Modeling and Estimating the Loss Given Default of Leasing Contracts

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2016

vorgelegt
von

Dipl.-Wirt.-Math. Patrick Michel Miller

aus

Leverkusen

Referent:	Univ.-Prof. Dr. Thomas Hartmann-Wendels, Universität zu Köln
Korreferent:	Univ.-Prof. Dr. Dieter Hess, Universität zu Köln
Vorsitz:	Univ.-Prof. Dr. Heinrich R. Schradin, Universität zu Köln

Tag der Promotion: 11. Januar 2017

Acknowledgments

The completion of this thesis would not have been possible without being supported by many people, whom I would like to thank.

Most of all, I would like to express my gratitude to my advisor Professor Dr. Thomas Hartmann-Wendels for his support, motivation and incredibly helpful suggestions during my research. Moreover, I would also like to thank him for the opportunity to present my studies at several meetings and, in particular, international conferences. Another word of thanks is addressed to Professor Dr. Dieter Hess for serving as co-referee and to Professor Dr. Heinrich R. Schradin for chairing the dissertation committee.

My sincere gratitude also goes to Dr. Eugen Töws for the great collaboration. We had countless extensive and productive discussions on several topics that resulted in two co-authored projects. I really appreciate the joint research with him.

I would also like to thank my former colleagues at the University of Cologne. I would like to mention Patrick Wohl, Dr. Tobias Schlüter, Benjamin Döring, David Fritz, Elke Brand, Dr. Claudia Wendels and, in particular, Dr. Wolfgang Spörk for their support and their valuable suggestions and advices on various topics.

Further thanks go to all of my friends for their motivation. In particular, I would like to thank Benjamin Rombelsheim for several discussions and for his support.

Finally, I would like to express my deepest gratitude to my parents Horst and Bettina Miller for the continuous encouragement and tremendous support that made the completion of this thesis possible. I would also like to thank my girlfriend Monika Wieczorek for her patience, motivation and support. Moreover, my thanks go to the rest of my family and in particular I would like to say thank you to my grandfather Franz Josef Bock, who passed away way too early, for his unbounded confidence in every single life situation.

Cologne, January 2017 Patrick Michel Miller

This thesis consists of the following works:

Thomas Hartmann-Wendels, Patrick Miller, and Eugen Töws, 2014. Loss given default for leasing: Parametric and nonparametric estimations. Journal of Banking & Finance 40, 364–375.

Patrick Miller, 2015. Does the Economic Situation Affect the Loss Given Default of Leases?. Working paper.

Patrick Miller and Eugen Töws, 2016. Loss Given Default-Adjusted Workout Processes for Leases. Working paper.

Contents

List of Abbreviations

AIC	Akaike information criterion
ALGD	Asset-related loss given default
ARR	Asset-related recovery rate
BDL	Federal Association of German Leasing Companies
BIC	Bayesian information criterion
CRR	Capital Requirement Regulation
EAD	Exposure at default
FMM	Finite mixture model
ICT	Information and communication technology
IRBA	Internal Ratings Based Approach
kNN	k-nearest neighbor
LGD	Loss given default
MAE	Mean absolute error
MLGD	Miscellaneous loss given default
MRR	Miscellaneous recovery rate
MSE	Mean squared error
MURD	Moody's Ultimate Recovery Database
NAREC	Normalized area under the regression error characteristic curve
OLS	Ordinary least squares
PD	Probability of default

REC	Regression error characteristic
REC Area	Area under the regression error characteristic curve
RF	Random forest
RMSE	Root mean squared error
RR	Recovery rate
RT	Regression tree
TIC	Theil inequality coefficient

List of Figures

List of Tables

1 Introduction

Risk management is of crucial importance for financial institutions. Identifying and quantifying the risk related to financial instruments is essential to make economically and strategically reasonable decisions. In particular, the knowledge of potential losses of financial assets is substantial for properly allocating regulatory and economic capital.

One of the major concerns of financial institutions' risk management is credit risk. Together with the probability of default (PD) and the exposure at default (EAD), the credit risk of a financial asset is particularly determined by the loss given default (LGD) respectively its counterpart, the recovery rate. The LGD is defined as the percentage of the EAD the financial institution loses if a debtor defaults.

According to Article 107 (1) of the Capital Requirement Regulation (CRR), financial institutes shall apply either the Standardised Approach or the Internal Ratings Based Approach (IRBA) to calculate their regulatory capital requirements for credit risk. To implement the advanced IRBA requires financial institutes to develop internal models for estimating PD, EAD, and LGD. One of the main objectives of the IRBA, which was introduced by the banking regulation within Basel II, the predecessor of the CRR, is to achieve risk-adjusted capital requirements (see Basel Committee on Banking Supervision (2003)). However, accurate estimates of PD, EAD, and LGD are also beneficial for pricing financial instruments and may lead to competitive advantages in general, as Gürtler and Hibbeln (2013) mention.

For many years, research on credit risk was mainly focused on analyzing the

PD. As a result, up to now, various elaborated methods for estimating the PD have been established (for an overview see, e. g., Saunders and Allen (2002)). In contrast, despite the importance of the LGD not only from the regulatory perspective, but also from the economic perspective, its detailed analysis for various asset classes has just started with the announcement of Basel II. In the recent years, several studies have analyzed and modeled the LGD. However, as yet, there are doubts about which methods are suitable for estimating the LGD. In addition, the typical drivers of the LGD have not yet been unambiguously identified.

In view of the key importance of the LGD for financial institutions' credit risk management, this thesis contributes to the recent literature on LGD research by focusing on the modeling and estimation of the LGD. In particular, the main attention is paid to the LGD of leasing contracts. The findings obtained in this thesis are evaluated from a practical point of view and are discussed in the light of the results of previous studies.

Several of the recently published studies that have addressed the modeling and estimation of the LGD have focused on bonds (see, e. g., Frye (2005), Dwyer and Korablev (2009), and Jankowitsch et al. (2014)). In addition, numerous researchers have analyzed the LGD of loans (see, e. g., Caselli et al. (2008), Bastos (2010), and Zhang and Thomas (2012)). However, despite the particular importance of the leasing business for economies like Germany, only a few studies have investigated the LGD of leases (see, e. g., Laurent and Schmit (2005), De Laurentis and Riani (2005), and Hartmann-Wendels and Honal (2010)). According to the Federal Association of German Leasing Companies (BDL), about 50% of the externally financed investments and nearly 25% of the total investments are currently lease-financed in Germany.

Typically, loans and leases feature a higher seniority level than bonds. Moreover, what is even more important for the calculation of the LGD, unlike bonds, loans and leases are not tradeable in general. Therefore, it is generally not possible to apply the concept of market LGDs to loans or leases. For a financial asset, the market LGD is calculated as one minus the ratio of the observed market price of the asset soon after its default to the trading price at the time of default. For loans and leases it is rather necessary to rely on the concept of workout LGDs. Workout LGDs can basically be applied to all types of financial assets and are calculated as one minus the ratio of the discounted cash flows after default to the EAD.

In almost all empirical studies, the density function of workout LGDs is reported to be bimodal with peaks around 0 and 1 (see, e. g., Hartmann-Wendels and Honal (2010), Zhang and Thomas (2012), and Calabrese (2014)). This applies in particular for both loans and leasing contracts. High concentrations of realized LGDs around 0 and 1 imply that frequently the recovered amount of the EAD is either quite high or fairly small. This unusual shape of the LGD density is particularly challenging from the econometric perspective, because, as Qi and Zhao (2011) argue, it is at least questionable whether standard statistical methods such as the ordinary least squares (OLS) linear regression are suitable for estimating the LGD.

The bimodal nature of the LGD density represents one of the considerable commonalities of loans and leases. Basically, also the drivers of the LGD might be to some extent similar for loans and leasing contracts. Nevertheless, leases are also characterized by some distinctive features. For the modeling and estimation of leasing LGDs it is crucial to consider these specific characteristics of leases. Any leasing contract is obligatory collateralized by its leased asset. In particular, the fundamental characteristic of all leases is that the lessor retains the legal owner of the leased asset. Therefore, repossession of the leased asset is easier than foreclosure on the collateral for a secured loan, which is a key advantage of the leasing business according to Eisfeldt and Rampini (2009). In particular, different to the collateral realization for a secured loan, the lessor can retain any recovered value of the leased asset's disposal. In fact, Schmit and Stuyck (2002) observe lower LGDs for leases than for loans in general, which emphasizes that leasing companies potentially enjoy a competitive advantage.

Taking into account the specific characteristics of leases is crucial for economic reasons but also from the econometric perspective. With regard to defaulted loans, it is frequently assumed that the LGD is bounded within the interval $[0, 1]$ (see, e. g., Dermine and de Carvalho (2006), Bastos (2010), and Calabrese (2014)). This implies that independent from the course of the workout process, the lender cannot recover respectively lose more than the EAD. While the predefinition of the lower limit is justified for bank loans, it is not generally applicable to leases. To be precise, when disposing the leased asset, the lessor, as the legal owner of the leased asset, may in particular retain revenues which exceed the EAD. In fact, LGDs smaller than 0 are a frequently observed phenomenon in the leasing business (see, e. g., Schmit and Stuyck (2002), Laurent and Schmit (2005), and Hartmann-Wendels and Honal (2010)). In addition, with regard to the upper limit of the LGD, the assumption that the LGD does not exceed 1 is inappropriate for both loans and leases. According to Article 5 (1) of the CRR, workout costs are required to be included in the LGD calculation. However, if workout costs are considered, the lender's loss may exceed the EAD. Analyzing commercial real estate loans, Johnston Ross and Shibut (2015) observe several LGDs above 1 and thus highlight that capping the LGD at 1 might understate true losses. Regarding leasing contracts, De Laurentis and Riani (2005) also stress the impact of workout costs when determining the LGD. Studying Italian leasing contracts, the authors find workout costs amounting to more than 5% of the EAD on average.

1.1 Estimating the loss given default: an initial literature review

The recent literature that addresses the modeling and estimation of the LGD basically focuses on two different aspects of research. One major research stream covers the analysis of different approaches for estimating the LGD. The other main area of research concentrates on the examination of the drivers of the LGD.

As yet, numerous different methods for estimating the LGD have already been investigated in the literature. Remarkably, also the basic ideas of the studied methods differ occasionally.

A couple of studies aim on reproducing the LGD's density function in order to extrapolate accurate LGD estimations in this way. For this purpose, Calabrese and Zenga (2010) model the LGD on the unit interval by a mixed random variable and apply this concept to a dataset of defaulted Italian loans. Similar approaches were also pursued by Hlawatsch and Ostrowski (2011) and Altman and Kalotay (2014). Hlawatsch and Ostrowski (2011) suggest a mixture of two beta distributions to approximate the LGD distribution. The authors generate accurate LGD estimations when employing their model to synthesized loan portfolios. Altman and Kalotay (2014) present an approach based on the mixture of Gaussian distributions and likewise report successful LGD predictions using Moody's Ultimate Recovery Database (MURD).

Several other surveys investigate the suitability of parametric and nonparametric methods for estimating the LGD. It must be stressed that the obtained findings do not always conform to the results of the mentioned studies which focus on reproducing the LGD's density function. Using MURD, Qi and Zhao (2011) estimate the LGD by different parametric and nonparametric methods and analyze the results. They note that the nonparametric methods generally outperform the parametric methods. In particular, the authors find regression trees to be a suitable nonparametric method to estimate the LGD. The predictions generated by the regression trees are noticeably more accurate than those obtained by fractional response regression, the best performing parametric method. They argue that the good performance of the nonparametric methods is related to their ability to model nonlinear relationships between the LGD and continuous explanatory variables. Moreover, Qi and Zhao (2011) find no evidence for a correlation between a model's ability to reproduce the LGD distribution and its estimation accuracy. They conclude that reproducing the LGD distribution is only of secondary importance when modeling the LGD. Li et al. (2014) utilize the same dataset as Qi and Zhao (2011) to further analyze the performance of some parametric methods for estimating the LGD, including recently proposed gamma regressions and different transformation regressions such as inverse Gaussian regression. Their results confirm the findings of the earlier study as they find none of the used methods performing at least as good as the nonparametric methods investigated by Qi and Zhao (2011). In another large study, Loterman et al. (2012) compare several regression techniques for modeling and predicting the LGD using data of six different banks. The results of their benchmarking study correspond to the conclusions of Qi and Zhao (2011). They notice a clear trend that the nonlinear methods, and in particular support vector machines and neural networks, perform better than the linear methods. In this context Bastos (2010) conducts another noteworthy study estimating the LGD of Portuguese bank loans by regression trees and fractional response regression. While the latter was successfully used in some earlier studies (see, e. g., Dermine and de Carvalho (2006) and Chalupka and Kopecsni (2009)), he finds in line with the results obtained by Qi and Zhao (2011) fractional response regression to be outperformed by the regression trees.

In fact, until now most of the methods that have been used for estimating the LGD are so-called single-stage models. This means that the LGD is directly modeled using a set of explanatory variables. Recently, some studies have proposed to predict the LGD by two-stage models. The basic idea of most two-stage models is to split the observations ex ante according to a specific key feature. In particular, the applied splitting criterion depends on the characteristics of the used data. Leow and Mues (2012) introduce a two-stage approach to forecast the LGD of mortgage loans. In a first step, they estimate the probability of a defaulted mortgage account undergoing repossession. In order to finally obtain an estimated LGD, they subsequently calculate the loss in the event of repossession using a haircut value. The latter is defined as the ratio of the forced sale price and the market valuation of the repossessed property. Another two-stage model was successfully implemented by Gürtler and Hibbeln (2013). Analyzing defaulted private and commercial loans of a German bank, the authors find that recovered and written off loans feature different characteristics. Hence, they first estimate the probability that a loan will be recovered or written off. In a second step, they predict the LGD for recovered and written off loans separately to combine these predictions to a final LGD estimate using the probability-weighted average. A similar approach was proposed by Johnston Ross and Shibut (2015) investigating commercial real estate loans. The authors suggest to differentiate between loans with zero and non-zero losses.

Recently, the fairly new concept of ensemble learning has been applied in different areas of credit risk research. The concept of ensemble learning provides a complement to the development of single procedures for estimating the LGD, as the basic idea of this approach is to combine predictions of several individual models in order to generate more precise estimates in this way. Bastos (2013) focuses on analyzing different ensemble learning strategies for estimating the LGD using MURD. He finds that in particular an ensemble learning strategy based on regression trees exhibits a high predictive power in forecasting the LGD.

With regard to the methods that have already been successfully used to estimate the LGD it is important to emphasize that proper LGD predictions have also been generated by OLS linear regression, although this may not be the best suited method to forecast the LGD from the econometric point of view. Actually, some studies show that the OLS linear regression is able to generate more accurate LGD predictions than more advanced estimation techniques. Zhang and Thomas (2012) estimate the LGD using various approaches, including some straight-forward two-stage models, and compare the outcomes with the results produced by OLS linear regression. Using a dataset of defaulted personal loans, the authors find that the OLS linear regression achieves the best LGD estimates in general. In particular, they find that the predictions of the OLS linear regression are more accurate than those obtained by first identifying loans with an expected LGD of 0 or 1 and explicitly estimating the LGD value only if the value is expected to be within the interval $(0, 1)$. A similar result is obtained by Bellotti and Crook (2012) when investigating the LGD of UK credit cards. They find OLS linear regression outperforming several other methods, including various transformation regressions.

Despite the wide range of different concepts for estimating the LGD that has been analyzed, as yet, no single approach could be established, neither for loans and particularly not for leases. This can probably be ascribed to the fact that the findings of different studies are to some extent contradictory. In fact, up to now the linear regression is the most commonly used method for estimating the LGD.

This thesis specifically addresses leasing contracts and examines which methods are particularly suitable for estimating the LGD in this context. In particular, this means that a method's ability to forecasting the LGD is strictly discussed against the background of the specific nature of the leasing business. Already at this early stage it has to be stressed that some of the methods for predicting the LGD which have been presented in the literature are not applicable to leases. Based on the assumption that the LGD is bounded within the interval $[0, 1]$, a few studies focused particularly on methods generating estimates that are likewise restricted to a corresponding range of values. Beside some other approaches such as inverse Gaussian regression, one of these methods is in particular the frequently used fractional response regression. While restricting the LGD estimates to the interval $[0, 1]$, is generally already questionable from the regulatory perspective, for leasing contracts such a restriction is basically inappropriate, because the specific characteristics of leases would be neglected in this way. As highlighted previously, a typical feature of defaulted leasing contracts is that its LGD certainly exceeds both limits of the interval $[0, 1]$.

Within this thesis, it is also explicitly analyzed under which circumstances, methods generate proper LGD estimates. This investigation is of particular importance because it might provide an explanation why the findings of recent studies are to some extent contradictory. For instance, the performance of a method potentially depends on the scope and quality of the available information.

The second main area in LGD research, which covers the analysis of the drivers of the LGD, contains in particular a large number of studies that dealt with loans (see, e. g., Grunert and Weber (2009), Chalupka and Kopecsni (2009), and Khieu et al. (2012)). By contrast, for leases the drivers of the LGD were rarely analyzed, notable exceptions are Schmit and Stuyck (2002), Laurent and Schmit (2005), and De Laurentis and Riani (2005).

Of course, when analyzing the drivers of the LGD for leases it is essential taking into account the specific characteristics of the leasing business. Nevertheless, factors driving the LGD of loans are probably partly also drivers of the LGD for leases. This assumption is reasonable, because loans and leases feature several common characteristics and often serve the same market segments. Regarding this, it should in particular be noted that, apart from the obligatory collateralization of a leasing contract by the leased asset, the seniority level of loans and leases is similar. Consequently, especially given the low number of studies that have analyzed the drivers of the LGD explicitly for leases, it is useful to consider the findings of recent studies that have covered loans.

Previous studies that have dealt with loans focused in particular on the impact of contract characteristics and customer characteristics on the LGD. In this regard, several studies examined the relationship between the LGD and, e. g., the customer type or the type of the loan. Moreover, in this context it was also analyzed to what extent the LGD depends on factors such as the creditworthiness of the debtor or the length of the business relationship between the financial institute and the customer.

Basically, all studies that have analyzed the LGD of loans emphasize that the LGD tends to be lower if the loan is secured by collateral. Caselli et al. (2008) and Grunert and Weber (2009) observe this correlation investigating defaulted loans issued by Italian respectively German banks. Khieu et al. (2012) confirm this finding using MURD. Moreover, among the other analyzed determinants, there are also some factors that have frequently been identified as drivers of the LGD of loans. Several authors note, e. g., that debtors with a poor creditworthiness exhibit higher LGDs. This dependence is found by Grunert and Weber (2009) in their study on German loans and the analysis of Portuguese loans by Bastos (2010) reveals a similar result.

Nevertheless, for plenty of the analyzed factors the results of recent studies on loans are quite controversial. As a result, apart from some exceptions, as yet, there has been no general consensus concerning the key drivers of the LGD for loans. The determinant with the most ambiguous results regarding its influence on the LGD is probably the size of the loan. Bastos (2010) finds that large loans feature higher LGDs and a similar result is obtained by Hurt and Felsovalyi (1998) investigating defaulted loans in Latin America. In contrast, e. g. Khieu et al. (2012) observe no significant relationship between the LGD and the size of the loan. Moreover, for corporate loans, Acharya et al. (2007) actually argue from a theoretical point of view that large loans could also exhibit lower LGDs due to the high bargaining power of big corporates that typically take those large loans. Beside the size of the loan, the results in the literature are also heterogeneous, e. g. concerning the link between the LGD and factors such as the customer type or the intensity of the business relationship between the financial institute and the customer. Grunert and Weber (2009) note higher LGDs for large companies, Khieu et al. (2012), however, do not confirm a significant correlation between the LGD and the size of the borrowing company. With regard to the dependence of the LGD on the intensity of the business relationship between the financial institute and the customer, Grunert and Weber (2009) and Chalupka and Kopecsni (2009) obtain quite contradictory results, whereas Bastos (2010) finds no evidence for such a dependency. While Grunert and Weber (2009) observe that an intensive business relationship between the financial institute and the customer leads to lower LGDs, Chalupka and Kopecsni (2009) note for Czech loans that customers having a long business relationship with the financial institute feature higher LGDs.

The few studies that covered leases concentrated in particular on the relationship between the LGD and determinants which are associated with the collateralization of the leasing contract by the leased asset. The authors highlight unanimously that the LGD of leases depends on the type of the leased asset. Schmit and Stuyck (2002) obtain this result investigating defaulted leasing contracts from 12 European financial institutions in six different countries. De Laurentis and Riani (2005) and Hartmann-Wendels and Honal (2010) confirm this finding analyzing Italian respectively German leases. Moreover, Schmit and Stuyck (2002) note that the LGD of leases depends on the loan to value ratio and therefore on the age of the contract at default relative to its term to maturity.

In order to identify the key drivers of the LGD specifically for leases, within this thesis a comprehensive analysis of factors that potentially influence the LGD of leasing contracts is conducted. Bearing in mind, in particular, the specific characteristics of the leasing business, this analysis considers numerous idiosyncratic factors which include first of all determinants that are related to the leased asset but also, e. g., contract characteristics and customer characteristics. Moreover, some macroeconomic factors are also taken into account within the analysis. Especially with regard to leases, so far only very few studies investigated the influence of macroeconomic factors on the LGD and, in particular, the few existing investigations on this topic were commonly not carried out within the context of a general analysis of the key drivers of the LGD (see, e. g., Hartmann-Wendels and Honal (2010)).

Moreover, bearing in mind that the findings of recent studies concerning the key drivers of the LGD are at least controversial for loans, potential reasons for such divergent outcomes are also discussed within this thesis. In particular, referring to the results of the conducted analysis on leasing contracts, this discussion also evaluates whether it is actually possible to determine the key drivers of the LGD globally for the entire leasing business. For instance, the LGD and its drivers potentially depend on a company's organization of the workout process.

1.2 Contents and structure of the thesis

This thesis consists of three essays dealing with the modeling and estimation of the LGD for leasing contracts. The workout LGDs are standardized calculated by taking into account the regulatory requirements, which means in particular that workout costs are incorporated. Basically, all models for estimating the LGD that are introduced in this thesis meet crucial requirements of the CRR respectively Basel II. This may involve, e. g., that LGD estimates that were carried out at the execution of a contract are updated in case of default.

The first essay (Hartmann-Wendels, Miller, and Töws, 2014, Loss given default for leasing: Parametric and nonparametric estimation) focuses on the methodological aspects of estimating the LGD and extends the related literature by comparing different approaches for predicting the LGD of leasing contracts. Using a dataset with a total of 14,322 defaulted leasing contracts provided by three major German leasing companies for several parametric and nonparametric estimation methods the quality of the LGD predictions is analyzed in-sample and out-of-sample. In particular, with finite mixture models (FMMs), on the one hand, an approach aiming on reproducing the LGD's density function is implemented and, conversely, with the model tree M5', which represents an extension of a classical regression tree, in addition a method is used that does not require any assumptions concerning the distribution of the underlying data. The results of the applied models are benchmarked against the historical average and the outcomes generated by the so far frequently used OLS linear regression.

The results stress that it is crucial to execute in-sample and out-of-sample testing to reliably evaluate a model's suitability for estimating the LGD. The in-sample estimation accuracy of a model turns out to be only a weak indicator for its out-of-sample estimation accuracy, and, in particular, a model operating well in-sample does not necessarily perform well out-of-sample. Accounting for the bimodal or rather multimodal nature of the LGD density, the FMMs produce precise predictions in-sample. Out-of-sample, however, the estimates generated by the FMMs are quite poor, although the LGD density is still properly reproduced. In contrast, by mainly outperforming a classical regression tree, the model tree M5' achieves robust LGD estimates in-sample, but, in addition, generally provides the most accurate LGD predictions out-of-sample. Moreover, while OLS linear regression is particularly outperformed on datasets with a large number of observations, the results show some indications that OLS linear regression might be a suitable method for estimating the LGD given datasets containing only a small number of observations. For all implemented methods, the quality of the LGD predictions differs significantly between the analyzed companies, but in general the prediction accuracy improves by using additional information that are only available at default of the contract.

The findings of the first essay emphasize that the quality of LGD estimates essentially depends on the applied estimation method. Moreover, taking into account the improved estimation accuracy at default of the contract, the results additionally suggest that a method's ability to forecast the LGD is considerably determined by the available set of information. Therefore, the second essay (Miller, 2015, Does the Economic Situation Affect the Loss Given Default of Leases?) uses data from two different leasing companies to analyze the drivers of the LGD. Bearing in mind that the results of the first essay point out significant differences between the companies with regard to the accuracy of the LGD predictions that could be achieved, it is particularly investigated whether and to what extent the drivers of the LGD differ for the two lessors. In order to obtain an overview of the drivers of the LGD which is as complete as possible, the analysis contains various idiosyncratic factors and additionally several macroeconomic factors. Referring to the specific characteristics of the leasing business, the considered idiosyncratic factors include substantial information about the leased asset and also details, e. g., about the contract structure and the customer. Based on an observation period covering defaults between 2002 and 2009 for the macroeconomic factors, it is also evaluated whether a potential impact on the LGD is stable over the economic cycle. To ensure that the obtained findings are not biased by the use of a specific estimation methods, the outcomes of two different estimation approaches, namely OLS linear regression and a nonlinear regression spline model, are considered for the analyzes. Moreover, in-sample and out-of-time testing is performed to validate the results.

