Center of Finance

Universität Regensburg

Center of Finance – Dissertation series No. 2 (2018)

Resolution of Defaulted Loan Contracts

An Empirical Analysis of Default Resolution Time and Loss Given Default

Jennifer Betz

Center of Finance

Faculty of Business, Economics,
and Management Information Systems

Universität Regensburg

Resolution of Defaulted Loan Contracts

An Empirical Analysis of Default Resolution Time and Loss Given Default

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

University Regensburg

submitted by

Jennifer Betz, M.Sc. in Business Administration

Advisors
Prof. Dr. Daniel Rösch
Prof. Dr. Gregor Dorfleitner

Date of disputation
July, 11th 2018

Acknowledgments

First and foremost, I would like to express my true gratitude to Prof. Dr. Daniel Rösch for his continuous support, his huge enthusiasm, and unending patience. I would like to thank him for always having an open door and for being the most encouraging supervisor I could ever imagine. It is a tremendous privilege and pleasure to have him as my first advisor. I am grateful to Prof. Dr. Gregor Dorfleitner for being my second advisor. Furthermore, I would like to thank him for raising my enthusiasm in research and for encouraging me to pursue this path.

I am grateful to all of my colleagues for any statistical and non-statistical discussion and, more than anything, for an amazing time. In particular, I would like to thank my two co-authors, Dr. Ralf Kellner and Dr. Steffen Krüger. Steffen – thank you for one or another lesson in math and for sharing your office with me. Ralf – thank you for your hard deadlines, each and every coffee at math, and for representing an example worth following.

Last but not least, I would like to thank my wonderful family and awesome friends. Above all, I am grateful to Sebastian Ertl for his love, his constant encouragement, and for listening to weird statistical problems in the middle of the night.

I would like to dedicate this thesis to the most wonderful parents in this world, Suzanne and Helmut Betz. *All I am, I am because of you*. Thank you for your unconditional love, your care, and your continuous support during my entire life. Thank you for teaching me not to quit if goals are worth reaching. Dad – thank you for reminding me to take a break from time to time. Mom – thank you for your eternal belief in me, even in times I have lost it.

Contents

List of Figures

List of Tables

Introduction

Motivation and research questions

Financial stability is indispensable for a robust economic system. However, the financial system is exposed to systemic risk. Extensive cascade effects might expand initially financial crises to the entire economic system as in the Global Financial Crisis (GFC). Although causes seem to be limited to the financial sector, effects on the real economy were serve. After the collapse of Lehman Brothers in September 2008, tightened credit conditions in the financial sector transmitted to the real economy. Companies struggled to roll over debt and were confronted with higher interest rates and shorter terms. Campello et al. (2010) suggest that 57% of U.S. American corporations were concerned by tighten credit conditions. Due to limited access to credit markets, smaller companies were severely affected. These corporations are critical for the U.S. American labor market as they employ 40% of the workforce. As consequence, reported unemployment rate increased to 10.1% in October 2009. Combined with the depressed housing and stock markets, household net wealth decreased by 17 trillion USD. In this context, Brian Moynihan (CEO, Bank of America) stated that "[...] *we, as an industry, caused a lot of damage. Never has it been clearer how poor business judgments we have made have affected Main Street.*" (see Financial Crisis Inquiry Commission, 2011). In the final report, the Financial Crisis Inquiry Commission concludes that the GFC was avoidable. Inter alia, they mention failures in financial regulation and supervision and malfunctions in the risk management of systemically important banks as causes of the GFC (see Financial Crisis Inquiry Commission, 2011).

As reaction to the crisis, the Basel Committee on Banking Supervision and the Bank of International Settlement adapted regulations of capital requirements (Basel II, see Basel Committee on Banking Supervision, 2006) to prevent future crises. These regulations are referred to as

Basel III (see Basel Committee on Banking Supervision, 2010). Just recently, the post-crisis reform was finalized (see Basel Committee on Banking Supervision, 2017). Failures causing crises might not only be found in the regulatory framework but also in the risk management of systemically important banks. Credit risk is the most substantial type of risk for the majority of financial institutions (see European Banking Authority, 2016b). In the advanced Internal Rating Based (IRB) approach of the Basel regulations, banks are permitted to use own empirical models to quantify capital requirements for credit risk. These capital requirements are calculated based on three central credit risk parameters – the *Probability of Default* (PD), the *Loss Given Default* (LGD), and the *Exposure At Default* (EAD). While own empirical models for the PD are allowed under the foundation IRB approach, own LGD and EAD estimates are reserved for the advanced IRB approach. Compared to PD modeling (see, e.g., Altman, 1968; Martin, 1977; Campbell et al., 2008; Hilscher and Wilson, 2017; Das et al., 2007; Duffie et al., 2009), the topic of LGD and EAD modeling is rather sparse in academic literature.

This thesis aims to shed light on the topic of LGD modeling. Most of the LGD literature is based on *market-based* LGDs (see, e.g., Qi and Zhao, 2011; Loterman et al., 2012, for comparative studies). These are calculated as one minus the price of defaulted debt instruments 30 days after default (share to par value). Market-based LGDs are available for traded debt such as bonds. Considering loan contracts, only *workout* LGDs are available in most of the cases. Workout LGDs are calculated based on actual recovery cash flows during the resolution process. Thus, characteristics of workout LGDs differ considerably compared to market-based LGDs. First, workout LGDs are shaped by an even more extreme distributional form. Multi-modality seems to be more pronounced and, thus, higher probability masses at the extremes of no loss (LGD = 0) and total loss (LGD = 1) occur. Both modes are shaped by bindings, i.e., LGD values which are exactly zero or exactly one. Second, systematic effects among average LGDs differ due to the underlying process. While market-based LGDs arise at a certain point in time, i.e., 30 days after default, workout LGDs develop over a longer time period – the *Time To Resolution* (TTR) or *Default Resolution Time* (DRT).[1] This hardens the identification of economically and statistically significant (evident) systematic variables, e.g., macro(-economic) variables, as the economic surrounding during the entire resolution process might impact workout LGDs. However, the identification of systematic variables is crucial as LGD predictions are required to reflect economic downturn conditions (see Basel Committee on Banking Supervision, 2006, 2005). Third, workout LGDs are shaped by the resolution bias. Assuming positive dependencies of DRTs and LGDs, bad loan contracts are characterized by long DRTs and high LGDs. These

[1] In this thesis, both terms – TTR and DRT – are used as synonyms.

loans are underrepresented at the end of the observation period as only LGDs of loans with short DRTs and, thus, low LGDs, are observable. This might entail parameters distortions and, consequently, an underestimation of LGDs.

Given the characteristics of workout LGDs, the consideration of the resolution process seems to be crucial in the context of defaulted loan contracts. This thesis pursues an comprehensive empirical analysis of the two central parameters of the resolution process – the DRT and the LGD. Hereby, it aims to answer following research questions.

Research question I | *What are the drivers of DRTs?*

The first paper of this thesis (see Chapter 1, *What drives the time to resolution of defaulted bank loans?*), aims to answer the question concerning general drivers of DRTs. Loan-specific and macro(-economic) variables are considered as covariates. Positive dependency structures of DRTs and LGDs emphasize the importance of DRTs in credit risk management. Long DRTs are accompanied with higher uncertainty regarding the timing of recovery cash flows. Furthermore, long resolution processes increase liquidity and interest rate risk.

Research question II | *How are systematic effects among DRTs?*

The second paper of this thesis (see Chapter 2, *Macroeconomic effects and frailties in the resolution of non-performing loans*), aims to answer the question concerning systematic effects among DRTs. Observable and unobservable systematic factors are considered. Systematic movements in DRTs might imply time-dependent correlation structures, i.e., averagely higher (lower) DRTs at certain points in time. Financial institutions might be able to compensate single defaulted loan contracts with high DRTs, however, correlations might increase the systematic risk of credit portfolios and, thus, further burden liquidity in crises periods.

Research question III | *How are systematic effects among LGDs?*

The third paper of this thesis (see Chapter 3, *Systematic effects among LGDs and their implications on downturn estimation*), aims to answer the question concerning systematic effects among LGDs. Following Basel Committee on Banking Supervision (2006, 2005), LGD predictions are required to reflect economic downturn conditions. Thus, the identification of systematic factors is crucial. However, common macro(-economic) variables might not be suited due to the complexity of identifying reasonable observable variables considering workout LGDs.

Unobservable systematic factors in terms of random effects are applied to analyze their ability to generate sufficiently conservative downturn predictions.

Research question IV | *How are the dependency structures among DRTs and LGDs?*

The forth paper of this thesis (see Chapter 4, *Time matters: How default resolution times impact final loss rates*), aims to answer the question concerning the dependency structures among DRTs and LGDs. A combined modeling approach is developed to deeply examine the dependence structure among DRTs and LGDs allowing for a direct and an indirect channel. Furthermore, effects of the resolution bias are quantified on an in sample and out of sample perspective comparing a pure (standard) LGD model with the combined modeling approach.

Literature

Although the DRT is crucial considering default resolution and, thus, the LGD, of defaulted loan contracts, most of the literature refers to the resolution of bankruptcy. Furthermore, a majority of publications relate to U.S. data, i.e., the resolution of Chapter 7 and Chapter 11 bankruptcies. Bandopadhyaya (1994) applies a hazard rate model to analyze the time spend under Chapter 11, whereas, Helwege (1999) uses Ordinary Least Square (OLS) regression to examine borrower specific influences on the DRT. Bris et al. (2006) run OLS and Heckman models to compare Chapter 7 and Chapter 11 resolutions. Partington et al. (2001), Denis and Rodgers (2007), and Wong et al. (2007) apply survival analysis.[2]

In contrast, the literature on LGD modeling has widened considerably in the last decades (see, e.g., Qi and Zhao, 2011; Loterman et al., 2012, for comparative studies). Given the demand for LGD predictions which reflect economic downturn conditions (see Basel Committee on Banking Supervision, 2006, 2005), the identification of systematic variables is crucial in an LGD modeling context. A common tool to consider systematic effects are observable, i.e., macro(-economic), variables. However, the identification of economically and statistically significant (evident) variables is ambiguous. Thus, some authors completely neglect systematic variables (see Bastos, 2010; Bijak and Thomas, 2015; Calabrese, 2014; Gürtler and Hibbeln, 2013; Matuszyk et al., 2010; Somers and Whittaker, 2007). In other publications, univariate significance (evidence) can not be confirmed in a multivariate context (see Acharya et al., 2007; Brumma et al., 2014; Caselli et al., 2008; Dermine and Neto de Carvalho, 2006; Grunert and Weber, 2009). Reasons might be found in non-linear impacts of observable systematic variables on LGDs. Acharya et al.

[2] A comprehensive literature review regarding the DRT can be found in Chapter 1 (Section 1.1, Introduction) and Chapter 2 (Section 2.1, Introduction).

(2007) find statistical significance of industry distress dummies, but not for continuous variables. Using quantile regression techniques, Krüger and Rösch (2017) identify statistically significant macro(-economic) variables on the inner quantiles of the LGD distribution. Again, this can be traced back to non-linear influences. In parts of the literature, statistical significance (evidence) is not reported (see Altman and Kalotay, 2014; Tobback et al., 2014; Yao et al., 2015). Statistical significance (evidence) might be found where data sets of bonds (see Jankowitsch et al., 2014; Nazemi et al., 2017; Qi and Zhao, 2011), credit cards (Bellotti and Crook, 2012; Yao et al., 2017), or mortgages (Leow et al., 2014; Qi and Yang, 2009) are applied. Bond data sets are usually characterized by market-based LGDs, thus, the identification of economically and statistically significant (evident) macro(-economic) variables might be more straightforward as the LGD is not the result of a complex resolution process, but determined 30 days after default occurred. Credit cards and mortgages belong to the bulk businesses of financial institutions. Hence, resolution processes might be standardized to a higher extend compared to corporate loans.[3]

To the best of my knowledge, no publication exists so far which covers the dependency structures of DRTs and LGDs. However, the importance of DRTs in an LGD modeling context is indicated in the related literature. Dermine and Neto de Carvalho (2006) apply mortality analysis on a data set of defaulted loan contracts, whereas, Gürtler and Hibbeln (2013) are, inter alia, concerned with the resolution bias. They suggest to restrict the data set to avoid biased estimates. Chapter 4 (*Time matters: How default resolution times impact final loss rates*) of this thesis analyzes the possibility to diminish effects of the resolution bias by considering censored observations, i.e., unresolved loan contracts, by a combined modeling approach for DRTs and LGDs. In the credit risk literature, joint modeling approaches are common for PDs and LGDs. Hereby, multivariate random effects are applied to consider time-dependent comovements. (see Bade et al., 2011; Rösch and Scheule, 2010, 2014)

Contributions

Related to research questions I, II, III, and IV which are stated above, the main contributions of this thesis can be structured by the independent research papers which are presented in the individual chapters of this thesis (see Chapter 1, 2, 3, and 4).

Contribution I | *What drives the time to resolution of defaulted bank loans?*

Research question I refers to general drivers of the DRT – i.e., the aim of the first research paper *What drives the time to resolution of defaulted bank loans?* is to analyze which loan-specific

[3] A comprehensive literature review regarding the LGD can be found in Chapter 3 (Section 3.2, Literature review).

and macro(-economic) variables impact the duration of the resolution process. Using OLS regressions, collateralization, seniority, industry, nature of default, and the macro(-economic) environment are identified as important drivers of the DRT. The analysis is conducted for Germany, the United States, and Great Britain. By this means, two major bankruptcy regimes are compared. The German insolvency codes are creditor friendly, whereas, the Anglo-American regulations are rather debtor orientated. The most striking deviations in causalities refer to collateralization and seniority. This can be traced back to the insolvency codes of the considered countries. While the access to collateral in the event of default is straightforward in Germany, it is more complicated in the United States and Great Britain. As consequence, collateralization reduces the DRT to a higher degree in Germany. Creditors seem to be aware of this divergence and seek for the best security mechanism in the limits set by local insolvency codes. Thus, collateralization appears to be more common in Germany, while Anglo-American creditors demand higher ranks in the seniority order.

By the inclusion of year fixed effects, first indications of time-dependent comovement in DRTs arise. Average DRTs seem to be higher (lower) during downturns (upturns). This is also true considering the LGD as dependent variable. Furthermore, LGDs are driven by similar variables, particularly, collateralization and macro(-economic) variables. Significant year fixed effects indicate similar time patterns in LGDs compared to DRTs.

Contribution II | *Macroeconomic effects and frailties in the resolution of non-performing loans*

Research question II refers to systematic effects among DRTs. As first indications of time-dependent comovement in DRTs arise (see **Contribution I** or Chapter 1), the second research paper *Macroeconomic effects and frailties in the resolution of non-performing loans* aims to analyze systematic effects among DRTs in more detail. Using Cox Proportional Hazard (PH) regressions, three model specifications are compared. In the first specification (model I), loan-specific variables are applied to explain cross-sectional variation in DRTs. The second specification (model II) additionally includes observable systematic effects, i.e., macro(-economic) variables, to account for time-dependent variations. In the third specification (model III), unobservable systematic effects, i.e., frailty effects, are introduced. These unobservable systematic effects seem to impact DRTs to a rather high extend even after controlling for common loan-specific and macro(-economic) variables. Stated differently, DRTs seem to be clustered in the time line.

Economic consequences of clustered DRTs might be considerable. First, the liquidity of financial institutions is burdened in downturn periods as DRTs are systematically longer. Aside from the direct availability of liquidity, the implementation of the Net Stable Funding Ratio (NSFR), where additional medium and long term liquidity is demanded for certain facilities, e.g., non-performing loans (see Board of Governors of the Federal Reserve System, 2016), adds additional pressure on the liquidity of financial institutions. Second, portfolio loss distributions of non-performing loan contracts are affected by clustered DRTs. Due to dependency structures of DRTs and LGDs, systematic patterns in DRTs are transfered to the loss side. The portfolio loss distribution is shifted towards higher values in downturn periods. Furthermore, the range of the distribution is broadened independent of the economic surrounding. This indicates that expected portfolio losses rise in downturns, while unexpected losses are constantly increased.

Contribution III | *Systematic effects among LGDs and their implications on downturn estimation*

Research question III refers to systematic effects among LGDs. Observable and, in particular, unobservable systematic effects have considerable impacts on DRTs (see **Contribution II** or Chapter 2). Thus, LGDs might be shaped by time-dependent comovements. This is of high relevance considering the demand for LGD predictions which reflect economic downturn conditions (see Basel Committee on Banking Supervision, 2006, 2005). Using a Bayesian Finite Mixture Model (FMM) with a probabilistic substructure, LGD distributions depending on covariates are estimated in the third research paper *Systematic effects among LGDs and their implications on downturn estimation*. Time-dependent random effects are included in the modeling framework to quantify the systematic nature of LGDs. By this means, deviations among the considered regions, i.e., the United States and Europe, arise. While realizations of the random effect seem to origin independently from an identical distribution in the United States, systematic patterns are characterized by cyclical nature expressed by an autoregressive (AR) process in Europe. The realizations of the random effects are compared with time patterns of common macro(-economic) variables. Thereby, considerable discrepancies occur. These deviations are emphasized when macro(-economic) variables are included in the modeling framework as their impacts seem not to be evident or limited regarding their magnitude.

Furthermore, a methodology to generate appropriate downturn estimations based on random effects is suggested and compared to approaches in the literature. The most common approach refers to the use of macro(-economic) variables in the modeling framework. Due to the limited evidence and magnitude, downturn estimates based on macro(-economic) variables seem to underestimate the probability mass of high losses, while the suggested methodology delivers sufficiently conservative estimates. Further approaches proposed by Calabrese (2014) and Bijak and Thomas (2015) tend to be over-conservative.

Contribution IV | *Time matters: How default resolution times impact final loss rates*

Research question IV refers to dependency structures of DRTs and LGDs. As an interconnection of DRTs and LGDs appears descriptively (see **Contribution I** or Chapter 1 and **Contribution II** or Chapter 2), dependencies of the two parameters are conceivable. In the fourth research paper *Time matters: How default resolution times impact final loss rates*, a joint modeling approach for DRTs and LGDs is developed. Therefor, an Accelerated Failure Time (AFT) model for the DRT is combined with an FMM with probabilistic substructure for the LGD (see **Contribution III** or Chapter 3). This approach allows for direct and indirect dependencies. To reflect direct dependencies, the DRT is included into the LGD model. Multivariate normally distributed random effects are implemented to display indirect dependencies, i.e., comovements in the time line. Positive dependency structures of DRTs and LGDs arise which are even more pronounced in extreme economic surroundings. In downturns (upturns), DRTs are longer (shorter) which burdens financial market liquidity. Moreover, LGDs are even higher in downturns due to the strengthened dependence.

The dependence of DRTs and LGDs introduces the resolution bias. In more recent time periods, only LGDs of loan contracts with short DRTs are observable. These loans tend to exhibit high LGDs due to the positive dependence. Thus, loans with long DRTs and high LGDs are underrepresented towards the end of the observation period. Applying a pure (standard) LGD model causes parameter distortions and, consequently, an underestimation of average LGDs on an out of sample perspective. Effects of the resolution bias are diminished by the combined approach. Thus, LGD predictions are adequate on an out of sample perspective.

Structure

This thesis consists of four independent research papers with varying co-authors.[4] Chapter 1 presents the first paper (*What drives the time to resolution of defaulted bank loans?*). The second paper (*Macroeconomic effects and frailties in the resolution of non-performing loans*) is subject to Chapter 2. In Chapter 3, the third paper (*Systematic effects among LGDs and their implications on downturn estimation*) is propound. The fourth and last paper (*Time matters: How default resolution times impact final loss rates*) is comprised in Chapter 4. The Conclusion summarizes, discusses, and provides an outlook.

[4] The co-authors and the current state of the research papers are mentioned at the beginning of each chapter.

Chapter 1

What drives the time to resolution of defaulted bank loans?

This chapter is joint work with Ralf Kellner* and Daniel Rösch† published as:

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

https://doi.org/10.1016/j.frl.2016.03.013

Abstract

Using a unique data base of Global Credit Data with individual loan information from small and medium sized entities in Germany, Great Britain and the United States, we evaluate the time to resolution of defaulted loans. A comparison across countries reveals country specific drivers for the resolution time which can be explained fairly well by differences in the regulatory and legal framework. Lenders seem to be aware of these differences and adjust their lending behavior in the limits set by these bankruptcy systems of the countries.

Keywords: credit risk; bankruptcy; resolution of financial distress; time to resolution; resolution bias

JEL classification: C51, G01, G28

* University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: ralf.kellner@ur.de.

† University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: daniel.roesch@ur.de.

1.1 Introduction

The time to resolution (TTR) of defaulted loan contracts is of great relevance to all kinds of creditors. Following Hotchkiss et al. (2008), indirect costs – e.g. opportunity costs and reputational losses – are characterized by a considerable magnitude and importance. However, these cost are not directly captured in the loss rate. As they are challenging to measure, the TTR might serve as a proxy (see, e.g., Franks and Torous, 1989; Bris et al., 2006; Annabi et al., 2012). Furthermore, the TTR seems to be positively correlated with the loss given default (LGD) of loan contracts. High TTRs increase uncertainty regarding the timing of cash flows during resolution and, therefore, liquidity and interest rate risk. To the best of our knowledge, no analysis exists so far which deeply examines drivers for the TTR on a transnational basis even though a profound understanding of it seems crucial in an international setting. A reason for missing studies in this field of literature might be found in the lack of data availability. Thus, our paper uses access to a unique loss database provided by Global Credit Data (GCD)[1] and conducts a detailed and comprehensive analysis of the TTR across Germany, Great Britain, and the United States. We provide insights which components of loans contribute to a short TTR and to which extent it depends on external factors such as the macroeconomic environment. Thereby, substantial differences among Germany, Great Britain, and the United States arise. These might be ascribed to discrepancies in the insolvency regimes.

While there exists a variety of analyses which examine drivers and estimation methods for the loss rate (see, e.g., Grunert and Weber, 2009; Qi and Yang, 2009; Bastos, 2010; Qi and Zhao, 2011; Loterman et al., 2012), the literature regarding the TTR is more limited even though its importance is indicated in the related analyses. Using a database from a Portuguese bank, Dermine and Neto de Carvalho (2006) analyze recovery rates (RRs) by means of survival time analysis. Their results show the importance of the resolution process as the impact of determinants of RRs change over time. In addition, they point out the importance of the timing of cash flows during the resolution process in the presence of interest rate risk. Gürtler and Hibbeln (2013) empirically find LGDs to be positively correlated with the TTR and quicker resolution times for defaulted loans which return back to performance. Davydenko and Franks (2008) detect transnational discrepancies caused by varying legislations with respect to the LGD. This might hold true for the TTR. However, previous analyses are restricted to individual

[1] 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.

countries while data sets are usually specific regarding their origin, i.e, banks or courts. Most studies consider the bankruptcy system in the United States – in particular Chapter 11. Helwege (1999) analyzes the TTR of junk bonds using ordinary-least-square (OLS) regressions. In contrast to our analysis, he focuses on borrower specific characteristics and uses a bankruptcy portfolio of minor debtor quality. Bris et al. (2006) use a data set of bankruptcies in Arizona and New York. They run OLS and Heckman models with the log transformation of the TTR as dependent variable and find that the outcome of bankruptcy (reorganization vs. liquidation) has no influence on the TTR. Denis and Rodgers (2007) and Wong et al. (2007) apply survival methods for examining the TTR of Chapter 11 bankruptcies. Overall they detect firm size, pre-default performance and the macroeconomic environment to be important drivers for TTR while accounting information seems to be less relevant. Few analyses have been made with respect to the TTR in other countries. Focusing on Portugal, the main interest of Bonfim et al. (2012) is on access to credit after default. They find that large firms tend to have shorter TTR. This contradicts with results in the United States from Denis and Rodgers (2007) and indicates country specific differences regarding the TTR. Dewaelheyns and Van Hulle (2009) examine bankruptcies in Belgium and observe that, among others, secured debt and industry conditions play an important role for the TTR. However, non of these analyses examine transnational differences with respect to the TTR.

Hence, we try to fill this gap and contribute to the literature in three ways. First, we investigate important drivers for the TTR of loan contracts using a database containing loans of small and medium sized entities (SMEs) from Germany, Great Britain, and the United States. Thereby, we cover two major bankruptcy regimes, i.e., Germany being traditionally creditor friendly and the Anglo American area more debtor orientated. In a second step, we examine whether deviations in the insolvency codes impact the determinants of the TTR on an inner-country basis and, thus, whether adjustments of lending regularities arise. This approach is motivated by Davydenko and Franks (2008) who show how differences in creditors' rights impact the general lending behavior in France, Germany, and the United Kingdom. Third, we analyze effects of the macroeconomic environment on the TTR. By including defaults from 2000 until 2014, our analysis covers at least one complete economic cycle.

After controlling for explanatory variables, we find that the resolution process in Germany is shortest compared to Great Britain and the United States. The TTR of American (British) loan contracts is c.p. on average 0.1 (0.5) years longer. This is in line with a higher degree of efficiency regarding the resolution of insolvency in Germany that is assigned by the World Bank.

For the overall dataset, we find seniority, nature of default, collateralization, industry, and the macroeconomic environment to be important determinants for the TTR. Additionally, we examine significant differences across countries. The most important results refer to collateralization and its impact on the seniority order. This seems to be driven by varying regulations regarding access to collateral during resolution and the importance of the seniority order during the bankruptcy proceeding given a certain access to collateral. Assuming easy access to collateral during resolution, its existence should reduce the TTR. In Germany, *real estate* backed loans as well as loans secured by *other collateral* types are resolved faster by 0.1 and 0.2 years, respectively. The access to collateral seems to be more complicated in the Anglo American counties as *other collateral* enhances TTR by 0.1 years and *real estate* is insignificant in the United States.

It seems that creditors are aware of these country specific features as they seek for the best safety mechanism in the limits set by the insolvency code by adjusting their lending behavior. As the access to collateral is easier in Germany compared to Great Britain and the United States, a majority of German loans is collateralized (72% general, 53% *real estate*). On the contrary, collateralization in Great Britain (67% general, 42% *real estate*) and the United States (64% general, 12% *real estate*) is less usual. Creditors in the Anglo American area seem to compensate this by demanding the highest rank in the seniority order (*super senior*). In Great Britain (44%) and the United States (84%), a higher fraction of loans exhibit the *super senior* status while German creditors do not depend on being the one and only preferred claimant (5% *super senior*). This seems to work for creditors in Great Britain as the *super senior* status reduces TTR by 1.2 years. In the United States it suffices to be among preferred claimants as no significant difference regarding TTR occurs among situations with one single and more than one (*pari-passu*) preferred claimants. However, being not among preferred claimants increases the TTR by 1.4 years.

The remainder of this paper is structured as follows. Section 1.2 provides the data description and descriptive statistics. Section 1.3 presents our main analysis. Section 1.4 provides several robustness tests. Section 1.5 concludes.

1.2 Data description

Our data set consists of a subsample of the unique loss data base composed by GCD. The data base contains historical loss data from 44 member banks. In this paper, we analyze loans of SMEs whose jurisdiction is located in Germany, Great Britain, and the United States. We focus on these countries as they represent two major bankruptcy regimes, i.e., Germany being traditionally creditor friendly and the Anglo American countries more debtor orientated. The time span of the entire data base reaches from September 1971 until Juli 2014. To ensure a consistent default definition and, thereby, unbiased estimation results, we refer to the default definition set by the Basel Committee on Banking Supervision (2006). A default occurs if an obligor is "unlikely to pay" or "past due more than 90 days on any material credit obligation" (§452).[2] Pursuant to Brumma et al. (2014), this definition is implemented since the year 2000 which is why we restrict the time period from 2000 to 2014. Furthermore, we eliminate loans with EADs smaller than 500 EUR to satisfy the materiality threshold of the European Banking Authority (2016a). Despite the reference to borrower level, we apply the rule on loan level as loans of this size seem unreasonable. To correct for minor input errors, we follow Höcht and Zagst (2010) and Höcht et al. (2011) who developed selection criteria on cash flow basis. We apply this approach with the distinction that we separately consider payments during and after the resolution process. Hence, the first criterion is calculated as the sum of all relevant transactions (including charge-offs) divided by the outstanding amount of the loan. Loans falling below 90% or exceeding 110% are sorted out. To verify post resolution payments, we evolve a second criterion. Thereby, the sum of all post resolution payments is divided by a fictional outstanding amount at the resolution date. The barriers are set to −10% and 110%. Finally, we eliminate loans with abnormal low and high LGDs ($< -50\%$ and $> 150\%$).[3] A subset of 24,870 individual loans remains.

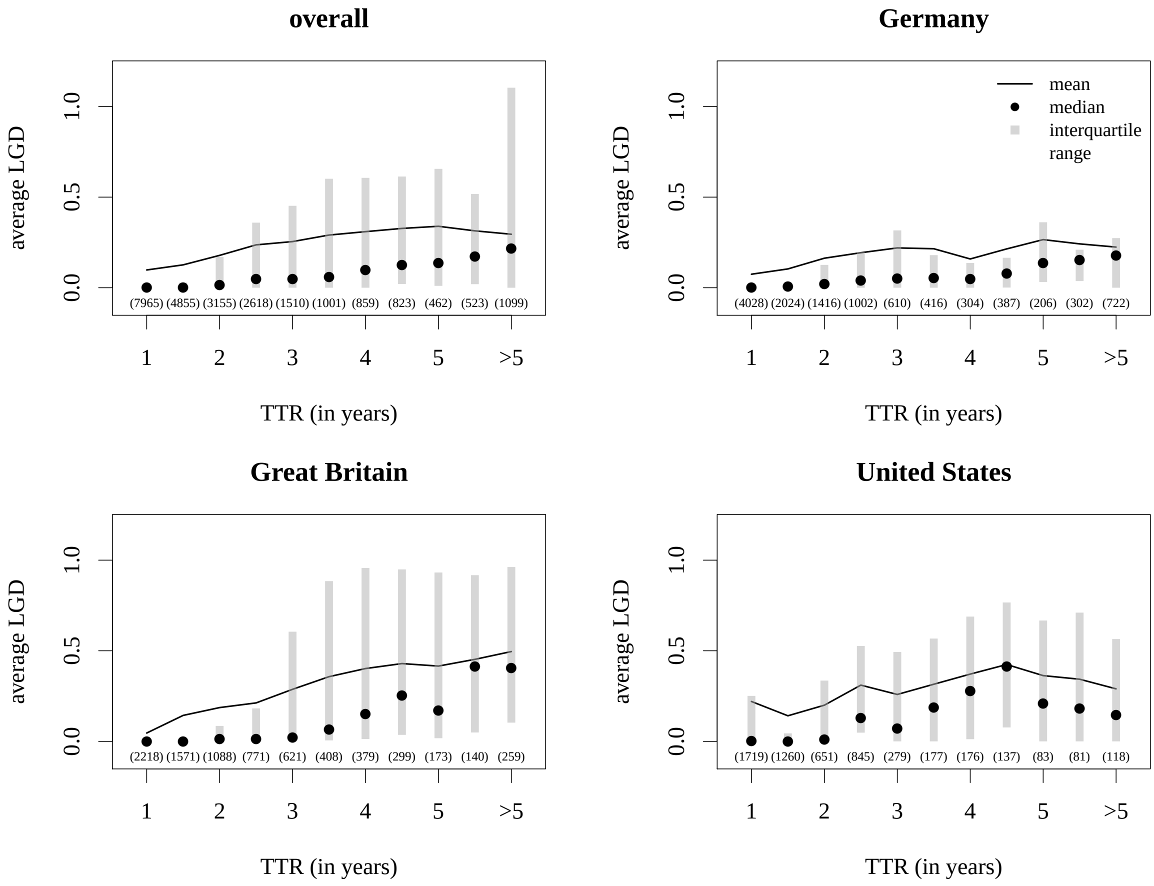
Figure 1.1 shows the relation between the average LGD and the TTR (in years) for the entire data set and the three country subsets. The black lines display the average LGD for the specified TTR buckets. The interquartile ranges are represented by the gray boxes, whereby, the black dots are the medians. The quantity of loans in the buckets is given in brackets. Investigating the general relation of credit losses and the TTR, the upper left panel of Figure 1.1 shows a positive dependence between the average LGD and the TTR in the overall data set. A longer TTR comes along with higher LGDs, and vice versa. The results differ if the same relationship is regarded on country level. While the link between TTR and average LGD seems to increase monotonously in Great Britain, the peak in the United States arises around four and a half years after default occurs. In Germany, a slight drop in the average LGD appears for a TTR of four years.

Figure 1.2 shows histograms and corresponding kernel density estimates of the TTR for the

[2] Note that we use default and financial distress in a synonymous way. The data set contains both, loans which are subject to common resolution mechanisms (e.g., restructuring, liquidation) and *cured* loans which returned back to performance after they got into financial distress.

[3] Economic LGDs are employed. The calculation of the stated LGDs are summarized in the appendix.

Figure 1.1: Relationship between the TTR and average LGD for the overall dataset and on country level



Notes: Figure of the average LGD with respect to the resolution time.

overall sample and on country level. The distribution of the TTR is asymmetric and extremely skewed to the right. This indicates that most of the loans exhibit a rather short TTR, but very long resolution times are probable with a certain (small) likelihood. Differences between the countries under investigation can be observed. The overall sample, Germany, and Great Britain show a unimodal distribution while the United States are characterized by bimodality with the second modus around 1.7 years after default.

Descriptive statistics of considered quantities for the overall data set and on country level are displayed in Table 1.1. The first row reports the sample size. Most of the 24,870 loans are located in Germany (11,417 loans) followed by Great Britain (7,927 loans) and the United States (5,526 loans). Hence, the overall data set may be dominated by German loans. Furthermore, Table 1.1 contains descriptive statistics regarding dependent and independent variables of the subsequent regression analysis.[4] The mean of the TTR is lowest for the United States (1.26 years) followed by Germany (1.52 years) and Great Britain (1.54 years).

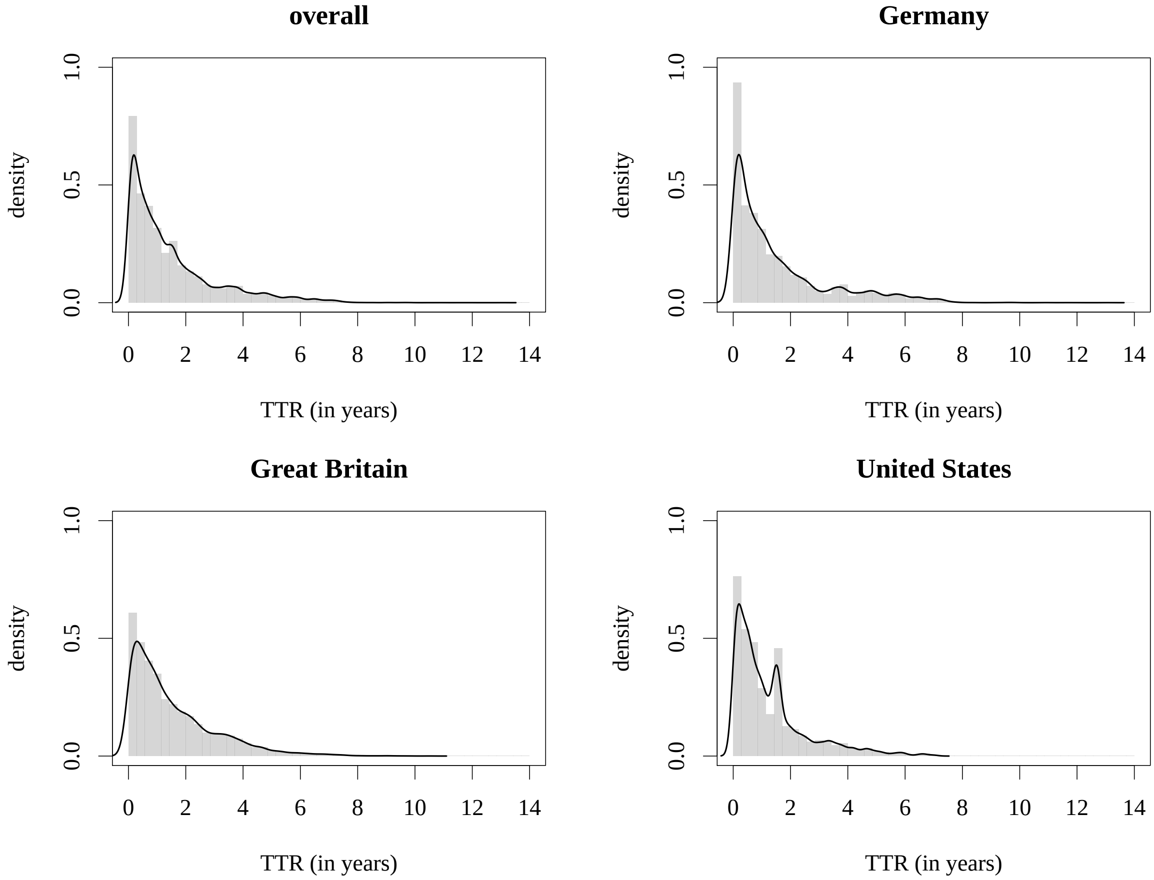
[4] See Section 1.3.

Table 1.1: Descriptive statistics for the overall data set and on country level

		Overall	Germany	Great Britain	United States
	n	24,870	11,417	7,927	5,526
	Dependent variable				
TTR	Mean	1.4696	1.5225	1.5386	1.2616
	Median	0.9333	0.8944	1.0639	0.8472
	Standard deviation	1.5484	1.7079	1.4659	1.2755
	Independent variables | Metric				
EAD	Mean	748,248.29	337,505.03	528,730.33	1,911,761.64
	Median	100,326.80	61,487.05	89,000.41	477,962.81
	Standard deviation	3,511,570.91	3,158,226.29	2,824,157.23	4,656,323.68
Number of collateral	Mean	2.2944	2.2142	3.2623	1.0717
	Median	1.0000	1.0000	1.0000	1.0000
	Standard deviation	4.8586	2.9864	7.5875	1.5452
Equity index	Mean	6.19%	9.84%	3.77%	2.14%
	Median	12.71%	18.38%	9.01%	7.58%
	Standard deviation	21.39%	23.45%	17.52%	20.74%
GDP	Mean	2.99%	2.78%	3.59%	2.58%
	Median	3.58%	2.26%	4.25%	3.59%
	Standard deviation	2.35%	1.70%	2.69%	2.77%
	Independent variables | Categoric				
Facility type	Medium term	53.19%	54.71%	48.67%	56.55%
	Short term	30.60%	28.61%	45.45%	13.41%
	Other / Unknown	16.21%	16.69%	5.88%	30.04%
Seniority code	Pari-passu	60.34%	86.25%	55.48%	13.79%
	Super senior	35.24%	5.47%	44.34%	83.71%
	Non senior	1.62%	2.90%	0.16%	1.09%
	Unknown	2.79%	5.39%	0.01%	1.41%
Nature of default	90 days past due	34.53%	37.71%	23.45%	43.85%
	Unlikely to pay	14.69%	8.50%	14.75%	27.40%
	Bankruptcy	8.79%	9.66%	11.81%	2.64%
	Charge-off / provision	22.17%	31.79%	23.10%	0.94%
	Sold at material credit loss	0.53%	0.20%	0.04%	1.92%
	Distressed restructuring	5.34%	10.48%	1.27%	0.56%
	Non accrual	11.39%	1.54%	18.10%	22.11%
	Unknown	2.56%	0.11%	7.48%	0.58%
Guarantee indicator	NO	68.05%	75.55%	65.89%	55.66%
	YES	31.88%	24.44%	34.05%	44.14%
	Unknown	0.07%	0.01%	0.06%	0.20%
Collateral indicator	NO	31.13%	27.67%	32.75%	35.98%
	Other collateral	28.58%	19.70%	25.15%	51.86%
	Real estate	40.27%	52.63%	42.08%	12.12%
	Unknown	0.01%	0.00%	0.01%	0.04%
Cured indicator	NO	63.60%	64.26%	57.56%	70.88%
	YES	36.40%	35.74%	42.44%	29.12%
Industry	Finance, insurance, RE	18.42%	12.24%	28.98%	16.05%
	Agriculture, foresty, fishing	1.33%	1.15%	1.67%	1.21%
	Mining	0.32%	0.14%	0.37%	0.62%
	Construction	8.79%	5.19%	13.18%	9.92%
	Manufacturing	13.54%	11.93%	13.89%	16.38%
	Transp., commu., sanitary services	3.68%	3.67%	3.38%	4.14%
	Wholesale and retail trade	16.12%	13.73%	20.99%	14.04%
	Services	24.94%	34.67%	16.34%	17.17%
	Unknown	12.86%	17.27%	1.21%	20.47%

Notes: The table contains mean, median, and standard deviation for variables of metric nature and proportions for variables of categoric nature, respectively.

Figure 1.2: Histograms and kernel densities of the TTR for the overall data set and on country level



In Section 1.3, we additionally control for various potential input parameters. Classical metric determinants are the exposure at default (EAD) and macroeconomic variables. We further include the number of collateral as multiple types of collateral can be assigned to a single loan contract. In our data set, loans located in the United States are on average considerably larger (1,911,761.64 EUR) than their counterparts from Great Britain (528,730.33 EUR) and Germany (337,505.03 EUR). The reason for this difference may be found in the heterogeneous ease of getting credit. Since 2004, the World Bank published a score evaluating this topic. Amongst other legal subjects, the series *Doing Business* covers – under the section *Getting Credit* – the ease of receiving credit lines (see World Bank, 2015a,b,c,d). Thereby, not only the case of getting credit is analyzed but also its achievable quantity. The score is survey based and expressed as a distance to frontier with 100 representing the optimality. Regarding this score, the United States reaches the second place with a score of 95.00 followed by Great Britain (75.00, 17th place) and Germany (70.00, 24th place). Thus, the access to credit seems to be easier in the United States compared to Great Britain and Germany. The differences in the average EADs might reflect this consideration.

Corresponding to the number of collateral, loans located in Great Britain are on average collateralized with more assets (3.26) than those from Germany (2.21) and the United States (1.07). However, since the median reveals one for all, the majority of loans seems to be secured by a single collateral. To embed macroeconomic variables, several factors were tested with the return of the equity index and of the gross domestic product (GDP) being top of the range regarding the explanatory power in the affiliating analysis.[5]

In addition, several loan specific categoric variables are included – such as seniority, guarantee, collateral type, and industry indicators. These are also common determinants in modeling LGD (see, e.g., Acharya et al., 2007; Bastos, 2010). It is also controlled for the facility type, nature of default, and if the loan is cured or not, i.e., if the loan returned to performing after being in financial distress.[6] Table 1.1 displays the shares of the corresponding categories for the overall data set and divided by country. There are solely minor country differences concerning the facility type. Approximately, half of the loans are medium term facilities, whereas the other half is grouped amongst short term and other facilities. Great Britain exhibits the greatest share of short term facilities in the data set (45.45%). Loans located in Germany and the United States are characterized trough a slightly higher proportion in other facilities.

Regarding the seniority code, *pari-passu* is the prevailing type in the overall data set as well as in Germany (86.25%) and Great Britain (55.48%). In contrast, loans from the United States are mainly *super senior* (83.71%) and only a small fraction (13.79%) appears to be *pari-passu*.[7] With respect to the nature of default, only minor country differences are observable. A majority of the loans entered default status due to the *90 days past due* criterion (around one third). Whereas, the proportion is highest in the United States (43.85%), followed by Germany (37.71%) and Great Britain (23.45%). The default condition *unlikely to pay* in general is most common in Great Britain. The United States are shaped through the rather precise payment delay. Thereby, different default preconditions might be a reason. We return to this topic in more detail in Section 1.3.

Guarantees seem to be more common in the United States (44.14%) and Great Britain (34.05%) compared to Germany (24.44%). Overall, 31.88% of all loans are characterized by some kind of guarantee. Generally, collateralization appears to be more familiar in all considered countries.

[5] The market indices are represented by the DAX, the FTSE, and the Dow Jones for Germany, Great Britain, and the United States.

[6] *Return to performing* indicates that the loan continues to exist after default because the obligor is back to a sound rating.

[7] *Super senior* describes a priority order where only one creditor has prior claims. If there is at least another claimant of equal rank, the seniority is defined as *pari-passu*.

While *real estate* represents the dominating asset class in Germany (52.63%) and Great Britain (42.08%), it seems considerably less important in the United States (12.12%). A reason for this could lie in different legal and regulatory determinations regarding the handling of collateral during the resolution process. A detailed analysis of this observation follows in Section 1.3.

Overall, around one third of the defaulted loans achieve the cured state, i.e., returned to performing. This proportion is highest for loans located in Great Britain (42.44%), followed by Germany (35.74%) and the United States (29.12%). Again, this might be ascribed to different default preconditions (see Section 1.3). Regarding the industry, the overall data set is dominated by loans in the *service* sector (24.94%) which is driven by great proportions in Germany (34.67%) and the United States (17.17% – highest share of all industries expect *unknown*). However, Great Britain is marked by high amounts in the industries *finance*, *insurance*, *real estate (RE)* (28.98%) and *wholesale and retail trade* (20.99%). This partly corresponds with general industry proportions. Table 1.2 contains the sector shares of the benchmark indices as of the 01.01.2014.

Table 1.2: Industry sectors regarding to equity indices

	DAX	**FTSE**	Dow Jones (**DJ**)
Oil and gas	0%	4%	7%
Basic materials	20%	9%	3%
Industrials	13%	17%	17%
Consumer goods	20%	13%	10%
Health care	10%	5%	13%
Consumer services	3%	20%	13%
Telecommunications	3%	2%	3%
Utilities	7%	5%	0%
Financials	17%	24%	17%
Technology	7%	2%	17%
	MDAX	**FTSE** 250	**DJ** mid cap
Oil and gas	0%	4%	8%
Basic materials	16%	5%	5%
Industrials	30%	19%	26%
Consumer goods	14%	6%	16%
Health care	6%	4%	10%
Consumer services	16%	20%	14%
Telecommunications	0%	2%	0%
Utilities	0%	1%	0%
Financials	16%	36%	3%
Technology	2%	4%	18%
	SDAX	**FTSE** small cap	**DJ** small cap
Oil and gas	0%	2%	5%
Basic material	0%	3%	8%
Industrials	38%	25%	25%
Consumer goods	18%	6%	8%
Health care	2%	3%	16%
Consumer services	22%	12%	15%
Telecommunications	0%	1%	0%
Utilities	0%	0%	0%
Financials	20%	42%	2%
Technology	0%	7%	21%

Notes: Proportions of the industry sector with respect to the equity indices in terms of the ICB (Industry Classification Benchmark) code.

Since we focus on SMEs it may be misleading to consider solely the DAX, FTSE, and the Dow Jones as they contain the largest companies of the economies. Therefore, the corresponding small and medium cap indices are taken into account. In Great Britain, *financials* is the dominating sector (FTSE: 24%, FTSE 250: 36%, FTSE small cap: 42%) which meets the sector specification *finance*, *insurance*, *real estate (RE)*. Whereas, it is less marked in Germany and the United States. In the GCD data base, Germany and the United States are characterized by a high proportion in the *services* sector. However, Table 1.2 indicates that Germany and the United States are strongly shaped by *industrials* with respect to SMEs.

1.3 Determinants of the TTR

In this section, regression models are applied to study determinants of the TTR in general and to examine deviations on country level. Furthermore, qualitative analyses purpose to find reasons for this differences. By this means, we aim to provide new insights into the resolution process and what significant disparities might be relevant to creditors and regulators in distinct countries.

In the multiple regression model, the TTR serves as the dependent variable[8] while the quantities of Table 1.1 are used as regressors. Standard errors are clustered by year. Table 1.3 and 1.5 contain the results of the regression models for the overall data set and on country level. Independent variables are illustrated in the first column and according types (in case of categorical variables) are given in the second column.[9] In a first step, general drivers of the TTR are derived. In the second part of this section, we focus on country specific differences.

1.3.1 General drivers of the TTR

Starting with the EAD, we find a significantly positive impact on the TTR in the overall data set. This implies that loans of larger size demand on average a longer TTR. The impact could be based on a higher level of complexity and administrative effort accompanied with the resolution of larger loans. In the model, the natural logarithm of the EAD is implemented. Hence, the relationship between the TTR and the loan size is characterized by a non-linear component.

[8] Despite the non-negative restriction of the TTR, we apply the level specification of the linear regression due to its higher explanatory power with respect to the adjusted R-squared compared to the log transformation. Appendix 1.A shows that our general results remain stable when regressing on log transformed TTR.

[9] Noteworthy, not all categories are integrated in the subset models since some types are nonexistent in these samples. For example, there are no loans of German origin which defaulted in the year 2014.

Table 1.3: Regression results of the TTR for the overall data set

		Coef.		SE
Intercept		3.694	***	(0.0782)
log(**EAD**)		0.062	***	(0.0045)
Time	2001	-1.401	***	(0.0723)
(2000)	2002	-1.706	***	(0.0638)
	2003	-2.361	***	(0.0591)
	2004	-2.330	***	(0.0622)
	2005	-2.670	***	(0.0547)
	2006	-2.555	***	(0.0544)
	2007	-2.526	***	(0.0570)
	2008	-2.279	***	(0.0710)
	2009	-2.840	***	(0.0786)
	2010	-2.492	***	(0.0599)
	2011	-2.988	***	(0.0587)
	2012	-3.243	***	(0.0585)
	2013	-3.300	***	(0.0604)
	2014	-3.602	***	(0.1069)
Country	United States	0.115	*	(0.0447)
(Germany)	Great Britain	0.460	***	(0.0337)
Facility	Short term	-0.104	***	(0.0190)
(Medium term)	Other	-0.095	***	(0.0244)
Seniority	Super senior	-0.359	***	(0.0254)
(Pari-passu)	Non senior	-0.032		(0.0641)
	Unknown	-0.583	***	(0.0592)
Nature of default	Sold at material credit loss	-0.432	***	(0.1023)
	Unlikely to pay	0.078	**	(0.0248)
	Charge-off / provision	0.205	***	(0.0227)
	Non accrual	0.483	***	(0.0277)
	Distressed restructuring	0.526	***	(0.0384)
	Bankruptcy	0.625	***	(0.0352)
	Unknown	-0.175	***	(0.0369)
Guarantee	YES	0.096	***	(0.0177)
(NO)	Unknown	-1.295	***	(0.3912)
Collateral	Other collateral	-0.259	***	(0.0215)
(NO)	Real estate	-0.215	***	(0.0205)
	Unknown	-0.061		(0.6613)
Number of collateral		-0.004	**	(0.0014)
Cured (NO)	YES	-0.569	***	(0.0176)
Industry	Mining	-0.415	**	(0.1299)
(Finance, insurance, RE)	Transp., commu., sanitary services	-0.199	***	(0.0436)
	Services	-0.078	**	(0.0239)
	Wholesale and retail trade	-0.046	.	(0.0255)
	Manufacturing	0.021		(0.0290)
	Agric., forestry, fishing	0.155	**	(0.0580)
	Construction	0.221	***	(0.0307)
	Unknown	0.209	***	(0.0381)
Equity Index		-0.534	***	(0.0605)
GDP		-12.171	***	(0.7551)
Adjusted R-squared				43.80%
F-statistic				421.92
p-value				0.0000

Notes: Results of the multiple linear regression regarding the overall data set. Significance codes: *** 0.001, ** 0.01, * 0.05, · 0.1. Standard errors (SE) are clustered by year.

We include dummies for the default year of the loan contract to address two issues. Firstly, they control for time varying effects on the TTR, such as level shifts even if they are of non-linear nature. Additional impacts due to the macroeconomic environment that fail to be integrated trough variables might be indirectly taken into account as adverse conditions could have a worse impact than actually indicated by the GDP or equity indices. Secondly, time dummies might absorb effects of a potential resolution bias.[10] In the overall data set, the coefficients of these dummies have significantly negative impacts implying that the TTR decreases compared to the base year 2000. With exception of the years 2004, 2010, and the time period from 2006 to 2008, they are monotonously decreasing. This indicates a shorter resolution of distressed loans in recent years. The years 2004, 2010, and the period from 2006 to 2008 are marked by economic turmoil.[11] The breaks of the monotony in these years seems to refer to a rather longer TTR in a hard economic environment.

To analyze country specific differences, we primarily focus on the country dummies in this section.[12] In Section 1.2, we determined the on average shortest TTR in the United States (1.26 years) and a rather longer one in Germany (1.52 years) and Great Britain (1.54 years). After controlling for other variables, a different picture appears. With Germany being the reference category, the coefficients of the country dummies show a significantly positive sign. This suggests that the resolution time is *ceteris paribus* (c.p.) on average 0.1 years longer in the United States and even 0.5 years longer in Great Britain compared to Germany. This results seems to relate to the efficiency of default resolution. Again, we refer to the *Doing Business* series of the World Bank (see World Bank, 2015a,b,c,d). Under the section *Resolving Insolvency*, the efficiency of the regulatory framework regarding the resolution of an insolvent company is evaluated. Thereby, a survey process is adopted and verified through a study of insolvency laws and regulations. Several assumptions about the insolvent company, the case, and the parties are made.[13] The score is inspired by the methodology in Djankov et al. (2008) and expressed as a distance to frontier with 100 representing optimality. It is calculated contingent on two equally weighted indicators, namely, *Recovery Rate* and *Strength of Insolvency Framework Index*. Thereby, the first is computed based on the reported time, costs, and outcome of the insolvency proceeding. The latter arises from several legal and regulatory conditions. According to this score, Germany reaches the third place with a score of 91.78 closely followed by the United

[10] The resolution bias belongs to the topic of sample selection. Excluding loans which are not completely resolved leads to averagely shorter TTR in the recent years and might cause distorted parameter estimates (see Section 1.4).

[11] The time period from 2006 to 2008 represents the global financial crisis. The year 2010 is shaped by the European debt crisis.

[12] A deeper insight into the deviations among the countries is presented in Section 1.3.2.

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

States with a score of 90.12 (fourth place). Amongst the considered countries, Great Britain is at the bottom of the range (82.04, 12th place). Thus, our results are confirmed by the World Bank score even though it is a purely judgmental survey process.

The variable facility distinguishes between *medium term* and *short term*. Whereas, *medium term* is defined as the reference. The *short term* category shows a significantly negative coefficient indicating that these loans tend to exhibit a shorter TTR compared to the reference category. This is noteworthy since we controlled for other more intuitive determinants, e.g., size, guarantees, and collateral. Possible explanations might be found in higher efforts with respect to resolving a *medium term* facility.

Regarding seniority, loans are divided in the categories *super senior*, *pari-passu*, and *non senior*. The category *non senior* combines the original categories *subordinated/junior* and *equity*. This pooling seems necessary owing to the quantity of both categories. The category *pari-passu* serves as the reference. The coefficients of *super senior* and *non senior* show negative signs. However, significance can only be observed for the first one. This indicates that loans of the category *non senior* do not show a significantly different TTR compared to *pari-passu*. The seniority type *super senior*, however, exhibits a shorter TTR which can be ascribed to the general nature of this category. Among the differences in the insolvency procedures of Germany, Great Britain, and the United States, a committee of creditors is involved leastwise at one point of the process.[14] Being the one and only preferred claimant might grant this creditor comprehensive rights in the considered committees which can result in a shorter TTR.

As stated in Section 1.2, our analysis is based on defaulted loan contracts according to the definition set by the Basel Committee on Banking Supervision (2006). Thereby, default occurs either if the debtor is "unlikely to pay" or "past due more than 90 days on any material credit obligation" (§452). Thus, the main categories *90 days past due* and *unlikely to pay* are integrated in the model. In the data base, five additional categories are indicated. These are graded as subcategories of the more general one *unlikely to pay*. Thereby, leeway in recording specific loan contracts is granted since it can be chosen between the general nature of default *unlikely to pay* or a more specific one (*bankruptcy*, *charge-off/provision*, *sold at material credit loss*, *distressed restructuring*, *non accrual*). We do not to summarize these categories since they might supply additional information in the model context. The default definition *90 days past due* serves as the reference category. All dummy variables show statistically significant coefficients. The signs are mostly positive, except the one of *sold at material credit loss*. This corresponds with

[14] See *Insolvenzordung* (§67, 69), Insolvency Act 1986 (4., 24., 98.), and Chapter 11 (§1102, 1126).

economic intuition: Selling an engagement should be linked to a rather short resolution process as the effort of either restructuring or liquidation is avoided. Loans defaulted due to *bankruptcy* tend to have on average the longest TTR since the involvement of public institutions – like courts – enhances the resolution process. Similar assumptions could be made regarding the category *distressed restructuring*.

The guarantee indicator shows a significantly positive sign indicating that loans provided with guarantees take on average longer to resolve. The fact that additional claims have to be established against the guarantor could lead to the observed increase. According Table 1.1, guarantees are more common in the United States (44.14%) than in Great Britain (34.05%) or Germany (24.44%).

However, collateralization seems to be a more common protection mechanism compared to guarantees. While 68.85% of all loan contracts exhibit segregation rights against one or more specific assets, only 31.88% include additional protection in form of guarantees. The collateral indicator is divided in the categories *NO*, *other collateral*, and *real estate*. Since the loans are partly secured by several assets, *real estate* is indicated if at least one of them is a property. Therefore, *other collateral* clarifies that there is no real estate among the related securities. The coefficients of the categories *other collateral* and *real estate* show a significantly negative sign implying that collateral in general yields to a shorter TTR. A reason for this could lie in the simplicity and velocity of selling an identified asset compared to winding up the company or engaging a restructuring procedure. The coefficient of the category *other collateral* is slightly lower than the one of *real estate*. Accordingly, other assets seem to have a more decreasing effect on the TTR. This is in line with the economic intuition since *other collateral* contains not only common assets, e.g., machinery, but also cash and general accounts which are easier and faster to liquidate compared to *real estate*. Investigating the impact for the number of collateral per loan, we find a significantly negative impact.

The *cured* indicator provides information regarding the outcome of the resolution process. If a loan contract is classified as *cured*, it returned to performing, i.e., the obligor is back to a sound rating. This implies that outstanding claims – in the form of principal and interest obligations – will be fulfilled nearby. In the model, *NO* serves as the reference category. The *cured* indicator shows a significantly negative sign indicating a shorter TTR of loan contracts which returned to performing. These findings are in line with the results found by Gürtler and Hibbeln (2013), whose analysis is based on a data set provided by a German bank. They detect loans returning back to performance exhibit a shorter TTR.

The industry is determined by eight sectors. In the model, *finance, insurance, real estate (RE)* serves as the reference category. The industry types *construction, agriculture, forestry, fishing*, and *manufacturing* show on average a longer TTR compared to the reference category whereas the coefficient of the latter one is not statistically significant. The industries *mining, transportation, communication, sanitary services, services*, and *wholesale and retail trade* exhibit on average a shorter TTR. This may be linked to industry specific features. Sectors such as *mining* or *transportation, communication, sanitary services* attend fixed assets, e.g., mines or communication networks, to a larger and more general extent. Selling these assets or whole divisions to one of the rather few competitors seems to be easier. In contrast, sectors such as *construction* and *manufactoring* are stamped by a rather wide range of specialization regarding the company and its assets. Although more competitors exist, to whom assets could be sold, it may be more difficult to find the appropriate counterpart. Furthermore, these sectors require a conversion of stock in trade into salable products to satisfy their current contracts. In this context, Davydenko and Franks (2008) show that a piecemeal liquidation is more likely for companies allocated in these sectors than a sale as a going concern.

To analyze the impact of the macroeconomic environment, the year-on-year (yoy) log-return of the equity index and the GDP as of the default date are included in the model.[15] Both show a significantly negative sign. This meets the economic intuition since liquidating assets or restructuring a company during favorable economic conditions seems to be more promising than during economic turmoil. Correspondingly, a stressed economic situation causes a longer TTR and exposes creditors to further interest rate and liquidation risks.

1.3.2 Country specific differences

In this section, we turn towards country specific deviations. Since the economic, legal, and regulatory environment varies among Germany, Great Britain, and the United States, differences regarding the determinants of the TTR possibly occur. This issue is motivated by Davydenko and Franks (2008) who show how the lending behavior among various countries is affected by their legal and regulatory framework. Table 1.4 gives an overview of the relevant legislations with respect to insolvency proceedings in Germany, Great Britain, and the United States. The examined legal systems show considerable differences.

[15] Several additional macroeconomic variables were tested. The equity index and the GDP supplied the best results with respect to the adjusted R-squared.

Table 1.4: Insolvency codes in Germany, Great Britain, and the United States

	Germany	Great Britain	United States
Code	*Insolvenzordnung* (IO)	Insolvency Act 1986 (IA 86) Insolvency Act 2000 (IA 00) Companies Act 2006 (CA 06)	Titel 11 (T11) (particularly Chapter 7,11)
Start	§13: Application by debtor/creditor §14: Creditor §15/15a: Duty of debtor – (Immanent) illiquidity – Over-indebtness	*CVA*: → (IA 86) 1. By comp./adm./liqu. *Administration*: → (IA 86) 9. By comp./creditor *Receivership*: → (IA 86) 32. By creditor *Liquidation*: [(IA 86) 73.] → Voluntary (IA 86) 84. ff. → By court (IA 86) 122. ff. ⇒ By comp./creditor	§301: Voluntary cases → By debtor §303: Involuntary cases → By creditor
Aim	§1: Satisfying creditor claims (1) Utilization/distribution (2) *Insolvenzplan*, alternatively → §217 Restructuring, going concern	*CVA*: (IA 86) 1. Supporting *Administration*: (IA 86) 8. Survival *Receivership*: Management, Liquidation *Liquidation*: (IA 86) 73. Winding up	*Chapter 11*: → Reorganization (§1101 ff.) *Chapter 7*: → Liquidation (§701 ff.)
Moratorium	§21(3): Moratorium → Not immovables → Not specific fin. assets	*Administration*: (IA 86) 10. and 11.	§362: Automatic stay → Moratorium included
Stay	§22: Temporary stay → Insolvency administrator §259: Temporary stay → Debtor (application)	No	§362: Automatic stay
Management	§27: Insolvency administrator → §67/69: Supervision by creditors → Creditor-in-possession *Insolvenzplan*: → Debtor-in-possession (§259)	*Administration* → (IA 86) 8. Adm. *Receivership* → (IA 86) 32. Receiver *Liquidation* → (IA 86) 91./100./135. Liqu.	*Chapter 11* → Debtor-in-possession (§1103/1107) *Chapter 7* → §702: Trustee (elected by creditors)
Financing	*Insolvenzplan* → §264: Super-priority-financing (limited)	No	§503/507: Super-priority-financing
Common	Liquidation → *Insolvenzplan* rarely used	*Administration* → Sale as going concern	*Chapter 11* → Reorganization
WB Code (2015)	*Overall*: 79.64 (rank 15) → *Resolving Insolvency*: 91.78 (rank 3) → *Getting Credit*: 70.00 (rank 24)	*Overall*: 82.18 (rank 6) → *Resolving Insolvency*: 82.04 (rank 12) → *Getting Credit*: 75.00 (rank 17)	*Overall*: 82.15 (rank 7) → *Resolving Insolvency*: 90.12 (rank 4) → *Getting Credit*: 95.00 (rank 2)

Notes: Common outcomes of the insolvency proceedings originate form Bonelli et al. (2014).

The corresponding code in Germany is the *Insolvenzordnung* (IO). The insolvency procedure starts with an application by either the debtor or one of the creditors (IO §13). A special feature of the German insolvency law is the explicit definition of the default event. According to §15 (IO) and §15a (IO), a debtor has the duty to file for insolvency if a company has to deal with (immanent) illiquitity or over-indebtness. Neither the British nor the American legal system includes such an exact definition or the debtor duty. In Germany, the main aim of the insolvency process is to satisfy the claims of the creditors (IO §1). The survival of the company by means of an *Insolvenzplan* is subordinated.[16] A moratorium exists during the insolvency proceedings whereas immovables and specific financial assets are excluded (IO §21). Furthermore, it is possible for corporations to remain in a temporary stay (IO §22). However, the management is held by the (temporary) insolvency administrator and a stay is only adopted by the application of the debtor (IO §259). During the insolvency proceeding, the corporation is managed by the insolvency administrator who is supervised by the creditors (IO §67 and §69). Such a process is called *creditor-in-possession*. The construct of a *debtor-in-possession* generally exists under

[16] An *Insolvenzplan* is a program to restructure the company.

the assumption of an *Insolvenzplan* (IO §259). Super-priority-financing is limited to §264 (IO) during restructuring. According to Bonelli et al. (2014) liquidation is the most common outcome of the insolvency proceedings in Germany. Furthermore, the concept of the *Insolvenzplan* is relatively new and rarely used in the legal practice.

