Advanced Dependency Modeling in Credit Risk

Lessons for Loss Given Default, Lifetime Expected Loss and Bank Capital Requirements

A dissertation in partial fulfillment of the requirements for the degree of

Doktor der Wirtschaftswissenschaft (Dr. rer. pol.)

submitted to the

Faculty of Business, Economics, and Management Information Systems

Universität Regensburg

submitted by Steffen Krüger

Advisors:

Prof. Dr. Daniel Rösch

Prof. Dr. Alfred Hamerle

Date of the Disputation: August 23, 2017

Acknowledgments

I would like to thank Prof. Dr. Daniel Rösch for his continuous support of my doctoral thesis and his motivation, enthusiasm, and patience. His guidance and immense knowledge have been invaluable. I gratefully acknowledge Prof. em. Dr. Alfred Hamerle to be my second advisor.

The chapters of this thesis would not have been possible without the fruitful collaboration with several co-authors and colleagues. I want to express my gratitude to Prof. Dr. Harald Scheule for his expertise and the invitation to his university. I would like to thank Jennifer Betz for the great time we had when we shared an office.

Last but not least, I would like to thank my family and friends for their ongoing support. I am forever indebted to my mother and to Juli for all their love and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Motivation

Financial institutions play a major role for the economy since they ensure the supply of money by lending to corporations, sovereign entities and consumers (cf. Kishan and Opiela (2000) and Gambacorta and Shin (2016)). As financial intermediaries they fulfill at least three functions to match conflicting needs of borrowers and lenders (see e.g., Freixas and Rochet (2008) p. 4). Maturity transformation involves the conversion of short-term liabilities to long-term assets, e.g., deposits to loans. Size transformation encompasses the pooling of small amounts mainly from savers to large amounts for borrowers. Furthermore, financial institutions perform risk transformation, e.g., by diversification, and help to reduce the risk for single lenders. The matching of conflicting needs is subject to the risk of imbalances, e.g., payment difficulties of borrowers can prevent banks to fulfill their own liabilities. Therefore, financial institutions need to hold adequate capital reserves to protect themselves against credit risk (cf. Kim and Santomero (1988)). From an overall economic perspective, a distressed financial sector can lead to a reduction of lending. This can extend or intensify economic downturns, as corporations, sovereign entities and consumers are particularly in these times dependent on intermediaries. In extreme cases, a credit crunch can even cause a recession (cf. Akhtar (1994), Sharpe (1995) and Ivashina and Scharfstein (2010)). A professional measurement of the risk to which financial institutions are exposed to is a substantial basis to determine adequate capital reserves (cf. Sharpe (1978) and Dewatripont and Tirole (1994) p. 218 f.).

The management of financial institutions' capital reserves is subject to several difficulties. For instance, there are incentives to lower reserves in expansions, e.g., to invest more money in order to increase returns, which can promote capital shortfalls in recessions (cf. Gambacorta and Mistrulli (2004)). In addition, the measurement of credit risk is a challenging task and requires the use of sophisticated statistical methods. Even small mistakes in the evaluation of systematic risk, which substantially affects the credit risk with respect to recessions, can lead to severe misjudgments (cf. Kuo and Lee (2007), Duffie et al. (2009), Bade et al. (2011) and Rösch and Scheule (2014)).

Financial regulation can improve the capitalization of financial institutions by minimum capital requirements and minimum supervisory requirements to risk management. The Basel Committee on Banking Supervision (1988, 2006, 2011) subsequently elaborated three international frameworks to standardize and extend risk precaution across countries as well as institutions. The Basel Committee was founded as a response to the Herstatt liquidation in 1974, and in 1988 it recommended capital requirements on debt instruments based on the underlying credit risk (Basel I). The first Accord distinguished risk weights between different types of assets but did not account for differences within a given category. As a result, institutions were able to hold riskier assets without having to fulfill higher capital requirements (cf. Basel Committee on Banking Supervision (1999), Hull (2015) p. 336, and Baesens et al. (2016) p. 7). This was one of the reasons for the substantial revision of the framework. Basel II introduced the internal ratings-based approach that enables financial institutions to estimate credit risk by their own statistical models. In addition, it extended capital requirements to operational and market risk, and introduced rules for supervisory review and market discipline. The latest revision (Basel III) is currently being introduced as a response to the global financial crises starting in 2007. Current discussion and consultation papers of financial regulatory authorities show that the development of the regulatory framework is still far from being completed, e.g., Basel Committee on Banking Supervision (2015b, 2016a, 2017) and European Banking Authority (2017).

Credit risk provides the largest share on regulatory capital requirements. For instance, the latest Risk Assessment Report of the European Banking Authority (2016b) identifies that 80.5 % of the risk-weighted assets of 131 major EU banks were attributable to credit risk as of June 30, 2016. Credit risk of a financial instrument can be characterized by three risk parameters that are generally modeled as stochastic variables. The probability of default (PD) denotes the probability that a borrower will not fulfill his payment obligations in a given future time period. The loss given default (LGD) specifies the share of the outstanding debt that is lost due to default. The exposure at default (EAD) denotes the outstanding amount. Increased credit losses during economic downturns empirically show co-movements between risk parameters of a single instrument and those of several borrowers. This thesis focuses on the modeling of dependencies in credit risk which is particularly crucial for adequate risk precaution prior to recessions and discusses implications on bank capital requirements. The analyses cover, amongst others, the measurement and statistical modeling of systematic effects on the LGD and workout processes, and the co-movement of PDs and LGDs.

Literature

The statistical modeling and estimation of risk parameters play a major role in the literature on credit risk. On the one hand, recent studies examine determinants of credit risk which includes information on the underlying debt instrument and the borrower. However, clustered defaults and higher losses during recessions are caused by systematic risk factors, i.e., observable covariates such as macroeconomic variables and unobservable time-varying risk factors. This motivates research on downturn effects and co-movements of risk parameters. Workout processes of defaulted debt also are of particular interest as they are characterized by a high degree of heterogeneity and determine realized losses. Capital reserves basically are supposed to reduce the risk of financial distress for financial institutions. However, legal requirements can strengthen the burden to raise capital in recessions due to higher regulatory needs.

There are many papers dealing with the default risk of borrowers and financial instruments. The structural model of Merton (1974) assumes a stochastic process for the value of a company and provides a formula for the probability that its value is less than its debt. In contrast, the following studies use reduced form models and directly model the default event dependent on explanatory variables. The Z-score of Altman (1968) represents a formula to predict the probability of corporate bankruptcies based on firm characteristics. Categorical regression models are used by Martin (1977) and, recently, by Campbell et al. (2008), Campbell et al. (2011) and Hilscher and Wilson (2016). The authors study the PD for discrete time periods and account for various covariates, e.g., firm and instrument characteristics and macroeconomic information. Survival models extend the approach to the continuous default time, i.e., by considering default intensities as done by Lee and Urrutia (1996), Chava and Jarrow (2004), Das et al. (2007), Duffie et al. (2007) and Orth (2013).

There are at least two definitions for loss severity that are studied in the literature, i.e., workout and market-based LGDs. The first measure takes into account all post-default cash flows during the workout process of defaulted debt. A defaulted marketable debt instrument also provides a market-based LGD that characterizes the post-default decrease in the market price of the instrument. The literature examines both types of loss severity by several statistical models. The method of ordinary least squares is studied by Qi and Zhao (2011) and Jankowitsch et al. (2014) for comparative reasons and to analyze determinants of loss severity. Some models account for the property that LGDs are often bounded by zero (no loss) and one (total loss), e.g., fractional response models (Hu and Perraudin (2002), Dermine and Neto de Carvalho (2006) and Chava et al. (2011)) and beta regression (Gupton (2005) and Huang and Oosterlee (2012)). Further examined statistical methods are regression trees (Bastos (2010)) and mixture models (Altman and Kalotay (2014) and Calabrese (2014)). Comprehensive comparisons of the performance of LGD models are given by Qi and Zhao (2011), Loterman et al. (2012) and Yashkir and Yashkir (2013), amongst others.

Recessions generally increase the credit risk for single debt instruments because borrowers are systematically exposed to poor economic conditions in these times. Economic downturn periods are empirically characterized by clustered defaults and high losses, and thus indicate dependencies between the credit risk of financial instruments. Hu and Perraudin (2002) and Altman et al. (2005) empirically find a positive correlation of default rates and realized LGDs on aggregated data. Observable macroeconomic and industry-specific information can be included as covariates in standard regression models to account for systematic risk (e.g., Chava and Jarrow (2004), Acharya et al. (2007) and Bellotti and Crook (2012)). Duffie et al. (2009) and Lando and Nielsen (2010) find evidence for additional unobservable systematic effects in default risk which substantially increase credit risk in economic downturns. The dependency between default rates and LGDs is caused by observable systematic factors such as macroeconomic covariates (Chava et al. (2011)) and unobservable systematic factors that can be modeled by random effects (Bruche and González-Aguado (2010), Bade et al. (2011), Bellotti and Crook (2012) and Rösch and Scheule (2014)). The Basel Committee on Banking Supervision (2005) requires that LGD estimates shall reflect economic downturn conditions for regulatory purposes and account for the positive co-movement of PDs and LGDs. This is currently emphasized by the European Banking Authority (2017) which also advises to account for the bimodality of losses, i.e., the high number of total losses and recoveries.

Besides the question whether a borrower or debt instrument defaults, the literature also deals with the time *in* default. For instance, Bandopadhyaya (1994), Helwege (1999), Bris et al. (2006) and Denis and Rodgers (2007) study the time in bankruptcy and its determinants. Dermine and Neto de Carvalho (2006) and Gürtler and Hibbeln (2013) mention that the length of workout processes is empirically positively correlated with LGDs, i.e., the longer a debt is in default the more severe the loss is. The consideration of workout processes goes beyond the realized loss and provides additional information on the emergence of losses. Betz et al. (2016) find increased LGDs for long workout processes and time-varying levels in the length of defaults which indicates systematic effects in workout processes and workout LGDs.

Another field of study examines whether and to what extent bank capital requirements burden institutions, particularly in recessions. Although the Basel Accords are intended to prevent procyclicality, Gordy and Howells (2006) and Repullo and Suarez (2013) show that regulatory capital requirements increase in recessions. In addition to regulatory capital, institutions are obliged to hold loan loss provisions which are identified to increase during recessions due to the underlying incurred loss model (Laeven and Majnoni (2003), Bikker and Metzemakers (2005) and Fonseca and González (2008)). The revised loan loss provisioning of the International Accounting Standards Board (2014) and the Financial Accounting Standards Board (2016) is based on an expected loss model and is intended to reduce procyclical effects and increase transparency of provisioning. However, the European Banking Authority (2016a) and the Basel Committee on Banking Supervision (2017) propose a transition period to provide institutions sufficient time to raise capital. The Basel Committee on Banking Supervision (2016b) additionally points to the volatility of provisions due to the new standards.

Contribution

This thesis contributes to the literature on credit risk modeling and focuses on co-movements of risk parameters that intensify losses during recessions. The models provide more precise estimates of credit risk and a better understanding of systematic risk. This can improve risk-based capital reserves and can help to avoid a severe underestimation of risk and capital shortfalls in economic downturn periods. Furthermore, the discussion of regulatory requirements and the supervision of internal risk models can benefit from empirical results.

The literature and current discussions show that a closer look on the impact of economic downturns on LGDs is still necessary. At the same time, the bimodality of losses, i.e., the high number of total losses and recoveries, must be recognized. Although the literature mentions a positive dependency between the length of workout processes and LGDs, its significance and the role of systematic effects have not been analyzed. Furthermore, the revised loan loss provisioning will be based on lifetime expected losses and raises the following questions amongst others. First, the sample selection of loss data causes a positive dependency between single-period default risk and loss severity and must be examined with respect to the maturity of financial instruments, i.e., a multi-period modeling is necessary. Second, the impact of the new accounting standards on bank capital requirements that are given by loan loss provisioning and regulatory frameworks must be analyzed. This covers differences between both approaches of expected losses and how market participants and regulatory authorities may react. The research questions of the four following studies can be summarized as:

- How can the bimodality of losses, i.e., the high number of total losses and recoveries, adequately be modeled? Do covariate effects vary over the probability distribution of LGDs and, in particular, does the impact of economic downturns depend on quantiles as well as firm- and instrument-specific information?
- Are there systematic co-movements in the length of workout processes? How can the dependency be explained and to what extent does it affect credit risk (i.e., workout LGDs) and liquidity risk (i.e., the reduction of non-performing loans)?
- How does the positive dependency between default risk and loss severity evolve over multiple periods? What conclusions are necessary for the modeling and estimation of lifetime expected loss which is required by revised loan loss provisioning?
- What are the differences and similarities in the model requirements of the revised loan loss provisioning and the regulatory framework with respect to expected losses? What is the impact of the new accounting standards on bank capital requirements?

The research questions are examined by advanced statistical methods. First, the scope of LGD modeling is extended by proposing the quantile regression to separately regress each quantile of the distribution. This approach enables a new look on covariate and particularly downturn effects that vary over quantiles. Second, the length of workout processes is a time variable and thus modeled by a Cox proportional hazards model, which is similar to existing approaches on default times. Systematic effects are examined by the inclusion of time-varying frailties. Third, lifetime expected losses are modeled by a copula approach that combines accelerated failure time models for the default time with a beta regression of the LGD. The use of copulas provide continuous-time LGD forecasts and flexible dependence structures between default risk and loss severity. The fourth approach combines a Probit model for the PD and a fractional response model for the LGD to demonstrate the impact of revised loan loss provisioning on bank capital requirements. In addition, goodness-of-fit measures enable to validate these approaches. Simulation studies and analyses of representative portfolios provide implications and demonstrate the significance of empirical results.

The use of comprehensive data strengthens the validity of empirical results. The first two studies use unique loss data from several banks on defaulted corporate loans with jurisdiction in the United States, Great Britain and Canada. The database provides workout LGDs and information on workout processes between 2000 and 2013, i.e., post-default cash-flows including repayments and costs. The other two studies use information on US American corporate bonds that provides defaults and market-based LGDs between 1982 and 2014. Both databases contain comprehensive covariate information with respect to the underlying instruments, borrowers and macroeconomic conditions.

This thesis presents four studies which separately examine the proposed research questions in Chapters 2 to 5. The remaining part of the introduction summarizes each study with respect to motivation, data, statistical method and contribution. Chapter 6 presents a conclusion and provides findings, a discussion and an outlook.

Chapter 2: Downturn LGD Modeling using Quantile Regression

The aim of capital reserves to reduce the risk of capital shortfalls is particularly pronounced in recessions. The internal ratings-based approach of Basel II, therefore, requires LGD estimates that reflect economic downturn conditions. The study contributes to the literature by analyzing covariate and downturn effects on the entire bimodal shape of losses that is characterized by the high number of total losses and recoveries. The quantile regression of Koenker and Bassett (1978) and Koenker (2005) is used to separately model each quantile of the distribution and to allow for quantile-specific effects of covariates. The study is based on US American data of workout loan LGDs of small and medium enterprises with defaults between 2000 and 2013. LGD estimates are evaluated by goodness-of-fit measures that evaluate the entire distributional fit. The validation reveals advantages of quantile regression in comparison with standard regression techniques.

The analysis of quantile-varying covariate effects shows that the bimodality of losses can only partly be explained by firm- and loan-specific information as well as macroeconomic conditions. The paper concludes with a discussion on the impact of economic downturns on the distribution of loss rates and shows implications for the determination of downturn LGDs for regulatory purposes.

Chapter 3: Macroeconomic Effects and Frailties in the Resolution of Non-Performing Loans

The workout LGD of defaulted debt is determined by incoming cash flows and direct as well as indirect costs after default. The literature indicates that delayed workout processes are empirically correlated with higher losses. The study analyzes determinants of the length of workout processes and its significance for credit and liquidity risk of financial institutions. The time from default to resolution of defaulted debt is modeled by a Cox proportional hazards model with stochastic time-varying frailties. This approach is comparable to Duffie et al. (2009) who investigate the time to default of US American bonds. The study uses data of defaulted loans from small and medium enterprises and large corporates with jurisdiction in the United States, Great Britain and Canada with defaults between 2004 and 2013. Besides firm- and loan-specific information, the analysis discusses the role of observable (e.g., macroeconomic covariates) and unobservable (frailties) systematic factors. The latter are empirically identified to cause a significant co-movement of workout periods. In addition to a descriptive analysis of the correlation between resolution times and loss severity, a simulation study shows that economic downturns delay workout processes and thereby increase single-loan LGDs and portfolio losses. Furthermore, a second simulation study demonstrates that systematic effects between workout processes increase stable funding needs.

Chapter 4: A Copula Sample Selection Model for Predicting Multi-Year LGDs and Lifetime Expected Losses

The revised loan loss provisioning will be based on lifetime expected losses and raises the question how the positive co-movement of default risk and loss severity for the 12-month horizon behaves over the lifetime of financial instruments. The study develops a copula-based approach that enables a simultaneous modeling of default risk and loss severity for arbitrary time horizons. The rationale behind this is to combine regression models for the default time and the LGD by flexible dependence structures. The empirical work is done on US American corporate bond data for the years between 1982 and 2014. It controls for firm- and bond-specific as well as macroeconomic covariates. Several accelerated failure time models are examined to regress default times as in Lee and Urrutia (1996), Das et al. (2007) and Orth (2013). The LGD is modeled by beta regression which is motivated by Gupton (2005) and Huang and Oosterlee (2012). The analysis reveals that the positive dependency of PDs and LGDs cause a decreasing term structure of loss severity, i.e., the longer a bond survives the lower the expected LGD is. The use of several copulas also demonstrates that correlation measures are not able to adequately capture the identified dependence structure. Finally, the empirical results indicate that standard credit risk models generally underestimate lifetime expected losses.

Chapter 5: The Impact of Loan Loss Provisioning on Bank Capital Requirements

The replacement of the incurred loss model for loan loss provisioning by an expected loss approach provides a convergence to the regulatory approach. The study discusses the new standards and remaining differences between expected losses for accounting and regulatory purposes with respect to the rating philosophy (through-the-cycle vs. point-in-time) and the required time period of possible losses (12-month vs. lifetime expected losses). Standard regression techniques are used to estimate the risk parameters PD and LGD, i.e., the Probit model (cf. Puri et al. (2017)) and a fractional response model (cf. Chava et al. (2011)). The study uses data of US American corporate bonds that were originated between 1991 and 2013. Expected losses for regulatory and accounting purposes are estimated and provisions as well as capital requirements are computed for stylized portfolios that are motivated by Gordy (2000) and Gordy and Howells (2006). The study reveals a procyclical impact of the new accounting standards on regulatory capital requirements. The impact is analyzed for different portfolio qualities, reinvestment strategies and states in the economic cycle. In addition, the criterion of a significant increase in default risk that determines the classification of instruments and its role for the transparency of provisioning is discussed.

Chapters 2 to 5 consist of independent studies with varying sets of co-authors. The first study is published as an article in a peer-reviewed academic journal. The remaining three studies are under review at the submission date of this thesis. Since the thesis consists of independent studies submitted to journals with different style requirements, minor formal differences exist across the chapters.

Chapter 2

Downturn LGD Modeling using Quantile Regression

This chapter is joint work with Daniel Rösch[1] and published as:

Krüger, S., Rösch, D. (2017). Downturn LGD modeling using quantile regression.
Journal of Banking & Finance 79, 42–56. DOI: 10.1016/j.jbankfin.2017.03.001

Abstract: Literature on Losses Given Default (LGD) usually focuses on mean predictions, even though losses are extremely skewed and bimodal. This paper proposes a Quantile Regression (QR) approach to get a comprehensive view on the entire probability distribution of losses. The method allows new insights on covariate effects over the whole LGD spectrum. In particular, middle quantiles are explainable by observable covariates while tail events, e.g., extremely high LGDs, seem to be rather driven by unobservable random events. A comparison of the QR approach with several alternatives from recent literature reveals advantages when evaluating downturn and unexpected credit losses. In addition, we identify limitations of classical mean prediction comparisons and propose alternative goodness of fit measures for the validation of forecasts for the entire LGD distribution.

JEL classification: C51; G20; G28

Keywords: downturn; loss given default; quantile regression; recovery; validation

[1]Chair of Statistics and Risk Management, Faculty of Business, Economics, and Business Information Systems, University of Regensburg, 93040 Regensburg, Germany. E-Mail: daniel.roesch@ur.de

2.1 Introduction

2.1.1 Motivation and Literature

Practitioners and academics have investigated several statistical models for the measurement and management of credit risk. Initially, the focus was on the probability of default, whereas in recent years more attention was given to the loss in case of default. In addition, financial institutions need to evaluate their risk exposures for regulatory capital requirements. Basel II introduced the internal ratings-based approach which enables institutions to provide their own estimates for the Loss Rate Given Default (LGD). However, because realized losses show a strong variation, particularly in recessions, the Basel Committee on Banking Supervision (2005) points to the importance of adequate estimates for economic downturns and unexpected losses. Thus, recent literature has started to extend the focus on expected LGDs to modeling an entire LGD distribution, but then usually aggregates the results to predictions of the mean.

Probably the most convenient LGD models use Ordinary Least Squares (OLS) (e.g., Qi and Zhao (2011)). Due to the non-normality of data, Dermine and Neto de Carvalho (2006) point to the need of data transformation and take into account the property of rates being between zero and one. An alternative to capture the shape of random variables which are ratios is presented by Gupton and Stein (2005) which assume a beta distribution for losses. Regression Trees are a non-parametric regression technique which splits the sample into groups and uses the groups' means as estimates (e.g., Bastos (2010)). With respect to Downturn LGDs, Altman and Kalotay (2014) and Calabrese (2014) propose methods to decompose losses into possible components for different stages of the economic cycle. While the latter models take into account some aspects of Downturn LGDs, they usually do not model the entire distribution of LGD and do not provide determinants over the entire spectrum of LGDs.

An approach which is able to model the impact of covariates for LGDs over the entire distribution is Quantile Regression (QR). Somers and Whittaker (2007) use QR to model the asymmetric distributed ratio of house sale recoveries to indexed values and estimate mean LGDs of secured mortgages by aggregating these results. Siao et al. (2015) apply QR to directly predict mean workout recoveries of bonds and show an improved performance compared to several methods by evaluating mean predictions. The present paper extends these approaches by (1) studying a unique and comprehensive loan database of small and medium enterprises, (2) using Quantile Regression results for risk measures in a direct way without aggregating the distributions to mean predictions, and (3) extending the validation of mean predictions to evaluating the entire distribution of losses. This enables us to consider different parts of economic cycles and to provide an alternative approach for Downturn LGDs and unexpected losses.

Generally, the definition of Loss Rates Given Default differs depending on the kind of financial instrument. Market-based LGDs are given by one minus the ratio of bond prices after default and the instrument's par value. It is the loss due to the immediate drop of bond prices after default and represents market beliefs of future recoveries. Jankowitsch et al. (2014) investigate determinants of market-based LGDs for US corporate bonds. Their data are almost evenly distributed due to the weighting of extreme low and high final recoveries by mean expectations. In contrast, ultimate or workout LGDs reflect the finally realized payoff depending on the entire resolution process of the instrument and all corresponding cash flows and costs after default. Altman and Kalotay (2014) exhibit the high number of total losses and recoveries for workout processes of defaulted bonds and the difficulties of an adequate statistical modeling. Yashkir and Yashkir (2013) show that loan LGDs are bimodal, i.e., U-shaped for personal loans in the United Kingdom. Dermine and Neto de Carvalho (2006) as well as Bastos (2010) identify similar properties for Portuguese loans of small and medium entities.

Some risk factors cause contrasting recoveries and LGDs during the workout process. Examples are the nature of default, the type of resolution, the recovery of collateral, the macroeconomic conditions during resolution and also the length of the resolution process. The prediction of these variables is a difficult task and, thus, U-shaped loss distributions result prior to default. In addition, administrative, legal and liquidation expenses or financial penalties (fees and commissions) and high collateral recoveries cause a significant number of LGDs lower than zero or higher than one. Most of the above-mentioned approaches are not able to capture both properties. Therefore, we suggest Quantile Regression which is suited to model extremely distorted distributions with bimodal shape without specific distributional assumptions or value restrictions of realizations.

This paper contributes to the literature in the following ways. First, our approach models the entire distribution of LGDs by Quantile Regression rather than predicting mean values. Thus, adequate measures for downturn scenarios and unexpected losses result can easily be derived. Second, this approach gives new insights into the impact of covariates over the entire spectrum of LGDs. For bimodal loan loss data, we distinguish between influences on extreme low, median and extreme high LGDs. Third, we propose several validation approaches to compare our model to most common and capable methods by an in-sample and out-of-sample analysis.

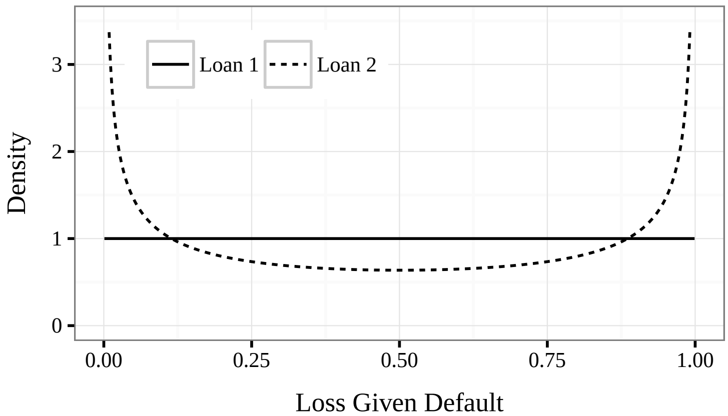
Following this introductory remarks, we motivate the relevance of adequate distributional estimates by an example. Section 2.2 contains a description of the QR approach and reference models. A simulation study explains theoretical advantages of the proposed method. Section 2.3 introduces the data and shows descriptive statistics. In Section 2.4, we show the model results and compare in-sample as well as out-of-sample performances of all models by different validation approaches. Afterwards, we show practical implications in Section 2.5 and propose downturn measures. Section 2.6 summarizes the results.

2.1.2 Introductory Example

Most academic and practical credit risk models focus on mean LGD predictions (cf. above mentioned literature). In addition, the Board of Governors of the Federal Reserve System (2006) proposes the computation of Downturn LGD measures by a linear transformation of means. This introductory example shows potential misleading results and interpretations.

Consider two loans with different stylized LGD distributions with the same mean. Let the loss of loan 1 be uniformly distributed between zero and one and the LGD of loan 2 be beta distributed with mean = 0.5 and variance = 0.125. Figure 2.1 shows the probability density functions. The means are equal, but the shapes differ. The LGD of loan 2 (dashed line) has more mass in the tails, particularly the probability of extreme losses is higher. Therefore, unexpected downturns may have a greater impact on loan losses compared to loan 1 (solid line). The concept of the Board of Governors of the Federal Reserve System (2006) does not account for this behavior and postulates the same downturn risk for both loans. The corresponding Downturn LGDs are given by $0.08 + 0.92 \cdot E(LGD) = 0.54$.

Figure 2.1: Stylized LGD Distributions



Notes: The figure shows two stylized LGD distributions with same means and, thus, same Downturn LGDs proposed by the Board of Governors of the Federal Reserve System (2006). However, real quantiles and downturn as well as unexpected risk differ.

This paper proposes an alternative approach of downturn considerations by using specific quantiles of the distribution, i.e., the Value at Risk (VaR). Thus, unexpected downturn effects are modeled more accurately. For example, the 75 % - VaR for loan 1 is 0.75 due to uniformity. Loan 2 has a 75% - VaR of 0.85 and therefore higher risk for this unexpected downturn. In this example, differences are found for all quantiles above the 50 % - level and, thus, for potential downturn risk (see Table 2.1). This shows the importance of adequate LGD models for regulatory purposes and the practical assessment of credit risk by considering the entire spectrum of the LGD distribution and not the mean only. This can easily be achieved by Quantile Regression which we briefly describe in the next section.

Table 2.1: Exemplary Downturn Loss Rates Given Default

	DLGD	Value at Risk			
	(FED)	50 %	75 %	90 %	95 %
Loan 1: LGD in %	54.00	50.00	75.00	90.00	95.00
Loan 2: LGD in %	54.00	50.00	85.36	97.55	99.38
ΔLGD in %	0.00	0.00	13.81	8.39	4.62

Notes: The table shows several LGD quantiles for two exemplary loans of Figure 2.1. Both distributions have the same mean and, thus, the same Downturn LGD proposed by the Board of Governors of the Federal Reserve System (2006): $0.08 + 0.92 \cdot E(LGD) = 0.54$. However, the first loan has a uniform distributed LGD, whereas the second follows a beta distribution. This results in different Values at Risk.

2.2 Modeling Loss Given Default

2.2.1 Quantile Regression

Standard regressions, e.g., the method of ordinary least squares, model the mean and do not adequately consider the entire distributional nature. They usually fail making adequate VaR forecasts when the LGD distribution is, for instance, bimodal. A useful alternative is Quantile Regression (QR) proposed by Koenker and Bassett (1978) and Koenker (2005). For $Y = LGD$ the VaR to a certain level $\tau \in (0, 1)$ is the (unconditional) τ-quantile

$$Q_Y(\tau) = \inf\{y \in \mathbb{R} | F_Y(y) \geq \tau\}, \tag{2.1}$$

which is the maximum loss that is exceeded in at least $(1 - \tau) \cdot 100\,\%$ of all cases. QR allows modeling each quantile individually by separate regressions. Suppose we want to model the (conditional) τ-quantile given some control variables, the regression equation is

$$Y_i = \beta(\tau)' x_i + \varepsilon_{i\tau} \tag{2.2}$$

with loans $i = 1, \ldots, n$ and x_i as covariate vector including a one for the intercept and $\beta(\tau)$ as unknown parameter vector. With $Q_\tau(\varepsilon_{i\tau}) = 0$, we have the τ-quantile of the LGD given by $Q_\tau(Y_i) = \beta(\tau)'x_i$. There is no need to make any more model assumption for the errors $\varepsilon_{i\tau}$ except of uncorrelatedness. Thus, non-normal, skewed and bimodal behaviors of LGDs can easily be handled. Hence, we allow for a high level of error heterogeneity. The unknown parameter vector is estimated by minimizing the following objective function over $\beta(\tau)$:

$$\sum_{i=1}^{n} \rho_\tau(y_i - \beta(\tau)'x_i) \quad \text{with} \quad \rho_\tau(x) = \begin{cases} \tau x & , \text{ if } x \geq 0, \\ (1-\tau)|x| & , \text{ else.} \end{cases} \tag{2.3}$$

Equation (2.3) shows the minimization of weighted residuals by the asymmetric loss function ρ_τ. In contrast, the method of Ordinary Least Squares uses a quadratic loss function to model the mean. Here, for the right tail of the distribution positive residuals are stronger weighted by $\tau \in (0.5, 1)$ than negative residuals by $1 - \tau$. This leads to the regression of the τ-th quantile with an unbiased, consistent and asymptotically normally distributed estimator.

2.2.2 Simulation Study

In contrast to standard regression techniques, Quantile Regression is able to capture several statistical issues that may cause non-normality. To illustrate this property, we consider two exemplary simple linear models with different reasons for bimodality. Given the data generating processes of Table 2.2, we simulate two datasets and identify serious issues when estimating the linear models with OLS and show the capability of Quantile Regression. For each model, a normally distributed covariate x is connected over a linear function with an intercept β_1, slope β_2 and an error term u.

The first model is given by an intercept of $\beta_1 = 0.5$, a slope of $\beta_2 = 1$ and a bimodal error that is generated by the sum of a Gaussian and a Bernoulli distribution. The second model uses a normally distributed error but assumes varying values for the intercept and

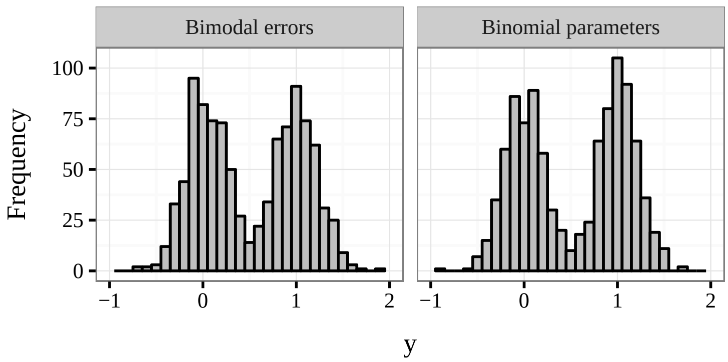
Table 2.2: Simulation – Data generation

	Bimodal errors	Binomial parameters
Model	$y_i = \beta_1 + \beta_2 x_i + u_i$	$y_i = \beta_{1\tau_i} + \beta_{2\tau_i} x_i + u_i$
Covariate	x_i as sample of $X \sim N(0,\, 0.01)$	
Parameters	$\beta_1 = 0.5$ $\beta_2 = 1$	$\beta_{1\tau_i} = z_{\tau_i}$ $\beta_{2\tau_i} = z_{\tau_i} - 0.5$ $Z_{\tau_i} \sim B(0.5)$
Error term	$U_i = V_i + W_i$ $V_i \sim N(0,\, 0.04)$ $W_i \sim B(0.5)$	$U_i \sim N(0,\, 0.04)$

Notes: We simulate bimodal data with the models shown here. All random variables are generated independently. Afterwards, we estimate both models with OLS and Quantile Regression.

the slope. Lower conditional quantiles, i.e., given the covariate information, are generated with $\beta_1 = 0$ and $\beta_2 = -0.5$. In contrast, upper conditional quantiles are calculated with $\beta_1 = 1$ and $\beta_2 = 0.5$. A Bernoulli distribution determines independently for each individual whether a lower or upper quantile is generated. We refer to both models as 'bimodal errors' and 'binomial parameters' model and independently simulate 1,000 observations each. Figure 2.2 shows the resulting bimodal distribution of the generated data.

Figure 2.2: Simulation – Data generation

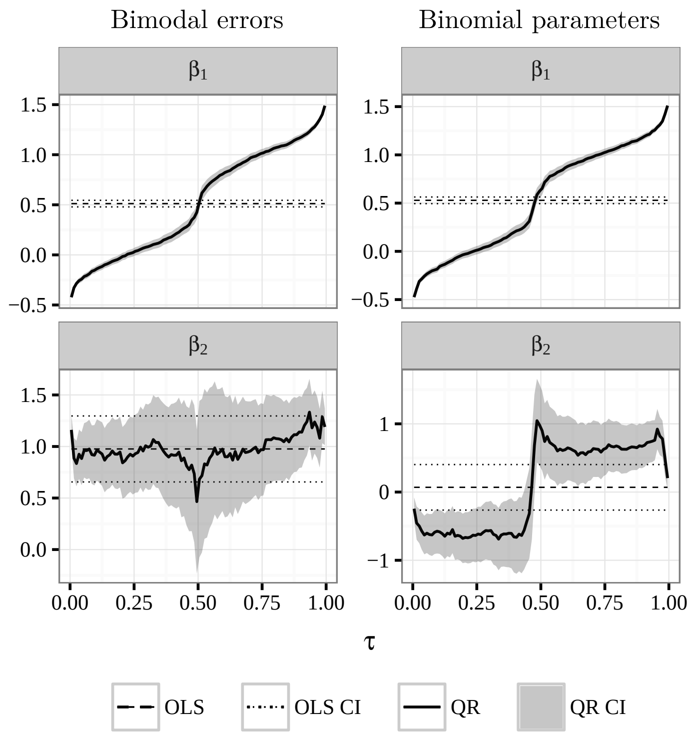


Notes: The figure shows simulated observations for both linear models of Table 2.2. The left panel is generated by bimodal errors, whereas in the right panel parameters vary over quantiles.

After generating bimodal data, we estimate linear models for the dependent variable y with an intercept β_1, the covariate x and the corresponding slope β_2. From this point of view, i.e., for estimation, the source of bimodality is unknown for both datasets.

Figure 2.3 shows parameter estimates for OLS and Quantile Regression. For the model with bimodal errors, which is shown on the left panel, OLS leads to plausible and constant parameter estimates. QR shows varying intercepts for different quantiles, because it aggregates the information of a regression constant and a bimodal error term. Overall, the slope is reliably estimated by QR. For the second model with normally distributed errors but binomial parameters, OLS is unable to capture non-constant parameters. This leads to distorted estimates of the intercept and the slope. The latter parameter is not even statistically significant different from zero, i.e., OLS does not capture the significant role of the covariate. In contrast, Quantile Regression is able to capture the differences in lower and upper quantiles. The intercept contains the two distinct values for the level as well as the normally distributed error. The slope is identified as a statistically significant covariate with varying influence.

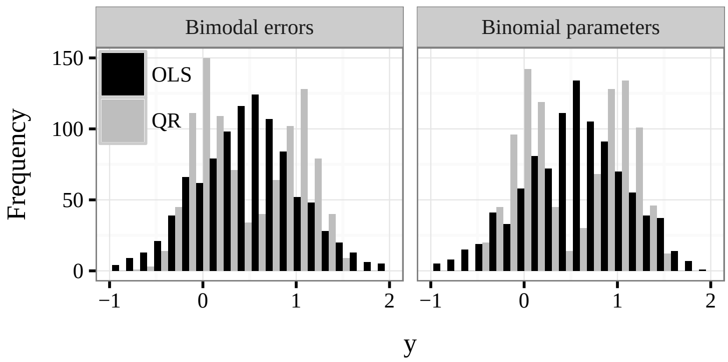
Figure 2.3: Simulation – Parameter estimation



Notes: The figure shows parameter estimates for the data shown in Figure 2.2 and generated by the models of Table 2.2. The left panel contains estimates for the model with a bimodal error term, whereas the right panel presents results for the model with binomial parameters. The confidence intervals (CI) are shown for the 95 % - level and are calculated for Quantile Regression with standard errors that are estimated by kernel estimates.

For both data generating processes, Figure 2.4 shows 1,000 independently simulated observations each using OLS and QR parameter estimates of Figure 2.3. Neither the bimodal error nor the binomial parameters are captured by OLS due to the assumptions of normally distributed errors and constant parameters. Thus, only QR is able to capture the bimodality of the data generating processes.

Figure 2.4: Simulation – Forecasting



Notes: The figure shows simulated data using the parameter estimates of Figure 2.3. The bimodality of the data generating process is only captured by Quantile Regression.

2.2.3 Comparative Methods

This paper compares QR to the most popular LGD models. We use the OLS method as the first benchmark model as proposed in Qi and Zhao (2011) and calculate quantiles by fitting a normal distribution for estimated errors. Secondly, a Fractional Response Model (FRM) is used which transforms the dependent variable before doing an OLS to bring it closer to normality (cf. Dermine and Neto de Carvalho (2006)). Here, we use the inverse normal regression by transforming the LGD with the inverse cumulative density function of the normal distribution as proposed by Hu and Perraudin (2002).[2]

The FRM and most LGD models assume the dependent variable to be a rate, i.e., to be bounded by zero and one. Here, we have data with plausible values out of this range, which is also relevant for creditors. Reasons are administrative, legal and liquidation expenses or financial penalties (fees and commissions) and high collateral recoveries. In

[2]We tested several transformation, but only present the best performing alternative, i.e., the inverse normal regression.

order to apply the FRM and other models for comparative reasons, we transform the observed LGD to lie in this interval for these models.[3] In addition, we add an adjustment parameter of 10^{-9} to avoid values on the bounds as proposed by Altman and Kalotay (2014). After estimation we re-transform predicted values to the observed LGD scale.

Gupton and Stein (2005) propose Beta Regression (BR) which is our third benchmark model.

The fourth model is Regression Trees (RT) which are a non-parametric regression technique, which splits data into groups and uses the groups' averages of the dependent variable as their mean prediction (see Bastos (2010)). Similar to OLS, we fit normal distributions per group for quantile predictions.

Table 2.3: Comparative methods

Method	Exemplary literature
Ordinary Least Squares (OLS)	Qi and Zhao (2011), Loterman et al. (2012)
Fractional Response Model (FRM)	Bastos (2010), Chava et al. (2011), Qi and Zhao (2011), Altman and Kalotay (2014)
Beta Regression (BR)	Huang and Oosterlee (2012), Loterman et al. (2012), Yashkir and Yashkir (2013)
Regression Tree (RT)	Bastos (2010), Qi and Zhao (2011), Loterman et al. (2012), Altman and Kalotay (2014)
Finite Mixture Model (FMM)	Calabrese and Zenga (2010), Loterman et al. (2012), Altman and Kalotay (2014)

Notes: We compare the performance of each of the five models with Quantile Regression.

Finally we apply Multistage Models which estimate probabilities of low, middle or high LGDs. Afterwards, the distribution of each component is modeled (see Altman and Kalotay (2014)). Some of these methods directly model probabilities of obtaining values of exactly zero or one and do not allow for values outside these bounds. Thus, these models are not able to cover the kind of bank loan LGDs that we observe, i.e., with several observations lower than zero and higher than one. In contrast, a Finite

[3] The transformation is done by $LGD_{[0,1]} = \frac{LGD - \min(LGD)}{\max(LGD) - \min(LGD)}$.

Mixture Model (FMM) assumes a mixed distribution of several margins. This paper uses a normal mixture distribution for the LGD.[4] Quantiles are directly observed from mixture distributions. Table 2.3 shows a summary of comparative methods from the literature.

2.3 Data

This paper uses a dataset of US American defaulted loans of small and medium enterprises from Global Credit Data (GCD) which provides the world largest LGD database. The association consists of 52 member banks from all over the world including global systemically important banks. Members exchange default data to develop and validate their credit risk models. We restrict the sample to all defaults since year 2000 to ensure the consistent default definition of Basel II. Since workout processes of recent defaults are not necessarily completed, we do not account for defaults after 2013. Cures are not considered because they do not provide default data with actual losses. We remove loans with exposures at default of less than 500 US $ because of minor relevance and to satisfy a materiality threshold.

Since we use loan data, we consider workout Recovery Rates (RR) including post-default cash flows. The RR is the difference of all discounted incoming cash flows (CF^+) and discounted direct as well as indirect costs (CF^-), divided by the exposure at default (EAD). The LGD is given as one minus the RR:

$$LGD = 1 - \frac{\sum CF^+ - \sum CF^-}{EAD}. \tag{2.4}$$

Incoming cash flows (CF^+) are: principal, interest and post-resolution payments, the recorded book value of collateral, received fees and commissions, and waivers. Direct and indirect costs (CF^-) are: legal expenses, administrator and receiver fees, liquidation expenses, and other workout costs. In addition to the exposure at default, the nominator

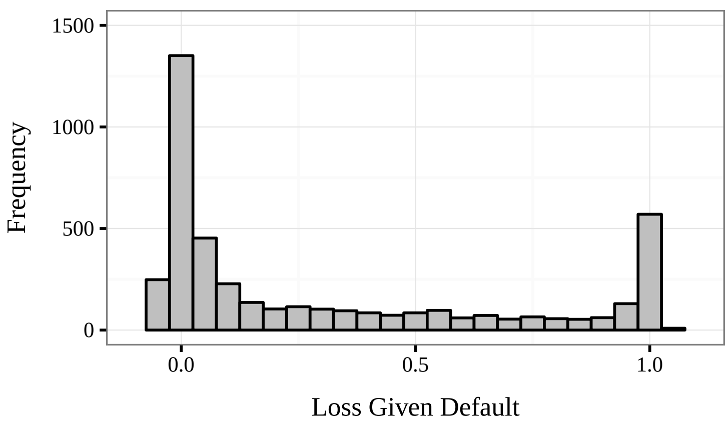
[4] In contrast to Altman and Kalotay (2014), we do not transform the dependent variable because we observe values out of the interval between zero and one. Actually, a transformation would lower the goodness of fit for mixture models for our dataset. Thus, we use raw data to allow a fair comparison. In addition, we tested a beta mixture distribution, which we do not report because of a worse performance compared to the normal mixture model.

of Equation (2.4) contains discounted principal advance, financial claim, interest accrual, and fees and commissions charged. All numbers are discounted by LIBOR to the default date.

We use two selection criteria due to Höcht and Zagst (2010), Höcht et al. (2011) and Betz et al. (2016) to ensure consistent as well as plausible data and to detect outliers. First, we consider loans for which the sum of all cash flows and further transactions, e.g, charge-offs, are between 90 % and 110 % of the outstanding exposure at default. Second, we regard loans with payments between -10 % and 110 % of the outstanding exposure at resolution. Actually, this selection procedure is less important for Quantile Regression because it is less sensitive to outliers. Nevertheless, we remove these loans as well as all observations with LGDs outside of $[-0.1, 1.1]$ to allow a fair comparison to standard regression techniques which are typically more sensitive to outliers.

Figure 2.5 shows a histogram of the final dataset. Most LGDs are nearly total losses or total recoveries which yields to a strong bimodality. The mean is given by 0.31 and the median by 0.09, i.e., LGDs are highly skewed. Both properties of the distribution may favor the application of QR because most standard methods do not adequately capture bimodality and skewness. Furthermore, many LGDs are lower than 0 and higher than 1 due to administrative, legal and liquidation expenses or financial penalties (fees and commissions) and high collateral recoveries.

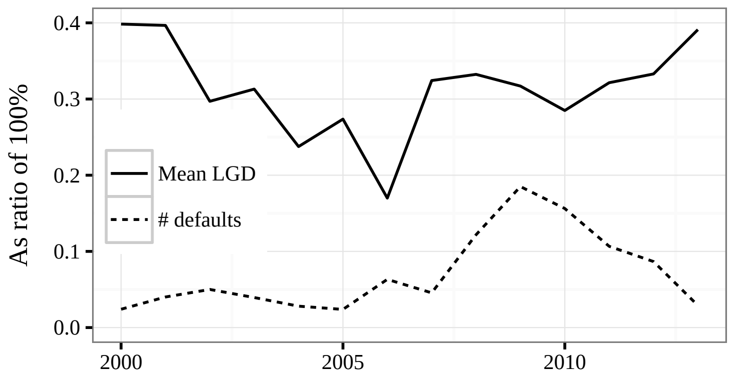
Figure 2.5: Observed Loss Rates Given Default



Notes: LGD calculation includes all cash flows and direct as well as indirect costs. All components are discounted to the loan's default date.

The yearly mean LGD and the distribution of default over time are visualized in Figure 2.6. The number of observed defaults strongly increased during the Global Financial Crises and slowly decreased in the aftermath. Loss rates were high due to the recession in 2000 and 2001 and subsequently decreased. In the last crises, the loss severity returned to a high level, where it remains since then.

Figure 2.6: Loss Rates Given Default over time



Notes: The figure shows yearly mean LGDs and the ratio (number of defaults per year) / (total number of defaults).

Table 2.4 shows descriptive statistics for the covariates of which we can make use in the database. For metric variables we report their means and several quantiles. For each level of categorical variables we show means and category-specific quantiles of the LGD as well as the number of observations per group. Medium term facilities imply higher LGDs than other facility types. For example, the median realized LGD for medium term facilities is 0.18 compared to 0.03 for short term facilities and 0.01 for other types. The seniority plays an important role for losses. On average non senior secured loans result in high LGDs of 0.64 compared to the super senior category with 0.30. We distinguish between loans with and without guarantees, i.e., whether there is an additional guarantor for a loan. In the mean, there is only a small difference for loans with or without guarantees. However, in the median we see a difference for loans with (0.16) and without (0.07) guarantees. Although a guarantee should reduce future losses, it can provide asymmetric information problems like moral hazard which can lead to higher losses. In addition, we distinguish between loans without any collateral, with real estate and with other collateral. Loans with real estate collateral produce lowest mean LGDs with 0.17 compared to 0.44 for loans without any collateral. The firms' industry affiliations may determine the magnitude of losses. For most industries, the mean LGD is around 0.30 - 0.35. The maximum value is given by the group of agriculture, etc. (0.46) and the minimum is given by the mining industry (0.21). Financial affiliation results in low mean LGDs of 0.25.

Table 2.4: Descriptive statistics

Variable	Level	Quantiles					Mean	Obs.
		0.05	0.25	0.50	0.75	0.95		
LGD		-2.99	0.32	9.01	61.56	100.00	31.27	4,308
log(EAD)		9.52	11.38	12.68	13.82	15.68	12.61	4,308
Facility type	Medium term	-0.90	2.40	17.52	67.46	100.00	35.58	2,372
	Short term	-1.30	0.87	3.28	51.87	100.00	26.57	559
	Other	-4.45	-1.02	1.31	49.74	100.00	25.76	1,377
Seniority code	Pari-passu	2.60	6.92	17.84	65.41	102.32	35.77	713
	Super senior	-3.38	0.09	5.07	58.75	100.00	29.85	3,539
	Non senior	-0.17	16.38	91.02	100.75	100.79	63.64	56
Guarantee indicator	N (no guarantee)	-3.65	0.00	6.92	62.99	100.00	30.20	2,587
	Y (guarantee)	-0.73	1.53	15.79	60.48	100.00	32.88	1,721
Collateral indicator	N (no collateral)	-2.46	0.42	33.28	96.29	100.00	44.01	1,152
	Other	-2.24	0.72	8.02	54.76	100.00	29.33	2,464
	Real estate	-4.97	-0.58	1.56	22.97	95.97	16.99	692
Number of collateral		0.00	0.00	1.00	1.00	4.00	1.32	4,308
Industry								
Finance, insurance, real estate	(FIRE)	-4.10	-0.29	5.96	43.87	100.00	25.38	646
Agriculture, foresty, fishing, hunting	(AFFH)	-3.80	-0.14	61.04	84.96	100.33	46.43	36
Mining	(MIN)	-0.42	0.71	2.06	20.83	98.94	21.43	21
Construction	(CON)	-3.87	-0.52	8.51	56.53	100.00	29.23	459
Manufacturing	(MAN)	-2.71	0.12	4.96	68.33	100.00	31.55	639
Transp., commu., elec., gas, sani. serv.	(TCEGS)	-1.33	0.36	22.13	59.72	100.00	34.62	173
Wholesale and retail trade	(WRT)	-3.41	0.05	4.55	62.60	100.00	30.93	590
Services	(SERV)	-3.77	0.13	5.42	81.29	100.00	33.76	704
Other	(Other)	0.66	3.93	14.71	59.90	101.53	33.29	1,040
S&P 500 (rel. change)		-0.28	0.01	0.11	0.18	0.33	0.08	4,308
TED spread (abs. spread in p. p.)		0.13	0.19	0.26	0.46	1.32	0.44	4,308
Term spread (abs. spread in p. p.)		0.07	1.94	2.63	3.28	3.50	2.40	4,308
VIX (abs. in p. p.)		12.70	16.52	18.00	25.61	42.96	21.90	4,308

Notes: The table shows means and quantiles of empirical LGDs for categorical variables with corresponding numbers of observation. In addition, the means and quantiles of metric variables, i.e., LGD, log(EAD) and number of collateral are given.

We also use two macroeconomic control variables. For the overall real and financial environment, we use the relative year-on-year growth of the S&P 500. Stock exchange performances are identified as general LGD drivers, e.g., by Qi and Zhao (2011) and Chava et al. (2011). Both papers study bond data and identify the three-month treasury-bill rate as a significant variable to consider expectations of future financial and monetary conditions. This paper uses loan data for which we find the absolute term spread between 10-years and 3-months US treasury rates to be significant drivers for LGDs which is also identified by Lando and Nielsen (2010) as a driver of credit losses in general. Macroeconomic information corresponds to each loan's default year. Both variables result in most plausible and significant results when testing different lead and lag structures. We also tested other popular macroeconomic variables, e.g., inflation, the volatility index VIX, industry production, gross domestic product, interest rates and the TED spread, which result in less explanatory power.[5]

2.4 Empirical Analysis

2.4.1 Quantile Regression

In this section the results of QR are provided. Table 2.5 shows the estimation results of QR for the 5th, 25th, 50th, 75th and 95th percentile and the corresponding OLS estimates. A full picture for all percentiles is given in the Appendix (Figure 2.A.1). As can be seen the level and significance of the parameter estimates strongly depend on the respective quantile.

The regression constant shows the behavior of the dependent variable when keeping covariates at zero. The course does not suggest normally distributed error terms, which would result in a smoothed s-shape. For low quantiles the intercept is near total recovery and starts to increase monotonically around the median until it reaches its maximum around total loss at higher quantiles. Loan losses generally exhibit a wide range regardless of control variables, which is due to the bimodal nature of LGDs. At the tails (5th and 95th percentile) most covariates are not statistically significant different from zero. This means that the extreme events, e.g., very high LGDs, are not explainable by our covariates and mainly due to unpredictable events.

[5] We would like to thank an anonymous referee for suggesting these variables. Section 2.5.2 contains a more detailed discussion of the choice of macroeconomic indicators on the example of the TED spread and the VIX.

Table 2.5: Parameter estimates Quantile Regression

		$\tau = 0.05$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.95$	OLS
(Intercept)		0.0332 **	0.0508 *	0.2238 ***	1.0682 ***	1.0373 ***	0.5189 ***
		(0.0152)	(0.0284)	(0.0607)	(0.0934)	(0.0516)	(0.0493)
log(EAD)		-0.0013	0.0009	0.0043	-0.0269 ***	-0.0015	-0.0103 ***
		(0.0010)	(0.0019)	(0.0033)	(0.0069)	(0.0034)	(0.0032)
Facility type	Short term	-0.0050	-0.0086	-0.0337 **	-0.0701 **	-0.0045	-0.0713 ***
(Medium term)		(0.0050)	(0.0095)	(0.0154)	(0.0341)	(0.0178)	(0.0178)
	Other	-0.0231 ***	-0.0218 ***	-0.0443 ***	-0.0452 *	-0.0014	-0.0618 ***
		(0.0040)	(0.0074)	(0.0119)	(0.0262)	(0.0120)	(0.0135)
Seniority	Super senior	-0.0368 ***	-0.0565 ***	-0.0748 ***	-0.1764 ***	-0.0193	-0.0829 ***
(Pari-passu)		(0.0060)	(0.0107)	(0.0167)	(0.0444)	(0.0206)	(0.0183)
	Non senior	-0.0164	0.1046	0.5896 ***	-0.0590	-0.0117	0.1818 ***
		(0.0351)	(0.0747)	(0.0541)	(0.0605)	(0.0254)	(0.0533)
Guarantee	Y	0.0134 ***	0.0207 ***	0.0474 ***	0.0296	0.0020	0.0574 ***
(N)		(0.0041)	(0.0074)	(0.0109)	(0.0221)	(0.0126)	(0.0123)
Collateral	Other	0.0034	-0.0127	-0.2021 ***	-0.2756 ***	0.0042	-0.1252 ***
(N)		(0.0056)	(0.0104)	(0.0301)	(0.0379)	(0.0149)	(0.0161)
	Real estate	-0.0094	-0.0183	-0.2279 ***	-0.5073 ***	-0.0310	-0.2098 ***
		(0.0063)	(0.0115)	(0.0303)	(0.0331)	(0.0317)	(0.0199)
Number of collateral		-0.0044 ***	-0.0024	-0.0056 **	-0.0468 ***	-0.0062	-0.0188 ***
		(0.0011)	(0.0021)	(0.0028)	(0.0048)	(0.0076)	(0.0044)
Industry	AFFH	-0.0025	0.0049	0.3885 ***	0.1387 **	0.0036	0.1674 ***
(FIRE)		(0.0250)	(0.0473)	(0.1055)	(0.0589)	(0.0464)	(0.0629)
	MIN	0.0142	0.0006	-0.0324	-0.1902 **	-0.0166	-0.0747
		(0.0182)	(0.0336)	(0.0488)	(0.0885)	(0.0869)	(0.0815)
	CON	0.0020	0.0024	0.0155	0.0830 **	0.0016	0.0335
		(0.0066)	(0.0122)	(0.0184)	(0.0414)	(0.0248)	(0.0226)
	MAN	0.0092	0.0021	0.0137	0.1433 ***	0.0001	0.0471 **
		(0.0061)	(0.0113)	(0.0180)	(0.0392)	(0.0207)	(0.0212)
	TCEGS	0.0097	0.0118	0.0390	0.1107 *	0.0032	0.0608 *
		(0.0107)	(0.0195)	(0.0290)	(0.0590)	(0.0278)	(0.0317)
	WRT	0.0006	-0.0010	0.0214	0.2177 ***	0.0074	0.0734 ***
		(0.0061)	(0.0114)	(0.0180)	(0.0347)	(0.0201)	(0.0212)
	SERV	-0.0028	-0.0015	0.0573 ***	0.1988 ***	0.0020	0.0744 ***
		(0.0059)	(0.0109)	(0.0192)	(0.0390)	(0.0190)	(0.0204)
	Other	0.0165 ***	0.0195 *	0.0631 ***	0.1517 ***	0.0058	0.0674 ***
		(0.0060)	(0.0109)	(0.0159)	(0.0395)	(0.0208)	(0.0199)
S&P 500		-0.0150	-0.0221	-0.0818 ***	-0.0505	-0.0070	-0.0763 **
		(0.0096)	(0.0176)	(0.0272)	(0.0561)	(0.0299)	(0.0304)
Term spread		0.0008	0.0049	0.0223 ***	0.0724 ***	-0.0003	0.0331 ***
		(0.0017)	(0.0031)	(0.0048)	(0.0128)	(0.0065)	(0.0060)
R^1 resp. R^2		0.0522	0.0386	0.0708	0.1158	0.0067	0.0949
Obs.		4,308	4,308	4,308	4,308	4,308	4,308

Notes: The table shows parameter estimates of Quantile Regression for the 5th, 25th, 50th, 75th and 95th percentiles. Standard errors are given in parentheses by kernel estimates. Significance is indicated by '*' (10 %), '**' (5 %) and '***' (1 %). Further percentile estimates are shown in Appendix Figure 2.A.1. The last column shows OLS estimates. Reference groups of categorical variables are given in parentheses.