Showing some remarkable differences between the lessors studied, the results point out that identifying the relevant drivers of the LGD individually for each leasing company is substantial in order to develop an appropriate model for estimating the LGD. In particular, the differences noted among the lessors refer to both the set of idiosyncratic factors influencing the LGD and the determined relationship between the LGD and macroeconomic factors. With regard to the idiosyncratic factors the outcomes differ in detail between the leasing companies investigated. Nonetheless, in summary the results support that the LGD of leases generally depends in particular on determinants that are related to the leased asset. Moreover, there are also indications that contract characteristics significantly influence the LGD of leases, whereas, e. g., details about the customer have only a marginal impact. Referring to the relationship between the LGD and macroeconomic factors, the findings vary considerably depending on whether the LGD estimates are carried out at the contract's execution or its default. For both leasing companies, the outcomes regarding the macroeconomic factors expose that the economic situation at the point in time of contract's execution drives the LGD. In contrast, a relationship between the LGD and the economic situation at the point in time of contract's default is revealed only for one of the lessors studied.

The findings of the first two essays show that the quality of LGD predictions depends on the used estimation method as well as on the available set of information. Furthermore, the results of the second essay attest that the LGD of leases generally depends on determinants that are related to the leased asset. However, the outcomes additionally emphasize significant differences between lessors, in particular with regard to the factors driving the LGD. Therefore, in order to obtain reliable LGD predictions, it is indispensable to calibrate a method for estimating the LGD for each leasing company individually, taking into account the company's specific characteristics. Moreover, with the objective to forecast the LGD as accurately as possible, developing advanced approaches for estimating the LGD which explicitly address the specific characteristics of the respective company seems to be reasonable. In particular, when designing advanced models for estimating the LGD of leases, it appears to be of crucial importance that these models consider the peculiarities of the leasing business.

Consequently, the third essay (Miller and Töws, 2016, Loss Given Default-Adjusted Workout Processes for Leases) contributes to the related literature on LGD research by using a dataset of a German lessor to develop an advanced approach for estimating the LGD of leases that explicitly considers the specific characteristics of the leasing business. Based on the economic consideration that the revenues received during the workout process of defaulted leasing contracts come from two different payment sources, the LGD is initially separated into two distinct parts. On the one hand, the asset-related LGD (ALGD) includes all asset-related payments, such as the asset's liquidation proceeds and incurred liquidation costs. In addition, the so-called miscellaneous LGD (MLGD) summarizes all remaining revenues, such as customer payments and indirect workout costs. Based on this separation of the LGD, subsequently, a multi-step model for estimating the LGD of leasing contracts is designed. In a first step, the respective parts of the LGD are estimated. Then, in a second step, a classification procedure is applied to predict whether a contract's LGD is expected to be below or above its ALGD, because the evaluation of the data reveals that this feature is important in order to distinguish the contracts. The implemented classification model in particular includes the previously calculated estimates of ALGD and MLGD. Following the classification, in a third step, two LGD estimates are generated for every contract, each under the assumption that the contract's LGD is below or above its ALGD. The final LGD estimation for each contract is obtained as a linear combination of these two estimated LGDs weighted with the contract's classification probability.

To evaluate the performance of the introduced multi-step estimation model, in-sample, out-of-sample and out-of-time testing is performed. The results prove that LGD estimates for leases clearly benefit from developing advanced estimation models that explicitly consider the peculiarities of the leasing business. Compared to the benchmarking results of established estimation approaches, the predictions generated by the proposed multi-step estimation model are significantly more accurate. Moreover, the developed multi-step model provides valuable interim results that can be used as a decision support for actions to be taken during the workout process. It turns out that the ALGD frequently exceeds the LGD which implies that the collection of miscellaneous payments generates losses due to incurred workout costs. Consequently, in case a contract's LGD is expected to be below its ALGD, the workout process should be restricted to the disposal of the leased asset in order to improve the resulting LGD of the contract.

2 Loss given default for leasing: Parametric and nonparametric estimations

2.1 Introduction

The loss given default (LGD) and its counterpart, the recovery rate, which equals one minus the LGD, are key variables in determining the credit risk of a financial asset. Despite their importance, only a few studies focus on the theoretical and empirical issues related to the estimation of recovery rates.

Accurate estimates of potential losses are essential to efficiently allocate regulatory and economic capital and to price the credit risk of financial instruments. Proper management of recovery risk is even more important for lessors than for banks because leases have a comparative advantage over bank loans with respect to the lessor's ability to benefit from higher recovery rates in the event of default. In their empirical cross-country analysis, Schmit and Stuyck (2002) note that the average recovery rate for defaulted automotive and real estate leasing contracts is slightly higher than the recovery rates for senior secured loans in most countries and much higher than the recovery rates for bonds. Moreover, the recovery time for defaulted lease contracts is shorter than that for bank loans. Because the lessor retains legal title to the leased asset, repossession of a leased asset is easier than foreclosure on the collateral for a secured loan. Moreover, the lessor can retain any recovered value in excess of the exposure at default. Repossessing used assets and maximizing their return through disposal in secondary markets are aspects of normal leasing business and are not restricted to defaulted contracts. Therefore, lessors have a good understanding of the secondary markets and of the assets themselves. Because the lessor's claims are effectively protected by legal ownership, the high recoverability of the leased asset may compensate for the poor creditworthiness of a lessee. Lasfer and Levis (1998) find empirical evidence for the hypothesis that lower-rated and cash-constrained firms have a greater propensity to become lessees. To leverage their potential lower credit risk, lessors must be able to accurately estimate the recovery rates of defaulted contracts.

This paper compares the in-sample and out-of-sample accuracies of parametric and nonparametric methods for estimating the LGD of defaulted leasing contracts. Employing a large data set of 14,322 defaulted leasing contracts from three major German lessors, we find in-sample accuracy to be a poor predictor of out-of-sample accuracy. Methods such as the hybrid finite mixture models (FMMs), which attempt to reproduce the LGD distribution, perform well for in-sample estimation but yield poor results out-of-sample. Nonparametric models, by contrast, are robust in the sense that they deliver fairly accurate estimations in-sample, and they perform best out-of-sample. This result is important because out-of-sample estimation has rarely been performed in other studies – with the notable exceptions of Han and Jang (2013) and Qi and Zhao (2011) – although out-of-sample accuracy is critical for proper risk management and is required for regulatory purposes.

Analyzing estimation accuracy separately for each lessor, our results suggest that the number of observations within a data set has an impact on the relative performance of the estimation methods. Whereas sophisticated nonparametric estimation techniques yield, by far, the best results for large data sets, simple OLS regression performs fairly well for smaller data sets.

Finally, we find that estimation accuracy critically depends on the available set of information. We estimate the LGD at two different points in time, at the execution of the contract and at the point of contractual default. This procedure is of particular importance for leasing contracts because the loan-to-asset value changes during the course of a leasing contract. Furthermore, the Basel II accord requires financial institutions using the advanced internal ratings-based approach (IRBA) to update their LGD estimates for defaulted exposure. To the best of our knowledge, an analysis of this type of update has been neglected in the literature thus far.

The remainder of our study is organized as follows. We review the related literature in Section 2.2. Section 2.3 provides an overview of the data set, defines the LGD measurement, and presents some descriptive statistics. In Section 2.4, we introduce the methods used in this study. Section 2.5 reports the empirical results, and Section 2.6 presents the conclusions of the study.

2.2 Literature review

There are two major challenges in estimating recovery rates for leases with respect to defaulted bank loans or bonds. First, estimates of LGD on loans or bonds take for granted that the recovery rate is bounded within the interval $[0, 1]$, which assumes that the bank cannot recover more than the outstanding amount (even under the most favorable circumstances) and that the lender cannot lose more than the outstanding amount (even under the least favorable circumstances). Although the assumption of an upper boundary is justified for bank loans, it does not apply to leasing contracts. As the legal owner of the leased asset, the lessor may retain any value recovered by redeploying the leased asset, even if the recoveries exceed the outstanding claim. In fact, there is some empirical evidence that recovery rates greater than 100% are by no means rare. For example, Schmit and Stuyck (2002) report that up to 59% of all defaulted contracts in their sample have a recovery rate that exceeds 100%. Using a different data set, Laurent and Schmit (2005) find that recovery rates are greater than 100% in 45% of all defaulted contracts. The lower boundary of the recovery rate rests on the implicit assumption of a costless workout procedure. In fact, most empirical studies neglect workout costs (presumably) because of data limitations. Only Grippa et al. (2005) account for workout costs in their study of Italian bank loans and find that workout costs average 2.3% of total operating expenses. The Basel II accord, however, requires that workout costs are included in the LGD calculation. Thus, when workout costs are incorporated, there is no reason to assume that workout recovery rates must be non-negative. The second challenge in estimating recovery rates is the bimodal nature of the density function, with high densities near 0 and 1. This property of workout recovery rates is well documented in almost all empirical studies, whether of bank loans or leasing contracts (e. g., Laurent and Schmit (2005)).

Because of the specific nature of the recovery rate density function, standard econometric techniques, such as OLS regression, do not yield unbiased estimates. Renault and Scaillet (2004) apply a beta kernel estimator technique to estimate the recovery rate density of defaulted bonds, but they find that it is difficult to model its bimodality. Calabrese and Zenga (2010) extend this approach by considering the recovery rate as a mixed random variable obtained as a mixture of a Bernoulli random variable and a continuous random variable on the unit interval and then apply this new approach to a large data set of defaulted Italian loans. Qi and Zhao (2011) compare fractional response regression to other parametric and nonparametric modeling methods. They conclude that nonparametric methods – such as regression trees (RTs) and neural networks – perform better than parametric methods when overfitting is properly controlled for. A similar result is obtained by Bastos (2010), who compares the estimation accuracy of fractional response to RTs and neural networks.

Despite the growing interest in the modeling of recovery rates, little empirical evidence is available on this topic. Several studies (e. g., Altman and Ramayanam

(2007), Friedman and Sandow (2005), and Frye (2005)) rely on the concept of market recoveries, which are calculated as the ratio of the price for which a defaulted asset is traded some time after default to the price of that asset at the time of default. Market recoveries are only available for bonds and loans issued by large firms. Workout recoveries are used by Khieu et al. (2012), Dermine and Neto de Carvalho (2005), and Friedman and Sandow (2005). However, Khieu et al. (2012) find evidence that the post-default price of a loan is not a rational estimate of actual recovery realization, i. e., it is biased and/or inefficient. According to Frye (2005), many analysts prefer the discounted value of all cash flows as a more reliable measurement of defaulted assets because: (1) cash flows ultimately become known with certainty, whereas the market price is derived from an uncertain forecast of future cash flows; (2) the market for defaulted assets might be illiquid; (3) the market price might be depressed; and (4) the asset holder might not account for the asset on a market-value basis.

Schmit et al. (2003) analyze a data set consisting of 40,000 leasing contracts, of which 140 are defaulted. Using bootstrap techniques, they conclude that the credit risk of a leasing portfolio is rather low because of its high recovery rates. Similar studies are conducted by Laurent and Schmit (2005) and Schmit (2004). Schmit and Stuyck (2002) find considerable variation in the recovery rates of 37,000 defaulted leasing contracts of 12 leasing companies in six countries. Average recovery rates depend on the type of the leased asset, country, and contract age. De Laurentis and Riani (2005) find empirical evidence that leasing recovery rates are inversely correlated with the level of exposure at default. However, recovery rates increase with the original asset value, contract age, and existence of additional bank guarantees. Applying OLS regressions to forecast LGDs in that study leads to rather poor results: the unit interval is divided into three equal intervals, and only 31–67% of all contracts are correctly assigned in-sample. With a finer partition of five intervals, the portion of correctly assigned contracts decreases even further. These results clearly indicate that more appropriate estimation techniques are needed to accurately estimate recovery rates.

Our study differs from the LGD literature in several crucial aspects. First, we calculate workout LGDs and consider workout costs. Second, we perform out-of-sample testing at contract execution and default, which meets the Basel II requirements for LGD validation. Third, by separately analyzing the data sets of three lessors, we gain insight into the robustness of the estimation techniques.

2.3 Data set

This study uses data sets provided by three German leasing companies, which shall be referred to herein as companies A, B, and C. All three companies use a default definition consistent with the Basel II framework. According to Table 2.1, the data set from lessor A contains 9,735 leasing contracts with 5,811 different customers and default dates between 2002 and 2010. The data set from lessor B contains 2,995 leasing contracts with 2,344 different lessees who defaulted between 1994 and 2009, with the majority of defaults occurring between 2001 and 2008. The data set for leasing company C consists of 1,592 leasing contracts with 964 different lessees who defaulted between 2002 and 2009.

Company	# Contracts	# Lessees
A	9,735	5,811
B	2,995	2,344
C	1,592	964

Table 2.1: Numbers of contracts and lessees in the data sets of companies A–C in descending order of the number of contracts.

For the defaulted contracts, we calculate the LGD as one minus the recovery rate. The recovery rate is the ratio of the present value of cash inflows after default to the exposure at default (EAD). For leasing contracts, the cash flows consist of the revenues obtained by redeploying the leased asset and other collateral combined with other returns and less workout expenses. The cash flows are discounted to the time of default using the term-related refinancing interest rate.[1] The EAD is the sum of the present value of the outstanding minimum lease payments, compounded default lease payments, and the present residual value. All values refer to the time of default. A contract is classified as defaulted when at least one of the triggering events set out in the Basel II framework has occurred.

Before the data was collected, all three companies agreed to use identical definitions for all the elements that are entered into the LGD calculation, and for all details of the leasing contract, lessee, and leased asset. Thus, for every contract, we have detailed information about the type and date of payments that the lessor received after the default event. Moreover, we incorporate expenses arising during the workout into the LGD calculation, to meet Basel II requirements. Workout costs are rarely considered in empirical studies.

The workouts have been completed for all the observed contracts. Gürtler and Hibbeln (2013) recommend restricting the observation period of recovery cash flows to avoid the under-representation of long workout processes, which might result in an underestimation of LGDs. Because we do not see a similar problem in our data, we do not truncate our observations based on that effect.

All three companies also provide a great deal of information about factors that might influence the LGD, which we divide into four categories:

1. contract information;
2. customer information;
3. object information; and
4. additional information at default.

[1] Only a few studies (such as Gibilaro and Mattarocci (2007)) address risk-adjusted discounting. We use the term related refinancing interest rate to discount cash flows at the time of default, independently of the time span of the workout and the risk of each type of cash flow.

Contract information is elementary information about the contract, such as its type, e. g., whether it was a full payment lease, partial amortization, or hire-purchase; its duration; its calculated residual value or prepayment rents; and information about collateralization and/or purchase options. Customer information mainly identifies retail and non-retail customers. The category object information consists of basic information about the object of the lease, including its type, initial value, and supplementary information, such as the asset depreciation range. Whereas all the information in the first three groups is available from the moment the contract is concluded, the last category consists of information that only becomes available after the contract has defaulted, such as the exposure at default and the contract age at default.

Descriptive statistics

The LGD is clearly not restricted to the interval $[0,1]$. As presented in Table 2.2 and Figure 2.1, negative LGDs are not only theoretically possible but also occur frequently in the leasing business. Hartmann-Wendels and Honal (2010) argue that such cases mainly occur if a defaulted contract with a rather low EAD yields a high recovery from the sale of the asset. Because we incorporate the workout expenses, LGDs greater than one are also feasible. Thus, we do not bound LGDs within the $[0,1]$ interval, as is common for bank loans and as is done by Bastos (2010), by Calabrese and Zenga (2010), and by Loterman et al. (2012).

Company	Mean	Std	P5	P25	Median	P75	P95
A	0.52	0.40	−0.11	0.19	0.52	0.88	1.05
B	0.35	0.42	−0.18	0.00	0.25	0.72	1.01
C	0.39	0.42	−0.23	0.03	0.32	0.77	1.03

Table 2.2: Loss given default (LGD) density information for companies A–C. Std is the standard deviation and P5–P95 are the respective percentiles.

An LGD of 45%, as specified in the standard credit risk approach, is considerably higher than the median LGDs observed for companies B and C. In general, we emphasize that the shape of the LGD distribution varies significantly among these three companies. As presented in Figure 2.1, only the LGD distribution of company C exhibits the frequently mentioned bimodal shape, whereas those of companies A and B feature three maxima. These differences continue to prevail when we account for differences in the leasing portfolio. Thus, we trace these variations back to differences in workout policies. Because the requirements for the pooling of LGD data, set out in section 456 of the Basel II accord, are clearly violated, we construct individual estimation models to account for institution-specific characteristics and differences in LGD profiles among the companies.

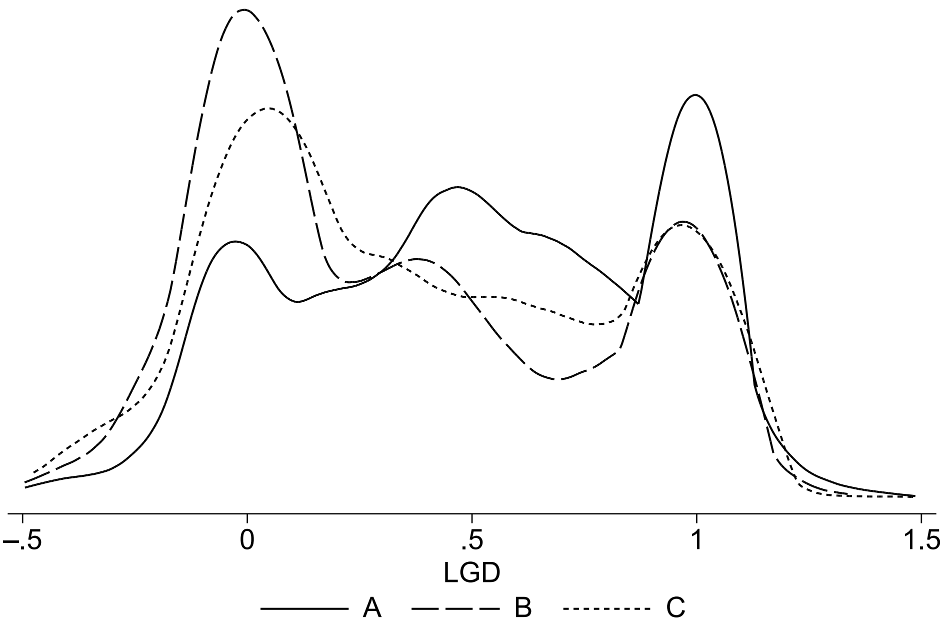


Figure 2.1: Density of the realized loss given default (LGD) by company. The realized LGD concentrates on the interval $[-0.5, 1.5]$. The figures describe a loss severity of -50% on the left end, which indicates that 150% of the exposure at default (EAD) was recovered. On the right end, the loss severity is 150%, indicating a loss of 150% of the EAD. Consequently, a realized LGD of 0 or 1 indicates the following: in case of 0, full coverage of the EAD (included workout costs); or, in case of 1, total loss of the EAD.

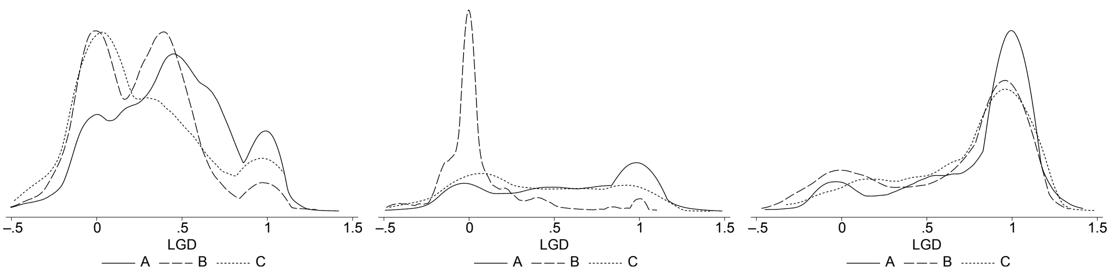
Previous studies on the LGD of defaulted leasing contracts consistently show that the LGD distribution depends largely on the underlying asset type. We categorize the contracts according to the underlying asset using five classes: vehicles, machinery, information and communications technology (ICT), equipment, and

Asset type	Company	# Contracts	Mean	Std	Median
	A	4,578	0.44	0.35	0.45
Vehicles	B	1,111	0.26	0.31	0.27
	C	599	0.28	0.37	0.21
	A	4,140	0.55	0.43	0.61
Machinery	B	779	0.06	0.27	0.00
	C	646	0.39	0.42	0.32
	A	606	0.77	0.38	0.96
ICT	B	1,062	0.64	0.43	0.84
	C	201	0.72	0.38	0.87
	A	353	0.61	0.44	0.74
Equipment	B	26	0.26	0.44	0.09
	C	26	0.38	0.41	0.15
	A	58	0.56	0.43	0.54
Other	B	17	0.39	0.44	0.26
	C	120	0.46	0.43	0.45

Table 2.3: Loss given default (LGD) density information by asset type for companies A–C. For each asset type, # Contracts is the number of contracts containing this type of asset, Mean is its mean, Std is its standard deviation, and Median is its median. ICT is information and communications technology. The displayed asset types vary in the numbers of their contracts and even further in the characteristics of their realized LGD.

other. Table 2.3 summarizes the key statistical figures of the distributions for each company. We can unambiguously rank the three companies with respect to their mean LGD. Company B achieves the lowest average LGD for all asset types, company C is second best, and company A bears the highest losses. Contracts in ICT have the highest average LGD. Examining the median of ICT, we find that companies A, B, and C retrieve only 4%, 16%, and 13% of the EAD, respectively, in half of the cases. The key statistical figures for equipment and other assets are seemingly less meaningful because of the small sample sizes for these classes, but the trends are consistent across all three companies.

Figure 2.2 presents the LGD distributions for vehicles, machinery, and ICT for each company. The shape of the LGD distributions differs tremendously with respect to the different asset types. Whereas for ICT, the LGD density in Fig-



(a) Vehicles. (b) Machinery. (c) ICT.

Figure 2.2: Densities of realized loss given default (LGD) by company for the three major asset types: vehicles, machinery, and information and communications technology (ICT). Depending on the asset type, the realized LGD density appears in completely different shapes. For machinery (Figure b), even the difference between companies is enormous.

ure 2.2c is right-skewed toward high LGDs with only weak bimodality throughout all of the companies, the density of machinery runs partly the opposite direction. For machinery, in Figure 2.2b, we see a higher concentration around 0, but for company A, larger LGDs again outweigh this effect. The LGD for contracts with vehicles varies greatly from company to company. We observe a strong multimodality for all of the companies with an additional peak at approximately 0.5, and most of the density lies in the lower LGD range.

2.4 Methods

This section describes the various approaches that we use to estimate the LGD and its density. According to section 448 of the Basel II regulations, institutes are required to base their estimations on a history of defaults and to consider all relevant data, information, and methods. Furthermore, a bank using the advanced IRBA must be able to break down its experience with respect to the probability of default (PD), LGD, and the IRBA conversion factor. This breakdown is to be based on the factors that are identified as drivers of the respective risk parameters.

The basic method used to identify these drivers is to partition the data according to a certain attribute (e. g., the type of object). Differences in the means of the partitions are then captured by setting the inducing factor as the driver. The average value is then the (naive) estimator of the LGD for the corresponding subclass. As Gupton and Stein (2005) note, this traditional look-up table approach is static and backward-looking, even if considerable variation is observed in the LGD distributions for different types of objects. An alternative method of verifying the impact of potential factors and developing an estimation model is to conduct a regression analysis. Linear regressions always estimate the (conditional) expectation of the target variable, but this average is not a reasonable parameter under mixed distributions, so it is not an adequate approach from a statistical perspective. However, regression analyses for LGD estimation are successfully implemented by Bellotti and Crook (2012) and by Zhang and Thomas (2012).

Table 2.2 reports the median LGDs as 52%, 25%, and 32% for companies A, B, and C, respectively. Considering the LGD distribution in Figure 2.1, its heterogeneity suggests that the overall portfolio is composed of several subclasses, which are less heterogeneous in terms of the LGD. This implies that each subclass has its own characteristic LGD distribution. We use FMMs to reveal these unknown classes (cluster analysis), to fit a reasonable model to the data and to classify the observations into these classes. Furthermore, we apply two different regression/model tree algorithms to the data. These tree-based models also have the basic function of dividing the portfolio into homogeneous partitions; by contrast to the FMMs, however, the number of subclasses is endogenously determined rather than exogenously specified.

At the end of this section, we present an overview of how to select the explanatory variables for tree-based methods. We also describe our methodology for out-of-sample testing.

2.4.1 Finite mixture models and classification

Modeling the probability density of realized LGDs as a mixed distribution allows us to use different potential LGD drivers for different clusters and to capture differences in the effects of these drivers on the LGD in various subclasses. We adapt an approach originally proposed by Elbracht (2011). FMMs are described by Frühwirth-Schnatter (2006).

