Plot analysis

We now plot all variables against time to get an impression of how the data has evolved over the period of study.

We start out with the firm specific variables that are directly observable in the market, hence equity value, debt value defined as short term debt + one half of long term debt, the 5-year CDS spread and the recovery rate.

Firm specific variables observable



The CDS spread

As expected the CDS spread peaks in the period defined as the financial crisis, December 2007 – June 2009 (observation 1-19), it peaks in March 2009 (observation 16) with a value of 883 basis points (rounded). We notice that the CDS spread takes on rather low values in the start of the financial crisis, and for the first 11 months the CDS spread does not get above 160 basis points. That is, more than 50% of the observations in the period defined as the financial crisis fall below the average and median of the CDS-spread for the whole time-period, which are 294 and 275 respectively.

The Recovery rate

The recovery rate does not change during the period and is fixed at 40%.

The Debt value

The debt value, here calculated as short-term debt plus one half of long term debt has an average value of 5193, a standard deviation of 450, a median of 5246, all denoted in millions of dollars.

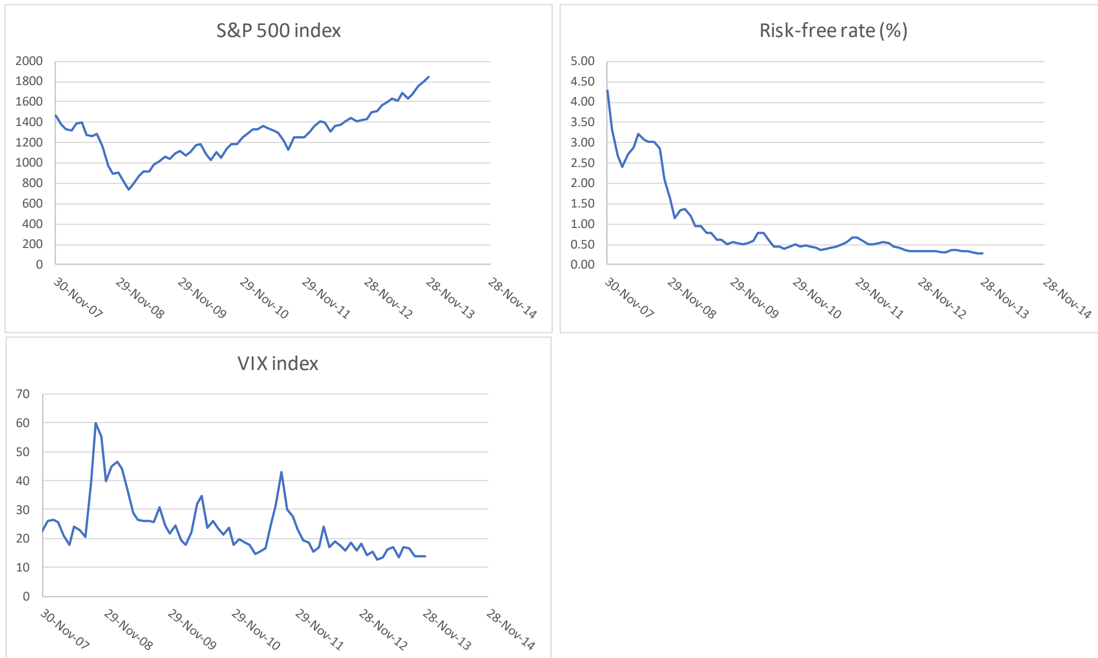
The minimum debt value is 4509, whereas the maximum is 6324. We note how the maximum debt value is within the financial crisis.

The Equity value

The equity value has an average of 13520 and a standard deviation of 6616, denoted in millions of $. The minimum value is 4995 and which is observation number 15, which is February 2009, the date falls in the interval of the financial crisis. The maximum value is of 33084 and is in the start of the period studied, hence both the minimum and maximum value of equity is in the financial crisis. Comparing the CDS spread and equity plot it appears that there is a negative relationship between the CDS spread and the Equity value, which intuitively makes perfect sense.

Macro variables

Macro variables observable



The S&P 500 index

The S&P 500 index has an average of 1272 and a standard deviation of 243. The S&P 500 index has, as expected, its minimum value in the financial crisis, the value is 735 and falls on the February 2009. This is the same date as Equity value of Alcon has it minimum, and one month before the CDS spread has it maximum. From its minimum value the S&P500 index shows a clear upward trend, and the index has a maximum value of 1848, on the last observation in the dataset.

The risk-free rate

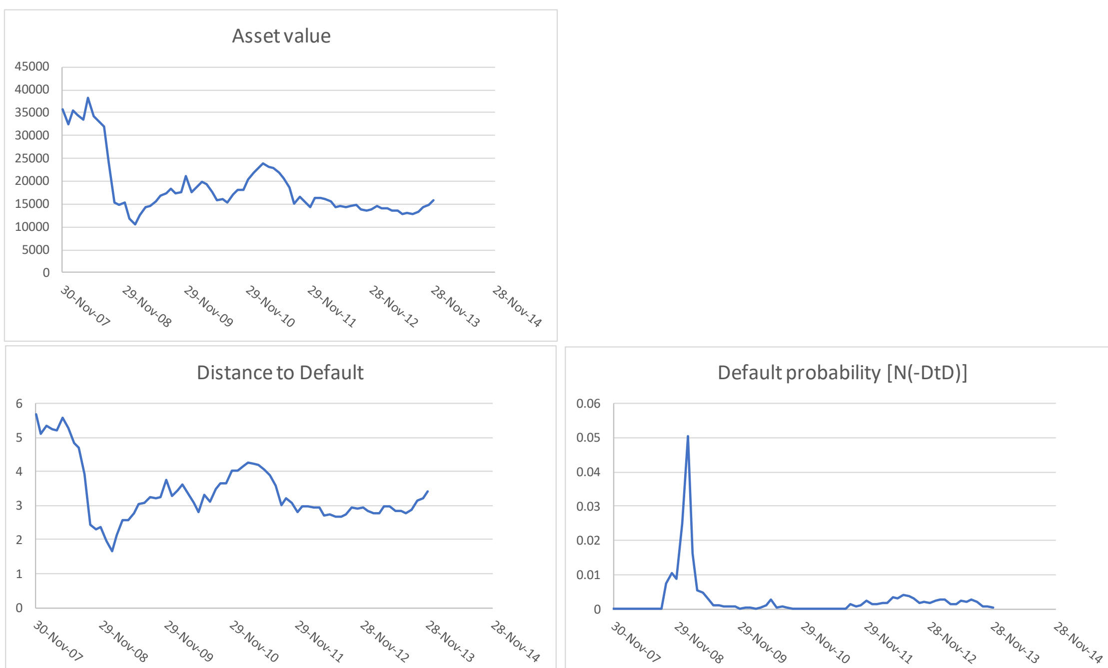
The first observation is the maximum for the risk-free rate, and it is 4.29%. The risk-free decreases down to a minimum of 0.39% on the last observation. The average value of the risk-free rate is 0.95%, whereas the median is 0.52% and the 0.75 quantile is 0.95%. The risk-free rate shows a significant decrease from observation 1 to 4. From 4 to 7 it increases and from 7 to 18 it again decreases significantly from 3.2% to 0.9%. From observation 18 and forward it decreases slowly, with a few of exceptions of positive increases.

The VIX index

The VIX index has an average value of 24 which is close to the median value of 22. The VIX index is rather volatile with a standard deviation of 10. The minimum value is 13 versus the maximum value of 60. The maximum value is the 11th observation, that is October 2008. At first it seems surprising the that VIX index peaks 4 months before the S&P500 index hits its low. But a chock hits the financial world at midnight Monday, September 15, 2008, when the investment bank Lehman Brothers filed for bankruptcy. And it seems natural the volatility measure of the VIX index increase tremendous given the news, it increased by around 300%. After this it starts to decline, but from June 2011 to September 2011 it increases from 25 to 43, and hereafter it declines. The explanation could be that investors had a renewed fear about the health of the global economy and may have feared a new recession. From the plot of the S&P500 index, we see that the S&P500 index level drops in the same period, but quickly recover. The US economy did not go in to recession and the VIX index decreased again.

Firm specific variables estimated from the Merton DD model

Firm specific variables, estimated from Merton DD model



The Asset value

Using the Merton DD model, we obtain an estimated asset value for each of 73 observations in the sample. The average asset value for the period is 18661 and it has a standard deviation of 6631 with a median of 16192, and it has a minimum and maximum value of 10535 and 38181 all values are in millions. It is noted that both the maximum and minimum value fall within the financial crisis, and that combined with a visual inspection of the time-series plot makes it clear that asset value has a higher variance in the financial crisis compared to the period after.

The distance to default estimate

The estimated distance to default is obtained for each observation in the sample. It has an average value of 3.35 with a standard deviation of 0.85. It has a minimum of 1.64 and a maximum of 5.66, both minimum and maximum are contained in the period defined as the financial crisis. As expected DtD has its minimum value in the financial crisis, indicating that the probability of default is highest in the financial crisis.
By visual inspection of the plot we see that variance for the DtD is greatest in the financial crisis. DtD is a function of the asset value, and increases when asset value increases, all else equal. The relation is confirmed by the plots and, to some degree, DtD resembles the pattern of the asset value.

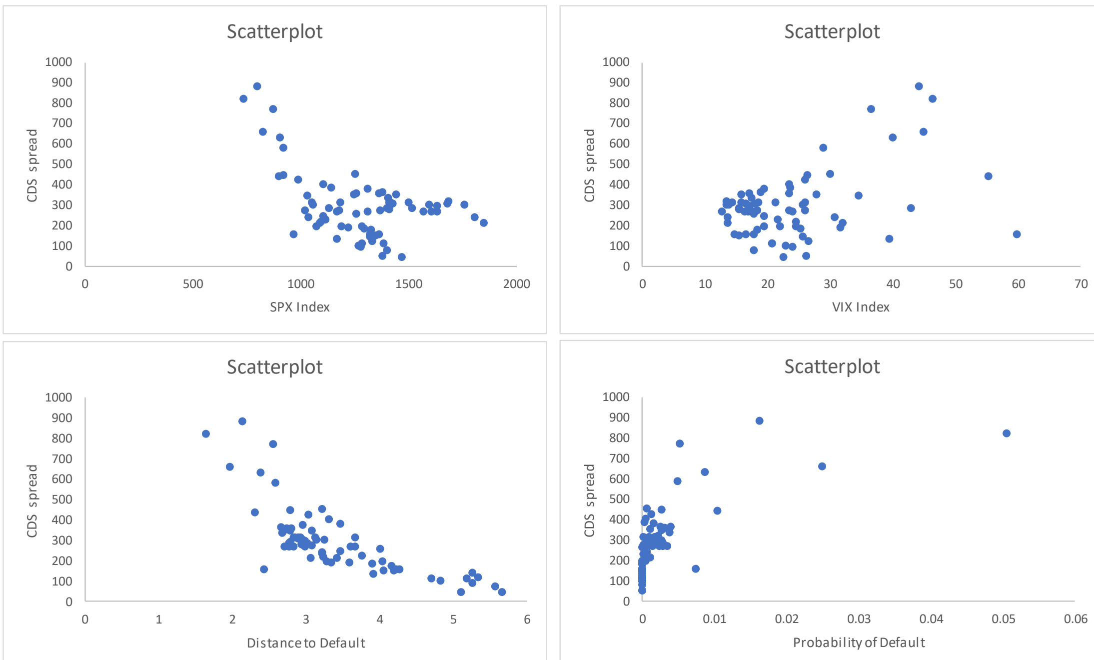
The probability of default

The PD estimate from the Merton DD model contains the same information as the DtD measure. However, it is easier to communicate a probability of default than the number of standard deviations the expected asset value is from the default point at maturity in one year. The probability of default is given by $1 - N(DtD) = N(-DtD)$, hence, it is just a question about how the credit risk is expressed. The maximum default probability is on the same date as when the DtD has its minimum value, the maximum default probability is N(-1.64) =5.0421% (note rounding). The minimum PD is $N(-5.66) \approx 0.0000\%$. PD has an average value of 0.2716% and a median of 0.1009%. Over all Alcoa Inc has a relatively low probability of default.

9.1.2 Relationship between the CDS spread and the potential explanatory variables.

We examine the potential relationship by plotting data. Four scatterplots are produced with the CDS spread on the vertical axis, and the variables: S&P 500 index, VIX index, Distance to default and the probability of default are on the horizontal axis.

Scatterplots



CDS spread against S&P 500 index

From the scatterplot it appears that there is a negative relationship between the CDS spread of Alcoa Inc and the S&P500 index level. When the S&P 500 index level increases it is a sign that the American economy is growing, and then the credit risk of firms in general decrease. From the scatterplot the relationship appears rather linear.

CDS spread against VIX index

The scatterplot indicates that there is a positive relationship between the VIX index level and the CDS spread of Alcoa. The relationship appears to be somewhat linear, but as the VIX index increases so does the variance of the datapoints. The relationship does not appear to be as strong as the relationship between the CDS spread and the S&P500

CDS spread against distance to default

A clear relationship appears when we plot the CDS spread against distance to default, as expected and predicted by theory the relationship is negative. We see low values of the CDS spread when distance to default is large and we see the CDS spread take on large values when distance to default takes on small values. The relationship appears to be linear and the variance of the datapoints does not increase significantly for large values of distance to default.

CDS spread against default probability

By visual inspection of the scatterplot with the CDS spread against the probability of default it appears that there is a positive relationship, just as we expect. However, in contradiction to the other variables the relationship does not appear to be linear. Hence, not to violate the model from assumption about the relation being linear we need to carry out a linear transformation of the variables before we run the OLS regression model.

We now present the correlation matrix. The matrix shows how each variable correlate with the CDS spread and how the potential explanatory correlate with each other.

Correlation Matrix

	CDS spread	S&P500	VIX	DtD	PD(%)
CDS spread	1.0000				
S&P500	-0.4970	1.0000			
VIX	0.4235	-0.7632	1.0000		
DtD	-0.7463	0.2364	-0.2990	1.0000	
PD(%)	0.6768	-0.4355	0.5230	-0.4983	1.0000

The correlation matrix does confirm what the inspection of the scatterplot indicated, and we see that all the correlation coefficients have the expected signs. The CDS spread correlates most with distance to default, the correlation is -0.7463, which is in line with plot analysis where the CDS spread has a strong relationship with distance to default.

A bit surprising is the correlation between CDS and PD, it is 0.6768. The relationship did not appear that strong and linear from the scatterplot. The correlation between the CDS spread and S&P500 index is -0.4970, whereas the CDS spread has a correlation of 0.4235 with the VIX index.

When we look at how the potential explanatory variables correlate with each other, then we see that the S&P500 and the VIX is highly correlated, with a correlation coefficient of -0.7636. We also see that the default probability is more correlated with S&P500 and the VIX than distance to default is. PD has correlations of -0.4355 and 0.5230 with S&P500 and the VIX respectively, compared to distance to default which has correlations of 0.2364 and -0.2990 with S&P500 and VIX respectively.

The fact, that DtD shows smaller correlations with S&P500 and the VIX than PD, indicates that it might be more optimal to use DtD over PD in the multivariate regression setting. However, there is not a problem with perfect multicollinearity, but imperfect multicollinearity could mean that one or more regression coefficients are estimated imprecisely (Stock & Watson, 2012).

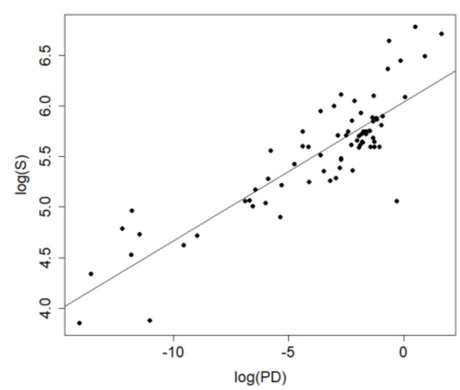
10. Regression

10.1 Regression, CDS spread regressed on default probability

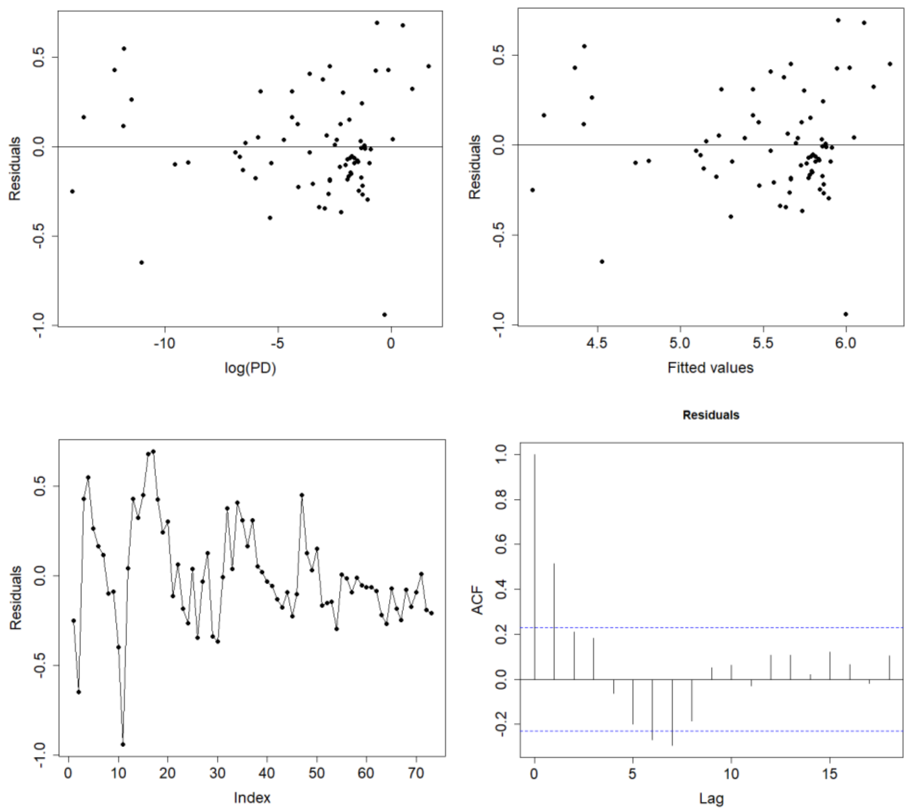
10.1.1 Level regression and checking the assumptions

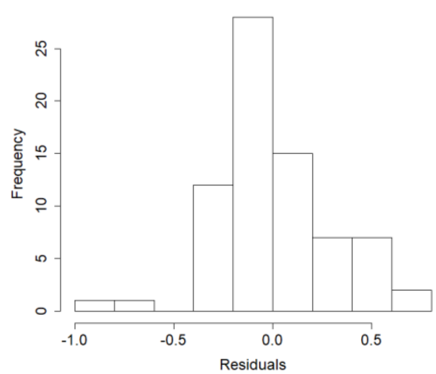
The relationship between the CDS spread and the default probability appear non-linear in the scatterplot in the previous section, to make a linear transformation we take the logarithms of both variables.

We run the OLS regression model: $\log(S_t) = \beta_0 + \beta_1 \log(PD_t) + \epsilon_t$, and plot the least squares regression line in a scatterplot with log(S) plotted against log(PD).



The relationship now appears linear. Before we analyze the regression model output, we examine if the 4 assumptions put forth in “OLS regression model assumptions” section are violated, and we check if both series are stationary.





