A semi-parametric Probability of Default model

Master Thesis submitted to the Stockholm School of Economics

by

Can Ertan[1] and Alex Gansmann[2]

Stockholm School of Economics

Winter semester 2015

[1] 40804@student.hhs.se

[2] 40738@student.hhs.se

A semi-parametric Probability of Default model

Can Ertan and Alex Gansmann
Stockholm School of Economics

Master Thesis submitted for the degree of Master of Science

Winter semester 2015

Abstract

With the implementation of IFRS 9, a new set of impairment rules will be effective as of 1st January 2018. We analyse alternative models for probability of default (PD) estimation that are in accordance with IFRS 9. In our model, PD is dependent on idiosyncratic firm-specific factors and systematic macroeconomic conditions. In order to identify the macroeconomic conditions that affect PD, we fit a semi-parametric Cox Proportional Hazards model to default data in a similar fashion to Figlewski et al. (2012). Subsequently, in line with Chen et al. (2005) and Kim and Partington (2014), we use a SAS macro programme to calculate PDs. We found that the inclusion of macroeconomic covariates in the regression increases explanatory power and improves the regression results. The regression results were transformed into PDs and we calculated PDs for each business partner. With an 'Area Under the Curve' value of 0.87, our model is able to accurately predict the business partners that will default within the next 12 months. With this study, we present a good foundation for the implementation of a new model in line with IFRS 9.

We would like to thank Maria Kim at the University of Wollongong, Australia, for providing us with her SAS macro programme and test data. Her support helped us refine our thinking on Survival Analysis and probability of default modelling. Furthermore, we are grateful to Alexander Kliushnyk from Nordea who never got tired to align our thinking with the thinking of actual practitioners in the field of credit loss provisioning. Lastly, we would like to thank Michael Halling from the Stockholm School of Economics for his invaluable advice and support.

Contents

List of Figures

List of Tables

Abbreviations

ADF	Actual Default Frequency
AUC	Area Under the (Receiver Operating Characteristic) Curve
EAD	Exposure at Default
ECL	Expected Credit Losses
FN	False Negative
FP	False Positive
IASB	International Accounting Standards Board
IFRS	International Financial Reporting Standards
LGD	Loss Given Default
PD	Probability of Default
PIT	Point-in-Time
POCI asset	Purchased or Originated Credit-Impaired asset
ROC	Receiver Operating Characteristic
Sensitivity	True Positive Rate
Specificity	True Negative Rate
TN	True Negative
TNR	True Negative Rate
TP	True Positive
TPR	True Positive Rate
TTC	Through-the-Cycle
TVC	Time-varying covariate

1 Introduction

The International Accounting Standards Board (IASB) issued the final version of the new accounting standard IFRS 9 on the 24th of July 2014 in order to improve the overall performance of the International Financial Reporting Standards (IFRS). IFRS 9 will become effective as of January 2018 and one objective of the new standard is to adapt the impairment rules from the old accounting standard IAS 39 to the lessons learned during the financial crisis of 2008. If financial institutions want to comply with the new standard, they need to revise existing credit loss models or conceive new modelling approaches. Due to the novelty of the standard and the strong differences to the old accounting standard, practitioners have not agreed on best practices in credit loss modelling yet.

In this study, we set out to enrich the academic literature by presenting a credit risk model which is in line with IFRS 9.

IFRS 9 introduces an expected loss impairment model that mandates more timely recognition of expected credit losses. In order to do so, probability of default (PD) calculations are expected to include information about past events, current conditions and expectations on future economic conditions (IASB, 2015). PDs that are in line with IFRS 9 show a strong dependence on short-term factors such as the current credit cycle and are contingent on a specific point in time. Therefore, IFRS

9 PDs are often referred to as point-in-time (PIT) PDs and are in contrast to regulatory PDs (Althoff and Lee, 2015). Regulatory PDs such as the ones set forth by the Internal Ratings Based Approach in the Basel II accords are through-the-cycle (TTC) PDs and show long-term trends in default probabilities.
Engelmann and Porath (2012) connect the general absence of models that incorporate macroeconomic information with the Internal Ratings Based Approach. They argue that there was no need to account for macroeconomic variables because the regulatory capital calculation is based on internal ratings. In contrast, they advocate using macroeconomic variables to calculate PIT PDs. With the implementation of Basel III and IFRS 9 on the horizon, the need to incorporate macroeconomic data into PD estimation has emerged.

Up to the present day the single-period logistic regression model has been a popular and widespread forecasting model among practitioners. Logistic regression models are popular because they are intuitive and simple to use. At the same time, there is some empirical evidence that suggests dynamic hazard models are able to improve or at least replicate the accuracy of static models (see e.g. Shumway, 2001; Bellotti and Crook, 2009).
As Shumway points out, bankruptcy occurs at rare intervals and the majority of static models disregard available information on healthy firms that eventually default further down the road. Instead, he argues that hazard models exhibit higher explanatory power for forecasting default than static models because hazard models can incorporate non-event observations, e.g. a business partner that does not default. This is especially true for a credit loss model that is in line with IFRS 9. One of the premises of IFRS 9 is the calculation of probability of default over the

lifetime of a business partner, the lifetime PD. When using small datasets, static logistic regression models have difficulties including enough data observations to build reasonably big estimation windows for these lifetime PDs.

Moreover, Bellotti and Crook (2009) directly compared the logistic regression with the Cox Proportional Hazards model and found that the Cox Proportional Hazards model delivered superior results compared to the logistic regression model, while the addition of macroeconomic variables further improved the predictions. A hazard model incorporates the most recent credit data and accounts for time to approximate a business partner's current financial health. These properties make a hazard model more dynamic and flexible in comparison to a static regression model.

In the light of these discussions we decided to develop a Cox Proportional Hazards model, that relies both on idiosyncratic and systematic explanatory variables.[1] The model is applied to credit default data obtained from the Stockholm-based bank Nordea. Our model shows high predictive power with an Area under the curve (AUC) value of 0.87.

The remainder of this study is structured as follows. In Section 2, we present the general structure of impairment models and give a more explicit overview of IFRS 9. Subsequently, we outline the semi-parametric hazard model and highlight other potentially suitable PD models in Section 3. The implementation of our PD calculation is presented in Section 4. In Section 5, we describe the dataset before we show the results of our model analysis in Section 6. In Section 7, we conclude and summarise our findings.

[1] The Cox Proportional Hazards model is referred to as the Cox model and Cox regression refers to the regression of the Cox Proportional Hazards model throughout the text.

2 Impairment models

In order to ensure continuity, a financial institution typically builds loan loss provisions for expected losses due to default or impairment. As set forth by the Basel Committee (2004) in its Explanatory Note on the Basel II IRB Risk Weight Functions, Expected Credit Losses are calculated using the following formula:

$$ECL = EAD \; * \; PD \; * \; LGD \tag{2.1}$$

These three input factors correspond to the risk parameters of the regulatory Basel II Internal Ratings Based Approach and are essential for the calculation of Expected Credit Losses (ECL). The Exposure at Default (EAD) is an estimation of the amount outstanding in case of default. The Probability of Default (PD) indicates the estimated average percentage of obligors that default per rating grade. The Loss Given Default (LGD) gives the estimated percentage of exposure that the bank loses in case of default.

In the course of this study we focus on probability of default models. Since the shift from IAS 39 to IFRS 9 affects PD to a large extent, such an analysis is justified.

2.1 Impairment according to IFRS 9

In order to understand the International Accounting Standards Board's (IASB's) shift towards IFRS 9, the historical development of the accounting standard should be considered. In 2005, together with IAS 39 an incurred loss approach to loss provisioning became effective in the European Union. In an incurred loss model approach, asset impairments are recognised when there is reasonable certainty that future cash flows cannot be collected. As pointed out by Gebhardt et al. (2011) the IASB introduced IAS 39 with the intention of stipulating unified rules for reporting of financial instruments and creating transparency as well as consistency. Before the introduction of IAS 39, more principle-based local GAAP allowed considerable management discretion and the standard aimed to reduce the extent of this expert judgment and management discretion.

During the financial crisis of 2008 and the subsequent European sovereign debt crisis, the deficiencies of IAS 39 surfaced. A reduction in discretionary behaviour limited management's ability to signal private information. At the same time, banks did not recognise loan portfolio losses in a reasonably short space of time. As a result, public markets were not informed about asset deteriorations triggered by expected events on a timely basis. Moreover, critics emphasised conceptual inconsistencies in IAS 39. Risk premia, incorporated in interest rates, were promptly recognised in net income, whereas the loan loss recognition was postponed until the actual default event. This practice aggravated the procyclicality of banks' earnings because delayed recognition resulted in higher earnings early on and in lower earnings in later years, especially during busts (Gebhardt et al., 2011). In addition to the aforementioned shortcomings, the IASB argues that preparers of financial statements found reporting financial instruments complex due to the IAS 39's rule-based requirements (IASB, 2015).

As a consequence of the experiences with IAS 39, the IASB set forth a new principle-based standard, IFRS 9. The new accounting standard aims to adapt the accounting standards to the realisations made during the financial crisis as outlined above. IFRS 9 is partitioned into three sections: Classification & Measurement, Impairment and Hedge Accounting. For the purpose of this study we focus on section two, Impairment, since it is most relevant for credit loss provisioning.

From January 2018 onwards, risk managers are expected to prepare an expected loss impairment model that mandates more timely recognition of expected credit losses. IFRS 9 PDs should include information about past events, current conditions and expectations on future economic conditions (IASB, 2015). Consequently, PDs will become more forward-looking and future losses can be accounted for proactively, before the actual impairment occurs.

While we give a short outline here, we refer to the project summary by the IASB from July 2014 for a detailed exposition of IFRS 9.

IFRS 9 provides a three-stage approach to impairment loss provisioning: At initial recognition non-credit-impaired assets are recognised with 12-month Expected Credit Losses (ECL) and interest revenues based on the gross carrying amount of the asset. As shown in the beginning of this section, the 12-month ECL is calculated as the total credit loss weighted by the probability that the loss will materialise within the next 12 months. This stage is commonly referred to as stage 1.

If at any time during the lifecycle of the financial asset a significant deterioration of

credit risk is recognised, lifetime ECL are to be estimated.[1] In case the financial asset does not have objective evidence of impairment post asset deterioration, the asset is allocated to stage 2. In stage 2, interest revenue is still based on the gross carrying amount of the asset. If, on the other hand, an objective evidence of impairment is recognised, the asset is assigned to stage 3 and the interest revenue is calculated net of credit losses (Althoff and Lee, 2015).
If a significant increase in credit risk reverses in a subsequent reporting period, expected credit losses revert to the 12-month ECL.

In addition to the framework described above, a number of exemptions and special accounting rules are stipulated in IFRS 9. Due to its immediate relevance for the three-stage approach one exemption is presented in greater detail:
Assets which are credit-impaired at initial recognition are so-called Purchased or Originated Credit-Impaired (POCI) financial assets. POCI assets are treated differently than non-POCI assets. At initial recognition and throughout the lifecycle of the financial assets, POCI assets are always recognised at lifetime ECL. Any changes in lifetime ECL are recognised as a loss allowance and affect profit or loss.
IFRS 9 features many more exemptions or special rules as the ones described above. The purpose of this study is not to specify proper implementation of all the details set forth in IFRS 9. Instead, we focus on a conceptual implementation of IFRS 9. In the following section we describe the generalisations we made in order to be able

[1] According to Appendix A of IFRS 9, a significant increase in credit risk can be due to any of the following: significant financial difficulty of the borrower or high likelihood of bankruptcy, a breach of contract (e.g. past-due event), the borrower was granted concessions that would not be considered in normal business conditions, a disappearance of an active market for a financial asset or a purchase or origination of a financial asset at a deep discount that reflects incurred losses.

to come up with a conceptual PD model.