In the left part of the distribution (5th and 25th quantile) only the following small effects can be observed: Short and medium term facilities have 2.2 - 2.3 percentage points higher losses than other facilities. The super senior status of loans lowers even small LGDs by 3.7 to 5.7 percentage points due to the priority in resolution processes. Interestingly, the existence of guarantees increases losses by 1.3 to 2.1 percentage points which might be caused by additional administrative costs when the guarantor becomes active. This effect is even more pronounced by a rise up to 4.7 percentage points in the median case.

Most control variables show significant effects in the median and third quarter. Here, short term facilities provide lowest LGDs and medium term loans the highest. According to the 75th percentile the facility type could cause a variation of 7.0 percentage points. The seniority of a loan determines to a large extent the median. Non senior loans might create a 59.0 percentage point higher loss in the median case compared to a pari-passu status. In the third quartile super senior loans might be responsible for a reduction of losses by 17.6 percentage points compared to pari-passu. Here, another advantage of Quantile Regression can be seen: OLS provides a variation of 8.3 + 18.2 = 26.5 percentage points due to seniority without distinction of the quantile. In contrast, QR shows a varying influence up to 66.5 percentage points.[6] We confirm that collateralization is an important factor for the recovery of defaulted loans. Real estate collateral decreases LGDs up to 50.1 percentage points. Other collateral types yield reductions of 27.6 percentage points. OLS underestimates the effects by more than 50 percent with values of 12.5 and 21.0 percentage points. The effect of additional collateral is small (4.7 percentage points) but still significant. In recent literature, the debtor's industry affiliation has been identified as a significant factor of mean LGDs. We observe industry effects mainly in the third quartile. The affiliation may cause a variation up to 40.8 percentage points with lowest LGDs for the mining industry and highest values for wholesale as well as retail trade and services. In contrast, the OLS estimates are misleading, because the behavior of the mining industry would not be identified as statistically different to FIRE. Regarding the macroeconomic variables we see that in the median case a decrease of the S&P 500,

[6] 7.5 (super senior) + 59.0 (non senior) = 66.5 percentage points at the 50th percentile.

i.e., a poor development on the stock market, significantly increases losses and a higher term spread also yields high losses. Again, OLS underestimates covariate influences. For the third quartile only the term spread effect is significant. For extreme quantiles we do not identify any macroeconomic effects.

We have seen that parameter estimates of the Quantile Regression vary over quantiles and that there are substantial differences to OLS estimates. In order to test the statistical distinction of parameter estimates, we perform tests of Koenker and Bassett (1982) with the estimated covariance matrix of Hendricks and Koenker (1992). The null hypothesis of the first test states that for no covariate of the models in Table 2.5 parameters differ over the 5th, 25th, 50th, 75th and 95th percentiles. The test rejects this hypothesis with an F value of 139 for 76 and 21,464 degrees of freedom resulting in an p-value of lower than $2.2 \cdot 10^{-16}$. This result shows the usefulness of Quantile Regression to allow for varying covariate influences. We repeat the test separated for each covariate to get a deeper insight. The constancy of parameters is rejected on the 1 % - level for all covariates expect the dummy variables for short term facilities and the industries MIN, SON and TCEGS. Since these tests focus on the 5th, 25th, 50th, 75th and 95th percentiles and some parameter effects can clearly be seen between those quantiles (see Appendix, Figure 2.A.1), we repeat the tests including all percentiles between 1 % and 99 %. Now the hypothesis of constant values is rejected for all parameters on the 1 % - level.

In summary, we conclude that there are statistically significant differences between quantile models. As seen in the simulation study in Section 2.2.2, this results in non-normal distributions and may cause bimodal LGDs. Quantile Regression extends results of standard regressions in a much more comprehensive view. We see that covariate impacts strongly vary over quantiles and this enables new insights on economic effects of control variables.

2.4.2 Model Comparison

In-sample

In this section, we compare the goodness of fit of the QR approach with the five alternative models. Hence, we analyze the fit of the entire distribution and do not restrict on the mean. The parameter estimates of the comparative methods are shown in Appendix 2.B.

Traditional LGD models and validation focus on mean predictions. For example, the literature (see Section 2.2.3, Table 2.3) uses the standard coefficient of determination and the mean absolute or root mean squared error. All these measures only focus on the mean. They do not fully capture the distributional behavior of LGDs. Therefore we report R^2-values only for the sake of completeness and evaluate the goodness of fit for the entire distribution using other measures.

First, we test the goodness of fit for the complete sample. We estimate all models with the entire dataset, i.e., in-sample. Table 2.6 shows low R^2 values between 0.077 and 0.101 because of the bimodal nature of LGDs.[7] Regression Trees are known for their good in-sample mean fit and provide the best R^2. QR provides a medium fit with 0.089. However, the introductory example of Section 2.1.2 shows the potential of misleading results when considering the mean.

Table 2.6: In-sample goodness of fit

	OLS	FRM	BR	RT	FMM	QR
R^2	0.0949	0.0883	0.0770	0.1008	0.0950	0.0893
HMI	0.1544	0.1372	0.1514	0.2017	0.0480	0.0030
HWMI	0.0168	0.0132	0.0169	0.0316	0.0019	$< 10^{-4}$
KS test	0.0883	0.0790	0.0741	0.0880	0.0549	0.0048
(p-value)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.9999)

Notes: The table shows measures for the goodness of fit for all models. A high value of R^2 and low values of HMI and HWMI indicate a good fit. High values for the KS test imply a bad fit. A low p-value leads to rejection of the null hypothesis of an adequate distributional fit. Models are estimated with the dataset of the entire time period.

As a counterpart to the coefficient of determination, we use an equivalent measure for

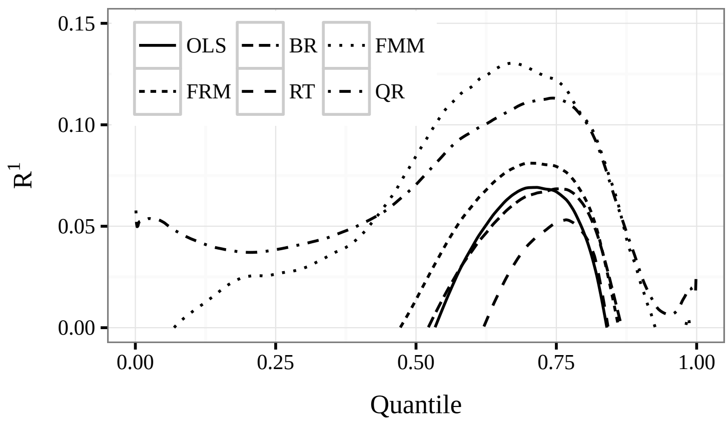
[7] Regardless of the method, the R^2 compares the predicted and the realized LGD, i.e., it is defined as the ratio of the variation explained by the model and the overall variation in LGD data.

the goodness of fit separated for each quantile:

$$\mathrm{R}^1(\tau) = 1 - \frac{\sum_{i=1}^{n} \rho_\tau(y_i - \widehat{\beta}(\tau)'x_i)}{\sum_{i=1}^{n} \rho_\tau(y_i - \widehat{Q}_y(\tau))}, \tag{2.5}$$

where ρ_τ is the asymmetric loss function defined in Equation (2.3) and $\widehat{Q}_y(\tau)$ is the τ-quantile of realizations $y_1, \ldots, y_n$. High values of R^1 indicate a good fit of the specific quantile. Figure 2.7 exhibits the values for all models over the entire range of the distribution and shows that the Quantile Regression and the Finite Mixture Model dominate other methods. The FMM is better suited to fit the middle part of the distribution between the 43th and the 80th percentile. The superiority of the QR is given in the left tail up to the 43th percentile and in the right tail for the last decile. Other methods even show negative values and thus a worse fit. This is a first indication for a better fit of the bimodality using QR.

Figure 2.7: In-sample goodness of fit (R^1)



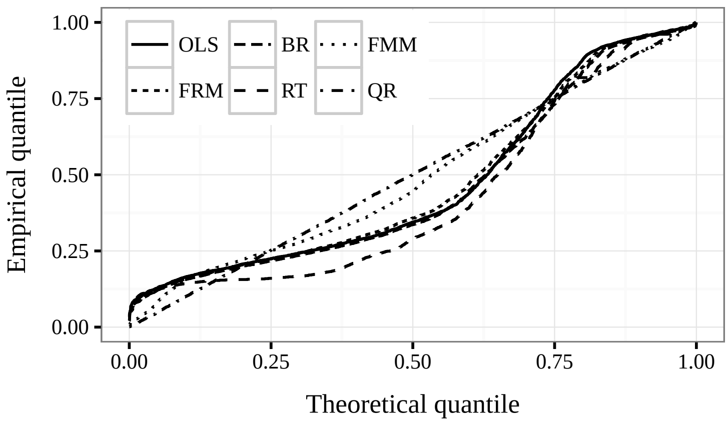
Notes: For each quantile the specific R^1 is given. Models are estimated with the dataset of the entire time period.

The probability–probability (P-P) plot of Michael (1983) is a graphical alternative that shows the accuracy level of the distributional fit. It compares theoretical and empirical quantiles by fitting the following two values against each other:

$$p_{1i} = \frac{i - 0.5}{n} \quad \text{and} \quad p_{2i} = \widehat{F}(y_i) \quad \text{for} \quad i = 1, \ldots, n, \tag{2.6}$$

where loans $i = 1, \ldots, n$ are ordered to ensure monotonic increasing quantiles $\widehat{F}(y_i)$. The lower the differences between both values for fixed loans the better is the fit. Figure 2.8 shows the P-P plot. The benchmark is given by the bisector line, where the theoretical quantile matches the empirical. All methods – except QR – show a systematic bias at the tails and in the middle. The in-sample fit of the Quantile Regression is almost perfect which is clear as the model is fitted for each quantile in-sample. In addition, we present three measures to summarize the P-P plot to one comparative number, which will be particularly relevant for the out-of-sample analysis.

Figure 2.8: In-sample goodness of fit (P-P plot)



Notes: The figure shows the P-P plot for the dataset of the entire time period.

Wagenvoort (2006) defines the Harmonic Mass Index (HMI) and the Harmonic Weighted Mass Index (HWMI) as the mean absolute and mean squared difference to the perfect P-P plot by

$$\text{HMI} = \frac{2}{n}\sum_{i=1}^{n}|p_{1i} - p_{2i}| \quad \text{and} \quad \text{HWMI} = \frac{2}{n}\sum_{i=1}^{n}(p_{1i} - p_{2i})^2. \tag{2.7}$$

Both measures are standardized to get values between zero and one, the lower the better. The HMI and HWMI (see Table 2.6) show the best results for QR with values of 0.0030 and less than 0.0001 compared to other methods with values of 0.0480 - 0.2017 and 0.0019 - 0.0316. Again, this is due to the in-sample fit for every quantile and in-sample validation.

The Kolmogorov-Smirnov (KS) test summarizes the P-P plot by evaluating the maximum difference to the perfect case. The null hypothesis claims that the data follow the estimated distribution of the regression method. The test statistic D and the critical value c for a level α are given by

$$D = \frac{1}{2n} + \max_{i=1,\ldots,n} |p_{1i} - p_{2i}| \quad \text{and} \quad c = \sqrt{\frac{\ln 2 - \ln \alpha}{2n}} \quad \text{for} \quad n > 35. \tag{2.8}$$

The KS test validates the maximum deviance but not the overall goodness of fit. Results should be treated with caution. In-sample, the KS test rejects the null hypothesis for all models except QR (see Table 2.6), which supports this method.

In summary, Quantile Regression leads to a moderate in-sample fit of the mean. However, considering the entire distributional LGD behavior the in-sample goodness of fit shows a superiority of this method as every quantile is exactly fitted.

Out-of-sample

The partition of the dataset into a subsample for estimation and a subsample for the out-of-sample validation is a modified version of Altman and Kalotay (2014). In a first step, we estimate all models for the data subsample from 2000 until 2009. This corresponds to approximately 60 % of the entire dataset. Afterwards, we randomly draw 10,000 times a subportfolio out of the remaining data from 2010 until 2013. Each step consists of 300 defaulted loans which is approximately the average number of defaults per year in our dataset. We evaluate the goodness of fit for each subportfolio out-of-sample (and out-of-time) and report the averages over all 10,000 steps in Table 2.7.

Table 2.7: Out-of-sample goodness of fit

	OLS	FRM	BR	RT	FMM	QR
R^2	0.0313	0.0451	0.0503	0.0196	0.0421	0.0344
HMI	0.1798	0.1672	0.1736	0.3933	0.1426	0.0955
HWMI	0.0217	0.0191	0.0210	0.1034	0.0182	0.0065
KS test	0.1673	0.1606	0.1538	0.4226	0.0740	0.0989
(p-value)	(0.0001)	(0.0001)	(0.0002)	(0.0001)	(0.1887)	(0.0417)

Notes: The table shows measures for the goodness of fit for all models. A high value of R^2 and low values of HMI and HWMI indicate a good fit. Models are estimated on the subsample 2000 - 2009. We report the mean goodness of fit measures for 10,000 subsamples of 300 randomly chosen loans which defaulted in 2010 - 2013.

Again, QR provides a medium fit of the mean, now with an R^2 of 0.0344 compared to other methods with values between 0.0196 and 0.0503. The mean absolute as well as the mean squared deviations to the perfect P-P plot (HMI and HWMI) are best (lowest) for the Quantile Regression with values of 0.0955 and 0.0065 compared to 0.1426 - 0.3993 and 0.0182 - 0.1034 for the other measures. In 99.5 % of all simulation steps the HMI was best for QR and in 99.8 % the HMWI, which shows a strong out-of-sample superiority.

The KS test tests the maximum deviance to the perfect P-P plot. Here, all methods except the Finite Mixture Model and Quantile Regression do not perform well with average p-values lower than 1 %. The KS test identifies a small benefit of avoiding large differences for the FMM.[8]

2.5 Implications

2.5.1 Loss Distribution

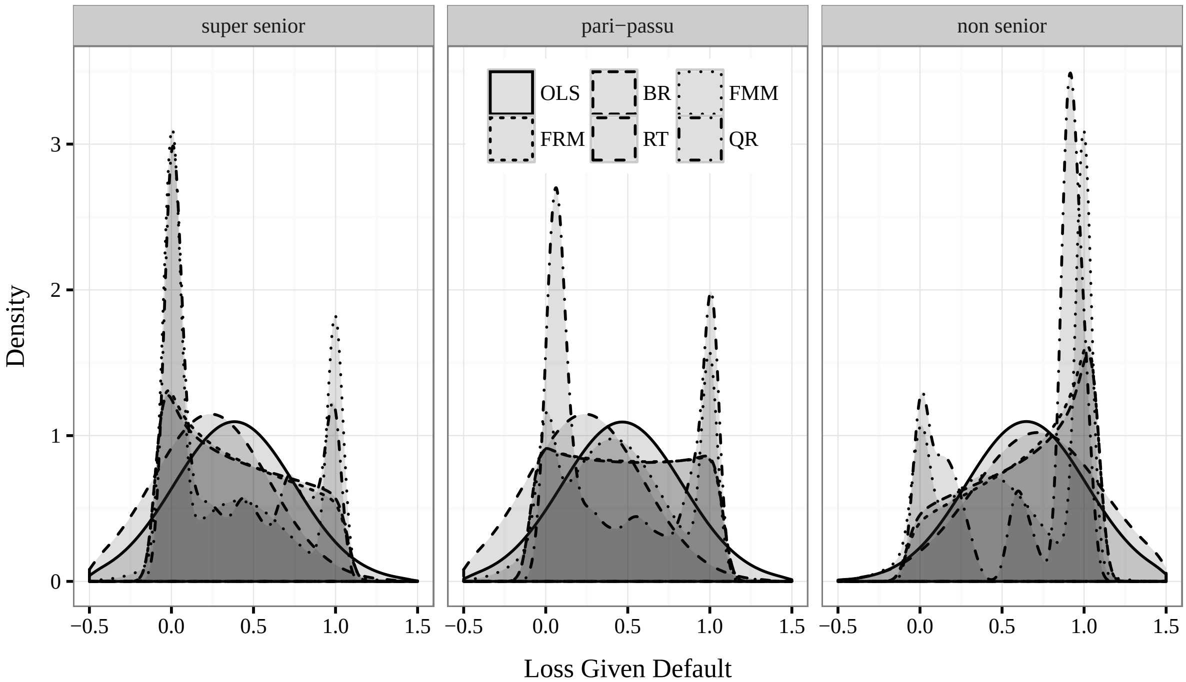
For the internal ratings-based (IRB) approach according to the Basel regulatory framework a financial institution is allowed to use internal probability of default (PD) and LGD models to calculate regulatory capital. For the latter it has to use so called Downturn LGD estimates, which reflect economic downturns (see Basel Committee on Banking Supervision (2005)). The Board of Governors of the Federal Reserve System (2006) proposes a simple formula based on the mean: $0.08 + 0.92 \cdot E(LGD)$. However, Calabrese (2014) and Rösch and Scheule (2014) model systematic unobservable risk factors and suggest using their quantiles to determine a Downturn LGD, similarly to the one-factor PD model. Here, we propose an alternative measure based on quantiles of LGDs in a direct way.

We start with an analysis of the LGD probability density functions in Figure 2.9 which we estimate for each model. Because the specific choice of covariates influences the distribution of a loan LGD, we compute the densities for three exemplary loans

[8] For the FMM, the null hypothesis is rejected in 55.7 %, 41.4 % and 18.8 % of all bootstrap steps for significance levels of 10 %, 5 % and 1 %. The rejection rates for the QR are 89.0 %, 81.6 % and 57.3 %.

which only differ in the seniority in order to allow visualization. Other covariates are chosen from representative cases. All three loans are medium term facilities, have neither any guarantee nor collateral and are from entities with FIRE industry affiliation. The macroeconomic information is chosen as median scenario with a year-on-year growth of 11 % for the S&P 500 and a term spread of 2.63 percentage points.

Figure 2.9: Densities of Loss Rates Given Default

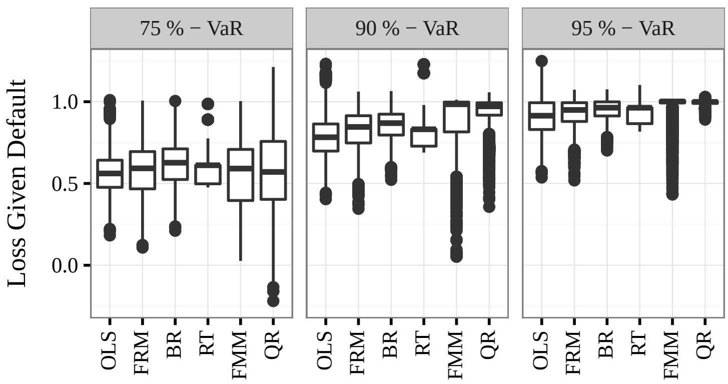


Notes: The figure shows density predictions for representative loans of different seniorities. Models are estimated with the dataset of the entire time period.

As can be seen, standard OLS techniques are not able to capture the non-normality of LGDs, even after transformation (FRM). Both methods predict too much probability mass in the middle and on the tails but too little on total losses. Beta Regression exhibits these issues to a smaller extent, but is also not able to capture the realized extreme bimodal shape. Regression Trees produce results similar to OLS because of the assumption of normality per group. The Finite Mixture Model provides a flexible density shape. It captures bimodality and predicts extreme values around zero and one. A flexible weighting of the three components can be seen, i.e., the left, middle and right component. However, there is too much weight on rare negative values and in the middle. Quantile Regression seems to deliver the most individual forecasts. Flexible shapes in the tails and especially in the middle are the result. The minimum and maximum realizations imply a plausible wide range of density estimates. The tails around zero and one are well fitted, very extreme values are rare and the middle is adequate compared to the tails. It can also be seen that the seniority has an impact on losses. The left peak, i.e., total recovery is most pronounced for the best seniority (super senior), while the FMM overestimates the chance of total losses. In contrast, for loans with pari-passu status the peaks are more balanced for QR and FMM, but the last mentioned method underestimates the chance of total recoveries (zero losses) and total losses. Finally, non senior loans imply the highest probability of total losses which results in a strong right peak.

For the Quantile Regression the Value at Risk (VaR) can be simply estimated over the corresponding quantile. We predict the VaR for each loan in our dataset for three exemplary quantiles (75 %, 90 %, 95 %). Figure 2.10 shows the specific box plots for each method and each quantile. The extreme bimodality of loan losses increases tail risk measures but the means do not change, because the peaks of total losses and recoveries increase to the same extent (see example in Section 2.1.2). The 75 % - VaR is widely spread over the sample. QR differentiates most with an average of 57.3 %. Other methods show a comparable average of 75 % - VaR estimates but with a lower deviation.

Figure 2.10: Predictions of the Value at Risk (VaR)



Notes: The figure shows various VaR estimates for all investigated methods. Models are estimated with the dataset of the entire time period.

Finally, we analyze the fit of VaR estimates. The VaR for a level α is expected to be exceeded by the realized loss in $(1-\alpha)\cdot 100\,\%$ of all cases. The closer the difference is at zero, the better is the fit. Table 2.8 shows the performance of predictions by their realized hit rate. QR results in a hit rate of 24.79 % for the 75 % - VaR. Thus, the difference between the theoretical and the empirical hit rate is only 0.8 % which is the lowest for all methods under consideration. For the 90 % - level the deviation is lowest for the FMM and QR. In the extreme tail of the 95 % - VaR we can see a further superiority of the Quantile Regression with the lowest deviation of 1.2 %, followed by the FMM with 23.4 %. Note that OLS results in a severe underestimation of the 95 % - VaR with a hit rate of more than 10 %.

Table 2.8: Hit rates of VaR based Downturn LGD

	OLS	FRM	BR	RT	FMM	QR
75 % - VaR	0.2614	0.2519	0.2423	0.2481	0.253	0.2479
	(4.6)	(0.8)	(-3.1)	(-0.8)	(1.2)	(-0.8)
90 % - VaR	0.1864	0.1755	0.1685	0.1500	0.1005	0.1026
	(86.4)	(75.5)	(68.5)	(50.0)	(0.5)	(2.6)
95 % - VaR	0.1028	0.1021	0.0915	0.0994	0.0383	0.0506
	(105.6)	(104.2)	(83.0)	(98.8)	(-23.4)	(1.2)

Notes: The table shows hit rates for the VaR, i.e., how often estimated values for the VaR are exceeded by real data. Models are estimated with the dataset of the entire time period. For the VaR to an α - level the expected hit rate is $1 - \alpha$. The realized percentage difference to this value is given in parentheses.

2.5.2 Downturn LGDs

The implementation of the internal ratings-based approach requires to use LGD estimates that reflect economic downturn conditions to calculate the required regulatory capital (see Basel Committee on Banking Supervision (2005)). In this section, we discuss to which extent macroeconomic covariates capture downturn risk and propose downturn LGD measures.

So far, we used the return of the S&P 500 and the treasury term spread as macroeconomic indicators. Here, we will present results for two alternative and very popular systematic covariates that reflect system-wide uncertainty: the CBOE volatility index VIX and the TED spread between the three-month LIBOR and the US treasury-bill rate. In addition, we consider a recession dummy that captures possible downturn effects that can not be covered by metric covariates. This downturn dummy takes the value one if there is a recession in the quarter of default or in the following two quarters. A recession is given by the definition of the National Bureau of Economic Research (NBER). We identify the chosen indicator as the best choice to capture the role of a downturn during the resolution process of a defaulted loan.

In the Appendix, Table 2.C.1 shows regression results for different combinations of macroeconomic variables (S&P 500, TED spread, term spread, VIX, downturn dummy) which we will summarize here. A high TED spread or VIX imply a high uncertainty and reasonably increases the LGD. However, the improvement of the explanatory power is limited and in connection with other macroeconomic indicators the parameter estimates of the TED spread becomes insignificant. The VIX even shows a significantly implausible sign. The S&P 500 and the term spread result in plausible as well as statistically significant parameter estimates. The results for both covariates are stable in combination with other macroeconomic indicators.

We also tested different lead and lag structures of systematic covariates. The presented economic conditions in the year after default seem to have the most explanatory power. If we consider the economy in the quarter prior to default and the three following quarters, the explanatory power decreases. Using the economic conditions of the year prior to default further deteriorates the results.[9] This may be caused by the fact that economic indicators often show neutral or positive conditions prior to clustered defaults. In contrast, the loan recovery depends mainly on the resolution process after default.

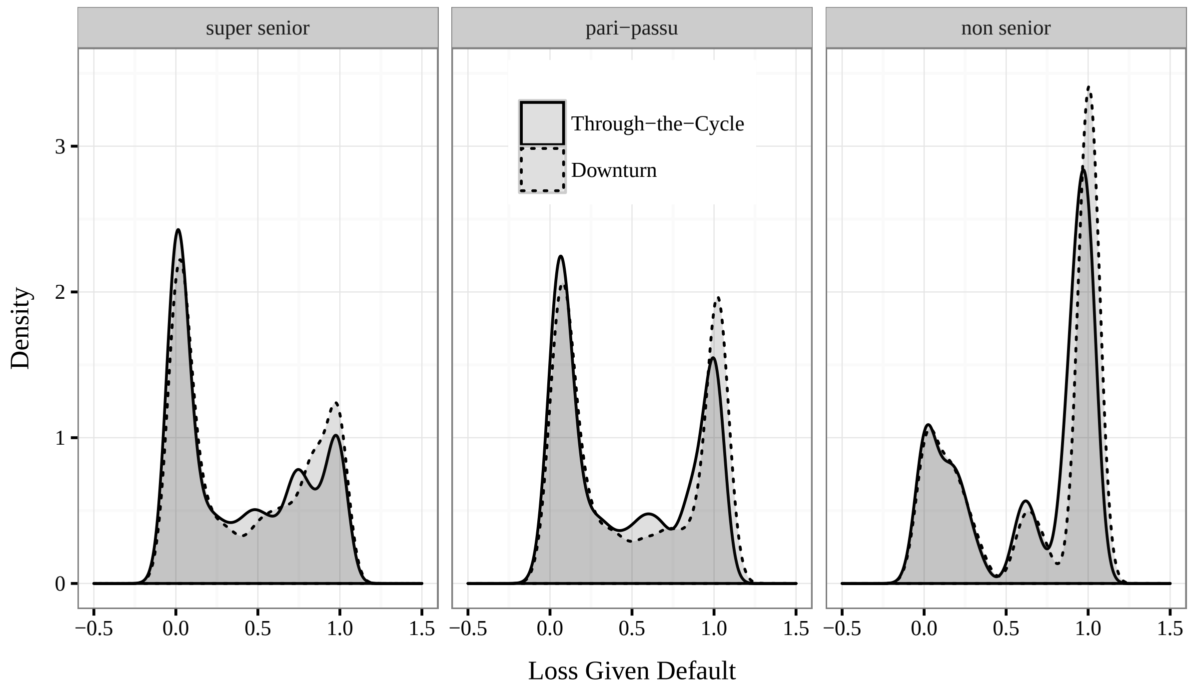
In order to calculate LGDs that reflect economic downturn conditions, one may set the considered macroeconomic variables to specific downturn values. Due to plausibility, significance and explanatory power, we use the model including the return of the S&P 500, the term spread and the recession dummy for downturn estimates (model (8) of Appendix Table 2.C.1). For our data one may consider the realized values of the macroeconomic variables during the Global Financial Crises. This results in a 40 % decrease of the S&P 500, a term spread of 3.6 percentage points and an indicated downturn by the

[9]The R^2 for Quantile Regression subsequently decreases from 0.0910 to 0.0890 and 0.0864 using the four considered macroeconomic variables and the downturn dummy.

recession dummy. In addition, we consider a model without macroeconomic information to get average LGD estimates during the cycle, known as Through-the-Cycle (TTC) approach (model (1) of Appendix Table 2.C.1).

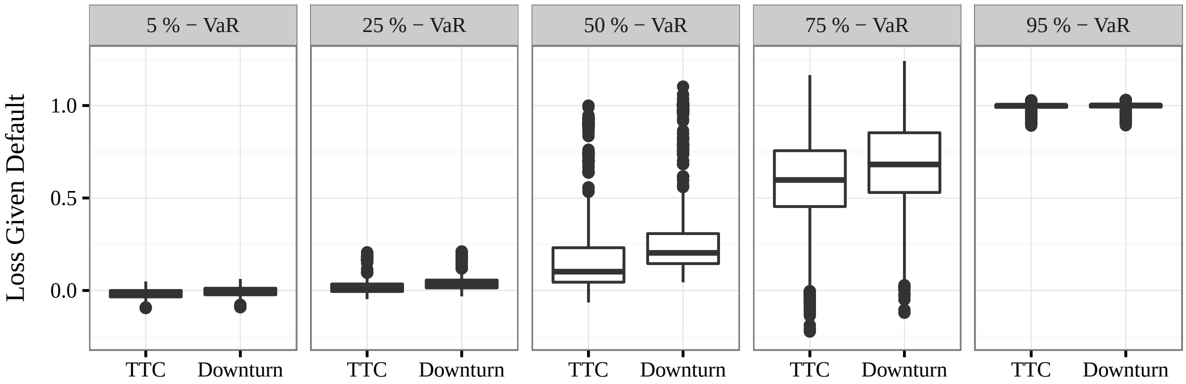
Reconsidering the three exemplary loans of Figure 2.9, we predict the resulting densities using the QR estimates through the economic cycle and during a downturn (Figure 2.11). There are diverse effects over quantiles. In a downturn the chance of total recoveries decreases, median losses shifts to higher and the chance of total losses increases. We generalize the results of the three exemplary loans to the entire dataset by calculating VaR estimates for the TTC approach and during a downturn. Figure 2.12 shows a shift to higher losses for several quantiles. Lower quantiles are slightly affected. The effect on the median and the third quarter are most significant. The Quantile Regression adequately models quantile-specific downturn effects.

Figure 2.11: Densities of Loss Rates Given Default during the economic cycle



Notes: The figure shows density predictions for representative loans of different seniorities. Models are estimated with the dataset of the entire time period. The Through-the-Cycle model does not contain macroeconomic information and is similar to an average through the economic cycle. In addition, we consider a downturn scenario by including the S&P 500, the term spread and a downturn dummy.

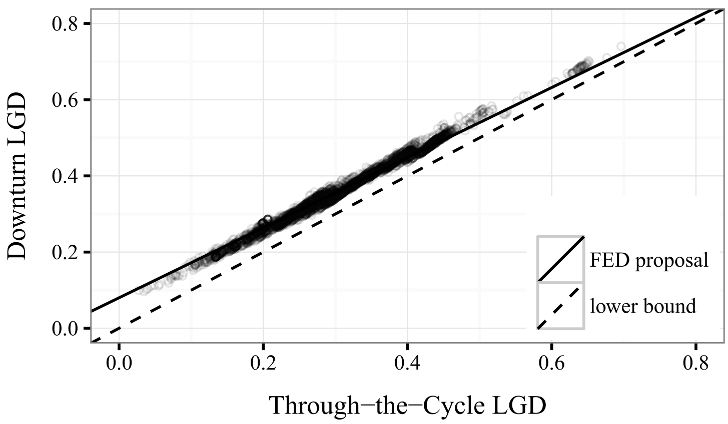
Figure 2.12: Downturn effects on LGD quantiles



Notes: The figure shows various QR estimates of quantiles for a Through-the-Cycle (TTC) model and a downturn scenario. Models are estimated with the dataset of the entire time period.

Finally, we propose Downturn LGD estimates for the internal ratings-based approach by mean predictions of the downturn densities. Figure 2.13 compares these estimates with their corresponding mean through the economic cycle. The Board of Governors of the Federal Reserve System (2006) proposal $0.08 + 0.92 \cdot E(LGD|TTC)$ is indicated by the solid line and ignores covariate information. In contrast, the downturn estimates of Quantile Regression consider loan characteristics due to quantile-specific downturn effects. The Board of Governors of the Federal Reserve System (2006) proposal results on average in higher values for low LGDs and lower values for high LGDs.

Figure 2.13: Predictions of Quantile Regression based Downturn LGDs



Notes: The figure compares the expected LGD through the economic cycle and in a downturn scenario. The solid lines corresponds to the FED proposal $0.08 + 0.92 \cdot E(LGD|TTC)$ by the Board of Governors of the Federal Reserve System (2006).

In summary, Downturn LGDs defined by transformations of mean predictions do not capture unexpected risk as well as loan characteristics and quantile-specific downturn effects. The bimodality of LGDs shows the need of VaR based downturn measures. Quantile Regression models the distributional behavior of LGDs more individual than standard models, in particular during downturns.

2.6 Conclusion

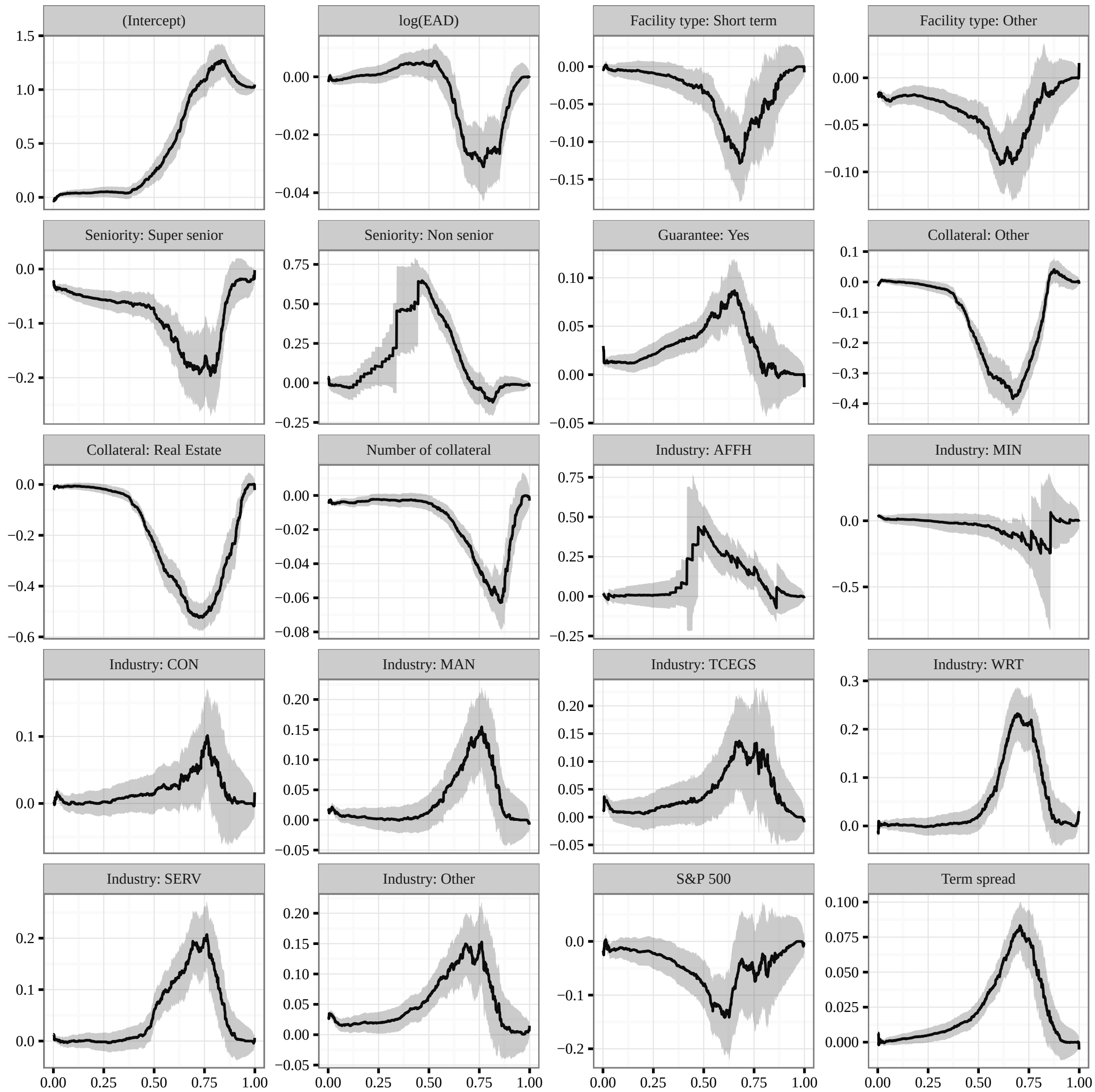
Recent loss models mainly focus on mean predictions. This is important because even banking authorities propose the computation of Downturn LGDs as a linear transformation of expected values. However, loan loss data show a strong variation and are extremely skewed and bimodal. We find that the random nature even remains after controlling for observable risk factors. Most existing models and the validation of mean predictions do not account for this behavior.

This paper proposes an alternative LGD model based on Quantile Regression. It allows covariate effects to vary over the entire distributional range of LGDs and reflects the strong probabilistic nature of losses. Our results indicate a high sensitivity of low LGDs to loans with super senior status, guarantees and the number of collateral. In the median case, LGDs are additionally affected by the kind of collateral, facility type as well as industry affiliation and the macroeconomic environment. High LGDs appear not to be influenced by any of our covariates. This indicates a remaining tail risk of a total loss which can not be controlled by bank practices or regulation.

We present alternative goodness of fit measures for the validation of the random behavior of LGDs. Most comparative methods fail to model the loss distribution. An in-sample and out-of-sample analysis shows the superiority of the Quantile Regression over other credit risk models. Macroeconomic effects are identified to vary over LGD quantiles. Thus, Quantile Regression based Downturn LGDs are more accurate and able to capture loan- and quantile-specific downturn effects in contrast to standard approaches.

Appendix 2.A Quantile Regression

Figure 2.A.1: Parameter estimates Quantile Regression



Notes: Each subfigure shows the corresponding quantile parameter estimates for one covariate. The grey area indicates a confidence interval to the 95 % - level. Standard errors are estimated by kernel estimates. Section 2.4.1 Table 2.5 shows the numerical results for the 5th, 25th, 50th, 75th and 95th percentiles.

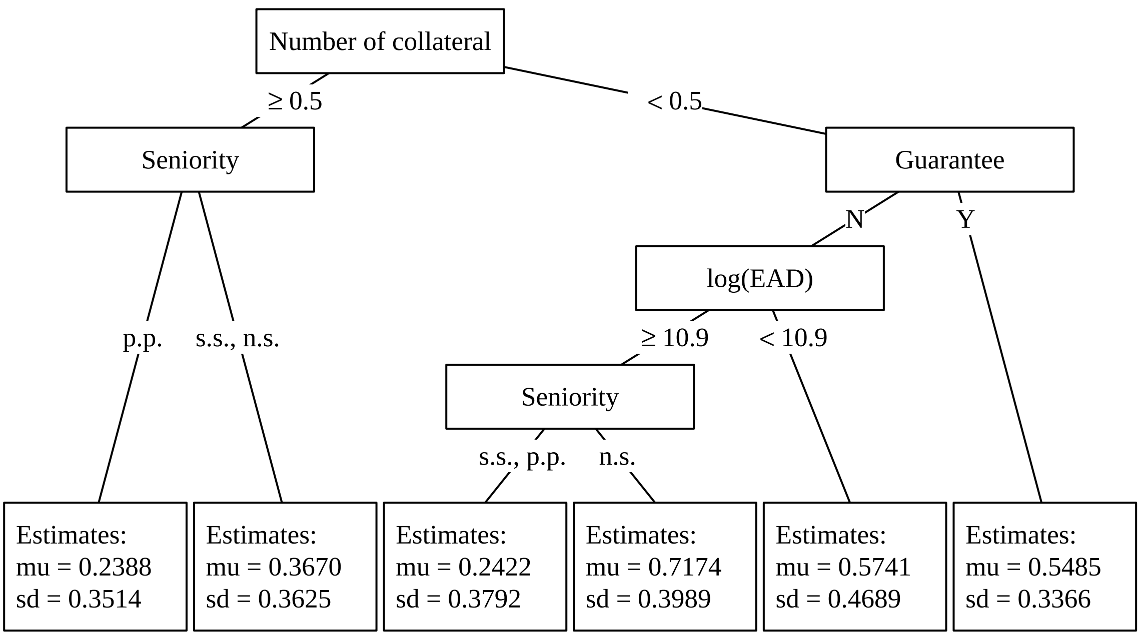
Appendix 2.B Comparative Methods

In this Appendix, we present the parameter estimates of comparative methods that are introduced in Section 2.2.3. Estimates are given for transformed data where modification was necessary, i.e., for the Fractional Response Model and Beta Regression. Table 2.B.1 shows the corresponding regression results. In summary, all methods detect the covariate effects similar to the Quantile Regression with only minor differences in significance.

We tested several normal mixture distributions with a varying number of components similar to Altman and Kalotay (2014). Likelihood based selection criteria show a superiority of three components. The corresponding estimated means with standard errors in parentheses are given by 0.4077 (0.0099), 0.9777 (0.0008) and 0.0037 (0.0008); and the estimated variances by 0.0863 (0.0035), 0.0001 ($< 10^{-4}$) and 0.0007 ($< 10^{-4}$).

Figure 2.B.1 shows the resulting Regression Tree. Other parameterizations lead to a higher goodness of fit for the mean, but impair the overall fit of the entire distribution. Thus, we choose a tree with less nodes but a better fit for the distributional LGD behavior to allow a fair comparison (cf. Bastos (2010)).

Figure 2.B.1: Estimated Regression Tree



Notes: This is our final estimated tree for comparative reasons. Seniority codes are given by super senior (s.s.), pari-passu (p.p) and non senior (n.s).

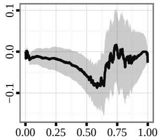
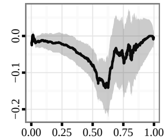
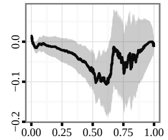
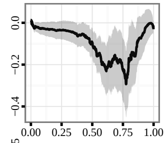
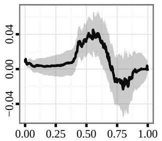
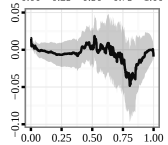
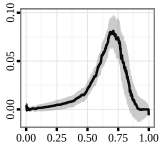
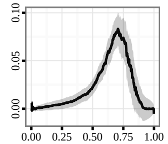
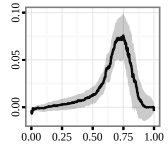
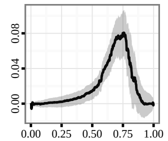
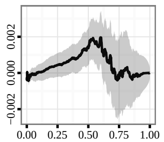
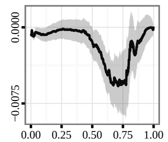
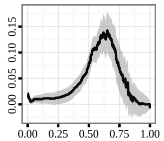
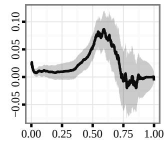
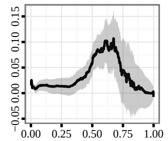
Table 2.B.1: Parameter estimates comparative methods

		OLS	FRM	BR	FMM (1)	FMM (2)
(Intercept)		0.5189 ***	0.1467	0.1776	-2.2914 ***	1.8167 ***
		(0.0493)	(0.1404)	(0.1514)	(0.4086)	(0.4518)
log(EAD)		-0.0103 ***	-0.0301 ***	-0.0313 ***	0.2479 ***	-0.1329 ***
		(0.0032)	(0.0092)	(0.0099)	(0.0260)	(0.0307)
Facility type	Short term	-0.0713 ***	-0.2160 ***	-0.2067 ***	-0.5223 ***	-0.8134 ***
(Medium term)		(0.0178)	(0.0507)	(0.0547)	(0.1259)	(0.1737)
	Other	-0.0618 ***	-0.2136 ***	-0.2123 ***	-0.4538 ***	-0.5571 ***
		(0.0135)	(0.0384)	(0.0415)	(0.0986)	(0.1255)
Seniority	Super senior	-0.0829 ***	-0.2877 ***	-0.2871 ***	-1.8925 ***	-1.1275 ***
(Pari-passu)		(0.0183)	(0.0522)	(0.0564)	(0.2076)	(0.2336)
	Non senior	0.1818 ***	0.5150 ***	0.5088 ***	-0.3036	0.7218
		(0.0533)	(0.1518)	(0.1639)	(0.5102)	(0.5074)
Guarantee	Y	0.0574 ***	0.1780 ***	0.1777 ***	0.7434 ***	0.2617 **
(N)		(0.0123)	(0.0349)	(0.0377)	(0.0909)	(0.1209)
Collateral	Other	-0.1252 ***	-0.3534 ***	-0.3478 ***	-0.6111 ***	-0.8619 ***
(N)		(0.0161)	(0.0459)	(0.0496)	(0.1200)	(0.1540)
	Real estate	-0.2098 ***	-0.6074 ***	-0.5719 ***	-0.7763 ***	-1.7850 ***
		(0.0199)	(0.0566)	(0.0612)	(0.1432)	(0.2369)
Number of collateral		-0.0188 ***	-0.0586 ***	-0.0543 ***	-0.0804 **	-0.2573 ***
		(0.0044)	(0.0125)	(0.0134)	(0.0318)	(0.0620)
Industry	AFFH	0.1674 ***	0.4067 **	0.3858 **	0.6026	-0.2195
(FIRE)		(0.0629)	(0.1792)	(0.1934)	(0.4472)	(0.7292)
	MIN	-0.0747	-0.1937	-0.1614	-1.0910	-0.5905
		(0.0815)	(0.2319)	(0.2495)	(0.6466)	(0.8389)
	CON	0.0335	0.0842	0.0863	0.1225	0.2035
		(0.0226)	(0.0644)	(0.0694)	(0.1603)	(0.2306)
	MAN	0.0471 **	0.1321 **	0.1324 **	-0.0170	0.2193
		(0.0212)	(0.0605)	(0.0652)	(0.1526)	(0.2053)
	TCEGS	0.0608 *	0.1988 **	0.2000 **	0.2876	0.8787 ***
		(0.0317)	(0.0902)	(0.0973)	(0.2429)	(0.2937)
	WRT	0.0734 ***	0.2168 ***	0.2046 ***	0.0159	0.6622 ***
		(0.0212)	(0.0605)	(0.0652)	(0.1530)	(0.2008)
	SERV	0.0744 ***	0.2069 ***	0.2038 ***	0.0047	0.5877 ***
		(0.0204)	(0.0582)	(0.0627)	(0.1500)	(0.1926)
	Other	0.0674 ***	0.2273 ***	0.2284 ***	0.6282 ***	0.7181 ***
		(0.0199)	(0.0565)	(0.0609)	(0.1529)	(0.2029)
S&P 500		-0.0763 **	-0.2320 ***	-0.2269 **	-1.3546 ***	-0.4486
		(0.0304)	(0.0864)	(0.0933)	(0.2246)	(0.3008)
Term spread		0.0331 ***	0.0891 ***	0.0833 ***	0.4799 ***	0.1653 ***
		(0.0060)	(0.0172)	(0.0186)	(0.0484)	(0.0567)

Notes: The table shows parameter estimates for comparative methods. Standard errors are given in parentheses. Significance is indicated by '*' (10 %), '**' (5 %) and '***' (1 %). For Beta Regression (BR) the precision parameter ϕ is estimated by 1.8536 (0.0344). FMM (1) describes the probit model for component 1, whereas FMM (2) is used for component 2.

Appendix 2.C Macroeconomic Variables

Table 2.C.1: Effects of macroeconomic information in the first year of resolution, i.e. after default

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
S&P 500									
TED spread									
Term spread									
VIX									
Downturn									

Continued on next page

Table 2.C.1 (continued): Effects of macroeconomic information in the first year of resolution, i.e. after default

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
R^2	0.0831	0.0837	0.0834	0.0881	0.0835	0.0870	0.0893	0.0898	0.0910
$R^1(0.05)$	0.0512	0.0519	0.0519	0.0513	0.0512	0.0524	0.0520	0.0528	0.0536
$R^1(0.25)$	0.0367	0.0373	0.0369	0.0376	0.0373	0.0383	0.0385	0.0390	0.0393
$R^1(0.50)$	0.0663	0.0674	0.0674	0.0691	0.0670	0.0716	0.0705	0.0727	0.0731
$R^1(0.75)$	0.1061	0.1061	0.1063	0.1126	0.1061	0.1084	0.1129	0.1129	0.1147
$R^1(0.95)$	0.0064	0.0065	0.0064	0.0064	0.0064	0.0064	0.0065	0.0065	0.0067
HMI	0.0102	0.0052	0.0108	0.0087	0.0069	0.0038	0.0030	0.0034	0.0027
HWMI·100	0.0165	0.0030	0.0190	0.0107	0.0052	0.0014	0.0007	0.0009	0.0006
KS test	0.0305	0.0120	0.0331	0.0263	0.0170	0.0083	0.0048	0.0062	0.0045
(p-value)	(0.0007)	(0.5835)	(0.0002)	(0.0053)	(0.1677)	(0.9999)	(0.9999)	(0.9999)	(0.9999)

Notes: The table shows several model estimated by the QR. Each subfigure shows the corresponding quantile parameter estimates for one covariate. The grey area indicates a confidence interval to the 95 % - level. Standard errors are estimated by kernel estimates. Each model includes the non-macroeconomic risk factors of Table 2.5, which we do not present here for clarity and due to an observed stability of the parameter estimates.

Chapter 3

Macroeconomic Effects and Frailties in the Resolution of Non-Performing Loans

This chapter is joint work with Jennifer Betz[1], Ralf Kellner[2] and Daniel Rösch[3], and corresponds to a working paper with the same name.

Abstract: Resolution of non-performing loans is a key determinant of bank credit default losses. This paper analyzes macroeconomic and systematic frailty effects of the default resolution time for a sample of 17,395 defaulted bank loans in USA, Great Britain, and Canada. We find that frailties have a huge impact on the resolution times. In a representative sample portfolio, median resolution times more than double in a recession when compared to an expansion. This leads to highly skewed distributions of losses and considerable systematic risk of the bank portfolio.

JEL classification: C23; G21; G33

Keywords: bank loans; credit risk; default resolution time; latent factors; systematic effects

[1]Chair of Statistics and Risk Management, Faculty of Business, Economics, and Business Information Systems, University of Regensburg, 93040 Regensburg, Germany. E-Mail: jennifer.betz@ur.de

[2]Chair of Statistics and Risk Management, Faculty of Business, Economics, and Business Information Systems, University of Regensburg, 93040 Regensburg, Germany. E-Mail: ralf.kellner@ur.de

[3]Chair of Statistics and Risk Management, Faculty of Business, Economics, and Business Information Systems, University of Regensburg, 93040 Regensburg, Germany. E-Mail: daniel.roesch@ur.de

3.1 Introduction

The default resolution time (DRT) of non-performing loans[4] is an important quantity owing to several reasons. Firstly, longer resolution processes are empirically related to higher losses (see Dermine and Neto de Carvalho, 2006; Gürtler and Hibbeln, 2013).[5] This effect is driven by higher discounting effects of later post-default payments and a negative relationship between the length of DRT and the sum of non-discounted recovery cash flows. While banks can compensate single outliers with long DRTs, systematic co-movements among DRTs might heavily increase the risk of credit portfolios if the above effects simultaneously occur for a multitude of non-performing loans in a portfolio. Secondly, high DRTs will burden the liquidity of financial institutions due to supplemental funding needs emerging from future legal requirements. Non-performing loans increase the required stable funding by definition and, thus, charge institutions additional burden to fulfill the Net Stable Funding Ratio. This paper emphasizes to take into account systematic effects on DRTs for predicting the reduction of clustered non-performing loans during downturns, which is relevant for credit portfolio risk and future liquidity management and regulation.

In the previous literature, most findings regarding DRTs stem from analyses which examine different workout schemes. Helwege (1999) analyzes the length of time a junk bond spends in default during the 1980s. He finds that the workout procedure as well as the bargaining power of firms are main drivers for quick resolutions. Even though the DRT is often assumed to vary with respect to the workout process, Bris et al. (2006) find no significant differences between the time of Chapter 7 liquidations and Chapter 11

[4]In this paper, we use a database that includes loans in default which is defined as “unlikely to pay” or “past due more than 90 days on any material credit obligation”. The loans can either be resolved informally or through the usage of a formal process and they can be reorganized/restructured or resolved by means of bankruptcy or insolvency, respectively. Throughout the paper, the terms non-performing loans, defaulted loans and loans in financial distress are used as synonyms.

[5]Reasons for this might be diverse. Previous literature mostly holds increasing costs stemming from higher liquidity and interest rate risks in combination with higher discounting effects accountable. In addition, Gürtler and Hibbeln (2013) find that loans which return back to performance after default usually cause lower losses. At the same these loans are typically the ones which can be resolved quickly. We also observe a negative relationship between default resolution times and recovery rates for our data set (see Section 3.2 for more details).

reorganizations. Bandopadhyaya (1994) uses a hazard rate model and examines the time spent until a firm exits Chapter 11 protection. He finds that firms spend less time under Chapter 11 if they have high interest amounts outstanding and high capacity utilization. Moreover, he includes two macroeconomic variables (interest on short term loans and rate of growth of the gross national product) which do not exhibit a significant impact on the time spent under Chapter 11. Further contributions which, among other issues, analyze the time spent under Chapter 11 filings are given by Partington et al. (2001), Wong et al. (2007), and Denis and Rodgers (2007). They find firm size and pre-default performance to be important drivers for the DRT. The authors also incorporate industry specific as well as two macroeconomic variables (credit and term spreads), and detect significant influences. Most of these papers use techniques from survival time analysis which seem to be natural choices for DRTs.

However, the common systematic behavior of DRTs is rarely analyzed in the literature. Few studies which consider systematic effects, e.g., in the form of macroeconomic variables, lead to diverging conclusions regarding their impact. The aim of the paper is to close this gap and examine common components in DRTs of non-performing loans. A profound understanding of systematic effects is crucial as co-movements among DRTs originate from joint determinants. Hence, DRTs are collectively higher or lower during certain time periods and, thus, might exert additional pressure in downturn periods. Higher DRTs accompany with higher losses. The systematic behavior is, thus, transfered to the recovery side. As the presence of non-performing loans entails further funding needs in the future, high DRTs maintain the increased liquidity burden on firm level. This depresses lending which might summit in a credit crunch if a majority of financial institutions is affected. We analyze a data base of 17,395 non-performing loans in the US, Great Britain, and Canada to deeply examine observable and unobservable systematic effects among DRTs. We find that these common factors determine DRTs and demonstrate their inference on the DRT itself and the loss involved.

A common approach in recent literature to incorporate systematic effects in risk models is to include macroeconomic variables which impact all debtors at the same time.