Model from

When we plot the residual against the variable $\log(PD)$, then the residuals appear random and independent of $\log(PD)$, we conclude that the assumption about model form being linear is met.

Independence of the residuals

From the plot where the residuals are plotted against observation number, the residuals appear to be dependent of each other, with positive autocorrelation.

Investigation of the autocorrelation function plot indicates that the residuals have significant autocorrelation at lag 1, 6 and 7, where especially the spike for lag 1 is far from the 95% limit. The overall conclusion is that the assumption about independence of the residuals is violated.

Homoscedasticity

The residuals appear to be independent of the value of fitted values from the regression model. We conclude that the assumption about constant variance of the residuals is not a problem.

Normal distributed residuals

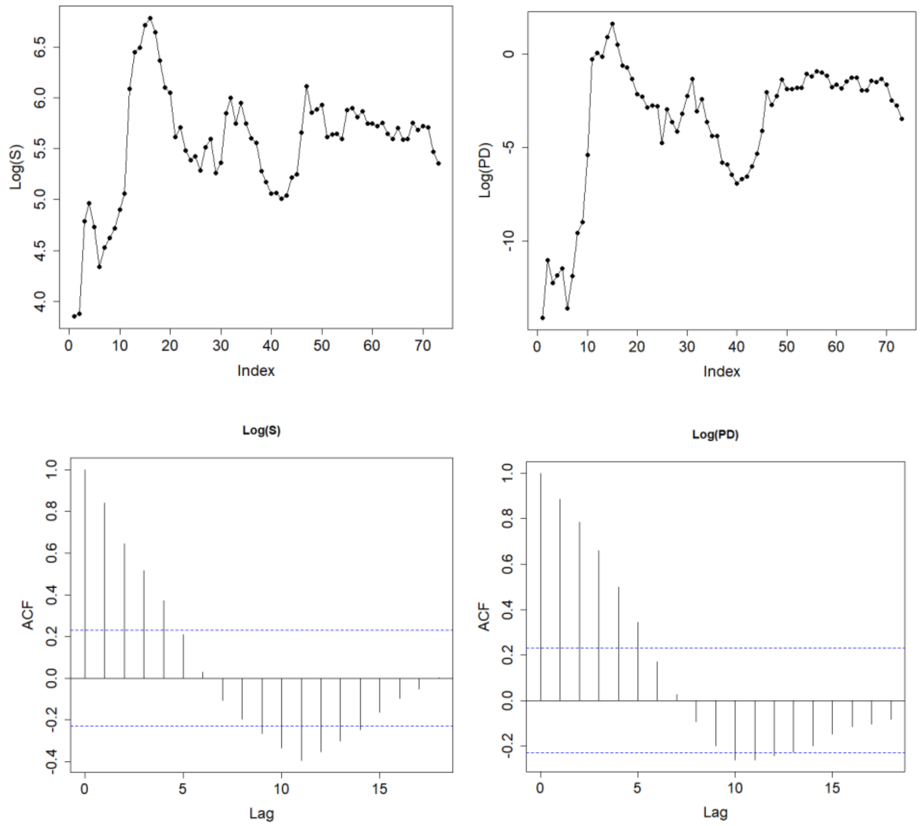
The histogram of the residuals resembles the probability density function of a normal distribution to some degree, but the match is not perfect. Observations near the mean are more frequent that observations far from the mean. The assumption is not seriously violated, and we conclude that the assumption about normality is not a problem.

Stationary

We recall that stationarity means that there is no decline or growth in data, that is the observations must be roughly placed horizontal along the time axis. That is, data fluctuate around a constant mean independent of time and the variance of the fluctuations is essentially constant over time.

Autocorrelation function plot can expose non-stationarity, for stationary data the autocorrelations drop to zero relatively fast, where a non-stationary series have significant autocorrelation for numerous lags (Makridakis et al., 1998).

We now plot the log(S) and log(PD) and plot the autocorrelation function to examine if the two series are stationary or non-stationary.



It is clear from both time series plots that the series do not fluctuate around a constant mean. The autocorrelation function plots also show two series that are non-stationary. The autocorrelation at lag 1 is significant and positive. For both series the autocorrelations at lag 1 are above 0.8. The autocorrelations do decrease but the autocorrelation for the first 4 lags are significantly different from zero at a 5% significance level for $\log(S_t)$. For $\log(PD_t)$ the autocorrelation for the first 5 lags are significant.

$\log(S_t)$ shows negative and significant autocorrelation for lag 9 to 14, where $\log(PD_t)$ shows negative and significant autocorrelation for lag 10-12, however these spikes are close to the 95% limit. The plot analysis indicates that both series are non-stationary.

In addition to the plot analysis we run the more formal test for stationarity, the augmented Dickey-Fuller test. The test yields the following results for $\log(PD_t)$

Variable	Dickey-Fuller test stat	Lag order	P-value
$\log(PD_t)$	-3.8298	4	0.0223

For the series $\log(S_t)$, we get the following results,

Variable	Dickey-Fuller test stat	Lag order	P-value
$\log(S_t)$	-3.1452	4	0.1088

In the augmented Dickey-Fuller test we have the following, hypothesis 0: the series is non-stationary against the alternative hypothesis that the series is stationary.

For the series $\log(PD_t)$, we reject hypothesis 0 at a 5% significance level, with a test stat of -3.8298 and a corresponding p-value of 0.02227, hence the ADF test indicates that the series is stationary.

For the $\log(S_t)$, we do not reject hypothesis 0 at a 5% or 10% significance level, the test yields a test stat of -3.1452 with a p-value of 0.1088. The ADF test indicates that the series is non-stationary. But if we accept significance level of 15% then ADF test indicates stationarity of the series.

For $\log(S_t)$ both the plot analysis and the augmented Dickey-Fuller test indicate that the series is non-stationary. For the $\log(PD_t)$ the ADF indicate that the series is stationary, but the plot analysis of $\log(PD_t)$ indicates that the series is non-stationary.

All though the assumptions are far from perfectly met, we report the regression model output to compare our results with the findings of (Bharath & Shumway, 2008) who, among many tests, also regress $\log(S_t)$ on $\log(PD_t)$.

Regression model results for model: $\log(S_t) = \beta_0 + \beta_1 \log(PD_t) + \epsilon_t$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	6.0396	0.0482	125.2	<2e-16
$\log(PD_t)$	β_1	0.1372	0.0095	14.41	<2e-16
R^2	0.7452				
$Adjusted\ R^2$	0.7416				

R square

The model yields a R-square of 0.7452, hence the model explains 74.52% of the variability in $\log(S_t)$. The adjusted R-square of the model is 0.7416.

Slope analysis for varibale $\log(PD_t)$, β_1

β_1 has an estimated value of 0.1372, the coefficient is significant at a 1% significance level, with a t-test stat of 14.41 with a corresponding p-value of 2e-16.

Intercept analysis, β_0

The intercept has an estimated value of 6.0396, it is significant at a 1% level with a t-test value of 125.20 yielding a p-value that is smaller than 2e-16.

Our findings differ from the resutls that Bharath & Shumway (2008) obtain. They find that $\log(PD)$ explains 26% of the variability in the $\log(S)$. Their model finds that the estimated coefficient to $\log(PD)$ is significant at a 1% significance level and has a value of 0.1737. Their results indicate that the CDS spread is more sensitive to default probability than our model predicts. However, their model has a significantly lower R-square than our model.

The explanation for the difference in results can be multiple. First, they have a large sample, where they investigate 3833 firms, furthermore their study covers the period 1998-2003, whereas our study only covers one firm for a period of 6 years, from 2007-2013.

The high R-square in our model, might be a result of spurious regression. From the plot analysis both of the variables appeared non-stationary, hence we are not sure that we can trust the R-square of our model.

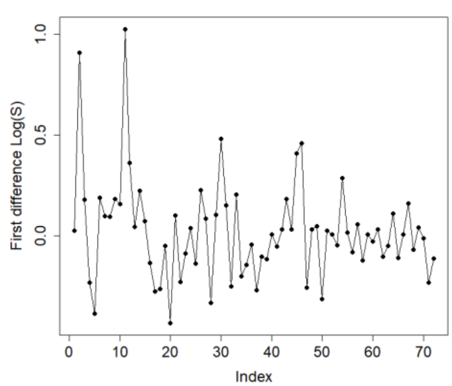
10.1.2 Removing non-stationarity

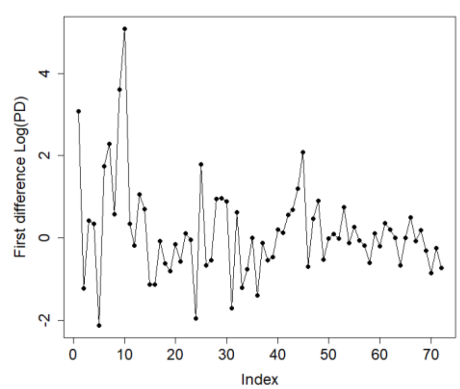
To obtain more trustworthy results, we want to remove the non-stationarity of both variables. One way to do so is trough differencing (Makridakis et al., 1998). We take the first difference of both series, the difference series for the two variables become:

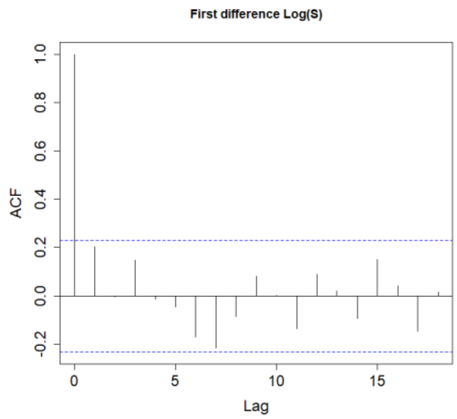
$\Delta\log(PD_t) = \log(PD_t) - \log(PD_{t-1})$

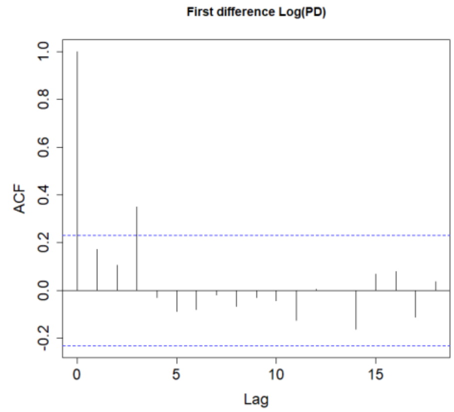
$\Delta\log(S_t) = \log(S_t) - \log(S_{t-1})$

We now plot both series and their autocorrelation function plots to investigate if they have become stationary.









The time series plots show two series that do not appear to wander of in any direction, hence indicting that the series are stationary in mean. However, we do notice that both time series plots have some larger absolute values within the first 11 observations compared to the rest of the period. The autocorrelation function plots confirm the impression of stationarity for both series. No autocorrelations are outside the 95% limits for $\Delta\log(S_t)$ and only the autocorrelation at lag 3 is significant for $\Delta\log(PD_t)$.

Running the augmented Dickey-Fuller test for the series $\Delta log(PD_t)$ yields the following results:

Variable	Dickey-Fuller test stat	Lag order	P-value
$\Delta\log(PD_t)$	-3.5626	4	0.0428

For the series $\Delta log(S_t)$, we get:

Variable	Dickey-Fuller test stat	Lag order	P-value
$\Delta\log(S_t)$	-3.3086	4	0.0773

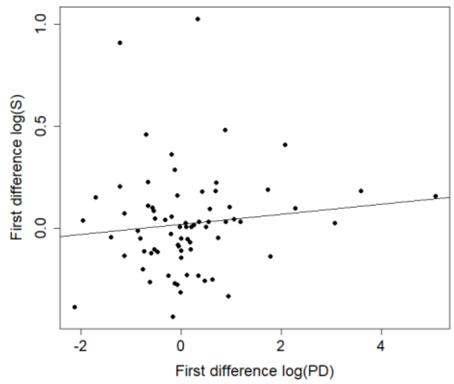
For the series $\Delta\ log(PD_t)$, we reject hypothesis 0 that the series is non-stationary at a 5% significance level, because p-value <0.05. Hence the ADF test indicates that the series is stationary at 5% level.

For the series $\Delta log(S_t)$, we do not reject hypothesis 0 at a 5%, because p-value >0.05, hence we cannot reject that the series is non-stationary at 5%. However, we notice that the p-value has dropped to 0.0734, and if we accept a significance level of 10% then we reject non-stationary and conclude that the series is stationary.

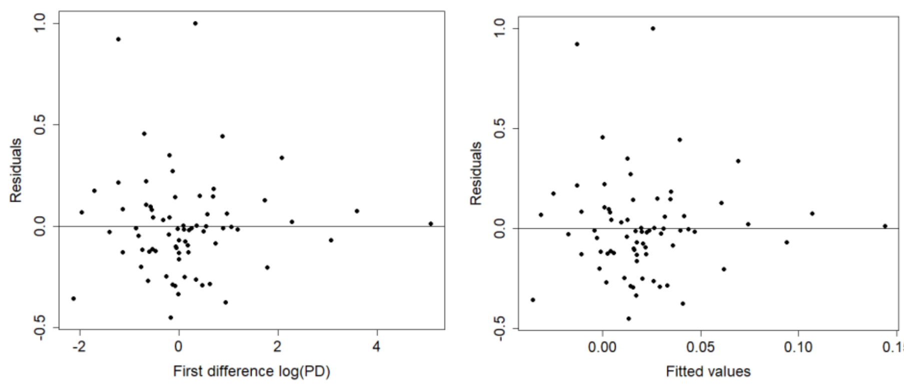
A way to make the $\Delta\log(S_t)$ stationary at a 5% significance level could be to take the second-order difference. Doing so would demand that we also take the second-order difference of $\Delta\log(PD_t)$ as well before running the regression model. The differences become less interpretable and the results of such a regression model are harder to interpret sensibly. So, we now assume that both series are stationary and continue without taking the second-order difference.

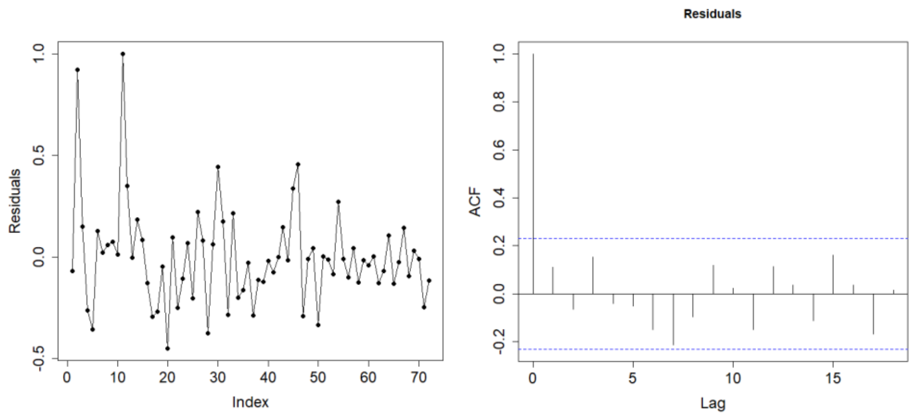
10.1.3 First difference regression

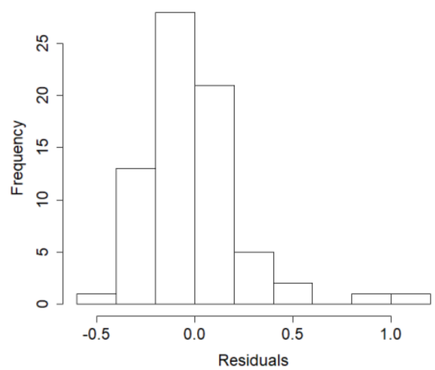
We run the following model: $\Delta\log(S_t) = \beta_0 + \beta_1 \Delta\log(PD_t) + \epsilon_t$ and plot the regression line in a scatterplot where we plot the $\Delta\log(S_t)$ against $\Delta\log(PD_t)$.



From the plot the relationship between the $\Delta\log(S_t)$ and $\Delta\log(PD_t)$ appears insignificant. Before we analyze the regression model results, we check if the assumptions are violated.







Model form

When the residuals are plotted against $\Delta\log(PD_t)$ no obvious pattern appears, and we conclude that the assumption about a linear relationship between $\Delta\log(PD_t)$ and $\Delta\log(S_t)$ is not violated. We do note the two *outliers*, observation 2 and 11, both observations are within the financial crisis and do not seem to come from measurement error. During the financial crisis both the CDS spread and the default probability took on big values, relative to the rest of the period. Therefore, we hope to “catch” these errors when we include the interaction-term and the financial crisis dummy.

Independence of residuals

The autocorrelation function plot shows that there is no problem with autocorrelation of the residuals, as all spikes are within the 95% limits. The time series plot of the residuals confirms this, and no pattern appear. We conclude that the assumption about independence of the residuals is not violated.

Homoscedasticity

The residuals appear to be independent of fitted values obtained from the regression model, thus the assumption about homoscedasticity is not a problem.

Normal distribution of residuals

In the histogram of the residuals mass is concentrated around the mean of 0, and the distribution appear to be somewhat symmetric. However, there are some outliers in right tail, meaning that the residuals do not resemble a normal distribution perfect, but the assumption is not severely violated.

All 4 assumptions are met at a reasonable level, we now analyze regression model results for:

$\Delta \log(S_t) = \beta_0 + \beta_1 \Delta \log(PD_t) + \epsilon_t$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	0.0172	0.0295	0.582	0.563
$\Delta\log(PD_t)$	β_1	0.0249	0.0252	0.989	0.326
R^2	0.0137				
$Adjusted\ R^2$	-0.0003				

R-square

The R-square is equal to 0.01378, hence 1.38% of the variability in the $\Delta\, \mathrm{dlog}(S)$ is explain by the variability in $\Delta\, \mathrm{dlog}(PD)$.

Slope analysis for variable $\Delta log(PD_t)$, β_1

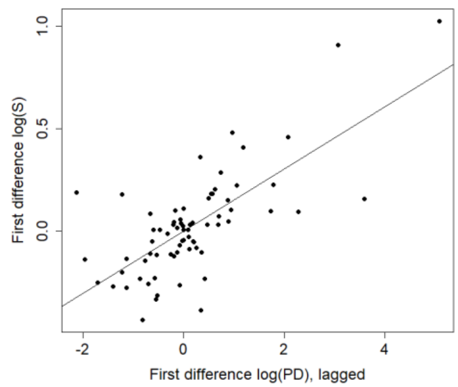
β_1 has an estimated value of 0.02492, a t-test stat of 0.989 with a corresponding p-value equal to 0.326. Hence, we cannot reject that β_1=0 on a 5%, 10% or even 15% significance level. We conclude that β_1 is insignificant.