The relevant codes in Great Britain involve the Insolvency Act 1986 (IA 86), the Insolvency Act 2000 (IA 00) and the Companies Act 2006 (CA 06). In general, there are four independent solution mechanism with respect to insolvency. Most of them can be employed separately or in combination. The first one is described by the *Company Voluntary Arrangement* (CVA).[17] This voluntary agreement can be applied by the debtor, the creditor, or the administrator if the company is situated under *Administration* (IA 86, 1.). A CVA only has supporting nature and does not have specific aims regarding the outcome of the insolvency proceeding. The concept of *Administration* describes the second insolvency mechanism which starts by the application of either the debtor or the creditor (IA 86, 9.). The main aim of this instrument is the survival of the company (IA 86, 8.). The remaining two mechanisms are the *Receivership* and the *Liquidation*, respectively. They are indicated either by a creditor (*Receivership*: IA 86, 32.) or by one of the affected parties (*Liquidation*: IA 86, 73.). The aim of both procedures is the satisfaction of the creditors' claims. While *Receivership* focuses on the management of a company during insolvency and the liquidation as a going concern, the *Liquidation* procedure is directing to wind up the company (IA 86, 73.). The possibility of a moratorium is only given under *Administration* (IA 86, 10. and 11.). However, there exists no regulation regarding a stay of a company after filing for insolvency. The most common outcome of the insolvency proceeding in Great Britain is a sale of the insolvent company as a going concern under the *Administration* regulations (see Bonelli et al., 2014).

The corresponding laws in the United States are recorded under Title 11 of the US Code. With respect to the resolution of a defaulted loan contract, particularly, Chapter 7 (*Liquidation*) and Chapter 11 (*Reorganization*) are of interest. An insolvency proceeding starts by filing for insolvency under a specific chapter. It is distinguished between voluntary cases (T11 §301), initiated by the debtor, and involuntary cases (T11 §303) where one of the creditors files for insolvency. In general, involuntary cases are considerably outnumbered by voluntary ones. Under Chapter 7, the main aim is the liquidation of the company (T11 §701 ff.). Therefore, a trustee is appointed by the creditors. In contrast, Chapter 11 focuses the reorganization of the corporation (T11 §1101 ff.) where the debtor stays in possession (*debtor-in-possession*: T11 §1103

[17] Alternatively, a *Scheme of Arrangement* (CA 06, 834.) could by applied. However, a company might not be actually insolvent to adopt such a scheme.

and §1107). Generally, there is an unlimited automatic stay during the insolvency proceeding (T11 §362) and super-priority-financing is possible (T11 §503 and §507). The most common outcome in the United States is the reorganization of the corporation under Chapter 11 (see Bonelli et al., 2014).

With the legal and regulatory framework in each country in mind, we turn towards the results on country level. Therefore, the data base is divided in three country subsets to investigate whether the coefficients of the variables are subject to changes. The results are presented in Table 1.5.[18] We find the impact of the EAD, facility type, guarantees, number of collateral, cured indicator, and macroeconomic variables to be similar compared to the overall model. While the TTR tends to increase for higher values of the EAD and the availability of a guarantee, it seems to decrease along with the remaining.[19] With regards to the time dummies, seniority, nature of default, collateral, and the industry type, considerable differences arise. These affect the magnitude, the statistic significance and even the sign of the coefficients.

Figure 1.3 displays the progression of the time coefficients regarding the overall data set and on country level. A general time trend is visible. However, the coefficients do not show a monotonous decreasing course. In the year 2008, there is a considerable peak with respect to the whole data base. This might be driven by the Anglo American subsets as the corresponding coefficient for Germany does not increase. Although significance cannot be observed in the United States regarding this particular year, these results indicate a stronger effect of the global financial crisis on the TTR in the Anglo American area. Analogously, the second peak in the overall data base in the year 2010 might be driven by the German subset. Whereas, the Anglo American countries do not show an increase, the coefficients for Germany in the years 2010 and 2011 are significant and higher compared to previous years.[20]

Besides this general effect, the time coefficients may also partly capture the resolution bias.[21] This phenomenon belongs to the issue of sample selection and describes the fact that the data in the current time periods has not been completed yet. For example, only a proportion of the loans which defaulted in the year 2013 has been fully resolved. Loans still being in the resolution process are characterized by a rather long TTR. Excluding such loans may lead to

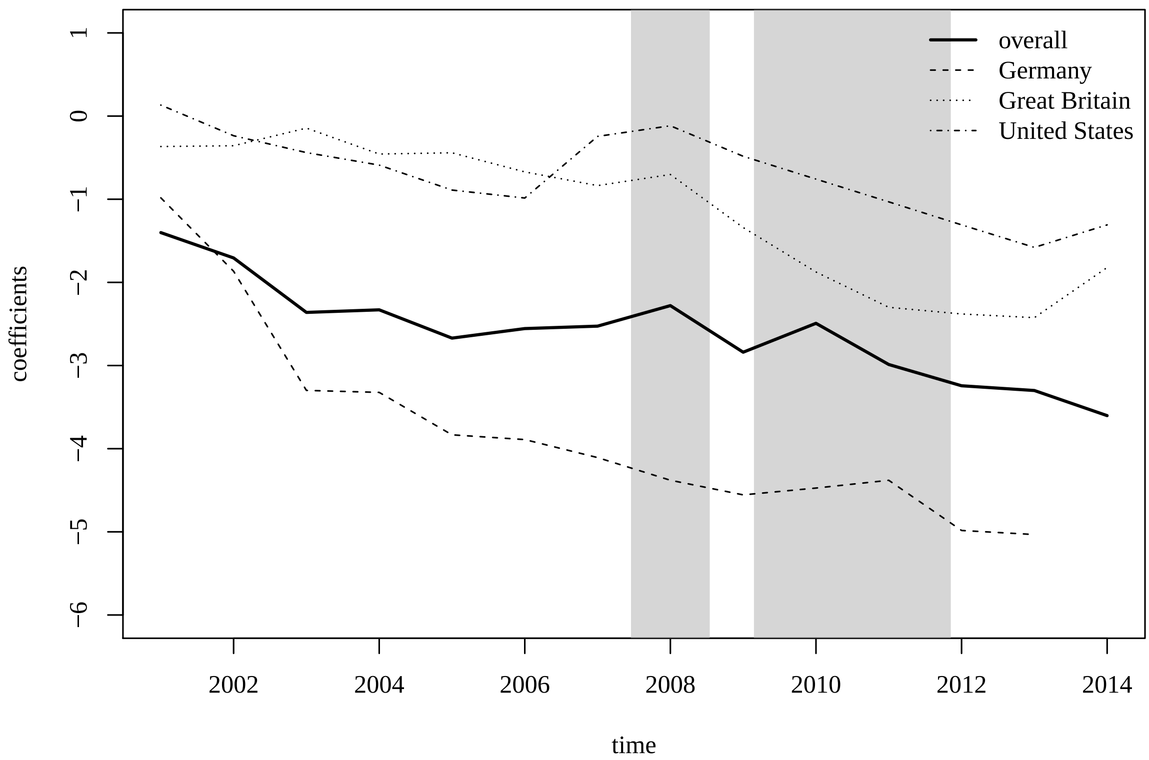
[18] Note that not all categories appear in the country models. This is due to their non-existence in the subsets. There are, e.g., no loan contracts in Germany which defaulted in the year 2014. Apart from that, the representation meets the one of Table 1.3.
[19] See Section 1.3.1.
[20] These years are marked by the European debt crisis.
[21] The inclusion of the finalization rate per cohort year as an additional variable does not affect the general trend of the time dummies.

Table 1.5: Regression results of the TTR on country level

		Germany			Great Britain			United States		
		Coef.		SE	Coef.		SE	Coef.		SE
Intercept		4.707	***	(0.0906)	3.574	***	(0.1768)	0.887	***	(0.2100)
log(**EAD**)		0.030	***	(0.0052)	0.036	***	(0.0080)	0.076	***	(0.0097)
Time	2001	-0.983	***	(0.0814)	-0.366	**	(0.1161)	0.132		(0.1536)
(2000)	2002	-1.866	***	(0.0683)	-0.357	**	(0.1173)	-0.236	.	(0.1379)
	2003	-3.299	***	(0.0630)	-0.145		(0.1381)	-0.439	***	(0.1306)
	2004	-3.323	***	(0.0664)	-0.457	***	(0.1269)	-0.590	***	(0.1339)
	2005	-3.833	***	(0.0616)	-0.441	***	(0.1216)	-0.890	***	(0.1472)
	2006	-3.890	***	(0.0674)	-0.671	***	(0.1092)	-0.985	***	(0.1295)
	2007	-4.107	***	(0.0735)	-0.836	***	(0.1117)	-0.244		(0.1519)
	2008	-4.380	***	(0.0749)	-0.703	***	(0.1208)	-0.117		(0.1554)
	2009	-4.556	***	(0.1805)	-1.340	***	(0.1576)	-0.483	*	(0.1916)
	2010	-4.474	***	(0.1373)	-1.875	***	(0.1121)	-0.758	***	(0.1323)
	2011	-4.380	***	(0.1900)	-2.300	***	(0.1087)	-1.032	***	(0.1325)
	2012	-4.984	***	(0.1991)	-2.380	***	(0.1160)	-1.308	***	(0.1286)
	2013	-5.033	***	(0.3256)	-2.424	***	(0.1155)	-1.579	***	(0.1392)
	2014				-1.821	***	(0.1988)	-1.308	***	(0.1801)
Facility	Short term	-0.050	*	(0.0223)	-0.082	*	(0.0319)	-0.093	.	(0.0505)
(Medium term)	Other	-0.108	***	(0.0306)	-0.760	***	(0.0781)	-0.115	**	(0.0406)
Seniority	Super senior	0.049		(0.0453)	-1.153	***	(0.0493)	-0.017		(0.0635)
(Pari-passu)	Non senior	0.197	***	(0.0569)	-1.246	***	(0.3597)	1.354	***	(0.1646)
	Unknown	-1.576	***	(0.0590)				0.026		(0.1561)
Nature of default	Sold at material credit loss	0.312		(0.1958)	0.239		(0.6464)	-1.184	***	(0.1326)
	Unlikely to pay	0.755	***	(0.0392)	-0.348	***	(0.0398)	-0.146	**	(0.0459)
	Charge-off / provision	0.665	***	(0.0218)	-0.048		(0.0464)	0.691	***	(0.1464)
	Non accrual	-0.030		(0.0772)	0.444	***	(0.0467)	-0.157	***	(0.0473)
	Distressed restructuring	0.756	***	(0.0340)	0.729	***	(0.0981)	0.265		(0.1922)
	Bankruptcy	1.223	***	(0.0399)	-0.214	***	(0.0584)	-0.057		(0.1125)
	Unknown	0.705	**	(0.2406)	-0.374	***	(0.0431)	-0.168		(0.1894)
Guarantee	YES	0.008		(0.0223)	0.059	.	(0.0312)	0.227	***	(0.0360)
(NO)	Unknown	-0.125		(1.2012)	0.253		(0.7136)	-1.325	***	(0.3996)
Collateral	Other collateral	-0.189	***	(0.0283)	0.085	*	(0.0402)	0.111	*	(0.0469)
(NO)	Real estate	-0.144	***	(0.0263)	-0.063	.	(0.0349)	0.074		(0.0593)
	Unknown							0.531		(0.9635)
Number of collateral		-0.003		(0.0036)	-0.005	**	(0.0017)	-0.061	***	(0.0117)
Cured (NO)	YES	-0.467	***	(0.0193)	-0.697	***	(0.0355)	-0.259	***	(0.0412)
Industry	Mining	-1.252	***	(0.2641)	-0.176		(0.2313)	-0.040		(0.1801)
(Finance, insurance, RE)	Transp., commu., san. ser.	-0.191	***	(0.0476)	-0.435	***	(0.0860)	0.187	*	(0.0875)
	Services	-0.095	***	(0.0267)	-0.069		(0.0444)	0.220	***	(0.0559)
	Wholesale and retail trade	-0.063	*	(0.0311)	-0.064	.	(0.0391)	0.150	**	(0.0582)
	Manufacturing	-0.010		(0.0345)	-0.153	**	(0.0525)	0.329	***	(0.0588)
	Agric., forestry, fishing	0.202	**	(0.0764)	0.010		(0.0833)	-0.049		(0.1410)
	Construction	0.156	***	(0.0425)	0.058		(0.0426)	0.449	***	(0.0678)
	Unknown	-0.217	***	(0.0536)	0.140		(0.1292)	-0.184	**	(0.0591)
Equity Index		-0.318	***	(0.0750)	-0.115		(0.1429)	-0.253	*	(0.1205)
GDP		-8.329	***	(1.2542)	-4.671	**	(1.5663)	-0.049		(1.8697)
Adjusted R-squared				73.49%			38.38%			22.90%
F-statistic				761.54			117.94			37.39
p-value				0.0000			0.0000			0.0000

Notes: Results of the multiple linear regression regarding the country subsets. Significance codes: *** 0.001, ** 0.01, * 0.05, · 0.1. Standard errors (SE) are clustered by year.

Figure 1.3: Progression for the time coefficients in the regression models for the overall data set and on country level



a systematic distortion in the parameters. By including time dummies, these effects can be absorbed. This could be a reason for the decreasing nature of the coefficients, particularly in the current years.[22]

Regarding the seniority type, considerable differences can be observed among Germany, Great Britain, and the United States. While loans of the category *super senior* show a significant shorter TTR compared to the reference category in the overall data set, this is only true with respect to Great Britain. In general, the *super senior* status gives a single preferred claimant creditor wide powers in the resolution proceeding which could lead to a shorter TTR. The insignificance in Germany and the United States seems to be related to the authorization of super-priority-financing during the insolvency process. This involves a dilution of *super senior* claims by additional new debt which weakens the power base of the initially preferred creditors. Still, being among preferred claimants seems to be vital in both countries as the non senior status increases the TTR in both countries.

Overall, the TTR does not differ between the categories *non senior* and *pari-passu*. On the contrary, *non senior* shows a significantly longer TTR in Germany and the United States, and a shorter one in Great Britain. A longer resolution process regarding subordinated claims seems

[22] Concerning the resolution bias, see Section 1.4.

to be intuitive since these creditors are the last ones to be satisfied out of the insolvency assets. However, the category *pari-passu* demands on average the longest TTR in Great Britain which might be ascribed to a more complex approval structure that accompanies with claims of equal rank.

The collateral types *other collateral* and *real estate* exhibit significantly negative signs in the overall data set. In Germany, the coefficients show the same direction and a similar scale. This corresponds with the legal framework. Although the possibility of a temporary stay exists according to the regulations, it is only adaptable during an *Insolvenzplan* and associated with an application by the debtor. Furthermore, restructuring is a rather rare outcome of the insolvency proceeding. In the standard case of liquidation, the secured creditors have direct access to the securities. A switch of the sign can be observed in Great Britain regarding *other collateral*. While there is no stay regulation in the British insolvency law, a moratorium takes place during the standard proceeding – the *Administration*. With the main aim of selling the company as going concern, it might be more difficult to gain access on some of the assets. In the United States, the coefficient of *other collateral* shows a significantly positive sign whereas the one of *real estate* is insignificant. Considering the insolvency law with its automatic stay regulation, some time may elapse until a creditor could liquidate securities. This seems to explain the insignificance with respect to *real estate* and the significantly positive impact of *other collateral*.

Regarding seniority, guarantees, and collateral, differences among the countries already arise in the descriptive statistics. In Germany, where we observe a positive impact of collateral in general, 72.33% of all loans are additionally secured by collateral. This proportion is lower in Great Britain (67.23%) and the United States (63.98%). Besides collateralization, guarantees and seniority can be classified as further security mechanisms. Compared to Germany (24.44%), a rather high proportion of guarantees can be found in the United States (44.14%) and Great Britain (34.05%). Analogously, a considerably higher seniority level can be observed in these countries.[23] This indicates that creditors try to compensate the harder and longer liquidation proceeding by alternative security mechanisms. However, rather unfavorable or negligible effects of guarantees and seniority can be observed regarding the TTR. The coefficients of the guarantee indicator are significantly positive in the Anglo American subsets and the one of *super senior* is insignificant in the United States. While positive effects are limited with respect

[23] Proportion of *super senior* among the considered countries: Germany (5.47%), Great Britain (44.34%), and the United States (83.71%).

to the TTR, they arise considering the LGD as the dependent variable (see Section 1.3.3).[24]

With respect to the nature of default, differences among the considered countries can be observed. In the overall data set, all coefficients – except the one of *sold at material credit loss* – are significantly positive. In Germany, the results tend to be consistent. Contrary to the overall data set, the coefficients of *unlikely to pay* and *bankruptcy* show a significantly negative sign in Great Britain.[25] Generally, the separate interpretation of *unlikely to pay* is rather ambiguous since it is constructed as a main category for the remaining default types – except *90 days past due*. Since relatively more loans are characterized by this category, there it seems possible that systematic differences compared to Germany occur. The term "bankruptcy" only applies to individuals (including sole proprietors) in Great Britain. It seems intuitive that the resolution of individual entrepreneurs is less entangled and, therefore, shorter. In the United States, sign switches can be observed for *unlikely to pay* and *non accrual*.[26] The separate interpretation of *unlikely to pay* remains ambiguous. The latter may also be related to general business practices. Since restructuring is the main outcome of insolvency proceedings, waivers on accruals seem to be rather common during financial distress to quickly reduce the pressure on the debtor.

While the industry dummies regarding Germany and Great Britain are rather similar to those of the overall data set, differences can be observed in the United States. The impacts of the categories *transportation*, *communication*, *sanitary services*, *services*, and *wholesale and retail trade* exhibit a significantly positive impact. These results reflect the divergence of import industrial sectors among countries.

Regarding Table 1.3 and 1.5, the impact of the cured indicator is similar among countries. However, discrepancies in terms of descriptive statistics in Table 1.1 are remarkable. Overall, 36.40% of all loans are *cured*. The highest proportion can be observed in Great Britain (42.44%), followed by Germany (35.74%) and the United States (29.12%). At first glance, this seems to be counterintuitive since restructuring – or at least going concern – is the main aim of the insolvency law in the United States. However, one should not mix up the concept of *cured* loan contracts with the survival of a corporation. Its relatively high proportion may be an indication that rather proficient debtors are pushed into insolvency. In this context, Germany with its

[24] For example, while the guarantee indicator shows a coefficient of −0.077 (***) in Great Britain with the transformed LGD as the dependent variable, the one of *super senior* is with −0.052 (**) also significantly negative in the United States (LGDs are transformed by the inverse of normal distribution function, see Section 1.3.3). Results are available from the authors upon request.

[25] The significance of *sold at material credit loss* and *charge-off / provision* disappears. This might be due to a inferior importance of these categories in Great Britain.

[26] The significance of *distressed restructuring* and *bankruptcy* disappears. This might be due to a inferior importance of these categories in the United States.

explicitly formalized default definition might serve as a reference. In the United States, rather less debtors may enter insolvency since the voluntary case is the norm and there is no duty to file for insolvency. Contrary, the application might be done more often by the creditors in Great Britain. They have minor insight in the corporations and may be rather inclined to file for insolvency.

1.3.3 Comparison to general drivers of LGD

In addition to the TTR as a proxy for the indirect costs, we analyze drivers for the LGD. Noteworthy, the loss rate only captures parts of the indirect cost as opportunity costs and reputational losses might not fully be included in the LGD. Significant differences regarding the determinants of both cost categories may arise. A consideration of the TTR and the LGD is thus essential to correctly evaluate the effects of specific loan characteristics, e.g., collateralization or guarantees. Otherwise, positive impacts might be overvalued and negative ones underestimated.

Figure 1.4: Boxplots of LGDs on country level

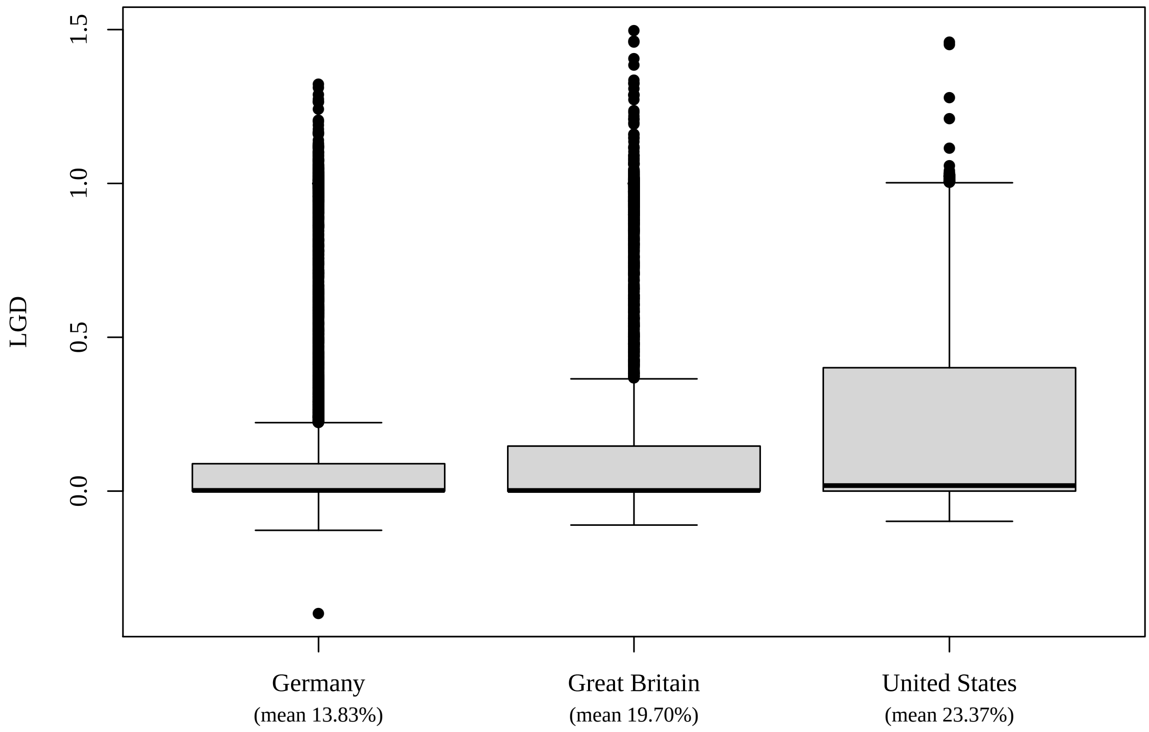


Figure 1.4 displays boxplots of the LGD divided by country[27] while means are given below country names. The average LGD of German loans is 13.83% and, thus, lower than in Great Britain (19.70%) and the United States (23.37%). Analogously to the TTR, deviations might arise

[27] As we are dealing with economic LGDs, i.e., *workout* LGDs, values lower than 0 and higher than 1 arise.

Table 1.6: Determination of ϵ

ϵ	Adjusted R-sqared
1.00E-50	27.4208%
1.00E-45	27.4208%
1.00E-40	27.4208%
1.00E-35	27.4208%
1.00E-30	27.4208%
1.00E-25	27.4208%
1.00E-20	27.4208%
1.00E-15	27.4208%
1.00E-10	27.4208%
1.00E-05	27.4202%
1.00E-04	27.4140%
0.001	27.3521%
0.01	26.7287%
0.1	20.1767%

Notes: ϵ of the Inverse Gaussian Regression with the corresponding adjusted R-squared. Resulting ϵ is written in bold.

when comparing descriptive statistics for LGDs and results from a modeling framework.[28] Thus, we apply a multiple linear regression model for the LGD using similar regressors as in the case of the TTR model. LGD values are caped to the interval [0,1]. To address common characteristics of the distribution, the LGDs are transformed by the inverse of the normal distribution function. Since the inverse is not defined at the points 0 and 1, a small positive value ϵ is determined which is added to 0 and subtracted from 1. Regarding the choice of ϵ, a trade off arises. A very small ϵ leads to values close to $-\infty$ (LGD = 0) and ∞ (LGD = 1) on the transformed scale and, thus, probably to a poor model fit. Higher values of ϵ imply larger deviations from the actual LGDs which could come along with a poor model fit. Accordingly, sensitivity has to be examined in the final choice of ϵ (see, e.g., Qi and Zhao, 2011; Hu and Perraudin, 2006). Table 1.6 contains various possible values of ϵ and the adjusted R-squared of the corresponding transformed regression. We elect ϵ = 1.00E-11 since the adjusted R-squared remains constant as ϵ is further decreased.

The results of the transformed multiple linear regression are displayed in Table 1.7. In general, analogy in the sign and significance of the coefficients with respect to the TTR (see Table 1.3) emerge. Large parts of the time dummies also follow a decreasing course. However, some of the coefficients are subject to sign switches. These arise particularly in years of financial turmoil indicating higher LGDs during harsh economic times.[29] Furthermore, we observe a different order of the country dummies. Germany shows the lowest TTR followed by Great Britain and

[28] According descriptive statistics in Table 1.1 the shortest TTR on average is observed in the United States. In the model context, loans located in Germany appear to exhibit the shortest resolution process (see Table 1.3).

[29] The sign switch regarding the year 2014 might be due to randomness as this year is rarely represented in the sample.

Table 1.7: Regression results of the LGD for the overall data set

		Coef.		SE
Intercept		0.446	***	(0.0162)
log(**EAD**)		-0.008	***	(0.0011)
Time	2001	0.005		(0.0116)
(2000)	2002	-0.020	.	(0.0108)
	2003	-0.033	**	(0.0108)
	2004	-0.065	***	(0.0112)
	2005	-0.025	**	(0.0098)
	2006	-0.027	**	(0.0097)
	2007	-0.065	***	(0.0097)
	2008	0.036	**	(0.0126)
	2009	-0.002		(0.0157)
	2010	0.016		(0.0116)
	2011	0.008		(0.0117)
	2012	-0.017		(0.0119)
	2013	-0.009		(0.0143)
	2014	0.407	***	(0.0666)
Country	United States	0.114	***	(0.0098)
(Germany)	Great Britain	0.074	***	(0.0077)
Facility	Short term	-0.007		(0.0045)
(Medium term)	Other	-0.135	***	(0.0057)
Seniority	Super senior	-0.026	***	(0.0060)
(Pari-passu)	Non senior	0.142	***	(0.0151)
	Unknown	-0.045	***	(0.0123)
Nature of default	Sold at material credit loss	0.125	***	(0.0259)
	Unlikely to pay	-0.042	***	(0.0063)
	Charge-off / provision	0.050	***	(0.0054)
	Non accrual	0.025	***	(0.0069)
	Distressed restructuring	0.048	***	(0.0090)
	Bankruptcy	-0.014	.	(0.0074)
	Unknown	-0.041	**	(0.0129)
Guarantee	YES	-0.049	***	(0.0041)
(NO)	Unknown	-0.172	*	(0.0709)
Collateral	Other collateral	-0.060	***	(0.0050)
(NO)	Real estate	-0.137	***	(0.0051)
	Unknown	-0.347	*	(0.1687)
Number of collateral		0.001		(0.0004)
Cured (NO)	YES	-0.280	***	(0.0043)
Industry	Mining	-0.112	**	(0.0341)
(Finance, insurance, RE)	Transp., commu., sanitary services	0.011		(0.0107)
	Services	0.053	***	(0.0061)
	Wholesale and retail trade	0.035	***	(0.0065)
	Manufacturing	0.013	.	(0.0070)
	Agric., forestry, fishing	-0.012		(0.0162)
	Construction	-0.005		(0.0078)
	Unknown	-0.018	*	(0.0084)
Equity Index		0.059	***	(0.0141)
GDP		-0.609	***	(0.1678)
Adjusted R-squared				27.42%
F-statistic				204.41
p-value				0.0000

Notes: Results of the multiple linear regression regarding the overall data set with the LGD as dependent variable. Significance codes: *** 0.001, ** 0.01, * 0.05, · 0.1. Standard errors (SE) are clustered by year.

the United States. Although the resolution is faster in the United States, it is accompanied with higher losses compared to Great Britain. This divergence could be ascribed to the insolvency proceeding. In the United States, Chapter 11 characterized by its high borrower friendliness might significantly reduce the TTR compared to the complex approval structures in the British insolvency law. However, effects on the loss seem to be contrary. Results with respect to facility types are similar, whereas, the significance of *short term* disappears. In addition, no significant deviation occurs regarding collateral, the cured indicator and macroeconomic variables.[30] The significance of number of collateral vanishes.

Major breaches are detected in the remaining variables. The coefficient of *non senior* is significantly negative indicating that seniority obtains a stronger effect on the LGD compared to the TTR. Furthermore, statistically significant sign switches arise regarding the natures of default *sold at material credit loss*, *unlikely to pay* and *bankruptcy*. The first category shows a positive, the latter ones a negative, sign. Regarding *sold at material credit loss*, this meets the economic intuition. Although, restructuring or the winding up of a company can result in a longer TTR, waivers may be linked to the selling of an engagement. This would raise LGDs. In spite of increasing the TTR, guarantees reduce the loss rate. This seems consistent as the assertion of additional claims might enhance the resolution process. However, further payments may be generated. Regarding industry, significant sign switches can be observed for the categories *services* and *wholesale and retail trade*. Although reducing the TTR, they lead to an increase in the LGD.

Generally, the negligence of the TTR as a considerable cost component might result in inaccurate conclusions. For example, we find that guarantees seem to engage opposite effects. Solely focusing on the LGD leads to an overestimation regarding benefits of guarantees as they reduce the loss rate. However, they also result in a longer resolution process and additional risk sources involved – such as opportunity cost and reputational losses or interest and liquidation risks – may be obliterated.

[30] Although, the GDP shows the expected sign, the coefficient of the equity index is significantly negative. As both macroeconomic variables are implemented as of the default date, reasons for this discrepancy might be found in the cyclical nature of the variables. While the GDP captures the main effect of the macroeconomic environment, the remaining variable might reflect the economic framework during the resolution process which may be inverse compared to the default date. Furthermore, this could catch an additional influence structure which can not be measured by a single variable.

1.4 Robustness

In this section, we examine the robustness of results presented in Section 1.3. We investigate deviations from previous findings when applying varying methods. First, we use a log transformation of the TTR. In contrast to the level specification, the TTR is restricted to values greater than zero. As stated in Section 1.3, the resolution bias is an issue of concern. Estimates of the multiple linear regression might be biased due to the overrepresentation of loans with a rather short TTR. If the exact number of missing observations, i.e., loans not completely resolved yet, is known, censored regressions are adaptable to adress the resolution bias. However, as we are not aware of the exact quantity, we apply truncated regressions. The results are compared with the linear model in order to draw conclusions regarding possible distortions in the parameters. Finally, we return to the multiple linear regression but use an alternative measure of the TTR as the dependent variable.

Log transformation

The results regarding the log transformation are displayed in Table 1.A.1 (see Appendix 1.A). Compared to the level specification, we find similar signs for almost all determinants. Differences arise for the seniority *non senior*, the collateral *real estate* and the industries *services* and *wholesale and retail trade*. However, these might be attributable to the modification of the dependent variable and, particularly, to the variation of its distribution.[31] Generally, the results seem to be robust regarding the transformation of the TTR as the directions of the coefficients are similar for most of the variables. Changes of the impacts can be explained through the log transformation itself.

Truncated regression

Excluding incompletely resolved loan contracts entails a selection bias. At the time of writing of this paper, the year 2015 marks the present and we do not have any information beyond this point. Loans defaulted in, e.g., 2013 could exhibit a maximum TTR of two years while those which are going to be resolved later on are neglected. Hence, including only resolved loan contracts might yield in distorted parameter estimates. To analyze this bias, we adopt truncated regressions on subsamples on yearly basis. The application on subsets seems necessary as the limiting value of the data has to be unique in the samples. E.g., the limit regarding loans defaulted in 2014 is one year, in 2013 two years, in 2012 three years, and so on. For the reason

[31] See Appendix 1.A for further information.

of comparability, truncated regressions on yearly basis are compared to their counterparts of the linear regression.[32]

Generally, the coefficients in both models are almost similar regarding all years. For simplicity, only the subsets with the highest deviations in the parameter estimates are displayed.[33] Table 1.B.1 and 1.B.2 (see Appendix 1.B) contain the results with respect to the years 2008 and 2009, where the coefficients of the year 2008 exhibit slightly higher deviations. Noteworthy, these years mark the summit of the global financial crisis. This indicates that the resolution bias is particularly pronounced during financial turmoil. At first glance, this may be counterintuitive since we would expect the major manifestation in the most recent years. However, the global financial crisis might have entailed a market shakeout. Rather poor debtors have defaulted in crisis years yielding to a better credit quality regarding debtors afterwards. Therefore, short resolution times are more frequent in the recent years and distortions due to longer TTR become less likely.

Comparing the truncated and the linear regression, almost no variations emerge regarding the sign or significance of the coefficients. This gives rise to the conjecture that the resolution bias might not lead to misjudgments regarding the direction of the determinants. In the truncated regression, the absolute values of the coefficients increase compared to the multiple linear model. Therefore, neglecting the resolution bias can lead to an undervaluation of the impacts but not to a misapprehension of the signs. Generally, the deviations are negligible indicating that either the resolution bias does not distort the parameter estimates or its effect is absorbed by the time dummies in the overall data set and the intercepts in the subsamples.

Alternative dependent variable

Besides the application of various models, the TTR is replaced by a different dependent variable inspired by the bond duration. Generally, it is specified by the cash-flow-weighted average of the payment dates and, therefore, expressed in years.[34] We will refer to this measure as loan duration:

$$D = \frac{1}{P}\left[\sum_{t=1}^{T} t\frac{C_t}{(1+r)^t}\right],$$

whereby P denotes the present value of the cash flows and is calculated as $P = \sum_{t=1}^{T}\frac{C_t}{(1+r)^t}$. The parameter C_t describes the cash flow at time t and r the discount rate. Generally, only positive

[32] Comparing regressions of the overall data set with the ones on yearly basis, several deviations among the parameters arise. See Appendix 1.B for further information.
[33] The results of the remaining subsets are available from the authors upon request.
[34] The implied statement of price sensitivity to the interest rate is neglected.

and cash-flow-related transactions enter the loan duration.

Table 1.C.1 (see Appendix 1.C) displays the results of the multiple linear regression with the loan duration as the dependent variable. Compared to the results of Table 1.3, several deviations among the coefficients arise. However, reverse signs are insignificant regarding the TTR or the loan duration.[35]

1.5 Conclusion

This paper analyzes the TTR of defaulted loan contracts. By this means, we aim to determine general drivers of the resolution time. Particularly, country specific deviations with respect to the determinants are examined.

Our results show that the TTR is determined by loan specific characteristics as well as the economic environment. While bad macroeconomic conditions extend the resolution process in general, we find that medium term loans with high exposures exhibit the longest TTR. Equally ranked claimants additionally decelerate the resolution process. Overall, collateralization abbreviates the procedure, whereas guarantees seem to entail opposite effects. Furthermore, we find Great Britain and the United States to reveal a higher TTR after controlling for additional input parameters. American loans take on average 0.1 years longer to resolve compared to loans located in Germany, British loans even 0.5 years longer. These findings are supported by a World Bank score. Regarding the TTR, the insolvency code in Germany is most efficient amongst the considered countries.

Based on these general findings, we deepen our analysis on country level. We find deviations in the impact of seniority, nature of default, collateral, and industry among Germany, Great Britain, and the United States. Several of these differences, e.g., regarding seniority and collateral, seem to be caused by dissimilar regulations. The effect of collateralization is an important driver regarding country specific differences. While both *real estate* and *other collateral* shorten the resolution process in Germany by averagely 0.1 and 0.2 years, *real estate* is insignificant in the United States and *other collateral* even exhibits an opposite impact in the Anglo American countries, thus, extending the TTR on average by 0.1 years. This may relate to the creditor friendly nature of the German insolvency law. While easy access to collateral is provided in Germany, moratorium (Great Britain) and automatic stay regulations (United States) complicate

[35] See Appendix 1.C for further information.

the liquidation of individual securities in the Anglo American area. Additionally, general business practices seems to be influenced by the legal frameworks. Such regulatory differences seem to be known to creditors which leads to country specific lending adjustments, e.g., a more frequent use of additional security mechanisms in Anglo American countries. Thus, higher fractions of loan contracts secured by guarantees and / or preferred seniority arise in Great Britain and the United States.

Our findings contribute to a better understanding of the TTR. Particularly, attributes are determined leading to faster resolutions of defaulted loan contracts. Risk accompanied with long resolution times might be avoided. Moreover, we find that considerable differences in the legal frameworks lead to adjustments in the general lending behavior of creditors which impacts the TTR of loans.

1.A Appendix | Log transformation

Table 1.A.1 displays the results of the multiple linear regression with the natural logarithm of the TTR as dependent variable. We find similar signs for almost all determinants compared to the level specification (see Table 1.3). However, differences can be examined regarding the seniority *non senior*, the collateral *real estate*, and the industries *services* and *wholesale and retail trade*.

In the level specification, *non senior* exhibits an insignificantly negative sign whereas a significantly positive impact can be observed in the log specification. The latter indicates a longer TTR of *non senior* compared to the reference category. This might be due to the modification of the dependent variable. Figure 1.A.1 presents the boxplots for the TTR and its log transformation divided by seniority.[36] Thereby, the altering nature of the natural logarithm is demonstrated. While high values are reduced, negative ones appear wherever the level variable is below one. This might be beneficial considering determinants discriminating amongst rather small values. The category *non senior* might be of such kind. Focusing on the mean values, a positive impact as revealed in the log specification might be expected since an averagely longer TTR is indicated. Compared to the level variable, the difference in the means seems rather large in the log specification. Therefore, the sign switch and the significance might be reasonable as the linear regression minimizes the squared differences with respect to the mean.

A considerable variation can be observed regarding the collateral *real estate*. While the sign is significantly negative in the level specification, the impact becomes significantly positive applying the log transformation. Figure 1.A.2 displays the boxplots of the level variable and its natural logarithm divided by collateral.[37] Again, the log transformation reduces high values and creates negative ones. Considering the means, a negative impact of *real estate* as revealed in the level specification might be expected. In contrast to seniority, the difference in the means is more pronounced regarding the level specification.

Regarding the industry, deviating signs can be observed in the categories *services* and *wholesale and retail trade*. While both exhibit a significantly negative sign in the level specification, the impacts are positive applying the log transformation. However, the coefficient of *wholesale and retail trade* is not statistically significant. Figure 1.A.3 displays the boxplots divided by the

[36] See Figure 1.A.1, p. 42.
[37] See Figure 1.A.2, p. 42.

considered industries.[38] Regarding the mean values, a negative sign of *services* and a positive (or insignificant) one of *wholesale and retail trade* might be expected in the level specification. Considering the log transformation, a positive impact of both categories might be indicated. While the expectation of *services* can be confirmed, we observe a significantly negative impact of *wholesale and retail trade* in the level specification and an insignificant one applying the natural logarithm. However, the coefficient in the level specification is quite small and only significant on a rather high level. The deviating signs might be attributable to the modification of the dependent variable and its distribution.

[38] See Figure 1.A.3, p. 42.

Figure 1.A.1: Boxplots of the TTR and its logarithm divided by seniority

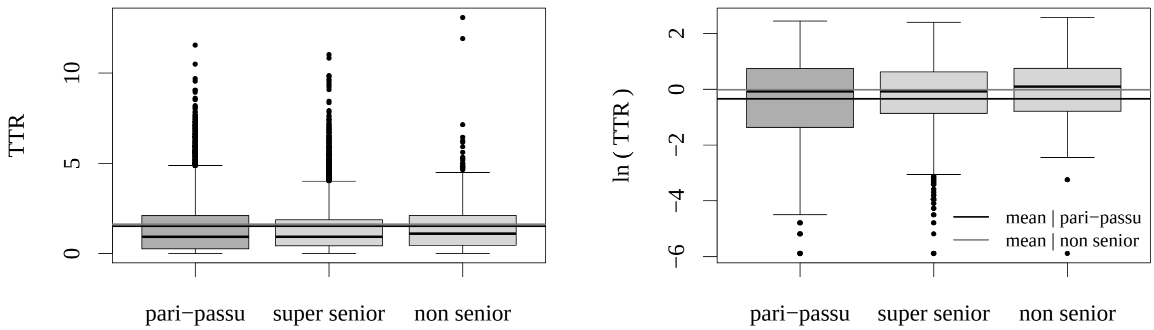


Figure 1.A.2: Boxplots of the TTR and its logarithm divided by collateral

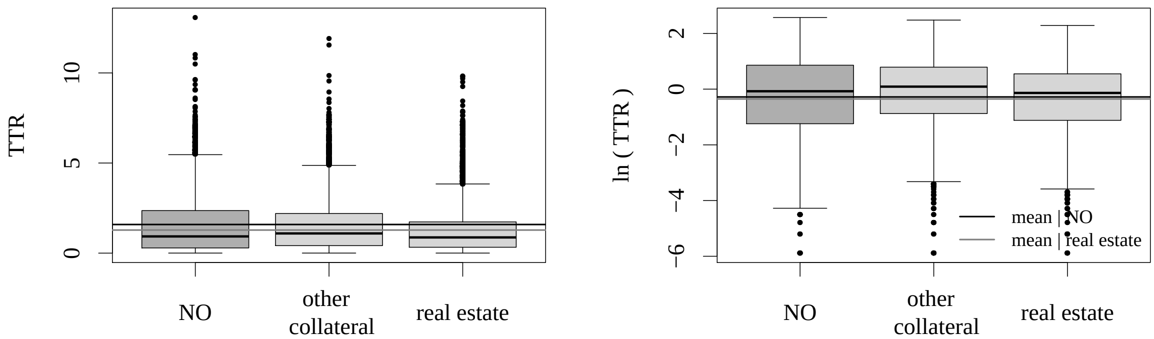


Figure 1.A.3: Boxplots of the TTR and its logarithm divided by industry

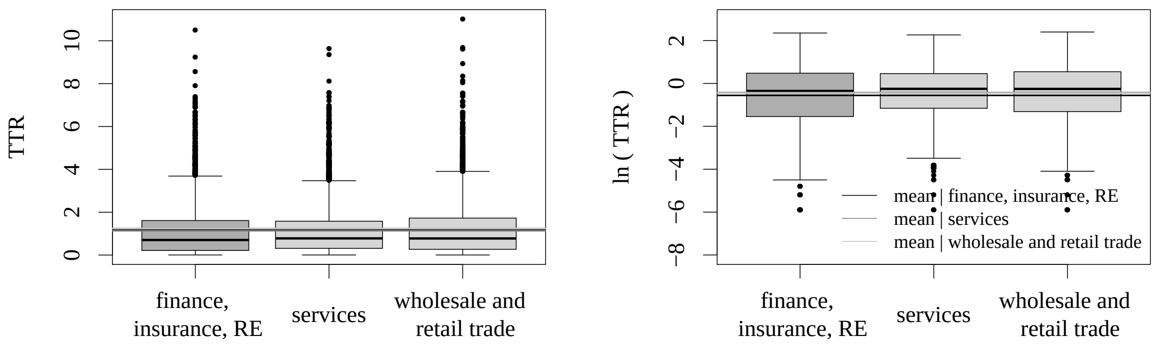


Table 1.A.1: Regression results of the log transformation of the TTR for the overall data set

		Coef.		SE
Intercept		0.771	***	(0.0585)
log(**EAD**)		0.034	***	(0.0039)
Time	2001	-0.638	***	(0.0391)
(2000)	2002	-0.696	***	(0.0335)
	2003	-1.056	***	(0.0341)
	2004	-1.078	***	(0.0374)
	2005	-1.424	***	(0.0341)
	2006	-1.299	***	(0.0320)
	2007	-1.387	***	(0.0350)
	2008	-1.354	***	(0.0461)
	2009	-1.509	***	(0.0561)
	2010	-1.115	***	(0.0387)
	2011	-1.732	***	(0.0448)
	2012	-2.159	***	(0.0464)
	2013	-2.497	***	(0.0643)
	2014	-2.715	***	(0.2074)
Country	United States	0.366	***	(0.0361)
(Germany)	Great Britain	0.711	***	(0.0291)
Facility	Short term	-0.276	***	(0.0169)
(Medium term)	Other	-0.118	***	(0.0211)
Seniority	Super senior	-0.170	***	(0.0221)
(Pari-passu)	Non senior	0.165	**	(0.0543)
	Unknown	0.061		(0.0428)
Nature of default	Sold at material credit loss	-1.169	***	(0.0998)
	Unlikely to pay	0.068	**	(0.0241)
	Charge-off / provision	0.440	***	(0.0201)
	Non accrual	0.339	***	(0.0258)
	Distressed restructuring	0.525	***	(0.0334)
	Bankruptcy	0.442	***	(0.0268)
	Unknown	-0.485	***	(0.0541)
Guarantee	YES	0.073	***	(0.0155)
(NO)	Unknown	-0.956	***	(0.2450)
Collateral	Other collateral	-0.107	***	(0.0188)
(NO)	Real estate	0.040	*	(0.0192)
	Unknown	-0.403		(0.6898)
Number of collateral		-0.009	***	(0.0017)
Cured (NO)	YES	-0.672	***	(0.0163)
Industry	Mining	-0.503	***	(0.1247)
(Finance, insurance, RE)	Transp., commu., sanitary services	-0.172	***	(0.0405)
	Services	0.044	.	(0.0234)
	Wholesale and retail trade	0.003		(0.0250)
	Manufacturing	0.077	**	(0.0265)
	Agric., forestry, fishing	0.046		(0.0649)
	Construction	0.178	***	(0.0297)
	Unknown	0.259	***	(0.0305)
Equity Index		-0.403	***	(0.0480)
GDP		-9.736	***	(0.6225)
Adjusted R-squared				38.77%
F-statistic				342.72
p-value				0.0000

Notes: Results of the multiple linear regression with log transformation regarding the overall data set. Significance codes: *** 0.001, ** 0.01, * 0.05, · 0.1. Standard errors (SE) are clustered by year.

1.B Appendix | Truncated regression

To consider the general nature of the data, we apply truncated regressions to subsets on yearly basis. The subsampling seems necessary as the limiting value has to be unique. Comparing the overall multiple linear regression (see Table 1.3) to the regressions on the subsamples (see Table 1.B.1 and 1.B.2), deviations with respect to the signs and significances arise.[39]

In the year 2008, sign switches can be observed regarding the facility *short term*, the nature of default *bankruptcy*, the collateral *other collateral*, several industries, and the GDP. However, only *other collateral* is statistically significant. This might be due to the considerably higher proportion of loan contracts located in Great Britain (34.92%) and the United States (49.22%) in this particular year.[40] Accordingly, the subsample might be driven by the Anglo American countries where the coefficient of *other collateral* exhibits a significantly positive sign.

In the year 2009, significant sign switches can be observed regarding the natures of default *unlikely to pay*, *charge off / provision*, and *bankruptcy*, and the collateral *other collateral*. Analogously, reasons may be found in the varying proportions of the loan origins as the signs correspond to the ones in the country subsets.

[39] See Table 1.3, p. 20, Table 1.B.1, p. 45 (subset 2008), and Table 1.B.2, p. 46 (subset 2009).

[40] Overall, these fractions amount to 31.87% (Great Britain) and 22.22% (United States).

Table 1.B.1: Regression results of the TTR for the linear and the truncated regression (TR) regarding 2008

		linear			TR		
		Coef.		SE	Coef.		SE
Intercept		-0.444		(0.3298)	-0.501		(0.3461)
log(**EAD**)		0.076	***	(0.0214)	0.080	***	(0.0223)
Country	United States	2.958	***	(0.2032)	3.062	***	(0.2127)
(Germany)	Great Britain	2.426	***	(0.1404)	2.528	***	(0.1472)
Facility	Short term	0.011		(0.0919)	0.014		(0.0968)
(Medium term)	Other	-0.498	***	(0.1232)	-0.512	***	(0.1273)
Seniority	Super senior	-1.796	***	(0.1270)	-1.908	***	(0.1345)
(Pari-passu)	Non senior	-0.092		(0.3351)	-0.089		(0.3533)
	Unknown	-1.369	***	(0.3947)	-1.444	***	(0.4068)
Nature of default	Sold at material credit loss	-2.125		(1.3811)	-2.169		(1.4049)
	Unlikely to pay	0.097		(0.1222)	0.098		(0.1273)
	Charge-off / provision	0.359	**	(0.1301)	0.378	**	(0.1368)
	Non accrual	0.432	***	(0.1121)	0.469	***	(0.1183)
	Distressed restructuring	0.168		(0.4170)	0.154		(0.4296)
	Bankruptcy	-0.146		(0.1453)	-0.159		(0.1507)
	Unknown	-1.876	.	(0.9820)	-1.985	*	(1.0100)
Guarantee	YES	0.107		(0.0869)	0.112		(0.0907)
(NO)	Unknown	-0.960		(1.3901)	-1.006		(1.4249)
Collateral	Other collateral	0.184	.	(0.1020)	0.199	.	(0.1069)
(NO)	Real estate	-0.233	*	(0.1186)	-0.248	*	(0.1243)
	Unknown	-0.394		(1.4348)	-0.404		(1.4672)
Number of collateral		-0.016		(0.0138)	-0.015		(0.0143)
Cured (NO)	YES	-0.254	**	(0.0924)	-0.268	**	(0.0960)
Industry	Mining	-1.546		(1.3787)	-1.543		(1.4022)
(Finance, insurance, RE)	Transp., commu., sanitary services	0.265		(0.1986)	0.290		(0.2111)
	Services	0.043		(0.1296)	0.045		(0.1353)
	Wholesale and retail trade	0.055		(0.1202)	0.062		(0.1260)
	Manufacturing	0.189		(0.1282)	0.195		(0.1341)
	Agric., forestry, fishing	-0.199		(0.4439)	-0.221		(0.4606)
	Construction	0.311	**	(0.1204)	0.337	**	(0.1278)
	Unknown	-1.609	***	(0.1777)	-1.663	***	(0.1843)
Equity Index		-0.587		(0.5311)	-0.605		(0.5565)
GDP		3.892		(3.8913)	4.103		(4.1340)
Adjusted R-squared				37.48%			
F-statistic				28.50			
p-value				0.0000			
sigma					1.394	***	(0.0280)
p-value							-2515.79 (34)

Notes: Results of the multiple linear regression and the truncated regression regarding the year 2008. Significance codes: *** 0.001, ** 0.01, * 0.05, · 0.1.

Table 1.B.2: Regression results of the TTR for the linear and the truncated regression (TR) regarding 2009

		linear			TR		
		Coef.		SE	Coef.		SE
Intercept		1.100	***	(0.3066)	1.069	***	(0.3199)
log(**EAD**)		0.054	***	(0.0153)	0.056	***	(0.0160)
Country	United States	0.598	*	(0.2465)	0.644	*	(0.2565)
(Germany)	Great Britain	0.812	***	(0.2292)	0.876	***	(0.2388)
Facility	Short term	-0.099		(0.0660)	-0.108		(0.0697)
(Medium term)	Other	-0.301	***	(0.0878)	-0.305	***	(0.0907)
Seniority	Super senior	-0.882	***	(0.1003)	-0.934	***	(0.1048)
(Pari-passu)	Non senior	-0.345		(0.2517)	-0.357		(0.2637)
	Unknown	-0.891	**	(0.3215)	-0.935	**	(0.3312)
Nature of default	Sold at material credit loss	-1.067	***	(0.1677)	-1.052	***	(0.1720)
	Unlikely to pay	-0.339	***	(0.0815)	-0.347	***	(0.0849)
	Charge-off / provision	-0.371	***	(0.1119)	-0.395	***	(0.1173)
	Non accrual	0.306	***	(0.0725)	0.348	***	(0.0771)
	Distressed restructuring	-0.209		(0.5166)	-0.218		(0.5461)
	Bankruptcy	-0.662	***	(0.1241)	-0.692	***	(0.1286)
	Unknown	0.444	*	(0.1990)	0.474	*	(0.2151)
Guarantee	YES	0.258	***	(0.0654)	0.276	***	(0.0692)
(NO)	Unknown	-2.674	**	(0.8516)	-2.779	**	(0.8735)
Collateral	Other collateral	0.119	.	(0.0683)	0.128	.	(0.0718)
(NO)	Real estate	-0.008		(0.0821)	-0.004		(0.0864)
Number of collateral		-0.013		(0.0119)	-0.015		(0.0126)
Cured (NO)	YES	-0.545	***	(0.0574)	-0.574	***	(0.0599)
Industry	Mining	0.578		(0.4798)	0.585		(0.4966)
(Finance, insurance, RE)	Transp., commu., sanitary services	-0.053		(0.1452)	-0.050		(0.1515)
	Services	-0.031		(0.0933)	-0.032		(0.0981)
	Wholesale and retail trade	-0.234	**	(0.0825)	-0.246	**	(0.0863)
	Manufacturing	0.225	**	(0.0867)	0.241	**	(0.0914)
	Agric., forestry, fishing	-0.107		(0.2930)	-0.116		(0.3080)
	Construction	0.002		(0.0796)	0.000		(0.0839)
	Unknown	-0.467	***	(0.1313)	-0.492	***	(0.1367)
Equity Index		-0.702	***	(0.1683)	-0.727	***	(0.1751)
GDP		-3.460		(2.4724)	-3.926		(2.5794)
Adjusted R-squared				28.44%			
F-statistic				27.76			
p-value				0.0000			
sigma					1.169	***	(0.0198)
p-value							-3208.28 (33)

Notes: Results of the multiple linear regression and the truncated regression regarding the year 2009. Significance codes: *** 0.001, ** 0.01, * 0.05, · 0.1.

1.C Appendix | Alternative dependent variable

Within the robustness analysis, the TTR is replaced by a different dependent variable – the loan duration (See Table 1.C.1). As previously stated, several deviations among the coefficients arise in comparison with the initial dependent variable.[41]

Sign switches can be observed regarding the facility *short term*, the nature of default *unlikely to pay*, and the guarantee indicator. However, non of the mentioned is statistically significant with respect to the loan duration. In contrast, significance emerge regarding the seniority *non senior*. The coefficient is significantly positive indicating a backward shift of the cash-flow-centroid. High payments might, therefore, reveal rather at the end of the resolution process.

In addition, discrepancies can be observed in the country coefficients. While Great Britain exhibits the longest TTR, the highest loan duration arises in the United States. To examine the significance of these findings, the regressions are re-estimated with Great Britain serving as the reference category. With respect to the resolution time, the United States exhibit a coefficient of -0.345 (***) indicating a significantly shorter resolution process. The corresponding coefficient regarding the loan duration is 0.115 (***). This implies that the centroid of cash flows is rather later in the resolution process in the United States compared to Great Britain.

[41] See Table 1.3, p. 20.

Table 1.C.1: Regression results of the loan duration for the overall data set

		Coef.		SE
Intercept		2.164	***	(0.0927)
log(**EAD**)		0.078	***	(0.0051)
Time	2001	-0.807	***	(0.0966)
(2000)	2002	-0.917	***	(0.0860)
	2003	-1.489	***	(0.0722)
	2004	-1.440	***	(0.0742)
	2005	-1.755	***	(0.0685)
	2006	-1.716	***	(0.0675)
	2007	-1.762	***	(0.0694)
	2008	-1.549	***	(0.0806)
	2009	-1.947	***	(0.0909)
	2010	-1.668	***	(0.0749)
	2011	-2.001	***	(0.0726)
	2012	-2.230	***	(0.0720)
	2013	-2.362	***	(0.0730)
	2014	-2.509	***	(0.0878)
Country	United States	0.377	***	(0.0519)
(Germany)	Great Britain	0.262	***	(0.0369)
Facility	Short term	0.014		(0.0213)
(Medium term)	Other	-0.383	***	(0.0275)
Seniority	Super senior	-0.080	**	(0.0288)
(Pari-passu)	Non senior	0.123	.	(0.0710)
	Unknown	-0.366	***	(0.0736)
Nature of default	Sold at material credit loss	-0.676	***	(0.1061)
	Unlikely to pay	-0.001		(0.0267)
	Charge-off / provision	0.290	***	(0.0251)
	Non accrual	0.142	***	(0.0295)
	Distressed restructuring	0.486	***	(0.0446)
	Bankruptcy	0.148	***	(0.0417)
	Unknown	-0.120	**	(0.0365)
Guarantee	YES	-0.022		(0.0200)
(NO)	Unknown	-1.149	*	(0.5134)
Collateral	Other collateral	-0.221	***	(0.0243)
(NO)	Real estate	-0.181	***	(0.0226)
	Unknown	0.100		(0.6333)
Number of collateral		-0.009	***	(0.0014)
Cured (NO)	YES	-0.251	***	(0.0192)
Industry	Mining	-0.335	*	(0.1490)
(Finance, insurance, RE)	Transp., commu., sanitary services	-0.215	***	(0.0473)
	Services	-0.071	**	(0.0253)
	Wholesale and retail trade	-0.062	*	(0.0270)
	Manufacturing	0.043		(0.0318)
	Agric., forestry, fishing	0.155	**	(0.0601)
	Construction	0.120	***	(0.0320)
	Unknown	0.036		(0.0444)
Equity Index		-0.371	***	(0.0686)
GDP		-8.399	***	(0.8175)
Adjusted R-squared				21.75%
F-statistic				150.35
p-value				0.0000

Notes: Results of the multiple linear regression with an alternative dependent variable – the loan duration. Significance codes: *** 0.001, ** 0.01, * 0.05, · 0.1. Standard errors (SE) are clustered by year.

1.D Appendix | LGD concepts

Member banks and external partners receive two different LGD concepts (LGD_1 and LGD_2) from GCD. Both of them show inexplicably extreme values. Therefore, we evolve a third concept (LGD_3) which is less shaped by outliers. Generally, LGDs are calculated as $LGD_i = 1 - RR_i$ regarding all concepts, whereby, RR_i denotes the corresponding recovery rate:

$$RR_i = \frac{P_i}{OA_i}.$$

Hereby, P_i indicates the sum of the present values of the transactions and OA_i the outstanding amount of the loan. The three LGD concepts differ in the transaction types included in P_i and OA_i. Table 1.D.1 summarizes these types and how they influence the three measures. The quote $+_P$ ($-_P$) indicates an increase (decrease) in the present value P_i and $+_{OA}$ ($-_{OA}$) an increase (decrease) in the outstanding amount OA_i. The key feature of the first LGD concept is that

Table 1.D.1: GCD transaction types

Transaction type	LGD_1	LGD_2	LGD_3
Principal payment	$+_P$	$+_P$	$+_P$
Interest payment	$+_P$	$+_P$	$+_P$
Recorded book value of collateral	$+_P$	$+_P$	$+_P$
Post-resolution payment	$+_P$	$+_P$	$+_P$
Charge-off			
Provision			
Principal advance	$-_P$	$+_{OA}$	$+_{OA}$
Cash out on guarantee			
Financial claim	$-_P$	$+_{OA}$	$+_{OA}$
Interest accrual			$+_{OA}$
Fees and commissions charged			$+_{OA}$
Fees and commissions received	$+_P$	$+_P$	$+_P$
Legal expenses	$-_P$	$-_P$	$-_P$
Administrator or receiver fees	$-_P$	$-_P$	$-_P$
Liquidation expenses	$-_P$	$-_P$	$-_P$
Other external workout costs	$-_P$	$-_P$	$-_P$
Waiver (off B/S commitment)	$+_P$	$+_P$	$+_P$

Notes: Transactions types associated with the resolution of a defaulted loan.

the outstanding amount remains unaffected by the transactions and equals the EAD. In the second concept, two transaction types increase the outstanding amount instead of reducing the present value. This corresponds to the economic intuition, since these types are actually non-cash effective. Two additional transaction types are included in the third concept. This seems consistent as their cash flow counterparts are included in the present value.

Chapter 2

Macroeconomic effects and frailties in the resolution of non-performing loans

This chapter is joint work with Steffen Krüger[*], Ralf Kellner[†], and Daniel Rösch[‡] published as:
Betz, J., S. Krüger, R. Kellner, D. Rösch (2017). Macroeconomic effects and frailties in the resolution of non-performing loans. *Journal of Banking and Finance*, forthcoming (available online since September, 2017).

https://doi.org/10.1016/j.jbankfin.2017.09.008

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.

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

JEL classification: C23, G21, G33

[*] BMW Financial Services, 80939 Munich, Germany, email: steffen.krueger@ur.de.
[†] University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: ralf.kellner@ur.de.
[‡] University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: daniel.roesch@ur.de.

2.1 Introduction

The default resolution time (DRT) of non-performing loans[1] 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).[2] 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 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,

[1] 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.

[2] 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 2.2 for more details).

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.

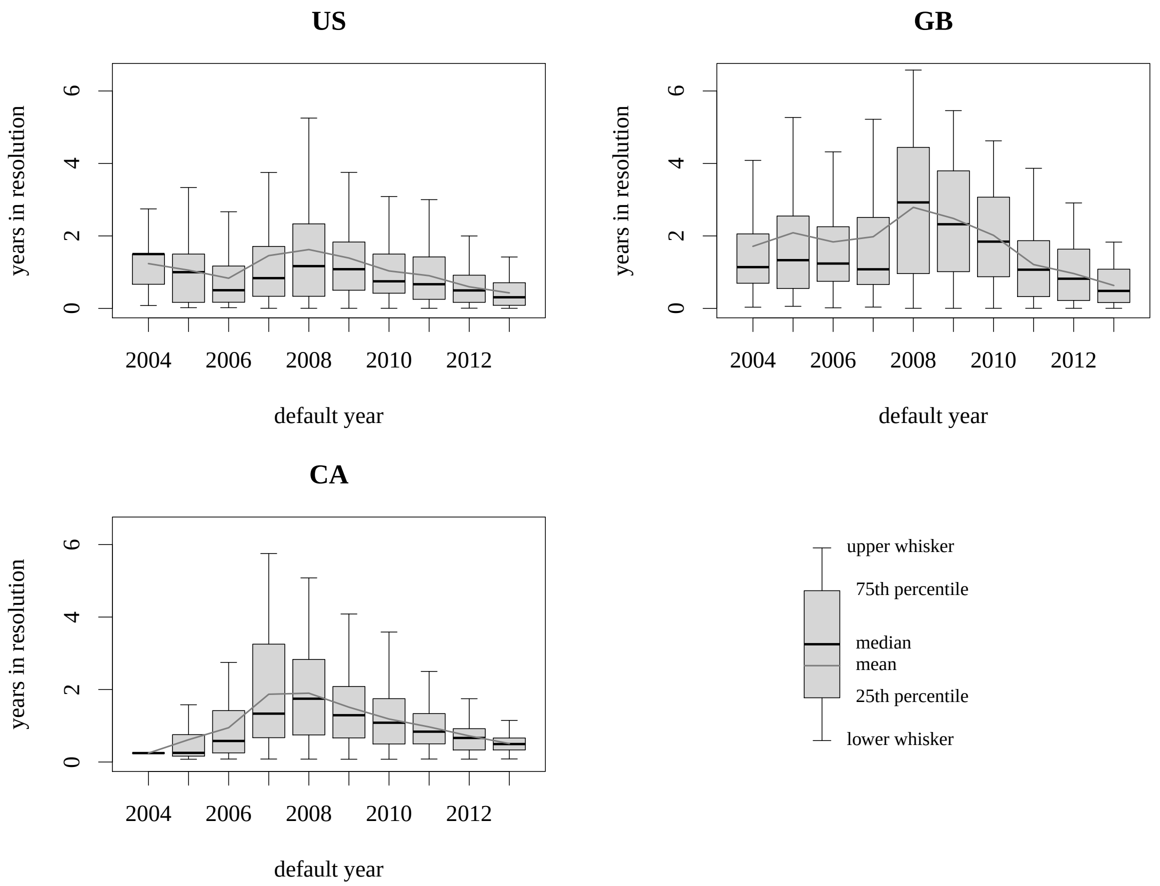
Our paper provides the following contributions. First, we thoroughly examine systematic effects among DRTs of non-performing loans. To the best of our knowledge, we are the first to extend the modeling scheme of DRTs to unobservable systematic factors using a doubly-stochastic continuous-time hazard 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 2.2 provides an example for the potential impact of systematic effects among default resolution times. Section 2.3 describes our data and methodology. Section 2.4 provides the empirical results. Section 2.5 shows the practical implications of these results. Section 2.6 concludes.