While these variables might capture parts of the co-movement in DRTs, they might not be enough to model unobservable stochastic shocks. Hence, we use continuous-time hazard rate models with a stochastic frailty to include unobservable systematic effects besides observable variables. These models have been successfully used for estimating *default times*, i.e., the *time up to the default* of a bond or loan. Das et al. (2007) focus on the doubly-stochastic assumption for default time models and find that defaults are clustered to a higher extent than expected. This might be due to the impact of unobservable variables. Duffie et al. (2009) show that latent factors have a significant impact on the default time even if macroeconomic and firm specific determinants are included in the model. Thus, neglecting unobservable variables results in downward biased assessments of credit portfolio risk. Applying a frailty approach for a credit risk model incorporating market risk, Kuo and Lee (2007) underline the potential downward bias of risk assessment when ignoring dependencies between market and credit risk. Koopman et al. (2011) use a default rate model including a dynamic latent (frailty) factor. They show that its impact does not vanish even if the model already incorporates a large amount of macro-financial covariates. Overall, their model shows improvements in out-of-sample forecasts in comparison to models not allowing for unobservable covariates. In addition, Lee and Poon (2014) consider global, parental-sector and sector specific factors in a portfolio loss model. Their results show that also sector specific frailty effects determine loan defaults and that their impact on the aggregated portfolio loss is greater compared to macroeconomic variables.

Our paper provides the following contributions. First, we thoroughly examine systematic effects among DRTs of non-performing loans. To the best of our knowledge, we are the first to extend the modeling scheme of DRTs to unobservable systematic factors using a doubly-stochastic continuous-time hazard rate framework, similar to those used in default time modeling. Second, we empirically measure these unobservable (frailty) factors for a unique and comprehensive data set. Our results show that DRTs are significantly driven by common unobservable factors even after controlling for individual specific and macroeconomic variables. This leads to collectively higher DRTs in downturn periods.

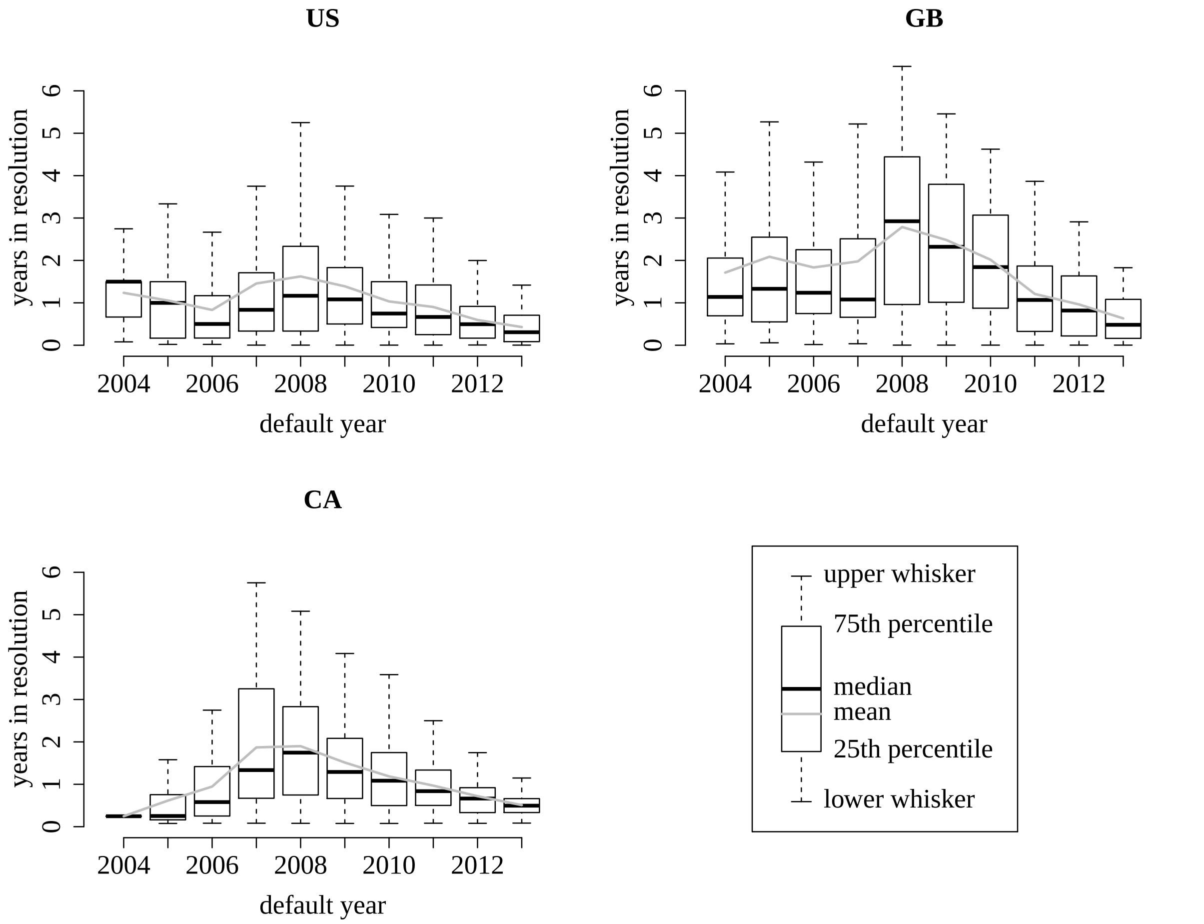
The average DRT of an exemplary portfolio consisting of non-performing loans increases from 1.59 to 2.42 years. Third, we evaluate potential effects of clustered DRTs on the loss of an exemplary portfolio consisting of non-performing loans. Long resolution processes are empirically related to higher losses. This might be due to ascending direct and indirect costs. Direct outlays, such as legal or liquidation expenses, increase either as they are charged during longer time periods or as these costs are higher due to long and, thus, complex resolutions. Indirect costs (administration expenses and opportunity costs) are also likely to rise with time. The increase of the average portfolio DRT by about 0.83 years yields to a rise in the average portfolio loss by about 5.05 percentage points which correspond to an increase of about 17%. The effect is even more pronounced in the outer tail of the portfolio loss distribution where the VaR (95%) increases by 32%. Forth, clustered DRTs are identified to put additional pressure on banks' liquidity in the future, because upcoming legal requirements will demand for higher stable funding needs for non-performing loans.

The remainder of this paper is structured as follows. Section 3.2 provides an example for the potential impact of systematic effects among default resolution times. Section 3.3 describes our data and methodology. Section 3.4 provides the empirical results. Section 3.5 shows the practical implications of these results. Section 3.6 concludes.

3.2 Why Care about Systematic Effects among Default Resolution Times?

Systematic effects involve longer or shorter average DRTs in certain time periods for non-performing loans. This co-movement across time is due to the joint dependency on common systematic factors. Banks have to deal with longer DRTs for all loans defaulted in crisis periods. This is particularly problematic as default rates are higher in recessions. The co-movement of DRTs has mainly two consequences for banks' profitability. First, long DRTs have negative impacts on the resolution process and, hence, on the loss. Besides from discounting effects, this might be due to dependencies between the DRT and non-discounted recovery payments. The DRT might serve as an indicator for the ease of resolution with long DRTs expressing complexity associated with high losses. Systematic effects among DRT lead, therefore, to higher losses in the aftermath of crisis periods. Second, liquidity will be restricted as loans stuck in the resolution process and increase the upcoming legally required amount of liquidity. In this section, we aim to quantify both – the impact of systematic effects among DRTs on the loss and liquidity.

Figure 3.1: Systematic movements in DRTs



Notes: The figure illustrates the systematic movements of DRTs for resolved loans. Box plots of the DRTs per year for the US, Great Britain, and Canada are displayed, whereas, outliers are hidden due to presentational purpose. The black horizontal lines within the box plots mark the medians. The means are separately displayed by gray lines.

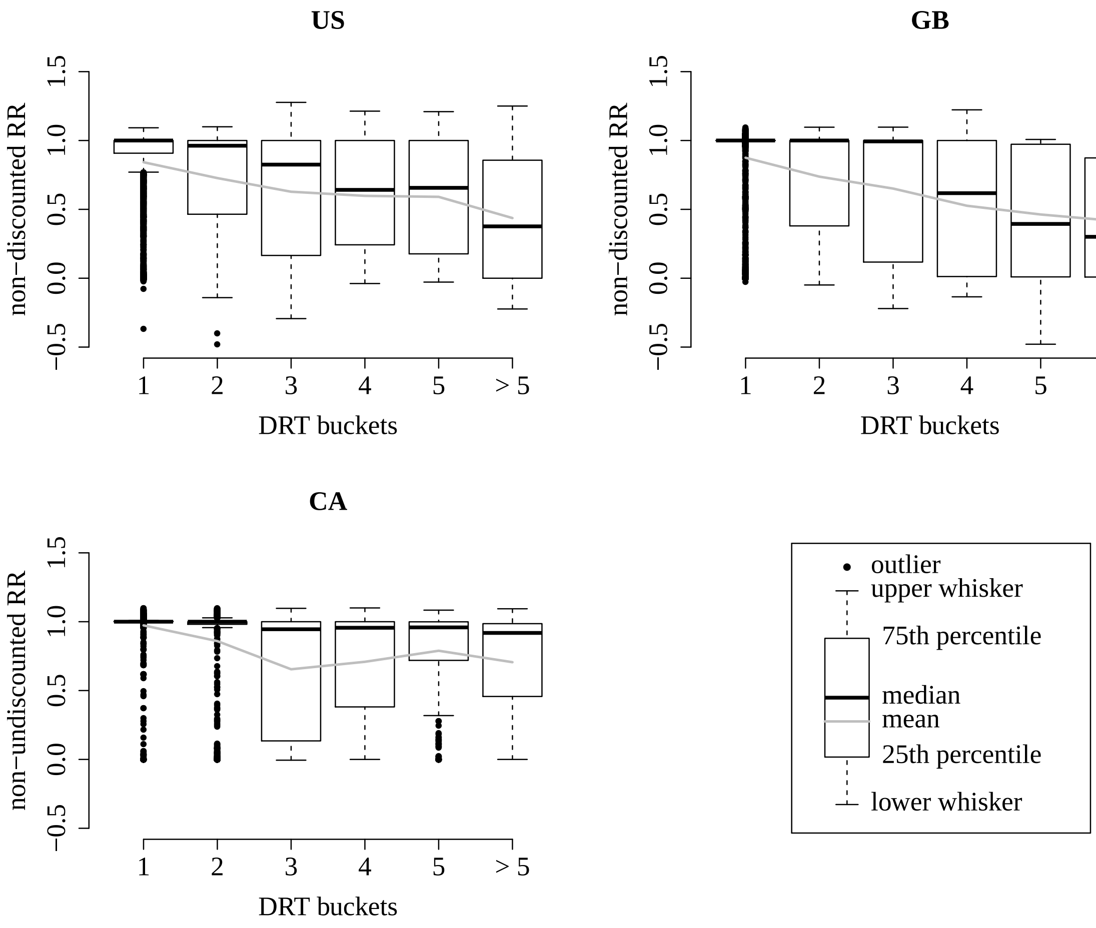
Figure 3.1 displays box plots of yearly final DRTs in the US, Great Britain and Canada. Systematic movements in accordance with economic conditions can be observed. Loans defaulted in crisis periods, e.g., 2008, are characterized by rather long resolutions.[6] This

[6]The increase of DRTs in Canada already starts in 2007 for at least some loans, which is caused

affects the recovery in the aftermath of crisis periods.

Long lasting resolutions of non-performing loans might negatively impact realized recoveries. Firstly, achievable recovery payments are more uncertain the further they reach into the future, as general conditions may change over time. In addition, DRTs could be seen as an indicator for the ease of resolution processes with complex resolutions exhibiting long DRTs and low feasible recovery payments. Figure 3.2 shows box plots of the non-discounted recovery rate (RR) divided by DRT buckets. The first bucket includes

Figure 3.2: Relation of DRT and non-discounted RR



Notes: The figure illustrates the relation of the DRT and non-discounted RR. Box plots of the non-discounted RR per bucket of DRT for the US, Great Britain, and Canada are displayed. The first bucket (marked with 1 on the x-axis) includes loans with DRTs up to one year. The second bucket (marked with 2 on the x-axis) includes loans with DRTs longer than one year up to two years and so on. In the last bucket (marked with > 5), loans with DRTs greater than five years are summarized. The black horizontal lines within the box plots mark the medians. The means are separately displayed by gray lines.

by longer resolution processes in general that push forward crises effects. A detailed description of the dataset is given in Section 3.3.

all loans with DRTs up to one year, the second one covers loans with DRTs higher than one but not higher than two years, and so on. In the last bucket, loans with DRT higher than five years are summarized. A rather monotonous, negative relation between DRTs and non-discounted RRs can be observed. Higher DRTs accompany with lower mean and median non-discounted RRs. Secondly, DRTs are directly considered in the final RR by discounting individual recovery cash flows.

In our comprehensive data set, the average DRT of US American loans amounts to 1.59 years. Our empirical results show that systematic effects increase the average DRT to 2.42 years during crisis periods. By assuming an exemplary loan with an exposure of default (EAD) of 1,000,000 USD and a constant risk adjusted interest rate of 5%, we evaluate the impacts of the DRT on the loss. A DRT of 1.59 years implies an average, non-discounted RR of 72.72%, whereas it amounts to 62.80% regarding a DRT of 2.42 years.[7] Thus, the consideration of systematic effects leads to an additional loss of 99,200.00 USD. After including discounting effects, the additional loss increases to 114,858.30 USD .

Aside from the direct restriction of available liquidity by longer DRTs, the US American implementation of Basel III requires to fulfill the Net Stable Funding Ratio (NSFR) and oblige financial institutions to provide additional amounts of medium and long term liquidity for certain facilities, e.g., non-performing loans, from 2018 on (see Board of Governors of the Federal Reserve System, 2016). The NSFR is defined by the ratio of the acquired stable funding (ASF) divided by the required stable funding (RSF), where RSF is calculated as the sum of banks' weighted assets. The weighting of corporate loans varies between 10% and 85%. However, it is enhanced to 100% if loans are rated as non-performi ng. As long as the non-performing status persists, banks need to provide additional liquidity for the affected assets.

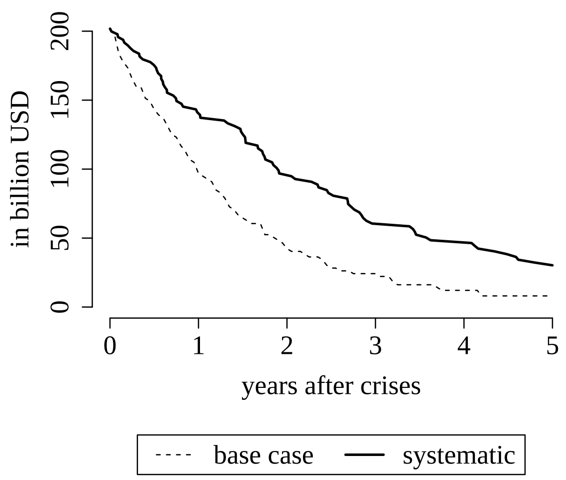
The total loan volume of the US American banks amounts to approximately 8.17 trillion USD.[8] The minimum ratio of non-performing loans was 0.70% while its maximum of 5.64% was reached during the Global Financial Crisis. Assuming an average RSF factor

[7] The above figures refer to an exemplary case in the US.

[8] All numbers are taken from the public research data base of https://fred.stlouisfed.org/. The time series are indexed by USTLLNUI and USNPTL.

of 50% for performing loans, the RSF would have increased from its pre-crisis level of 4.11 trillion USD to 4.32 trillion USD after the peak of the crisis due to a higher rate of non-performing loans. Differently stated, the US American banking system would have to face additionally required liquidity of 201.80 billion USD, if the NSFR requirements had been active. Again, we refer to an average DRT without (1.59 years) and with (2.42 years) the consideration of systematic effects. Hence, non-performing loans retain their status longer leading to a slower reduction in additionally required liquidity. Figure 3.3 displays the profile of additional RSF for an average DRT of 1.59 (when neglecting systematic risk) and 2.42 (when regarding systematic risk) years. The time a loan retains the non-performing status is assumed to follow an exponential distribution. We imply two stylized tendencies to resolution – without and with the consideration of systematic effects. Without systematic effects (average DRT of 1.59 years) the additional required liquidity declines to 96.86 billion USD after one year and to 42.34 billion USD after two years. After four years the additional amount shrinks to 12.11 billion USD. Allowing for systematic effects among DRTs (average DRT of 2.42 years), the reduction of additional

Figure 3.3: Additional required amount of stable funding



Notes: The figure illustrates the development of the aggregated additional required amount of stable funding for US American banks after the Global Financial Crisis. Assuming an average RSF of 50% for performing loans, the additional required amount of stable funding amounts to 201.80 billion USD after the peak of crisis due to an increased number of non-performing loans. Depending on the assumed DRTs, the development of additional required liquidity is displayed. In the base case (dashed line), an average DRT of 1.59 years is assumed, whereas, in the systematic case (solid line) the average DRT amounts to 2.42 years.

RSF decelerates. It amounts to 139.24 billion USD after one, 94.85 billion USD after two, and even 46.41 billion USD after four years.

These mechanisms might be one reason for the constantly increased RSF after the Global Financial Crisis as observed in the introductory observation phase for future liquidity requirements (see Gobat et al., 2014). Higher RSF in the aftermath of financial crises might burden the real economy and extend recessions. Confronting financial institutions with higher liquidity needs might restrict corporate lending and favor the investment in save havens, e.g., government bonds or gold. These not only exhibit a low weighting factor and, thus, reduce RSF, but also increase acquired stable funding (ASF). Furthermore, King (2013) and Dietrich et al. (2014) both identify possible negative effects of the NSFR on the profitability of banks. Systematic effects among DRTs might introduce procyclicality of liquidity regulation standards.

In the light of the above, a profound understanding of systematic effects in the time loans maintain non-performing is crucial. Systematic effects among DRTs entail co-movements, i.e., DRT are longer in certain time periods – namely during financial crises – for the entirety of non-performing loans. Firstly, we find a lower RRs for long DRTs on average. This is why the focus of this paper is laid on DRT and their systematic drivers instead of examining RRs directly. Modeling this can be compared to latent impact factors leading to higher default rates, longer DRTs, and smaller RRs during adverse economic scenarios and opposite effects during good economic times. Secondly, this will burden the liquidity of financial institutions during downturns due to future regulatory requirements and may even dampen upswings.

3.3 Methods and Data

3.3.1 Methods

This section derives a formal model for the DRT T, which we define as the length of the time period from a default date of a loan to its final resolution. Survival analysis provides established methods for modeling lengths of time up to a certain event and is, therefore, well suited for our purposes. A continuous-time approach takes into account that resolution may take place at any time after default. Thus, we define the intensity of defaulted loan i that represents the instantaneous tendency of resolution in the infinitesimally small interval $[t; t + \Delta t]$, conditional on no resolution up to t, as

$$\lambda_{it} = \lim_{\Delta t \to 0,\ t>0} \frac{P(\ t \leq T < t + \Delta t \mid T \geq t\ )}{\Delta t}, \tag{3.1}$$

In order to regress the intensity of resolution, we apply the Cox proportional hazards model. Let x_i be a vector containing a set of loan specific characteristics. The Cox model then takes the following functional form

$$\lambda_{it} = \lambda_{0t} \exp\left(x_i \beta\right), \tag{3.2}$$

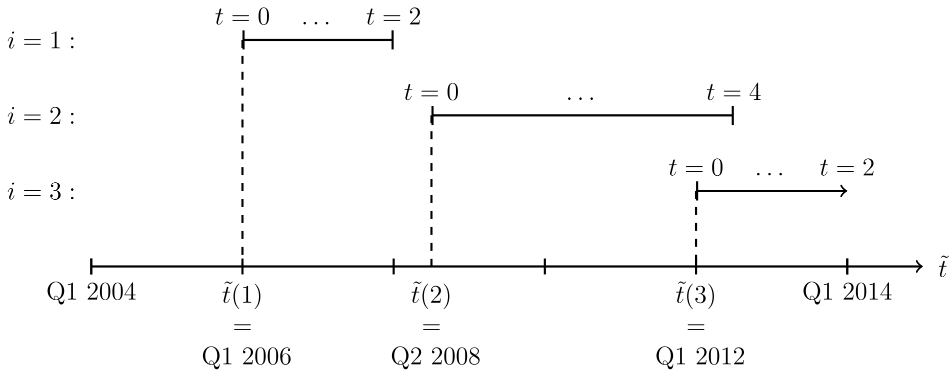
where, λ_{0t} is the baseline hazard rate representing an underlying tendency in the hazard function λ_{it} at baseline levels of the covariates x_i. The baseline hazard rate has an arbitrary functional form. Thus, the Cox model is a semi-parametric approach. The vector β includes the unknown parameters of the covariates x_i. In contrast to the method of ordinary least squares, we can include censored, i.e., not yet completed resolutions in our model. The statistical background and estimation procedure for the Cox model is given in Appendix 3.A. In the following, we refer to Equation (3.2) as Model I. Since it only contains loan specific characteristics as covariates, it serves as a reference model.

Next, we allow resolution processes to be additionally affected by the macroeconomy and include observable macroeconomic variables in the Cox model. We add a calender time level $\tilde{t}$ to the model and extend the Cox model to

$$\lambda_{it} = \lambda_{0t} \exp\left(x_i \beta + z_{\tilde{t}(i)} \gamma\right), \tag{3.3}$$

where $z_{\tilde{t}(i)}$ are macroeconomic variables at default time $\tilde{t}(i)$ of loan i and γ is an unknown parameter vector. In the following, we refer to Equation (3.3) as Model II. Figure 3.4 shows the two time levels that we take into account. First, we control for the time t since default because the tendency of resolution may change during the resolution process itself as we can see in the baseline hazard rate λ_{0t}. Second, we take into account the calendar time $\tilde{t}$ to investigate the role of macroeconomic covariates over (calender) time. As the macroeconomy changes over time, this model controls for some common co-movements in DRTs.

Figure 3.4: Resolution time levels



Notes: The figure illustrates the applied time stamps. Consider, e.g., loan $i = 1$ (upper part of figure). This loan defaulted at time $\tilde{t}(1)$ which corresponds to Q1 2006. Generally, the default time $\tilde{t}$ depends on the individual loan i. Thus, systematic variables (i.e., macroeconomic and frailties) are indexed at the loan depended default time $\tilde{t}(i)$. Afterwards, the loan $i = 1$ remains two years in resolution. The resolution intensity λ_{1t} depends on the time spend in resolution t. The index of the time spend in default t and the default time $\tilde{t}(i)$ are, thus, two deviating time scales which is indicated by the different notation.

Finally, we extend the model to *unobservable stochastic* common factors which play an important role in the credit risk literature for modeling default risk, in addition to observable common factors. The unobservable factors yield the dependent variable to be stochastically correlated, in contrast to only deterministic co-movements driven by observable factors. Let $U_{\tilde{t}(i)}$ be a normally distributed random variable with mean zero and variance σ^2, i.e.,

$$U_{\tilde{t}(i)} \sim N\left(0, \sigma^2\right), \tag{3.4}$$

commonly termed as frailty in the Cox model. Then Model III becomes

$$\lambda_{it} = \lambda_{0t} \exp\left(x_i\beta + z_{\tilde{t}(i)}\gamma + U_{\tilde{t}(i)}\right), \tag{3.5}$$

where, σ^2 is an additional parameter to be estimated. Frailties introduce stochastic correlation into the modeling framework, i.e., a negative time $\tilde{t}(i)$ realization of the frailty reduces the hazard rate of all loans simultaneously and, thus, increases their DRT, et vice versa.

3.3.2 Data

This paper uses a subsample of the unique loss data base provided by GCD.[9] This data base pools loss information of 50 member banks around the world, including several global systemically important banks.

To correct for minor input errors we apply the procedure of Höcht and Zagst (2010) and Höcht et al. (2011) with the distinction that we evolve a second selection criterion for post-resolution payments. The pre-resolution criterion is calculated as the sum of all relevant transactions (including charges-offs) divided by the outstanding amount of the loan at default. The post-resolution criterion, in contrast, is the sum of all post-resolution payments divided by a fictional outstanding amount at resolution. The barriers are set to $[\,90\%\,,\,110\%\,]$ for the pre-resolution criterion for resolved loans and to $[\,-50\%\,,\,400\%\,]$ for unresolved loans. The barriers of the post-resolution criterion are set to $[\,-10\%\,,\,110\%\,]$. This criterion is only adapted for resolved loans. Loans found outside these intervals are excluded because of an extraordinary structure of cash flows. Hereby, 2.0% of resolved loans in the overall data base are sorted out due to the pre-resolution criterion and 0.2% due to the post-resolution criterion. Regarding unresolved loans, we excluded 0.2%. Finally, we eliminate loans with abnormal high and low LGDs (< –50% and >150%). Thereby, less than 0.1% of the overall data base is excluded.

We use a subsample of the corrected overall data base consisting of small and medium sized entities (SMEs) and large corporates (LCs) from the US, Great Britain, and Canada. We remove loans with exposures at default (EADs) smaller than 500 USD, which corresponds to 11.6% of the subsample data. With respect to corporate loans (SMEs and LC), loans of this size seem negligible and might distort results.[10] We further restrict

[9] GCD is a non profit initiative which aims to help banks to measure their credit risk by collecting and analyzing historical loss data. They are formally known as the Pan-European Credit Data Consortium (PECDC). See http://www.globalcreditdata.org/ for further information.

[10] The magnitude is inspired by the materiality threshold of the European Banking Authority (2014).

the time period from 2004 until 2013 to ensure a consistent default definition due to the Basel accords and a minimum quantity of data per year. Thereby, we exclude 16.5%. A subset of 17,395 loan remains.

DRT Data

The data show country specific differences regarding the DRT. Table 3.1 displays descriptive statistics for the DRT.

Table 3.1: Descriptive statistics of DRT

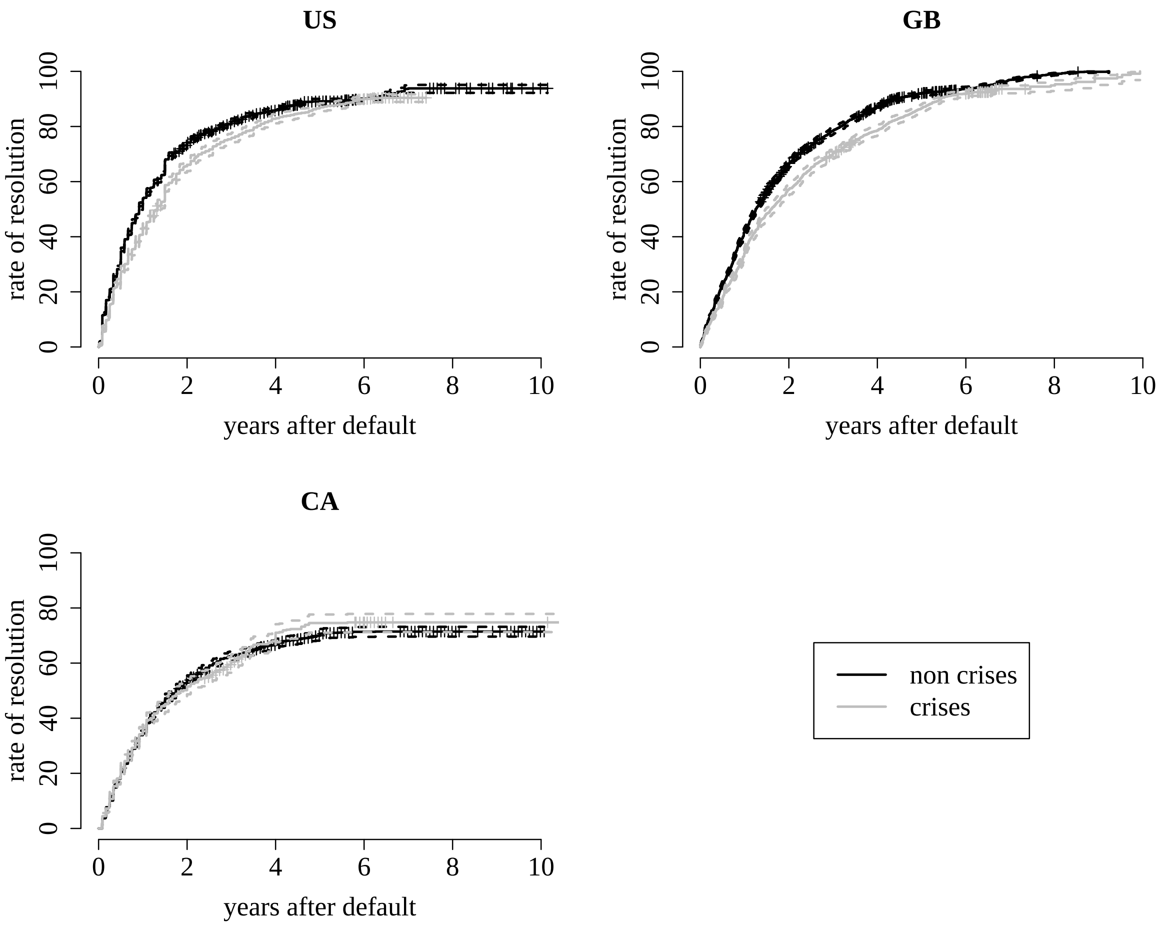
	Overall	US	GB	CA
n	17,395	7,133	5,780	4,482
Resolved	83.07%	86.26%	92.65%	65.64%
Mean	1.40	1.17	1.74	1.24
Median	0.99	0.83	1.23	0.92
Standard deviation	1.37	1.18	1.61	1.10
Unresolved	16.93%	13.74%	7.35%	34.36%
Mean	4.39	4.20	4.50	4.48
Median	3.99	3.90	4.38	3.90
Standard deviation	2.07	1.90	1.56	2.28

Notes: The table summarizes the descriptive statistics (mean, median, and standard deviation) of DRTs for the overall data set and separated for the US, Great Britain, and Canada. The presentation additionally distinguishes between resolved and unresolved loans. The latter indicates that these loans are still in the resolution process.

The mean for resolved loans is 1.40 years compared to a median of 0.99 years which indicates a skewness of resolution processes. The present mean and median of the unresolved loans are substantially higher with 4.39 and 3.99 years. From US American loans 13.7% are defaulted but not entirely resolved until March 2015. We include these censored information by using a survival model in order to avoid a resolution bias. In Great Britain 7.4% and in Canada 34.4% are not resolved yet.

Figure 3.5 shows the cumulative rate of resolution in years after default by using the inverse Kaplan-Meier estimator and reveals country specific differences. For example, 71.5% of all defaulted loans in the US are resolved after 24 months in contrast to 63.1% in Great Britain and 52.7% in Canada. It also provides evidence of a systematic component in resolution processes. The figure distinguishes between loans defaulted within and out of economic downturns, defined by the indicator of the OECD. Up to 24 months after default, the tendency of resolution is lower for loans that defaulted in a crisis. In the US and Great Britain the effects are still valid for the following years. Especially in Great Britain, recession based defaults imply higher DRT. For example, 66.4% of loans that did not default in recessions are resolved in two or less years. The same proportions takes 23.1% longer for crises based defaults. In contrast, the ratio of censorships for Canada is high with 34.4% and, thus, flat and close-lying courses result.

Figure 3.5: Observed resolution rates



Notes: The figure illustrates the inverse Kaplan-Meier estimators of resolution separated for the US, Great Britain, and Canada. Generally, the inverse Kaplan-Meier estimator displays the rate of resolution, i.e., the proportion of loans which are resolved depending on the time spent in resolution. Defaulted but unresolved loans are included as censorships by vertical marks. Confidence intervals are indicated at the 95%-level by close dashed lines due to the data set's size. Furthermore, it is distinguished between non-crisis (black lines) and crisis (gray lines) periods, whereby, crises are defined by the monthly recession dummy of the Organisation for Economic Co-operation and Development (OECD).

Loan Specific Variables

Table 3.2 summarizes loan specific characteristics in our data set. In the following analysis, they are included to control for variation in the intensities caused by loan specific attributes. As metric determinants, we include the logarithm of EAD in USD and the number of collateral. The latter specifies the exact quantity of security assets which are assigned to the loans. Furthermore, various categoric variables are used. The asset class defines whether the debtor is an SME or LC. With the facility type, we distinguish between medium term, short term, and other facilities. The seniority level is divided into the categories super senior, pari-passu, and non senior. Super senior implies that the considered loan is the only preferred claimant. Pari-passu, in contrast, indicates that more creditors share the highest rank to the debtor. Nature of default controls for the two main default definitions set by Basel Committee on Banking Supervision (2006), namely if a debtor is "unlikely to pay" or "past due more than 90 days on any material credit obligation" (§452). The remaining categories – bankruptcy, charge-off / provision, sold at material credit loss, distressed restructuring, and non accrual – may be seen as subcategories of the more general one unlikely to pay. We do not summarize these categories as they might supply additional explanation power. Furthermore, we include a guarantee indicator stating if the loan is additionally secured by some form of guarantee. The collateral indicator is divided into the categories NO, other collateral, and real estate. As loans might exhibit more than one security, real estate indicates that there is at least one among the collaterals. The cured indicator states if a debtor returned back to performance after entering default, i.e., if the debtor is back to a sound rating. Finally, we control for various industries.

Macroeconomic Variables

As explained in the previous subsection, observable macroeconomic risk factors are included in Model II and III (see Equation (3.3) and (3.5)). All macroeconomic variables are defined on a country specific quarterly basis. We include the year-on-year log return of the equity index to capture the financial economy. The S&P 500 is used for the US, the

Table 3.2: Descriptive statistics of loan specific characteristics

		Overall	US	GB	CA
	n	17,395	7,133	5,780	4,482
	Metric				
EAD	Mean	2,131,389.29	4,238,878.41	925,714.09	332,210.97
	Median	129,086.24	470,859.53	72,533.24	53,880.04
	Standard deviation	20,826,299.90	31,945,773.20	5,773,926.64	2,042,389.08
Number of collateral	Mean	1.53	1.09	2.85	0.55
	Median	1.00	1.00	1.00	0.00
	Standard deviation	3.99	1.71	6.01	2.62
	Categoric				
Borrower	SME	76.97%	77.23%	89.95%	59.82%
	LC	23.03%	22.77%	10.05%	40.18%
Facility type	Medium term	51.78%	53.51%	45.38%	57.30%
	Short term	27.47%	11.62%	46.04%	28.74%
	Other / Unknown	20.75%	34.87%	8.58%	13.97%
Seniority code	Pari-passu	39.00%	13.16%	63.91%	47.99%
	Super senior	48.86%	82.80%	35.76%	11.76%
	Non senior	0.37%	0.64%	0.33%	0.00%
	Unknown	11.76%	3.39%	0.00%	40.25%
Nature of default	90 days past due	20.48%	31.11%	23.01%	0.31%
	Unlikely to pay	18.76%	27.97%	15.73%	8.01%
	Bankruptcy	7.27%	3.01%	12.49%	7.30%
	Charge-off / provision	6.35%	1.51%	16.90%	0.45%
	Sold at material credit loss	0.62%	1.50%	0.00%	0.00%
	Disstressed resturcturing	1.21%	0.64%	2.85%	0.00%
	Non accural	39.25%	31.61%	25.19%	69.52%
	Unknown	6.06%	2.64%	3.82%	14.41%
Guarantee indicator	NO	69.84%	61.14%	65.85%	88.84%
	YES	30.00%	38.47%	34.15%	11.16%
	Unknown	0.16%	0.39%	0.00%	0.00%
Collateral indicator	NO	32.57%	36.65%	37.84%	19.30%
	Other collateral	37.15%	47.55%	26.73%	34.02%
	Real estate	18.29%	14.43%	35.43%	2.34%
	Unknown	11.99%	1.37%	0.00%	44.33%
Cured indicator	NO	78.59%	75.84%	75.12%	87.42%
	YES	21.41%	24.16%	24.88%	12.58%
Industry	Finance, insurance, RE	11.84%	15.97%	10.57%	6.92%
	Agriculture, foresty, fishing	3.48%	1.33%	2.79%	7.81%
	Mining	0.79%	0.91%	0.38%	1.14%
	Construction	11.69%	10.63%	14.79%	9.37%
	Manufacturing	16.90%	18.52%	15.10%	16.64%
	Transp., commu., sanitary services	5.96%	6.27%	4.71%	7.07%
	Wholesale and retail trade	22.17%	13.94%	30.31%	24.77%
	Services	19.81%	16.87%	18.91%	25.66%
	Unknown	7.36%	15.58%	2.44%	0.62%

Notes: The table summarizes the descriptive statistics (mean, median, and standard deviation) of metric independent variables. For categoric independent variables, proportions of the categories are displayed. Generally, the variable name is stated in the first column. For categoric variables, the categories are presented in the second column. The presentation is done for the overall data set and separated for the US, Great Britain, and Canada.

FTSE 100 for Great Britain, and the S&P TSX for Canada. The year-on-year log growth of industry production is included as an indicator for the real economy. In order to capture long-term monetary expectations we include the year-on-year change in term spread between 10-year long-term government bonds and 3-month government securities. The stock market volatility index captures market expectations of future economic conditions. We use the volatility indices of the CBOE for the US, the FTSE for Great Britain and the S&P TSX for Canada.[11] Figure 3.6 shows the macroeconomic variables. The Global Financial Crisis results in increasing term spreads, volatility indices and lower industry production and equity indices. Table 3.3 reports the corresponding pairwise correlations which appear to be comparably high in absolute terms. This is important for the interpretation of regression results in the next section. In general, interactions between signs and significances may result when correlated independent variables are simultaneously included in a model.

Furthermore, we include a World Bank score measuring the efficiency of default resolution (see World Bank, 2015). This score evaluates the efficiency of the regulatory framework regarding the resolution of an insolvent company by adopting a survey process. The methodology is inspired by Djankov et al. (2008).[12]

Table 3.3: Pairwise correlations of macroeconomic variables

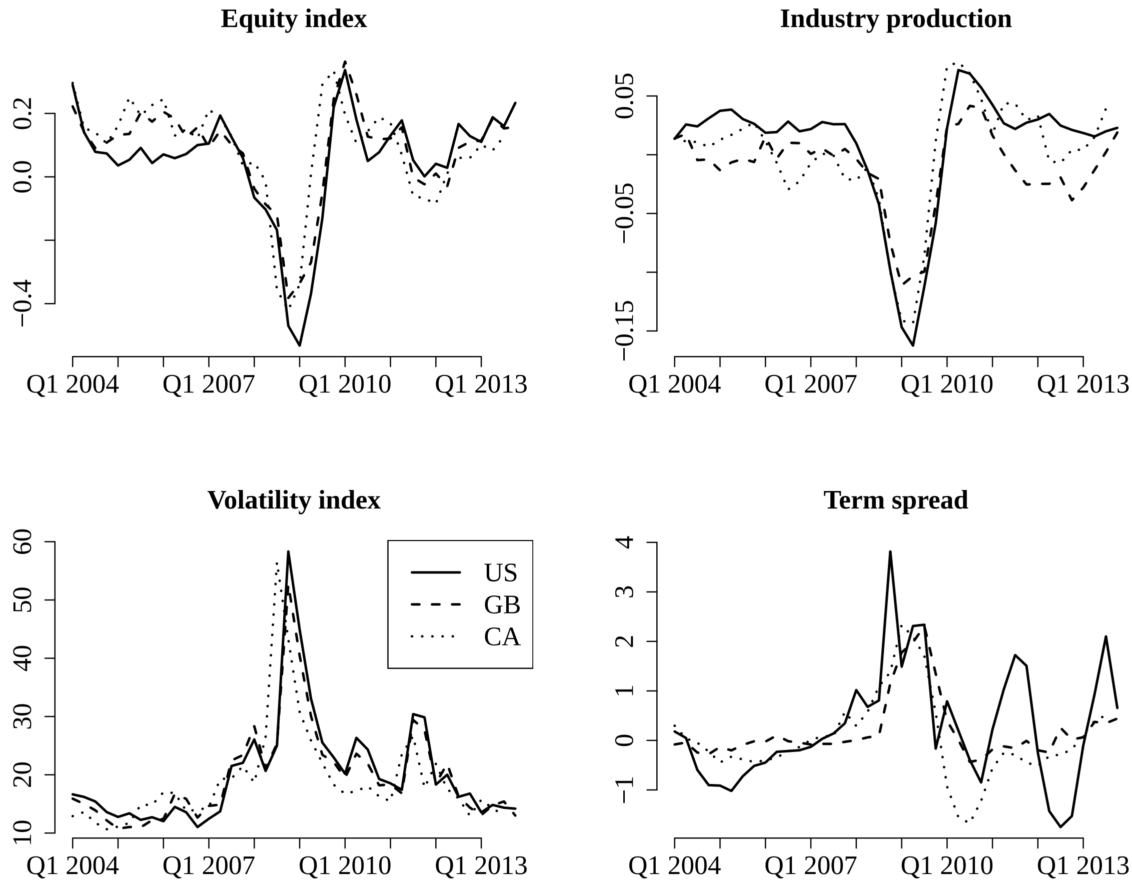
	US			Great Britain			Canada		
	IP	VIX	TS	IP	VIX	TS	IP	VIX	TS
Equity Index	0.79	-0.79	-0.68	0.76	-0.81	-0.56	0.50	-0.79	-0.47
Industry Production (IP)		-0.64	-0.70		-0.65	-0.86		-0.51	-0.91
Volatility Index (VIX)			0.77			0.59			0.58

Notes: The table summarizes the pairwise correlations of macroeconomic variables which are included in Model II as of Equation (3.3) and Model III as of Equation (3.5). Thereby, the year-on-year (yoy) log return of the country specific equity index, the yoy log return of the country specific industry production (IP), the level specification of the country specific volatility index (VIX), and the yoy log return of the country specific term spread are considered.

[11]We have also tested other popular macroeconomic indicators, e.g., government bond rates, gross domestic product, house price indices, inflation, the ratio of non-performing loans and the unemployment rate. None of these performed as good as the chosen four variables in terms of goodness of fit, plausibility and significance. (see Section 3.4.3)

[12]See http://www.doingbusiness.org/methodology/resolving-insolvency for further information.

Figure 3.6: Descriptive statistics of macroeconomic variables



Notes: The figure illustrates the time-series of macroeconomic variables which are included in Model II as of Equation (3.3) and Model III as of Equation (3.5). Thereby, the year-on-year (yoy) log return of the country specific equity index, the yoy log return of the country specific industry production (IP), the level specification of the country specific volatility index (VIX), and the yoy log return of the country specific term spread are considered.

3.4 Results

3.4.1 Overview of Formal and Informal Proceedings of Resolution

Before we turn towards the results of Model I, II, and III, a brief introduction of formal insolvency proceedings and informal resolution mechanisms in the US, Great Britain and Canada is presented to derive creditor incentives and determine bargaining powers in resolution processes. The relevant codes can be found in Title 11 of the United States Codes for the US, in the Insolvency Act of 1986 (IA 86) and 2000 for Great Britain, and in the Company Creditor Agreement Act (CCAA) and the Bankruptcy and Insolvency Act (BIA) for Canada (see Table 3.4). Those statutory regulations are primary relevant in formal insolvency proceedings, however, conclusions can be drawn towards informal resolution of non-performing loans, in particular towards bargaining powers in negotiations.

Table 3.4: Overview of formal and informal proceedings

	United States	**Great Britain**	**Canada**
Focus	Debtor	Senior secured creditor	Creditor
Stay	Automatic stay (§362)	Restriction of proceedings by court (IA 86 285.)	Stay regulations (CCAA §11.02 and BIA §69 ff) → Unlimited with court approval (CCAA §11.02) → Unlimited (BIA §69 ff)
Enforcement	Avoiding enforcement → Automatic stay (§362)	Secured claims enforceable (IA 86 285. (4))	Avoiding enforcement → Stay (CCAA §11.02) → Stay (BIA §69)
Management	*Debtor-in-possession* (§1103 and §1107)	→ Administrator (IA 86 8.) → Receiver (IA 86 32.) → Liquidator (IA 86 91. ff)	*Debtor-in-possession* (CCAA §11.03 and BIA §69.31) → Monitor (CCAA §23 ff) → Trustee (BIA §43)
Financing	*Super-priority-financing* (§503 and §507)	No	*Super-priority-financing* due to court permission
Informal	Implemented by contract (e.g., debt refinancing)	→ Consensual → Receivership (fixed charges)	Consensual (approved by court)

Notes: The table gives an overview of the formal and informal proceedings of resolution for the US, Great Britain, and Canada. The regulations regarding formal insolvency can be found in Title 11 of the United States Code for the US, in the Insolvency Act 1986 (IA 86) and 2000 for Great Britain, and in the Companies Creditor Agreement Act (CCAA) and the Bankruptcy and Insolvency Act (BIA) for Canada. The focus which is displayed in the first row of the table is derived on basis of the insolvency codes.

The US insolvency code strongly focuses on the debtor. A comprehensive automatic stay (§362) avoids enforcement of secured claims in formal insolvency. The debtor stays in possession (*debtor-in-possession*, §1103 and §1107) during the process. Additional claims of higher priority to existing debt (*super-priority-financing*) are possible by law. Informal resolution mechanisms are commonly implemented by contract. Conceivable are among others debt refinancing, debt for equity swaps, or exchange offers.

Historically, the British insolvency law favors senior secured creditors. Therefore, a restriction of proceedings has to be approved by court (IA 86 285.). Generally, secured claims are enforceable at any time during formal and informal workouts (IA 86 285. (4)).

The management is transfered to an Administrator (IA 86 8.), Receiver (IA 86 32.), or Liquidator (IA 86 91. ff) depending on the kind of formal insolvency proceeding. There is no option of *super-priority-financing* in insolvency. Informal resolutions are typically the result of negotiations post default. Except the enforcement of fixed charges where a receiver can be appointed at any time during formal and informal proceedings.

Although the Canadian insolvency code is rather similar to the US, the focus is shifted towards creditors. Comprehensive stay regulations to avoid enforcement are also implemented by law (CCAA §11.02 and BIA §69 ff), however, under the CCAA an unlimited stay has to be approved by court (CCAA §11.02). In contrast to the US, the *debtor-in-possession* is subject to stronger supervision by a monitor (CCAA §23 ff) or trustee (BIA §43). *Super-priority-financing* is not implemented by law, but can be granted by court. Informal workouts are achieved on consensual basis. Usually, the result of negotiations is additionally approved by court.

According to Haugen and Senbet (1978) and Haugen and Senbet (1988), all affected parties, i.e., debtors and creditors, have incentives to prefer informal proceedings compared to formal insolvency as informal workouts are less costly and more efficient. However, there exist conditions lowering these incentives and induce debtors or creditors to file for formal insolvency. This conditions are (i) dispersion of creditors, (ii) incomplete contracts, and (iii) information asymmetries (see, e.g., Blazy et al., 2014). These conditions do not only increase incentives to formal workout but also indirectly increase DRTs as they complicate informal negotiations and, thus, lead to longer resolution processes. In the following, we discuss these three conditions regarding the US, Great Britain, and Canada to conclude a profound theory regarding resolution intensities among these countries.

Dispersion of creditors – or more broadly formulated dispersion of affected parties – lowers incentives to informal proceedings which increases DRTs. Hereby, not the number of parties involved but the bargaining powers of a single party or a homogeneous group of parties is of relevance (see Blazy et al., 2014). In the US, the focus of formal insolvency proceedings is strongly set in favor of the debtor. The debtor might, thus, have strong bargaining powers in pre-insolvency negotiations as the opportunity of threatening gestures exists. Not much is going to change from a debtor perspective in formal insolvency proceedings (due to, e.g., automatic stay and *debtor-in-possession*). Therefore, debtors are able to threaten with filing for Chapter 11 as creditors aim to avoid formal proceedings. In Great Britain, secured creditors are historically favored by formal insolvency proceedings and informal resolution mechanisms as the enforcement of claims is assured at any time during resolution. Therefore, bargaining powers are concentrated on this rather homogeneous group. Although the Canadian insolvency law is similar to the US, it is more creditor orientated (due to, e.g., more court involvement and supervision to a higher extent). Bargaining powers are, thus, more disperse as not a homogeneous group of creditors is focused.

Incomplete contracts further decrease incentives to informal resolution and increase DRTs. In the US, informal proceedings are implemented by contract. Thus, rather complete contracts can be assumed. This is not the case in Great Britain and Canada where informal workouts are negotiated on consensual basis. Furthermore, negotiation results are usually approved by court in Canada. Asymmetric information might be present in all considered countries to a rather low extent as the quality of accounting standards is high in the US, Great Britain and Canada (see La Porta et al., 1998). Creditors are, thus, informed in similar and adequate manner.

In summary, the US should be characterized by comparatively high resolution intensities and, thus, short DRTs as the bargaining powers are concentrated in favor of the debtor and contracts are rather complete. Contrary, Canada is shaped by dispersion of creditors as creditors in general hold bargaining powers in formal insolvency proceedings and informal resolution mechanisms. Furthermore, contracts are rather incomplete and courts are involved to approve negotiated informal resolutions which additionally increases DRTs. Due to concentrated bargaining powers in favor of secured creditors but rather incomplete contracts, Great Britain should exhibit longer DRTs than the US, but shorter DRTs than Canada.

Besides the three discussed obstacles of informal resolution, systematic effects might influence incentives of debtors and creditors. Although all affected parties exhibit strong incentives to fast and efficient resolution, this may change in crises periods. Confronted with harsh market conditions, debtors and creditors tend to let time pass by to realize better prices in liquidations or to ensure more favorable conditions for restructuring efforts. Secondly, default rates are high in crises periods. Thus, creditors are confronted with a considerably higher amount of non-performing loans. The affected divisions might be at their capacity limits leading to decelerated internal proceedings. Therefore, resolution intensities should be lower during crises periods and, thus, DRT increase.

3.4.2 Loan Specific Impacts on Resolution

In the first part of our analysis, we investigate the role of loan specific characteristics in modeling the DRT. Table 3.5 shows the estimation results for Model I, i.e., a Cox proportional hazards model including loan specific covariates. A positive parameter estimate indicates a higher intensity to resolution and, thus, a tendency to shorter resolution processes. Dermine and Neto de Carvalho (2006) and Grunert and Weber (2009) do not study determinants of the length of workout processes but of the resulting loan LGDs and give a general behavior of loan-specific covariates on the workout process. The intuition behind those impacts can be mostly transferred to DRT modeling.

The loan size lowers the resolution tendency and, thus, increases DRT. This can be explained by more efforts and coordination problems for larger loans. Regarding seniority, the medium case pari-passu indicates the lowest resolution rate compared to super senior and non senior loans. Equally ranked debt seems to complicate the processing after default. In general, collateral leads to a shorter resolution process. A higher resolution rate for real estate collateral can only be identified for Great Britain. This may be justified by the comprehensive stay regulations in the US (§362) and Canada (CCAA §11.02 and BIA §69 ff). The strong focus on senior secured creditors in Great Britain ensures the enforcement of fixed charges at any time prior and throughout formal insolvency (IA 86 285. (4)). The number of collateral results in an acceleration in the US only. Some loans in default return to performing. Usually, these cures occur shortly after default. Thus,

Table 3.5: Regression results for Model I

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
log(EAD)		-0.0524	***	(0.0070)	-0.0855	***	(0.0077)	-0.0212	*	(0.0118)
Asset Class (SME)	Large Corporates	0.2054	***	(0.0377)	0.1713	***	(0.0495)	-0.1023	*	(0.0586)
Facility	Short term	0.1372	***	(0.0411)	-0.1645	***	(0.0308)	-0.0952	**	(0.0446)
(Medium term)	Other	-0.0495		(0.0333)	0.2238	***	(0.0546)	0.0711		(0.0582)
Seniority	Super senior	0.1361	***	(0.0485)	0.7606	***	(0.0357)	0.6476	***	(0.1007)
(Pari-passu)	Non senior	0.3548	**	(0.1581)	0.5463	**	(0.2348)			
	Unknown	0.2012	**	(0.0963)				0.7364	***	(0.1212)
Nature of default	Unlikely to pay	-0.2662	***	(0.0426)	-0.0059		(0.0521)	-0.5571	*	(0.3077)
(90 days past due)	Bankruptcy	0.1183		(0.0753)	-0.1522	***	(0.0512)	-0.3545		(0.3022)
	Charge-off / provision	0.3608	***	(0.1064)	-0.1769	***	(0.0451)	0.9064	**	(0.3713)
	Sold at material credit loss	1.6991	***	(0.1023)						
	Distressed restructuring	0.3102	**	(0.1532)	0.0135		(0.0945)			
	Non accrual	-0.0090		(0.0366)	-0.2103	***	(0.0444)	-0.4389		(0.2937)
	Unknown	-0.1240		(0.1288)	1.0943	***	(0.0796)	-0.8234	***	(0.2983)
Guarantee	Unknown	0.4043		(0.2485)						
(NO)	YES	0.1300	***	(0.0289)	-0.1182	***	(0.0311)	0.1780	**	(0.0738)
Collateral	Other collateral	0.0722	**	(0.0321)	0.0613	*	(0.0371)	0.3963	***	(0.0799)
(NO)	Real estate	0.0314		(0.0465)	0.1085	***	(0.0372)	0.2085		(0.1889)
	Unknown	-1.7936	***	(0.2424)				1.2859	***	(0.1139)
Number of collateral		0.0309	***	(0.0085)	0.0024		(0.0023)	0.0095		(0.0067)
Cured (NO)	YES	0.4780	***	(0.0325)	0.8825	***	(0.0355)	0.9426	***	(0.0536)

Continued on next page

Table 3.5 (continued): Regression results for Model I

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
Industry	Agric., forestry, fishing	-0.2189	*	(0.1211)	0.2637	***	(0.0951)	-0.1141		(0.1008)
(Finance, insurance, RE)	Mining	-0.0450		(0.1414)	-0.0530		(0.2291)	0.5440	***	(0.1773)
	Construction	-0.3230	***	(0.0540)	-0.3110	***	(0.0604)	0.2099	**	(0.0964)
	Manufacturing	-0.2453	***	(0.0464)	0.0959	*	(0.0575)	0.2785	***	(0.0848)
	Transp., commu., sanitary services	0.0242		(0.0617)	0.2421	***	(0.0773)	0.2322	**	(0.1019)
	Wholesale and retail trade	-0.1076	**	(0.0485)	0.0930	*	(0.0512)	0.2460	***	(0.0833)
	Services	-0.0841	*	(0.0462)	0.1479	***	(0.0551)	0.2070	***	(0.0798)
	Unknown	0.2876	***	(0.0474)	0.2375	**	(0.0970)	0.0589		(0.2960)
LL				-49,365			-40,620			-22,046
AIC				98,787			81,290			44,142
McFadden's adjusted R^2				0.0112			0.0234			0.0486
Cox & Snell's R^2				0.0223			0.0386			0.0486
Resolved Loans				6,153			5,355			2,942
Loans				7,133			5,780			4,482

Notes: The table summarizes regression results for country specific impacts of loan characteristics on the tendency of resolution. The model specification fulfill Equation (3.2), i.e., neither frailties nor macroeconomic variables are included. The first column contains covariate names and the second includes corresponding categories if the variable is of categoric nature. The reference category is given in parenthesis. Significance is indicated at 10% (*), 5% (**) and 1% (***). Standard errors (SE) are given in parenthesis. For completeness, results for a regression that uses all observations jointly are given in Appendix Table 3.B.3.

they are identified as a factor for short resolution processes.

Some differences across countries in the effects of the covariates can be seen. While LCs show significant higher resolution intensities in the US and Great Britain, the effect is negative in Canada. However, the significance in Canada vanishes in Model III, i.e., after considering systematic components.[13] Short term facilities have lower resolution tendencies in Great Britain and Canada, whereas, the influence of this category is significantly positive in the US. This might be ascribed to country specific lending behavior. Other facility types significantly lead to higher resolution rates in Great Britain. Furthermore, guarantees accelerate the resolution process in the US and Canada but lead to a deceleration in Great Britain. Both directions might be explained either by the possibility of direct access to a third party or the necessity to establish additional claims in the resolution process. The actual causality might depend on the type of guarantee. The nature of default and the debtor's industry affiliation have several country specific particularities. For example, FIRE affiliation seems to accelerate resolution processes in the US but decelerates it in Great Britain and Canada.

In summary, loan specific characteristics seem to have a great impact on resolution processes. Identified decelerators are the loan size and an equally seniority weighting of the loan. Collateralization is detected as an accelerator. The effects of nature of default and industry affiliation strongly depend on the loan's country of origin.