The approach consists of three steps: (1) cluster the total data set into finite classes by finite mixture distributions using all available information; (2) classify the data set into the resulting classes using only the information available at the execution or default of the contract by the k-nearest neighbors (kNN) or the classification tree algorithm J4.8; and (3) perform OLS regressions for each class.

Step (1) can be adjusted between the two extremes of nonparametric and parametric modeling, thus providing a flexible method of data adaptation. We use normal distributions to construct the mixing distributions. We estimate unknown model parameters using the expectation maximization algorithm, which also provides a probabilistic classification of the observations. The accuracy of classification step (2) can be measured for in-sample testing. However, in out-of-sample testing, the goal is to classify observations that do not belong to any class initially – because these objects are not part of the training sample used to form classes – into exactly one of the given classes.

We compare two different approaches to classifying contracts into previously established classes. The nonparametric kNN approach assigns an observation to the class with the majority of its k nearest neighbors, whereas the distance between observations is determined as the Euclidean distance. This approach is described by Hastie et al. (2009). We also apply the tree algorithm J4.8 for classification.

The J4.8 algorithm generates pruned C4.5 revision 8 decision trees, as illustrated by Witten et al. (2011) and originally implemented by Quinlan (1993). The decision tree is constructed by dividing the sample according to certain threshold values. The optimal split in terms of maximized gain ratio is performed until additional splits yield no further improvement, or a minimum of instances per subset is reached. Every partition results in a node. To prevent overfitting, we prune back the fully developed tree to a certain level. According to Quinlan (1993), these deleted nodes shall not contribute to the classification accuracy of unseen cases.

2.4.2 Regression and model trees

RTs are classified as nonparametric and nonlinear methods. Similar to other regression methods, they can be applied to analyze the underlying data set and to predict the (numeric) dependent variable. An essential difference between RTs and parametric methods, such as linear or logistic regressions, is that ex ante no assumption is made concerning the distribution of the underlying data, and no functional relationship is specified.

These characteristics are particularly beneficial in case of LGD estimation because it is typically not possible to describe the distribution of the LGD suitably with a single distribution, such as the normal distribution. In addition, the distribution of the LGD varies significantly according to the underlying data. Thus, as described in Section 2.3, the LGD distributions of the three companies studied here are all multimodal, although there are appreciable differences between companies, such as the number of maxima. In particular, more types of distributions are observed for bank loans (for an overview, see Dermine and Neto de Carvalho (2005)).

The basic idea of regression and model trees is to partition the entire data set into homogeneous subsets by a sequence of splits, which creates a tree consisting of logical if-then conditions. Starting with the root node of the tree that contains all instances of the underlying data, each leaf covers only a fraction of the data.

In an RT, the prediction of the dependent variable is given by a constant for all instances belonging to a leaf, typically defined as the average value of these instances. Model trees are an extension of RTs in the sense that the target variable of instances belonging to a leaf is estimated by a linear regression model. Therefore, model trees are hybrid estimation methods combining RTs and linear regression. Model trees are clearly applicable for LGD estimation because RTs are successfully used in previous studies such as Bastos (2010) and Qi and Zhao (2011). Linear regression models are also applied to analyze and predict LGDs, and these models may deliver comparable or better results than those of more complex models, as shown by Bellotti and Crook (2012) and Zhang and Thomas (2012).

For our LGD estimation, we apply the M5' model tree algorithm and the corresponding RT algorithm that is introduced by Wang and Witten (1997) and described by Witten et al. (2011). This algorithm is a reconstruction of Quinlan's M5 algorithm that was published in 1992. In the case of the M5' algorithm, the underlying data set is divided step by step, each time using the binary split based on the explanatory variables with the greatest expected reduction in the standard deviation. The constructed tree is subsequently pruned back to obtain an appropriately sized tree to control overfitting, which can influence out-of-sample performance negatively.

The resulting tree essentially depends on the explanatory variables used, particularly with respect to the M5' model tree algorithm; selecting appropriate variables is a complex issue because of ex ante relevance and effectiveness not always being known. In the first step, we consider the potential application of a large number of parameters. However, it might be preferable to include only a fraction of the available variables, which we account for in the second step.

There are various algorithmic approaches for variable selection; two frequently applicable greedy algorithms are forward selection and backward elimination. Bellotti and Crook (2012) use forward selection for their LGD estimations of retail credit cards with OLS regression. However, forward selection has a significant disadvantage neglecting variable interactions.

Instead of forward selection, we employ backward elimination, initiating all available variables and step by step eliminating the variables without which the best value in terms of the respective fit criterion is achieved. This procedure continues until a stop condition is reached, or all the variables are eliminated.

A typical fit criterion for regression models is the F-score. However, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) are used for forecasting, both of which are based on the log-likelihood function.

Analogous to the approximation of the AIC used by Bellotti and Crook (2012), the BIC can be approximated by

$$\mathrm{BIC} = n \cdot ln(\mathrm{MSE}) + p \cdot ln(n), \tag{2.1}$$

where n denotes the number of observations, p is the number of input variables, and MSE is the mean squared error of the observations.

We use the BIC, which penalizes the complexity of the model more than the AIC. This complexity is measured by the number of input variables. In addition to the number of explanatory variables, regression and model trees offer another complexity feature: the number of leaves in the computed tree. This aspect is among those included by Gray and Fan (2008) when designing the TARGET RT algorithm. The more leaves that are present in the computed tree, the greater the risk is that a contract will be misclassified, which negatively influences the estimation.

We find that the number of leaves is determined not only by the pruning procedure but also by the input variables. Thus, we modify the BIC and penalize the size of the computed tree

$$\text{BIC}^* = n \cdot ln(\text{MSE}) + p \cdot ln(n) + |T| \cdot ln(n), \tag{2.2}$$

where $|T|$ denotes the number of leaves of the computed tree.

For BIC*, lower values are preferred. As with our data, MSE $\in (0,1)$, $n \gg p$ and $n \gg |T|$ holds; thus, we have BIC$^* < 0$. We set the stop condition for our backward elimination such that a variable in the i-th iteration can only be eliminated if the BIC* value increased by an absolute value of at least one, which implies that the following constraint must be fulfilled

$$\text{BIC}^*_{i-1} - \text{BIC}^*_i \geq 1. \tag{2.3}$$

2.4.3 Out-of-sample testing

We calibrate our models on randomly divided training sets of 75% and validate their performance on the remaining 25% of the total data set. Division and calibration are repeated 25 times. The final results are averaged. Our out-of-sample validation combines the advantages of k-fold cross validation and the approach of splitting the data set into training and test sets, and is particularly suitable for large data sets.

Bastos (2010) and Qi and Zhao (2011) employ k-fold cross validation – using $k = 10$ – to evaluate the out-of-sample performance of their models. This method relies on partitioning the data set randomly into k equal-sized subsets. While the model is calibrated on $k-1$ subsets, the models predictive performance is validated on the remaining subset. This procedure is performed k times, with each of the k subsets used exactly once for validation. Therefore each observation contained in the total data set is used exactly once for validation. By contrast, we draw the 25 divisions in training and test data randomly. With a small k in the k-fold cross validation there are fewer performance estimates, but the size of the subsets, and therefore the amount of the total data set which is used for each validation, is larger. As k increases, the number of performance estimates increases, however, the size of the validation subset decreases rapidly. Given larger data sets, the data can be split into some training and test sets. Here, the validation is restricted to the unseen cases of the test set. Gürtler and Hibbeln (2013) randomly shuffle and divide their data as 70% training and 30% validation. Consequently, our out-of-sample validation combines the advantages of these two approaches. In particular we make use of large test sets and still generate multiple estimations.

2.5 Results

We present both in-sample and out-of-sample results in terms of LGD estimation – using different error measurements – and compare the results. These error parameters reflect the performance of our methods. Naturally, a low parameter outcome is preferable. We calculate the mean absolute error (MAE) and root mean squared error (RMSE) for each applied method according to the following definitions

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\mathrm{LGD}_i - \mathrm{LGD}_i^*|, \tag{2.4}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{LGD}_i - \mathrm{LGD}_i^*)^2}, \tag{2.5}$$

where LGD denotes the realized LGD, LGD^* is the predicted LGD, and n is the number of observations.

In addition to these measurements we calculate the Theil inequality coefficient

(TIC), presented by Theil (1966)

$$\mathrm{TIC} = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(\mathrm{LGD}_i - \mathrm{LGD}_i^*)^2}{\sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}\mathrm{LGD}_i^2} + \sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}(\mathrm{LGD}_i^*)^2}}. \tag{2.6}$$

TIC sets the mean squared error relative to the sum of the average quadratic realized and estimated LGD and thereby accounts for both the model's goodness of fit and robustness. The factor is bound to $[0, 1]$ with $\mathrm{TIC} = 0$ being the perfect estimator. Theil finds that a useful forecast can be made up to $\mathrm{TIC} \approx 0.15$.

For a better interpretation of the results, we also show the results of the historical average and two simple OLS regression models as benchmarks. We use identical explanatory variables for OLS regression as for the M5' algorithm and RT before applying the variable selection procedure. Similar to M5' and RT, we further apply a backward elimination to the OLS regression according to the BIC in Equation (2.1).

We estimate the LGD at two different points in time: once at the execution of the contract and once at the time of default. Typically, more information is available at default, which should theoretically yield better predictions.

The in-sample and out-of-sample results are evaluated by calculating the Janus quotient introduced by Gadd and Wold (1964)

$$\mathrm{Janus} = \sqrt{\frac{\frac{1}{n}\sum\limits_{i=1}^{n}(\mathrm{LGD}_i - \mathrm{LGD}_{i,\mathrm{Oos}}^*)^2}{\frac{1}{m}\sum\limits_{i=1}^{m}(\mathrm{LGD}_i - \mathrm{LGD}_{i,\mathrm{Is}}^*)^2}}, \tag{2.7}$$

with the in-sample estimation $\mathrm{LGD}_{\mathrm{Is}}^*$ in the denominator and the out-of-sample estimation $\mathrm{LGD}_{\mathrm{Oos}}^*$ in the numerator. $\mathrm{Janus} = 1$ for equally large prediction errors for both estimations. A value close to 1 indicates a stable model and data structure.

At the end of the chapter, we also provide quality features of the identified finite mixture distributions and the error rates of classification for robustness reasons, and we interpret these results.

2.5.1 In-sample results

Beginning with the in-sample outcomes presented in Table 2.4, our models largely produce better estimations with the additional information available at default.

	Company A				Company B				Company C			
Method	Lv.	MAE	RMSE	TIC	Lv.	MAE	RMSE	TIC	Lv.	MAE	RMSE	TIC
Hist. avg.		0.3418	0.4018	0.1381		0.3646	0.4205	0.1988		0.3662	0.4195	0.1822
At execution												
OLS		0.3240	0.3868	0.1268		0.2706	0.3451	0.1235		0.3282	0.3889	0.1519
OLS$_{\text{BIC}}$		0.3246	0.3874	0.1272		0.2719	0.3471	0.1251		0.3307	0.3909	0.1537
FMM$_{\text{3NN}}$		0.2806	0.3713	0.1099		0.2209	0.3354	0.1044		0.2571	0.3601	0.1176
FMM$_{\text{J4.8}}$		0.2919	0.3916	0.1043		0.2589	0.3911	0.1192		0.3028	0.3914	0.1255
M5'	13	0.3142	0.3786	0.1209	1	0.2711	0.3459	0.1241	2	0.3272	0.3874	0.1504
M5'$_{\text{BIC}^*}$	17	0.3132	0.3774	0.1201	9	0.2640	0.3388	0.1185	9	0.3148	0.3751	0.1400
RT	34	0.3183	0.3817	0.1231	7	0.2726	0.3464	0.1248	9	0.3279	0.3871	0.1510
RT$_{\text{BIC}^*}$	26	0.3197	0.3829	0.1240	11	0.2687	0.3423	0.1217	7	0.3314	0.3898	0.1531
At default												
OLS		0.3114	0.3761	0.1191		0.2692	0.3435	0.1211		0.3238	0.3858	0.1490
OLS$_{\text{BIC}}$		0.3123	0.3768	0.1195		0.2709	0.3451	0.1234		0.3238	0.3858	0.1527
FMM$_{\text{3NN}}$		0.2550	0.3468	0.0955		0.2148	0.3280	0.1001		0.2432	0.3437	0.1091
FMM$_{\text{J4.8}}$		0.2588	0.3594	0.0900		0.2408	0.3693	0.1056		0.2835	0.3723	0.1190
M5'	6	0.3014	0.3680	0.1134	2	0.2650	0.3399	0.1193	1	0.3274	0.3883	0.1513
M5'$_{\text{BIC}^*}$	12	0.2997	0.3666	0.1126	12	0.2539	0.3277	0.1101	3	0.3244	0.3858	0.1490
RT	49	0.3032	0.3689	0.1143	25	0.2642	0.3373	0.1181	13	0.3247	0.3844	0.1487
RT$_{\text{BIC}^*}$	39	0.3046	0.3699	0.1150	10	0.2674	0.3422	0.1215	7	0.3294	0.3886	0.1523

Table 2.4: In-sample estimation errors at the execution and default of contracts by company. The best results are underlined for each company and type of error. Hist. avg. is the historical average loss given default (LGD) used as estimation of the LGD. OLS represents the ordinary least squares regression, and FMM is the finite mixture model in combination with 3-nearest neighbors (3NN), or J4.8. OLS is also performed with the variable selection BIC algorithm and the M5' algorithm and the RT are performed with the variable selection BIC* algorithm. Lv. defines the number of leaves on the tree. MAE is the mean absolute error defined in Equation (2.4) and RMSE is the root mean squared error defined in Equation (2.5). TIC is the Theil inequality coefficient defined in Equation (2.6). For MAE, RMSE, and TIC, lower outcomes are preferable.

Our results clearly show the superiority of the FMMs for in-sample testing. The MAE, RMSE, and TIC of the FMM$_{\text{3NN}}$ are mostly far from their counterparts of the other models and even farther from the historical averages. The OLS regressions are outperformed in all cases.

Upon closer inspection, we note a large gap between the MAE and RMSE of the FMMs of approximately 10 percentage points, which is thus much larger than for the OLS regressions and tree-based models. We discuss this effect in more detail in Section 2.5.3. Our findings are consistent with Elbracht (2011) and with the discrepancy between MAE and RMSE noted by Loterman et al. (2012).

Proceeding with the tree-based models, we determine that by application of the variable selection procedures, $M5'_{BIC^*}$ strictly outperforms all RT models and the OLS regressions, except for company C at default. However, the variable selection is beneficial because, without it, the algorithm partly divides the contracts into only one or two classes, leading to estimation errors close to those of the OLS regression models. Furthermore, by applying the variable selection, we observe that more underlying contracts tend to be associated with more classes. Compared to the FMMs, the M5' models yield significantly higher MAEs, but they are somewhat competitive in terms of the RMSE.

The RTs tend to divide the contracts into significantly more classes than the M5' models. Nonetheless, the results are predominantly worse than those for the M5' models. The RTs, with all available explanatory variables RT and in combination with variable selection RT_{BIC^*}, outperform the OLS regressions for most companies. As expected, we notice that punishing the number of classes in RT_{BIC^*} results in a model with fewer classes and thereby reduces the prediction quality. Unlike the M5' algorithm, the RT can reduce its error only by increasing the number of classes. Likewise, OLS regression performs better when using all available variables.

The TIC of all considered models remains well below its values of the historical average and mainly less than the value of OLS. Additionally, all of the values are within the range of the suggested threshold value of TIC ≈ 0.15 or less, which confirms that the methods used are worth being considered for estimating LGDs.

2.5.2 Out-of-sample results

Most of the studies in this field report in-sample findings but not out-of-sample results, although the latter are crucial for proper risk management and are required for regulatory purposes. Certainly, our out-of-sample findings, summarized in Table 2.5, differ significantly from the in-sample results. Accordingly, to evaluate the method's efficiency and robustness, out-of-sample testing is essential because in-sample results can be misleading.

	Company A			Company B			Company C		
Method	MAE	RMSE	TIC	MAE	RMSE	TIC	MAE	RMSE	TIC
Hist. avg.	0.3437	0.4022	0.1383	0.3657	0.4221	0.1999	0.3679	0.4200	0.1828
At execution									
OLS	0.3257	0.3891	0.1282	0.2722	0.3469	0.1246	0.3348	0.3959	0.1576
OLS$_{\mathrm{BIC}}$	0.3262	0.3893	0.1285	0.2734	0.3479	0.1256	0.3369	0.3962	0.1583
FMM$_{\mathrm{3NN}}$	0.3539	0.4479	0.1600	0.2917	0.4178	0.1621	0.3593	0.4755	0.2056
FMM$_{\mathrm{J4.8}}$	0.3424	0.4422	0.1544	0.2749	0.4004	0.1453	0.3313	0.4193	0.1720
M5'	0.3235	0.3879	0.1271	0.2723	0.3475	0.1250	0.3365	0.3957	0.1576
M5'$_{\mathrm{BIC}^*}$	0.3215	0.3873	0.1264	0.2711	0.3467	0.1242	0.3384	0.4004	0.1607
RT	0.3245	0.3890	0.1280	0.2751	0.3490	0.1266	0.3386	0.3961	0.1587
RT$_{\mathrm{BIC}^*}$	0.3243	0.3888	0.1278	0.2746	0.3480	0.1259	0.3386	0.3961	0.1595
At default									
OLS	0.3132	0.3786	0.1206	0.2702	0.3447	0.1229	0.3319	0.3949	0.1560
OLS$_{\mathrm{BIC}}$	0.3143	0.3790	0.1209	0.2730	0.3494	0.1260	0.3334	0.3945	0.1565
FMM$_{\mathrm{3NN}}$	0.3204	0.4214	0.1410	0.2832	0.4085	0.1549	0.3345	0.4504	0.1870
FMM$_{\mathrm{J4.8}}$	0.2958	0.3988	0.1253	0.2741	0.3996	0.1464	0.3297	0.4260	0.1732
M5'	0.3100	0.3767	0.1191	0.2678	0.3433	0.1220	0.3332	0.3951	0.1565
M5'$_{\mathrm{BIC}^*}$	0.3069	0.3757	0.1178	0.2659	0.3428	0.1206	0.3341	0.3977	0.1579
RT	0.3136	0.3804	0.1216	0.2710	0.3462	0.1244	0.3370	0.3958	0.1583
RT$_{\mathrm{BIC}^*}$	0.3142	0.3811	0.1220	0.2727	0.3474	0.1252	0.3384	0.3979	0.1598

Table 2.5: Out-of-sample estimation errors at the execution and default of contracts by company. The best results are underlined for each company and type of error. Hist. avg. is the historical average loss given default (LGD) used as estimation of the LGD. OLS represents the ordinary least squares regression, and FMM is the finite mixture model in combination with 3-nearest neighbors (3NN), or J4.8. OLS is also performed with the variable selection BIC algorithm and the M5' algorithm and the RT are performed with the variable selection BIC* algorithm. MAE is the mean absolute error defined in Equation (2.4) and RMSE is the root mean squared error defined in Equation (2.5). TIC is the Theil inequality coefficient defined in Equation (2.6). For MAE, RMSE, and TIC, lower outcomes are preferable.

Our findings indicate that, in general, $M5'_{BIC^*}$ generates the best out-of-sample results, although the performance seems to depend on the size of the underlying data set. Consistent with the in-sample results, we observe predominately more accurate estimations using the additional information available at default.

Concerning the FMMs, we first note that the in-sample favorable FMM_{3NN} is now outperformed by all of the other models and also partly by the historical averages. This outcome is unexpected because the in-sample results are good and sturdy. The $FMM_{J4.8}$ generates isolated good MAE values but the RMSE and TIC values are worse than their counterparts from the tree-based models and OLS regressions. As we did in-sample, we continue to note a large gap between the MAE and RMSE for both FMMs; furthermore, the TIC values exceed the suggested value of 0.15 by several times.

Our results clearly demonstrate that by application of the variable selection procedure, the model tree $M5'_{BIC^*}$ is the best choice for companies A and B. For these two companies, applying the variable selection procedure to the model tree algorithm is beneficial without exception. Whereas both model tree methods outperform the RT models, $M5'_{BIC^*}$ also generates consistently better MAE, RMSE, and TIC values than both OLS regressions. For company C, we obtain a slightly different picture. Considering the performance measures in total, the OLS regression – particularly using all available explanatory variables – is favorable now. At the very least, at execution, the M5' algorithm generates equally good or even slightly better RMSE and TIC values as the OLS regressions. However, the results of the $M5'_{BIC^*}$ and both RT models are worse for company C. Consistent with the in-sample results, we find that the variable selection procedure almost throughout worsens the results of the OLS regression and RT for all companies.

The out-of-sample results suggest to some extent a link between the numbers of observations and the relative performances of the estimation methods considered. Containing the LGD data from three different companies, our data set provides us with a particularly good opportunity to analyze this relationship in greater detail. Bearing in mind the ranking of the data set sizes, company A delivered the largest number of observations (9,735), followed by company B (2,995) and company C (1,592).

First of all, we note that the TIC exceeds the suggested value of 0.15 for all of the methods in the case of company C. For companies A and B, TIC values remain well below 0.15, at least with respect to the tree-based models and the OLS regressions. This finding indicates that the prediction accuracy in general becomes weaker if the underlying data set contains fewer observations. Furthermore, we note that the performances of the model trees relative to the regression model OLS improve with an increasing data set size. At default of the contracts, M5' performs 0.51% worse than OLS for company C concerning the MAE. But, for company B, M5' performs 0.89% better than OLS and the improvement increases to 1.02% for company A. Analogously, this tendency applies to the RMSE. At execution of the contract the relative performance improves with an increasing sample size only regarding the MAE. Actually, for M5'$_{\mathrm{BIC}^*}$, the estimation accuracy relative to OLS improves throughout monotonically with an increasing sample size. Moreover, the link between the performance relative to OLS and the number of observations included in the underlying data set is even more distinctive for M5'$_{\mathrm{BIC}^*}$. At default M5'$_{\mathrm{BIC}^*}$ performs 0.66% worse than OLS for company C concerning the MAE, but 1.59% (2.01%) better than OLS for company B (A).

With respect to the RT models we cannot establish an unambiguous link between the performances relative to OLS and the sample size. Compared with OLS, the performances of RT and RT$_{\mathrm{BIC}^*}$ improve with an increasing data set size only with regard to the MAE. Whereas, regarding the RMSE the results of RT deteriorate relative to OLS with an increasing data set size at default of the contract and RT$_{\mathrm{BIC}^*}$ obtains the relatively worst outcomes at execution and default of the contract on the data set of company B. Also concerning the FMMs and OLS_{BIC} we could not identify a link between the performances relative to OLS and the number of observations contained in the underlying data set. Both FMMs and OLS_{BIC} obtain relative to OLS the worst results almost throughout for company B.

We further compare the results for each of the random divisions of the respective data set used for out-of-sample testing. The findings support the link between the performances of M5' and particularly $M5'_{BIC^*}$ relative to OLS and the sample size. For company C, $M5'_{BIC^*}$ yields better results than OLS on only about 5 partitions. With an increasing sample size, $M5'_{BIC^*}$ performs better than OLS significantly more often, to be precise, for company B on at least 60% of the divisions and for company A on more than 90%. Actually, at execution of the contract, $M5'_{BIC^*}$ yields throughout better MAE values than OLS for company A.

Although there might be several factors influencing the estimation accuracy of the models, such as idiosyncratic firm characteristics, we find the sample size to be of particular importance. We apply an additional test to confirm the link between estimation accuracy and sample size. Pooling the three data sets generates a large sample that contains 14,322 contracts (100%). We randomly draw 7,161 contracts (50%) out of the large sample to generate a medium sized sample. For a small sample, we randomly draw 1,432 contracts (10%) out of the large sample. We repeat these random drawings ten times, leaving us with a total of 21 data sets. Again, for out-of-sample testing we split the data sets randomly into 75% training sample and 25% test sample. This step is also done ten times. We showed before, that the $M5'_{BIC^*}$ seems to be particularly sensible to small sample sizes. Also, bearing in mind that $M5'_{BIC^*}$ and the regression model OLS perform best for companies A and B, respectively for company C, we focus on testing the impact of sample size for these two models. All results are averaged with respect to the sample size.