2.2 Why care about systematic effects among DRTs?

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 2.1: Systematic movements in DRTs

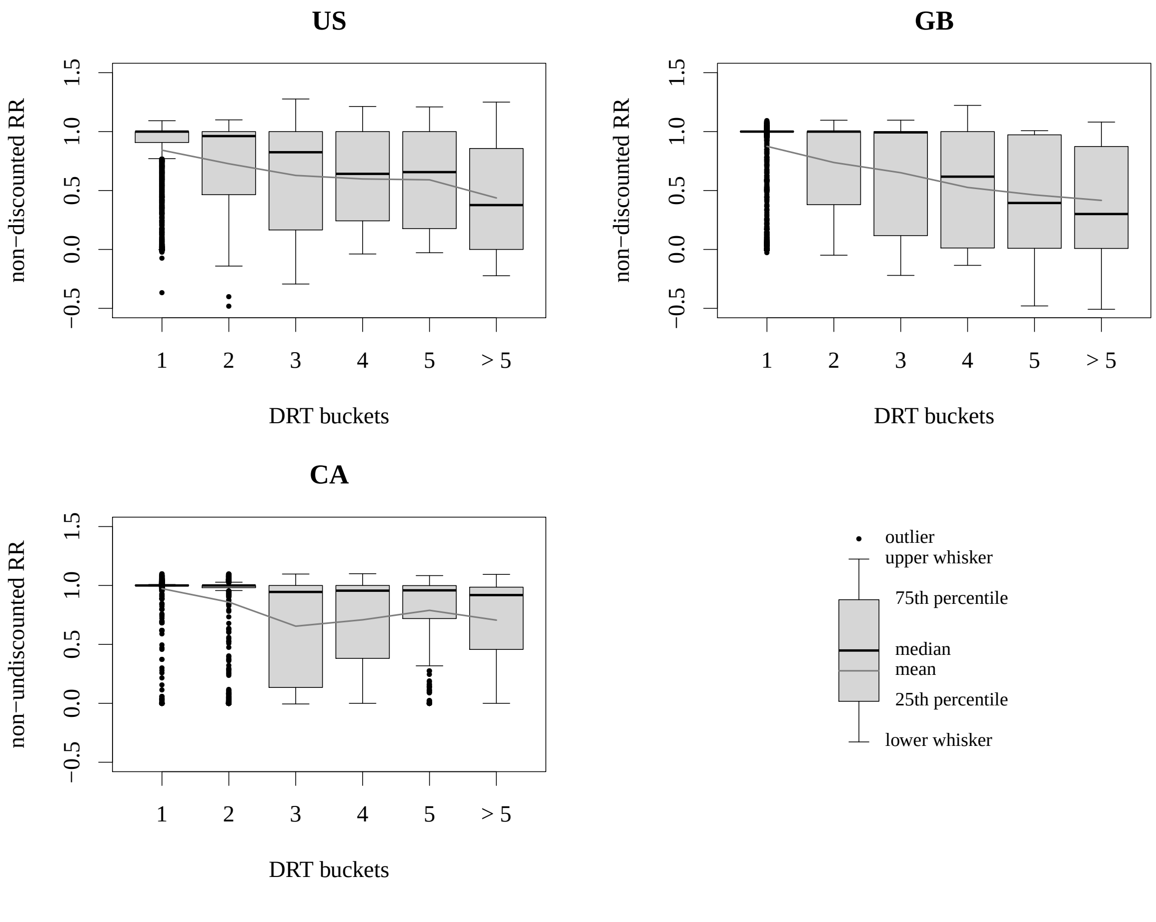


Notes: The figure illustrates the systematic movements of DRTs for resolved loans. Box plots of the DRTs per year for the US, Great Britain, and Canada are displayed, whereas, outliers are hidden due to presentational purpose. The black horizontal lines within the box plots mark the medians. The means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

Figure 2.1 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.[3] This 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 2.2 shows box plots of the non-discounted recovery rate (RR) divided by DRT buckets. The first bucket includes all loans with DRTs up to one year, the second one

Figure 2.2: 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.

[3] The increase of DRTs in Canada already starts in 2007 for at least some loans, which is caused by longer resolution processes in general that push forward crises effects. A detailed description of the dataset is given in Section 2.3.

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.[4] 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, DRTs will play an important role with respect to future regulation standards. 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-performing. As long as the non-performing status persists, banks need to provide additional liquidity for the affected assets. Thus, these additional liquidity demands persist, the longer defaulted loan resolutions last.

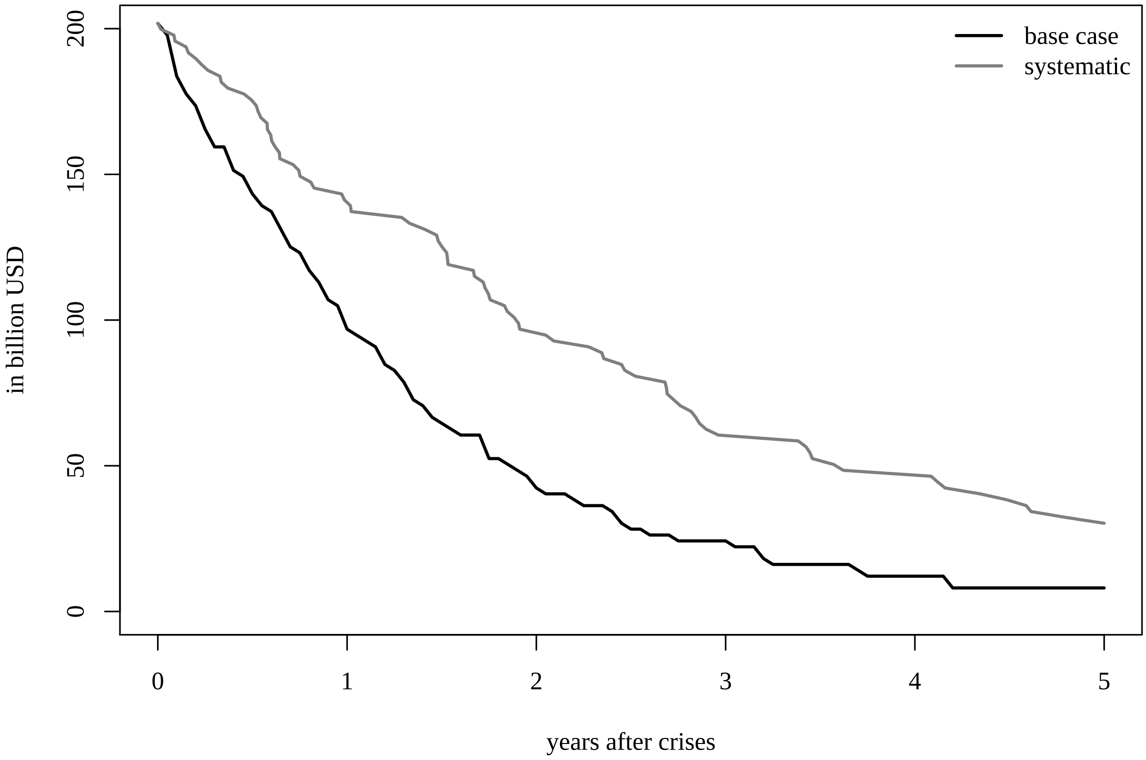
The analysis in Section 2.4.3 shows that DRTs systematically increase during times of financial crisis. In addition, the ratio of non-performing loans is likely to increase at the same time.[5] Both effects burden the banking industry with additional liquidity demands which are more pronounced, the higher systematic effects among DRTs are. To exemplary show this effect that may occur after the implementation of the NSFR, picture the following example. Given additional required liquidity of 201.80 billion USD due to an increase in the ratio of non-

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

[5] E.g., during the GFC the non-performing loan ratio in the banking industry reached with 5.64% its maximum since the time series data started in 1984.

performing loans.[6] As before, we compare 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 2.3 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

Figure 2.3: 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 (black line), an average DRT of 1.59 years is assumed, whereas, in the systematic case (gray line) the average DRT amounts to 2.42 years.

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

[6] This number is based on a scenario linked to the GFC. Even though discussions on the NSFR have been introduced after the GFC, we treat this counterfactual analysis as if future regulation standards would have already been present during the GFC. This is done to create an example for the impact of DRTs with representative values of a severe financial crisis. Assuming an average RSF factor of 50% for performing loans, after the start of the GFC, 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. These numbers are taken from the public research data base of `https://fred.stlouisfed.org/`. The time series are indexed by USTLLNUI and USNPTL.

years), the reduction of additional 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.

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. This example shows that 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 in the future during downturns due to changes in the regulatory environment and may even dampen upswings.

2.3 Methods and data

2.3.1 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

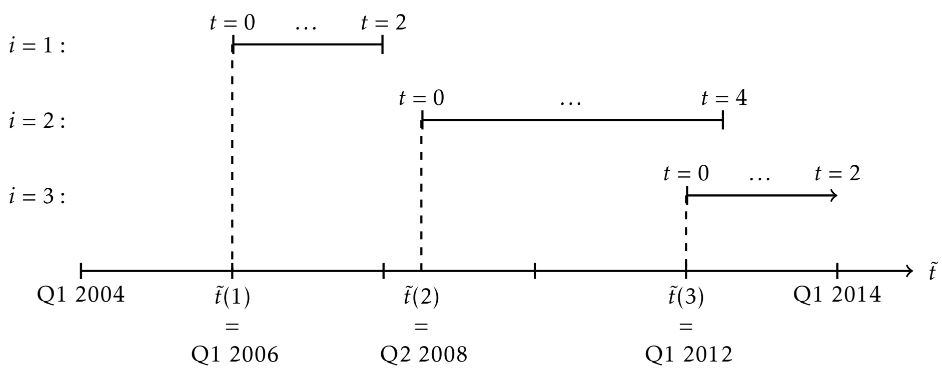
$$\lambda_{it} = \lim_{\Delta t \to 0,\ t>0} \frac{P(\ t \leq T < t + \Delta t \mid T \geq t\ )}{\Delta t} \ , \tag{2.1}$$

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

$$\lambda_{it} = \lambda_{0t} \exp\left(x_i \beta\right) \ , \tag{2.2}$$

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 2.A. In the following, we refer to Equation (2.2) as Model I. Since it only contains loan specific characteristics as covariates, it serves as a reference model.

Figure 2.4: 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.

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

$$\lambda_{it} = \lambda_{0t} \exp\left(x_i \beta + z_{\tilde{t}(i)} \gamma\right) \ , \tag{2.3}$$

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 (2.3) as Model II. Figure 2.4 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.

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.,

$$U_{\tilde{t}(i)} \sim N(0, \sigma^2)\,, \tag{2.4}$$

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

$$\lambda_{it} = \lambda_{0t} \exp\left(x_i\beta + z_{\tilde{t}(i)}\gamma + U_{\tilde{t}(i)}\right)\,, \tag{2.5}$$

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.

2.3.2 Data

This paper uses a subsample of the unique loss data base provided by GCD.[7] 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

[7] 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.

(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.[8] We further restrict 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 2.1 displays descriptive statistics for the DRT. 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 2.5 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

[8] The magnitude of 500 USD is inspired by the materiality threshold of the European Banking Authority (2016a).

Table 2.1: 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

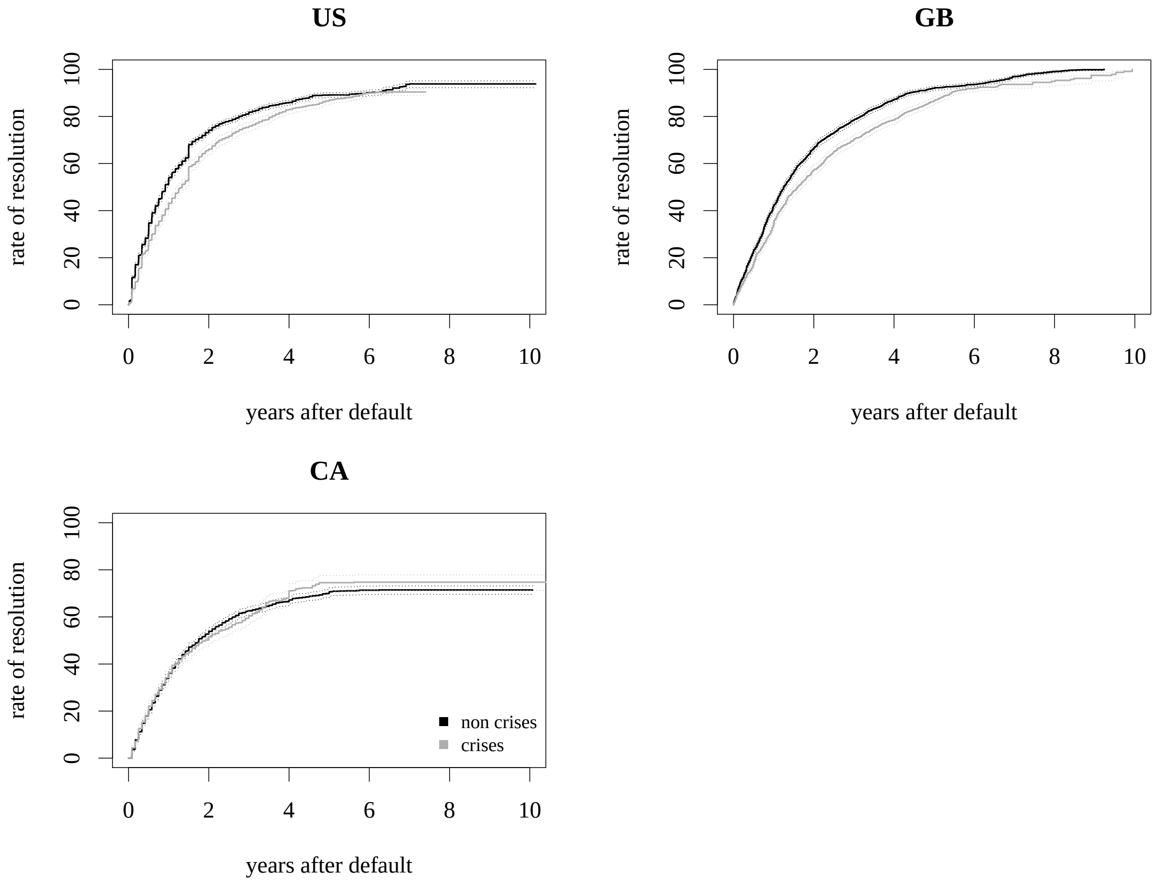
Notes: The table summarizes the descriptive statistics (mean, median, and standard deviation) of DRTs for the overall data set and separated for the US, Great Britain, and Canada. The presentation additionally distinguishes between resolved and unresolved loans. The latter indicates that these loans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

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.

Loan specific variables

Table 2.2 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

Figure 2.5: 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 dotted 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).

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 (2.3) and (2.5)). 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 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

Table 2.2: 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	1	1	0
	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.

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.[9] Figure 2.6 shows the macroeconomic variables. The Global Financial Crisis results in increasing term spreads, volatility indices and lower industry production and equity indices. Table 2.3 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

[9] 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 2.4.3)

signs and significances may result when correlated independent variables are simultaneously included in a model.

Table 2.3: 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 (2.3) and Model III as of Equation (2.5). 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.

Furthermore, we include a World Bank score measuring the efficiency of default resolution (see World Bank, 2015a). 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).[10]

2.4 Results

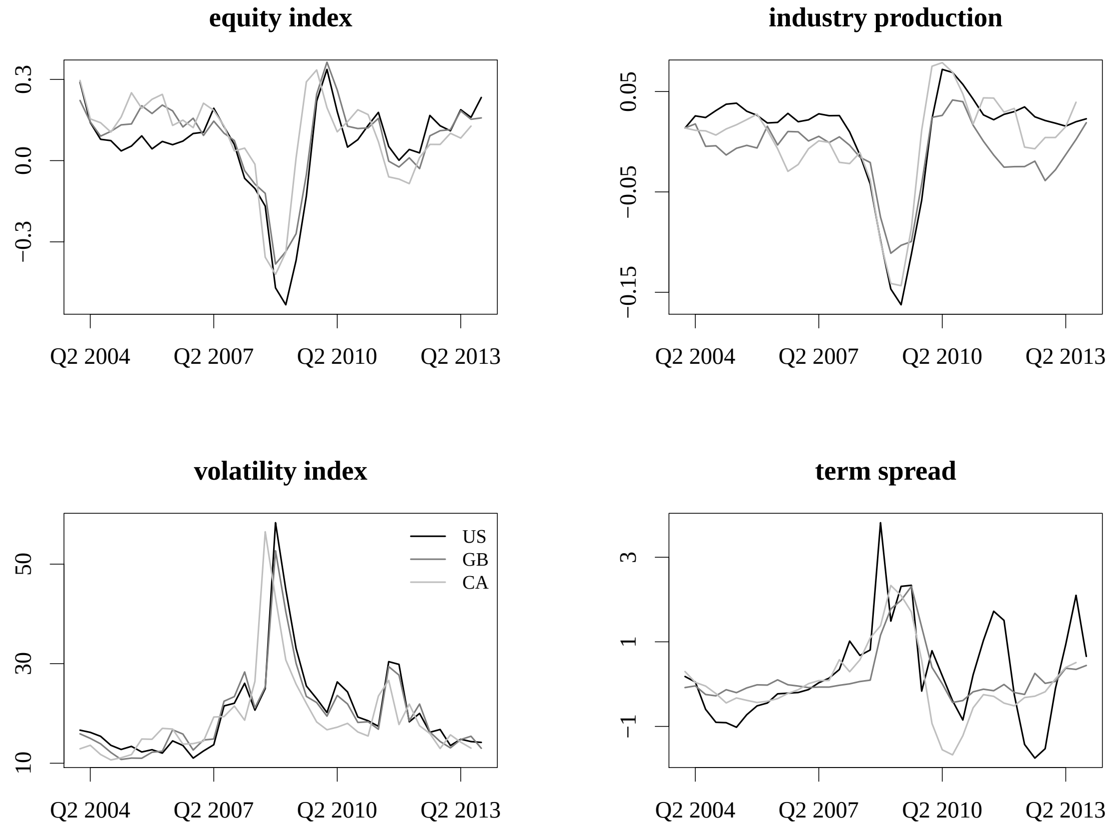
2.4.1 Overview of formal and informal proceedings of resolution

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 2.4). 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.

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

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

Figure 2.6: Descriptive statistics of macroeconomic variables



Notes: The figure illustrates the time-series of macroeconomic variables which are included in Model II as of Equation (2.3) and Model III as of Equation (2.5). 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.

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

Table 2.4: 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.

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. In addition, creditors want to avoid or delay recognizing losses to prevent rating downgrades, and thus, more expensive access to funding or supervisory intervention.[11] 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.

2.4.2 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 2.5 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

Table 2.5: 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)
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 (2.2), 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 2.B.3.

intensity to resolution and, thus, a tendency to shorter resolution processes. Dermine and Neto

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

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, 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.[12] 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.

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

2.4.3 The systematic movement of resolution processes

Legal and administrative reasons

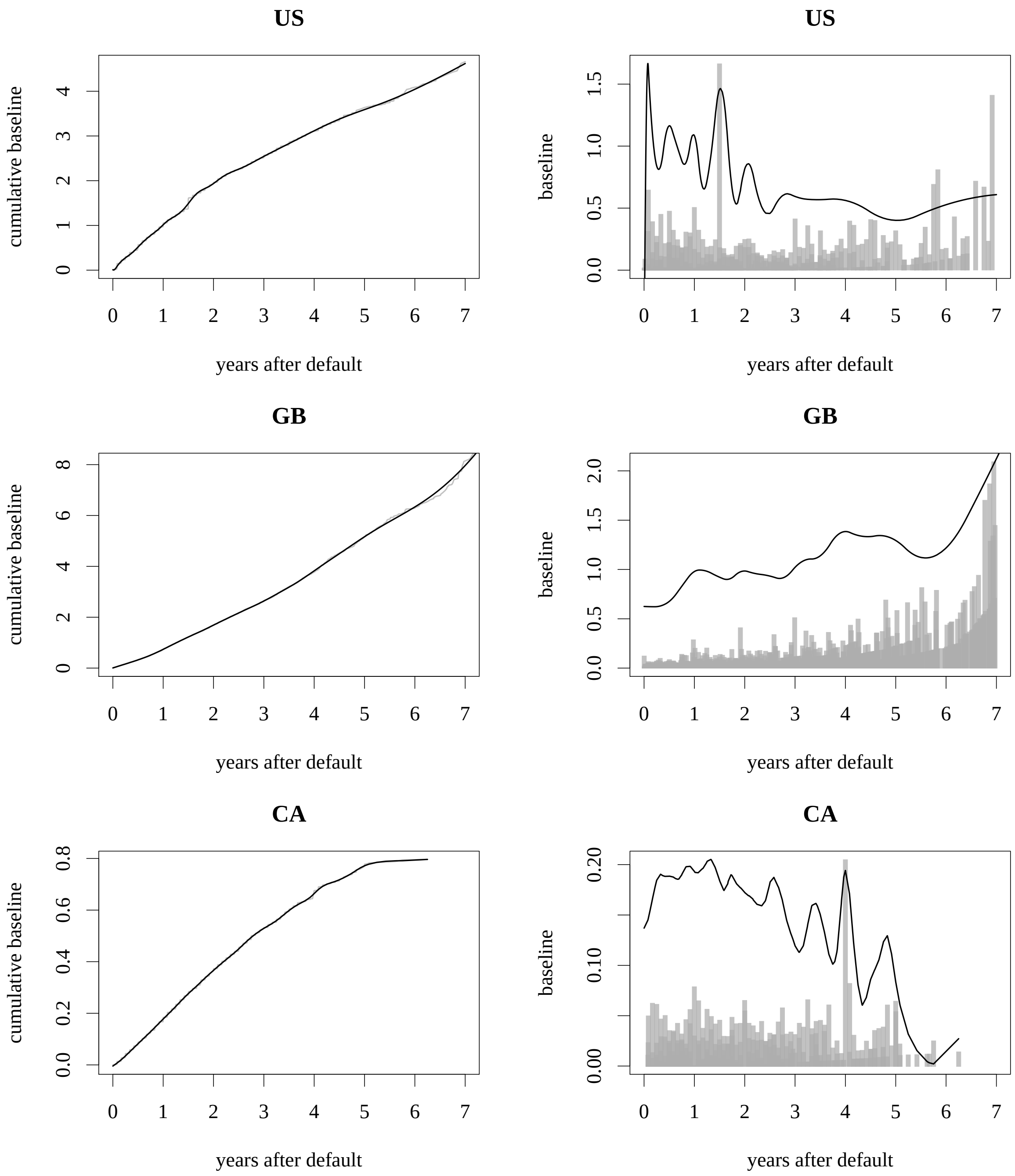
In addition to loan specific characteristics, country specific legal conditions and bank practices affect the time to resolution (see Section 2.4.1 for an overview). As stated in Section 2.3.1, the Cox proportional hazard model is a semi-parametric approach because the baseline hazard rate λ_{0t} has 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 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 2.7 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 2.2, 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

Figure 2.7: 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.

considerably lower. This corresponds with the theoretical considerations in Section 2.4.1. 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 informal 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 2.6 summarizes the country specific estimation results. Compared to the results of Model I only minor differences regarding loan specific impacts arise (see Table 2.5). 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[13] 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 2.6).

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. 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

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

Table 2.6: 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)
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 (2.3), 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 2.B.3.

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.[14] This suggests a systematic co-movement of DRTs caused by the macroeconomy in both countries. In

[14] 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.

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 2.4.1, 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 individually (see Table 2.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 2.7 summarizes the regression results. The changes in parameter estimates for loan specific variables are minor.

Table 2.7: Regression results for Model III

		United States			Great Britain			Canada		
		Coef.		SE	Coef.		SE	Coef.		SE
log(EAD)		-0.0545	***	(0.0072)	-0.0766	***	(0.0077)	-0.0252	**	(0.0120)
Asset Class (SME)	Large Corporates	0.2478	***	(0.0386)	0.3473	***	(0.0509)	-0.0885		(0.0611)
Facility (Medium term)	Short term	0.1087	***	(0.0415)	-0.0812	**	(0.0316)	-0.0949	**	(0.0449)
	Other	-0.0534		(0.0343)	0.4482	***	(0.0569)	0.0718		(0.0588)
Seniority (Pari-passu)	Super senior	0.2164	***	(0.0497)	0.9269	***	(0.0436)	0.6240	***	(0.1016)
	Non senior	0.4365	***	(0.1597)	0.8704	***	(0.2366)			
	Unknown	0.1056		(0.0993)				0.7252	***	(0.1222)
Nature of default (90 days past due)	Unlikely to pay	-0.3034	***	(0.0428)	-0.0501		(0.0535)	-0.5747	*	(0.3096)
	Bankruptcy	0.2071	***	(0.0767)	0.0394		(0.0546)	-0.3583		(0.3044)
	Charge-off / provision	0.4992	***	(0.1083)	-0.1163	**	(0.0463)	0.8721	**	(0.3742)
	Sold at material credit loss	1.9401	***	(0.1159)						
	Distressed restructuring	0.2621		(0.1618)	-0.3562	***	(0.0988)			
	Non accrual	-0.0340		(0.0374)	-0.0793	*	(0.0466)	-0.4583		(0.2955)
	Unknown	-0.1681		(0.1292)	1.0076	***	(0.0827)	-0.7400	**	(0.3026)
Guarantee (NO)	Unknown	0.5355	**	(0.2496)						
	YES	0.1121	***	(0.0293)	-0.0927	***	(0.0319)	0.2030	***	(0.0748)
Collateral (NO)	Other collateral	0.1264	***	(0.0338)	0.0888	**	(0.0383)	0.3846	***	(0.0805)
	Real estate	0.0724		(0.0474)	0.1036	***	(0.0378)	0.2006		(0.1895)
	Unknown	-1.6879	***	(0.2472)				1.3329	***	(0.1151)
Number of collateral		0.0240	***	(0.0086)	0.0062	**	(0.0026)	0.0090		(0.0070)
Cured (NO)	YES	0.5411	***	(0.0334)	0.9685	***	(0.0366)	0.9615	***	(0.0552)
Industry (Finance, insurance, RE)	Agric., forestry, fishing	-0.2295	*	(0.1223)	0.0343		(0.0971)	-0.0599		(0.1026)
	Mining	-0.1980		(0.1452)	-0.2885		(0.2307)	0.5405	***	(0.1783)
	Construction	-0.2932	***	(0.0547)	-0.3495	***	(0.0612)	0.2109	**	(0.0976)
	Manufacturing	-0.2609	***	(0.0468)	0.0705		(0.0584)	0.2935	***	(0.0865)
	Transp., commu., sanitary services	-0.0042		(0.0624)	0.1649	**	(0.0783)	0.2455	**	(0.1036)
	Wholesale and retail trade	-0.1456	***	(0.0490)	-0.0311		(0.0521)	0.2392	***	(0.0842)
	Services	-0.1041	**	(0.0470)	0.0345		(0.0558)	0.2112	***	(0.0810)
	Unknown	0.2803	***	(0.0482)	0.0999		(0.0988)	0.0979		(0.2975)
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 (2.5), 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 2.B.3.

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.[15] 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

[15] Even though differences in macroeconomic variables are less pronounced in Figure 2.6, systematic tendencies regarding credit dynamics in Canada seem to differ more clearly in comparison to the US and Great Britain. This can be seen in Figure 2.B.1 in Appendix 2.B which shows that the lending behavior and resolution practices seem to deviate among countries.

& Snell's R^2 – improves for the US 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

$$\begin{aligned}
\frac{E(T|U_{\tilde{t}(i)} = u_0 + \Delta u)}{E(T|U_{\tilde{t}(i)} = u_0)} - 1 &= \frac{\lambda_{it|U_{\tilde{t}(i)}=u_0}}{\lambda_{it|U_{\tilde{t}(i)}=u_0+\Delta u}} - 1 \\
&= \frac{\lambda_{0t}\exp\left(x_i\beta + z_{\tilde{t}(i)}\gamma + u_0\right)}{\lambda_{0t}\exp\left(x_i\beta + z_{\tilde{t}(i)}\gamma + u_0 + \Delta u\right)} - 1 \\
&= \exp(-\Delta u) - 1 \,, \qquad (2.6)
\end{aligned}$$

when assuming a constant baseline hazard rate for the definition of Model III in Equation (2.5). Table 2.8 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 2.8: 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%

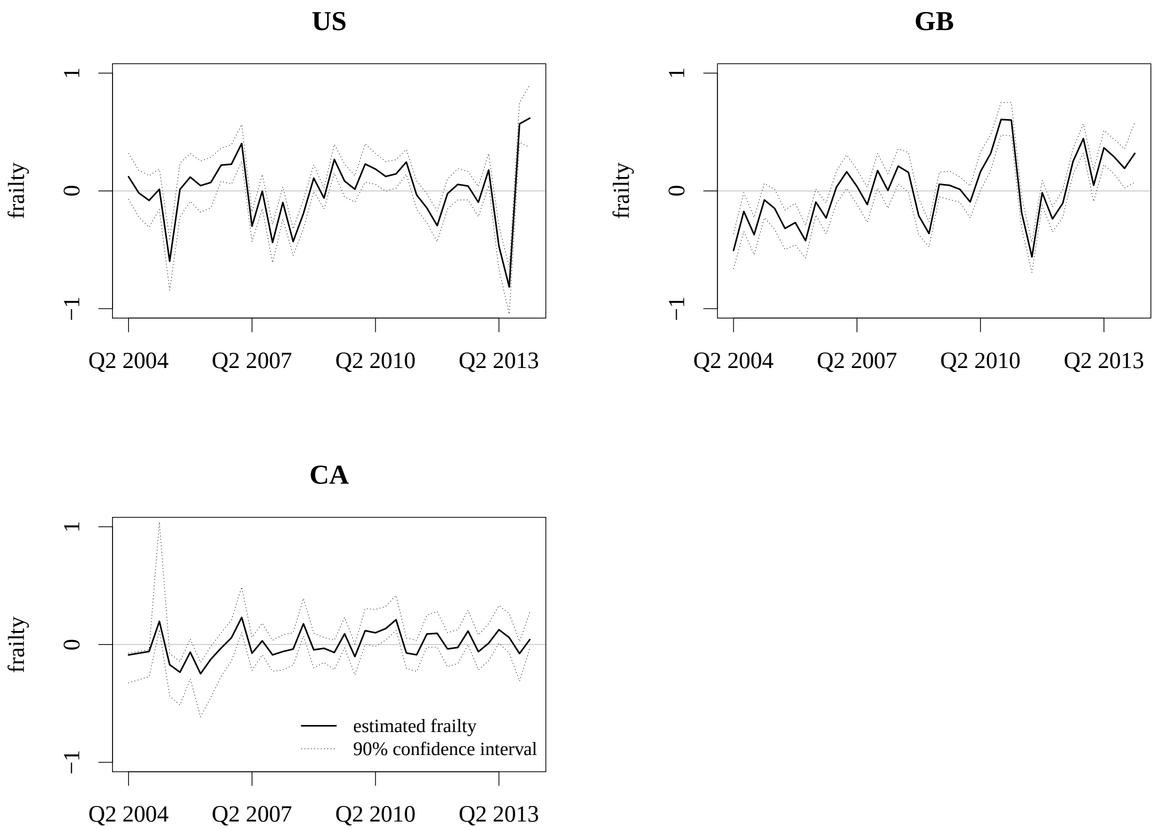
Notes: The table summarizes the impact of the frailty on the DRT as of Equation (2.6). The relative change of the mean DRT for a frailty change of one standard deviation is displayed.

Generally, frailties represent the development of the systematic components over time. Figure 2.8 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

[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%.

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.

Figure 2.8: 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 dotted 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 2.B.3.

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 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 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 2.B.2, Table 2.B.1 and Table 2.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

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

[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).

relevant for the creditor itself, but also for regulators who nowadays often 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 2.2, there seems to be a negative 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]

2.5 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.

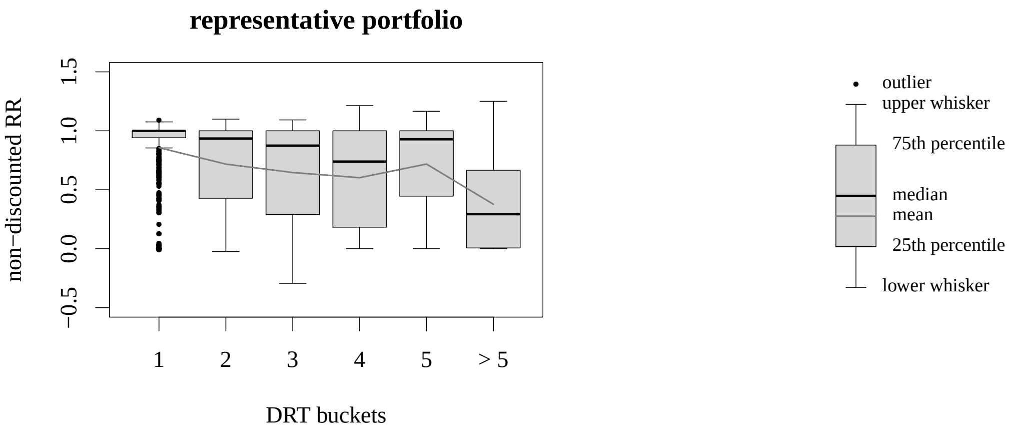
The representative portfolio consists of 1,000 US American loans which are randomly sampled from our data set. Firstly, we consider one of these loans to derive implications of systematic factors on loan level. Secondly, we extend the analysis to portfolio level by including all of the 1,000 loans. The relation between DRTs and non-discounted RRs is explicitly taken into account

[19] For instance, risk assessment under stressed market conditions is demanded for market risks under Basel III (Basel Committee on Banking Supervision, 2010) 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.

in the second part of the analysis. Figure 2.9 shows the relation for the randomly sampled, representative portfolio. Compared to the complete data sample only minor differences arise (see Figure 2.2). Overall, the mean as well as the median of the non-discounted RRs follow a decreasing course.

Figure 2.9: Relation of DRT 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.

2.5.1 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

$$f_T^I(t) = \lambda^I \exp(-\lambda^I t)\,, \qquad t \geq 0\,. \tag{2.7}$$

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,

[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.

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

$$f^{II}_{T,\tilde{t}}(t) = \lambda^{II}(\tilde{t})\ \exp(-\lambda^{II}(\tilde{t})\ t)\,, \qquad t \geq 0\,. \tag{2.8}$$

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

$$f^{III}_{T,\tilde{t}|U=u}(t) = \lambda^{III}(\tilde{t}, u)\ \exp(-\lambda^{III}(\tilde{t}, u)\ t)\,, \qquad t \geq 0\,. \tag{2.9}$$

The unconditional probability density function can be derived by the integral of the joint probability density function over the frailty realizations u

$$f^{III}_{T,\tilde{t}}(t) = \int_{-\infty}^{+\infty} f^{III}_{T,\tilde{t}|U=u}(t)\ f_U(u)\ du\,, \qquad t \geq 0\,, \tag{2.10}$$

where $f_U(u)$ is the density of the Normal distribution with mean 0 and variance σ^2 (see Equation (2.4)). Equation (2.10) can be solved by numerical integration.

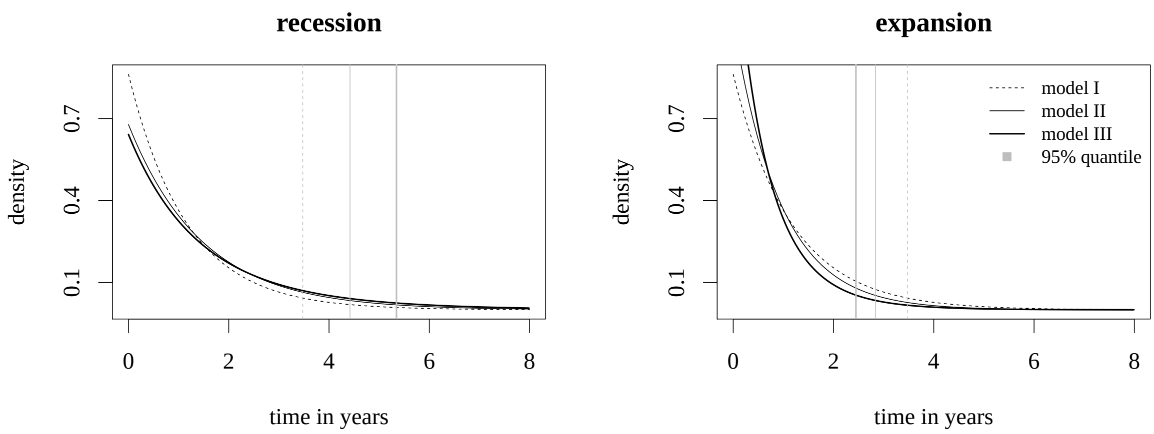
As the baseline hazard rate λ_0 directly impacts the distribution of DRTs and, thus, its mean, we calibrate it on the average DRT of 1.59 years (see Table 2.1). This ensures an average simulated portfolio DRT in accordance with the empirical data. Thus, the average portfolio DRT corresponds to 1.59 years 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 $\lambda^{I}_0 = 1.08$ as well as $\lambda^{II}_0 = 0.12$ for Model II and $\lambda^{III}_0 = 0.07$ for Model III.[22]

Figure 2.10 shows the probability density functions of the DRT in Model I, II, and III as of Equation (2.7), (2.8), and (2.10) for an exemplary recession and expansion period.[23] The left panel of Figure 2.10 displays the probability density functions for a recession period. The

[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.

Figure 2.10: Density of DRT



Notes: The figure illustrates the probability density function of the DRT for Model I, II, and III according to Equation (2.7), (2.8), and (2.10) 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.

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 2.10 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 2.9 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.

Table 2.9: 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 (2.7), (2.8), and (2.10) 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 2.10.

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.

2.5.2 Implications on portfolio level

Systematic effects in modeling DRT might not only affect the DRT itself, but also the loss involved. In Section 2.2 (see Figure 2.2), 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

$$\mathrm{LGD} = 1 - \frac{RR_T}{(1+r)^T} \,, \tag{2.11}$$

where RR_T denotes the time dependent non-discounted RR. Its value is set to the mean of the related DRT bucket. Table 2.10 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 2.10: 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 2.2.

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

[24] Figure 2.9 shows that this relation also holds for the representative portfolio.

to Model I, II, and III, respectively.[25] The corresponding LGDs are calculated by Equation (2.11). Finally, the portfolio loss is given by:

$$LGD_{PF} = \frac{1}{n\,\overline{EAD}} \sum_{i=1}^{n} (LGD_i\ EAD_i) \ , \tag{2.12}$$

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

Table 2.11: 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

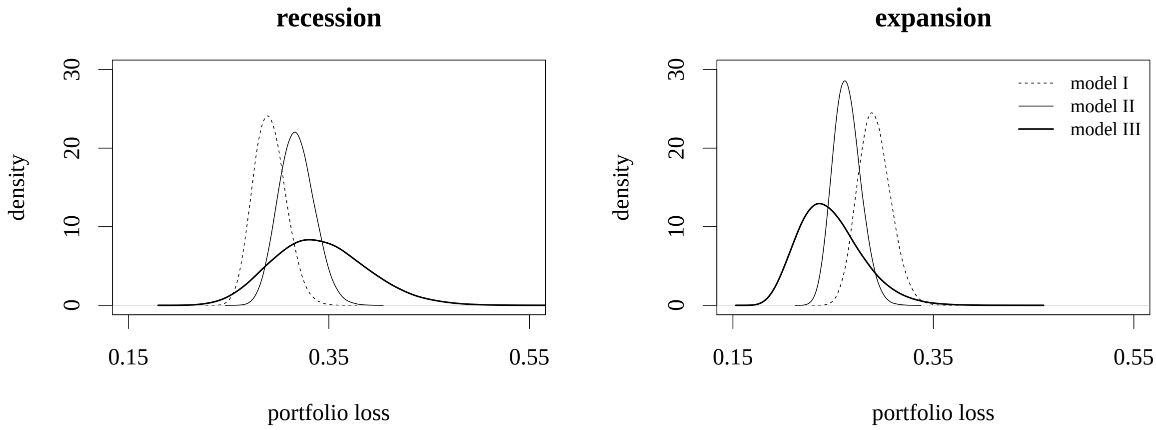
Notes: The table summarizes the mean and 95% quantile of the portfolio DRT for Model I, II, and III in an exemplary recession (realizations of macroeconomic variables as of Q1 2009) and expansion (realizations of macroeconomic variables as of Q2 2011) period. For every loan in the representative portfolio 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.

Table 2.11 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 2.1), 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.

The portfolio loss distribution is simulated based on the portfolio DRTs and the non-discounted RRs as of Table 2.10. Figure 2.11 displays the portfolio loss distributions for 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

[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., $T, \tilde{t} \mid U{=}u \sim \text{Exp}(\lambda^{III}(\tilde{t}, u))$.

Figure 2.11: Kernel density estimates of loss on portfolio level



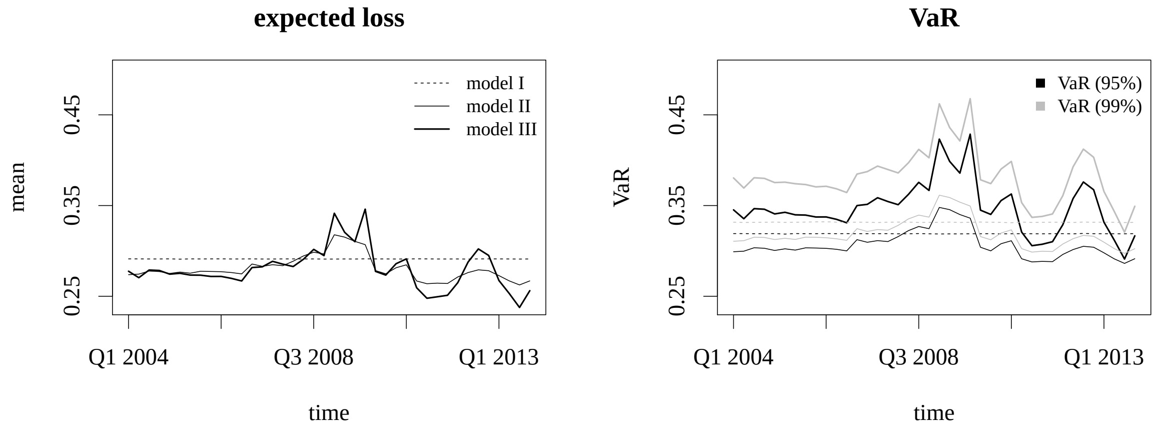
Notes: The figure illustrates the kernel density estimates of the exposure weighted portfolio loss distribution based on simulated DRT of Model I, II, and III as of Equation (2.7), (2.8), and (2.9) in an exemplary recession (realizations of macroeconomic variables as of Q1 2009) and expansion (realizations 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 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 2.11, the portfolio loss distribution of an expansion is displayed. The portfolio loss distribution of Model II 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 2.12 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 2.12 shows the VaR(95%) and the VaR(99%) of the portfolio loss distribution. As the resolution intensity and, thus, the

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

Figure 2.12: 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 (2.7), (2.8), and (2.9) 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.

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.

2.6 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.

2.A Appendix | 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 (2.1) it follows for Model I:

$$\lambda_{it} = \frac{f_T(t|x_i)}{1 - F_T(t|x_i)}, \tag{2.13}$$

($i = 1, \ldots, n$), 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

$$L(\beta|x,\delta) = \prod_{i=1}^{n} \left[f_T(t_i|x_i)^{\delta_i} (1 - F_T(t_i|x_i))^{1-\delta_i} \right], \tag{2.14}$$

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 (2.13) into Equation (2.14) yields

$$L(\beta|x,\delta) = \prod_{i=1}^{n} \left[\lambda_{it}^{\delta_i} (1 - F_T(t_i|x_i)) \right] \tag{2.15}$$

$$= \prod_{i=1}^{n} \left[(\lambda_{0t} \exp(x_i\beta))^{\delta_i} \exp\left(-\exp(x_i\beta) \int_0^{t_i} \lambda_{0v} dv \right) \right] \tag{2.16}$$

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 (2.16) to

$$L(\beta|x,\delta) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \left[\frac{\exp(x_i\beta)}{\sum_{j=i}^{n} \exp(x_j\beta)} \lambda_{0t_i} \sum_{j=i}^{n} \exp(x_j\beta) \right] \prod_{i=1}^{n} \exp\left(-\exp(x_i\beta) \int_0^{t_i} \lambda_{0v} dv \right) \tag{2.17}$$

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

$$PL(\beta|x,\delta) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \frac{\exp(x_i\beta)}{\sum_{j=i}^{n}\exp(x_j\beta)} \,. \tag{2.18}$$

Afterwards, the baseline hazard rate can be estimated by

$$\widehat{\lambda}_{0t_i} = \left(\sum_{j=i}^{n}\exp(x_j\widehat{\beta})\right)^{-1} \,. \tag{2.19}$$

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

$$CPL(\beta,\gamma,\sigma|x,\delta,U=u) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \frac{\exp(x_i\beta + z_{\tilde{t}(i)}\gamma + u_{\tilde{t}(i)})}{\sum_{j=i}^{n}\exp(x_j\beta + z_{\tilde{t}(j)}\gamma + u_{\tilde{t}(i)})} \,, \tag{2.20}$$

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:

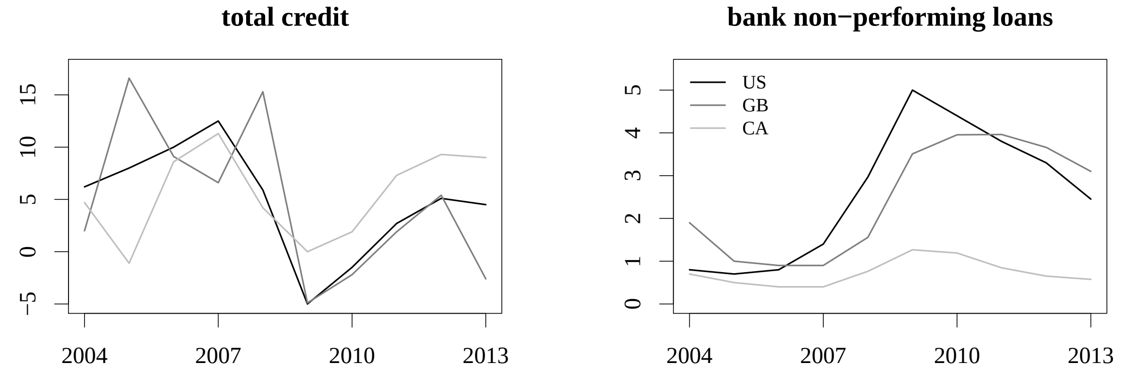
$$UPL(\beta,\gamma,\sigma|x,\delta) = \prod_{\substack{i=1 \\ \delta_i=1}}^{n} \int_{-\infty}^{\infty} \frac{\exp(x_i\beta + z_{\tilde{t}(i)}\gamma + u)}{\sum_{j=i}^{n}\exp(x_j\beta + z_{\tilde{t}(j)}\gamma + u)} \, f_{U_{\tilde{t}(i)}}(u)\, du \,. \tag{2.21}$$

Equation (2.21) is maximized to estimate the unknown model parameters β,γ,σ. Afterwards, we use Equation (2.20) 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 (2.19) extended by the frailty term.

2.B Appendix | Further analyses

Credit dynamics in the US, Great Britain and Canada

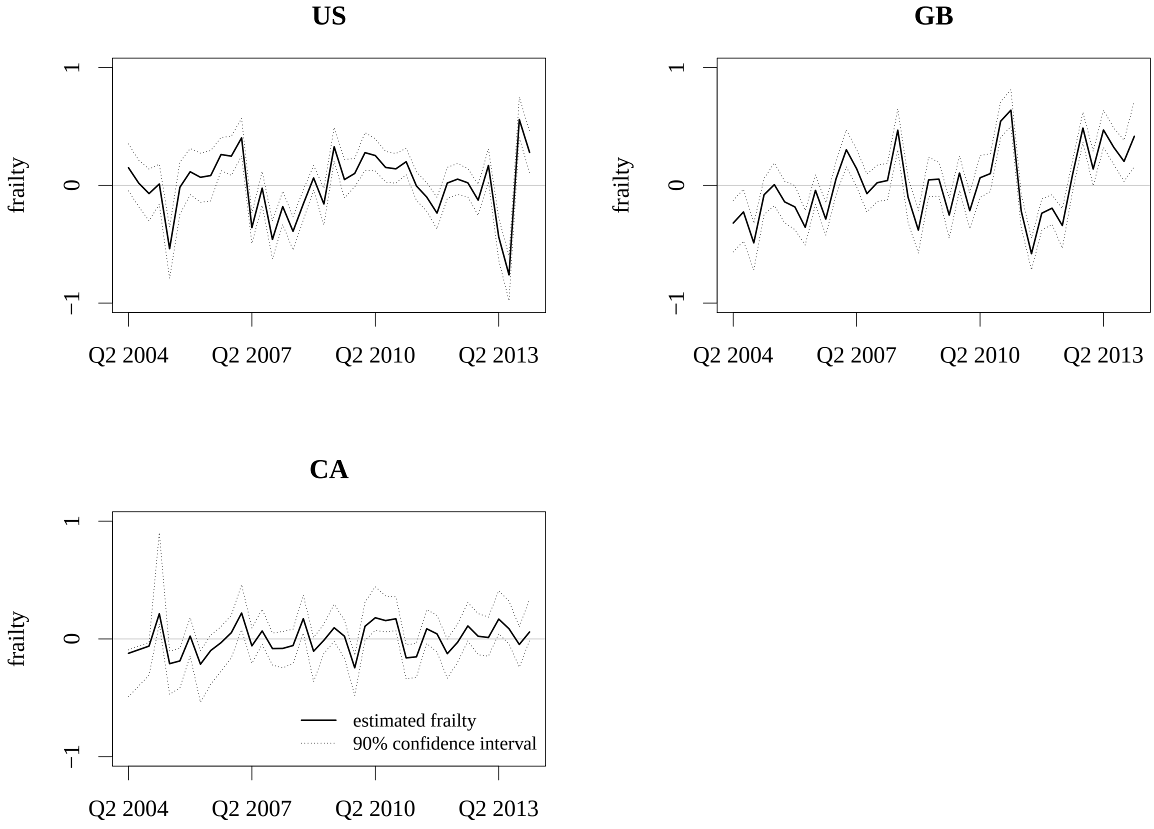
Figure 2.B.1: Behavior of loan portfolios and default incidence



Notes: The figure shows the time-series of the year-on-year change in total credit volume to non-financial corporations and the ratio of bank non-performing loans to gross loans. All data is taken from the public research data base of `https://fred.stlouisfed.org/`. The time series are indexed by CRDQUSANABIS_PC1, CRDQGBANABIS_PC1, CRDQCAANABIS_PC1, DDSI02USA156NWDB, DDSI02GBA156NWDB and DDSI02CAA156NWDB.

Covariate influences over time

Figure 2.B.2: 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 dotted 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 2.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)
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 (2.5), 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 2.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).

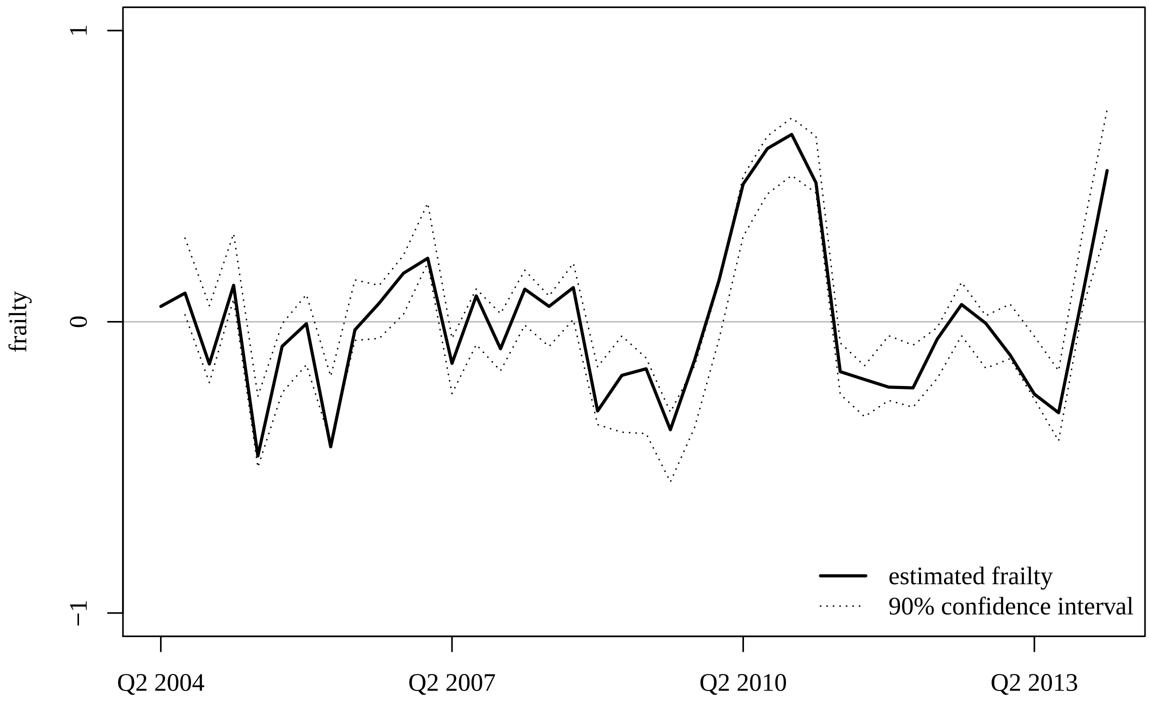
Table 2.B.2: Regression results for Model III when including interactions of covariates and recessions (interactions to recessions)

		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 (Medium term)	Short term	-0.1403		(0.0964)	-0.1865	***	(0.0667)	0.0409		(0.1031)
	Other	-0.1345	*	(0.0794)	-0.0327		(0.1180)	-0.0081		(0.1379)
Seniority (Pari-passu)	Super senior	0.2144	*	(0.1177)	-0.2620	***	(0.0892)	-0.2454		(0.2318)
	Non senior	0.5222		(0.3691)	-0.3479		(0.4836)			
	Unknown	0.8804	***	(0.2261)				-0.1100		(0.2783)
Nature of default (90 days past due)	Unlikely to pay	0.1933	*	(0.1043)	0.0930		(0.1138)	0.0293		(0.8743)
	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 (NO)	Unknown	-0.0180		(0.5018)						
	YES	-0.2070	***	(0.0691)	0.1204	*	(0.0691)	0.2990	*	(0.1683)
Collateral (NO)	Other collateral	0.0535		(0.0867)	-0.0915		(0.0833)	-0.4795	**	(0.2021)
	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)
Industry (Finance, insurance, RE)	Agric., forestry, fishing	0.8419	**	(0.3644)	-0.6409	***	(0.2193)	0.1436		(0.2407)
	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 (2.5), 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 2.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).

Joint country regression

Figure 2.B.3: 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 dotted 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 2.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)
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 (2.2) (Model I), (2.3) (Model II) resp. (2.5) (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.

Regression results for additional macroeconomic variables

Table 2.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 (2.3), 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 3

Systematic effects among loss given defaults and their implications on downturn estimation

This chapter is joint work with Ralf Kellner* and Daniel Rösch† published as:

Betz, J., R. Kellner, D. Rösch (2018). Systematic effects among loss given defaults and their implications on downturn estimation. ***European Journal of Operational Research*****, forthcoming** (available online since June, 2018).

https://doi.org/10.1016/j.ejor.2018.05.059

Abstract

Banks are obliged to provide downturn estimates for loss given defaults (LGDs) in the internal ratings-based approach. While downturn conditions are characterized by systematically higher LGDs, it is unclear which factors may best capture these conditions. As LGDs depend on recovery payments which are collected during varying economic conditions in the resolution process, it is challenging to identify suitable economic variables. Using a Bayesian Finite Mixture Model, we adapt random effects to measure economic conditions and to generate downturn estimates. We find that systematic effects vary among regions, i.e., the US and Europe, and strongly deviate from the economic cycle. Our approach offers straightforward supportive tools for decision makers. Risk managers are enabled to select their individual margin of conservatism based on their portfolios, while regulators might set a lower bound to guarantee conservatism. In comparison to other approaches, our proposal appears to be conservative enough during downturn conditions and inhibits over-conservatism.

Keywords: risk management; bank loans; credit risk; random effects

JEL classification: C23, G21, G33

* University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: ralf.kellner@ur.de.

† University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: daniel.roesch@ur.de.

3.1 Introduction

Banks are obliged to provide conservative estimates for the probability of default (PD) and the loss given default (LGD) in the advanced internal ratings-based (IRB) approach of the Basel regulations (see Basel Committee on Banking Supervision, 2006). While through-the-cycle, i.e., average, PDs are translated into (conservative) conditional PDs by a supervisory mapping function, LGDs are required to reflect conditions of economic downturn. Economic variables might, thus, be a natural choice to identify downturns and to generate consistent downturn estimates. Recently, the European Banking Authority (EBA) published new technical standards (see European Banking Authority, 2017) which emphasize the use of economic or credit risk factors. However, the identification of economic variables seems to be ambiguous in the literature.[1] This might be due to the nature of *workout* LGDs. Recovery payments are collected during the whole resolution process which usually takes multiple years ahead of the default event. Thus, economic or credit risk factors at a specific point in time (e.g., the default time) may not be able to represent the systematic patterns inherent in LGDs.

Although the identification of suitable economic variables seems to be ambiguous, financial institutions are confronted with the need to generate consistent downturn estimates for loss rates and, thus, identify and measure downturns. Following the literature, this may be a challenging task. Economic variables are either not significant or deliver limited explanatory power. To resolve this, we suggest to use time-specific random effects to measure the systematic movement inherent in LGDs over time. Consistent downturn estimates can be generated by considering a conservative quantile of the random effect. This entails a comprehensive supportive tool for decision makers. On the one side, risk managers are enabled to determine the margin of conservatism by selecting an according quantile based on the characteristics of their loan portfolio. On the other side, regulators might set a lower bound to this quantile to guarantee overall conservatism and, thus, the stability of the financial system during crises. Furthermore, our approach bypasses identification issues that may occur to decision makers in terms of selecting variables for LGD modeling among a set of observable variables.

This paper contributes to the literature in multiple ways. First, we use access to the unique loss database of Global Credit Data (GCD)[2] to reveal the systematic nature inherent in LGDs.

[1] See Section 3.2 for a comprehensive literature review.

[2] GCD is a non profit initiative which aims to support banks to measure their credit risk by collecting and analyzing historical loss data. See `http://www.globalcreditdata.org/` for further information. Currently, 52 member banks from all around the world share their loss information.

Thereby, we find considerable differences among regions, i.e., the US and Europe. Second, we show that random effects strongly deviate from the economic cycle measured by common economic variables. We compare the estimated random effects to macro variables common in the LGD literature. Thereby, not only descriptive differences come to light. The impact of macro variables seems not to be evident or limited regarding its magnitude. Third, we suggest a methodology to generate consistent downturn estimates and compare it to a variety of alternatives.

The remainder of this paper is structured as follows. Section 3.2 provides a literature overview, while general background information to the topic of downturn LGD modeling is presented in Section 3.3. Methods used in this paper are explained in Section 3.4. The data description and the empirical results are presented in Section 3.5. Results with respect to downturn LGDs are discussed in Section 3.6 and compared to other approaches in this field. Section 3.7 concludes.

3.2 Literature review

The literature regarding loss rates can be divided into two streams. The first one focuses on separate methods, whereby, the LGD is the only dependent variable. The second one applies joint modeling approaches for the PD and the LGD to consider the dependence structure between the two central credit risk parameters.

Table 3.D.1 in Appendix 3.D summarizes the first stream of literature, i.e., separate methods, focusing systematic effects. Typically, macro variables are included to display synchronism in time line. However, some authors completely renounce macro variables in their analysis (see, e.g., Bastos, 2010; Bijak and Thomas, 2015; Calabrese, 2014; Gürtler and Hibbeln, 2013; Matuszyk et al., 2010; Somers and Whittaker, 2007). Others examine univariate significance which (partly) disappears in a multivariate context (see, e.g., Acharya et al., 2007; Brumma et al., 2014; Caselli et al., 2008; Dermine and Neto de Carvalho, 2006; Grunert and Weber, 2009). Acharya et al. (2007) find statistical significance for industry distress dummies but not for continuous variables. They trace this to non linear impacts of macro variables. Krüger and Rösch (2017) adapt quantile regression and report statistical significance of macro variables for the inner quantiles. Again, this can be interpreted as an indication for non linear impacts. In some papers, statistical significance or evidence is not reported (see, e.g., Altman and Kalotay, 2014; Tobback et al., 2014; Yao et al., 2015). Other authors state statistical significance for

a variety of macro variables. However, they apply data sets of bonds (see Jankowitsch et al., 2014; Nazemi et al., 2017; Qi and Zhao, 2011), credit cards (see Bellotti and Crook, 2012; Yao et al., 2017), or mortgages (see Leow et al., 2014; Qi and Yang, 2009). Bonds are typically characterized by *market-based* LGDs. This simplifies the identification of significant macro variables as the time a bond spends in default is standardized to 30 days, thus, the final LGD is certain shortly after default and there is no additional uncertainty regarding timing of cash flows. Credit cards and mortgages are among bulk businesses of banks. Resolution processes might be more standardized and related to less uncertainty compared to corporate loans. This also may simplify the identification of significant macro variables.

The second stream of literature models PDs and LGDs jointly and introduces time-specific systematic random effects. PD and LGD models are linked by joint random variables to measure correlation structures among the risk parameters. Bruche and Gonzalez-Aguado (2010) implements a binary variable indicating the state of the cycle. Chava et al. (2011) use a frailty in a hazard-type PD model, however, show that it has no significant impact on the recovery rate. Bade et al. (2011), Rösch and Scheule (2010), and Rösch and Scheule (2014) implement correlated random effects in both models. Thereby, Rösch and Scheule (2014) present a closed form expression for generating downturn LGD estimations which are based on an adverse realization of the random effect using a simple and parsimonious Merton-type model. Keijsersy et al. (2017) adopts random effects that are modeled by a VAR process to link default, loss, and the economic environment. The closest studies to ours are Keijsersy et al. (2017) and Rösch and Scheule (2014). In comparison to Keijsersy et al. (2017), we compare two model specifications for the random effect in order to analyze if cyclical behavior among LGDs is present at all. Due to their model specification, such a distinction is not possible. Furthermore, our analysis enables us to identify regional differences in systematic effects among LGDs. In addition, we provide a downturn LGD method and compare it to others. Their approach is focused on a comprehensive risk framework that models all credit risk components at once. Thus, their insights seem to be more relevant for internal modeling under Pillar II of the Basel regulations, while our approach is strongly driven towards Pillar I which combines credit risk components by a standardized formula. Rösch and Scheule (2014) use a random effect to capture systematic movements that impact PDs and LGDs. As stated by the EBA, downturn LGDs may additionally include systematic effects that are independent of those for PDs. Furthermore, their study is based on *market-based* LGDs. As explained above, their nature strongly differs from *workout* LGDs which is why inferences with respect to the systematic impact are not comparable. Moreover, they do not present results regarding the progress of systematic effects over time which inhibits inferences regarding similarities to the economic cycle.

Summarizing, the identification of systematic effects among LGDs seems to be not trivial in a LGD modeling context and results regarding macro variables are ambiguous in the literature. This applies in a special manner to corporate loan data sets which are characterized by *workout* LGDs and non standardized and, thus, complex resolutions.

3.3 Background

In general, it is a challenging task to model LGDs. This is mainly due to the complex form of the LGD distribution which can exhibit skewness and multi-modality. *Market-based* LGDs are typically bounded on the interval [0,1] and have less extreme modes, while *workout* LGDs can exhibit realizations outside this interval and are characterized by high probability masses at zero and one. LGDs smaller zero might occur if recovery payments exceed the original outstanding amount. This is the case if additional fees are demanded. LGDs greater than one are conceivable if banks are confronted with additional costs, e.g., if courts are involved in the resolution process. Hence, different approaches for modeling the LGD distribution exist. Usually explanatory variables are included as characteristics such as protection, industry affiliation, etc., which may impact LGDs. It is crucial to combine these explanatory variables with a distributional assumption that provides great flexibility. Such an approach is presented in Altman and Kalotay (2014) upon which we build up our model. Altman and Kalotay (2014) estimate a finite mixture model with four components based on *market-based* LGDs. Hence, they transform LGDs via the quantile function of the normal distribution and estimate the model based on the transformed values.

We adapt this approach to consider characteristics of *workout* LGDs, e.g., high probability masses at zero and one and values outside the unit interval. We estimate a mixture of five normally distributed components based on the untransformed scale, and hereby, fix the two outer components to identify loans with no (LGD = 0) and total (LGD = 1) loss. The parameters of the remaining components are estimated within the modeling framework. Component probabilities are derived with an ordered regression model. By this means, explanatory variables are included. Besides the adaption for *workout* LGDs, we extend the model presented in Altman and Kalotay (2014) by a random effect. This is an essential extension with respect to the analysis of systematic effects among LGDs. During poor (good) economic conditions, LGDs tend to be higher (lower) on average. Random effects can capture these time trends which cannot be covered by loan characteristics and macro variables alone.

Systematic effects are crucial in the light of downturn LGD estimation. Following Basel Committee on Banking Supervision (2006) and Basel Committee on Banking Supervision (2005), financial institutions have to provide estimates, that "[...] reflect economic downturn conditions where necessary to capture the relevant risks [...]". First, this ensures conservativeness and, thus, prevents LGD predictions being systematically too low. Second, it seems to be driven by the need for counter-cyclical safety buffers. These buffers are accumulated in favorable economic conditions and absorb unexpected economic losses during downturns. Hence, downturn LGDs should provide safety buffers that are (too) conservative during good and normal conditions, but are sufficient once critical conditions emerge. A consistent approach to derive downturn estimates does not exist so far. Different suggestions are even presented in more recent publications (see, e.g. Tobback et al., 2014; Calabrese, 2014; Bijak and Thomas, 2015). We will compare these approaches to our suggestion in Section 3.6.

3.4 Methods

We use a hierarchical model combining a Finite Mixture Model (FMM) with a probabilistic substructure (see Altman and Kalotay, 2014). In FMMs, the dependent variable is assumed to be divided into a finite number of latent classes. In these, the dependent variable follows a specific distribution, e.g., Normal distribution with parameters depending on the latent class. We refer to the FMM as the *component model*. The probability of belonging to a latent class is modeled by an Ordered Logit (OL) model. We refer to the probabilistic substructure in form of the OL model as the *probability model*.