3.4.3 The Systematic Movement of Resolution Processes

Legal and Administrative Reasons

In addition to loan specific characteristics, country specific legal conditions and bank practices affect the time to resolution (see Section 3.4.1 for an overview). As stated in Section 3.3.1, the Cox proportional hazard model is a semi-parametric approach because the baseline hazard rate λ_{0t} has an arbitrary functional form. Thus, it catches country specific particularities of resolution processes. However, this baseline indicates an underlying resolution tendency in the hazard function λ_{it} independent of covariates. In

[13]The results are available from the authors upon request.

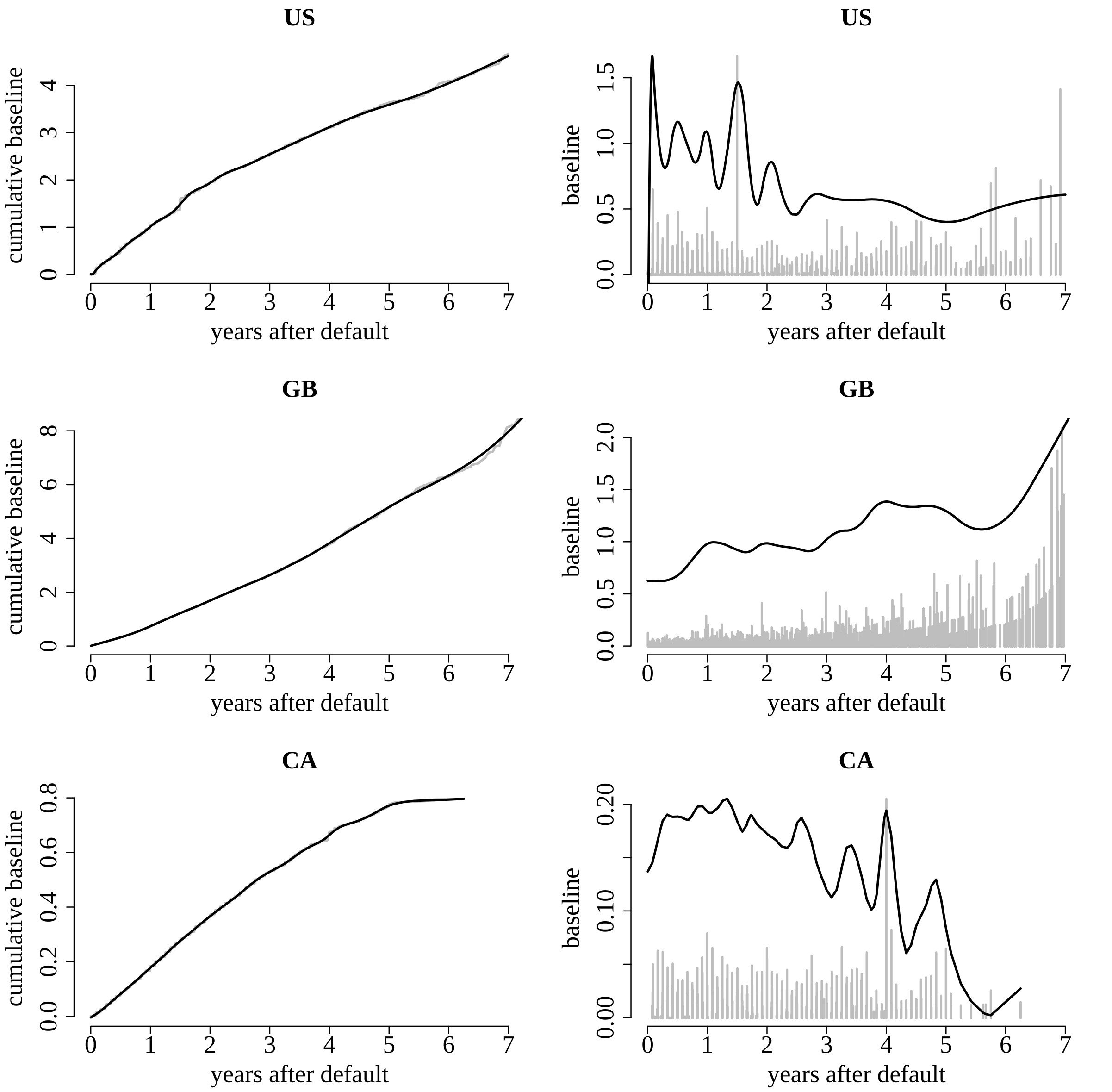
contrast to the frailty which expresses time-dependent effects based on calender dates, the baseline hazard rate provides information regarding effects that are caused by the resolution process itself.

Figure 3.7 displays the baseline hazard rates of Model I. The left panels show the step-like cumulative baseline hazard rate $\Lambda_{0t} = \int_0^t \lambda_{0v}\, dv$. The bars in the right panels are the corresponding non-cumulative baseline hazards, i.e., λ_{0t}. We smooth these discrete baselines in order to facilitate the interpretation. In a first step, we estimate cubic splines to approximate cumulative baselines which are displayed as black lines in the left panels. In a second step, we use the corresponding derivative as a smoothed (non-cumulative) baseline. The results show a background intensity, i.e., a basic tendency of resolution depending on the time spent in resolution.

Generally, more than 90% of all loans are resolved within five years after default. Thus, the results in this time period might be of the most interest when focusing on DRTs. Note that this statement can change when considering losses of loans at the same time. As already laid out at the end of Section 3.2, loans exhibiting long DRTs usually come along with low recovery rates. Among defaulted loans in our overall data set, we find that below the 10% quantile of recovery rates (which equals the 90% quantile of losses given default), the fraction of loans with DRTs higher than five years equals 19%, while this fraction corresponds to 7% for loans above the 10% quantile of recovery rates. Results derived for DRTs in this study should be kept in mind when studying recovery rates. Precise credit risk assessment only seems possible if both risk components and especially their linkage is taken into account holistically.

Regarding the analysis of DRTs, considerable differences across the US, Great Britain, and Canada arise. Comparing the level, the US exhibits the highest non-cumulative baseline intensity to resolution, closely followed by Great Britain. The baseline hazard rate of Canada is considerably lower. This corresponds with the theoretical considerations in Section 3.4.1. The US and Great Britain are characterized by concentrated bargaining powers since the debtor (US) or senior secured creditors (GB) are strongly favored by insolvency codes. This increases the concentration of negotiation powers even in infor-

Figure 3.7: Baseline intensities of resolution



Notes: This figure illustrates country specific baseline hazard rates of resolution for the US, Great Britain, and Canada. In the left panels the cumulative, in the right panels the non-cumulative baseline intensity is displayed. The estimated outcome is marked in gray. In the left panels, the black lines smooth the cumulative step functions by cubic splines. Thus, we compress the discrete baselines in the right panels to informative continuous baseline intensities which are derivatives of the smoothed cumulative baselines and represented by black lines.

mal proceedings due to threatening gestures. Nearly any harm occurs for US debtors in filling for insolvency due to the comprehensive automatic stay (§362) and the debtor-in-possession setting (§1103 and §1107). Whereas, US creditors aim to avoid long lasting formal proceedings. In Great Britain, senior secured creditors can enforce their claims at any time in formal insolvency proceedings and informal resolution mechanisms. Therefore, this rather homogeneous group holds comprehensive bargaining powers in negotiations compared with the debtor or unsecured creditors, who would prefer a going concern and, thus, avoid enforcements. Although the insolvency code of the US and Canada is quite similar, Canadian law is more creditor orientated. The debtor-in-possession is controlled to a higher extend (CCAA §11.05 and BIA §69ff). Unlimited stay (CCAA §11.02) and super-priority-financing have to be permitted by court. However, favors are not granted to one homogeneous group but to creditors in general as the enforcement of claims is avoided by the stay. Negotiations might be rather complex in Canada and the tendency to formal proceedings high. Completeness of contracts might also influence resolution intensities. An inclusion of informal proceedings in loan contracts as in the US seems to accelerate resolution processes. Negotiations follow a more prepackaged course and, thus, tendencies to informal workouts are high. In Canada, resolution might be further slowed down by the involvement of courts in informal proceedings.

Differences among the countries are also apparent regarding the course of the non-cumulative baseline hazard rates. In Great Britain, the baseline hazard is slightly rising during the first five years indicating an increasing resolution tendency with the time spend in resolution. I.e., the longer a loan spends in resolution the higher is its intensity of resolution. As loans have to be resolved at some point in time, this meets the economic intuition. However, the baseline hazard rate seems to decrease slightly in the US, i.e., the longer a defaulted loan is already in resolution the lower is its future tendency of resolution. This might reflect the high tendency to informal proceedings immediately after default in the US. The longer a loans stays in resolution, the lower this tendency might be – leading to lower resolution intensities the longer a loan stays in resolution. Furthermore, there is a peak 18 months after default. This might be caused by loans which directly entered Chapter 11. As the data base contains both – formal and informal workouts – the baseline hazard rate displays the average baseline intensity across proceedings. Under Chapter 11 (§1121), the maximum timespan for a debtor to file a restructuring plan is set to 18 months after default. I.e., 18 month after default the latest, a plan has to exist regulating the details of the restructuring procedure. If it succeeds, the debtor may exit common resolution mechanisms. If it fails, the debtor might enter Chapter 7 and will be liquidated. In Canada, the baseline hazard rate is rising during the first year after default. The rather low baseline tendency directly after default might reflect the dispersion of bargaining powers and the associated complex negotiations in informal proceedings. Thereafter, the baseline hazard rate is decreasing. This might display loans which entered formal insolvency proceedings.

Macroeconomic Conditions

In this section, we investigate the influence of macroeconomic conditions as a factor of synchronous resolution processes. Therefore, we estimate Model II on country subsets and add macroeconomic variables to the regression.

Table 3.6 summarizes the country specific estimation results. Compared to the results of Model I only minor differences regarding loan specific impacts arise (see Table 3.5). Where sign switches occur either the parameter estimates of Model I or II are not statistically significant. Statistically significant parameter estimates show the expected sign.

In general, it is a challenging task to choose macroeconomic variables for the analysis of the resolution process. First, good proxies for the economic environment need to be found and, second, the number of macroeconomic factors that are simultaneously taken into account have to be set reasonably. Hereby, a trade-off between parsimony and additional goodness of fit occurs. We examine twelve different macroeconomic variables for our analysis, starting with regressions only including a single variable at a time. Based on these results, we try various combinations of macroeconomic variables[14] which are simultaneously included in the regression. This procedure leads to a model including five macroeconomic factors and exhibiting the highest additional fit according to goodness of fit measures (see Table 3.6).

The industry production is a significant accelerator of resolution processes for each country. Good economic conditions (measured by high growth values for the industry production) increase the tendency of resolution and, thus, accelerate resolution processes.

[14] Variables which show significance on individual basis were preferred for the model selection including more than one variable.

Table 3.6: Regression results for Model II

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
log(EAD)		-0.0478	***	(0.0071)	-0.0779	***	(0.0077)	-0.0223	*	(0.0118)
Asset Class (SME)	Large Corporates	0.2384	***	(0.0382)	0.2349	***	(0.0501)	-0.1100	*	(0.0601)
Facility	Short term	0.1423	***	(0.0412)	-0.1392	***	(0.0310)	-0.0946	**	(0.0446)
(Medium term)	Other	-0.0641	*	(0.0335)	0.2942	***	(0.0547)	0.0742		(0.0584)
Seniority	Super senior	0.1809	***	(0.0488)	0.7605	***	(0.0391)	0.6377	***	(0.1007)
(Pari-passu)	Non senior	0.4018	**	(0.1582)	0.7276	***	(0.2353)			
	Unknown	0.1869	*	(0.0963)				0.7243	***	(0.1216)
Nature of default	Unlikely to pay	-0.2969	***	(0.0425)	-0.0683		(0.0524)	-0.5796	*	(0.3081)
(90 days past due)	Bankruptcy	0.1285	*	(0.0754)	-0.0925	*	(0.0520)	-0.3706		(0.3026)
	Charge-off / provision	0.3661	***	(0.1064)	-0.1757	***	(0.0453)	0.8387	**	(0.3717)
	Sold at material credit loss	1.6841	***	(0.1030)						
	Distressed restructuring	0.3014	**	(0.1533)	-0.2382	**	(0.0948)			
	Non accrual	-0.0249		(0.0367)	-0.1554	***	(0.0452)	-0.4495		(0.2941)
	Unknown	-0.1108		(0.1288)	0.8508	***	(0.0807)	-0.8324	***	(0.2988)
Guarantee	Unknown	0.4612	*	(0.2489)						
(NO)	YES	0.1141	***	(0.0289)	-0.1109	***	(0.0316)	0.1930	***	(0.0742)
Collateral	Other collateral	0.1072	***	(0.0327)	0.0906	**	(0.0377)	0.3943	***	(0.0799)
(NO)	Real estate	0.0649		(0.0466)	0.1115	***	(0.0373)	0.2238		(0.1889)
	Unknown	-1.7116	***	(0.2427)				1.3039	***	(0.1143)
Number of collateral		0.0279	***	(0.0084)	0.0034		(0.0024)	0.0095		(0.0067)
Cured (NO)	YES	0.5061	***	(0.0328)	0.9439	***	(0.0357)	0.9515	***	(0.0539)

Continued on next page

Table 3.6 (continued): Regression results for Model II

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
Industry	Agric., forestry, fishing	-0.2549	**	(0.1212)	0.1052		(0.0955)	-0.1070		(0.1011)
(Finance, insurance, RE)	Mining	-0.0780		(0.1417)	-0.3419		(0.2294)	0.5306	***	(0.1777)
	Construction	-0.3020	***	(0.0542)	-0.3592	***	(0.0605)	0.2011	**	(0.0971)
	Manufacturing	-0.2575	***	(0.0464)	0.0337		(0.0577)	0.2850	***	(0.0855)
	Transp., commu., sanitary services	0.0073		(0.0616)	0.1646	**	(0.0775)	0.2381	**	(0.1024)
	Wholesale and retail trade	-0.1313	***	(0.0486)	0.0009		(0.0514)	0.2457	***	(0.0837)
	Services	-0.1173	**	(0.0463)	0.0671		(0.0551)	0.2063	**	(0.0804)
	Unknown	0.2855	***	(0.0475)	0.1109		(0.0974)	0.0637		(0.2955)
Equity index		0.0495		(0.1117)	0.0242		(0.1694)	-0.3091	*	(0.1687)
Industry production		1.5835	***	(0.3177)	1.8201	*	(0.9372)	1.8491	**	(0.8651)
Volatility index		-0.0005		(0.0024)	-0.0043		(0.0028)	-0.0041		(0.0038)
Term spread		0.0439	**	(0.0171)	0.0920	*	(0.0486)	0.1171	**	(0.0546)
World Bank score		0.0256	***	(0.0064)	0.1988	***	(0.0103)	0.0293		(0.0228)
LL				-49,319			-40,317			-22,041
AIC				98,705			80,695			44,143
McFadden's adjusted R^2				0.0120			0.0305			0.0486
Cox & Snell's R^2				0.0240			0.0500			0.0488
Resolved Loans				6,153			5,355			2,942
Loans				7,133			5,780			4,482

Notes: The table summarizes regression results for country specific impacts of loan characteristics on the tendency of resolution. The model specification fulfill Equation (3.3), i.e., macroeconomic variables but no frailties are included. The first column contains covariate names and the second includes corresponding categories if the variable is of categoric nature. The reference category is given in parenthesis. Significance is indicated at 10% (*), 5% (**) and 1% (***). Standard errors (SE) are given in parenthesis. For completeness, results for a regression that uses all observations jointly are given in Appendix Table 3.B.3.

Stock market growth plays a minor role for resolution compared to industry production which already catches the general economic condition. We identify a significant but small effect for Canada which vanishes when including the frailty. Financial and monetary expectations are modeled by the volatility index and the term spread. The effect of the former is overlaid by the latter. In combination with the industry production as indicator for general economic conditions, the term spread captures the expectations of long-term economic conditions and is, therefore, important for resolution processes. In the US and Great Britain, low expectations are identified to result in significant lower resolution tendencies which lead to decelerated resolution processes.

The measures for the goodness of fit – AIC, McFadden's adjusted R^2, Cox & Snell's R^2 – improve with regards to the US and Great Britain when including macroeconomic variables.[15] This suggests a systematic co-movement of DRTs caused by the macroeconomy in both countries. In Canada, we do not find such evidence. The World Bank score, i.e., the efficiency of insolvency regulations indicates an accelerated resolution process in the US and Great Britain but not significantly for Canada.

As stated in Section 3.4.1, we expect lower default resolution intensities in crises periods due to capacity limits in resolution devisions and a wait-and-see strategy of creditors in harsh economic surroundings. This seems to be true for the US and Great Britain. However, DRTs in Canada seem less influenced by the macroeconomy. This might be due to a lower crises susceptibility. The Canadian banking system is rather homogeneous and, therefore, said to be less affected by crises (see, e.g., Bordo et al., 2015). Furthermore, DRTs are generally rather high in Canada due to disperse bargaining powers and court involvement even in informal proceedings. The economic environment seems to influence the DRT to a lower extent.

In summary, the measured impacts are plausible and significant. However, various macroeconomic variables do not exhibit a significant impact when including them individ-

[15] For instance, McFadden's adjusted R^2 increases from 1.12% to 1.20% (US) and from 2.34% to 3.05% (Great Britain). Cox & Snell's R^2 changes from 2.23% to 2.40% resp. from 3.86% to 5.00%. In general, absolute values for these measures should be treated with care. It is important to note that an increase can be considered as favorable because this indicates a model better capturing given realizations, e.g., higher values for McFadden's adjusted R^2 imply an increase in the likelihood of a given model in comparison to a benchmark model.

ually (see Table 3.B.4) and the improvement for goodness of fit seems to be moderate even when including macroeconomic variables simultaneously. Bandopadhyaya (1994) identify similar issues for bankrupt US American firms when determining systematic variables for the time spent under Chapter 11. A study of Grunert and Weber (2009) detects no significant effects of the macroeconomy on LGDs of defaulted loans from German companies. A reason for this behavior may be ascribed to the complexity that appears when capturing systematic impacts on DRTs and LGDs. Observations for those are treated as if they were known at default date at which the condition of the macroeconomic environment can be observed. However, DRTs are influenced by the macroeconomic environment at default date and macroeconomic conditions after the default date during the resolution process, what is also mentioned by Grunert and Weber (2009). This makes the quantification of systematic effects on DRTs very complex and challenging as one needs to be aware of future macroeconomic conditions during default resolution to fully capture all systematic effects. In addition, this might explain why macroeconomic variables can only capture systematic effects on DRTs up to a certain degree.

Frailties as Unobservable Factors

Next, we estimate Model III and include stochastic time-dependent frailties. These capture stochastic co-movements of resolution intensities by common unobserved factors. Table 3.7 summarizes the regression results. The changes in parameter estimates for loan specific variables are minor and, therefore, not provided in detail.

The frailty can be investigated by its estimated volatility. In the US and Great Britain, the values are similar with around 0.30. In Canada, the frailty effect is considerably smaller with a volatility of 0.15 but still greater than 0. This is in line with earlier findings. Generally, systematic patterns seem to have less influence on the DRTs in Canada, compared to the US and Great Britain. Reasons may be found in a rather low crises susceptibility and in the fact that resolution intensities are rather low in Canada due to a high rate of court involvement even in informal proceedings. The model fit – measured by AIC, McFadden's adjusted R^2, and Cox & Snell's R^2 – improves for the US

Table 3.7: Regression results for Model III

	United States			Great Britain			Canada		
	Coef.		SE	Coef.		SE	Coef.		SE
	Loan specific variables are dropped in this presentation.								
Equity index	-0.2939		(0.2419)	0.4607	*	(0.2700)	-0.3109		(0.2389)
Industry production	3.3437	***	(0.9261)	-0.4466		(1.2467)	1.0775		(1.1402)
Volatility index	-0.0107		(0.0088)	0.0022		(0.0079)	-0.0023		(0.0059)
Term spread	0.1579	**	(0.0639)	0.0529		(0.1079)	0.0763		(0.0773)
World Bank score	0.0357		(0.0247)	0.2337	***	(0.0291)	0.0378		(0.0377)
Frailty volatility	0.3035	***	(0.0249)	0.2959	***	(0.0190)	0.1530	***	(0.0530)
LL			-49,179			-40,141			-22,016
AIC			98,453			80,370			44,108
McFadden's adjusted R^2			0.0130			0.0326			0.0472
Cox & Snell's R^2			0.0275			0.0548			0.0491
Resolved Loans			6,153			5,355			2,942
Loans			7,133			5,780			4,482

Notes: The table summarizes regression results of country specific impact of frailties on the tendency of resolution. The model specification fulfill Equation (3.5), i.e., with loan specific characteristics, macroeconomic information, and frailties. Significance is indicated at 10% (*), 5% (**) and 1% (***). Using the likelihood ratio test for the frailty where the null model is given by Model II. The standard error of the frailty is computed by bootstrapping with resampling and replacement for 10,000 steps. For completeness, results for a regression that uses all observations jointly are given in Appendix Table 3.B.3.

and Great Britain when including frailties.[16] In addition, we run a likelihood ratio test to check whether Model III increases the likelihood compared to Model II. The null hypothesis of no improvement is rejected with p-values of lower than 10^{-4}. Thus, the results show clear evidence for an improvement in all three countries. This indicates systematic dependencies among resolution intensities which can not be explained by covariates.

Next we analyse DRT changes due to varying frailty realizations to study the effect of unobservable factors. Starting from a realization of the systematic frailty factor u_0 the relative change of the expected DRT due to a change of the frailty Δu is given by

$$\begin{aligned}\frac{E(T|U_{\tilde{t}(i)} = u_0 + \Delta u)}{E(T|U_{\tilde{t}(i)} = u_0)} - 1 &= \frac{\lambda_{it|U_{\tilde{t}(i)}=u_0}}{\lambda_{it|U_{\tilde{t}(i)}=u_0+\Delta u}} - 1 \\ &= \frac{\lambda_{0t} \exp\left(x_i\beta + z_{\tilde{t}(i)}\gamma + u_0\right)}{\lambda_{0t} \exp\left(x_i\beta + z_{\tilde{t}(i)}\gamma + u_0 + \Delta u\right)} - 1 \\ &= \exp(-\Delta u) - 1, \end{aligned} \tag{3.6}$$

[16] For instance, McFadden's adjusted R^2 increases from 1.20% to 1.30% (US) and from 3.05% to 3.26% (Great Britain). Cox & Snell's R^2 changes from 2.40% to 2.75% resp. from 5.00% to 5.48%.

when assuming a constant baseline hazard rate for the definition of Model III in Equation (3.5). Table 3.8 shows the relative change for one standard deviation changes of the frailty. A decrease of the unobservable factor by one standard deviation increases the mean DRT by approximately 35% in the US and in Great Britain, whereas, a one standard deviation rise decreases the mean DRT by about 26%. In Canada, the impact is lower by 17% and -14%.

Table 3.8: Frailty impact on mean DRT

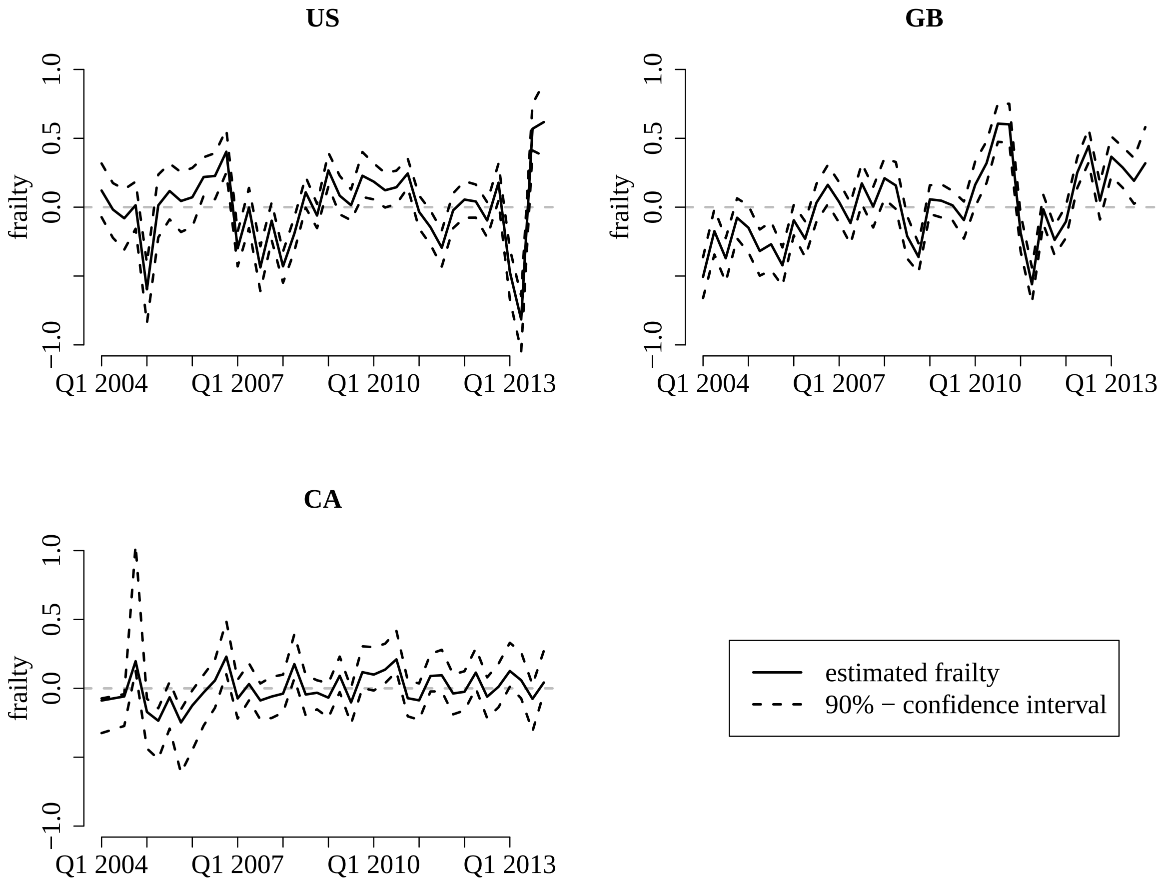
	US	GB	CA
$\Delta u = -1\sigma$	35.46%	34.43%	16.53%
$\Delta u = +1\sigma$	-26.18%	-25.61%	-14.19%

Notes: The table summarizes the impact of the frailty on the DRT as of Equation (3.6). The relative change of the mean DRT for a frailty change of one standard deviation is displayed.

Generally, frailties represent the development of the systematic components over time. Figure 3.8 illustrates the curves of the estimated frailty realizations over time. Country specific differences arise. In the US, we can see a decline in the crises years 2007 and 2008. This implies lower resolution rates for loans defaulting at the beginning of the Global Financial Crisis and, thus, longer resolution processes. The upwards shift for 2013 is caused by a low number of defaults in this year and is thus less meaningful. In Great Britain, the unobservable systematic component shows a different pattern and the crisis affects it with a delay and a weaker impact. The Canadian frailty is rather evenly spread. The correlations between country frailties show a link of systematic resolutions in Great Britain and Canada with a correlation coefficient of 0.418 which is significantly different from zero with a p-value of 0.008. Neither the link of the US to Great Britain is significantly different from zero (ρ=0.026, p-vlaue=0.873) nor to Canada (ρ=0.098, p-vlaue=0.551). This shows that the unobservable systematic risk strongly depends on country specific patterns.

There are country specific changes in the significance of macroeconomic variables after including the frailty. In the US, the effect of the industry production and the term spread are still significant and even more pronounced. For Great Britain, the significance of both previously significant variables vanishes. However, after including all systematic factors

Figure 3.8: Time-dependent frailties as systematic components of resolution



Notes: The figure illustrates the course of estimated frailties over time for the US, Great Britain, and Canada. The solid black lines displays the frailty, whereas, the dashed black line shows the 90%-confidence interval. The confidence interval is computed by bootstrapping with resampling and replacement for 10,000 steps. We check the assumption of normal distribution by a Kolmogorov–Smirnov test. As required, the null hypotheses of normal distribution is not rejected for each country with p-values of 0.371 (US), 0.982 (GB) and 0.319 (CA). For completeness, the estimated frailty for a regression that uses all observations jointly are given in Appendix Figure 3.B.2.

the stock market is the only macroeconomic variable identified as significant. The picture is different for Canada where we do not identify any observable systematic risk factor as significant trigger. The entire systematic risk in Canada is driven by the frailty.

The frailty represents unobservable systematic effects, but can also be triggered by time-varying influences of loan- and borrower-specific as well as macroeconomic variables.[17] For instance, industry-specific effects may hold in recessions for certain industries, but not in expansions. In order to check the robustness of unobservable systematic effects in this context, we studied the divergence of parameter estimates and frailties for

[17]We would like to thank an anonymous referee for pointing this out.

periods of economic recessions and expansions as defined by country-specific recession dates of the Organisation for Economic Co-operation and Development (OECD). In the Appendix, regression results and frailties for a modification of Model III that includes interactions of all covariates to recessions at default date (Figure 3.B.1, Table 3.B.1 and Table 3.B.2) are shown. Some parameter estimates significantly change in recession. However, the course and the volatility of the frailty are not substantially divergent, i.e., the measured unobservable systematic effects seem to be not triggered by time-varying covariates.

In summary, the systematic co-movement of resolution processes can only partially be explained by macroeconomic variables and the frailty represents a more important systematic component.[18] This observation leaves us with implications regarding the occurrence of observed DRTs and with respect to risk quantification and forecasting. Creditors are in need to determine risk of loans that default in the future. First, our results show that systematic factors play an important role with this respect. However, only using observable factors may not be enough as there seems to be some kind of systematic behavior among defaulted loans and their resolution which can not be captured by contemporaneous macroeconomic observations. A reason for this may be due to the fact that DRTs do not only depend on contemporaneous but also on future conditions of the economic environment. Such future conditions are unobservable from todays perspective which at first sight may seem disencouraging from a practitioners perspective. Nevertheless, simply being aware of this attribute may improve risk assessment. In addition, the estimation of frailty observations through Model III can be used to derive conservative forecasts for future DRTs. Given the estimated frailty volatility, we can define a critical state of the systematic environment, e.g., the 5% quantile of frailty distribution. Determining DRTs under such a hypothetical critical scenario provides us insights what we need to expect during critical conditions regarding possible realizations for DRTs. This type of analysis is not only relevant for the creditor itself, but also for regulators who nowadays often

[18]This is in line with the study of Khieu et al. (2012) on loan LGDs that identifies systematic effects in addition to macroeconomic influences by significant year dummies in its regressions. The literature on the correlation between the probability of default and the LGD also mentions systematic effects on LGDs (e.g., Düllmann and Trapp, 2004; Altman et al., 2005).

demand risk forecasts under downturn or stressed economic scenarios.[19]

The Relationship between DRTs and Recovery Rates

As already indicated by the descriptive analysis at the end of Section 3.2, there seems to be a positive relationship between DRTs and recovery rates. In order to examine this in more detail, additional analyses are conducted in which we change the point of view and use recovery rates as dependent variables, and, among others DRTs as an explanatory variable. First, a linear regression model is applied, second, we apply a logistic regression differentiating between no losses and losses. Different set ups regarding the inclusion of loan specific and macro variables are employed. Under each set up, we find DRTs to have a positive significant impact on recovery rates. This indicates that loans with longer DRTs are more likely to come along with higher losses. While such an analysis is of no use for forecasting purposes as both variables are unknown at the time of default, it emphasizes the relationship between DRTs and recovery rates and may motivate future research to model both variables simultaneously.[20]

3.5 Implications of Systematic DRTs

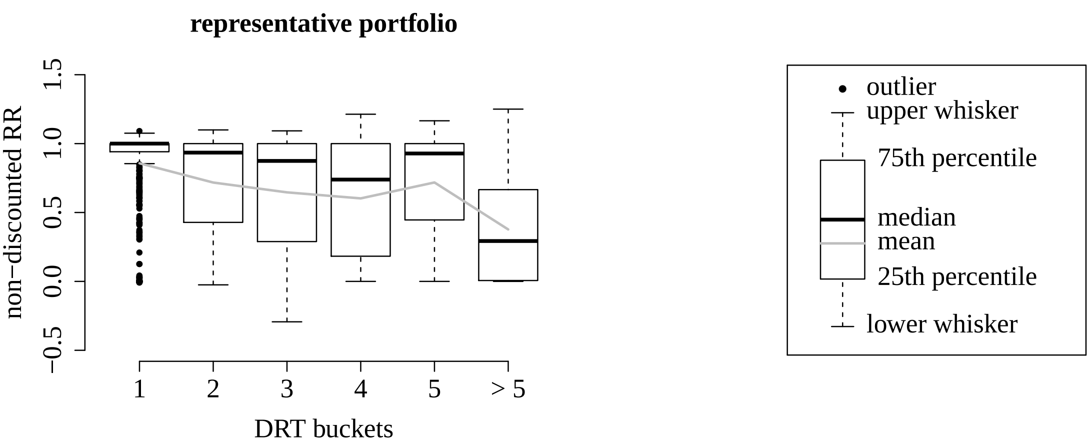
Creditors need to determine the risk of their portfolio for their internal risk assessment as well as for pricing and regulatory purposes. As stated earlier, DRTs represent a possible stochastic risk factor for the overall credit loss which may be modeled according to approaches presented in this paper. The better the model captures the nature of DRTs and their occurrence, the more precise their risk assessment should be. Results from the previous sections give rise to the conjecture that underestimating or neglecting systematic impacts (observable and especially unobservable) of DRTs may lead to a false or poor evaluation of those. This section uses a representative portfolio and a simulation analysis to show the possible extent of this misspecification.

[19]For instance, risk assessment under stressed market conditions is demanded for market risks under Basel III (Basel Committee on Banking Supervision, 2011) and under economic downturn conditions for loss given defaults (Basel Committee on Banking Supervision, 2005; Board of Governors of the Federal Reserve System, 2006).

[20]All additional analyses are available from the authors upon request.

The representative portfolio consists of 1,000 US American loans which are randomly sampled from our data set. Firstly, we consider one of these loans to derive implications of systematic factors on loan level. Secondly, we extend the analysis to portfolio level by including all of the 1,000 loans. The relation between DRTs and non-discounted RRs is explicitly taken into account in the second part of the analysis. Figure 3.9 shows the relation for the randomly sampled, representative portfolio. Compared to the complete data sample only minor differences arise (see Figure 3.2). Overall, the mean as well as the median of the non-discounted RRs follow a decreasing course.

Figure 3.9: Relation of default resolution time and non-discounted RR (representative sample)



Notes: The figure illustrates the relation of the DRT and non-discounted RR for the representative sample which is adapted in the impact study. Box plots of the non-discounted RR per bucket of DRT for the US, Great Britain, and Canada are displayed. The first bucket (marked with 1 on the x-axis) includes loans with DRTs up to one year. The second bucket (marked with 2 on the x-axis) includes loans with DRTs longer than one year up to two years and so on. In the last bucket (marked with > 5), loans with DRTs greater than five years are summarized. The black horizontal lines within the box plots mark the medians. The means are separately displayed by gray lines.

3.5.1 Implications on Loan Level

Consider a single loan with resolution intensity according to Model I (λ^I) which is time constant because the linear predictor of the loan specific variables ($x\beta$) is constant over time. In the Cox Model, the time to an event follows an exponential distribution with rate parameter λ if a constant baseline hazard rate λ_0 is assumed.[21] Thus, the probability density function of the DRT in Model I is determined by

$$f_T^I(t) = \lambda^I \exp(-\lambda^I t), \qquad t \geq 0. \tag{3.7}$$

In contrast, the resolution intensity in Model II depends on the default time $\tilde{t}$ of the considered loan. Therefore, the resolution intensity might be lower in recessions and higher in expansions depending on the linear predictor of the macroeconomic variables $(z_{\tilde{t}}\gamma)$. Given the default time $\tilde{t}$ of the loan, the resolution intensity of Model II is fully specified because the realizations of the macroeconomic variables are known at time of default. The DRT in Model II is, therefore, exponentially distributed with a constant rate parameter $\lambda^{II}(\tilde{t})$ for a given time of default $\tilde{t}$ and its probability density function is

$$f_{T,\tilde{t}}^{II}(t) = \lambda^{II}(\tilde{t})\ \exp(-\lambda^{II}(\tilde{t})\ t), \qquad t \geq 0. \tag{3.8}$$

As the resolution intensity of Model II varies over calendar time, longer DRTs might arise during weak economic conditions and shorter DRTs in a favorable environment.

In Model III there is not such a simple expression for the probability density function of the DRT as in Model I and II as the realization of the frailty is unknown at the time of default. Conditioning on the frailty factor $U = u$, the conditional intensity of Model III $\lambda^{III}(\tilde{t}, u)$ is constant, given the quarter of default $\tilde{t}$. Thus, the conditional probability density of the DRT is determined by

$$f_{T,\tilde{t}|U=u}^{III}(t) = \lambda^{III}(\tilde{t}, u)\ \exp(-\lambda^{III}(\tilde{t}, u)\ t), \qquad t \geq 0. \tag{3.9}$$

The unconditional probability density function can be derived by the integral of the joint

[21]To check for robustness, we derive the simulation also with the estimated time varying hazard rates following Bender et al. (2005) and receive similar results. We would like to thank an anonymous referee for this remark.

probability density function over the frailty realizations u

$$f^{III}_{T,\tilde{t}}(t) = \int_{-\infty}^{+\infty} f^{III}_{T,\tilde{t}|U=u}(t)\ f_U(u)\ du, \qquad t \geq 0, \tag{3.10}$$

where $f_U(u)$ is the density of the Normal distribution with mean 0 and variance σ^2 (see Equation (3.4)). Equation (3.10) can be solved by numerical integration.

As the baseline hazard rate λ_0 directly impacts the distribution of DRTs and, thus, its mean, we calibrate it on the average DRT of 1.59 years (see Table 3.1). This ensures an average simulated portfolio DRT in accordance with the empirical data. Thus, the average portfolio DRT corresponds to 1.59 years for Model I. Regarding Model II and III, it amounts to 1.59 years in an average economic scenario. The simulated DRTs might be higher relating to recessions and lower in expansions. The calibration yields in a baseline hazard rate for Model I of $\lambda_0^I = 1.08$ as well as $\lambda_0^{II} = 0.12$ for Model II and $\lambda_0^{III} = 0.07$ for Model III.[22]

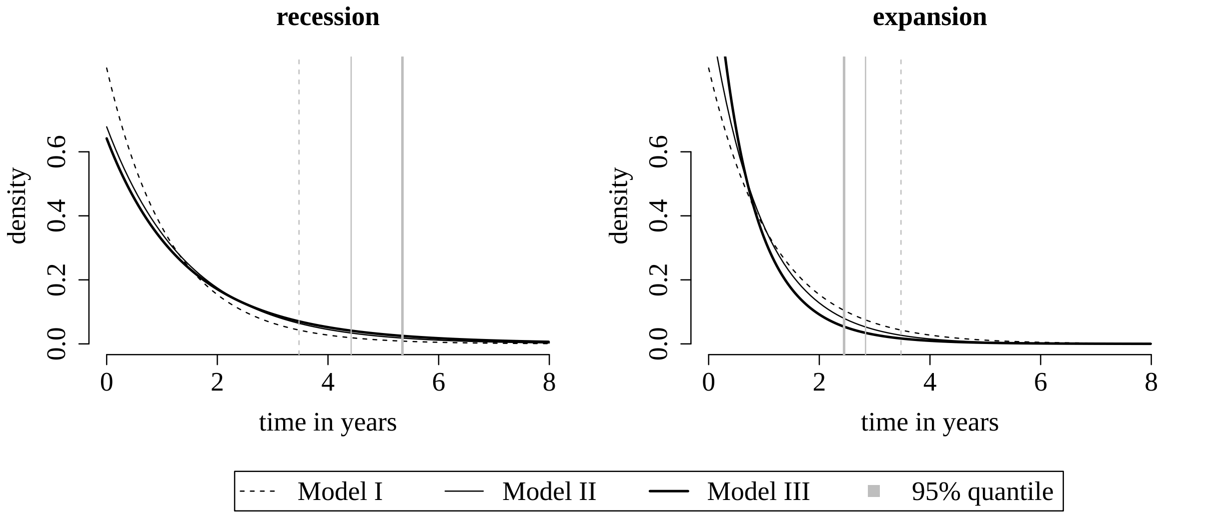
Figure 3.10 shows the probability density functions of the DRT in Model I, II, and III as of Equation (3.7), (3.8), and (3.10) for an exemplary recession and expansion period.[23] The left panel of Figure 3.10 displays the probability density functions for a recession period. The underlying quarter (Q1 2009) is shaped by the Global Financial Crisis and includes inter alia the crash of Lehman Brothers. Compared to Model I, the density of Model II is lower for short DRT and higher for longer ones. The distribution is, thus, shifted towards higher DRTs. This tendency is even more pronounced considering Model III as the frailty intensifies the impact of the economic surrounding. Firstly, an unobservable systematic factor widens the distribution of DRT. Secondly, impacts of the observable systematic factors are enhanced due to the consideration of the frailty. The right panel of Figure 3.10 shows the probability density functions for an expansion period. Considering favorable economic surroundings, opposite effects appear. The distribution of DRT for Model II is shifted towards lower values compared to Model I. Table 3.9

[22]The deviations in the baseline hazard rates among the models seems adequate as the difference in levels also emerges in the estimation of the models.

[23]The realizations of the macroeconomic variables are assumed to match their values as of Q1 2009 for the recession and Q2 2011 for the expansion period.

summarizes the median and 95% quantile of the distributions. Whereas the difference is less pronounced in the median, it is apparent considering the 95% quantile. In a recession period, there is an increase of this quantile by 54% comparing Model I and III.

Figure 3.10: Density of DRT



Notes: The figure illustrates the probability density function of the DRT for Model I, II, and III according to Equation (3.7), (3.8), and (3.10) in an exemplary recession (realizations of macroeconomic variables as of Q1 2009) and expansion (realizations of macroeconomic variables as of Q2 2011) period. Under the assumption of constant baseline hazard rates, the DRTs of Model I and II follows an exponential distribution with rate parameter λ^I for Model I and λ^{II} for Model II. The density of Model III is derived by numerical integration.

Table 3.9: Inferences of systematic factors on the distribution of DRTs

		Recession	Expansion
Model I	mean	1.16	1.16
	95% quantile	3.48	3.48
Model II	mean	1.47	0.95
	95% quantile	4.42	2.84
Model III	mean	1.70	0.78
	95% quantile	5.35	2.45

Notes: The table summarizes the mean and 95% quantile of the DRT for Model I, II, and III according to Equation (3.7), (3.8), and (3.10) in an exemplary recession (realizations of macroeconomic variables as of Q1 2009) and expansion (realizations of macroeconomic variables as of Q2 2011) period. The values arise from the probability density functions illustrated in Figure 3.10.

Generally, the distribution of DRTs for Model I is independent of the economic surrounding at the time of default. In Model II, favorable economic conditions shift the distribution towards lower values, adverse economic conditions shift it towards higher values indicating shorter DRTs in expansions and longer ones in recessions. This effect is enhanced in Model III.

3.5.2 Implications on Portfolio Level

Systematic effects in modeling DRT might not only affect the DRT itself, but also the loss involved. In Section 3.2 (see Figure 3.2), the relation between DRTs and the non-discounted RRs has been shown.[24] This indicates that recovery cash flows are lower the longer the resolution process takes. Furthermore, the DRT directly enters the calculation of the LGD by discounting the recovery cash flows. To put it simple, assuming a constant risk adjusted interest rate of 5% and recovery cash flows being paid at the end of the resolution process, the LGD of a single loan is derived as

$$\mathrm{LGD} = 1 - \frac{RR_T}{(1+r)^T}, \tag{3.11}$$

where RR_T denotes the time dependent non-discounted RR. Its value is set to the mean of the related DRT bucket. Table 3.10 summarizes the six DRT buckets and the corresponding means. For example, a loan with a DRT of two years is assigned with a non-discounted RR of 72.72%.

Table 3.10: Non-discounted RR by DRT buckets

DRT bucket	non-discounted RR
$0 < \mathrm{DRT} \leq 1$	84.23%
$1 < \mathrm{DRT} \leq 2$	72.72%
$2 < \mathrm{DRT} \leq 3$	62.80%
$3 < \mathrm{DRT} \leq 4$	59.82%
$4 < \mathrm{DRT} \leq 5$	59.09%
$\mathrm{DRT} > 5$	43.70%

Notes: The table summarizes the mean of the non-discounted RR per bucket of DRTs. The first row of the table ($0 < \mathrm{DRT} \leq 1$) includes loans with DRTs up to one year. The second row ($1 < \mathrm{DRT} \leq 2$) includes loans with DRTs longer than one year up to two years and so on. In the last row ($\mathrm{DRT} > 5$) loans with DRTs longer than five years are summarized. The means meet the ones illustrated in the upper left panel of Figure 3.2.

[24] Figure 3.9 shows that this relation also holds for the representative portfolio.

We further study the representative portfolio of the randomly sampled 1,000 loans by considering implications on portfolio level. The exposure weighted portfolio loss distribution is generated via Monte-Carlo simulation. DRTs for the 1,000 loans are randomly drawn according to Model I, II, and III, respectively.[25] The corresponding LGDs are calculated by Equation (3.11). Finally, the portfolio loss is given by:

$$LGD_{PF} = \frac{1}{n\,\overline{EAD}} \sum_{i=1}^{n} \left(LGD_i \; EAD_i \right), \tag{3.12}$$

where, $\overline{EAD}$ indicates the average EAD of the portfolio. The procedure is repeated 100,000 times to generate the portfolio loss distribution.

Table 3.11 summarizes the mean and the 95% quantile of the simulated portfolio DRTs. As the baseline hazard rates are calibrated on the empirical mean of the DRTs (see Table 3.1), the average portfolio DRT in Model I corresponds to this value for both economic scenarios. In Model II, the mean is higher in a recession and lower in an expansion period. This effect is more pronounced in Model III.

Table 3.11: Inferences of systematic factors on the distribution of portfolio DRTs

		Recession	Expansion
Model I	mean	1.59	1.59
	95% quantile	1.69	1.69
Model II	mean	2.01	1.29
	95% quantile	2.14	1.37
Model III	mean	2.42	1.10
	95% quantile	3.83	1.75

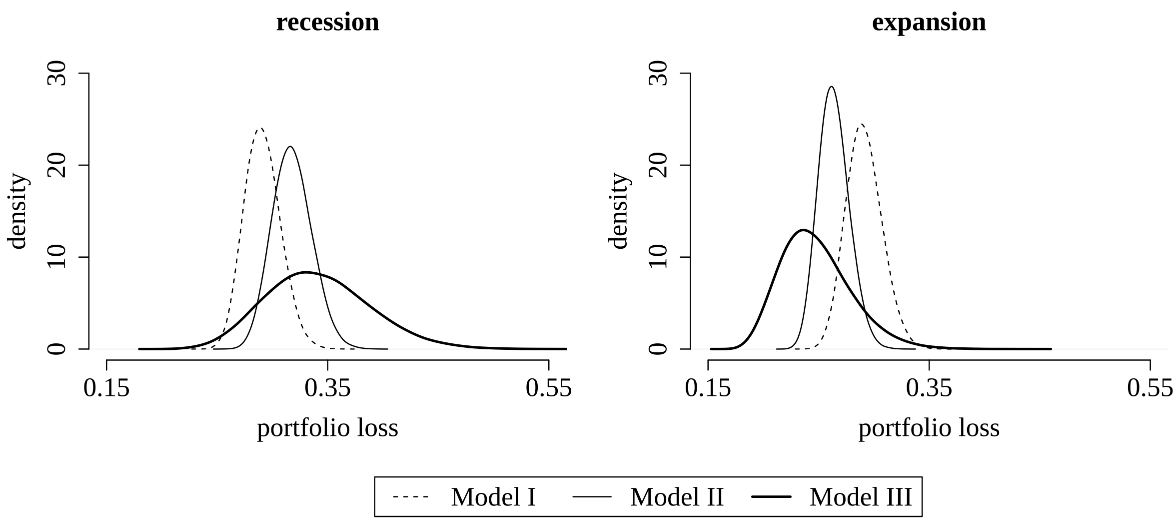
Notes: The table summarizes the mean and 95% quantile of the portfolio DRT for Model I, II, and III in an exemplary recession (realizations of macroeconomic variables as of Q1 2009) and expansion (realizations of macroeconomic variables as of Q2 2011) period. For every loan in the representative portfolio a DRTs is drawn according to the underlying model. Afterwards, the mean of the random draws is calculated to generate the average portfolio DRT. The procedure is repeated 100,000 times to generate the distribution of portfolio DRTs.

The portfolio loss distribution is simulated based on the portfolio DRTs and the non-discounted RRs as of Table 3.10. Figure 3.11 displays the portfolio loss distributions for

[25] In Model III, we initially draw a frailty from the Normal distribution with mean 0 and variance σ^2. This frailty realization u is constant for all 1,000 loans. Given this realization, the resolution intensity is constant among the loans in the homogeneous portfolio and we then draw the DRTs from the conditional distribution of the DRT, i.e., $T, \tilde{t} \mid U{=}u \sim \mathrm{Exp}(\lambda^{III}(\tilde{t}, u))$.

Model I, II, and III for the exemplary recession and expansion period. In the left panel the portfolio loss distribution of a recession is shown. Compared to Model I, the portfolio loss distribution of Model II is shifted to the right and slightly wider. This indicates that not only the mean of the portfolio loss but also its variation increases compared to Model I. This is mainly due to the exponential distribution of the DRT. Since it is fully specified by one parameter, mean and variance of the DRT are solely driven by this parameter and, thus, move in parallel. This effect is also reflected in the portfolio loss. However, the difference to Model III is much more pronounced than the difference between Model I and II. Through the frailty effect substantially more uncertainty is introduced into the model and the portfolio loss distribution is characterized by a higher mean and a much wider range. This indicates that not only the expected loss but also extreme quantiles of the portfolio loss distribution rise. In the right panel of Figure 3.11, the portfolio loss distribution of an expansion is displayed. The portfolio loss distribution of Model II

Figure 3.11: Kernel density estimates of loss on portfolio level



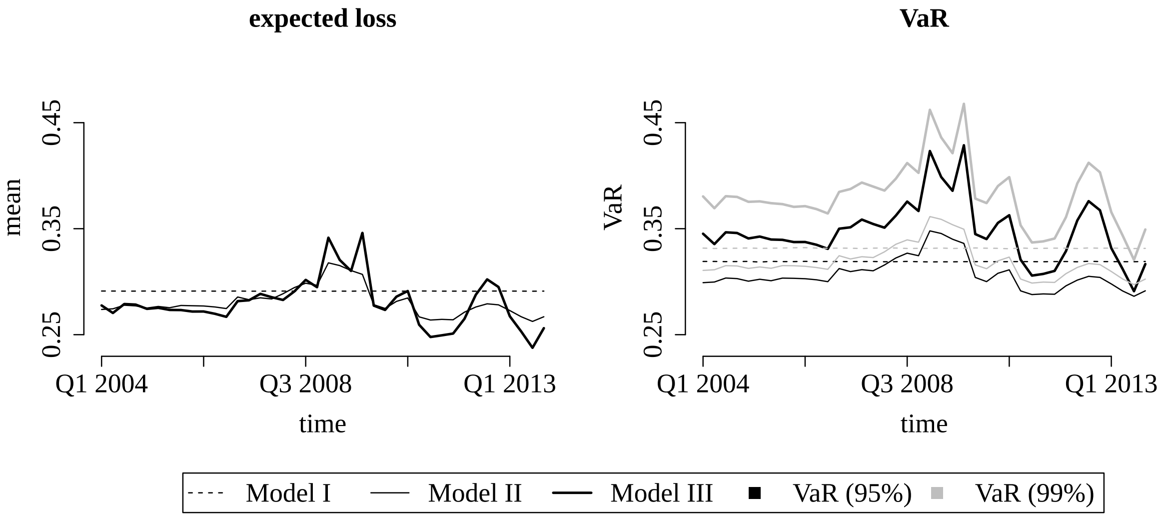
Notes: The figure illustrates the kernel density estimates of the exposure weighted portfolio loss distribution based on simulated DRT of Model I, II, and III as of Equation (3.7), (3.8), and (3.9) in an exemplary recession (realizations of macroeconomic variables as of Q1 2009) and expansion (realizations of macroeconomic variables as of Q2 2011) period. For every loan in the representative portfolio a DRTs is drawn according to the underlying model. In Model III, a frailty is drawn from the Normal distribution with mean 0 and variance σ^2 for each run. The corresponding loss is calculated. Afterwards, the mean of the losses is calculated to generate the average loss. The procedure is repeated 100,000 times to generate the distribution of portfolio losses.

is shifted to the left and is narrower compared to Model I. Comparing Model I and II with Model III, major differences arise. Although the distribution is shifted to the left of Model II, its wide range persists. The expected loss of Model III is lower compared to Model I and II. However, extreme quantiles are still higher.

While the former analysis considered exemplary portfolio loss distributions in a recession and expansion period, we now extend it to all possible scenarios in the estimation sample to analyze potential portfolio risk if similar scenarios arise in the future. Figure 3.12 displays the mean (left panel) and the the VaR(95%) as well as the VaR(99%)[26] (right panel) of the portfolio loss distribution for all macroeconomic scenarios in the estimation sample for the three models. In Model I, the mean is constant over time as the resolution intensity λ^I is constant. The mean of Model II lies above the one of Model I in quarters characterized by adverse economic conditions, e.g., Q1 2009. In favorable economic surroundings, e.g., Q2 2011, it lies below the one of Model I. Comparing Model II and III, the mean of Model III seems to be more extreme in the majority of the cases (e.g., Q1 2009 and Q2 2011). The right panel of Figure 3.12 shows the VaR(95%) and the VaR(99%) of the portfolio loss distribution. As the resolution intensity and, thus, the portfolio loss distribution is constant over time regarding Model I, the corresponding extreme quantiles are time-invariant. The course of the VaR(95%) and the VaR(99%) in Model II seems strongly related to the course of its mean. In recessions, the extreme quantiles of Model II lie above the ones of Model I, whereas, they lie below in expansions. This might be due to the rather similar shape of the portfolio loss distributions of Model I and II. Although, the range of the distribution of Model II slightly increases (decreases) if it is shifted to the right (left), the deviation seems marginal. A clearer contrast emerges considering Model III where the extreme quantiles are shifted upwards throughout. Generally, this shows that the stochastic frailty introduces non diversifiable systematic risk and co-movement between DRTs. This could have a substantial impact on losses on portfolio level.

[26] The VaR(95%) and VaR(99%) are the 95% and 99% quantiles of the portfolio loss distribution.

Figure 3.12: Mean and VaR(95%) of loss on portfolio level



Notes: The figure illustrates mean, VaR(95%) and VaR(99%) of the exposure weighted portfolio loss distribution based on simulated DRTs of Model I, II and III as of Equation (3.7), (3.8), and (3.9) for all quarters in the estimation sample. For every loan in the representative portfolio a DRTs is drawn according to the underlying model. In Model III, a frailty is drawn from the Normal distribution with mean 0 and variance σ^2 for each run. The corresponding loss is calculated. Afterwards, the mean of the losses is calculated to generate the average loss. The procedure is repeated 100,000 times to generate the distribution of portfolio losses.

3.6 Conclusion

This paper analyzes DRTs of defaulted loan contracts. The emphasis is laid on systematic effects among DRT intensities - both observable and unobservable. The observable systematic factors shift DRT intensities through time while the unobservable factors (frailties) lead to stochastic correlations.

We use access to a large data base and analyze DRTs of 17,395 loans located in the US, Great Britain, and Canada. Three models are taken into account, including loan specific; loan specific and macroeconomic; and loan specific, macroeconomic as well as unobservable variables. Our results show that unobservable factors impact default resolution intensities and that this influence remains when macroeconomic variables are additionally included in the model. The impact of systematic effects leads to more skewed distributions of DRTs. Thus, given good or adverse systematic conditions, a higher magnitude of more extreme DRTs occurs. An implication exercise shows that this can lead to higher credit risk regarding the credit portfolio loss distribution. In other words, neglecting systematic effects among DRTs might lead to a flawed and poor risk assessment of the credit portfolio.

