

# What Drives the Difference in Probability of Default from Reduced Form- and Structural Approaches

Lund Institute of Technology

Madelene Larsson & Albin Ädel

May 30, 2014

## Abstract

This Master Thesis successfully explains the difference in probability of default implied by Credit Default Swaps, traded by the market, and the benchmark Moody's EDF<sup>TM</sup>. The difference is explained by the market price of risk, related to the Girsanov kernel, allowing us to transform the risk neutral measure  $\mathbb{Q}$  to the physical measure  $\mathbb{P}$ . This market price of risk is modeled with a log-linear multivariate regression model combined with elastic net, using market data. The predictability of the model is examined. The market price of risk is seen to be mostly dependent on market sentiment, in front of firm specific factors and liquidity. The analysis is made for AB Volvo, Stora Enso Oyj and TeliaSonera AB on data from 2006 - 2014. The work was carried out at Swedbank.

Keywords: Credit Default Swaps, CDS, probability of default, reduced form model, market price of risk, risk neutral measure, physical measure, elastic net.



## **Acknowledgements**

We would like to thank Erik Lindström at Lund Institute of Technology for valuable insights and guidance throughout this thesis. We would also like to thank Krister Sundholm and Babak Soltani at Swedbank for their time and patience and for giving us an interesting subject to base our study on.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Purpose . . . . .	3
1.2	Delimitations . . . . .	3
1.3	Outline . . . . .	3
1.4	Related Previous Work . . . . .	4
1.5	Problem Formulation . . . . .	4
<b>2</b>	<b>Theory and Concepts</b>	<b>7</b>
2.1	The Credit Default Swap . . . . .	7
2.2	Existing Models of Default Risk . . . . .	8
2.2.1	Structural Models . . . . .	8
2.2.2	Reduced-Form Models . . . . .	9
2.3	The Benchmark Probability of Default - Moody's EDF <sup>TM</sup> . . . . .	9
2.4	The Pricing of a Credit Default Swap using a Reduced Form Approach . . . . .	10
2.4.1	Discretization of the Premium- and Protection Leg Assuming Constant Hazard Rates . . . . .	11
2.4.2	Implied PD, a Practical Example . . . . .	12
2.5	The Lasso Method . . . . .	12
2.6	The Elastic Net . . . . .	14
2.7	Cross-Validation . . . . .	14
2.7.1	K-Fold Cross-Validation . . . . .	15
2.8	Goodness of Fit . . . . .	15
2.9	Change of Probability Measure . . . . .	16
2.9.1	Hazard rates under $\mathbb{Q}$ . . . . .	17
2.9.2	The Interpretation of $\Gamma$ . . . . .	17
<b>3</b>	<b>The Data</b>	<b>19</b>
3.1	Factors Evaluated . . . . .	19
3.1.1	Market Sentiment . . . . .	19
3.1.2	Firm and Sector Specific Variables . . . . .	20
3.1.3	Liquidity . . . . .	20
3.2	The Discount Factor . . . . .	21
<b>4</b>	<b>Method</b>	<b>23</b>
<b>5</b>	<b>Modeling and Assumptions</b>	<b>25</b>
5.1	The Model . . . . .	25
5.2	Underlying Assumptions . . . . .	25
5.2.1	Recovery Rate, RR . . . . .	26
5.2.2	Counterparty Risk . . . . .	26
<b>6</b>	<b>Results</b>	<b>27</b>
6.1	Variable Selection . . . . .	27
6.2	Estimated $\Gamma$ and the implied PD under $\mathbb{P}$ . . . . .	29
6.3	Model Sensitivity . . . . .	32

<b>7</b>	<b>Discussion</b>	<b>35</b>
7.1	Overall Succession . . . . .	35
7.2	Weaknesses in the Analysis . . . . .	35
7.3	Suggested Future Work . . . . .	36
<b>8</b>	<b>Conclusion</b>	<b>37</b>
<b>A</b>	<b>Appendix</b>	<b>43</b>
A.1	The Model Underlying Hypothesis 1 . . . . .	43
A.2	The GARCH Model . . . . .	43
A.3	Bootstrapping Confidence Intervals for Elastic Net-Parameters . . . . .	44
A.4	Time Varying Parameters in Linear Regression Using the Kalman Filter . . . . .	44
A.5	Deriving Premium- and Protection leg . . . . .	46
A.6	Summary of Data . . . . .	46

# 1 Introduction

The basic purpose of the Credit Default Swap (CDS) is to insure the buyer against default risk. The CDS spreads have historically been seen as a pure measure of default risk since the cash-flows exchanged between the buyer and seller of protection represent the price of such an insurance. After the invention of the CDS in the late 1990's this was what researchers assumed. However, the probability of default (PD) implied by the market is observed to diverge from the PD calculated from more fundamental models. This suggests that the CDS spread contains more than a compensation for a possible default. After the contracts started trading more actively in 2004 this view was absorbed by the research community who started to try to explain the difference in implied credit risk from CDS and bond spreads.

While researchers have discussed drivers of CDS spreads, apart from the true credit risk, not many have discussed its impact on the probability of default. As the CDS contracts should insure you for a possible default, the comparison between the PD implied by the market and a "true" PD should contain the same determinants as the extra compensation in spread, since the "true" credit risk should be ascribed to the "true" PD. This difference is attributed to the *market price of risk* related to the Girsanov kernel via the Radon-Nikodym derivative for the change between a risk neutral measure,  $\mathbb{Q}$ , and the physical measure,  $\mathbb{P}$ .

## 1.1 Purpose

This thesis tries to evaluate and explain the difference in the probability of a firm to default from the two main methods available; Structural and Reduced form models.

## 1.2 Delimitations

The goal of this thesis is not to perfectly model the difference in implied PD from the reduced form models and structural models. It rather tries to *explain* the difference using market data.

## 1.3 Outline

The thesis is split into four main parts. A theoretical background on the subjects handled is given after which the methodology and model is specified. It continues to present the results and ends with a discussion and potential drawbacks. The reader is expected to be familiar with basic probability theory.

## 1.4 Related Previous Work

Longstaff, Neis and Mithal (2003) investigate how much of yield spread from corporate bonds that are attributed to default risk and how much that comes from a non-default component, which they attribute to liquidity and taxes. In contrast to previous work they use CDS spreads. However, they use the CDS spread as a pure measure of default risk. The same authors move along to examine whether credit protection is priced consistently in the corporate bond market and CDS market using a reduced form model. They base their analysis on raised concerns that the insurance is not priced fairly in the credit derivative market and exemplifies by rumors that hedge funds should have driven up the CDS spreads in an effort to induce rating agencies to downgrade specific firms [Longstaff *et al.* , 2003].

The view of CDS contracts by the research community largely differs from the inception of the first contracts, it also differs before and after the financial crisis. It has been suggested that the CDS could cause, rather than insure against, default. The CDS contracts played large roles in both the collapse of Lehman Brothers in 2008 and the Greek debt crisis in 2010 as an increase in CDS spread can lead to an increased borrowing cost misrepresenting to the underlying credit risk. Tang and Yan (2013) discusses the lack of empirical evidence on the determinants of spread changes and investigate what moves CDS spreads through a regression model of daily changes [Tang & Yan, 2013] .

When looking at default, there is a difference between physical and risk neutral default measures. Hull, White and Predescu (2005) show through analysis of bonds that the ratio of risk neutral to real world default intensities decreases as the credit quality, measured by ratings from Moody's, declines. They also show that the difference between the two default intensities increases as the credit quality declines. The size between the two default intensities is sometimes referred to as the *credit spread puzzle*. Many researchers have tried to solve this puzzle with more or less qualitative arguments. Later research, that has included CDS contracts instead of bonds, has mainly focused on structural frameworks [Hull *et al.* , 2005a].

In this thesis we take on the task to try and solve this puzzle in a more quantitative way. The reader should thus be aware of the difficulty of the problem, which at a first look might seem simple.

## 1.5 Problem Formulation

In difference to models from the beginning of the 2000's this thesis assumes that the CDS spread observed in the market contains a demanded compensation for default and a compensation for a non-default component that could include other risks in the market.

$$CDS_{\text{spread}} = f(\text{compensation for default}, \text{compensation for non-default}) \quad (1)$$

**Hypothesis 1** An initial hypothesis was to think of the non-default component as noise. In the market this noise causes the investor to pay a different premium for the insurance than the actual premium for the risk of default. This was assumed to be the largest cause of the difference between the calculated implied PD and Moody's EDF<sup>TM</sup>. An initial approach to explain and "wash" away this extra noise in the spreads to gain a more accurate estimated implied PD was

made, see Appendix A.1 for further reading. While this initial analysis was shown not to be fruitful, analysis of the correlation between market data and the CDS spread of the selected firm was useful.

**Hypothesis 2** A second hypothesis was made under the discovery that the implied PD and Moody's EDF<sup>TM</sup> are priced under different probability measures. Given that Moody's EDF<sup>TM</sup> is priced under the physical measure  $\mathbb{P}$  while the implied PD is priced under the risk neutral measure  $\mathbb{Q}$ , there should be a fundamental difference between the two. This difference is commonly attributed to the market price of risk, i.e. the risk premium demanded by the market for the investment. It is next assumed that it is this difference that causes the divergence between the implied PD and EDF<sup>TM</sup>. This thesis next tries to evaluate if it is possible to quantify this market price of risk and further, is it possible to explain it and find its biggest drivers and possible predictive abilities using market data?





## 2 Theory and Concepts

To assure the reader's understanding of the problem and the results, some underlying theory need to be specified.

### 2.1 The Credit Default Swap

A Credit Default Swap (CDS) is a contract providing insurance against a default of a certain company, the reference entity. It is designed to transfer credit risk. The default can be triggered by a number of events such as the failure to pay, bankruptcy or restructuring. The latter was subject to change after years of debate in the ISDA "Big Bang" in 2009, where some North American contracts started trading without the restructuring clause [Markit, 2009]. The protection buyer pays a periodic fee, a premium, to the protection seller in return to be compensated in the case of a triggered default. The fee is paid until the end of the life of the CDS or until a credit event occurs. The swap is settled by cash or physical delivery, where the actual underlying bond is delivered in exchange for its par value. In the case of a cash settlement the protection seller pays  $(1 - RR)$  of the principal, where  $RR$  is the recovered amount from the reference borrower as a percentage of the face value, the *recovery rate*, see Figure 1 [Berndt *et al.* , 2005, ISDA, 2013].

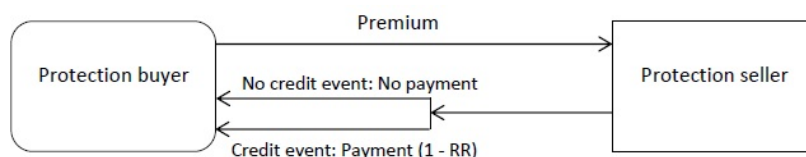


Figure 1: The figure shows the cash flow exchanged by the buyer and the seller of protection

The instruments are used to hedge and trade credit risk and are traded over the counter, OTC. The contracts could for long be tailor made to the buyer preferences. Although, since after the 2008 financial crisis, where shortcomings in the CDS market were revealed, the derivatives have moved to become more standardized with standard maturities and coupons. CDS contracts are available in a number of maturities where the 5 year maturity is by far the most liquid. The CDS has contributed to market completeness through an easier way to express a short view on bonds, since shorting bonds can be difficult as it required the short-seller to borrow the assets through repo [Weistoffer, 2009].

Since the seller agrees to compensate the buyer in the case of default in return for a premium the market price of the payment is an indication of the perceived default risk for the reference entity. The premium is quoted as an annual spread on the market. The value of a swap at time zero is usually assumed to be fair, the present value of the exchanged cashflows are zero [Hull & White, 2000].

$$PV_{\text{Premium leg}} = PV_{\text{Protection leg}} \quad (2)$$

## 2.2 Existing Models of Default Risk

Two main approaches behind the estimation of default probability has emerged; structural models and reduced form models.

### 2.2.1 Structural Models

Structural models includes a set of models that model the PD through the firm value of a specific firm. The default event is assumed to be triggered when the firm value drops below a random or non-random barrier. This definition results in the default time being a predictable stopping time with respect to the reference filtration, modeling the flow of information of the whole market. Since the model models default time in terms of firm value the method is closely linked to the economic fundamentals of the firm, the structure. Some well known models are the Merton model from 1974, Black and Cox model and Moody's KMV model. The latter is currently most commonly used by market practitioners as a benchmark and will continue to serve this purpose in this thesis [Bielecki & Rutkowski, 2002].