Component model | FMM

We apply FMM to model the distribution of LGDs due to their flexibility in modeling distributions of unknown shape (see, e.g., McLachlan and Peel, 2000). As stated earlier (see Section 3.1 and Section 3.3) workout LGDs are not compulsory limited to the interval $[0,1]$. Thus, common distribution applied to rates, such as the Beta distribution, are not applicable. Loan LGDs are often characterized by bindings at no (LGD = 0) and total loss (LGD = 1). Again, this is hardly applicable by, e.g., a Beta distribution. In the component model, the LGD as dependent variable Y is assumed to be divided into a finite number of K latent classes. In each class k, the probability density function (PDF) for observation y given k is $f_k(y \mid \theta_k)$, e.g., Normal density functions, with parameters θ_k depending on the latent class k. These constituent PDFs are weighted by $p_1, \ldots, p_K$ to generate the PDF of a finite mixture distribution $g(y \mid \theta_1, \ldots, \theta_K)$:

$$g(y|\theta_1, \ldots, \theta_K) = \sum_{k=1}^{K} p_k \, f_k(y|\theta_k) \,, \tag{3.1}$$

where, $f_1(y|\theta_1), \ldots, f_K(y|\theta_k)$ are the PDFs with parameters $\theta_1, \ldots, \theta_K$ of the constituent classes $k \in \{1, \ldots, K\}$. To ensure the general properties of a PDF, i.e., $g(y) \geq 0$ for all $y \in \mathbb{R}$ and $\int_{-\infty}^{\infty} g(y) = 1$, $p_k \geq 0$ and $\sum_k p_k = 1$ must hold.

In the following, we assume Gaussian FMMs. The Normal distribution is an appropriate choice due to its computational transparency (see, e.g., McLachlan and Peel, 2000). The constituent PDFs $f_k(y \mid \theta_k)$ correspond to Normal density functions and the dependent variable Y given a latent class k follows a Normal distribution with parameters μ_k and σ_k. Assuming conditional independence, the likelihood of a Gaussian FMM $\phi(Y_1, ..., Y_N \mid \mu, \sigma, p)$ is the product of the individual likelihood contributions, which arise from the above densities:

$$\phi(Y_1, ..., Y_N \mid \mu, \sigma, p) = \frac{1}{(2\pi)^{\frac{N}{2}}} \prod_{i=1}^{N} \left(\sum_{k=1}^{K} \frac{p_k}{\sigma_k} \exp\left[-\frac{(Y_i - \mu_k)^2}{2\sigma_k^2} \right] \right) , \tag{3.2}$$

where, μ is a $(1 \times K)$ vector of component means, σ is a $(1 \times K)$ vector of component standard deviations, and p is a $(1 \times K)$ vector of component weights. N is the number of observations.

As the model is estimated via Bayesian inference, posterior distributions are generated via a Markov Chain Monte Carlo (MCMC) sampler. It generates samples from the posterior distribution by constructing reversible Markov-chains. The equilibrium distribution corresponds to the target posterior distribution. The solution via an MCMC sampler is necessary due to the complexity of the model, i.e., the priors are partly non conjugate. Thus, direct sampling from posterior distributions is not possible as they do not have an analytical solution.

To adapt data augmentation in the MCMC sampler, the component weight p_k is replaced with an indicator variable e_{ik} in the likelihood specification of Equation (3.2):

$$\phi(Y_1, ..., Y_N \mid \mu, \sigma, e) = \frac{1}{(2\pi)^{\frac{N}{2}}} \prod_{i=1}^{N} \left(\sum_{k=1}^{K} \frac{e_{ik}}{\sigma_k} \exp\left[-\frac{(Y_i - \mu_k)^2}{2\sigma_k^2} \right] \right) . \tag{3.3}$$

If loan i is a random draw of component k, $e_{ik} = 1$ and zero otherwise. In each step of the MCMC sampler, loan $i \in N$ is assigned to one specific component k and, thus, follows one specific probability density distribution $f(y \mid \mu_k, \sigma_k)$. However, changes of component affiliation remain possible within a chain.

Probability model | OL model

For the probabilistic substructure of the component model we use an OL approach. Hereby, observable covariates are included. To rely on the classical formulation of OL models, we define the component affiliation z_i for each loan i:

$$z_i = k \quad \text{if } e_{ik} = 1 \,, \tag{3.4}$$

where, e_{ik} is the indicator as of Equation (3.3). The latent component variable z_i describes the affiliation to individual components k for every loan i in the data. Thus, the variable Z_i follows a categorical distribution with component probabilities p_k:

$$Z_i \sim \text{Cat}(p_i) \,, \tag{3.5}$$

where, p_i is a $(1 \times K)$ vector of component probabilities. The categorical distribution is a generalization of the Bernoulli distribution for more than two outcomes ($k > 2$) and a special case of the multinomial distribution for one trial. It is a prominent distribution for categorical data in Bayesian inference (see, e.g., Plummer, 2017).

Following Altman and Kalotay (2014), an underlying variable Z_i^* is defined which represents the true but unobservable dependent variable. This latent variable approach is formulated to include independent variables into the FMM as of Equation (3.1) and (3.3). The latent variable Z^* follows a linear model:

$$Z_i^* = x_i\beta + F_{t(i)} + \epsilon_i, \qquad \epsilon_i \sim \text{logistic}, \tag{3.6}$$

whereby, x_i is a $(1 \times J)$ vector of J standardized independent variables and β the $(J \times 1)$ vector of coefficients. The expression ϵ_i describes the error term. A random effect $F_{t(i)}$ with time stamp $t(i)$ is introduced in the modeling framework to capture systematic effects among LGDs. The time stamp $t(i)$ indicates the default time t in quarters which depends on the loan i as every loan defaulted in a specific quarter, e.g. 2007 Q4. Two loans i and i' which defaulted in the same quarter ($t(i) = t(i') = t$) are characterized by the same realization of the random effect ($f_{t(i)} = f_{t(i')} = f_t$). A positive value of F_t in a specific quarter t leads to a higher value of the latent dependent variable Z_i^* for all loans defaulted in this quarter and vice versa.

We consider two alternative specifications for the random effect. In specification I, the random effect is modeled as an i.i.d. normally distributed random variable and implemented in terms of a random intercept:

$$F_t \sim N(\alpha, \sigma^F)\,, \tag{3.7}$$

where, α is its mean and σ^F its standard deviation. In other words, the mean α is the intercept in the linear model as of Equation (3.6). The standard deviation σ^F can be interpreted as the impact of the systematic effect. A higher value of σ^F allows for more extreme realizations of the random effect and, thus, for more extreme time dependent shifts in Z_i^*. Specification I, i.e., the normal distribution, is the most common way to implement random effects (see, e.g., Rösch and Scheule, 2014).

In specification II, the random effect follows an autoregressive process of order 1, i.e., an AR(1) process, to allow for cyclical movements in the realizations of the random effect:

$$\begin{aligned} F_t &= a + \varphi F_{t-1} + \varepsilon_t \\ \mu_u^F &= \frac{a}{1-\varphi} \\ \sigma_u^F &= \frac{\sigma_c^F}{\sqrt{1-\varphi^2}}\,, \end{aligned} \tag{3.8}$$

where, a is the constant and φ the parameter of the AR(1) process. The unconditional mean and standard deviation are denoted by μ_u^F and σ_u^F. The conditional standard deviation σ_c^F corresponds to the standard deviation of the errors ε_t. The stationary condition of an AR(1) process is $|\varphi| < 1$. Duffie et al. (2009) apply an Ornstein-Uhlenbeck process to a frailty in a default model. This might be interpreted as time continuous AR(1) process.[3]

The component affiliations Z_i are determined by the location of the latent variable Z_i^* as of Equation (3.6) to corresponding cut points c_k:

$$Z_i = \begin{cases} 1 & \text{if } Z_i^* \leq c_1 \\ 2 & \text{if } c_1 < Z_i^* \leq c_2 \\ \vdots & \\ K & \text{if } c_{K-1} < Z_i^*\,. \end{cases} \tag{3.9}$$

Thus, loan i is assigned to component 1 ($Z_i = 1$) if Z_i^* is smaller or equal than c_1. If Z_i^* lies in

[3] Frailties are random effects in survival models.

between c_1 and c_2, loan i is assigned to component 2 ($Z_i = 2$) and so on. Finally, component K ($Z_i = K$) is assigned if Z_i^* is greater than c_{K-1}. Generally, there are $K-1$ cut points to estimate within the OL model. Consider loan i and i' which defaulted in the same quarter ($t(i) = t(i') = t$) and, thus, share the same realization of the random effect ($f_{t(i)} = f_{t(i')} = f_t$). For rather high (low) values of F_t the component allocation is shifted towards higher (lower) components for all loans defaulted in the corresponding quarter t.

The corresponding component probabilities as of Equation (3.5) can be derived by the cumulative distribution function of the Logistic distribution:

$$\begin{aligned}
\mathbb{P}\left(Z_i = 1 \mid x_i, f_{t(i)}\right) &= \mathbb{P}\left(Z_i^* \leq c_1 \mid x_i, f_{t(i)}\right) \\
&= \frac{1}{1+\exp\left[-\left(c_1 - Z_i^*\right)\right]} \\
\mathbb{P}\left(Z_i = 2 \mid x_i, f_{t(i)}\right) &= \mathbb{P}\left(Z_i^* \leq c_2 \mid x_i, f_{t(i)}\right) - \mathbb{P}\left(Z_i^* \leq c_1 \mid x_i, f_{t(i)}\right) \\
&= \frac{1}{1+\exp\left[-\left(c_2 - Z_i^*\right)\right]} - \frac{1}{1+\exp\left[-\left(c_1 - Z_i^*\right)\right]} \\
&\;\vdots \\
\mathbb{P}\left(Z_i = K \mid x_i, f_{t(i)}\right) &= \mathbb{P}\left(Z_i^* > c_{K-1} \mid x_i, f_{t(i)}\right) \\
&= 1 - \frac{1}{1+\exp\left[-\left(c_{K-1} - Z_i^*\right)\right]}\,.
\end{aligned} \tag{3.10}$$

To guarantee the non negativity of component probabilies p_{ik} ($p_{ik} \geq 0$ for $k \in \{1, \ldots, K\}$), $c_1 \leq c_2 \leq \cdots \leq c_{K-1}$ must hold. In analogy to the component allocation in Equation (3.9), the random effect as of Equation (3.6), (3.7), and (3.8) introduces systematic movement into the component probabilities. Again, consider loan i and i' which defaulted in the same quarter ($t(i) = t(i') = t$) and, thus, share the same realization of the random effect ($f_{t(i)} = f_{t(i')} = f_t$). For high values of F_t ($F_t > \alpha$ in specification I and $F_t > \mu_u^F$ in specification II), the probability of the first component decreases as Z_i^* increases. Simultaneously, the probability of the K-th component increases. The probabilities of the remaining components might be affected to a minor extent. Summarizing, a high realization of the random effect leads to a systematically lower probability for the first component and higher probability for the last component while a low realization of the random effect implies the opposite effect. Loans defaulted in the same quarter are, thus, characterized by systematically lower probabilities of the first and systematically higher probabilities of the last component or systematically higher probabilities of the first and systematically lower probabilities of the last component.

The above OL model suffers from over specification, i.e., the parameters of the model are not identified. Thus, an infinite number of solutions exists. Three solution mechanisms are common to solve this problem: (i) fixation of the variance parameter of Z_i^* and fixation of one cut point, (ii) dropping of intercept and fixation of the variance parameter of Z_i^*, and (iii) fixation of two cut points (see Jackman, 2009). All three solution mechanisms aim to fix the range of the latent variable Z_i^* and can be transfered into each other by choosing according values. We select (iii) and fix the two outer cut points c_1 and c_{K-1} as the variability of the latent variable itself allows the use of conjugate priors for the random effect in specification I. However, results are reproducible by alternative identification restrictions.

From a pure methodical perspective, the contribution of our paper lies in the adaption of the mixture model from Altman and Kalotay (2014) to *workout* LGDs, i.e., to loan data, and the inclusion of two different specifications for the random effect. The model of Altman and Kalotay (2014) is suitable for *market-based* LGDs, i.e., bond data, where, LGD values are limited to the interval [0,1]. *Workout* LGDs, however, are not compulsory restricted to this interval. Furthermore, *workout* LGDs are characterized by bindings at no (LGD = 0) end total (LGD = 1) loss. Common transformations are not applicable for these values. We consider the characteristics of *workout* LGDs and directly estimate the LGD distribution by a FMM. By this means, biased estimates are prevented which might occur if *workout* LGDs are reduced to the interval (0,1) before transformation. Furthermore, it is essential to consider the high probability masses at zero and one as these account for major parts of the data (see left panels of Figure 3.1 in Section 3.5.1). Therefore, we fix the parameters of the two outer components to identify loans with no and total loss. The flexibility of the FMM is retained for the remaining data range. These modifications lead to flexible LGD distributions which can capture specific characteristics of *workout* LGDs. The major methodical contributions refers to the inclusion of different specifications for the random effect in a complex LGD model. Thus, we are able to analyze if LGDs are shaped by systematic patterns. Furthermore, the nature of systematic patterns can be determined, i.e., if systematic effects arise independently in the time line or if they are driven by cyclicality. The most common approach in the literature is to include normally distributed random effects (see, e.g, Rösch and Scheule, 2010, 2014). Under this setting, systematic effects are random over time. Thus, current systematic trends are irrelevant for expectations in the future. By extending the random effect to an autoregressive process, cyclicality can be considered. The inference of the AR(1) parameter φ provides additional benefits in terms of the economic interpretation. If this parameter is statistically different from zero, cyclicality is present among LGDs. Thus, LGDs are persistent, i.e., high LGDs today indicate high LGDs tomorrow. In the light of the need for downturn LGD estimates, the nature of systematic patterns is of special relevance. Downturn estimations are based on individual downturn periods. Finding persistence over time implies the need for higher capital buffers not only in crises but also in subsequent time periods.

As the models are estimated via Bayesian inference, prior distributions have to be specified for every parameter of the model.[4] Detailed information can be found in Appendix 3.A. Bayesian inference bears several advantages compared to maximum likelihood estimation techniques. First, the implementation of complex models seems to be more stable. In many cases, the expectation maximization (EM) algorithm has to be applied to solve complex models in a maximum likelihood setting. However, the EM algorithm delivers only a point estimate, whereas, the posterior distribution of the parameter is generated via Bayesian inference. By this means, parameter uncertainty can be considered intuitively. This is the second benefit of Bayesian inference. Third, statistical inference becomes more intuitive as the interpretation of credibility intervals is straight forward.

Downturn estimation

As random effects capture systematic patterns, i.e., comovement in the time line, they are suited to generate downturn estimates. The random effects of specification I as of Equation (3.7) and specification II as of Equation (3.8) are introduced into the modeling framework with time stamp $t(i)$. This time stamp indicates the default time t which depends on the loan i. Thus, all loans defaulted in the same quarter ($t(i) = t(i') = t$) share the same realization of the random effect f_t. For $f_t > \alpha$ ($f_t < \alpha$), the latent variable Z^* is higher (lower) for all loans defaulted in t. Thus, probabilities for higher components are increased (decreased) and probabilities for lower components are decreased (increased) resulting in higher (lower) average LGDs in t (see Section 3.4). This mimics the time patterns observed in average LGDs (see Figure 3.C.2). Random effects are a common method to introduce comovement in the time line (see, e.g., Rösch and Scheule, 2014; Duffie et al., 2009) Hence, these unobservable variables might be an appropriate control factor for downturn estimates. Downturn LGDs are based on negative systematic conditions, i.e., on assuming a high realization f_t ($>> \alpha$ in specification I, $>> \mu_u^f$ in specification II). Thus, we do not sample the random effect in the MCMC chains as in posterior

[4] The MCMC samples are drawn via the Gibbs sampler JAGS. JAGS is *"Just Another Gibbs Sampler"* and a widely applied program for Bayesian inference using MCMC chains (see Plummer, 2017). It uses a dialect of the BUGS (*"Bayesian inference Using Gibbs Sampling"*) language to generate a model file. The model file is compiled in C++. Figure 3.C.1 in Appendix 3.C illustrates an exemplary model file.

predictive distributions, but set it to a conservative quantile for all iterations instead.[5] By doing so, we generate LGD distributions reflecting unfavorable systematic conditions.

We compare this downturn concept to others presented in the literature. Hereby, we first replace the random effect by macroeconomic variables and use a conservative quantile of these variables to generate downturn estimates.[6] Second, we adapt the concept presented by Bijak and Thomas (2015) who also apply Bayesian inference. They suggest to use a conservative quantile of the posterior predictive distribution. Third, the concept of Calabrese (2014) is considered. She proposes an upper, i.e., rather conservative, component based on a frequentistic mixture model to reproduce a downturn distribution. Forth, Board of Governors of the Federal Reserve System (2006) proposed a linear mapping function to generate downturn estimates ($\text{LGD}_{\text{downturn}}$) based on through-the-cycle estimates (LGD_{TTC}). We will refer to the mapping function as FED proposal:

$$\text{LGD}_{\text{downturn}} = 0.08 + 0.92 \cdot \text{LGD}_{\text{TTC}} \,. \tag{3.11}$$

To adapt the approach of Bijak and Thomas (2015) we use a conservative quantile of the individual posterior predictive distributions of each loan.[7] To incorporate the suggestion of Calabrese (2014) in the adapted modeling framework, we employ component 4 as downturn distribution. As Calabrese (2014) excludes real zeros and ones and estimates a mixture of Beta distributions on the remaining data, we neglect component 5. This might most likely reproduce her approach.[8] To adopt the FED proposal, we generate posterior predictive distributions based on median realizations of random effects on which we apply Equation (3.11).

3.5 Data and results

3.5.1 Data

We use access to the unique loss data base of Global Credit Data (GCD) to build the subsample adopted in the paper. The data base includes detailed loss information on transaction basis of

[5] In Section 3.6, we apply the 90% quantile and 95% quantile as conservative quantile.
[6] In Section 3.6, we apply the 99% quantile of the macro variable.
[7] We select a conservative quantile of 99%, however, alternative conservative quantiles can be applied.
[8] As an alternative, a weighted mixture of several components can be adopted, e.g., component 3 and 4.

52 member banks all around the world. The LGD is determined by:

$$\mathrm{LGD}_i = 1 - \mathrm{RR}_i, \tag{3.12}$$

where, LGD_i denotes the LGD of loan i and RR_i the corresponding recovery rate (RR). The RR is calculated as the sum over the present values of relevant transactions divided by the outstanding amount.[9]

To check for the appropriateness of data, an expanded version of the procedure as in Höcht and Zagst (2010) and Höcht et al. (2011) is applied. Two selection criteria are evolved to identify defaulted loans with an extraordinary payment structure pre- and post-resolution. The first criterion (pre-resolution criterion) is calculated as the sum of all relevant transactions including charges-offs divided by the outstanding amount at default. In the second criterion (post-resolution criterion), the sum of post-resolution payments is divided by a fictional outstanding amount at resolution. The barriers are set to $[90\%, 110\%]$ for the pre-resolution and $[-10\%, 110\%]$ for the post-resolution criterion. Loans with realizations of these selection criteria (pre- and post-resolution criterion) outside the intervals are sorted out. Höcht and Zagst (2010) and Höcht et al. (2011) use barriers of $[90\%, 105\%]$ for the pre-resolution criterion. The structure of post-resolution payments is not considered. We decided to widen the barrier interval of the pre-resolution criterion to allow for pre-resolution payments being 10% greater than the outstanding amount at default. Regarding the post-resolution criterion, post-resolution payments might exceed a fictional outstanding amount after resolution by 10%. Transactions subtracted from the outstanding amount are allowed up to a relative height of -10%. Finally, loans with abnormal low and high LGDs ($< -50\%$ and $> 150\%$) are eliminated.[10] A rather homogeneous subsample of defaulted US American and European term loans and lines to small and medium sized enterprises (SMEs) is selected. Some further adjustment mechanisms are applied. We remove loans with an EAD less than 500 USD (5.8% of subsample data) and non resolved loans (13.0% of subsample data). The latter might entail resolution bias as resolved loans in current time periods are often characterized by systematically lower LGDs compared to unresolved loans. Thus, we restrict the time period from 2006 to 2012. We further exclude loans with incomplete observations (7.1% of subsample data). A data set of 2,987 US American and 16,924 European loans including 3,958 British loans remains.

[9] A more detailed description of the LGD calculation can be found in Betz et al. (2016) and is available from the authors upon request.

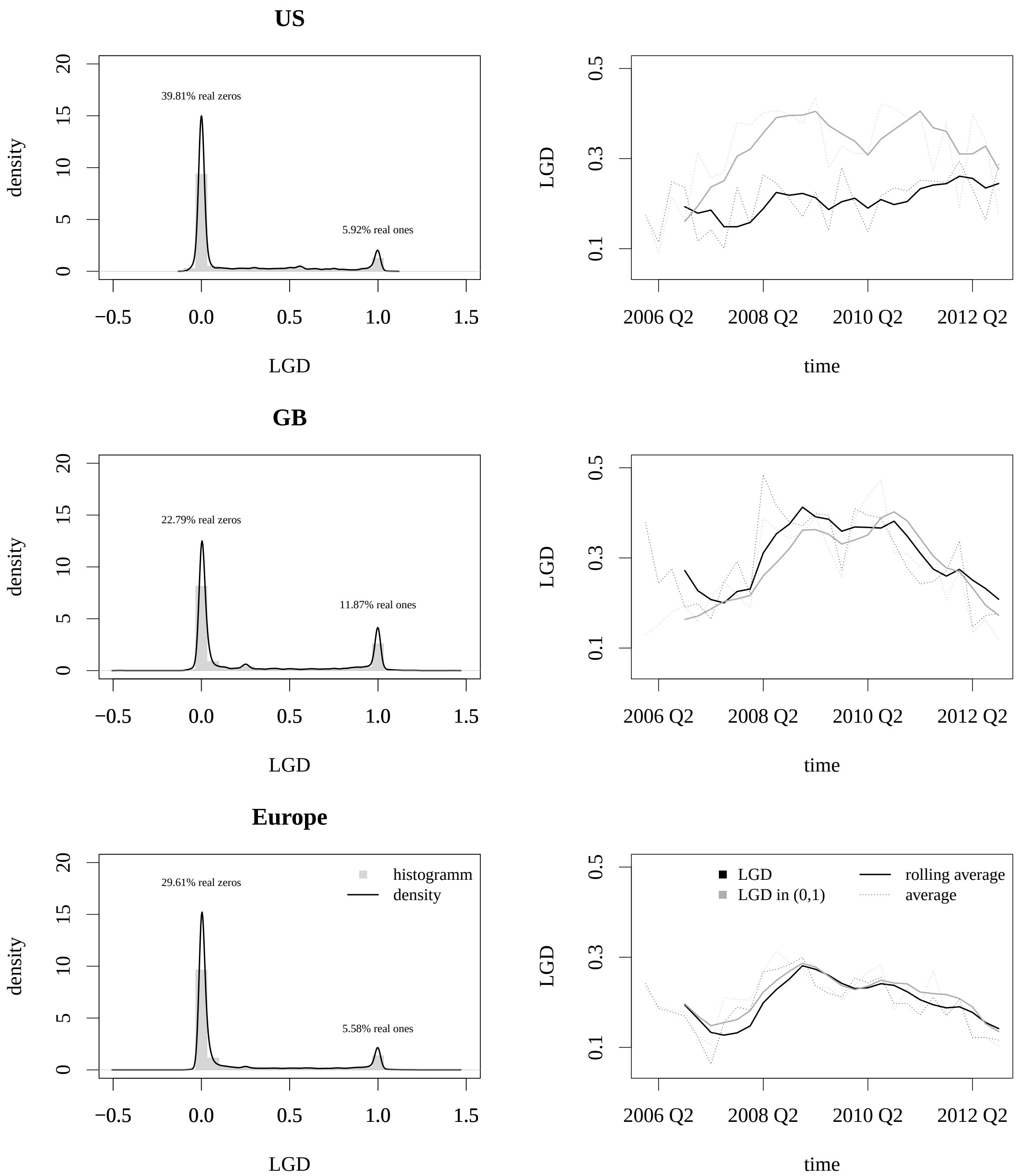
[10] Overall, 2.0% of the data is sorted out due to the pre-resolution criterion. The remaining data is reduced by 0.2% based on the post-resolution criterion. Less than 0.1% is eliminated due to abnormal LGD values.

Analyses are performed separately for the three regions – US, GB, and Europe – as fundamental differences might arise regarding the regulatory and economic setting in the US and Europe. The European sample consists of the twelve most common European countries in the data set.[11] Table 3.D.2 in Appendix 3.D summarizes the structure of the European sample. The second column includes the proportion of the corresponding country in the European sample. It might be expected that countries such as Portugal, Ireland, Latvia, and Estonia, were severely affected by the GFC compared to Norway and Poland. However, Norway and Poland account for just 5.82% of the European sample. Among the European countries, a special status is attributed to GB since it shares similarities with the US, e.g., the origin of legal systems. Hence, we focus the analysis on US, GB, and Europe.

The left panels of Figure 3.1 illustrate the distribution of LGDs for the US American, British and European data set. Histograms are presented by gray bars, the black lines show kernel density estimates. In all three regions, LGDs are characterized by a rather extraordinary distributional form. The distribution is extremely bimodal with its modi at values of 0% and 100%. The density mass in between these extremes is rather low. Unlike *market-based* LGDs, the range of *workout* LGDs is not compulsory restricted to $[0\%, 100\%]$ as workout resolutions can recover more than the EAD (LGD < 0%) due to additional charges or cause costs despite failed workout resolution efforts (LGD > 100%) on the basis of additional expenses. Furthermore, the distribution is shaped by ties. Depending on the data set, we observe between 23% and 40% of LGDs that are exactly zero and between 6% and 11% that are exactly one. These characteristics complicate modeling approaches as values outside the range]0%,100%[hinder common transformations and ties are challenging to estimate by one-stage models. An adequate modeling framework which considers common characteristics of LGDs is crucial for estimation and prediction intentions. The right panels of Figure 3.1 display the average LGDs in time line. The black lines represent averages for all LGD observations, whereas, LGDs are limited to the interval $(0,1)$, i.e., zeros and ones are excluded, for the gray lines. Considerable deviations between the US and Europe arise. While average LGDs exhibit the maximum during the GFC and slowly decrease to pre-crisis levels afterwards in Europe, average loss rates remain on high levels in the US post-crisis and even further increase in 2011 and 2012. This is even more pronounced considering averages of the limited LGDs (LGD $\in (0,1)$). Furthermore, the limited averages are shifted upwards in the US. This is due to a higher proportion of loans with no loss

[11] The European sample consists of Germany, Great Britain, Portugal, Ireland, Denmark, Norway, Sweden, Finland, Latvia, Estonia, France, and Poland. We tested an alternative European sample including Great Britain, Portugal, Ireland, Denmark, Norway, and Sweden. Results are similar and may be available upon request.

Figure 3.1: Histogram and time patterns of LGDs



Notes: The left panels of the figure illustrate the distribution of LGDs in the data divided by region. The histograms are represented by gray bars, the black lines show the kernel density estimates. Furthermore, the proportions of values which are exactly zero and one are stated above the modi of the distribution. The right panels of the figure illustrate the time patterns of average LGDs divided by region. The black dotted lines display the average LGD in the quarters of default, whereas, the black solid line show the rolling average. The corresponding values for LGDs $\in (0,1)$ are displayed in gray.

(LGD = 0).[12] The right panels of Figure 3.1 seem to be a first indication that systematic patterns regarding LGDs differ between the US and Europe, whereas, the systematic patterns of GB seem to be similar to continental Europe.

Table 3.D.3 in Appendix 3.D summarizes descriptive statistics of the dependent and independent variables divided by region. Mean, median, and standard deviation are presented for metric quantities. For variables of categoric nature, proportions of categories are stated. The extraordinary distributional form of LGDs is reflected in its descriptive statistics. While the means amount to round about 20% to 30% depending on the region, medians are close to 0%. In addition, the standard deviation is relatively large compared to the maximal range $[-50\%, 150\%]$. The exposure at default (EAD) is stated in USD and restricted to values > 0. Thus, its distribution is skewed to the right (median < mean). In the affiliating analysis, we implement the standardized logarithm of the EAD. Facility types are nearly balanced in the data sets, i.e., half of all loans are lines. A majority of loans (70% to 84%, depending on region) is protected by either collateral or guarantee or both. In total, 9% to 15% of loans are granted to corporations of FIRE (finance, insurance, and real estate) affiliation – the rest to other industry sectors.

To capture the systematic effect among LGDs over time, we use a stochastic latent variable.[13] However, in the literature, it is often controlled for observable macro variables when analyzing systematic effects. This is why we include five macro variables for comparison.[14] We consider the year-on-year (yoy) percentage change of the seasonally adjusted Gross Domestic Product (ΔGDP) to examine the influence of the real economy on LGDs. For Europe, we use the weighted average of country specific GDPs. The quarterly average of yoy log returns of major stock indices (ΔEI) and the level of the volatility index (VIX) are applied to investigate impacts of the financial economy. For the US, we use the S&P 500 and the CBOE Volatility Index. The FTSE and the VSTOXX Volatility Index are applied for GB. Weighted average equity returns and VSTOXX Volatility Index are adopted for Europe. Further common determinants in credit risk are the yoy percentage change of house prices indices (ΔHPI) and the ratio of non-performing loans (NPL ratio). As proxy for house prices, we use real residential property prices. For Europe, the weighted average is adopted. The ratio of non-performing loans is non-performing total loans to total loans for the US and bank non-performing loans to gross loans for GB and Europe

[12] We would like to thank an anonymous referee for the remark on presenting the time patterns for LGDs $\in (0,1)$.

[13] To capture systematic effects, Chava et al. (2011) use a frailty in a hazard-type PD model. Rösch and Scheule (2014) implement correlated random effects in a combined PD and LGD model.

[14] At least one of these macro variables is also considered in Acharya et al. (2007), Altman and Kalotay (2014), Brumma et al. (2014), Caselli et al. (2008), Dermine and Neto de Carvalho (2006), Grunert and Weber (2009), Jankowitsch et al. (2014), Krüger and Rösch (2017), Qi and Yang (2009), Qi and Zhao (2011), and Tobback et al. (2014).

(weighted average). The first data type is on quarterly basis, however, not available for European countries. The second one is on yearly basis.[15]

3.5.2 Results

In this section, we state the results of the model presented in Section 3.4. Subsequently, we focus on systematic effects among LGDs in the US American, British, and European data set by analyzing the implied realizations of the random effects within the modeling approach.

Component model

In the component model, i.e., the Gaussian FMM, component 1 and 5 are fixed to zero ($\mu_1 = 0$) and one ($\mu_5 = 1$), respectively, in order to capture the number of LGDs which are exactly zero and one (see Section 3.3). Thus, standard deviations of these components are set to reasonably small values ($\sigma_1 = \sigma_5 = 0.001$). Three further components are estimated within the modeling approach. Posterior means and common Bayesian figures – i.e., highest posterior density intervals (HPDIs), posterior odds ratios[16], naive and time-series standard errors – for the parameters of the component model are shown in Table 3.1. In specification I, the random effect is normally distributed as in Equation (3.7), whereas, it follows an AR(1) process as in Equation (3.8) in specification II. However, the specification of the random effect does not impact the parameter estimates of the component model. We, thus, present only one specification for the US, GB, and Europe.[17]

In the US, components 2 and 4 are centered close to zero and one, respectively. According to their comparably small standard deviations, they cover LGD values around the fixed components 1 and 5. Component 3 exhibits a mean of round about 0.44 and a comparably high standard deviation ($\sigma_3 \approx 0.25$). Thus, it contains LGD values in between the extremes zero and one. Comparing the results of GB with the US, components are slightly shifted. Components 2 and 4 are not as close to the extremes. Component 3 is with a mean of round about 0.24 shifted towards the lower end of the value range. However, it exhibits still the highest standard deviation ($\sigma_3 \approx 0.18$) and, thus, covers large parts of the area in between the extremes. In Europe, results are more similar to GB than to the US as components 2 and 4 are not as close to the

[15] The variables ΔGDP, ΔHPI, and NPL ratio originate from Federal Reserve Economic Data (FRED, see `https://fred.stlouisfed.org/`), whereas, the variable ΔEI and VIX steams from Thomson Reuter's EIKON.

[16] See results of the probability model for a detailed explanation on how to interpret the posterior odds.

[17] Due to statistical evidence of the AR(1) parameter, we apply specification I for the US and specification II for GB and Europe (see results of the probability model for detailed information). Remaining results are available from the authors upon request.

Table 3.1: Results of component model

	posterior mean	HPDI (90%)		posterior odds	naive standard error	time-series standard error
				US | specification I		
μ_1	0.0000			*not estimated*		
μ_2	0.0027	0.0010	0.0045	174.4386	0.0000	0.0000
μ_3	0.4388	0.4199	0.4589	∞	0.0001	0.0001
μ_4	0.9657	0.9589	0.9725	∞	0.0000	0.0000
μ_5	1.0000			*not estimated*		
σ_1	0.0010			*not estimated*		
σ_2	0.0245	0.0229	0.0261	∞	0.0000	0.0000
σ_3	0.2542	0.2405	0.2682	∞	0.0001	0.0001
σ_4	0.0306	0.0244	0.0365	∞	0.0000	0.0000
σ_5	0.0010			*not estimated*		
				GB | specification II		
μ_1	0.0000			*not estimated*		
μ_2	0.0153	0.0144	0.0162	∞	0.0000	0.0000
μ_3	0.2427	0.2164	0.2655	∞	0.0001	0.0002
μ_4	0.9039	0.8867	0.9202	∞	0.0001	0.0001
μ_5	1.0000			*not estimated*		
σ_1	0.0010			*not estimated*		
σ_2	0.0147	0.0138	0.0157	∞	0.0000	0.0000
σ_3	0.2016	0.1839	0.2186	∞	0.0001	0.0001
σ_4	0.1205	0.1066	0.1345	∞	0.0001	0.0001
σ_5	0.0010			*not estimated*		
				Europe | specification II		
μ_1	0.0000			*not estimated*		
μ_2	0.0154	0.0149	0.0159	∞	0.0000	0.0000
μ_3	0.1158	0.1073	0.1244	∞	0.0001	0.0001
μ_4	0.7132	0.6974	0.7292	∞	0.0001	0.0002
μ_5	1.0000			*not estimated*		
σ_1	0.0010			*not estimated*		
σ_2	0.0128	0.0124	0.0133	∞	0.0000	0.0000
σ_3	0.0875	0.0817	0.0931	∞	0.0000	0.0001
σ_4	0.2515	0.2403	0.2621	∞	0.0001	0.0001
σ_5	0.0010			*not estimated*		

Notes: The table summarizes the results of the component model. The first column presents the posterior means of the component means μ_k and standard deviations σ_k. The second and third column contain the lower and upper bound of the HPDI to a credibility level of 90%. In the last two columns, the naive and time-series standard error of the chains are presented, whereas, the time-series standard error is calculated based on the effective (N^*_{MCMC}) instead of the real (N_{MCMC}) sample size. Hereby, $N^*_{\text{MCMC}} < N_{\text{MCMC}}$ holds for autocorrelated chains.

extremes as in the US. However, component 4 exhibits the highest standard deviation ($\sigma_4 \approx 0.25$) and, thus, covers the highest proportion of the value range.

Probability model

Loan specific component probabilities are derived based on a probability model, i.e., the underlying OL model as of Equation (3.6). The results of the probability model are shown in Table 3.2 for specification I of the random effect and Table 3.3 for specification II of the random effect. Parameter estimates of the independent variables, i.e., β_{EAD}, β_{Facility}, $\beta_{\text{Protection}}$, and β_{Industry}, should be interpreted in Bayesian terms. In Bayesian inference, posterior distributions of β_j are assumed to be continuous. Thus, specific values in the posterior distributions exhibit

Table 3.2: Results of probability model (specification I)

	posterior mean	HPDI (90%)		posterior odds	naive standard error	time-series standard error
			US			
β_{EAD}	-0.0215	-0.0804	0.0408	2.5186	0.0004	0.0005
β_{Facility}	-0.0832	-0.1983	0.0404	6.6864	0.0007	0.0011
$\beta_{\text{Protection}}$	-0.4442	-0.6009	-0.2790	∞	0.0010	0.0019
β_{Industry}	0.2592	0.1002	0.4216	237.0952	0.0010	0.0013
α	2.1402	1.9252	2.3466	∞	0.0013	0.0023
σ^F	0.3522	0.2400	0.4604	∞	0.0007	0.0007
c_1	1.5000			*not estimated*		
c_2	2.4103	2.3529	2.4658	∞	0.0003	0.0003
c_3	3.9466	3.8677	4.0274	∞	0.0005	0.0005
c_4	4.5000			*not estimated*		
			GB			
β_{EAD}	-0.3792	-0.4378	-0.3190	∞	0.0004	0.0004
β_{Facility}	0.2138	0.1112	0.3237	2499.0000	0.0006	0.0009
$\beta_{\text{Protection}}$	-0.4271	-0.5451	-0.3216	∞	0.0007	0.0011
β_{Industry}	-0.1882	-0.3591	-0.0188	26.4725	0.0010	0.0013
α	2.7511	2.5321	2.9818	∞	0.0014	0.0017
σ^F	0.6040	0.4579	0.7499	∞	0.0009	0.0009
c_1	1.5000			*not estimated*		
c_2	2.7395	2.6818	2.7929	∞	0.0003	0.0003
c_3	3.6138	3.5480	3.6771	∞	0.0004	0.0005
c_4	4.5000			*not estimated*		
			Europe			
β_{EAD}	-0.0295	-0.0527	-0.0025	38.0625	0.0002	0.0002
β_{Facility}	0.2686	0.2188	0.3211	∞	0.0003	0.0006
$\beta_{\text{Protection}}$	-0.4176	-0.4715	-0.3655	∞	0.0003	0.0007
β_{Industry}	-0.1811	-0.2462	-0.1180	∞	0.0004	0.0007
α	2.1206	1.9521	2.2807	∞	0.0010	0.0011
σ^F	0.4929	0.3738	0.6038	∞	0.0007	0.0007
c_1	1.5000			*not estimated*		
c_2	2.6594	2.6289	2.6885	∞	0.0002	0.0003
c_3	3.2201	3.1839	3.2555	∞	0.0002	0.0004
c_4	4.5000			*not estimated*		

Notes: The table summarizes the results of the probability model with a latent variable specification as of specification I (Equation (3.6) and (3.7)). The first column presents the posterior means of the coefficients (β_j), the parameters of the random effect, and the cut points (c_k). The second and third column contain the lower and upper bound of the HPDI to a credibility level of 90%. In the fourth column, posterior odds are displayed. In the last two columns, the naive and time-series standard error of the chains are presented, whereas, the time-series standard error is calculated based on the effective (N^*_{MCMC}) instead of the real (N_{MCMC}) sample size. Hereby, $N^*_{\text{MCMC}} < N_{\text{MCMC}}$ holds for autocorrelated chains.

Table 3.3: Results of probability model (specification II)

	posterior mean	HPDI (90%)		posterior odds	naive standard error	time-series standard error
				US		
β_{EAD}	-0.0217	-0.0858	0.0353	2.6245	0.0004	0.0005
β_{Facility}	-0.0850	-0.2057	0.0367	6.8555	0.0007	0.0013
$\beta_{\text{Protection}}$	-0.4451	-0.6097	-0.2792	∞	0.0010	0.0022
β_{Industry}	0.2656	0.1030	0.4209	269.2703	0.0010	0.0013
a	2.7054	1.6139	3.8989	∞	0.0070	0.0289
φ	-0.2672	-0.7882	0.2520	0.2694	0.0032	0.0131
σ_c^F	0.3221	0.1934	0.4471	∞	0.0008	0.0016
c_1	1.5000			*not estimated*		
c_2	2.4099	2.3503	2.4639	∞	0.0003	0.0003
c_3	3.9469	3.8657	4.0275	∞	0.0005	0.0005
c_4	4.5000			*not estimated*		
				GB		
β_{EAD}	-0.3825	-0.4439	-0.3232	∞	0.0004	0.0004
β_{Facility}	0.2093	0.1078	0.3176	2499.0000	0.0006	0.0009
$\beta_{\text{Protection}}$	-0.4314	-0.5464	-0.3236	∞	0.0007	0.0011
β_{Industry}	-0.1930	-0.3579	-0.0223	32.1126	0.0010	0.0013
a	0.6779	0.0944	1.2111	999.0000	0.0036	0.0072
φ	0.7504	0.5576	0.9534	∞	0.0013	0.0025
σ_c^F	0.4023	0.2865	0.5188	∞	0.0007	0.0008
c_1	1.5000			*not estimated*		
c_2	2.7397	2.6838	2.7959	∞	0.0003	0.0003
c_3	3.6150	3.5517	3.6807	∞	0.0004	0.0005
c_4	4.5000			*not estimated*		
				Europe		
β_{EAD}	-0.0291	-0.0553	-0.0045	31.7869	0.0002	0.0002
β_{Facility}	0.2685	0.2196	0.3217	∞	0.0003	0.0006
$\beta_{\text{Protection}}$	-0.4184	-0.4685	-0.3634	∞	0.0003	0.0007
β_{Industry}	-0.1840	-0.2525	-0.1210	∞	0.0004	0.0006
a	0.2874	0.0094	0.5458	237.0952	0.0018	0.0044
φ	0.8565	0.7359	0.9862	∞	0.0009	0.0020
σ_c^F	0.2678	0.1982	0.3338	∞	0.0004	0.0005
c_1	1.5000			*not estimated*		
c_2	2.6597	2.6315	2.6907	∞	0.0002	0.0003
c_3	3.2201	3.1841	3.2560	∞	0.0002	0.0004
c_4	4.5000			*not estimated*		

Notes: The table summarizes the results of the probability model with a latent variable specification as of specification II (Equation (3.6) and (3.8)). The first column presents the posterior means of the coefficients (β_j), the parameters of the random effect, and the cut points (c_k). The second and third column contain the lower and upper bound of the HPDI to a credibility level of 90%. In the fourth column, posterior odds are displayed. In the last two columns, the naive and time-series standard error of the chains are presented, whereas, the time-series standard error is calculated based on the effective (N^*_{MCMC}) instead of the real (N_{MCMC}) sample size. Hereby, $N^*_{\text{MCMC}} < N_{\text{MCMC}}$ holds for autocorrelated chains.

a probability of zero. In frequentistic terms, one *true* value is assumed. A null hypothesis for β_j is set up accordingly to reach a yes-no-decision. In Bayesian terms, posterior distributions of parameters β_j are adopted to examine if the results are in favor of a positive or negative impact or if there is no clear one-sided influence. Two concepts might be applied. First, credible intervals, e.g., HPDIs, are intervals in the domain of the posterior distribution. If zero is not included in the credible interval, the domain of the posterior is located in the positive or negative value range – indicating positive or negative impact. Second, Bayes factors might be applied to evaluate statistical evidence and are defined as the relation between posterior and prior odds. Posterior odds are the ratio of the posterior probability mass favoring the sign of the posterior mean to the posterior probability mass of the opposite sign

$$\text{posterior odds}_{\,\mathrm{E}[\beta_j]<0} = \frac{\mathbb{P}(\beta_j < 0 \mid \text{data})}{\mathbb{P}(\beta_j \geq 0 \mid \text{data})}$$
$$\text{posterior odds}_{\,\mathrm{E}[\beta_j]>0} = \frac{\mathbb{P}(\beta_j > 0 \mid \text{data})}{\mathbb{P}(\beta_j \leq 0 \mid \text{data})}\,,$$

whereas, prior odds are the corresponding ratio of the prior distribution.[18] Thus, the Bayes factor complies with the posterior odds if the prior odds equal to one. This is true for symmetric prior distributions around zero. We set such a prior for the parameters vector β.[19] Hence, the corresponding posterior odds are interpretable in terms of Bayes factors. Following Kass and Raftery (1995), a Bayes factor exceeding 3.2 is deemed as substantial evidence. Values above 10 are assigned with strong evidence, whereas, values above 100 are related to decisive evidence.

The upper panels of Table 3.2 (specification I) and 3.3 (specification II) summarize the results of the probability model for the US. The specification of the random effect seems to have no impact on the remaining parameter estimates which seems to be important for the overall model consistency, i.e., changing the assumptions regarding the unobservable variable does not alter any conclusion regarding observable predictor variables. Considering the posterior odds, only β_{EAD} exhibits no clear evidence for the sign of the posterior mean (posterior $\text{odds}_{\mathrm{E}[\beta_{\text{EAD}}<0]} \approx \{2.5, 2.6\} < 3.2$). Thus, it can not be stated with conviction whether loans with higher EADs lead to lower or higher LGDs. The remaining variables are of categoric nature. The reference categories are term loan for facility, non protected for protection, and non FIRE for industry. The posterior odds indicate substantial evidence for a negative impact of β_{Facility} (posterior $\text{odds}_{\mathrm{E}[\beta_{\text{Facility}}<0]} \approx \{6.7, 6.9\} > 3.2$), i.e., lines exhibit lower values for

[18] Prior odds equal the ratio of prior probabilities for two states of the world. For instance, the states are $\beta_j < 0$ and $\beta_j \geq 0$.
[19] The prior of the parameter vector β is a Multivariate Normal distribution with mean vector zero.

Z_i^* and, thus, lower losses, compared to term loans. The evidence for a negative impact of $\beta_{\text{Protection}}$ and a positive impact of β_{Industry} is decisive (posterior odds$_{\text{E}[\beta_{\text{Protection}}<0]} \to \infty > 100$ and posterior odds$_{\text{E}[\beta_{\text{Industry}}>0]} \approx \{237.1, 269.3\} > 100$). Protection is, thus, associated with lower values for Z_i^* and lower losses, whereas, the FIRE industry affiliation seems to be shaped by higher losses. Generally, higher values of Z_i^* imply lower probabilities for lower components and higher probabilities for higher components. Higher values of Z_i^* can, thus, be directly associated with higher losses.

The middle panels of Table 3.2 (specification I) and 3.3 (specification II) contain the results of the probability model for GB. Again, the specification of the random effect seems to have no influence on the remaining parameter estimates. Regarding the posterior odds, the evidence for a negative impact of EAD, a positive impact of facility, and a negative impact of protection is decisive (posterior odds$_{\text{E}[\beta_{\text{EAD}}<0]} \to \infty > 100$, posterior odds$_{\text{E}[\beta_{\text{Facility}}>0]} \approx 2499.0 > 100$ and posterior odds$_{\text{E}[\beta_{\text{Protection}}<0]} \to \infty > 100$). Higher EADs are associated with lower losses. Lines generate higher losses compared to term loan and protection leads to lower losses. The evidence of a negative sign for industry is strong (posterior odds$_{\text{E}[\beta_{\text{Industry}}<0]} \approx \{26.5, 32.1\} > 10$). Thus, FIRE affiliation entails lower losses compared to other industries. Comparing these results with the US, two deviations in signs of parameters arise. First, lines are associated with lower losses compared to term loans in the US, however, higher losses arise for lines in GB. This might be caused by different business practices, e.g., lines might be closer monitored in the US and, thus, information asymmetries might be reduced to a higher extent. Second, loans granted to corporations of FIRE affiliation are characterized by higher losses in the US, whereas, lower losses occur in GB. Reasons may be found in deviating industry standards. While the financial system in GB is strongly shaped by banks, the US American system is more market-orientated – i.e., corporations might rather fund debt by traded instruments than loans. Thus, corporations depending on bank financing might worse in the first place. The lower panels of Table 3.2 (specification I) and 3.3 (specification II) displays the results of the probability model for Europe. The evidence for a negative impact of EAD is strong (posterior odds$_{\text{E}[\beta_{\text{EAD}}<0]} \approx \{38.1, 31.8\} > 10$), whereas, the evidence for a positive impact of facility (posterior odds$_{\text{E}[\beta_{\text{Facility}}>0]} \to \infty > 100$), for the impact of protection (posterior odds$_{\text{E}[\beta_{\text{Protection}}<0]}$ $\to \infty > 100$) and industry (posterior odds$_{\text{E}[\beta_{\text{Industry}}<0]} \to \infty > 100$) is decisive. The signs and magnitudes of the posterior means correspond to GB.

Besides the summary of the posterior estimates for β_{EAD}, β_{Facility}, $\beta_{\text{Protection}}$, and β_{Industry}, Table 3.2 and 3.3 present the results regarding the parameters of the random effect for specification I (α and σ^F in Table 3.2) and specification II (a, φ, and σ_c^F in Table 3.3). It is essential to state that the specification of the random effect (specification I in Table 3.2 and specification II in Table 3.3) has no impact in the posterior distributions of the β_j parameters and the cut points c_k. Thus, the specifications are consistent. However, it is substantial for the understanding of the systematic nature among LGDs. An i.i.d. random effect as of specification I implies unconditional behavior of the LGDs in the time line, i.e., LGDs of loans defaulted in previous quarters have no impact on LGDs in the current quarter. Assuming a time dependent AR(1) process as of specification II entails conditional behavior, i.e., if average LGDs of loans defaulted in the current quarter are high, average LGDs tend to be high in the subsequent quarter. To evaluate the specifications, we consider the evidence of the AR(1) parameter φ and select specification I if φ is not evident ($0 \in$ HPDI and posterior odds ¡ 3.2) and specification II otherwise.[20] In the US, the AR(1) parameter is not statistically evident as the HPDI includes the zero and the posterior odds amount to round about 0.3. This does not indicate conditional behavior of the random effect and, thus, no cyclical nature of its realizations. Contrary, φ is decisively evident in GB and Europe. Thus, we consider specification I for the US and specification II for GB and Europe. This indicates that the systematic impact on LGDs in GB and Europe is rather of cyclical nature.[21]

Systematic effects among LGDs

In the following, the results regarding the random effect F_t are further analyzed. We introduced F_t to capture systematic effects impacting LGDs of all loans defaulted in the same quarter. In specification I, the random effect is normally distributed with mean α an standard deviation σ^F. Thus, the latent variable Z_i^* is shifted with a constant level of α. Variations in the random effect and, thus, its impact, are captured by the standard deviation σ^F. In specification II, the random effect follows an AR(1) process with intercept a and AR(1) parameter φ. The latent variable Z_i^* is shifted with the unconditional mean of the AR(1) process $\mu_u^F = a/(1-\varphi)$. Its impact is expressed by the unconditional standard deviation $\sigma_u^F = \sigma_c^F / \sqrt{1-\varphi^2}$.

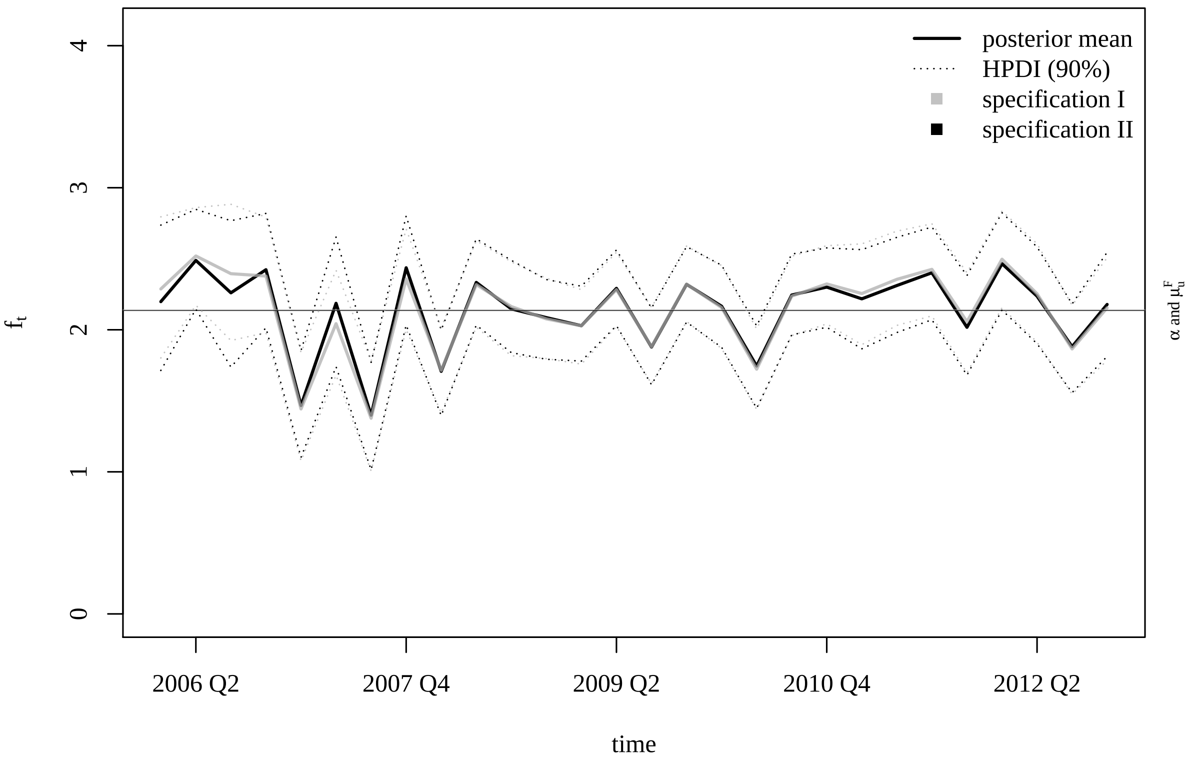
In both specifications, realizations f_t vary through time. For realizations $f_t > \alpha$ (specification I) or $f_t > \mu_u^F$ (specification II), z_i^* is shifted upwards for all loans i defaulted in the same time ($t(i) = t$). Thus, probabilities of low components decrease and probabilities of high components increase. This implies higher average LGDs at time t. If realizations f_t lie below its mean α (specification I) or μ_u^F (specification II), probabilities of low (high) components are increased

[20] Please note, that the posterior odds of φ are interpretable in terms of a Bayes factor as we set a symmetric prior around zero (normal distribution with mean zero and symmetric truncation $[-1,1]$).
[21] Convergence diagnostics of the selected model specifications can be found in Appendix 3.B.

(decreased) resulting in lower average LGDs at time t. Hence, we expect low realization of F_t in favorable systematic conditions and high realization during adverse systematic conditions.

Figure 3.2 illustrates the course of the random effect over time in the US. Specification I is plotted in gray, specification II in black. Posterior means are displayed by thick lines, HPDIs at credibility levels of 90% by dotted lines, α (specification I, thin gray line) and μ_u^F (specification II, thin black line) are nearly identical. The specification of the random effect rarely influences

Figure 3.2: Random effect in time line (US)



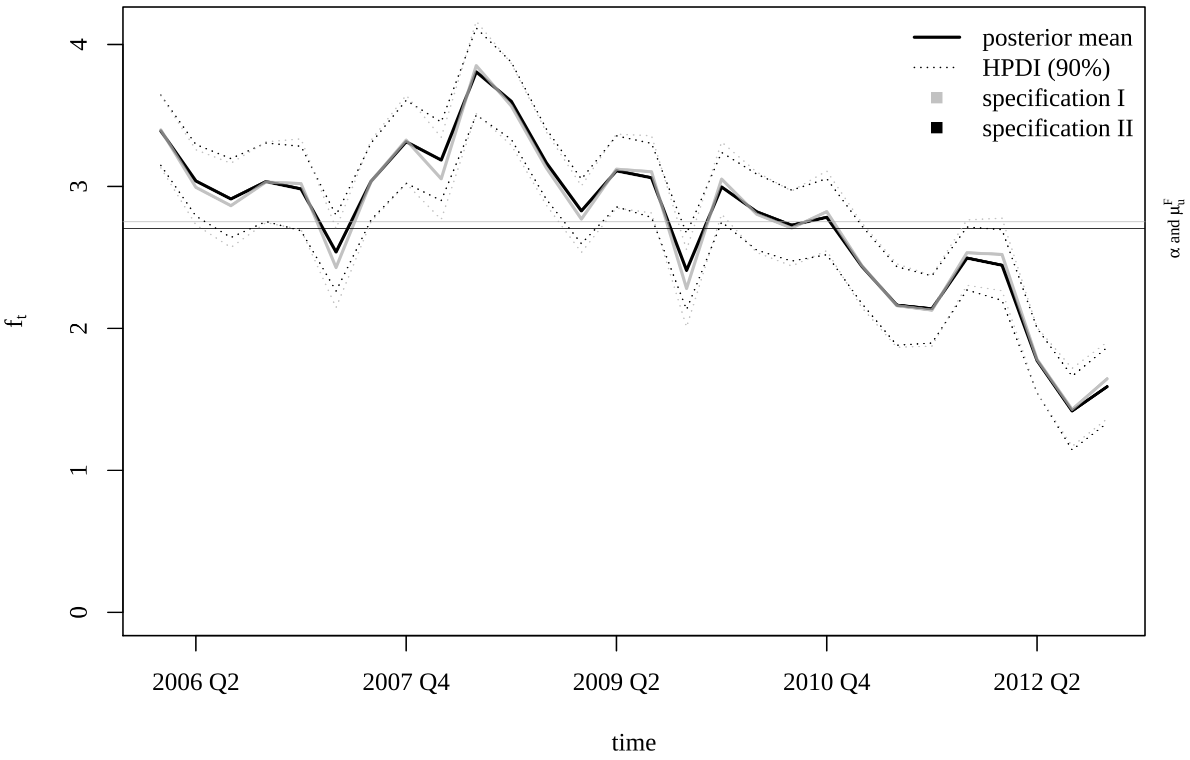
Notes: The figure illustrates the course of the random effect over time. The thick gray line represents the posterior means for specification I, whereas, the HPDI to a credibility of 90% is expressed by dotted gray lines. The thick black line shows the posterior means for specification II. The corresponding HPDI is displayed by dotted black lines. The thin horizontal lines represent the mean of the random effect (α for specification I and μ_u^F for specification II).

its individual realizations. Realizations f_t seem to independently proceed around the means α (specification I) and μ_u^F (specification II). This supports the evidence of an i.i.d. random effect. Despite its rather random process, the systematic effect is characterized by low realizations pre crisis indicating lower average LGDs. An upward shift during the GFC is less pronounced. After a slight drop in 2010 Q2, the random effect remains on a rather high level indicating high average LGDs in the recent time periods. Thus, the random effect does not seem to display the time patterns of common macro variables. However, it simulates the time series of average LGDs in the data. The upper panels of Figure 3.C.2 in Appendix 3.C contrasts average LGDs (black lines) with the random effect (gray lines). In the left panel, average LGDs and the realizations

f_t for specification I per quarter are displayed.[22] The rolling averages of both time series are plotted in the right panel.

Figure 3.3 illustrates the course of the random effect over time in GB. The presentation corresponds to Figure 3.2. The GFC is clearly identifiable in the random effect of GB. The highest

Figure 3.3: Random effect in time line (GB)

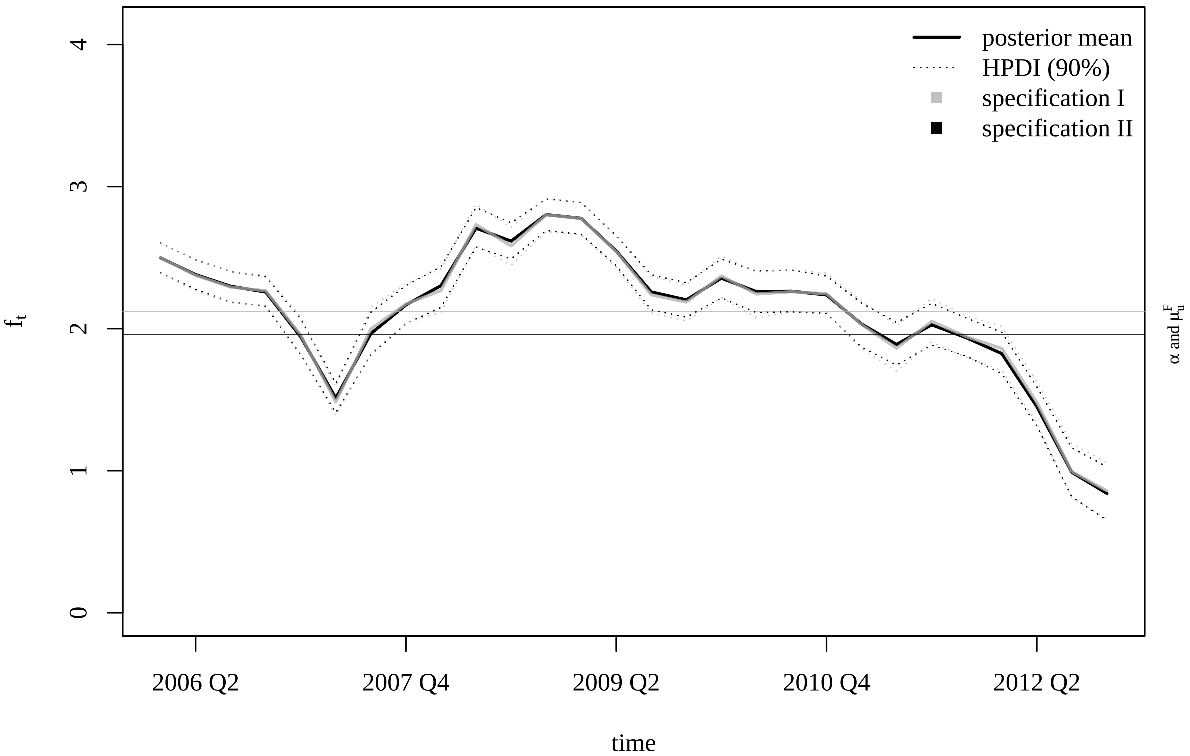


Notes: The figure illustrates the course of the random effect over time. The thick gray line represents the posterior means for specification I, whereas, the HPDI to a credibility of 90% is expressed by dotted gray lines. The thick black line shows the posterior means for specification II. The corresponding HPDI is displayed by dotted black lines. The thin horizontal lines represent the mean of the random effect (α for specification I and μ_u^F for specification II).

realization f_t is in 2008 Q2 during the summit of the crisis. Afterwards, the realizations of the random effect constantly decline until the minimum is reached in recent time periods. In analogy to the US, the random effect reproduces the time patterns of average LGDs. The middle panels of Figure 3.C.2 in Appendix 3.C contrast the two time series. However, the British random effect seems to exhibit a rather cyclical behavior compared to the US. This is already indicated by the evidence of φ in Table 3.3. Figure 3.4 illustrates the course of the random effect over time in Europe. The presentation corresponds to Figure 3.2. The European random effect shows strong similarity to GB. The GFC is clearly observable, however, the crisis seems prolonged compared to GB as f_t is still near its maximum in 2009 Q1. Following the GFC, the random effect slowly declines. Again, the random effect simulates the time series behavior of

[22] Specification II is skipped for presentational purposes. However, the realizations f_t are similar.

Figure 3.4: Random effect in time line (Europe)



Notes: The figure illustrates the course of the random effect over time. The thick gray line represents the posterior means for specification I, whereas, the HPDI to a credibility of 90% is expressed by dotted gray lines. The thick black line shows the posterior means for specification II. The corresponding HPDI is displayed by dotted black lines. The thin horizontal lines represent the mean of the random effect (α for specification I and μ_u^F for specification II).

average LGDs in the data. The lower panels of Figure 3.C.2 in Appendix 3.C contrasts average LGDs and f_t.

Systematic effects vs. macro variables

The time patterns of the random effects (see Figure 3.2 for the US, Figure 3.3 for GB, and Figure 3.4 for Europe) might be a first indication that macro variables might not be suitable to capture the intrinsic systematic effects among LGDs. To examine this in more detail, we consider macro variables.

In this context, Figure 3.C.3, 3.C.4, and 3.C.5 in Appendix 3.C contrast the course of the considered macro variables – ΔGDP, ΔEI, VIX, ΔHPI, and NPL ratio – to the course of the random effects. The y-axis is reversed for ΔGDP, ΔEI, and ΔHPI to ensure an intuitive interpretation. If macro variables imply the same information as the random effects, their time patterns should be congruent with the random effect. However, this does not seem to be the case in the US (see Figure 3.C.3). While stronger deteriorations are indicated by the macro variables in the GFC (upward movement), macro variables return to pre-crisis levels at the end of 2009. Contrary, the random effect further increases after the GFC. Just the course of the NPL ratio seems to capture parts of this movement as a rather slow recovery post crisis is indicated by this variable. In GB and Europe, the time series patterns of macros are more similar to the random effects (see Figure 3.C.4 and 3.C.5). However, the random effect lies above macros variables post crisis. While sharp rebound pushes the economic indicators back to their pre crisis levels, the easing is slower in terms of the random effects and, thus, average LGDs. In addition, the NPL ratio seems incapable of capturing this behavior as it remains on its crisis levels until the end of 2012, whereas, slow recovery is indicated by the random effects. Table 3.D.4 in Appendix 3.D summarizes the pairwise correlations of the displayed time series. Perfect multicollinearity would be indicated by a correlation coefficient of 100%. However, the correlation is negative for most of the macro variables (ΔGDP, ΔEI, VIX, and ΔHPI) in the US. Just the course of the NPL ratio exhibits a certain collinearity to the random effect. In GB and Europe, more macro variables exhibit positive correlations to the random effect (ΔGDP and ΔEI in GB and ΔGDP, ΔEI, VIX, and ΔHPI in Europe).