Overall

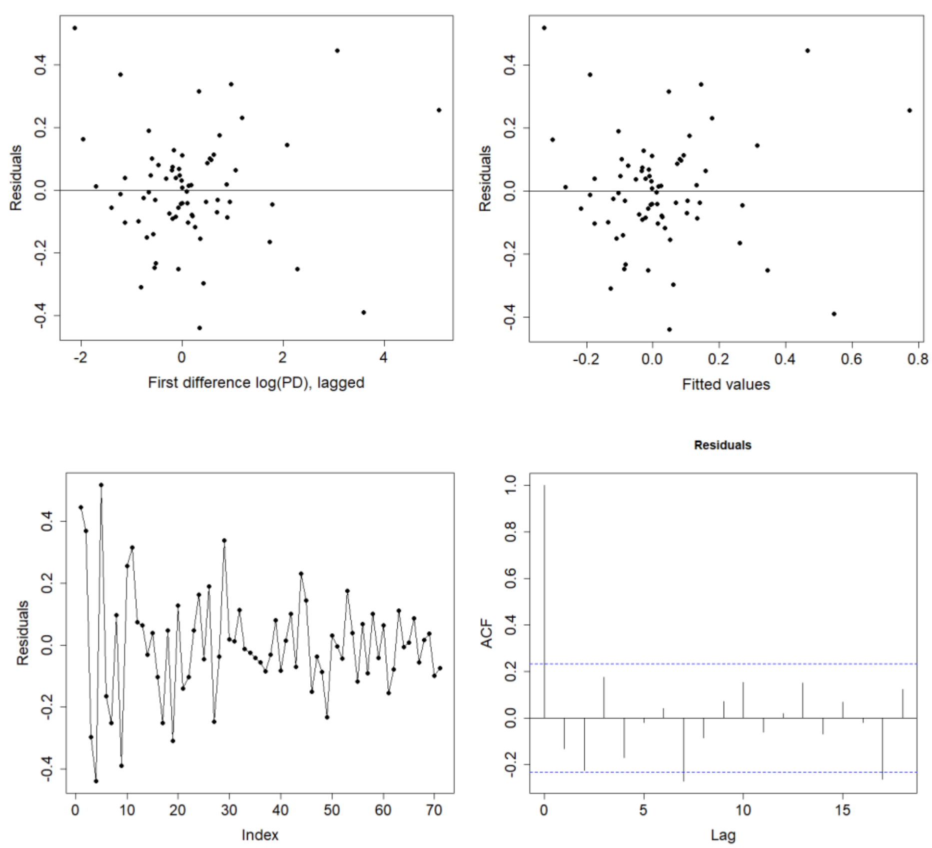
The model results indicate that the changes in $\Delta\log(PD_t)$ has no effect on the changes in $\Delta\log(S_t)$. An explanation could be that the market is not fast enough to incorporating the small changes of $\Delta \log(PD_t)$ into $\Delta\log(S_t)$ just as they occur. It might be that it takes some time the CDS spread to react. In the next session we test if this is the case by lagging $\Delta\log(PD_t)$ by one period.

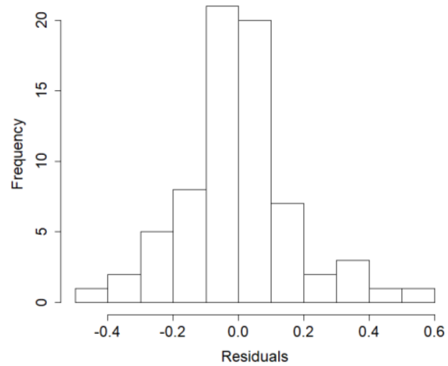
10.1.4 First difference regression, default probability lagged one period

We now set up the following regression model: $\Delta \log(S_t) = \beta_0 + \beta_1 \Delta \log(PD_{t-1}) + \epsilon_t$ and we plot the OLS regression line in a scatter plot with $\Delta\log(S_t)$ plotted against $\log(PD_{t-1})$



By inspection of the plot and it appears that the relationship between the $\Delta \log(S_t)$ and $\Delta\log(PD_{t-1})$ is stronger than the relationship between $\Delta \log(S)_t$ and $\Delta \log(PD)_t$. Before going into the regression results we check if any of the assumptions has been violated.





Model form

No pattern in the plot where the residuals are plotted against $\Delta\log(PD_{t-1})$, we conclude that the assumption about linearity is not violated.

Homoscedasticity

The residuals have come closer to mean value of zero when we plot them against the fitted values, compared to the former model. The residuals are now more *well-behaved*. We see no big outliers, and no clear curvature, hence the assumption about constant variance is not violated.

Normal distribution of residuals

From the histogram we conclude that the residuals now have come closer to resemble a normal distribution. It is still not perfect, but there is no extreme values and the highest concentration of the mass is concentrated around the mean, thus the assumption is met.

Independence of the residuals

From the residuals plotted against the observation number it appears that the residuals are independent of each other. Investigation of the autocorrelation function plot, however, shows that there are significant and negative autocorrelation at lag 7 and 16. But the spikes are just outside the 95% limit. The spikes are close to the limit and we therefor conclude that the assumption of independence of residuals is not severely violated, and we continue as the residuals are independent of each other for all lags.

We now report the results for the regression model: $\Delta\log(S)_t = \beta_0 + \beta_1\Delta\log(PD)_{t-1} + \epsilon_t$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	-0.0036	0.0211	-0.171	0.865
$\Delta\log(PD_{t-1})$	β_1	0.1526	0.0179	8.52	2.24E-12
R^2	0.5127				
$Adjusted\ R^2$	0.5056				

R-square

The model yields a R-square of 0.5127 and an adjusted R-square of 0.5056, that is the $\Delta\log(PD_{t-1})$ explains 50.56% of the variability $\Delta\log(S_t)$.

Slope analysis for variable $\Delta log(PD_{t-1})$, β_1

The estimate slope, β_1, has a value of 0.1526 , a t-test stat value equal to 8.52, and a corresponding p-value of 2.24e-12. We conclude that the β_1 is significant different from 0, at a 1% level. Hence, there is a positive and significant relationship between the $\Delta\log(S)_t$ and $\Delta\log(PD)_{t-1}$.

Intercept analysis, β_0

The estimated intercept, β_0, takes on a negative value of 0.0036, but the intercept is insignificant with a t-test stat value of -0.171 yielding a p-value of 0.865.

Overall

The result is surprising, when we compare it to the existing literature. It is often concluded that the Merton DD model is insufficient in explaining the credit spread and/or its movements. The first question in the problem statement is “How well does the Merton DD model explain the variability in CDS spread in a linear regression setup?” The answer is that the Merton DD model explains 50.56% of the variability of the changes in first difference of the CDS spread, when the estimate from the Merton DD model is lagged by one period. However, the Merton DD model explains basically zero % when not lagged, and the relationship is insignificant.

We conclude that the estimate of the default probability obtained from the Merton DD model is an important variable when predicting the changes in first difference of the CDS spread. We also conclude that the CDS spread does not react immediately when the estimated default probability change, and we conclude that it takes some time before change is incorporated into the CDS spread.

10.1.5 Adding an interaction-term and a financial crisis dummy

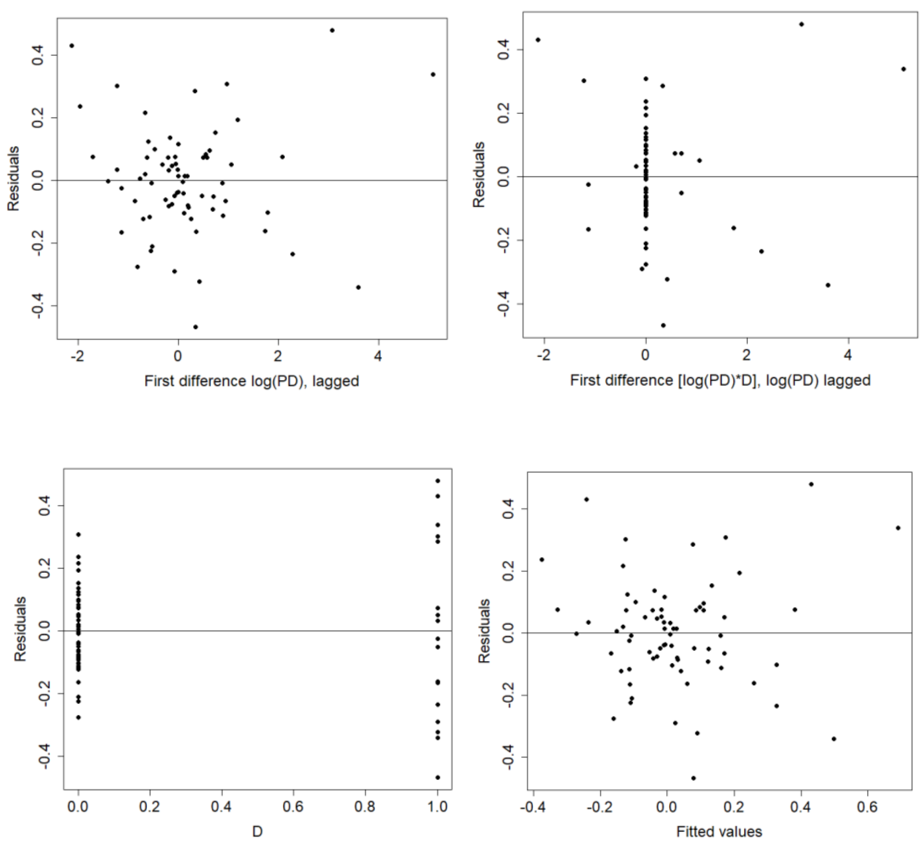
The second question addressed by the thesis is: “Does the relationship between the Merton DD model and the CDS spread depend on the state of the economy”, we answer the question by adding an interaction-term and a financial crisis dummy, which yield the following regression model:

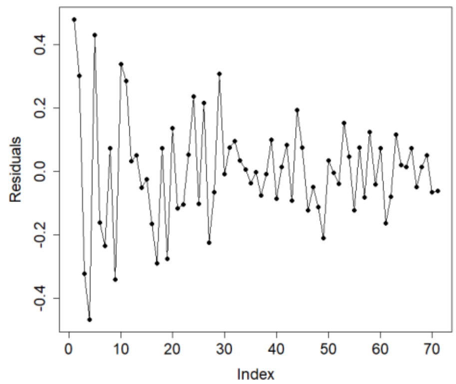
$$\Delta\log(S)_t = \beta_0 + \beta_1\Delta\log(PD)_{t-1} + \beta_2 D + \beta_3\Delta\log(PD_{t-1} * D) + \epsilon_t$$

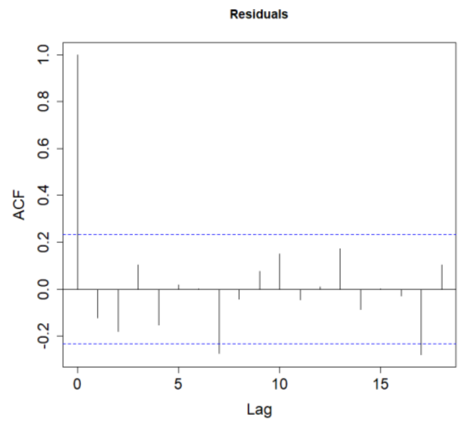
The regression model allows for different intercept during the financial crisis, if β_2 is significantly different from zero then the intercept differs by β_2 and becomes $(\beta_0 + \beta_2)$ when in crisis, given β_0 is significant else it would be β_2.

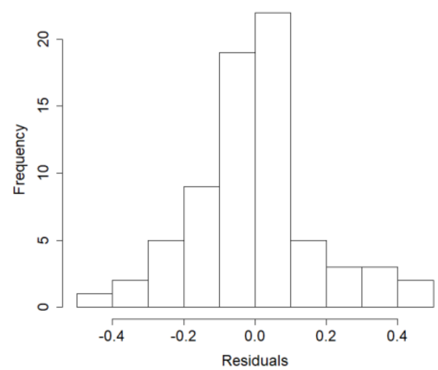
The interaction-term allows the sensitivity of $\Delta\log(S_t)$ to $\Delta\log(PD_{t-1})$ to change when in crisis relative to normal times. If β_3 is significantly different from zero, then the slope of the regression line changes by β_3 when in crisis relative to normal times. $\Delta\log(S_t)$ changes by $(\beta_1 + \beta_3)$ when $\Delta\log(PD_{t-1})$ changes by one unit when in crisis, all else equal. When not in crisis $\Delta\log(S_t)$ changes by β_1 when $\Delta\log(PD_{t-1})$ changes by one unit given β_1 is significant, all else equal.

Before we analyze the regression model results we go over how well the model meets its assumptions.









Model form

When the residuals are plotted against $\Delta\log(PD_{t-1})$, then they scatter random around the mean and we detect no clear pattern, hence the assumption of linearity is not violated.

When the residuals are plotted against $\Delta\log(PD_{t-1} * D)$, then we ignore the data points where $\Delta\log(PD_{t-1} * D)$ is equal to zero, the non-crisis period. No clear pattern in the other datapoints are detected, hence we conclude that the assumption about linearity is not violated.

When we plot the residuals against the dummy D, then we cannot expect the residuals to be random in the scatterplot. The object of the financial crisis dummy is to catch the potential change in intercept, if the intercept changes for an economy in crisis, then we would expect the residuals to be close to zero for $D = 1$.

Independence of residuals

The residuals plotted against observation number show no clear pattern, besides taking on bigger absolute values for the first observations compared to the rest of the study. The autocorrelation function plot, however, shows significant autocorrelation at lag 7 and 17, but the spikes are not far away from the 95% limit. This is autocorrelations as in the former regression model. When comparing the autocorrelation function plots, we do notice that the spikes on average have decreased. The comparison indicates the regression model is an improvement. We conclude that the assumption is not severely violated.

Homoscedasticity

There appear to be pattern of curvature in the residuals when they are plotted against the fitted values of the model. Therefore, we conclude that the assumption about constant variance of the residuals is not violated.

Normal distribution of residuals

The residuals are close to resemble a normal distribution, when plotted as a histogram. It is not perfect but close enough not to be a concerned.

Regression model results for: $\Delta\log(S)_t = \beta_0 + \beta_1\Delta\log(PD)_{t-1} + \beta_2 D + \beta_3\Delta\log(PD_{t-1} * D) + \epsilon_t$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	-0.0081	0.0241	-0.3340	0.7390
$\Delta\log(PD_{t-1})$	β_1	0.1877	0.0309	6.0660	6.81E-08
D	β_2	0.0413	0.0512	0.8060	0.4230
$\Delta\log(PD_{t-1} * D)$	β_3	-0.0586	0.0387	-1.5120	1.35E-01
R^2	0.5302				
$Adjusted\ R^2$	0.5092				

R-Square

The model yields an adjusted R-Square of 0.5092 versus 0.5056 of the former model which did not include the interaction-term and the financial-crisis dummy. Hence the increase in explanatory power is insignificant.

Intercept analysis, β_0

The estimated intercept, β_0, has a value of -0.008054, with a t-test stat of -0.334 yielding a corresponding p-value of 0.739. That is, we cannot reject that $\beta_0 = 0$ at a 1%,5%,10% or even 15% significant level, we conclude that the intercept is insignificant.

Slope analysis for $\Delta\log(PD_{t-1})$*,* β_1

The estimated beta coefficient to the variable $\Delta\log(PD)_{t-1}$, β_1, has a value of 0.1877, a t-test stat of 6.066, with a corresponding p-value of 6.81e-08. We conclude that the estimated β_1 is significantly different from zero at a 1% level. It is noted that estimated coefficient to $\Delta\log(PD_{t-1})$ has increased compared to the model where we do not include the financial crisis dummy and the interaction-term.

Intercept financial crisis, D*,* β_2

β_2, has an estimated value 0.0413, a t-test stat of 0.806, with a corresponding p-value of 0.423. Hence, we cannot reject that $\beta_2 = 0$ at 1%,5%,10% or even 15% significance level. We conclude that the regression model intercept does not change from normal times to crisis.

Interaction-term analysis, variable $\Delta log(PD_{t-1} * D)$, β_3

The estimated beta coefficient, β_3, has a value of -0.0586, with a t-test stat of -1.512, with a corresponding p-value of 0.135. We cannot reject that $\beta_3 = 0$ at a 1%,5%, or 10% level. However, if we accept a significance level of 15% then we reject $\beta_3 = 0$ and conclude that the sensitivity of $\Delta\log(S_t)$ to $\Delta\log(PD_{t-1})$ does changes when the economy is in crisis relative to normal times. For the following part we assume that that β_3 is significant.

The sign of the estimated beta coefficient for the interaction-term is negative. It indicates that the CDS-spread becomes less sensitive to changes in the estimated default probability, when the economy is in crisis relative to normal times, because β_1 is positive. The difference in the regression line is given by β_3, thus -0.0586. Then "*the financial crisis beta coefficient"* for the variable $\Delta\log(PD_{t-1})$ becomes: $(\beta_1 + \beta_3) = (0.187666 - 0.051202) = 0.136464$.

Default probability is a non-systematic risk factor and the result indicates that first difference of CDS spread during times of emergency becomes less sensitive to the changes in it.

The question that arise is, does the CDS spread in our case study depend on systematic risk factors, and do the potential relationships between theses and CDS spread then become stronger during the financial crisis relative to normal times?

To test if the CDS spread does depend on systematic risk, we include the two systematic risk factors: the S&P 500 index and the VIX index.

Adding the first difference of the SP500 index and the VIX index variables to the regression yield the following model:

$$\Delta\log(S)_t = \beta_0 + \beta_1\Delta\log(PD)_{t-1} + \beta_2 D + \beta_3\Delta\,(\log(PD) * D)_{t-1} + \beta_4\Delta SP500_t + \beta_5\Delta VIX_t + \epsilon_t$$

But the model causes concerns, first it is questionable that $\Delta\log(S)_t$ is a linear function of $\Delta SP500_t$ and ΔVIX_t, hence we might violate the assumption about model form. A way to by-pass this is to remember that it is possible to express the riskiness of default as distance to default instead of default probability. Distance to default contains the same information as probability of default and we remember that they are linked together by $PD = N(-DtD)$.