We see in Table 2.6 that both methods perform better on larger data sets. Un-

	100% (large)		50% (medium)		10% (small)	
Method	MAE	RMSE	MAE	RMSE	MAE	RMSE
At execution						
OLS	0.3327	0.3977	0.3311	0.3943	0.3359	0.4012
$M5'_{BIC^*}$	0.3186	0.3877	0.3194	0.3869	0.3323	0.4002
At default						
OLS	0.3316	0.3964	0.3285	0.3943	0.3334	0.4001
$M5'_{BIC^*}$	0.3069	0.3783	0.3098	0.3814	0.3285	0.3998

Table 2.6: Out-of-sample estimation errors at the execution and default of contracts by sample size. The sample sizes are 100% (large), 50% (medium), and 10% (small) of all contracts. OLS is the ordinary least squares regression without variable selection and $M5'_{BIC^*}$ is the M5' algorithm with the variable selection BIC^* algorithm. MAE is the mean absolute error defined in Equation (2.4) and RMSE is the root mean squared error defined in Equation (2.5). For MAE and RMSE lower outcomes are preferable.

like OLS regression, the estimation accuracy of $M5'_{BIC^*}$ increases almost monotonically with increasing sample size. In particular, the degree of accuracy improvement is clearly higher for the $M5'_{BIC^*}$. The $M5'_{BIC^*}$ performs consistently better than OLS regression. The improvement of $M5'_{BIC^*}$ over OLS regression increases with increasing sample size. At default of the contract $M5'_{BIC^*}$ performs 1.5% better than OLS regression on small data sets concerning the MAE. The improvement increases to 5.7% and 7.4% on medium sized and large data sets. With regard to the RMSE the improvement over OLS regression is 0.1% (2.8%, 4.6%) for small (medium, large) data sets.

Furthermore, we compare the results of $M5'_{BIC^*}$ and OLS regression for each of the 210 randomly drawn subsamples. We find that $M5'_{BIC^*}$ outperforms OLS regression in all drawings on large and medium sized data sets, whereas OLS regression achieves better results on small samples in 28% of the drawings concerning the MAE and in over 40% concerning the RMSE.

We conclude that the $M5'_{BIC^*}$ should be based on an adequately large data set to process the information more efficiently than the OLS regression. Moreover, the performed test confirms the link between prediction accuracy and sample size.

Considering that the data set of company C contains the fewest observations, and the M5'$_{\text{BIC}^*}$ improves with additional observations, we conclude that the M5'$_{\text{BIC}^*}$ in general is the best choice for out-of-sample predictions.

At first glance, the differences between the values of the accuracy measurements of the sophisticated estimation methods compared to OLS regression without variable selection seem to be negligible in most cases and are consistent with the results of Zhang and Thomas (2012). This finding raises the question as to whether it is worth the effort to implement more demanding estimation methods. For a more illustrative interpretation of our results, we use the average aggregated EAD of our test sample, which is €133,671,554 (€34,762,061) for company A (B) to estimate the total loss for the test sample. Using the M5'$_{\text{BIC}^*}$ yields an estimation that is in expectation up to €220,000 more accurate than the OLS regression for company B and for company A, the estimation is even up to €1,340,000 more accurate. Thus, improvements of an even few percentage points matter in terms of the parameter outcomes.

Our results indicate that in-sample results are an insufficient indicator of a method's out-of-sample performance. In particular, for the in-sample outperforming FMM$_{\text{3NN}}$, the results are obviously misleading because the out-of-sample predictions are worst. Hence, we study the stability of our models using the Janus quotient, as shown in Table 2.7. According to the Janus quotient, we can partition our methods into stable and unstable methods. A Janus quotient close to 1 indicates a stable model and data structure, which holds for the tree-based models, the OLS regressions, and the historical averages mainly with quotients less than 1.05. Exclusively taking into account the stable models, we observe more or less the same order concerning the estimation accuracy in-sample and out-of-sample. In particular, the M5'$_{\text{BIC}^*}$ performs in-sample conspicuously better than the other stable methods for companies A and B. This finding remains valid for the out-of-sample results without exception, only the advantage is smaller. As expected,

	Company A		Company B		Company C	
Method	Exec.	Dflt	Exec.	Dflt	Exec.	Dflt
Hist. avg.	1.0011	1.0011	1.0037	1.0037	1.0013	1.0013
OLS	1.0059	1.0066	1.0052	1.0035	1.0180	1.0236
OLS_{BIC}	1.0049	1.0058	1.0023	1.0125	1.0136	1.0226
FMM_{3NN}	1.2062	1.2151	1.2456	1.2453	1.3202	1.3105
$\text{FMM}_{\text{J4.8}}$	1.1291	1.1095	1.0238	1.0818	1.0713	1.1442
M5'	1.0246	1.0236	1.0046	1.0100	1.0214	1.0175
$\text{M5'}_{\text{BIC}^*}$	1.0262	1.0248	1.0233	1.0461	1.0674	1.0308
RT	1.0191	1.0312	1.0075	1.0264	1.0232	1.0297
RT_{BIC^*}	1.0154	1.0303	1.0167	1.0152	1.0162	1.0239

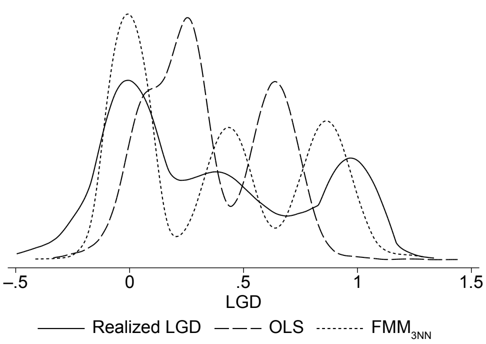
Table 2.7: Janus quotient for in-sample and out-of-sample estimations of loss given default (LGD) for each method and company at execution (Exec.) and default (Dflt) of contracts. The quotient is calculated according to Equation (2.7) and is constant for the historical average. A Janus quotient greater than 1 indicates that the error for the out-of-sample estimation is greater than the error for the in-sample estimation. OLS represents the ordinary least squares regression, FMM is the finite mixture model in combination with 3-nearest neighbors (3NN), or J4.8. OLS is also performed with the variable selection BIC algorithm and the M5' algorithm and the RT are performed with the variable selection BIC* algorithm.

for the FMMs, a Janus quotient that is mainly considerably greater than 1 indicates that these models are unstable. For the FMM_{3NN}, the quotient consistently exceeds 1.20. Hence, if out-of-sample testing is impossible, e. g. due to an insufficiently large data set, the in-sample results can be used as a prime indicator of the out-of-sample performance for stable methods, but this relationship obviously does not apply for unstable methods.

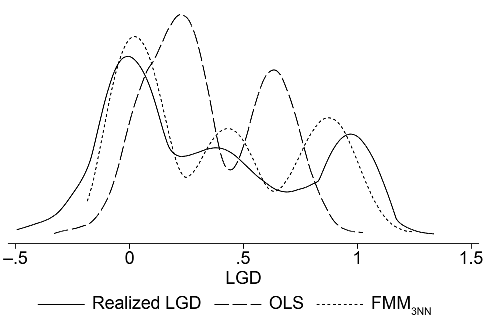
2.5.3 Validation and interpretation

To analyze the models' performances in detail and to elaborate on the several steps of FMMs, we present some key figures of our methods in this section.

FMMs produce accurate in-sample results by aiming to reproduce the distribution density. This relationship is true for both of our FMMs and is independent of the choice of the classification method in step (2). Figure 2.3a displays the realized



(a) In-sample LGD density.

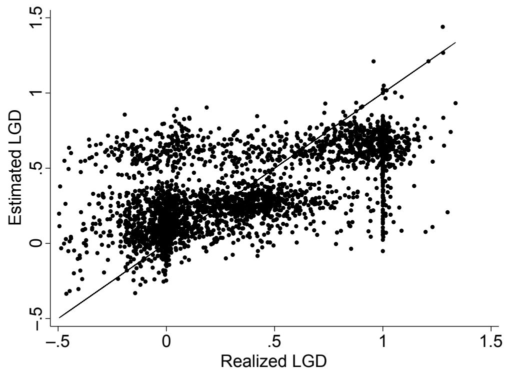


(b) Out-of-sample LGD density.

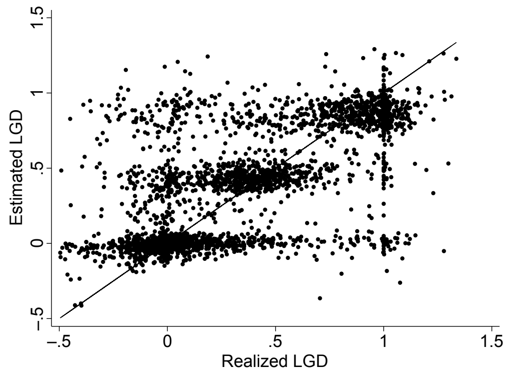
Figure 2.3: Densities of realized loss given default (LGD), LGD estimated by ordinary least squares (OLS) regression without variable selection, and LGD estimated by finite mixture combined with 3-nearest neighbors (FMM_{3NN}) for company B. The in-sample approximation of the realized LGD distribution by FMM is already good (Figure a) and it even improves in the out-of-sample estimation (Figure b). OLS regression, by contrast, is visibly only slightly changing and is not necessarily improving from in-sample to out-of-sample estimation.

and estimated LGDs for company B. Whereas OLS regression is not capable of properly accounting for the multimodality of the realized LGD distribution, the FMM's estimation is a good approximation. However, such density representations could be misleading because they are not capable of showing the deviation of an estimate from its realized value. This effect becomes particularly clear when we consider the out-of-sample results of the FMMs. For misclassified observations during either the clustering or classification process, the RMSE increases rapidly, while the approximation of the density remains accurate (Figure 2.3b).

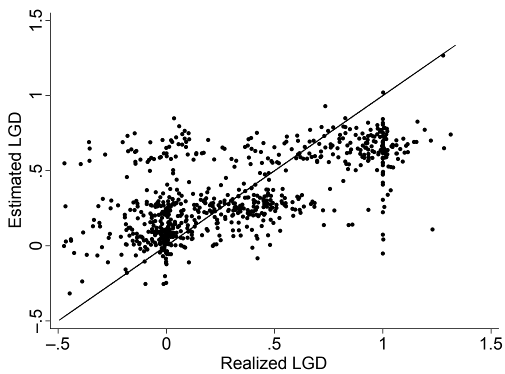
The effect can also be observed regarding the scatter plots in Figure 2.4. For both OLS regression and FMM, the in-sample estimation of LGD is rather concentrated around the diagonal in Figures 2.4a and 2.4b. Out-of-sample, we notice for OLS regression in Figure 2.4c that the LGD estimates are thinned out uniformly, which leaves most of its density close to the diagonal. The FMM, by contrast, retains a relatively large amount of its estimates that are far from the diagonal, thus far from the realized value of the LGD. These large deviations consequently result in a larger RMSE. The MAE remains at an acceptable level because most



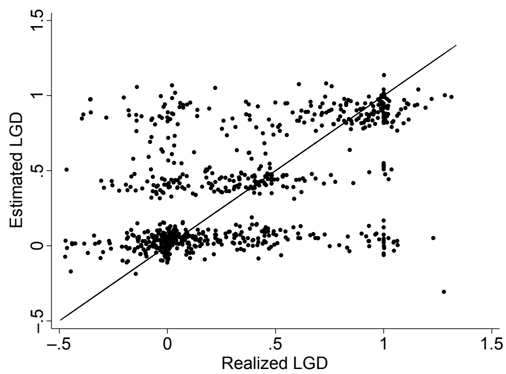
(a) In-sample: realized LGD versus estimated LGD by OLS regression.



(b) In-sample: realized LGD versus estimated LGD by FMM_{3NN}.



(c) Out-of-sample: realized LGD versus estimated LGD by OLS regression.



(d) Out-of-sample: realized LGD versus estimated LGD by FMM_{3NN}.

Figure 2.4: In-sample and out-of-sample: realized loss given default (LGD) versus estimated LGD by ordinary least squares (OLS) regression without variable selection (Figures a and c) and finite mixture combined with 3-nearest neighbors (FMM_{3NN}) (Figures b and d) for company B. Each figure has a simple diagonal line to illustrate the deviation.

of the density stays on the diagonal.

We analyze the quality of FMMs by examining the density of the a posteriori probability of belonging to a certain class, as proposed by Grün and Leisch (2007). The classification becomes more unambiguous as the probability approaches one, which indicates the quality of the adaptation.

For mixing distributions with two clusters, the average in-sample probability that observations are classified into a particular class is at least 88%, whereas the median is close to one. Poorer performance is observed with three clusters because of the larger overlap caused by additional clusters, resulting in lower classification

	Company A		Company B		Company C	
Method	Exec.	Dflt	Exec.	Dflt	Exec.	Dflt
3NN	0.2176	0.1925	0.2018	0.1834	0.1943	0.1808
J4.8	0.3740	0.3371	0.3693	0.3085	0.2833	0.2634

Table 2.8: In-sample classification errors for the 3-nearest neighbors (3NN) and J4.8 methods at execution (Exec.) and default (Dflt) of the contracts. Given the clustering of the finite mixture model in step (1) (see Section 2.4.1), a contract is classified incorrectly if the classification algorithm (3NN or J4.8) in step (2) assigns this contract to a different cluster. The classification error is then the relative number of falsely assigned contracts.

probabilities. However, three-quarters of all observations are classified during the clustering process in step (1), with a minimum probability of 58%.

Validating the classification methods is even more important than validating the clustering in step (1) of the procedure. Although clustering works well for all of the companies, classifying the observations with the information available at the contract execution and default is critical. By reviewing the classification errors of our classification methods, we analyze the performance of these methods. Thus, we can assess the percentage of incorrectly assigned observations. This process only works in-sample because, for unseen cases the true class is unknown. Table 2.8 demonstrates an improvement when we compare classification errors at the execution and default of the contract for both methods. The 3NN approach clearly results in a more accurate classification, which is attributable to the 2-clustered mixing distribution. J4.8 distinguishes among three clusters. These clusters naturally overlap to a significant extent, which results in higher classification errors. The errors are in line with the MAE and RMSE in Table 2.4.

The number of mixing distributions is an exogenous parameter. Our model with three mixing normal distributions constantly produces the smallest AIC and BIC. However, the lower classification error in step (2) might suggest a model with two mixing normal distributions. In terms of MAE and RMSE, neither the parameters, such as the AIC and BIC of the mixing models, nor the in-sample classification error is a consistently good performance indicator for the composed method. Our results in Section 2.5.2 show that the in-sample classification error and out-of-sample MAE and RMSE do not behave proportionally. By contrast, AIC and BIC work well with the out-of-sample results of the $\mathrm{FMM}_{\mathrm{J4.8}}$.

Overall, the large difference between the MAE and RMSE arises from the entire procedure of FMMs, which are focused on accurately mapping the LGD density. Out-of-sample in particular, the classification is problematic, which becomes obvious in Figure 2.4d, as a large number of estimations is far from the realized value. Therefore, reproduction results in comparatively robust MAEs, but the RMSE rises quadratically and penalizes these outliers.

Reproducing the LGD distribution to yield accurate estimations is proposed by Hlawatsch and Ostrowski (2011). Qi and Zhao (2011), however, conclude that mapping the density is only of minor importance for precisely predicting the LGD. Using transformation regressions under different parameters, they cannot establish a link between the ability to map the density properly and the estimation accuracy, neither in-sample nor out-of-sample. To some extent the results of the FMMs support this conclusion. Nonetheless, there is a significant difference. For instance, the $\mathrm{FMM}_{\mathrm{3NN}}$ generates accurate predictions in-sample and only performs worse out-of-sample. This finding suggests that the $\mathrm{FMM}_{\mathrm{3NN}}$ adapts well to the training data by reproducing the density, but it also indicates that overfitting might be a severe problem.

With regard to out-of-sample predictions, a high level of adaptation to the underlying training data is only reasonable if the training and test data are exceedingly homogeneous. Given inhomogeneous data sets, a good adaptation to the training data basically involves potential overfitting. This relationship is also supported by the results of the model trees. The data set of company C contains notably fewer observations than those of companies A and B. Moreover, the TIC for company C exceeds the suggested value of 0.15 out-of-sample for all of the methods. Thus, it can reasonably be concluded that the training and test data are comparatively inhomogeneous. M5'$_{BIC^*}$ performs strictly better than M5' for all three companies in-sample; in other words, the M5'$_{BIC^*}$ attains a superior adjustment to the underlying data set. Out-of-sample, however, M5'$_{BIC^*}$ yields better results only for companies A and B whereas M5' is beneficial for company C. Hence, given inhomogeneous training and test data, the superior adaptation to the training data is not transferred into sturdy out-of-sample predictions.

For the FMMs, the classification is obviously of prime importance, and out-of-sample in particular, it is problematic. However, classification is also relevant for the tree-based models because the observations are also partitioned into different classes. Certainly, by contrast to the FMMs, the tree-based models use more classes.[2] This increased number of classes indicates that in case of the M5' models, the different classes considerably overlap with one another. For the RTs, the classes are spread over the entire observation interval. In-sample, the number of misclassified contracts is manageable for both the tree-based models and the FMMs. As a result, the latter method mainly yields accurate in-sample estimations (Figure 2.4b), whereas the predictions of the tree-based models, particularly in terms of the MAE, are not as accurate. Naturally, it is more difficult to classify unseen observations correctly. This fact also holds for the tree-based models, although the out-of-sample predictions are significantly better than those of the FMMs. However, based on the tree model's class structure, a misclassified contract tends to be placed into an adjacent class; thus, the resulting error remains low. By contrast, the classes of the FMMs are largely disjointed; thus, the error for a misclassified observation tends to be more significant.

2.6 Conclusion

We use contracts of three leasing companies separately to evaluate various models in-sample and out-of-sample at two different points in time. Our findings prove

[2]The number of classes is chosen by the algorithm and is not defined ex ante.

that out-of-sample testing is essential for evaluating a model for LGD estimation. In-sample results might be significantly misleading when estimating out-of-sample LGDs, which are crucial for proper risk management and are required for regulatory purposes.

FMMs account for the multimodality of the LGD density. Combined with the classification algorithm 3NN, this method achieves the lowest in-sample MAE, RMSE, and TIC values. In particular, it outperforms the historical average and the OLS regressions, which were used as benchmarks. Along with the FMMs, the model tree with variable selection $M5'_{BIC^*}$ yields the best results for in-sample estimation.

Out-of-sample, a clear trend can be observed that model trees and particularly $M5'_{BIC^*}$ generate the best results. Compared with OLS regression the performance of $M5'_{BIC^*}$ improves notably with an increasing data set size. We confirm this result by applying an additional test, in which we eliminate idiosyncratic features by pooling the three data sets. Furthermore, for the company with the fewest observations, the TIC values indicate that all applied methods have difficulties predicting the LGD of unseen contracts accurately. As opposed to in-sample results, FMMs now are outperformed even by the OLS regression; in particular, FMM_{3NN} performs worst.

The Janus quotient determines the stability of our models, dividing them into stable and unstable methods. In particular, the in-sample results of unstable methods, namely the FMMs, cannot be used as indicators for out-of-sample estimation errors.

3 Does the Economic Situation Affect the Loss Given Default of Leases?

3.1 Introduction

One of the elementary components of risk management in financial institutes is the quantification of credit risk. The knowledge of the expected loss of a financial asset is essential for a proper allocation of regulatory and economic capital. Beside the probability of default (PD) and the exposure at default (EAD), the credit risk is in particular determined by the loss given default (LGD) respectively its counterpart, the recovery rate. The LGD represents the fraction of the EAD that a financial institute looses if a debtor defaults.

According to the Basel II accord and the Capital Requirement Regulation (CRR), financial institutes may choose between the Standardised Approach and the Internal Ratings Based Approach (IRBA) in order to calculate their capital requirements for credit risk. To implement the advanced IRBA, it is necessary to develop internal models for the estimation of PD, EAD, and LGD which are in line with the regulatory framework. Despite the importance of the LGD estimation, previously the academic literature mainly dealt with the calculation of the PD. Just in the recent past, several studies have addressed the estimation of the LGD. However, the majority of these studies focused on bonds and partly on loans, whereas only a few studies covered the LGD of leasing contracts (see, e. g., Schmit and Stuyck (2002), Laurent and Schmit (2005), De Laurentis and Riani (2005), Hartmann-Wendels and Honal (2010), and Hartmann-Wendels et al. (2014)). Although loans and leases have similar characteristics, there exist some crucial differences. A specific characteristic of leasing contracts is that the lessor keeps the legal ownership of the leased asset during the entire contract period. This permits the leasing company to repossess and dispose the leased asset if a debtor defaults. In particular, the lessor can retain all returns from disposing the leased asset. Therefore, leasing companies benefit from comparatively low LGDs.

Especially in the light of the regulatory requirements, one crucial aspect of LGD research should be the dependency between the LGD and the economic situation. According to section 468 of the Basel II framework, LGD estimations must reflect economic downturn conditions wherever necessary to capture the relevant risks. Moreover, from a practical point of view, the substantial influence of the economic situation on credit risk in general became particularly evident during the financial crisis. Serious losses, which led to difficulties in several financial institutes, clearly demonstrated that credit risk is significantly higher during a recession. However, the number of empirical studies that dealt with the influence of the economic conditions on the LGD of loans or leases is limited. As already stated by Bruche and González-Aguado (2010), for the PD time-variations are taken for granted, whereas the LGD is often assumed to be constant over time.

The existing literature that addressed the LGD of loans and leasing contracts has primarily either investigated different estimations methods for the LGD or analyzed the factors that influence the LGD. In order to identify the key drivers of the LGD, most studies focused on analyzing the relationship between the LGD and idiosyncratic factors, such as contract characteristics and customer characteristics (see, e. g., De Laurentis and Riani (2005), Grunert and Weber (2009), Bastos (2010), Gibilaro and Mattarocci (2011), and Khieu et al. (2012)). Notably, the findings of the studies are different on some factors and these differences cannot be solely attributed to differences between loans and leases. Studying Portuguese loans, Bastos (2010) found, e. g., that the LGD increases with an increasing size of the loan. On the other hand, Thorburn (2000) could not identify a significant relationship between the LGD and the size of the loan when analyzing Swedish data. Furthermore, Chalupka and Kopecsni (2009), investigated Czech loans and found, e. g., a positive relationship between the LGD and the EAD. In contrast, both Gibilaro and Mattarocci (2011) for bank transactions and Elbracht (2011) for leases observed a negative link between the LGD and the EAD.

Incorporating macroeconomic factors in the analysis of the drivers of the LGD has been less common so far. In fact, sporadically a single macroeconomic factor was taken into account, but in particular the influence of the economic situation as a whole was not considered. On the other hand, in the recent past, some studies on loans concentrated specifically on the relationship between the LGD and the economic situation, but paid only little attention to the influence of the idiosyncratic factors on the LGD. Araten et al. (2004) and Emery (2008) both observed that LGDs are on average higher during recessions and lower during expansions. Moreover, e. g., Caselli et al. (2008) and Leow et al. (2014) studied the relationship between the LGD and various macroeconomic factors. The authors found that the LGD of loans depends among others on the growth rate of the gross domestic product and the level of interest. Nevertheless, relating to leases, studies that dealt with the relationship between the economic situation and the LGD are extremely rare. It is important to consider that for leases, the relationship between the LGD and macroeconomic factors might be different than to loans, as the lessor keeps the legal ownership of the leased asset. The corresponding disposal of the leased asset during the workout generates an additional source of payments which can be particularly influenced by the lessor. Early studies by Schmit and Stuyck (2002) and Laurent and Schmit (2005) actually argued that the LGD of European leasing contracts is essentially independent on the economic situation. However, a recent study by Hartmann-Wendels and Honal (2010) exclusively considered a set of macroeconomic factors and found evidence that the LGD of German leases indeed depends on the economic cycle.

To the best of our knowledge, so far no study examined at once the influence of the economic situation on the LGD and its idiosyncratic key drivers with regard to leasing contracts. In order to fill this gap, this paper analyzes the influence of the economic situation, modeled by various macroeconomic factors, on the LGD of defaulted leases while simultaneously taking into account essential idiosyncratic factors. As the LGD undoubtedly depends on idiosyncratic factors, it is crucial to identify and incorporate the idiosyncratic key drivers of the LGD when analyzing the relationship between the LGD and the economic situation. Considering exclusively macroeconomic factors and neglecting these idiosyncratic drivers of the LGD may result in an overestimation respectively underestimation of the dependency of the LGD on the economic situation. Consequently, such models may prove to be unreliable.

The data we use for our study are provided by two German leasing companies and cover a wide range of economic conditions including the impact of the recent financial crisis. Bearing in mind the observed differences in the existing literature concerning the relationship between the LGD and some idiosyncratic factors, the separate analysis of two lessors is of particular interest. In line with the Basel II framework, we estimate the LGD at two different points in time, once at execution of the contract and additionally at default of the contract. To validate our results we perform in-sample and out-of-time testing. Although out-of-time testing is of major importance from a practical point of view and is mandatory to meet the regulatory requirements, it has rarely been done by other studies (e. g., Bastos (2010) and Bellotti and Crook (2012)).

The detailed analysis of the relationship between the LGD and macroeconomic factors over a long observation period across different economic conditions is particularly essential for the proper calculation of a downturn LGD. Several studies,

e. g., Hartmann-Wendels and Honal (2010), Bellotti and Crook (2012), and Leow et al. (2014), proposed to use macroeconomic stress tests for calculating a downturn LGD. Assuming that the LGD is higher during an economic downturn, the LGD is modeled depending on macroeconomic factors. The downturn LGD is then obtained by the use of unfavorable realizations of the macroeconomic factors. However, this approach requires a stable relationship between the LGD and the macroeconomic factors, which in particular consists during a downturn period. Ex ante it is generally unknown whether this condition is satisfied.