In the light of the above, macro variables might not be suited to capture the true systematic effects among LGDs. However, we estimate the models with macro variables instead of random effects to analyze their impact.[23] Variable selection is not trivial considering highly correlated time series such as macro variables as multicollinearity might arise. This endangers model stability. Small changes in model specifications or on data side might provoke huge alterations in parameter estimates. Furthermore, parameter estimates tend to be less precise and standard errors large. We, thus, decide to include just one of the considered macro variables at a time. Table 3.4 summarizes the results of the models with macro variables. The presentation of the remaining parameters (β_{EAD}, β_{Facility}, $\beta_{\text{Protection}}$, and β_{Industry}) is skipped as no changes in the signs and magnitudes arise.[24] We assume favorable economic conditions to be associated with lower LGDs, thus, negative impacts of ΔGDP, ΔEI, and ΔHPI and positive signs for VIX and NPL ratio. Comparing the signs of the posterior means to the expected signs of the macro variables, discrepancies arise. In the US, only the posterior mean of the NPL ratio exhibits the expected sign. In GB and Europe, all considered macro variables except the NPL ratio show intuitive signs. However, there is no statistical evidence for the impact of VIX and ΔHPI in GB (posterior odds$_{\text{E}[\beta_{\text{VIX}}>0]} \approx 1.4 < 3.2$ and posterior odds$_{\text{E}[\beta_{\text{HPI}}<0]} \approx 2.0 < 3.2$).[25] Counterintuitive

[23] Please note that these models, i.e., the macro models, are very similar to the model presented in Altman and Kalotay (2014). However, the authors transform LGDs. As we explicitly aim to allow for LGD ≤ 0 and LGD ≥ 1, we do not transform the observations.

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

[25] Please note, that the posterior odds of β_{GDP}, β_{EI}, β_{VIX}, β_{HPI}, and $\beta_{\text{NPL ratio}}$ are interpretable in terms of a Bayes factor as we set symmetric priors around zero, i.e., normal distributed priors with means zero.

Table 3.4: Results of macro models

		posterior mean	HPDI (90%)		posterior odds	naive standard error	time-series standard error
					US		
β_{GDP}	−	0.0266	-0.0291	0.0821	3.4703	0.0003	0.0003
β_{EI}	−	0.0062	-0.0529	0.0617	1.3175	0.0003	0.0003
β_{VIX}	+	-0.0319	-0.0880	0.0255	4.5866	0.0003	0.0004
β_{HPI}	−	0.0483	-0.0065	0.1069	11.8041	0.0003	0.0003
$\beta_{\text{NPL ratio}}$	+	0.0661	0.0078	0.1218	33.6021	0.0003	0.0004
					GB		
β_{GDP}	−	-0.0908	-0.1381	-0.0429	768.2308	0.0003	0.0003
β_{EI}	−	-0.1012	-0.1501	-0.0548	2499.0000	0.0003	0.0003
β_{VIX}	+	0.0063	-0.0397	0.0550	1.4213	0.0003	0.0003
β_{HPI}	−	-0.0134	-0.0603	0.0348	2.0331	0.0003	0.0003
$\beta_{\text{NPL ratio}}$	+	-0.3264	0.0078	0.1218	∞	0.0003	0.0003
					Europe		
β_{GDP}	−	-0.1781	-0.2020	-0.1556	∞	0.0001	0.0001
β_{EI}	−	-0.1845	-0.2083	-0.1621	∞	0.0001	0.0001
β_{VIX}	+	0.1694	0.1459	0.1918	∞	0.0001	0.0001
β_{HPI}	−	-0.2314	-0.2535	-0.2073	∞	0.0001	0.0001
$\beta_{\text{NPL ratio}}$	+	-0.0853	-0.1088	-0.0619	∞	0.0001	0.0001

Notes: The table summarizes the results of the macro models. The presentation is reduced to the results regarding the macro variables itself. Every macro variable was separately included in a model without random effect. The first column includes the expected signs of the posterior means as bad macro economic environment should entail higher LGDs. The second column presents the posterior means of the coefficients of the macro variables (β_{GDP}, β_{EI}, β_{VIX}, β_{HPI}, $\beta_{\text{NPL ratio}}$). The third and fourth column contain the lower and upper bound of the HPDI to a credibility level of 90%. The fifth column includes posterior odds, while, in the last two columns, the naive and time-series standard error of the chains are presented, whereas, the time-series standard error is calculated based on the effective (N^*_{MCMC}) instead of the real (N_{MCMC}) sample size. Hereby, $N^*_{\text{MCMC}} < N_{\text{MCMC}}$ holds for autocorrelated chains.

signs and the lack of statistical impact cast doubt on the use of macro variables for LGD modeling. These results do not claim to universal validity. However, the identification problem of macro variables is emphasized.

To check the robustness of these findings, we reestimate the models considering macro variables and random effects. Among the macro variables with intuitive signs, we select those offering the highest statistical evidence.[26] Thus, the NPL ratio is selected for the US as it is the only macro variable with an intuitive sign, ΔEI for GB due to its high posterior odds ratio, and ΔHPI for Europe as its HPDI is furthest from zero. Table 3.D.5 in Appendix 3.D summarizes the results of the combined models.[27] Statistical evidence vanishes for all considered macro variables. First, posterior odds ratios are smaller than 3.2. Second, the corresponding HPDIs to a credibility level of 90% include zero. In contrast, the parameters of the random effect – α and σ^F for specification I and a, φ, and σ^F_c for specification II – remain nearly unchanged compared to the original model specification without the inclusion of macro variables (see Table 3.2 and 3.3).

[26] Results for the remaining macro variables are available from the authors upon request.

[27] The presentation of the remaining parameters (β_{EAD}, β_{Facility}, $\beta_{\text{Protection}}$, and β_{Industry}) is skipped. Results are available from the authors upon request.

Summarizing the above results, the identification of appropriate macro variables seems challenging in an LGD modeling context. Macro variables in general seem to be not entirely suitable to capture the true systematic effects deriving LGDs. This might be due to three reasons. First, LGD observations are treated as they arise at the default time t. However, LGD realizations are not available until the defaulted loan is completely resolved. In our data sets, default resolution takes typically between one and five years. During the resolution process, recovery payments are processed. Thus, the realized LGD in t does not only depend on the economic condition in t but on the conditions during the whole resolution process ($t+\Delta t$). As consequence, estimated random effects with time stamp t are rather an aggregated proxy of the economic conditions during $t+\Delta t$, where, Δt corresponds to the resolution time. Second, financial institutions have to deal with high stocks of non-performing loans post crises. This may enforce fast and potentially cost-intensive resolutions as institutions want to settle open claims. Slower recovery as indicated by economic proxies such as macro variables might be the consequence. Third, systematic effects on LGDs might not be purely of economic nature. In the US, rather high realizations of the random effect occur since 2010 Q2. Even through the time period from 2010 Q2 to 2013 Q2 is considered as crisis by the OECD, it is surprising that LGDs are averagely higher compared to the GFC. Thus, it is possible that something beyond economic conditions systematically increasing LGDs. As the implementation of Basel II into US law proceeds post crisis, regulatory effects on LGDs could be conceivable. Besides, regulations might lead to adjustments in banking practice which also could influence loss rates.

Posterior predictive distribution

In the following, we briefly analyze the posterior predictive distributions of LGDs as generated by the model. We focus on its ability to capture the patterns of the empirical LGD distribution (see, e.g., left panels of Figure 3.1 in Section 3.5.1). As we do not compare different models, we concentrate on graphical tools.

Figure 3.C.6 in Appendix 3.C illustrate the characteristics of the posterior predictive distribution for the US, GB, and Europe. The left panels contrast kernel density estimates of the posterior predictive distribution (gray line) to the empirical LGD distribution (dotted line). As the band width is fixed to 0.015 for both kernel density estimates, their height is comparable despite ties in the empirical data. As the particular shape of LGD distributions is challenging to illustrate via density estimates due to the extreme modi at zero and one, the right panels present the quantile-quantile (qq) plot contrasting the quantiles of the empirical distribution (x-axis) to the quantiles of the posterior predictive distribution (y-axis). The bisector (black line) represents

optimality as it indicates that the quantiles of both distributions correspond. The models seem to be characterized by a good fit regarding the distributional form of the posterior predictive distribution in-sample in all three considered regions (see left panels of Figure 3.C.6). Thus, the assumption of a Gaussian FMM, i.e., five normally distributed components, as component model seems adequate. The right panels of Figure 3.C.6 support this impression as the dots in the qq plot almost perfectly lie on the optimality line indicating that the quantiles of the posterior predictive distribution comply with the quantiles of the empirical distribution.

3.6 Analysis of downturn LGDs

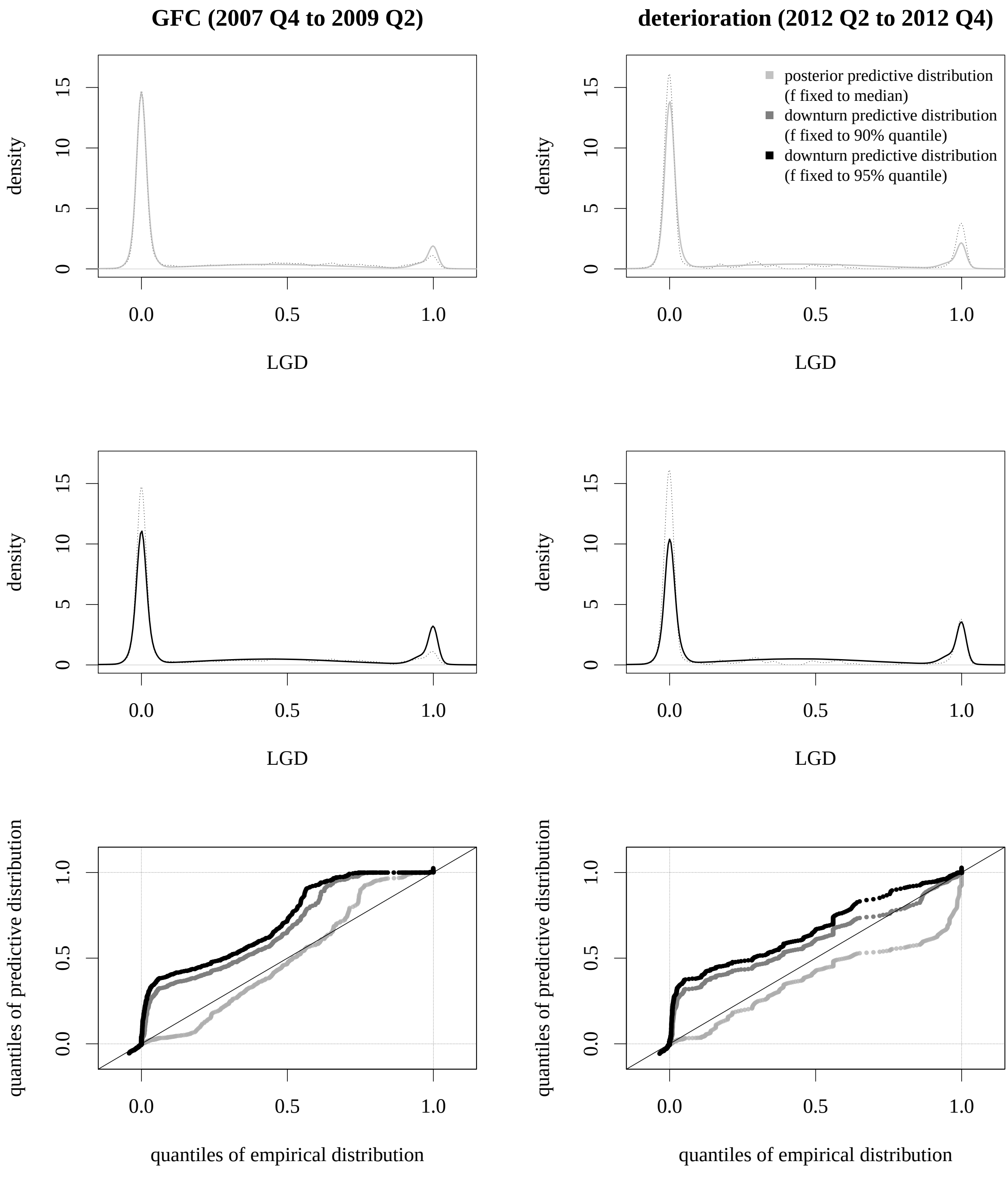
We analyze downturn LGD distributions for two time periods in each region. In all regions, we consider the GFC. According to the OECD, the financial crisis is terminated in the period from 2007 Q4 to 2009 Q2 in the US, whereas, it is shifted by one quarter in GB and Europe (2008 Q1 to 2009 Q3). Considering the time patterns of average LGDs as in Figure 3.1, we evaluate additional time periods which are characterized by high average losses. We will refer to this periods as deterioration periods. In the US, we select the time span from 2012 Q2 to 2012 Q4.[28] The year right after the GFC is applied in GB and Europe (2009 Q4 to 2010 Q3).

Downturn estimation via random effects

Figure 3.5 contrast the posterior predictive and downturn predictive distributions to the empirical LGD distribution in the GFC and deterioration period for the US. The upper panels display the kernel density estimates of the data (dotted lines), the posterior predictive distribution (light gray lines), and a downturn predictive distribution (black lines), whereby, the random effect is set to its 95% quantile. The lower panels show the corresponding qq plots, whereby, the downturn predictive distribution based on the 90% quantile of the random effect is added. The posterior predictive distribution is reflected by light gray dots, the downturn predictive distribution based on the 90% quantile of the random effect by dark gray dots and the posterior predictive distribution based on the 95% quantile of the random effect by black dots. During the GFC, the posterior predictive distribution seems to be sufficiently conservative as it fits the empirical LGD distribution quite well. However, this time period is characterized by rather low probability mass at total loss compared to the US American data set as a whole. Thus, the posterior predictive distribution implying average systematic conditions already overesti-

[28] Under the terms of the OECD, the time period from 2012 Q2 to 2013 Q2 is classified as recession period in the US.

Figure 3.5: Posterior and downturn distribution for the GFC and a deterioration period (US)



Notes: The figure contrasts the empirical LGD distribution to the posterior (light gray) and downturn (dark gray and black) predictive distribution for the GFC (2007 Q4 to 2009 Q2, left panels) and a deterioration period (2012 Q2 to 2012 Q4, right panels). The upper panels display the kernel density estimates, the lower panels the quantile-quantile (qq) plots. The kernel density estimates for the downturn predictive distribution for the 90% quantile of the random effect is skipped and available form the authors upon request.

mates high LGD realizations up to a certain degree. Accordingly, quantiles of the downturn predictive distributions lie always higher than empirical quantiles. In the deterioration period, the posterior predictive distribution underestimates the probability mass of high and total losses. However, the downturn predictive is capable of capturing this systematic higher fraction. Figure 3.C.7 and 3.C.8 in Appendix 3.C illustrate the corresponding analytics for GB and Europe. The presentation corresponds to Figure 3.5. Results are almost similar. However, the posterior predictive distribution does not seem to be sufficiently conservative during the GFC as the empirical LGD distribution is characterized by higher probability masses at total loss in GB and Europe regarding this time period.[29]

Generally, these downturn LGD distributions may be considered too conservative. However, the distance between the downturn predictive distribution and the optimality line is directly impacted by the selected conservative quantile. Confidence levels smaller than 90% result in a lower gap. Hence, downturn estimates are adjustable according to the needs of the risk manager or regulator. The main advantage of the presented model and the implied downturn approach via random effects is the option to generate conservative estimates in accordance with the characteristic nature of LGDs.

Downturn estimation via macro variables

In the literature, several suggestions to generate downturn estimations exist (see Section 3.4). The most common one refers to the inclusion of macro variables in the modeling context (see, e.g., Altman and Kalotay, 2014; Krüger and Rösch, 2017). Hence, we derive posterior and downturn predictive distributions based on the model including a macro variable instead of the random effect. In each region, we select the macro variable with the highest statistical evidence. Thus, we apply the NPL ratio for the US, ΔEI for GB, and ΔHPI for Europe.[30] In analogy to the downturn generation based on random effects, a conservative quantile of the macro variable is selected to generate the downturn predictive distribution.[31]

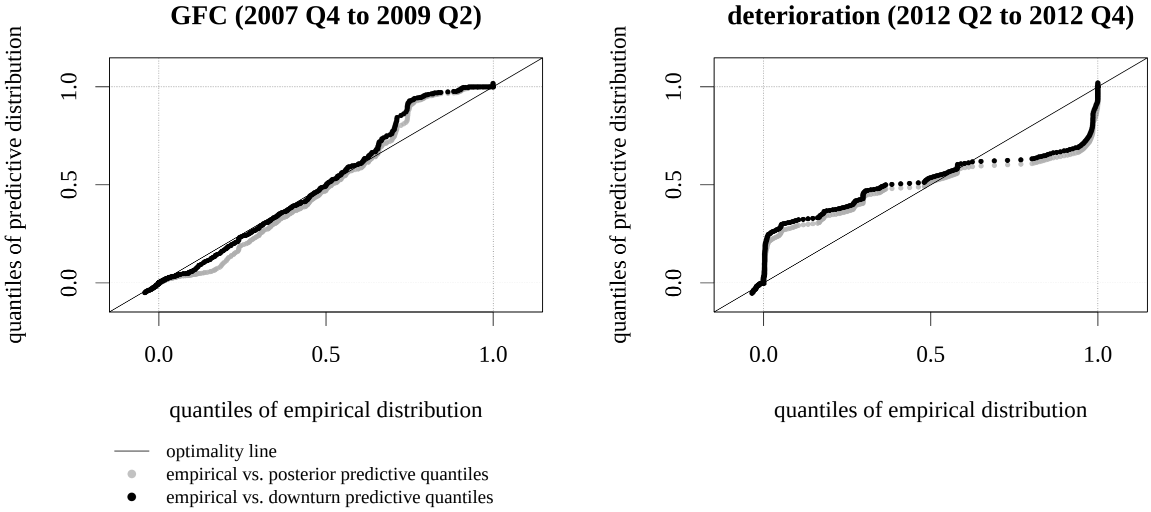
Figure 3.6 contrasts the quantiles of the empirical distribution in the GFC and the deterioration period to the corresponding posterior and downturn predictive distributions for the US. The course of quantiles regarding the posterior predictive distributions based on the macro model

[29] To examine robustness, we evaluate the downturn distributions on an out-of-time basis. The training set contains the time period from 2006 Q1 to 2010 Q1. The test set consists of the time period from 2010 Q2 to 2012 Q4. Results are presented in Figure 3.C.9 in Appendix 3.C. The downturn distributions are still conservative if fitted on a out-of-time basis.

[30] We only include just one macro variable at a time due to their high correlation (see Section 3.5.2). Results regarding the remaining macro variables are available form the authors upon request.

[31] We select a conservative quantile of 99%.

Figure 3.6: Posterior and downturn distribution for the GFC and a deterioration period based on the macro model containing the NPL ratio (US)



Notes: The figure contrasts the empirical LGD distribution to the posterior (light gray) and downturn (dark gray and black) predictive distribution of the macro model containing the NPL ratio instead of a random effect for the GFC (2007 Q4 to 2009 Q2, left panel) and a deterioration period (2012 Q2 to 2012 Q4, left panel).

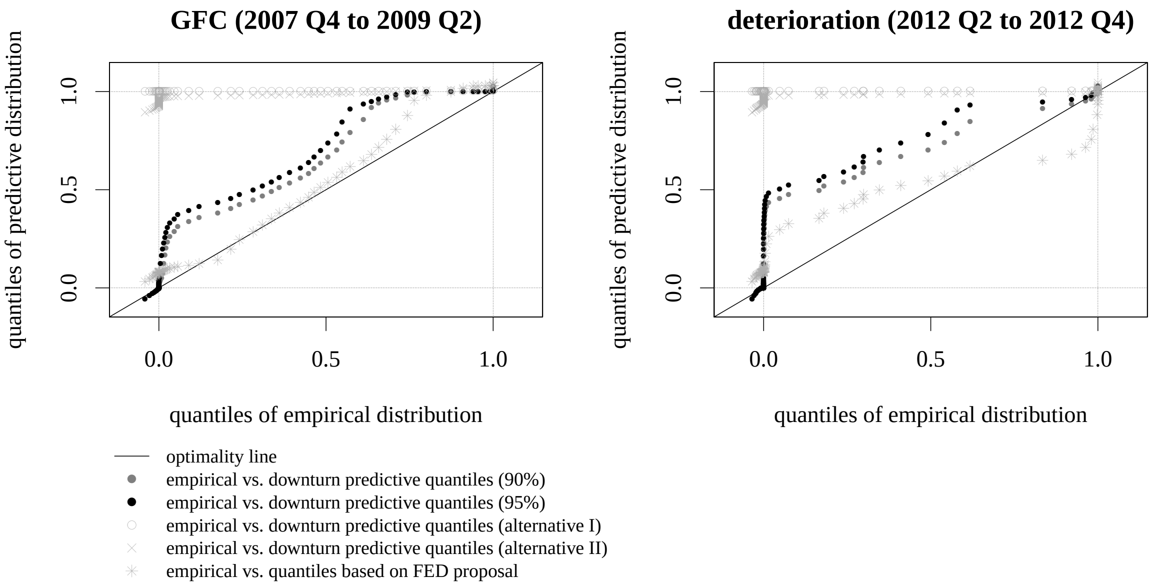
is rather similar compared to the random effect model (see Figure 3.5). However, the corresponding downturn distribution does not sufficiently capture the quantiles of the empirical LGD distribution in the time period from 2012 Q2 to 2012 Q4. Comparing the downturn distributions of the macro model with the one of the random effect model in Figure 3.5, the gap to the corresponding posterior predictive distributions seems undersized. This might be due to a rather small posterior mean of the coefficient ($\beta_{\text{NPL ratio}} \approx 0.07$, see Table 3.4) and a HPDI (90%) which nearly reaches zero (HPDI (90%) = $[0.0078, 0.1218]$, see Table 3.4). As macro variables do not seem to be sufficient to capture the true systematic effect impacting LGDs, generating downturn distributions via macro variables might be less effective. Figures 3.C.10 and 3.C.11 in Appendix 3.C illustrate the corresponding analytics for GB and Europe. As the posterior means of the macro variables are of higher magnitude ($\beta_{\text{EI}} \approx -0.10$ for GB and $\beta_{\text{HPI}} \approx -0.23$ for Europe), the downturn predictive distribution differs clearer from the posterior predictive distribution. In Europe, the resulting downturn distribution is sufficiently conservative in both considered periods. However, the downturn distribution underestimates high-loss components in GB during the GFC. Furthermore, even though reasonable downturn LGD estimates can be derived by the inclusion of selected macro variables, the identification issue which macro variable is most appropriate to use is avoided by the use of the random effect model.

Alternative downturn concepts

Lastly, we compare our downturn concept with three other approaches presented in Section 3.4.

Figure 3.7 contrasts the downturn predictive distribution via a random effect (gray and black dots) to the approaches of Bijak and Thomas (2015) (alternative I, gray cycles), Calabrese (2014) (alternative II, gray crosses), and the FED proposal (gray stars) for the US. Both alternative

Figure 3.7: Downturn distribution for the GFC and a deterioration period based on alternative concepts (US)



Notes: The figure contrasts the empirical LGD distribution to downturn predictive distributions. The black and gray dots represent the downturn approach via a random effect, whereas, the suggestion of Bijak and Thomas (2015) (alternative I) is displayed by gray cycles and the proposal of Calabrese (2014) (alternative II) by gray crosses. The FED approach is displayed by gray stars. The figure refers to the GFC (2007 Q4 to 2009 Q2, left panel) and a deterioration period (2012 Q2 to 2012 Q4, right panel).

approaches generate far more conservative downturn distributions compared to applying random effects as the resulting downturn distributions do not capture the whole range of LGDs $[-50\%, 150\%]$. The probability mass is highly centered around one, i.e., total loss. However, the empirical LGD distribution is shaped by high probability masses around zero, i.e., no loss, even in quarters with high average LGDs. This pattern seems to be neglected by the approaches suggested by Bijak and Thomas (2015) and Calabrese (2014). In the GFC, the FED proposal seems to fit the empirical LGD distribution quite well. However, the posterior predictive distribution almost succeeds as well in this time period (see Figure 3.5). The FED proposal does not produce sufficiently conservative estimates in the deterioration period. Figure 3.C.12 and 3.C.13 in Appendix 3.C confirm these findings for GB and Europe.

In general, fundamental deviations among the approaches arise. While downturn distributions generated based on a random effect still cover the whole LGD range, the alternatives result in rather constant conservative values (Bijak and Thomas, 2015) or in restricted and conservative single component distributions (Calabrese, 2014). The FED proposal does not seem to be able to constantly generate sufficiently conservative estimates. The approach suggested in this paper is based on a shift in component probabilities in critical systematic conditions. Fixing the random effect to a conservative quantile decreases probabilities of low, i.e., *good*, components and increases the probability of high, i.e., *bad*, components. Although the deteriorated systematic surrounding is reflected in the downturn distribution, the whole range of LGDs remains covered. This functionality is based on the empirical observations of LGDs as the bi-modality remains during crises periods.

Dependent on the intention, the presented approach may be more or less favorable compared to the alternative suggestions by Bijak and Thomas (2015) and Calabrese (2014). Restricting downturn distributions to a certain value range might result in overestimation of risk and, thus, in extreme safety buffers. While this is in line with the conservativity principle, holding too much capital creates new risk and burdens the solvency of banks as opportunity cost increase and operating business is hampered. The approach presented in this paper generates reasonable safety buffers and is adoptable to the needs of risk managers and regulators. By illustrating this, we aim to contribute to the ongoing discussion on how to define and estimate downturn LGDs.

3.7 Conclusion

The central contribution of this paper lies in the analysis of systematic effects among LGDs. By this means, a novel approach for estimating downturn LGDs is suggested. We use a random effect to capture the systematic comovements in LGDs and apply realizations of this latent variable to calibrate downturn conditions. This enables us to generate downturn estimates which are consistent with the true systematic time patterns of average LGDs. A drawback of this approach may be the less intuitive time patterns of the latent variable in some instances as the random effect is straightly driven by the underlying data. However, conservative downturn estimates can be generated via random effects which is not true in all cases applying macro variables. The main obstacle regarding macro variables is the difficulty to find statistically evident proxies for economic conditions.

We find that systematic effects in LGDs differ among regions and to the macroeconomic cycle. Cyclical patterns can be observed in Europe and Great Britain, while the US seems to be characterized by time independent systematic patterns. Either way, these systematic effects strongly deviate from the economic cycle measured by common macro variables. Reason for this might be found in the collection process of recovery payments. Resolution processes take multiple years, thus, varying economic conditions impact final LGDs. This might hamper the use of economic variables with a specific time stamp. The random effect in our model may be interpreted as an average systematic impact on LGDs during the resolution process. We believe that multiple time lags and leads of macro variables are indispensable to consistently stimulate the impact of economic conditions by observable variables. This might complicate regulation mechanisms for risk managers and regulators as further issues rise.

In this paper, we contribute to the ongoing discussion how to provide consistent downturn estimates for LGDs. In comparison to other approaches, ours is fundamentally different. We set a critical state for a latent variable which impacts the probability of belonging to different segments of the LGD distribution. This leads to an increase in the fraction of high to low LGDs during critical times, a behavior we descriptively observed in crises periods. Nevertheless, the validation of downturn LGDs is still an open topic in the academic literature which complicates the final evaluation of downturn approaches. Thus, it cannot be finally stated which approach is superior. However, downturn estimation based on random effects offers straightforward regulation mechanisms for decision makers. Risk managers are enabled to adapt the critical state based on their portfolios, while regulators might set a lower bound to guarantee overall conservatism.

Lastly, data quality and quantity is substantial in the context of LGD modeling in general and regarding downturn estimation in particular. The presented results heavily depend on the economic cycle. The adapted data set covers only one economic crisis, i.e., the GFC, and slight expansions. Thus, analyses based on extended data set would be desirable. This may constitute potential weaknesses as our results are strongly data driven. However, we believe that our approach generates downturn estimates that entail capital buffers which are sufficient during crises periods similar to or worse than the GFC. Therefore, an open discussion among regulators and financial institutions is required in order to set reasonable critical levels for the random effect in our or similar approaches.

3.A Appendix | Bayesian model specification

The approach consisting of the component and probability model is estimated via Bayesian inference.[32] Thus, prior distributions have to be specified for every parameter of the model.

Component model

In the component model, the number of components is fixed to five ($K = 5$). As the distribution of LGDs is extremely bimodal and characterized by a high amount of ties at zero and one, we set the parameters of the first and fifth component to identify loans with no and total loss. The means are fixed to $\mu_1 = 0$ and $\mu_5 = 1$, whereas, the standard deviations are set to small values, i.e., $\sigma_1 = 0.001$ and $\sigma_5 = 0.001$. Due to convergence reasons, the remaining parameters (μ_k and σ_k for $k \in \{2, 3, 4\}$) are provided with weakly informed priors:

$$\begin{aligned} \mu_k &\sim N(\ \mu = 0\ ,\ \tau = 0.00001\ ) \quad [\ -0.5\ ,\ 1.5\ ] \\ \sigma_k &\sim \Gamma(\ s = 0.001\ ,\ r = 0.001\ ) \quad [\ 0.0001\ ,\ 1\ ]\ , \end{aligned} \tag{3.13}$$

where, the squared brackets indicate truncation. The priors of component means are Normal distributions with means zero ($\mu = 0$) and low precisions ($\tau = 0.00001$).[33] The parameter τ denotes the precision and is calculated as $\tau = \frac{1}{\sigma^2}$. To consider the maximal range of LGDs $[-50\%, 150\%]$, the priors for component means are truncated. The priors of component standard deviations are Gamma distributions with low values for the scale ($s = 0.001$) and rate ($r = 0.001$) parameters. This corresponds to an uninformed specification of the Gamma distribution. To achieve weakly informed priors, the Gamma distribution is truncated. The lower bounds of the truncation ensure convergence as the density of the Gamma distribution tends to infinity for values near zero. This can break Gibbs sampling as the sampler gets stuck at values of infinite density. The upper bound, again, takes into account the maximal range of LGDs $[-50\%, 150\%]$. Considering this range, a standard deviation of one is still rather high.

Probability model

In the probability model, two of the cut points are fixed to solve the over specification problem. The values are selected such that the scale of the latent variable Z_i^* is comparable to the

[32] The MCMC samples are drawn via the Gibbs sampler JAGS.
[33] A truncated normal distribution with low precession corresponds to a Uniform distribution. Thus, the choice of the mean is irrelevant.

magnitude of the latent class Z_i ($c_1 = 1.5$ and $c_4 = 4.5$). The remaining cut points (c_k for $k \in \{2,3\}$) are provided with uninformed prior distributions:

$$c_k \sim N(\ \mu = 3\ ,\ \tau = 0.00001\ )\ [\ c_1 = 1.5\ ,\ c_4 = 4.5\ ]\ . \tag{3.14}$$

Normal distributions with means in between the outer cut points ($\mu = 3$) and low precisions ($\tau = 0.00001$) are provided. The truncation ensures that $c_1 \leq c_k \leq c_4$ for $k \in \{2,3\}$ holds.

To enable block sampling in the MCMC chains, the prior of the coefficients β_j is set to an uninformed J-dimensional Multivariate Normal distribution:

$$\beta \sim N_J\left(\ \mu_{(J\times1)} = 0_{(J\times1)}\ ,\ \tau_{(J\times J)} = 1_{(J\times J)} \cdot 0.00001\ \right), \tag{3.15}$$

where, $\mu_{(J\times1)}$ is a $(J\times1)$ vector containing the prior means. These are set to zero and, thus, $\mu_{(J\times1)}$ corresponds to the J-dimensional zero vector ($0_{(J\times1)}$). The term $\tau_{(J\times J)}$ refers to the $(J\times J)$ precision matrix. This matrix contains precisions of 0.00001 on its diagonal and zero on its the non diagonal elements and, thus, corresponds to the $(J\times J)$ identity matrix ($1_{(J\times J)}$) multiplied by 0.00001.

The random effect in specification I is i.i.d. and follows a Normal distribution with mean α and standard deviation σ^F. As uniformed priors, the conjugate distributions are provided:

$$\begin{gathered}\alpha \sim N(\ \mu = 0\ ,\ \tau = 0.00001\ )\\ \left(\sigma^F\right)^2 \sim \Gamma^{-1}(\ s = 0.001\ ,\ r = 0.001\ )\ .\end{gathered} \tag{3.16}$$

Thus, the prior of α is a Normal distribution with mean zero ($\mu = 0$) and low precision ($\tau = 0.00001$) and the prior of $\left(\sigma^F\right)^2$ is an Inverse Gamma distribution with uniformed specification for the scale ($s = 0.001$) and rate ($r = 0.001$) parameter.

The random effect in specification II follows an AR(1) process where its realization at time t depends on its realization at time $t-1$. To generate the posterior distribution of $f_{t=1}$, a prior distribution for $f_{t=0}$ has to be determined. Furthermore, priors for the parameters a, φ, and the conditional variance $\left(\sigma_c^F\right)^2$ have to be defined:

$$\begin{aligned} f_0 &\sim N\left(\mu = \mu_u^F \, , \, \tau = \frac{1}{\left(\sigma_u^f\right)^2} \right) \\ a &\sim N(\mu = 0 \, , \, \tau = 0.00001) \\ \varphi &\sim N(\mu = 0 \, , \, \tau = 0.00001) [-1 \, , \, 1] \\ \left(\sigma_c^F\right)^2 &\sim \Gamma^{-1}(s = 0.001 \, , \, r = 0.001) \, . \end{aligned} \tag{3.17}$$

The adopted priors for f_0 and φ guaranty the stationarity of the process as the prior of f_0 corresponds to the unconditional posterior of f_t (Normal distribution with unconditional mean and variance) and the prior of φ is a truncated Normal distribution in the range $[-1,1]$. The prior of a is an uninformed Normal distribution with mean zero ($\mu = 0$) and low precision ($\tau = 0.00001$). The prior of the conditional variance $\left(\sigma_c^F\right)^2$ is a Inverse Gamma distribution with uninformed specification for the scale ($s = 0.001$) and rate ($r = 0.001$) parameter.

Adaption and burn-in

The model is sampled with two MCMC chains. Adaption and burn-in are set to 25,000 iterations. The posterior samples contain 100,000 iterations with a thinning of 20, resulting in chain lengths of 5,000 iterations per chain. The two chains are combined. Finally, the MCMC sample contains 10,000 iterations.[34]

[34] For the European sample, thinning is reduced to 10 with 50,000 iterations. This results in 5,000 per chain and 10,000 in the combined sample as in the sample for the US and GB.

3.B Appendix | Bayesian convergence diagnostics

To evaluate the convergence of the estimated models, we primarily adduce trace plots. Trace plots illustrates the progression of parameters in the chains. Stable trace plots indicate that the chains converge to a steady state and, thus, that priors are well calibrated and adaption, burn-in, and thinning is sufficient. Furthermore, we examine two prominent figures in Bayesian inference – the Gelman-Rubin and Heidelberger-Welch diagnostic. Both are hypotheses tests in frequentistic terms, however, applied widely to evaluate the length of burn-in (Gelman-Rubin) and the length of chains (Heidelberger-Welch).[35]

Component model

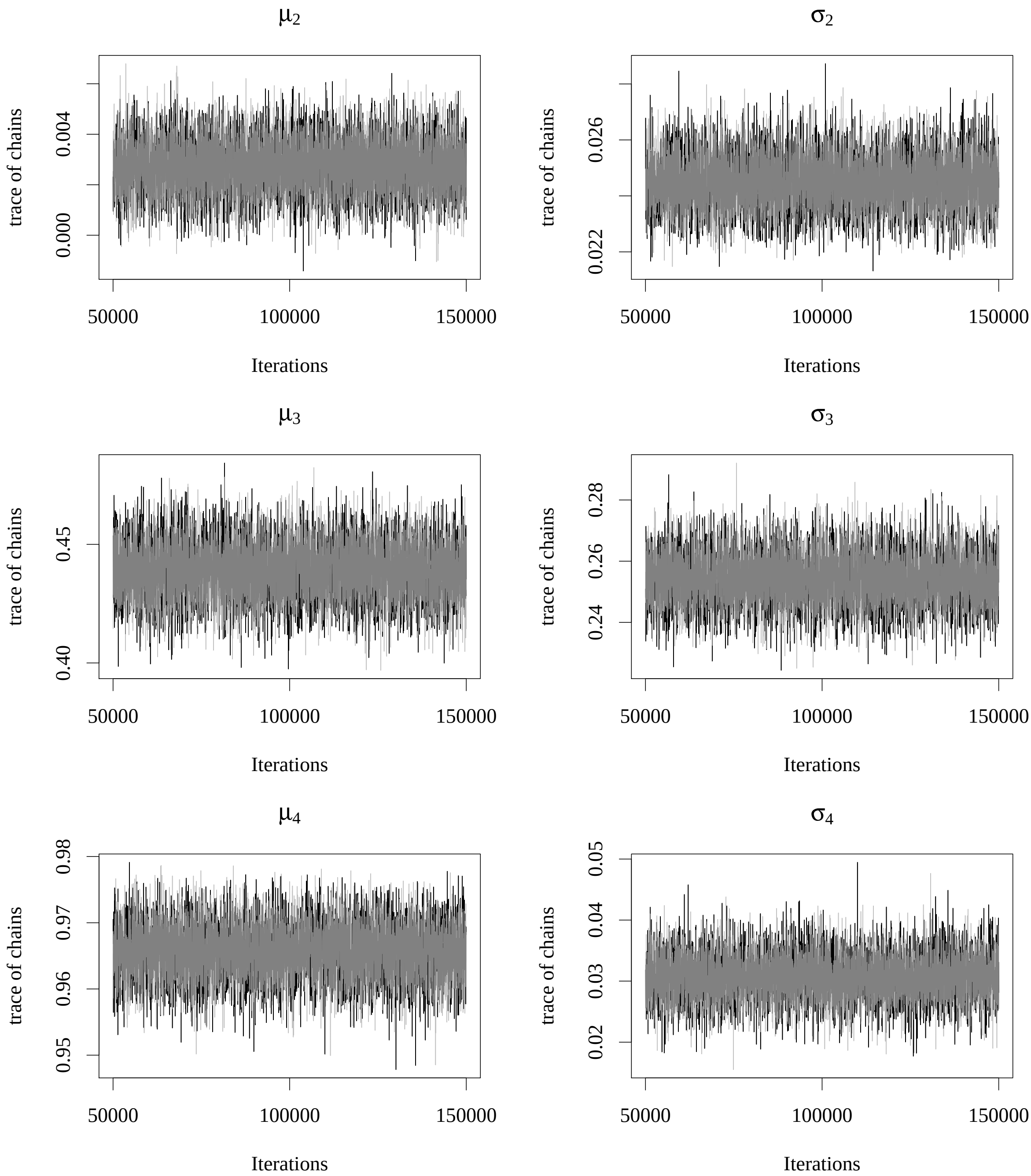
Figure 3.B.1, 3.B.2, and 3.B.3 illustrate the trace plots of the component models for the US, GB, and Europe. As the parameters of the first and fifth component are fixed, their presentation is skipped. The first chain is displayed in black, the second in gray. The trace plots seem to be stationary for every parameter. The evidence for convergence is supported by the Gelman-Rubin (see Table 3.B.1) and Heidelberger-Welch (see Table 3.B.2) diagnostic.

Probability model

Figure 3.B.4, 3.B.5, and 3.B.6 illustrate the trace plots of the probability model for the US, GB, and Europe. The presentation corresponds to Figure 3.B.1, 3.B.2, and 3.B.3. In analogy to the component model, convergence is indicated by the trace plots which is underpinned by the Gelman-Rubin (see Table 3.B.3) and Heidelberger-Welch (see Table 3.B.4) diagnostic.

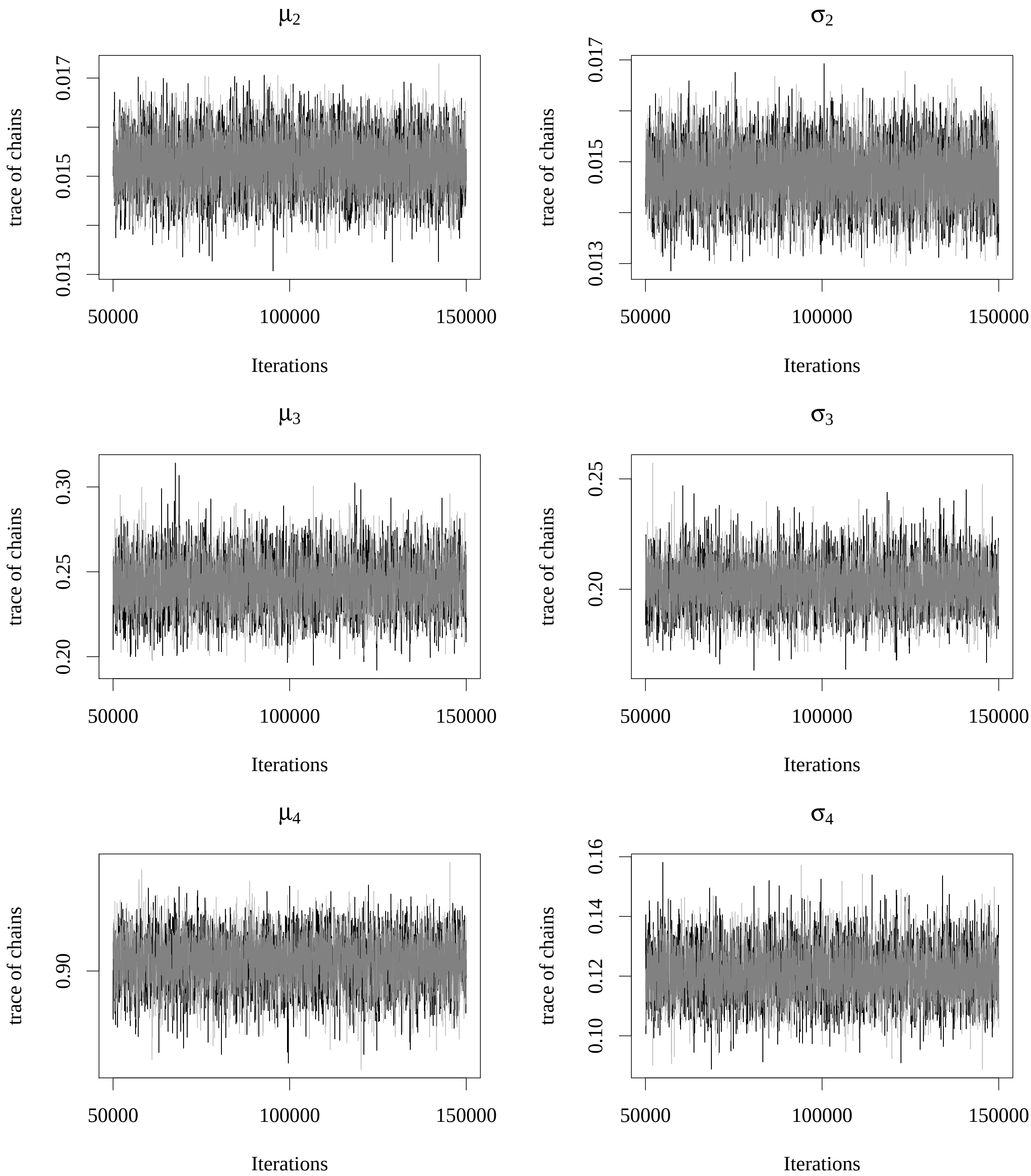
[35] We present trace plots and diagnostics for the in the paper selected model specifications – namely, specification I for the US and specification II for GB and Europe. The convergence diagnostics of the remaining model specifications are available from the authors upon request.

Figure 3.B.1: Trace plots of component model (US)



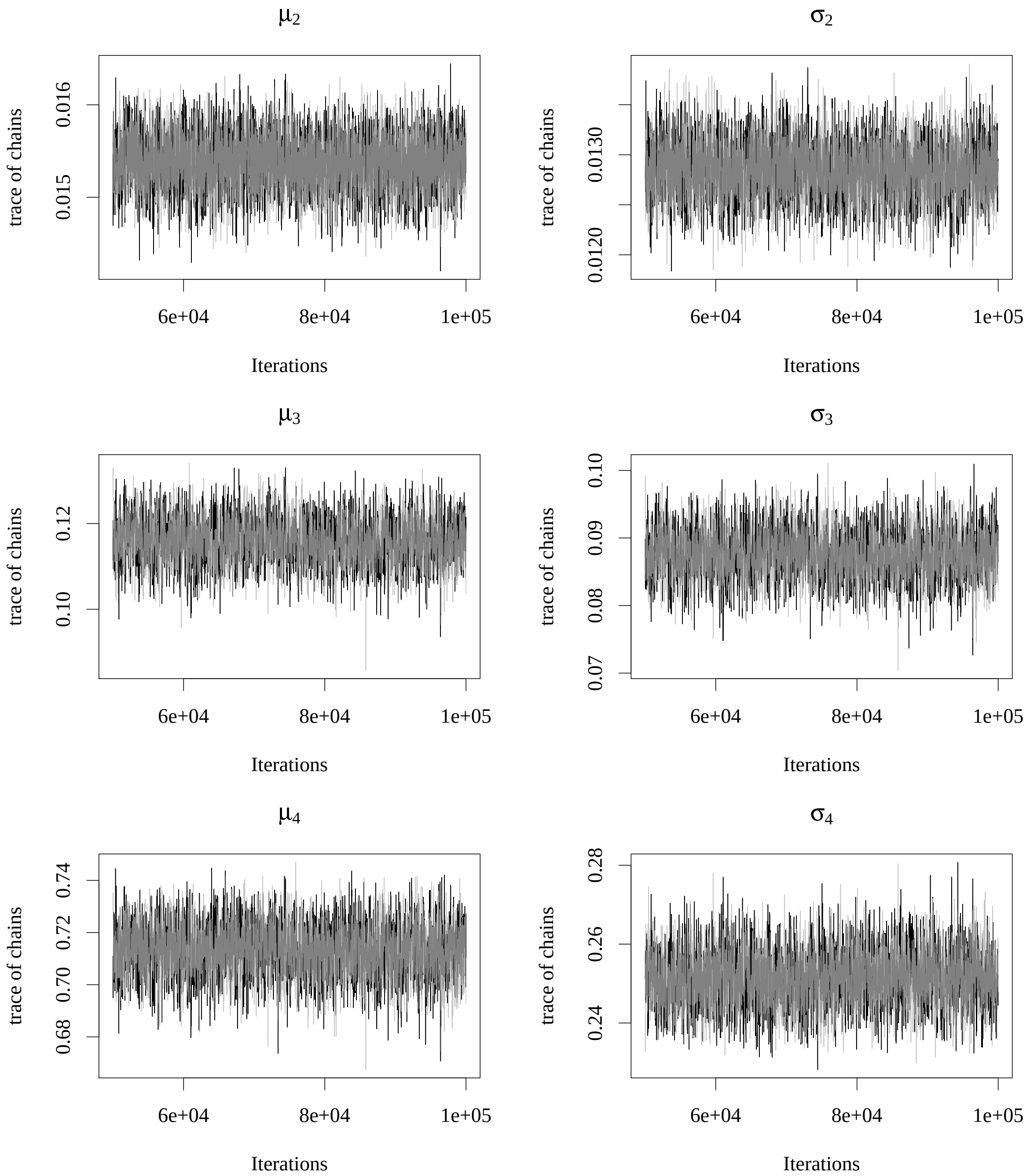
Notes: The figure illustrates the trace of MCMC chains for the estimated parameters of the component model (μ_k and σ_k for $k \in \{2,3,4\}$). The first chain is represented in black, the second in gray. Adaption phase and burn-in are dropped.

Figure 3.B.2: Trace plots of component model (GB)



Notes: The figure illustrates the trace of MCMC chains for the estimated parameters of the component model (μ_k and σ_k for $k \in \{2, 3, 4\}$). The first chain is represented in black, the second in gray. Adaption phase and burn-in are dropped.

Figure 3.B.3: Trace plots of component model (Europe)



Notes: The figure illustrates the trace of MCMC chains for the estimated parameters of the component model (μ_k and σ_k for $k \in \{2,3,4\}$). The first chain is represented in black, the second in gray. Adaption phase and burn-in are dropped.

Table 3.B.1: Gelman-Rubin diagnostic of component model

	Point estimate	Upper confidence limits (90%)
	US	
μ_2	1.0008	1.0034
μ_3	0.9999	0.9999
μ_4	1.0000	1.0000
σ_2	0.9999	0.9999
σ_3	0.9999	1.0001
σ_4	1.0010	1.0018
	GB	
μ_2	1.0001	1.0006
μ_3	0.9999	1.0000
μ_4	0.9999	1.0000
σ_2	1.0009	1.0021
σ_3	0.9999	0.9999
σ_4	1.0005	1.0024
	Europe	
μ_2	1.0001	1.0008
μ_3	1.0017	1.0049
μ_4	1.0014	1.0024
σ_2	1.0003	1.0013
σ_3	1.0015	1.0026
σ_4	1.0006	1.0013

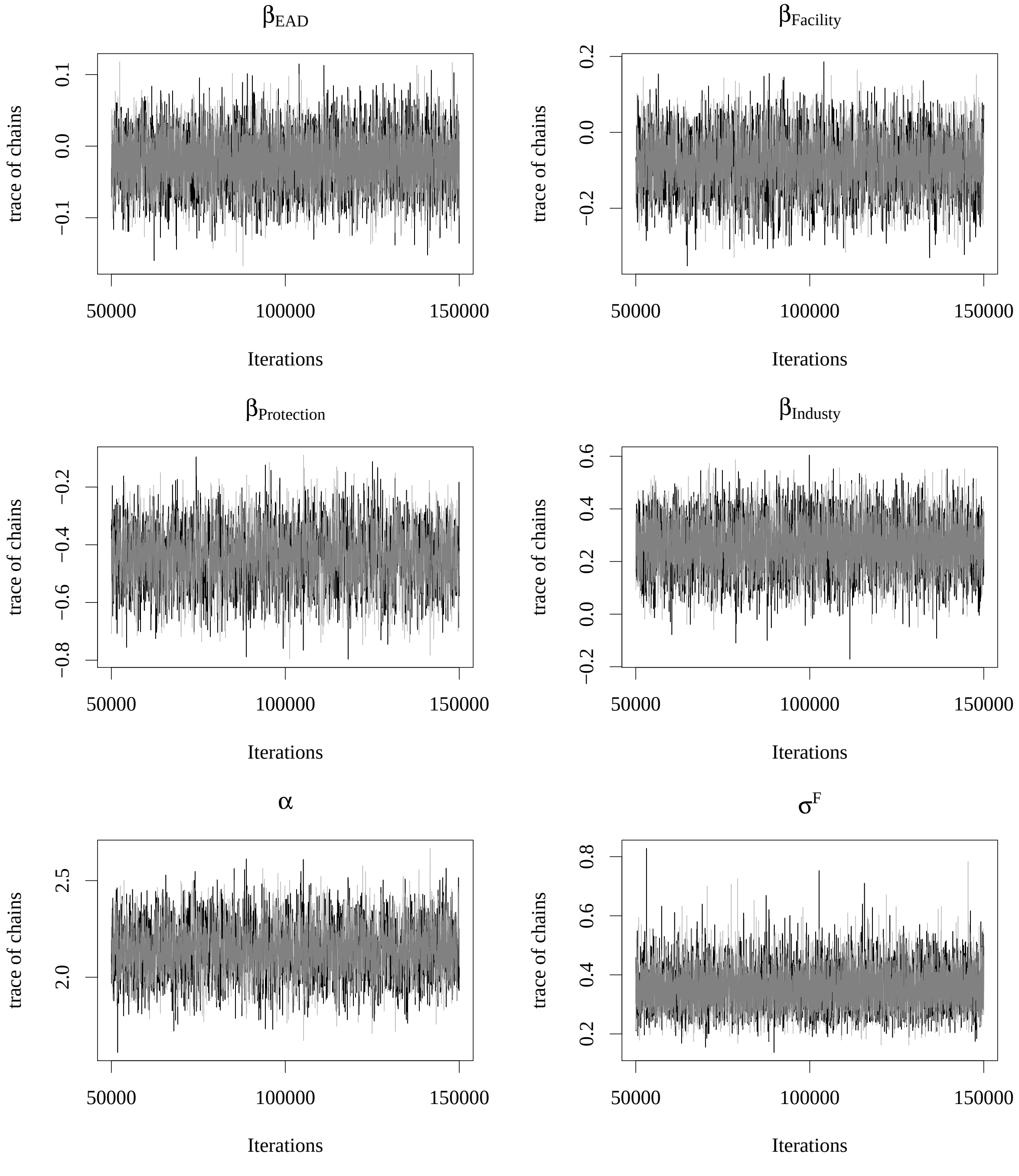
Notes: The table summarizes the results of the Gelman-Rubin diagnostic for the component models (US with specification I, GB and Europe with specification II). In the Gelman-Rubin diagnostics, the *potential reduction factor* and its upper and lower confidence limits are calculated for each variable. Convergence is diagnosed if chains have ”forgotten” their initial values, thus, for upper limits close to one (see Gelman and Rubin, 1992). A rule of thump assumes 1.1 as critical value. Generally, the Gelman-Rubin diagnostic examines the length of the burn-in.

Table 3.B.2: Heidelberger-Welch diagnostic of component model

	Stationary test	Start	p-value	Halfwidth mean test	Mean	Halfwidth
			US			
μ_2	passed	1	0.0835	passed	0.0027	0.0000
μ_3	passed	1	0.2402	passed	0.4388	0.0002
μ_4	passed	1	0.8241	passed	0.9657	0.0001
σ_2	passed	1	0.9700	passed	0.0245	0.0000
σ_3	passed	1	0.8521	passed	0.2542	0.0002
σ_4	passed	1	0.4306	passed	0.0306	0.0001
			GB			
μ_2	passed	1	0.6975	passed	0.0153	0.0000
μ_3	passed	1	0.9433	passed	0.2427	0.0004
μ_4	passed	1	0.5247	passed	0.9039	0.0003
σ_2	passed	1	0.3990	passed	0.0147	0.0000
σ_3	passed	1	0.7053	passed	0.2016	0.0003
σ_4	passed	1	0.2619	passed	0.1205	0.0002
			Europe			
μ_2	passed	1	0.1212	passed	0.0154	0.0000
μ_3	passed	1001	0.0798	passed	0.1157	0.0002
μ_4	passed	1	0.1004	passed	0.7132	0.0004
σ_2	passed	1	0.1024	passed	0.0128	0.0000
σ_3	passed	1	0.1068	passed	0.0875	0.0001
σ_4	passed	1	0.1728	passed	0.2515	0.0003

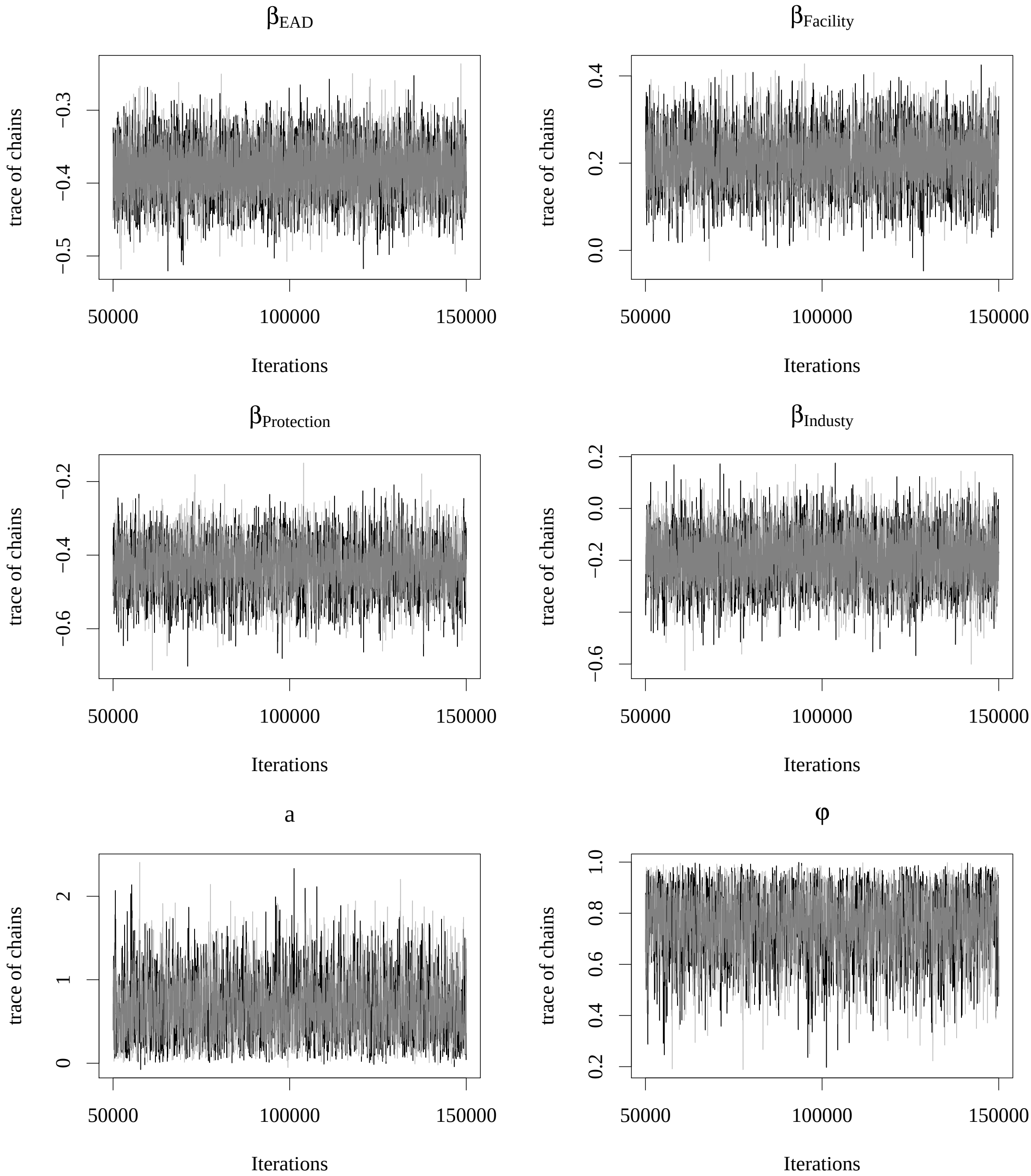
Notes: The table summarizes the results of the Heidelberger-Welch diagnostic for the component model (US with specification I, GB and Europe with specification II). Therefore, the two chains are combined. In the Heidelberger-Welch diagnostic, a criterion of relative accuracy for the posterior means is calculated. The frequentistic stationary test adopts the Cramer-von-Mises statistic to test the null hypotheses that the sampled values origin from a stationary process (see Heidelberger and Welch, 1981, 1983). Generally, the Heidelberger-Welch diagnostic examines the length of the chain.

Figure 3.B.4: Trace plots of probability model (US)



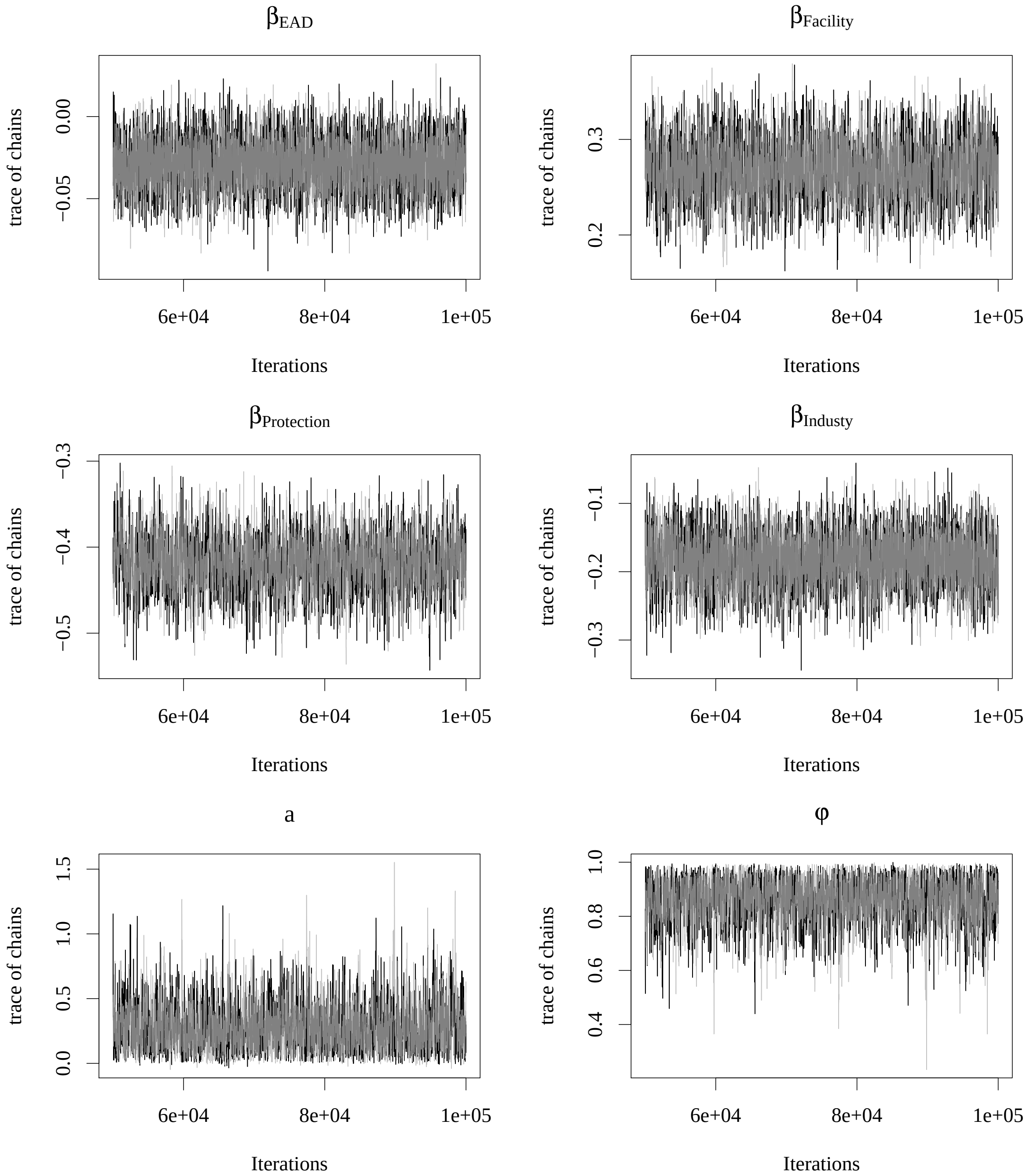
Notes: The figure illustrates the trace of MCMC chains for the estimated parameters of the probability model (β_{EAD}, β_{Facility}, $\beta_{\text{Protection}}$, β_{Industry}, α, and σ^F for specification I). The first chain is represented in black, the second in gray. Adaption phase and burn-in are dropped.

Figure 3.B.5: Trace plots of probability model (GB)



Notes: The figure illustrates the trace of MCMC chains for the estimated parameters of the probability model (β_{EAD}, β_{Facility}, $\beta_{\text{Protection}}$, β_{Industry}, a, and φ for specification II). The first chain is represented in black, the second in gray. Adaption phase and burn-in are dropped.

Figure 3.B.6: Trace plots of probability model (Europe)



Notes: The figure illustrates the trace of MCMC chains for the estimated parameters of the probability model (β_{EAD}, β_{Facility}, $\beta_{\text{Protection}}$, β_{Industry}, a, and φ for specification II). The first chain is represented in black, the second in gray. Adaption phase and burn-in are dropped.

Table 3.B.3: Gelman-Rubin diagnostic of probability model

	Point estimate	Upper confidence limits (90%)
	US	
β_{EAD}	1.0005	1.0017
β_{Facility}	1.0018	1.0066
$\beta_{\text{Protection}}$	1.0016	1.0019
β_{Industry}	1.0005	1.0008
α	0.9999	1.0000
σ^F	1.0005	1.0020
c_2	1.0003	1.0006
c_3	1.0000	1.0004
	GB	
β_{EAD}	0.9999	1.0000
β_{Facility}	1.0012	1.0027
$\beta_{\text{Protection}}$	1.0011	1.0044
β_{Industry}	1.0002	1.0010
a	1.0001	1.0001
φ	1.0002	1.0003
c_2	1.0001	1.0001
c_3	0.9999	1.0000
	Europe	
β_{EAD}	1.0018	1.0037
β_{Facility}	1.0002	1.0002
$\beta_{\text{Protection}}$	1.0038	1.0141
β_{Industry}	1.0000	1.0000
a	1.0008	1.0008
φ	1.0009	1.0009
c_2	1.0010	1.0042
c_3	1.0008	1.0030

Notes: The table summarizes the results of the Gelman-Rubin diagnostic for the probability model (US with specification I, GB and Europe with specification II). In the Gelman-Rubin diagnostic, the *potential reduction factor* and its upper and lower confidence limits are calculated for each variable. Convergence is diagnosed if chains have ”forgotten” their initial values, thus, for upper limits close to one (see Gelman and Rubin, 1992). A rule of thump assumes 1.1 as critical value. Generally, the Gelman-Rubin diagnostic examines the length of the burn-in.