We show that the DRT can be of great importance in direct and indirect ways. While it immediately impacts liquidity of financial institutions it also plays an important role with regards to credit costs, such as discounting costs and lower non-discounted RRs due to longer resolution processes. Hence, the analysis of DRT helps us in better understanding the occurrence of credit losses and, thus, improves risk assessments. Future research might lie in the development of credit risk models which simultaneously determine DRTs as well as default and loss given default estimates.

Appendix 3.A Estimation of the Cox Model

This section describes the theoretical background of the Cox proportional hazards model. First, we show a likelihood approach to estimate Model I and II, i.e., without frailties. Afterwards, we extend this by a time-dependent frailty.

From the definition of resolution intensities in Equation (3.1) it follows for Model I:

$$\lambda_{it} = \frac{f_T(t|x_i)}{1 - F_T(t|x_i)}, \tag{3.A.1}$$

($i = 1, \dots, n$), where $f_T(t|x_i)$ is the probability density function of the resolution time at t and $1 - F_T(t|x_i)$ is the probability that there is no resolution prior to time t.

The general likelihood for survival data is given by

$$L(\beta|x,\delta) = \prod_{i=1}^{n} \left[f_T(t_i|x_i)^{\delta_i} \left(1 - F_T(t_i|x_i)\right)^{1-\delta_i} \right], \tag{3.A.2}$$

with observed times after default t_i and censor indicators δ_i (1, if i was resolved at time t_i, 0 else). The first part describes the likelihood contributions of all resolved loans and the second part the contribution of the censored observations.

Inserting Equation (3.A.1) into Equation (3.A.2) yields

$$\begin{aligned} L(\beta|x,\delta) &= \prod_{i=1}^{n} \left[\lambda_{it}^{\delta_i} \left(1 - F_T(t_i|x_i)\right) \right] \\ &= \prod_{i=1}^{n} \left[\left(\lambda_{0t} \exp(x_i\beta)\right)^{\delta_i} \exp\left(- \exp(x_i\beta) \int_0^{t_i} \lambda_{0v} dv \right) \right]. \end{aligned} \tag{3.A.3}$$

The Cox model is a semi-parametric approach, i.e., the baseline rate λ_{0t} is not specified. Thus, Cox (1972) extends the likelihood of Equation (3.A.3) to

$$L(\beta|x,\delta) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \left[\frac{\exp(x_i\beta)}{\sum_{j=i}^{n} \exp(x_j\beta)} \lambda_{0t_i} \sum_{j=i}^{n} \exp(x_j\beta) \right] \prod_{i=1}^{n} \exp\left(- \exp(x_i\beta) \int_0^{t_i} \lambda_{0v} dv \right) \tag{3.A.4}$$

where the observations $i = 1, \dots, n$ are ordered so that $t_1 < \cdots < t_n$. For estimation of the unknown parameters the following partial likelihood is maximized

$$PL(\beta|x,\delta) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \frac{\exp(x_i\beta)}{\sum_{j=i}^{n}\exp(x_j\beta)}. \tag{3.A.5}$$

Afterwards, the baseline hazard rate can be estimated by

$$\widehat{\lambda}_{0t_i} = \left(\sum_{j=i}^{n}\exp(x_j\widehat{\beta})\right)^{-1}. \tag{3.A.6}$$

Model II additionally contains macroeconomic variables and the unknown parameter vector γ. Thus, the partial likelihood and the baseline estimate only changes by extending the linear predictor to $x_i\beta + z_{\tilde{t}(i)}\gamma$.

Including a frailty leads to higher computational effort because a frailty is unknown. For a more detailed description see Therneau et al. (2003). For Model III the conditional partial likelihood given fixed frailty realizations changes to

$$CPL(\beta,\gamma,\sigma|x,\delta,U=u) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \frac{\exp(x_i\beta + z_{\tilde{t}(i)}\gamma + u_{\tilde{t}(i)})}{\sum_{j=i}^{n}\exp(x_j\beta + z_{\tilde{t}(j)}\gamma + u_{\tilde{t}(i)})}, \tag{3.A.7}$$

where U denotes a vector including all frailties for all default times. Because the frailty realization u is unknown, we need to consider the conditional likelihood by integrating out the normally distributed frailty:

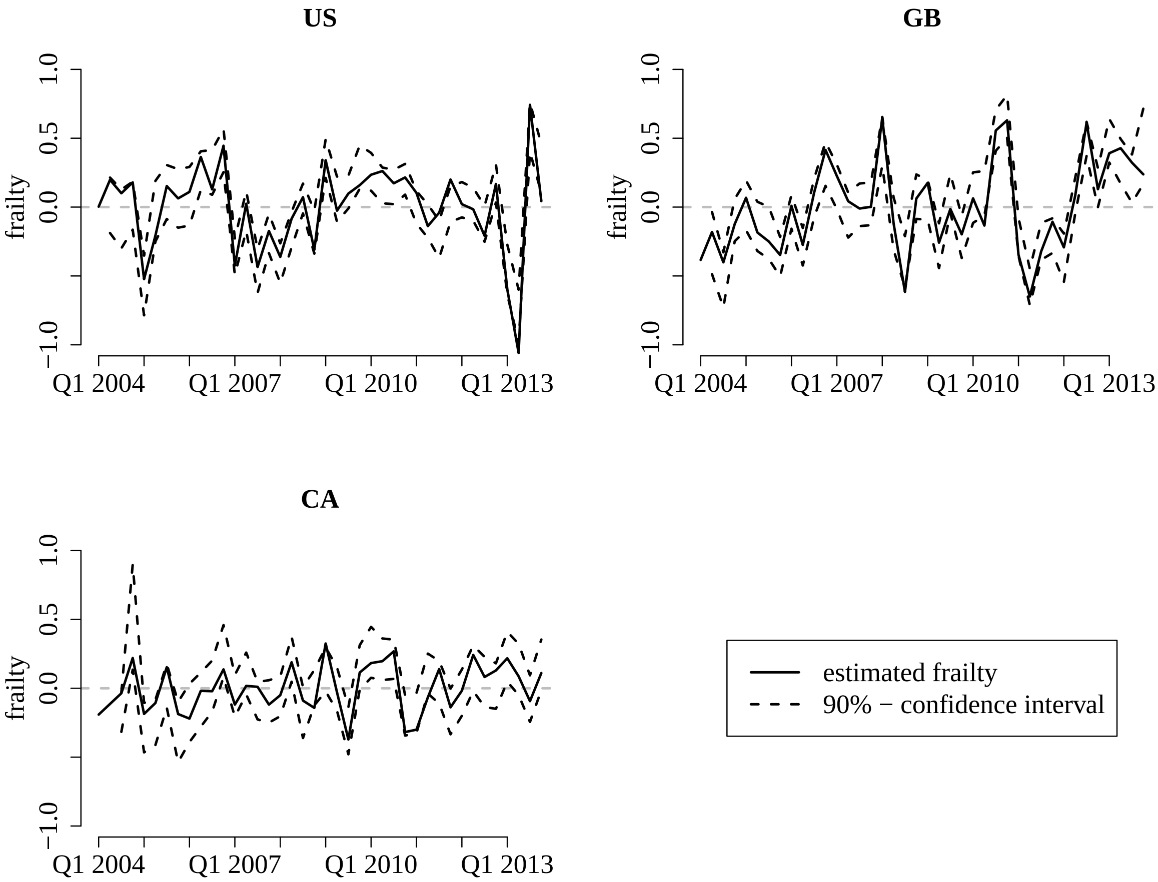
$$UPL(\beta,\gamma,\sigma|x,\delta) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \int_{-\infty}^{\infty} \frac{\exp(x_i\beta + z_{\tilde{t}(i)}\gamma + u)}{\sum_{j=i}^{n}\exp(x_j\beta + z_{\tilde{t}(j)}\gamma + u)}\ f_{U_{\tilde{t}(i)}}(u)\ du. \tag{3.A.8}$$

Equation (3.A.8) is maximized to estimate the unknown model parameters β,γ,σ. Afterwards, we use Equation (3.A.7) as $CPL(u|x,\delta,\widehat{\beta},\widehat{\gamma},\widehat{\sigma})$ to estimate the frailty vector. The baseline hazard rate is again estimated by Equation (3.A.6) extended by the frailty term.

Appendix 3.B Further Outputs

3.B.1 Covariate Influences over Time

Figure 3.B.1: Frailties when including interactions of covariates and recessions



Notes: The figure illustrates the course of estimated frailties over time for the US, Great Britain, and Canada. In the estimation covariate interactions to recessions are included. The solid black lines displays the frailty, whereas, the dashed black line shows the 90%-confidence interval. The confidence interval is computed by bootstrapping with resampling and replacement for 2,000 steps. We check the assumption of normal distribution by a Kolmogorov–Smirnov test. As required, the null hypotheses of normal distribution is not rejected for each country with p-values of 0.3044 (US), 0.9555 (GB) and 0.8235 (CA).

Table 3.B.1: Regression results for Model III when including interactions of covariates and recessions: parameter estimates non-recession

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
log(EAD)		-0.0668	***	(0.0086)	-0.0820	***	(0.0096)	-0.0071		(0.0143)
Asset Class (SME)	Large Corporates	0.2223	***	(0.0466)	0.3790	***	(0.0688)	-0.0994		(0.0728)
Facility	Short term	0.1591	***	(0.0483)	-0.0198		(0.0391)	-0.0979	*	(0.0526)
(Medium term)	Other	-0.0134		(0.0403)	0.4570	***	(0.0763)	0.0790		(0.0687)
Seniority	Super senior	0.1526	***	(0.0588)	1.0580	***	(0.0541)	0.6728	***	(0.1215)
(Pari-passu)	Non senior	0.3005		(0.1858)	1.0925	***	(0.3123)			
	Unknown	-0.1236		(0.1168)				0.7179	***	(0.1445)
Nature of default	Unlikely to pay	-0.3597	***	(0.0490)	-0.0949		(0.0667)	-0.5946	*	(0.3382)
(90 days past due)	Bankruptcy	0.1295		(0.1007)	-0.0243		(0.0676)	-0.3780		(0.3322)
	Charge-off / provision	0.4855	***	(0.1370)	-0.0986	*	(0.0574)	0.7535	*	(0.4317)
	Sold at material credit loss	1.8696	***	(0.1175)						
	Distressed restructuring	0.1434		(0.1824)	-0.5476	***	(0.1168)			
	Non accrual	-0.0774	*	(0.0439)	-0.1193	**	(0.0565)	-0.5467	*	(0.3213)
	Unknown	-0.4503	***	(0.1732)	0.8871	***	(0.0964)	-0.8556	***	(0.3315)
Guarantee	Unknown	0.4749		(0.3567)						
(NO)	YES	0.1752	***	(0.0344)	-0.1237	***	(0.0391)	0.1225		(0.0882)
Collateral	Other collateral	0.1675	***	(0.0394)	0.1098	**	(0.0476)	0.4236	***	(0.0951)
(NO)	Real estate	0.1104	**	(0.0543)	0.1758	***	(0.0464)	-0.0654		(0.2365)
	Unknown	-0.1382		(0.3201)				1.3079	***	(0.1362)
Number of collateral		0.0275	***	(0.0090)	0.0053	*	(0.0028)	0.0069		(0.0075)
Cured (NO)	YES	0.6652	***	(0.0388)	10.575	***	(0.0455)	1.0739	***	(0.0637)

Continued on next page

Table 3.B.1 (continued): Regression results for Model III when including interactions of covariates and recessions: parameter estimates non-recession

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
Industry	Agric., forestry, fishing	-0.3250	**	(0.1324)	0.2515	**	(0.1175)	-0.0904		(0.1178)
(Finance, insurance, RE)	Mining	-0.0603		(0.1601)	-0.2497		(0.2695)	0.5625	***	(0.2122)
	Construction	-0.2633	***	(0.0664)	-0.1474	*	(0.0790)	0.1770		(0.1163)
	Manufacturing	-0.2846	***	(0.0552)	0.2848	***	(0.0763)	0.1971	*	(0.1019)
	Transp., commu., sanitary services	0.0901		(0.0724)	0.1218		(0.1094)	0.2664	**	(0.1240)
	Wholesale and retail trade	-0.1612	***	(0.0569)	0.1321	*	(0.0685)	0.2023	**	(0.0990)
	Services	-0.0815		(0.0540)	0.0972		(0.0718)	0.2156	**	(0.0949)
	Unknown	0.1881	***	(0.0580)	0.3429	***	(0.1232)	0.0122		(0.3443)
Equity index		-0.4123		(0.2774)	1.1414	***	(0.3379)	-0.3421		(0.2997)
Industry production		3.6166	***	(1.0587)	-0.9898		(1.3909)	-0.9121		(1.3353)
Volatility index		-0.0134		(0.0117)	0.0355	***	(0.0104)	0.0193		(0.0123)
Term spread		0.1341	**	(0.0651)	-0.0144		(0.1165)	-0.0753		(0.0922)
World Bank score		0.0320		(0.0253)	0.2509	***	(0.0305)	0.0425		(0.0415)
Frailty volatility		0.2875	***	(0.0243)	0.3090	***	(0.0261)	0.1637	***	(0.0520)
LL				-49,102			-40,070			-21,980
AIC				98,363			80,287			44,093
McFadden's adjusted R^2				0.0139			0.0336			0.0474
Cox & Snell's R^2				0.0304			0.0574			0.0505
Resolved Loans				6,153			5,355			2,942
Loans				7,133			5,780			4,482

Notes: The table summarizes regression results of country specific impacts of covariates on the tendency of resolution. The model specification fulfill Equation (3.5), i.e., loan specific characteristics, macroeconomic information and frailties. In addition, the interaction between covariates and recession is taken into account. This table shows parameter estimates for periods of no recession. The interaction parameter estimates are shown in Table 3.B.2. Significance is indicated at 10% (*), 5% (**) and 1% (***), using the likelihood ratio test for the frailty where the null model is given by the model without frailty. The standard error of the frailty is computed by bootstrapping with resampling and replacement for 2,000 steps. Recessions are defined by the monthly recession dummy of the Organisation for Economic Co-operation and Development (OECD).

Table 3.B.2: Regression results for Model III when including interactions of covariates and recessions: interactions to recessions

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
log(EAD)		0.0320	**	(0.0158)	0.0164		(0.0163)	-0.0692	***	(0.0267)
Asset Class (SME)	Large Corporates	-0.0394		(0.0874)	-0.0634		(0.1040)	0.0291		(0.1385)
Facility	Short term	-0.1403		(0.0964)	-0.1865	***	(0.0667)	0.0409		(0.1031)
(Medium term)	Other	-0.1345	*	(0.0794)	-0.0327		(0.1180)	-0.0081		(0.1379)
Seniority	Super senior	0.2144	*	(0.1177)	-0.2620	***	(0.0892)	-0.2454		(0.2318)
(Pari-passu)	Non senior	0.5222		(0.3691)	-0.3479		(0.4836)			
	Unknown	0.8804	***	(0.2261)				-0.1100		(0.2783)
Nature of default	Unlikely to pay	0.1933	*	(0.1043)	0.0930		(0.1138)	0.0293		(0.8743)
(90 days past due)	Bankruptcy	0.2562		(0.1595)	0.2522	**	(0.1147)	0.0683		(0.8644)
	Charge-off / provision	0.2076		(0.2338)	0.0210		(0.0986)	0.2572		(1.0172)
	Sold at material credit loss	0.4918		(1.0149)						
	Distressed restructuring	0.7304	*	(0.4022)	0.7219	***	(0.2245)			
	Non accrual	0.1421		(0.0870)	0.1045		(0.1015)	0.3083		(0.8482)
	Unknown	0.7039	***	(0.2593)	0.3766	*	(0.1933)	0.6191		(0.8585)
Guarantee	Unknown	-0.0180		(0.5018)						
(NO)	YES	-0.2070	***	(0.0691)	0.1204	*	(0.0691)	0.2990	*	(0.1683)
Collateral	Other collateral	0.0535		(0.0867)	-0.0915		(0.0833)	-0.4795	**	(0.2021)
(NO)	Real estate	-0.0162		(0.1172)	-0.2436	***	(0.0852)	0.5530		(0.4082)
	Unknown	-2.7862	***	(0.4937)				0.1950		(0.2586)
Number of collateral		-0.0520	*	(0.0266)	0.0061		(0.0087)	0.3212	***	(0.0651)
Cured (NO)	YES	-0.4532	***	(0.0791)	-0.1891	**	(0.0775)	-0.3333	**	(0.1305)

Continued on next page

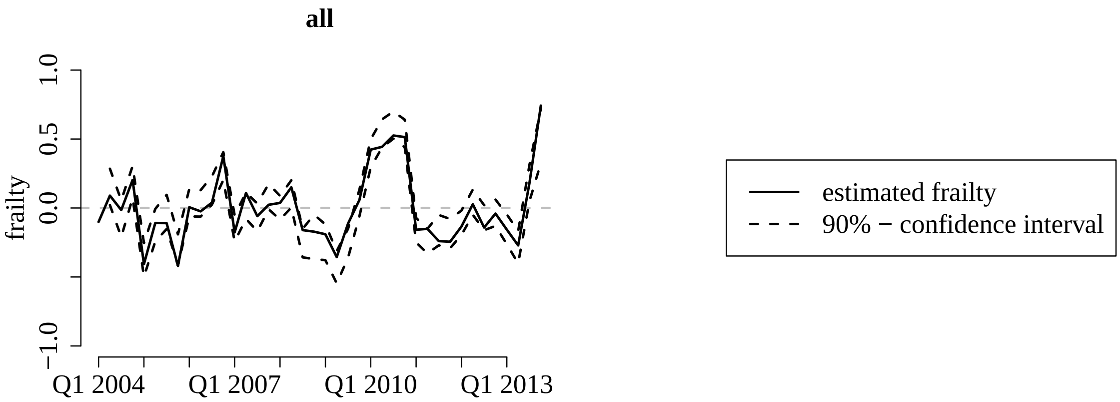
Table 3.B.2 (continued): Regression results for Model III when including interactions of covariates and recessions: interactions to recessions

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
Industry	Agric., forestry, fishing	0.8419	**	(0.3644)	-0.6409	***	(0.2193)	0.1436		(0.2407)
(Finance, insurance, RE)	Mining	-0.7570	*	(0.4007)	0.4475		(0.5364)	-0.1882		(0.3999)
	Construction	-0.0945		(0.1178)	-0.4977	***	(0.1264)	0.0346		(0.2147)
	Manufacturing	0.0974		(0.1071)	-0.5229	***	(0.1201)	0.2449		(0.1956)
	Transp., commu., sanitary services	-0.2918	*	(0.1495)	0.0497		(0.1585)	0.0060		(0.2297)
	Wholesale and retail trade	0.0671		(0.1155)	-0.3799	***	(0.1064)	0.0275		(0.1892)
	Services	-0.1159		(0.1136)	-0.1379		(0.1167)	-0.1046		(0.1814)
	Unknown	0.2604	**	(0.1076)	-0.7117	***	(0.2100)	-0.4007		(0.7251)
Equity index		1.5816	**	(0.7059)	-2.3853	***	(0.6126)	0.0621		(0.6974)
Industry production		-4.4882	**	(2.0843)	-1.4552		(2.9498)	6.0515	**	(2.8523)
Volatility index		-0.0026		(0.0217)	-0.0755	***	(0.0130)	-0.0293	**	(0.0149)
Term spread		0.1655		(0.2029)	0.0987		(0.1603)	0.4440	**	(0.2006)
World Bank score		-0.0050		(0.0056)	0.0208	***	(0.0041)	0.0102		(0.0104)

Notes: The table summarizes regression results of country specific impacts of covariates on the tendency of resolution. The model specification fulfill Equation (3.5), i.e., loan specific characteristics, macroeconomic information and frailties. In addition, the interaction between covariates and recession is taken into account. This table shows parameter estimates interaction of covariates and recessions. The parameter estimates for periods out of recessions are shown in Table 3.B.1. Significance is indicated at 10% (*), 5% (**) and 1% (***). Recessions are defined by the monthly recession dummy of the Organisation for Economic Co-operation and Development (OECD).

3.B.2 Joint Country Regression

Figure 3.B.2: Frailty for joint country regression



Notes: The figure illustrates the course of estimated frailty over time across countries. The solid black lines displays the frailty, whereas, the dashed black line shows the 90%-confidence interval. The confidence interval is computed by bootstrapping with resampling and replacement for 2,000 steps. We check the assumption of normal distribution by a Kolmogorov–Smirnov test. As required, the null hypotheses of normal distribution is not rejected with a p-value of 0.3751.

Table 3.B.3: Regression results for Model I, II and III across countries

		Model I			Model II			Model III		
		Coef.		SE	Coef.		SE	Coef.		SE
log(EAD)		-0.0285	***	(0.0046)	-0.0237	***	(0.0046)	-0.0269	***	(0.0047)
Asset Class (SME)	Large Corporates	0.0431	*	(0.0246)	0.0785	***	(0.0248)	0.1072	***	(0.0253)
Facility	Short term	-0.0717	***	(0.0212)	-0.0680	***	(0.0212)	-0.0707	***	(0.0213)
(Medium term)	Other	-0.0725	***	(0.0252)	-0.0995	***	(0.0252)	-0.0946	***	(0.0255)
Seniority	Super senior	0.4317	***	(0.0241)	0.4481	***	(0.0244)	0.4909	***	(0.0257)
(Pari-passu)	Non senior	0.5000	***	(0.1283)	0.5468	***	(0.1284)	0.5978	***	(0.1289)
	Unknown	0.3938	***	(0.0593)	0.3830	***	(0.0589)	0.4224	***	(0.0590)
Nature of default	Unlikely to pay	-0.1950	***	(0.0298)	-0.2365	***	(0.0298)	-0.2517	***	(0.0298)
(90 days past due)	Bankruptcy	-0.0349		(0.0374)	-0.0200		(0.0374)	0.0230		(0.0381)
	Charge-off / provision	-0.0499		(0.0381)	-0.0596		(0.0381)	-0.0206		(0.0385)
	Sold at material credit loss	1.6320	***	(0.0931)	1.6745	***	(0.0934)	1.8792	***	(0.0984)
	Distressed restructuring	0.1389	*	(0.0787)	0.0016		(0.0790)	-0.0444		(0.0808)
	Non accrual	-0.0827	***	(0.0249)	-0.0768	***	(0.0249)	-0.0646	***	(0.0250)
	Unknown	-0.4547	***	(0.0496)	-0.4786	***	(0.0496)	-0.3559	***	(0.0519)
Guarantee	Unknown	0.9537	***	(0.2158)	1.0700	***	(0.2159)	0.9109	***	(0.2167)
(NO)	YES	0.0291		(0.0194)	0.0257		(0.0193)	0.0260		(0.0195)
Collateral	Other collateral	0.0904	***	(0.0214)	0.1288	***	(0.0216)	0.1322	***	(0.0219)
(NO)	Real estate	0.1217	***	(0.0275)	0.1228	***	(0.0275)	0.1134	***	(0.0277)
	Unknown	1.0801	***	(0.0628)	1.1938	***	(0.0629)	1.1915	***	(0.0628)
Number of collateral		0.0033	*	(0.0020)	0.0031		(0.0020)	0.0039	*	(0.0022)
Cured (NO)	YES	0.7335	***	(0.0210)	0.7788	***	(0.0211)	0.7862	***	(0.0213)

Continued on next page

Table 3.B.3 (continued): Regression results for Model I, II and III across countries

		Model I			Model II			Model III		
		Coef.		SE	Coef.		SE	Coef.		SE
Industry	Agric., forestry, fishing	-0.1529	***	(0.0549)	-0.1964	***	(0.0549)	-0.1766	***	(0.0552)
(Finance, insurance, RE)	Mining	0.1914	**	(0.0960)	0.1253		(0.0961)	0.0965		(0.0965)
	Construction	-0.4465	***	(0.0367)	-0.4320	***	(0.0366)	-0.4020	***	(0.0370)
	Manufacturing	-0.0814	**	(0.0324)	-0.1068	***	(0.0324)	-0.0980	***	(0.0326)
	Transp., commu., sanitary services	0.0928	**	(0.0426)	0.0678		(0.0426)	0.0721	*	(0.0427)
	Wholesale and retail trade	-0.0380		(0.0309)	-0.0726	**	(0.0309)	-0.0736	**	(0.0311)
	Services	0.0231		(0.0313)	-0.0255		(0.0313)	-0.0280		(0.0315)
	Unknown	0.4950	***	(0.0393)	0.4675	***	(0.0394)	0.4788	***	(0.0398)
Equity index					0.0139		(0.0779)	-0.4188	***	(0.1358)
Industry production					0.9906	***	(0.2576)	-1.1185	***	(0.4291)
Volatility index					-0.0024		(0.0016)	-0.0098	**	(0.0046)
Term spread					0.0481	***	(0.0138)	0.1189	***	(0.0181)
World Bank score					0.0894	***	(0.0048)	0.1547	***	(0.0122)
Great Britain		0.8523	***	(0.0348)	1.2751	***	(0.0407)	1.5186	***	(0.0622)
United States		0.6538	***	(0.0366)	1.7938	***	(0.0687)	2.5488	***	(0.1464)
Frailty volatility								0.2601	***	(0.0197)
LL				-128,336			-128,096			-127,894
AIC				256,734			256,264			255,895
McFadden's adjusted R^2				0.0177			0.0195			0.0203
Cox & Snell's R^2				0.0310			0.0341			0.0359
Resolved Loans				14,472			14,472			14,472
Loans				17,420			17,420			17,420

Notes: The table summarizes regression results across countries of impacts of covariates on the tendency of resolution. The model specifications fulfill Equations (3.2) (Model I), (3.3) (Model II) resp. (3.5) (Model III), i.e., loan specific characteristics, macroeconomic information and/or frailties. Significance is indicated at 10% (*), 5% (**) and 1% (***). Using the likelihood ratio test for the frailty where the null model is given by Model II. The standard error of the frailty is computed by bootstrapping with resampling and replacement for 2,000 steps.

3.B.3 Regression Results for Additional Macroeconomic Variables

Table 3.B.4: Regression results for Model II with different macroeconomic variables

		3-month government securities	1-year government bonds	Consumer price index	Gross domestic product	House price index	Equity index	Industry production	Ratio non-performing loans	Term spread	Unemployment rate	Volatility index
United States	Coef.	0.0368	0.3952	1.7057	2.4381	0.7566	0.2800	1.2620	0.0122	-0.0188	2.5877	-0.0036
		***	***	**	***	***	***	***		*	***	***
	SE	(0.0113)	(0.0611)	(0.7361)	(0.5143)	(0.2542)	(0.0595)	(0.1977)	(0.0104)	(0.0110)	(0.7065)	(0.0014)
	LL	-49,338	-49,323	-49,341	-49,332	-49,339	-49,332	-49,323	-49,343	-49,342	-49,337	-49,340
	AIC	98,738	98,707	98,744	98,726	98,740	98,727	98,708	98,748	98,746	98,735	98,742
	McFadden's adjusted R^2	0.0117	0.0120	0.0116	0.0118	0.0116	0.0118	0.0120	0.0116	0.0116	0.0117	0.0116
	Cox & Snell's R^2	0.0233	0.0239	0.0232	0.0235	0.0233	0.0235	0.0239	0.0231	0.0231	0.0233	0.0232
Great Britain	Coef.	0.0072	-0.1111	-5.7128	0.8262	0.4247	0.3181	1.2428	0.0231	-0.0404	14.8656	-0.0063
				***		**	***	***	*	*	***	***
	SE	(0.0216)	(0.0701)	(1.6865)	(0.5461)	(0.1878)	(0.0833)	(0.3777)	(0.0121)	(0.0243)	(2.0106)	(0.0017)
	LL	-40,327	-40,326	-40,322	-40,326	-40,325	-40,320	-40,322	-40,326	-40,326	-40,300	-40,321
	AIC	80,709	80,706	80,697	80,707	80,704	80,694	80,698	80,705	80,706	80,654	80,695
	McFadden's adjusted R^2	0.0303	0.0304	0.0305	0.0304	0.0304	0.0305	0.0305	0.0304	0.0304	0.0310	0.0305
	Cox & Snell's R^2	0.0496	0.0497	0.0498	0.0497	0.0497	0.0499	0.0498	0.0497	0.0497	0.0506	0.0499
Canada	Coef.	-0.0407	-0.2097	-1.7865	-0.4313	-0.1682	-0.1570	-0.1268	0.2110	0.0256	4.8003	0.0024
		*	*						***		**	
	SE	(0.0214)	(0.1200)	(2.0927)	(0.4796)	(0.3856)	(0.0975)	(0.3482)	(0.0725)	(0.0218)	(2.3041)	(0.0022)
	LL	-22,043	-22,044	-22,045	-22,045	-22,045	-22,044	-22,045	-22,041	-22,045	-22,043	-22,045
	AIC	44,141	44,141	44,144	44,144	44,144	44,142	44,144	44,136	44,143	44,140	44,143
	McFadden's adjusted R^2	0.0486	0.0486	0.0486	0.0486	0.0486	0.0486	0.0486	0.0487	0.0486	0.0487	0.0486
	Cox & Snell's R^2	0.0487	0.0487	0.0487	0.0487	0.0487	0.0487	0.0487	0.0488	0.0487	0.0487	0.0487

Notes: The table summarizes regression results of impacts of macroeconomic covariates on the tendency of resolution. The model specifications fulfill Equation (3.3), i.e., loan specific characteristics and macroeconomic variables. Significance is indicated at 10% (*), 5% (**) and 1% (***). The house price indices are taken from S&P Case Shiller (US), Nationwide (Great Britain), Teranet-National Bank (Canada). The ratio non-performing loans (to gross loans of banks), the unemployment rate and the volatility index are included as levels. All other macroeconomic variables are given by log returns.

Chapter 4

A Copula Sample Selection Model for Predicting Multi-Year LGDs and Lifetime Expected Losses

This chapter is joint work with Toni Oehme[1], Daniel Rösch[2] and Harald Scheule[3], and corresponds to a working paper with the same name.

Abstract: Recent credit risk literature has proposed (i) sample selection models for dependencies between the one-year Probability of Default (PD) and Loss Given Default (LGD), and (ii) multi-year approaches which are limited to default risk. This paper provides a model for the simultaneous prediction of continuous default times and multi-year LGDs. These measures are paramount to predict term structures of LGDs and Lifetime Expected Losses under the revised loan loss provisioning framework of IFRS 9 and US GAAP. The model includes a variation of copulas and corrects for sample selection bias of LGDs, which are only observed given a default event. We find empirical evidence that bonds which default closer to origination tend to generate higher LGDs. The model enables more precise estimates of Lifetime Expected Losses and prevents a severe underestimation in contrast to more restricted credit risk models.

JEL classification: C51; G20; G28

Keywords: continuous time-to-default; IFRS 9; lifetime expected loss; loss given default; multi-period; term structure

[1]Risk Research GmbH, Am Biopark 11, 93053 Regensburg, Germany. E-Mail: toni.oehme@risk-research.de

[2]Chair of Statistics and Risk Management, Faculty of Business, Economics, and Business Information Systems, University of Regensburg, 93040 Regensburg, Germany. E-Mail: daniel.roesch@ur.de

[3]Finance Discipline Group, UTS Business School, University of Technology, Sydney, PO Box 123, Broadway NSW 2007, Australia. E-Mail: harald.scheule@uts.edu.au

4.1 Introduction

Stochastic modeling of future cash flows of financial instruments and correlations between these instruments is paramount in quantitative finance. In credit risk, structural models which follow Black and Scholes (1973) and Merton (1974) suggest implicit dependence between the stochastic event of default of debt (triggered by the stochastic firm value) and the stochastic Loss Given Default (LGD), as it assumes a recovery of the firm value in case of default. Reduced form models can provide for more precise credit risk predictions (see e.g., Uhrig-Homburg (2002)), but empirical models for the dependence between continuous default time and future LGD have not been provided by the literature to date. This paper suggests how this dependence could be modeled precisely and how it can be quantified empirically.

The issue is important to practice as a higher propensity to default early during the lifetime of a credit risky financial instrument tends to be associated with a higher loss rate from the default. This has also been recognized for the estimation of the expected credit loss over the entire lifetime of a credit risky financial instrument (Lifetime Expected Loss) under the revised loan loss provisioning of the International Financial Reporting Standards (IFRS 9) and the US Generally Accepted Accounting Principles (GAAP) becoming effective in 2018 and 2020 respectively (see International Accounting Standards Board (2014) and Financial Accounting Standards Board (2016)). The paper provides a novel approach for modeling and empirically estimating dependencies between multi-year defaults and LGDs consistent with these requirements.

Much empirical literature on credit risk analyses the relationship between defaults or LGDs and their risk drivers in separate models. Defaults and the Probability of Default (PD) in a discrete-time setting are analyzed by Altman (1968), Shumway (2001), Campbell et al. (2008) and Kiefer (2009), whereas default intensities in a continuous-time setting are modeled and examined in Duffie et al. (2007), Duffie et al. (2009), Bellotti and Crook (2009), Lando and Nielsen (2010) and Orth (2013). LGDs and their risk drivers are the matter of subject in Carey (1998), Pan and Singleton (2008), Loterman et al.

(2012), Huang and Oosterlee (2012) and Jankowitsch et al. (2014).

Only few papers, however, have modeled and analyzed the dependence between LGDs and default. Altman et al. (2005) and Chava et al. (2011) empirically find a positive correlation between the one-year PD and the LGD of US corporates bonds. Bade et al. (2011) and Rösch and Scheule (2014) show that independence assumption can lead to a severe underestimation of credit risk.

Hence, we identify and close the following gaps in the existing empirical literature on credit risk. While there are continuous-time multi-year intensity models for the time-to-default on the one hand, and LGD models on the other hand, we combine both sides to a simultaneous model for multi-year continuous default times and LGDs. This simultaneous model allows us to infer an association between time-to-default and LGD, e.g., whether debts that default closer to origination tend to produce higher LGDs. The existing models using linear correlations, see Bade et al. (2011), are extended by using a copula approach, rather than simple linear correlations, because copulas allow for more sophisticated and complex dependence structures. We also explicitly take a potential sample selection following Heckman (1979) into account as LGDs can only be observed in the case of default. The models for dependence between stochastic default events and LGDs in a one-period setting are extended to a continuous default time multi-period setting. This allows modeling the entire term structure of losses of credit risky financial instruments from which, for example, Lifetime Expected Losses can be derived that could be used by financial institutions under IFRS 9 and US GAAP.

For the empirical analysis, we use default and loss data of 48,828 US corporate bonds with origination between 04/26/1982 and 03/07/2014 provided by Moody's Default & Recovery Database (DRD). We find a significantly negative stochastic dependence between the time-to-default and LGD after controlling for a large information set of observable covariates. Thus, bonds that default close to origination tend to generate higher LGDs while bonds approaching maturity tend to generate lower LGDs after controlling for covariates. These dependencies generally increase Lifetime Expected Losses and need to be considered for loan loss provisioning in order to prevent capital shortfalls in recessions. Traditional one-year selection models underestimate losses. The analysis shows that the dependence between default times and LGDs can not be adequately modeled by correlation measures and is more precisely measured by Archimedean copulas.

The rest of the paper is organized as follows. In Section 4.2, we derive a general approach for a simultaneous, multivariate time-to-default and LGD model, both on a multi-year perspective. Section 4.3 provides the data description and shows the empirical results. In Section 4.4, we compare Lifetime Expected Loss estimates and show implications for multi-year predictions of Losses Given Default. Section 4.5 concludes.

4.2 Model

4.2.1 Multi-Year Losses over Lifetime

A financial instrument has a maturity (lifetime) m until which a default can happen at any time. Therefore, it is necessary to consider potential losses at various default times and to weight those with a suitable discount factor. Let $T \geq 0$ be the stochastic time-to-default of an instrument since its issuance or origination. The corresponding loss L_T up to maturity m is given by the product of a default indicator $\mathbb{1}_{\{T \leq m\}}$ (1, if default up to maturity m; 0, otherwise), the loss rate given default at the time of default (LGD_T) which is typically between 0 and 1, and the exposure at default (EAD_T). As we deal with bonds where the exposure can be assumed to be deterministic, we set $EAD_T = 1$ without loss of generality.[4] The present value of lifetime loss at bond origination is then

$$L_T \cdot b(T) = \mathbb{1}_{\{T \leq m\}} \cdot LGD_T \cdot b(T), \tag{4.1}$$

where $b(T) = \exp\left(-\int_0^T r(u)du\right)$ is a continuous discount rate with time dependent discount factor $r(t)$ ($t \geq 0$). All variables relate to the random time-to-default T. By including a discount rate, we take into account the time value of money. The choice of discount rates does not affect the *estimation* procedure of the model as we model term

[4]The model can be modified for other financial instruments like loans or derivatives in future work.

structures in a first stage. Discount rates affect the *prediction* of expected losses over lifetime when term structures are aggregated in a second stage. Discount rates, e.g., risk-free or risk-adjusted rates, as well as, option of stochastic or non-stochastic rates can then be chosen according to individual preferences or requirements.

Given stochastic process specifications for the variables on the right hand side of Equation (4.1), e.g., hazard rates for the time-to-default, densities for the LGDs and dependence measures between the variables, the easiest way of obtaining the distribution of losses over lifetime is to use a Monte-Carlo simulation. Moreover, moments from this distribution can be computed, and a particularly important moment is the Expected Loss over Lifetime (or Lifetime Expected Loss).

A general definition of the Lifetime Expected Loss (LEL) of a financial instrument is then given by the expectation of Equation (4.1) as

$$LEL = E\left(\mathbb{1}_{\{T\leq m\}} \cdot LGD_T \cdot b(T)\right). \tag{4.2}$$

Denote a realization of LGD_T by l, then a more explicit expression under deterministic discount rates can be given by the usual formula for the expectation as

$$\begin{aligned} LEL &= \int_0^\infty \int_0^1 f_{LGD_T,T}(l,t) \cdot l \cdot \mathbb{1}_{\{t\leq m\}} \cdot b(t)\ dl\ dt \\ &= \int_0^m \int_0^1 f_{LGD_T,T}(l,t) \cdot l \cdot b(t)\ dl\ dt, \end{aligned} \tag{4.3}$$

where $f_{LGD_T,T}$ denotes the joint density of the time-to-default t and the LGD l. The integration is from time 0 to maturity m for the default time and from 0 to 1 for the LGD as it is typically expressed as a fraction.

Therefore, the lifetime loss is determined by the joint distribution of the time-to-default and the LGD. Another representation is given by using the definition of conditional random variables and their expectations:

$$
\begin{aligned}
LEL &= \int_0^m f_T(t) \cdot b(t) \cdot \int_0^1 f_{LGD_T|T=t}(l) \cdot l \; dl \; dt \\
&= \int_0^m f_T(t) \cdot b(t) \cdot E(LGD_T|T=t) \; dt,
\end{aligned}
\tag{4.4}
$$

where $f_T(t)$ denotes the marginal density of the time-to-default, $f_{LGD_T|T=t}(l)$ the conditional density of the LGD given the time-to-default and $E(LGD_T|T=t)$ the conditional expected LGD given the time-to-default. The LEL is determined by these conditional expectations of LGDs. In fact the LEL is just the integral of these means weighted by the density of the time-to-default and the discount rate.

Therefore, we need the joint distribution of default times and LGDs for modeling the lifetime loss distribution and LEL. A copula is a means for modeling multivariate joint distributions based on the marginal distributions. In the following we first describe the marginal distributions for the time-to-default and the LGD, and then their link via copulas.

4.2.2 Default Time Model

Hazard rate or survival time models are well suited for investigating influences on censored metric variables describing times up to an event. In finance two methods are widely used: the Cox regression and Accelerated Failure Time (AFT) models and both approaches share properties (see Lee and Urrutia (1996), Shumway (2001), Chava and Jarrow (2004), Duffie et al. (2007), Duffie et al. (2009) and Bellotti and Crook (2009)). The Cox model is a semi-parametric approach, i.e., there exists no fully-specified density to compute the loss distribution. Hence, we use AFT models because they give a fully specified density to model the co-movement of the time-to-default and the LGD.

AFT models are a regression method for investigating the influence of covariates on the time-to-default. Let $x_i^T = (1, x_{1i}^T, \cdots, x_{p_T i}^T)'$ be a covariate vector, including p_T covariates, which are known or predictable at origination for bond i ($i = 1, ..., n$).[5] The

[5] The covariates for the time-to-default are measured at origination, i.e., $t = 0$. The subscript T

regression equation of the AFT model is then given by

$$\log T_i = \beta_T' x_i^T + \sigma \varepsilon_i, \tag{4.5}$$

with unknown parameter vector β_T, shape parameter $\sigma > 0$ and error terms ε_i. Various kinds of distributions of the error terms imply a variety of AFT models and distributions of failure times. In our empirical analysis, we use the Weibull, log-logistic and log-normal models.[6]

We consider right-censoring as we do not observe default events after maturity, call or the end of the observation period. Let δ_i be a censoring indicator for instrument i that takes the value zero for censoring and one for default. The end of the observation period for an instrument is defined by $t_i \leq m_i$ with maturity m_i.

In a stand-alone (marginal) approach, the parameters β_T and σ are estimated by the Maximum-Likelihood method using the likelihood function

$$L(\beta_T, \sigma) = \prod_{i:\delta_i=0} \left(1 - F_{T_i}(t_i) \right) \prod_{i:\delta_i=1} f_{T_i}(t_i), \tag{4.6}$$

where the density $f_{T_i}(t_i)$ and cumulative density function $F_{T_i}(t_i)$ depend on the unknown parameters β_T and the covariates x_i^T. The first product describes all censored cases by their probabilities of survival up to the last observation at t_i. The second product describes all defaults by the densities of default at time t_i.

4.2.3 Loss Given Default Model

As mentioned earlier, the LGD is a loss *rate* given default and usually bounded by zero and one. A common and widely used distribution for proportions is the beta distribution (see Gupton et al. (1997) (CreditMetrics), Gupton (2005) (Moody's KMV) and Huang and Oosterlee (2012)). In order to include covariates we use a beta regression where the

indicates that the parameter vector β_T and the covariate vector x_i^T differ from the later presented LGD counterparts, i.e., for the mean of loss rates with index μ.

[6] The corresponding distributions for the error terms are exponential, logistic and normal. For a more detailed introduction see Hosmer et al. (2008).

mean LGD is modeled by a logit transformation (see Ferrari and Cribari-Neto (2004)). In addition, the dispersion is modeled by a precision parameter.

Let the LGD be denoted by the beta distributed random variable $Y \sim Beta(\alpha, \beta)$ with parameters $\alpha, \beta > 0$. The density of Y is given by

$$f_Y(y) = \frac{1}{B(\alpha,\beta)} y^{\alpha-1}(1-y)^{\beta-1}, \qquad y \in (0,1), \tag{4.7}$$

with the beta function $B : (0,\infty) \times (0,\infty) \to \mathbb{R} \times \mathbb{R}$. The standard parameters α and β are replaced by $0 < \mu < 1$ and $\phi > 0$, such that

$$\mu = \frac{\alpha}{\alpha+\beta} \quad \text{and} \quad \phi = \alpha + \beta, \tag{4.8}$$

where μ is the mean of the random variable Y which is the expected LGD of bond i with covariate vector $x_i^\mu = (1, x_{1i}^\mu, \cdots, x_{p_\mu i}^\mu)'$ that contains p_μ idiosyncratic and systematic risk factors. The score of the logit transformation is defined by $\beta_\mu' x_i^\mu$ with an unknown parameter vector β_μ. Hence, the resulting regression model equation for the mean LGD is given by

$$\mu_i = \frac{1}{1+\exp(-\beta_\mu' x_i^\mu)}. \tag{4.9}$$

The randomness of the LGD is modeled by the second distribution parameter ϕ, which can be interpreted as precision by the equation for the variance $Var(Y) = \frac{\mu(1-\mu)}{1+\phi}$, i.e., the higher the precision the lower the variance of the LGD is, and vice versa. In a stand-alone (marginal) approach, the unknown parameters β_μ and ϕ are estimated by the Maximum-Likelihood method using the likelihood function

$$L(\beta_\mu, \phi) = \prod_i f_{Y_i}(y_i) = \prod_i \frac{1}{B(\mu_i\phi, (1-\mu_i)\phi)} y_i^{\mu_i\phi-1}(1-y_i)^{(1-\mu_i)\phi-1}, \tag{4.10}$$

where the mean $\mu_i = \mu(\beta_\mu' x_i^\mu)$ depends on covariates, unknown parameters and the logit link function.

4.2.4 Copula Selection Model

Earlier literature shows that one-period default event and LGDs are dependent. Ignoring dependencies may cause biased and inconsistent parameter and LGD estimates due to sample selection. We propose various copulas to model the dependencies between times-to-default and LGDs.

Copulas offer flexible dependence structures and have become a widely used method in finance (see Embrechts et al. (2003), Nelsen (2006) and Cherubini et al. (2012)). For a brief introduction to copulas we refer to Appendix 4.A.1. In general, copulas are defined for any number of univariate random variables. Here, we consider the two-dimensional case in which a copula describes the stochastic co-movement of the time-to-default and the LGD after controlling for covariates via marginal regression models. This means that some of our covariates control for observed dependencies, e.g., a credit rating may affect default risk *and* loss rates. A copula then describes additional stochastic effects to these shifts which can not be measured by covariates and hence marginal models.

A common copula is the Gaussian copula based on the bivariate Gaussian (normal) distribution, which is simple, but often does not match empirical observations in finance (see Embrechts et al. (2003) and Cherubini et al. (2012)), e.g. it does not exhibit fat tails and models only linear dependence, i.e. correlation. Similar limitations apply to Student's t copula. We consider both in this paper, but also study the class of Archimedean copulas, which are also simple but more flexible because they can model non-linear dependence and asymmetric distributions including fat tails. In this class the marginals are connected by a so-called generator function $\varphi : [0, 1] \to [0, \infty]$, which is continuous, strictly decreasing and convex. This copula family is given by

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)), \qquad u, v \in [0, 1], \tag{4.11}$$

where $u = F_T(t)$ and $v = F_Y(y)$ denote transformations of the time-to-default and LGD depending on the unknown parameters β_T and β_μ in Equation (4.5) and (4.9) and their marginal distributions. In addition, the generator depends on a parameter θ, which identifies the strength of interdependence between the marginals. We can compare this strength by the standardized rank correlation measure Kendall's $\tau \in [-1, 1]$. A positive value of this measure indicates positive dependence and vice versa. Genest and Rivest (1993) derive a simple formula for Archimedean Kendall's τ, which can be computed by the generator via

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt. \tag{4.12}$$

In our empirical analysis we use the popular Archimedean copulas Ali-Mikhail-Haq (AMH), Clayton, Frank, Gumbel and Joe. Archimedean copulas generate a variety of distributional behavior including skewness and tail shape. Table 4.A.1 in the appendix shows the copula functions, generators and corresponding ranges for their dependence parameters as well as Kendall's τ. The Product copula is also shown which serves as a base case and implies the stand alone approach, i.e., the independent modeling of time-to-default and LGD. In addition, we use Gaussian and Student's t copulas as benchmarks.

The general lifetime loss likelihood follows from Equation (4.6) and (4.10):

$$L(\beta_T, \sigma, \beta_\mu, \phi, \theta) = \prod_{i:\delta_i=0} \left(1 - F_{T_i}(t_i) \right) \prod_{i:\delta_i=1} f_{(T_i,Y_i)}(t_i, y_i). \tag{4.13}$$

Equation (4.13) extends the sample selection approaches by Tobin (1958) and Heckman (1979) to distributions other than Gaussian and the model by Smith (2003) to continuous times-to-default. The first product contains all information of all bonds without default by using their probabilities of survival. The second product contains all information of defaulted bonds, i.e., their times-to-default, LGDs and dependencies via the joint density of the marginals. Expressed in terms of copulas the joint density becomes

$$f_{(T_i,Y_i)}(t_i, y_i) = c\left(F_{T_i}(t_i), F_{Y_i}(y_i)\right) \cdot f_{T_i}(t_i) \cdot f_{Y_i}(y_i), \qquad t_i \geq 0,\ 0 < y_i < 1, \tag{4.14}$$

with the copula density $c(u, v) = \partial^2 C(u, v)/(\partial u \partial v)$, which is the derivative of the copula function $C(u,v)$ by its marginals.[7] This copula representation of the joint density changes the likelihood in Equation (4.13) to

$$L(\beta_T,\sigma,\beta_\mu,\phi,\theta)=\prod_{i:\delta_i=0}\left(\,1-F_{T_i}(t_i)\,\right)\prod_{i:\delta_i=1}\left[\,c\Big(F_{T_i}(t_i),F_{Y_i}(y_i)\Big)\cdot f_{T_i}(t_i)\cdot f_{Y_i}(y_i)\,\right] \quad (4.15)$$

which is the lifetime loss likelihood we use for the parameter estimation.

So far, the presented approach is unrestricted and depends on the distributions of default times and LGDs and their dependencies. We study several copulas and in particular whether the dependence is negative or positive. Because some copulas only allow for positive dependencies, i.e., Joe and Gumbel (see Appendix Table 4.A.1), the estimation procedure explicitly tests both signs. For this purpose, we estimate the model twice, i.e., with loss and recovery data based on the recovery rate (RR) = 1 − LGD.

Finally, the procedure provides a copula-based expression for the LEL by combining Equation (4.3) and (4.14) via

$$LEL=\int_0^m\int_0^1 c\left(F_T(t),F_Y(y)\right)\cdot f_Y(y)\cdot f_T(t)\cdot y\cdot b(t)\;dy\;dt. \quad (4.16)$$

The choice of the discount rate $b(t)$ does not affect the estimation procedure in Equation (4.15) which is fully determined by the default process, the loss rate and the link between both.

4.3 Empirical Analysis

4.3.1 Data

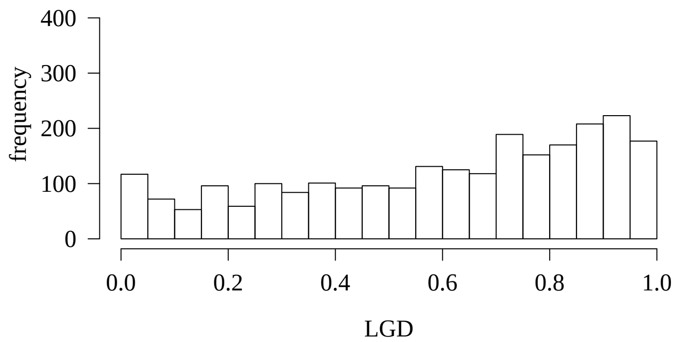
Our dataset is based on Moody's Default & Recovery Database (DRD). It provides information on US corporate bonds with long-term ratings between 01/01/1970 and 03/07/2014. We use bonds with origination after the last significant change in Moody's

[7] A proof for Equation (4.14) is given in Appendix 4.A.2. For the Product copula it holds $c(u,v)=1$ and for the Gaussian $c(u,v)=\phi_2(\Phi^{-1}(u),\Phi^{-1}(v);\theta)/[\phi_1(\Phi^{-1}(u))\phi_1(\Phi^{-1}(v))]$ with the one- and two-dimensional standard Gaussian densities ϕ_1 and ϕ_2. Archimedean copula densities are represented by their generator via $c(u,v)=-\frac{\varphi''(C(u,v))\varphi'(u)\varphi'(v)}{[\varphi'(C(u,v))]}$ (see Embrechts et al. (2003)).

rating methodology on 04/26/1982. We exclude incomplete as well as redundant data and obtain 48,828 bonds including 2,455 defaults. For each of these observations we have a closing date, i.e., the day when a bond is issued, and a censor date, i.e., the day of default or the last observed day if there was no default.[8] In the Moody's rating methodology a default event is recorded if (i) interest and/or principal payments are missed or delayed, (ii) Chapter 11 or Chapter 7 bankruptcy is filled, or (iii) a distressed exchange such as a reduction of the financial obligation occurs.

Moody's reports the recovery rate (RR) for defaulted bonds, which is measured as the ratio of the price of a bond 30 days after default and the par value of the bond. We cap the recovery of a small number of 45 bonds to the highest observed value lower than 1, i.e., 0.9985, to get LGDs between 0 and 1. Figure 4.1 shows a histogram of the realized LGDs. The observed losses seem to be almost uniformly distributed, but with a slight tendency to higher values.

Figure 4.1: Histogram of sample LGDs



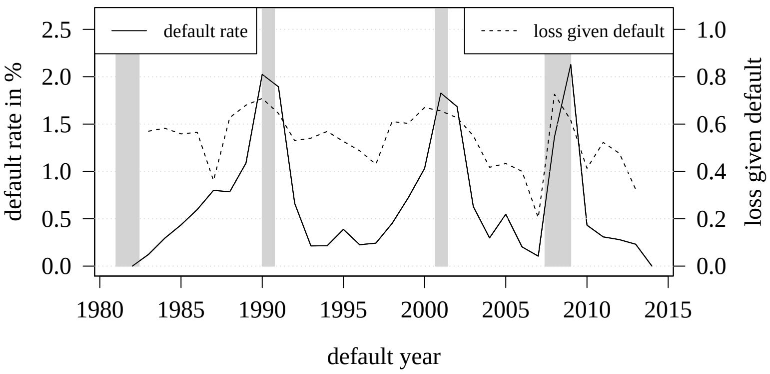
Notes: The loss rate given default is measured as 1 - the ratio of the price of a bond 30 days after default and the par value of the bond. Dependences between LGDs and times-to-default are shown in Figure 4.4.

Figure 4.2 shows the time series of yearly default rates (DR), i.e., the ratio of defaulted bonds to all observed bonds, and yearly average LGDs. Both are volatile through time and increase for recessions.

We control for a variety of information that is available at the origination of each bond. Moody's provides bond- and issuer-specific information on ratings, seniorities and

[8]The latter can be determined by the end of maturity, the date a bond is no longer observed in the Moody's database or the last day included in the dataset, i.e., 03/07/2014.

Figure 4.2: Sample default rates and mean LGDs



Notes: The figure shows yearly default rates and LGDs by the year of default. The solid line is the default rate with values at the left axes. The dashed line is the mean of realized LGDs with values at the right axes. The grey bars indicate recessions as defined by the National Bureau of Economic Research (NBER).

the issuers' industry affiliations. Table 4.1 shows default rates, means and standard deviations of LGDs and the number of observations in corresponding categories. We divide the ratings at the closing dates into 7 groups (Aaa, Aa, A, Baa, Ba, B and C). The default rates in the rating classes increase strictly monotonic in this order. The mean LGDs tend to be higher for poorer ratings. The seniorities are given by the categories senior secured (SS), senior unsecured (SU), senior subordinated (SR) and subordinated (SB). Observed mean LGDs and DRs are higher for weaker seniorities. The industry affiliations of bond issuers are grouped to financial institutions (financial), industrial corporates (industrial) and others.

Furthermore, we use the face amounts (in natural logarithm of millions US $) and maturities (in natural logarithm of years) as metric covariates. The importance of the above information with respect to credit risk is consistent with Acharya et al. (2007) and Jankowitsch et al. (2014) among others. In addition we control for market-based and balance sheet data of the issuers. We use weighted lagged excess return (EXRETAVG), market-to-book ratio (MB), weighted lagged net income to total assets (NIMTAAVG), the relative size as the log ratio of firms' market capitalization to the overall capitalization of the S&P500 (RSIZE), the volatility of firms' stock return (SIGMA), and total liabilities to market value of total assets (TLMTA). All measures are collected from COMPUSTAT and CRSP and computed following Campbell et al. (2008), where the significance of the

Table 4.1: Descriptive statistics of bond- and issuer-specific information (1)

Rating	Aaa	Aa	A	Baa	Ba	B	C	all
Default rate in %	0.03	0.20	2.02	2.88	6.72	17.33	21.25	5.03
Mean LGD in %	47.00	52.16	63.48	55.32	53.70	59.65	70.76	59.82
Std. dev. LGD in %	0.00	24.70	29.82	29.10	28.28	28.33	28.56	28.92
Observations	2,977	8,636	13,903	9,721	4,937	7,520	1,134	48,828

	Seniority				Industry		
	SS	SU	SB	SR	financial	industrial	other
Default rate in %	4.48	3.62	11.89	18.78	1.99	8.20	2.32
Mean LGD in %	42.62	60.57	69.91	71.03	63.56	59.80	54.26
Std. dev. LGD in %	29.39	28.09	22.75	25.19	30.38	28.60	28.68
Observations	11,848	31,567	1,981	3,432	16,351	23,423	9,054

Notes: This table shows the number of observations (=number of bond issues), the default rate and the mean as well as the standard deviation of the LGDs in our dataset and in various bond categories. These are the rating groups Aaa, Aa, A, Baa, Ba, B or C. In addition, we take into account the seniorities senior secured (SS), senior unsecured (SU), senior subordinated (SR) and subordinated (SB). Finally, we distinguish between bonds of financial institutions, industrial corporates and other industries.

variables for default and loss models has been shown. Descriptive statistics for metric variables are summarized in Table 4.2.