In short, the Merton model and structural models derived from it includes three steps:

1. Estimate the market value  $A_t$  and volatility  $\sigma_A$  of the firm in question's assets.
2. Calculate the distance-to-default  $DD$  given in number of asset standard deviations  $\sigma_A$ .
3. Transform the distance-to-default into a PD.

The Merton model sees the firm's value of equity  $E_t$  as a call option written on its value of assets  $A_t$  with strike price equal to the debt payment  $D$  promised i.e.

$$E_T = \max[A_T - D, 0] \quad (3)$$

here  $E_T$  is the payment to the shareholders at time  $T$  [Hull *et al.* , 2005b]. The firm defaults when the value of assets  $A_t$  goes below the promised debt payment  $D$ . The Merton model operates under Black & Schole's framework. To estimate the actual probability of default, it is assumed that under the real-world probability  $\mathbb{P}$  the value process  $A_t$  has the following dynamics

$$dA_t = \mu A_t dt + \sigma_A A_t dW_t \quad (4)$$

where  $\mu$  is the expected growth rate and  $W_t$  is a Brownian motion under  $\mathbb{P}$  [Bielecki & Rutkowski, 2002]. Under these dynamics the asset values are assumed to be lognormal distributed

$$\ln(A_t) \sim N \left( \ln(A_0) + \left( \frac{\mu - \sigma_A^2}{2} \right) t, \sigma_A^2 t \right) \quad (5)$$

The probability of default,  $PD$ , can now be calculated

$$\mathbb{P}[\tau < T | \mathcal{F}_0] = \mathbb{P}[A_T < D | \mathcal{F}_0] = \mathbb{P}[\ln(A_T) < \ln(D) | \mathcal{F}_0] = \Phi \left[ -\frac{\ln(\frac{A_0}{D}) + (\mu - \frac{\sigma_A^2}{2})T}{\sigma_A \sqrt{T}} \right] = \Phi[-DD] \quad (6)$$

[Bielecki & Rutkowski, 2002].  $\Phi$  is the standard normal cumulative density function and  $DD$  the distance to default according to Merton [Hull *et al.* , 2005b].

### 2.2.2 Reduced-Form Models

In difference from the structural models the reduced-form models or hazard rate models do not model the firm value but are instead only concerned with the modeling of the default time. The set of reduced form models model a random default time, nonpredictable stopping time, defined as the jump time of some one-jump process, usually a Poisson process. This leads to the modeling of the default intensity process, the *hazard rate* [Bielecki & Rutkowski, 2002].

Given that default has not occurred before time  $t$  the conditional risk neutral probability of time of default is

$$\mathbb{Q}[\tau < t + dt | \tau \geq t] = h(t)dt \quad (7)$$

Where  $h(t)$  is the intensity of the Poisson process, the hazard rate, which is assumed to be deterministic (independent of interest rates and recovery rates) and time varying [O’Kane & Turnbull, 2003]. The hazard rate is the instantaneous rate of default and using it, the survival probability can be derived. Defining  $H(t) = \int_0^t h(u)du$ , the survival probability under the risk neutral measure  $\mathbb{Q}$  can be written as

$$\mathbb{Q}\{\tau \geq t\} = \mathbb{Q}\{H(\tau) \geq H(t)\} = \mathbb{Q}\left\{H(\tau) \geq \int_0^t h(u)du\right\} = e^{-\int_0^t h(u)du} \quad (8)$$

where the survival probability is the probability of no default before time  $t$ . The last step of equation (8) uses that under the Poisson assumption the waiting time until the default is exponentially distributed. The risk neutral probability of default ( $PD_t$ ) before time  $t$  can now be expressed as

$$PD_t = \mathbb{Q}(\tau < t) = 1 - \mathbb{Q}(\tau \geq t) = 1 - e^{-\int_0^t h(u)du} \quad (9)$$

[Brigo & Mercurio, 2006].

### 2.3 The Benchmark Probability of Default - Moody’s EDF<sup>TM</sup>

Moody’s KMV (MKMV) is a structural model which is used by Moody’s to calculate the *Expected Default Frequency*, EDF<sup>TM</sup>. This is a widely used benchmark measure of PD. This thesis uses this EDF<sup>TM</sup> to test the explanatory power of the model [Hull *et al.* , 2005b].

As mentioned, MKMV is an extension to Merton’s default probability framework. However, it is extended to better match the real world behavior. For instance Moody’s uses proprietary routines to calculate asset value and asset volatility, they also have a large database of historical defaults.  $DD$  is calculated in terms of standard deviations and then they look at the empirical distribution for default and from that they calculate the EDF<sup>TM</sup>. With this model they calculate the "physical probability of default" which is a probability under the physical measure  $\mathbb{P}$ .

The survival probability,  $S(t)$ , is related to Moody’s annualized EDF<sup>TM</sup> as

$$EDF_t^{\text{TM}} = 1 - S(t)^{\frac{1}{t}} \quad (10)$$

See [Sun *et al.* , 2012] for details.

## 2.4 The Pricing of a Credit Default Swap using a Reduced Form Approach

In the absence of arbitrage and under the risk neutral probability measure  $\mathbb{Q}$  one can derive the risk neutral valuation formula of a defaultable claim.

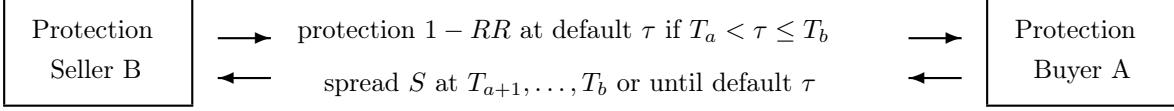


Figure 2: The cashflows in a CDS contract valid between times  $[T_a, T_b]$  where  $\tau$  is the time of default and  $RR$  is the recovery rate for the reference entity. Premium payments are to be made at  $T_{a+1}, \dots, T_b$  unless the reference entity defaults somewhere between  $[T_a, T_b]$ , then B is required to cover the losses for A due to the default [Brigo & Mercurio, 2006].

From the sellers perspective, the present value at time  $t$  of the CDS contract can be expressed according to

$$\begin{aligned} \Pi_{a,b}(t) = & S \frac{B(t)}{B(\tau)} (\tau - T_{\beta(\tau)-1}) \mathbb{1}_{\{T_a < \tau < T_b\}} + \\ & + S \sum_{i=a+1}^b \frac{B(t)}{B(T_i)} \Delta_i \mathbb{1}_{\{\tau \geq T_i\}} - (1 - RR) \frac{B(t)}{B(T_i)} \mathbb{1}_{\{T_a < \tau \leq T_b\}} \end{aligned} \quad (11)$$

where  $\tau$  is the time of default,  $T_{\beta(t)}$  is the first date among the  $T_i$ 's that follows  $t$ ,  $t \in [T_{\beta(t)-1}, T_{\beta(t)})$ ,  $\Delta_i$  is the year fraction between  $T_{i-1}$  and  $T_i$ .  $B(t)/B(T)$  is the stochastic discount factor at time  $t$  for maturity  $T$ .  $RR$  is the recovery rate which is assumed to be deterministic and  $S$  is the CDS spread. The risk neutral valuation formula for the CDS price is

$$CDS_{a,b}(t, S, RR) = \mathbb{E}^{\mathbb{Q}} (\Pi_{a,b}(t) | \mathcal{G}_t) \stackrel{\text{def.}}{=} PV_{\text{PremiumLeg}}^{a,b} - PV_{\text{ProtectionLeg}}^{a,b} \quad (12)$$

Where  $\mathcal{G}_t$  is the default free filtration that includes default monitoring i.e. it is known whether the underlying name of the CDS has defaulted so far or not. As equation (12) shows, the evaluation formula can be expressed using the present values of the premium leg and protection leg mentioned in Section 2.1. These legs can now be expressed as equation (13) and (14) respectively. See the Appendix A.5 for derivation.

$$\begin{aligned} PV_{\text{PremiumLeg}}^{a,b} = & S \int_{t=T_a}^{T_b} \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(t)} \right] (\tau - T_{\beta(\tau)-1}) \mathbb{Q}(\tau \in [t, t + dt]) + \\ & + S \sum_{i=a+1}^b \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(T_i)} \right] \Delta_i \mathbb{Q}(\tau \geq T_i) \end{aligned} \quad (13)$$

$$PV_{\text{ProtectionLeg}}^{a,b} = (1 - RR) \int_{t=T_a}^{T_b} \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(t)} \right] \mathbb{Q}(\tau \in [t, t + dt]) \quad (14)$$

The fair CDS spread  $S$  is calculated using that the expected value of future exchanged cash flows should be zero which is the same as equalising the premium- and protection leg as in Section 2.1 i.e.

$$\begin{aligned}
CDS_{a,b}(t, S, RR) = 0 &\Rightarrow PV_{\text{PremiumLeg}}^{a,b} = PV_{\text{ProtectionLeg}}^{a,b} \Rightarrow \\
S = \frac{(1 - RR) \int_{t=T_a}^{T_b} \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(t)} \right] \mathbb{Q}(\tau \in [t, t + dt])}{\int_{t=T_a}^{T_b} \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(t)} \right] (\tau - T_{\beta(\tau)-1}) \mathbb{Q}(\tau \in [t, t + dt]) + \sum_{i=a+1}^b \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(T_i)} \right] \Delta_i \mathbb{Q}(\tau \geq T_i)} & \quad (15)
\end{aligned}$$

If  $S$  is known from the market, equation (2) can be used to calculate the default probability implied by it. One way of solving for this probability is by discretizing the expressions in equation (13) and (14) by assuming a hazard rate model [Brigo & Mercurio, 2006].

#### 2.4.1 Discretization of the Premium- and Protection Leg Assuming Constant Hazard Rates

One way of computing the implied PD is to calculate the implied hazard rate. Assume that CDS spreads  $S_{0,t_N}$  for different maturities  $t_N$  written on the same underlying entity is accessible,  $t_i$ ,  $i = 0, 1, 2, \dots, N$  is the premium payments dates. If the hazard rate is assumed to be constant and deterministic between the  $t_N$ 's the premium- and protection leg given in equation (13) can be approximated as follows

$$\begin{aligned}
PV_{\text{PremiumLeg}} &= S_{0,t_N} \left\{ \sum_{i=1}^N \Delta_{t_{i-1}, t_i} DF_{t_0, t_i} \left[ \mathbb{Q}(t_0, t_{i-1}) + \frac{\mathbb{1}_{PA}}{2} (\mathbb{Q}(t_0, t_{i-1}) - \mathbb{Q}(t_0, t_i)) \right] \right\} \\
PV_{\text{ProtectionLeg}} &= (1 - RR) \sum_{m=1}^{M \cdot t_N} DF_{t_0, t_m} (\mathbb{Q}(t_0, t_{m-1}) - \mathbb{Q}(t_0, t_m)) & (16)
\end{aligned}$$

Where  $\Delta_{t_{i-1}, t_i}$  is the year fraction between  $t_{i-1}$  and  $t_i$ ,  $DF_{t_0, t_i}$  is the risk free discount factor from  $t_0$  to  $t_i$  which is assumed to be known,  $\mathbb{Q}(t_0, t_i)$  is the survival probability between  $t_0$  and  $t_i$  for the underlying entity that the CDS is written on. The accrued interest i.e. the integral in the expression for the premium leg, equation (13), is here approximated using the trapezoidal rule. Accrued interest is not always specified in the CDS contract, the function  $\mathbb{1}_{PA}$  is therefore a function that is 1 if accrued interest is specified and 0 if not. The integral in the protection leg, equation (14) is here discretized by assuming that default only can occur on a finite number of discrete points per year,  $M$ , called the discretization frequency. A CDS with maturity  $t_N$  has  $M \cdot t_N$  possible default times  $m = 1, 2, \dots, M \cdot t_N$  in equation (16). The larger the value for  $M$  the more accurate the calculations will be but larger  $M$  also means more calculations.