Table 3.B.4: Heidelberger-Welch diagnostic of probability model

	Stationary test	Start	p-value	Halfwidth mean test	Mean	Halfwidth
			US			
β_{EAD}	passed	1	0.3352	passed	-0.0215	0.0009
β_{Facility}	passed	1	0.1967	passed	-0.0832	0.0022
$\beta_{\text{Protection}}$	passed	1	0.6406	passed	-0.4442	0.0038
β_{Industry}	passed	1	0.6237	passed	0.2592	0.0025
α	passed	1	0.7012	passed	2.1402	0.0045
σ^F	passed	1	0.2197	passed	0.3522	0.0014
c_2	passed	1	0.2220	passed	2.4103	0.0007
c_3	passed	1	0.3240	passed	3.9466	0.0010
			GB			
β_{EAD}	passed	1	0.2285	passed	-0.3825	0.0009
β_{Facility}	passed	1	0.3353	passed	0.2093	0.0018
$\beta_{\text{Protection}}$	passed	1	0.1387	passed	-0.4314	0.0021
β_{Industry}	passed	1	0.4385	passed	-0.1930	0.0025
a	passed	1	0.8255	passed	0.6779	0.0144
φ	passed	1	0.8304	passed	0.7504	0.0051
c_2	passed	1	0.8060	passed	2.7397	0.0007
c_3	passed	1	0.5804	passed	3.6150	0.0009
			Europe			
β_{EAD}	passed	1	0.1500	passed	-0.0291	0.0004
β_{Facility}	passed	1	0.9691	passed	0.2685	0.0012
$\beta_{\text{Protection}}$	passed	1	0.0581	passed	-0.4184	0.0014
β_{Industry}	passed	1	0.8401	passed	-0.1840	0.0013
a	passed	1	0.9515	passed	0.2874	0.0084
φ	passed	1	0.9346	passed	0.8565	0.0040
c_2	passed	4001	0.0638	passed	2.6592	0.0007
c_3	passed	2001	0.0859	passed	3.2198	0.0009

Notes: The table summarizes the results of the Heidelberger-Welch diagnostic for the probability model (US with specification I, GB and Europe with specification II). Therefore, the two chains are combined. In the Heidelberger-Welch diagnostic, a criterion of relative accuracy for the posterior means is calculated. The frequentistic stationary test adopts the Cramer-von-Mises statistic to test the null hypotheses that the sampled values origin from a stationary process (see Heidelberger and Welch, 1981, 1983). Generally, the Heidelberger-Welch diagnostic examines the length of the chain.

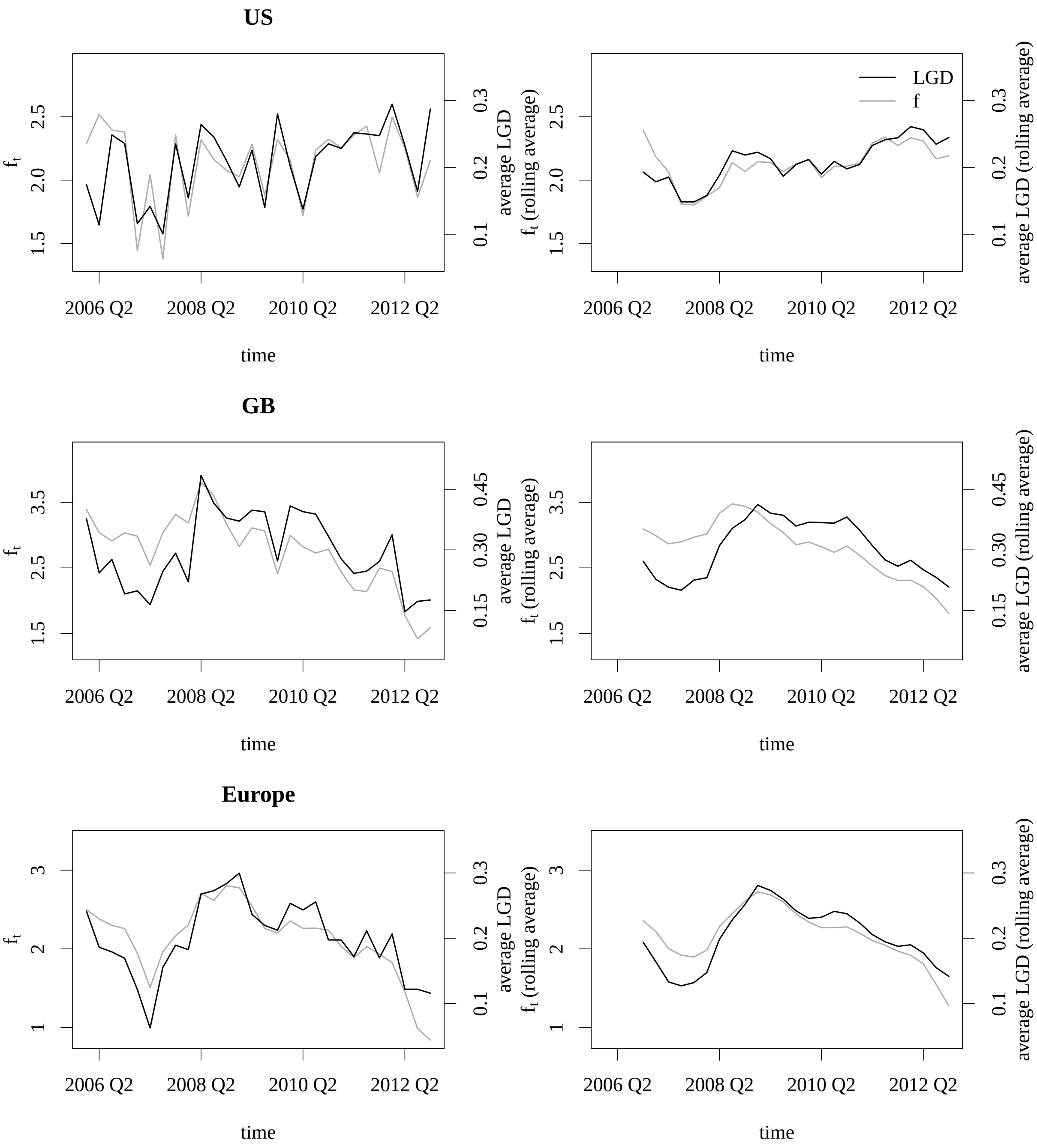
3.C Appendix | Further figures

Figure 3.C.1: Exemplary JAGS model file

```
data{
muComp[1]                                    <- 0
muComp[Ncomp]                                <- 1
sigmaComp[1]                                 <- 0.001
sigmaComp[Ncomp]                             <- 0.001
}
model{
for( i in 1:N ){
y[i]                                         ~ dnorm( mu[i] , tau[i] )
mu[i]                                        <- muComp[ z[i] ]
tau[i]                                       <- pow( sigma[i] , -2 )
sigma[i]                                     <- sigmaComp[ z[i] ]
z[i]                                         ~ dcat( p[i,] )
logit(Q[i,1])                                <- c[1] - zstern[i]
p[i,1]                                       <- Q[i,1]
for( j in 2:(Ncomp-1) ){ logit(Q[i,j])       <- c[j] - zstern[i]
                         p[i,j]              <- Q[i,j] - Q[i,j-1] }
p[i,Ncomp]                                   <- 1 - Q[i,(Ncomp-1)]
zstern[i]                                    <- inprod( X[i,] , beta ) + f[time[i]]
}
beta[1:Npara]                                ~ dmnorm(betamu[],betaTau[,])
c[1]                                         <- 1.5
for( j in 1:(Ncomp-3) ){ c0[j]               ~ dnorm(3,0.00001)I(c[1],c[Ncomp-1]) }
c[2:(Ncomp-2)]                               <- sort(c0[1:(Ncomp-3)])
c[Ncomp-1]                                   <- 4.5
#####################################################################################################
for( t in 1:Ntime ){ f[t]                    ~ dnorm( fmu[t] , ftau ) }
fmu[1]                                       <- ar0 + ar1 * f0
for( t in 2:Ntime ){ fmu[t]                  <- ar0 + ar1 * f[t-1] }
f0                                           ~ dnorm( fmuU , ftauU )
ar0                                          ~ dnorm( 0 , 0.00001 )
ar1                                          ~ dnorm( 0 , 0.00001 )I( -1 , 1 )
fsigma                                       ~ dgamma( 0.001 , 0.001 )
ftau                                         <- pow( fsigma , -2 )
#####################################################################################################
for( j in 1:(Ncomp-2) ){ muComp0[j]          ~ dnorm( 0 , 0.00001 )I( -0.5 , 1.5 ) }
muComp[2:(Ncomp-1)]                          <- sort(muComp0[1:(Ncomp-2)])
for( j in 2:(Ncomp-1) ){ sigmaComp[j]        ~ dgamma( 0.001 , 0.001 )I( 0.0001 , 1 ) }
}
```

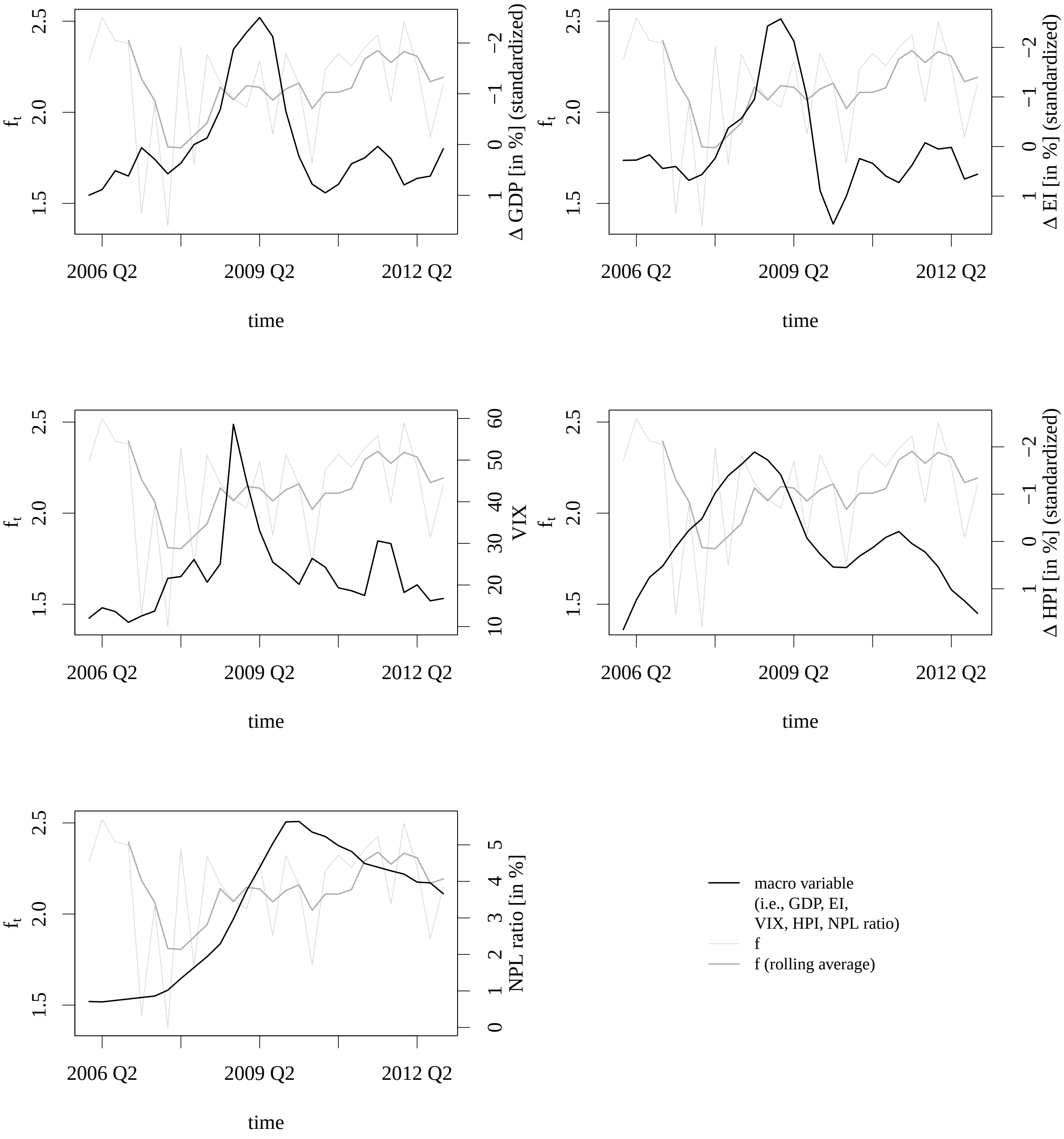
Notes: The figure illustrates an exemplary JAGS model file for specification II. For the implementation of specification I, the random effect has to be modified to a normal distribution (lines surrounded by ### in the code).

Figure 3.C.2: Random effect vs. average LGD



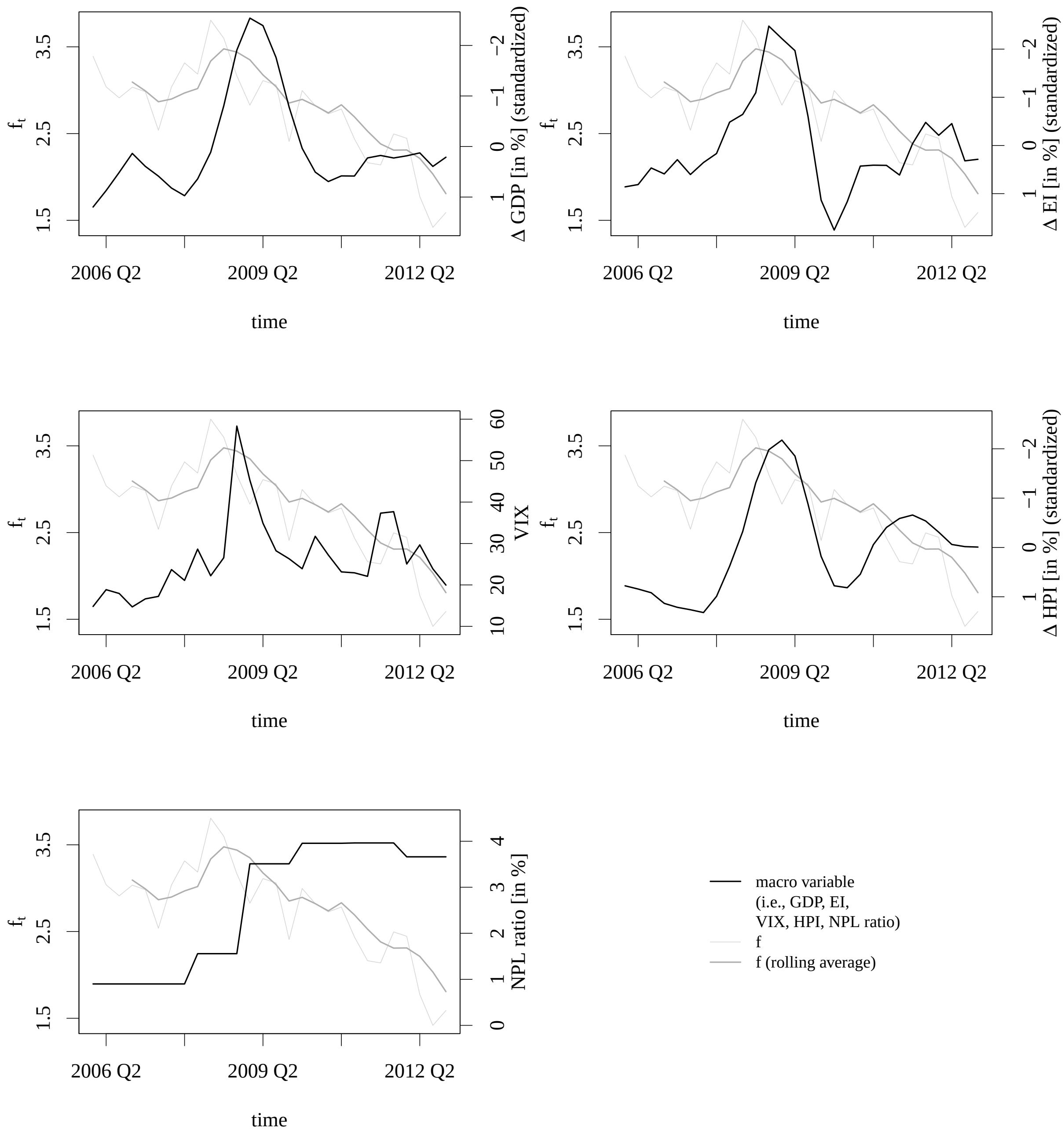
Notes: The figure illustrates the synchronism of the estimated posterior means of the random effect (specification I for the US and specification II for GB and Europe) and the average LGD in time line. In the left panels of the figure, the course of the random effect (gray line) and the quarterly average LGD (black line) is displayed. For representational purpose the rolling average of the random effect (gray line) and the rolling quarterly average of the LGD (black line) is plotted in the right panels.

Figure 3.C.3: Random effect vs. macro variables (US)



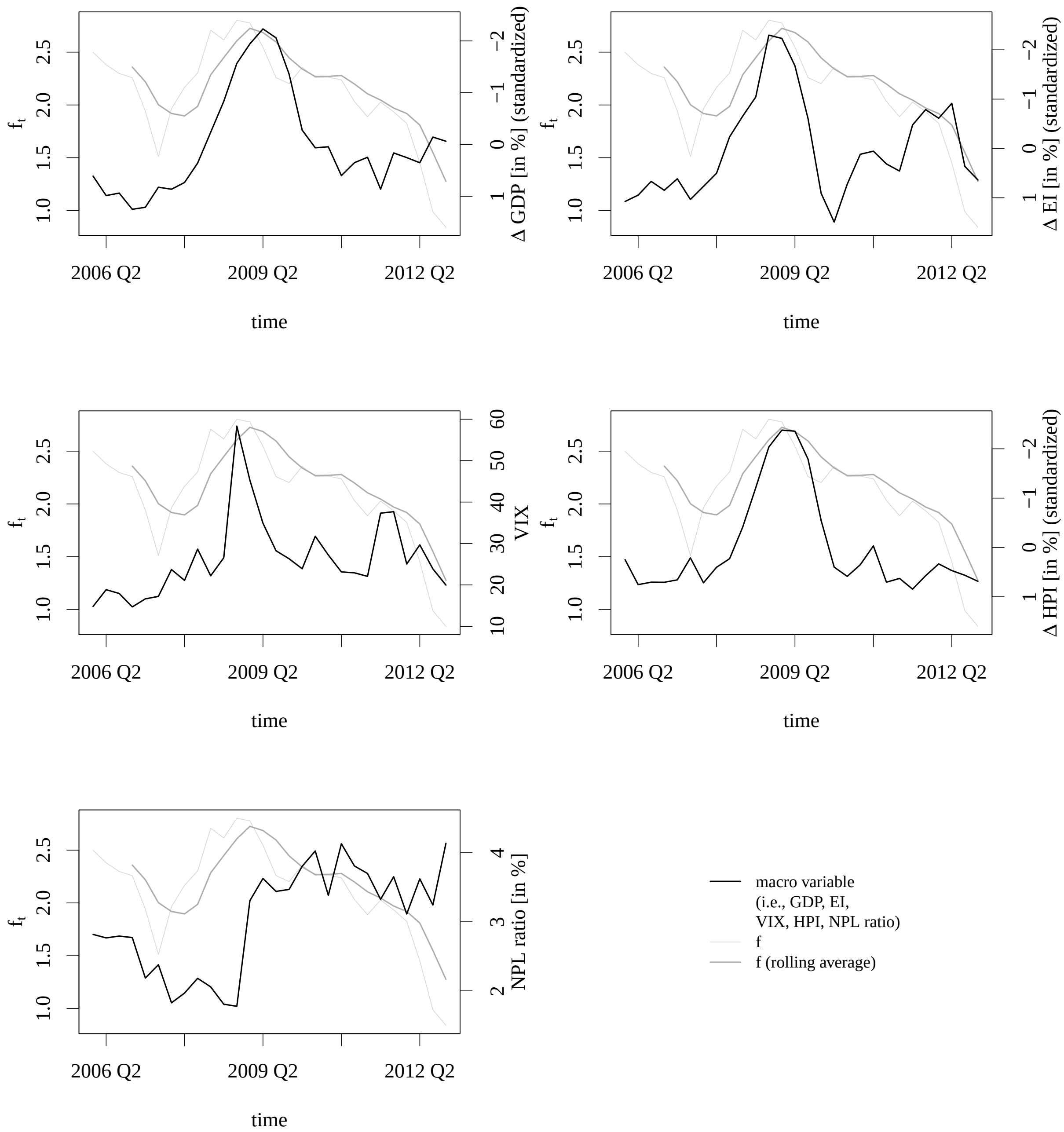
Notes: The figure contrasts the course of the macro variables and the random effect over time. The macro variables (i.e., GDP, S&P 500, VIX for the US, HPI, and NPL ratio) are represented by the black lines. The thin gray lines map the posterior means of the random effect (specification I). For representation purpose, the rolling average of the posterior means is illustrated by thick gray lines.

Figure 3.C.4: Random effect vs. macro variables (GB)



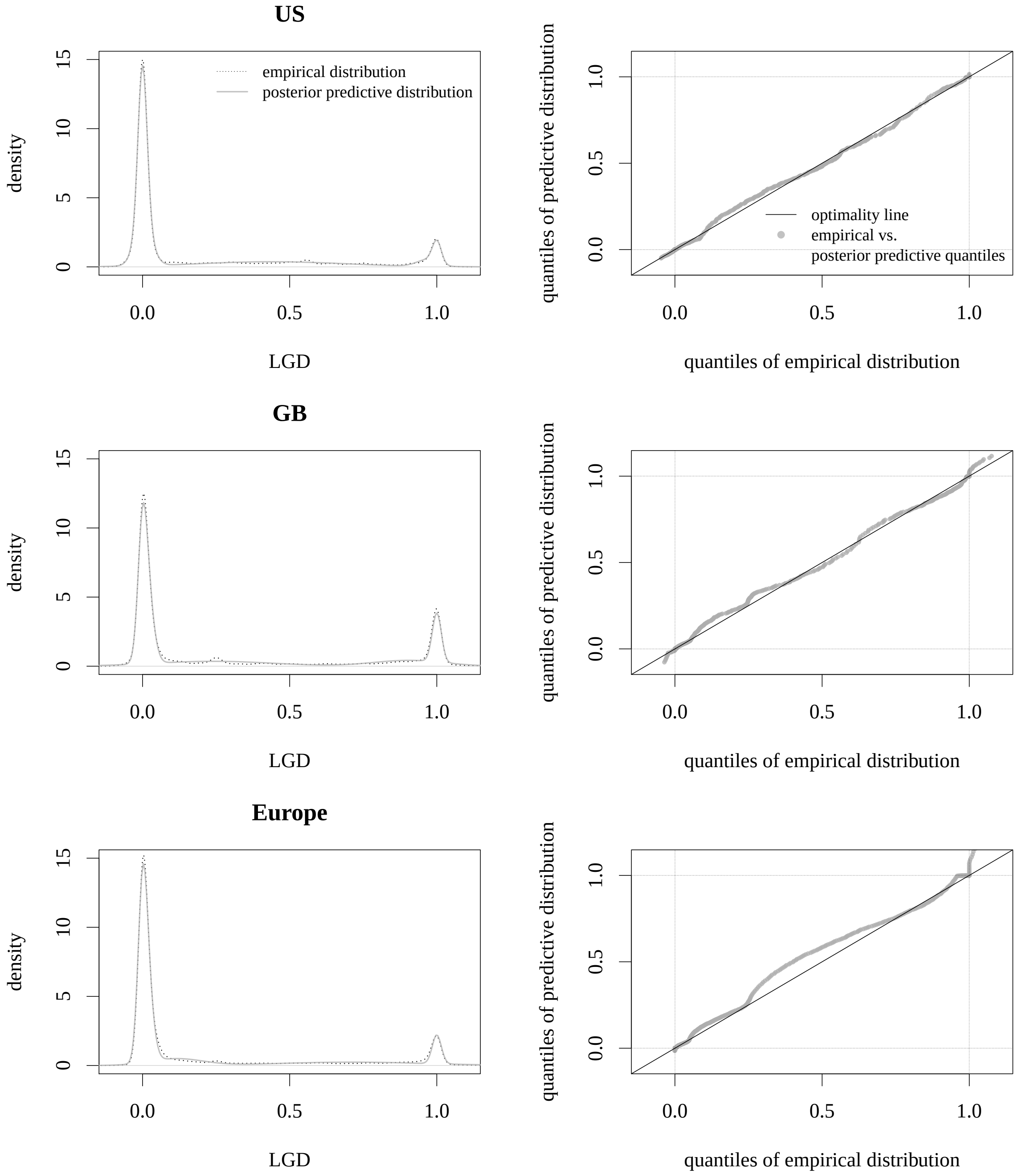
Notes: The figure contrasts the course of the macro variables and the random effect over time. The macro variables (i.e., GDP, FTSE, VIX for Europe, HPI, and NPL ratio) are represented by the black lines. The thin gray lines map the posterior means of the random effect (specification II). For representation purpose, the rolling average of the posterior means is illustrated by thick gray lines.

Figure 3.C.5: Random effect vs. macro variables (Europe)



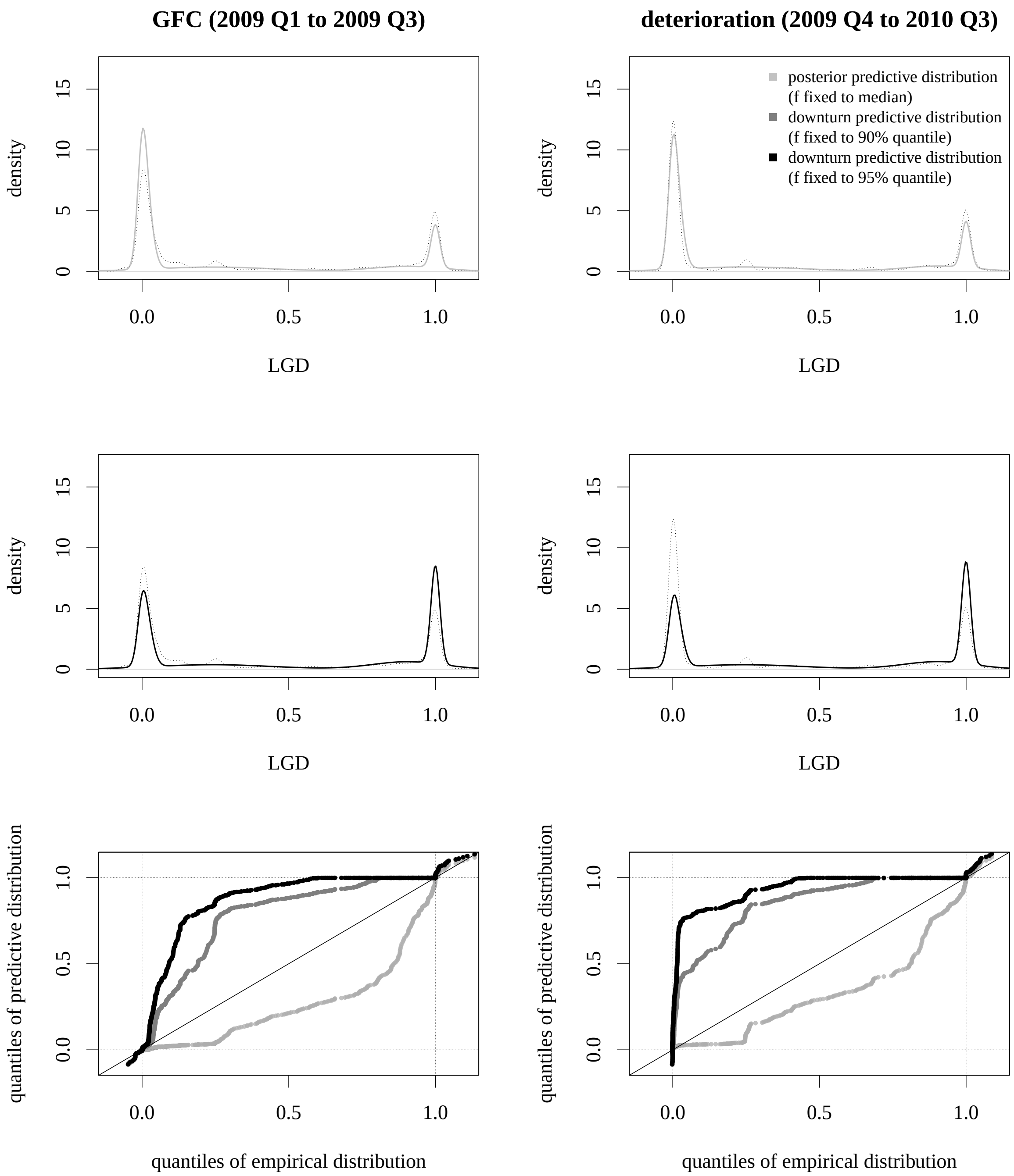
Notes: The figure contrasts the course of the macro variables and the random effect over time. The macro variables (i.e., the weighted average of GDP, EIs, HPIs, and NPL ratio as well as the VIX for Europe) are represented by the black lines. The thin gray lines map the posterior means of the random effect (specification II). For representation purpose, the rolling average of the posterior means is illustrated by thick gray lines.

Figure 3.C.6: Posterior predictive distribution



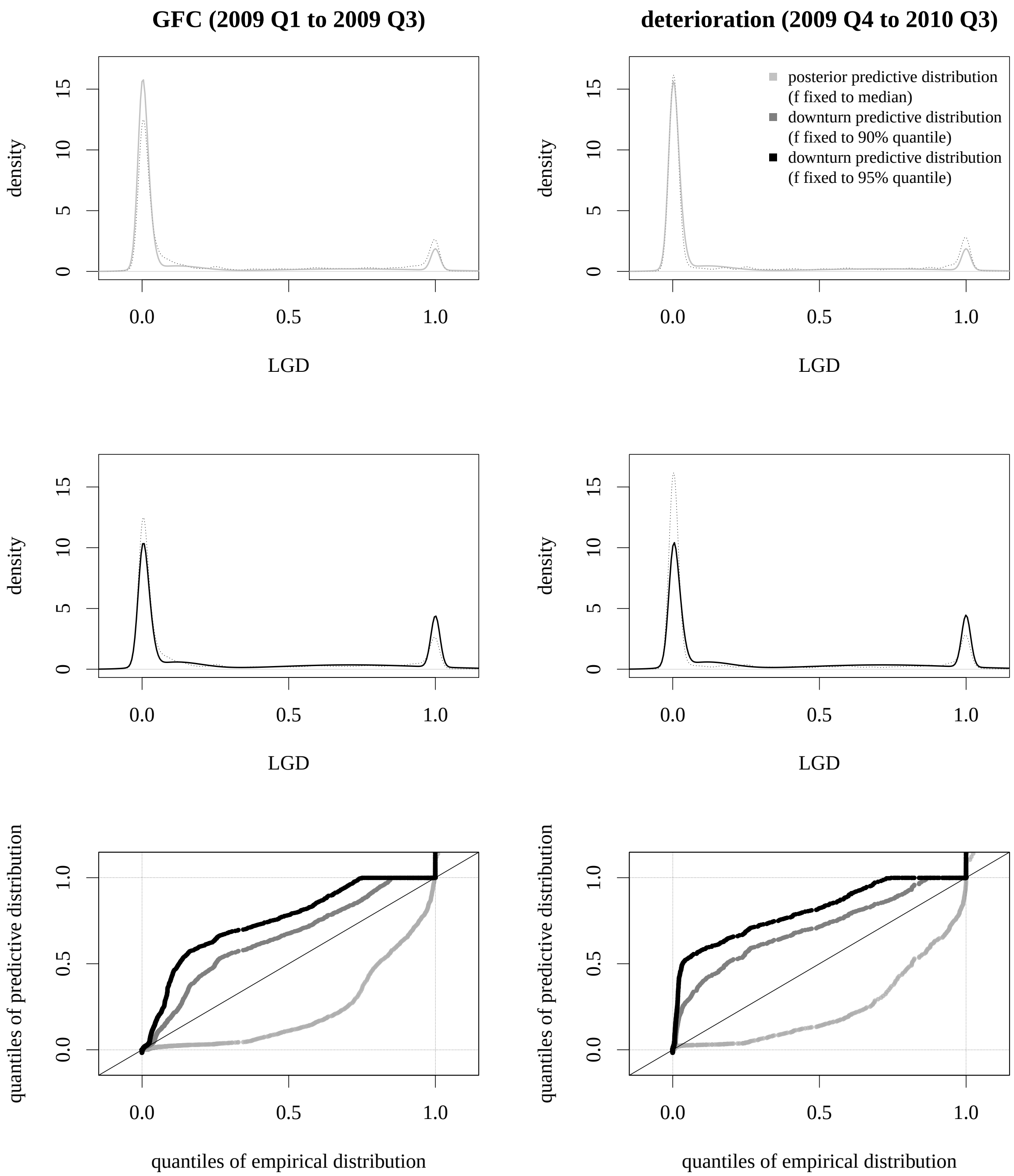
Notes: The figure illustrates the posterior predicted distribution for the US American, British, and European data set. In the left panels, the kernel density estimates of the posterior predicted (gray line) is contrasted to the kernel density estimate of the empirical data (dotted line). The band widths are fixed to 0.015 for both kernel density estimates to ensure comparability regarding the height of the density. Thus, the curves are comparable in spite of ties. As differences are hard to identify due to the high probability masses at the two modi, quantile-quantile (qq) plots are presented in the right panels. The gray dots represent the quantiles of the empirical data plotted against the quantiles of the posterior predicted distribution. The bisector (black line) represents optimality.

Figure 3.C.7: Posterior and downturn distribution for the GFC and a deterioration period (GB)



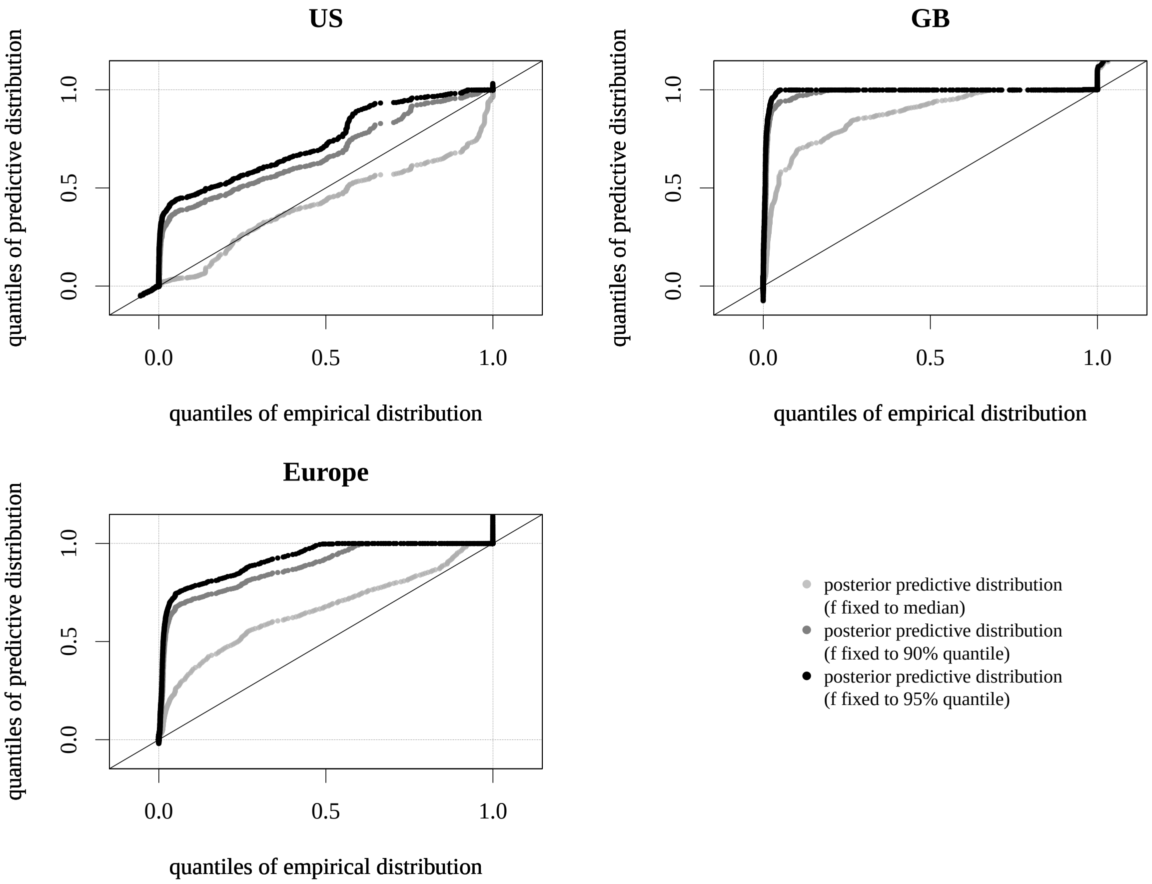
Notes: The figure contrasts the empirical LGD distribution to the posterior (light gray) and downturn (dark gray and black) predictive distribution for the GFC (2009 Q1 to 2009 Q3, left panels) and a deterioration period (2009 Q4 to 2010 Q3, right panels). The upper panels display the kernel density estimates, the lower panels the quantile-quantile (qq) plots. The kernel density estimates for the downturn predictive distribution for the 90% quantile of the random effect is skipped and available form the authors upon request.

Figure 3.C.8: Posterior and downturn distribution for the GFC and a deterioration period (Europe)



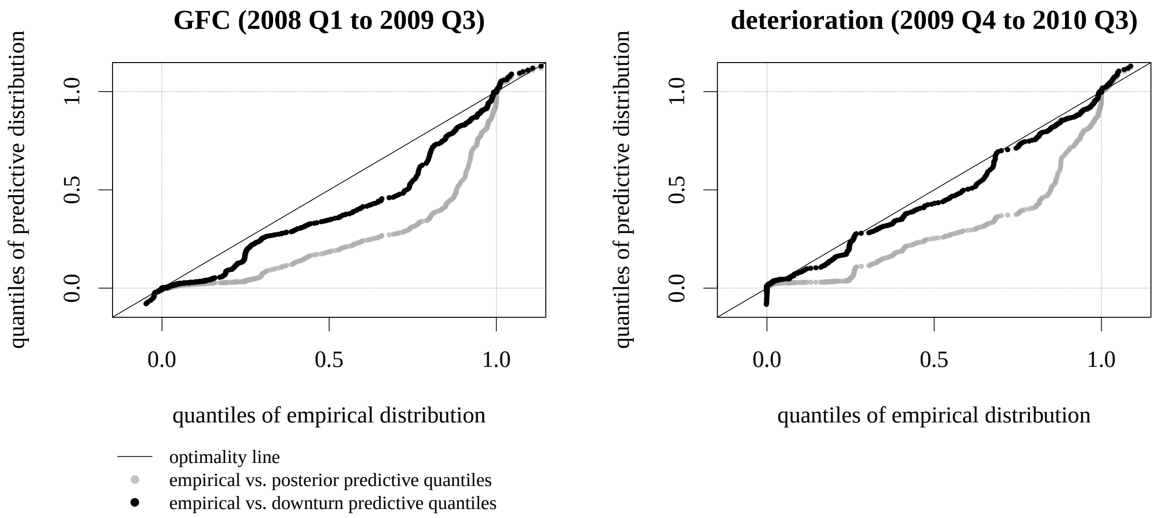
Notes: The figure contrasts the empirical LGD distribution to the posterior (light gray) and downturn (dark gray and black) predictive distribution for the GFC (2009 Q1 to 2009 Q3, left panels) and a deterioration period (2009 Q4 to 2010 Q3, right panels). The upper panels display the kernel density estimates, the lower panels the quantile-quantile (qq) plots. The kernel density estimates for the downturn predictive distribution for the 90% quantile of the random effect is skipped and available form the authors upon request.

Figure 3.C.9: Posterior and downturn distribution based on out-of-time estimation



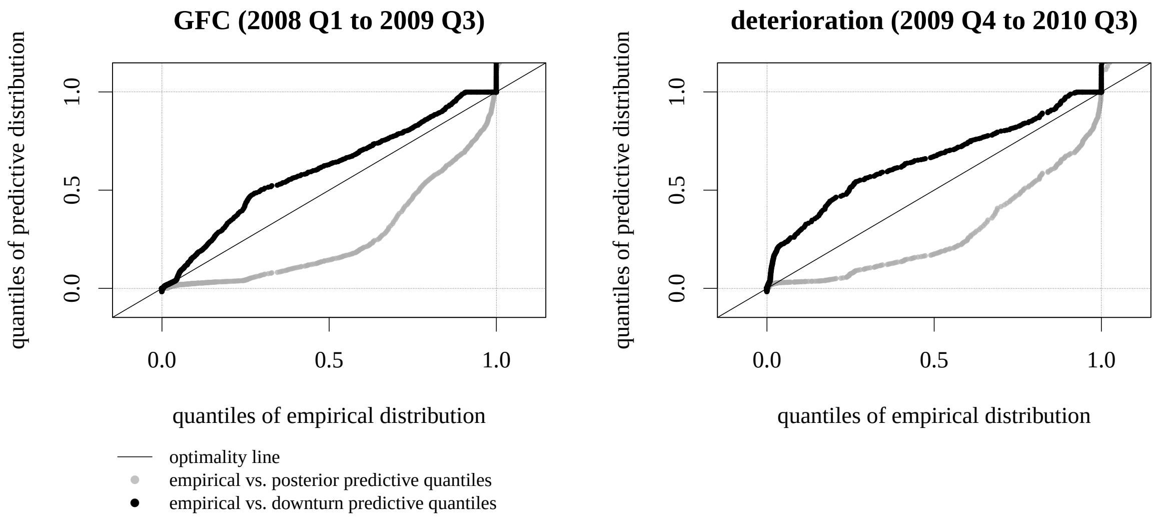
Notes: The figure contrasts the empirical LGD distribution to the posterior (light gray) and downturn (dark gray and black) predictive distribution based on out-of-time estimation. The training set contains the time period from 2006 Q1 to 2010 Q1. The test set consists of the time period from 2010 Q2 to 2012 Q4.

Figure 3.C.10: Posterior and downturn distribution for the GFC and a deterioration period based on the macro model containing the EI (GB)



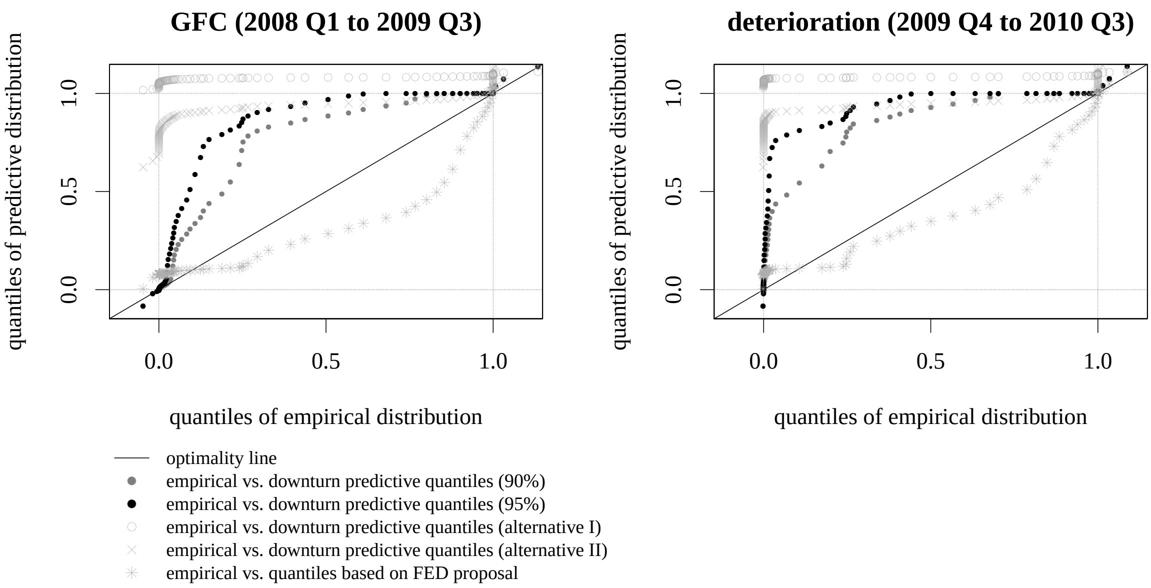
Notes: The figure contrasts the empirical LGD distribution to the posterior (light gray) and downturn (dark gray and black) predictive distribution of the macro model containing the EI instead of a random effect for the GFC (2008 Q1 to 2009 Q3, left panel) and a deterioration period (2009 Q4 to 2010 Q3, left panel).

Figure 3.C.11: Posterior and downturn distribution for the GFC and a deterioration period based on the macro model containing the HPI (Europe)



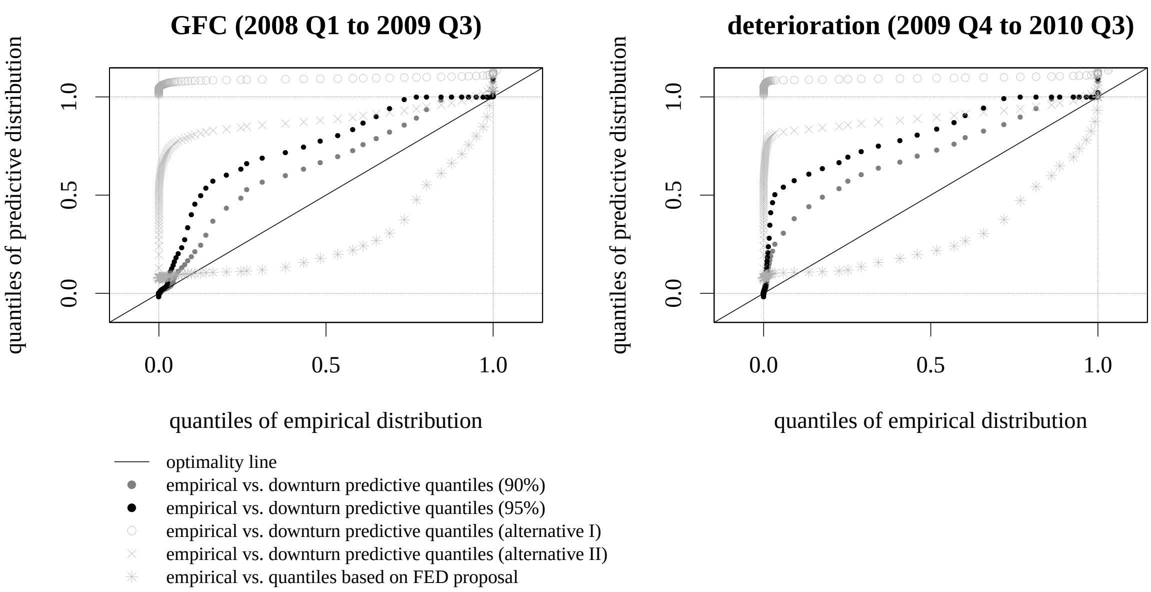
Notes: The figure contrasts the empirical LGD distribution to the posterior (light gray) and downturn (dark gray and black) predictive distribution of the macro model containing the HPI instead of a random effect for the GFC (2008 Q1 to 2009 Q3, left panel) and a deterioration period (2009 Q4 to 2010 Q3, left panel).

Figure 3.C.12: Downturn distribution for the GFC and a deterioration period based on alternative concepts (GB)



Notes: The figure contrasts the empirical LGD distribution to downturn predictive distributions. The black and gray dots represent the downturn approach via a random effect, whereas, the suggestion of Bijak and Thomas (2015) (alternative I) is displayed by gray cycles and the proposal of Calabrese (2014) (alternative II) by gray crosses. The FED approach is displayed by gray stars. The figure refers to the GFC (2008 Q1 to 2009 Q3, left panel) and a deterioration period (2009 Q4 to 2010 Q3, right panel).

Figure 3.C.13: Downturn distribution for the GFC and a deterioration period based on alternative concepts (Europe)



Notes: The figure contrasts the empirical LGD distribution to downturn predictive distributions. The black and gray dots represent the downturn approach via a random effect, whereas, the suggestion of Bijak and Thomas (2015) (alternative I) is displayed by gray cycles and the proposal of Calabrese (2014) (alternative II) by gray crosses. The FED approach is displayed by gray stars. The figure refers to the GFC (2008 Q1 to 2009 Q3, left panel) and a deterioration period (2009 Q4 to 2010 Q3, right panel).

3.D Appendix | Further tables

Table 3.D.1: Literature review

	Data	Method	Country	Time	Macro variables	Result
Acharya et al. (2007)	Bonds and loans (market-based RRs)	Regression	US	1982–1999	• Industry macro variables • Aggregates default rates	• Industry distress dummies (–***) • Aggregated macro variables lose significance
Altman and Kalotay (2014)	Bonds and loans (workout RRs)	Bayesian FMM	US	1987–2011	Industry specific default likelihood	Statistical evidence not reported
Bastos (2010)	Loans (workout RRs)	• Regression • Regression trees	Portugal	1995–2000	No	
Bijak and Thomas (2015)	Loans (workout LGDs)	Bayesian FMM	UK	1987–1998	No	
Brumma et al. (2014)	Loans (workout LGDs)	Regression	International	2000–2010	Varity of macro variables	Significance only for cash flow weighted LGDs
Calabrese (2014)	Loans (workout RRs)	FMM	Italy	1975–1999	No	
Caselli et al. (2008)	Loans (workout LGDs)	Regression	Italy	1990–2004	Varity of macro variables	SMEs: Macro variables statistically or economically non significant
Bellotti and Crook (2012)	Personal credit cards (workout RRs)	• Regression (Tobit) • Regression trees	UK	1999–2005	• Interest rate • Unemployment • Earnings growth	• Interest rate (–***) • Unemployment (–***) Not economically significant • Earnings growth (+*)
Dermine and Neto de Carvalho (2006)	Loans (workout RRs)	Regression	Portugal	1995–2000	• GDP • Default rate	No statistical significance
Grunert and Weber (2009)	Loans (workout RRs)	Regression	Germany	1992-2003	• GDP • Loss provisions • Unemployment	No statistical significance
Gürtler and Hibbeln (2013)	Loans (workout RRs)	Regression (model improvements)	Germany	2006–2008	No	
Jankowitsch et al. (2014)	Bonds (market-based RRs)	Regression	US	2002–2010	• Market default rate • Industry default rate • Federal funds rate • Slope of term structure	• Market default rate (–***) • Industry default rate (–***) • Federal funds rate (+***) • Slope of term structure (+***)
Krüger and Rösch (2017)	Loans (workout LGDs)	Quantile regression	US	2000–2014	• S&P 500 • TED spread • Term spread • VIX	• S&P 500 (significant for inner quantiles) • TED spread (significant for inner quantiles)
Leow et al. (2014)	Mortgages and loans (workout LGDs)	• Two-stage model • Regression	UK	1990–2002 (mortgages) 1989–1999 (loans)	Varity of macro variables (with different time stamps)	Mixed results regarding significance (might be due to non linear impacts)
Matuszyk et al. (2010)	Loans (workout LGDs)	Two-stage model	UK	1989–2004	No	
Nazemi et al. (2017)	Bonds (market-based LGDs)	Fuzzy decision fusion	US	2002–2012	Principal components of 104 macro variables	Principal components have high contribution to R^2
Qi and Yang (2009)	Mortgages (workout LGDs)	Regression	US	1990–2003	• HPI • Stress dummy • Current LTV	• HPI (not included) • Stress dummy (+***) • Current LTV (+***)
Qi and Zhao (2011)	Bonds (market-based and workout LGDs)	Varity of methods	US	1985–2008	• (Industry) distance to default • (Industry) default rate • Market return • Interest rate	• Industry distance to default (–***) • Default rate (+***) • Market return (–***) • Interest rate (+***)
Somers and Whittaker (2007)	Mortgages (workout LGDs)	Quantile regression	EU	since 1990	No	
Tobback et al. (2014)	Loans (workout LGDs)	• Regression • Regression trees • Non linear models	US	1984–2011	Varity of macro variables	Statistical significance not reported
Yao et al. (2015)	Bonds (workout LGDs)	Support vector regression	US	1985–2012	• GDP • Unemployment • S&P 500 • Interest rate	Statistical significance not reported
Yao et al. (2017)	Credit cards (workout LGDs)	Support vector machines	UK	2009–2010	• Unemployment • CPI • HPI	• Unemployment (+***) • CPI (–***) • HPI (–***)

Notes: The table summarizes LGD related literature with focus on systematic effects (i.e., impact of macro variables).

Table 3.D.2: Composition of European sample

Country	percentage
Germany	27.48%
Great Britain	23.39%
Portugal	12.83%
Ireland	8.30%
Denmark	7.66%
Norway	5.63%
Sweden	4.79%
Finland	3.52%
Latvia	2.71%
Estonia	2.68%
France	0.83%
Poland	0.19%

Notes: The table summarizes the composition of the European sample. The second column displays the percentage of the corresponding country in the European sample.

Table 3.D.3: Descriptive statistics

		US	**GB**	**Europe**
		Dependent		
LGD	mean	0.2057	0.2949	0.1958
	median	0.0000	0.0201	0.0080
	standard deviation	0.3438	0.4121	0.3466
		Loan specific | metric		
EAD	Mean	1,785,234.75	648,425.48	708,777.43
	Median	506,889.22	94,707.10	115,412.68
	Standard deviation	3,893,459.81	3,508,848.60	5,292,547.89
		Loan specific | categoric		
Facility	term loan	43.63%	46.34%	60.22%
	line	56.37%	53.66%	39.78%
Protection	no	16.13%	28.88%	30.45%
	yes	83.87%	71.12%	69.55%
Industry	non FIRE	85.88%	91.26%	84.62%
	FIRE	14.12%	8.74%	15.38%
		Macro variables		
Δ GDP	mean	1.11%	0.70%	1.24%
	median	1.69%	1.31%	1.92%
	standard deviation	2.07%	2.67%	2.75%
Δ EI	mean	1.91%	1.49%	2.35%
	median	7.60%	6.74%	11.86%
	standard deviation	20.54%	17.11%	26.22%
VIX	mean	22.7187	26.4292	26.4292
	median	20.4113	23.8916	23.8916
	standard deviation	10.1888	9.5692	9.5692
Δ HPI	mean	-5.70%	-1.56%	-1.88%
	median	-4.74%	-1.65%	0.42%
	standard deviation	7.68%	7.52%	5.75%
NPL ratio	mean	3.20%	2.63%	3.04%
	median	3.85%	3.51%	3.27%
	standard deviation	1.79%	1.36%	0.76%

Notes: The table summarizes descriptive statistics of dependent and independent variables. For metric variables the mean, median, and standard deviation is presented. Proportions are given for variables of categoric nature. FIRE is an abbreviation for corporations which are active in the finance, insurance or reals estate industry.

Table 3.D.4: Pairwise correlations of macro variables and random effect

	US	GB	Europe
$-\Delta$[GDP (standardized)]	-11.64%	12.39%	34.44%
$-\Delta$[EI (standardized)]	-0.50%	20.91%	38.41%
VIX	-11.17%	-1.04%	33.84%
$-\Delta$[HPI (standardized)]	-17.46%	-2.03%	54.41%
NPL ratio	12.26%	-58.74%	-32.32%

Notes: The table summarizes the pairwise correlations of the macro variables (i.e., GDP, EI, VIX, HPI, and NPL ratio) with the random effect posterior means.

Table 3.D.5: Results of combined models

	posterior mean	HPDI (90%)		posterior odds	naive standard error	time-series standard error
			US			
$\beta_{\text{NPL ratio}}$	0.0480	-0.0736	0.1701	3.0306	0.0007	0.0011
α	2.1408	1.9280	2.3576	∞	0.0013	0.0024
σ^F	0.3559	0.2344	0.4642	∞	0.0007	0.0007
			GB			
β_{EI}	-0.0641	-0.2917	0.1627	0.4637	0.0014	0.0044
a	0.6977	0.0973	1.2440	768.2308	0.0037	0.0073
φ	0.7439	0.5412	0.9515	∞	0.0013	0.0026
σ_c^F	0.4120	0.2901	0.5282	∞	0.0008	0.0009
			Europe			
β_{HPI}	-0.1909	-0.3989	0.0232	0.0769	0.0013	0.0082
a	0.3435	0.0166	0.6503	343.8276	0.0022	0.0045
φ	0.8347	0.6924	0.9851	∞	0.0010	0.0021
σ_c^F	0.2644	0.1946	0.3319	∞	0.0004	0.0007

Notes: The table summarizes the results of the combined models where a macro variable and the random effect are included. The presentation is reduced to the results regarding the macro variables itself and the parameters of the random effect. The first column presents the posterior means of the parameters. The second and third column contain the lower and upper bound of the HPDI to a credibility level of 90%. The fourth column includes posterior odds, while, in the last two columns, the naive and time-series standard error of the chains are presented, whereas, the time-series standard error is calculated based on the effective (N^*_{MCMC}) instead of the real (N_{MCMC}) sample size. Hereby, $N^*_{\text{MCMC}} < N_{\text{MCMC}}$ holds for autocorrelated chains.

Chapter 4

Time matters: How default resolution times impact final loss rates

This chapter is joint work with Ralf Kellner* and Daniel Rösch† and corresponds to a working paper with the same name.

Abstract

The Loss Given Default (LGD) and the Default Resolution Time (DRT) are two important credit risk parameters which are treated on an isolated basis in academia so far. However, dependency between the two seems to be obvious as complex default resolutions may be accompanied with long DRTs and high LGDs. We propose a hierarchical Bayesian model with multiple random effects for joint estimation of DRTs and LGDs. Our approach explicitly takes unresolved loans into account and eliminates a potential resolution bias which arises when LGD distributions are estimated using resolved cases only, as is standard in common approaches. Using a European data set, we find strong positive dependencies between DRTs and LGDs. We show that neglecting unresolved cases leads to systematic underestimation of average LGDs on an out of sample perspective. Thus, our approach leads to a better out of sample fit than a pure LGD model.

Keywords: credit risk; default resolution time; loss given default; random effects; resolution bias

JEL classification: C23, G21, G33

* University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: ralf.kellner@ur.de.

† University Regensburg, Chair of Statistics and Risk Management, 93040 Regensburg, Germany, email: daniel.roesch@ur.de.

4.1 Introduction

The Default Resolution Time (DRT) and the Loss Given Default (LGD) of defaulted loan contracts are outcomes of the same random process – the resolution process. Thus, an interconnection of DRTs and LGDs is plausible as complex default resolutions might be accompanied with longer resolution processes and higher losses. To the best of our knowledge, no study exists so far which deeply examines the relation of these credit risk parameters. However, its consideration may be beneficial for generating accurate predictions. Right after default, DRTs and LGDs are unknown and need to be estimated. As resolution proceeds, loans might still be stuck in the resolution process. However, additional information in terms of the time a loan has already spent in the resolution process is available. This information is not included in traditional LGD models, but might improve LGD predictions – particularly, in the presence of an interconnection of DRTs and LGDs. Moreover, its consideration controls for the so called resolution bias. In the more recent time periods, only LGDs of loans with short DRTs are observable. Assuming a positive dependence of DRTs and LGDs which is indicated by economic intuition and supported by descriptive analyses (see Betz et al., 2016, 2017), these loans tend to exhibit lower LGDs. This leads to downward biased LGD estimates.

As a numerical illustration, assume three loan contracts which defaulted at the same time. The first loan is almost completely recovered in half a year after default ($\text{LGD}_1 = 7\%$ and $\text{DRT}_1 = 0.5$). The second loan generates a medium loss in the first two years after default ($\text{LGD}_2 = 21\%$ and $\text{DRT}_2 = 2.0$). Finally, the resolution of the third loan is terminated after four years causing a high loss ($\text{LGD}_3 = 42\%$ and $\text{DRT}_3 = 4.0$).[1] Considerable distortions arise if expected LGDs are estimated at varying points in time. Thus, the estimated expected LGD half a year after default amounts to 7% ($\widehat{\text{E}}[LGD_i \mid t = 0.5] = 7\%$) as only the LGD of the first loan is observable. The estimated expected LGD yields to 14% ($\widehat{\text{E}}[LGD_i \mid t = 2.0] = 14\%$) two years after default. An unbiased estimate of the expected LGD can foremost be made four years after default ($\widehat{\text{E}}[LGD_i \mid t = 4.0] = 23\%$) when all loans are resolved. The expected LGD is underestimated by 16 percentage points half a year after default and still by 9 percentage points two years after default. The inclusion of censored observations, i.e., unresolved cases, might improve the estimation. After the resolution of the second loan, it is known that $\text{DRT}_3 > 2.0$ and, thus, $\text{LGD}_3 > 21\%$ might be expected. This leads to the conclusion that $\widehat{\text{E}}[\text{LGD}_i \mid t = 2.0] > 16\% \left(= \frac{7\%+21\%+21\%}{3}\right)$

[1] The figures reflect the descriptive statistics in our data set (see Figure 4.1). Loans with DRTs $\in [0, 0.5]$ generate average LGDs of 7%, wheres, DRTs $\in (1.0, 2.0]$ (DRTs $\in (3.0, \dots])$ correspond to average LGDs of 21% (42%). LGDs are calculated based on discounted recovery payments (see Section 4.2.1).

which might be used as lower limit for LGD prediction. Considering the dependence structure of DRTs and LGDs in more detail, the estimation might be further adjusted.

In recent years, the literature regarding LGD modeling has widened considerably. Comparative studies can be found in, e.g., Qi and Zhao (2011) and Loterman et al. (2012). However, literature considering *workout* LGDs is still limited. Most of the publications refer to *market-based* LGDs, whereby, the corresponding Recovery Rate (RR) is defined as ratio of the market price 90 days after default to the outstanding amount. Hence, market-based LGDs are only observable for traded securities such as bonds. Workout LGDs are based on actual recovery payments collected during the resolution process and, thus, usually applied for loans. The distribution of workout LGDs is more extreme compared to market-based LGDs and, typically, high probability masses at no loss (LGD = 0) and total loss (LGD = 1) arise (see, e.g., Krüger and Rösch, 2017; Betz et al., 2018). Thus, the consideration of the distributional form is essential for workout LGDs. Altman and Kalotay (2014) develop a Bayesian Finite Mixture Model (FMM) with a probabilistic substructure in terms of an ordered logit (OL) model to estimate the probability of the mixture components depending on explanatory variables. A frequentistic version of this model is presented by Kalotay and Altman (2017). The model of Altman and Kalotay (2014) and Kalotay and Altman (2017) is applied by Bijak and Thomas (2015) and enhanced by Betz et al. (2018). Calabrese (2014) estimates a mixture of Beta distributions, whereas, Krüger and Rösch (2017) apply quantile regression on the LGD distribution. The literature regarding DRTs is more sparse and mainly refers to the duration of Chapter 7 and Chapter 11 resolutions (see, e.g., Helwege, 1999; Partington et al., 2001; Bris et al., 2006; Denis and Rodgers, 2007; Wong et al., 2007). Betz et al. (2016) and Betz et al. (2017) analyze DRTs of defaulted loan contracts and descriptively find impacts of DRTs on LGDs. The interconnection of DRTs and LGDs is also indicated in the LGD literature. Dermine and Neto de Carvalho (2006) apply mortality analysis on a data set of defaulted bank loans, whereas, Gürtler and Hibbeln (2013) focus on the resolution bias. They suggest to restrict the data set to avoid biased estimates. However, LGD data is sparse so constraints might be unfavorable. The inclusion of the DRT into the LGD modeling framework might diminish the effect of the resolution bias without restrictions in the data set. Common ways to implement dependence structures of credit risk parameters are random effects. By this means, joint time patterns of these parameters are considered in the modeling context. Rösch and Scheule (2010), Bade et al. (2011), and Rösch and Scheule (2014) apply random effects to model the dependence of probabilities of default (PDs) and LGDs. Furthermore, Lee and Poon (2014) state that frailties, i.e., random effects in survival models, make more sigificant risk contributions than macroeconomic factors in a credit risk context.

Using a unique European data set provided by Global Credit Data (GDC), we develop a hierarchical Bayesian modeling approach for joint estimation of DRTs and LGDs combining a Finite Mixture Model (FMM) with a probabilistic substructure for the LGD and an Accelerated Failure Time (AFT) model for the DRT. Thereby, we allow for direct and indirect dependency structures. The first is attained by the inclusion of the DRT in the LGD model. This allows LGD predictions for censored observations, i.e., non-performing loans, within the modeling framework as final DRTs are estimated based on censored observations in the DRT model. The second refers to common time patterns of DRTs and LGDs. We implement two correlated random effects in the DRT model and the LGD model.

We contribute to the literature in three ways. First, we deeply examine the dependence structure of DRTs and LGDs allowing for a direct and an indirect channel and find impacts of DRTs on LGD distributions which are even more pronounced in boom and crisis periods. Thus, DRTs are longer (shorter) in crisis (boom) periods. In crisis periods, this burdens financial market liquidity as systematically more loans are stuck in the resolution process. On top of that, losses are systematically higher during such periods due to the stronger positive dependence. Second, we analyze and quantify the impacts of the resolution bias. We compare a pure (traditional) LGD model with the proposed hierarchical approach and find parameter distortions in the pure LGD model. These result in biased predictive LGD distributions on an out of sample and out of sample out of time perspective. Impacts of the resolution bias are diminished in the hierarchical approach leading to appropriate predictive LGD distributions. Neglecting unresolved cases leads to an underestimation of average LGDs on a out of sample perspective. Third, we are able to generate intuitive LGD predictions for non-performing loans. As final DRTs are estimated within the AFT model, these can be applied to generate final predictive LGD distributions for censored observations. These considerably outperform predictions based on a pure LGD model.