In order to account for macroeconomic influences, we consider the changes of the total US industrial production and the term spread between 10-year and 3-month US treasury rates, published by the Federal Reserve Bank of St. Louis (see Das et al. (2007) and Lando and Nielsen (2010) and compare Figure 4.3).

For modeling multi-year macroeconomic effects, we also use an average macro scenario, i.e., for each bond we use the mean change over lifetime up to maturity in total US industrial production in % per annum and term spread in % - points per annum.

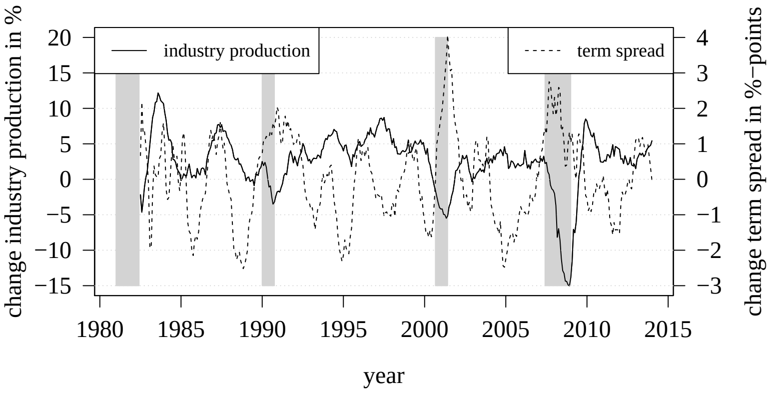
Finally we use a recession indicator following the crises definition of the National Bureau of Economic Research (NBER). It marks a recession at the closing date by the dummy variable ‘vintage’ (1, if recession; 0, otherwise). For robustness checks we also tested other macroeconomic variables like inflation, gross domestic product, financial indices, total loan volume and unemployment, but did not include these in our final model due to low informational values.

Table 4.2: Descriptive statistics of bond- and issuer-specific information (2)

	Min	25%	Mean	Median	75%	Max	SD
Face amount in million $	0.00	60.00	376.67	167.00	400.00	27,420.00	769.00
Maturity in years	0.01	5.00	10.57	8.01	11.01	100.04	9.10
EXRETAVG	-0.0299	-0.0036	0.0017	0.0017	0.0069	0.0322	0.0096
MB	0.3865	1.2370	2.0451	1.7359	2.5248	5.5349	1.1879
NIMTAAVG	-0.0445	0.0023	0.0053	0.0062	0.0098	0.0216	0.0078
RSIZE	-9.1109	-4.0161	-3.2056	-2.6691	-2.1450	-2.1450	1.2763
SIGMA	0.1672	0.2166	0.3443	0.2959	0.4029	1.3836	0.1815
TLMTA	0.0433	0.3826	0.5706	0.5611	0.7826	0.9269	0.2339

Notes: This table shows descriptive statistics of metric covariates. Bond-specific information are taken into account by face amount and maturity. Market-based and balance sheet data using COMPUSTAT and CRSP are included to control for issuer-specific effects. Following the definitions of Campbell et al. (2008), we use weighted lagged excess return (EXRETAVG), market-to-book ratio (MB), weighted lagged net income to total assets (NIMTAAVG), the relative size as the log ratio of firms' market capitalization to the overall capitalization of the S&P500 (RSIZE), the volatility of firms' stock return (SIGMA), and total liabilities to market value of total assets (TLMTA).

Figure 4.3: Macroeconomic variables



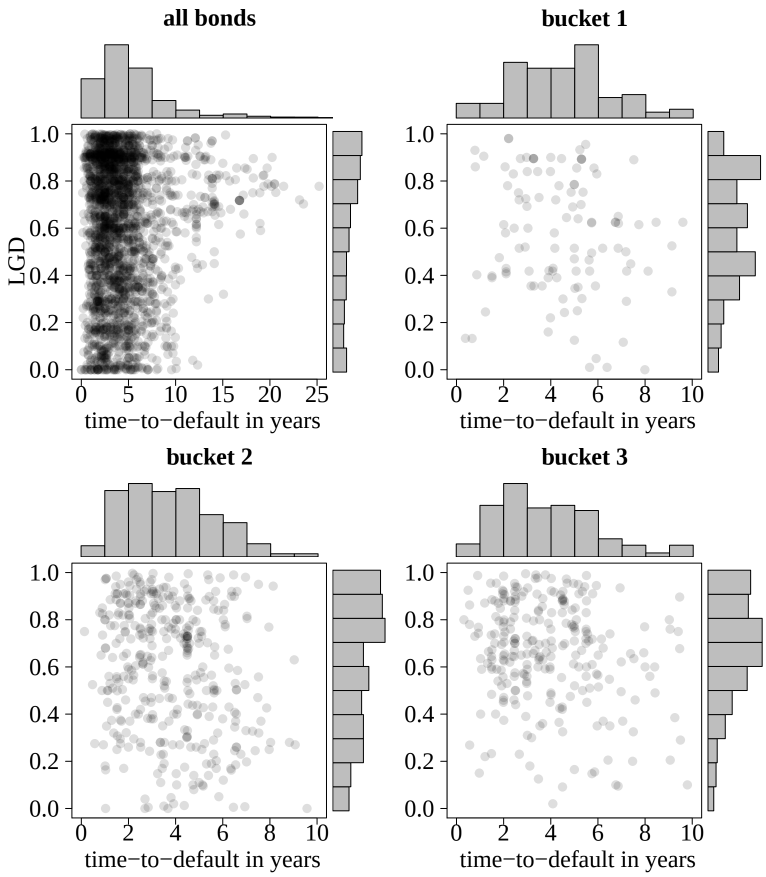
Notes: The figure shows the yearly changes in industry production and term spreads. The growth in industry production is given as percentage changes with values at the left axes. The term spread is measured as the difference between the 10-year and 3-month US treasury rate in %-points with absolute changes at the right axes. The grey bars indicate recessions as defined by the National Bureau of Economic Research (NBER).

4.3.2 Univariate Dependencies

The dependent variables are time to default (and the censor variable) next to LGD. For a defaulted bond it is given by the time in years from origination to the default date on a daily basis. For a non-defaulted bond it is given by the time in years from the closing date to the last observed date on a daily basis, i.e., the end of maturity, the date a bond left Moody's database or the last day included in the dataset, which is 03/07/2014. The distribution of times-to-default and the relation between times-to-default and LGDs is shown in Figure 4.4, both unconditionally and grouped for three specific 'buckets' of bonds. At first, unconditionally, Pearson's correlation coefficient ρ between times-to-default and LGDs takes the value 0.0576 for the entire dataset, which indicates a slight positive dependence between both variables. The more general dependence measure Kendall's τ is 0.0035 and does not indicate any significant link. This, however, represents only an 'average' or unconditional dependence which mixes over various states.

Our approach, however, investigates dependencies which are not triggered by observable factors, i.e., the remaining conditional dependence after controlling for covariates.

Figure 4.4: Scatter plots of realized LGDs and corresponding times-to-default



Notes: The upper left figure shows for all defaulted bonds the time-to-default since closing date in combination with the realized LGD. The empirical correlation between both is 0.0576 and Kendall's τ is 0.0035. The other figures shows the same data for three exemplary risk-buckets (see Table 4.3). For these bonds the empirical correlations are -0.0967, -0.2072 and -0.1230. The values of Kendall's τ are -0.0810, -0.1300 and -0.0737.

To illustrate the distinction between unconditional and conditional dependence graphically for descriptive purposes, we next exemplarily consider three different risk-buckets with specific values for the categorical variables rating, seniority and industry affiliation to control (see Table 4.3). This matches the set of control variables of the Basic Model. The first bucket is comprised of senior unsecured industrial bonds with Ba ratings. The second bucket is more risky including the same bonds but with B ratings. The last bucket contains even riskier industrial subordinated bonds with B ratings. The selected bonds represent 30.4% of all defaults and generate negative correlations between times-to-defaults and LGDs of -0.2072 up to -0.0967 as well as negative values for Kendall's τ between -0.1300 and -0.0737. Thus, the dependence changes after controlling for covariates which shows the need of an extensive analysis including various control factors. Figure 4.4 also shows the existence of a non-linear dependence structure, which supports the use of non-linear copulas.

Table 4.3: Descriptive statistics of exemplary risk-buckets

	Bucket 1	Bucket 2	Bucket 3
Industry	industrial	industrial	industrial
Rating	Ba	B	B
Seniority	SU	SU	SB
Observations	1,382	1,805	1,091
Defaults	109	378	259
Correlation ρ	-0.0967	-0.2072	-0.1230
Kendall's τ	-0.0810	-0.1300	-0.0737

Notes: The risk-buckets represent 30.4% of all defaults. ρ denotes the correlation between times-to-default and LGDs in the corresponding bucket. Kendall's τ is the more general rank correlation measure.

4.3.3 Marginal Models

We estimate the models for two different sets of covariates which are built on each other. In the *Basic Model* we use bond- and issuer-specific information for ratings, seniorities and industry affiliations. In the time-to-default model we exclude seniorities because of limited influence. We aggregate A - Aaa rated bonds for the LGD part due to less observation numbers. In the *Extended Model* we add metric bond-specific (maturity, face amount), issuer-specific (EXRETAVG, MB, NIMTAAVG, RSIZE, SIGMA, TLMTA) and macroeconomic (vintage, industry production and term spread) information as explanatory variables. This specification provides the best results with respect to plausibility, significance and overall goodness-of-fit. Finally, we provide some robustness checks. In our study, the coupon rate, cash and short-term investments as well as the stock price variable as defined in Campbell et al. (2008) did not provide further plausible insights or increased overall goodness-of-fit, either individually or jointly. Thus, we exclude them from the Extended Model.

Furthermore, we consider two dependence models. The *Unrestricted Model* is based on the Frank copula in combination with a log-logistic distribution for the time-to-default, because this combination provides the best performance, measured by the Log-Likelihood and McFadden's R^2 (see also Appendix 4.B). In contrast, the *Restricted Model* is based on the Product copula with a log-logistic distributed default time. This is equivalent to a stand-alone approach with a separate modeling of the times-to-default and LGDs which corresponds to a simple multiplication of marginal likelihoods and hence assumption of independence.

Table 4.4 shows the parameter estimates of the marginal time-to-default (AFT) model based on the log-logistic distribution (see Table 4.B.2 in the Appendix for other distributions).[9] Differences in estimates and standard errors depending on the covariate sets are small. The parameter estimates decrease with lower rating grades and imply stronger tendencies to default for bonds with poorer ratings as well as from issuers with financial or industrial industry affiliation. Longer maturities seem to result in less frequent and later defaults. This could be due to the fact that bonds with longer maturities are able to attract investors if the average risk is low. In contrast, higher face amounts result in higher default rates for given time periods. Issuers with higher default risk may have poor access to capital markets. Thus, they may have to originate bonds less frequently with higher face amounts. Consistent with Campbell et al. (2008), firms with higher excess re-

[9]Expected signs are opposite to standard PD models like probit or logit regression as the dependent variable is the time to default. For example, a negative sign indicates a decreased time-to-default, and thus, an increased probability of default.

Table 4.4: Parameter estimates for the time-to-default

Model	Basic		Extended	
	Restricted	Unrestricted	Restricted	Unrestricted
(Intercept)	7.4130***	7.4731***	6.8052***	6.6360***
	(0.6423)	(0.6469)	(0.7047)	(0.7075)
Aa	-1.1902*	-1.1909*	0.1764	0.2116
	(0.6554)	(0.6600)	(0.7162)	(0.7215)
A	-2.5775***	-2.5781***	-1.4572**	-1.4633**
	(0.6391)	(0.6436)	(0.6568)	(0.6619)
Baa	-2.8552***	-2.8556***	-1.4096**	-1.4034**
	(0.6394)	(0.6439)	(0.6579)	(0.6629)
Ba	-3.7157***	-3.7164***	-2.4167***	-2.4502***
	(0.6402)	(0.6446)	(0.6600)	(0.6651)
B	-4.4755***	-4.4597***	-3.1752***	-3.1433***
	(0.6404)	(0.6448)	(0.6615)	(0.6663)
C	-4.7144***	-4.6617***	-3.1241***	-3.1012***
	(0.6421)	(0.6465)	(0.6704)	(0.6744)
log(maturity)			0.0507	0.0658
			(0.0495)	(0.0489)
log(face amount)			-0.0355	-0.0279
			(0.0228)	(0.0227)
financial	-0.5041***	-0.5400***	-0.4203***	-0.4536***
	(0.0600)	(0.0603)	(0.1051)	(0.1039)
industrial	-0.3679***	-0.4090***	-0.1015	-0.1573
	(0.0524)	(0.0526)	(0.0982)	(0.0971)
EXRETAVG			9.0536***	8.1409***
			(2.0623)	(2.0111)
MB			0.0294	0.0340*
			(0.0201)	(0.0197)
NIMTAAVG			-0.5095	0.9829
			(2.7982)	(2.6478)
RSIZE			0.0255	0.0101
			(0.0234)	(0.0229)
SIGMA			-1.2091***	-1.0527***
			(0.1261)	(0.1227)
TLMTA			-1.1619***	-1.1254***
			(0.1465)	(0.1451)
vintage			0.5691***	0.5127***
			(0.1087)	(0.1062)
industry prod.			1.8500***	1.9028***
			(0.1202)	(0.1204)
term spread			-0.8572***	-0.7144***
			(0.1433)	(0.1359)
σ	1.5711	1.5590	1.5390	1.5296
McFadden's R^2	0.1837	0.2066	0.2408	0.2588
Obs.	48,828	48,828	15,313	15,313

Notes: Standard errors are in parenthesis, significance is indicated at 10% (*), 5% (**) and 1% (***) levels. All results are presented for log-logistic distributed times-to-default. The Restricted Model is given by the Product copula, the Unrestricted Model is given by the Frank copula. McFadden's R^2 is the pseudo R^2 defined in McFadden (1974). Signs are opposite to PD models like probit or logit, i.e., negative signs indicate an increased tendency to default.

turns as well as market-to-book ratios have lower default risk and firms with more volatile stock prices and total liabilities have higher default risk. The influence of net income and relative size seems to be covered by other covariates. We identify a vintage effect, i.e., bonds issued in a recession are more robust to future defaults, because they fulfill higher requirements of investors. As expected, a good macroeconomic scenario over a bond's lifetime – measured by increasing industrial production and decreasing term spreads – lowers default risk.

Table 4.5 shows the results for the marginal LGD model. The goodness-of-fit increases from the Basic to the Extended Model as well as from the Restricted to the Unrestricted Model – measured by R^2.[10] In contrast to the time-to-default model we find substantial differences between the Restricted and Unrestricted Model. The reason is the sample selection in the Restricted Model which ignores dependencies between default times and LGDs and yields biased and inconsistent parameter estimates. In contrast, the Unrestricted Model allows for dependencies and identifies influences of the covariates more properly. In particular, this can be seen for ratings. The Unrestricted (Basic and Extended) Model implies higher LGDs for lower ratings, whereas the Restricted Model is somewhat ambiguous. Regarding seniority, low seniority always implies higher losses. Longer maturities increase LGDs, which is consistent with Chava et al. (2011). The effect of face amounts and financial or industrial industry affiliation is not or only weakly significant. Market-based and balance sheet data have similar influences as in the time-to-default model, i.e., an increased default risk with respect to these variables implies higher loss rates given default. Note also the difference between the estimates and their significances of the Restricted vs. the Unrestricted Model, e.g., for SIGMA, which is significantly positively related to LGDs in the Unrestricted Model only. Concerning macroeconomic covariates, we see that a good economic scenario over a bond's lifetime lowers its LGD, as industry production and term spread have a negative and positive sign, respectively. Moreover, the results show strong evidence for a vintage effect, i.e., the LGDs for bonds originated in recessions are low.

[10] An alternative performance measure that we tested with consistent results is $\mathrm{R}^1 = 1 -$ (MAE of the considered model) / (MAE of model only with intercept) with MAE as the mean absolute error.

Table 4.5: Parameter estimates for the loss given default

Model	Basic		Extended	
	Restricted	Unrestricted	Restricted	Unrestricted
(Intercept)	-0.6764***	-1.7875***	0.1187	-2.5640***
	(0.1136)	(0.1114)	(0.4221)	(0.4089)
Baa	-0.3201***	-0.3221***	-0.4443***	-0.3379**
	(0.0997)	(0.0915)	(0.1686)	(0.1537)
Ba	-0.3040***	-0.1788*	-0.2357	0.0613
	(0.1010)	(0.0923)	(0.1841)	(0.1690)
B	-0.3532***	0.0087	-0.4978***	0.2622
	(0.0912)	(0.0842)	(0.1872)	(0.1760)
C	0.2689**	0.5895***	-0.1518	0.4837**
	(0.1121)	(0.1018)	(0.2395)	(0.2240)
SU	0.8428***	0.7758***	0.7699***	0.6504***
	(0.0647)	(0.0614)	(0.1249)	(0.1153)
SR	1.3373***	1.2231***	1.6013***	1.2088***
	(0.0794)	(0.0704)	(0.1634)	(0.1490)
SB	1.2949***	1.2265***	1.3483***	1.1389***
	(0.0817)	(0.0726)	(0.1525)	(0.1394)
log(maturity)			0.2466***	0.3108***
			(0.0820)	(0.0744)
log(face amount)			-0.0253	0.0300
			(0.0361)	(0.0318)
financial	0.3396***	0.1926*	-0.2135	-0.0266
	(0.1089)	(0.0990)	(0.1944)	(0.1895)
industrial	0.2509***	0.1253	0.2098	0.2729*
	(0.0908)	(0.0831)	(0.1624)	(0.1554)
EXRETAVG			-2.6994	-5.5798**
			(2.7070)	(2.4758)
MB			-0.0782**	-0.0630**
			(0.0306)	(0.0287)
NIMTAAVG			-6.7995*	-0.1353
			(3.9426)	(3.7384)
RSIZE			0.1071***	0.0691**
			(0.0378)	(0.0347)
SIGMA			-0.1176	0.5015***
			(0.1732)	(0.1667)
TLMTA			-0.2892	0.2479
			(0.2412)	(0.2195)
vintage			-0.6524***	-0.7871***
			(0.1746)	(0.1619)
industry prod.			-0.8273***	-1.1417***
			(0.1437)	(0.1370)
term spread			0.8525***	0.7699***
			(0.1578)	(0.1448)
ϕ	1.5906***	1.3043***	1.846***	1.3969***
R^2	0.0835	0.1103	0.1677	0.2232
Obs.	2,455	48,828	886	15,313

Notes: Standard errors are in parenthesis, significance is indicated at 10% (*), 5% (**) and 1% (***) levels. All results are presented for log-logistic distributed times-to-default. The Restricted Model is given by the Product copula, the Unrestricted Model is given by the Frank copula. R^2 denotes the well known coefficient of determination.

4.3.4 Copula Models

Comparing different sets of covariates, negative dependence between times-to-default and losses is identified for both the Basic and the Extended Model, as can be seen in Table 4.6 for the Frank copula based on the log-logistic distribution (see Table 4.B.1 in the Appendix for other copulas). The model goodness-of-fit measures increase from the Basic to the Extended Model as well as from the Restricted to the Unrestricted Model, respectively. For the Extended Model, McFadden's R^2 is 0.2588 for the Unrestricted Model, and thus, it is higher than 0.2408 of the Restricted Model.

Table 4.6: Parameter estimates for the copula

Model	Basic		Extended	
	Restricted	Unrestricted	Restricted	Unrestricted
Log(Likelihood)	-11,368.4	-11,049.8	-3,715.5	-3,627.7
McFadden's R^2	0.1837	0.2066	0.2408	0.2588
τ (LGD)	0	-0.4642***	0	-0.4767***

Notes: Significance for Kendall's τ is indicated at 10% (*), 5% (**) and 1% (***) levels and computed as in Kendall (1938). McFadden's R^2 is the pseudo R^2 defined in McFadden (1974). All results are presented for log-logistic distributed times-to-default. The Restricted Model is given by the Product copula, the Unrestricted Model is given by the Frank copula. The copula parameter θ is given for RR modeling and the rank correlation measure Kendall's τ is given for LGD modeling, see Table 4.B.1.

4.4 Applications

4.4.1 The Term Structure of LGDs

In this section we discuss the practical implications of our approach. In the Restricted Model the distribution of LGDs is determined by idiosyncratic and macroeconomic risk but independent from the default time. For the Unrestricted Model the LGD distribution additionally varies over the 'lifecycle' of the bond and thus follows a term structure.

After estimation, we predict conditional LGDs for a given default time t following the definition of conditional random variables and Equation (4.14) where the conditional loss density is given by

$$f_{LGD_T|T=t}(y) = c(F_T(t), F_{LGD}(y)) \cdot f_{LGD}(y). \tag{4.17}$$

Thus, the mean LGD given a default time t is predicted by

$$E(LGD_T|T = t) = \int_0^1 c\left(F_T(t), F_Y(y)\right) \cdot f_Y(y) \cdot y \; dy. \tag{4.18}$$

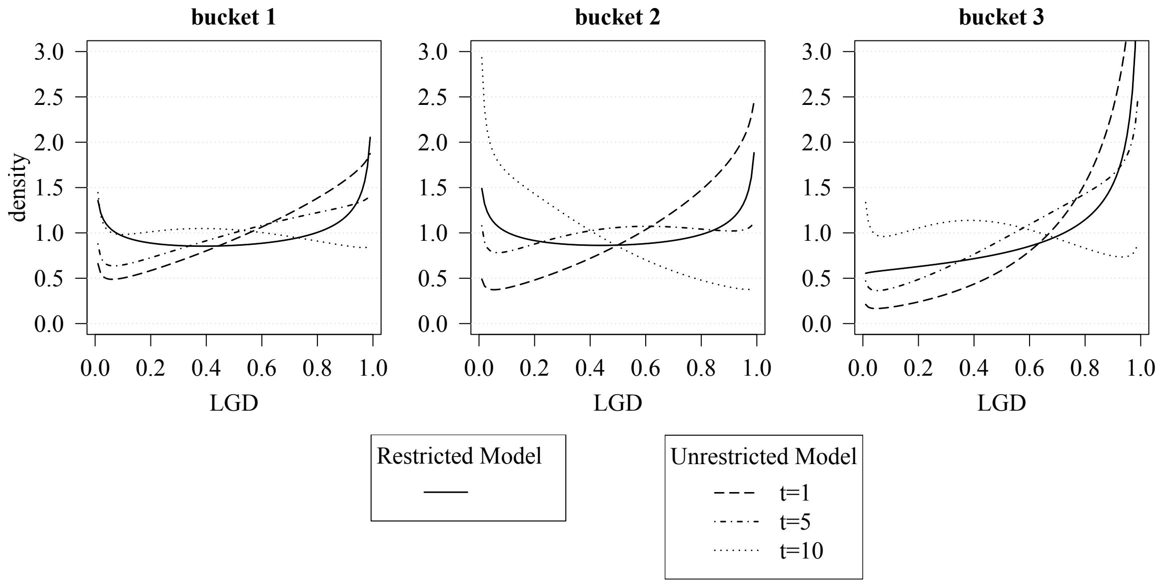
For both model specifications we observe a significant increase of the LGD goodness-of-fit when moving from the Restricted to the Unrestricted Model (see Table 4.5). The R^2 increases in-sample on average by 4.1 percentage points, i.e., by 32.6 %. In addition to the in-sample goodness-of-fit measurement, we perform an out-of-sample evaluation by the following bootstrap procedure. For each of 1000 independent steps we randomly divide the dataset into an 80 % estimation sample – where we estimate model parameters for the Restricted and Unrestricted Model for each specification of covariates (basic and extended) – and a 20 % prediction sample – for which we predict LGDs. In the Basic Model, the out-of-sample R^2 is higher in 95.59 % of all steps for the Unrestricted Model. The ratio for the Extended Model is 91.28 %. An analysis whether the changes are positive on average, by a t-test or a percentile bootstrap hypothesis test, provides p-values of less than 10^{-16} supporting the copula model.

In order to get insights into the impact on LGD and Lifetime Expected Loss (LEL) predictions, we revisit the three risk-buckets of Table 4.3, which represent 30.4% of all defaults and show general implications for LEL predictions.

We distinguish between the Restricted and Unrestricted Model by using the parameter estimates from Table 4.4, 4.5 and 4.6. For ease of exposition we conduct the analysis using the Basic Model.

The Restricted Model assumes independence between default time and LGD. Hence, for each default time the same LGD density applies and the term structure of mean LGDs is flat. The densities for the three risk-buckets are shown in Figure 4.5 as solid lines. For the first two buckets we can see a flat u-shaped function of the Beta distribution, i.e., LGDs are almost uniformly distributed but with a slightly increased chance of small and high values. The third bond bucket with a lower seniority tends to result in higher losses. The corresponding mean LGDs are 52.8%, 51.6% and 63.6%.

Figure 4.5: Densities of LGDs



Notes: The figure shows estimated LGD densities of three exemplary risk-buckets (see Table 4.3). All results are presented for log-logistic distributed times-to-default. The density of the Restricted Model (Product copula) is independent from the default time. For the Unrestricted Model (Frank copula) the density varies over the lifetime.

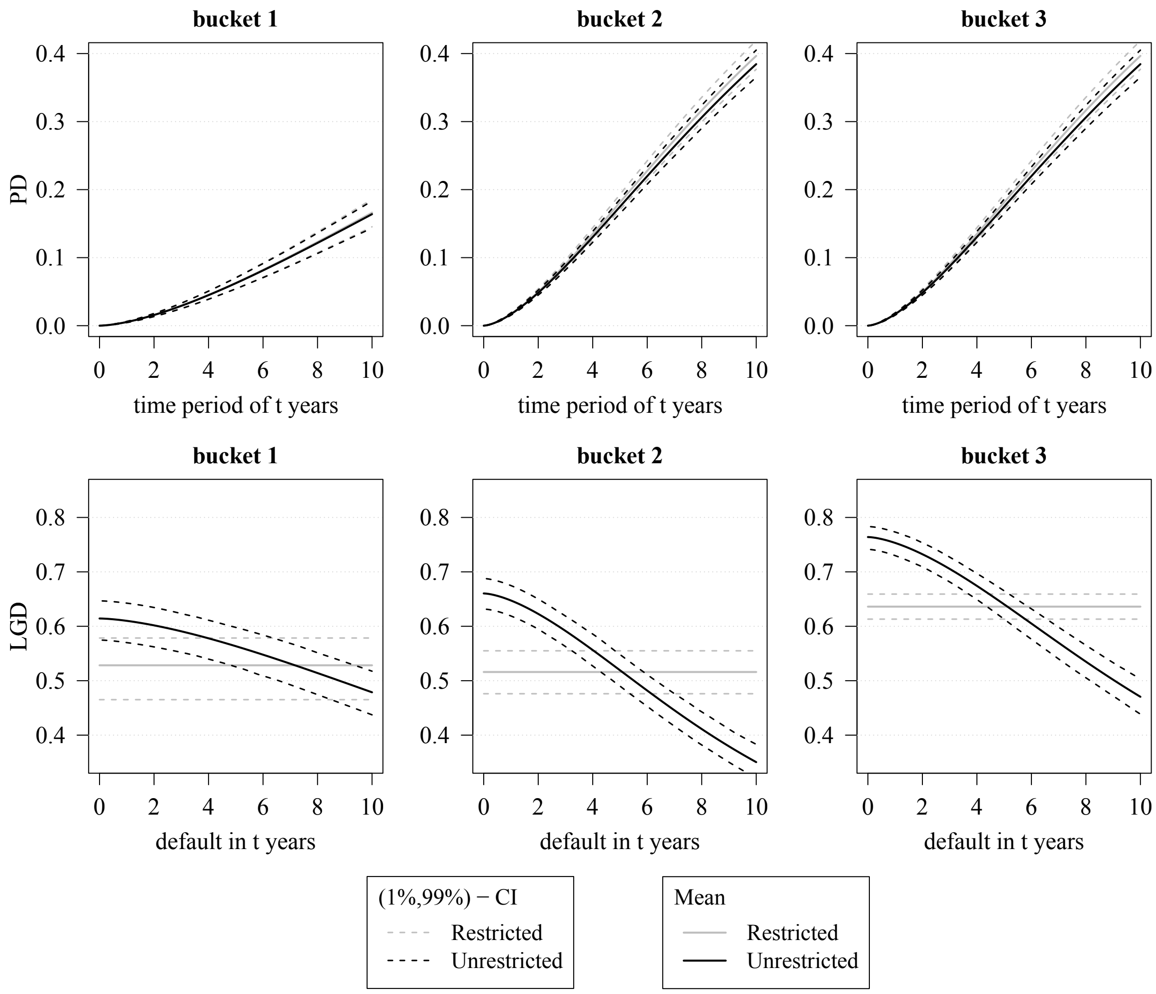
For the Unrestricted Model, the loss depends on the exact time-to-default, as shown in the Equations (4.17) and (4.18). Figure 4.5 shows the term structure of the density for defaults after one, five and ten years since origination as dotted and dashed lines. Due to the negative dependencies we obtain a propensity to large LGDs for early defaults and smaller LGDs for later defaults. For a default after one year since issuance the mean LGDs for the three buckets are 60.8%, 64.6% and 75.3%, and therefore higher than those of the Restricted Model. The longer a bond survives, the further the probability mass of the loss distribution shifts to the left, i.e., the propensity for higher losses decreases. At a certain time, the mean falls below that of the Restricted Model for each of the buckets. For example, five years after issuance the buckets have mean LGDs of 56.1%, 51.7% and 64.0%. The means decrease to 47.5%, 34.4% and 46.9% after five further years.

4.4.2 Estimation of Lifetime Expected Losses

Losses in the Restricted Model are default-time independent and the mean LGDs are constant over time. In the Unrestricted Model, the conditional mean LGD is given by Equation (4.18), which depends on the realized time-to-default. The negative dependence implies decreasing mean LGDs.

Figure 4.6 revisits the term structure of mean LGDs and also shows multi-year PDs as a function of the default time (in years after origination). The PD (in the panel in the

Figure 4.6: Term structure of PD and LGD



Notes: The solid lines show predictions for multi-year PDs and LGDs. The dashed lines show percentile confidence intervals which are constructed by bootstrapping 1,000 independent samples with replacement. The solid lines represent corresponding averages. All results are presented for log-logistic distributed times-to-default. The Restricted Model is given by the Product copula, the Unrestricted Model is given by the Frank copula.

upper row) is cumulated, i.e., shows that probability of defaulting within t years since origination, and increases over time. Across the risk-buckets it can be clearly seen that the riskier buckets have steeper PD curves, e.g., the ten year PD of bucket 1 is half as high as for the others due to the better rating (Ba instead of B). For all types of bonds there are only minor differences in PDs between the Restricted and the Unrestricted Model.

The difference between the Restricted and the Unrestricted Model becomes visible for mean LGDs (panel in the lower row of the figure). The mean LGDs are constant for the Restricted Model, and a decreasing function of the default time for the Unrestricted Model for every bucket with a unique shape for each bucket. This also impacts predictions for Lifetime Expected Losses (LEL).

LEL predictions connect multi-year PDs and default-time dependent losses by Equation (4.16). For computation of LELs we use US treasury rates as discount rates.[11] Due to higher potential LGDs in the first years of the lifetime of a bond, the Unrestricted Model generates higher expected lifetime losses for short- and medium-term bonds. For example, the second bucket provides LEL predictions of 1.09, 9.05 and 15.30 for bonds with maturities of one, five and ten years, instead of 0.87, 7.99 and 15.09 for the Restricted Model, which underestimates lifetime losses of short- and medium-term bonds. For example, the Unrestricted Model provides LEL predictions that are up to 30 % higher than in the Restricted Model. Even for maturities of up to 5 and 10 years, the increases are around 14 % and 7 %. Table 4.7 shows for other buckets that senior secured bonds potential lifetime losses can be more than 30 % higher in the Unrestricted Model. In addition, the underestimation seems to be independent of the maturity, but more relevant for short- and medium-term bonds. In summary, simple multi-year models generally underestimate lifetime losses. In our data, we identify for the Extended Model that the unrestricted estimates of expected losses over lifetime are on average 7.95 % higher than in the Restricted Model with single deviations of more than 30%.

We have seen that standard models produce biased parameter estimates for control

[11]The kind of discounts can be chosen according to individual preferences or requirements and represents an individual evaluation of the time value of payments. Other discount rates can be used as well and do not substantially change the conclusions of the paper. We have also investigated risk-adjusted rates and constant rates. The resulting differences of results are negligible.

Table 4.7: LEL predictions for industrial bonds

	Aaa		Aa		A		Baa		Ba		B		C	
Model	R.	Unr.	R.	Unr.	R.	Unr.	R.	Unr.	R.	Unr.	R.	Unr.	R.	Unr.
m=1														
SS	<0.01	<0.01	<0.01	<0.01	0.03	0.04	0.04	0.06	0.17	0.25	0.53	0.82	1.12	1.38
SU	<0.01	<0.01	<0.01	<0.01	0.05	0.06	0.07	0.08	0.27	0.33	0.87	1.09	1.62	1.77
SR	<0.01	<0.01	<0.01	<0.01	0.06	0.07	0.09	0.10	0.33	0.38	1.06	1.26	1.84	1.95
SB	<0.01	<0.01	<0.01	<0.01	0.06	0.07	0.09	0.10	0.33	0.38	1.07	1.26	1.86	1.95
m=5														
SS	<0.01	<0.01	0.04	0.06	0.37	0.47	0.47	0.66	1.74	2.43	4.87	6.77	9.63	10.47
SU	<0.01	0.01	0.06	0.07	0.57	0.63	0.76	0.85	2.82	3.19	7.99	9.05	13.90	14.02
SR	0.01	0.01	0.08	0.08	0.66	0.72	0.92	1.00	3.41	3.75	9.70	10.66	15.83	15.87
SB	0.01	0.01	0.08	0.08	0.67	0.72	0.93	1.00	3.46	3.75	9.85	10.65	15.99	15.85
m=10														
SS	0.02	0.02	0.10	0.13	0.90	1.12	1.11	1.54	3.86	5.13	9.20	11.98	16.89	16.50
SU	0.02	0.03	0.16	0.18	1.37	1.48	1.81	1.98	6.27	6.67	15.09	15.30	24.37	22.50
SR	0.03	0.03	0.19	0.20	1.60	1.72	2.19	2.34	7.57	7.91	18.32	18.34	27.75	26.12
SB	0.03	0.03	0.19	0.20	1.62	1.71	2.23	2.33	7.68	7.90	18.61	18.31	28.04	26.10

Notes: Left columns show LEL predictions (in %) for the Restricted (R.), right columns for the Unrestricted (Unr.) Model. If we take into account dependences between times-to-default and LGDs, LEL predictions tend to increase. The corresponding rises are marked by $0\% - 10\%$, $10\% - 20\%$, $20\% - 30\%$ and $\geq 30\%$ and characterize a potential underestimation of risk by ignoring the term structure of losses.

variables and significantly underestimate expected losses for a one-year horizon and over lifetime which leads to the following issues. First, one-year LGDs which are used to compute capital requirements are significantly underestimated by models that ignore the term structure of losses. This can lead to capital shortfalls of financial institutions when ignoring the sample selection of loss data. Second, the new accounting rules of IFRS 9 require the prediction of lifetime expected losses for balance sheet impairments of some financial assets. Due to the construction of IFRS 9, the group of assets for which lifetime expected losses are required – when the risk significantly increases since origination – is noticeable larger in recessions. Hence, standard models distort impairments and balance sheet data especially in recessions and thus, negatively affect already challenging investor considerations. In addition, inadequate provisions undermine the precautionary purpose of impairments.

4.5 Conclusion

The literature has proposed (i) sample selection models for dependencies between the one-year PD and LGD, and (ii) multi-year approaches which are limited to default risk. The paper provides a novel approach for simultaneous modeling and empirically estimating dependencies between multi-year defaults and LGDs. These measures are paramount to predicting term structure of LGDs and Lifetime Expected Losses under the revised loan loss provisioning of IFRS and US GAAP.

The results show strong evidence for negative stochastic dependencies between times-to-default and LGDs after controlling for covariates. Bonds with early defaults have a tendency for higher losses. In contrast, a bond which proves its financial strength in the beginning of its maturity implies lower future losses. A separate stand-alone modeling of times-to-default and LGDs ignores these dependencies and results in biased parameter and loss estimates due to sample selection issues. The consequences are shortfalls with respect to regulatory capital (Basel) and loan loss provisions (IFRS 9 and US GAAP). The proposed models may help to strengthen capital cushion and hence prevent institutions as well as investors from being affected by financial risk.

In addition, we identify significant risk triggers for future defaults and losses. These are bond-specific information (rating, seniority, maturity and face amount), market-based as well as balance sheet data (excess returns, stock volatility, market-to-book ratio and liabilities), industry affiliation and macroeconomic information (industry production and term spread). Finally, we measure vintage effects and find that bonds issued in past crises seem to be more robust against future trouble.

The contribution is confirmed by various robustness checks. First, the choice of the time-to-default model does not affect our results. Second, the negative stochastic default time dependence of the LGD is scalable for different copula types. Third, the results are robust for different combinations of control variables.

Appendix 4.A Copula Theory

4.A.1 Introduction

X and Y (also called marginals) are arbitrary but fixed univariate random variables with marginal cumulative density functions F_X and F_Y and joint distribution function $F_{(X,Y)}$. A two-dimensional copula is a function $C : [0,1]^2 \to [0,1]$ with

$$F_{(X,Y)}(x,y) = C(F_X(x), F_Y(y)) = C(u,v), \qquad u,v \in [0,1], \tag{4.A.1}$$

with $u = F_X(x)$ and $v = F_Y(y)$ and further properties for C.[12] Thus, a copula describes the entire dependence structure between the one-dimensional random variables and is hence more complex than standard correlation measures. Sklar (1959) shows the uniqueness of copulas, i.e., for arbitrary but fixed continuous random variables there exists exactly one copula. Just as there are a variety of distributions for random variables, there are a variety of copulas. In case of independence of two random variables the interdependence is given by the Product copula

$$C(u,v) = uv, \qquad u,v \in [0,1], \tag{4.A.2}$$

because $F_{(X,Y)}(x,y) = F_X(x)F_Y(y)$ in case of independence. This assumption is equivalent to a separate modeling of times-to-default and LGDs and ignores interdependences between both. In contrast, the bivariate standard Gaussian distribution with the cumulative density function Φ_2 depends on the correlation parameter $\theta \in [-1,1]$ and offers another well known copula. The marginals itself are transformed by the univariate standard Gaussian distribution. The corresponding interdependences are described by the

[12]These properties are: (i) $C(u,0) = C(0,v) = 0$, (ii) $C(u,1) = u, C(1,v) = v$ and (iii) $C(u_1,v_1) - C(u_1,v_2) - C(u_2,v_1) + C(u_2,v_2) \geq 0$ for any $u,v,u_1,u_2,v_1,v_2 \in [0,1]$ with $u_1 > u_2$ and $v_1 > v_2$.

Gaussian copula

$$C(u,v) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \theta), \qquad u, v \in [0,1]. \tag{4.A.3}$$

4.A.2 Densities

Let X and Y be univariate random variables and C the copula that describes the interdependence between both. In addition, define $U = F_X(X)$ and $V = F_Y(Y)$ as margins. It holds:

$$\begin{aligned}
f_{(X,Y)}(x,y) &= \frac{\partial^2 F_{(X,Y)}(x,y)}{\partial x \partial y} && \text{(density as derivative of c.d.f.)} \\
&= \frac{\partial}{\partial y} \cdot \frac{\partial C\Big(F_X(x), F_Y(y)\Big)}{\partial x} && \text{(double as two simple derivatives)} \\
&= \frac{\partial}{\partial y} \cdot \frac{\partial C(u,v)}{\partial u} \cdot \frac{\partial u}{\partial x} && \text{(chain rule)} \\
&= \frac{\partial}{\partial y} \cdot \frac{\partial C(u,v)}{\partial u} \cdot f_X(x) && \text{(density as derivative of the c.d.f.)} \\
&= \frac{\partial}{\partial v} \cdot \frac{\partial C(u,v)}{\partial u} \cdot f_X(x) \cdot f_Y(y) && \text{(repeat both last steps)} \\
&= c(u,v) \cdot f_X(x) \cdot f_Y(y) && \text{(copula density as derivative of the copula).}
\end{aligned} \tag{4.A.4}$$

4.A.3 Properties

Table 4.A.1: Copula properties

Copula	$C_\theta(u,v),\ u,v \in [0,1]$	Generator $\varphi_\theta(t)$	Parameter space for θ	Possible ranges for τ
AMH	$\frac{uv}{1-\theta(1-u)(1-v)}$	$\log\frac{1-\theta(1-t)}{t}$	$[-1,1)$	$[-0.1817,\frac{1}{3})$
Clayton	$(u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}}$	$\frac{1}{\theta}(t^{-\theta}-1)$	$(-\infty,0)\cup(0,\infty)$	$(-1,0)\cup(0,1)$
Frank	$-\frac{1}{\theta}\log\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{(e^{-\theta}-1)}\right)$	$-\log\frac{e^{-\theta t}-1}{e^{-\theta-1}}$	$(-\infty,\infty)$	$(-1,1)$
Gaussian	$\Phi_2(\Phi^{-1}(u),\Phi^{-1}(v);\theta)$	-	$[-1,1]$	$[-1,1]$
Gumbel	$\exp\left(-\left((-\log(u))^\theta+(-\log(v))^\theta\right)^{\frac{1}{\theta}}\right)$	$(-\log(t))^\theta$	$[1,\infty)$	$[0,1)$
Joe	$1-\left((1-u)^\theta+(1-v)^\theta-(1-u)^\theta(1-v)^\theta\right)^{\frac{1}{\theta}}$	$-\log(1-(1-t)^\theta)$	$[1,\infty)$	$[0,1)$
Product	uv	$-\log(t)$	-	$[0,0]$
Student's t	$t_2(t^{-1}(u),t^{-1}(v);\theta)$	-	$[-1,1]$	$[-1,1]$

Notes: $C_\theta(u,v)$ denotes the cumulative probability function of a copula. For Archimedeans it is defined by a generator φ_θ via $C_\theta(u,v)=\varphi_\theta^{-1}(\varphi_\theta(u)+\varphi_\theta(v))$. The strength of dependence is given by the parameter θ and can be measured by Kendall's τ (rank correlation). Student's t copula is given by the Student's t cumulative probability function t (univariate) and t_2 (bivariate).

Appendix 4.B Empirical Robustness Tests

We estimated different combinations of time-to-default (AFT) models and dependence structures (copulas). Table 4.B.1 shows the goodness-of-fit measures Log-Likelihood, McFadden's R^2, and the resulting dependence measures by using the extended set of covariates and a log-logistic distribution in the AFT model. The best model choice is the *Frank copula* in combination with a log-logistic distribution for the time-to-default, because this combination provides the best performance, measured by the Log-Likelihood and McFadden's R^2.

Table 4.B.1: Copula performance

Copula	AMH	Clayton	Frank	Gaussian
Log(Likelihood)	-3,659.0	-3,707.9	-3,627.7	-3,694.0
McFadden's R^2	0.2524	0.2424	0.2588	0.2452
θ (RR)	0.8721	0.3564	5.3281	0.3711
τ (LGD)	-0.2650***	-0.1513***	-0.4767***	-0.5505***

Copula	Gumbel	Joe	Student's t	Product
Log(Likelihood)	-3,687.1	-3,710.1	-3,703.9	-3,715.5
McFadden's R^2	0.2466	0.2419	0.2432	0.2408
θ (RR)	1.6777	1.6624	0.5915	-
τ (LGD)	-0.4039***	-0.2697***	-0.8011***	0

Notes: Significance for Kendall's τ is indicated at 10% (*), 5% (**) and 1% (***) levels and computed as in Kendall (1938). McFadden's R^2 is the pseudo R^2 defined in McFadden (1974). All results are presented for log-logistic distributed times-to-default and with covariates of the Extended Model. We estimate the model with recovery rate (RR) data, as some copulas do not allow for negative interdependence. Thus, the copula parameter θ is given for RR modeling and the rank correlation measure Kendall's τ is given for LGD modeling. Testing for positive and negative interdependences show evidence for the negative case.

The Unrestricted Model shows strong evidence for negative dependencies between times-to-default and LGDs, e.g., the rank correlation measure Kendall's τ for the extended set of control variables is -0.4767. As stated in Section 4.2.4, the Joe and the Gumbel copula do not allow for negative dependencies. Thus, we report copula parameter estimates (θ) for the dependencies between default times and recoveries. The rank measure Kendall's τ and the beta regression results are transformed for interpretation of loss data. The negative dependence between default times and LGDs implies that the earlier a bond defaults, the higher its loss and vice versa. A default that happens early after the origination of a bond and after controlling for covariates is usually triggered by a surprising and severe cause, which implies higher losses. In contrast, a later default demonstrates a greater financial strength of the issuer. Thus, a lower loss for future defaults is plausible.

The choice of the distribution in the marginal AFT model does not substantially affect the results. Table 4.B.2 shows the Log-Likelihood, McFadden's R^2, the copula parameter estimate and Kendall's τ for the Frank copula using the extended set of covariates and Weibull, log-logistic and log-normal distributions in the AFT model. The best fit is provided by the log-logistic distribution with similar results for the Weibull or log-normal specification of the dependence measures.

Table 4.B.2: Default model performance

AFT model	Weibull	Log-logistic	Log-normal
Log(Likelihood)	-3,649.3	-3,627.7	-3,614.9
McFadden's R^2	0.2558	0.2588	0.2546
θ (RR)	5.2650	5.3281	5.5278
τ (LGD)	-0.4730***	-0.4767***	-0.4883***

Notes: Significance for Kendall's τ is indicated at 10% (*), 5% (**) and 1% (***) levels and computed as in Kendall (1938). McFadden's R^2 is the pseudo R^2 defined in McFadden (1974). All results are presented for the Frank copula and with covariates of the Extended Model. The copula parameter θ is given for RR modeling and the rank correlation measure Kendall's τ is given for LGD modeling, see Table 4.B.1.

Chapter 5

The Impact of Loan Loss Provisioning on Bank Capital Requirements

This chapter is joint work with Daniel Rösch[1] and Harald Scheule[2], and corresponds to a working paper with the same name.

Abstract: This paper shows that the revised loan loss provisioning based on the International Financial Reporting Standards (IFRS) and the Generally Accepted Accounting Principles (GAAP) implies a reduction of Tier 1 capital which levies an additional burden on banks. The paper finds in a counterfactual analysis that these changes are more severe (i) during economic downturns, (ii) for credit portfolios of low quality, (iii) for banks that do not tighten capital standards during downturns, and (iv) under a more lenient definition of significant increase in credit risk (SICR) under IFRS. Hence, the provisioning rules further increase the procyclicality of bank capital requirements. Adjustments of the SICR threshold or capital buffers are suggested as ways to mitigate negative effects on the banking industry.

JEL classification: C51; G28; M48

Keywords: GAAP 326; IFRS 9; lifetime expected loss; loan loss provisioning; regulatory capital; SICR

[1]Chair of Statistics and Risk Management, Faculty of Business, Economics, and Business Information Systems, University of Regensburg, 93040 Regensburg, Germany. E-Mail: daniel.roesch@ur.de

[2]Finance Discipline Group, UTS Business School, University of Technology, Sydney, PO Box 123, Broadway NSW 2007, Australia. E-Mail: harald.scheule@uts.edu.au

5.1 Introduction

Loan loss provisioning has historically been based on the incurred loss model and increases following economic downturns (Laeven and Majnoni (2003) and Bikker and Metzemakers (2005)). Gunther and Moore (2003), Fonseca and González (2008) and Cummings and Durrani (2016) find that this approach has led to a non-transparent management of loss reserves and income smoothing. Hence, the International Accounting Standards Board (2014) and the Financial Accounting Standards Board (2016) decided to replace the existing standards with a more forward looking approach based on expected losses of financial instruments. The International Financial Reporting Standards 9 (IFRS 9) and Generally Accepted Accounting Principles Topic 326 (GAAP 326) are intended to ensure more transparency and less procyclicality (§ BC 16 and § BC 79 of International Accounting Standards Board (2011) and Financial Accounting Standards Board (2011)).

On the other hand, Basel's *regulatory capital requirements* under pillar I are designed to cover unexpected losses because expected losses have been recognized by loan loss provisioning and hence deducted from bank capital. The Basel Committee on Banking Supervision (2011, 2015a) acknowledges that the computation of risk measures differ in the regulatory and accounting definition. Basel defines loan loss provisions as the 12-month expected losses, whereas IFRS 9 defines loan loss provisions as the 12-month expected loss for unimpaired assets and as expected losses for the entire remaining lifetime for financial instruments that have experienced a significant increase in credit risk (SICR). GAAP 326 applies the expected lifetime loss concept to all assets regardless of whether they have experienced significant changes in credit risk. Furthermore, Basel excludes macroeconomic risk factors, while IFRS 9 and GAAP 326 consider the current economic state and forecasts of future states for the instruments that have experienced a SICR.

The European Banking Authority (2016a) and the European Commission (2016) expect a decrease of the Core Tier 1 capital (CET 1) ratio due to IFRS 9 and GAAP 326 and propose in accordance with the Basel Committee on Banking Supervision (2017) a transition phase of five years to lower the additional burden on banks. The Basel

Committee on Banking Supervision (2016b) points to the volatility of the new provisioning approach. This paper quantifies the magnitude of Tier 1 capital changes and the cyclicality of capital.

The paper offers the following contributions. First, it shows the link between IFRS 9 and GAAP 326 loan loss provisioning and Basel bank capital regulation.[3] Second, the impact on the eligible regulatory capital of IFRS 9 and GAAP 326 is analyzed in a counterfactual analysis by studying the IFRS 9 and GAAP 326 rules for US American bonds between 1991 and 2013, it being a period in which these rules were not applied. The analysis includes different economic periods, portfolio credit qualities, SICR thresholds as well as reinvestment strategies.

The paper explores the procyclical reduction of Tier 1 capital levels due to loan loss provisioning and how institutions might mitigate the impact in dependence of several factors: (i) portfolio quality, (ii) portfolio reinvestment strategy, and (iii) SICR criterion. The paper further analyzes how regulators may assist banks in these efforts.

The remainder of the paper is organized as follows. Section 5.2 describes theoretical requirements of IFRS 9 as well as GAAP 326 and the regulatory handling of provisions. Section 5.3 provides the data description. Section 5.4 estimates probabilities of default (PD) and loss rates given default (LGD) and computes 12-month expected losses as well as lifetime expected losses. A formula for the lifetime expected loss is developed and requirements on the SICR criterion are discussed. Finally, Section 5.5 shows the impact of expected loss based loan loss provisioning on the eligible regulatory capital and discusses implications for institutions, regulators and supervisors.

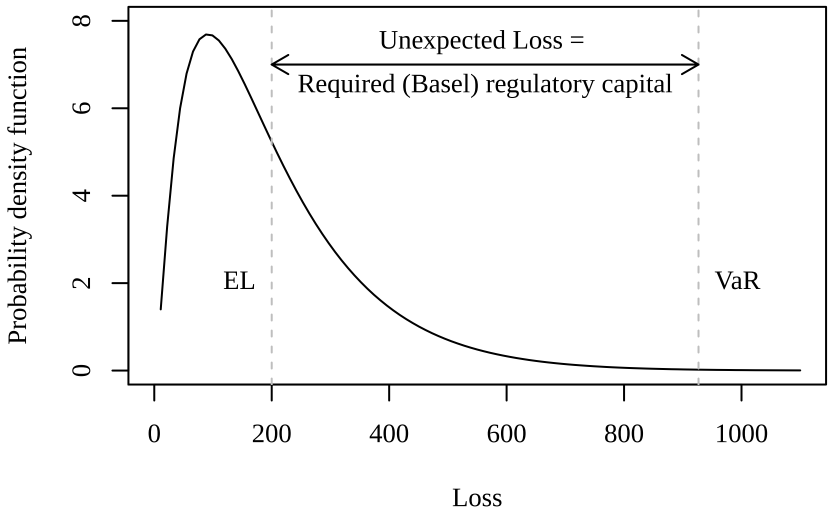
5.2 Capital Requirements and Provisioning

This paper analyzes the interaction between loan loss provisioning and bank capital. Figure 5.1 shows that financial institutions hold loan loss provisions for expected credit losses and capital for unexpected losses, i.e., the difference between the 99.9 % Value at

[3] We focus on institutions that use the internal ratings-based (IRB) approach. The framework for institutions using the standardized approach is different and will be revised in the future as discussed by the Basel Committee on Banking Supervision (2016b, 2017).

Risk and the expected losses.

Figure 5.1: The meaning behind capital and provisions



Notes: This exemplary loss distribution shows the principal links between provisions and regulatory capital. Loan loss provisioning represents the expected loss (EL) of an institution due to credit risk. Regulatory capital shall cover unexpected losses in 99.9 % of all possible future cases. The overall required amount is given by the Value at Risk (VaR).

5.2.1 Accounting Provisions

The International Accounting Standards Board (2011) and the Financial Accounting Standards Board (2011) propose to replace the incurred loss model for loan loss provisioning by an approach that recognizes expected losses to reflect the economic value of financial instruments. Two basic accounting regimes exist: The International Financial Reporting Standards (IFRS) and the United States Generally Accepted Accounting Principles (GAAP).

The International Accounting Standards Board (2014) introduces IFRS 9 and stipulates a three-stage model that will be mandatory from 2018 on. Financial instruments generally start in Stage 1 where the required provision is based on the 12-month expected loss, i.e., "the expected credit losses that result from default events on a financial instrument that are possible within the 12 months after the reporting date" (§ 5.5.5 and p. 53 IFRS 9). If the instrument's credit risk for the remaining lifetime significantly increases since initial recognition, it will be classified in Stage 2. Section 5.4.3 discusses the criterion for a significant increase in credit risk (SICR). In this second stage, the provision is calculated by the lifetime expected loss that is given by the "expected credit losses that result from all possible default events over the expected life of a financial instrument" (§ 5.5.3 and p. 56 IFRS 9). If an instrument becomes credit-impaired (i.e., is in default), it will be assigned to Stage 3 where the lifetime expected loss must also be recognized (p. 191 IFRS 9). If the conditions of Stage 2 or 3 are no longer met an instrument shifts back to Stage 1.

The Financial Accounting Standards Board (2016) updates GAAP on Topic 326 (GAAP 326). Thereby, institutions are obliged from 2020 on to recognize the "current estimate of all expected credit losses" (p. 3 GAAP 326) which is consistent with the lifetime expected loss of IFRS 9 in Stage 2. The board rejected the three-stage model of IFRS 9 due to lack of clarity of the SICR criterion, concerns about different measurements of identical instruments and potential for earnings management as well as cliff effects (p. 250 GAAP 326).

5.2.2 Basel Expected Loss and Capital Requirements

As mentioned above, Basel assumes that provisions cover expected losses whereas the required regulatory capital covers unexpected losses. The loan loss provisioning of IFRS 9 and GAAP 326 is based on expected loss computations which differ from the expected loss amount under the Basel regulation for a number of reasons. First, the time horizon differs on which possible losses need to be considered. The Basel framework is based on a 12-month horizon (e.g., § 285 Basel II, see Basel Committee on Banking Supervision (2006)) whereas accounting standards consider the entire remaining lifetime of at least some or even all financial assets. Second, economic conditions are differently treated. § 5.5.17 of IFRS 9 and § 20-30-9 of GAAP 326 oblige institutions to account for current economic conditions. In contrast, in the Basel regulation loan loss provisions are considered to abstract from macroeconomic risk. This section analyzes the implications of a difference between expected loss based provisions and the Basel expected loss on the calculation of the eligible regulatory capital.