The relationship between the CDS spread and the DtD appears to be rather linear from the scatterplot in previous sections, so do the relationships between CDS spread and the two variables SP500 index and VIX index. Hence a potential regression model where we do not expect to violate the model from assumption is:

$$S_t = \beta_0 + \beta_1 DtD_{t-1} + \beta_2 D + \beta_3 (DtD * D)_{t-1} + SP500_t + VIX_t + \epsilon_t$$

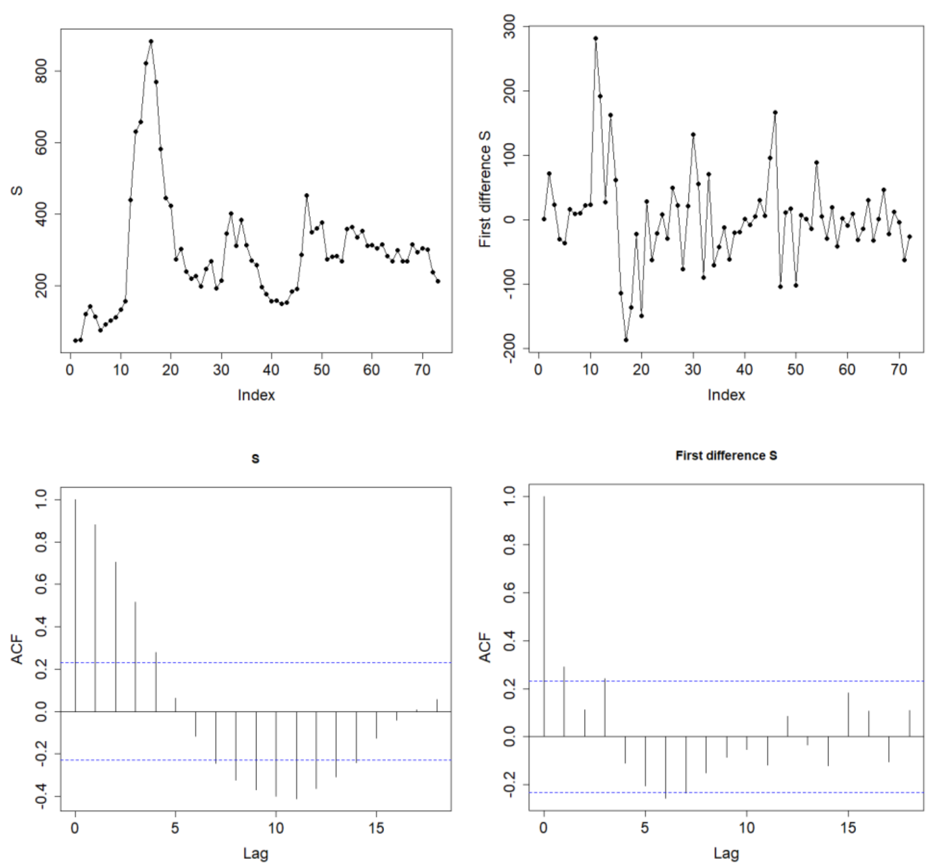
But before running the model, we would have to check if the variables are stationary, and we might end up with a model of the form:

$$\Delta S_t = \beta_0 + \beta_1 \Delta DtD_{t-1} + \beta_2 D + \beta_3 \Delta (DtD * D)_{t-1} + \Delta SP500_t + \Delta VIX_t + \epsilon_t$$

It is not unreasonable to assume that both the dependent and independent variables are stationary after taking first difference. In the following section we examine if the variables are stationary.

10.2 Are the CDS spread, DtD, the S&P500 index, and the VIX index stationary?

The CDS spread level and first difference



Level

Investigation of the plot with the CDS spread as a function of the observation number does not appear to be stationary in mean. The conclusion is supported by the autocorrelation function plot, where we see spikes far from the 95% limits, autocorrelation at lag 1-4 are positive and significant, where 7-13 are negative and significant and lag 14 is just on the limit.

The augmented Dickey-Fuller test yields the following results:

Variable	Dickey-Fuller test stat	Lag order	P-value
S_t	-3.6042	4	0.0392

A bit surprisingly, the test reject H0 that the series is non-stationary at a 5% significance level, hence the augmented Dickey-Fuller test indicates that the series is stationary. The plot analysis and the ADF test point in different directions for the CDS spread at levels.

First difference

The first difference of the CDS spread appears stationary when plotted against observation number. The autocorrelation function plot shows significant autocorrelation at lag 1, but the spike is relatively close to the 95% limit, and the autocorrelations at lag 3, 6 and 7 are just at the 95% limit.

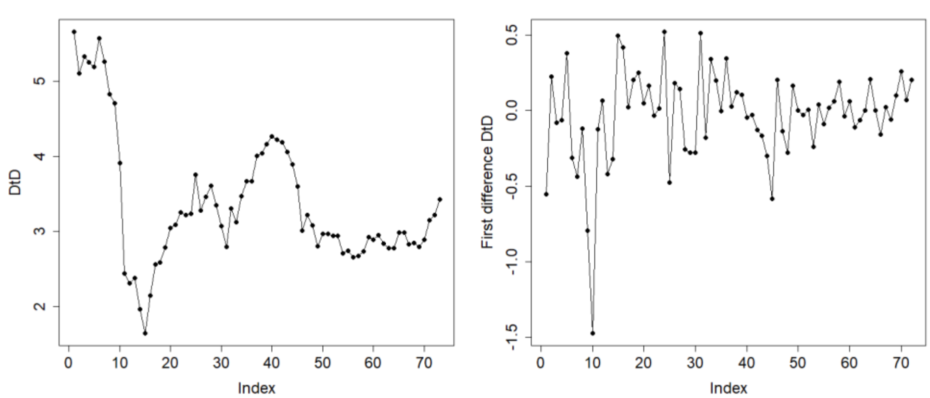
The augmented dickey-fuller test for the first difference of the CDS spread yields the following results:

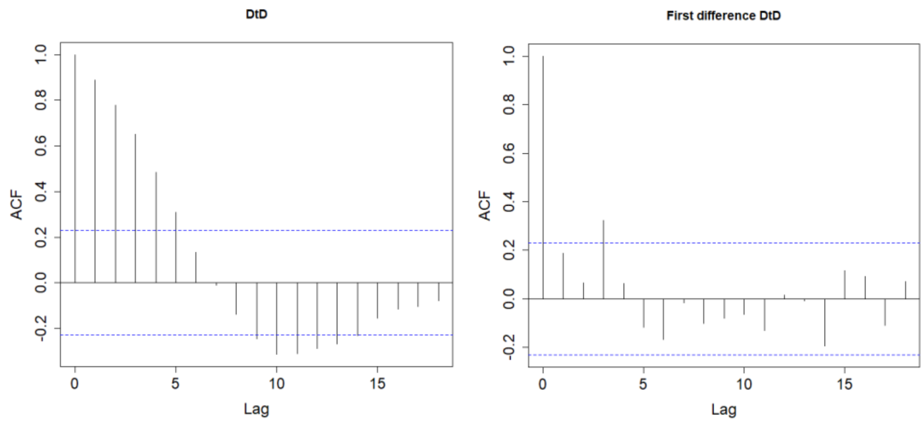
Variable	Dickey-Fuller test stat	Lag order	P-value
ΔS_t	-4.0302	4	0.0133

We reject H0, that the series is non-stationary at a 5% significance, hence the test indicates that the first difference of the CDS spread is stationary.

The results of the augmented Dickey-Fuller test and only a few of the autocorrelation being significant, but close the 95% limit, makes it reasonable to assume that the first difference of the CDS spread is stationary.

Distance to default level and first difference





Level

From the time series plot DtD does not appear to be stationary. Investigation of the autocorrelation function plot also indicates that the series is non-stationary, with autocorrelations significant for lag 1-5 and 10-13. The augmented Dickey-Fuller test yields the following results:

Variable	Dickey-Fuller test stat	Lag order	P-value
DtD_t	-3.8535	4	0.0212

With a p-value of 0.02119, that is, we reject non-stationary at a 5% significance level, and concludes that the series is stationary. The plot analysis and the augmented Dickey-Fuller test point in different direction.

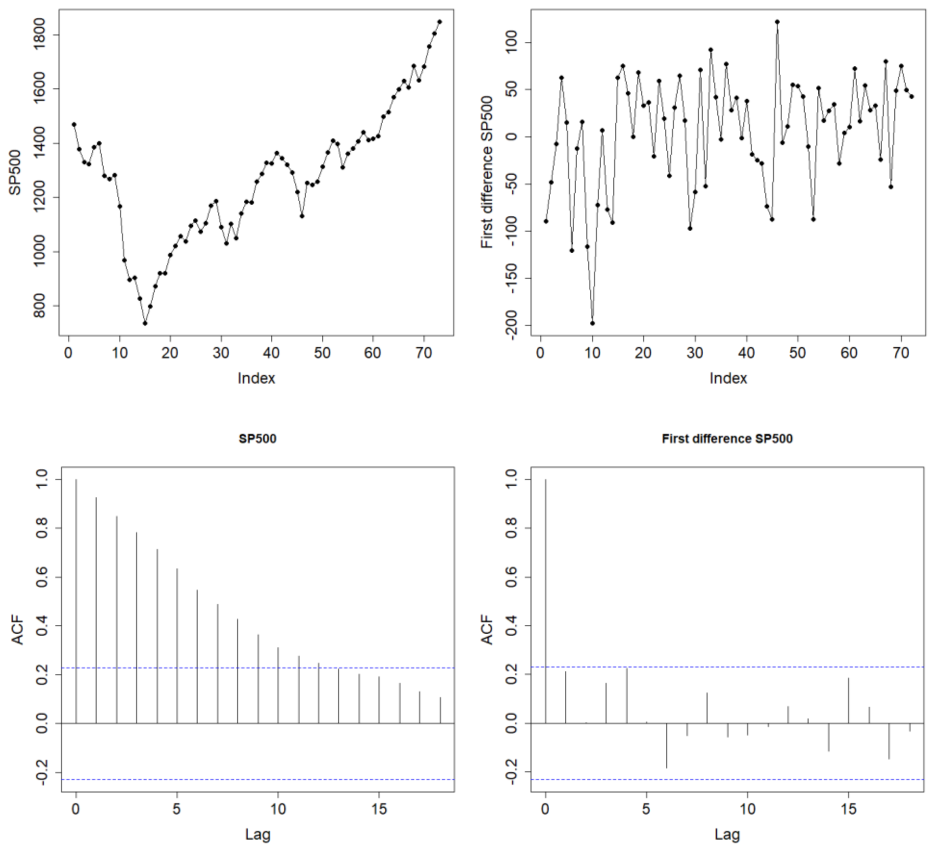
First difference

The time series plot of ΔDtD show a series that appear to be stationary, however it is noted that the variance appears to be a bit larger in the start of the series. The autocorrelation function plot also indicates that the series is stationary with only the autocorrelation at lag 3 being significant. The augmented Dicky-Fuller test yields the fowling results:

Variable	Dickey-Fuller test stat	Lag order	P-value
ΔDtD_t	-3.3731	4	0.0670

The test yields a p-value of 0.06699, that is, we cannot reject that the series is non-stationary at 5% level. If we accept a 10% significance level, then we reject non-stationarity and to conclude that the series is stationary. Keeping this significance level in mind we continue as if the ΔDtD is stationary.

The S&P500 index level and first difference



Level

From the time series plot we conclude that the series S&P500 index is not stationary. The autocorrelation function plot shows a text book example of a series that is non-stationary in mean with a slow decay of the autocorrelation for numerous lags. The first 11 autocorrelations are positive and significantly different from zero at a 5% significance level.

The augmented Dickey-Fuller yields the following results:

Variable	Dickey-Fuller test stat	Lag order	P-value
$SP500_t$	-3.3354	4	0.0729

Which indicates that the series is non-stationary at a 5% significance level, due to the p-value of 0.07292. If we accept a significance level of 10% then the test rejects H0 and indicate that series is stationarity. This is bit surprising considering the time series plot and autocorrelation function plot.

First difference

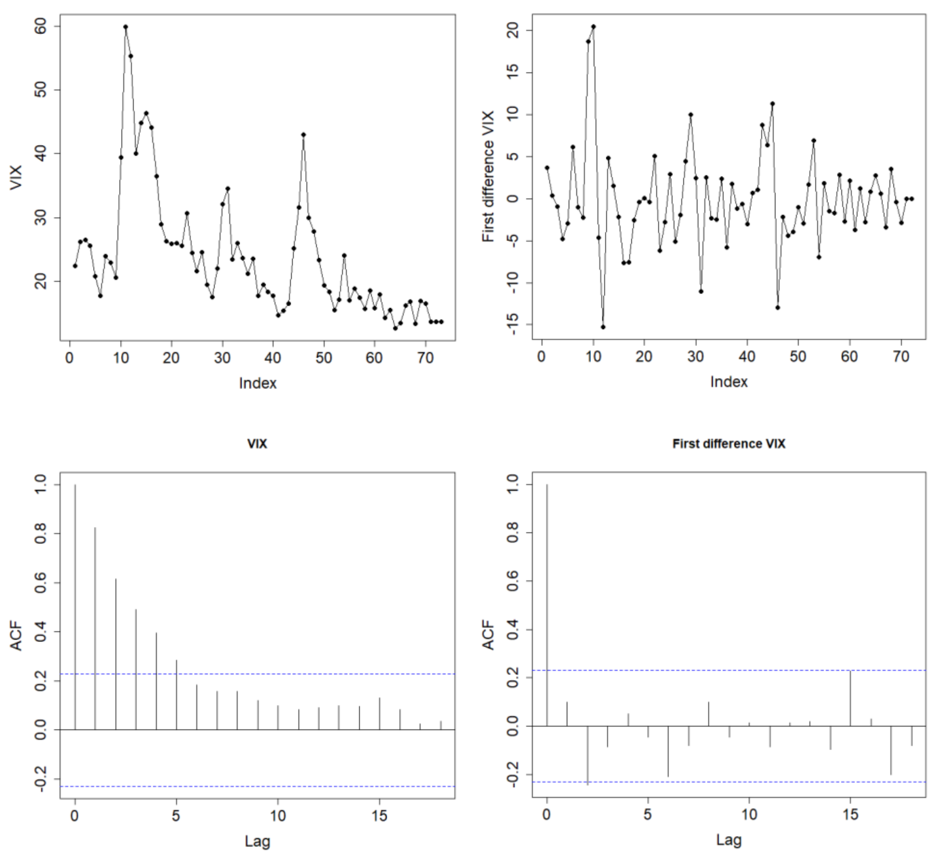
Judging form time series plot the series now appear stationary, however we do note a slightly upward trend. The autocorrelation function plot shows no significant autocorrelations, although the autocorrelation at lag 4 is just at the 95% limit.

The augmented Dickey-Fuller yields the following results:

Variable	Dickey-Fuller test stat	Lag order	P-value
$\Delta SP500_t$	-3.2365	4	0.0889

The augmented Dickey-Fuller test result is surprising, the p-value has increased to 0.08891, hence to reject that the series is non-stationary, we must accept a significance level of 10%. We now move on as if the series $\Delta SP500_t$ is stationary.

The VIX index level and first difference



Level

From the time series plot, the series does not appear stationary, the autocorrelation function plot confirms the impression with the first 5 autocorrelations being significantly different from zero at a 5% level. The augmented Dickey-Fuller test yields the following results:

Variable	Dickey-Fuller test stat	Lag order	P-value
VIX_t	-3.9346	4	0.0175

The test has a p-value of 0.01754, indicating that the series is stationary, hence the plot analysis and the augmented Dickey-Fuller test point in different directions.

First difference

From the time series plot the series now appear to be stationary, we do note that the variance appears to be bigger in start of the period. The autocorrelation function plot also shows a series that appear to be stationary. Only the autocorrelation at lag 2 is significant, but it is close to the 95% limit. And the autocorrelation at lag 15 is just on the 95% limit. The plot analysis indicates that it is not unreasonable to assume that the series is stationary. Running the augmented Dickey-Fuller test yields the following results:

Variable	Dickey-Fuller test stat	Lag order	P-value
ΔVIX_t	-4.1972	4	<0.01

The test has a p-value smaller than 0.01, hence we reject that the series is non-stationary at 1% significance level and conclude that the series is stationary.

Over all conclusion:

We have taken the first difference of all variables and examined both time series plots and autocorrelation function plots. And in addition, we run the augmented Dickey-Fuller test for all the variables.

When running the augmented Dickey-Fuller test on levels, we reject that the CDS spread, distance to default and the VIX index are non-stationary at a 5%, hence the test indicates that all three variables are stationary. The plot analysis, however, indicates that the three series are non-stationary.

For the S&P500 index we cannot reject non-stationarity at a 5% level, but only if we accept a 10% significance level. The plot analysis also indicates that series S&P500 index is non-stationary.

The question that arise is, should we run the regression models on levels rather than on first difference? The plot analysis indicates it would be wrong, all 4 series have significant autocorrelation for several lags and none of the time series plots seem to resemble series that are stationary.

After taking the first difference only few of the autocorrelations are significant for the series CDS spread, and for DtD only one autocorrelation is significant. For the first difference of the S&P500 index a few of the lags are at the 95% limit but none of the spikes cross the limit. For the first difference of the VIX index one lag is significant, but close to the 95% limit. The time series plots of the first difference of all the variables are closer to resemble series that are stationary compared to the plots on levels.

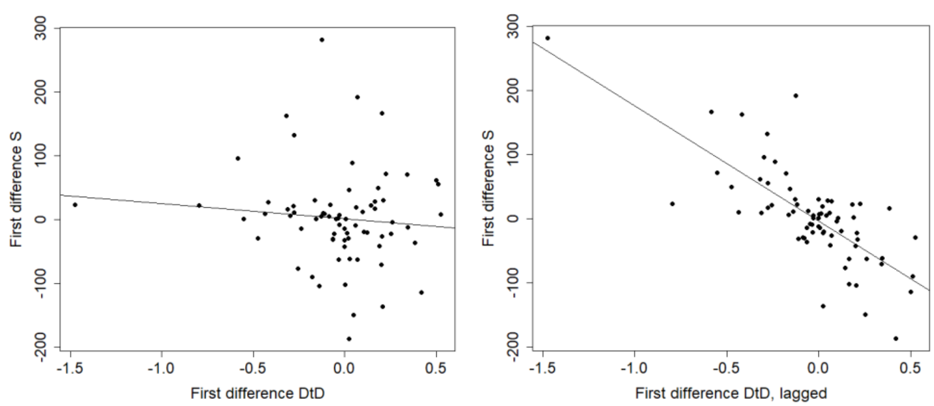
The augmented Dickey-Fuller test indicates that ΔS_t and ΔVIX_t are stationary at 5% level, whereas we must accept a significance level of 10% to reject non-stationarity for $\Delta SP500_t$ and ΔDtD_t. Combing the plot analysis with augmented Dickey-Fuller tests we believe that it is most correct to run the regression on first difference rather than on levels.