2.2 Generalised IFRS 9 impairment

We attempt to identify a model that is able to solve the conceptual difficulties of impairment according to IFRS 9. Therefore, we disregard any special accounting rules from IFRS 9 that do not directly affect PD calculations.

To facilitate understanding we use a two-stage approach (stages 1 and 2) and disregard stage 3 of IFRS 9. The difference in stages 2 and 3 is in the interest revenue calculation and the probability of default calculation in these stages remains unchanged. Therefore, in order to identify an appropriate PD model, a two-stage approach is sufficient. For a new business partner, the 12-month PD is calculated at initial recognition. In the subsequent period, it is analysed whether the business partner suffered a significant deterioration of credit risk. If the business partner experienced a significant deterioration of credit risk, the model calculates the lifetime PD. In the next and all future periods the credit risk is reassessed and the respective PD is calculated. Due to their low creditworthiness, we identified all business partners that entered the database in the lowest three out of 18 rating groups as being POCI assets. Additionally, we argue that business partners that enter the database at default should also be treated as POCI assets. In order to include as much data as possible into our 12-month PD model validation, we decided to treat POCI assets as any other asset class and recognised these business partners at 12-month PD at initial recognition.

For a visualisation of the process, see Figure 2.1 below.

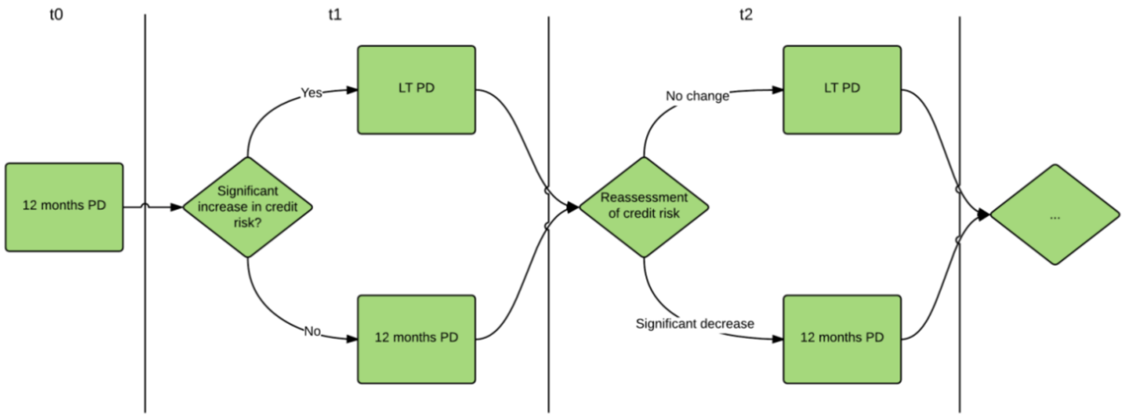

Figure 2.1: A simplification of IFRS 9 is presented here. A business partner that enters the dataset receives a 12-month PD. After one period a threshold criterion assesses whether a significant deterioration in credit risk occurred. If a deterioration occurred, the business partner's expected credit losses are calculated based on a probability of default over the entire lifetime. In the third period, the threshold criterion reevaluates the business partner's current creditworthiness. The same process takes place in all future periods.

While we outline a valid implementation of both the threshold for significant deterioration of credit risk and the lifetime PD in Section 4, we focus our analysis on the validation of the 12-month PDs due to the limited sample size of our dataset.

3 Modeling of default risk

In this section, we outline why we decided to implement a Cox model and elaborate on some of the alternative models that have emerged in the literature of default risk. With regard to IFRS 9, there are at least three general PD model concepts that are qualified for use in credit loss provisioning. These are dynamic models, static models and transformation models. Under dynamic models, there are two broad classes of survival analysis models that are suitable for analysing credit default: fully parametric proportional hazards models and semi-parametric proportional hazards models. In Section 3.1, we first provide an overview of hazard models, discuss their advantages and present the semi-parametric Cox model. In Section 3.2 to Section 3.4 we discuss fully parametric hazard models, static logistic regression models and transformation models.

3.1 Cox Proportional Hazards Model

Survival analysis is a branch of statistics that deals with collecting data and analysing the period of time until one or more events happen. The aim of survival analysis is to estimate the time to event, which is called survival time, as well as identifying a relationship between a cause and the event. Historically, the event of interest was death and the first known applications of survival analysis were mortality tables,

which showed, given that one had survived until the current period, the probability of surviving to the next period. Hazard models are a category of models in survival analysis and have the feature to account for time, making them dynamic models as opposed to static models. They have the ability to incorporate time-varying covariates, i.e. explanatory variables that can assume different values for different time periods. Event analysis in hazard models is not based on calendar time but on event time intervals, which are referred to as 'spells'. We define spells as the amount of time t a business partner spends in a rating class. A spell ends when there is a change in the rating class, a default occurs or when the business partner leaves the dataset. When the business partner enters another rating group as a result of any of the aforementioned, a new spell starts in the given time period.
A visualisation of the differences between event time and calendar time analysis (without sorting by duration) is shown in Figure 3.1 below.

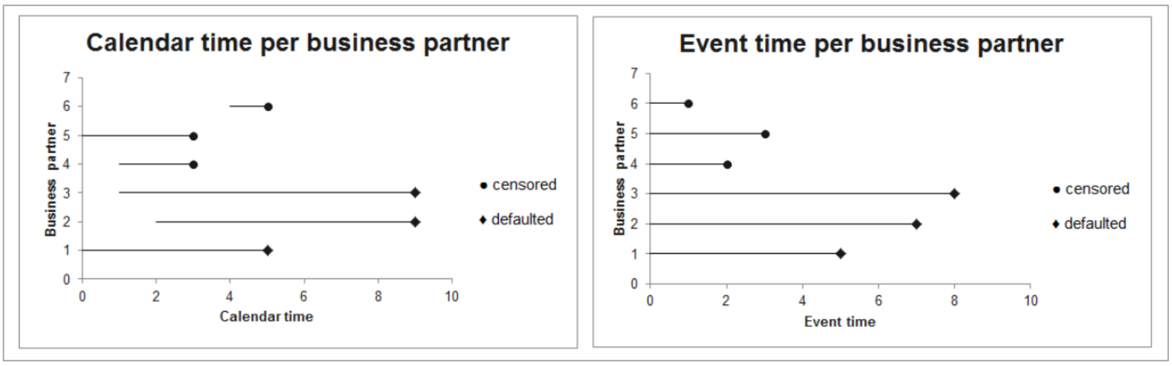

Figure 3.1: Calendar and event time

In Figure 3.1, we see that a spell can end either by a default event or by being censored. The latter is classified as right-censored data, which indicates that the subject has left the dataset without the occurrence of an event. This can be either due to the end of the observation period or because the subject is not in the risk set anymore, e.g. a loan was fully repaid or the rating of the subject has changed. Hazard models are distinctive in that they are able to incorporate these non-event observations. For credit data, the number of periods in which companies are healthy composes the vast majority of data points when compared to the number of default periods. Therefore, the ability to incorporate non-default data significantly increases the explanatory power of the model and thus improves results. This is especially true for a credit loss model that is in line with IFRS 9. One of the premises of IFRS 9 is the calculation of the lifetime PD for business partners with a significant deterioration of credit risk. When using small datasets, static logistic regression models have difficulties including enough data observations to build reasonably big estimation windows for longer PD horizons.

Moreover, the use of contemporary variables in the model increases the 'point-in-timeness' of PD estimation, as it adjusts to the changes of the variables throughout the past until now. In other words, as the factors of a business partner change over time, a dynamic model incorporates the most recent data to estimate the business partner's current financial health. Thus, among the advantages of hazard models over static models are their ability to build relationships between age and credit impairment as well as to account for changes over time. In static models, business partners that have the same factors would receive the same PD, as the PD solely depends on the coefficients of the factors. However, hazard models extend to time-varying covariates and incorporate the longevity of the current event horizon. Both of these considerations have a significant effect on the default probability of a business partner. Consider two business partners with the same covariates but different event horizons (e.g. one business partner entered the dataset one year ago, the other two years ago). Using a hazard model for estimating the 12-month PD, we get distinct PDs for each business partner. Static models are not able to account for this effect.

A wide literature has been published in the field of credit modelling that advocates the use of hazard models. Due to the above mentioned advantages, Shumway (2001) argues that hazard models are more appropriate for forecasting default than static models. Engelmann and Porath (2012) advocate the necessity of using macroeconomic variables to calculate point-in-time (PIT) PDs, which incorporate the current state of the economy. In order to show the effect of macroeconomic variables, they compare a logit model, a Bayesian approach incorporating macroeconomic information into the logit model and finally a hazard model with macroeconomic covariates. The Bayes model with macroeconomic information outperforms the simple logit model without the macroeconomic variables, whereas the hazard model further improves the estimation of PDs. Consequently, Engelmann and Porath (2012) argue that including macroeconomic information adds value to the PD estimation, and that hazard models deliver superior results to static models. Moreover, Bellotti and Crook (2009) directly compare the logistic regression with the Cox model. In their study, the Cox model delivered superior results compared to the logistic regression model, while the addition of macroeconomic variables further improved the predictions. In the light of this discussion, we have decided to implement the Cox model.

The Cox Proportional Hazards Model was introduced by Cox (1972) in the field of biomedicine. Since then it has been widely used in biomedicine, medicine, social sciences, engineering and economics. Furthermore, it has recently been popular in the field of finance. The expression of the model is as follows:

$$h_i(t) = h_0(t) * exp\left\{\sum_{j=1}^{p} \beta_j z_j^i(t)\right\}, \tag{3.1}$$

where $h(t)$ is the hazard function that describes the rate of change in the probability of failure at time t. $h_0(t)$ is the baseline hazard function, $z_j^i(t)$ denotes the value of the jth explanatory variable of business partner i at time t and β_j describes the respective regression coefficient.

The Cox model is defined as semi-parametric because it contains the unspecified non-parametric time-dependent part $h_0(t)$ and the parametric part $exp\left\{\sum_{j=1}^{p} \beta_j z_j^i(t)\right\}$. The latter is referred to as the hazard ratio and can be derived by rearranging terms:

$$\frac{h_i(t)}{h_0(t)} = exp\left\{\sum_{j=1}^{p} \beta_j z_j^i(t)\right\}, \tag{3.2}$$

The hazard ratio is the ratio of the hazard rate over the baseline hazard rate. It can be noted that the hazard ratio does not depend on the baseline hazard function and if all the covariates $z_j^i(t)$ are constant, i.e. not time-dependent, the hazard ratio is constant over time. Due to this property the conventional Cox Hazard model is commonly referred to as the Cox Proportional Hazards model. However, for time-varying covariates (TVC) the hazard ratio changes with t and the proportionality assumption of the Cox model is no longer maintained. The hazard ratio can be interpreted as the instantaneous risk of an event occurrence with the effect of a variable over the control arm, i.e. without the effect of the variable. For instance, if we keep all other variables equal, the one-unit increase in a variable with a hazard ratio of two indicates that the hazard rate would be twice as high as in the case without the unit increase of the variable. The estimation of the regression coefficients β_j is done with the partial likelihood approach, a powerful tool introduced by Cox with the following form:

$$PL(\beta_j) = \prod_{i=1}^{m} \frac{exp\left\{\sum_{j=1}^{p} \beta_j z_j^i(t)\right\}}{\sum_{k \in R_i(t)} exp\left\{\sum_{j=1}^{p} \beta_j z_j^k(t)\right\}} \tag{3.3}$$

In Equation 3.3, t denotes distinct lengths of spells that end with a credit event, i is the business partner at default and k is the business partner in the risk set at time t. So, all ts represent time periods with a distinct survival time. If we initially assume that there are no simultaneous credit events, then each t would refer to a single firm's credit event. The Cox model takes time as continuous and therefore it is assumed that any number of events D occurring at a particular instant t is in the form $D(t) = (0, 1)$. However, in the case of defaults, data may be collected daily, weekly or, as in our case, monthly, which leads to tied data where several firms experience an event at the same time t and $D(t) \geq 0$. There are different ways in the literature to deal with this issue, but we mainly considered four alternatives: the Breslow approximation, the Efron approximation, the Kalbfleisch and Prentice exact expression and the discrete method. Borucka (2014) analyses the different methods for handling ties and argues that the Breslow method causes a severe bias in cases when tied data is high, whereas the Efron approximation and the exact expression are relatively unaffected. According to Borucka, the exact expression method presented by Kalbfleisch and Prentice (2011) has yielded overall the best results, followed by the discrete model. However the computational power required for both of these methods is not justified by only slightly better results. In our regressions, we use the Efron approximation method, proposed in Efron (1977), due to the amount of tied data being large. Even though it might be interesting to run multiple regressions and compare the results of the four different methods, it is not in our interest to go into the detail of each of these methods. For further empirical background and for a comparison between these methods, we refer the reader to the detailed work of Borucka (2014).