The required regulatory capital under pillar I is based on unexpected credit losses that are caused by the credit risk on the asset side for a 12-month horizon and does not depend on current economic conditions. Any provisioning directly lowers the Common Equity Tier (CET) 1 on the liability side. However, the Basel Committee on Banking Supervision (2011, Basel III) makes an adjustment for possible shortfalls in provisioning. If the Basel expected loss is higher than the provisions, the difference must be deducted from the eligible CET 1 (§ 73 Basel III). The exact amount of provisions does not affect the eligible regulatory capital as long as there is a shortfall. The excess directly lowers the eligible CET 1 if provisions exceed Basel expected losses.[4] This case mainly occurs in recessions due to higher provision levels. As a result, the new accounting standards may require additional Core Tier 1 capital during downturns which we empirically study in Section 5.5.

Table 5.1 shows the treatment of shortfalls and excesses in the calculation of regulatory capital. The Basel framework distinguishes between three levels of capital that are built on each other: Core Tier 1 (CET 1), CET 1 capital plus additional Tier 1 capital, Tier 1 capital plus Tier 2 capital. Let the regulatory expected loss in both cases be 200 monetary units. The provisions for financial instruments may be 150 in an economic upturn, i.e., 50 less than required by Basel. The provisioning level in a downturn may be 250, i.e., 50 units more than required by Basel.

The example assumes that the initial CET 1 before the deduction of provisions is 1,000. The remaining CET 1 after provisioning is 850 (shortfall) in an economic upturn and 750 (excess) in an economic downturn. In the first case, the deficit of 50 must be deducted so that the eligible CET 1 amounts to 800 and is equal to the initial capital minus the Basel expected loss. However, an excess of the provisions directly lowers the eligible CET 1 to 750.

The additional Tier 1 capital is not affected by provisions and exemplary amounts to 100. Let the initial Tier 2 capital also be 100. The excess in provisions (which was deducted from CET 1) is added to Tier 2 capital. Whilst the total regulatory capital

[4]The excess may be added up to an amount of 0.6 % in terms of risk weighted assets (RWA) to Tier 2 capital (§ 61 Basel III).

Table 5.1: Exemplary calculation of regulatory capital

	Eligible capital		Required
	Upturn	Downturn	capital
Accounting provisions for financial instruments	150	250	
Basel regulatory expected loss	(200)	(200)	
CET 1 before provisions for financial instruments	1000	1000	
provisions for financial instruments	(150)	(250)	
CET 1 before regulatory adjustments due to provisions	850	750	
regulatory adjustments due to provisions	*(50)*	–	
CET 1 (Tier 1a)	800	750	4.5% of RWA
additional Tier 1	100	100	
Tier 1	**900**	**850**	6% of RWA
Tier 2 before regulatory adjustments due to provisions	100	100	
regulatory adjustments due to provisions	–	*50*	
Tier 2	**100**	**150**	
Tier 1 + Tier 2	**1000**	**1000**	8% of RWA

Notes: This table shows the calculation of the three regulatory capital amounts (CET 1, Tier 1, Tier 1 plus Tier 2). A positive difference between accounting provisions and the Basel expected losses (excess) leads to a deduction of the CET 1 and a potential addition to Tier 2 capital up to 0.6 % of RWA. A deficit must be deducted of the CET 1 to calculate the eligible regulatory capital.

(Tier 1 plus Tier 2) equals in both cases (1,000) the composition differs.

In summary, the amount of the required regulatory capital does not depend on a shortfall or excess of the provisions, whereas the amount of eligible CET 1 does.

Financial institutions generally need to hold in relation to the risk weighted assets 4.5 % CET 1, 6 % Tier 1 capital and 8 % Tier 1 plus Tier 2 capital. In addition to these requirements, institutions must provide three additional CET 1 buffers that are currently phased in: capital conversion buffer (2.5 percentage points), countercyclical capital buffer (0 - 2.5 percentage points, depending on the current economic state) and systemic risk buffer (0 - 3.5 percentage points, depending on the institution's systematic relevance). The results of Carlson et al. (2013) and Repullo (2013) show the need of cyclical capital adjustments due to procyclical effects of Basel regulatory capital requirements on lending.

We contribute to this discussion by clarifying the role of future provisioning.

In summary, IFRS 9 and GAAP 326 may increase the pressure to raise high-quality capital for banks. Since the upcoming regulatory capital buffers are currently introduced, the new accounting standards may strengthen the existing pressure to raise high-quality capital.

5.3 Data

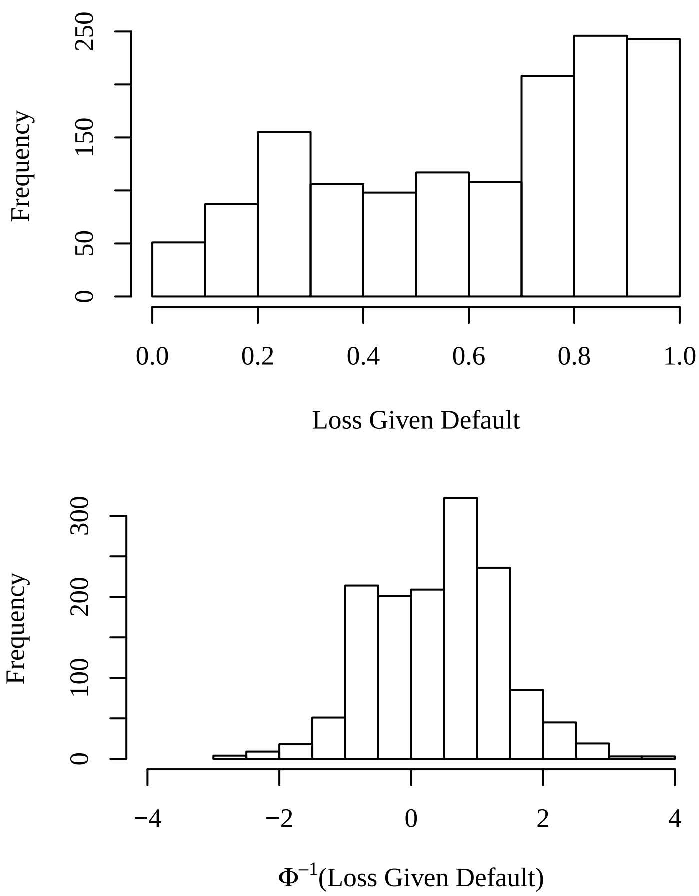
Our credit risk models are estimated using the Moody's Default and Recovery Database and macroeconomic risk factors provided by the FRED database from the Federal Reserve Bank in St Louis. US American bonds with issuance after 1990 are selected and a yearly panel dataset is set up, covering all years until 2013. After removal of observations with missing information in any of the variables used in this study 181,066 bond-years remain including 35,300 bonds and 1,419 defaults.

Figure 5.2 shows in the upper panel the empirical distribution of the expected loss rate at default (LGD) that is computed by one minus the ratio of the bond price 90 days after default and the par value. The mean and the median LGD are 61.69 % and 70.00 % and indicate a left-skewed distribution with a standard deviation of 27.82 %. Consistent with Chava et al. (2011) and Altman and Kalotay (2014) we transform the LGD (that is a rate) by the inverse Gaussian cumulative distribution function Φ^{-1} to provide a dependent variable on the full range of the OLS model (lower panel of Figure 5.2).

5.3.1 Issuer- and Bond-Specific Covariates

We account for several issuer- and bond-specific covariates that are shown in Table 5.2 with corresponding means of realized LGDs and yearly default rates. Moody's long-term ratings are included as key proxies for default and loss risk and are categorized into four groups, i.e., Aaa - Baa for investment grade bonds, Ba, B, and Caa - C. The historical default rate increases when creditworthiness decreases, e.g., from 0.02 % for investment grade to 13.16 % for the lowest ratings. This tendency can also be observed for the

Figure 5.2: Empirical distribution of Losses Given Default



Notes: Losses given default are calculated by 1 minus the ratio of the bond price 90 days after default and the par value (left panel). These values are transformed by the inverse Gaussian cumulative distribution function Φ^{-1} for a better regression handling (right panel).

loss severity of speculative grade ratings with mean LGDs of 47.97 % - 65.63 %. The lower number of investment grade defaults of 25 limits the interpretation of LGD for this category.

Moody's rating adjustments are caused by significant changes in a bond's credit risk and may indicate a significant increase in credit risk (SICR) for IFRS 9. Following a downgrade, the default rate of a bond increases from 0.43 % to 2.44 %. Thus, we include a downgrade dummy variable that equals one if there was a downgrade of at least one notch in Moody's granular ratings in the past.

The seniority characterizes the position in a bond's post default order of payments. Senior secured bonds are first repaid and have a first lien on collateral and have lowest mean LGDs of 48.36 %. They are then followed by senior unsecured, senior subordinated,

Table 5.2: Descriptive statistics for issuer- and bond-specific covariates

	Default rate in %	Mean LGD in %	# obs.
Rating			
Aaa - Baa	0.02	68.14	100,549
Ba	0.41	47.97	58,339
B	2.30	61.75	16,238
Caa - C	13.16	65.63	5,940
Downgrade			
No	0.43	64.54	148,959
Yes	2.44	59.36	32,107
Seniority			
Senior Secured	0.95	48.36	13,970
Senior Unsecured	0.63	59.14	147,351
Senior Subordinated	3.85	72.69	8,827
Subordinated	0.20	79.33	10,918
Industry			
Banking	0.09	72.72	28,387
Capital Industries	1.59	67.03	22,963
Consumer Industries	1.28	63.33	18,675
Energy & Environment	0.73	61.49	14,375
Finance. Insurance & Real Estate	0.37	38.95	49,480
Media & Publishing	1.96	57.82	6,800
Retail & Distribution	1.35	63.48	7,638
Technology	1.40	72.46	13,391
Transportation	1.70	67.39	3,707
Utilities	0.08	21.29	15,650
Total maturity			
Short-term	0.19	31.54	7,711
Medium-term	1.00	58.02	73,717
Long-term	0.67	66.40	99,638
Time to maturity in years (TTM)			
$0<\text{TTM}\leq 1$	0.29	32.09	17,300
$1<\text{TTM}\leq 2$	0.64	48.07	17,078
$2<\text{TTM}\leq 3$	0.78	54.49	16,184
$3<\text{TTM}\leq 4$	0.89	55.50	15,047
$4<\text{TTM}\leq 5$	1.07	57.98	15,431
$5<\text{TTM}$	0.83	68.07	100,026

Notes: This table shows for each categorical explanatory variable the default rate, the mean realized LGD and the number of observations.

and subordinated bonds that result in a loss of 79.33 %. Default rates are driven by other issuer- and bond-specific information next to seniority and security.

Industries have been identified as key credit risk drivers (Acharya et al. (2007)). The default rate is lowest for the Utilities sector with 0.08 % and highest for Media & Publishing with 1.96 %. The loss rate varies between 21.29 % for Utilities and 72.72 % for Banking.

The credit risk of a bond generally depends on two time components that are particularly relevant in the context of lifetime expected losses: (i) the total maturity that is the timespan from issuance to maturity date, and (ii) the stage in the life of a financial instrument. First, we split the sample into three categories of total maturity: short-term (up to three years), medium-term (more than three but less than or equal to ten years) and long-term (more than ten years). The lowest default rate is realized by short-term bonds with 0.19 % in contrast to 1.00 % of medium-term bonds and 0.67 % for long-term bonds. The LGD varies between 31.54 % (short-term) and 66.40 % (long-term).

We take into account a possible term structure of credit risk by the inclusion of the remaining time to maturity (TTM) that is given by the time in years from the beginning of the observation year up to the last day of maturity. As the given metrics are conditional, i.e., given a bond does not default prior to the observation year, the credit risk seems to decrease with maturity. In other words, surviving bonds have lower default rates and LGDs at the end of their maturity.

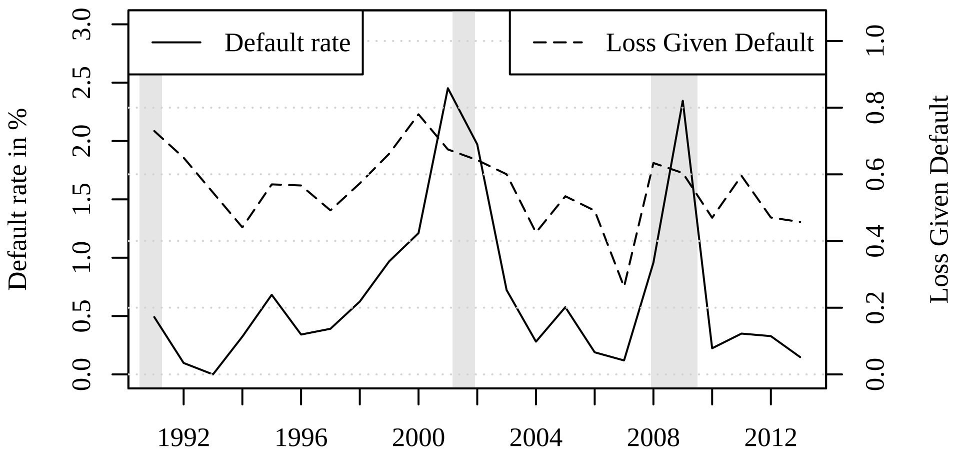
5.3.2 Cyclical Behavior

In addition to issuer- and bond-specific covariates, macroeconomic conditions affect credit risk. Figure 5.3 shows the cyclical behavior of yearly default rates and LGDs over time. The shaded areas indicate economic downturns as indicated by the National Bureau of Economic Research. Defaults are clustered in the crisis of 2001 and the Global Financial Crises (2008/2009).

The computation of provisions requires estimates of the expected loss based on the current economic state (§ 5.5.17 and § B5.5.49 IFRS 9, § 20-30-9 GAAP 326). This approach is also known as Point-in-Time (PIT) rating philosophy.[5] In contrast, the Basel Committee on Banking Supervision (2006) aims to avoid procyclical patterns of regulatory requirements. The risk parameters of the Basel formula under Pillar 1 must be modeled using the Through-the-Cycle (TTC) philosophy (§ 447 Basel II). This implies the exclu-

[5]The rating philosophies Point-in-Time and Through-the-Cycle are commonly used terms for the handling of macroeconomic conditions in credit risk models. This paper follows the classification of the International Accounting Standards Board (2014), the Financial Accounting Standards Board (2016) and the Basel Committee on Banking Supervision (2015a, 2016c).

Figure 5.3: Default rates and mean realized Losses Given Default



Notes: The solid line shows the default rate and the dashed line shows the mean realized LGD for each year. The shaded areas indicate recession dates of the National Bureau of Economic Research.

sion of macroeconomic risk factors. The remaining time-variation of risk is exclusively driven by time-varying idiosyncratic risk factors and changes in the risk population.

As the requirements for the computation of expected losses differ with respect to the inclusion of macroeconomic variables, we distinguish between a PIT and a TTC model. We study the impact of several macroeconomic variables in order to provide a PIT model as required for accounting purposes. Macroeconomic information of the financial year is used to estimate the expected loss for IFRS 9 and GAAP 326.

The literature proposes a variety of macroeconomic variables for modeling credit risk. Economic upturn (downturn) conditions result in lower (higher) default rates and LGDs. This paper investigates the role of the growth in gross domestic product (GDP), the historic default rate (of the total dataset without timely restriction), the TED spread (difference between three-month LIBOR and three-month US trasury bill), US treasury rates for the one year and ten year horizon, the treasury term spread between both treasury rates, the unemployment rate and the CBOE volatility index VIX. Appendix 5.A shows descriptives statistics. The suitability of the variables is mentioned in Section 5.4.1.

Appendix 5.A discusses descriptives and the suitability of those macroeconomic variables (for the latter see also Section 5.4.1).

5.4 Loan Loss Provisioning

5.4.1 12-month Expected Loss for Basel and IFRS 9 (Stage 1)

We model the risk parameters probability of default and loss rate given default for a 12-month horizon for Basel and accounting purposes.

Probability of Default (PD)

The default behavior of financial instruments was considerably investigated by the Z-score of Altman (1968), the firm value model of Merton (1974) and the categorical default model as discussed in Campbell et al. (2008) and Hilscher and Wilson (2016) amongst others. In accordance with these approaches, we model the PD by a Probit model which follows, e.g., Puri et al. (2017). The regression equation for the PD of bond i in year t is given by

$$\mathrm{PD}_{it} = \mathrm{P}(\mathrm{D}_{it} = 1 | x_{it-1}) = \Phi(x_{it-1}\beta), \tag{5.1}$$

where x_{it-1} is the vector of covariates (including an intercept) of the previous year and unknown parameter vector β. The default indicator D_{it} equals one for defaults and zero for non-defaults. We estimate this model with two different sets of variables in order to meet the different requirements of accounting standards and the Basel framework. In a first setting, we include all issuer- and bond-specific information in order to provide a TTC approach for Basel purposes. The PIT model for provisioning is extended by including macroeconomic information.

Table 5.3 shows the parameter estimates for the PD models. The PDs increase with credit ratings from Aaa-Baa to Caa-C. A downgrade of at least one rating notch significantly increases the PD.

The issuer's industry affiliation captures industry-specific effects. Although parameter estimates are not statistically significantly different from zero, the corresponding variables increase the goodness of fit. The Utilities sector implies the lowest PDs else being equal. In contrast, the Transportation sector leads to the highest PDs.

Table 5.3: Parameter estimates Probability of Default

	Through-the-Cycle	Point-in-Time
(Intercept)	-3.1573 ***	-4.5761 ***
	(0.2501)	(0.4601)
Ba	0.2429	0.2806
	(0.1885)	(0.1844)
B	0.8954 ***	0.9550 ***
	(0.1000)	(0.0917)
Caa - C	1.6202 ***	1.6717 ***
	(0.0961)	(0.1084)
Downgrade	0.3934 ***	0.3883 ***
	(0.1123)	(0.1069)
Capital Industries	0.2787	0.2972
	(0.2936)	(0.3124)
Consumer Industries	0.1851	0.1955
	(0.2384)	(0.2521)
Energy & Environment	0.2033	0.2224
	(0.3236)	(0.3434)
Finance, Insurance & Real Estate	0.2063	0.1772
	(0.2734)	(0.2759)
Media & Publishing	0.3599	0.3638
	(0.3077)	(0.3256)
Retail & Distribution	0.1735	0.1970
	(0.3107)	(0.3251)
Technology	0.3965	0.3810
	(0.3524)	(0.3684)
Transportation	0.4618	0.5252
	(0.3353)	(0.3500)
Utilities	-0.1991	-0.2110
	(0.3754)	(0.3846)
Short-term	0.1238	0.0746
	(0.0868)	(0.0981)
Long-term	-0.1443 ***	-0.1625 ***
	(0.0415)	(0.0471)
$0<\text{TTM}\leq1$	-0.2851 ***	-0.2907 ***
	(0.0942)	(0.1057)
$1<\text{TTM}\leq2$	-0.0279	-0.0260
	(0.0586)	(0.0632)
$2<\text{TTM}\leq3$	-0.0243	-0.0228
	(0.0671)	(0.0714)
$3<\text{TTM}\leq4$	-0.0739	-0.0787
	(0.0707)	(0.0734)
$4<\text{TTM}\leq5$	-0.0156	-0.0219
	(0.0542)	(0.0572)
VIX		0.0625 ***
		(0.0143)
Accuracy Ratio	0.8285	0.8729
McFadden's adjusted R^2	0.2896	0.3282
# observations	181,066	181,066

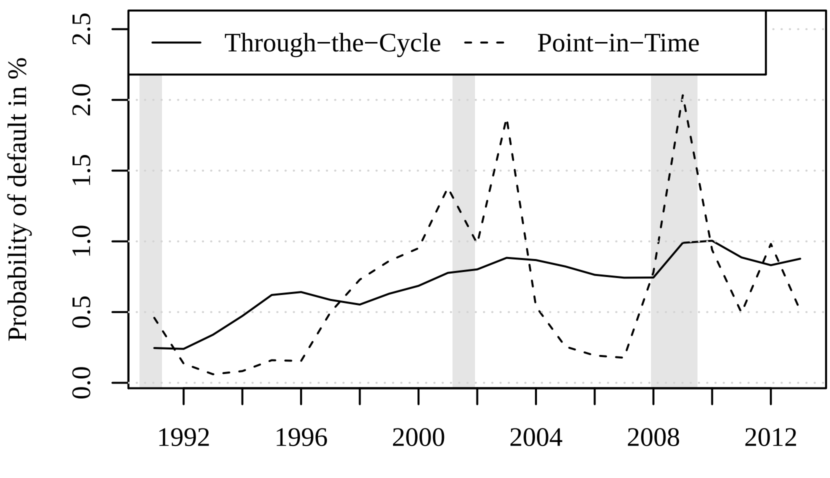
Notes: The table shows regression results for the PD models that are used for the computation of regulatory (TTC) and accounting (PIT) expected losses. These Probit models are based on Equation (5.1). Standard errors are given in parentheses and clustered for issuer- and year-specific fixed effects as proposed in Petersen (2009). The significance is indicated for the 1 % (***), 5 % (**) and 10 % (*) level.

The total length of maturity and the remaining time to maturity affect the PD. Corporates with high creditworthiness generally issue bonds with longer maturities due to the higher trust of lenders. Risky borrowers are generally forced to issue bonds with shorter maturities. The default risk declines over time and is particularly low in the year prior to maturity.

We test several macroeconomic variables for inclusion in the PIT model of accounting standards (see Appendix 5.A). We do not include more than one macroeconomic variable because correlations between variables are high and the marginal improvement of the fit is low while the complexity of forecasting multiple variables for multiple periods and hence the model risk is substantially greater. Bloom (2009) and Jo and Sekkel (2017) show that the VIX predicts future economic states. Consistent with this literature, the PDs increase with VIX. This study empirically identifies that the VIX has the highest goodness of fit for the PD model. In comparison to the TTC model, the Accuracy Ratio (McFadden's adjusted R^2) increases from 82.85 % (28.96 %) to 87.29 % (32.82 %).

Figure 5.4 shows the mean estimated PD for each year. The PIT model provides more cyclical PD estimates than the TTC model as it includes a macroeconomic variable, next to idiosyncratic risk factors and changes in the population over time. The remaining variation is caused by the changing composition of the dataset.

Figure 5.4: Mean predicted Probability of Default



Notes: The PD for each observation is predicted by the Probit model of Table 5.3. The figure shows the resulting means for each year. The shaded areas indicate recession dates of the National Bureau of Economic Research.

Loss Rate Given Default (LGD)

Acharya et al. (2007), and Jankowitsch et al. (2014) amongst others use OLS regression models for recovery and LGD models. Consistent with Chava et al. (2011) and Altman and Kalotay (2014), we transform the LGD (that is rate) by the inverse Gaussian cumulative distribution function Φ^{-1} to provide a dependent variable on the full range of the OLS model. The regression equation of a bond i in year t is given by

$$\Phi^{-1}(\mathrm{LGD}_{it}) = z_{it-1}\gamma + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0, \sigma^2), \tag{5.2}$$

with a covariate vector z_{it-1} that includes an intercept and information of the previous year. The unknown components of the model are the parameter vector γ and the standard deviation σ.

Similar to the PD modeling we consider a TTC and a PIT model for Basel and accounting requirements. Table 5.4 shows the corresponding estimation results. Covariate effects on LGDs are generally consistent with the ones of PDs.

The seniority determines the order of the borrower's payments after default and has a significant effect on LGDs. The results show higher losses for lower seniority and security levels. Industry-specific effects are significant in comparison to the reference group Banking that provides the highest LGDs. The Utilities sector shows the lowest loss rates in addition to the lowest default risk. The total length of maturity does not cause significant variation in recoveries. LGDs significantly decrease over lifetime due to survivorship.

A high uncertainty—measured by an increased VIX—strengthens loss severity. The advantages of the VIX for inclusion in the LGD model in terms of goodness of fit is discussed in greater detail in Appendix 5.A. The PIT model shows an adjusted R^2 of 25.36 % and dominates the TTC model with an adjusted R^2 of 21.74 %.

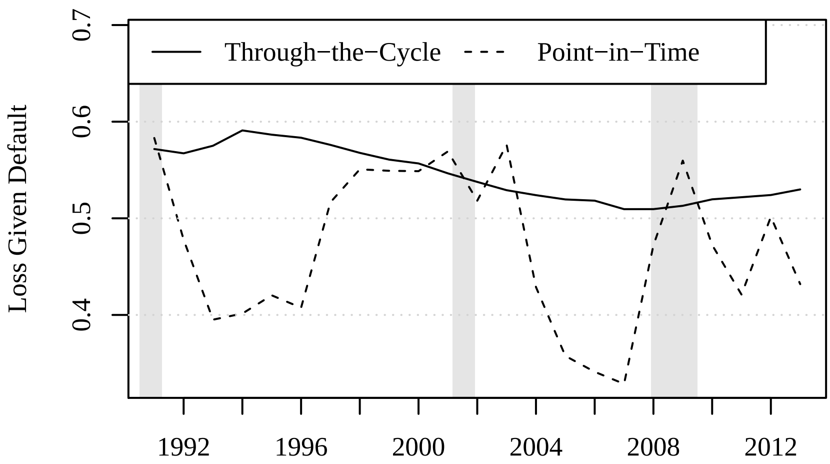
Figure 5.5 shows the mean estimated LGDs for each year. The PIT model shows procyclical patterns whereas the TTC model does not.

Table 5.4: Parameter estimates Loss Rate Given Default

	Through-the-Cycle	Point-in-Time
(Intercept)	0.5032 ***	-0.4139 *
	(0.1139)	(0.2282)
Ba	-0.1197	-0.0740
	(0.1266)	(0.1755)
B	0.0442	0.1218
	(0.1049)	(0.0917)
Caa - C	0.2528	0.3173 **
	(0.1721)	(0.1568)
Senior Unsecured	0.4363 ***	0.3661 ***
	(0.1067)	(0.1087)
Senior Subordinated	0.7526 ***	0.6808 ***
	(0.0980)	(0.0980)
Subordinated	1.0520 ***	0.9693 ***
	(0.3091)	(0.3350)
Capital Industries	-0.4545 **	-0.5508 ***
	(0.2181)	(0.1873)
Consumer Industries	-0.6057 ***	-0.6791 ***
	(0.2144)	(0.1822)
Energy & Environment	-0.6039 **	-0.6296 **
	(0.2854)	(0.2588)
Finance, Insurance & Real Estate	-1.0137 ***	-1.1640 ***
	(0.1791)	(0.1990)
Media & Publishing	-0.5922 **	-0.6818 ***
	(0.2740)	(0.2350)
Retail & Distribution	-0.5127 **	-0.5525 ***
	(0.2282)	(0.2010)
Technology	-0.3037	-0.4221 *
	(0.2677)	(0.2330)
Transportation	-0.3688 *	-0.3156
	(0.2131)	(0.2372)
Utilities	-1.5577 ***	-1.7220 ***
	(0.3416)	(0.3127)
Short-term	0.0892	0.0696
	(0.2413)	(0.2371)
Long-term	-0.0142	-0.0133
	(0.0602)	(0.0509)
$0<\text{TTM}\leq1$	-0.9018 ***	-0.8457 ***
	(0.1158)	(0.1176)
$1<\text{TTM}\leq2$	-0.5174 ***	-0.4972 ***
	(0.1053)	(0.0971)
$2<\text{TTM}\leq3$	-0.3113 ***	-0.2775 **
	(0.1177)	(0.1170)
$3<\text{TTM}\leq4$	-0.2791 **	-0.2689 **
	(0.1130)	(0.1051)
$4<\text{TTM}\leq5$	-0.2370 ***	-0.2155 ***
	(0.0908)	(0.0831)
VIX		0.0417 ***
		(0.0083)
Adjusted R^2	0.2174	0.2536
# observations	1,419	1,419

Notes: The table shows regression results for the LGD models that are used for the computation of regulatory (TTC) and accounting (PIT) expected losses. These OLS models are based on Equation (5.2). Standard errors are given in parentheses and clustered for issuer- and year-specific fixed effects as proposed in Petersen (2009). The significance is indicated for the 1 % (***), 5 % (**) and 10 % (*) level.

Figure 5.5: Mean predicted Loss Given Default



Notes: The LGD for each observation is predicted by the OLS model of Table 5.4. The figure shows the resulting means for each year. The shaded areas indicate recession dates of the National Bureau of Economic Research.

5.4.2 Lifetime Expected Loss for GAAP 326 and IFRS 9 (Stage 2)

Macroeconomic Forecasts

Lifetime expected losses for GAAP 326 and IFRS 9 (Stage 2) must contain information on the current economic state, which changes over the remaining lifetime of an instrument and multi-period forecasts are necessary (§ B5.5.49 IFRS 9 and § 20-30-9 GAAP 326). This paper uses an autoregressive (AR) process for forecasting.

Figure 5.6 shows the time-series plot of the VIX in the upper panel. The autocorrelation and the partial autocorrelation function (lower panel) suggest an AR process of order one. Hence, the difference of the VIX in year t to the mean φ_0 is modeled by

$$\mathrm{VIX}_t - \varphi_0 = \varphi_1(\mathrm{VIX}_{t-1} - \varphi_0) + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_\epsilon^2), \tag{5.3}$$

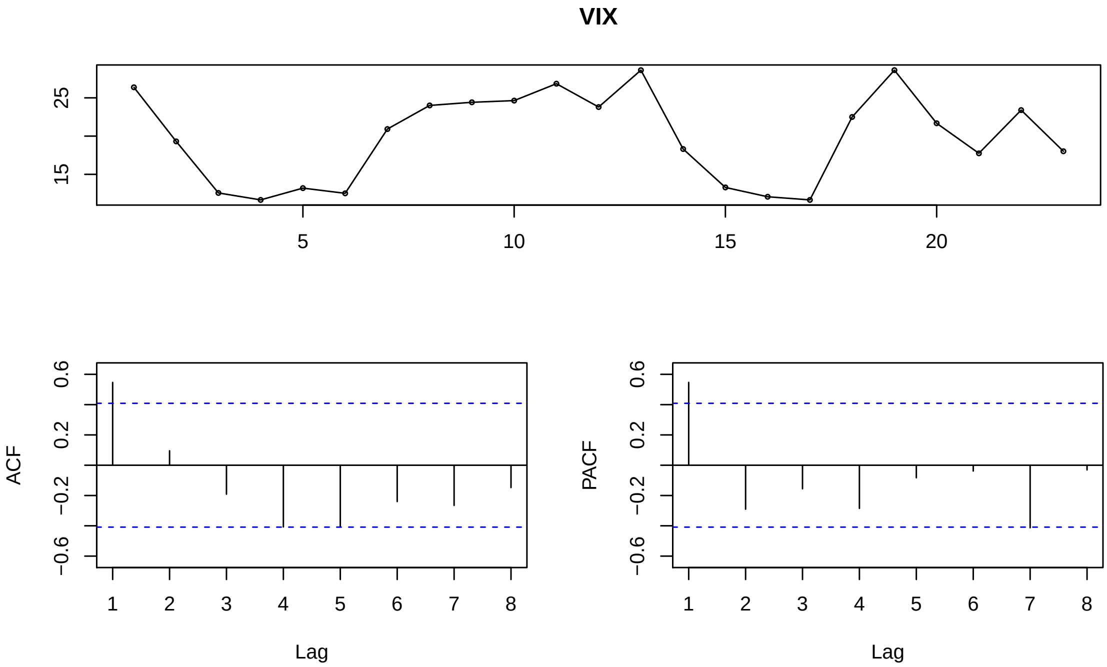
with unknown parameters φ_0, φ_1 and σ_ϵ. Note that AR processes converge to the long run mean over time. Table 5.5 shows the estimation results.

The estimated long-run average of the VIX is 20.07 percentage points. The AR parameter estimate for the lag amounts to 0.5580. It is statistically significantly different from zero and indicates stationarity. The forecast of the VIX for s years ahead given a realization in year t is given by

$$\widehat{\mathrm{VIX}}_{t+s} = (1 - \hat{\varphi}_1^s)\hat{\varphi}_0 + \hat{\varphi}_1^s \, \mathrm{VIX}_t, \tag{5.4}$$

where $\hat{\varphi}_1^s$ is the s-th power of the estimated AR parameter.

Figure 5.6: VIX and corresponding ACF and PACF plot



Notes: The figure shows the time-series of the VIX between 1990 and 2012 (upper panel). The autocorrelation function (ACF) and the partial autocorrelation function (PACF) suggest a time lag of one year for the autoregressive model.

Table 5.5: Parameter estimates AR model

	VIX
$\hat{\varphi}_0$	20.0687 ***
	(2.1094)
$\hat{\varphi}_1$	0.5580 ***
	(0.1705)
AIC	142.83

Notes: The table shows regression results for the VIX model that is used for the computation of lifetime expected losses. This AR model is based on Equation (5.3) and estimates the variance by 22.08. Standard errors are given in parentheses. The significance is indicated for the 1 % (***), 5 % (**) and 10 % (*) level.

Prediction of Lifetime Expected Losses

The upcoming accounting standards require the computation of 12-month and lifetime expected losses. Both measures must account for current economic conditions. Hence, we use risk parameters based on PIT models to provide accounting expected losses.

In Stage 1 of IFRS 9, provisions are given by the 12-month expected loss. If the time to maturity of an instrument is less than 12 months, the remaining lifetime is crucial for the computation (§ B5.5.43 IFRS 9). The expected loss of a regular bond is principally given by the product of the PD and the LGD.[6] We denote the information that is available up to year t by $\mathcal{F}_t$. Hence, the estimated Stage 1 expected loss of instrument i for year t is

$$\widehat{\mathrm{EL}}_{it}^{PIT} = \widehat{\mathrm{P}}(\mathrm{D}_{it} = 1|\mathcal{F}_t) \cdot \widehat{\mathrm{E}}(\mathrm{LGD}_{it}\,|\mathcal{F}_t) \cdot \min(1, \mathrm{TTM}_{it}), \tag{5.5}$$

where $\widehat{\mathrm{P}}(\mathrm{D}_{it} = 1|\mathcal{F}_t)$ is the estimated PD using Equation (5.1) and $\widehat{\mathrm{E}}(\mathrm{LGD}_{it}\,|\mathcal{F}_t)$ is the estimated LGD using Equation (5.2). Both calculations use lagged covariates, i.e., provisions in a financial year $t-1$ are based on the available information of that year and correspond to the expected loss for the following year t. TTM_{it} denotes the time to maturity that is left at the reporting date.

In GAAP 326 as well as Stage 2 of IFRS 9 the provision for an instrument shall represent the lifetime expected loss. This amount is the sum of the expected losses of all remaining years up to maturity. The loss contribution of future years must be discounted to account for the time value of money. Accounting standards require the consideration of “the contractual terms of the financial instrument” (p. 55 IFRS 9) and “the financial asset’s effective interest rate” (§ 20-30-4 GAAP 326). Consistent with this we use the contractual coupon rate r_i of bond i as discount rate.

[6] The exposure of a regular bond is deterministic. For the empirical study we assume a constant exposure of one monetary unit.

Hence, the lifetime expected loss for instrument i in year t can be calculated by

$$\widehat{\mathrm{LEL}}_{it} = \sum_{\Delta t=0}^{\lfloor \mathrm{TTM}_{it} \rfloor} \Bigg[\widehat{\mathrm{P}}(\mathrm{D}_{it+\Delta t} = 1, \mathrm{D}_{it+s} = 0 \,\forall\, s \in \mathbb{Z} : 0 \leq s \leq \Delta t - 1 | \mathcal{F}_t) \cdot \frac{\widehat{\mathrm{E}}(\mathrm{LGD}_{it+\Delta t} | \mathcal{F}_t)}{(1+r_i)^{\Delta t}} \cdot \min(1, \mathrm{TTM}_{it} - \Delta t) \Bigg], \tag{5.6}$$

where $\widehat{\mathrm{P}}(\mathrm{D}_{it+\Delta t} = 1, \mathrm{D}_{it+s} = 0 \,\forall\, s \in \mathbb{Z} : 0 \leq s \leq \Delta t - 1 | \mathcal{F}_t)$ is the estimated probability that an instrument defaults in and not prior to year $t + \Delta t$. The time-varying LGDs are included by the term $\widehat{\mathrm{E}}(\mathrm{LGD}_{it+\Delta t} | \mathcal{F}_t)$ and again calculated by Equation (5.2). In contrast to the 12-month expected loss, it is essential here to use predictions for the VIX, i.e., for the Δt-th year in the future we forecast the VIX Δt years ahead by Equation (5.4). Furthermore, we subsequently lower the time to maturity over a bond's lifetime. Again, the last year is only partly considered by the factor $\mathrm{TTM}_{it} - \lfloor TTM_{it} \rfloor$ where $\lfloor \mathrm{TTM}_{it} \rfloor$ is the largest integer less than or equal to the remaining time to maturity at reporting date.

We replace the estimated probability that an instrument defaults in and not prior to year $t + \Delta t$ by the product of the (unconditional) survival probability prior to that year (which is the product of (conditional) survival probabilities) and the (conditional) probability of default in $t + \Delta t$, i.e.,

$$\widehat{\mathrm{LEL}}_{it} = \sum_{\Delta t=0}^{\lfloor \mathrm{TTM}_{it} \rfloor} \Bigg[\Bigg(\prod_{s \in \mathbb{Z} : 0 \leq s \leq \Delta t - 1} \Big(1 - \widehat{\mathrm{P}}(\mathrm{D}_{it+s} = 1 | \mathcal{F}_t) \Big) \Bigg) \cdot \widehat{\mathrm{P}}(\mathrm{D}_{it+\Delta t} = 1 | \mathcal{F}_t) \cdot \frac{\widehat{\mathrm{E}}(\mathrm{LGD}_{it+\Delta t} | \mathcal{F}_t)}{(1+r_i)^{\Delta t}} \cdot \min(1, \mathrm{TTM}_{it} - \Delta t) \Bigg], \tag{5.7}$$

where $\widehat{\mathrm{P}}(\mathrm{D}_{it+s} = 1 | \mathcal{F}_t)$ and $\widehat{\mathrm{P}}(\mathrm{D}_{it+\Delta t} = 1 | \mathcal{F}_t)$ are the estimated PDs from Equation (5.1). We apply the same methodology for LGD computations and aggregate PDs and LGDs for future years following Equation (5.6).

5.4.3 Significant Increase in Credit Risk (SICR)

The classification of financial instruments in IFRS 9 depends on the credit risk at reporting date compared to the initial level. Technically, an instrument shifts from Stage 1 to Stage 2 if the default risk significantly increases (§ 5.5.9 IFRS 9). Instruments with low credit risk are excluded from this rule (§ 5.5.10 and § B5.5.23 IFRS 9). Note that this significance has to be interpreted as "substantial" as it is not applied in a statistical sense. This exception holds for investment grade bonds for those we estimate a PD of less than 19.38 basis points. The standard suggests changes of external as well as internal ratings and economic states as SICR indicators (§ B5.5.13 IFRS 9). We define SICR based on estimated PDs (i.e., ratings, other borrower controls and macroeconomic factors).

IFRS 9 requires consideration of the same time period for the SICR evaluation (§ B5.5.13). We consider an exemplary financial instrument to clarify this requirement. Let the instrument be initially recognized in year $t_0 = 2000$ with maturity ending in year $t_0 + 10 = 2010$. The (conditional) PD for each year is assumed to be 1 %. Thus, the probability of default for the total remaining lifetime is $1 - (1 - 0.01)^{10} = 9.56\,\%$ from initial recognition, i.e., it is one minus the product of (conditional) survival probabilities For the SICR evaluation after four years, i.e., at reporting date in 2004, the probability of default for the remaining lifetime of six years might be computed as 8 % (including new information, e.g., economic conditions). The false comparison would be between the 10-year PD at initial recognition (9.56 %) and the 6-year PD after four years (8 %). Instead, from the view of the initial recognition, the 6-year PD, given no default in the first four years of the initial remaining maturity, was $1 - (1 - 0.01)^6 = 5.85\,\%$. Thus, the relevant remaining lifetime PD deteriorates by 8 - 5.85 = 2.15 percentage points, i.e., 36.71 % and indicates a risk deterioration.

Under certain conditions, IFRS 9 allows use of the 12-month PD for the SICR criterion, but only if default risk changes are comparable over time horizons. We emphasize two main aspects as to why these changes are principally not similar and, thus, bonds should be evaluated using their remaining lifetime PD. First, short-term changes (e.g., caused by macroeconomic shocks) may significantly deteriorate the 12-month PD but the influence vanishes over lifetime. A naive consideration of the 12-month horizon may thus amplifies a possible procyclicality. Second, long-term changes (e.g, caused by bond- or issuer-specific fundamentals) may negligibly increase the 12-month PD but sum up over the long-term to a significant risk deterioration over lifetime. These changes may not be identified by a 12-month SICR criterion.

We call the probability of default for the remaining time to maturity the lifetime probability of default (LPD). IFRS 9 demands computation of the LPD of an instrument i from the point of initial recognition or a reporting year t_1. The crucial time horizon is starting at a reporting year $t_2 \geq t_1$ and ends with maturity. The LPD is given by one minus the (unconditional) survival probability, i.e., the product of the (conditional) survival probabilities and estimated by

$$\widehat{\mathrm{LPD}}_{it_2}(t_1) = 1 - \prod_{\Delta t=0}^{\lfloor \mathrm{TTM}_{it_2} \rfloor} \left[1 - \widehat{\mathrm{P}}(\mathrm{D}_{it_2+\Delta t} = 1|\mathcal{F}_{t_1}) \ \cdot \ \min(1, \mathrm{TTM}_{it_2} - \Delta t) \right], \tag{5.8}$$

where $\widehat{\mathrm{P}}(\mathrm{D}_{it_2+\Delta t} = 1|\mathcal{F}_{t_1})$ is the estimated PD for year $t_2 + \Delta t$ using the information set of year t_1 and Equation (5.1). For this calculation, the VIX forecast is done $t_2 - t_1 + \Delta t$ years ahead. Again, the time to maturity is subsequently decreased and the last year only partly recognized.

At reporting date t the current estimate of the lifetime PD $\mathrm{LPD}_{it}(t)$ of instrument i must be compared to the estimated $\mathrm{LPD}_{it}(t_0)$ from the point of initial recognition t_0. The evaluation of a significant risk increase shall be made in relative terms (§ B5.5.9 IFRS 9). An asset is classified to Stage 2 under IFRS 9 (i.e., the formal SICR criterion is fulfilled) if

$$\frac{\widehat{\mathrm{LPD}}_{it}(t)}{\widehat{\mathrm{LPD}}_{it}(t_0)} - 1 \geq \alpha \tag{5.9}$$

with a threshold $\alpha > 0$. IFRS 9 does not suggest a specific value and leaves room for interpretation. This paper discusses three thresholds: 5 %, 20 % and 50 %, and analyzes the sensitivity to the SICR criterion.

5.5 Impact on Regulatory Capital

5.5.1 Stylized Asset Portfolios

This section discusses several portfolio qualities and reinvestment strategies to allow for a comprehensive impact study. Institutions may manage their asset portfolio risk profile based on internal ratings. We follow four different stylized portfolios of different credit qualities that are given by the rating distributions of Table 5.6 over time. The higher fraction of assets with a better credit rating (e.g., Aaa-Baa) and a lower fraction of assets with a lower credit rating (e.g., Caa-C) implies a better portfolio quality. Each portfolio consists of 2,000 assets (represented by bonds) and is based on representative bank data of the Federal Reserve as presented in Gordy (2000). The impact on the eligible regulatory capital of IFRS 9 and GAAP 326 is analyzed in a counterfactual analysis by studying the IFRS 9 and GAAP 326 rules for US American bonds between 1991 and 2013, which is a period where these rules have not been applied. As assets mature or default we replace these following one of five reinvestment strategies following an adaption of Gordy and Howells (2006).

Table 5.6: Credit quality distributions of stylized portfolios

Rating	Credit quality			
	High	Average	Low	Very low
Aaa	76	58	20	10
Aa	118	100	31	21
A	585	268	74	63
Baa	758	623	331	264
Ba	382	649	761	712
B	55	222	647	740
Caa - C	26	80	136	190
Total	2,000	2,000	2,000	2,000

Notes: The paper uses these stylized portfolios for further analysis. The four cases stand for representative banks based on internal FED data and reported in Gordy (2000).

The consideration of cyclicality leads to one of three basic strategies. The idea of the first type is to keep the average portfolio PIT PD constant and to account for current economic conditions. This ‘cyclical’ reinvestment strategy requires a tightening of lending standards in recessions in order to compensate for the decreasing quality of the existing portfolio. For the derivation of the corresponding ratings, we estimate PDs of all assets by the PIT model of Table 5.3. Then we order these risk measures to assign internal ratings. The classification follows the frequencies of Moody's ratings in the dataset: 4.76 % Aaa, 17.06 % Aa, 33.71 % A, 25.33 % Baa, 6.88 % Ba, 8.97 % B, and 3.29 % Caa - C.

The contrary 'non-cyclical' reinvestment strategy aims to keep constant the average long-run default risk. Here institutions keep the long-term risk constant and do not adapt their lending standards according to economic surroundings. This strategy uses estimated PD of the TTC model of Table 5.3 for the classification of internal ratings.

In practice, institutions choose a mix of both above mentioned reinvestment strategies as they tighten their lending standards during downturns. However, poor market conditions may prevent a full adjustment. This 'semi-cyclical' approach uses internal ratings that are based on the average of PIT and TTC estimates. This intermediate case is used as the base case for the empirical results. Both extreme strategies show the sensitivity and robustness of implications due to portfolio management.

Results are presented for each combination of the four different portfolio qualities and the three above mentioned reinvestment strategies. For each combination we consider 100 independent portfolios that represent 100 different banks to ensure that results do not depend on one specific choice. The procedure for one bank is as follows. The bank portfolio consists of initially 2,000 randomly chosen bonds from all bonds that are in the dataset in year 1991, clustered by ratings according to the portfolio quality in this first year. For the following years all bonds of the first year principally stay in the portfolio. The portfolio needs, however, to be actively managed over time to restore the initial portfolio size and quality. Some bonds drop out due to maturity or default.[7] In addition, bond ratings, and thus the initial portfolio quality, change. To restore the initial portfolio size and rating distribution we subsequently add and replace bonds year by year. First, bonds are randomly removed for rating classes that are over-represented due to rating migration. Second, some rating classes are under-represented over time

[7] Similar to Gordy and Howells (2006) bonds are excluded after default to analyze the impact of loan loss provisioning of non-defaulted bonds.

because of bonds' default or maturity, or bonds' rating changes to other classes. Bonds for the specific year from the dataset to the portfolio are randomly added if rating classes are over-represented. These bonds also principally stay in the portfolio for following years but may be removed for further restoring. The procedure is subsequently performed year by year until 2013. This procedure is repeated for each bank separately, i.e., sampling is carried out independently. The first five years of the data are treated as a burn-in phase to setup representative portfolios. Each bond is equally weighted by the same exposure and are sampled with replacement.

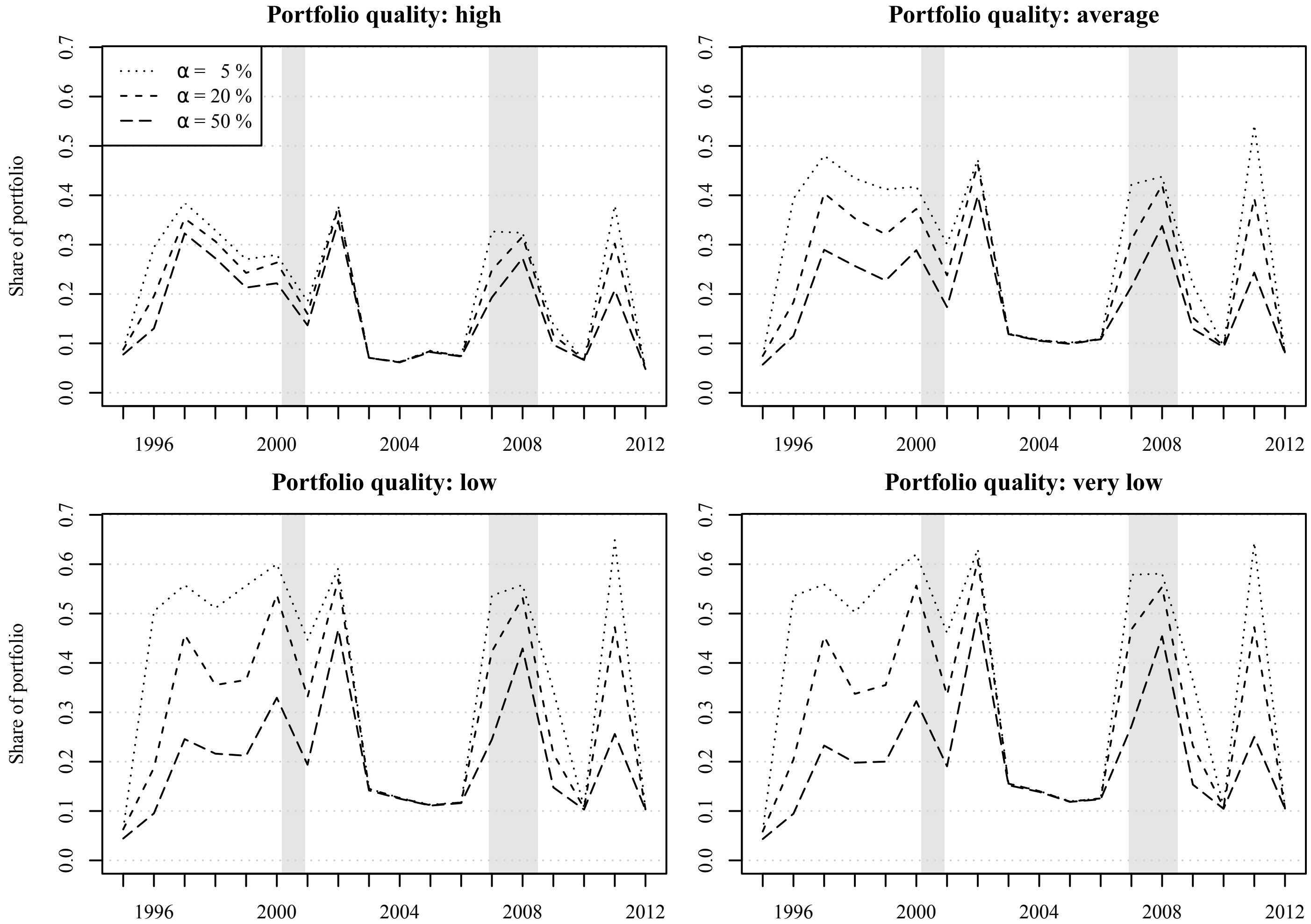
This paper also considers two reinvestment strategies of Gordy and Howells (2006) for further robustness. The 'fixed' strategy does not restore the initial portfolio quality. New bonds are added from the initial distribution following default/maturity but no bond is removed due to over-representation in a rating class. In the 'passive' strategy new bonds are added each year following the current rating distribution in the portfolio for that year. Both approaches are optimized for pure simulation studies over a long time-horizon. The first years of the dataset will cause a shift in the portfolio qualities.

5.5.2 Significant Increase in Credit Risk (SICR)

The instrument classification in IFRS 9 is based on the evaluation of the change in default risk. The current LPD estimate (for the remaining lifetime) must be compared to the initially estimated LPD (for the same time horizon) for each financial instrument at the reporting date.

Figure 5.7 shows the mean for each portfolio quality and SICR threshold per year for the base case of a semi-cyclical reinvestment strategy. The portfolios are initialized in the financial year 1990 with reporting date 31.12.1990, i.e., starting with default risk and expected losses from 1991 on. The first five years of the data (1991 - 1995) are treated as a burn-in phase to setup representative portfolios and excluded for the results.

Figure 5.7: Portfolio share of Stage 2 instruments in IFRS 9



Notes: The figure shows the share of bonds in Stage 2 for different portfolio qualities using the semi-cyclical reinvestment strategy. The dashed and dotted lines represent three different SICR thresholds α in IFRS 9: 5 %, 20 % resp. 50 %. Each year (e.g., 2007) represents the financial year ending on the 31th December of the corresponding year (e.g., 31.12.2007). Each line corresponds to the average share over 100 sampled portfolios that represent 100 independent banks. The shaded areas indicate recession dates of the National Bureau of Economic Research.

Bonds shift over time from Stage 1 to Stage 2 due to significant increases in default risk and shift back if the SICR criterion does not longer apply. In addition, some instruments leave the portfolio due to default or maturity and new instruments are added to restore the portfolio quality and size. The minimum mean share of Stage 2 bonds in expansions is between 5 % and 15 %., e.g., in year 2005 approximately 10 % of all instruments in a portfolio with an average credit risk are in Stage 2. A lower portfolio quality is more likely to cause an exceedance of the SICR threshold and increases the share of Stage 2 instruments. The choice of α does not seem to cause differences for good economic conditions, e.g., in 2003 - 2006. Downturns increase the systematic default risk and, thus, the share of Stage 2 bonds. The maximum strongly depends on the SICR threshold and the portfolio quality. For the average credit risk the 50 % threshold leads to 39 % of bonds in Stage 2, the 20 % threshold leads to 45 % and the 5 % threshold leads to 53 %. A high portfolio quality leads to a low number of bonds in Stage 2 due to a lower risk sensitivity and the exception of low risk assets from the lifetime loss requirement. The corresponding maximum varies between 34 and 39 % depending on the threshold. In contrast, for banks with very low credit quality the maximum share is between 50 and 65 %.

In the following sections, we study the resulting impact of IFRS 9 on provisions as well as regulatory capital and compare those to GAAP 326 requirements.

5.5.3 Computation of Basel and Accounting Expected Losses

After the stage classification of IFRS 9 in the previous section, the corresponding provisions can be calculated: the 12-month expected loss for Stage 1 and the lifetime expected loss for Stage 2. GAAP 326 uses the latter in all instances. The 12-month and lifetime expected losses are computed by using the PIT PD and LGD models (see Section 5.4). This section compares the corresponding provisions and Basel expected losses. Furthermore, Section 5.5.4 analyzes the impact on the eligible regulatory capital of IFRS 9 and GAAP 326 in a counterfactual analysis for the data, for which the rules have not been applied. Again, we present results for different portfolio qualities and SICR thresholds using the semi-cyclical reinvestment strategy.