To ensure that our results are not caused by a specific estimation method, we conduct our analysis with both a linear and a nonlinear model. Selecting a suitable method for estimating the LGD is principally challenging due to the characteristic density function of the LGD. The LGD of loans and leases is typically bimodal respectively multimodal distributed with high concentrations near zero and one and in some cases additionally around zero point five (see, e. g., Laurent and Schmit (2005), Bastos (2010), Hartmann-Wendels and Honal (2010), Qi and Zhao (2011), Zhang and Thomas (2012), and Hartmann-Wendels et al. (2014)). It means that predominately either low recoveries or nearly complete recoveries occur. Furthermore, especially with regard to loans it is often assumed that the LGD is bounded within the interval [0,1] (see, e. g., Bastos (2010)). This implies that the lender cannot recover respectively lose more than the outstanding debt. However, the LGD of leases exceeds both limits certainly (see, e. g., Laurent and Schmit (2005) and Hartmann-Wendels and Honal (2010)). In addition to the linear regression that was successfully used for predicting the LGD, e. g., by Bellotti and Crook (2012) and Zhang and Thomas (2012), previous studies applied various advanced methods for the estimation of the LGD (see, e. g., Bastos (2010), Qi and Zhao (2011), Loterman et al. (2012), and Hartmann-Wendels et al. (2014)). As it has been noted, reproducing the LGD distribution is only of secondary importance for estimating the LGD. Comparing different estimation methods, Qi and Zhao

(2011) rather concluded that the outcomes differ depending on whether linear or nonlinear relationships between the LGD and the explanatory variables are established. Similar was stated by Loterman et al. (2012). Therefore, it seems to be particularly reasonable to use both a linear and a nonlinear model for our analysis.

Summarized, our study makes the following crucial contributions to the existing literature. First, we analyze the influence of macroeconomic factors on the LGD while simultaneously taking into account essential idiosyncratic factors, by which we meet the requirements of the Basel II accord. In particular, we analyze under which circumstances the consideration of macroeconomic factors could lead to significantly better estimations of the LGD. We find that the benefit of using macroeconomic factors for the LGD estimation of leases crucially depends on the point in time the estimates are carried out. In particular at execution of the contract we observe a clear link between the LGD and the economic situation. Second, we study the relationship between the LGD and idiosyncratic factors to identify its idiosyncratic key drivers. By separately analyzing data from two lessors with both a linear and a nonlinear model we obtain insight why the observed effects of idiosyncratic factors on the LGD are partly inconsistent in the existing literature. Our results show that the idiosyncratic key drivers of the LGD substantially differ between the investigated leasing companies. However, we observe that the LGD generally depends more on object characteristics and contract characteristics than on customer characteristics.

The remainder of this study is structured as follows. Section 3.2 contains an overview of the data used in this study, describes the factors that potentially influence the LGD, and introduces some descriptive statistics. In Section 3.3 we discuss the methods used in this study. Section 3.4 comprises the analysis of the in-sample results, followed by a discussion of the out-of-time results in Section 3.5. Finally, in Section 3.6 we present the conclusion of our study.

3.2 Data

This study uses data originally provided by three major German leasing companies. In our analysis, however, we only consider two lessors, as we noted inconsistencies regarding the default dates for the third dataset. According to Table 3.1, the remaining datasets contain 2,908 defaulted leasing contracts of 2,270 different lessees for lessor A and 9,171 defaulted leasing contracts of 5,430 different customers for company B. For both companies the datasets consider leases that have defaulted between 2002 and 2009. The dates of execution of the leasing contracts are between 1995 and 2009.[3] Consequently, the datasets of both companies meet the requirements set out in section 472 and 473 of the Basel II accord. According to this the historical data the LGD estimation is based on has to cover an observation period of seven years respectively five years in case of retail contracts.

Company	#Contracts	#Lessees	∅ LGD	σ(LGD)	Min. LGD	Max. LGD
A	2,908	2,270	0.3518	0.4192	-0.4955	1.3359
B	9,171	5,430	0.5041	0.3996	-0.4948	1.4866

Table 3.1: Number of contracts and lessees in the datasets of the companies A and B. In addition loss given default (LGD) density information for both companies: average LGD, standard deviation of the LGD, minimum LGD and maximum LGD.

Concerning the general availability of data, it has to be stated that even major companies suffer difficulties in providing data sufficient for an adequate LGD estimation. Only in the recent years financial institutions started to establish detailed loss databases, due to the increasing regulatory requirements. For this reason, datasets for LGD estimation including a long timespan are rare. With regard to our datasets we emphasize that our observation period does not only meet the requirements of the Basel II accord but in particular cover two crises that have had a significant impact on the German economy. Firstly this is the Dotcom crisis in 2000. The Dotcom crisis was triggered by the bursting of a

[3]The observation period ends at the end of the year 2009 because for the following years, no further data have been provided.

speculative bubble in March 2000 and led to recessionary tendencies in Germany. The second significant crisis is the financial crisis, beginning in the second half of 2007. The financial crisis began with the subprime crisis in the summer of 2007, and developed into a global economic crisis in 2008. In the course of the financial crisis, there were recessions in many industrial countries, including in Germany. In particular, it should be noted that even the supposedly short observation period of considered contract defaults covers an entire economic cycle. This is sufficient to analyze whether there is a stable relationship between the LGD and the economic situation.

The default definition both companies use meet the requirements of the Basel II accord. Moreover, the workout of all defaulted contracts is completed and we calculated the LGD as one minus the workout recovery rate. The recovery rate is the ratio of the recovery amount to the EAD. Corresponding to the requirements of the Basel II accord, we calculated the recovery amount as the sum of all cash inflows, discounted to the time of default using the term-related refinancing interest rate, reduced by the expenses incurred during the workout. The EAD is given by the sum of the present value of the outstanding minimum lease payments, compounded lease payments, and the present value of the calculated residual value. In particular, both companies use identical definitions for the data which entered into the LGD calculation.

As shown in Table 3.1, the average realized LGD of the two companies differs significantly. Company A realizes on average a LGD of 35.18 %, whereas the average LGD of company B is 50.41 %. This significant difference in the average LGD cannot be attributed to a different field of activity, as both companies offer a wide range of leasing objects and make contracts with both retail and non-retail customers. Although the average LGD differs, we observe uniformly for both companies that the LGDs of the individual contracts scatter strongly over the interval $[-0.5,1.5]$. This can be also recognized by means of the LGD distribution,

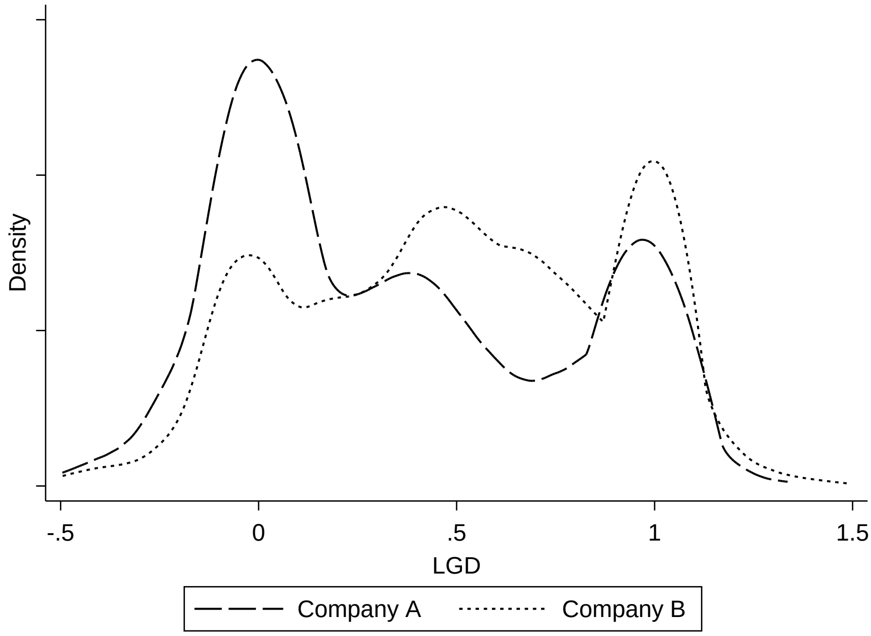


Figure 3.1: Density of the realized loss given default (LGD) for company A and B. The realized LGD concentrates on the interval [−0.5,1.5]. The figure describes a loss severity of −50 % (−0.5) on the left end, which indicates that 150 % of the exposure at default (EAD) was recovered. On the right end, the loss severity is 150 % (1.5), indicating a loss of 150 % of the EAD.

illustrated in Figure 3.1. For both companies the LGD distribution exhibits a multimodal shape. However, there are differences in the LGD distribution of the two companies. For company A we basically observe the frequently mentioned two maxima around zero and one, while company B features a third maximum around zero point five. In addition, we note that several of the realized LGDs exceed the interval [0,1]. Negative LGDs for leasing contracts have been also observed, e. g., by Laurent and Schmit (2005) and Hartmann-Wendels and Honal (2010). The latter cited that negative LGDs in particular result from contracts with a rather small EAD that yield high recoveries from the disposal of the leased asset. LGDs greater than one can occur as we consider workout expenses.

With regard to the apparent differences in the LGD profiles among the two companies, we do not pool the LGD data. In fact, in line with the regulatory requirements, we construct individual estimation models for both companies to account for institution specific characteristics.

3.2.1 Fluctuations of the LGD

To get a first impression of the evolution of the LGD over time, we analyze the fluctuations of the average LGD of contracts that have been executed or defaulted in a year, presented in Table 3.2.

	Company A				**Company B**			
	Executed contracts		Defaulted contracts		Executed contracts		Defaulted contracts	
Year	#Contracts	∅ LGD	#Contracts	∅ LGD	#Contracts	∅ LGD	#Contracts	∅ LGD
1995					2	0.5155		
1996	10	0.2355			4	0.4519		
1997	19	0.3119			10	0.4525		
1998	71	0.3150			101	0.5426		
1999	169	0.3777			318	0.4776		
2000	362	0.2971			671	0.5228		
2001	627	0.2681			981	0.4864		
2002	453	0.2820	363	0.2932	1,140	0.4692	1,115	0.5281
2003	333	0.3001	666	0.3197	1,133	0.4527	1,128	0.5161
2004	394	0.4688	570	0.3183	1,046	0.4408	1,167	0.4638
2005	274	0.4931	572	0.3725	953	0.5266	1,041	0.4258
2006	123	0.5056	419	0.3765	1,074	0.5478	1,057	0.4731
2007	66	0.6205	266	0.4601	969	0.5430	789	0.5172
2008	5	0.4665	46	0.6450	607	0.5825	1,268	0.5481
2009	2	-0.0598	6	-0.1218	112	0.6457	1,606	0.5381

Table 3.2: Number of executed respectively defaulted contracts and related average loss given default (LGD), by year and company. All of the executed contracts listed defaulted later between 2002 and 2009.

We first examine the evolution of the average LGD according to the year of execution. Measured by the total number of observations, the majority of the contracts was executed between 1998 and 2007 for company A and between 1999 and 2008 for company B. It should be noted that contracts executed during the last year of the observation period are possibly underrepresented, because we only consider contracts with completed workout. With respect to the remaining periods, we observe in particular for company A significant differences in the average LGD for each year. Here, the LGD varies between 26.81 % in 2001 and 62.05 % in 2007, while company B realized the lowest average LGD in 2004 with 44.08 % and the highest average LGD of 58.25 % in 2008. If we additionally consider for company B the year 2009 with an average LGD of 64.57 %, the scatter of the average LGDs becomes more apparent, but still not as pronounced as for company A.

We note that the evolution of the average LGDs differs partly for both companies. Around the Dotcom crisis in 2000, we observe higher LGDs for contracts signed in 1999 (2000) for company A (B). Furthermore, we notice that the average LGDs are comparatively low subsequent to the Dotcom crisis for both companies, but increase significantly particularly in the years before the financial crisis. However, for company A the increase of the average LGD is more pronounced and starts already in 2004 while company B realized higher average LGDs only from 2005 onwards. Moreover, we observe that the average LGD continuously increased during the financial crisis for company B.

We further continue analyzing the evolution of the average LGDs with regard to the year of default. The number of all defaulted contracts is distributed over the entire observation period for company B, while for company A the majority of the contracts defaulted between 2002 and 2008. As at execution of the contracts, for company A we observe significant differences in the average LGD for the individual years. In the period from 2002 to 2008, the average LGD varies between 29.32 % and 64.50 %. However, in contrast to the execution of the contracts, this wide spread results almost exclusively by a significant increase of the averge LGD in the years 2007 and especially 2008. For company B, the average LGD fluctuates moderately between 42.58 % in 2005 and 54.81 % in 2008.

At default of the contracts we observe notable differences in the evolution of the average LGDs for both companies, in particular during the financial crisis. At first glance the average LGD started to increase in 2007 for both companies, although the change is more pronounced for company A. In fact, for company A the average LGD already exceeded the previous highest value considerably in 2007. In 2008, the average LGD then continued to increase significantly by almost 20 %. Company B, however, realized higher LGDs than before the financial crisis only from 2008 onwards, and the LGDs exceeded the highest pre-crisis values by at most 2 %.

3.2.2 Workout Characteristics

In the preceding sections we have seen that the realized LGD differs significantly for the two companies. Company A exhibits on average about 15 % lower LGDs than company B and additionally, the average LGDs of company A are subjected to noticeable higher fluctuations over time. In the following, we examine the revenues of the lessors during the workout to reveal possible structural differences between the two companies. Table 3.3 presents the share of the revenues from disposing the leased asset respectively the payments made by the lessee in relation to the total payments the lessor receives during the workout. We focus on these two sources of revenues, because commonly they represent the major portion of the income the leasing company receives during the workout. Moreover, in contrast to the revenues from disposing the leased asset respectively the payments made by the lessee, other sources of revenues, such as guarantees, are typically associated with considerable costs for the lessor.

	Company A				**Company B**			
	Share of asset resale		Share of customer payments		Share of asset resale		Share of customer payments	
Year of default	∅	σ	∅	σ	∅	σ	∅	σ
2002	0.4297	0.3854	0.4977	0.3934	0.6849	0.4278	0.1155	0.2627
2003	0.5746	0.3879	0.3566	0.3788	0.7314	0.4025	0.1187	0.2648
2004	0.5582	0.3943	0.3758	0.3879	0.7628	0.3749	0.1219	0.2578
2005	0.5411	0.4188	0.3708	0.4051	0.7617	0.3727	0.1144	0.2445
2006	0.5888	0.4085	0.3471	0.3867	0.7506	0.3857	0.1105	0.2467
2007	0.6083	0.4186	0.2243	0.3302	0.6964	0.4300	0.1000	0.2530
2008	0.5318	0.4774	0.0986	0.2536	0.7242	0.4066	0.1006	0.2380
2009	0.5683	0.4788	0.0284	0.0440	0.6778	0.4093	0.1567	0.2861

Table 3.3: Average share and standard deviation of the inflows from disposing the leased asset respectively through customer payments in relation to the total inflows received during the workout, by company and year of default. The inflows by the customer include direct payments by the customer and inflows from the dispose of additional collateral.

In general, the revenues from disposing the leased asset represent a substantially higher proportion of the total payments for company B compared to company A. The share of the inflows from disposing the leased asset amounts on average around 70 % for company B, whereas for company A the proportion averages around 55 %. In contrast, the payments made by the lessee are significantly higher for company

A, representing partially more than 35 % of the total payments received during the workout. For company B, the share of the lessee payments amounts just over 10 %. This composition of the essential inflows the lessors retain during the workout suggests that company A cooperates closer with the lessee and manages the workout more actively.

Considering the development of the recoveries over time, it is noticeable that the proportion of the revenues from disposing the leased asset is relatively constant for both companies. This indicates that the inflows from the realization of the leased asset are only minor affected by the economic situation. In fact, the leasing companies are able to obtain adequate revenues also during downturns due to their good knowledge of the secondary markets. This argumentation is also supported by the fact that the LGD, which is based exclusively on the returns from disposing the leased asset, increases only slightly in the wake of the financial crisis. In contrast, with regard to the evolution of the payments made by the lessee, we note significant differences between the two companies. For company B the comparatively low proportion of payments made by the lessee remains largely constant over time, whereas for company A the share of the lessee payments depends considerably on the economic situation. In the wake of the financial crisis, the proportion of the payments made by the lessee starts to decline in 2007 by more than 10 % for company A. One possible explanation for the observed decline is the typically decreasing order situation during an economic downturn. Consequently, the lessee obtains less inflows which complicates paying the lessor. As the share of the inflows from disposing the leased asset does not increase during the financial crisis, company A apparently tries to compensate the lack of lessee's payments by other sources of revenues, such as guarantees. However, these other revenues are typically associated with considerable costs for the lessor, such as legal costs, which leads to significant higher LGDs, as discussed in Section 3.2.1.

3.2.3 Explanatory Idiosyncratic Factors

Concerning the leasing contracts, both companies provided a great deal of specific information about factors that might influence the LGD. We have classified these idiosyncratic factors into four groups: (1) object characteristics, (2) contract characteristics, (3) customer characteristics, and (4) additional information at default. In the following we discuss the importance of the individual factors (variable definitions see appendix in Section 3.7) in terms of the potential impact on the LGD. In particular, we take reference to the results of existing empirical studies.

Previous studies, e. g., Schmit and Stuyck (2002), De Laurentis and Riani (2005), and Hartmann-Wendels and Honal (2010), have shown that the LGD of leasing contracts essentially depends on the type of the leased asset. As the legal owner of the leased asset, the lessor can dispose the leased asset if the contract defaults. Laurent and Schmit (2005) and Hartmann-Wendels and Honal (2010) pointed out that the value of machines is relatively stable over time while liquid secondary markets exist for vehicles. Both indicate comparatively low LGDs. In contrast, information and communication technology (ICT) facilities are characterized by a rapid loss of value, which has a negative effect on the LGD. For our study, we distinguish between vehicles, machinery, ICT, equipment, and other objects.

The original value, which can be considered as an indicator of the contract volume, could also effect the LGD. With regard to loans, the evidence on the link between loan size and LGD is different, which possibly can be traced back to the consideration of different credit classes. Whereas among others, Thorburn (2000) and Khieu et al. (2012) could not observe any significant association between the size of a loan and the LGD, Bastos (2010) and Dermine and de Carvalho (2006) have found a positive relationship between these variables. The latter argued that banks possibly delay the foreclosure of larger loans, which finally results in higher

LGDs. In contrast, Acharya et al. (2007) have determined a negative relation between the debt size and the LGD, which they attributed to the fact that high debts are particularly taken by large borrowers that have a higher bargaining power in bankruptcy. Based on leasing contracts, De Laurentis and Riani (2005) have also found that the original value is negatively related to the LGD. A reason for this outcome might be that the lessor monitors a lessee with a higher contract volume more closely.

Another factor we consider is whether the procured leased asset has already been used. A used object may lose excessively fast in value, e. g. by unpredictable arising damage. This entails a lower income from the asset resale and therefore a higher LGD.

In terms of contract characteristics, we consider the potential impact of the contract type on the LGD. We distinguish full pay-out lease contracts, partial amortization contracts, hire-purchase contracts and other types of contracts. Compared to partial amortization contracts, full pay-out contracts are characterized by a more favorable loan to value ratio for the lessor because the return of capital is constant during the term to maturity while for partial amortization contracts the majority of the lessee's payments occur at maturity. For hire-purchase contracts the transfer of ownership of the leased asset to the lessee at maturity is obligatory. As Elbracht (2011) pointed out, this may affect a more careful handling of the leased asset by the lessee which results in lower LGDs.

Further, we take into account the information whether the lessee has a purchase option. To the best of our knowledge the influence of this factor on the LGD has not been studied so far. However, we only have the appropriate information for company B.

With respect to loans, Grunert and Weber (2009) and Bastos (2010) have found that a poor creditworthiness of the debtor implies a higher LGD. As we have no direct information about the creditworthiness of the lessees, we consider the interest rate implicit in the lease. In addition to the available interest rate level at execution of the contract and the term to maturity, the interest rate implicit in the lease includes, in particular, the risk premium. Therefore a high interest rate implicit in the lease is more likely assigned to a riskier lessee.

Furthermore we consider the term to maturity of the lease as a factor potentially influencing the LGD. De Laurentis and Riani (2005) have exhibited that the LGD of leasing contracts decreases with an increasing term to maturity. However, it should be noted that the studies of Schmit and Stuyck (2002) and Schmit et al. (2003) showed evidence that for vehicle leasing, a short term to maturity indicates slightly lower LGDs.

We also include the ratio of rent prepayments to the original value of the leased asset, which can be considered as an indicator of the proportion of the contract volume that is already paid at the beginning of the contract period. Rent prepayments have a positive effect on the loan to value ratio and Elbracht (2011) has empirically shown that the LGD decreases with the existence of rent prepayments.

Moreover, we take into account the potential impact of the calculated residual value on the LGD whereas we explicitly look at the ratio of the calculated residual value to the original value of the leased asset. This ratio can be considered as a broad indicator of the proportion of the outstanding payments at default of the contract, which can be covered by the disposal of the leased asset. Theoretically, the LGD should decrease with an increase of this ratio.

De Laurentis and Riani (2005) have shown that the presence of buy-back agreements with asset suppliers imply significantly lower LGDs. However, we consider this factor only for company B, as we do not have the appropriate information for company A.

Previous studies on loans, e. g., Caselli et al. (2008), Grunert and Weber (2009), and Khieu et al. (2012), have shown that the existence of collateral typically implies lower LGDs. Compared to loans, leases are already secured by the underlying leased asset, nevertheless, we take this factor into account as the results of Elbracht (2011) indicated that also the LGD of leases potentially decreases with the existence of additional collateral.

With regard to the customer characteristics, we distinguish among retail and non-retail customers. In studying the LGD, this distinction has been rarely considered by now. Nonetheless, the distinction among retail and non-retail customers could influence the LGD, e. g. the study of Grunert and Weber (2009) on loans attested higher LGDs to large companies.

In addition we use the information whether the contract is a subsequent contract as an indicator of the existence of a longterm contractual relationship between the leasing company and the lessee. To the best of our knowledge, this aspect has not been considered in terms of leasing so far. For loans, Grunert and Weber (2009) have outlined that the existence of a longterm contractual relationship is significant and leads to lower LGDs, but this result is not generally confirmed by other studies.

After default of the contract, we obtain information about the EAD. This could influence the LGD, although the empirical evidence is mixed. Elbracht (2011) for leases and Gibilaro and Mattarocci (2011) for bank transactions have both found that the LGD decreases with an increasing EAD. They argued that the lender monitors the recovery process more closely if the potential losses are high. On the other hand, it should be considered that a high EAD, which also significantly exceeds the potential selling value of the leased asset, is more difficult to compensate for the debtor, which consequently could lead to a higher LGD. Laurent and Schmit (2005) cited a similar argument, moreover Chalupka and Kopecsni (2009) empirically observed increasing LGDs with increasing EADs for loans.

In addition to the absolute EAD, we also take into account the ratio of the EAD to the original value of the leased asset. In contrast to the absolute EAD, this ratio is an indicator of the proportion of the initial contract volume the lessee has already paid. Zhang and Thomas (2012) have found that the LGD of loans increases if the ratio of EAD to loan size increases. We expect the same for leasing contracts.

Moreover, the relative contract age, defined as the term between execution and default of the contract relative to the term to maturity of the contract, could influence the LGD. De Laurentis and Mattei (2009) argued theoretically that the LGD should decrease with an increasing relative contract age because the loan to value ratio develops beneficial for the leasing company. This explanation is consistent with the assumed relation between the LGD and the ratio of the EAD to the original value of the leased asset. Empirically, however, the results of Schmit and Stuyck (2002), Schmit et al. (2003), and Elbracht (2011) have shown that at least no linear relationship between the relative contract age and the LGD can be established.

In case of company B, we have information on whether the leasing company has transferred the monitoring of the lease to another company. Such contracts may show a divergent structure of the LGDs.

Finally, after default of the contract, we replace the ratio of the calculated residual value to the original value of the leased asset by the ratio of the calculated residual value to the EAD. The latter should be a more precise indicator of the proportion of the outstanding payments after the default of the contract, which can be covered by the disposal of the leased asset.

3.2.4 Explanatory Macroeconomic Factors

As we have seen in Section 3.2.1, the average LGD for each year varies considerably over time. This is an indication that the LGD does not only depend on idiosyncratic factors but in particular on the economic situation. Therefore it is obvious to use macroeconomic factors for the estimation of the LGD in addition to the mentioned idiosyncratic factors (variable definitions see appendix in Section 3.7). These macroeconomic factors map the respective economic situation at execution and default of the contract.