10.3 Regression, the CDS regressed on distance to default

10.3.1 First difference regression, distance to default lagged one period

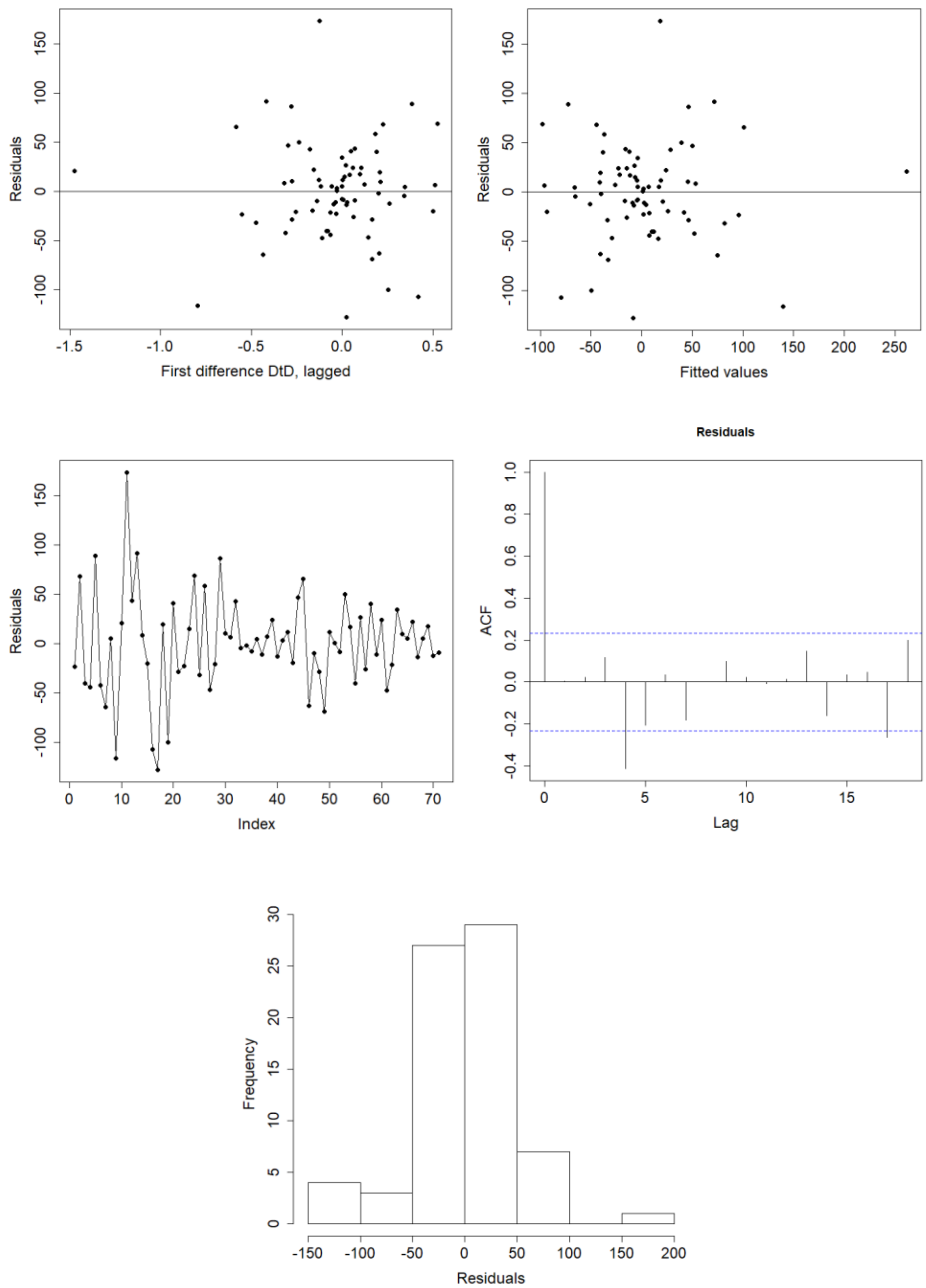
From theory and the empirical tests we have carried out so far, a negative and significant relationship between the CDS spread and the distance to default is expected. When we regressed $\Delta \log(S_t)$ on $\Delta(PD_t)$ we had to lag $\Delta(PD_t)$ by one period for the relationship to be significant, we expect the same to be true for ΔDtD_t.

We now run two regression models: $\Delta S_t = \beta_0 + \beta_1 \Delta DtD_t + \epsilon_i$ and $\Delta S_t = \beta_0 + \beta_1 \Delta DtD_{t-1} + \epsilon_i$ and plot the estimated regression lines in scatterplots with ΔS_t plotted against ΔDtD_t and ΔDtD_{t-1}.



It is clear from the two plots that the relatinship between ΔS_t against ΔDtD_{t-1} is much stronger than the relationship between ΔS_t against ΔDtD_t, just as expected. Hence, we continue the analysis using ΔDtD_{t-1} to explain ΔS_t.

We now check how well the regression model: $\Delta S_t = \beta_0 + \beta_1 \Delta DtD_{t-1} + \epsilon_i$ meets the 4 assumptions:



Model form

We examine the plot where the residuals are plotted against ΔDtD_{t-1}. The linear model specification seems somewhat correct, but it is noted that there appear to be a bit of a spread when ΔDtD_{t-1} deviate from 0, both in positive and negative direction. It may be a problem for the accuracy of the F-test, t-tests and the confidence intervals, but the assumption about linearity does not appear to be severely violated. The explanation for the spread in the residuals might be that the relationship between ΔS_t and ΔDtD_{t-1} differ for the financial crisis, where ΔDtD_{t-1} also takes on bigger absolute values relative to normal times. If this is the explanation, then including the interaction-term and the financial crisis dummy might make residuals independent of the values of ΔDtD_{t-1}.

Independence of residuals

The autocorrelation at lag 4 is significant, and the autocorrelation at lag 17 is significant, however much closer to the 95% limit. We conclude that the assumption is violated.

Homoscedasticity

In the plot with the residuals against the fitted values, the variance increases for bigger absolute values of the fitted values. The explanation might be the same as for the violation of the model form, and if so the solution would be the same.

Investigation of the residuals plotted against observation number shows that the variance is bigger in the start of the period relative to later observations. Hence, we may solve the problem of lack of homoscedasticity by including the interaction-term and the financial crisis dummy to catch the potential different relationship that may be between the CDS spread and distance to default for the financial crisis, relative to the rest of the period of study. But it is noted that the assumption about homoscedasticity is not severely violated.

Normality of the residuals

From the histogram we see that the distribution of the residuals to some degree resembles a normal distribution. We conclude that the assumption about normality is not a problem.

Nevertheless, there are *extreme* values in both ends of the tails, with a relatively high concentration in the left tail. This indicates that the model overestimates the predicted change in the CDS spread for certain values of ΔDtD_{t-1}, these may be for the financial crisis. If this is the case, then the beta coefficient should be smaller in absolute terms when in crisis relative to normal times. This means that we expect the estimated beta coefficient for the interaction-term to have the opposite sign than the estimated beta coefficient ΔDtD_{t-1} when included in the model.

The regression model results for: $\Delta S_t = \beta_0 + \beta_1 \Delta DtD_{t-1} + \epsilon_i$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	-3.885	5.954	-0.653	0.516
$\Delta \mathrm{DtD}_{t-1}$	β_1	-179.904	19.276	-9.333	7.41E-14
R^2	0.5580				
$Adjusted\ R^2$	0.5516				

R-Square

The model yields a R-square of 0.558 and an adjusted R-Square of 0.5516. This is a slight improvement compared to the regression model: $\Delta\log(S)_t = \beta_0 + \beta_1\Delta\log(PD_{t-1}) + \epsilon_t$, which had an adjusted R-Square of 0.5056. Hence, using the ΔDtD_{t-1} instead of $\Delta\log(PD_{t-1})$ makes it possible to explain a bigger fraction of the variability in the first difference of the CDS spread.

Slope analysis for variable ΔDtD_{t-1}, β_1

The estimated beta coefficient is -179.904, it has a t-test stat of -9.333 with a corresponding p-value of 7.41e-14, therefore we conclude that the estimated beta coefficient is significant. β_1 has the expected sign.

Intercept analysis β_0,

The intercept has an estimated value of -3.885, but it is insignificant with a t-test stat of -0.653 and a corresponding p-value of 0.516.

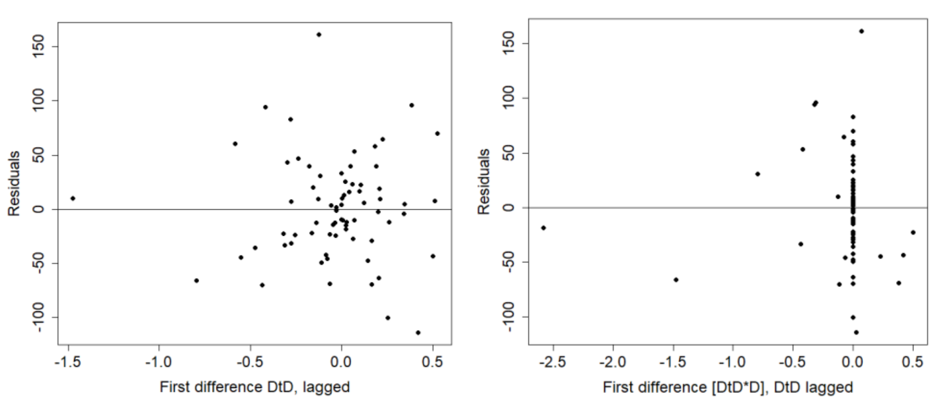
10.3.2 Adding an interaction-term and a financial crisis dummy

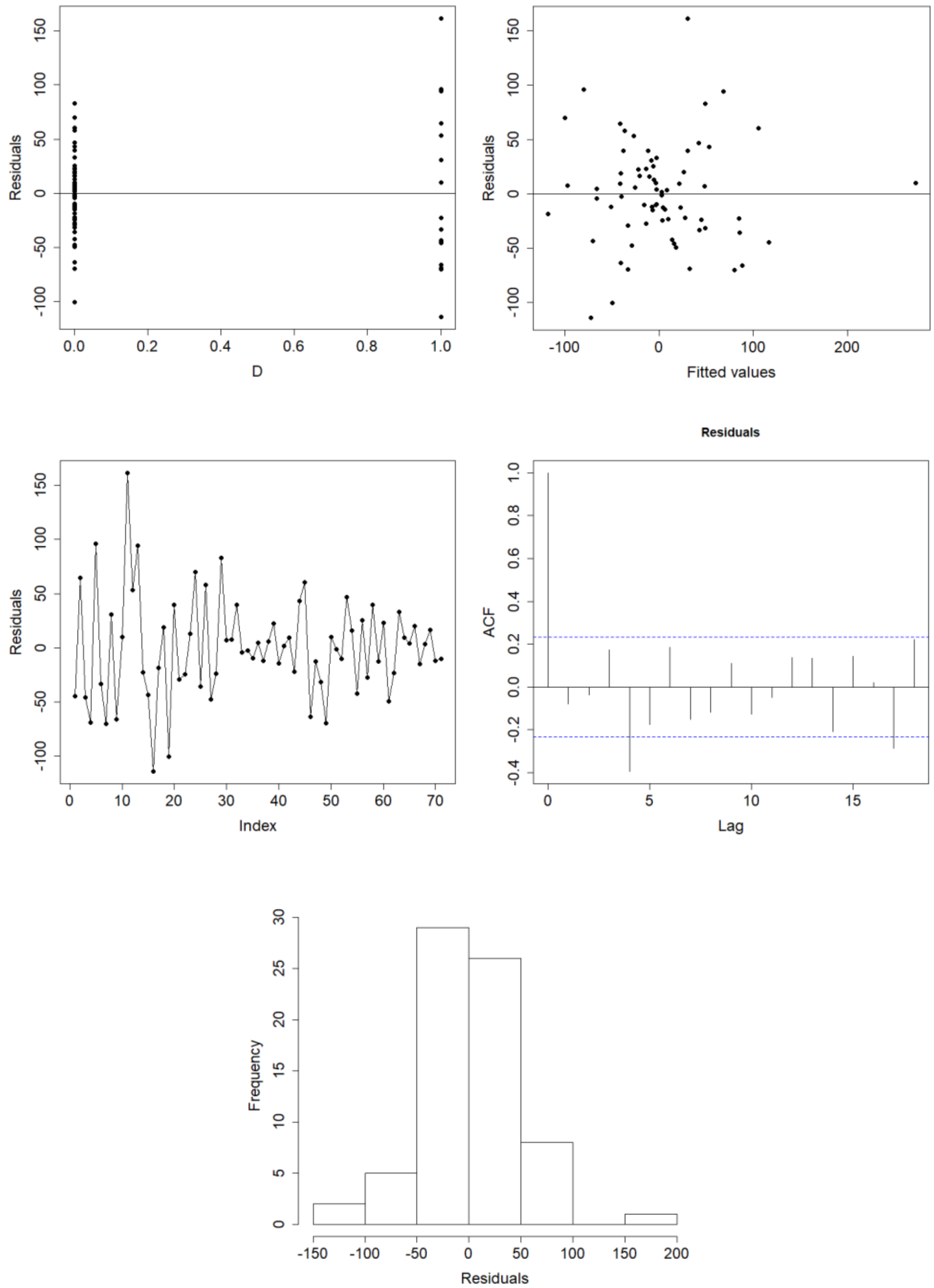
We now set up the regression model:

$$\Delta S_t = \beta_0 + \beta_1\Delta DtD_{t-1} + \beta_2\Delta(DtD_{t-1} * D) + \beta_3 D + \epsilon_t$$

Where D is an indicator variable that takes on 1 in value if we are in the financial crisis, and otherwise 0. $(DtD_{t-1} * D)$ is the interaction term, is β_2 significantly different from zero then we conclude that the slope of the regression line between ΔS_t and ΔDtD_{t-1} depends on the state of the economy. From the analysis of the residuals from the previous regression model, $\Delta S_t = \beta_0 + \beta_1\Delta DtD_{t-1} + \epsilon_t$, makes us expected the estimated β_2 to be positive given it is significant.

We now check the assumptions to by examining the residuals.





Model from

The residuals still appear to be a somewhat dependent of the of the value of the explanatory variable ΔDtD_{t-1}. Hence including the interaction-term has not solved the issue perfectly. Nevertheless, the assumption about linearity is not severely violated.

For the residuals plotted against Interaction-term $\Delta(DtD_{t-1} * D)$ no clear pattern appears in the datapoints, and the residuals do not appear to depend on $\Delta(DtD_{t-1} * D)$. We conclude that the assumption about the model form being linear is not a problem for this variable either. For the residuals plotted against D, we see that the residuals deviate more from zero, when $D = 1$, relative to $D = 0$.

Independence of residuals

We see that the assumption about independence of the residuals is violated, the residuals show significant autocorrelation at lag 4, and we see that the spike of lag 17 is just outside the 95% limit. Therefore we conclude that the the assumption is still violated.

Homoscedasticity

After including the interaction-term and the time dummy the residuals still seem to depend on the fitted values and tend to increase in absolute terms when the fitted values deviate from zero. But it is noted that the assumption is not severely violated.

Normal distribution of residuals

The residuals are relatively close to resemble a normal distribution, with a high concentration of the mass around the mean 0. The distribution does not resemble a normal distribution perfectly, but close enough not to be a concern.

The regression model results for $\Delta S_t = \beta_0 + \beta_1 \Delta DtD_{t-1} + \beta_2 \Delta(DtD_{t-1} * D) + \beta_3 D + \epsilon_t$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	-2.4860	6.4880	-0.3830	0.7028
ΔDtD_{t-1}	β_1	-185.558	19.3950	-9.5670	3.73E-14
$\Delta(DtD_{t-1} * D)$	β_2	42.9230	15.0870	2.8450	0.0059
D	β_3	6.6320	14.1850	0.4680	0.6417
R^2	0.6059				
$Adjusted\ R^2$	0.5883				

R-square

Including the two variables, make the adjusted R-square increase a bit. The adjusted R-Square of the former regression model was 0.5516, whereas the new regression model yields an adjusted R-Square of 0.5833.

Slope analysis for variable ΔDtD_{t-1}, β_1

The estimated β_1 coefficient is -185.558, hence the slope has increased compared to the former model. The t-test stat is -9.567 yielding p-value of 3.73e-14, thus we conclude that beta coefficient is significant at a 1% level significance level.

The interaction-term analysis for $\Delta(DtD_{t-1} * D)$, β_2

The estimated beta coefficient, β_2, has a value of 42.923, a t-test stat value equal to 2.845 which yields a p-value of 0.0059. Hence, we reject that the β_2 is equal to zero at 1% significance level. The estimated beta is positive as expected. The result indicates that the relationship between the ΔS_t and the ΔDtD_{t-1} depends on the state of the economy.

β_2 estimates the change in the beta coefficient for ΔDtD_{t-1} when in crisis. When not in crisis ΔS_t changes by β_1 when ΔDtD_t changes by one unit, all else equal. When in crisis ΔS_t changes by $(\beta_1 + \beta_2)$ when ΔDtD_t changes by one unit, all else equal.

β_1 is negative and β_2 is positive, hence the ΔS_t becomes less sensitive to ΔDtD_{t-1} when in crisis.

The finding indicates that the unsystematic risk that DtD capture is less important for the CDS spread when in crisis relative to normal times.

We also note that the former regression model yields an estimated coefficient for ΔDtD_{t-1} of -179.904 whereas the new regression model yields an estimated coefficient for ΔDtD_{t-1} of -185.558.

The estimated coefficient for ΔDtD_{t-1} increases (in absolute terms) when we include the interaction-term. The result indicates that the estimated coefficient for ΔDtD_{t-1} in the model where the interaction-term and financial crisis dummy were not included, was positively biased for the non-crisis period (not steep enough), whereas it negatively biased for crisis period (too steep).

Intercept analysis, β_0

The intercept is estimated to a value of -2.486, but is insignificant with a t-test stat of -0.383 and a corresponding p-value of 0.7028.

Intercept financial crisis, D, β_3

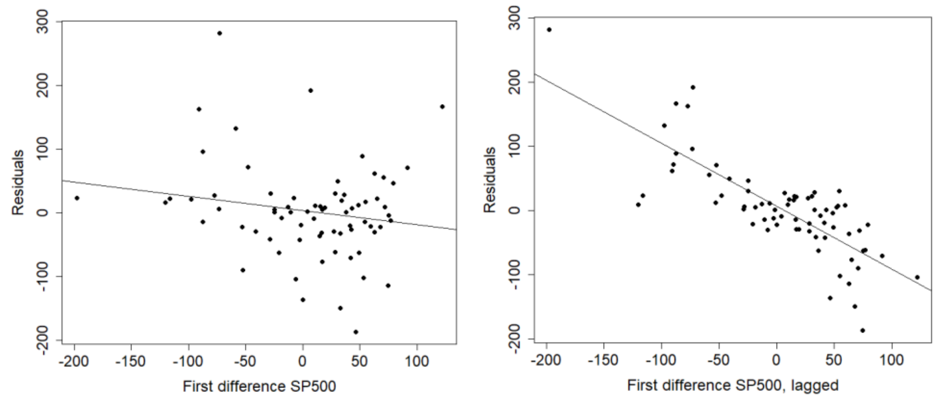
The estimated beta coefficient for the financial crisis dummy variable, β_3 is equal to 6.632, it is insignificant with a t-test stat value of 0.468 with a corresponding p-value of 0.64166. We conclude that the intercept does not change as a function of state of the economy.

10.4 Regression, CDS spread regressed on S&P500 index

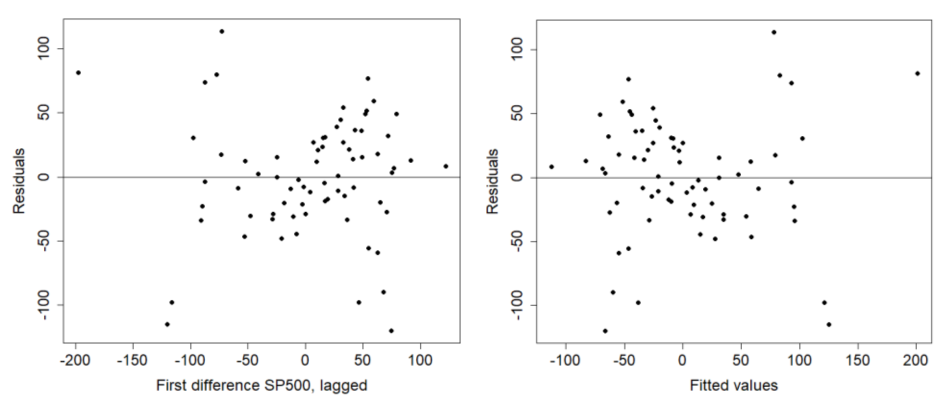
10.4.1 First difference regression, the S&P500 index lagged one period

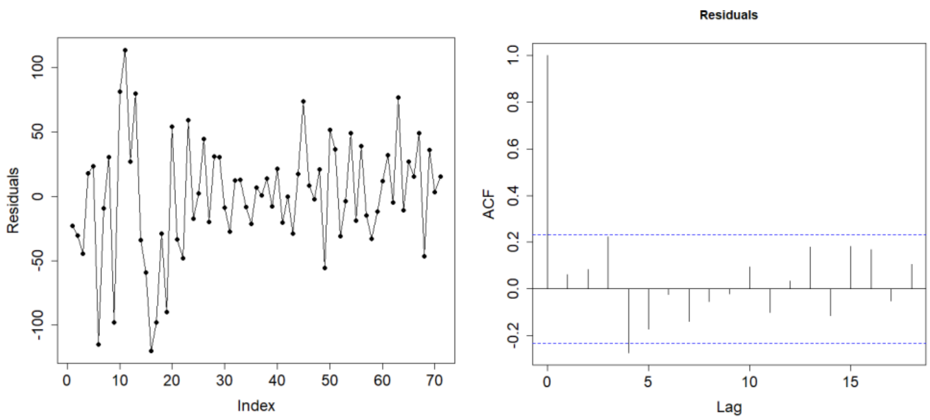
The third question of the problem statement is, "Is it possible to improve the regression model by including non-firm-specific variables such as the S&P 500 index and the volatility measure from the VIX index?". To answer the question, we start out by investigating the relationship between the CDS spread and the SP500 index.