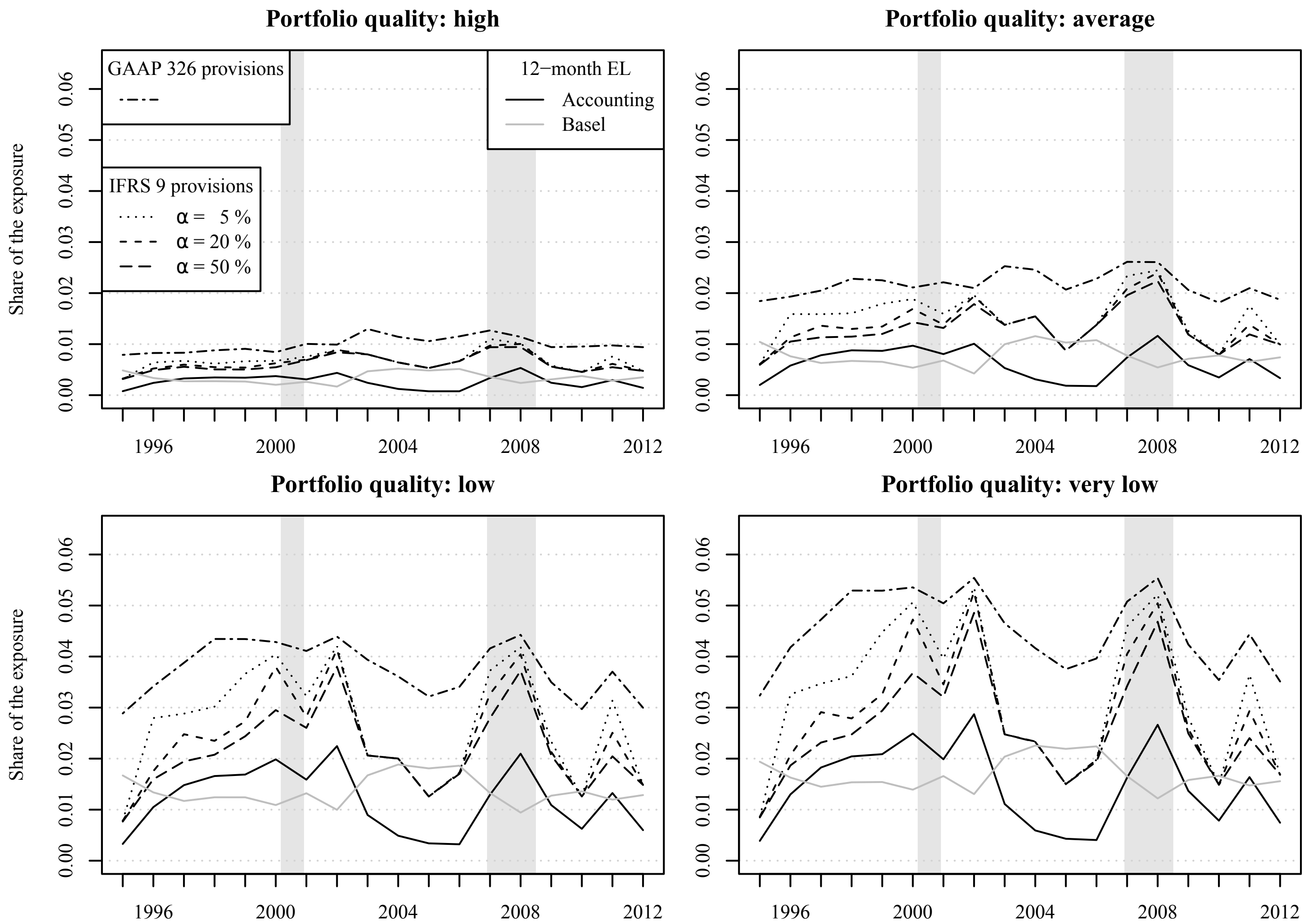
The Basel expected loss is generally the product of the estimated PD and LGD of the TTC models in Table 5.3 and Table 5.4. Current information of the financial year is used and no VIX forecast is included. However, Basel requires several corrections to both risk parameters. First, the LGD must reflect economic downturn conditions (§ 468 Basel II). We account for those by the adjustment $0.08 + 0.92 \cdot \text{LGD}$ as proposed by the Board of Governors of the Federal Reserve System (2006). Second, regulators apply a floor for the parameter estimates for the internal ratings-based approach (see Basel Committee on Banking Supervision (2016a)). The PD estimate must be greater or equal to 5 basis points which affects approximately 15.6 % of all observations. In addition, the LGD parameter minimum of 25 % affects 6.7 % of all observations.[8]

For each bank, i.e., sampled portfolio, we calculate the portfolio sum of the Basel expected loss and the sum of all provisions. Figure 5.8 shows the time-series of means for all banks with the same portfolio quality. All measures are reported as a fraction of the portfolio exposure. We additionally consider the provisions depending on the accounting standard and the SICR threshold α for IFRS 9. The Basel expected loss (gray line) is less volatile due to the underlying TTC approach. The solid black line characterizes the PIT 12-month expected loss and is the lower bound for IFRS 9 provisioning that holds if all instruments are in Stage 1. The upper bound is given by GAAP 326 provisions (dash-dotted line) that are generally calculated by the lifetime expected loss (what equals Stage 2) and, thus, are less volatile. The corresponding provisions are on average approximately 1 % for the high portfolio quality, 2 % for the average case and up to 5 % for very risky portfolios.

The IFRS 9 provisions (the three middle dashed and dotted lines) are by definition lower than GAAP 326 provisions. In expansions, the SICR threshold plays a minor role and overall provisions are closer to the 12-month expected losses. The IFRS 9 require-

[8]The proposed 25 % floor holds for unsecured bonds. As the data does not contain sufficient information on collateral we also use the 25 % floor for secured bonds. This do not affect the contributions because (i) the affected secured bonds have on average estimated LGDs of 19.0 %, and (ii) lower proposed floors lower Basel expected losses and thus even increase the impact of IFRS 9 and GAAP 326 on regulatory capital.

Figure 5.8: Provisions and expected losses



Notes: The figure shows the Basel expected loss and the IFRS 9 as well as GAAP 326 provisions as share of the exposure for different portfolio qualities using the semi-cyclical reinvestment strategy. Each year (e.g., 2007) represents the financial year ending on the 31th December of the corresponding year (e.g., 31.12.2007). Each line corresponds to the average deduction over 100 sampled portfolios that represent 100 independent banks. The shaded areas indicate recession dates of the National Bureau of Economic Research.

ments are closer to GAAP 326 requirements in downturns and for lower SICR thresholds.

Although GAAP 326 requires more provisions in general, it is less procyclical than IFRS 9, i.e., the additional burden from upturn to downturn periods is lower in GAAP 326.

5.5.4 Impact on Common Equity Tier (CET) 1

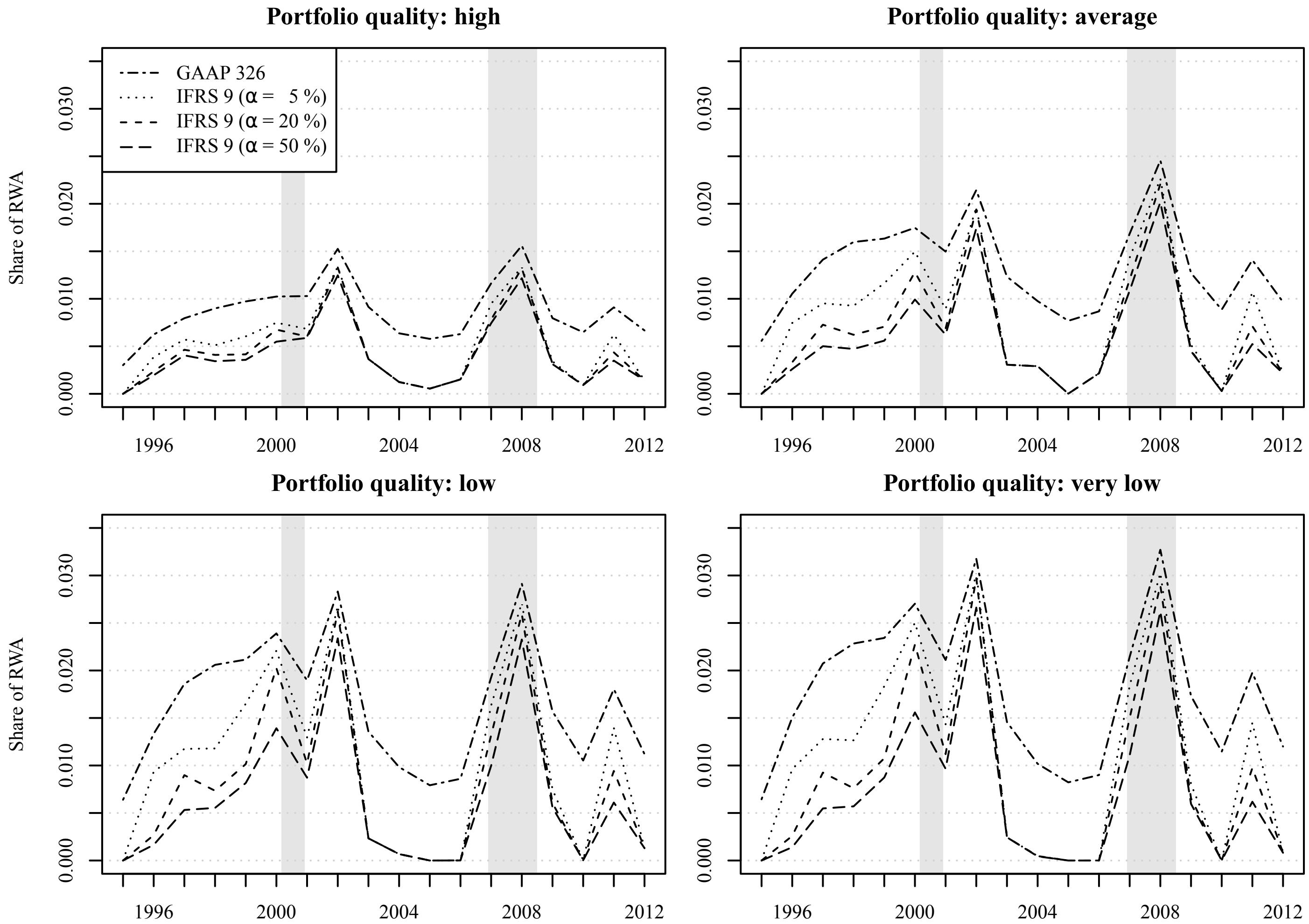
The previous section shows what the provisions would have been, had the accounting standards been mandatory in the past. Here we discuss the corresponding impact on CET 1 that is directly lowered by the deduction implied by provisioning.[9] We present the deduction in regulatory capital in percentages of the exposure and the RWA. These are calculated according to the Basel II formula (§ 272) and take into account the parameter adjustments as previously mentioned for the Basel expected loss. Again, we present results for the four portfolio qualities and the three SICR thresholds using the semi-cyclical reinvestment strategy (Figure 5.9). We aggregate the mean capital for four time horizons: (i) through the economic cycle, (ii) for recessions as given by the National Bureau of Economic Research, (iii) expansions (times of no recession), and (iv) the Global Financial crises (GFC). Table 5.7 shows the mean capital deduction distinguished for portfolio qualities, accounting standards (including SICR threshold) and time horizon.

For the average portfolio quality, the average deduction of the CET 1 ratio due to GAAP 326 is 134 bps (= 1.34 % of RWA). This is the average additional amount of CET 1 institutions need to hold due to differences between Basel expected losses and GAAP 326 provisions. Due to the risk sensitivity of the lifetime expected loss, this gap behaves procyclically. The additional requirement lowers in expansions to 126 bps but increases on up to 198 bps in recession and would have been 246 bps in the GFC. This is more than a half of the required minimum CET 1 ratio of 4.5 %.

The portfolio quality influences the capital deduction because higher risk increases lifetime expected losses. For low overall credit risk the gap decreases to 82 bps in expansions and 156 bps in the GFC. Very risky portfolios result in capital needs of 170 bps and 327 bps receptively.

[9] This paper focuses on the impact on the higher-quality CET 1 and does not further consider the Tier 2 component as Tier 2 capital does not provide a binding constraint for most banks.

Figure 5.9: Deduction of the Common Equity Tier 1 due to provisioning



Notes: The figure shows the difference of provisions and the Basel expected losses given as share of the risk-weighted assets for different portfolio qualities using the semi-cyclical reinvestment strategy. A positive difference (excess) leads to a deduction of the CET 1 and addition to Tier 2 capital. A deficit must be deducted of the CET 1 to calculate the eligible regulatory capital. Each year (e.g., 2007) represents the financial year ending on the 31th December of the corresponding year (e.g., 31.12.2007). Each line corresponds to the average deduction over 100 sampled portfolios that represent 100 independent banks. The shaded areas indicate recession dates of the National Bureau of Economic Research.

Table 5.7: Deduction of Common Equity Tier 1 due to provisioning

		GAAP 326 Accounting provisions – expected loss under Basel				IFRS 9 Accounting provisions – expected loss under Basel											
						SICR threshold $\alpha = 5\,\%$				SICR threshold $\alpha = 20\,\%$				SICR threshold $\alpha = 50\,\%$			
	Portfolio quality	Expansion	Average	Recession	GFC	Expansion	Average	Recession	GFC	Expansion	Average	Recession	GFC	Expansion	Average	Recession	GFC
In % of RWA	High	0.82	0.87	1.29	1.56	0.44	0.50	1.01	1.33	0.37	0.44	0.96	1.32	0.34	0.40	0.90	1.22
	Average	1.26	1.34	1.98	2.46	0.71	0.80	1.58	2.26	0.56	0.66	1.45	2.21	0.48	0.57	1.32	2.02
	Low	1.54	1.64	2.40	2.91	0.88	1.00	2.00	2.71	0.68	0.80	1.81	2.59	0.52	0.64	1.60	2.33
	Very low	1.70	1.81	2.69	3.27	0.96	1.10	2.23	3.03	0.74	0.88	2.01	2.90	0.57	0.70	1.79	2.62
In % of exposure	High	0.63	0.65	0.82	0.90	0.31	0.35	0.63	0.77	0.27	0.30	0.60	0.77	0.24	0.28	0.56	0.70
	Average	1.36	1.41	1.80	2.06	0.71	0.78	1.41	1.90	0.56	0.64	1.28	1.86	0.48	0.55	1.17	1.69
	Low	2.29	2.38	3.14	3.48	1.23	1.38	2.57	3.24	0.94	1.09	2.30	3.10	0.72	0.87	2.03	2.78
	Very low	2.72	2.85	3.85	4.32	1.46	1.65	3.15	4.00	1.11	1.29	2.81	3.82	0.85	1.03	2.51	3.46

Notes: This table shows the difference of provisions and the Basel expected losses given as share of the risk-weighted and non-weighted assets for different portfolio qualities using the semi-cyclical reinvestment strategy. A positive difference (excess) leads to a deduction of the CET 1 and addition to Tier 2 capital. A deficit must be deducted of the CET 1 to calculate the eligible regulatory capital. Each number corresponds to the average deduction over 100 sampled portfolios that represent 100 independent banks. Recession dates are those of the National Bureau of Economic Research.

IFRS 9 generally results in lower provisions due to the recognition of the 12-month expected loss for Stage 1 instruments. The lower the SICR threshold α, the more sensitive the transition from Stage 1 (12-month expected loss) to Stage 2 (lifetime expected loss) and the higher provisions and the capital deduction are. The 20 % threshold serves as a median case, where the average gap for the average portfolio quality is 66 bps of the RWA and, thus, 50.75 % less than the corresponding amount of an GAAP 326 institution.[10] The difference between both accounting standards is greater in expansions and lower in recessions. The IFRS 9 gap is with 145 bps only 26.77 % lower in recessions (than in GAAP 326) and with 221 bps in the GFC 10.16 % less than GAAP 326 requirements.

The results indicate that GAAP 326 requires more high-quality regulatory capital and burdens institutions through the economic cycle. IFRS 9 results in lower provisions and reacts with a lag to recessions and may challenge institutions substantially more in downturns due to procyclicality.[11]

The choice of the SICR threshold affects the share of Stage 2 instruments in IFRS 9 (see Section 5.5.2). The effect remains for the capital deduction but to a minor extent. This is caused by instruments that are already classified in Stage 2 for high SICR thresholds and have large lifetime expected losses. The average capital deduction for the average portfolio quality is 57 bps - 80 bps for SICR thresholds between 5 % and 50 %. The gap in the GFC increases up to 202 bps - 226 bps. The lower the threshold, the higher IFRS 9 provisions are. This may stimulate institutions to non-transparent SICR management to lower provisions.

The portfolio reinvestment strategy affects provisioning and capital deduction. As previous results correspond to the representative semi-cyclical approach, we will briefly summarize the results for other strategies (see Table 5.8). A cyclical approach, i.e., a tightening of lending standards during downturns, results in lower gaps in general. In contrast, the non-cyclical management with constant long-term credit risk leads to higher

[10] The survey of the European Banking Authority (2016a) under European banks shows an expected capital deduction of 59 bps due to IFRS 9 and supports the findings of this empirical and more comprehensive study. It ensures robust and representative conclusions for further results

[11] In the transition from expansion to recession the additional capital deduction due to GAAP 326 was 198 - 126 = 72 bps whereas it was 145 - 56 = 89 bps. for IFRS 9 ($\alpha = 20\,\%$).

Table 5.8: Deduction of CET 1 for different reinvestment strategies

In % of RWA		GAAP 326 Accounting provisions – expected loss under Basel				IFRS 9 Accounting provisions – expected loss under Basel			
						SICR threshold $\alpha = 20\,\%$			
Reinvestement strategy	Portfolio quality	Expansion	Average	Recession	GFC	Expansion	Average	Recession	GFC
Cyclical	High	0.86	0.92	1.40	1.79	0.32	0.40	1.06	1.55
	Average	1.18	1.25	1.76	2.22	0.39	0.49	1.24	1.86
	Low	1.54	1.59	2.05	2.39	0.52	0.62	1.42	1.94
	Very low	1.70	1.76	2.31	2.68	0.59	0.70	1.54	2.09
Semi-cyclical	High	0.82	0.87	1.29	1.56	0.37	0.44	0.96	1.32
	Average	1.26	1.34	1.98	2.46	0.56	0.66	1.45	2.21
	Low	1.54	1.64	2.40	2.91	0.68	0.80	1.81	2.59
	Very low	1.70	1.81	2.69	3.27	0.74	0.88	2.01	2.90
Non-cyclical	High	0.91	0.98	1.56	1.93	0.53	0.60	1.22	1.69
	Average	1.38	1.48	2.32	3.01	0.75	0.86	1.78	2.64
	Low	1.65	1.78	2.82	3.66	0.81	0.95	2.07	3.07
	Very low	1.79	1.93	3.06	3.93	0.87	1.03	2.27	3.31
Fixed	High	1.04	1.17	2.21	3.24	0.65	0.78	1.85	3.01
	Average	1.33	1.48	2.65	3.71	0.71	0.86	2.08	3.32
	Low	1.64	1.80	3.10	4.15	0.79	0.96	2.33	3.58
	Very low	1.75	1.92	3.27	4.32	0.79	0.96	2.38	3.63
Passive	High	0.97	1.10	2.17	3.22	0.53	0.65	1.67	2.67
	Average	1.07	1.21	2.31	3.37	0.58	0.71	1.77	2.79
	Low	1.18	1.32	2.45	3.51	0.62	0.76	1.85	2.89
	Very low	1.19	1.33	2.45	3.49	0.63	0.76	1.85	2.88

Notes: This table shows the difference of provisions and the Basel expected losses given as share of the risk-weighted assets for different portfolio qualities and reinvestment strategies. A positive difference (excess) leads to a deduction of the CET 1 and addition to Tier 2 capital. A deficit must be deducted of the CET 1 to calculate the eligible regulatory capital. Each number corresponds to the average deduction over 100 sampled portfolios that represent 100 independent banks. Recession dates are those of the National Bureau of Economic Research.

capital gaps. For both strategies in combination with the average portfolio quality the average gap over time is 125 bps - 148 bps (GAAP 326) and 49 bps - 86 bps (IFRS 9) instead of 134 bps and 66 bps for the semi-cyclical strategy. In the GFC, this gap is 222 bps - 301 bps (GAAP 326) and 186 bps - 264 bps (IFRS 9) instead of the median case with 246 bps resp. 221 bps. In summary, institutions that tighten lending standards during downturns have to hold less capital with less procyclicality. Institutions that do not or cannot tighten lending standards during downturns have to keep more capital, which is also more sensitive to the economic cycle.[12]

We report results for the fixed and passive reinvestment strategy in Table 5.8. The size of the gap partly differs but the main conclusions are similar as before (with respect to CET 1 deduction, portfolio quality and procyclicality). Appendix 5.A shows results for other macroeconomic variables. They are not able to capture cyclicality as well as the VIX.

The results show that IFRS 9 and GAAP 326 loan loss provisioning that is based on expected losses causes procyclicality. We will briefly summarize the main conclusions with respect to level and procyclicality here. GAAP 326 reduces eligible regulatory CET 1 more than IFRS 9. However, the latter causes more procyclicality. The SICR criterion is subject to a trade-off effect as a conservative approach (low α) causes higher capital needs. This is particularly interesting for institutions' earnings management and transparency. Institutions with stressed portfolios are more affected by downturns, due to the higher sensitivity of eligible capital. Despite the implications for institutions, reinvestment and management decisions, the conclusions are important for regulators and supervisors. A transition phase as proposed by the European Commission (2016) may help to raise the general level of high-quality capital. However, further discussions need also to focus on procyclicality aspects and a possible adjustment of regulatory requirements with respect to the handling of provisions. Furthermore, the introduction of parameters floors in IFRS 9 and GAAP 326 consistent with Basel may reduce procyclicality and variation between institutions. Finally, the determination of the counter-cyclical capital buffer needs to account for the procyclicality of provisioning. The additional buffer for a systematically important bank might account for the accounting standard and portfolio quality of the institution.

[12] For the cyclical reinvestment strategy the additional burden due to downturns is 176 - 118 = 58 bps (GAAP 326) resp. 124 - 39 = 85 bps (IFRS 9, $\alpha = 20\,\%$). Is contrast, the non-cyclical numbers are 232 - 138 = 94 bps resp. 178 - 75 = 103 bps.

5.6 Discussion

The accounting standards IFRS 9 and GAAP 326 replace the existing incurred loss model. The new approach is intended to increase transparency and reduce procyclicality. This paper discusses both standards and shows that the objectives are not fully met. A counterfactual analysis on US American bonds between 1991 and 2013 shows the impact of future loan loss provisioning and explores the cyclicality of eligible regulatory capital and net income.

For representative portfolios, we find GAAP 326 leads on average to a future deduction of CET 1 of 1.34 % in terms of risk-weighted assets (RWA) which needs to be seen in relation to the minimum required capital ratio of 4.5 %. This gap behaves procyclically and would have been 1.98 % during past recessions and 2.46 % in the Global Financial Crisis. Due to the SICR criterion IFRS 9 leads to lower loan loss provisioning and capital deduction. For a median threshold, i.e., a significant increase in credit risk is given by a 20 % increase in default risk, the average CET 1 gap is 0.66 %. However, due to the high number of threshold excesses during downturns, IFRS 9 is more procyclical than GAAP 326 and would have led to a capital deduction of 1.45 % during past recessions and 2.21 % in the Global Financial Crisis. As banks are constrained in Tier 1 capital and required to hold even more Tier 1 capital in upcoming years IFRS 9 and GAAP 326 are highly likely to require banks to raise additional Tier 1 capital.

Finally, we discuss several aspects how future loan loss provisioning may be managed. The following factors are identified to reduce provisions in general and the procyclical impact on net income and regulatory capital deduction: (i) a portfolio with low credit risk, and (ii) a constant risk profile by tightening lending standards during economic downturns. A higher SICR threshold increases provisioning and must be viewed critically in combination with the objective to increase transparency as institutions may have incentives to lower provisions. Regulators might also take into account the following aspects for the debate on how to treat future loan loss provisioning. The variability of risk parameters due to varying statistical approaches may cause a high variation of loan loss provisioning for institutions with similar credit portfolios. In addition, regulators may dampen the additional burden during downturns by lowering counter-cyclical capital buffers in economic downturns or changing the treatment of provisioning when banks are close to failure (e.g., revert to 12-month provisioning during economic downturns).

Appendix 5.A The Performance of Macroeconomic Variables

IFRS 9 and GAAP 326 require an accounting of macroeconomic information for the computation of expected losses. This appendix discusses several macroeconomic variables in order to capture the cyclicality of credit risk (see Table 5.A.1).

Table 5.A.1: Macroeconomic variables

	Quantiles					Mean
	10 %	25 %	50 %	75 %	90 %	
GDP growth (in %)	-0.24	1.68	2.39	4.13	4.45	2.46
Historic default rate (in %)	0.19	0.28	0.56	0.95	2.36	0.80
TED spread (in %-points)	0.18	0.22	0.49	0.67	1.32	0.56
Treasury rate 1 year (in %)	0.18	1.24	3.49	5.05	5.63	3.06
Treasury rate 10 years (in %)	2.78	3.66	4.61	5.26	6.35	4.58
Treasury term spread (in %-points)	-0.08	0.57	1.63	2.60	2.79	1.46
Unemployment (in %)	4.00	4.40	5.20	7.10	9.10	5.83
Volatility index VIX (in %-points)	12.07	13.29	21.68	24.42	28.62	20.40

Notes: Macroeconomic variables are lagged one year and winsorized to the 5 % and 95 % level.

For all variables we evaluate the goodness of fit for the PD and LGD model (see Table 5.A.2). Almost all variables indicate higher credit risk for poor economic surroundings. The VIX provides the best goodness of fit and shows that today's uncertainty adequately forecasts future macroeconomic conditions as discussed by Bloom (2009) and Jo and Sekkel (2017). Thus, our preferred PIT model includes the VIX for presenting results in the main part (cf. Section 5.4.1 and Section 5.5).

For a robustness analysis, we consider the variables with the next smallest goodness of fit: TED spread, 1-year treasury rate and treasury term spread. Table 5.A.3 shows the impact of GAAP 326 and IFRS 9 provisions on regulatory capital using those macroeconomic variables for the PIT and AR model (cf. Section 5.5.4). The variables are not able to capture cyclicality as well as the VIX and they over- and undervalue cyclicality. These additional empirical results are biased and do not affect the main conclusions.

Table 5.A.2: Regression results for additional macroeconomic variables

	Issuer- and bond-specific variables	Probability of Default			Loss Rate Given Default	
		Coefficient	Accuracy Ratio	McFadden's adjusted R^2	Coefficient	Adjusted R^2
GDP growth	✓	-0.0535 (0.0704)	0.8370	0.2921	-0.0200 (0.0334)	0.2180
Historic default rate	✓	0.1703 (0.1273)	0.8424	0.2963	0.0816 (0.0659)	0.2199
TED spread	✓	0.5447 *** (0.2090)	0.8460	0.3091	0.3542 *** (0.1341)	0.2374
Treasury rate 1 year	✓	0.1045 ** (0.0445)	0.8444	0.3060	0.0650 *** (0.0242)	0.2305
Treasury rate 10 years	✓	0.1520 ** (0.0721)	0.8375	0.3016	0.0689 (0.0474)	0.2216
Treasury term spread	✓	-0.1613 * (0.0842)	0.8447	0.3010	-0.1523 *** (0.0434)	0.2380
Unemployment	✓	-0.1058 * (0.0615)	0.8422	0.3003	-0.0534 (0.0351)	0.2223
Volatility index VIX	✓	0.0625 *** (0.0143)	0.8729	0.3279	0.0417 *** (0.0083)	0.2536

Notes: In the PIT models of the PD (Table 5.3) and the LGD (Table 5.4) we replace the VIX by the given macroeconomic variables. Each row represents one PD and one LGD model. The issuer- and bond-specific variables are included in each model but are not presented due to clarity. The table shows the parameter estimate of the corresponding macroeconomic variable and the goodness of fit. Standard errors are given in parentheses and clustered for issuer- and year-specific fixed effects as proposed in Petersen (2009). The significance is indicated for the 1 % (***), 5 % (**) and 10 % (*) level.

Table 5.A.3: Deduction of CET 1 using alternative macroeconomic variables

In % of RWA		GAAP 326 Accounting provisions – expected loss under Basel				IFRS 9 Accounting provisions – expected loss under Basel, SICR threshold $\alpha = 20\,\%$			
Macroeconomic variable	Portfolio quality	Expansion	Average	Recession	GFC	Expansion	Average	Recession	GFC
Volatility index VIX	High	0.82	0.87	1.29	1.56	0.37	0.44	0.96	1.32
	Average	1.26	1.34	1.98	2.46	0.56	0.66	1.45	2.21
	Low	1.54	1.64	2.40	2.91	0.68	0.80	1.81	2.59
	Very low	1.70	1.81	2.69	3.27	0.74	0.88	2.01	2.90
TED spread	High	1.24	1.27	1.44	1.72	0.54	0.58	0.89	1.32
	Average	1.82	1.86	2.17	2.79	0.75	0.81	1.36	2.30
	Low	2.13	2.17	2.50	3.25	0.69	0.77	1.41	2.40
	Very low	2.30	2.34	2.71	3.53	0.72	0.81	1.55	2.63
Treasury rate 1 year	High	1.40	1.39	1.28	1.24	0.76	0.75	0.62	0.65
	Average	2.06	2.04	1.91	1.92	1.07	1.05	0.90	1.09
	Low	2.42	2.39	2.15	2.05	1.14	1.12	0.91	0.88
	Very low	2.63	2.60	2.33	2.18	1.20	1.18	1.02	0.98
Treasury term spread	High	1.40	1.40	1.39	1.41	0.75	0.75	0.75	0.79
	Average	2.02	2.01	1.96	2.15	1.00	1.00	0.97	1.26
	Low	2.28	2.27	2.14	2.29	0.93	0.92	0.88	1.01
	Very low	2.47	2.45	2.30	2.43	0.98	0.98	0.99	1.11

Notes: This table shows the difference of provisions and the Basel expected losses given as share of the risk-weighted assets for different portfolio qualities and macroeconomic variables using the semi-cyclical reinvestment strategy. A positive difference (excess) leads to a deduction of the CET 1 and addition to Tier 2 capital. A deficit must be deducted of the CET 1 to calculate the eligible regulatory capital. Each number corresponds to the average deduction over 100 sampled portfolios that represent 100 independent banks. Recession dates are those of the National Bureau of Economic Research.

Chapter 6

Conclusion

Summary

This thesis contributes to the literature on credit risk modeling and reveals new insights on systematic credit losses during recessions. The models provide more precise estimates of credit risk and can help to avoid a severe underestimation and capital shortfalls in economic downturn periods. The four studies deal with downturn risk parameters, the reduction of non-performing loans, multi-period sample selection of losses, and the procyclicality of bank capital requirements. The analyses cover several dimensions of credit risk. The first study focuses on modeling aspects of bimodal distributed single-loan loss rates. It examines determinants of the high number of total losses and recoveries. The empirical results provide a sophisticated view on the impact of economic downturns on LGDs. The second study goes beyond the realized loss and explores workout processes. The role of systematic effects on workout periods is analyzed and the corresponding impact on loss severity and the reduction of non-performing loans is discussed. The third study extends the positive dependency between default risk and loss severity from a single-period perspective to a multi-period approach. The corresponding significance of an adequate modeling of expected losses over the remaining lifetime of financial instruments is examined in the context of the revised loan loss provisioning. The fourth study further contributes to lifetime expected loss modeling. It discusses the convergence of the new accounting standards to the regulatory approach. Furthermore, the impact of the revised loan loss provisioning on bank capital requirements is examined.

The chapter “Downturn LGD Modeling using Quantile Regression” examines the bimodal distributional shape of loss severities and the impact of the economic cycle on the loss distribution. The importance of an adequate modeling of workout LGDs is demonstrated for US American loans from small and medium enterprises with defaults between 2000 and 2013. The study reveals covariate effects that vary between quantiles by introducing the quantile regression for LGD modeling. Covariates can explain the bimodal nature of losses only to a minor extent. A comprehensive validation study reveals advantages of the approach in comparison to standard regression techniques with respect to the entire distributional fit. The empirical results show that the effect of economic downturns on the loss distribution varies between quantiles and firm- and loan-specific covariates. The proposed downturn LGDs can be used for the internal ratings-based approach of the regulatory framework which are more precise than standard measures.

The first study examines systematic co-movements of the LGD. In contrast, the chapter "Macroeconomic Effects and Frailties in the Resolution of Non-Performing Loans" analyzes workout processes that directly determine losses by cash flows and direct as well as indirect costs. A Cox proportional hazards model is used to model the length of workout processes of defaulted corporate loan data from the United States, Great Britain and Canada between 2004 and 2013. As a result, the explanatory power of macroeconomic information is identified as limited which partly can be explained by the dynamics of economic conditions during length-varying resolution processes. Unobservable systematic effects are modeled as frailties and statistically significantly cause a co-movement of workout periods. A simulation study demonstrates that these frailties cause higher single-loan LGDs and portfolio losses in recessions. Furthermore, systematically delayed workouts charge higher stable fundings needs.

The revised loan loss provisioning requires the estimation of lifetime expected losses and raises questions about the impact on bank capital requirements. The chapter "A Copula Sample Selection Model for Predicting Multi-Year LGDs and Lifetime Expected Losses" deals with the first aspect. It examines the long-term impact of the positive dependency between default risk and loss severity. The copula-based approach allows to combine an accelerated failure time model of the default time and a beta regression of the LGD. The data consist of US American corporate bonds between 1982 and 2014.

Besides the identification of covariate effects for both marginal models, the study reveals a decreasing term structure of loss severity in case of default. The longer a bond survives the lower the expected LGD is. If the selection is not taken into account it results in a severe underestimation of lifetime expected losses.

In the chapter "The Impact of Loan Loss Provisioning on Bank Capital Requirements", the impact of the new accounting standards on bank capital is examined. The introduced expected loss approach is compared to the current regulatory framework. The study estimates 12-month and lifetime expected losses by combining a Probit model for the PD and a fractional response model for the LGD of US American corporate bonds between 1991 and 2013. A counterfactual simulation study of representative portfolios shows a procyclical impact on bank capital requirements due to the new accounting standards. Financial institutions are proposed to reduce the additional burden by holding portfolios with low credit risk, tightening lending standards during recessions, and using the US Generally Accepted Accounting Principles. Regulators and supervisors may dampen the procyclicality by adjusting the countercyclical capital buffer and systemic risk buffers as well as a revision of provision handling in the regulatory framework. Furthermore, the definition of a significant increase in credit risk provides incentives for non-transparent earnings and provisions management.

The results of this thesis have the following implications for financial institutions, financial regulatory authorities and researchers. First, they help financial institutions to understand and account for downturn effects on losses and workout processes of defaulted debt. This is significant for the measurement and management of credit and liquidity risk. Furthermore, the studies demonstrate important aspects for the determination of the revised loan loss provisioning. The results provide options how institutions can manage their capital needs under the new standards. Second, financial regulatory authorities may improve the supervision of financial institutions' risk models with respect to downturn parameters and lifetime expect losses. This helps to ensure an early detection of financial dangers due to the management of credit and liquidity risk and capital reserves. In addition, authorities may learn for future revisions on the regulatory framework or the determination of required capital and liquidity buffers. Besides the practical implications above, this thesis contributes to the academic literature. On the one hand, it helps to improve the understanding of the bimodality of LGDs and the role of workout processes for losses and the amount of non-performing loans. On the other hand, it discusses several issues on the new measure of lifetime expect loss with respect to estimation and implications on bank capital requirements. The results also reveal several aspects for future research that will be discussed in the next section.

Discussion and Outlook

Besides the contributions to the literature, this thesis provides incentives for further research. The next paragraph starts with general aspects that are valid for all four studies. The remaining part of the section separately discusses each chapter and presents possible further developments.

First, the significance of the results may be strengthened for other data, e.g., for other types of financial instruments, borrowers, and countries. In addition, the studies should be repeated for recent data in the future. Second, the stability of the statistical methods and estimates may be subject of further research. As the model performance varies with the sample size, in particular small institutions with limited data need to evaluate the practicability of statistical approaches. However, it should be noted that small institutions can extend their data by joining data pools in order to increase model stability. The more recent discussion of model uncertainty raises questions about the variation of estimates and should be further examined for the evaluation of risk measures. Third, future revisions of regulatory requirements may affect the significance of the results and the applicability of the proposed models. However, regulatory modifications can also account for implications of this thesis.

The consultation paper of the European Banking Authority (2017) demonstrates that the discussion on downturn LGD is still ongoing. Future research may evaluate the advantages of the quantile regression for future discussion results and possible revised regulatory requirements. There is also the necessity for the development of goodness-of-fit measures for the validation of downturn parameter predictions. Furthermore, the quantile regression may be extended to improve the coverage of recession-specific effects by time-varying random effects to account for unobservable systematic effects. A copula-approach similar to Chapter 4 could include the positive dependency between PD and LGD.

The impact of the length of workout processes on the LGD and the role of economic conditions over multiple years in default may be subject to further research. The length of workout processes and the LGD could jointly be modeled by a copula-based approach similar to Chapter 4. Furthermore, the feasibility of estimating macroeconomic and unobservable systematic effects must be evaluated with respect to prediction and management purposes. Competing risk models may reveal insights into the workout process itself by focusing on the intensity of single incoming cash flows and costs.

The results on the lifetime expected loss may stimulate further research on multi-period sample selection modeling. An extension would be to simultaneously study three time dimensions, i.e., the default time, the remaining time to maturity, and the total maturity. The term structure also needs to be analyzed for different categories such as industries and rating classes in order to examine possible differences. This could be done by separated estimates or the regression of the dependence parameter. Practitioners may also be interested in simplified models that account for loss term structures but reduce model complexity and uncertainty.

In order to asses the revised loan loss provisioning more comprehensively, the future practical implementation of the new standards must be evaluated. Although the counterfactual simulation study does not account for the reaction of financial institutions to the new standards, it identifies serious issues why institutions have to account for the modifications. It presents options how to react with respect to portfolio and capital management. Future research may analyze the realized provisioning and the reaction of financial institutions under the new standards. From the regulatory point of view, there may be need to develop dynamic regulatory countercyclical buffers that account for the procyclicality of the revised loan loss provisioning.

In summary, the measurement of credit risk places high demands on statistical methods. The precise modeling of dependencies in credit risk is crucial to particularly ensure adequate capital reserves in recessions. This thesis provides significant implications for the implementation of credit risk models and the development of regulatory requirements. Nevertheless, there are at least three sources that will create new challenges in the future. First, the regulation of the financial sector has to be constantly developed in order to close gaps that are exploited or provide disincentives. Second, new financial instruments may require new modeling approaches. Third, future crises may reveal yet unknown aspects of credit risk.

References

Acharya, V. V., Bharath, S. T., Srinivasan, A. (2007). Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries. Journal of Financial Economics 85 (3), 787–821.

Akhtar, M. A. (1994). Causes and consequences of the 1989-98 credit slowdown: Overview and perspective. Studies on Causes and Consequences of the 1989-98 Credit Slowdown, Federal Reserve Bank of New York, 1–37.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance 23 (4), 589–609.

Altman, E. I., Brady, B., Resti, A., Sironi, A. (2005). The link between default and recovery rates: Theory, empirical evidence, and implications. Journal of Business 78 (6), 2203–2228.

Altman, E. I., Kalotay, E. A. (2014). Ultimate recovery mixtures. Journal of Banking & Finance 40, 116–129.

Bade, B., Rösch, D., Scheule, H. (2011). Default and recovery risk dependencies in a simple credit risk model. European Financial Management 17 (1), 120–144.

Baesens, B., Rösch, D., Scheule, H. (2016). Credit risk analytics: Measurement techniques, applications, and examples in SAS. Wiley, Hoboken, NJ.

Bandopadhyaya, A. (1994). An estimation of the hazard rate of firms under Chapter 11 protection. Review of Economics and Statistics 76 (2), 346–350.

Basel Committee on Banking Supervision (1988). International convergence of capital measurement and capital standards. Bank for International Settlements.

Basel Committee on Banking Supervision (1999). Capital requirements and bank behaviour: The impact of the Basle Accord. Working Paper. Bank for International Settlements.

Basel Committee on Banking Supervision (2005). Guidance on paragraph 468 of the framework document. Bank for International Settlements.

Basel Committee on Banking Supervision (2006). International convergence of capital measurement and capital standards. A Revised Framework. Bank for International Settlements.

Basel Committee on Banking Supervision (2011). Basel III: A global regulatory framework for more resilient banks and banking systems. Bank for International Settlements.

Basel Committee on Banking Supervision (2015a). Guidance on accounting for expected credit losses. Consultative Document. Bank for International Settlements.

Basel Committee on Banking Supervision (2015b). Revisions to the standardised approach for credit risk. Second Consultative Document. Bank for International Settlements.

Basel Committee on Banking Supervision (2016a). Reducing variation in credit risk-weighted assets – constraints on the use of internal model approaches. Consultative Document. Bank for International Settlements.

Basel Committee on Banking Supervision (2016b). Regulatory treatment of accounting provisions. Discussion Paper. Bank for International Settlements.

Basel Committee on Banking Supervision (2016c). Regulatory treatment of accounting provisions – interim approach and transitional arrangements. Consultative Document. Bank for International Settlements.

Basel Committee on Banking Supervision (2017). Regulatory treatment of accounting provisions – interim approach and transitional arrangements. Standards. Bank for International Settlements.

Bastos, J. A. (2010). Forecasting bank loans loss-given-default. Journal of Banking & Finance 34 (10), 2510–2517.

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

Bellotti, T., Crook, J. (2012). Loss given default models incorporating macroeconomic variables for credit cards. International Journal of Forecasting 28 (1), 171–182.

Bender, R., Augustin, T., Blettner, M. (2005). Generating survival times to simulate Cox proportional hazards models. Statistics in Medicine 24 (11), 1713–1723.

Betz, J., Kellner, R., Rösch, D. (2016). What drives the time to resolution of defaulted bank loans? Finance Research Letters 18, 7–31.

Bikker, J. A., Metzemakers, P. A. J. (2005). Bank provisioning behaviour and procyclicality. Journal of International Financial Markets, Institutions & Money 15 (2), 141–157.

Black, F., Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy 81 (3), 637–654.

Blazy, R., Martel, J., Nigam, N. (2014). The choice between informal and formal restructuring: The case of French banks facing distressed SMEs. Journal of Banking & Finance 44, 248–263.

Bloom, N. (2009). The impact of uncertainty shocks. Econometrica 77 (3), 623–685.

Board of Governors of the Federal Reserve System (2006). Risk-based capital standards: Advanced capital adequacy framework and market risk. Proposed Rules. Federal Register 71 (185), 55829–55958.

Board of Governors of the Federal Reserve System (2016). Net stable funding ratio: Liquidity risk measurement standards and disclosure requirements. Proposed Rules. Federal Register 81 (105), 35123–35183.

Bordo, M. D., Redish, A., Rockoff, H. (2015). Why didn't Canada have a banking crisis in 2008 (or in 1930, or 1907, or ...)?. Economic History Review 68 (1), 218–243.

Bris, A., Welch, I., Zhu, N. (2006). The costs of bankruptcy: Chapter 7 liquidation versus Chapter 11 reorganization. Journal of Finance 61 (3), 1253–1303.

Bruche, M., González-Aguado, C. (2010). Recovery rates, default probabilities, and the credit cycle. Journal of Banking & Finance 34 (4), 754–764.

Calabrese, R. (2014). Downturn loss given default: Mixture distribution estimation. European Journal of Operational Research 237 (1), 271–277.

Calabrese, R., Zenga, M. (2010). Bank loan recovery rates: Measuring and nonparametric density estimation. Journal of Banking & Finance 34 (5), 903–911.

Campbell, J. Y., Hilscher, J., Szilagyi, J. (2008). In search of distress risk. Journal of Finance 63 (6), 2899–2939.

Campbell, J. Y., Hilscher, J., Szilagyi, J. (2011). Predicting financial distress and the performance of distressed stocks. Journal of Investment Management 9 (2), 1–21.

Carey, M. (1998). Credit risk in private debt portfolios. Journal of Finance 53 (4), 1363–1387.

Carlson, M., Shan, H., Warusawitharana, M. (2013). Capital ratios and bank lending: A matched bank approach. Journal of Financial Intermediation 22 (4), 663–687.

Chava, S., Jarrow, R. A. (2004). Bankruptcy prediction with industry effects. Review of Finance 8 (4), 537–569.

Chava, S., Stefanescu, C., Turnbull, S. (2011). Modeling the loss distribution. Management Science 57 (7), 1267–1287.

Cherubini, U., Mulinacci, S., Gobbi, F., Romagnoli, S. (2012). Dynamic copula methods in finance. Wiley, Chichester.

Cox, D. R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society 34 (2), 187–220.

Cummings, J. R., Durrani, K. J. (2016). Effect of the Basel Accord capital requirements on the loan-loss provisioning practices of Australian banks. Journal of Banking & Finance 67, 23–36.

Das, S. R., Duffie, D., Kapadia, N., Saita, L. (2007). Common failings: How corporate defaults are correlated. Journal of Finance 62 (1), 93–117.

Denis, D. K., Rodgers, K. J. (2007). Chapter 11: Duration, outcome, and post-reorganization performance. Journal of Financial and Quantitative Analysis 42 (1), 101–118.

Dermine, J., Neto de Carvalho, C. (2006). Bank loan losses-given-default: A case study. Journal of Banking & Finance 30 (4), 1219–1243.

Dewatripont, M., Tirole, J. (1994). The prudential regulation of banks. MIT Press, Cambridge.

Dietrich, A., Hess, K., Wanzenried, G. (2014). The good and bad news about the new liquidity rules of Basel III in Western European countries. Journal of Banking & Finance 44, 13–25.

Djankov, S., Hart, O., McLiesh, C., Shleifer, A. (2008). Debt enforcement around the world. Journal of Political Economy 116 (6), 1105–1149.

Duffie, D., Eckner, A., Horel, G., Saita, L. (2009). Frailty correlated default. Journal of Finance 64 (5), 2089–2123.

Duffie, D., Saita, L., Wang, K. (2007). Multi-period corporate default prediction with stochastic covariates. Journal of Financial Economics 83 (3), 635–665.

Düllmann, K., Trapp, M. (2004). Systematic risk in recovery rates - An empirical analysis of US corporate credit exposures. Deutsche Bundesbank, Discussion Paper, Series 2: Banking and Financial Supervision, No 02/2004.

Embrechts, P., Lindskog, F., McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. Handbook of Heavy Tailed Distributions in Finance 1, 329–384.

European Banking Authority (2014). Draft regulatory technical standards on materiality threshold of credit obligation past due under article 178 of regulation (EU). Consultation Paper.

European Banking Authority (2016a). On results from the EBA impact assessment of IFRS 9. Report.

European Banking Authority (2016b). Risk assessment of the European banking system.

European Banking Authority (2017). Draft regulatory technical standards on the specification of the nature, severity and duration of an economic downturn in accordance with articles 181(3)(a) and 182(4)(a) of regulation (EU) No 575/2013. Consultation Paper.

European Commission (2016). Proposal for a regulation of the European Parliament and of the Council, COM/2016/0850 final - 2016/0360 (COD).

Ferrari, S., Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. Journal of Applied Statistics 31 (7), 799–815.

Financial Accounting Standards Board (2011). Accounting for financial instruments and revisions to the accounting for derivative instruments and hedging activities: Impairment. Supplementary Document.

Financial Accounting Standards Board (2016). Financial instruments - Credit losses (Topic 326). Accounting Standards Update.

Fonseca, A. R., González, F. (2008). Cross-country determinants of bank income smoothing by managing loan-loss provisions. Journal of Banking & Finance 32 (2), 217–228.

Freixas, X., Rochet, J.-C. (2008). Microeconomics of banking, 2nd Edition. MIT Press, Cambridge.

Gambacorta, L., Mistrulli, P. E. (2004). Does bank capital affect lending behavior? Journal of Financial Intermediation 13 (4), 436–457.

Gambacorta, L., Shin, H. S. (2016). Why bank capital matters for monetary policy. Journal of Financial Intermediation (forthcoming).

Genest, C., Rivest, L.-P. (1993). Statistical inference procedures for bivariate archimedean copulas. Journal of the American Statistical Association 88 (423), 1034–1043.

Gobat, J., Yanase, M., Maloney, J. (2014). The net stable funding ratio: Impact and issues for consideration. International Monetary Fund, Working Paper 14 (106).

Gordy, M. B. (2000). A comparative anatomy of credit risk models. Journal of Banking & Finance 24 (1), 119–149.

Gordy, M. B., Howells, B. (2006). Procyclicality in Basel II: Can we treat the disease without killing the patient? Journal of Financial Intermediation 15 (3), 395–417.

Grunert, J., Weber, M. (2009). Recovery rates of commercial lending: Empirical evidence for german companies. Journal of Banking & Finance 33 (3), 505–513.

Gunther, J. W., Moore, R. R. (2003). Loss underreporting and the auditing role of bank exams. Journal of Financial Intermediation 12 (2), 153–177.

Gupton, G. M. (2005). Advancing loss given default prediction models: How the quiet have quickened. Economic Notes 34 (2), 185–230.

Gupton, G. M., Finger, C. C., Bhatia, M. (1997). CreditMetrics - Technical Document. JP Morgan.

Gupton, G. M., Stein, R. M. (2005). LossCalc v2: Dynamic prediction of LGD. Moody's KMV Company.

Gürtler, M., Hibbeln, M. (2013). Improvements in loss given default forecasts for bank loans. Journal of Banking & Finance 37 (7), 2354–2366.

Haugen, R. A., Senbet, L. W. (1978). The insignificance of bankruptcy costs to the theory of optimal capital structure. Journal of Finance 33 (2), 383–393.

Haugen, R. A., Senbet, L. W. (1988). Bankruptcy and agency costs: Their significance to the theory of optimal capital structure. Journal of Financial and Quantitative Analysis 23 (1), 27–38.

Heckman, J. J. (1979). Sample selection bias as a specification error. Econometrica 47 (1), 153–161.

Helwege, J. (1999). How long do junk bonds spend in default? Journal of Finance 54 (1), 341–357.

Hendricks, W., Koenker, R. (1992). Hierarchical spline models for conditional quantiles and the demand for electricity. Journal of the American Statistical Association 87 (417), 58–68.

Hilscher, J., Wilson, M. (2016). Credit ratings and credit risk: Is one measure enough? Management Science (forthcoming).

Höcht, S., Kroneberg, A., Zagst, R. (2011). Explaining aggregated recovery rates. Working Paper.

Höcht, S., Zagst, R. (2010). Loan recovery determinants – A Pan-European study. Working Paper.

Hosmer, D. W., Lemeshow, S., May, S. (2008). Applied survival analysis: Regression modeling of time to event data, 2nd Edition. Wiley Series in Probability and Statistics. Wiley, Hoboken, NJ.

Hu, Y.-T., Perraudin, W. (2002). The dependence of recovery rates and defaults. Working Paper.

Huang, X., Oosterlee, C. W. (2012). Generalized beta regression models for random loss given default. Journal of Credit Risk 7 (4), 45–70.

Hull, J. C. (2015). Risk management and financial institutions, 4th Edition. Wiley, Hoboken, NJ.

International Accounting Standards Board (2011). Financial instruments: Impairment. Supplement to ED/2009/12.

International Accounting Standards Board (2014). IFRS 9 Financial instruments. International Financial Reporting Standards.

Ivashina, V., Scharfstein, D. (2010). Bank lending during the financial crisis of 2008. Journal of Financial Economics 97 (3), 319–338.

Jankowitsch, R., Nagler, F., Subrahmanyam, M. G. (2014). The determinants of recovery rates in the US corporate bond market. Journal of Financial Economics 114 (1), 155–177.

Jo, S., Sekkel, R. (2017). Macroeconomic uncertainty through the lens of professional forecasters. Federal Reserve Bank of Dallas, Working Paper 1702.

Kendall, M. G. (1938). A new measure of rank correlation. Biometrika 30 (1-2), 81–93.

Khieu, H. D., Mullineaux, D. J., Yi, H.-C. (2012). The determinants of bank loan recovery rates. Journal of Banking & Finance 36 (4), 923–933.

Kiefer, N. M. (2009). Default estimation for low-default portfolios. Journal of Empirical Finance 16 (1), 164–173.

Kim, D., Santomero, A. M. (1988). Risk in banking and capital regulation. Journal of Finance 43 (5), 1219–1233.

King, M. R. (2013). The Basel III net stable funding ratio and bank net interest margins. Journal of Banking & Finance 37 (11), 4144–4156.

Kishan, R. P., Opiela, T. P. (2000). Bank size, bank capital, and the bank lending channel. Journal of Money, Credit and Banking 32 (1), 121–141.

Koenker, R. (2005). Quantile regression. Cambridge University Press, Cambridge.

Koenker, R., Bassett, G. (1978). Regression quantiles. Econometrica 46 (1), 33–50.

Koenker, R., Bassett, G. (1982). Robust tests for heteroscedasticity based on regression quantiles. Econometrica 50 (1), 43–61.

Koopman, S. J., Lucas, A., Schwaab, B. (2011). Modeling frailty-correlated defaults using macroeconomic covariates. Journal of Econometrics 162 (2), 312–325.

Kuo, C.-K., Lee, C.-W. (2007). Integrating market and credit risk using a simplified frailty default correlation structure. Journal of Fixed Income 17 (1), 48–58.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., Vishny, R. W. (1998). Law and finance. Journal of Political Economy 106 (6), 1113–1155.

Laeven, L., Majnoni, G. (2003). Loan loss provisioning and economic slowdowns: Too much, too late? Journal of Financial Intermediation 12 (2), 178–197.

Lando, D., Nielsen, M. S. (2010). Correlation in corporate defaults: Contagion or conditional independence? Journal of Financial Intermediation 19 (3), 355–372.

Lee, S. H., Urrutia, J. L. (1996). Analysis and prediction of insolvency in the property-liability insurance industry: A comparison of logit and hazard models. Journal of Risk and Insurance 63 (1), 121–130.

Lee, Y., Poon, S.-H. (2014). Forecasting and decomposition of portfolio credit risk using macroeconomic and frailty factors. Journal of Economic Dynamics & Control 41, 69–92.

Loterman, G., Brown, I., Martens, D., Mues, C., Baesens, B. (2012). Benchmarking regression algorithms for loss given default modeling. International Journal of Forecasting 28 (1), 161–170.

Martin, D. (1977). Early warning of bank failure: A logit regression approach. Journal of Banking & Finance 1 (3), 249–276.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (ed.), Frontiers in Econometrics. Academic Press, New York, 105–142.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. Journal of Finance 29 (2), 449–470.

Michael, J. R. (1983). The stabilized probability plot. Biometrika 70 (1), 11–17.

Nelsen, R. B. (2006). An introduction to copulas, 2nd Edition. Springer, New York.

Orth, W. (2013). Multi-period credit default prediction with time-varying covariates. Journal of Empirical Finance 21, 214–222.

Pan, J., Singleton, K. J. (2008). Default and recovery implicit in the term structure of sovereign CDS spreads. Journal of Finance 63 (5), 2345–2384.

Partington, G., Russel, P., Stevenson, M., Torbey, V. (2001). Predicting return outcomes to shareholders from companies entering Chapter 11 bankruptcy. Managerial Finance 27 (4), 78–96.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. Review of Financial Studies 22 (1), 435–480.

Puri, M., Rocholl, J., Steffen, S. (2017). What do a million observations have to say about loan defaults? Opening the black box of relationships. Journal of Financial Intermediation (forthcoming).

Qi, M., Zhao, X. (2011). Comparison of modeling methods for loss given default. Journal of Banking & Finance 35 (11), 2842–2855.

Repullo, R. (2013). Cyclical adjustment of capital requirements: A simple framework. Journal of Financial Intermediation 22 (4), 608–626.

Repullo, R., Suarez, J. (2013). The procyclical effects of bank capital regulation. Review of Financial Studies 26 (2), 452–490.

Rösch, D., Scheule, H. (2014). Forecasting probabilities of default and loss rates given default in the presence of selection. Journal of the Operational Research Society 65 (3), 393–407.

Sharpe, S. A. (1995). Bank capitalization, regulation, and the credit crunch: A critical review of the research findings. Finance and Economics Discussion Series 95-20. Federal Reserve Board.

Sharpe, W. F. (1978). Bank capital adequacy, deposit insurance and security values. Journal of Financial and Quantitative Analysis 13 (4), 701–718.

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

Siao, J.-S., Hwang, R.-C., Chu, C.-K. (2015). Predicting recovery rates using logistic quantile regression with bounded outcomes. Quantitative Finance 16 (5), 777–792.

Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231.

Smith, M. D. (2003). Modelling sample selection using archimedean copulas. Econometrics Journal 6 (1), 99–123.

Somers, M., Whittaker, J. (2007). Quantile regression for modelling distributions of profit and loss. European Journal of Operational Research 183 (3), 1477–1487.

Therneau, T. M., Grambsch, P. M., Pankratz, V. S. (2003). Penalized survival models and frailty. Journal of Computational and Graphical Statistics 12 (1), 156–175.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. Econometrica 26 (1), 24–36.

Uhrig-Homburg, M. (2002). Valuation of defaultable claims - A survey. Schmalenbach Business Review 54, 24–57.

Wagenvoort, R. (2006). Comparing distributions: The harmonic mass index: Extension to m samples. European Investment Bank, Economic and Financial Report 2006/03.

Wong, B., Partington, G., Stevenson, M., Torbey, V. (2007). Surviving Chapter 11 bankruptcies: Duration and payoff? Abacus 43 (3), 363–387.

World Bank (2015). Doing business 2015: Going beyond efficiency. A World Bank Group Flagship Report.

Yashkir, O., Yashkir, Y. (2013). Loss given default modelling: A comparative analysis. Journal of Risk Model Validation 7 (1), 25–59.