As a result of calculating the parametric part of the model, we get values for the hazard ratios, which represent the relative risk over the study time period and must be non-negative. Counterintuitively, the hazard ratios can be greater than 1, as they do not represent absolute probability. In order to calculate absolute probability of default, we need to estimate the survival function of a given business partner. An estimation of the survival function was first introduced in Kaplan and Meier (1958) as a Product-Limit step function. Breslow (1972) used the exact analogue and formulated the integrated cumulative hazard function, which is the minus log of the survival function. In the Cox model, the survival function, predicting the survival of the ith subject at time t, has the following form:

$$\hat{S}_i(t) = exp(-\int_0^t \hat{h}_i(u)\,\mathrm{d}u) \tag{3.4}$$

However, in order to calculate the survival function $\hat{S}_i(t)$, and the absolute probabilities of default, we need to estimate the baseline hazard function $h_0(t)$ from function 3.1. The baseline hazard function is challenging to calculate because it is in the form of a step function, due to the discontinuous characteristics of the event data. Royston and House (2011) discuss approaches to calculate a smooth estimate of the baseline hazard function, however we have found that the computational effort does not justify the use. Another way to estimate the baseline hazard function is to assume a specific distribution for it. In the literature, the Weibull distribution has commonly been used to estimate the baseline hazard function when analysiing death rates. However, as the effect of time on the hazard rate is unknown, we prefer to implement a parametric estimate of the baseline hazard function in the form of a step function, following Chen et al. (2005), as described in Section 4.

In conclusion, we chose the Cox model due to its ability to incorporate time-varying macroeconomic covariates, specific time-dependent variables for each business partner and the period-at-risk of a business partner. Consequently, the Cox model is suitable for estimating PIT PDs in accordance with IFRS 9 by taking into account the current business cycle and the business partner's current creditworthiness. Moreover, our model generates a distinct PD for each business partner and therefore the output can be analysed from multiple dimensions, e.g. average PDs in a rating class, in an industry, in a country etc. Finally, from a statistical point of view, our model is particularly suitable for event data with large amounts of observations and only a small fraction of events occurring, as the model uses every spell as an observation and thereby increases the statistical power.

In the following subsections, we will describe alternative models that we considered for estimating PIT PDs and list their strengths, drawbacks as well as the reasons why we have opted for the Cox model instead of an alternative model.

3.2 Fully parametric hazard models

Fully parametric proportional hazards models are suitable for many types of dynamic models. These models can be expressed in terms of the hazard function's dependence on time and the explanatory variables. One of the specifications Allison (2014) presents in his book on survival analysis is the Gompertz parametric model, which allows the log of the hazard rate $h(t)$ to increase linearly over time:

$$log(h(t)) = b_0 + b_1x_1 + b_2x_2 + ct \tag{3.5}$$

The constants b_0, b_1 and b_2 have to be estimated in a previous step. $h(t)$ is a function of the explanatory variables and time t. c is a constant.

There are many other fully parametric proportional hazards models that differ in the way that time is incorporated. Allison (2014) presents other classes of fully parametric event history models such as Accelerated Failure Time models. However, the general advantages and disadvantages of fully parametric hazard models apply to all of these models. Fully parametric proportional hazards models are well suited for probability of default modelling because predictions on time to event and probability of default can be easily computed as a direct regression output. However, due to two material drawbacks we decided against a fully parametric model. On the one hand, these models were historically not able to incorporate time-varying variables and only recently statistical software packages were released that account for time-varying variables. For a dataset like ours, which contains variables that change over time (e.g. unemployment levels), it is therefore difficult to implement a fully parametric model. On the other hand, any fully parametric model makes assumptions regarding the effect of time on the hazard rate and consequently the probability of default and it is unclear how this effect should look like in reality. Instead, we decided to focus on the semi-parametric proportional hazards model, which allows a higher level of generalisation and flexibility, because it does not assume an explicit effect of time on the hazard rate.

3.3 Static models

Many static PD models are based on probit or logistic regression analysis. Due to their relative simplicity and predictive power, they have emerged as the standard models in the area of credit default modelling. In order to understand these regressions, one can first look at the simple linear regression:

$$y_i^* = \beta\,' \cdot x_i + u_i \tag{3.6}$$

The methodology for all regression models is similar in its essence. A relationship between the dependent variable y and the independent factors x, is established on the values both can assume, denoted as y_i and x_i, respectively. β is the coefficient of the independent factors, indicating the linear relationship established between x_i and y_i whereas u_i denotes the residuals, i.e. error term, in the regression. In the field of credit default modelling, the dependent variable is binary, it can either be $y = 1$, indicating default, or $y = 0$, indicating no default has occurred. Building on the linear regression definition above, the binary default variable y becomes:

$$y_i = \begin{cases} 1 \text{ if } y_i^* > 0 \\ 0 \end{cases}, \tag{3.7}$$

meaning that if y_i^* exceeds the threshold 0, the default event occurs. Deriving the probability of default leads to:

$$P(y_i = 1) = P(u_i > -\beta\ ' \cdot x_i) = 1 - F(-\beta\ ' \cdot x_i) = F(\beta\ ' \cdot x_i), \tag{3.8}$$

where $F()$ is the distribution function, assuming it is symmetric to 0. The distribution function then depends on the distribution assumption of the residuals u_i. If we assume a normal distribution for the residuals, we would have the following probit function where $F(\beta\ ' \cdot x_i)$ becomes:

$$F(\beta\ ' \cdot x_i) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\beta' \cdot x_i} \mathrm{e}^{\frac{-t^2}{2}}\, \mathrm{d}t \tag{3.9}$$

Whereas if we assume a logistic distribution for the residuals, the function $F(\beta\ ' \cdot x_i)$ takes the following form, widely known as the logit or logistic model:

$$F(\beta\ ' \cdot x_i) = \frac{\mathrm{e}^{\beta' \cdot x_i}}{1 + \mathrm{e}^{\beta' \cdot x_i}} \tag{3.10}$$

The difference between the models comes from the assumption regarding the distribution of the residual term, however the outcome is interpreted similarly. The outcomes of $F(\beta\ ' \cdot x_i)$ are probabilistic numbers and can be readily defined as the PDs. Recently, with the effect of the business cycle becoming more apparent, especially following the recent financial crisis, there have been a number of studies where logistic regressions were applied with macroeconomic covariates. In a study by Yurdakul (2014), different macroeconomic factors such as GDP, interest rates, unemployment, etc. are used to identify a relationship between these factors and non-performing loans.

In conclusion, there are a few important reasons why static models have become the industry standard. Due to their nature, the logit and probit models are particularly suitable for analysing binary variables and are statistically accepted. Due to their wide application, they are included in most statistical software packages and are relatively easy to implement. Furthermore, the outcome of the regression gives absolute default probabilities and can be validated with alternative methods. However, we decided against a static model due to some conceptual limitations. Recent discussions in the academic literature suggest that when analysing corporate loan data with very few default events (as opposed to retail credit data), logistic and probit regressions have been found unable to accurately predict default probabilities. For further discussion concerning this subject, please refer to Kennedy et al. (2010) and Wagner (2008). Another distinction between static models and hazard models is the notion of period-at-risk, which is defined as the time spent in a spell. When the sampling period is long, it is important to control for the time a business partner spends in that risk set before he defaults, otherwise a selection bias may result. Period-at-risk is ignored by static models due to their inability to account for time. Static models are also unable to incorporate time-varying coefficients and therefore result in low 'point-in-timeness'. This limitation is further complicated by the IFRS 9 requirement to calculate lifetime PDs for significantly deteriorated assets, because static models have difficulties to create enough data observations with sufficiently long forecasting periods to calculate lifetime PDs.

Finally, as mentioned in Hayden and Porath (2011), the parameter coefficients of a static regression don't provide an intuitive interpretation. Even though the end result is easily applicable as PDs, these PDs are not differentiated among business partners, so that PDs generated for category sets (such as rating) will be the same for all business partners in that set.

3.4 Transformation models

Due to regulatory requirements many financial institutions in Europe report through-the-cycle (TTC) PDs. TTC PDs attempt to balance the need for accurate default estimates and the desire for rating stability. These probabilities represent a long-term trend in ratings. In contrast, IFRS 9 PDs tend to be more point-in-time. Since financial institutions typically have models to calculate TTC PDs, researchers have proposed transformation models that transform TTC PDs into PIT PDs. A very recent transformation model was suggested by Perederiy (2015). Perederiy suggests a transformation model that estimates the business cycle endogenously using TTC PDs and a measure of systematic dependence. On top of this general framework an autoregressive process is added to obtain PIT PDs:

$$PD_{i,Tf}^{PIT,AR(1)} = \Phi\left(\frac{\Phi^{-1}(PD_{i,Tf}^{TTC}) - \Psi_{T_0}\sqrt{\rho\alpha_1^{2(T_f-T_0)}}}{\sqrt{1-\rho\alpha_1^{2(Tf-T0)}}}\right) \tag{3.11}$$

According to Perederiy the PIT probability of default PD_{PIT} of business partner i at time T_f is a function of the TTC probability of default PD_{TTC}, the endogenously calculated macroeconomic factor Ψ at time T_0, the systematic correlation coefficient ρ and the autoregressive component α_1 for the period $T_f - T_0$.

Other TTC-PIT PD transformation models are presented in Aguais (2008) and Carlehed and Petrov (2012). Transformation models are generally attractive because they build on existing probability of default models and the added complexity of a transformation model is moderate and intuitive. However, the quality of the PIT PDs is directly dependent on the quality of the TTC PDs and the assumptions made in the model calibration. For the course of this study, we did not consider a TTC-PIT PD transformation model, because detailed information on TTC PDs was not readily available.

4 Model implementation

Until now we have discussed the advantages of Cox models when analysing event data. We have argued that due to its ability to incorporate information about time and contemporaneous factors it is particularly suitable for an IFRS 9 setting. In this section we will outline how the practical reconciliation and implementation within the context of IFRS 9 can take place.
Any model that attempts to account for impairment according to IFRS 9 needs to capture at least two distinct features. On the one hand, an appropriate model needs to provide PD values for 12 month as well as for the entire lifetime of a financial asset. On the other hand, a mechanism is required that determines whether a significant deterioration in credit risk has taken place. In the following, we outline how our model satisfies both of these requirements.

4.1 Survival function calculation

One of the properties of the Cox model is that statistical inferences can be made without specification of the effect of time on the hazard rate. However, this remarkable characteristic is also one of the biggest drawbacks of a time-varying semi-parametric hazard model because making predictions of absolute probability is challenging. In 2005, Chen et al. presented a SAS macro that extended the traditional

Cox regression. They were able to estimate independent survival functions for patients with hepatocellular carcinoma. These survival functions built on the parameter outputs of a traditional Cox regression. Kim and Partington (2014) applied Chen et al.'s approach to bankruptcy prediction models for the first time.