The remainder of this paper is structured as follows. Section 4.2 describes the data and introduces the hierarchical modeling framework. Results are presented in Section 4.3. In Section 4.4, the model is validated on an in sample and out of sample perspective. Section 4.5 concludes.

4.2 Data and methods

4.2.1 Data

We use access to the unique loss data base of Global Credit Data (GCD). The data base includes detailed loss information on transaction basis of 53 member banks all around the world. In the data base, the LGD is determined by:

$$\mathrm{LGD}_i = 1 - \mathrm{RR}_i, \tag{4.1}$$

whereby, LGD_i is the loss rate of loan i and RR_i is the corresponding RR. The RR is calculated as the sum over the present values of all relevant transactions divided by the outstanding amount.[2]

We follow Höcht and Zagst (2010) and Höcht et al. (2011) and develop two selection criteria to eliminate loans with extraordinary payment structures. Both criteria relate all relevant transactions including charge-offs (which are not included in the LGD calculation) to the outstanding amount. The first criterion, to which we refer as *pre-resolution criterion*, relates transactions arising pre resolution to the outstanding amount at default. We set the barriers of the pre-resolution criterion to $[90\%, 110\%]$ for resolved and $[-50\%, 400\%]$ for unresolved loans. In the second criterion, i.e., *post-resolution criterion*, transactions occurring post resolution are related to a fictional outstanding amount at resolution. The barriers are set to $[-10\%, 110\%]$ for the post-resolution criterion. The post-resolution criterion applies for resolved loans only. Subsequently, loans with abnormal low and high LGDs ($< -25\%$ and $> 125\%$) are eliminated.[3] We consider a subsample of defaulted European term loans and lines to small and medium sized enterprises (SMEs).[4] We further exclude loans which defaulted before 2004 and after 2016 (10.02% of subsample data). A subsample of 38,165 loans remains.

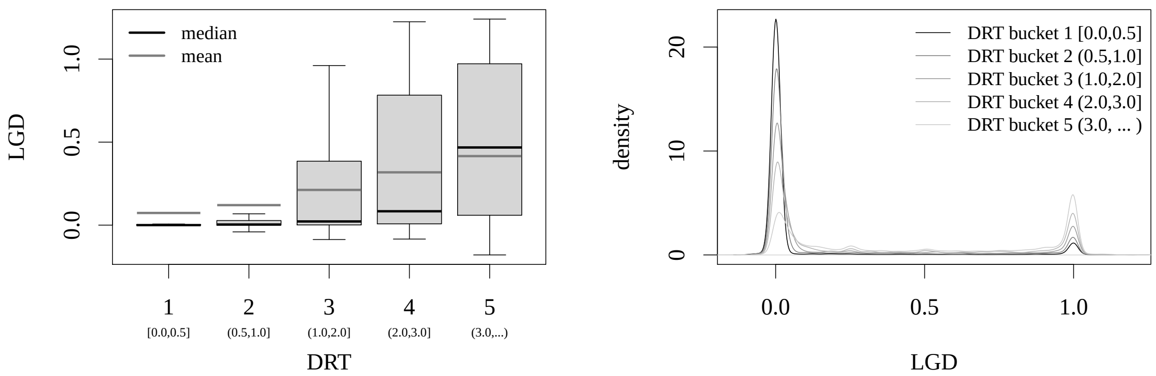
Figure 4.1 illustrates the interconnection of the two dependent variables, i.e., the DRT and the LGD. Therefore, the data set is divided into DRT buckets based on DRTs. The first bucket includes all loans with DRTs $\in [0, 0.5]$ years. The second bucket contrains all loans with DRTs $\in (0.5, 1.0]$ years, and so on (see x-axis of left panel and legend of right panel). In the left

[2] See Betz et al. (2018). More detailed information of the LGD calculation can be found in Betz et al. (2016).

[3] Overall, 0.50% of resolved loans are eliminated due to the pre-resolution criterion and 0.19% due to the post-resolution criterion, whereas, 0.23% of unresolved loans are eliminated based on the pre-resolution criterion. Subsequently, 0.13% are sorted out due to abnormal low and high LGD values.

[4] We restrict the data base to reduce noise and generate a rather homogeneous sample. We consider the twelve most common European countries in the data base, i.e., Great Britain, Germany, Denmark, Portugal, Ireland, France, Finland, Sweden, Norway, Latvia, Estonia, and Poland.

Figure 4.1: Relation of DRT and LGD



Notes: The figure illustrates the relation of DRTs and LGDs. The data is divided into DRT buckets based on the realized DRTs. Thus, the first bucket includes all loans with DRTs $\in [0, 0.5]$ years. The second bucket contains all loans with DRTs $\in (0.5, 1.0]$ years, and so on (see x-axis of left panel and legend of right panel). In the left panel, box plots of LGDs for the DRT buckets are displayed. Outliers are hidden. The thick black lines mark the medians, whereas, the thick gray lines are the means. In the right panel, kernel density estimates of LGDs for the DRT buckets are illustrated. The band width is fixed to 0.015 to ensure comparability.

panel, box plots of LGDs divided by DRT buckets are displayed. The thick black lines mark the medians, whereas, the thick gray lines are the means. Considering the latter, average LGDs seem to linearly increase in the DRT buckets. To examine the origin of this increase, the right panel displays kernel density estimates for the DRT buckets. The LGD distribution of higher DRT buckets is shifted towards higher LGD values, i.e., probability masses of lower LGD values decrease and probability masses of higher LGD values increase. Thus, average values increase.

Table 4.1 summarizes the descriptive statistics of the dependent and independent variables. Figures are stated for all loans (resolved and unresolved cases) and for resolved and unresolved loans separately. The upper panel of the table includes descriptive statistics for the LGD and the DRT. For unresolved cases, incurred LGDs are applied. Incurrent LGDs are computed as the sum over the present values of all relevant transactions, which occurred up to the end of the observation period (end of 2016), divided by the outstanding amount. As the resolution process is not terminated, incurrent LGDs are higher than final LGDs. DRTs for unresolved cases are censored to the end of the observation period (end of 2016), e.g., for unresolved loans defaulted at the end of 2015, a censored DRT of one year is assigned. Censored DRTs are lower than final DRTs as the resolution process is not terminated. In the table, average values of LGDs and DRTs for unresolved cases are higher compared to resolved cases as unresolved cases are shaped by rather *bad* loans, i.e., loans exhibiting high DRTs and high LGDs. In the middle panels of the table, descriptive statistics of loan specific independent variables are stated. We use the EAD to control for the size of the loan. It is further distinguished between term loans and lines, if a

Table 4.1: Descriptive statistics

		all	**resolved**	**unresolved**
n		38,165	35,272	2,893
	dependent variables			
LGD	mean	0.2534	0.2099	0.7839
	median	0.0133	0.0082	0.9780
	standard deviation	0.3810	0.3531	0.3017
DRT	mean	1.9882	1.7342	5.0839
	median	1.2621	1.1335	4.9090
	standard deviation	2.0756	1.7509	3.0147
	loan specific (metric)			
EAD	mean	533,118.89	516,582.05	734,739.20
	median	102,987.29	100,237.48	155,097.53
	standard deviation	3,624,711.52	3,610,978.60	3,782,983.43
	loan specific (categoric)			
Facility	*term loan*	62.00%	60.39%	81.68%
	line	38.00%	39.61%	18.32%
Protection	*no*	25.61%	26.03%	20.46%
	yes	74.39%	73.97%	79.54%
Industry	*non FIRE*	83.13%	82.13%	95.30%
	FIRE	16.87%	17.87%	4.70%
	macro variables			
Δ HPI	mean	-1.6966	-1.7765	-0.7229
	median	0.1662	0.1662	0.8360
	standard deviation	6.0221	5.9928	6.2878
VIX	mean	24.4762	24.3681	25.7947
	median	22.9249	22.6771	23.3451
	standard deviation	9.4105	9.4699	8.5465

Notes: The table summarizes descriptive statistics for dependent and independent variables in the data set. For metric variables, means, medians, and standard deviations are stated. Proportions are presented for variables of categoric nature. The sample size is denoted by *n*. The abbreviation FIRE means *Finance, Insurance, Real Estate* and denotes corporations of this industries. The macro variable Δ HPI is the is the yoy percentage chance of the *House Price Index*, whereas, the VIX is the *Volatility Index*.

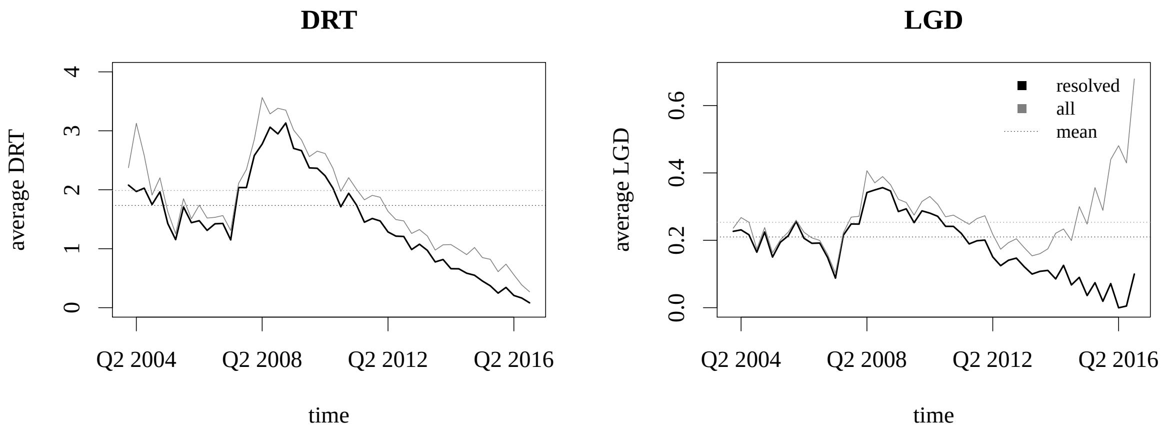
loan is protected by collateral or guarantee or not, and if the debtor has Finance, Insurance, Real Estate (FIRE) industry affiliation. Reference categories in the subsequent models are written italic in the table.[5] The lower panel of the table contains descriptive statistics of the applied macro variables. The year-on-year (yoy) percentage change of weighted average real residential prices (Δ HPI) is employed as explanatory variable for the LGD, whereas, we use the VSTOXX Volatility Index (VIX) for the DRT.[6]

Figure 4.2 illustrates the time patterns of average DRTs in the left panel and average LGDs in the right panel for resolved loans (thick black line) and all loans (resolved and unresolved loans, thin gray line). Regarding the latter, values for unresolved loans, i.e., censored observations,

[5] The reference category is *term loan* for Facility, *no* for protection, and *non-FIRE* for industry.
[6] We tested further macro variables, e.g, the yoy percentage change of weighted average seasonally adjusted GDPs and the quarterly average yoy percentage change of weighted average equity indices. However, Δ HPI and VIX exhibit the highest statistical evidence. Remaining results are available form the authors upon request.

have to be calculated. Thus, DRTs are censored to the end of the observation period (end of 2016) and incurrent LGDs are considered for unresolved cases. Incurrent LGDs are computed as sum over the present values of all relevant transactions, which occurred up to the end of the observation period (end of 2016), divided by the outstanding amount. The relation of DRTs

Figure 4.2: Time patterns of average DRTs and average LGDs



Notes: The figure illustrates time patterns of average DRTs in the left panel and average LGDs in the right panel. The black lines display the average values for resolved loans, whereas, the gray lines are average values for all loans, i.e., resolved and unresolved cases. Thus, the latter include censored values. Means over the entire time period are illustrated by dotted lines.

and LGDs (see Figure 4.1) might partly be driven by analogous time patterns. Both dependent variables sharply increase prior to the Global Financial Crisis (GFC, 2007 Q2) and reach the maximum during the climax of the GFC. The rebound in the aftermath of the crisis seems gradual. There are only minor deviations between resolved loans and all loans considering the average DRTs. The graph for all loans is slightly shifted upwards by the censored observations. Regarding average LGDs, this spread is severe particularly in the most recent time periods. This is mainly due to incurrent LGDs, i.e., LGDs based on transactions which occur up to the end of the observation period, in the averaging. Final LGDs will be lower. However, final LGDs of all loans will still lie above the black line (final LGDs of resolved loans). This mismatch is due to the resolution bias. Due to the censoring, final LGDs are only observable for defaults with short DRTs in the more recent time periods. Due to the interconnection of DRTs and LGDs, these tend to be lower implying an underestimation of LGDs in the more recent time periods.

In this paper, we aim to analyze the effects of the resolution bias on an in sample and out of sample perspective. Therefore, we divide the data set as of Table 4.1 into subsamples. The first subsample serves as *estimation sample*. It includes all loans defaulted between 2004 Q1 and 2010 Q4. Thus, it comprises times of rather sound economic surrounding, the GFC, and parts of the rebound phase. As we aim to analyze effects of the resolution bias, we treat loans which are not resolved until 2010 Q4 as censored observations, i.e., unresolved loans. The second subsample, to which we refer to as *validation sample I*, includes the final observations to the censored observations as of the estimation sample. We apply validation sample I to perform an *out of sample validation* of LGDs. The third sample, i.e., *validation sample II*, includes all loans defaulted between 2011 Q1 and 2016 Q4. It is used to perform an *out of sample out of time validation* of LGDs. Table 4.2 summarizes the estimation sample and the validation samples. In

Table 4.2: Estimation sample and validation samples

		estimation sample	**validation sample I** (*out of sample*)	**validation sample II** (*out of sample out of time*)
n	all	31,988	10,171	6,177
	resolved	21,817	8,447	5,008
	unresolved	10,171	1,724	1,169
		dependent variables		
average LGD	all	0.2586	0.4270	0.2267
	resolved	0.1801	0.3511	0.1017
	unresolved	0.4270	0.7987	0.7622
average DRT	all	1.5763	4.2566	1.0495
	resolved	1.1964	3.6851	0.7869
	unresolved	2.3911	7.0568	2.1743

Notes: The table summarizes the applied samples. The number *n*, the average LGD, and the average DRT of all loans, resolved loans, and unresolved loans are presented for the estimation sample and the two validation samples. The models are estimated based on the estimation sample. This sample includes all loans defaulted between 2004 Q1 and 2010 Q4. Loans which are not resolved until 2010 Q4 are treated as censored observations, i.e., unresolved cases, in the estimation. Validation sample I contains the final observations of these unresolved cases. However, observations exist which are still censored at the end of 2016 (unresolved cases in validation sample I). In validation sample II, loans which defaulted between 2011 Q1 and 2016 Q4 are included. Thus, validation sample I is applied for the *out of sample* validation, whereas, the *out of sample out of time* validation is performed on validation sample II.

the upper panel, the sample sizes are stated. Validation sample II consists of the 10,171 loans which are treated as unresolved cases in the estimation sample. Some of these loans (1,724) are still unresolved at the end of 2016. However, the proportion of unresolved loans is lower in validation sample I compared to the estimation sample. In the lower panel, average values of LGDs and DRTs are stated. These are rather similar comparing the estimation sample and validation sample II, however, considerably higher in validation sample I. This is due to the fact that validation sample I contains final observations to censored cases in the estimation sample, thus, observations with higher DRTs and higher LGDs.

4.2.2 Methods

This paper aims to compare a pure LGD model with a hierarchical approach combining a model for the DRT with a LGD model. For the LGD model, we adapt the model presented in Betz et al. (2018). In the hierarchical approach, this LGD model is combined with an *Accelerated*

Failure Time (AFT) model for the DRT. Survival time models such as the AFT model bear the advantage of considering censored observations, i.e., unresolved loans. This can be applied to include censored observations in an LGD modeling context. By this means, we are able to diminish effects of the resolution bias. To investigate the direct dependence of DRTs on LGDs, the DRT serves as an explanatory variable in the LGD model. Two correlated random effects are included to study comovements of DRTs and LGDs in the time line (indirect dependence). In the following, we briefly review the LGD model of Betz et al. (2018) and discuss the extensions in the context of the hierarchical model.

LGD model

A Normal Finite Mixture Model (FMM) combined with a probabilistic substructure in terms of an Ordered Logit (OL) model is applied to the loss rate L.[7] In FMMs, the dependent variable is assumed to be divided into a finite number of K latent classes. In each class k, L follows a normal distribution with parameters θ_k depending on the latent class k. Thus, the probability density function (PDF) of an FMM $g(L|\theta_1,\ldots,\theta_K)$ is the p_k weighted sum of the component PDFs $f_k(L|\theta_k)$:

$$g(L\,|\,\theta_1,\ldots,\theta_K) = \sum_{k=1}^{K} p_k\; f_k(L\,|\,\theta_k)\,. \tag{4.2}$$

To ensure the general properties of a PDF, i.e., $g(l) \geq 0$ for all $l \in \mathbb{R}$ and $\int_{-\infty}^{\infty} g(l) = 1$, $p_k \geq 0$ and $\sum_k p_k = 1$ must hold. Assuming conditional independence, the likelihood of a Normal FMM $\phi(L_1,...,L_N\,|\,\mu,\sigma,p)$ is the product of the individual likelihood contributions:

$$\phi(L_1,...,L_N\,|\,\mu,\sigma,p) = \frac{1}{(2\pi)^{\frac{N}{2}}} \prod_{i=1}^{N} \left(\sum_{k=1}^{K} \frac{p_k}{\sigma_k} \exp\left[-\frac{(L_i-\mu_k)^2}{2\sigma_k^2} \right] \right), \tag{4.3}$$

where, μ_k and σ_k are the parameters of the latent class k and N is the number of observations. To adapt data augmentation, the component weight p_k is replaced by an indicator variable d_{ik} which takes the value one if l_i is a random draw of component k and zero otherwise:

$$\phi(L_1,...,L_N\,|\,\mu,\sigma,d) = \frac{1}{(2\pi)^{\frac{N}{2}}} \prod_{i=1}^{N} \left(\sum_{k=1}^{K} \frac{d_{ik}}{\sigma_k} \exp\left[-\frac{(L_i-\mu_k)^2}{2\sigma_k^2} \right] \right). \tag{4.4}$$

A probabilistic substructure is formulated to include covariates in the FMM. To rely on the classical formulation of the OL model, we define the component affiliation y_i:

$$y_i = k \quad \text{if } d_{ik} = 1\,, \tag{4.5}$$

[7] We use the notation L for the LGD due to aesthetic reasons.

where, d_{ik} is the indicator as of Equation (4.4). The component affiliation Y_i is categorically distributed and determined by the location of a metric latent variable Y_i^* to corresponding cut points c_k ($k \in \{1, \ldots, K-1\}$):

$$Y_i = \begin{cases} 1 & \text{if } Y_i^* \leq c_1 \\ 2 & \text{if } c_1 < Y_i^* \leq c_2 \\ \vdots & \\ K & \text{if } c_{K-1} < Y_i^* \,. \end{cases} \tag{4.6}$$

The latent variable Y_i^* follows a linear model:

$$Y_i^* = z_i \zeta + F_{t(i)} + e_i, \qquad e_i \sim \text{logistic}\,, \tag{4.7}$$

where, z_i is a $(1 \times J)$ vector of independent variables and ζ is the $(J \times 1)$ vector of coefficients. The term e_i describes the errors. A random effect $F_{t(i)}$ is introduced into the modeling framework to control for comovement in the time line. It originates from a Normal distribution with mean zero and standard deviation σ:

$$F_t \sim \mathrm{N}(0, \sigma)\,. \tag{4.8}$$

The time stamp $t(i)$ in Equation (4.7) indicates the default time t in quarters of loan i. Two loans i and i' which defaulted in the same quarter ($t(i) = t(i') = t$) share the same realization of the random effect ($f_{t(i)} = f_{t(i')} = f_t$). For $f_t > 0$ ($f_t < 0$), both loans exhibit higher (lower) values of y_i^* and, thus, higher (lower) probabilities of high component affiliations y_i. Higher component affiliations y_i are accompanied with higher loss rates and vice versa. Thus, the random effect displays the comovement in time line, i.e., higher or lower average loss rates in specific default quarters which can not be explained by observable variables included in z_i.

Hierarchical model

In the following, the hierarchical model for the joint estimation of DRTs and LGDs is discussed. We apply an AFT model for the DRT. Thus, the logarithm of the resolution time $\ln(T_i)$ can be expressed by a linear model:

$$\ln(T_i) = \beta_0 + x_i \beta + F^T_{t(i)} + s\,\varepsilon_i\,, \qquad \varepsilon_i \sim \text{negative Gumbel}\,, \tag{4.9}$$

where, x_i is a $(1 \times J_T)$ vector of independent variables, β is the $(J_T \times 1)$ vector of coefficients, and β_0 is the intercept. We assume the errors ε_i to follow a negative Gumbel distribution and, thus, the DRT to be Weibull distributed.[8] The term s is the scale parameter. A random effect $F^T_{t(i)}$ is introduced into the modeling framework to control for comovement in the time line. Equation (4.9) applies to non censored, i.e., final, observations. For censored observations, final observations are estimated within the Bayesian modeling framework. By this means, we are able to predict final DRTs for censored data points, i.e., unresolved loans.

In the hierarchical approach, the AFT model for the DRT is simultaneously estimated with the LGD model. As LGD model the Normal FMM with the probabilistic substructure is adapted, whereby, the logarithm of the DRT is included as explanatory variable to account for the direct dependence of DRTs on LGDs. Equation (4.7) modifies to:

$$\mathcal{Y}_i^* = z_i \gamma + \ln(T_i)\gamma_T + F^L_{t(i)} + \epsilon_i\,, \qquad \epsilon_i \sim \text{logistic}\,, \tag{4.10}$$

where, z_i is the $(1 \times J_L)$ vector of independent variables, γ is the $(J_L \times 1)$ vector of coefficients, and γ_T is the coefficient of the logarithm of the DRT. Again, a random effect $F^L_{t(i)}$ is introduced into the modeling framework to control for comovement in the time line.[9]

The random effects $F^T_{t(i)}$ as of Equation (4.9) and $F^L_{t(i)}$ as of Equation (4.10) originate from a bivariate normal distribution:

$$\begin{pmatrix} F_t^T \\ F_t^L \end{pmatrix} \sim \mathrm{N}_2\left(0_2\,,\Sigma\right)\,, \tag{4.11}$$

where, 0_2 is the two dimensional zero vector ($0_k = (\,0 \quad 0\,)^T$) and Σ is the (2×2) covariance matrix. Prior distributions are provided for the individual standard deviations (σ_T and σ_L) and the (2×2) correlation matrix Ω (see Appendix 4.A):

$$\begin{aligned} \Sigma &= \operatorname{diag}(\sigma_T, \sigma_L)\ \Omega\ \operatorname{diag}(\sigma_T, \sigma_L) \\ &= \begin{pmatrix} \sigma_T^2 & \sigma_T \sigma_L \omega_{L,T} \\ \sigma_T \sigma_L \omega_{T,L} & \sigma_L^2 \end{pmatrix}, \end{aligned} \tag{4.12}$$

whereby, $\omega_{T,L} (= \omega_{L,T})$ is the correlation of F_t^T and F_t^L. By the inclusion of the random effects, we control for joint comovements of loss rates and resolution times in the time line. Two loans i and i' defaulted in the same quarter ($t(i) = t(i') = t$) share the same realizations of the random effects ($f^T_{t(i)} = f^T_{t(i')} = f_t^T$ and $f^L_{t(i)} = f^L_{t(i')} = f_t^L$, however, $f_t^T \neq f_t^L$ in most of the cases). For $f_t^T > 0$

[8] We tested various distributional assumptions regarding the resolution time, e.g., log Normal, log Logistic, Exponential, and Weibull. The Weibull distribution seems to have the best fit.

[9] Equation (4.2), (4.3), (4.4), (4.5), and (4.6) apply in analogy to Y_i*.

($f_t^T < 0$), average DRTs are higher (lower). Assuming a positive correlation between the random effects and a positive parameter estimate of the logarithm of the DRT in the LGD model ($\gamma_T > 0$), the corresponding LGDs are effected through two channels:[10] Directly, as higher (lower) DRTs are inserted in the LGD model. Indirectly, as positive (negative) realizations of f_t^T tend to imply positive (negative) realizations of f_t^L due to the positive correlation. Thus, LGDs are even higher. However, negative realizations of f_t^L remain possible for a stochastic process as of Equation (4.11) which might reduce LGDs. Both scenarios are conceivable. Confronted with tense economic surrounding, financial institutions might decide to follow a wait-and-see strategy and relocate resolution efforts in the future. This might provide benefits and reduce the LGD ($f_t^L < 0$). However, LGDs might be further increased ($f_t^L > 0$) if financial institutions are forced to resolve defaulted loans at a certain point in time, e.g., if there is no further option to wait.

Estimation

The parameters of the LGD model and hierarchical model are estimated via Bayesian inference. We use a Markov Chain Monte Carlo (MCMC) sampler to derive the posterior distributions of the parameters. The MCMC sampler generates samples by constructing reversible Markov chains. The equilibrium distribution corresponds to the target posterior distribution. The solution via MCMC sampling is necessary due to the model complexity of the LGD model and the hierarchical model, i.e., priors are partly non conjugate. Direct sampling from the posterior distributions is not possible as there is no analytical solution. The LGD model and the hierarchical model are sampled with two MCMC chains. Burn-in is set to 500. Posterior samples contain 25,000 iterations per chain with a thinning of 5. Metric dependent variables are standardized to ease convergence.[11]

Most of the model parameters are provided with weakly informative prior distributions. See Appendix 4.A for detailed information on the Bayesian model specifications. Common convergence diagnostics can be found in Appendix 4.B.

[10] These assumptions correspond to the empirical results (see Section 4.3).

[11] As MCMC sampler, we adapt Stan which is a Hamiltonian Monte Carlo (HMC) sampler. It overcomes some of the problems inherent in Gibbs sampling, e.g., regarding highly correlated posterior distributions.

4.3 Empirical results

In Bayesian inference, posterior distributions of parameters are assumed to be continuous. Thus, a single value of the posterior distribution has a probability of zero. On the contrary, one *true* parameter estimate is assigned in frequentistic terms. A null hypothesis is set up to reach a yes-or-no decision. Under the Bayesian approach, estimates are provided by posterior distributions which offer an intuitive consideration of parameter uncertainty. Thus, other concepts are indicated to quantify if results are in favor of an impact, i.e., if an impact is statistical evident. We apply two of them – *credible intervals* and *Bayes factors*.

Credible intervals, e.g., Highest Probability Density Intervals (HPDIs), specify intervals in the domain of the posterior distribution in which the unobservable parameter lies with a certain probability. The HPDI denotes the narrowest credible interval. If $0 \notin$ HPDI, the domain of the posterior distribution is located in the positive (negative) value range indicating statistical evidence of the positive (negative) sign. Besides credible intervals, we apply Bayes factors to evaluate statistical evidence. Bayes factors are the relation of posterior odds to prior odds. Posterior odds are defined as the ratio of posterior probability masses under the null hypotheses and the alternative hypothesis. As we are interested in the evidence of the signs, posterior odds are derived as the ratio of posterior mass favoring the sign of the posterior mean to posterior mass of the opposite sign:

$$\text{posterior odds}_{\mathrm{E}[\theta]<0} = \frac{\mathbb{P}(\theta < 0 \,|\, \text{data})}{\mathbb{P}(\theta \geq 0 \,|\, \text{data})}$$
$$\text{posterior odds}_{\mathrm{E}[\theta]>0} = \frac{\mathbb{P}(\theta > 0 \,|\, \text{data})}{\mathbb{P}(\theta \leq 0 \,|\, \text{data})},$$

whereby, θ denotes an arbitrary parameter. Assuming a positive posterior mean ($\mathrm{E}[\theta] > 0$), posterior odds$_{\mathrm{E}[\theta]>0} = 3$ indicates that a positive impact is three times as likely as a negative impact. Prior odds are the corresponding ratio of the prior distribution. Assuming a symmetric prior distribution around zero, posterior odds are equivalent to the Bayes factor. A Bayes factor exceeding 3.2 is deemed as substantial evidence. Values above 10 are assigned with strong evidence, whereas, values above 100 are related to decisive evidence (see Kass and Raftery, 1995).

LGD model

The LGD model is estimated based on the estimation sample (see Table 4.2). However, it offers no possibility to include censored observations, i.e., unresolved loans, in the estimation process. Thus, the 21,817 resolved cases are included, whereas, 10,171 unresolved defaults are neglected. As these unresolved loans tend to exhibit higher LGDs due to the resolution bias, parameter estimates might be distorted.

Table 4.3 summarizes the results of the LGD model. Parameters are stated in the first column, whereas, the second column presents posterior means. In the FMM within the LGD model,

Table 4.3: Results of the LGD model

	posterior mean	HPDI (95%)		posterior odds	naive standard error	time series standard error
				LGD model		
μ_1	0.0000			*not estimated*		
μ_2	0.0067	0.0064	0.0070	∞	0.0000	0.0000
μ_3	0.0290	0.0277	0.0303	∞	0.0000	0.0000
μ_4	0.5114	0.5004	0.5229	∞	0.0000	0.0000
μ_5	1.0000			*not estimated*		
σ_1	0.0010			*not estimated*		
σ_2	0.0045	0.0042	0.0048	∞	0.0000	0.0000
σ_3	0.0249	0.0237	0.0261	∞	0.0000	0.0000
σ_4	0.3364	0.3295	0.3436	∞	0.0000	0.0000
σ_5	0.0010			*not estimated*		
c_1	-0.6959	-0.9773	-0.4082	∞	0.0006	0.0012
c_2	-0.0349	-0.3203	0.2510	1.4857	0.0006	0.0012
c_3	0.8952	0.6087	1.1777	∞	0.0006	0.0012
c_4	2.7509	2.4649	3.0421	∞	0.0007	0.0012
ζ_{EAD}	-0.1099	-0.1357	-0.0824	∞	0.0001	0.0001
ζ_{Facility}	0.2038	0.1495	0.2584	∞	0.0001	0.0001
$\zeta_{\text{Protection}}$	-0.4147	-0.4751	-0.3559	∞	0.0001	0.0001
ζ_{Industry}	-0.2355	-0.3000	-0.1683	∞	0.0002	0.0002
ζ_{HPI}	0.0590	-0.2183	0.3311	2.0243	0.0006	0.0010
				random effect		
σ	0.8191	0.6200	1.0329	∞	0.0005	0.0005

Notes: The table summarizes the results of the LGD model. Parameters are stated in the first column. Categorical variables are included via dummy coding. The reference categories are term loan for facility, no for protection, and non FIRE for industry. The second column presents the posterior means. In the third and fourth column, lower and upper bounds of the corresponding HPDIs to a credibility level of 95% are displayed. The fifth column contains the posterior odds. Naive and time series standard errors are shown in the last two columns. Time series standard errors are calculated based on the effective chain length (N^*_{MCMC}) instead of the actual chain length (N_{MCMC}), whereby, $N^*_{\text{MCMC}} < N_{\text{MCMC}}$ holds for autocorrelated chains.

parameters of the outer components (μ_1 and σ_1 for the first component, μ_5 and σ_5 for the fifth component) are fixed to identify loans with no (LGD = 0) and total (LGD = 1) loss. The second and third component are located nearby the first component ($\mu_2 = 0.0067$ and $\mu_3 = 0.0290$) and have rather small standard deviations ($\sigma_2 = 0.0045$ and $\sigma_3 = 0.0249$), whereas, the forth component seems to cover the range in between the extremes of no and total loss ($\mu_4 = 0.5114$ and $\sigma_4 = 0.3364$). The posterior distributions of the cut points (c_k for $k \in \{1, 2, 3, 4\}$) are not directly interpretable as they depend on the range of the latent variable (Y^*).

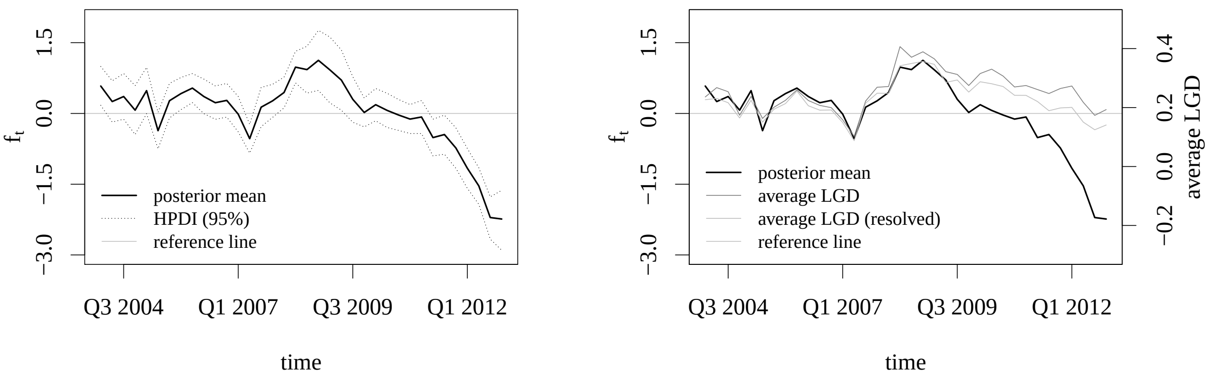
Component probabilities are derived based on the OL model within the LGD model. The parameter of the EAD (ζ_{EAD}) exhibits a negative posterior mean indicating a lower value of the latent variable (Y^*) for higher EADs and, thus, lower LGDs. This impact is characterized by decisive evidence as the posterior odds are tending to infinity (posterior $\text{odds}_{\text{E}[\zeta_{\text{EAD}}]<0} \rightarrow \infty$) and the HPDI ($\text{HPDI}_{\zeta_{\text{EAD}}} = [-0.14, -0.08]$) excludes zero. Reasons for the negative impact of the EAD might be found in higher resolution efforts and, thus, lower loss rates, for loans of major size. The posterior mean of lines (ζ_{Facility}) is positive. Thus, lines are characterized by higher LGDs compared to term loans. This positive influence is decisively evident (posterior $\text{odds}_{\text{E}[\zeta_{\text{Facility}}]>0} \rightarrow \infty$ and $0 \notin \text{HPDI}_{\zeta_{\text{Facility}}} = [0.15, 0.26]$). Protection ($\zeta_{\text{Protection}}$) exhibits a negative posterior mean with decisive evidence (posterior $\text{odds}_{\text{E}[\zeta_{\text{Protection}}]<0} \rightarrow \infty$ and $0 \notin \text{HPDI}_{\zeta_{\text{Protection}}} = [-0.48, -0.36]$). This indicates lower loss rates for protected loans which corresponds to the economic intuition. According to the negative sign of the industry FIRE (ζ_{Industry}), LGDs for loans of this industry affiliation are lower compared to other industries. This impact is decisively evident (posterior $\text{odds}_{\text{E}[\zeta_{\text{Industry}}]<0} \rightarrow \infty$ and $0 \notin \text{HPDI}_{\zeta_{\text{Industry}}} = [-0.30, -0.17]$). The applied macro variable, i.e., the HPI (ζ_{HPI}), exhibits a positive sign indicating higher LGDs for higher values of the HPI. This contradicts the economic intuition as a sound economic surrounding should be accompanied with lower loss rates. However, the positive sign is not statistical evident (posterior $\text{odds}_{\text{E}[\zeta_{\text{HPI}}]>0} = 2.02 < 3.2$ and $0 \in \text{HPDI}_{\zeta_{\text{HPI}}} = [-0.22, 0.33]$). The last row of the table summarizes the posterior distribution of the random effect parameter.[12]

Figure 4.3 illustrates the realizations of the random effect f_t in the LGD model. Higher realizations of the random effect ($f_t > 0$) indicate higher values of the latent variable Y^* for all loans defaulted in t and, thus, higher average LGDs in this quarter. The left panel of the figure presents the time patterns of f_t. The path of f_t seems to be related to the economic cycle. While the realizations of the random effect scatter around zero prior to the crisis, increased values occur since 2007 Q2. In the climax of the GFC, f_t reaches its maximum. The rebound in the aftermath of the crisis instates gradually. The right panel of the figure contrasts these time patterns of f_t to average LGDs in the time line as of Figure 4.2. Thus, the latter include observations which are not considered in the estimation.[13] Up to the more recent time periods, the random effect seems to mimic the path of average LGDs. The time series disperse afterwards, whereby, the spread further increases in the time line. This deviation might be attributed to the resolution bias. In the LGD model, unresolved cases are neglected due to the impossibility

[12] Results are similar to Betz et al. (2018).

[13] The LGD model is based on the estimation sample, whereas, the average LGDs include information of validation sample I, i.e., final LGD observations which are treated as unresolved cases in the estimation (see Table 4.2). We include final observations of unresolved cases to point out effects of the resolution bias.

Figure 4.3: Random effect of the LGD model



Notes: The figure illustrates the course of the random effect in the LGD model over time. In the left panel, the posterior means (thick line) and the HPDI (95%, dotted lines) of the random effect realizations, i.e., f_t, are displayed. In the right panel, the random effect (black line) is contrasted with the time patterns of average LGDs for all loans (dark gray line) and for resolved loans (light gray line). Final and incurrent LGDs as of validation sample I are included in the averaging. The thin lines mark zero and serves as a reference line.

of observing final LGDs. Thus, observations are excluded which tend to have higher LGDs. This distorts the estimated realizations of the random effect in the more recent time periods. The resolution bias and the associated distortion worsen in the time line, i.e., the distortion of f_t enlarges for higher t. Furthermore, a distortion of the random effect parameter σ has to be considered as the downward distortions in the random effect realizations might erroneously increase the underlying standard deviation of the random effect. We will come back to this later on (see subsequent paragraph).

Hierarchical model

In analogy to the LGD model, the hierarchical approach is estimated based on the estimation sample (see Table 4.2). Due to DRT model in the hierarchical approach, it is possible to include censored observations, i.e., unresolved loans, in the estimation process. By this means, we are able to generate posterior predictive distributions for the DRT of unresolved cases and, thus, posterior predictive distributions for the LGD of unresolved loans. Furthermore, effects of the resolution bias as in the pure LGD model (see Figure 4.3) might be diminished.

Table 4.4 summarizes the results of the hierarchical model. Parameters are stated in the first column, whereas, the second column presents posterior means. Posterior distributions for the estimated component parameters (μ_k and σ_k for $k \in \{2,3,4\}$) and loan specific covariate parameters of the LGD model in the hierarchical approach (γ_{EAD}, γ_{Facility}, $\gamma_{\text{Protection}}$, and γ_{Industry}) are similar to their counterparts in the pure LGD model (see Table 4.3, μ_k and σ_k for $k \in \{2,3,4\}$ and ζ_{EAD}, ζ_{Facility}, $\zeta_{\text{Protection}}$, and ζ_{Industry}). A deviation arises for the parameter of the

Table 4.4: Results of the hierarchical model

	posterior mean	HPDI (95%)		posterior odds	naive standard error	time series standard error
		LGD model in the hierarchical approach				
μ_1	0.0000			*not estimated*		
μ_2	0.0064	0.0062	0.0067	∞	0.0000	0.0000
μ_3	0.0279	0.0268	0.0290	∞	0.0000	0.0000
μ_4	0.5033	0.4923	0.5144	∞	0.0000	0.0000
μ_5	1.0000			*not estimated*		
σ_1	0.0010			*not estimated*		
σ_2	0.0043	0.0040	0.0045	∞	0.0000	0.0000
σ_3	0.0234	0.0223	0.0244	∞	0.0000	0.0000
σ_4	0.3384	0.3314	0.3453	∞	0.0000	0.0000
σ_5	0.0010			*not estimated*		
c_1	-1.4391	-1.5803	-1.3004	∞	0.0003	0.0005
c_2	-0.5848	-0.7242	-0.4422	∞	0.0003	0.0006
c_3	0.5728	0.4306	0.7090	∞	0.0003	0.0005
c_4	2.6716	2.5262	2.8169	∞	0.0003	0.0005
γ_{EAD}	-0.1952	-0.2233	-0.1667	∞	0.0001	0.0001
γ_{Facility}	0.3259	0.2700	0.3840	∞	0.0001	0.0001
$\gamma_{\text{Protection}}$	-0.6291	-0.6932	-0.5676	∞	0.0001	0.0002
γ_{Industry}	-0.2736	-0.3437	-0.2036	∞	0.0002	0.0002
γ_{HPI}	-0.0061	-0.1287	0.1170	1.1847	0.0003	0.0005
γ_T	0.9996	0.9711	1.0280	∞	0.0001	0.0001
		DRT model in the hierarchical approach				
β_0	0.7341	0.6112	0.8521	∞	0.0003	0.0006
β_{EAD}	0.0512	0.0343	0.0678	∞	0.0000	0.0000
β_{Facility}	-0.0903	-0.1238	-0.0555	∞	0.0001	0.0001
$\beta_{\text{Protection}}$	0.1345	0.0981	0.1718	∞	0.0001	0.0001
β_{Industry}	-0.1555	-0.1954	-0.1141	∞	0.0001	0.0001
β_{VIX}	0.2731	0.1514	0.3946	7141.8571	0.0003	0.0004
s	0.8488	0.8395	0.8583	∞	0.0000	0.0000
		random effect				
σ_T	0.3424	0.2627	0.4327	∞	0.0002	0.0002
σ_L	0.3615	0.2696	0.4634	∞	0.0002	0.0003
$\omega_{T,L}$	0.1863	-0.1398	0.5031	6.3057	0.0007	0.0008

Notes: The table summarizes the results of the hierarchical model. Parameters are stated in the first column. Categorical variables are included via dummy coding. The reference categories are term loan for facility, no for protection, and non FIRE for industry. The second column presents the posterior means. In the third and fourth column, lower and upper bounds of the corresponding HPDIs to a credibility level of 95% are displayed. The fifth column contains the posterior odds. Naive and time series standard errors are shown in the last two columns. Time series standard errors are calculated based on the effective chain length (N^*_{MCMC}) instead of the actual chain length (N_{MCMC}), whereby, $N^*_{\text{MCMC}} < N_{\text{MCMC}}$ holds for autocorrelated chains.

HPI (γ_{HPI}). In comparison the corresponding parameter in the pure LGD model (ζ_{HPI}) it exhibits an intuitively negative sign, thus, indicating lower LGDs in sound economic surroundings which is displayed by an increasing HPI. However, the parameter of the macro variable is still characterized by a lack of statistical evidence (posterior odds$_{\text{E}[\gamma_{\text{HPI}}]<0} = 1.18 < 3.2$ and $0 \in \text{HPDI}_{\gamma_{\text{HPI}}} = [-0.13, 0.12]$). The sign switch of γ_{HPI} compared to ζ_{HPI} might be due to the inclusion of the logarithmized DRT as explanatory variable in the LGD model of the hierarchical approach (γ_T) as further systematic variables, i.e., the VIX and the random effect of the DRT model, enter the LGD model through the DRT. The posterior mean of γ_T has a positive sign indicating higher LGDs for loans with higher DRTs. In Section 4.2.1 (see Figure 4.1), we determined this relation descriptively. The impact of the DRT is decisively evident (posterior odds$_{\text{E}[\gamma_T]>0} \to \infty$ and $0 \notin \text{HPDI}_{\gamma_T} = [0.97, 1.03]$).

In the DRT model of the hierarchical approach, loan specific covariates and a macro variable, i.e., the VIX, are included. The posterior mean of the EAD (β_{EAD}) exhibits a positive sign. Thus, loans of major size are accompanied with longer DRTs. This supports the thesis we stated in the previous paragraph. Financial institutions might undertake higher resolution efforts for loans of major size. This might increase the DRTs and simultaneously lower LGDs. Decisive evidence can be stated for the positive impact of the EAD in the DRT model (posterior odds$_{\text{E}[\beta_{\text{EAD}}]>0} \to \infty$ and $0 \notin \text{HPDI}_{\beta_{\text{EAD}}} = [0.03, 0.07]$). According to the negative posterior mean of lines (β_{Facility}), this facility type is accompanied with shorter DRTs compared to term loans. This impact is decisively evident (posterior odds$_{\text{E}[\beta_{\text{Facility}}]<0} \to \infty$ and $0 \notin \text{HPDI}_{\beta_{\text{Facility}}} = [-0.12, -0.06]$). In analogy to the EAD, the impact of facility is opposite in the LGD and DRT model of the hierarchical approach. While lines are characterized by shorter DRTs, they result in higher LGDs. Reasons may be found in divergent resolution efforts related to the size of the loan and its protection. The posterior mean of protection ($\beta_{\text{Protection}}$) exhibits a positive, decisively evident (posterior odds$_{\text{E}[\beta_{\text{Protection}}]<0} \to \infty$ and $0 \notin \text{HPDI}_{\beta_{\text{Protection}}} = [-0.69, -0.57]$), sign indicating longer DRTs for protected loans. The impact of protection is divergent among the models in the hierarchical approach ($\gamma_{\text{Protection}} < 0$ and $\beta_{\text{Protection}} > 0$). This might be due to the nature of protection itself. If loans are secured either by collateral or guarantees, efforts have to be taken to realize the protection value. This might extent DRTs, however, reduce LGDs when the protection value is realized. The industry affiliation FIRE (β_{Industry}) reveals a negative posterior mean, thus, it is connected to shorter DRTs. The sign is decisively evident (posterior odds$_{\text{E}[\beta_{\text{Industry}}]<0} \to \infty$ and $0 \notin \text{HPDI}_{\beta_{\text{Industry}}} = [-0.69, -0.57]$) and corresponds to the sign of the LGD model in the hierarchical approach ($\gamma_{\text{Industry}} < 0$ and $\beta_{\text{Industry}} < 0$). Resolution prospects in the FIRE industry might be limited compared to other industries due to less tangible assets. Thus, DRTs are short and LGDs high. To control for the impact of the macro economy, the VIX (β_{VIX}) is included in the DRT model of the hierarchical approach. Its posterior mean is positive and decisively evident (posterior odds$_{\text{E}[\beta_{\text{VIX}}]>0} = 7,141.86 > 100$ and $0 \notin \text{HPDI}_{\beta_{\text{VIX}}} = [0.15, 0.39]$). This entails longer DRTs in bad economic surroundings which corresponds to the economic intuition.

The parameters of the multivariate random effect as of Equation (4.11) are stated in the lower panel of Table 4.4. As the DRT is included in the LGD model of the hierarchical approach, the random effect of the DRT model (F_t^T) enters the LGD model. Thus, the aggregated systematic impact of the random effects on LGDs ($\mathcal{F}_t$) is the linear combination of $\gamma_T F_t^T$ and F_t^L :

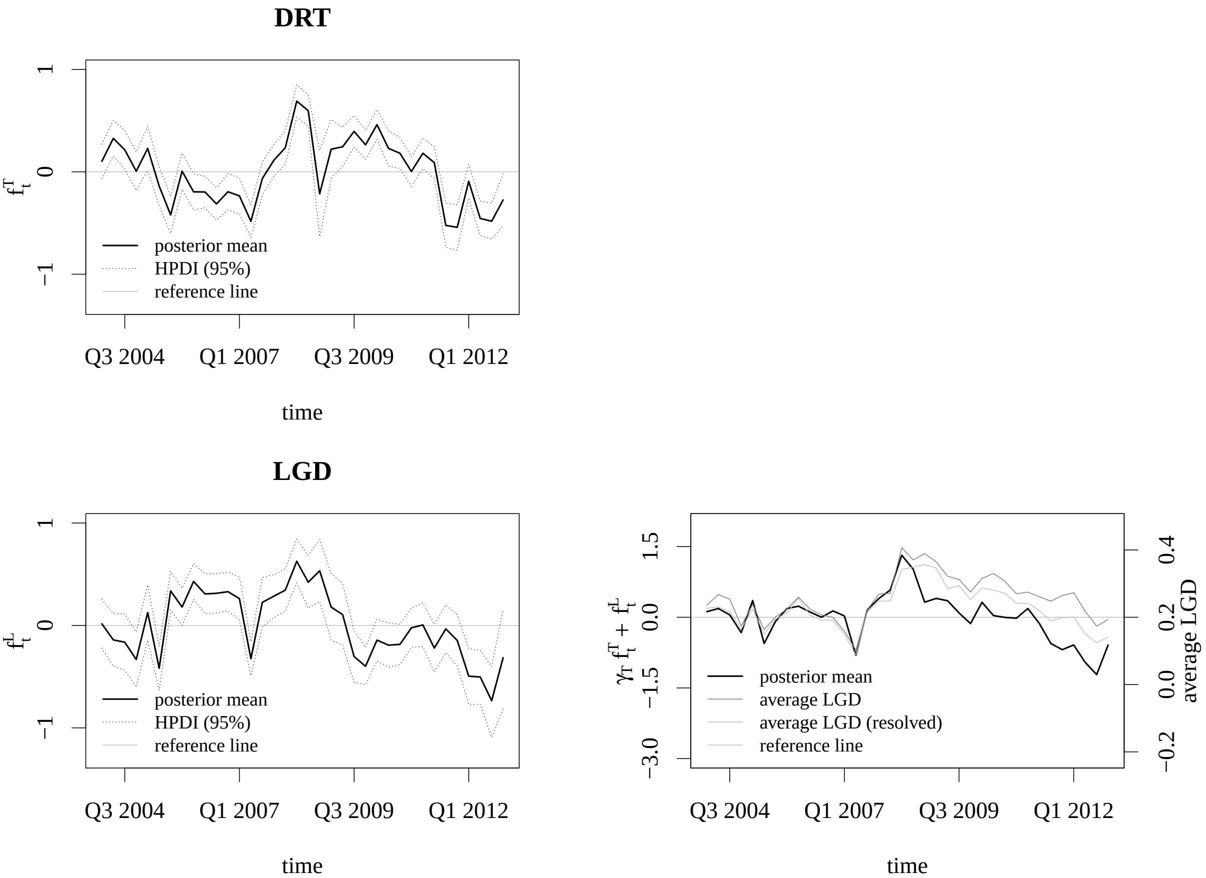
$$\begin{aligned} \mathcal{F}_t &= \gamma_T F_t^T + F_t^L \\ \sigma_{\mathcal{F}}^2 &= \gamma_T^2 \sigma_T^2 + \sigma_L^2 + 2\gamma_T \sigma_T \sigma_L \omega_{T,L}, \end{aligned} \tag{4.13}$$

whereby, $\sigma_{\mathcal{F}}^2$ is the variance of the aggregated systematic effect. Considering the results of Table 4.4, the standard deviation $\sigma_{\mathcal{F}}$ of $\mathcal{F}_t$ amounts to 0.54. This standard deviation is considerably smaller compared to the standard deviation of the random effect in the pure LGD model (see Table 4.3, $\sigma = 0.82$). As suspected in the previous paragraph, the estimated standard deviation of the random effect in the pure LGD model seems to be distorted due to the resolution bias. Neglecting censored observations, i.e., unresolved loans, leads to distorted realizations of the random effect (f_t) and, thus, subsequently to distorted parameters (σ).[14] However, the standard deviation of the random effect should be reduced, if additional explanatory variables are included in the model.

Figure 4.4 illustrates the realizations of the random effects of the DRT model f_t^T (upper left panel) and the LGD model f_t^L (lower left panel) in the hierarchical approach. Higher realizations of the random effect in the DRT model ($f_t^T > 0$) imply higher DRT for all loans defaulted in t, whereas, higher realizations of the random effect in the LGD model ($f_t^L > 0$) lead to higher values of the latent variable $\mathcal{Y}^*$ for all loans defaulted in t and, thus, to higher average LGDs in this quarter. Hence, DRTs impact LGDs through two channels (see Section 4.2.2). Directly, as higher DRTs are inserted in the LGD model. Indirectly, as positive realizations of f_t^T tend to imply positive realizations of f_t^L due to the positive correlation ($\omega_{T,L}$). However, the indirect channel might also weaken the impact of DRTs on LGDs as negative realizations of f_t^L are still possible. Considering the time patterns of the random effects as of Figure 4.4, four settings

[14] Based on our data set, we find $\sigma > \sigma_{\mathcal{F}}$. However, $\sigma < \sigma_{\mathcal{F}}$ is conceivable if the most recent time period is characterized by crisis conditions. Considering the time period from 2004 Q1 to 2008 Q3, f_t would be lower during the crisis and presumable σ would be lower compared to $\sigma_{\mathcal{F}}$.

Figure 4.4: Random effect of the hierarchical model



Notes: The figure illustrates the course of the random effects in the hierarchical model over time. In the left panels, the posterior means (thick lines) and the HPDI (95%, dotted lines) of the random effect realizations, i.e., f_t^T (DRT) and f_t^L (LGD), are displayed. In the right panel, the combined systematic effect on the LGDs according to the random effects of the hierarchical model ($\gamma_T f_t^T + f_t^L$, black line) is contrasted with the time patterns of average LGDs for all loans (dark gray line) and for resolved loans (light gray line). Final and incurrent LGDs as of validation sample I are included in the averaging. The thin lines mark zero and serves as a reference line.

of the indirect channel are apparent. In the first setting prior to the GFC, $f_t^T < 0$ and $f_t^L > 0$ are valid. Thus, average DRTs of loans defaulted in t are shorter. The positive realization of f_t^L, however, increases average LGDs. Resolutions of these loans at least partly take place during the crisis. This might depress recovery payments at the end of the resolution process and, thus, increase LGDs. The second setting in the climax of the GFC is characterized by positive realizations of both random effects ($f_t^T > 0$ and $f_t^L > 0$) indicating longer DRT and simultaneously higher LGDs of loans defaulted in t. In the third setting in the aftermath of the GFC, signs of the random effects are contrary ($f_t^T > 0$ and $f_t^L < 0$). Hence, average DRTs of loans defaulted in t are longer, whereas, average LGDs are lower. This might be due to the time delay as of the first setting. Analogously, parts of the recovery payments take place during the rebound period which favors recovery collection and decreases LGDs. The fourth setting is located in the most recent time period. The realizations of both random effects exhibit negative signs ($f_t^T < 0$ and $f_t^L < 0$) indicating shorter DRTs and simultaneously lower LGDs for loans defaulted in t. These settings illustrate the impacts of systematic effects in the resolution process. The positive correlation of the random effects ($\omega_{T,L}$) seems to be driven by extreme economic surroundings as synchronism appears in crises and boom periods. Furthermore, reasoning for the gradual rebound in the aftermath of the GFC can be provided (see Figure 4.2). While the random effect of the LGD model f_t^L indicates the rebound in the aftermath of the crisis (third setting), the random effect of the DRT model f_t^T remains on its high level. This might be due to the high stock of non-performing loans in the aftermath of the GFC which decelerated resolution proceedings. Average LGDs increase due to the direct channel.

The right panel of Figure 4.4 contrasts the aggregated systematic impact of the random effects ($\mathcal{F}_t$) to average LGDs in the time line. The latter include observations which are not considered in the estimation.[15] The aggregated systematic effect seems to mimic the path of average LGDs. However, slight dispersions are apparent in the more recent time periods. Reasons might be found in a less accurate estimation of the random effect realizations of the LGD model (f_t^L) in the more recent time periods. Although censored observations are included through the DRT model, unresolved loans do not directly enter the LGD model in the hierarchical approach. Comparing the dispersions of the hierarchical model with the pure LGD model (see Figure 4.3), improvements are apparent. While the spread extremely increases in the time line for the pure LGD model, the deviation is considerably less pronounced in the hierarchical approach. Thus, the hierarchical approach succeeds in reducing distortions due to the resolution bias.

4.4 Validation

The validation is conducted on an in sample, out of sample, and out of sample out of time perspective. As stated in Section 4.2.1 (see Table 4.2), the models are estimated based on the estimation sample. In the *in sample* validation, the posterior predictive distributions based on the estimation sample are compared to the empirical distributions of completely resolved loans in the estimation sample. The *out of sample* validation examines the distributional fit for censored observations, i.e., loans which have defaulted till the end of the estimation period but are still unresolved. Thus, Posterior predictive distributions based on validation sample I are compared to the corresponding empirical distribution. The posterior predictive distributions are generated based on the estimated realizations of the random effect. In the *out of sample out*

[15] The presentation corresponds to the right panel of Figure 4.3.

of time validation, loans which defaulted after the end of the estimation period are considered. As no random effect realizations are available for those loans, posterior predictive distributions are generated on the means of the random effects, i.e., zero, and compared to the corresponding empirical distribution.

We adapt two graphical tools to evaluate the distributional fit of the models. First, kernel density estimates of the posterior predictive distributions are compared to kernel density estimates of the empirical data. The bandwidth is fixed to 0.015 to ensure comparability. So heights of the kernel density estimates are comparable despite ties. Second, quantile-quantile plots are applied. Hereby, the quantiles of the posterior predictive distributions are contrasted to the quantiles of the empirical distribution. In the case of optimality, i.e., if the distributions correspond to each other, the points of the quantile-quantile plot are on the bisector line. If the probability of low loss components is overestimated and the probability of high loss components is underestimated, the points are below the bisector line as the the quantiles of the posterior predictive distributions are smaller than the quantiles of the empirical distribution. This corresponds to an underestimation of average LGDs.

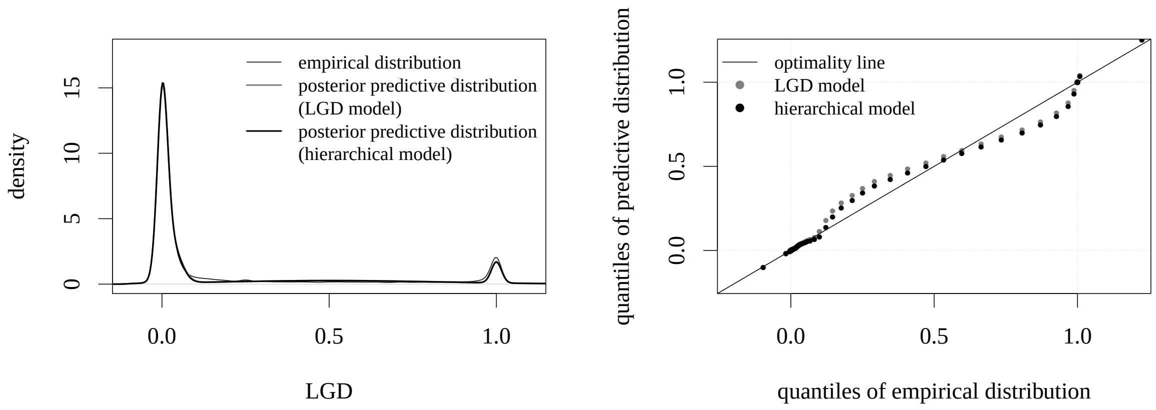
In sample

Figure 4.5 illustrates the *in sample* validation of the LGD model and the hierarchical model. In the left panel, kernel density estimates of the empirical distribution (thin black line), the posterior predictive distribution of the LGD model (thick gray line), and the posterior predictive distribution of the hierarchical model (thick black line) are presented. However, lines lie directly on top of each other, thus, the black line of the posterior predictive distribution of the hierarchical model overlays the remaining two. To get a more detailed impression, the right panel of the figure illustrates quantile-quantile plots, whereby, the quantiles of the posterior predictive distributions are contrasted to the quantiles of the empirical distribution. The gray dots mark the LGD model, whereas, the hierarchical model is represented by black dots. The dots are near the optimality, i.e., bisector, line for both models. Thus, the in sample fit of the posterior predictive distributions is quite good with respect to the LGD model and the hierarchical model. This indicates that the applied FMM with five mixture components seems to deliver satisfactory results regarding the distributional fit.

Out of sample

Figure 4.6 illustrates the *out of sample* validation for the LGD model and the hierarchical model.

Figure 4.5: Validation (*in sample*)



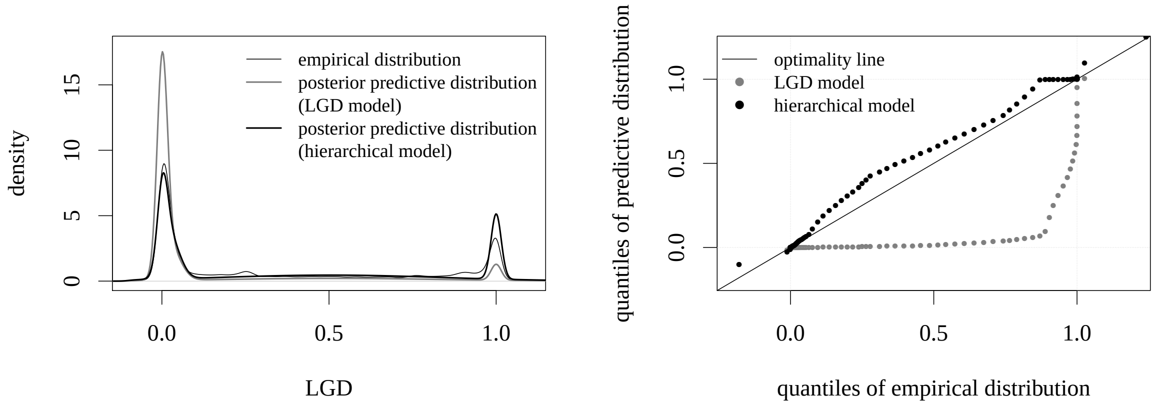
Notes: The figure illustrates the *in sample* validation of the LGD model and the hierarchical model. In the left panel, the kernel density estimates of the empirical distribution (resolved loans as of estimation sample, thin black line), the posterior predictive distribution of the LGD model (thick gray line), and the posterior predictive distribution of the hierarchical model (thick black line) are displayed. The band width is fixed to 0.015 to ensure comparability. The kernel density estimates lie directly on top of each other. Thus, differences are not identifiable. In the right panel, the corresponding quantile-quantile plots are presented. The quantiles of the empirical distribution (x-axis) are plotted against the quantiles of the posterior predictive distributions (y-axis). The gray (black) dots mark the quantiles of the empirical distribution vs. the quantiles of the posterior predictive distribution of the LGD model (hierarchical model). The black line represents optimality.

The presentation corresponds to Figure 4.5 (in sample validation). The posterior predictive distribution of the LGD model (gray) seems to overestimate the probability mass of the low loss components, whereas, it underestimates the probability mass of the high loss components. In the left panel, its kernel density estimate lies above the kernel density estimate of the empirical distribution for no loss (LGD = 0) and below for total loss (LGD = 1). This appears even clearer considering the quantile-quantile plots in the right panel. The dots are considerably below the optimality line indicating an underestimation of average LGDs. The underestimation is caused by the parameter distortion of the random effect due to the resolution bias. The random effect realizations f_t are characterized by a downward bias, thus, leading to downward biased estimates for Y_i and downward biased loss rates (see Section 4.3). In contrast, the distributional fit of the hierarchical model is good on an out of sample perspective. This is due to two reasons. First, the parameter distortions caused by the resolution bias are diminished (see Section 4.3). Second, the additional information of how long a loan is in resolution is utilized to improve predictions on an out of sample perspective. Final DRTs for censored observations, i.e., unresolved cases, are estimated within the hierarchical approach. These can be applied to generate predictions of final LGDs for unresolved loans.

Out of sample out of time

Figure 4.7 illustrates the *out of sample out of time* validation of the LGD model and the hier-

Figure 4.6: Validation (*out of sample*)



Notes: The figure illustrates the *out of sample* validation of the LGD model and the hierarchical model. In the left panel, the kernel density estimates of the empirical distribution (resolved loans as of validation sample I, thin black line), the posterior predictive distribution of the LGD model (for resolved loans, thick gray line), and the posterior predictive distribution of the hierarchical model (for resolved loans, thick black line) are displayed. The band width is fixed to 0.015 to ensure comparability. In the right panel, the corresponding quantile-quantile plots are presented. The quantiles of the empirical distribution (x-axis) are plotted against the quantiles of the posterior predictive distributions (y-axis). The gray (black) dots mark the quantiles of the empirical distribution vs. the quantiles of the posterior predictive distribution of the LGD model (hierarchical model). The black line represents optimality.

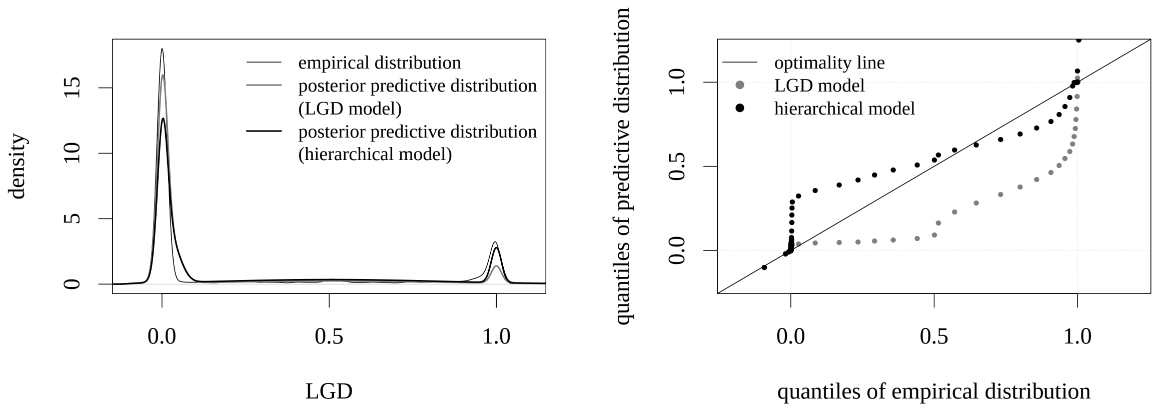
archical model. The presentation corresponds to Figure 4.5. In analogy to the out of sample validation, an overestimation of low loss components and an underestimation of high loss components arises for the LGD model. However, it is not as striking as in the out of sample validation (see Figure 4.6). This might be due to the use of the random effect in average terms – instead of the individual realizations f_t as in the out of sample validation – to generate the posterior predictive distribution. However, the poor distributional fit of the LGD model on the out of sample out of time perspective suggests that there are additional distortions beyond the realizations of the random effect and its standard deviation. These might be found in the cut points which represent the intercepts in an OL model.[16] In contrast to the LGD model, the distributional fit of the hierarchical model is quite good on an out of sample out of time perspective.