Based on the fact that hazard rates between the maturity times  $t_N$  are assumed to be constant, given a set or subset of maturities  $t_{N_1}, t_{N_2}, \dots, t_{N_n}$ , the survival probability on  $\alpha$  years basis can be written as

$$\mathbb{Q}(t_0, t_{N_\alpha}) = \exp(-(h_{0,1} t_{N_1} + h_{1,2} (t_{N_2} - t_{N_1}) + \dots + h_{n-1,n} (t_{N_\alpha} - t_{N_{\alpha-1}}))) \quad (17)$$

Where  $h_{i-1,i}$  is the constant hazard rate between maturities  $t_{N_{i-1}}$  and  $t_{N_i}$ . Using this, it is possible to iterate a hazard rate term structure by taking the shortest CDS maturity and use equation (2) to calculate the constant hazard rate  $h_{0,1}$  for that maturity. When this is done the procedure is repeated and solved for  $h_{1,2}$  and so on [O’Kane & Turnbull, 2003].

## 2.4.2 Implied PD, a Practical Example

Assume given spreads  $S_{0,t_{N_i}}$  for maturities  $i = 1, 3$  and 5 years. A monthly discretization frequency  $M = 12$  and quarterly premium payments are set, which mean that a premium is paid at  $i = 3, 6, 9, \dots, t_{N_i}$ . It is also assumed that no accrued interest is specified, see equation (16). The value of  $h_{0,1}$  can now be calculated using the 1 year CDS spread and equation (2).

$$PV_{\text{PremiumLeg}} = PV_{\text{ProtectionLeg}} \Rightarrow$$

$$S_{0,t_{N_1}} \sum_{i=3,6,9,12} \Delta_{t_{i-3},t_i} DF_{t_0,t_i} e^{-h_{0,1} \frac{i}{12}} = (1 - RR) \sum_{m=1}^{12} DF_{t_0,t_m} (e^{-h_{0,1} \frac{m-1}{12}} - e^{-h_{0,1} \frac{m}{12}}) \quad (18)$$

This equation can now be solved for  $h_{0,1}$  by using a numerical method such as a one dimensional root searching algorithm e.g. Newtons method. Next step is to repeat the procedure to solve for  $h_{1,2}$  and so on until the final maturity is reached, in this case  $t_{N_5}$ . When the iteration is done the piecewise constant hazard rate term structure is evaluated [O’Kane & Turnbull, 2003].

## 2.5 The Lasso Method

The *Least Absolute Shrinkage and Selection Operator Method* or the lasso method, is a method for variable selection for regression models introduced by Tibshirani in 1996 [Tibshirani, 1996]. The lasso method minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant. It improves prediction accuracy by shrinking some coefficients or setting others to zero by imposing a  $L_1$  - norm penalty on the regression coefficients. The more commonly used Ordinary Least Squares, OLS, regression finds an unbiased linear combination of the  $x_i$ ’s that minimizes the residual sum of squares. However, if the number of parameters are large or if the parameters are highly correlated through multicollinearity, the OLS method may yield results with large variance, thus reducing the accuracy of the prediction. The lasso method sacrifices a bit of the bias in order to reduce the variance of the parameters to improve the overall prediction accuracy [Tibshirani, 1996]. In order to understand the lasso method it is relevant to start with something that can be solved in closed form, which lasso cannot. Therefore a short description of *ridge regression* follows.

**Definition 2.1** (Ridge Regression). Consider the data  $(\mathbf{x}_i, y_i)$ ,  $i = 1, 2, \dots, n$ , where  $\mathbf{x}_i = (x_{1i}, \dots, x_{pi})$  are the prediction variables and  $y_i$  are the responses. Assume that the  $x_{ij}$  are standardized so that  $\sum_i x_{ij}/n = 0$ ,  $\sum_i x_{ij}^2/n = 1$ . If  $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)$ , the ridge estimate on Lagrangian form is defined as

$$\hat{\beta} = \arg \min_{\beta} \sum_{t=1}^N (y_t - \beta_0 - \sum_{j=1}^p x_{tj} \beta_j)^2 + \lambda \sum_{j=1}^p \beta_j^2, \quad \lambda \geq 0 \quad (19)$$

[Hastie *et al.* , 2009]

Here  $\lambda$  is the shrinkage parameter, the larger  $\lambda$  the greater the shrinkage. The parameters are shrunk towards zero due to the  $L^2$ -norm penalty. Writing equation (19) on matrix form

$$\arg \min_{\beta} (\mathbf{y} - \mathbf{X}\beta)^T(\mathbf{y} - \mathbf{X}\beta) + \lambda\beta^T\beta, \lambda \geq 0 \quad (20)$$

it is possible to get a closed form solution

$$\hat{\beta} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}. \quad (21)$$

Although ridge regression does not have the property of setting coefficients to zero, the lasso method, which possess this property, is very similar [Hastie *et al.* , 2009].

**Definition 2.2** (Lasso). Consider the data  $(\mathbf{x}_i, y_i)$ ,  $i = 1, 2, \dots, n$ , where  $\mathbf{x}_i = (x_{1i}, \dots, x_{pi})$  are the prediction variables and  $y_i$  are the responses. Assume that the  $x_{ij}$  are standardized so that  $\sum_i x_{ij}/n = 0$ ,  $\sum_i x_{ij}^2/n = 1$ . If  $\hat{\beta} = (\hat{\beta}_1, \dots, \hat{\beta}_p)$ , the lasso estimate  $\hat{\beta}$  is defined by

$$\hat{\beta} = \arg \min \frac{1}{2} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 \text{ subject to } \sum_{j=1}^p |\beta_j| \leq t \quad (22)$$

Here  $t \geq 0$  is a tuning parameter controlling the amount of shrinkage that is applied to the estimates. Values  $t < \sum |\beta_j^o|$ , where  $\beta_j^o$  are the OLS estimates, will cause shrinkage of the solution towards zero, some coefficients may be exactly zero. If  $t > \sum |\beta_j^o|$  the lasso method yields the same estimate as the OLS [Tibshirani, 1996]. This expression can also be written on an equivalent *Lagrangian form*

$$\hat{\beta} = \arg \min \frac{1}{2} \sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij})^2 + \lambda \sum_{j=1}^p |\beta_j|, \lambda > 0 \quad (23)$$

[Hastie *et al.* , 2009].

A fundamental step is to chose the value of  $\lambda$  optimally. As mentioned this parameter controls the amount of shrinkage on the coefficients resulting in a subset of selected variables to include in the final model. A too large  $\lambda$  leads to excessive shrinkage on the parameters and the selected variables may lack some informative variables. On the other hand a too small  $\lambda$  could lead to a scenario where the amount of shrinkage is too small and the set of selected variables may contain some uninformative variables. For these reasons the  $\lambda$  parameter should be chosen optimally. This is often done through cross-validation.

The lasso has shown to have excellent abilities, however some drawbacks should be mentioned. In the case of multicollinearity among the variables, the penalty from the  $L_1$  norm is not sufficient. Methods as ridge regression or elastic net, that uses the  $L_2$ -norm respectively both the  $L_1$  and  $L_2$  norm, succeeds the lasso in prediction performance. Therefore the issue of multicollinearity should be carefully evaluated before a method is selected. A common critique is also found in the case when  $p > n$ , as the lasso cannot select more than  $n$  variables [Tibshirani, 1996].

The *relaxed lasso* first uses lasso to select non-zero predictors, then applies the lasso method again using only the selected variables from the first step. Since the variables in the second step have



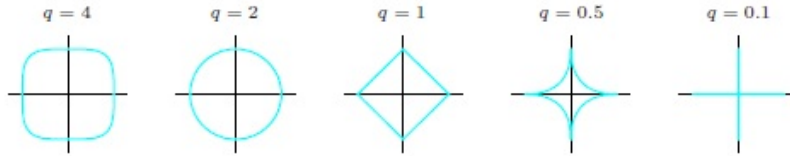


Figure 3: Contours of constant value of  $\sum |\beta_j|^q$  for given values of  $q$  [Hastie *et al.* , 2009].

less competition from noise variables, cross validation will tend to pick smaller values for lambda. This results in the coefficients to be shrunken less than those in the initial estimate. This is the method used for parameter estimating in the thesis but in combination with the *elastic net* [Hastie *et al.* , 2009].

## 2.6 The Elastic Net

The elastic net is a related method to lasso which should be used when the variables are highly correlated. In general the elastic net tends to retain more non-zero variables than lasso but with smaller magnitudes. As previously mentioned the elastic net penalize using both the  $L_1$  and  $L_2$  norm. The penalty has the form

$$(1 - \alpha)\|\beta\|_2^2 + \alpha\|\beta\|_1 = \sum_{j=1}^p (\alpha|\beta_j| + (1 - \alpha)\beta_j^2) \quad (24)$$

When  $\alpha = 1$  the elastic net is the same as lasso and when  $\alpha = 0$  the elastic net approaches ridge regression. The  $\alpha$  is often chosen qualitatively and is set to 0.5 for the purpose of this thesis [Hastie *et al.* , 2009].

As seen in Figure 3, depending on the choice of norm,  $q$ , the constraint region has different shapes. In the two-dimensional case the lasso has the shape of a diamond while ridge regression has the shape of a disk. In the case when  $q > 1$ ,  $|\beta_j|^q$  is differentiable at 0 and does not share the ability of the lasso for setting coefficients exactly to zero. The elastic net mixes the  $L_1$ - and  $L_2$ -norm depending on the choice of  $\alpha$ , thus resulting in a diamond shape with smooth edges but with non-differentiable corners, see Figure 4.

## 2.7 Cross-Validation

A simple and commonly used method for estimating prediction errors is cross validation. The method estimates the expected extra-sample error  $\epsilon = E[L(Y, f(X))]$ , the average generalization error when the method  $f(X)$  is applied to an independent test sample from the joint distribution of  $X$  and  $Y$  [Hastie *et al.* , 2009]. The cross-validation method is here used to choose an optimal penalizing  $\lambda$ .

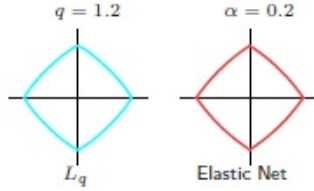


Figure 4: Contours of constant value of  $\sum |\beta_j|^q$  for  $q = 1.2$  (left plot), and the elastic-net penalty  $\sum (\alpha \beta_j^2 + (1 - \alpha) |\beta_j|)$  for  $\alpha = 0.2$  (right plot). Although visually very similar, the elastic-net has sharp (non-differentiable) corners, while the  $q = 1.2$  penalty does not [Hastie *et al.*, 2009].

### 2.7.1 K-Fold Cross-Validation

K-Fold cross-validation uses parts of the available data to fit the model and a different part to test it. The data is split into  $K$  roughly equal sized parts and the  $k$ th part is fitted to the other  $K - 1$  parts of the data. The prediction error of the fitted model is calculated when predicting the  $k$ th part of the data. This is done for  $k = 1, 2, \dots, K$  and the  $K$  estimates of the prediction error are combined [Hastie *et al.*, 2009].

**Definition 2.3.** Let  $\kappa : \{1, \dots, N\} \mapsto \{1, \dots, K\}$  be an indexing function indicating the partition to which observation  $i$  is allocated by the randomization. Denote by  $f^{-\kappa(i)}(x)$  the fitted function, computed with the  $k$ th part of the data removed. Then the cross-validation estimate of prediction error is

$$CV(f) = \frac{1}{N} \sum_{i=1}^N L(y_i, f^{-\kappa(i)}(x_i)) \quad (25)$$

[Hastie *et al.*, 2009]

In the case when  $K = N$  the method is known as *leave-one-out* cross-validation and the estimator is approximately unbiased for the true predictor error, but can have high variance due to the  $N$  training sets being so "similar" to each other. With a lower  $K$ , the cross-validation has lower variance but depending on the size of the training set, bias could be a problem. A common choice for  $K$  is 10 and is usually recommended as a good compromise between bias and variance. This is also the choice for this thesis [Hastie *et al.*, 2009].

## 2.8 Goodness of Fit

To test how much of the variability of the response that can be explained by the variables, the *coefficient of determination* or the  $R^2$  can be used.  $R^2$  ranges between 0 – 1 where a  $R^2 = 1$  implicates a perfect fit.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (26)$$

where  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ , and  $\hat{y}$  is the model estimate of  $y$  [Rawlings *et al.*, 2001].