The first macroeconomic factor we use for explaining the LGD is the growth rate of the gross domestic product in comparison to the preceding quarter. The growth rate of the gross domestic product is the most common way to represent the economic situation. If the economic environment is represented just by the growth rate of the gross domestic product, we intuitively expect that the LGD is higher in periods of a low growth rate. In particular the manufacturing sector has to face the impact of a poor order situation in those times of economic downturn. Hence, there are only low inflows available the lessee can use to repay the debt. This results in a higher default risk and moreover, in lower LGDs in case of default, because during the workout the lessee is only able to make small payments. Some empirical studies, e. g., Araten et al. (2004) and Emery (2008), have observed that LGDs are on average higher during economic downturns. However, the empirical evidence about the relationship between the growth rate of the gross domestic product and the LGD is comparatively weak. Although the expected relationship could be observed by Khieu et al. (2012) on loans, e. g., Caselli et al. (2008) pointed out that other macroeconomic factors than the growth rate of the gross domestic product may be more suitable for explaining the LGD.

Another factor we use to depict the economic development is the index of the business climate monthly provided by the Ifo Institute for Economic Research. If the economic situation is mapped solely by the business climate, we likewise expect an inverse relationship between the LGD and the business climate. In addition, the Ifo Institute for Economic Research also publishes monthly the index of the business expectations. This index has been already used by Hartmann-Wendels and Honal (2010), it provides information about the expected future development of the economic situation. Those information are relevant both at execution and default of the contract. In terms of the default of the contract, it should be noted that the workout typically lasts multiple months and therefore positive business expectations could indicate lower LGDs.

Furthermore, we consider the potential influence of the growth rate of the gross fixed asset investments on the LGD. This factor was also used in the study by Hartmann-Wendels and Honal (2010). The growth rate of the gross fixed asset investments provides in particular information about whether a contract was executed during an investment boom.

Finally, we also take into account the level of lending rates, which can be considered as an general indicator for financing costs. Leow et al. (2014) have found that the level of interest rates has substantial influence on the LGD of mortgage loans. Because no public available time series of the lending rates covers the entire observation period, we use the monthly yield on domestic bearer bonds. For the period from January 2003 until December 2009, the latter are correlated almost 90% with the interest rates on loans to non-financial corporations.

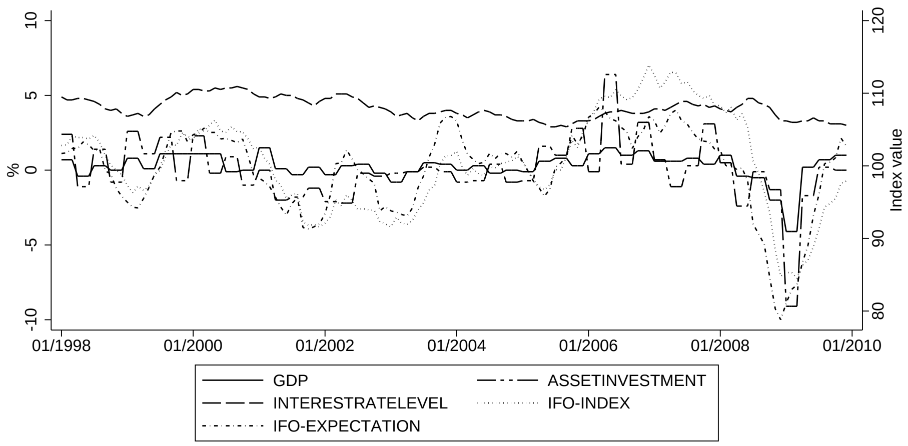


Figure 3.2: Evolution of the macroeconomic factors in the period from 01/1998 to 12/2009. GDP represents the growth rate of the gross domestic product in comparison to the preceding quarter, ASSETINVESTMENT stands for the growth rate of the gross fixed asset investments in comparison to the preceding quarter. INTERESTRATELEVEL is defined as the monthly average of the yield to maturity for domestic bearer bonds. These variables are given in percent. IFO-INDEX represents the monthly collected index business climate of the Ifo Institute for Economic Research and IFO-EXPECTATION stands for the monthly collected index business expectations of the Ifo Institute for Economic Research. These variables are given as an index value.

It is unavoidable that the considered macroeconomic factors are correlated to some degree. However, considering the evolution of the macroeconomic factors over time, shown in Figure 3.2, we still note differences between the individual factors. Hence, the different macroeconomic factors map different aspects of the economy. Consequently, as also Leow et al. (2014) have annotated, it is not sensible to remove some of the macroeconomic factors ex ante because important information would be lost. Instead, we consider in the following principally the influence of the macroeconomic factors on the LGD as a whole.

Nevertheless, a brief discussion of the evolution of the macroeconomic factors during the observation period is sensible. Moreover, analyzing particularly the growth rate of the gross domestic product we are able to identify the different phases of the economic cycles. Despite some differences, we recognize that the evolution of the growth rate of the gross domestic product, the index of the business climate, the index of the business expectations and the growth rate of the gross fixed asset investments is largely similar. Taking into account seasonal differences for the growth rate of the gross fixed asset investments, the years before the Dotcom crisis are characterized by a moderate increase of the gross domestic product and the gross fixed asset investments. The index of the business climate and the index of the business expectations also increase just before the Dotcom crisis, however, it should be noted that particularly compared to the growth rate of the gross domestic product both indices are generally much more volatile. The Dotcom crisis in 2000 led to a mild recession in Germany. We notice a significant decrease of all four mentioned factors. Partially even a decline of the gross domestic product and the gross fixed asset investments is observable. In 2003, particularly the index of the business situation and the index of the business expectations show first signs of a recovery. However, only from 2005 we observe a marked expansion of the economy. A notable rise of all four factors clearly indicates a period of good economic conditions. Subsequently, during the financial crisis, which led to a distinctive recession, we note considerable differences between the four factors. While the gross domestic product grew even at the beginning of the financial crisis before declining since the 2nd quarter of 2008, the index of the business situation and the index of the business expectations dropped already in the 3rd quarter of 2007. Furthermore, we recognize that the growth rate of the gross fixed asset investments experienced a first slump in the 2nd quarter of 2007, but that followed once more an upturn and a general decline of the gross fixed asset investments is observable from the 2nd quarter of 2008. Generally, we note that all four mentioned factors show at least first signs of recovery in the 2nd half of 2009.

Compared to the evolution of the mentioned macroeconomic factors, the shape of the yield curve is quite different. At the beginning of the observation period the level of interest rates decreased, but in advance to the Dotcom crisis this trend turned from the 2nd quarter of 1999. An increase respectively a high level of interest rates of about 5 % can be observed throughout the Dotcom crisis until mid-2002. Thereafter, the level of interest rates decreased steadily, reaching the lowest level of the observation period at the end of the 3rd quarter 2005. This low level of interest rates coincided basically with the beginning of the significant increase of the gross fixed asset investments. In the following, the interest rate level rose again, particularly also at the beginning of the financial crisis. Only from end of 2008 on, the level of interest rates started to decline markedly. However, the lowest pre-crisis level was not reached until the end of 2009.

3.3 Methods

Qi and Zhao (2011) have suggested that especially taking into account nonlinear relationships between the LGD and the explanatory variables contributes to more accurate LGD estimations. Furthermore, in line with the results of Hartmann-Wendels et al. (2014), the authors found that a methods ability to reproduce the LGD distribution does not necessarily lead to accurate LGD estimations. Therefore, in this study we focus in particular on the application of a linear as well as a nonlinear method to estimate the LGD. First, we apply a linear regression model, as it has been successfully used to estimate the LGD, e. g., by Bellotti and Crook (2012) and Zhang and Thomas (2012). The second method we apply is a regression spline model, which is essentially a nonlinear extension of the linear regression model. We focus on these two methods, because they do not differ in model-specific particularities, but notably in the fact that the regression spline model is able to depict nonlinear relationships between the LGD and the continuous explanatory variables.

The quite comparable outcomes of both models allow us to analyze the extent to which the assumed influences of the individual factors on the LGD differ in the linear and the nonlinear model. In particular, we are able to verify whether a general relationship between the LGD and the economic situation exists, which is not driven by a respective model. This verification is important, because the findings of Qi and Zhao (2011) may indicate that linear and nonlinear methods concentrate on different information. Additionally, we obtain particular insight whether the consideration of nonlinear effects actually implies a higher estimation accuracy.

The linear regression model, which we use to estimate the LGD is given by

$$\mathrm{LGD} = c + \sum_{i=1}^{k} a_i x_i + \sum_{i=k+1}^{m} a_i x_i + \varepsilon, \tag{3.1}$$

where c denotes the constant, $a_1, \ldots, a_m$ are the regression coefficients, $x_1, \ldots, x_m$ are the explanatory variables which represent the factors that might influence the LGD, and ε is the error term. Thereby the explanatory variables $x_1, \ldots, x_k$ are dummy variables whereas $x_{k+1}, \ldots, x_m$ represent continuous variables. Following the observation in Section 3.2 that the realized LGDs are not bounded within

[0,1], the linear regression and the regression spline model described below provide unbounded estimates of the LGD.

Subsequent to the explanation of the regression splines, this section also presents an overview of the performance measurements we use to evaluate the results of the different models. In addition, we describe our methodology for out-of-time testing.

3.3.1 Regression Splines

The application of regression splines allows depicting nonlinear relationships between the LGD and the continuous explanatory variables. For this study we use a model that was introduced by Royston and Sauerbrei (2007).

Generally, a spline of order p is a continuous function that is piecewise defined by polynomials of order p. Here, the points where two polynomials are connected are called knots. This means the range of values of a continuous explanatory variable is divided into intervals and within each interval, the respective influence of the explanatory variable on the LGD is modeled. Consequently, unlike as in the linear regression model, it is not implicitly assumed that the relationship between the explanatory variable and the LGD is identical over the whole range of values.

For a continuous explanatory variable x_i, $i \in \{k+1, \dots, m\}$ with corresponding knots κ_v, $v = 1, \dots, d$, the spline of order p is defined as follows

$$\mathrm{LGD}\left(x_i\right) = c_i + \sum_{s=1}^{p} a_{is} x_i^s + \sum_{v=1}^{d} b_{iv} \max\left(0, x_i - \kappa_{iv}\right)^p + \varepsilon_i, \tag{3.2}$$

where c_i is the constant, a_{is} and b_{iv} denote the regression coefficients and ε_i is the error term. For the estimation of the LGD we use exclusively linear splines of order $p = 1$. Hence, we focus on changes in the relationship between the LGD and an explanatory variable across the range of values. We find that linear splines represent a good tradeoff between an adequate consideration of nonlinear relationships and concurrently limiting potential overfitting. The results of Qi and Zhao (2011) and Hartmann-Wendels et al. (2014) indicate that the latter is a common concern of complex models, which may negatively affect the forecasting accuracy. Accordingly, the relationship between a continuous explanatory variable and the LGD is only modeled by a spline, if this nonlinear model provides a significantly better fit than a linear model. Otherwise the variable is depicted linearly as in the linear regression model.[4] That way the regression spline model forms a nonlinear extension of the linear regression.

Let $x_{k+1}, \ldots, x_l$ and $x_{l+1}, \ldots, x_m$ be the continuous explanatory variables that are considered linearly respectively nonlinearly. Then the regression spline model, which we use for estimating the LGD is given by

$$\mathrm{LGD} = c + \sum_{i=1}^{k} a_i x_i + \sum_{i=k+1}^{l} a_i x_i + \sum_{i=l+1}^{m} \mathrm{LGD}\left(x_i\right) + \varepsilon. \tag{3.3}$$

A close look at the equations (3.1) and (3.3) clarifies the extending character of the regression spline model compared to the linear regression.

3.3.2 Performance Measurements

We evaluate the estimation results of the different models by means of various performance measurements. In fact, we use the mean absolute error (MAE), the root mean squared error (RMSE), and the area under the regression error characteristic curve (REC Area). In-sample we also compute the adjusted coefficient of determination (R^2).

We calculate the MAE and the RMSE using the following definitions

[4] To avoid excessive complexity, we deviate from the recommendations of Royston and Sauerbrei (2007) and change the value of alpha, by which the statistical significance of modeling a variable nonlinearly instead of linearly is tested, to 0.01.

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n} |\mathrm{LGD}_j - \mathrm{LGD}_j^*|, \tag{3.4}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n} \left(\mathrm{LGD}_j - \mathrm{LGD}_j^*\right)^2}, \tag{3.5}$$

where LGD and LGD^* denote the realized respectively the predicted LGD and n is the number of observation. For both measures a low parameter outcome is preferable, because this implies a smaller difference between the realized and the predicted LGD.

Regression error characteristic (REC) curves were introduced by Bi and Bennett (2003) as a generalization of receiver operating characteristic curves to regression problems. The REC curve plots the error tolerance δ on the x-axis versus the accuracy of the regression model $\mathrm{acc}(\delta)$ on the y-axis. The accuracy $\mathrm{acc}(\delta)$ for a given tolerance δ is defined as the percentage of observations whose estimates are within this tolerance. We calculate the accuracy $\mathrm{acc}(\delta)$ according the following definition

$$\mathrm{acc}(\delta) = \frac{\#\left\{\mathrm{LGD}^*\colon |\mathrm{LGD}_j^* - \mathrm{LGD}_j| \leq \delta,\ j = 1, \ldots, n\right\}}{n}. \tag{3.6}$$

The REC Area provides a measure to evaluate the performance of a regression model. The larger the REC Area, the more accurate the estimates in total over all observations.

The adjusted R^2, we additionally calculate for the in-sample estimations measures the proportion of the variance that can be explained by the model. Consequently a higher outcome is preferable.

3.3.3 Out-of-time Testing

In the recent past, numerous studies have performed out-of-sample testing, but still very few studies, e. g., Bastos (2010) and Bellotti and Crook (2012), have conducted out-of-time tests. However, from a practical point of view in particular a method's ability to forecast the LGD of future contracts on the basis of historical data is of major importance. Moreover, with the implementation of out-of-time testing we meet the requirements of the Basel II accord. Because our observation period includes different economic situations, the out-of-time results provide insight on how the respective models perform in different circumstances. This is particularly important with regard to the macroeconomic factors. The relationship between the LGD and the explanatory variables might change according to movements in the economic environment.

	Model fitting		**Predictions at execution**		**Predictions at default**	
	Up to year	#Contracts	Year	#Contracts	Year	#Contracts
Company A	2004	517	2005	274	2005	572
	2005	1,119	2006	123	2006	419
	2006	1,696	2007-2009	73	2007-2009	318
Company B	2004	3,379	2005	953	2005	1,041
	2005	4,400	2006	1,074	2006	1,057
	2006	5,336	2007	969	2007	789
	2007	6,124	2008-2009	719	2008	1,268
	2008	7,237			2009	1,606

Table 3.4: Number of contracts with completed workout at the end of year t that are used for model fitting. Additionally, number of loss given default (LGD) estimations carried out at execution respectively default of the contract in year $t + 1$. Sorted by company and year.

For our out-of-time testing we adopt the walk-forward approach used by Gupton and Stein (2005). We fit the models each with the data of all contracts whose workout is completed by the end of year t and predict the LGD for all contracts that have been executed respectively defaulted in the subsequent year $t + 1$. Consequently, the number of contracts that are used to fit the models increase by time, which is also shown in Table 3.4. In order to ensure an adequate data base for fitting the models and to cover different macroeconomic conditions, we first estimate the LGD for the year 2005. In particular with regard to company A, we have to consider the decreasing number of observations at the end of the observation period that are additionally concentrated on certain times of the respective years. As set out in Table 3.4, we therefore aggregate the LGD estimations for the years 2007 to 2009 for company A and for company B we summarize the LGD estimations at execution of the contract for the years 2008 and 2009.

3.4 In-sample Analysis

In this section we analyze the in-sample results of our models for estimating the LGD. For both companies we have estimated the LGD with the nonlinear regression spline model and the linear regression. In each case we first performed the estimates once only using the idiosyncratic factors presented in Section 3.2.3 as explanatory variables and then with the additional use of the macroeconomic factors described in Section 3.2.4. We also considered the macroeconomic factors with lags, but as the results are generally slightly worse, they are not listed here. Since we are particularly interested in forecasting the LGD, we do not consider the macroeconomic factors with leads.

The in-sample performance measurements, presented in Table 3.5, consistently show that the additional use of macroeconomic factors leads to better estimations of the LGD, irrespective which estimation method is used. Accordingly, our outcomes are in line with the findings of, e. g., Bellotti and Crook (2012). Moreover, considering the specification of each model, illustrated in Table 3.6 and Table 3.7, we recognize that the macroeconomic factors provide additional information without having essential influence on the significance and effect of the idiosyncratic factors.

Bearing in mind the mentioned distinctive differences between the companies,

		Company A				Company B			
		Adj. R^2	MAE	RMSE	REC Area	Adj. R^2	MAE	RMSE	REC Area
Linear Regression									
At execution	FIRM	0.3260	0.2699	0.3432	0.7305	0.0733	0.3211	0.3843	0.6791
	FIRM+MACRO	0.3350	0.2677	0.3406	0.7328	0.0797	0.3197	0.3829	0.6806
At default	FIRM	0.3462	0.2648	0.3378	0.7356	0.1312	0.3061	0.3720	0.6941
	FIRM+MACRO	0.3494	0.2640	0.3367	0.7365	0.1381	0.3040	0.3704	0.6963
Regression Spline									
At execution	FIRM	0.3321	0.2676	0.3415	0.7328	0.0813	0.3190	0.3826	0.6813
	FIRM+MACRO	0.3433	0.2651	0.3384	0.7354	0.0912	0.3160	0.3803	0.6842
At default	FIRM	0.3577	0.2603	0.3343	0.7402	0.1459	0.3012	0.3687	0.6991
	FIRM+MACRO	0.3612	0.2601	0.3334	0.7403	0.1614	0.2972	0.3652	0.7031

Table 3.5: In-sample performance measurements at execution and default of the contracts by company. The estimates were carried out each with the nonlinear regression spline model and the linear regression model. FIRM represents that only idiosyncratic factors were used as explanatory variables, FIRM+MACRO implies that additionally macroeconomic factors were added, in each case the better result is underlined. R^2 represents the adjusted coefficient of determination. REC Area is defined as the area under the regression error characteristic curve, MAE is the mean absolute error, and RMSE is the root mean squared error. For R^2 and REC Area higher, for MAE and RMSE lower outcomes are preferable.

which are notably reflected in much more precise estimates of the LGD for company A, it is striking that the adding of the macroeconomic factors consistently has a positive effect on the LGD estimation for both companies. Unexpectedly, however, a close look on the adjusted R^2 reveals that it increases with the addition of the macroeconomic factors partly by more than 10 % for company B, but maximally around 3.5 % for company A. This observation can obviously not be attributed to the volatility of the realized LGDs over time. But it should be taken into account that the models for company A comprise more than twice as high adjusted R^2 than those for company B. This suggests that for company B the idiosyncratic factors can explain only a small part of the volatility of the LGDs and therefore additional factors are necessary.

Furthermore, comparing the results at execution and default of the contracts, we observe a noticeable higher value of the macroeconomic factors at execution of the contracts for company A. This result could be expected, since the fluctuations of the realized LGDs, analyzed in Section 3.2.1, are lower at default of the contract, which may be an indication of a reduced dependence on the economic situation.

		Company A				**Company B**			
		Regression Spline		Linear Regression		Regression Spline		Linear Regression	
	Variable	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO
Object	VEHICLE	-0.1835** (0.0892)	-0.1857** (0.0885)	-0.1730* (0.0896)	-0.1687* (0.0890)	-0.1442** (0.0598)	-0.1398** (0.0596)	-0.1463** (0.0600)	-0.1456** (0.0599)
	MACHINERY	-0.2705*** (0.0871)	-0.2719*** (0.0866)	-0.2844*** (0.0876)	-0.2895*** (0.0871)	-0.0310 (0.0596)	-0.0229 (0.0593)	-0.0340 (0.0598)	-0.0298 (0.0596)
	ICT	0.1609* (0.0875)	0.1652* (0.0869)	0.2079** (0.0875)	0.2010** (0.0869)	0.1265** (0.0618)	0.1314** (0.0615)	0.1334** (0.0621)	0.1322** (0.0619)
	EQUIPMENT	-0.1617 (0.1106)	-0.1776 (0.1098)	-0.1757 (0.1111)	-0.1820* (0.1104)	-0.0151 (0.0626)	-0.0054 (0.0624)	-0.0118 (0.0629)	-0.0055 (0.0627)
	ORIGINALVALUE	**1.0000***** (0.1633)	**1.0000***** (0.1441)	-0.0000*** (0.0000)	-0.0000*** (0.0000)	**1.0000***** (0.1591)	**1.0000***** (0.1485)	-0.0000** (0.0000)	-0.0000** (0.0000)
	USED	-0.0410 (0.0304)	-0.0413 (0.0302)	-0.0365 (0.0306)	-0.0364 (0.0304)	0.0183 (0.0121)	0.0119 (0.0122)	0.0196 (0.0121)	0.0150 (0.0122)
Contract	FULLPAYOUT	-0.0079 (0.0314)	0.0065 (0.0314)	0.0091 (0.0314)	0.0114 (0.0312)	-0.0055 (0.0268)	-0.0043 (0.0267)	0.0061 (0.0268)	0.0018 (0.0268)
	PARTIALAMORTISATION	0.0322 (0.0251)	0.0166 (0.0251)	0.0325 (0.0252)	0.0165 (0.0253)	-0.0035 (0.0217)	-0.0008 (0.0216)	0.0083 (0.0217)	0.0057 (0.0217)
	HIREPURCHASE	-0.0567** (0.0266)	-0.0387 (0.0266)	-0.0593** (0.0268)	-0.0445* (0.0267)	-0.0969*** (0.0231)	-0.0918*** (0.0231)	-0.0901*** (0.0230)	-0.0934*** (0.0231)
	PURCHASEOPTION					0.0522** (0.0245)	0.0472* (0.0244)	0.0571** (0.0246)	0.0568** (0.0245)
	INTEREST	-0.0247*** (0.0080)	0.0605*** (0.0231)	-0.0267*** (0.0080)	0.0583** (0.0232)	**1.0000***** (0.1308)	**1.0000***** (0.1526)	0.0070*** (0.0015)	0.0059*** (0.0015)
	MATURITY	0.0003 (0.0005)	-0.0002 (0.0005)	-0.0000 (0.0005)	-0.0006 (0.0005)	0.0006* (0.0004)	0.0002 (0.0004)	0.0001 (0.0004)	-0.0003 (0.0004)
	PRETOVALUE	-0.0214 (0.0908)	-0.0326 (0.0902)	-0.0085 (0.0912)	-0.0137 (0.0907)	-0.3998*** (0.0627)	-0.3432*** (0.0627)	-0.3476*** (0.0618)	-0.3097*** (0.0618)
	RESIDUALTOVALUE	0.0055 (0.0719)	0.0005 (0.0715)	0.0202 (0.0722)	0.0206 (0.0719)	-0.0300 (0.0391)	-0.0304 (0.0392)	-0.0507 (0.0391)	-0.0629 (0.0392)
	BUYBACK					-0.1463*** (0.0217)	-0.1520*** (0.0216)	-0.1608*** (0.0218)	-0.1623*** (0.0217)
	COLLATERAL	-0.0625*** (0.0156)	-0.0553*** (0.0155)	-0.0672*** (0.0156)	-0.0621*** (0.0156)	-0.0039 (0.0093)	-0.0039 (0.0093)	-0.0058 (0.0093)	-0.0065 (0.0096)
Customer	RETAIL	0.0360 (0.0464)	0.0456 (0.0461)	0.0471 (0.0466)	0.0562 (0.0463)	-0.0135 (0.0091)	-0.0122 (0.0090)	-0.0073 (0.0090)	-0.0069 (0.0090)
	EXTENSION	0.0265 (0.0296)	0.0286 (0.0295)	0.0241 (0.0298)	0.0313 (0.0296)	-0.0526* (0.0306)	-0.0660** (0.0305)	-0.0786** (0.0305)	-0.0889*** (0.0305)
Macroeconomic	GDP		-0.0223 (0.0148)		-0.0228 (0.0149)		-0.0006 (0.0091)		-0.0054 (0.0089)
	IFO-INDEX		-0.0013 (0.0028)		-0.0018 (0.0028)		**1.0000***** (0.1102)		0.0082*** (0.0011)
	IFO-EXPECTATION		0.0056** (0.0027)		0.0060** (0.0027)		**1.0000***** (0.1660)		-0.0077*** (0.0014)
	ASSETINVESTMENT		0.0112** (0.0057)		0.0105* (0.0057)		0.0039 (0.0029)		0.0027 (0.0029)
	INTERESTRATELEVEL		-0.0993*** (0.0246)		-0.0974*** (0.0247)		**1.0000***** (0.2150)		0.0048 (0.0066)
	CONSTANT	0.6409 (0.0955)	0.2449 (0.2074)	0.5673 (0.0950)	0.2459 (0.2046)	0.3896 (0.0748)	1.6884 (0.3250)	0.5580 (0.0700)	0.5156 (0.1257)

Table 3.6: In-sample coefficient estimates at execution of the contracts. The estimates were carried out for both companies each with the nonlinear regression spline model and the linear regression model. FIRM represents that only idiosyncratic factors were used as explanatory variables, FIRM+MACRO implies that additionally macroeconomic factors were added. In brackets the standard errors are reported. Coefficients printed in bold show that the the corresponding variables were considered nonlinear. * (**, ***) stands for the statistical significance at 10 % (5 %, 1 %) of the respective variable. Particularly surprising respectively new results are shown in grey shading.