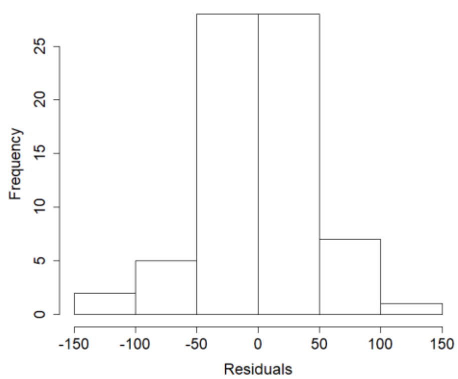
We plot the ΔS_t against $\Delta SP500_t$, and add the least square regression line, we also plot ΔS_t against $\Delta SP500_{t-1}$, and add the corresponding least square regression line.



For ΔS_t against the $\Delta SP500_t$ the relationship appears weak, whereas for ΔS_t against $\Delta SP500_{t-1}$ the relationship appears much stronger. We continue the analysis by running the following regression: $\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \epsilon_t$, but before we go into analyzing the regression model results we go over the assumptions.







Model form

When the residuals are plotted against the explanatory variable $\Delta SP500_{t-1}$ a pattern appears. When $\Delta SP500_{t-1}$ takes on big values, both positive and negative, then the residuals deviate more from zero than they do for small absolute values of $\Delta SP500_{t-1}$. Therefore, we conclude that the residuals are not independent of the $\Delta SP500_{t-1}$, hence model form is not perfect.

$\Delta SP500_{t-1}$ takes on bigger absolute values in the financial crisis relative to normal times. Consequently, it may be solved by including the interaction-term into the model.

Independence of the residuals

From the time series plot of the residuals we conclude that the residuals take on the biggest values, both positive and negative, within the 20 first observations, all observations that fall within the financial crisis. In the time series plot the residuals, however appear to be independent of each other. The autocorrelation function plot shows that the autocorrelation at lag 4 is significant, but just outside the 95% limit. Thus, it might just be by chance. It is noted that the autocorrelation at lag 3 is just on 95% limit.

The bigger values of the residuals for the period defined as the financial crisis may indicate that we can decrease the sum of squared errors by including the interaction-term, that allow the slope of the regression line to differ when in crisis relative to when not in crisis.

Homoscedasticity

To check assumption about homoscedasticity, we plot the residuals against the fitted values estimated from the regression model. The values of the residuals on average tend to increase for bigger absolute values of the fitted values, just as it did in the plot with the residuals plotted against $SP500_{t-1}$. However, the pattern is not that clear, and we conclude that the assumption is not severely violated.

Normality of residuals

From the histogram we conclude that there is no problem with the assumption about normality of the residuals.

The assumptions seem to be met at a reasonable level or at least not severely violated.

Regression model results for $\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \epsilon_t$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	6.9615	5.4674	1.2730	0.2070
$\Delta SP500_{t-1}$	β_1	-0.9800	0.0914	-10.7230	2.42E-16
R^2	0.6250				
$Adjusted\ R^2$	0.6195				

R-square

The model yields a R-Square of 0.625, and an Adjusted R-Square of 0.6195. That is significantly higher than model where we regressed ΔS_t on ΔDtD_{t-1}, indicating that the S&P500 index explains more of the variability in the CDS spread than DtD does.

Slope analysis for variable $\Delta SP500_{t-1}$, β_1

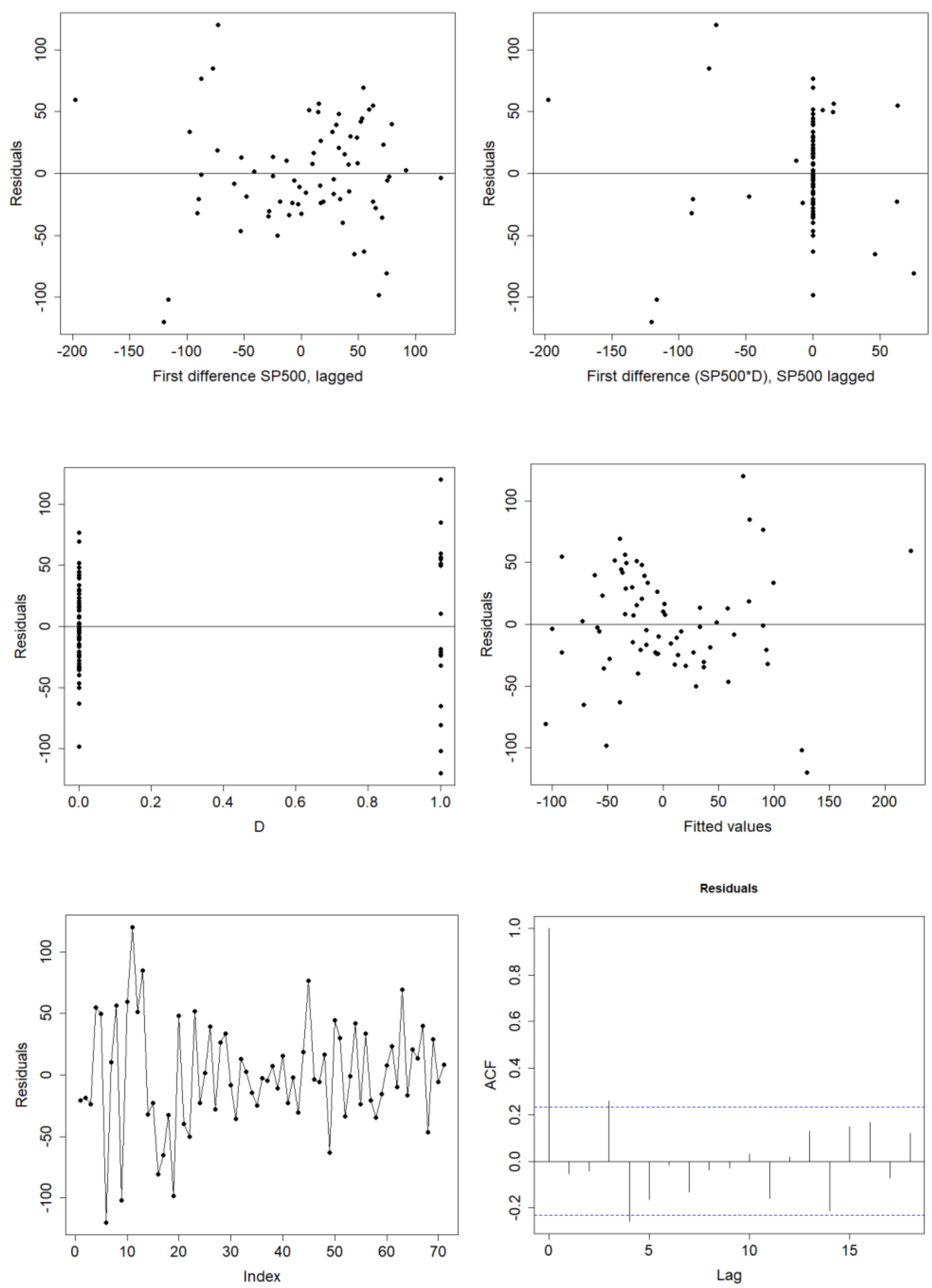
The estimated beta coefficient, β_1, is -0.97991, and has a t-test stat value of -10.723 with a corresponding p-value of 2.24e-16, hence we conclude that the coefficient is significant at a 1% level. The sign of the estimated beta coefficient is negative, just as expected.

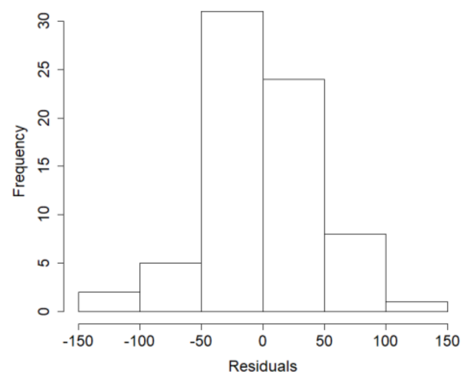
Intercept analysis, β_0

The intercept, β_0, has an estimated value of 6.96148, with a t-test value of 1.273 yielding p-value of 0.207, hence we conclude that it is insignificant.

10.4.2 Adding interaction-term and financial crisis dummy

We set up a regression model that allow for the relationship between ΔS_t and $\Delta S\&P500_{t-1}$ to change during the financial crisis by including the interaction-term and the financial crisis dummy. The regression model has the following form: $\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \beta_2 \Delta(SP500_{t-1} * D) + \beta_3 D + \epsilon_t$





Model form

When plotting the residuals against $\Delta SP500_{t-1}$, then the residuals still deviate more from zero for large absolute values of $\Delta SP500_{t-1}$ relative to smaller absolute values of $\Delta SP500_{t-1}$. The pattern is not as clear as in the former model but including the interaction-term has not solved the problem completely. The assumption about linearity is to some degree violated for the variable $\Delta SP500_{t-1}$.

In the plot where we have the residuals against the $\Delta(SP500_{t-1} * D)$ no clear pattern appears, and we accept the assumption about a linear relationship.

For the residuals plotted against D, we see that the residuals deviate more from zero, when $D = 1$, relative to $D = 0$

Homoscedasticity

The assumption about constant variance for the residuals is to some degree violated. For large absolute values of the fitted values the residuals show greater variance compared to small absolute values.

Independence of residuals

From the residuals against observation number, no autocorrelation in the residuals appears. However, when we look at the autocorrelation function plot, we conclude that we have significant autocorrelation at lag 3 and 4, but the two spikes are just outside 95% limits, so the assumption does not seem to be violated severely.

Normality of residuals

The residuals do to some degree resemble a normal distribution, but it is clear from the histogram that the distribution it positively skewed. But the distribution does not deviate so much from a normal distribution that it is unreasonable to continue as if the assumption has been met.

The regression model results for the model: $\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \beta_2 \Delta(SP500_{t-1} * D) + \beta_3 D + \epsilon_t$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	10.5979	6.4462	1.6440	0.1048
$\Delta SP500_{t-1}$	β_1	-0.9089	0.1262	-7.2000	6.64E-10
$\Delta(SP500_{t-1} * D)$	β_2	-0.2971	0.1923	-1.5450	0.1271
D	β_3	-26.1953	13.4857	-1.9420	0.0563
R^2	0.6526				
$Adjusted\ R^2$	0.6371				

R-Square

Including the financial crisis dummy and the interaction term do increase the adjusted R-Square to 63.71%, but the increase is small.

Slope analysis for variable $\Delta SP500_{t-1}$, β_1

The estimated beta coefficient, β_1, for the variable $\Delta SP500_{t-1}$ is still significant, with a t-test stat value of -7.20 and a corresponding p-value of 6.64e-10. The beta coefficient has increased in value to -0.9089, hence the slope has decreased relative to former regression model, where the interaction-term and the financial crisis dummy were not included.

Interaction-term analysis, variable $\Delta(SP500_{t-1} * D)$, β_2

The estimated beta, β_2, for the interaction-term is -0.2971, but it is insignificant at a 5% and 10% significance level, and we must accept a significance level of 15% to reject that it is insignificant. Although the interaction-term is in insignificant it can make sense to include it in the regression model to function as a control variable that makes β_1 less biased.

If we accept a significance level of 15%, then we can conclude that the regression line becomes steeper during the financial crisis relative to normal times, and it changes by, β_2, hence -0.2971. Meaning when the $\Delta S\&P500_{t-1}$ increases by one unit during the financial crisis then the ΔS_t changes by $(\beta_1 + \beta_2) = (-0.9089 - 0.2971) = -1.206$ whereas during normal times ΔS_t changes by -0.9089 (β_1) when $\Delta S\&P500_{t-1}$ increases by one unit. The interpretation can be that the changes in the CDS spread becomes more sensitive to systematic risk, during stages of emergencies. The finding goes well with the finding that the CDS spread become less sensitive to changes in the un-systematic risk estimate Distance to Default when in crisis relative to normal times.

Intercept, β_0

The estimated beta for the intercept, β_0, takes on the value 10.5979, it is insignificant at a 10% level, with a t-test stat of 1.644 and a corresponding p-value of 0.1048. If we accept a significance level of 15% then we conclude that the beta coefficient is significant.

Intercept financial crisis, β_3

The potential difference in intercept for the financial crisis, β_3, has an estimated value of -26.1953, a t-test stat value of -1.942 with a p-value of 0.0563. β_3 is insignificant at 5% significance level, but do we accept a significance level of 10% then we conclude that it is significant and conclude that the intercept changes from normal times to crisis.

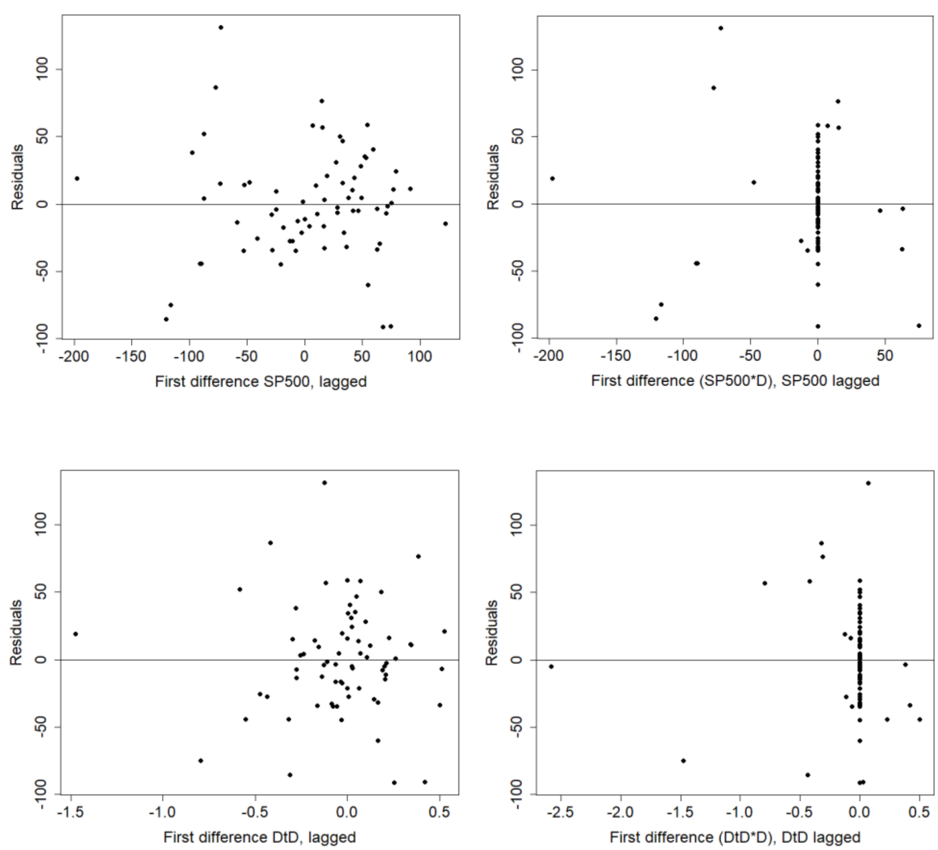
10.5 Regression, CDS spread regressed on DtD and the S&P500 index

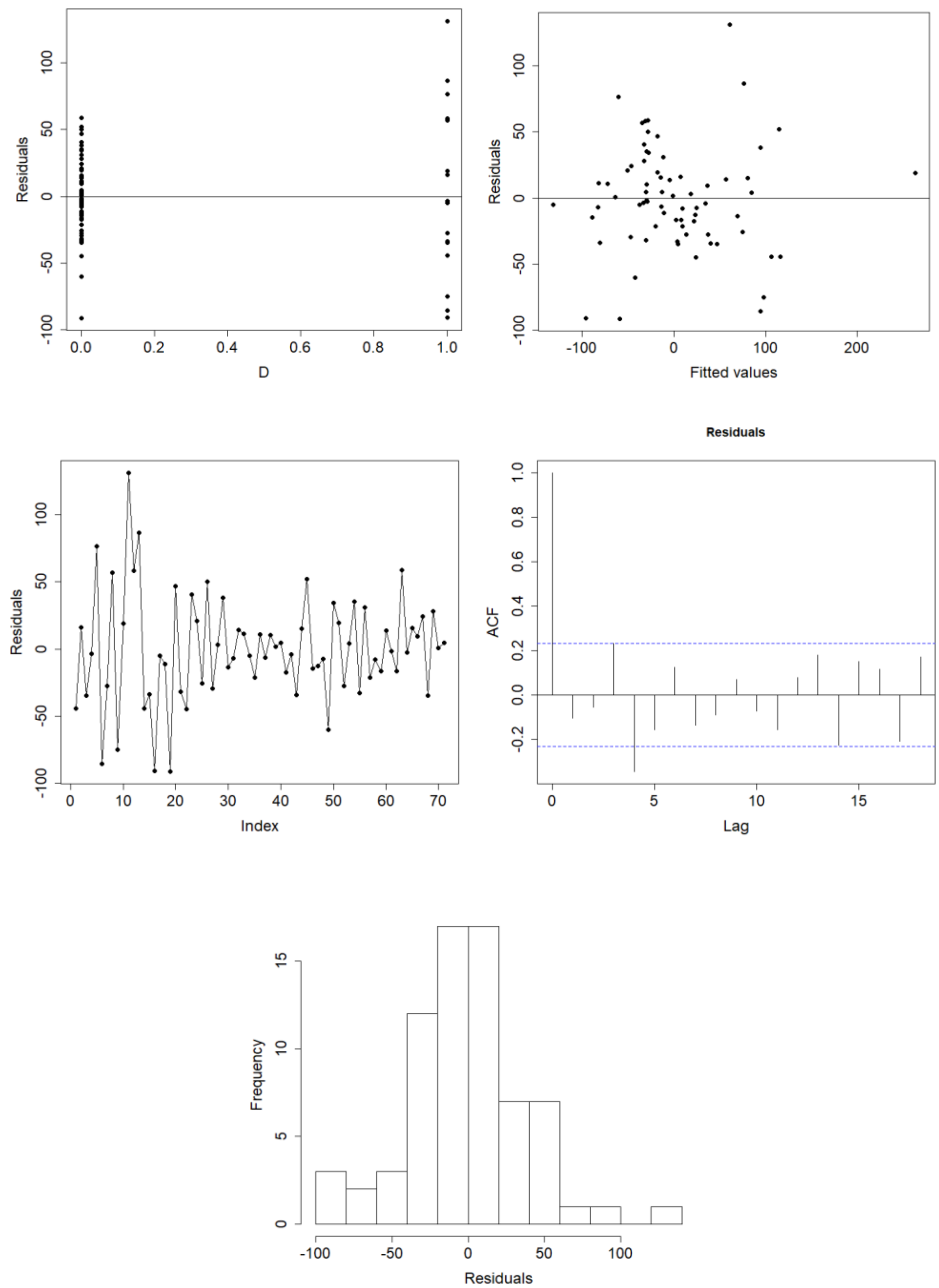
10.5.1 First difference regression, the DtD and S&P500 index lagged one period plus interaction-terms

Now we create a regression model that allows both the firm specific risk and systematic risk to explain ΔS_t, the model is of the fowling form:

$$\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \beta_2 \Delta(SP500_{t-1} * D) + \beta_3 \Delta DtD_{t-1} + \beta_4 \Delta(DtD_{t-1} * D) + \beta_5 D + \epsilon_t$$

Checking the assumptions:





Model from

$\Delta SP500_{t-1}$ and $\Delta(SP500_{t-1} * D)$

The residuals do not appear 100% random when they are plotted against $\Delta SP500_{t-1}$, we still see the same pattern where the residuals tend to increase when $\Delta SP500_{t-1}$ deviates from zero. Hence, we conclude that the assumption of a linear relationship between ΔS_t and $\Delta SP500_{t-1}$ is to some degree violated. When we plot the residuals against $\Delta(SP500_{t-1} * D)$, then they appear random and we conclude that the assumption of linearity is met.