## 2.9 Change of Probability Measure

To change measure in the setting of a general probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , the Radon-Nikodym Theorem is used. The theorem states that for a measure  $\mathbb{Q}$  absolutely continuous with respect to  $\mathbb{P}$  there is a stochastic variable  $L$  such that  $\mathbb{Q}$  is given by

$$\mathbb{Q} = E^{\mathbb{Q}}[\mathbf{1}_A L]$$

$L$  is the likelihood ratio of the measure change  $\mathbb{P} \rightarrow \mathbb{Q}$  [Åberg, 2010].

**Theorem 1** (Change of Measure and  $L$ ). *Consider a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  and an absolutely continuous probability measure  $\mathbb{Q}$ . Then there is a stochastic variable  $L$  with*

$$L \geq 0$$

and

$$E^{\mathbb{P}}[L] = 1$$

such that

$$\mathbb{Q}(A) = E^{\mathbb{P}}[\mathbf{1}_A L], \quad \forall A \in \mathcal{F}$$

$$E^{\mathbb{Q}}[X] = E^{\mathbb{P}}[XL], \quad \forall \mathcal{F} - \text{measurable } X$$

If  $\mathbb{P}$  and  $\mathbb{Q}$  are equivalent, i.e.  $\mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0, \forall A$ , the likelihood ratio  $L$  fulfills

$$L > 0$$

and

$$\mathbb{P}(A) = E^{\mathbb{P}}\left[\mathbf{1}_A \frac{1}{L}\right], \quad \forall A \in \mathcal{F}$$

[Åberg, 2010]

**Lemma 1.** Let  $g(t)$  be an  $\mathcal{F}_t$ -adapted process that satisfies

$$\mathbb{P}\left\{\int_0^T g^2(t) dt < \infty\right\} = 1$$

Then the equation

$$dL(t) = g(t)L(t)dX(t), \quad L(0) = 1 \tag{27}$$

has the unique and strictly positive solution

$$L(t) = \exp\left(\int_0^t g(s) dX(s) - \frac{1}{2} \int_0^t g^2(s) ds\right) \tag{28}$$

[Madsen *et al.*, 2004]

**Theorem 2** (The Girsanov Theorem). *Let  $X(t)$  be a  $(\mathbb{P}, \mathcal{F}_t)$ -Wiener process and let  $g(t)$  and  $L(t)$  be as defined in Lemma 1. Assume that  $E[L(t)] = 1$  and define the probability measure  $\mathbb{Q}$  by  $d\mathbb{Q} = L(T)d\mathbb{P}$  on  $\mathcal{F}_T$ . Then the process  $W(t)$ , defined by*

$$W(t) = X(t) - \int_0^t g(s) ds$$

becomes a  $(\mathbb{Q}, \mathcal{F}_t)$ -Wiener process. Or on differential form

$$dX(t) = g(t)dt + dW(t)$$

[Madsen *et al.*, 2004]

### 2.9.1 Hazard rates under $\mathbb{Q}$

If  $\mathbb{Q}$  is absolutely continuous with respect to the reference probability measure  $\mathbb{P}$ , then under mild assumptions, it is possible to derive a relationship between the hazard functions  $H^{\mathbb{Q}}$  and  $H^{\mathbb{P}}$  under  $\mathbb{Q}$  and  $\mathbb{P}$  respectively. If the filtration  $\mathcal{F}$  is generated by a Brownian motion, any  $\mathcal{F}$ -martingale is a  $\mathcal{G}$ -martingale under  $\mathbb{P}$  and the hazard process  $H(\tau)$  is a continuous increasing process, then it would be natural to conjecture that if a probability measure  $\mathbb{Q}$  is equivalent to  $\mathbb{P}$  on  $(\Omega, \mathcal{G}_t)$ , the  $\mathcal{F}$ -hazard process  $H^{\mathbb{Q}}$  is given by the following expression

$$H^{\mathbb{Q}} = \int_0^t (1 + \Gamma) dH_u^{\mathbb{P}}, \quad \forall t \in [0, T].$$

[Bielecki & Rutkowski, 2002]. In the case of a random time  $\tau$ , which admits the  $\mathcal{F}$ -intensity process  $h^{\mathbb{P}}$  under  $\mathbb{P}$  it would be natural to expect that  $\tau$  also admits the  $\mathcal{F}$ -intensity process  $h^{\mathbb{Q}}$  under  $\mathbb{Q}$  and that the relationship

$$h^{\mathbb{Q}} = h^{\mathbb{P}}(1 + \Gamma) \tag{29}$$

is satisfied for every  $t \in [0, T]$  [Bielecki & Rutkowski, 2002].

### 2.9.2 The Interpretation of $\Gamma$

The Bank for International Settlements suggest that the  $\Gamma$  is the market price of *default event risk* based on analysis of default intensities in bonds. In another publication they apply (29) to CDS where they suggest  $\Gamma$  to be the *jump-at-default risk*. However, these definitions were made in 2005-2006 before the financial crisis which very much changed the view of CDS contracts [Amato & Luisi, 2006, Amato, 2005]. Given that later research, such as Tang and Yan (2013), suggest that the CDS spreads contain other drivers than the true credit risk, this thesis defines  $\Gamma$  as *the market price of risk*, thus including more than default event risk/jump-at-default risk. Björk (2009) defines a market price of risk as  $-g$ , where  $g$  is the Girsanov kernel, and the market price of risk measures the aggregated risk aversion in the market [Björk, 2009]. It always appears in pricing equations when something that isn't traded is modeled, in this case Moody's EDF<sup>TM</sup>. It measures how much extra return you need to take unhedgeable risk [Ahmad & Wilmott, 2007]. A  $\Gamma = 0$  implies a risk neutral market where  $h^{\mathbb{Q}} = h^{\mathbb{P}}$ .



## 3 The Data

The data used was daily samples from January 3<sup>rd</sup> 2006 to December 31<sup>st</sup> 2013 obtained through Bloomberg. Moody's EDF was provided by Moody's Investors Service. The data was matched using an intersection of the time vectors. As the dataset included several crises with turbulent changes on a daily basis the intersection of the times was used instead of interpolation of missing data. Mid price CDS data was obtained for 1y, 3y and 5y CDS for AB Volvo, TeliaSonera AB and Stora Enso Oyj. Data points where the bid price was higher than the ask price were removed. See Appendix A.6 for further explanation of the data. Each set of factor data was divided with its mean in an attempt to neutralize the effect of different units across the different factors.

It is to be noted that the data obtained from another timezone can have a lagged correlation, investor reaction in the United States does not match reaction in the European market due to the difference in market opening times. As the US market has opening times that intersect the European market this offers some difficulty and the correlation plot need to be closely examined.

### 3.1 Factors Evaluated

The factors chosen for evaluation were based on previous research and discussion with experienced traders. Tang and Yan (2013) find that CDS prices are more sensitive to trading when option implied volatility is higher. They also find that changes in macroeconomic conditions, measured by changes in the five year swap rate and the VIX among else, and firm fundamentals, such as firm volatility, are important determinants for changes in the CDS spread [Tang & Yan, 2013]. The traders also indicated the importance of macroeconomic conditions but rejected the hypothesis that the change in exchange rate of the quotation currency of the CDS contracts has any impact on the actual CDS spread. After this initial analysis subgroups of potential factors that influence CDS spreads were selected.

#### 3.1.1 Market Sentiment

**Equity Based Volatility Measures** A measure of option implied volatility is the VIX, a volatility index well known by the market. It is based on implied volatility from options on the S&P 500, an equity index comprising of the 500 largest companies by market value in the United States [CBOE, n.d.b]. A European alternative is the V2X derived from Euro Stoxx 50 index options. Both of these indices measures market participants sentiment of 30 days forward volatility [CBOE, n.d.b, Eurex, n.d.]. These indices are also commonly used as a measurement for market risk aversion or the "fear factor" in the market.

**Volatility Skew** The Chicago Board Option Exchange (CBOE), has created an index derived from the price of S&P 500 tail risk, the risk of outlier returns of two or more standard deviations below the mean. An increased perceived tail risk results in an increased demand for low strike puts which increases the Skew [CBOE, n.d.a].

**Foreign Exchange Implied Volatility** The implied volatility for options on foreign exchange rates for USD/SEK and EUR/USD with expiry in 1 day, 1 week, 1 month, 6 months, 1 year, 2 year, 3 year and 5 year was analyzed. The selection was highly influenced by the number of data samples available in each set which were limited in some of the expires.

**Swaption Volatility** Implied volatility for 5 year swaptions for SEK, EUR and USD with expiry 3 months, 6 months and 1 year were analyzed.

**Rates** Germany is one of the leading economies in Europe with a substantial economic influence. When Europe faces distress the German Bund benefits from investors flight to quality and the rate could therefore potentially be used as a market sentiment factor for Europe. Data for a 5 year Germany government bond-rate index was used. Evaluation was also made with some of the European interbank rates for 3 months maturity; London Interbank Rate (LIBOR), Stockholm Interbank Rate (STIBOR) and European Interbank Rate (EURIBOR). These are all often used as a benchmark. Five year swap rates for EUR and SEK were also analyzed.

### 3.1.2 Firm and Sector Specific Variables

**Equity** The companies analyzed all had publicly traded equity. The equity traded on the Stockholm exchange was chosen to minimize the effect of lagged correlation from another timezone.

**Volatility from Equity Returns** A GARCH(1,1) with a student -  $t$  distribution was used to model the time-varying volatility from daily returns in equity, see Appendix A.2 for further reading.

**Sector index** A global equity based sector index was used to model changes in the market affecting the sector but not necessarily leading to an increase in firm specific credit risk. For AB Volvo the capitalization weighted Bloomberg World Industrials Index was used. The Bloomberg World Communications Index was used for TeliaSonera AB and the Bloomberg World Basic Material Index was used for Stora Enso Oyj.

### 3.1.3 Liquidity

Market liquidity is generally defined as the ability to trade quickly at a low cost. This usually transfers to low transaction costs from low bid-ask spreads, low price impact of trading large volumes and large market depth [Bielecki *et al.* , 2011]. Liquidity risk is usually defined as the risk of not being able to trade immediately in the market to liquidate or hedge ones position. This risk results from the fact that the financial market is not perfect at all times [Bevras, 2006]. One should note that liquidity level and liquidity risk are conceptually distinct from each other, although often correlated. Generally, market liquidity is a precondition for market efficiency and a sudden worsening of market liquidity may degenerate into a systematic crisis [ECB, 2009].

The market has yet to come to consensus for a measure of liquidity and liquidity risk. Many has evaluated different measures with different ratios of success but the difference in bid-ask spread is still the most commonly used as it expresses various components captured by the market microstructure, such as adverse selection, inventory cost and search frictions [Tang & Yan, 2013]. The choice of this thesis is the difference in ask and mid price as that is the extra cost of your insurance due to illiquidity in the market.

### 3.2 The Discount Factor

To get the "risk-free" discount factor the swap curve that ISDA uses was used. The data included annualized historical rates for maturities 2, 3, 4, 5, 6, 7, 8, 9, 10, 12 and 15 years. When the implied PD was calculated, monthly historical discount factors were needed, therefore the missing rates were interpolated linearly and then transformed into discount factors according to

$$DF_i = (1 + R_i T_i)^{-1} \tag{30}$$

where  $DF_i$  is the discount factor for time  $i$ ,  $R_i$  is the rate and  $T_i$  is the time to maturity [White, 2013].





## 4 Method

The methodology of the thesis consists of three main steps: implementation of an implied PD model, data mining of factors from market data and the estimation of the market price of risk.

The implied PD model was implemented in accordance with the methods and assumptions used by the market [O’Kane & Turnbull, 2003, Luo, 2005]. Sensitivity analysis was performed to conclude that the model’s behaviour was consistent with theory.

To evaluate which market factors that could impact the market price of risk, discussions with experienced traders were held and the factors were divided into three subsets; market sentiment, firm specific variables and liquidity. Given that Moody’s updated their model after critique of its slow reactions during the financial crisis, the data was split into two subsets. One including the full set, 2006 - 2014, and one from 2011 - 2014 [FCIR, 2011, Sun *et al.* , 2012].