		Company A				Company B			
		Regression Spline		Linear Regression		Regression Spline		Linear Regression	
	Variable	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO
Object	VEHICLE	-0.1952** (0.0867)	-0.1778** (0.0865)	-0.1723** (0.0870)	-0.1627* (0.0869)	-0.1420** (0.0577)	-0.1361** (0.0572)	-0.1439** (0.0581)	-0.1423** (0.0579)
	MACHINERY	-0.2702*** (0.0857)	-0.2594*** (0.0856)	-0.2927*** (0.0863)	-0.2872*** (0.0862)	0.0045 (0.0575)	0.0120 (0.0570)	-0.0129 (0.0579)	-0.0103 (0.0577)
	ICT	0.1746** (0.0858)	0.1921** (0.0858)	0.1996** (0.0862)	0.2066** (0.0861)	0.1758*** (0.0597)	0.1797*** (0.0592)	0.1578*** (0.0601)	0.1589** (0.0599)
	EQUIPMENT	-0.1683 (0.1088)	-0.1550 (0.1087)	-0.1893* (0.1095)	-0.1739 (0.1094)	0.0191 (0.0604)	0.0253 (0.0599)	0.0103 (0.0609)	0.0129 (0.0607)
	ORIGINALVALUE	**1.0000***** (0.2376)	**1.0000***** (0.1626)	-0.0000** (0.0000)	-0.0000** (0.0000)	**1.0000***** (0.2156)	**1.0000***** (0.2213)	0.0000 (0.0000)	0.0000 (0.0000)
	USED	-0.0443 (0.0297)	-0.0432 (0.0297)	-0.0395 (0.0299)	-0.0390 (0.0299)	0.0274** (0.0117)	0.0225* (0.0117)	0.0240** (0.0118)	0.0189 (0.0118)
Contract	FULLPAYOUT	0.0048 (0.0309)	0.0077 (0.0310)	0.0136 (0.0309)	0.0104 (0.0308)	0.0114 (0.0258)	0.0244 (0.0256)	0.0228 (0.0259)	0.0266 (0.0258)
	PARTIALAMORTISATION	0.0173 (0.0224)	0.0088 (0.0225)	0.0225 (0.0225)	0.0149 (0.0226)	0.0131 (0.0210)	0.0187 (0.0208)	0.0284 (0.0210)	0.0313 (0.0210)
	HIREPURCHASE	-0.0371 (0.0258)	-0.0371 (0.0258)	-0.0482* (0.0258)	-0.0492* (0.0258)	-0.0542** (0.0221)	-0.0363 (0.0220)	-0.0546** (0.0220)	-0.0458** (0.0221)
	PURCHASEOPTION					0.0501** (0.0221)	0.0467** (0.0235)	0.0489** (0.0238)	0.0536** (0.0237)
	INTEREST	-0.0068 (0.0083)	0.0097 (0.0101)	-0.0062 (0.0084)	0.0093 (0.0101)	**1.0000***** (0.1664)	**1.0000***** (0.1602)	0.0043*** (0.0014)	0.0041*** (0.0014)
	MATURITY	0.0005 (0.0005)	0.0004 (0.0005)	0.0002 (0.0005)	-0.0000 (0.0005)	**1.0000***** (0.2170)	**1.0000***** (0.2318)	0.0004 (0.0003)	0.0001 (0.0003)
	PRETOVALUE	0.0983 (0.0911)	0.1021 (0.0909)	0.1056 (0.0913)	0.1059 (0.0911)	-0.0748 (0.0623)	-0.0121 (0.0620)	-0.1243** (0.0606)	-0.0807 (0.0607)
	BUYBACK					-0.1552*** (0.0210)	-0.1591*** (0.0209)	-0.1585*** (0.0211)	-0.1606*** (0.0211)
Customer	COLLATERAL	-0.0714*** (0.0154)	-0.0673*** (0.0155)	-0.0764*** (0.0154)	-0.0739*** (0.0155)	-0.0009 (0.0090)	-0.0021 (0.0090)	0.0034 (0.0090)	0.0036 (0.0090)
	RETAIL	0.0382 (0.0455)	0.0391 (0.0454)	0.0415 (0.0458)	0.0430 (0.0457)	-0.0010 (0.0088)	-0.0000 (0.0087)	0.0064 (0.0088)	0.0083 (0.0088)
	EXTENSION	0.0478 (0.0327)	0.0396 (0.0327)	0.0885*** (0.0317)	0.0850*** (0.0316)	-0.0840*** (0.0310)	-0.0988*** (0.0308)	-0.1286*** (0.0292)	-0.1425*** (0.0293)
After Default	EAD	**1.0000***** (0.2093)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	**1.0000***** (0.1950)	**1.0000***** (0.2073)	-0.0000 (0.0000)	-0.0000 (0.0000)
	EADTOVALUE	**1.0000***** (0.1375)	**1.0000***** (0.1004)	0.1731*** (0.0284)	0.1654*** (0.0285)	**1.0000***** (0.0437)	**1.0000***** (0.0425)	0.3349*** (0.0158)	0.3472*** (0.0161)
	CONTRACTAGE	**1.0000***** (0.2376)	**1.0000***** (0.2694)	-0.0096 (0.0358)	-0.0234 (0.0363)	**1.0000***** (0.0664)	**1.0000***** (0.0635)	0.2022*** (0.0209)	0.2144*** (0.0211)
	SOLD					-0.0016 (0.0114)	-0.0098 (0.0114)	-0.0174 (0.0112)	-0.0206* (0.0112)
	RESIDUALTOEAD	0.0081 (0.0067)	0.0050 (0.0067)	0.0056 (0.0066)	0.0046 (0.0065)	-0.0880*** (0.0204)	-0.0824*** (0.0202)	-0.0916*** (0.0203)	-0.0902*** (0.0202)
Macroeconomic	GDP		-0.0422** (0.0185)		-0.0455** (0.0187)		**1.0000***** (0.0886)		0.0503*** (0.0076)
	IFO-INDEX		0.0080*** (0.0021)		0.0074*** (0.0021)		0.0012 (0.0011)		0.0040*** (0.0010)
	IFO-EXPECTATION		-0.0037 (0.0028)		-0.0028 (0.0028)		**1.0000***** (0.1051)		-0.0105*** (0.0013)
	ASSETINVESTMENT		0.0019 (0.0052)		0.0020 (0.0053)		-0.0104*** (0.0028)		-0.0108*** (0.0027)
	INTERESTRATELEVEL		-0.0235 (0.0149)		-0.0278* (0.0150)		**1.0000****** (0.1671)		0.0032 (0.0077)
	CONSTANT	0.5665 (0.1071)	0.0537 (0.2281)	0.3622 (0.0095)	-0.0263 (0.2237)	-0.3169 (0.0867)	-0.0922 (0.1388)	0.1783 (0.0711)	0.7990 (0.1152)

Table 3.7: In-sample coefficient estimates at default of the contracts. The estimates were carried out for both companies each with the nonlinear regression spline model and the linear regression model. FIRM represents that only idiosyncratic factors were used as explanatory variables, FIRM+MACRO implies that additionally macroeconomic factors were added. In brackets the standard errors are reported. Coefficients printed in bold show that the the corresponding variables were considered nonlinear. * (**, ***) stands for the statistical significance at 10 % (5 %, 1 %) of the respective variable. Particularly surprising respectively new results are shown in grey shading.

Unexpectedly, however, for company B at contract default only the adjusted R^2 increases less than at execution of the contract, but not the other performance measures. With regard to the realized LGDs, we already observed that these vary only marginally over time (see Section 3.2.1).

Beyond, we observe that in-sample the nonlinear regression spline model achieves a higher estimation accuracy than the linear regression. In particular, we note that the regression spline model mostly benefits more from the addition of the macroeconomic factors. A possible reason for this outcome, discussed by Miller and Töws (2014), is that the relation between continuous explanatory variables and the LGD cannot be adequately recognized linearly but nonlinearly. As already Leow et al. (2014) mentioned, this might apply especially to the macroeconomic factors.

3.4.1 Analysis of the Idiosyncratic Factors

Referring to Table 3.6 and Table 3.7, we analyze in detail the influence of the idiosyncratic factors on the LGD. We find that the LGD generally depends in particular on object characteristics and contract characteristics. Customer characteristics, however, have only little effect on the LGD. Moreover, the additional information only available at default of the contract also contain significant drivers of the LGD. The importance of these additional information at default for the LGD estimation can be already seen by means of the improved performance measurements at default, displayed in Table 3.5.

With a few exceptions the documented effects of the significant idiosyncratic factors on the LGD are traceable and consistent with our expectations as well as with the results of previous studies. Nevertheless, we note that the relevant idiosyncratic drivers of the LGD as well as their impact on the LGD differ markedly between the two companies and in some cases also between the used estimation method. This observation is particularly interesting with regard to the partly differing results on some factors of prior studies. Therefore, we conclude that the different results mentioned in the literature can possibly be attributed to the used data or applied estimation method. In this context, it should be particularly emphasized that the addition of the macroeconomic factors only marginally influences the significance and the effect of the idiosyncratic factors on the LGD.

In line with our expectations, object characteristics have substantial influence on the LGD. In particular, we can confirm the object type of the leased asset as one of the key drivers of the LGD, observing low LGDs for vehicles and machinery respectively high LGDs for ICT facilities. In addition, the original value of the leased asset is also identified as a key driver of the LGD.

In terms of contract characteristics, we find that some of these factors are also important drivers of the LGD. With regard to the contract type we observe, as expected, lower LGDs for hire-purchase contracts, but surprisingly the distinction between full pay-out lease contracts and partial amortization contracts has no significant effect on the LGD. Furthermore, we have to pay particular attention to the results concerning the interest rate implicit in the lease. For company B we observe the assumed relationship that the LGD increases with an increasing interest rate implicit in the lease. Interestingly, for company A the interest rate implicit in the lease is significant only at execution of the contract and contrary to our expectations the respective coefficients indicate decreasing LGDs with an increasing interest rate implicit in the lease. One possible reason might be in turn that a lessor monitors contracts with a high interest rate implicit in the lease more closely, because the high interest rate implicit in the lease is an indicator for a riskier lessee. However, different to the other significant variables, the sign of the corresponding coefficient changes for the interest rate implicit in the lease in the models including the macroeconomic factors. The reason might be that in particular the macroeconomic variable which represents the level of interest significantly influences the LGD in these models. We observe that a high level of interest rates implies lower LGDs. Therefore, in the model without the macroeconomic factors, the decrease of the LGD with an increase of the interest rate implicit in the lease may not be the result of a better monitoring of a risky lessee, but rather an indicator that contracts which are signed despite a high level of interest exhibit lower LGDs. If both the level of interest and the interest rate implicit in the lease are considered, we observe lower LGDs with a high level of interest and on the other hand, as expected, decreasing LGDs with an increasing interest rate implicit in the lease. Referring to the other contract characteristics our results confirm that both the existence of buy-back guarantees as well as the existence of additional collateral may significantly reduce the LGD. The latter result applies particularly to company A and supports the findings obtained in Section 3.2.2. Moreover, being the first study to consider the information whether the lessee has a purchase option, we observe slightly higher LGDs for contracts with a purchase option.

In general, customer characteristics have only little effect on the LGD. Interestingly, however, our results indicate that the existence of a longterm contractual relationship does not necessarily lead to lower LGDs, as stated by some prior studies on loans.

In turn, the additional information we obtain at default of the contract contain significant drivers of the LGD. We identify the ratio of the EAD to the original value of the leased asset as the most important driver of the LGD in this category. In line with our expectations we consistently observe a significant increasing LGD when the ratio of EAD to loan size increases. Additionally, at least if considered nonlinearly, the EAD and the relative contract age also effect the LGD. Concerning the first time studied information whether the leasing company has transferred the monitoring of the lease to another company, we do not find a remarkable relationship to the LGD.

In conclusion, our results enable us to identify some of the studied idiosyncratic factors as common key drivers of the LGD. Nevertheless, our findings emphasize that the relevant drivers of the LGD and their impact on the LGD depend in particular on the investigated company and partially also on the used estimation method. With this result, we provide in particular an explanation for the partly inconsistent results in the literature.

Firstly, it should be noted that, e. g., the ratio of rent prepayments to the original value of the leased asset and the ratio of the calculated residual value to the EAD are relevant drivers only for company B, whereas the LGD for company A crucially depends on whether additional collateral is available. Moreover, we observe several times that an idiosyncratic factor is a driver of the LGD for both companies, but its influence on the LGD is different for both lessors. We already addressed this issue regarding the interest rate implicit in the lease and the information whether a contract is a subsequent contract, and we note similar, e. g. concerning the original value of the leased asset, the EAD and the relative contract age. For instance, the LGD of company A increases with an increasing EAD, while the opposite holds for company B.

Further, comparing explicitly the outcomes for the linear regression and the nonlinear regression spline models, it is noteworthy that some of the continuous variables are only significant if considered nonlinearly. This applies, e. g., to the EAD, the relative contract age, the term to maturity and partly also to the original value of the leased asset. On the other hand, we state that for company A at default of the contract, e. g., the information whether a contract is a subsequent contract is only significant if the calculation of the LGD is performed using the linear regression.

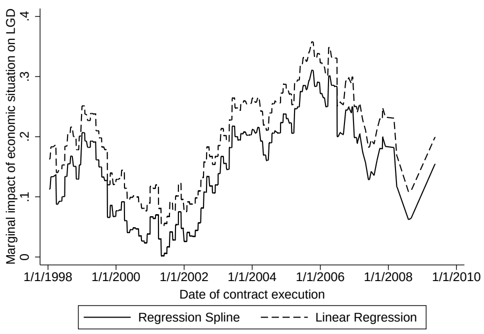
Summarized, our results supply an explanation for the partly inconsistent results regarding some factors in the literature. In view of our findings, it is absolutely essential that each company calibrates its applied LGD estimation method individually on their specific dataset.

3.4.2 Analysis of the Macroeconomic Factors

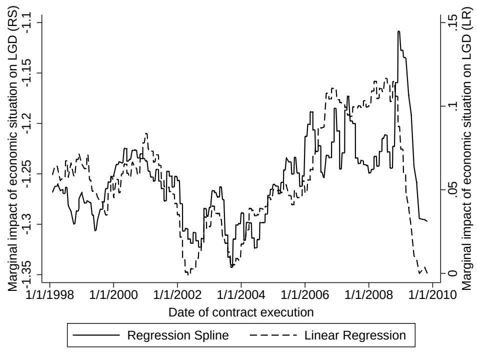
With regard to the analysis of the effect of the macroeconomic factors on the LGD, it should be taken into account that an independent interpretation of each factor could be misleading. Only the joint examination of the significant macroeconomic factors models the influence of the economic situation on the LGD as a whole. The significant macroeconomic factors are partially different for the nonlinear regression spline model and the linear regression model. However, the overall influence of the macroeconomic factors, presented in Figure 3.3, is not driven by the respective model.

At execution of the contracts, we observe a clear link between the economic situation and the LGD, which is even more distinctive for company A. In particular, this means that we can connect changes of the LGD to special events such as the Dotcom crisis and the financial crisis. At default of the contracts, however, we recognize only for company A a slight link between the economic situation and the LGD, but not for company B.

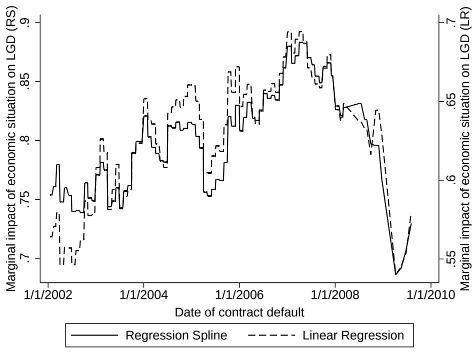
Beginning with the detailed analysis at execution of the contract, we notice that for company A for both models the significant variables are IFO-EXPECTATION, ASSETINVESTMENT and INTERESTRATELEVEL. These variables are considered linearly in each case. In detail, a positive sign of each respective coefficient indicates that LGDs are higher during periods with high business expectations and growth of gross fixed asset investments while a negative sign of the corresponding coefficient states that the LGD decreases if the level of interest rate increases. The combined effect of these variables on the LGD, displayed in Figure 3.3a, clearly outlines that contracts executed during periods of relatively good economic conditions realize higher LGDs, whereas the LGDs turn out to be comparatively low during the Dotcom crisis and the financial crisis. Based on the observation period, periods of good economic conditions are the years before the Dotcom crisis in 2000 and particularly before the financial crisis beginning at the end of 2007. Conse-



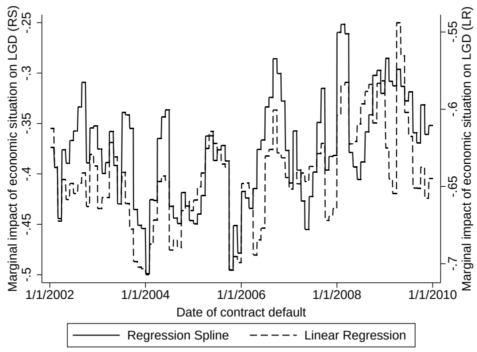
(a) Company A: At execution



(b) Company B: At execution



(c) Company A: At default



(d) Company B: At default

Figure 3.3: Influence of the economic situation represented by the aggregate impact of the significant macroeconomic factors in the respective model at execution and default of the contracts for both companies.

quently the increasing realized LGDs observed in 1999 and from 2004 to 2007 (see Table 3.2) can be traced back to the economic conditions at that time. One potential explanation for the possibly unexpected observed relationship could be that in times of economic upturn, particularly the manufacturing sector is enticed to invest because their capacities are temporarily fully stretched. In contrast, investments which are carried out in a weak economy may suggest a stable business environment with stable order situation and therefore stable inflows. This has a positive effect on the LGD in case of a default, as the lessor can expect payments by the lessee during the workout. Obviously, towards the end of the observation period we have to consider the lower number of observations, nevertheless we are convinced that our findings for the financial crisis are reliable as they correspond to the results for the Dotcom crisis. Basically, our observed relationship between the LGD and the economic situation corresponds in particular to the procyclical character of the lending activity in the banking sector. As, e. g., Borio et al. (2001) and Mérõ et al. (2002) point out, typically risk is underestimated in times of good economic conditions and overestimated during recessions. During an economic upturn also high-risk loans are granted, their potential losses, however, materialize only in the following downturn. Besides, we observe in general that the macroeconomic factors explain variations of 30 % of the LGD over the time, which represents a large part of the observed variability of the realized LGDs.

For company B at execution of the contract the joint influence of the macroeconomic factors on the LGD, displayed in Figure 3.3b, has to some extent the same structure as for company A. We recognize that the fluctuations of the LGD being explained by the macroeconomic factors are inferior for company B, but this corresponds to the changes of the realized LGDs over time for this company, outlined in Section 3.2.1. In contrast to company A, for the time of the Dotcom crisis no clear link between the economic situation and the LGD can be established. However, again corresponding to the procyclical character of the lending activity in the banking sector, in turn we observe lower LGDs from the beginning of 2002 and significantly increasing LGDs during the subsequent economic upturn, whereby the increase of the LGDs starts a little later than for company A. In the linear regression model the economic situation is represented by the variables IFO-INDEX and IFO-EXPECTATION. The sign of the respective coefficient suggests higher LGDs during times of good business climate, but otherwise lower LGDs in periods of high business expectations. In the nonlinear regression spline model also the INTERESTRATELEVEL is significant. Furthermore the influences of the variables are more sophisticated due to their nonlinear consideration. We note that for both business climate as well as for business expectations high or low expressions imply higher LGDs. The same applies to the level of interest rates, in particular the lowest LGDs are experienced at interests around 4 %, which roughly corresponds to the average over the observation period. This generally means the lowest LGDs are realized in quiet periods, which are characterized by neither boom nor downturn. This applies, e. g., to the year 2004.

Continuing with the analysis at default of the contract, we note that partly other macroeconomic factors are significant than at execution of the contract. For company A the linearly considered variables GDP and IFO-INDEX are significant in both models. The respective signs of the corresponding coefficients imply that on the one hand the LGD decreases with an increasing growth rate of the gross domestic product, but on the other hand that LGDs are higher during times of good business situation. In the linear regression model the variable INTERESTRATELEVEL is also significant, the corresponding coefficient indicates that LGDs increase with an increasing level of interest rate. Considering the impact of the economic situation on the LGD as a whole, displayed in Figure 3.3c, we find that in contrast to the execution of the contract, only a portion of the variation of the realized LGDs can be explained by the macroeconomic factors. Between early 2002 and late 2005 the development of the economic situation suggests generally rising LGDs. It should be noted that the values of the individual macroeconomic factors show only small fluctuations in this period (see Figure 3.2). In particular the increased realized LGDs in 2005 cannot be traced back directly to changes in the economic situation. What we do observe is that increasing realized LGDs in 2007 and 2008, within the context of the financial crisis, can be at least partially explained by the macroeconomic factors. Of course we have to consider again the lower number of observations towards the end of the observation period. However, the observed connection is commonly comprehensible, because a typical feature of crises is a decrease in the order situation, which complicates the repayment of the outstanding debt. This seems to be particularly evident for company A (see Section 3.2.2).

For company B at default of the contract in both models four macroeconomic factors are significant. In the linear regression model the LGD increases both with an increasing growth rate of the gross domestic product and an improving business climate. On the contrary, with an increasing growth of gross fixed asset investments and improving business expectations, the LGD decreases. In the nonlinear regression spline model the variables GDP, IFO-EXPECTATION and ASSETINVESTMENT are significant, too. The relationship between these variables and the LGD is essentially the same as in the linear model, only the influence of the two first-mentioned is more sophisticated because of their nonlinear consideration. Furthermore the level of interest rate is significant in the nonlinear model and we observe again that LGDs are higher during periods of either high and low interests. If we analyze the impact of the economic situation on the LGD as a whole, displayed in Figure 3.3d, we cannot see a clear structure. In particular it is not possible to clearly identify periods of higher or lower LGDs. Quite contrary to company A, for company B the joint influence of the macroeconomic factors on the LGD is extremely volatile over the entire observation period. Thereby, the linear (nonlinear) model reduces fluctuations of the realized LGDs up to 15 % (25 %) to changes in the economic situation. This outcome is unexpected since we observed significant lower fluctuations in the realized LGDs over time (see Table 3.2). Even during the financial crisis both models suggest both particularly high as well as particularly low LGDs, depending on changes in the macroeconomic factors. As can be also seen on basis of the improvement of the estimation accuracy, outlined in Table 3.5, the additional macroeconomic factors are apparently identified as drivers of the LGD. But with regard to the relatively low R^2 it should be considered that the reason might be a minor expressiveness of the idiosyncratic factors. The detailed analysis of the influence of the macroeconomic factors on the LGD outlines that a direct relationship between the LGD and the general economic situation is certaily not apparent.

3.5 Out-of-time Analysis

In this section we analyze and discuss the out-of-time results of our models for estimating the LGD. Analogous to the procedure of the in-sample testing, for both companies we have estimated the LGD with the nonlinear regression spline model and the linear regression. Again, we have performed the estimates once using only idiosyncratic factors and once with additionally taking into account macroeconomic factors. To avoid excessive complexity and possibly preventing potential overfitting which can influence the out-of-time performance negatively, for each model we only consider the explanatory variables that were at least significant at 10% in-sample (see Table 3.6 and Table 3.7).

The out-of-time results support the in-sample findings in many aspects. At execution of the contract both companies yield more accurate LGD estimations by incorporating macroeconomic factors, irrespective which estimation method is used. However, in line with the in-sample findings, we observe differences between the two companies as company A benefits more from the use of the macroeconomic factors. Moreover, at default of the contract we notice that exclusively company A yields more accurate LGD estimations by including macroeconomic factors while the opposite holds for company B. Analogous to the in-sample results, we note that even for company A the benefit of including macroeconomic factors is less pronounced at default of the contract than at its execution.

3.5.1 Results at Execution of the Contract

The out-of-time performance measurements at execution of the contract are presented in Table 3.8. The outcomes show that incorporating macroeconomic factors generally improves the LGD estimation for both companies, irrespective of whether the estimates are obtained by the linear regression or the nonlinear regression spline model.