ΔDtD_{t-1} and $\Delta(DtD_{t-1} * D)$

The residuals appear independent of both variables from the two plots, so we conclude that the assumption about a linear relationship between ΔS_t and the two variables is met.

For the residuals plotted against D, we see that the residuals deviate more from zero, when $D = 1$, relative to $D = 0$

Independence of residuals,

Investigation of the autocorrelation function plot shows that there is significant autocorrelation at lag 4 for the residuals, all other spikes are within the 95% or just right at the limit. The assumption about the residuals being 100% independent is not met. However, it is not strongly violated either, as the autocorrelation is around 0.3 at lag 4.

Homoscedasticity

We see that the variance of the residuals tends to increase as the fitted value deviate from zero. Hence including the interaction-term and the financial crisis dummy did not remove the pattern in the residuals, when plotted against the fitted values from the model. However, the pattern is not that clear, and we conclude the assumption about constant variance does not appear to be strongly violated.

Normal distribution

The distribution of the residuals does not resemble a normal distribution perfectly, but the assumption does not appear to be severely violated, hence the normality assumption is not a problem.

The regression model results for:

$$\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \beta_2 \Delta(SP500_{t-1} * D) + \beta_3 \Delta DtD_{t-1} + \beta_4 \Delta(DtD_{t-1} * D) + \beta_5 D + \epsilon_t$$

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	7.3445	5.9984	1.224	0.2252
$\Delta SP500_{t-1}$	β_1	-0.6487	0.141	-4.601	2.00E-05
$\Delta(SP500_{t-1} * D)$	β_2	-0.0875	0.1867	-0.469	0.6409
ΔDtD_{t-1}	β_3	-86.4915	26.7258	-3.236	0.0019
$\Delta(DtD_{t-1} * D)$	β_4	34.7972	13.7787	2.525	0.0140
D	β_5	-12.9736	13.1665	-0.985	0.32810
R^2	0.7164				
$Adjusted\ R^2$	0.6946				

R-Square

The regression model now yields an adjusted R-square of 69.49%, hence including both ΔDtD_{t-1} and $\Delta SP500_{t-1}$ into one model increase the explanatory power of the model, relative to having only one of the variables explain the CDS spread.

Slope analysis for variable $\Delta SP500_{t-1}$, β_1

We see that β_1 is significant at a 1%, with a t-test stat of -4.601 with a corresponding p-value of 2e-5. The estimated β_1 value is -0.6487, that is when $\Delta SP500_{t-1}$ increase by 1 unit then ΔS_t changes by -0.6487. The sign of the estimated beta coefficient is negative as expected. However, we do notice that in this multivariate setting the CDS spread is now less sensitive to S&P500 index, relative to the model where CDS spread is only regressed on the S&P500 and interaction-term.

Interaction-term analysis, variable $\Delta(SP500_{t-1} * D)$, β_2

β_2 is insignificant, with a t-test stat of -0.469 and a corresponding p-value of 0.64091, that is the model indicates that the relationship between the changes in ΔS_t and $\Delta SP500_{t-1}$ does not depend on the state of the economy.

Slope analysis for variable $\Delta DtD_{t-1}, \beta_3$

β_3 is significant at a 1% level, with a t-test stat of 3.236 with a corresponding p-value of 0.00191. The estimated value of β_3 is -86.4915, hence when ΔDtD_{t-1} increase by one unit then ΔS_t changes by -86.4915 basis points. The sign is negative as expected. But just as for $S\&P500$, the sensitivity of the CDS spread to the DtD has decreased, compared to the model where the CDS spread is only explained by the DtD and its interaction-term.

Interaction-term analysis, for variable $\Delta(DtD * D)$, β_4

β_4 is significant at a 5% level, with a t-test stat of -2.525 with a corresponding p-value of 0.014. β_4 has an estimated value of 34.7972, hence we conclude that the relationship between the ΔS_t spread and ΔDtD_{t-1} does change from normal times to crisis. The model results say that ΔS_t becomes less sensitive to the changes in ΔDtD_{t-1} during crisis relative to normal times, because β_4 is significant and has the opposite sign than β_3 has.

The model predicts that ΔS_t change by $(\beta_3 + \beta_4) = -51.6943$ when ΔDtD_{t-1} increase of one unit during the crisis, but when not in crisis ΔS_t changes by $\beta_3 = -86.4915$ when ΔDtD_{t-1} increases by one unit, all else equal. The finding indicates that the unsystematic risk that ΔDtD_{t-1} captures is less important during the financial crisis for the movements in ΔS_t relative to normal times.

Intercept analysis, β_0 and intercept analysis for financial crisis intercept, β_5

Both β_0 and β_5 are insignifincant, with t-test stats values of 1.224 and -0.985 with the corresponding p-values of 0.22521 and 0.32810 respectively. That is, we cannot reject that both beta coefficients are equal to zero at a

1%, 5%, 10% or even 15% significance level. The regression model indicates that the intercept is equal to zero in normal times and that the intercept does not change from zero when in the financial crisis.

Conclusion:

Letting both the $\Delta SP500_{t-1}$ and ΔDtD_{t-1} enter the regression model with both of their interaction-terms plus a dummy for the financial crisis, yields an adjusted R-Square of 0.6949. That is, we have achieved a regression model that explains almost 70% of the variability ΔS_t by only including two variables and their interaction-terms, leaving 30% of the variability to unexplained.

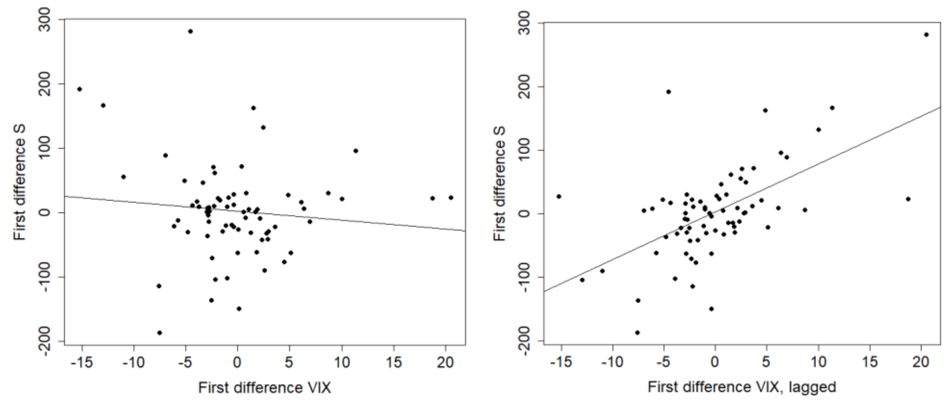
In the regression model, we see that both $\Delta SP500_{t-1}$ and ΔDtD_{t-1} are significant predictors for ΔS_t. We also see that that $\Delta(DtD_{t-1} * D)$ is a significant, hence the sensitivity of ΔS_t to ΔDtD_{t-1} dependents on the state of the economy and changes from non-crisis to crisis. The ΔS_t becomes less sensitive to changes in ΔDtD_{t-1} when in crisis, indicating that the CDS spread is less sensitive to the unsystematic risk captured by the DtD in crisis. The question then becomes, does it mean that the CDS spread becomes more sensitive to systematic risk during the crisis? The results from the regression model does not suggest so, as the estimated beta coefficient for the interaction-term, $\Delta(SP500_{t-1} * D)$, is insignificant, indicating that the relationship between the ΔS_t and $\Delta SP500_{t-1}$ does not change as a function of the state of the economy.

In the next session we examine if there is a relationship between the CDS spread and the systematic risk measure, the VIX index, and if the potential relationship changes as a function of the economy.

10.6 Regression, CDS spread regressed on the VIX index

10.6.1 First difference regression, the VIX index lagged one period

We now plot ΔS_t against ΔVIX_t and add the least square regression line. We also plot ΔS_t against ΔVIX_{t-1} and add the regression line.



From the two plots the relationship between ΔS_t and ΔVIX_{t-1} is much stronger than the relationship between ΔS_t and ΔVix_{t-1}, so we set up the following regression model: $\Delta S_t = \beta_0 + \beta_1 \Delta VIX_{t-1} + \epsilon_t$

Running the regression model: $\Delta S_t = \beta_0 + \beta_1 \Delta Vix_{t-1} + \epsilon_t$ yields the following results:

Term	**coeff.**	**value**	**se of β_j**	**t value**	**P value**
Intercept	β_0	3.2350	7.2180	0.4480	0.6550
$\Delta \mathrm{VIX}_{t-1}$	β_1	7.5100	1.2520	5.9960	8.30E-08
R^2	0.3425				
$Adjusted\ R^2$	0.3330				

R-Square

The simple regression model yields an adjusted R-Square of 0.333. The R-square is lower than the ones the univariate regression models with $\Delta SP500_{t-1}$ and ΔDtD_{t-1} as explanatory variables. They explained about 60% and 55% of the variability in ΔS_t respectably.

Slope analysis for variable ΔVIX_{t-1}, β_1

The estimated beta, β_1 is equal to 7.510, with a t-test stat value of 5.996 with a corresponding p-value of 8.3e-08, hence we reject that β_1 is equal to zero at a 1% significance level, and conclude that ΔVix_{t-1} is a significant predictor of ΔS_t

Intercept analysis, β_0

We see that the intercept is insignificant at a 1%, 5%, 10% and even 15% significance level, with a t-test stat value of 0.448 yielding a p-value of 0.655.

10.6.2 Adding interaction term and financial crisis dummy

To test if the potential relationship between ΔS_t spread and ΔVIX_{t-1} is a function of the state of the economy we include the financial crisis dummy and the interaction-term, yielding the following regression model:

$\Delta S_t = \beta_0 + \beta_1 \Delta Vix_{t-1} + \beta_2 \Delta (VIX_{t-1} * D) + \beta_3 D$

The regression model results are:

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	-2.3900	8.3136	-0.2870	0.7750
ΔVIX_{t-1}	β_1	7.6639	1.5685	4.8860	6.72E-06
$\Delta(VIX_{t-1} * D)$	β_2	-0.4195	1.6671	-0.2520	0.8020
D	β_3	23.0171	16.9702	1.3560	0.1800
R^2	0.3604				
$Adjusted\ R^2$	0.3318				

R-Square

Including both the financial crisis dummy and the interaction-term does not improve the model, the adjusted R-square has decreased to 0.3318, making it clear that the two new variables only add noise to the model.

Slope analysis for variable ΔVIX_{t-1}, β_1

β_1 does barely change by adding the two variables $\Delta(Vix_{t-1} * D)$ and D to the model, hence they do not serve a purpose as control variables. β_1 is still significant with a t-test value of 4.886 yielding a p-value of 6.72e-06.

Interaction-term analysis, $\Delta(VIX_{t-1} * D)$, β_2

β_2 is insignificant with a t-test stat of -0.252 and a corresponding p-value of 0.802, hence the relationship between the ΔS_t and ΔVIX_{t-1} does not changes as a function of the state of the economy.

Intercept analysis, β_0, and intercept analysis for financial crisis intercept, β_3

The estimated intercept, β_0, is insignificant with a t-test stat value of -0.287 yielding a p-value of 0.775. The change in the intercept, β_3, has a t-test stat value of 1.356 with a corresponding p-value of 0.180, hence it is concluded that the changes in intercept from normal times to crisis is insignificant.

10.7 Regression, CDS spread regressed on DtD and the VIX index

10.7.1 First difference regression, the DtD and VIX index lagged one period plus interaction-terms

We now add the ΔVIX_{t-1} to the DtD regression model to see if it adds any value, we run the following model:

$$\Delta S_t = \beta_0 + \beta_1 \Delta VIX_{t-1} + \beta_2 \Delta DtD_{t-1} + \beta_3 \Delta(DtD_{t-1} * D) + \beta_4 D + \epsilon_t$$

Regression model results:

Term	coeff.	value	se of β_j	t value	P value
Intercept	β_0	-1.9990	6.5300	-0.3060	0.7605
ΔVIX_{t-1}	β_1	1.1430	1.4470	0.7900	0.4323
ΔDtD_{t-1}	β_2	-170.1960	28.2000	-6.0350	8.05E-08
$\Delta(DtD_{t-1} * D)$	β_3	43.7370	15.8740	2.7550	0.0076
D	β_4	7.2490	14.8440	0.4880	0.6269
R^2	0.6090				
$Adjusted\ R^2$	0.5853				

R-Square

Adding the ΔVIX_{t-1} to the regression model does not improve explanatory power of the model. The model yields an adjusted R-Square of 0.5853 , compared to the regression model of the form:

$\Delta S_t = \beta_0 + \beta_1 \Delta DtD_{t-1} + \beta_2 \Delta(DtD_{t-1} * D) + \beta_3 D$ which yields an adjusted R-Square of 0.5883.

Slope analysis for variable ΔVIX_{t-1}, β_1

β_1 is insignificant, with a t-test stat value of 0.790 with a corresponding p-value of 0.43230.

Slope analysis for variable ΔDtD_{t-1}, β_2

β_2 is significant with a t-test stat value of -6.035 and a p-value if 8.05e-08, it has an estimated value of -170.196.

Interaction-term analysis for $\Delta(DtD_{t-1} * D)$, β_3

β_3 is significant with a t-test stat value of 2.755 and a p-value if 0.00757, it has an estimated value of 43.737

Intercept analysis, β_0, and intercept analysis for financial crisis intercept, β_4

Both β_0 and β_4 are insignificant with t-test stat values of -0.306 and 0.488 with corresponding p-values of 0.76063 and 0.62692

Over all, we conclude that it does not add value to include ΔVIX_{t-1} into the regression. We conclude that VIX does not capture variability in the CDS spread that is not captured by the DtD. Hence, there is no reason to check if the assumptions are violated.

11. Discussion

11.1 Implications to practice

The results of the regression models in the thesis show that it matters if one picks the DtD or PD when one wants to predict the CDS spread in an OLS regression setting. The results indicate that it is easier to capture the variability in the first difference of the CDS spread with DtD rather than the PD, although the two measures contain the exact same information. The regression model results also show that if one is interested in capturing the change in the relationship between the CDS spread the risk measure produced by the Merton DD model for an economy in crisis, then one must use DtD rather than PD. The results of the regression models also show, that one can explain a higher fraction of the variability in the first difference of the CDS spread by the first difference in the S&P500 index, than one can by the first difference in the DtD. This is relevant for practice where the first difference of the S&P500 index is easily obtained whereas much work must be done to calculate the DtD.

11.2 Possible to improve the model?

For several of the regressions run in the thesis there are some pattern left in the residuals, which can affect the validity of F-test, t-tests and the R-square. One way we could have attempted solve the violation of independence of residuals, is to allow for more time-related features of our data, in addition to the interaction-terms and financial crisis dummy.

We could for example have made use of seasonal dummy variables. We have monthly data, so we could use 11 dummy variables, one dummy variable for 11 months out the 12 months in a year, using the general rule of use (x-1) dummy variables to denote x different periods (Makridakis et al., 1998). However, we do note that there is no guarantee that this will solve the problem.

Another way to deal with the problem could have been running regressions with ARIMA errors after we have detected that there was some pattern in the residuals. However, it would require the use of more complicated methods of estimation such as the Maximum likelihood estimation instead of the simple OLS regression estimation (Makridakis et al., 1998).

12. Conclusion

The thesis addresses four questions. The first question is "how well does the Merton DD model explain the variability in the CDS spread in a linear regression setup?" First it was shown that the Merton DD model offers two measures for default risk, distance to default (DtD) and default probability (PD). The two measures contain the exact same information but express it differently. The two measures are linked together by $PD = N(-DtD)$.

Question one was answered by running regression models where the CDS spread (S) was regressed on the two risk measures from the Merton DD model. The first regression model took the form $\log(S_t) = \beta_0 + \beta_1 \log(PD_t) + \epsilon_t$, the model is similar to the one run in Bharath and Shumway (2008). Their regression model yielded a R-square 26% and the sign of the coefficient was positive. In our model the relationship was positive and significant and yielded an adjusted R-square of 74.16%. At first it indicates that the PD estimate from the Merton DD model explains 74,16% of the variability in the CDS spread. Checking the assumptions, however, made it clear that there was an imminent risk of a spurious regression.

To meet the challenge of the potential non-stationarity first difference of both variables was taken and a plot analysis of the two variables indicated that it was not too unreasonable to assume that both variables were stationary. The new regression model took the from $\Delta\log(S_t) = \beta_0 + \beta_1 \Delta\log(PD_t) + \epsilon_t$.

The new model yielded an adjusted R-square of -0.0003 and beta coefficient for the default probability was insignificant. The result indicates that the $\Delta\log(PD_t)$ has zero explanatory power on the $\Delta\log(S_t)$.
Lagging $\Delta\log(PD_t)$ by one period and running the regression model of the form $\Delta\log(S_t) = \beta_0 + \beta_1\Delta\log(PD_{t-1}) + \epsilon_t$, however, yields an adjusted R-square of 50.56% and the coefficient to $\Delta\log(PD_t)$ is positive and significant.
The result indicates that the changes in first difference of CDS-spread do not react immediately to changes in the first difference of the default probability. The answer to the first question the thesis addresses at this point is that the changes in the first difference of $\log(PD)$ is a significant predictor of the changes in the first difference of $\log(S)$, when $\log(PD)$ is lagged by one period. The model explains 50.56% of the variability in $\log(S)$. But when not lagged then the changes in first difference of $\log(PD)$ explains essentially zero % of the variability in $\log(S)$.
The second question the thesis addresses is "Does the potential relationship between the Merton DD model and the CDS spread depend on the state of the economy?" It was investigated by including a financial crisis dummy and an interaction-term. The regression model results show that the relationship between $\Delta\log(PD_{t-1})$ and $\Delta\log(S_t)$ did not change as a function of the state of the economy at a 10% significance level, however if we accept a significance level of 15% then $\Delta\log(S_t)$ does become less sensitive to the changes in $\Delta\log(PD_{t-1})$ when in crisis relative normal times. The regression model of the form $\Delta\log(S)_t = \beta_0 + \beta_1\Delta\log(PD)_{t-1} + \beta_2 D + \beta_3\Delta\log(PD_{t-1} * D) + \epsilon_t$ yields an adjusted R-square of 50.92%.