Starting of with data for AB Volvo, an expression for the market price of risk,  $\Gamma$ , was derived. Regression plots of the factors were observed and the style of relation was analyzed, i.e. linear, logistic, non-linear. The correlation and cross-correlation between  $\Gamma$  and the market factors were examined to determine the choice of regression method, lasso or elastic net, as well as the lag with the largest dependence. The chosen variable selection method was performed with different choices for  $\lambda$  and the relaxed elastic net method was used to determine the parameters. A Bootstrap method was used to gain confidence intervals for the parameters, see Appendix A.3 for details. The model estimated  $\Gamma$  was used to estimate a new implied PD, this time under  $\mathbb{P}$ , and it was once again compared to Moody’s EDF<sup>TM</sup>. The step was repeated for TeliaSonera AB and Stora Enso Oyj. To evaluate the behaviours of the estimated parameters over time, linear regression of time-varying parameters was done with a Kalman filter, see Appendix A.4. To evaluate the consistency of the estimated parameters regression of chosen factors was performed using rolling windows of sizes 30 - 600 sample points. To evaluate the predictability of the model the same windows were used to estimate the parameters, then used to predict the next point in time. The 2011 - 2014 data sample was used for this step to minimize the effect on  $\Gamma$  due to updates in Moody’s model. The results were discussed.



## 5 Modeling and Assumptions

Given the assumptions that

$$\begin{cases} h_t^{\mathbb{Q}} = h_t^{\mathbb{P}}(1 + \Gamma_t) \\ 1 - EDF_t^{\text{TM}} = e^{-\int_0^t h_s^{\mathbb{P}} ds} \end{cases} \quad (31)$$

an expression for the market price of risk,  $\Gamma$ , can be derived.

$$\begin{aligned} 1 - EDF_t^{\text{TM}} &= e^{-\int_0^t h_s^{\mathbb{P}} ds} = e^{-\int_0^t \frac{h_s^{\mathbb{Q}}}{(1+\Gamma_t)} ds} \Leftrightarrow \\ \ln(1 - EDF_t^{\text{TM}}) &= -\int_0^t \frac{h_s^{\mathbb{Q}}}{(1 + \Gamma_t)} ds \approx \frac{1}{(1 + \Gamma_t)} \sum_{s=0}^t h_s^{\mathbb{Q}} \Delta_s \end{aligned}$$

Resulting in

$$\Gamma_t = \frac{-\sum h_s^{\mathbb{Q}} \Delta_s}{\ln(1 - EDF_t^{\text{TM}})} - 1 \quad (32)$$

This  $\Gamma_t$  is for the purpose of this thesis assumed to be the *true*  $\Gamma_t$ . Here the first assumption in equation 29 is derived from Section 2.9.1 and the second assumption suggest that Moody's structural model can be written on a reduced form.

### 5.1 The Model

To model  $\Gamma$  using market data, a multivariate linear regression model was used. A multivariate linear regression model is on the form,

$$Y = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \epsilon \quad (33)$$

This model often assumes that  $\epsilon \in N(0, \sigma^2)$ . If the relationship is shown to be nonlinear one can transform the variables, the response or both to see if the transformation model can be linearized. A log-linear model is on the form

$$\ln(\Gamma) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n + \eta \quad (34)$$

If additive errors are assumed this can be written as

$$\Gamma = e^{\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n} + \epsilon \quad (35)$$

[Rawlings *et al.*, 2001]. In this case the  $\epsilon \in \ln N(0, \sigma^2)$ . This distribution is well known within finance from the framework of Black and Scholes, who assumes the logarithm of stock returns to be normally distributed [Åberg, 2010].

### 5.2 Underlying Assumptions

Absence of arbitrage is a fundamental assumptions for many mathematical models. To guarantee the existence of the risk neutral measure  $\mathbb{Q}$ , this is also an assumption for this thesis. However, a mechanism called *basis* is present in the CDS market, where an arbitrage opportunity can be present in the difference between CDS index and underlying firm specific CDS contracts. In order for  $\mathbb{Q}$  to be unique the market has to be complete. In spite of the unrealistic nature of the assumption, this is an assumption that is made. For the purpose of this thesis, it is also assumed that no accrued interest is paid if default occurs between payment dates.

### **5.2.1 Recovery Rate, RR**

Many models in literature assumes the recovery rate to be constant, commonly 40% in accordance with market practice. However, some argue that the recovery rate may be cyclically negatively correlated with the default rate, implying a positive correlation between loss given default and PD. Literature propose that RR can be modeled as a stochastic process or cheapest-to-deliver bond price can be used to estimate a more accurate RR [ECB, 2009].

However, IMF (2009) suggest that it could be difficult to disentangle both a stochastic PD and a stochastic RR [Singh & Spackman, 2009]. With this in mind and with the purpose of this paper the RR is assumed to be constant at 40% and the sensitivity of the model to RR is evaluated.

### **5.2.2 Counterparty Risk**

Since the collapse of Lehman Brothers in 2008 the concern about counterparty risk has increased in the context of valuing CDS contracts. However, as previously examined by Hull and White (2001) the counterparty risk only has a minor impact on the valuation when the correlation between the counterparty and reference entity is zero [Hull & White, 2001]. Therefore for the nature of this thesis the counterpart risk is assumed to be nonexistent.

## 6 Results

### 6.1 Variable Selection

Numerous factors from market data was used in the regression. The elastic net was chosen as a variable selection method due to high correlation between some of the data sets used when modeling. It penalizes the parameters in different extent due to the choice of  $\lambda$ . For that reason one must sacrifice a bit of the accuracy of the model in order to reduce the number of variables. Table 1 shows the coefficients for relaxed elastic net for different choices of  $\lambda$ . For the choice of  $\lambda$  that minimises the mean squared error plus one standard deviation,  $\lambda_{MSE+1\sigma}$ , for variable selection but with the choice of  $\lambda$  that minimises MSE,  $\lambda_{MSE}$ , for the fine tuning of the parameters through relaxed elastic net, 14 variables were chosen for AB Volvo. The 95% parameter confidence intervals can be seen in Table 2. These selected variables were used throughout the rest of the analysis.

Table 1: Variable selection results for AB Volvo 2006-2014 data for different choices of  $\lambda$ .

Factor	Elastic Net, $\lambda_{MSE}$	Elastic Net, $\lambda_{MSE+1\sigma}$
$\beta_0$	-1.6233	-1.1856
V2X	-0.0499	-0.1103
Vix	-0.2075	-0.1982
VixLagged	-0.0070	0
FXImpVolEUR1y	0.9009	0.5235
FXImpVolSEK1m	-0.2928	0
Ger5yChange	0.2485	0
Skew	-0.6534	-0.6490
Stibor3m	-0.4335	-0.3387
Libor3m	0.2539	0.2456
Euribor3m	0.8548	0.7849
SwapVolEUR6m	0.2791	0
SwapVolSEK6m	-0.5486	-0.3617
SwapVolUSD6m	0.8658	0.8320
Swap5ySEK	-1.2374	-1.0049
Swap5yEUR	0	0
Equity	0.9749	0.9886
EquityVol	0.6192	0.5867
Sector	1.2003	1.0324
Liquidity	-0.0343	-0.0365
$R^2$	0.7337	0.7293

Table 2: Bootstrapped 95% confidence intervals for the parameters given in Table 1 (AB Volvo), the result is based on 1000 bootstrap iterations. Significant parameters are bold, see Appendix A.3 for more details.

Factor	Elastic Net, $\lambda_{MSE}$		Elastic Net, $\lambda_{MSE+1\sigma}$	
$\beta_0$	<b>-2.1621</b>	<b>-1.3120</b>	<b>-1.5711</b>	<b>-0.6450</b>
V2X	-0.1561	0.0925	-0.2627	0.0341
Vix	<b>-0.3377</b>	<b>-0.0946</b>	<b>-0.3137</b>	<b>-0.0688</b>
VixLagged	-0.1274	0.0948	0	
FXImpVolEUR1y	<b>0.7831</b>	<b>1.1297</b>	<b>0.3762</b>	<b>0.6595</b>
FXImpVolSEK1m	<b>-0.4435</b>	<b>-0.1952</b>	0	
Ger5yChange	<b>0.1033</b>	<b>0.5361</b>	0	
Skew	<b>-0.9688</b>	<b>-0.3778</b>	<b>-1.0153</b>	<b>-0.2889</b>
Stibor3m	<b>-0.5558</b>	<b>-0.3331</b>	<b>-0.4250</b>	<b>-0.2352</b>
Libor3m	<b>0.2087</b>	<b>0.2937</b>	<b>0.1966</b>	<b>0.3015</b>
Euribor3m	<b>0.7518</b>	<b>0.9846</b>	<b>0.6476</b>	<b>0.8983</b>
SwapVolEUR6m	<b>0.1929</b>	<b>0.4365</b>	0	
SwapVolSEK6m	<b>-0.7107</b>	<b>-0.4548</b>	<b>-0.4806</b>	<b>-0.2346</b>
SwapVolUSD6m	<b>0.7635</b>	<b>1.0128</b>	<b>0.6947</b>	<b>0.9635</b>
Swap5ySEK	<b>-1.6813</b>	<b>-1.0168</b>	<b>-1.1587</b>	<b>-0.8197</b>
Swap5yEUR	0		0	
Equity	<b>0.8946</b>	<b>1.0607</b>	<b>0.8949</b>	<b>1.0921</b>
EquityVol	<b>0.5619</b>	<b>0.6831</b>	<b>0.5034</b>	<b>0.6652</b>
Sector	<b>1.0339</b>	<b>1.4893</b>	<b>0.7906</b>	<b>1.2773</b>
Liquidity	<b>-0.0482</b>	<b>-0.0183</b>	<b>-0.0561</b>	<b>-0.0167</b>

Table 3: Explanatory power of percentage weighted subgroups for  $\lambda_{MSE+1\sigma}$  with relaxed elastic net with  $\lambda_{MSE}$  for 2006 - 2014.

Company	Market Sentiment	Firm Specific	Liquidity	$R^2$
AB Volvo	0.656	0.339	0.005	0.729
TeliaSonera AB	0.772	0.225	0.003	0.588
Stora Enso Oyj	0.805	0.182	0.013	0.821

Table 4: Explanatory power of percentage weighted subgroups for  $\lambda_{MSE+1\sigma}$  with relaxed elastic net with  $\lambda_{MSE}$  for 2011 - 2014 datasample.

Company	Market Sentiment	Firm Specific	Liquidity	$R^2$
AB Volvo	0.703	0.289	0.008	0.730
TeliaSonera AB	0.709	0.286	0.005	0.703
Stora Enso Oyj	0.675	0.314	0.011	0.758

## 6.2 Estimated $\Gamma$ and the implied PD under $\mathbb{P}$

Figure 5 shows the estimated  $\Gamma$  through multivariate regression against the derived market price of risk and has a  $R^2$  of 0.729. The resulting implied PD under the probability measure  $\mathbb{P}$  can be seen in Figure 6 for AB Volvo. The result for equivalent analysis for Stora Enso Oyj and TeliaSonera AB can be seen in Figure 7 and 8. As seen in Table 3 and 4 the subgroups of variables named market sentiment is shown to have the largest percentage impact on all firms. Liquidity is shown to have the largest effect on Stora Enso Oyj. The time-varying parameters estimated with a Kalman filter was shown to not vary very much through time, motivating the choice of estimating constant parameters, see Appendix A.4, Figure 13.

As seen in Figure 5 there is a huge peak in  $\Gamma$  in March 2008 in the 2006 - 2014 dataset, where Moody's fail to react as quickly as the market. The model fail to incorporate this huge peak and for that reason the errors have a fat tail. For the 2011 - 2014 where these outliers are absent the errors are nicely normally distributed. As the model fail to explain the peak it also fails to predict it, resulting in huge squared errors for the prediction. For the 2011 - 2014 set the model show decent prediction abilities for almost all windows. Figure 9 shows the resulting implied PD under  $\mathbb{P}$  for windows of size 30, 100, 300 and 500 sample points. The corrected MSE for the prediction of  $\ln(\Gamma)$  for these windows are 0.0066, 0.0080, 0.0230 and 0.0377.