In the past, time-varying models have had difficulties when put into practise. This is mainly due to a violation of the proportionality property described in Section 3.1.

As described in Function 3.1, the Cox model has the following form:

$$h_i(t \mid z(t)) = h_0(t) * exp\left\{\sum_{j=1}^{p} \beta_j \; z_j^i(t)\right\} \tag{4.1}$$

If proportionality applies, the ratio of the hazards of two business partners i and m remains constant over time because the effect parameters $exp\left\{\sum_{j=1}^{p} \beta_j z_j^i(t)\right\}$ of business partners i and m remain constant over time:

$$\frac{h_i(t \mid z)}{h_m(t \mid z)} = \frac{h_0(t) * exp\left\{\sum_{j=1}^{p} \beta_j \; z_j^i\right\}}{h_0(t) * exp\left\{\sum_{j=1}^{p} \beta_j \; z_j^m\right\}} = exp\left\{\sum_{m=1}^{p} \beta_j (z_j^i - z_j^m)\right\} \tag{4.2}$$

$h_i(t \mid z)$ is the hazard rate of business partner i at time t, where z_j^i is the jth covariate of firm i. The baseline hazard rate $h_0(t)$ is an indicator of the effect of time on the hazard rate and we assume it is the same for all business partners. If the proportionality property applies, the baseline hazard rate can be estimated and cumulative survival functions be plotted. Many statistical software programmes are able to provide cumulative survival functions for proportional data. While the calculation is straightforward, the proportionality property is oftentimes unrealistic. Instead, we would like to build a model that is able to incorporate time-varying variables. Moreover, we are interested in receiving survival functions for individual business partners rather than cumulative survival functions.

For time-varying variables $z_j(t)$, the relationship between business partners i and m is no longer constant. In order to estimate the baseline hazard function $h_{0,i}(t)$ for all t of business partner i we estimate the integrated baseline hazard function $H_0(t)$ first. Based on Andersen (1992) the integrated baseline hazard function is estimated as follows:

$$\hat{H}_0(t) = \sum_{\tilde{T}_i \leq t} \frac{D_i}{\sum_{j \in R(\tilde{T}_i} exp(\hat{\beta}\ '\ z_j(\tilde{T}_i))} \tag{4.3}$$

D_i is a dummy event variable that can assume the values '0' and '1' and indicates whether the business partner i defaulted at any time T before or at t. b is the vector of parameter coefficients obtained in the Cox regression, where $z_j(T_i)$ is the jth covariate of firm i at default time T. With our dataset, the calculation of the integrated baseline hazard rate is a challenge, because default information is updated on a monthly basis. Thus, it is not uncommon that more than one business partner defaults in the same time interval. To resolve this problem of tied data we use a method suggested by Hosmer et al. (1999) and Borucka (2014). Ties can be broken by removing a small random value from each survival time. By doing so, each business partner that defaults has a unique time of default but the position in relation to the remaining business partners remains approximately unchanged.

It is also possible to describe the integrated baseline hazard rate as a step function, which is discontinuous at time t_m:

$$H_0(t) = \sum_{t_i \in t} \left[h_0(t_{m-1}) * (t_m - t_{m-1}) \right] \tag{4.4}$$

From equations 4.3 and 4.4 the baseline hazard h_0 for time t is derived.

Thus, the hazard rate $\hat{h}_i(t)$ of firm i at time t is:

$$\hat{h}_i(t) = \hat{h}_0(t) * exp(\hat{\beta}\,'\, z_i(t)) \tag{4.5}$$

Using the baseline hazard rates for any time t, we obtain the survival function of business partner i as the integral over the hazard functions of all time intervals between 0 and t:

$$\hat{S}_i(t) = exp(-\int_0^t \hat{h}_i(u)\, \mathrm{d}u) \tag{4.6}$$

The 12-month probability of default for business partner i at calendar time t^* is then calculated as the difference between the survival probability at event time t corresponding to calendar time t^* and event time $t + 12$.

To illustrate this concept of event-time-based PD calculation, we present the survival probabilities of a stylised business partner who has the following survival time information:

In order to calculate the 12-month probability of default at calendar time 4 we need to estimate the survival function of the business partner until event time 14, because at calendar time 4 the business partner is at event time 2.

From the earlier analysis it is obvious that any forward-looking probability of default can be calculated. Therefore, the shift from a 12-month PD to a lifetime PD is

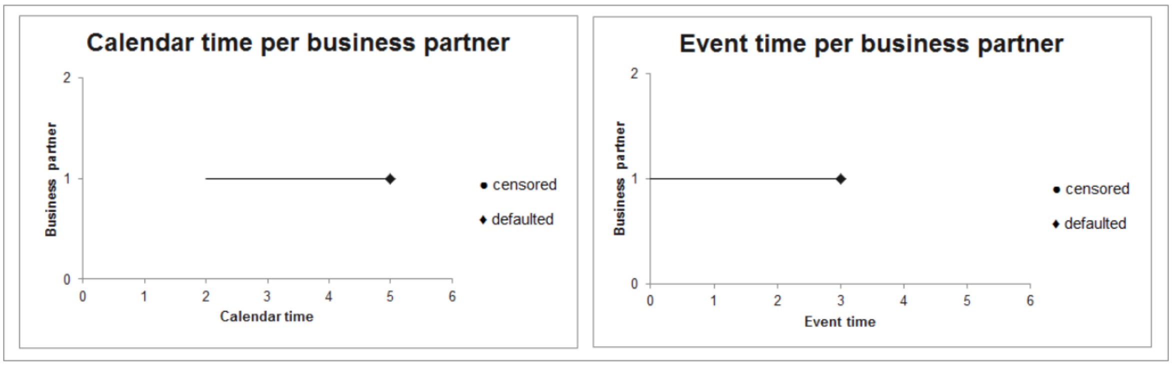

Figure 4.1: Individual calendar and event time

straightforward. However, it is not obvious how to define lifetime PD. In many cases the financial instrument of a business partner has a fixed maturity. Nonetheless, financial instruments such as credit cards do not have a fixed maturity. In these cases, assumptions on the average lifetime of a financial instrument need to be made. In either case, it is unlikely that any kind of available data is sufficient to calculate the probability of default of a business partner with a high expected lifetime. Our sample period includes five years of default data and it is not possible to make any direct probability of default predictions beyond that. Therefore, we suggest to pick an upper limit for the PD and to extrapolate for any PDs in excess of this limit. The amount of business partners with a survival time in excess of three years is high enough to justify calculating the lifetime PD with a 36-month PD and linear extrapolation. Since we decided to validate our model on the 12-month PDs only, we did not do further analysis on an adequate upper limit of the lifetime PD in our specific dataset.

4.2 Significant credit risk deterioration

A fully functioning PD model that is in line with IFRS 9 needs to allow for differentiated treatment of financial assets with a significant deterioration of credit risk. In such cases, IFRS 9 mandates usage of the lifetime probability of default for ECL calculation. ECL for financial assets which have not experienced a significant deterioration of credit risk are calculated using the 12-month PD. A financial asset's credit risk is considered low if the probability of default is low. The standard indicates that an 'Investment grade' rating may be considered a low credit risk signal and that a significant increase in credit risk is to be understood as an increase in the default probability since initial recognition (IASB, 2015). Thus, the standard setters were careful to specify a significant deterioration in credit risk as a relative measurement criterion.

In light of the details presented above, we feel it is appropriate to use a relative threshold that identifies downgrades from an 'Investment grade' rating to a 'Speculative grade'. Since most of the business partners in the available dataset are private companies, little information on external ratings is available. Instead, an internal rating scale with 18 rating levels is used to assess the firm-specific default probability of an individual business partner. Brunel (2015) showed that an adequate threshold balances the 'Hit Rate' and the 'False Alarm Rate'. The former is the proportion of defaulted financial assets that have received a marker for significant deterioration of credit risk, whereas the latter indicates the proportion of financial assets that were incorrectly recognised with a significant deterioration of credit risk. The transfer criterion obtained in the aforementioned paper quantifies the target Hit Rate to be between 70% and 80% for realistic scoring models. While more research should be undertaken to find an ideal threshold for significant deterioration, Brunel's quantitative criterion is a good approximation for a conceptual PD model that incorporates IFRS 9 impairment.

In order to validate the general attractiveness of our PD model we focus on 12-month PDs due to the higher amount of available data. Therefore, we do not implement a threshold for significant deterioration of credit risk in our analysis.

5 Default data

Our analysis is based on credit default data obtained from the Stockholm-based bank Nordea. The data contains information on [redacted] corporate business partners with domicile in Sweden. In order to capture the state of the macroeconomic environment in which these business partners operate, we include Swedish macroeconomic indicators in our analysis. Corporate business partners with activities in the shipping industry were excluded since the shipping industry is mostly affected by global industry-specific business cycles.

Due to the seniority of Nordea's operations, the historical data in the database is extensive. The estimation period is from January 2008 to December 2012. However, in order to avoid the bias introduced by left-truncated data, we included spells that were under way in January of 2008. As we had historical data available for each business partner that was in the dataset before our sample period, we know when a business partner entered the current spell and included this information in the model estimation. For additional information on left-truncated data, please refer to Section 5.3 Data Adjustments. The period from January 2013 to December 2013 is the out-of-sample validation period for our model.

Nordea uses a continuous internal rating grade model that assesses the relative risk of a business partner. A corporate business partner receives one of 18 risk grades.

These risk grades are translated into values from 3 to 20, 3 indicating the lowest, 20 the highest risk of default. By doing so, Nordea is able to maintain a consistent risk modelling approach across corporate, retail, institutional and sovereign customers because in some of these business verticals 20 risk grades are used.

Default is considered to have occurred with regard to a particular customer when any of the following three events have taken place: The business partner is performing, but based on objective evidence that has occurred it is considered unlikely that the business partner will be able to repay its debt obligation in full, and the situation cannot be satisfactorily remedied. The business partner is non-performing, i.e. no longer able or willing to fulfill its payment obligations, e.g. because material payments are past due more than 90 days on any debt obligation. The business partner has filed for bankruptcy or has been declared bankrupt.

Our model is based on event time analysis and, therefore, we had to incorporate event time, in addition to calendar time. One business partner may have several independent spells due to the occurrence of default or non-default events or because the business partner leaves the dataset and returns. A common reason why a business partner leaves the dataset for a given time period is because the exposure of the business partner changes from positive to negative, e.g. by overpaying a loan and thus turning credit into deposit. In our analysis, we disregard the amount of time a business partner stays in default because the event has already occurred.

The [redacted] business partners in our database translate into [redacted] spells. Out of these [redacted] spells, 1.03% ended with a default event.

5.1 Model covariates

In this section, we present the different factors we considered in order to identify relationships with the probability of default. As the Cox model is able to take time-varying-covariates into account, we included variables that change over time as well as constant variables. We can analyse the variables we have used under two dimensions, one being the variable's dependence on time and the other being their specific characteristics, namely firm-specific and macroeconomic. We have one time-constant variable, the 'Current rating' covariate, while all other variables can change over time for a specific business partner. Considering the firm-specific and macroeconomic covariates, our prior expectation is that the firm-specific covariates have higher explanatory power than the macroeconomic covariates, since the firm-specific factors are able to differentiate for idiosyncratic risk on a business partner level. Figlewski et al. (2012) argue that stability in firm-specific covariate regression parameters, irrespective of the number of added macroeconomic covariates, shows high explanatory power. As Figlewski et al. were able to show high stability in the firm-specific parameters, we expect a similar finding.