Validation in the time line

Thus far, the distributional fit of the LGD model and the hierarchical model are analyzed for the estimation sample and the validation samples. Figure 4.8 illustrates the time patterns of average LGD predictions based on the posterior predictive distributions for specific default quarters. The upper left panel contrast average LGDs (thin black line) to average LGD predictions based on the LGD model (thick gray line) and the hierarchical model (thick black line) on an *in*

[16] The cut points of the LGD model and the hierarchical model are not directly comparable as the logarithm of the DRT is included as additional variable. By this means, the mean of the latent variable $\mathcal{Y}^*$ and, thereby, the level of the cut points are shifted.

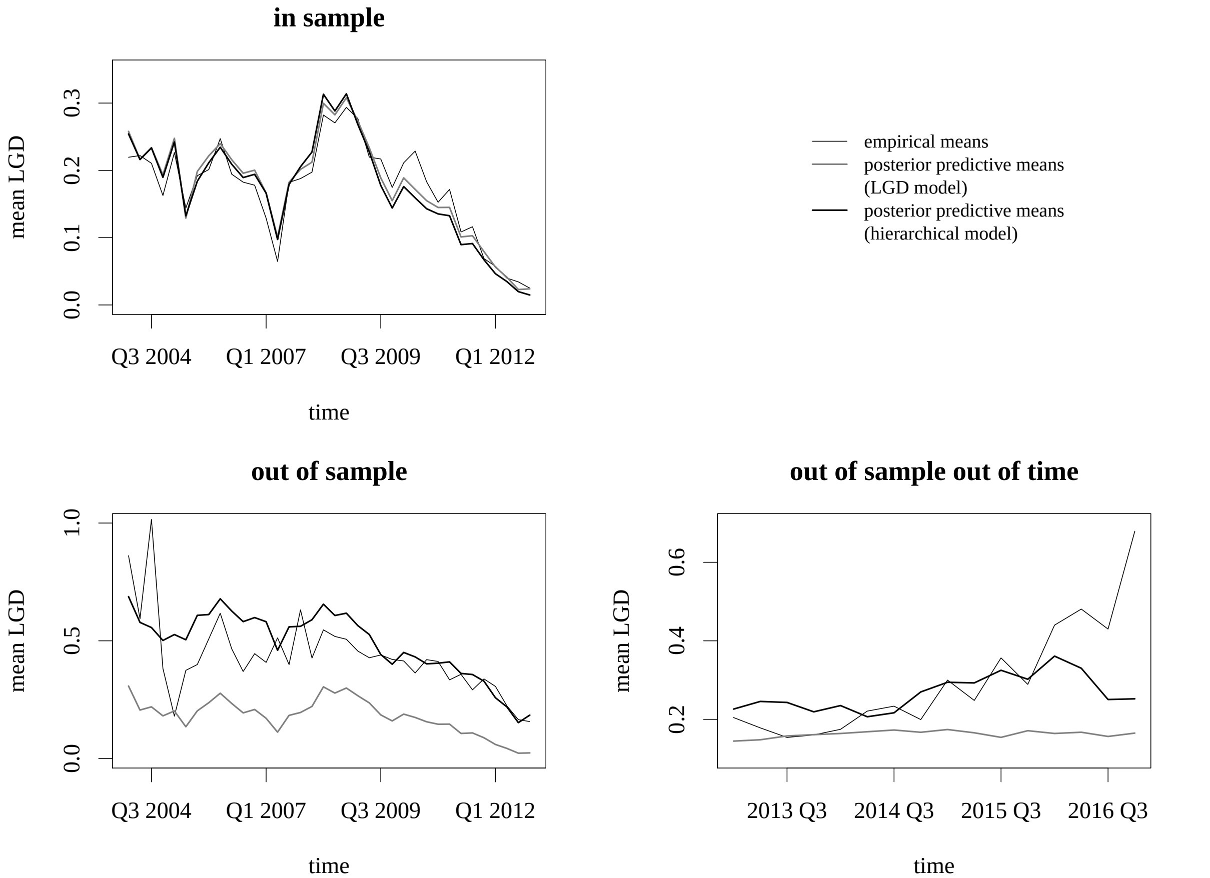
Figure 4.7: Validation (*out of sample out of time*)



Notes: The figure illustrates the *out of sample out of time* validation of the LGD model and the hierarchical model. In the left panel, the kernel density estimates of the empirical distribution (all loans as of validation sample II with incurrent LGDs, thin black line), the posterior predictive distribution of the LGD model (thick gray line), and the posterior predictive distribution of the hierarchical model (thick black line) are displayed. The band width is fixed to 0.015 to ensure comparability. In the right panel, the corresponding quantile-quantile plots are presented. The quantiles of the empirical distribution (x-axis) are plotted against the quantiles of the posterior predictive distributions (y-axis). The gray (black) dots mark the quantiles of the empirical distribution vs. the quantiles of the posterior predictive distribution of the LGD model (hierarchical model). The black line represents optimality.

sample perspective. In analogy to Figure 4.5, a good in sample fit for both models can be stated. The lower left panel illustrates the time patterns of average LGDs and LGD predictions on an *out of sample* perspective. Although the relative progressions of the LGD predictions based on the LGD model and the hierarchical model are similar, the predictions based on the LGD model are downward biased. Thus, average LGDs are underestimated by the LGD model in almost all quarters in validation sample I. This is not the case considering the predictions of the hierarchical model. The noisy behavior of average LGDs at the beginning of the time period is due to a lack of data as most loans defaulted in these quarters are resolved by the end of 2010 and, thus, not included in validation sample I. The lower right panel illustrates the time patterns of average LGDs and LGD predictions on an *out of sample out of time* perspective. The predictions based on the LGD model seem to be constant through time as the random effect is set to its mean, i.e., zero, and the macro variable is the only remaining systematic factor. However, the latter is not statistically evident (see Table 4.3). Furthermore, LGD predictions based on the LGD model seem to be systematically too low. LGD predictions based on the hierarchical model better fit average LGDs. Deviations at the end of the time period might be attributed to the inclusion of incurrent LGDs for unresolved cases (see Figure 4.2). Final LGDs will be lower and adjust the line downwards. In addition, LGD predictions based on the hierarchal model display systematic movement as the statistically evident macro variable of the DRT model is enclosed in the LGD model of the hierarchical approach (see Table 4.4).

Figure 4.8: Validation in the time line



Notes: The figure illustrates the validation in the time line. The means of the empirical distribution are displayed by a thin black line, whereas, the means of the posterior predictive distributions are marked by a thick gray line for the LGD model and a thick black line for the hierarchical model, respectively. In the upper panel, the *in sample* validation in the time line is presented (empirical means of resolved loans in estimation sample). The lower panels show the *out of sample* (empirical means of resolved loans in validation sample I) and *out of sample out of time* validation (empirical means of all loans in validation sample II with incurrent LGDs) in the time line.

4.5 Conclusion

In this paper, we deeply examine the dependence structure of DRTs and LGDs using a hierarchical modeling framework. We find direct and indirect dependencies among the credit risk parameters. First, LGDs seem to be directly impacted by DRTs, i.e., longer resolution processes are accompanied with higher losses. Second, the parameters are characterized by common time patterns as correlation of the random effects in the individual models is positive. Due to the random nature of these effects, the dependence of DRTs and LGDs might be intensified or weakened in certain time periods. We find similar signs of the random effect realizations during the GFC and deviating signs pre and post crisis. Due to the consideration of direct dependency structures, we are able to generate intuitive LGD predictions for censored cases, i.e., non-performing loans. As final DRTs for unresolved loans are estimated within the DRT model of the hierarchical approach, these estimations can be used to directly generate final LGD predictions for these cases. LGD predictions based on the hierarchical approach, thereby, outperform predictions based on a pure LGD model.

Furthermore, effects of the resolution bias are diminished in the hierarchical approach. While the parameters and, thus, the realizations of the random effect are biased in a pure LGD model, these distortions are eliminated in the hierarchical approach. The parameter distortions due to the resolution bias have considerable impacts on the *out of sample* and *out of sample out of time* performance of pure LGD models. Out of sample, a pure LGD model generates average LGD predictions underestimating actual average LGDs by up to 25 percentage points for loans defaulted during the GFC (2008 Q1 to 2009 Q3). The hierarchical approach delivers sufficiently conservative predictions for loans defaulted in the crisis (up to 16 percentage points above actual average LGDs). Assuming ten loans with an EAD of 500,000 EUR defaulted in 2008 Q4, a pure LGD model underestimates losses due to these loans by round about 1,05 million EUR. Out of sample out of time, these effects are less pronounced, however, still remarkable. A pure LGD model constantly underestimates actual average LGDs in the time period from 2013 Q1 to 2015 Q4 by up to 20 percentage points, while the hierarchical approach delivers slightly conservative predictions in most time periods (between 3 percentage points below and 8 percentage points above actual LGDs).[17] Assuming ten loans with an EAD of 500,000 EUR defaulted in 2015 Q4, a pure LGD model underestimates losses due to these loans by round about 600,000 EUR.

Concluding, the consideration of the dependency structure of DRTs and LGDs and the thereto entailed resolution bias is essential to generate suitable LGD predictions. The presented hierarchical model prevents the need of additional data constraints and provides fruitful insights into the dependence structure of DRTs and LGDs.

[17] The subsequent time period is hard to interpret due to incurrent LGDs.

4.A Appendix | Bayesian model specification

The LGD model and the hierarchical model are estimated via Bayesian inference.[18] Thus, prior distributions have to be specified for every parameter in the models. Most of the prior distributions are characterized by an uninformative specification. If weakly informative priors are a provided, this is done to avoid convergence reasons.

LGD model

In the FMM, the number of latent classes is set to five ($K = 5$). As the LGD distribution is extremely bimodal and characterized by high probability masses at zero and one, we fix the parameters of the outer components ($k = 1$ and $k = 5$). The means are set to $\mu_1 = 0$ and $\mu_5 = 1$ with rather small standard deviations ($\sigma_1 = \sigma_5 = 0.001$) to identify loans with no and total loss. The remaining component parameters are provided with uninformative prior distributions:

$$\begin{aligned}
&\mu_2 \sim \mathrm{N}\Big(\, \mu = 0.0 \,,\, \sigma = 1 \cdot 10^5 \,\Big) \\
&\mu_3 \sim \mathrm{N}\Big(\, \mu = 0.5 \,,\, \sigma = 1 \cdot 10^5 \,\Big) \\
&\mu_4 \sim \mathrm{N}\Big(\, \mu = 1.0 \,,\, \sigma = 1 \cdot 10^5 \,\Big) \\
&\sigma_k \sim \mathrm{N}\Big(\, \mu = 0.0 \,,\, \sigma = 1 \cdot 10^5 \,\Big)\, [\, 0 \,,\ \ ] \text{ for } k \in \{2,3,4\}\,,
\end{aligned} \tag{4.14}$$

where, the squared brackets indicate truncation. The prior distributions of the component means are Normal distributions with mean $\mu = 0.0$ for component 2, $\mu = 0.5$ for component 3, and $\mu = 1.0$ for component 4. The means of the components are ordered and, thus, truncated to the interval $[0,1]$ due to the fixed outer components. The standard deviations of the prior distributions are set to a rather high value which corresponds to a uninformative prior specification. The component standard deviations are provided with an uninformative truncated Normal prior distribution. We forgo the conjugate prior distributions (inverse Gamma distribution for the variance) as convergence problems might arise for small values of σ_k.

The cut points c_k for $k \in \{1,\dots K-1\}$ of the OL model are restricted to be ordered ($c_1 < \cdots < c_{K-1}$) and provided with uninformative prior distributions:

$$c_k \sim \mathrm{N}\Big(\, \mu = 0 \,,\, \sigma = 1 \cdot 10^5 \,\Big) \text{ for } k \in \{1,\dots,K-1\}\,. \tag{4.15}$$

[18] The MCMC samples are drawn via the sampler Stan.

The prior distributions of the coefficients ζ_j for $j \in \{1,\ldots,J\}$ are uninformative Normal prior distributions with mean 0 and high standard deviations:

$$\zeta_j \sim \mathrm{N}\Big(\ \mu = 0\ ,\ \sigma = 1 \cdot 10^5\ \Big) \text{ for } j \in \{1,\ldots,J\}\,. \tag{4.16}$$

The random effect F_t follows a normal distribution with mean zero and standard deviation σ. We provide the standard deviation of the random effect σ with uninformative prior distribution:

$$\sigma \sim \mathrm{Gamma}\,(\ 0.001\ ,\ 0.001\ )\ . \tag{4.17}$$

Hierarchical model

In the hierarchical model, we adopt the prior distributions of the components as of Equation (4.14) and the cut points as of Equation (4.15). The coefficients of the DRT model β_j for $j \in \{1,\ldots,J_T\}$ and the LGD model γ_j for for $j \in \{1,\ldots,J_L\}$ are also provided with uninformative Normal prior distributions:

$$\beta_j \sim \mathrm{N}\Big(\ \mu = 0\ ,\ \sigma = 1 \cdot 10^5\ \Big) \text{ for } j \in \{1,\ldots,J_T\} \tag{4.18}$$

$$\gamma_j \sim \mathrm{N}\Big(\ \mu = 0\ ,\ \sigma = 1 \cdot 10^5\ \Big) \text{ for } j \in \{1,\ldots,J_L\}\,. \tag{4.19}$$

The random effects F_t^T and F_t^L in the hierarchical model follow a bivariate normal distribution, whereby, the mean vector corresponds to the two dimensional zero vector ($0_2 = (\ 0 \quad 0\ )^T$). We provide the correlation matrix Ω of the bivariate normal distribution with an uninformative LKJ prior distribution (see Stan Development Team, 2016) and the standard deviations of the random effects σ_T and σ_L with uninformative gamma distributions:

$$\Omega \sim \mathrm{LKJ}\,(\ 1\ ) \tag{4.20}$$

$$\sigma_T \sim \mathrm{Gamma}\,(\ 0.001\ ,\ 0.001\ ) \tag{4.21}$$

$$\sigma_L \sim \mathrm{Gamma}\,(\ 0.001\ ,\ 0.001\ )\ . \tag{4.22}$$

The covariance matrix Σ of the bivariately distributed random effects might be calculated based on the correlation matrix Ω and the individual standard deviations σ_T and σ_L.

4.B Appendix | Bayesian convergence diagnostics

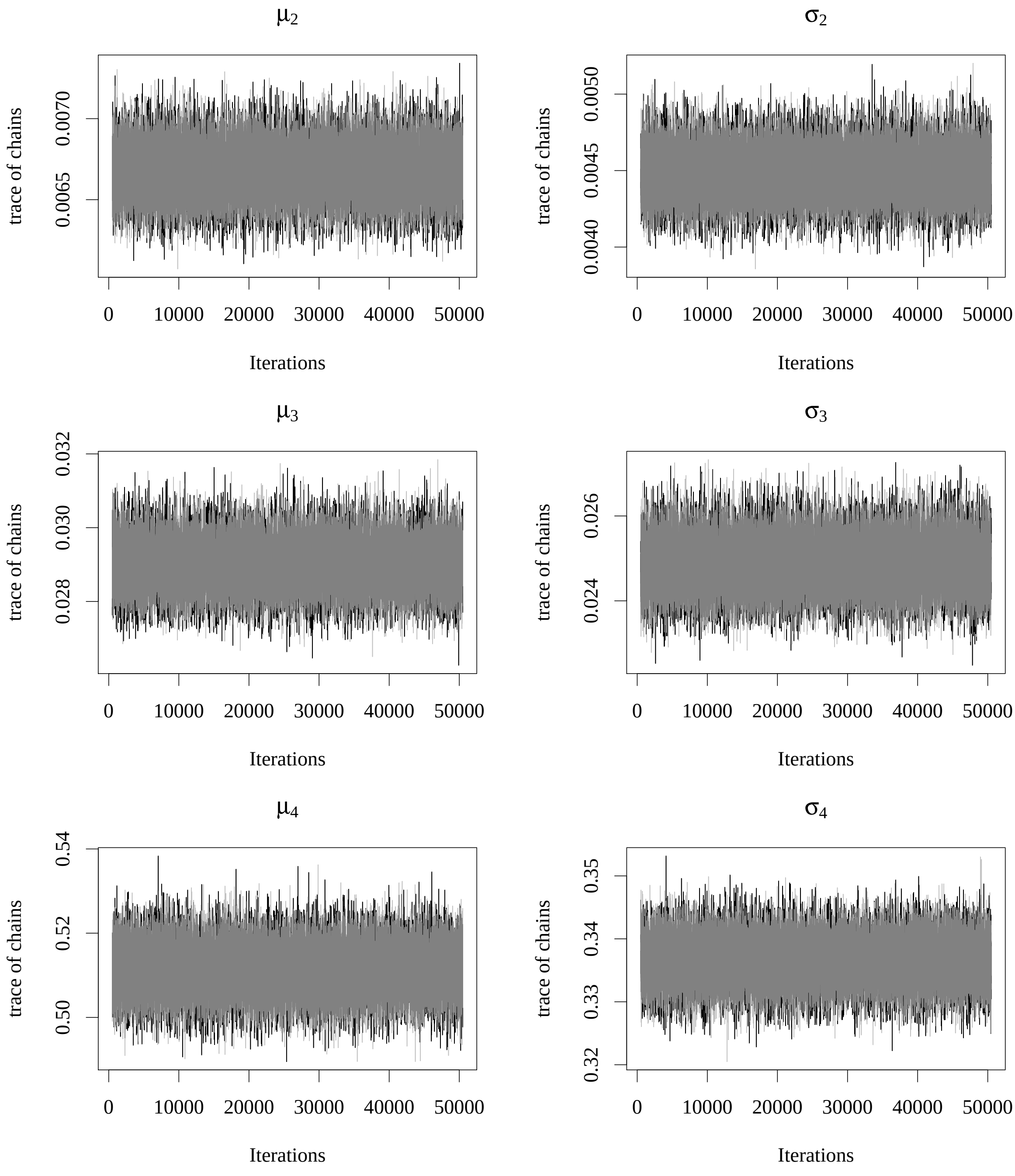
LGD model

Table 4.B.1: Convergence diagnostics of LGD model

	Gelman-Rubin diagnostic		**Heidelberger-Welch** diagnostic		
	point estimate	upper confidence limits (95%)	stationary test	start	p value
μ_2	1.0000	1.0002	passed	1	0.3857
μ_3	1.0000	1.0000	passed	1	0.2169
μ_4	1.0000	1.0002	passed	1	0.2180
σ_2	1.0000	1.0000	passed	1	0.6438
σ_3	1.0000	1.0000	passed	1	0.3903
σ_4	1.0001	1.0008	passed	1	0.4836
c_1	1.0001	1.0005	passed	1	0.4102
c_2	1.0000	1.0003	passed	1	0.4826
c_3	1.0001	1.0005	passed	1	0.4600
c_4	1.0001	1.0005	passed	1	0.3650
ζ_{EAD}	1.0002	1.0013	passed	1	0.0595
ζ_{Facility}	1.0000	1.0001	passed	1	0.4377
$\zeta_{\text{Protection}}$	1.0000	1.0001	passed	1	0.6311
ζ_{Industry}	1.0002	1.0013	passed	1	0.4295
ζ_{HPI}	1.0000	1.0000	passed	1	0.3440
sigma	1.0000	1.0000	passed	1	0.9538

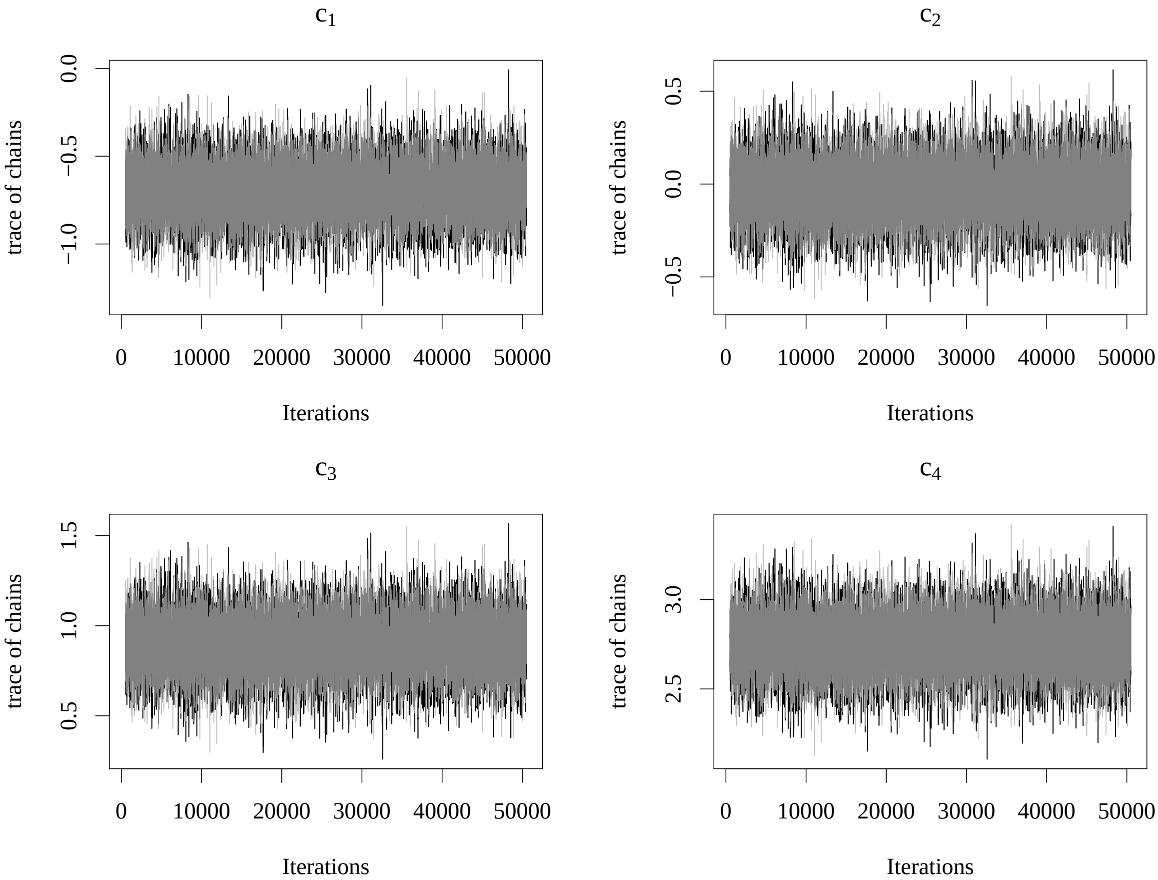
Notes: The table summarizes the convergence diagnostics of the LGD model. Parameters are stated in the first column. In the second and third column, the Gelman-Rubin diagnostics are displayed. The *potential reduction factor* and its upper confidence limit are calculated. Convergence is diagnosed if chains have ”forgotten” their initial values, i.e., for upper limit close to one (see Gelman and Rubin, 1992). A rule of thumb assumes 1.1 as critical value. The Gelman-Rubin diagnostic examines the length of burn-in. In the last three columns, the Heidelberger-Welch diagnostics are presented. The two chains are combined to calculate a criterion of relative accuracy for the posterior means. The frequentistic stationarity test adopts the Cramer-von-Mises statistic to test the null hypothesis of a stationary process in the chains (see Heidelberger and Welch, 1981, 1983). The Heidelberger-Welch diagnostic examines the length of chains.

Figure 4.B.1: Trace plots of the component parameters (LGD model)



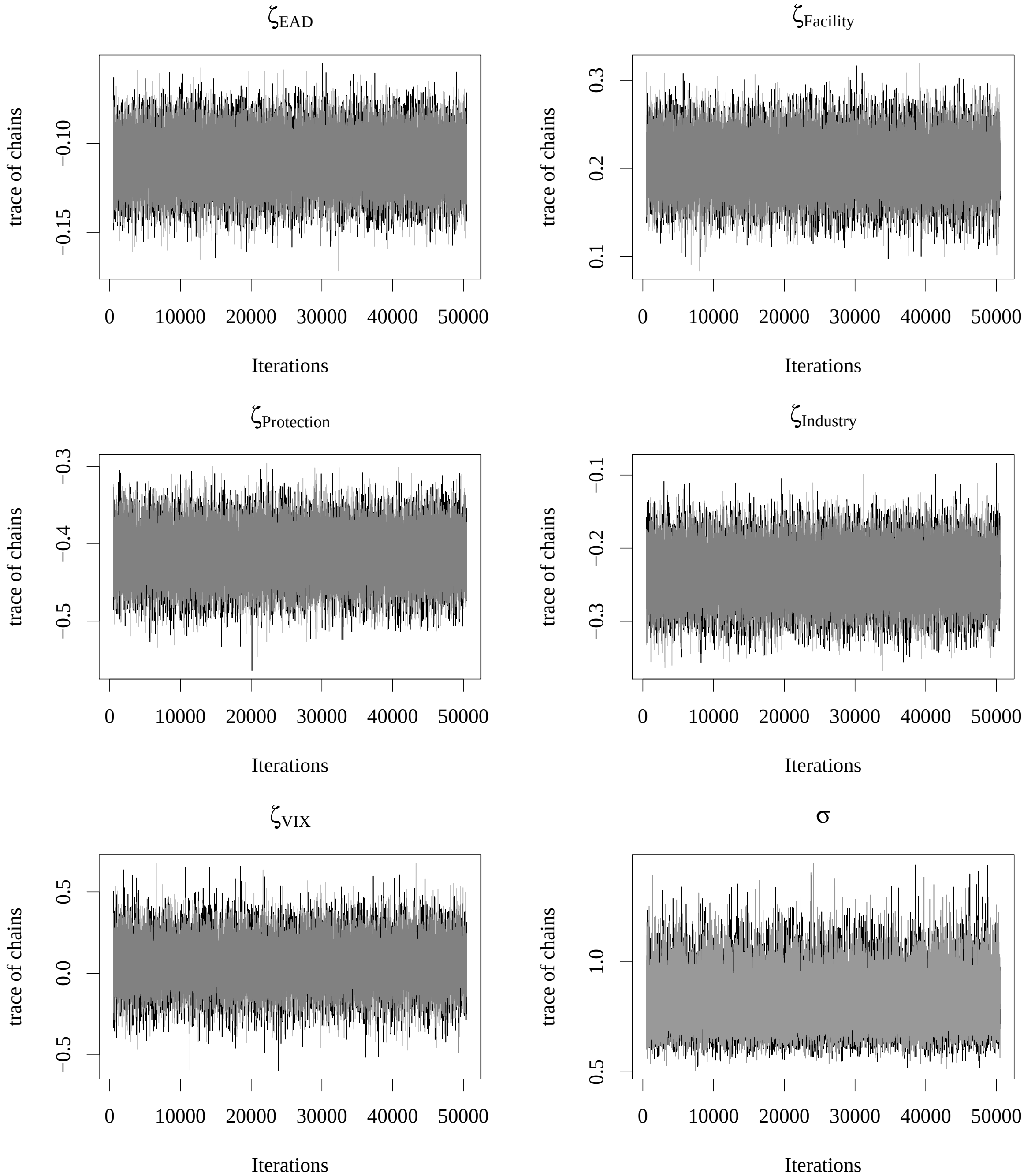
Notes: The figure illustrates the trace of MCMC chains for the component parameters (μ_k and σ_k for $k \in \{2,3,4\}$) in the LGD model. The first chain is displayed in black, the second in gray.

Figure 4.B.2: Trace plots of the cut points (LGD model)



Notes: The figure illustrates the trace of MCMC chains for the cut points (c_k for $k \in \{1, 2, 3, 4\}$) in the LGD model. The first chain is displayed in black, the second in gray.

Figure 4.B.3: Trace plots of the covariate parameters ζ_j and the random effect parameter σ (LGD model)



Notes: The figure illustrates the trace of MCMC chains for the covariate parameters (ζ_j for $j \in$ {EAD, Facility, Protection, Industry, HPI}) and the parameter of the random effect (σ) in the LGD model. The first chain is displayed in black, the second in gray.

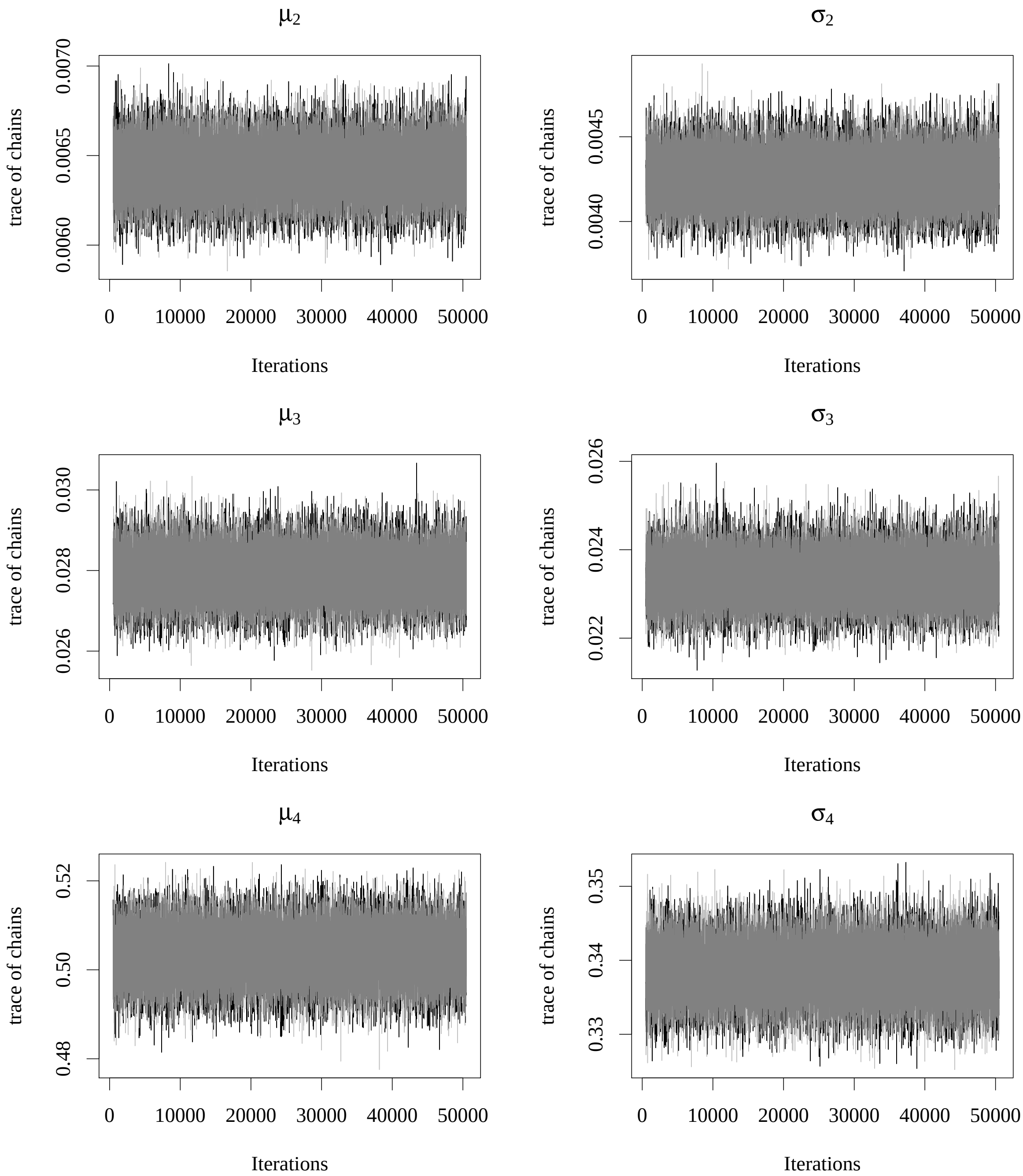
Hierarchical model

Table 4.B.2: Convergence diagnostics of hierarchical model

	Gelman-Rubin diagnostic		**Heidelberger-Welch** diagnostic		
	point estimate	upper confidence limits (95%)	stationary test	start	p value
μ_2	1.0000	1.0000	passed	1	0.17944
μ_3	1.0000	1.0002	passed	1	0.87543
μ_4	1.0000	1.0001	passed	1	0.44269
σ_2	1.0000	1.0000	passed	1	0.69836
σ_3	1.0002	1.0003	passed	1	0.54110
σ_4	1.0002	1.0008	passed	1	0.89167
c_1	1.0000	1.0000	passed	1	0.50214
c_2	1.0000	1.0000	passed	1	0.74251
c_3	1.0000	1.0002	passed	1	0.83016
c_4	1.0001	1.0004	passed	1	0.87461
γ_{EAD}	1.0000	1.0000	passed	1	0.61537
γ_{Facility}	1.0000	1.0001	passed	1	0.72393
$\gamma_{\text{Protection}}$	1.0001	1.0009	passed	1	0.20917
γ_{Industry}	1.0000	1.0001	passed	1	0.62439
γ_{HPI}	1.0004	1.0005	passed	1	0.27974
γ_T	1.0000	1.0000	passed	5001	0.23503
β_0	1.0001	1.0004	passed	1	0.32351
β_{EAD}	1.0000	1.0000	passed	1	0.70406
β_{Facility}	1.0000	1.0000	passed	1	0.38782
$\beta_{\text{Protection}}$	1.0000	1.0001	passed	1	0.42642
β_{Industry}	1.0001	1.0007	passed	1	0.23275
β_{VIX}	1.0003	1.0015	passed	10001	0.21704
s	1.0000	1.0001	passed	1	0.46664
σ_T	1.0000	1.0002	passed	1	0.99576
σ_L	1.0000	1.0000	passed	1	0.18011
$\omega_{T,L}$	1.0000	1.0003	passed	1	0.47216

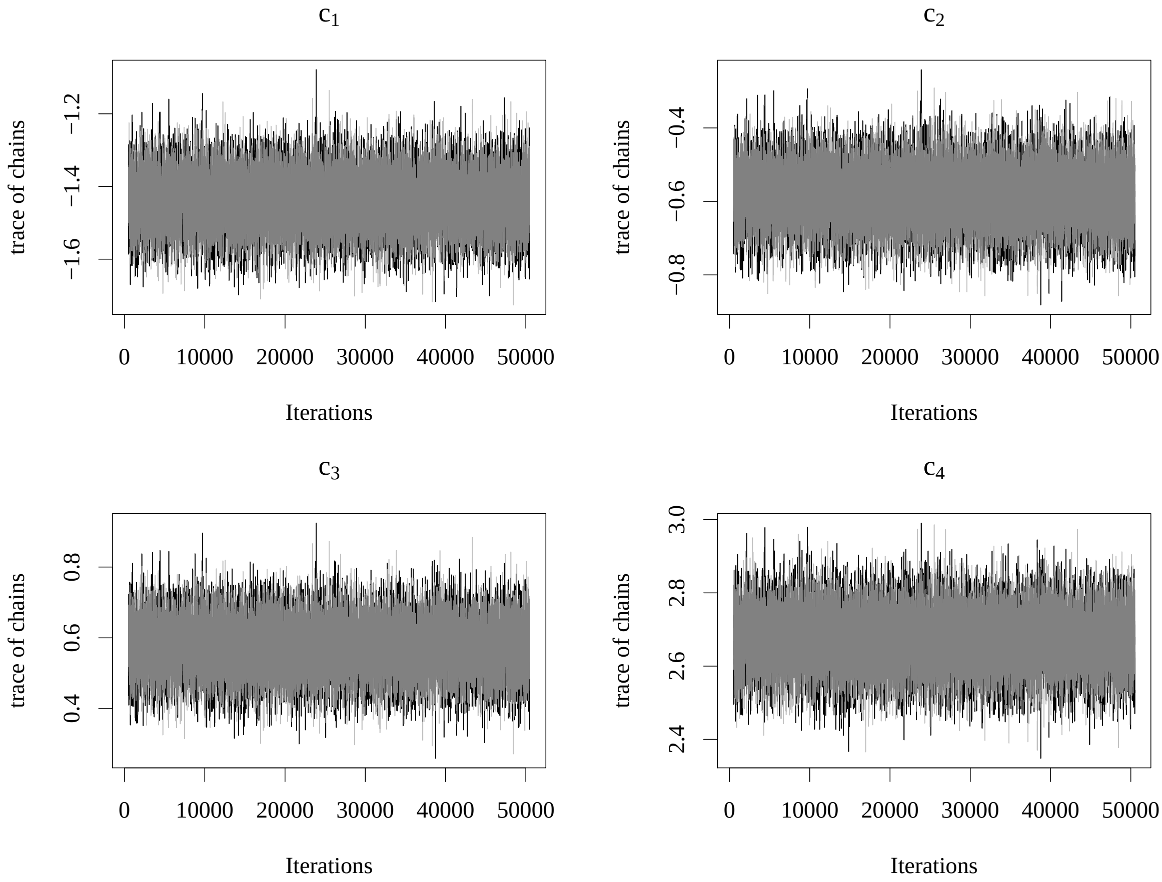
Notes: The table summarizes the convergence diagnostics of the hierarchical model. Parameters are stated in the first column. In the second and third column, the Gelman-Rubin diagnostics are displayed. The *potential reduction factor* and its upper confidence limit are calculated. Convergence is diagnosed if chains have ”forgotten” their initial values, i.e., for upper limit close to one (see Gelman and Rubin, 1992). A rule of thumb assumes 1.1 as critical value. The Gelman-Rubin diagnostic examines the length of burn-in. In the last three columns, the Heidelberger-Welch diagnostics are presented. The two chains are combined to calculate a criterion of relative accuracy for the posterior means. The frequentistic stationarity test adopts the Cramer-von-Mises statistic to test the null hypothesis of a stationary process in the chains (see Heidelberger and Welch, 1981, 1983). The Heidelberger-Welch diagnostic examines the length of chains.

Figure 4.B.4: Trace plots of the component parameters (hierarchical model)



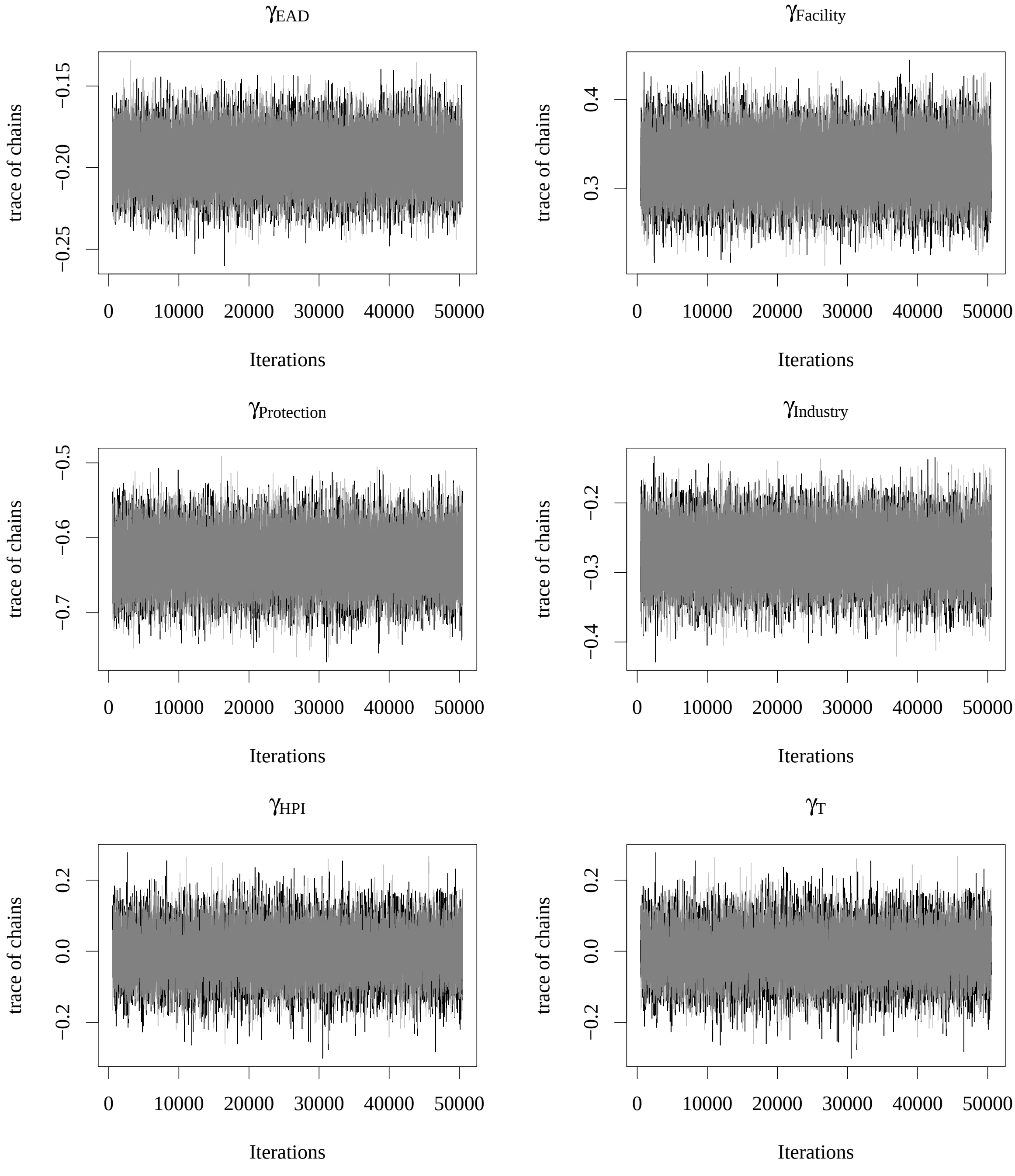
Notes: The figure illustrates the trace of MCMC chains for the component parameters (μ_k and σ_k for $k \in \{2,3,4\}$) in the hierarchical approach. The first chain is displayed in black, the second in gray.

Figure 4.B.5: Trace plots of the cut points (hierarchical model)



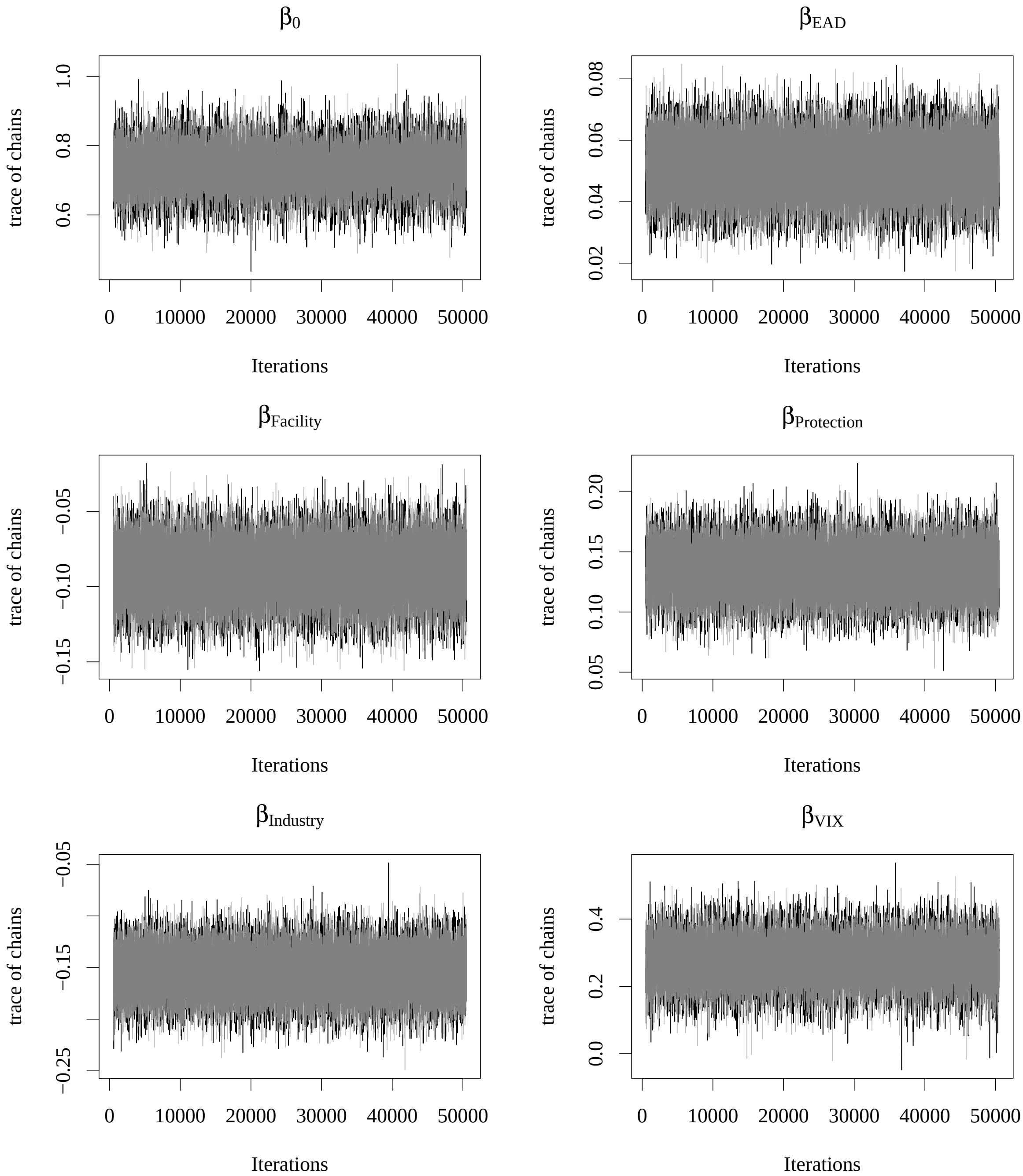
Notes: The figure illustrates the trace of MCMC chains for the cut points (c_k for $k \in \{1,2,3,4\}$) in the hierarchical approach. The first chain is displayed in black, the second in gray.

Figure 4.B.6: Trace plots of the covariate parameters γ_j and impact parameter γ_T (hierarchical model)



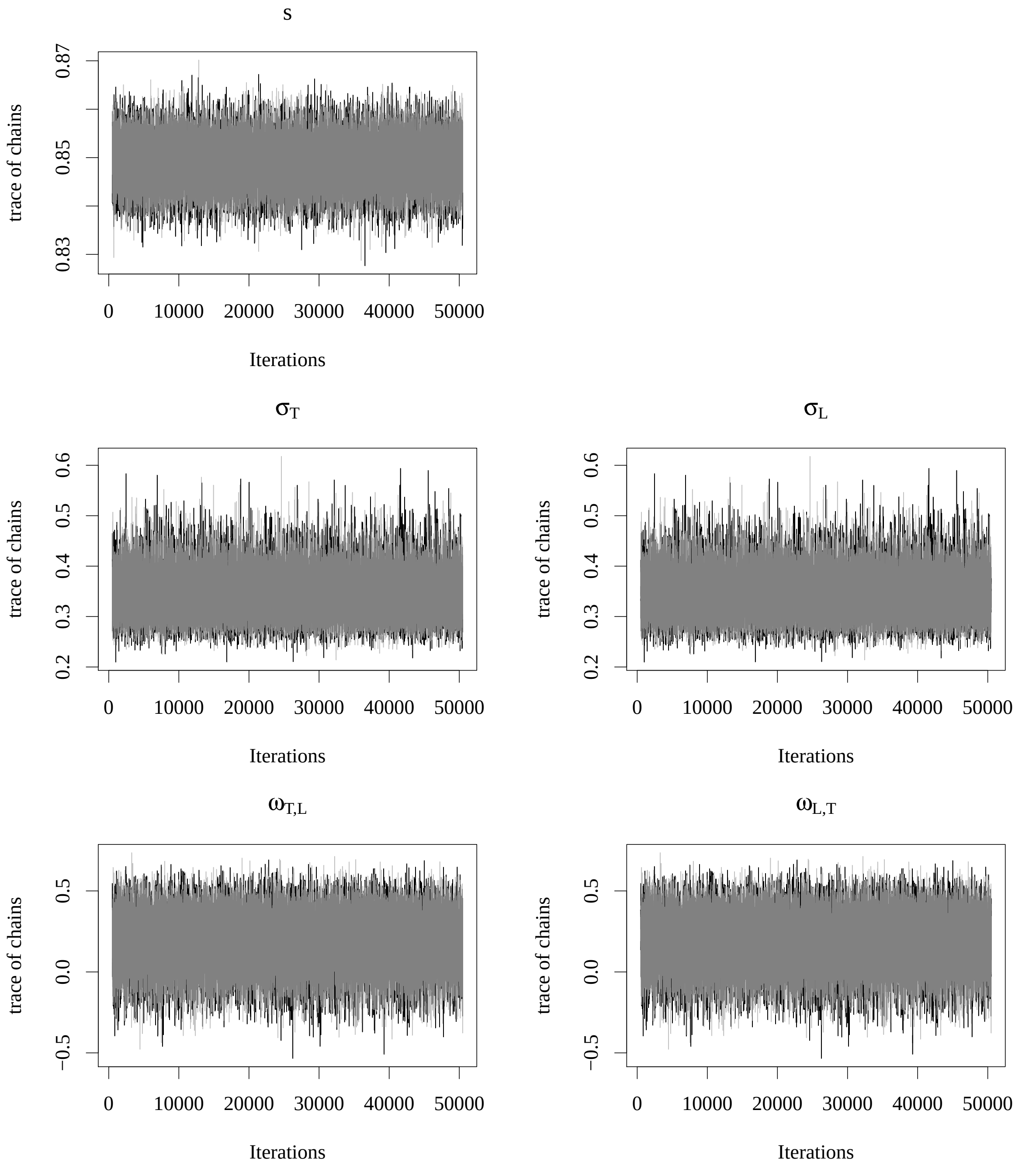
Notes: The figure illustrates the trace of MCMC chains for the covariate parameters (γ_j for $j \in$ {EAD, Facility, Protection, Industry, HPI}) and the impact parameter of the DRT (γ_T) of the LGD model in the hierarchical approach. The first chain is displayed in black, the second in gray.

Figure 4.B.7: Trace plots of the intercept β_0 and covariate parameters β_j (hierarchical model)



Notes: The figure illustrates the trace of MCMC chains for the intercept (β_0) and the covariate parameters (β_j for $j \in$ {EAD, Facility, Protection, Industry, VIX}) of the DRT model in the hierarchical approach. The first chain is displayed in black, the second in gray.

Figure 4.B.8: Trace plots of the standard error s and the random effect parameters σ_T, σ_L, $\omega_{T,L}$, and $\omega_{L,T}$ (hierarchical model)



Notes: The figure illustrates the trace of MCMC chains for the standard error (s) of the DRT model and the random effect parameters (σ_T, σ_L, $\omega_{T,L}$, and $\omega_{L,T}$) in the hierarchical approach. The first chain is displayed in black, the second in gray.

Conclusion

Summary

This thesis focuses on the resolution of defaulted loan contracts. It is composed of a profound empirical analysis of the DRT and the LGD – the two central parameters of the resolution process. In the first research paper *What drives the time to resolution of defaulted bank loans?* (see Chapter 1), general loan-specific and macro(-economic) drivers of the DRT are identified, whereas, the second research paper *Macroeconomic effects and frailties in the resolution of non-performing loans* (see Chapter 2) analyzes systematic effects among DRTs considering not only observable, i.e., macro(-economic), variables, but also unobservable systematic factors in terms of frailty effects. Both research papers emphasize the high relevance of the DRT considering LGDs of defaulted loan contracts by descriptive explorations.

The empirical analysis of LGDs is subject to the third research paper *Systematic effects among LGDs and their implications on downturn estimation* (see Chapter 3). In analogy to the second research papers, observable and unobservable systematic effects, i.e., macro(-economic) variables and random effects, lie in the heart of the analysis. The fourth and last research paper *Time matters: How default resolution times impact final loss rates* (see Chapter 3) connects the two parameters of the resolution process. As DRTs and LGDs are outcomes of the same random process – the resolution process, dependency structures among the two parameters are conceivable. A combined modeling approach is developed allowing for direct and indirect dependencies. By this means, parameter distortions due to the resolution bias are diminished and, thus, appropriate LGD predictions on an out of sample perspective are ensured which arise in pure (standard) LGD models.

Discussion and outlook

The topic of this thesis is of high relevance for financial institutions and regulators. As the financial stability is indispensable for a robust economic system (see Introduction), the regulatory framework regarding, for instance, capital requirements for credit risk, is subject to ongoing alterations (see Basel Committee on Banking Supervision, 2006, 2010, 2017). Thus, questions regarding the economic impact or the implementation of new capital standards arise. Some topics – e.g., the resolution of defaulted loan contracts – attained limited consideration in academic literature due to a lack of data availability. Data regarding loan contracts is of private nature and, thus, not publicly available. The research papers presented this thesis use access to the unique loss data base of Global Credit Data (GCD) which includes detailed loss information on transaction basis of defaulted loan contracts. Access to this data base made a profound analysis of the resolution process possible. However, the resolution of defaulted loan contracts is a wide research area. There are still open questions worth analyzing. First, the identification of economically and statistically significant (evident) macro(-economic) variables for workout LGDs attains ongoing attention in the recent literature (see, e.g., Lee and Poon, 2014; Krüger and Rösch, 2017). Applying a combined modeling approach as in Chapter 4 might simplify this task as variables can be integrated contemporary in the modeling framework. Second, alternative data – e.g., Moody's ultimate recovery data base – might be used to compare systematic effects of market-based and workout LGDs as deviations among the two concepts are conceivable. To the best of my knowledge, no study exists so far which examines the differences of market-based and workout LGDs. Third, the GCD data base offers information on transaction basis. A deeper analysis of individual recovery cash flows might shed additional light into the inherent structures of resolution processes.

Resolution of defaulted loan contracts is characterized by high levels of complexity as a series of random processes are liable for the resulting parameters – the DRT and the LGD. These processes not only depend on observable and unobservable variables but also on the regulatory framework and common business practices. To decently reflect such processes, statistical models should be as complex as necessary, but as simple as possible to ensure model stability. Some random processes – and I think the resolution of defaulted loan contracts belongs into this category – are shaped by noise and, thus, randomness, to a high degree. This hardens modeling efforts or as Kruschke (2015) stated *"[...] random variation is the researcher's bane. Noise is the nemesis."*. The reality seems to be more complex than any statistical model could reflect. However, we can do our very best to come close to it.

Bibliography

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

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

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

Altman, E. and E. Kalotay (2014). Ultimate recovery mixtures. *Journal of Banking and Finance 40*(1), 116–129.

Annabi, A., M. Breton, and P. Francois (2012). Resolution of financial distress under Chapter 11. *Journal of Economic Dynamics and Control 36*(12), 1867–1887.

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

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 (2005). Guidance on paragraph 468 of the framework document. *Bank of International Settlement*.

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

Basel Committee on Banking Supervision (2010). Basel III: A global regulatory framework for more resilient banks and banking systems. *Bank for International Settlements*.

Basel Committee on Banking Supervision (2017). Basel III: Finalising post-crisis reforms. *Bank of International Settlements*.

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

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

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

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

Betz, J., R. Kellner, and D. Rösch (2018). Systematic effects among loss given defaults and their implications on downturn estimation. *European Journal of Operational Research*, forthcoming.

Betz, J., S. Krüger, R. Kellner, and D. Rösch (2017). Macroeconomic effects and frailties in the resolution of non-performing loans. *Journal of Banking and Finance*, forthcoming.

Bijak, K. and L. Thomas (2015). Modelling LGD for unsecured retail loans using Bayesian methods. *Journal of the Operational Research Society 66*(2), 342–352.

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

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.

Bonelli, S., K. Frankfurth, and E. Eyeman (2014). Comparing major bankruptcy and insolvency regimes. *Fitch*.

Bonfim, D., D. Dias, and C. Richmond (2012). What happens after corporate default? Stylized facts on access to credit. *Journal of Banking and Finance 36*(7), 2007–2025.

Bordo, M., A. Redish, and H. Rockoff (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., I. Welch, and N. Zhu (2006). The costs of bankruptcy: Chapter 7 liquidation versus Chapter 11 reorganization. *Journal of Finance 61*(3), 1253–1303.

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

Brumma, N., K. Urlichs, and W. Schmidt (2014). Modeling downturn LGD in a Basel framework. *Working paper*.

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

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

Campello, M., J. Graham, and C. Harvey (2010). The real effects of financial constraints: Evidence from a financial crisis. *Journal of Financial Economics 97*(3), 470–487.

Caselli, S., S. Gatti, and F. Querci (2008). The sensitivity of the loss given default rate to systematic risk: New empirical evidence on bank loans. *Journal of Financial Services Research 34*(1), 1–34.

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

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

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

Davydenko, S. and J. Franks (2008). Do bankruptcy codes matter? A study of defaults in France, Germany, and the U.K. *Journal of Finance 63*(2), 565–608.

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

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

Dewaelheyns, N. and C. Van Hulle (2009). Filtering speed in a continental European reorganization procedure. *International Review of Law and Economics 29*(4), 375–387.

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

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

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

Düllmann, K. and M. Trapp (2004). Systematic risk in recovery rates – An empirical analysis of US corporate credit exposures. *Discussion paper* (Deutsche Bundesbank).

European Banking Authority (2016a). Draft regulatory technical standards on the materiality threshold for credit obligations past due under article 178 of regulation (EU) No 575/2013. *Final report*.

European Banking Authority (2016b). Risk assessment of the European banking system. *Final report*.

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*.

Financial Crisis Inquiry Commission (2011). The financial crisis inquiry report. *Final report*.

Franks, J. and W. Torous (1989). An empirical investigation of U.S. firms in reorganization. *Journal of Finance 44*(3), 747–769.

Gelman, A. and D. Rubin (1992). Inference from iterative simulation using multiple sequences. *Statistical Science 7*(4), 457–472.

Grunert, J. and M. Weber (2009). Recovery rates of commercial lending: Empirical evidence for German companies. *Journal of Banking and Finance 33*(3), 505–513.

Gürtler, M. and M. Hibbeln (2013). Improvements in loss given default forecasts for bank loans. *Journal of Banking and Finance 37*(7), 2354–2366.

Haugen, R. and L. Senbet (1978). The insignificance of bankruptcy costs to the theory of optimal capital structure. *Journal of Finance 33*(2), 383–393.

Haugen, R. and L. Senbet (1988). Bankruptcy and agency costs: Their significance to the theory of optimal capital structure. *Journal of Financial and Quantitative Analysis 23*(1), 27–38.

Heidelberger, P. and P. Welch (1981). A spectral method for confidence interval generation and run length control in simulations. *Communications of the ACM* (Special issue on *Simulation Modeling and Statistical Computing*) *24*(4), 233–245.

Heidelberger, P. and P. Welch (1983). Simulation run length control in the presence of an initial transient. *Operations Research 31*(6), 1109–1144.

Helwege, J. (1999). How long do junk bonds spend in default? *Journal of Finance 54*(1), 341–357.

Hilscher, J. and M. Wilson (2017). Credit ratings and credit risk: Is one measure enough? *Management Science 63*(10), 3414–3437.

Höcht, S., A. Kroneberg, and R. Zagst (2011). Explaining aggregated recovery rates. *Working paper*.

Höcht, S. and R. Zagst (2010). Loan recovery determinants – A Pan-European study. *Working paper*.

Hotchkiss, E., K. John, R. Mooradian, and K. Thorburn (2008). Bankruptcy and the resolution of financial distress. In *Handbook of Empirical Corporate Finance*, Volume 2, Chapter 14, pp. 235–287. Elsevier.

Hu, Y. and W. Perraudin (2006). The dependence of recovery rates and defaults. *Working paper*.

Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*, Volume 1. Wiley.

Jankowitsch, R., F. Nagler, and M. Subrahmanyam (2014). The determinants of recovery rates in the US corporate bond market. *Journal of Financial Economics 114*(1), 155–177.

Kalotay, E. and E. Altman (2017). Intertemporal forecasts of defaulted bond recoveries and portfolio losses. *Review of Finance 21*(1), 433–463.

Kass, R. and A. Raftery (1995). Bayes Factors. *Journal of the American Statistical Association 90*(430), 773–795.

Keijsersy, B., B. Diris, and E. Kole (2017). Cyclicality in losses on bank loans. *Working paper*.

Khieu, H., D. Mullineaux, and H. Yi (2012). The determinants of bank loan recovery rates. *Journal of Banking and Finance 36*(4), 923–933.

King, M. (2013). The Basel III net stable funding ratio and bank net interest margins. *Journal of Banking and Finance 37*(11), 4144–4156.

Koopman, S., A. Lucas, and B. Schwaab (2011). Modeling frailty-correlated defaults using macroeconomic covariates. *Journal of Econometrics 162*(2), 312–325.

Krüger, S. and D. Rösch (2017). Downturn LGD modeling using quantile regression. *Journal of Banking and Finance 79*(1), 42–56.

Kruschke, J. (2015). *Doing Bayesian Data Analysis – A Tutorial with R, JAGS, and Stan*, Volume 2. Elsevier.

Kuo, C. and C. Lee (2007). Integrating market and credit risk using a simplified frailty default correlation structure. *Journal of Fixed Income 17*(1), 48–58.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. Vishny (1998). Law and finance. *Journal of Political Economy 106*(6), 1113–1155.

Lee, Y. and S. Poon (2014). Forecasting and decomposition of portfolio credit risk using macroeconomic and frailty factors. *Journal of Economic Dynamics and Control 41*, 69–92.

Leow, M., C. Mues, and L. Thomas (2014). The economy and loss given default: Evidence from two UK retail lending data sets. *Journal of the Operational Research Society 65*(3), 363–375.

Loterman, G., I. Brown, D. Martens, C. Mues, and B. Baesens (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 and Finance 1*(3), 249–276.

Matuszyk, A., C. Mues, and L. Thomas (2010). Modelling LGD for unsecured personal loans: Decision tree approach. *Journal of the Operational Research Society 61*(3), 393–398.

McLachlan, G. and D. Peel (2000). *Finite Mixture Models*, Volume 1. Wiley.

Nazemi, A., F. Fatemi Pour, K. Heidenreich, and F. Fabozzi (2017). Fuzzy decision fusion approach for loss-given-default modeling. *European Journal of Operational Research 262*(2), 780–791.

Partington, G., P. Russel, M. Stevenson, and V. Torbey (2001). Predicting return outcomes to shareholders from companies entering Chapter 11 bankruptcy. *Managerial Finance 27*(4), 78–96.

Plummer, M. (2017). JAGS version 4.3.0 user manual. *User manual*.

Qi, M. and X. Yang (2009). Loss given default of high loan-to-value residential mortgages. *Journal of Banking and Finance 33*(5), 788–799.

Qi, M. and X. Zhao (2011). Comparison of modeling methods for loss given default. *Journal of Banking and Finance 35*(11), 2842–2855.

Rösch, D. and H. Scheule (2010). Downturn credit portfolio risk, regulatory capital and prudential incentives. *International Review of Finance 10*(2), 185–207.

Rösch, D. and H. Scheule (2014). Forecasting probabilities of default and loss rates given default in the presence of selection. *Journal of Operational Research Society 65*(3), 393–407.

Somers, M. and J. Whittaker (2007). Quantile regression for modelling distributions of profit and loss. *European Journal of Operational Research 183*(3), 1477–1487.

Stan Development Team (2016). Stan modeling language. *User manual*.

Therneau, T., P. Grambsch, and V. Pankratz (2003). Penalized survival models and frailty. *Journal of Computational and Graphical Statistics 12*(1), 156–175.

Tobback, E., D. Martens, T. Van Gestel, and B. Baesens (2014). Forecasting loss given default models: Impact of account characteristics and the macroeconomic state. *Journal of the Operational Research Society 65*(3), 376–392.

Wong, B., G. Partington, M. Stevenson, and V. Torbey (2007). Surviving Chapter 11 bankruptcies: Duration and payoff? *Abacus 43*(3), 363–387.

World Bank (2015a). Doing business 2015 – Going beyond efficiency. Particularly, *Getting credit* and *Resolving insolvency*.

World Bank (2015b). Doing business 2015 – Going beyond efficiency (Germany). Particularly, *Getting credit* and *Resolving insolvency*.

World Bank (2015c). Doing business 2015 – Going beyond efficiency (United Kingdom). Particularly, *Getting credit* and *Resolving insolvency*.

World Bank (2015d). Doing business 2015 – Going beyond efficiency (United States). Particularly, *Getting credit* and *Resolving insolvency*.

Yao, X., J. Crook, and G. Andreeva (2015). Support vector regression for loss given default modelling. *European Journal of Operational Research 240*(2), 528–538.

Yao, X., J. Crook, and G. Andreeva (2017). Enhancing two-stage modelling methodology for loss given default with support vector machines. *European Journal of Operational Research 263*(2), 679–689.