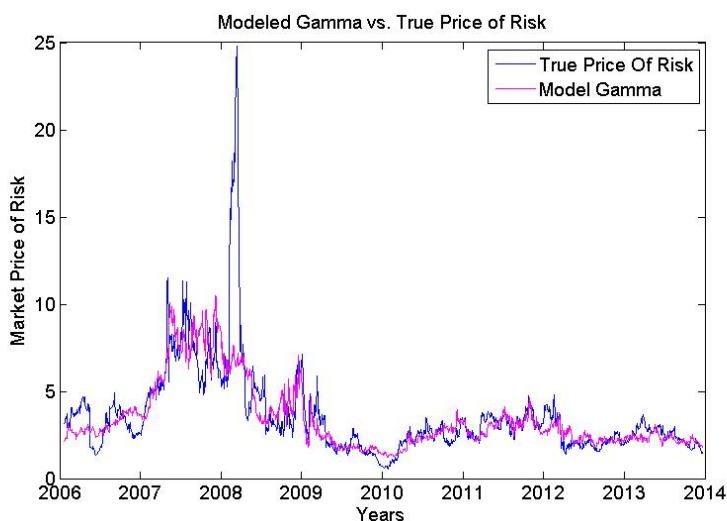


Figure 5: The figure shows the modeled  $\Gamma$  from market data against the true  $\Gamma$  for AB Volvo.



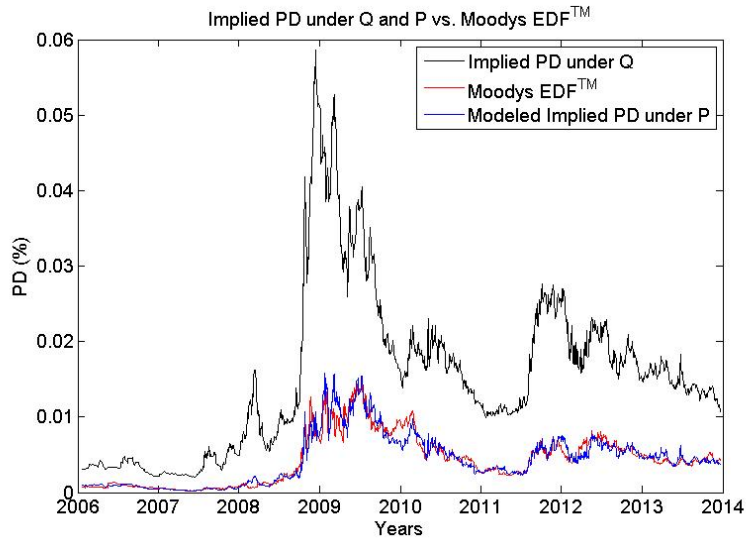


Figure 6: The figure shows implied PD under  $\mathbb{Q}$  and the implied PD under  $\mathbb{P}$  transformed with the modeled  $\Gamma$ , compared to the benchmark, Moody's EDF<sup>TM</sup> for AB Volvo.

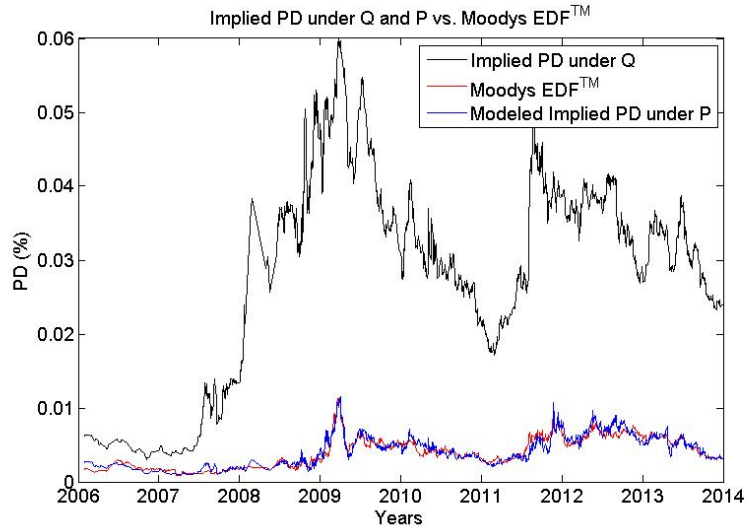


Figure 7: The figure shows implied PD under  $\mathbb{Q}$  and the implied PD under  $\mathbb{P}$  transformed with the modeled  $\Gamma$ , compared to the benchmark, Moody's EDF<sup>TM</sup> for Stora Enso Oyj.

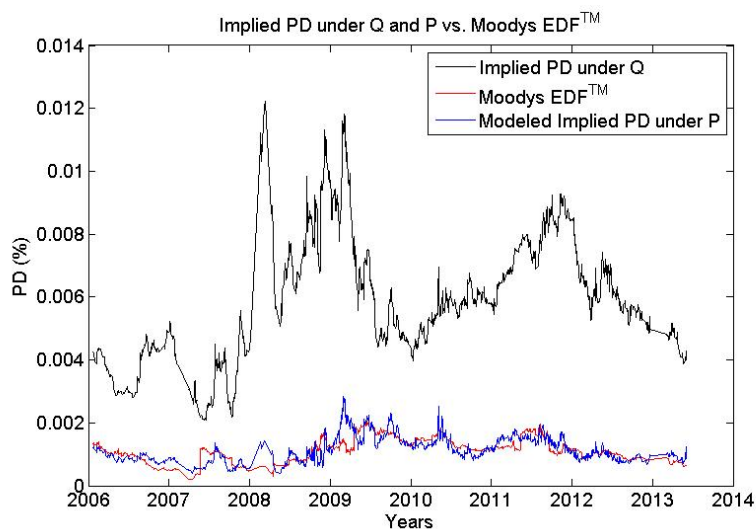


Figure 8: The figure shows implied PD under  $\mathbb{Q}$  and the implied PD under  $\mathbb{P}$  transformed with the modeled  $\Gamma$ , compared to the benchmark, Moody's EDF<sup>TM</sup> for TeliaSonera AB.

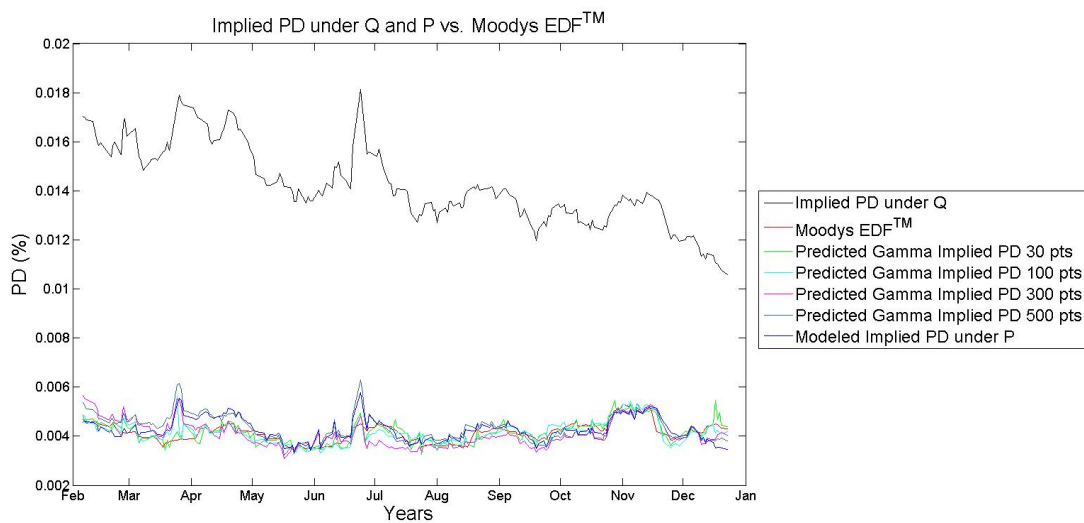


Figure 9: The figure shows implied PD under  $\mathbb{Q}$ , compared to Moody's EDF<sup>TM</sup> and the implied PD under  $\mathbb{P}$  transformed with the modeled and predicted  $\Gamma$  for windows of size 30, 100, 300 and 500 sample points for AB Volvo February - December 2013.

### 6.3 Model Sensitivity

To test how sensitive the model output was against changes in input data some tests were performed. The first test was to check how the model answers to a small change in the yield curve, second the reaction to CDS spread changes was analysed and lastly the implied PD was calculated for different recovery rates.

When the yield curve was moved upward 1 percent the difference between the implied PD's were very small. However, one can see though that the difference is larger during periods where the CDS-spreads change rapidly such as during the financial crisis, see Figure 10. This indicates that the model is a little bit more sensitive against changes in the yield curve during periods where the CDS-market is very volatile.

Pushing the credit curve up 1 basis point results in an average increase of  $9.5 \cdot 10^{-5}$  for the implied PD which is good because a higher spread should indicate a higher implied PD, see Figure 11.

The implied PD was calculated for recovery rates going from 0 - 0.95 with a distance of 0.01 between the calculations i.e. 96 implied PD's. All other parameters being held constant. The result shows that the model is more sensitive against large assumed recovery rates. As Figure 12 shows, a higher assumed recovery rate implies a higher implied PD and for very large recovery rates the increase seems to be almost exponential [Schönbucher, 2003].

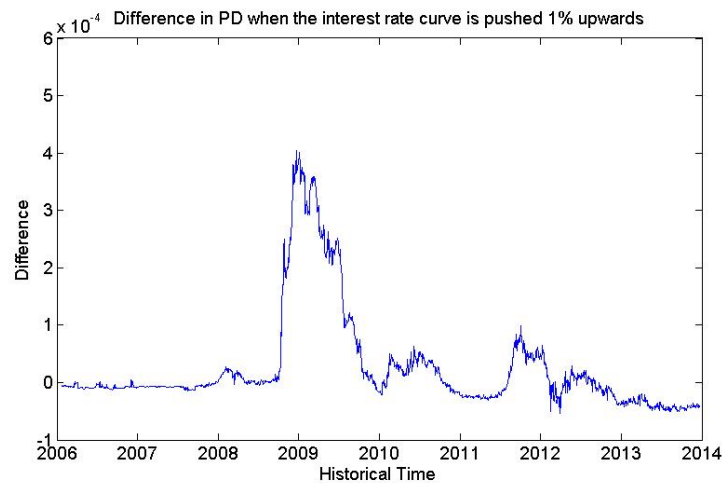


Figure 10: The Figure shows the original implied PD minus the implied PD when the interest rate curve is pushed 1 percent upwards, the calculation was made using CDS spreads from AB Volvo.

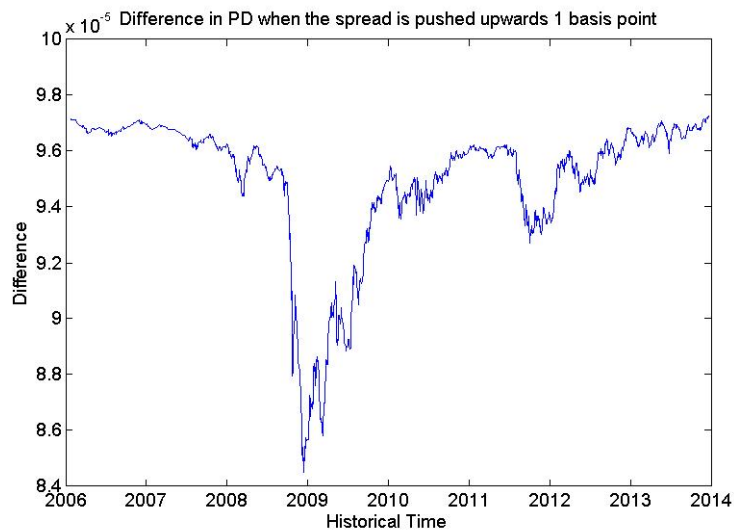


Figure 11: The figure shows the implied PD calculated when pushing the spread curve up 1 basis point minus the original implied PD, the calculation was made using CDS spreads from AB Volvo.