5.1.1 Firm-specific covariates

Current rating: The current rating indicates the expectations on the creditworthiness of a business partner. It reflects information concerning the business partner's financial stability and ability to repay. The current rating does not take into account the general macroeconomic situation or the current situation of the industry. Therefore, ratings are assumed to be independent from the other covariates and to not have a confounding effect on the regression.

The best rating is defined as 3, the worst rating as 20. We use a continuous variable that can take on any value between 3 and 20 and expect an increase in the 'Current rating' covariate to increase the probability of default. In Table 8.1 in the Appendix we show the results of a preliminary regression analysis with dummy variables for ratings 3 to 20. As is apparent from the table there is an unproportionally large jump in parameter coefficients from rating 6 to 7 and from rating 14 to 15. This finding is in line with an unproportionally large jump in actual default frequencies from rating 6 to 7 and from rating 14 to 15. Based on this information it is difficult to uphold a linearity argument for the continuous variable 'Current rating'. At the same time, practitioners prefer a continuous rating variable to dummy variables for each rating because of the simplicity of the former.

Moreover, in the specific context of our PD model, it is important to maintain a low number of covariates to keep the total computation time in check. Thus, instead of using dummy variables for each rating we captured the rating group differences in additional grouped covariates for our 'Recent downgrade' and 'Recent upgrade' covariates. Ratings 3 to 6 have very strong negative coefficients, which is in line with our expectations because the amount of defaults within these ratings is very low. Ratings 7 to 14 have negative coefficients, which are less strong than the coefficients from ratings 3 to 6. Even though rating 14 is positive, we allocate it to this bucket, because the absolute change from rating 13 to 14 is in line with the other rating changes in this bucket. Ratings 15 to 20 have strong positive coefficients, indicating that the probability of default increases for worse ratings. Based on these findings, we created three rating buckets with these rating groups, where the best ratings 3 to 6 are classified into bucket one, ratings 7 to 14 into bucket two and the worst ratings 15 to 20 into bucket three. These rating buckets were used for the 'Recent downgrade' and the 'Recent upgrade' covariates as outlined below.

Recent downgrade/upgrade: Using empirical evidence, the academic literature has identified a 'ratings drift', indicating a positive serial correlation in rating changes (see e.g. Altman and Kao, 1992; Hamilton and Cantor, 2004). Thus, a business partner that received a downgrade in a recent past period, is more likely to default than a business partner whose rating has not recently changed. Initially we used dummy variables to indicate a downgrade or upgrade within a recent period. However, we recognised that refining the downgrade and upgrade covariates could further improve regression results. Consequently, we defined a downgrade into the third rating bucket (ratings 15 to 20), which we expect to have a significant positive coefficient, and a downgrade into the second bucket (ratings 7 to 14), which we expect to have a positive coefficient. Vice versa, we defined an upgrade into the second bucket (ratings 7 to 14), which we expect to have a significant negative coefficient, whereas we expect an upgrade into the first bucket (ratings 3 to 6) not to be significant. For all potential downgrade and upgrade variables, we considered the recent 6 months, 1 year and 2 years. While both the covariates for 6 months and 2 years were significant, we decided to use the 6-month variable due to higher 'point-in-timeness'.

Months since first rated: It has been suggested in the literature that business partners that have been newly rated are less likely to change the rating than more seasoned firms (see e.g. Altman (1998)). To test and incorporate this finding we consider the number of months since the business partner has been rated initially. While more seasoned firms may be more likely to change the rating, we argue that one reason a company has survived longer than the majority of business partners could be its superior creditworthiness. Therefore, it is unclear whether the longevity of a business partner increases or decreases the probability of default.

If a business partner enters the dataset before the starting point of the sample dataset, we look back in the available historical data and take the data period in which the business partner appeared for the first time. If a business partner defaults at some point and receives a non-default rating subsequently, the covariate is reset in the period in which the business partner returns.

For a business partner that entered the dataset before January 2005, we start the covariate in January 2005.

Due to the different starting points of the 'Months since first rated' covariate, the variable reached up to 96 in some cases, which results in a particularly high weight in the regression. In line with Figlewski et al. (2012), we adjust for this bias by normalizing the 'Months since first rated' covariate by using the log of the variable, not allowing observations with very high values to have a significant effect on the regression.

For more information regarding the floor and the adjustments that were made to the raw data, refer to Section 5.3 Data Adjustments. We do not have a prior expectation on the months since first rated covariate because it is unclear which of the two effects outlined above is dominant.

5.1.2 Macroeconomic covariates

We have classified the macroeconomic covariates under two categories. Covariates under 'General direction of the economy' are taken into account as indicators of the general state of the economy in which the business partners operate. These are used to capture the business cycle in Sweden over the analysed period. Financial market covariates are providing information about the level of interest rates and are presumed to behave in line with the interest rates on the financial assets in the dataset, thus having an effect on the default probability of the business partners. We faced two main challenges when deciding on how to include these factors. First of all, we were left with the choice of including levels, first differences, log of first differences or, in the case of interest rates, the slope or curvature of the yield curve itself. When choosing the right form of these variables, we oriented ourselves towards the literature and mainly followed Figlewski et al. (2012) where possible. However, we are aware that the time-series of the covariates may differ, e.g. US interest rates versus Swedish Interest rates, and therefore show different characteristics. As it is not in the scope of this study, we did not test the time-series of these covariates for stationarity or unit-root. Neither did we perform robustness tests in order to decide whether using a log change would result in better indicative power. Further analysis on the time-series of the covariates might lead to a better selection of their form and improve model fit. Secondly, we decided against using the covariates as current variables, due to the intuitive assumption that indicators of the economy do not have an immediate effect on the defaults. This is further discussed in Section 5.2 Macroeconomic Lags.

5.1.2.1 General direction of the economy

Real GDP growth: Real GDP growth is commonly regarded as a measure of relative economic strength and we expect fewer defaults during times of economic expansion. However, as Figlewski et al. (2012) point out, a high real GDP growth might also be the consequence of a prior recession period. The real GDP growth figure is based on annual GDP growth data on a quarterly basis. For calculating the months between quarters, linear extrapolation was used.

Unemployment: Historically regarded, unemployment is a direct indicator of the general health of the economy. In times of high economic activity, unemployment is assumed to be low. Therefore, in times when unemployment is low, we expect the probability of default of business partners to be low as well. We argue that the change in unemployment is a relevant measure of the direction of general economic health and incorporate the change in unemployment.

Inflation: A high level of inflation affects incomes, profitability and savings negatively. Therefore, many central banks have taken active measures to control the inflation level. Subsequently, it may be argued that an increase in inflation increases the probability of default. At the same time, as outlined by Figlewski et al. (2012), the outstanding debt of a company is denominated in the local currency and an increase in inflation reduces the debt burden. Considering both of these effects, we do not have a strong opinion on the dominance of either one of these effects and have no prior expectations on the parameter sign.

5.1.2.2 Financial markets

Nasdaq OMXS30 index: Public stock markets are assumed to be direct indicators for the general health of the corporate sector. Thus, we expect that rising stock prices result in lower probabilities of default. Since our database consists of corporate business partners with domicile in Sweden, we decided to incorporate the change in the Nasdaq OMXS30 index.

Swedish Treasury Bills and Government Bonds: It is presumed that high interest rates should indicate an increased difficulty in raising the funds to service debt payments. However, in the light of the Financial crisis of 2008 and the subsequent European Sovereign Debt crisis it is unclear to which extent the aforementioned is true. Our dataset is largely affected by irregular and historically rare events such as the central bank response to the financial crisis. The crisis response resulted in historically low interest rates for an extended time period. As a result, the correlation between interest rates and the general health of the corporate sector may have been forced away from historical correlation levels and the term structure of the yield curve might be different in our sample period. We therefore initially identified interest rates on 3-month Swedish Treasury Bills as well as 2-year, 5-year and 10-year Swedish Government Bonds with no prior expectations on its coefficients. However, the correlations between rates on 10-year Treasury Bills and 2-year as well as 5-year Treasury Bills are 87% and 98% as shown in Figure 5.1 below. The high correlation between the interest rates would lead to a distortion in the regression, which is why we removed the 2-year and 5-year Treasury Bill rates from the analysis. By keeping the 3-month and 10-year T-Bill rates we keep a short-term and long-term interest rate to capture both the slope and level effects, respectively.

Correlation matrix of macroeconomic covariates							
	GDP	UNE	EQTY	S3M	S2Y	S5Y	S10Y
GDP	1	-0,04	0,24	0,08	0,24	0,2	0,18
UNE		1	0,22	-0,91	-0,82	-0,67	-0,58
EQTY			1	-0,33	-0,4	-0,52	-0,61
S3M				1	0,92	0,79	0,69
S2Y					1	0,94	0,87
S5Y						1	0,98
S10Y							1

Table 5.1: The correlation matrix shows correlation values for the analysed macroeconomic covariates. The real GDP growth rate shows moderate correlation values with all covariates. The equity index 'EQTY' and the Swedish T-Bill rates ('S3M', 'S2Y', 'S5Y', 'S10Y') show inverse correlations. The different interest rates are highly correlated with each other.

5.2 Macroeconomic lags

We add macroeconomic factors to our model in order to approximate the overall financial health of the corporate sector in Sweden. As Figlewski et al. (2012) remark, it is unlikely that changes in the macroeconomic factors immediately affect business partners. Many factors of corporate success such as order books and existing projects have limited flexibility. Moreover, in order for data to be used for prediction purposes, it is necessary that the data is readily available. For most macroeconomic variables, the data for a period is only available after a few months following that period. Taking these considerations into account, we included a 6-month, 12-month and 18-month lag to compare the significance of these factors in a preliminary regression analysis. The results for the 6- and 18-month lags demonstrated similar significance, with 6-month lags being slightly more significant and having expected coefficient signs. We conducted an analysis of fit for the historical time series of macroeconomic factors with the historical actual default events and found that the 6-month lag has a good fit with at least some of the historical covariates and a better fit in comparison to the 12- or 18-month lags. In addition, using macroeconomic variables with a 6-month lag increases the 'point-in-timeness' of the calculated PDs in comparison to using an 18-month lag. Therefore we included the macroeconomic variables with a 6-month lag.

5.3 Data adjustments

Business partners that have not been rated for 12 months are unrated. Among other things, this can happen if the exposure of a business partner is low. Since we cannot infer any rating information from an unrated business partner, we exclude all unrated business partners. Similarly, in order to conduct a meaningful analysis, it is important to know which rating the business partner had immediately before the default event. For a business partner that enters the dataset at default it is unclear which rating, if any, the business partner had prior to default. Due to this reason, we removed all default events at initial recognition. Therefore, business partners that enter the dataset at default are not recognised until they receive the first non-default rating after the default event. However, we differentiate business partners that already had a credit rating, left the dataset without default and returned to the dataset afterwards. Because these business partners had at least one historical rating, we use the last rating before the business partner left the dataset as the rating at default.

It is not uncommon that a business partner that experienced default remains in default for several periods. In survival analysis, time to event is decisive, not the length of the event. Since we cannot determine whether a business partner received a new rating and defaulted within the same period, we remove all default periods immediately following the first default event of a business partner.

Some of the business partners present in the beginning of our sample period have their spells already under way. These spells are considered left-truncated. For the spells already under way at the beginning of the sample period, we used historical information on these specific spells. Nonetheless, to keep the survival time of business partners within a reasonable range, the minimum possible calendar time a spell can have is January 2005. Accordingly, we removed all spells that had a spell under way prior to the floor level of January 2005. This was done in order to maintain a reasonable survival time and 'Months since first rated' covariate, as otherwise the model would have been biased towards the business partners that have survived since 2002 until now, i.e. generating a probability of default figure that is extraordinarily low. Without setting a floor level, business partners that have not undergone rating adjustments, e.g. due to low exposure or other internal issues, would have unproportionally long survival times and consequently a low probability of default that is not indicative of the real probability of default. Setting a floor at 2005 for spells under way reduces the bias in comparison to including all data since 2002 and still provides enough spells to be included in our analysis. However, it is important to note that there might still be a survivorship bias due to the small sample size of the data.