Question one and two addressed by the thesis were also answered using the risk measure distance to default from the Merton DD model. For ΔDtD_t to be a significant predictor of the ΔS_t, it had to be lagged by one period just as PD. The regression model of the form: $\Delta S_t = \beta_0 + \beta_1\Delta DtD_{t-1} + \beta_2\Delta(DtD_{t-1} * D) + \beta_3 D + \epsilon_i$ yields an adjusted R-square of 58.83%. Both β_1 and β_2 are significant at a 1% significance level, β_1 is negative and β_2 is positive. The model result shows that ΔDtD_{t-1} is a significant predictor of ΔS_t and that the relationship between the two changes as a function of the economy. β_2 being positive indicates that the changes in CDS spread becomes less sensitive to the changes in DtD lagged one period when in crisis relative to normal times.
The third question addressed in the thesis is "Is it possible to improve the regression model by including non-firm-specific variables such as the return on the S&P 500 index and the volatility measure from the VIX index?" and the forth question is: "Does the potential relationship between the non-firm-specific variables and the CDS spread change depending on the state of the economy?"

A plot analysis of the potential relationship between ΔS and $\Delta SP500$ indicates that changes in the first difference of CDS spread take some time to react to the changes in S&P500 index, just as it did for the default probability and the distance to default. Therefore, to answer the third question, we set up a univariate regression model of the form $\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \epsilon_t$. The model yields an adjusted R-square of 61.95%, and β_1 is negative and significant at a 1% level. Adding the interaction-term and the financial crisis dummy to the regression form the model: $\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \beta_2 \Delta(SP500_{t-1} * D) + \beta_3 D + \epsilon_t$. The model has an adjusted R-square of 63.71, hence only a slight increase compared to the simple model. Both β_2 and β_3 are insignificant at 5% level, at a 10% β_3 becomes significant whereas β_2 is still insignificant. The results show that the relationship between the CDS spread and the S&P500 does not change as function of the state of the economy.

We do the same analysis for the VIX index and we conclude that we must lag VIX by one period to have a significant relationship between VIX index and CDS spread. We run the model $\Delta S_t = \beta_0 + \beta_1 \Delta VIX_{t-1} + \beta_2 \Delta(VIX_{t-1} * D) + \beta_3 D$, and only β_1 is significant and we conclude that the relationship between the first difference of the VIX index and the first difference CDS spread does not change as a function of the state of the economy. The model explains 33% of the variability in ΔS_t, hence significantly lower than what the $\Delta SP500_{t-1}$ and the ΔDtD_{t-1} explains individually.

When we run $\Delta S_t = \beta_0 + \beta_1 \Delta VIX_{t-1} + \beta_2 \Delta DtD_{t-1} + \beta_3 \Delta(DtD_{t-1} * D) + \beta_4 D + \epsilon_t$, then β_1 becomes insignificant, whereas β_2 and β_3 are still significant at a 1% level, and the model has an adjusted R-square of 58.53%, hence the model does not improve by including VIX.

When we set up the regression model $\Delta S_t = \beta_0 + \beta_1 \Delta SP500_{t-1} + \beta_2 \Delta(SP500_{t-1} * D) + \beta_3 \Delta DtD_{t-1} + \beta_4 \Delta(DtD_{t-1} * D) + \beta_5 D + \epsilon_t$ then we see an increase in the adjusted R-square, the model now explains 69.46% of the variability in ΔS_t, the model results show that only β_1, β_3 and β_4 are significant in the final model. The overall conclusion is that the relationship between ΔS_t and ΔDtD_{t-1} is significant and does depend on the state of the economy. Adding $\Delta SP500_{t-1}$ to the regression does improve the model, but the relationship between the changes in the first difference of CDS spread and the changes in first difference of S&P500 index lagged one period does not changes from normal times to crisis.

14. Appendix

Table 1, data for the period December 2007 to November 2013, on monthly basis, start of the month values. Equity and debt are million USD, the CDS spread is in basis points.

Month/obs	Equity	Debt	Rf rate%	Recovery	Rating	CDS spread
1	30999.809	4812.5	4.29	0.4	AA	47.22
2	27378.732	5277.0	3.33	0.4	AA	48.38
3	30245.739	5277.0	2.69	0.4	AA	120.16
4	29366.218	5277.0	2.41	0.4	AA	143.49
5	28348.587	5246.0	2.71	0.4	AA	113.39
6	33084.219	5246.0	2.88	0.4	AA	76.74
7	29033.256	5246.0	3.22	0.4	AA	92.72
8	27452.081	5944.0	3.09	0.4	AA	102.22
9	26134.381	5944.0	3.03	0.4	AA	112.17
10	18366.459	5944.0	3.03	0.4	AA	134.46
11	9203.646	6323.5	2.86	0.4	AA	157.23
12	8611.411	6323.5	2.11	0.4	AA	439.29
13	9011.569	6323.5	1.62	0.4	AA	631.25
14	6234.469	5651.5	1.15	0.4	AA	658.33
15	4995.058	5651.5	1.33	0.4	AA	821.09
16	6975.327	5651.5	1.36	0.4	AA	882.86
17	8836.683	5571.5	1.19	0.4	AA	768.96
18	8982.825	5571.5	0.94	0.4	AA	582.12
19	10064.271	5571.5	0.95	0.4	A	445.55
20	11458.615	5539.5	0.78	0.4	AA	423.54
21	11741.183	5539.5	0.77	0.4	A	273.66
22	12783.761	5539.5	0.61	0.4	A	302.29
23	12101.775	5332.0	0.62	0.4	AA	240.11
24	12199.213	5332.0	0.50	0.4	BB	219.32
25	15706.973	5332.0	0.57	0.4	A	227.40
26	12403.845	5294.5	0.53	0.4	BB	197.73
27	13569.684	5294.5	0.51	0.4	A	247.51
28	14528.744	5294.5	0.53	0.4	A	269.33
29	13712.299	5659.5	0.59	0.4	A	192.76
30	11884.673	5659.5	0.78	0.4	AA	213.92

Month/obs	Equity	Debt	Rf rate%	Recovery	Rating	CDS spread
31	10271.461	5659.5	0.78	0.4	A	346.55
32	11406.960	4814.0	0.60	0.4	A	402.30
33	10434.254	4814.0	0.46	0.4	A	312.52
34	12366.902	4814.0	0.44	0.4	AA	383.28
35	13421.748	4744.0	0.38	0.4	A	313.04
36	13406.426	4744.0	0.44	0.4	A	270.66
37	15719.992	4744.0	0.50	0.4	AA	258.68
38	16934.971	5043.5	0.46	0.4	A	196.85
39	17906.310	5043.5	0.47	0.4	AA	176.87
40	18767.088	5043.5	0.44	0.4	AA	157.30
41	18083.155	4961.5	0.41	0.4	A	158.06
42	17845.143	4961.5	0.37	0.4	AA	149.58
43	16836.643	4961.5	0.39	0.4	AA	154.21
44	15675.416	4982.0	0.43	0.4	AA	184.81
45	13621.542	4982.0	0.44	0.4	AA	190.47
46	10184.231	4982.0	0.49	0.4	A	286.57
47	11451.911	5051.0	0.57	0.4	BB	452.97
48	10664.326	5051.0	0.66	0.4	AA	348.94
49	9206.230	5051.0	0.67	0.4	AA	359.68
50	10814.426	5434.0	0.58	0.4	A	376.97
51	10842.318	5434.0	0.50	0.4	AA	274.61
52	10682.282	5434.0	0.50	0.4	A	281.32
53	10379.020	5268.5	0.52	0.4	AA	282.78
54	9120.311	5268.5	0.55	0.4	AA	269.11
55	9333.651	5268.5	0.52	0.4	AA	358.02
56	9036.575	5349.0	0.45	0.4	A	363.42
57	9132.596	5349.0	0.42	0.4	BB	334.67
58	9447.329	5349.0	0.36	0.4	A	353.84
59	9145.921	4673.5	0.33	0.4	A	312.66
60	8975.169	4673.5	0.33	0.4	A	314.28

Month/obs	Equity	Debt	Rf rate%	Recovery	Rating	CDS spread
61	9263.313	4673.5	0.33	0.4	AA	305.49
62	9434.154	5052.5	0.33	0.4	A	314.69
63	9110.658	5052.5	0.33	0.4	BB	283.28
64	9110.658	5052.5	0.33	0.4	A	269.13
65	9089.875	4509.0	0.32	0.4	A	299.70
66	9089.875	4509.0	0.31	0.4	AA	267.84
67	8362.685	4509.0	0.35	0.4	AA	269.15
68	8502.930	4529.0	0.35	0.4	AA	315.47
69	8235.543	4529.0	0.34	0.4	AA	293.66
70	8684.754	4529.0	0.33	0.4	BB	305.54
71	9915.211	4515.5	0.31	0.4	A	301.06
72	10278.875	4515.5	0.28	0.4	A	238.01
73	11369.869	4515.5	0.29	0.4	A	211.97

Table 2, distance to default (DtD), probability of default (PD), asset values in million USD.

Month/obs	DtD	PD	Asset
1	5.7	0.0000%	35610.219
2	5.1	0.0000%	32482.902
3	5.3	0.0000%	35382.68
4	5.3	0.0000%	34517.563
5	5.2	0.0000%	33454.329
6	5.6	0.0000%	38181.289
7	5.3	0.0000%	34113.026
8	4.8	0.0001%	33215.22
9	4.7	0.0001%	31900.979
10	3.9	0.0046%	24133.037
11	2.4	0.7404%	15344.127
12	2.3	1.0395%	14795.975
13	2.4	0.8658%	15227.773
14	2.0	2.4975%	11804.881
15	1.6	5.0421%	10535.499
16	2.1	1.6221%	12540.324
17	2.6	0.5272%	14339.349
18	2.6	0.4911%	14499.482
19	2.8	0.2653%	15581.697
20	3.0	0.1181%	16954.491
21	3.1	0.1009%	17237.699
22	3.3	0.0573%	18289.302
23	3.2	0.0643%	17400.524
24	3.2	0.0614%	17504.339
25	3.8	0.0087%	21008.632
26	3.3	0.0518%	17670.126
27	3.5	0.0268%	18837.135
28	3.6	0.0156%	19795.192
29	3.3	0.0405%	19338.315
30	3.1	0.1072%	17499.663

Month/obs	DtD	DD	Asset
31	2.8	0.2608%	15885.595
32	3.3	0.0475%	16191.97
33	3.1	0.0889%	15225.785
34	3.5	0.0264%	17159.663
35	3.7	0.0123%	18147.709
36	3.7	0.0124%	18129.553
37	4.0	0.0031%	20440.321
38	4.0	0.0027%	21955.314
39	4.2	0.0016%	22926.155
40	4.3	0.0010%	23788.442
41	4.2	0.0012%	23024.35
42	4.2	0.0014%	22788.314
43	4.1	0.0025%	21778.822
44	3.9	0.0049%	20636.021
45	3.6	0.0162%	18581.606
46	3.0	0.1297%	15141.298
47	3.2	0.0650%	16473.92
48	3.1	0.1043%	15681.633
49	2.8	0.2562%	14222.28
50	3.0	0.1526%	16216.246
51	3.0	0.1509%	16248.47
52	2.9	0.1654%	16088.357
53	2.9	0.1630%	15619.41
54	2.7	0.3440%	14358.16
55	2.7	0.3033%	14573.295
56	2.7	0.3946%	14359.492
57	2.7	0.3734%	14457.229
58	2.7	0.3113%	14775.506
59	2.9	0.1718%	13803.286
60	2.9	0.1924%	13632.438

Month/obs	DtD	PD	Asset
61	3.0	0.159%	13920.737
62	2.8	0.227%	14468.931
63	2.8	0.278%	14145.175
64	2.8	0.278%	14145.175
65	3.0	0.143%	13583.883
66	3.0	0.143%	13584.332
67	2.8	0.236%	12854.929
68	2.8	0.220%	13015.175
69	2.8	0.265%	12748.033
70	2.9	0.194%	13198.017
71	3.1	0.082%	14416.409
72	3.2	0.065%	14781.499
73	3.4	0.031%	15872.177

Table 3, the S&P500 index level and the VIX index level, start of the month value.

Month/obs	S&P 500	Vix
1	1468.36	22.50
2	1378.55	26.20
3	1330.63	26.54
4	1322.70	25.61
5	1385.59	20.79
6	1400.38	17.83
7	1280.00	23.95
8	1267.38	22.94
9	1282.83	20.65
10	1166.36	39.39
11	968.75	59.89
12	896.24	55.28
13	903.25	40.00
14	825.88	44.84
15	735.09	46.35
16	797.87	44.14
17	872.81	36.50
18	919.14	28.92
19	919.32	26.35
20	987.48	25.92
21	1020.62	26.01
22	1057.08	25.61
23	1036.19	30.69
24	1095.63	24.51
25	1115.10	21.68
26	1073.87	24.62
27	1104.49	19.50
28	1169.43	17.59
29	1186.69	22.05
30	1089.41	32.07

Month/obs	S&P 500	Vix
31	1030.71	34.54
32	1101.60	23.50
33	1049.33	26.05
34	1141.20	23.70
35	1183.26	21.20
36	1180.55	23.54
37	1257.64	17.75
38	1286.12	19.53
39	1327.22	18.35
40	1325.83	17.74
41	1363.61	14.75
42	1345.20	15.45
43	1320.64	16.52
44	1292.28	25.25
45	1218.89	31.62
46	1131.42	42.96
47	1253.30	29.96
48	1246.96	27.80
49	1257.60	23.40
50	1312.41	19.44
51	1365.68	18.43
52	1408.47	15.50
53	1397.91	17.15
54	1310.33	24.06
55	1362.16	17.08
56	1379.32	18.93
57	1406.58	17.47
58	1440.67	15.73
59	1412.16	18.60
60	1416.18	15.87

Month/obs	S&P 500	Vix
61	1426.19	18.02
62	1498.11	14.28
63	1514.68	15.51
64	1569.19	12.70
65	1597.57	13.52
66	1630.74	16.30
67	1606.28	16.86
68	1685.73	13.45
69	1632.97	17.01
70	1681.55	16.60
71	1756.54	13.75
72	1805.81	13.70
73	1848.36	13.72

14. References

Augustin, P., Subrahmanyam, M. G., Tang, D. Y., & Wang, S. Q. (2014). Credit Default Swaps: A Survey. *Foundations and Trends® in Finance*, *9*(1–2), 1–196. https://doi.org/10.1561/0500000040

Berk, J. B., & DeMarzo, P. M. (2013). *Corporate Finance, Global Edition* (Third edit). United Kingdom: Pearson Education M.U.A.

Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies*, *21*(3), 1339–1369. https://doi.org/10.1093/rfs/hhn044

Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, *81*(3), 637–654. https://doi.org/10.1086/260062

Blanco, R., Brennan, S., & Marsh, I. W. (2005). An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps. *Journal of Finance*, *60*(5), 2255–2281. https://doi.org/10.1111/j.1540-6261.2005.00798.x

Brooks, C. (2008). *Introductory econometrics for finance* (2. ed.). Cambridge.

Cesare, A. Di, & Guazzarotti, G. (2010). An analysis of the determinants of credit default swap spread changes before and during the subprime financial turmoil. *Temi Di Discussione*, *749*, 45.

Chen, L., Lesmond, D. A., & Wei, J. (2007). Corporate Yield Spreads and Bond Liquidity. *Journal of Finance*, *62*(1), 119–149. https://doi.org/10.1111/j.1540-6261.2007.01203.x

Cohen, L., & Frazzini, A. (2008). Economic Links and Predictable Returns. *Journal of Finance*, *63*(4), 1977–2011. https://doi.org/10.1111/j.1540-6261.2008.01379.x

Crosbie, P., & Bohn, J. (2003). Modeling default risk: A new structural approach. *Moody's KMV*, 165–172. https://doi.org/10.1016/j.frl.2006.03.003

Dick-Nielsen, J. (n.d.). My Estimation. *Lecture Note*, 1–2.

Eom, Y. H., Helwege, J., & Huang, J.-Z. (2004a). Structural Models of Corporate Bond Pricing: An Empirical Analysis. *The Review of Financial Studies*, *17*(2), 499–544. https://doi.org/10.1093/rfs/hhg053

Eom, Y. H., Helwege, J., & Huang, J. Z. (2004b). Structural models of corporate bond pricing: An empirical analysis. *Review of Financial Studies*, *17*(2), 499–544. https://doi.org/10.1093/rfs/hhg053

Ericsson, J., Jacobs, K., & Oviedo, R. (2010). The Determinants of Credit Default Swap Spread. *Journal of Financial & Quantitative Analysis*, *44*(1), 109–132. Retrieved from http://bora.nhh.no/bitstream/2330/2939/1/Eliassen 2010.pdf

Feldhütter, P., & Lando, D. (2008). Decomposing swap spreads. *Journal of Financial Economics / Vol. 88, Nr. 2*.

Gündüz, Y., & Uhrig-Homburg, M. (2014). Does modeling framework matter? A comparative study of structural and reduced-form models. *Review of Derivatives Research*, *17*(1), 39–78. https://doi.org/10.1007/s11147-013-9090-8

Huang, J.-Z., & Huang, M. (2012). How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk? *Review of Asset Pricing Studies*, *2*(2), 153–202. https://doi.org/10.1093/rapstu/ras011

Huang, J.-Z., & Zhou, H. (2008). Specification Analysis of Structural Credit Risk Models. *Finance and Economics Discussion Series*, (2008–55).

Hull, J. C. (2012). *Options, futures, and other derivatives* (8. ed.). Harlow.

Jessen, C., & Lando, D. (2015). Robustness of distance-to-default. *Journal of Banking and Finance*, *50*, 493–505. https://doi.org/10.1016/j.jbankfin.2014.05.016

Lando, D. (2004). *Credit risk modeling: Theory and applications*. Princeton, NJ .

Longstaff, F. A., Mithal, S., & Neis, E. (2005). Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. *Journal of Finance*, *60*(5), 2213–2253. https://doi.org/10.1111/j.1540-6261.2005.00797.x

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications* (3. ed). New York, USA.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance*, *29*(2), 449–470.

Research, N. B. of economic. (2010). US Business Cycle Expansions and Contractions. *October*, *1938*(October 1945), 1–2. Retrieved from http://www.nber.org/cycles.html

Stock, J. H., & Watson, M. W. (2012). *Introduction to econometrics* (3. ed.). Harlow.

Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance*, *59*(2), 831–868.

Zhang, B. Y., Zhou, H., & Zhu, H. (2009). Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms. *The Review of Financial Studies*, *22*(12), 5099–5131. https://doi.org/10.1093/rfs/hhp004

Zhu, H. (2006). An Empirical Comparison of Credit Spreads between the Bond Market and the Credit Default Swap Market. *Journal of Financial Services Research*, *29*(3), 211–235. https://doi.org/10.1007/s10693-006-7626-x