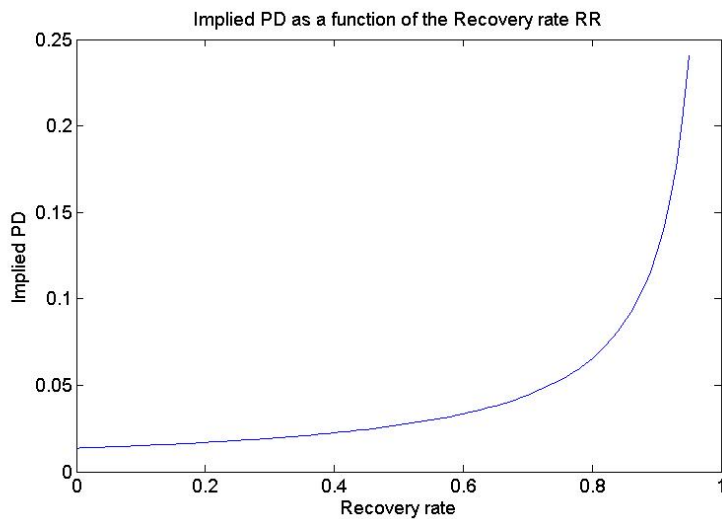


Figure 12: Implied PD as a function of the recovery rate calculated using 1, 3 and 5 year constant CDS spreads from AB Volvo, the values of the spreads where 94, 176 and 228 basis points.



## 7 Discussion

### 7.1 Overall Succession

The model was built with data for AB Volvo but tested on TeliaSonera AB and Stora Enso Oyj, rated A- respectively BB by S&P. The model performed very well with data from Stora Enso Oyj with a resulting  $R^2$  of 0.820 for the full set. However, the model was shown to be a worse fit for higher rated TeliaSonera AB with a large constant and a  $R^2 = 0.588$ . From this result one could argue that the model potentially could be better suited for lower rated companies. Since lower rating means larger risk with a higher risk premium demanded, these companies could be more effected by uncertainty in the market, thus having a larger  $\Gamma$  explainable by market data. This is supported by the fact that liquidity was shown to have the largest importance for high-yield rated Stora Enso Oyj. However, analysis should be made for more companies with a larger difference in rating before a conclusion of the argument can be made.

The time series of data analyzed contains much of the life of the CDS contracts and also contains a number of contractual changes during that period. This could potentially lead to some changes in the CDS spreads, not considered by the model. For example, the removal of restructuring as a credit event in some contracts in 2009 could lead to a lower risk neutral implied PD as the group of credit event you pay to be protected against are smaller. Also Moody's EDF<sup>TM</sup> has been subject to change during the time series. During the 2008 financial crisis the rating companies were subject to much critique and were accused of reacting to slowly. Moody's has since updated their EDF<sup>TM</sup> model, after which the measure can be seen to react more quickly to coming events [Markit, 2009]. For this thesis this means that the cross-correlation changes with time. However, for the full dataset the correlation is the highest in lag zero. Using this lag, a PD under  $\mathbb{P}$ , comparable to Moody's, was predicted with decent results.

### 7.2 Weaknesses in the Analysis

Points of weakness in the thesis include: the discretization of the integral in equation (31), the assumptions of a constant recovery rate of 40%, the absence of counterparty risk and the liquidity measure. As counterparty risk has been an increasing concern since after the 2008 financial crisis it is a major weakness that it, for this purpose, is assumed to be zero. Liquidity was shown not to have a very large impact on the market price of risk, although significant. However, as there lie some uncertainty about the quality of the the actual measure, the results might improve if a better measure was used. For example trading volume was discussed as a potential measure, but not available due to the OTC nature of the CDS market.

It is observed that some highly correlated variables gain similar non-zero parameters with the elastic net method but with different signs. Logically, these should add out in the linear-regression model thus limiting the use of them in the model. As elastic net chooses quite a large set of variables, this could be an additive method to reduce the complexity of the input although sacrificing a bit of the explanatory power.

By discretizing the integral in the implied PD calculations, it is assumed that default only can happen on equidistant monthly times. From a mathematical view one might seek to model default as an event that can happen at any time. However, the actual event of default is often not

something that happens from one day to another but rather at some specific dates. The thesis also assumes zero accruals, hence you will not get paid for the days between the premium payment days if a default should occur. Other has approximated the integral of the accruals with the trapezoidal rule. This basically implies one more payment date between the original dates, and the impact on the calculated implied PD should be minimal.

### **7.3 Suggested Future Work**

Future work on the subject would be to create an improved model including counterparty risk, able to forecast the market price of risk,  $\Gamma$ , using some of the factors evaluated in this thesis and potentially new ones. An attempt to make the model more general with less explanatory variables would also be of value as well as modeling of the recovery rate as a stochastic process. A good forecasting model would let you transfer from  $\mathbb{Q}$  to  $\mathbb{P}$ . This could open opportunities for trading as well as risk management.

## 8 Conclusion

This thesis has successfully explained the difference in implied PD from a reduced-form model, using CDS contracts, and a structural model, Moody's EDF<sup>TM</sup>. This is done by modeling the market price of risk,  $\Gamma$ , with a log-linear multivariate regression model and using it to transfer the implied PD from the risk neutral measure  $\mathbb{Q}$  to the physical measure  $\mathbb{P}$ .





## References

- [Ahmad & Wilmott, 2007] Ahmad, Riaz, & Wilmott, Paul. 2007. *The Market Price of Interest-rate Risk: Measuring and Modelling Fear and Greed in the Fixed-income Markets*. Wilmott magazine.
- [Amato, 2005] Amato, Jeffery D. 2005. *Risk Aversion and Risk Premia in the CDS Market*. BIS Quarterly Review.
- [Amato & Luisi, 2006] Amato, Jeffery D, & Luisi, Maurizio. 2006. *Macro Factors in the Term Structure of Credit Spreads, Working Paper*.
- [Berndt *et al.*, 2005] Berndt, Antje, Douglas, Rohan, Duffie, Darrell, Ferguson, Mark, & Schranz, David. 2005. *Measuring default risk premia from default swap rates and EDFs*. Bank for International Settlements, Monetary and Economic Department.
- [Bevras, 2006] Bevras, Arnaud. 2006. Market liquidity and its incorporation into risk management. *Financial Stability Review*, **8**.
- [Bielecki & Rutkowski, 2002] Bielecki, Tomasz R., & Rutkowski, Marek. 2002. *Credit Risk: Modeling, Valuation and Hedging*. Springer-Verlag Berlin Heidelberg.
- [Bielecki *et al.*, 2011] Bielecki, Tomasz R., Brigo, Damiano, & Patras, Frédéric. 2011. *Credit Risk Frontiers: Subprime Crisis, Pricing and Hedging, CVA, MBS, Ratings and Liquidity*. John Wiley and Sons Inc. Pages 587–615.
- [Björk, 2009] Björk, Tomas. 2009. *Arbitrage Theory in Continuous time*. Oxford University Press.
- [Brigo & Mercurio, 2006] Brigo, Damiano, & Mercurio, Fabio. 2006. *Interest Rate Models - Theory and Practice*. Springer.
- [CBOE, n.d.a] CBOE. *CBOE Skew Indexes*. [www.cboe.com/SKEW](http://www.cboe.com/SKEW). Retrived: 2014-04-08.
- [CBOE, n.d.b] CBOE. *Introduction to VIX Options and Futures*. [www.cboe.com/VIX](http://www.cboe.com/VIX). Retrived: 2014-04-08.
- [ECB, 2009] ECB, European Central Bank. 2009. *Credit Default Swaps and Counterparty Risk*.
- [Eurex, n.d.] Eurex. *Calculating the VSTOXX index*. [www.eurexchange.com/vstox/c\\_VSTOXX\\_Calculation.html](http://www.eurexchange.com/vstox/c_VSTOXX_Calculation.html). Retrived: 2014-04-08.
- [FCIR, 2011] FCIR, The Financial Crisis Inquiry Commission. 2011. *The Financial Crisis Inquiry Report*. Official United States Government Edition.
- [Fox, 2008] Fox, John. 2008. *Applied regression analysis and generalized linear models*. SAGE Publications. Chap. 21 - Bootstrapping Regression Models.
- [Hastie *et al.*, 2009] Hastie, Trevor, Tibshirani, Robert, & Friedman, Jerome. 2009. *The Elements of Statistical Learning - Data Mining, Inference and Prediction*. Second edn. Springer. Pages 61–73, 241–247.
- [Hull & White, 2000] Hull, John, & White, Alan. 2000. Valuing Credit Default Swaps I: No Counterparty Default Risk. *The Journal of Derivatives*, **8**(1), 29–40.

- [Hull & White, 2001] Hull, John, & White, Alan. 2001. Valuing Credit Default Swaps II: Modeling Default Correlations. *The Journal of Derivatives*, **8**(3), 12–21.
- [Hull *et al.* , 2005a] Hull, John, Predescu, Mirela, & White, Alan. 2005a. *Bond Prices, Default Probabilities and Risk Premiums*. Working Paper.
- [Hull *et al.* , 2005b] Hull, John C., Nelken, Izzy, & White, Alan D. 2005b. Merton’s Model, Credit Risk and Volatility Skews. *Journal of Credit Risk*, **1**.
- [ISDA, 2013] ISDA. 2013. *Product Description and Frequently Asked Questions*. <http://www.isda.org/educat/faqs.html>.
- [Longstaff *et al.* , 2003] Longstaff, Francis A., Mithal, Sanjay, & Neis, Eric. 2003. *The credit default swap market: is credit protection priced correctly?* Working paper.
- [Luo, 2005] Luo, Lawrence S. J. 2005. Bootstrapping default probability curves. *Journal of Credit Risk*, **1**(4), 174–175.
- [Madsen *et al.* , 2004] Madsen, Henrik, Nygaard, Jan, Lindström, Erik, Baadsgaard, Mikkel, & Holst, Jan. 2004. *Statistics in Finance*. Lund Institute of Technology, Centre for Mathematical Sciences. Pages 83, 105, 147.
- [Markit, 2009] Markit. 2009. *The CDS Big Bang: Understanding the Changes to the Global CDS Contract and North American Conventions*. [http://www.markit.com/cds/announcements/resource/cds\\_big\\_bang.pdf](http://www.markit.com/cds/announcements/resource/cds_big_bang.pdf).
- [O’Kane & Turnbull, 2003] O’Kane, Dominic, & Turnbull, Stuart. 2003. Valuation of Credit Default Swaps. *Lehman Brothers: Fixed Income Quantitative Credit Research*, **1**, 5–13.
- [Ravichandran & Prahmeshu, 2002] Ravichandran, S., & Prahmeshu. 2002. Time - varying state space regression model in fisheries. *Indian journal of fisheries*, **49**, 41–44.
- [Rawlings *et al.* , 2001] Rawlings, John O., Pantula, Sastry G., & Dickey, David A. 2001. *Applied Regression Analysis - A Research Tool*. Springer. Pages 9, 487.
- [Schönbucher, 2003] Schönbucher, Philipp J. 2003. *Credit derivatives pricing models*. Wiley.
- [Singh & Spackman, 2009] Singh, Manmohan, & Spackman, Carolyne. 2009. *The Use (and Abuse) of CDS Spreads During Distress*. Working Paper.
- [Sun *et al.* , 2012] Sun, Zhao, Munves, David, & Hamilton, David T. 2012. *Public Firm Expected Default Frequency (EDF) Credit Measures: Methodology, Performance, and Model Extensions*. Moody’s Analytics.
- [Tang & Yan, 2013] Tang, Dragon Yongjun, & Yan, Hong. 2013. *What Moves CDS Spreads*. Working Paper.
- [Tibshirani, 1996] Tibshirani, Robert. 1996. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*, **58**, 267–288.
- [Weistoffer, 2009] Weistoffer, Christian. 2009. *Credit Default Swaps - Heading towards a more stable system*. Deutsche Bank Research.

[White, 2013] White, Richard. 2013. *The Pricing and Risk Management of Credit Default Swaps, with a Focus on the ISDA Model*. OpenGamma Quantitative Research n.16.

[Åberg, 2010] Åberg, Sebastian. 2010. *Derivative Pricing*. Lund Univerisity, Centre for Mathematical Sciences Mathematical Statistics. Pages 176 – 182.