Lastly, the dataset features a general rating bias towards rating 15. This rating bias results in higher defaults than expected. In the risk grading process, there is a psychological bias towards rating 15 which is perceived as the best of the worst, and companies that are evidently part of the worst rating group tend to receive a rating which is in this category, even though they might be more suitable for worse ratings. Vice versa, business partners in rating 17 have an unusually small average default rate.

6 Analysis and results

In Section 6.1, we show the results of our regression analyses and discuss their implications. Subsequently, we describe how the coefficients we have obtained are linked to PDs. In Section 6.2, we present our final calculations of the 12-month PDs based on the available credit data and validate our results.

6.1 Cox regression and model selection

At first, we adapted our dataset to the requirements of an event time analysis. Allison (2014) presents the episode splitting method as an appropriate way to set up data with time-varying covariates. The episode splitting method sorts data observations according to spells and puts all observations per spell in the 'long' form. In the long form a separate data record for each interval of time is created in which all explanatory variables stay the same. The starting and ending times of these intervals are measured from the beginning of the spell. Each interval is considered censored if the event did not occur and uncensored if the interval ended with a default. The advantage of the episode splitting method is that it allows equal treatment for time-constant and time-varying covariates, because during an interval all covariates are constant.

In our regressions, we used backward selection in order to identify the most explanatory covariates for default. Backward selection is an iterative regression method that allows to run a regression analysis with all covariates of interest. In each run, the covariate producing the least significant F statistic is dropped and the process continues until all covariates remaining in the regression have F statistics significant at a stay significance level. We have specified a stay significance level of 0.05 in our analysis. The F statistic has the intuitive definition of explained variance over unexplained variance. [1]

We ran multiple regressions to analyse various effects of the covariates. In Section 6.1.1, we present the regression we ran with firm-specific covariates only. In Section 6.1.2, we outline the individual effects of the macroeconomic covariates. In Section 6.1.3, we include all the firm-specific and macroeconomic covariates and present the final regression results that were used for the estimation of PDs.

6.1.1 Firm-specific Cox regression

We ran initial regressions with several subsets of the covariates presented in Section 5 to identify the significance of our covariates on default events. As a starting point, we ran a regression on the firm-specific factors from Section 5.1.1 to understand the explanatory power of our non-macroeconomic covariates on default. The results of the regression are presented in Table 6.1.

---

[1] As described in the SAS/STAT user's guide (1999), if the current model has p parameters excluding the intercept, and if you denote its residual sum of squares by RSS_p, drop an effect with k degrees of freedom and denote the residual sum of squares of the resulting model by RSS_p , the F statistic for removal with k numerator degrees of freedom and $n-p-k$ denominator degrees of freedom is given by:

$$F = \frac{(RSS_{(p-k)} - RSS_p)/k}{RSS_p/(n-p-k)} \tag{6.1}$$

Cox Regression with firm-specific factors (Sample period: January 2008 - December 2012)						
		Total	Events		Total	Events
Number of observations:		●	●		●	●
	coefficient	p value	Hazard ratio	coefficient	p value	Hazard ratio
Firm-specific covariates:						
Current rating	0.5737	<.0001	1.775	0.5738	<.0001	1.775
Downgrade into rat. group 7-14 (last 6 months)	-10.9989	0.9886	0.000			
Downgrade into rat. group 15-20 (last 6 months)	1.4060	<.0001	4.080	1.4064	<.0001	4.081
Upgrade into rat. group 3-6 (last 6 months)	-10.3569	0.9908	0.000			
Upgrade into rat. group 7-14 (last 6 months)	0.0779	0.8772	1.081			
Months since first rated (log)	-0.1826	0.0010	0.833	-0.1825	0.0010	0.833
Model goodness of fit:						
-2 log L			13,137.59			13,137.89
AIC			13,175.90			13,143.89

Table 6.1: Cox regression output over all firm-specific covariates. The left regression presents the results for a fitted model over all firm-specific covariates. The right regression presents the final covariate selection after backward selection. Significant covariates at the 5% confidence level have a p-value below 0.05. For model goodness of fit the -2 log likelihood and the Akaike Information Criterion is presented.

In line with our expectations, the 'Current rating' covariate is highly significant and positive with a hazard ratio of 1.775. The hazard ratio can be interpreted as the chance of an event occurrence for a business partner with the covariate divided by the chance of the event occurrence for a business partner without the covariate. Thus, everything else equal, a one-unit increase in the credit rating results in a 78% higher hazard rate. Similarly, the hazard rate decreases by 17% for a one-unit increase in the log value of a business partner's 'Months since first rated'. This indicates that a 'Months since first rated' value that is 'e' times higher, has a 17% lower hazard rate. We had no prior expectation as to the sign of the coefficient as there are mixed views on this in the literature. Moreover, a business partner that received a downgrade from a rating better than 15 into the lowest rating group has a 308% higher hazard rate than a business partner who has always had the rating from the lowest rating group.[2] This result is in line with the empirical finding of a positive serial correlation in rating changes (see e.g. Altman and Kao, 1992; Hamilton and Cantor, 2004), which suggests that a rating downgrade is often followed by a further rating downgrade.

On the other hand, an upgrade from the lowest bucket into the second bucket is non-significant. Our prior expectation was that an upgrade should reduce the probability of default. This non-significance might have been due to the fact that 86% of all the business partners that upgraded into the second bucket were classified in the ratings 12-14. As these ratings in general have a higher probability of default than the bucket average, this might confound the rating drift effect of an upgrade. In a similar fashion, the downgrade from bucket one into bucket two might not be significant because the majority of these are classified in ratings 7 to 11. In addition to this effect, the number of default observations in the first and second rating buckets might be too low to show significant risk increases. This is particularly true for upgrades from the second bucket into the first bucket, where no default observation was observed for multiple periods. In order to improve the analysis of the upgrade and downgrade covariates, a larger dataset could be used to allow analysis of more refined buckets.

We calculated goodness of fit measures to choose the model with the optimal parameter selection. The measures are based on the likelihood function that can be defined for a set of data as the probability of obtaining that particular set of data, given the underlying probability distribution model. For a mathematical expression of the likelihood function, suppose a sample $x_1, x_2, \ldots, x_n$ of n independent and

[2] For the classification into rating groups, please refer to Section 5.1.1

identically distributed observations with an unknown probability density function $f_0(\cdot)$. Assuming that the function f_0 belongs to a certain model $f(\cdot|\theta), \theta \in \Theta$ and that θ is a vector of parameters of this model so that $f_0 = f(\cdot|\theta)$, where θ_0 is the true value of the parameter vector. The maximum likelihood estimation is used to make an estimation of θ that is as close to θ_0 as possible. For the sample described above, the joint density function becomes:

$$f(x_1, x_2, ..., x_n|\theta) = f(x_1|\theta) * f(x_2|\theta) * \cdots * f(x_n|\theta) \tag{6.2}$$

The probability density functions of each x_i, given θ, are multiplied with each other. If θ is not assumed as given and we control for the x_is, θ is the function's variable and the function results in the likelihood function of the parameter vector θ:

$$\mathcal{L}(\theta; x_1, \ldots, x_n) = f(x_1, x_2, \ldots, x_n|) = \prod_{i=1}^{n} f(x_i|\theta) \tag{6.3}$$

For convenience, generally the log of the likelihood function is used, because it is easier to calculate the maximum likelihood estimator given the derivative of the log likelihood. The log-likelihood measure is widely used in statistics as a goodness of fit measure in order to choose the appropriate model and parameters. In our analysis we have used two measures, the '$-2*$ log-likelihood' measure and the 'Akaike Information Criteria' (AIC). Lower values for these measures suggest a better fit of the model.

However, given a set of parameters, when more explanatory parameters are added the log-likelihood measure will necessarily increase and thus result in a lower log-likelihood measure. This results in a bias called overfitting, where additional variables with little explanatory power are added to the model. In order to prevent this, penalised-likelihood information criteria can be used. The AIC is defined as a penalised-likelihood information criterion that implements a penalty for each additional parameter in the regression. It is expressed as follows:

$$AIC = -2 * \text{ log-likelihood } + 2 * \text{ number of estimated parameters} \tag{6.4}$$

In conclusion, it is suggested that the model with the lowest AIC should be selected in order to optimise the trade off between increased explanatory power and overfitting. As we can see in Table 6.1, the parameters obtained by backward elimination decrease the AIC value despite the slight increase in the '$-2*$log-likelihood' figure.

6.1.2 Individual effects of macroeconomic Cox regression

In order to assess the absolute importance of our systematic covariates, we ran a Cox regression for each of our macroeconomic covariates as outlined in Section 5.1.2. We added the respective macroeconomic covariate to a regression with all the firm-specific covariates that survived backward selection in Section 6.1.1. By doing so we can assess the explanatory power of each of our macroeconomic covariates without a confounding effect of other macroeconomic covariates. The results of this analysis are presented in Table 6.2.

Individual Macro covariates (Sample period: January 2008 - December 2012)			
		Total	Events
Number of observations:		●	●
Firm-specific covariates:			
Current rating			
Downgrade into rat. group 15-20 (last 6 months)			
Months since first rated (log)			
	coefficient	p value	Hazard ratio
General macroeconomic covariates:			
Real GDP growth (6-month lag)	-4.8924	<.0001	0.008
Chg. in unemployment (6-month lag)	2.8779	0.0002	17.776
Chg. in CPI (6-month lag)	0.1401	0.0278	1.150
Financial market covariates:			
Chg. in Nasdaq OMXS30 index (6-month lag)	-0.8370	0.2564	0.433
Chg. in Swedish 3-month T-Bill rate (6-month lag)	0.0416	0.8385	1.043
Chg. In Swedish 10-year T-Bill rate (6-month lag)	-1.0908	0.0651	0.336

Table 6.2: For each macroeconomic covariate an independent Cox regression was run. In each run, all firm-specific covariates and the respective macroeconomic covariate were included in the regression. Significant covariates at the 5% confidence level have a p-value below 0.05. Model goodness of fit values are not presented because they are different for each regression.

While the results for the covariates that signal the direction of the economy are significant, all of the financial market covariates fail to reach significance. An increase in the GDP or the equity index reduces the hazard rate, which is expected because an increase in these factors is generally a sign of economic strength. In contrast, an increase in the unemployment rate or the Consumer Price Index increases the hazard rate. These findings are in line with our expectations, since an increase in these covariates typically signals a sign of economic weakness. The changes in interest rates are less clear to interpret. An increase in interest rates typically implies tight credit markets, which in turn should indicate an increased difficulty in raising the funds to service debt payments and therefore raise the probability of default. The change in the 3-month interest rate seems to confirm this effect, since an increase in the 3-month rate increases the hazard rate. However, the effect of the 10-year interest rate is negative. Because the 3-month and the 10-year interest rates have a correlation of 69%, as was shown in Figure 5.1, it is unclear why we should expect opposite effects.

We decided to run a preliminary Cox regression that includes both the 3-month and the 10-year interest rates (see Figure 8.2 in the Appendix). Because both interest rates maintained their parameter sign we decided to remove the 10-year Treasury Bill rate from our final analysis due to its lower significance.

6.1.3 Cox regression with all covariates

Finally, we ran a Cox regression model on all our covariates except the 10-year Swedish T-Bill rate. This was done in order to assess the marginal effect of the macroeconomic covariates on the firm-specific covariates and to find the final covariate selection for our PD calculation. The results are presented in Table 6.3. We estimated parameters using all covariates and backward selection.

Cox Regression (Sample period: January 2008 - December 2012)						
		Total	Events		Total	Events
Number of observations:		●	●		●	●
	All covariates			**Backward selection**		
	coefficient	p value	Hazard ratio	coefficient	p value	Hazard ratio
Firm-specific covariates:						
Current rating	0.5743	<.0001	1.776	0.5742	<.0001	1.776
Downgrade into rat. group 7-14 (last 6 months)	-11.6672	0.9913	0.000			
Downgrade into rat. group 15-20 (last 6 months)	1.3072	<.0001	3.696	1.3113	<.0001	3.711
Upgrade into rat. group 3-6 (last 6 months)	-11.0174	0.9930	0.000			
Upgrade into rat. group 7-14 (last 6 months)	0.0755	0.8809	1.078			
Months since first rated (log)	-0.1861	0.0009	0.830	-0.1903	0.0007	0.827
General macroeconomic covariates:						
Real GDP growth (6-month lag)	-6.9584	<.0001	0.001	-7.0958	<.0001	0.001
Chg. in unemployment (6-month lag)	0.6670	0.4509	1.948			
Chg. in CPI (6-month lag)	-0.0161	0.7888	0.984			
Financial market covariates:						
Chg. in Nasdaq OMXS30 index (6-month lag)	-1.2132	0.0873	0.297			
Chg. in Swedish 3-month T-Bill rate (6-month lag)	0.8181	0.0041	2.266	0.8012	0.0001	2.228
Model goodness of fit:						
-2 log L			13,096.07			13,099.83
AIC			13,118.07			13,109.83

Table 6.3: Cox regression output over all covariates without the 10-year Swedish T-Bill rate. The regression on the left presents the results for a fitted model over all covariates. The regression on the right presents the final covariate selection after backward selection. Significant covariates at the 5% confidence level have a p-value below 0.05. For model goodness of fit the -2 log likelihood and the Akaike Information Criterion is presented.

The firm-specific covariate coefficients changed only slightly in comparison to the regression with only firm-specific covariates. This is in line with the findings of Figlewski et al. (2012) who suggest that the stability in firm-specific covariates indicates that these covariates outweigh the macroeconomic covariates in importance. Nevertheless, the inclusion of macroeconomic covariates has improved the AIC significantly. Furthermore, the best AIC value is achieved through the backward elimination with macroeconomic covariates.

In total, five of our covariates showed statistical significance and all of these covariates showed coefficients in line with our prior expectations. All three firm-specific covariates, which were significant in the firm-specific analysis with backward selection, are still significant. Additionally, the lagged 6 months real GDP growth shows significance with the same parameter sign as before. Interestingly, the change in the 3-month interest rate is highly significant, even though it was insignificant in the individual analysis in Table 6.2. However, since the 3-month interest rate shows a sign that is in line with our expectations, we decided to keep the 3-month interest rate in our further analysis.
All other macroeconomic covariates were eliminated during the backward selection process. Nonetheless, as our earlier analysis of the explanatory power of our macroeconomic covariates showed, especially the 6-month lag values for unemployment and inflation have relevant explanatory power on an individual level.

Our results show that both firm-specific and macroeconomic factors have a significant influence on PDs. The five significant covariates we identified are used in our SAS macro for PD calculation.

6.2 Analysis of PD values

As we defined in Section 5, our in-sample estimation period is from the beginning of January 2008 to the end of December 2012, whereas our validation period is from the beginning of January 2013 to the end of December 2013. We calculated 12-month probabilities of default (PDs) for the [redacted] business partners that were in the dataset at the end of December of 2012 and did not default. For the same business partners, we calculated the 12-month actual default frequencies (ADFs). Subsequently, we were able to compare our predicted PDs with the respective ADFs for the same period. The ADF is a dummy variable that takes the value 0 or 1. If a business partner defaults at any time within one of the following 12 months, the dummy variable for actual default is set to 1, otherwise it is given a value of 0. In Section 6.2.1, we present our findings and compare the PDs in different rating classes with the respective ADFs. We go on to discuss why we cannot make a direct comparison between these two in the context of rating classes and present a more sophisticated validation method in Section 6.2.2.

6.2.1 PDs and ADFs within rating classes

Based on the calculated 12-month PDs and the ADFs we assess the predictive power of the proposed Cox model within rating classes. It is important to note that our model incorporates a series of covariates and the rating covariate is only one of several factors that affect the probability of default. Therefore, within each rating class the probabilities of default of the business partners vary. Nonetheless, in Table 6.4 we compared the average 12-month PDs and ADFs according to rating class.

Comparison of 12-Month Actual Defaults with 12-Month PDs				
Rating Group	**Total No. Of Observations**	**Defaults**	**ADF**	**PDs**
3	●	●	●%	0.002%
4	●	●	●%	0.004%
5	●	●	●%	0.007%
6	●	●	●%	0.012%
7	●	●	●%	0.024%
8	●	●	●%	0.042%
9	●	●	●%	0.079%
10	●	●	●%	0.147%
11	●	●	●%	0.261%
12	●	●	●%	0.456%
13	●	●	●%	0.793%
14	●	●	●%	1.358%
15	●	●	●%	5.109%
16	●	●	●%	8.028%
17	●	●	●%	14.804%
18	●	●	●%	22.267%
19	●	●	●%	24.311%
20	●	●	●%	31.908%
Total	●	●	●	●%

Table 6.4: Summary statistics are presented for the time interval from January 2013 to December 2013. If a business partner defaulted within this time interval, the 'Defaults' variable is increased by 1. The Actual Default Frequency (ADF) per rating class is calculated as the number of defaults over the total number of observations. The 'PDs' column presents the average Cox model output per rating.

We see that the average 12-month PD estimation increases as the rating class increases. This comes naturally, as 'Current rating' is the firm-specific covariate that changes for each rating class and has a positive coefficient, increasing the PDs, whereas the other firm-specific factors may have similar effects for each rating class. For example 'Log of months since first rated' is expected to have a similar effect for different rating classes as we do not expect this variable to have significantly different values across different rating classes. We see the effect of a rating downgrade in the ratings 15-20 as these business partners are mainly the ones that received a downgrade into bucket three in the last 6 months. Recent downgrade has a significantly increasing effect on the PDs for these ratings. This is in line with the jump in the ADFs from rating 14 to 15. However, the rating bias in rating 15 (for details, please refer to Section 5.1.1) further escalates the jump in the ADFs and results in a gap between the ADFs and the estimated PDs. When analysing the overall average PDs in one period and in one rating class, we are not able to observe the effect of macroeconomic covariates, as they are the same for every business partner in a given period. However, the macroeconomic covariates have a level effect, i.e. they have an effect on the overall levels of the PDs over different periods, enhancing the 'point-in-timeness' of the model. All in all, the estimated PDs show a general trend that is similar to the ADFs. However, we see a strong deviation in the ratings 17-20. Due to the limited number of observations, we encounter a highly volatile ADF figure in these ratings, in which a one-unit increase in default has an immense effect on the frequency. To give an example, if [redacted] more business partners with the rating 19 defaulted, the ADF would increase to [redacted]%, which would even surpass the predicted PD. In order to obtain more reliable results when comparing the PDs and ADFs of these ratings, we would need more observations. However, considering the total ADF figures, we see that our model slightly underestimates the number of defaults in comparison to what actually occurred and this is mainly associated with the high ADFs in ratings 12 and 13 that have a significant effect on the total figure due to their weight, stemming from the high number of total observations.

6.2.2 Validation of PDs

Due to the limited number of observations in some of the rating classes and the characteristic of our model to predict PDs for each business partner using multiple input factors, a more sophisticated validation approach should be considered. Chen et al. (2005) provide a macro to validate the predictive ability of their survival model over time, using Receiver Operating Characteristics (ROC) curves for time-dependent analysis. We decided against using this validation method, because the informative value of such an ROC curve for static logistic models is limited, since static models do not account for time.
In order to allow comparison of our Cox model with existing static logistic regression models, we decided to validate our results based on a ROC method that is especially useful for continuous variables such as predicted probabilities.

Gönen (2006) presents a method to compare our predicted probabilities with the observed default frequencies of the respective business partners. This method enables validation on the level of business partners rather than on the level of factors, such as rating.
Two important concepts are the 'True Positive Rate' (TPR) and the 'True Negative Rate' (TNR). The combination of these two measures presents an accuracy measure that is calculated by ROC curves. In a binary model the TPR is the ratio of the business partners that were predicted to default and actually defaulted ('True Positives') over all the business partners that defaulted. Similarly, the TNR is the ratio of the business partners that were predicted to not default and did not default ('True Negatives') over all the business partners that did not default. The TPR is

referred to as the 'sensitivity' and the TNR is the 'specificity'.

$$sensitivity = TPR = \frac{TP}{TP + FN}, \; specificity = TNR = \frac{TN}{TN + FP} \tag{6.5}$$

TP is the amount of 'True Positives', FP is the amount of 'False Positives', TN is the amount of 'True Negatives' and FN is the amount of 'False Negatives'.

In order to use the concepts of TPR and TNR for a continuous predictor such as our probability of default, we can use a threshold level to separate our predicted probabilities. If we set the threshold at 90%, we can group our predicted probabilities into TP, FP, TN and FN and compare the 'sensitivity' with the 'specificity'. Since it is unclear how to set this threshold level, we should use a number of threshold levels and analyse the results for each one. Using ROC curves, and specifically the Area under the ROC curve (AUC), we can test for many threshold levels and visualise the results.
The AUC assesses the model's accuracy to discriminate between event and non-event observations. The better the predictive power of the model, the higher the AUC. Random predictions have an AUC of 0.5 and models that are not significantly different from 0.5 have little predictive power.

In Figure 6.1 we present the AUC of our model. With an AUC of 0.87, the explanatory power of our model is significantly higher than a random model with an AUC of 0.5.

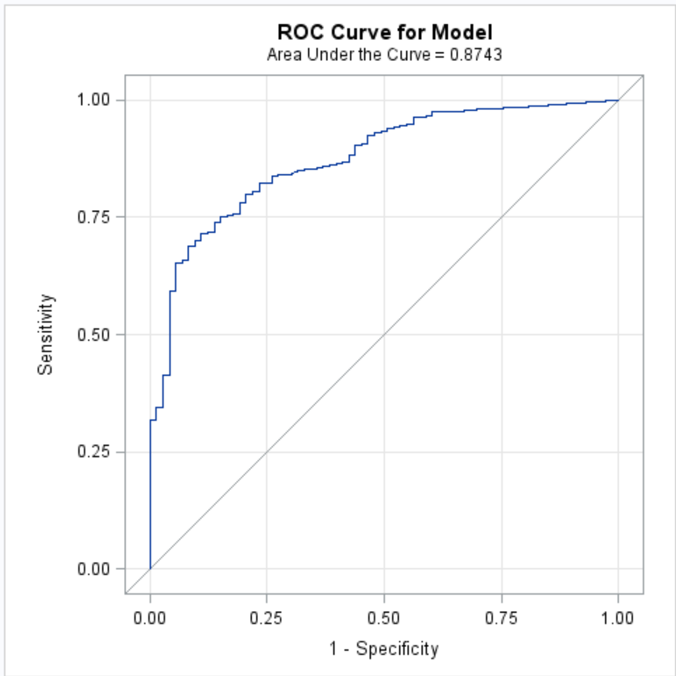

Figure 6.1: The Area Under the Curve (AUC) shows the predictive accuracy of the out-of-sample forecast of the dynamic Cox model. The PDs of each business partner are compared with the respective ADF. The AUC measures the discriminatory power of the model and random predictions have an AUC of 0.5. Models that do not beat this benchmark have no predictive power.

The high predictive power of our model emphasises the general suitability of hazard models for probability of default modelling according to IFRS 9. We showed that time-varying covariates such as a recent downgrade in a credit rating are significant determinants of default. Furthermore, we confirmed the relevance of macroeconomic factors for probability of default modelling. Lastly, we presented a PD model that is able to calculate both 12-month and lifetime PDs, which showed high absolute predictive power.