# A Appendix

## A.1 The Model Underlying Hypothesis 1

The initial model assumed that the noise in the CDS spread from the non-default component can explain the difference in PD from reduced form models, implied PD, and structured models, Moody's EDF<sup>TM</sup>. To evaluate which factors in the market that may explain the noise a linear regression model was used. The daily changes in the mid CDS spread was regressed on daily changes in the factors,  $\mathbf{X}$ , and the unexplained portion was assumed to be the daily changes in Credit Risk, i.e. the defaultable component. The same linear regression model was also tested on the absolute values of the spread and factors.

$$\Delta CDS_{Spread} = \Delta CR + \beta \Delta \mathbf{X}, \quad (36)$$

$$CDS_{Spread} = CR + \beta \mathbf{X}, \quad (37)$$

The correlation between each factor and the CDS spread was analyzed. In each data group the factors with the largest correlation to the data was chosen. Further variable selection was then performed with the relaxed lasso method and the  $R^2$  value was noted. A Kalman Filter was implemented to filter out a time varying  $\Delta CR$ . However, the linear regression model of the absolute spreads showed not to be sufficient to model the CDS spread as most of the variation was explained by the constant, thus indicating the bad fit of the model. The delta model performed slightly better but as the goal was to extract a "clean" spread to estimate a new PD, a transformation from daily changes to spread was needed. This offered many obstacles as the transformation depends on a start value as well as assumptions about the noise in the one year and three year CDS spreads, as they also are inputs to the PD-model. For these reasons, this model approach was discarded leading to *Hypothesis 2*.

## A.2 The GARCH Model

The GARCH model is a model describing conditional variance through an extension of ARMA-like structures for the squared process. While for the ARCH process, the conditional variance is specified as a linear function of past sample variances only, the GARCH process uses past values of conditional variance as well. The GARCH model is given by

$$\begin{aligned} \epsilon_t | \eta_{t-1} &\sim N(0, \sigma_t^2) \\ \sigma_t^2 &= \alpha_0 + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \end{aligned}$$

The coefficients  $(\alpha_i, \beta_i)$  must be non-negative to ensure positive variances, as well as

$$\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i < 1$$

to preserve stability [Madsen *et al.*, 2004].

### A.3 Bootstrapping Confidence Intervals for Elastic Net-Parameters

The only way of calculating confidence intervals for the parameters estimated by the elastic net method is to use the bootstrap. Bootstrapping can be performed in different kind of ways but for this thesis *residual bootstrap* was used [Fox, 2008].

Residual bootstrap is when you resample the estimated residuals of the model and add them to your estimated parameters and explanatory variables. The parameter estimation method is then done again calculating new parameters which are saved. If this procedure is repeated a lot of times (say 1000) then a histogram of the estimated parameters can be plotted and also bootstrapped confidence intervals can be calculated by taking the quantiles of the parameters. Assume you have a model such as

$$y = X\beta + \epsilon \quad (38)$$

where  $y$  is a  $(n \times 1)$  vector of target variables,  $\beta$  is a  $(k \times 1)$  vector of parameters,  $X$  is a  $(n \times k)$  matrix of explanatory variables and  $\epsilon$  is a  $(n \times 1)$  vector of residuals. The different steps in the residual bootstrap are

1. Calculate the parameters  $\hat{\beta}$  using the elastic net method.
2. Calculate the estimated residuals  $\hat{\epsilon} = y - \hat{\beta}X$ .
3. Resample  $\hat{\epsilon}$  with replacement creating  $\hat{\epsilon}_{\text{resampled}}$ .
4. Create a new vector of target variables  $y^*$  by using  $X$ ,  $\hat{\beta}$  and  $\hat{\epsilon}_{\text{resampled}}$ ,  $y^* = X\hat{\beta} + \hat{\epsilon}_{\text{resampled}}$ .
5. Use  $y^*$  to calculate and save new parameters  $\hat{\beta}^*$  using the elastic net method.
6. Repeat steps 3 - 5 lots of times (at least 1000).
7. The residual bootstrap is done and bootstrapped distributions for the parameters can be examined.

### A.4 Time Varying Parameters in Linear Regression Using the Kalman Filter

If the parameters in the regression model, equation (34), are assumed to be time varying, they can be estimated using a Kalman filter. The problem can be described in the following state space form

$$\begin{cases} y_t = X_t\beta_t + \epsilon_t \\ \beta_t = A\beta_{t-1} + \eta_t \end{cases} \quad (39)$$

Where  $y_t$  is the observed process and  $\beta_t$  is the assumed dynamics of the parameters which is a latent process,  $A$  is the identity matrix,  $\epsilon_t$  and  $\eta_t$  are assumed to be normally distributed [Ravichandran & Prahmeshu, 2002]. To estimate the parameters i.e. filter out the latent process the following Kalman filter was used

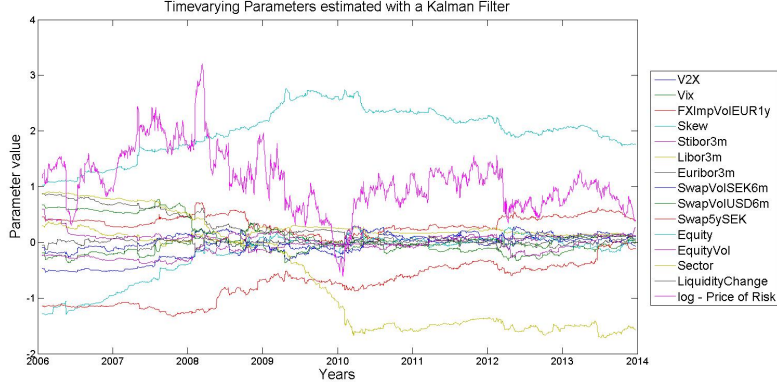


Figure 13: The estimated time varying parameters for AB Volvo 2006-2014 and the log-price of risk in pink.

*Prediction*

$$\beta_{t|t-1} = E[\beta_t | \mathcal{F}_{t-1}] = A\beta_{t-1|t-1}$$

$$P_{t|t-1} = AP_{t-1|t-1}A^T + Q$$

*Prediction error and its variance*

$$p_{err} = Y_t - X_t\beta_{t|t-1}$$

$$V_{err} = X_tP_{t|t-1}X^T + R$$

*Update*

$$K_t = P_{t|t-1}X_t^T V_{err}^{-1}$$

$$\beta_{t|t} = \beta_{t|t-1} + K_t p_{err}$$

$$P_{t|t} = P_{t|t-1} - K_t X_t P_{t|t-1}$$

Where  $P_t$  is the covariance matrix,  $K_t$  is the so called Kalman gain,  $Q$  and  $R$  comes from the assumption about  $\epsilon_t$  and  $\eta_t$  i.e.  $\epsilon_t \sim N(0, R)$  and  $\eta_t \sim N(0, Q)$ .



## A.5 Deriving Premium- and Protection leg

Using  $\mathcal{F}_t$  equation (12) is rewritten as

$$\begin{aligned}
CDS_{a,b}(t, S, RR) &= \frac{\mathbb{1}_{\{\tau > t\}}}{\mathbb{Q}(\tau > t | \mathcal{F}_t)} \mathbb{E}^{\mathbb{Q}} (\Pi_{a,b}(t) | \mathcal{F}_t) \\
&= \frac{\mathbb{1}_{\{\tau > t\}}}{\mathbb{Q}(\tau > t | \mathcal{F}_t)} \left\{ S \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(t)}{B(T)} (\tau - T_{\beta(\tau)-1}) \mathbb{1}_{\{T_a < \tau < T_b\}} | \mathcal{F}_t \right] \right. \\
&\quad + \sum_{i=a+1}^b \Delta_i S \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(t)}{B(T_i)} \mathbb{1}_{\{\tau > T_i\}} | \mathcal{F}_t \right] \\
&\quad \left. - (1 - RR) \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{T_a < \tau \leq T_b\}} \frac{B(t)}{B(\tau)} | \mathcal{F}_t \right] \right\}
\end{aligned} \tag{40}$$

If we were to calculate the the fair spread by setting equation (40) to zero the solution would technically be defined for  $\tau > t$ , but since the value of  $S$  does not matter if  $\tau < t$  the indicator function in de beginning of the above expression can be ignored [Brigo & Mercurio, 2006]. From equation (40) the premium- and protection leg can be identified. The premium leg corresponds to the cash flows coming from the buyer of the CDS contract and the protection leg corresponds to the ones from the seller. The premium- and protection leg is defined at time  $s < t$  in equations (41) and (42) respectively. This is done assuming independence between the stochastic discount factors  $B(s)/B(t)$  and the time of default  $\tau$

$$\begin{aligned}
PV_{\text{PremiumLeg}}^{a,b} &= S \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(\tau)} (\tau - T_{\beta(\tau)-1}) \mathbb{1}_{\{T_a < \tau < T_b\}} | \mathcal{F}_t \right] \\
&\quad + \sum_{i=a+1}^b \Delta_i S \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(T_i)} \mathbb{1}_{\{\tau > T_i\}} | \mathcal{F}_t \right] \\
&= S \int_{t=T_a}^{T_b} \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(t)} \right] (\tau - T_{\beta(\tau)-1}) \mathbb{Q}(\tau \in [t, t + dt]) \\
&\quad + S \sum_{i=a+1}^b \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(T_i)} \right] \Delta_i \mathbb{Q}(\tau \geq T_i)
\end{aligned} \tag{41}$$

$$\begin{aligned}
PV_{\text{ProtectionLeg}}^{a,b} &= (1 - RR) \mathbb{E}^{\mathbb{Q}} \left[ \mathbb{1}_{\{T_a < \tau \leq T_b\}} \frac{B(s)}{B(\tau)} | \mathcal{F}_t \right] \\
&= (1 - RR) \int_{t=T_a}^{T_b} \mathbb{E}^{\mathbb{Q}} \left[ \frac{B(s)}{B(t)} \right] \mathbb{Q}(\tau \in [t, t + dt])
\end{aligned} \tag{42}$$

With the definitions and statements above, equation (12) can now be defined according to

$$CDS_{a,b}(t, S, RR) \stackrel{\text{def.}}{=} PV_{\text{PremiumLeg}}^{a,b} - PV_{\text{ProtectionLeg}}^{a,b}. \tag{43}$$

For the more interested reader, see [Brigo & Mercurio, 2006] for details.

## A.6 Summary of Data

Table (5) is an overview over the firms analysed in the thesis, Table 6 is an overview of the market factor data used.

Table 5: The table shows a summary of the firm data used, source Bloomberg.

Company	S&P rating	Moody's rating	Sector	IG/HY
AB Volvo	BBB	Baa2	Industrials	IG
TeliaSonera AB	A-	A3	Telecom	IG
Stora Enso Oyj	BB	Ba2	Basic Materials	HY

Table 6: List of factor data including Bloomberg tickers used

Bloomberg ticker	Description	mean	std	min	max
VIX	Volatility index on S&P500	20.60	10.23	9.89	80.86
V2X	Volatility index on EuroStoxx 50	24.06	10.02	11.60	87.51
SKEW	CBOE tail risk from S&P500	120.20	5.34	106.43	143.20
USDSEKV1y	FX Implied Volatility USD/SEK, expiry 1y	12.84	3.35	7.40	22.33
USDSEKV1m	FX Implied Volatility USD/SEK, expiry 1m	12.60	4.30	6.38	32.28
EURUSDV1y	FX Implied Volatility EUR/USD, expiry, 1y	10.88	3.00	5.55	19.86
EURUSDV1m	FX Implied Volatility EUR/USD, expiry, 1m	10.25	3.61	4.65	28.63
EUSV0F5	Swaption Implied Volatility EUR, expiry 6m	30.43	14.12	10.70	66.04
SKSV0F5	Swaption Implied Volatility SEK, expiry 6m	26.40	9.25	11.20	55.80
USSV0F5	Swaption Implied Volatility USD, expiry 6m	36.26	14.49	11.90	72.90
GDBR5	5y Gernamny Govenment Bond Index	2.36	1.28	0.24	4.76
SKSW5	5y Swap Rate SEK	3.13	1.00	1.48	5.63
EUSA5	5y Swap Rate EUR	2.82	1.18	0.71	5.20
EUR003M	EURIBOR 3 months	2.01	1.56	0.18	5.39
US0003M	LIBOR USD 3 months	2.09	2.05	0.23	5.73
STBB3M	STIBOR 3 months	2.27	1.31	0.47	5.60
BWINDU:IND	Bloomberg World Industrial Index	167.81	29.51	83.34	239.87
BWBMAT:IND	Bloomberg World Basic Materials Index	199.85	41.81	96.39	316.80
BWCOMM:IND	Bloomberg World Communication Index	135.42	19.08	83.71	174.08
VOLVA.SS	Volvo A Equity	82.93	22.51	31.10	153.00
TLSN.ST	TeliaSonera Equity	44.96	6.92	29.82	59.76
STER.SS	Stora Enso Equity	69.43	25.18	27.36	124.73