In future research, a model as ours should be cross-validated with another PD model such as a logistic regression model to assess the relative strength of a dynamic semi-parametric PD model.

7 Conclusion

The purpose of this study was to identify and build a probability of default model that is in line with the changes in the new accounting standard IFRS 9. In our analysis, we considered static logistic regression models, transformation models and hazard models in order to identify a suitable PD model.

Influenced by the work of Figlewski et al. (2012) and Kim and Partington (2014), we developed a semi-parametric Cox model which is based on a SAS macro developed by Chen et al. (2005). Our model allows both the calculation of 12-month PDs as well as lifetime PDs and includes a threshold for significant deterioration of credit risk.

In line with the literature, we have identified 'Current rating' as the covariate with highest explanatory power and highest effect on the resulting PDs. We have also included a downgrade factor to capture the effect of rating drift as well as a 'Months since first rated' covariate to determine the effect that time since initial rating has on PDs. One notable feature of our study was that we confirmed the high relevance of macroeconomic factors in PD calculations. Two macroeconomic covariates were found to be significant and their inclusion increased the maximum likelihood of the model while reducing the AIC value.

Within rating classes, we compared ADF and PD figures and were able to explain some of the deviations. Overall, our model makes an accurate estimation of the PDs, with a slight understatement. The average PD over all business partners has been estimated to be [redacted]% whereas the ADF over the same period amounts to [redacted]%. In a more flexible analysis, we validated the model using ROC curves for the 12-month PDs. The AUC of our model is 0.87, which implies strong discriminatory power.

While we were able to show high absolute predictive power in our out-of-sample validation, we have identified a few limitations and drawbacks.

Due to the scope of our thesis, it is difficult to assess the relative strength of our model. In order to make inferences about the strength of the model in comparison to other models, a thorough cross-validation between our model and alternative models should be undertaken. Similarly, in order to fully assess the effect of including macroeconomic covariates, additional analyses involving PD calculations with only firm-specific variables should be performed.

For a practical implementation, the long computation time of our model might pose a problem.[1] We ran into frequent computational problems despite performing our analysis on a relatively small subset of data. In contrast, large financial institutions are active in numerous countries and have millions of customers. The amount of data that is used for PD models on an entity-level exceeds the amount of data in our dataset significantly.

Furthermore, using a semi-parametric hazards model with time-varying covariates for PD calculation requires several steps that are not as intuitive as a simple logistic regression model.

Lastly, the proposed advantage of the Cox model's semi-parametric nature can turn

[1] When run on a machine with a 2.6GHz processor and 8GB RAM, it takes 65 hours to complete.

into a disadvantage when analysing small datasets. A semi-parametric hazard model allows a higher level of generalisation and flexibility than a fully parametric hazard model because it estimates the effect of time on the hazard rate endogenously. However, it is inappropriate to assume that we can adequately estimate the distribution for small datasets in which few business partners default with very long survival times. To remedy the survivorship bias in a semi-parametric hazard model, risk managers could include as much longitudinal data as possible, which in turn would increase the computation time of the model.

In conclusion, the proposed semi-parametric Cox model exhibits conceptual suitability for an IFRS 9 implementation and accurately estimates PIT PDs. However, the model has a number of limitations that need to be considered and require further research.

Bibliography

Aguais, S. (2008). Designing and implementing a basel ii compliant pit-ttc ratings framework.

Allison, P. D. (2014). *Event History and Survival Analysis: Regression for Longitudinal Event Data*, volume 46. SAGE Publications.

Althoff, J. and Lee, Y.-B. (2015 (accessed December 6, 2015)). Ifrs 9 - expected credit losses. http://www.pwc.com/ca/en/accounting-advisory-services/publications/us2014-06-ifrs-9-expected-credit-losses.pdf.

Altman, E. I. (1998). The importance and subtlety of credit rating migration. *Journal of Banking & Finance*, 22(10):1231–1247.

Altman, E. I. and Kao, D. L. (1992). Rating drift in high-yield bonds. *The Journal of Fixed Income*, 1(4):15–20.

Andersen, P. K. (1992). Repeated assessment of risk factors in survival analysis. *Statistical Methods in Medical Research*, 1(3):297–315.

Basel Committee (2004). An explanatory note on the basel ii irb risk weight functions.

Bellotti, T. and Crook, J. (2009). Credit scoring with macroeconomic variables using survival analysis. *Journal of the Operational Research Society*, 60(12):1699–1707.

Borucka, J. (2014). Methods of handling tied events in the cox proportional hazard model. *Studia Oeconomica Posnaniensia*, 2(2):91–106.

Breslow, N. E. (1972). Contribution to the discussion of the paper by dr cox. *Journal of the Royal Statistical Society, Series B*, 34(2):216–217.

Brunel, V. (2015). Ifrs 9 provision analytics. *Available at SSRN 2606080*.

Carlehed, M. and Petrov, A. (2012). A methodology for point-in-time–through-the-cycle probability of default decomposition in risk classification systems. *The Journal of Risk Model Validation*, 6(3):3–25.

Chen, L.-S., Yen, M.-F., Wu, H.-M., Liao, C.-S., Liou, D.-M., Kuo, H.-S., and Chen, T. H.-H. (2005). Predictive survival model with time-dependent prognostic factors: development of computer-aided sas macro program. *Journal of evaluation in clinical practice*, 11(2):181–193.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 187–220.

Efron, B. (1977). The efficiency of cox's likelihood function for censored data. *Journal of the American statistical Association*, 72(359):557–565.

Engelmann, B. and Porath, D. (2012). Do not forget the economy when estimating default probabilities. *Willmott Mag.*

Figlewski, S., Frydman, H., and Liang, W. (2012). Modeling the effect of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics & Finance*, 21(1):87–105.

Gebhardt, G., Novotny-Farkas, Z., et al. (2011). Mandatory ifrs adoption and accounting quality of european banks. *Journal of Business Finance & Accounting*, 38(3-4):289–333.

Gönen, M. (2006). Receiver operating characteristic (roc) curves.

Hamilton, D. T. and Cantor, R. (2004). Rating transitions and defaults conditional on watchlist. *Outlook and Rating History, Moody?s Investors Service, Special Comment*.

Hayden, E. and Porath, D. (2011). Statistical methods to develop rating models. In *The Basel II Risk Parameters*, pages 1–12. Springer.

Hosmer, D. W., Lemeshow, S., and May, S. (1999). Assessment of model adequacy. *Applied Survival Analysis: Regression Modeling of Time-to-Event Data, Second Edition*, pages 169–206.

IASB (2014 (accessed December 6, 2015)). Project summary - ifrs 9 financial instruments. `http://www.ifrs.org/current-projects/iasb-projects/financial-instruments-a-replacement-of-ias-39-financial-instruments-recognitio/documents/ifrs-9-project-summary-july-2014.pdf`.

Institute, S. A. S. (1999). *SAS/STAT user's guide, version 8*, volume 2. SAS Institute.

Kalbfleisch, J. D. and Prentice, R. L. (2011). *The statistical analysis of failure time data*, volume 360. John Wiley & Sons.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481.

Kennedy, K., Mac Namee, B., and Delany, S. J. (2010). Learning without default: A study of one-class classification and the low-default portfolio problem. In *Artificial Intelligence and Cognitive Science*, pages 174–187. Springer.

Kim, M. H. and Partington, G. (2014). Dynamic forecasts of financial distress of australian firms. *Australian Journal of Management*, page 0312896213514237.

Perederiy, V. (2015). Endogenous derivation and forecast of lifetime pds. *arXiv preprint arXiv:1507.05415*.

Royston, P. and House, A. (2011). Estimating a smooth baseline hazard function for the cox model. Technical report, Research report.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model*. *The Journal of Business*, 74(1):101–124.

Wagner, N. (2008). *Credit risk: Models, derivatives, and management.* CRC press.

Yurdakul, F. (2014). Macroeconomic modelling of credit risk for banks. *Procedia-Social and Behavioral Sciences*, 109:784–793.

8 Appendix

Rating Dummies (Sample period: January 2008 - December 2012)			
		Total	Events
Number of observations		●	●
Firm-specific covariates	coefficient	p value	Hazard ratio
Current rating = 3	-14.22253	0.9709	0.000
Current rating = 4	-14.25941	0.9695	0.000
Current rating = 5	-14.27751	0.9657	0.000
Current rating = 6	-14.26846	0.9565	0.000
Current rating = 7	-2.67391	0.0002	0.069
Current rating = 8	-3.63816	0.0003	0.026
Current rating = 9	-2.08020	<.0001	0.125
Current rating = 10	-2.09622	<.0001	0.123
Current rating = 11	-1.20344	<.0001	0.300
Current rating = 12	-0.65970	0.0030	0.517
Current rating = 13	-0.22625	0.2100	0.798
Current rating = 14	0.46571	0.0012	1.593
Current rating = 15	2.71010	<.0001	15.031
Current rating = 16	2.62355	<.0001	13.785
Current rating = 17	2.74891	<.0001	15.626
Current rating = 18	3.69813	<.0001	40.372
Current rating = 19	4.09985	<.0001	60.331
Current rating = 20	4.12655	<.0001	61.964
Model goodness of fit:			
-2 log L			17,636.71
Likelihood ratio			<.0001

Table 8.1: A Cox regression was fitted to dummies for each rating class. In order to assess the relative effect of each rating group, a dataset was prepared that included all unrated business partners. It may be assumed that the unobserved ratings of all unrated business partners are evenly distributed across rating classes. Therefore, the unrated rating class '100' was set as a reference value for the Cox regression. The dummies are grouped according to the order of size of the coefficient. These three groups are the rating buckets that were defined. Significant covariates at the 5% confidence level have a p-value below 0.05. For model goodness of fit the -2 log likelihood and the Akaike Information Criterion is presented.

Cox Regression with all covariates (Sample period: January 2008 - December 2012)						
		Total	Events		Total	Events
Number of observations:		●	●		●	●
	All covariates			**Backward selection**		
	coefficient	p value	Hazard ratio	coefficient	p value	Hazard ratio
Firm-specific covariates:						
Current rating	0.57530	<.0001	1.778	0.57537	<.0001	1.778
Downgrade into rat. group 7-14 (last 6 months)	-11.55431	0.9908	0.000			
Downgrade into rat. group 15-20 (last 6 months)	1.30019	<.0001	3.670	1.29938	<.0001	3.667
Upgrade into rat. group 3-6 (last 6 months)	-10.90157	0.9926	0.000			
Upgrade into rat. group 7-14 (last 6 months)	0.07151	0.8872	1.074			
Months since first rated (log)	-0.19782	0.0004	0.821	-0.19842	0.0004	0.820
General macroeconomic covariates:						
Real GDP growth (6-month lag)	-7.44044	<.0001	0.001	-8.09713	<.0001	0.000
Chg. in unemployment (6-month lag)	0.97889	0.2799	2.661			
Chg. in CPI (6-month lag)	-0.00388	0.9509	0.996			
Financial market covariates:						
Chg. in Nasdaq OMXS30 index (6-month lag)	-0.03541	0.9653	0.965			
Chg. in Swedish 3-month T-Bill rate (6-month lag)	1.09538	0.0003	2.990	1.19523	<.0001	3.304
Chg. In Swedish 10-year T-Bill rate (6-month lag)	-2.29874	0.0014	0.100	-2.26006	0.0004	0.104
Model goodness of fit:						
-2 log L			13,086.14			13,087.64
AIC			13,110.14			13,099.64

Table 8.2: Cox regression output over all covariates with the 10-year Swedish T-Bill rate. The left regression presents the results for a fitted model over all covariates. The right regression presents the final covariate selection after backward selection. Significant covariates at the 5% confidence level have a p-value below 0.05. For model goodness of fit the -2 log likelihood and the Akaike Information Criterion is presented.