		Company A					Company B			
		Linear Regression		Regression Spline			Linear Regression		Regression Spline	
	Year	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO	Year	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO
MAE		0.2939	0.2717	0.2932	0.2698		0.3468	0.3461	0.3457	0.3470
RMSE	2005	0.3555	0.3408	0.3555	0.3408	2005	0.4094	0.4089	0.4085	0.4087
REC Area		0.7080	0.7307	0.7020	0.7324		0.6538	0.6545	0.6549	0.6535
MAE		0.2806	0.2679	0.2806	0.2640		0.3407	0.3362	0.3399	0.3412
RMSE	2006	0.3484	0.3369	0.3484	0.3334	2006	0.4032	0.4005	0.4023	0.4033
REC Area		0.7236	0.7238	0.7236	0.7238		0.6599	0.6644	0.6606	0.6594
MAE		0.5195	0.5048	0.5099	0.4962		0.3095	0.2992	0.3089	0.2989
RMSE	2007-2009	0.5804	0.5638	0.5695	0.5562	2007	0.3689	0.3629	0.3688	0.3627
REC Area		0.4877	0.5022	0.4971	0.5110		0.6911	0.7014	0.6916	0.7016
MAE							0.3126	0.3031	0.3088	0.3008
RMSE						2008-2009	0.3726	0.3663	0.3689	0.3641
REC Area							0.6885	0.6891	0.6921	0.7003

Table 3.8: Out-of-time performance measurements at execution of the contract by company and forecast period. The estimates were carried out each with the nonlinear regression spline model and the linear regression model. FIRM represents that only idiosyncratic factors were used as explanatory variables, FIRM+MACRO implies that additionally macroeconomic factors were included, in each case the better result is underlined. REC Area is defined as the area under the regression error characteristic curve, MAE is the mean absolute error, and RMSE is the root mean squared error. For the REC Area higher, for MAE and RMSE lower outcomes are preferable.

Upon closer inspection, the outcomes for company A clearly demonstrate that the respective models using macroeconomic factors yield throughout more accurate LGD estimations for all forecast periods. Consequently, at execution of the contract the out-of-time results for company A confirm the in-sample findings, discussed in Section 3.4, that the additional use of macroeconomic factors leads to better estimations of the LGD. In particular considering the poor predictions for the last forecast period, which includes the financial crisis, one may argue that these results are biased due to the low number of observations. However, it should be borne in mind that the respective training sample already includes information about the Dotcom crisis and according to the in-sample results the relationship between the LGD and the economic situation is similar for the Dotcom crisis and the financial crisis. Consequently, although the estimation accuracy is poor for this forecast period, improved predictions by using macroeconomic factor are quite comprehensible. Basically, by assuming an identical relationship between the LGD and the explanatory variables over the entire observation period, comparatively poor estimates for the final forecast period are indeed surprising, as most training data are available for this period. On the other hand, it should be considered that the few estimations carried out for this period focus on the year 2007. The observations in 2007 exhibit significantly higher LGDs than those of the previous years which are used for model fitting (see Table 3.2). On average, basically all models underestimate the LGD in particular for the period 2007-2009. However, this underestimation of the LGD is minor if macroeconomic factors are considered.

With regard to the outcomes for company B we recognize that the inclusion of the macroeconomic factors is overall less beneficial than for company A. The linear model achieves a higher estimation accuracy in all forecast periods by incorporating the macroeconomic factors, however, a distinct advantage over the respective model without macroeconomic factors can only be obtained in the last two periods. In addition, the nonlinear regression spline model only benefits by the inclusion of the macroeconomic factors from the year 2007 onwards. One reason for the use of the macroeconomic factors being clearly beneficial only in the last two forecast periods might be the requirement of more training data for an adequate model fitting, because, as stated in Section 3.4, the models for company B feature a significant lower coefficient of determination than those for company A. Still, altogether the out-of-time results at execution of the contract confirm also for company B largely the in-sample observed benefit of considering macroeconomic factors for estimating the LGD.

Analyzing the performance measurements for both companies, we observe that the LGD estimations for company B are less accurate in the early forecast periods. Nevertheless, in contrast to company A the quality of the estimations remains stable in the later periods. This can be explained, among other reasons, by the fact that for company B the level of the LGD increased only moderately in the last years of the observation period.

Besides, we have seen that the nonlinear regression spline model achieves a higher estimation accuracy than the linear regression in-sample. Out-of-time we again observe this tendency for company A, however, it is mainly reflected in the models that incorporate macroeconomic factors. Furthermore, for company B, out-of-time a slight advantage of the nonlinear model over the linear model only exists from the forecast period 2007 onwards. One possible explanation might be in turn the rising number of training data over time. As Hartmann-Wendels et al. (2014) have shown, more complex models, as the nonlinear regression spline model, typically yield better results in-sample. However, those complex models might require more training data to achieve an adequate model fitting, which is necessary to perform accurate out-of-time.

3.5.2 Results at Default of the Contract

The out-of-time performance measurements at default of the contract are shown in Table 3.9. In contrast to the execution of the contract the results differ significantly between both companies. While for company A both methods again yield more accurate LGD estimations by incorporating the macroeconomic factors, the opposite holds for company B.

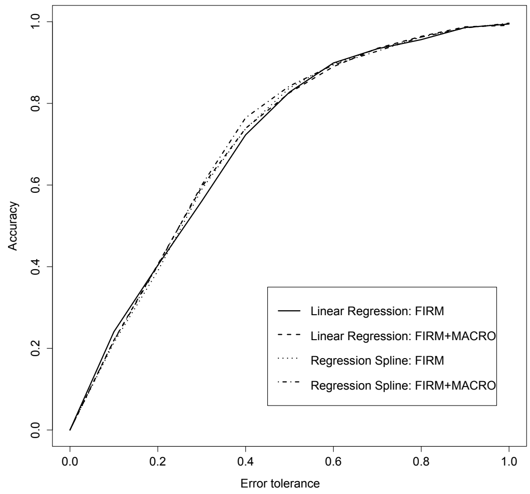
A closer study of the outcomes for company A shows that particularly the LGD estimations performed by the nonlinear regression spline model benefit from the inclusion of macroeconomic factors for all forecast periods. The predictions of the linear regression improve consistently less. In particular, for the linear regression the benefit of using macroeconomic factors decreases towards the end of the observation period. This is a possible indication that a linear estimation model for the LGD can hardly consider the serious changes of the macroeconomic factors that occurred in the wake of the financial crisis (see Section 3.2.4). In this situation it seems to be more reasonable to extrapolate the training data with a nonlinear model.

Comparing the out-of-time results for company A at execution and default of the contract, we recognize that the improvement of the estimation accuracy by in-

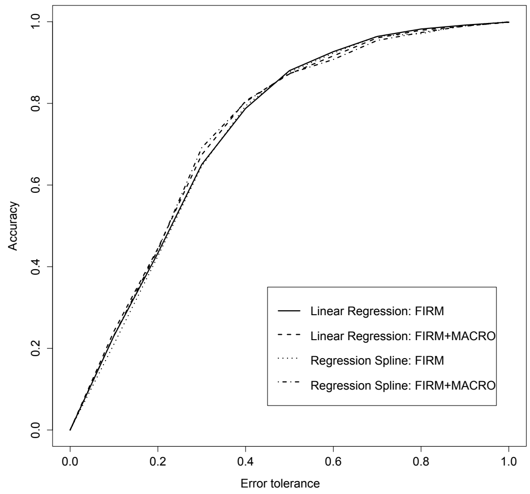
		Company A					Company B			
		Linear Regression		Regression Spline			Linear Regression		Regression Spline	
	Year	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO	Year	FIRM	FIRM+MACRO	FIRM	FIRM+MACRO
MAE		0.2958	0.2941	0.2965	0.2893		0.3064	0.3120	0.3018	0.2985
RMSE	2005	0.3731	0.3708	0.3701	0.3648	2005	0.3680	0.3738	0.3661	0.3674
REC Area		0.7058	0.7076	0.7047	0.7119		0.6711	0.6715	0.6973	0.7021
MAE		0.2634	0.2603	0.2667	0.2617		0.3179	0.3212	0.3167	0.3209
RMSE	2006	0.3291	0.3294	0.3328	0.3328	2006	0.3770	0.3827	0.3778	0.3784
REC Area		0.7378	0.7409	0.7179	0.7396		0.6826	0.6794	0.6846	0.6802
MAE		0.3321	0.3319	0.3334	0.3229		0.3634	0.3641	0.3646	0.3647
RMSE	2007-2009	0.4105	0.4112	0.4285	0.4208	2007	0.4352	0.4392	0.4573	0.4567
REC Area		0.6702	0.6706	0.6733	0.6836		0.6404	0.6395	0.6418	0.6414
MAE							0.3410	0.3444	0.3411	0.3558
RMSE						2008	0.4073	0.4109	0.4085	0.4258
REC Area							0.6595	0.6561	0.6595	0.6453
MAE							0.3213	0.3241	0.3184	0.3649
RMSE						2009	0.3838	0.3893	0.3825	0.4396
REC Area							0.6792	0.6764	0.6820	0.6365

Table 3.9: Out-of-time performance measurements at default of the contract by company and forecast period. The estimates were carried out each with the nonlinear regression spline model and the linear regression model. FIRM represents that only idiosyncratic factors were used as explanatory variables, FIRM+MACRO implies that additionally macroeconomic factors were included, in each case the better result is underlined. REC Area is defined as the area under the regression error characteristic curve, MAE is the mean absolute error, and RMSE is the root mean squared error. For the REC Area higher, for MAE and RMSE lower outcomes are preferable.

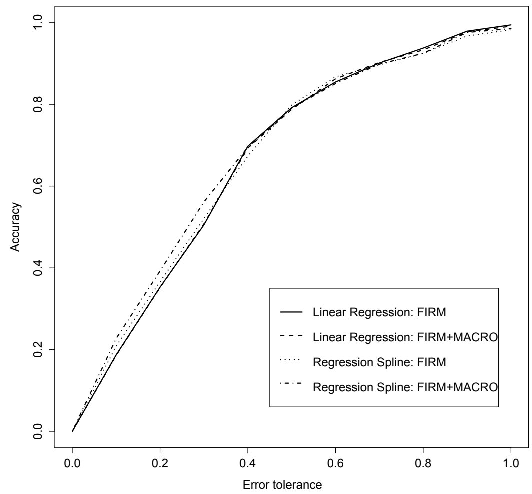
corporating macroeconomic factors is significantly more pronounced at execution of the contract. This observation is in line with the in-sample results. Moreover, it is noticeable that at default of the contract the improvement of the estimation accuracy by the inclusion of macroeconomic factors is mainly reflected by the outcomes of the MAE and the REC Area and only partially by the outcomes of the RMSE. One explanation for this observation is provided by the respective REC curves, displayed in Figure 3.4. In particular the shape of the REC curves for the forecast periods 2006 and 2007-2009, presented in Figure 3.4b and Figure 3.4c, shows that the proportion of observations with minor deviations between the realized and the predicted LGD is higher if macroeconomic factors are incorporated. This increasing proportion of observations with minor deviations between the realized and the predicted LGD is characteristic if the LGD estimation benefits from the use of macroeconomic factors. This feature is also observed at execution of the contract for both companies. However, in contrast to the situation at execution of the contract, the REC curves of the models that use macroeconomic factors undercut the respective REC curves of the models without macroeconomic factors



(a) Forecast period: 2005



(b) Forecast period: 2006



(c) Forecast period: 2007-2009

Figure 3.4: Regression error characteristic (REC) curves of the out-of-time loss given default (LGD) estimations at default of the contract for company A by forecast period. The estimates were carried out each with the nonlinear regression spline model and the linear regression model. FIRM represents that only idiosyncratic factors were used as explanatory variables, FIRM+MACRO implies that additionally macroeconomic factors were included.

with increasing error tolerance at default of the contract for company A. This implies a higher proportion of observations with substantial deviations between the realized and the predicted LGD for the models that incorporate macroeconomic factors. For this reason, the difference between the models with and without macroeconomic factors is less in terms of the RMSE compared to the MAE, because the RMSE penalizes in particular large deviations between the realized and the predicted LGD.

Analogously to the out-of-time results at execution of the contract, for company A at default of the contract the nonlinear regression spline model primary achieves a higher estimation accuracy than the linear regression if macroeconomic factors are considered. This might be another indication that the relationship between the LGD and the macroeconomic explanatory variables is nonlinear for company A. In-sample, however, the nonlinear model consistently performs better than the respective linear model.

For company B the out-of-time outcomes at default of the contract show a completely different picture than for company A. In contrast to the out-of-time estimation at execution of the contract, the consideration of macroeconomic factors does not lead to more accurate estimations of the LGD at default. In particular, it should be highlighted that this finding holds for all forecast periods. Consequently, the different results for both companies are consistent and are not the outcome of, e. g., biased data at the end of the observation period. The linear regression yields consistently better predictions without incorporating the macroeconomic factors. For the nonlinear regression spline model the inclusion of the macroeconomic factors worsens the LGD estimation particularly towards the end of the observation period. Even a simple historical average provides more accurate estimations than the regression spline model which considers macroeconomic factors for these forecast periods.

At first glance, for company B the out-of-time results seem to contradict the in-sample results, because the latter also indicate an improvement of the estimation accuracy by considering macroeconomic factors at default of the contract. An explanation is provided by the analysis of the effect of the macroeconomic factors on the LGD, discussed in Section 3.4.2. While a clear link between the economic situation and the LGD can be observed for both companies at execution of the contract, this applies only for company A at default of the contract. In contrast, for company B the influence of the macroeconomic factors on the LGD is extremely volatile at default of the contract and does not follow a clear structure. In particular, it is not possible to connect changes of the LGD directly to the impact of the financial crisis. In other words, although in-sample a higher estimation accuracy is achieved by the inclusion of the macroeconomic factors, a direct relationship between the LGD and the economic situation cannot be established. Consequently, it can be assumed that the relationship between the LGD and the macroeconomic explanatory variables is not identical over the entire observation period. This implies that extrapolating the training data is hardly possible, especially towards the end of the observation period when serious changes of the macroeconomic factors occurred in the wake of the financial crisis. In this regard, it should also be borne in mind that the serious changes of the macroeconomic factors are only reflected by moderate changes of the level of the LGD (see Table 3.2). The negative influence of the inclusion of the macroeconomic factors on the estimation accuracy especially towards the end of the observation period is more pronounced for the nonlinear regression spline model, because this model achieves a higher adaption to the training data and it can be assumed that it attaches more importance to the serious changes of the macroeconomic factors.

3.5.3 Interpretation and Implications

The out-of-time results, which are of major importance from a practical point of view, suggest that the benefit of considering macroeconomic factors for estimating the LGD depends on whether the estimates are performed at execution or default of the contract. In particular, we observe significant differences between the results of the two companies. We find that generally both companies yield more accurate estimates by incorporating macroeconomic factors at execution of the contract, although the advantage is more pronounced for company A. In contrast, at default of the contract exclusively company A benefits by the inclusion of the macroeconomic factors. In this section we point out differences between the companies and analyze why the economic situation affects the LGD more for company A. In addition, we deduce practical implications from our results. On the one hand we address the organization of the workout process and on the other hand we state crucial aspects that have to be taken into account in order to calculate a downturn LGD.

The improvement of the estimation accuracy at execution of the contract by incorporating macroeconomic factors apparently shows that the level of the LGD depends significantly on the time of the investment. It should be noted that the customer makes the investment decision mainly independent from the leasing company, as investment decisions are usually very complex and depend on several aspects. Referring to the in-sample results, set out in Section 3.4.2, we observe increasing LGDs for leases that are signed during an economic upturn. This relationship is particularly evident for the years before the financial crisis. As mentioned earlier, in times of economic upturn particularly the manufacturing sector is enticed to invest because their capacities are temporarily fully stretched. However, if the economic situation deteriorates, the order situation regresses. The resulting lower inflows might in particular not cover the lease payments of the investment made during the preceding economic upturn. Furthermore, the in-sample results indicate that the leases which are signed in times of a weak economy generally realize relatively low LGDs, even though the highest LGDs for company B are observed in 2009. Strictly speaking, the results point out that the level of the LGD during an economic downturn remains at least stable or even decreases. As stated above, investments which are carried out in a weak economy may suggest a stable business environment with stable order situation and therefore stable inflows. Consequently, it is likely that the lessee obtains inflows also in the event of default and hence is able to repay at least a portion of the debt to the lessor.

Payments by the lessee to the lessor are particularly important for company A. As shown in Table 3.3, for company A, the payments by the lessee represent a substantially higher proportion of the total payments the lessor receives during the workout than for company B. Possibly for this reason, the dependence of the LGD on the economy is more pronounced for company A.

Considering the outcomes at default of the contract, the inclusion of the macroeconomic factors leads to more accurate LGD estimations only for company A. As both companies are subjected to the same economic conditions, this indicates structural differences between the companies that might be related to the organization of the workout process. As shown in Table 3.1, company A exhibits on average about 15% lower LGDs than company B. Moreover, as mentioned earlier, for company A the payments by the lessee represent generally a higher proportion of the total payments the lessor receives during the workout. Combined with the finding that the existence of collateral is more important for company A, this indicates that lessor A monitores the workout more actively.

With regard to the composition of the recovery, it has been outlined in Section 3.2.2 that the proportion of the revenue from disposing the leased asset fluctuates only slightly over the observation period for both companies. This suggests that due to their knowledge of the secondary markets, both lessors are able to achieve reasonable revenues also during crises. In contrast, with respect to the evolution of the lessee payments, we observe significant differences between the two companies. For company B the proportion of the lessee payments is comparatively limited, but it is largely constant over time. For this reason we note only a slight increase of the LGD during the prolonged financial crisis for company B. For company A, however, we observe that the proportion of the lessee payments is more volatile and obviously depends considerably on the economic situation. While these payments represent a significant proportion of the total recovery in general, the share decreases significantly during the financial crisis.

The generally high proportion of lessee payments might be an outcome of an active workout management by the lessor, e. g., by calculating a new payment plan in the event of default of the lessee. However, as outlined, if the economic situation deteriorates, typically the order situation regresses. Consequently, despite an active management of the workout, during an economic downturn no surpassing lessee payments can be achieved. This explanation particularly corresponds to the in-sample findings in Section 3.4.2 that the LGD of company A increases in the wake of the financial crisis. In summary, it can be noted that the active workout management generally helps company A to achieve on average lower LGDs than company B, but it also contributes to greater dependency of the LGD on the economic situation.

Our out-of-time results show that the consideration of macroeconomic factors does not necessarily lead to more accurate LGD estimations. In fact, as the outcomes for company B at default of the contract show, it is likewise possible that the estimates become inaccurate. According to the in-sample results, this is because no clear relationship between the LGD and the economic situation exists. With regard to the development of an estimation model for the LGD this implies that leasing companies have to analyze individually to which extent a relationship between the LGD and the economic situation exists. Our findings suggest that this relationship depends in particular on the organization of the workout process. An active management of the workout seems to contribute to lower LGDs in general, however the level of the LGD depends more on the economic situation.

Moreover, our findings are of crucial importance with regard to the calculation of a downturn LGD. Several studies, e. g., Hartmann-Wendels and Honal (2010), Bellotti and Crook (2012), and Leow et al. (2014), have used macroeconomic stress tests to calculate a downturn LGD. Here, the LGD is modeled depending on macroeconomic factors. The downturn LGD is then obtained by the use of unfavorable realizations of the macroeconomic factors, e. g., the worst observed realization of the past 25 years. As Hartmann-Wendels and Honal (2010) cite, a prerequisite for this approach is a stable relationship between the LGD and the macroeconomic factors, which in particular consists during a downturn period. This requirement is obviously not satisfied for company B at default of the contract. While the in-sample results already point out that no stable relationship between the LGD and the macroeconomic factors exists, out-of-time we notice that the estimates are worse in particular during the financial crisis. One reason is that the macroeconomic factors vary strongly during the financial crisis, but these variations cannot be transfered equally to the level of the LGD. Instead of using macroeconomic stress tests to calculate a downturn LGD, for company B at default of the contract it is more appropriate to utilize a worst case scenario approach as mentioned by Chalupka and Kopecsni (2009). Before the financial crisis began in 2007, the contracts defaulted in 2002 realized the highest average LGD, which was about 4% higher than the overall average LGD up to 2007. Accordingly a downturn LGD can be calculated by adding 4% to the LGD estimates obtained by a model which exclusively uses idiosyncratic explanatory variables.

3.6 Conclusion

In this study we have analyzed the influence of macroeconomic factors on the LGD of defaulted leasing contracts while simultaneously taking into account idiosyncratic factors. The data we have used for our analysis are provided by two German leasing companies. With defaults between 2002 and 2009, the observation period covers a wide range of economic conditions including the impact of the recent financial crisis. We have estimated the LGD at execution and default of the contract both with a linear regression and a nonlinear regression spline model. To validate our results we have performed in-sample and out-of-time testing. The latter is of major importance from a practical point of view and is mandatory to meet the regulatory requirements.

We find that the relevance of macroeconomic factors for the LGD estimation depends in particular on whether the estimates are performed at execution or default of the contract. At execution of the contract, for both companies a clear link between the economic situation and the LGD can be established, irrespective which estimation method is used. The in-sample results show that changes in the level of the LGD are directly connected with special events, such as the financial crisis. Consistently, out-of-time the respective models which use macroeconomic factors yield throughout more accurate LGD estimations.

At default of the contract, the in-sample outcomes indicate only for company A a slight link between the LGD and the economic environment. This finding is confirmed by the results of our out-of-time tests. We observe that out-of-time the consideration of macroeconomic factors generally leads to more accurate LGD estimations for company A, whereas the opposite holds for company B. In particular, we state that the respective results are again not driven by the applied estimation method.

We find that the general dependency between the LGD and the macroeconomic factors at execution of the contract is only limitedly connected to the leasing company. We rather observe that the LGD is directly linked to the time of investment. In contrast, at default of the contract the influence of the economic situation on the LGD depends essentially on the organization of the workout process and thus differs between the leasing companies. Our findings are of crucial importance, especially with regard to the calculation of a downturn LGD. At least the results of macroeconomic stress tests could be misleading, because a stable relationship between the LGD and macroeconomic factors does not necessarily exist.

In addition, we have studied the relationship between the LGD and various idiosyncratic factors to identify the key drivers of the LGD. We find that the relevant drivers of the LGD and their impact on the LGD depend in particular on the investigated company and partially also on the used estimation method.

Nevertheless, some idiosyncratic factors can be identified as key drivers of the LGD. In general, we observe that the LGD depends more on object characteristics and contract characteristics than on customer characteristics. Moreover, the additional information at default of the contract contribute to the explanation of the LGD. The most significant driver of the LGD is the object type of the leased asset. We observe low LGDs especially for vehicles, whereas ICT facilities feature high LGDs. Additionally, the LGD depends on the original value of the leased asset. Our results further show that hire-purchase contracts generally realize lower LGDs. Besides, we find that an increasing ratio of the EAD to the original value of the leased asset throughout increases the LGD. Beyond this, the key drivers of the LGD differ markedly between the analyzed companies. Among others, the LGD of company A depends substantially on the existence of additional collateral, while the presence of buy-back agreements significantly reduces the LGD of company B.

We therefore conclude that the different results of previous studies on some factors are in particular the result of differences in the used datasets and methods. Contrary to our work, most of these studies used only data from one company or pooled data from different companies. Although there exist some common idiosyncratic key drivers of the LGD, our results clearly outline that the calibration of an estimation method has to be done by each company individually.

However, our study also provides evidence whether the consideration of nonlinear relationships between the LGD and the explanatory variables contributes to more accurate LGD estimations. In-sample we observe that the nonlinear regression spline model consistently achieves a higher estimation accuracy than the linear regression. Furthermore, the out-of-time results show that the detailed mapping of the dependencies between the LGD and the explanatory variables by the nonlinear regression spline model may also improve the forecast of the LGD. However, an advantage of the nonlinear model is not guaranteed, in particular to achieve more accurate LGD forecasts than the linear model, the nonlinear model requires at least a definite relationship between the LGD and the explanatory variables and additionally sufficient training data.

3.7 Appendix

	Variable definitions
Idiosyncratic factors	
Object characteristics	
VEHICLE	A dummy variable equal to one if the leased asset is a vehicle
MACHINERY	A dummy variable equal to one if the leased asset is a machine
ICT	A dummy variable equal to one if the leased asset belongs to the ICT area
EQUIPMENT	A dummy variable equal to one if the leased asset is an item of equipment
ORIGINALVALUE	The original value of the leased asset in euros
USED	A dummy variable equal to one if the procured leased asset has been used already
Contract characteristics	
FULLPAYOUT	A dummy variable equal to one if the contract type is a full pay-out lease contract
PARTIALAMORTISATION	A dummy variable equal to one if the contract type is a partial amortisation contract
HIREPURCHASE	A dummy variable equal to one if the contract type is a hire-purchase contract
PURCHASEOPTION	A dummy variable equal to one if the customer has a contractually agreed purchase option
INTEREST	The interest rate of the contract in percent
MATURITY	The maturity of the contract in month
PRETOVALUE	The prepayment rent (possible zero) divided by the original value of the leased asset
RESIDUALTOVALUE	The calculated residual value of the leased asset (possible zero) divided by its original value
BUYBACK	A dummy variable equal to one if the supplier commits to buy back the leased object in case of a contract disturbance
COLLATERAL	A dummy variable equal to one if the contract is secured by an additional collateral
Customer characteristics	
RETAIL	A dummy variable equal to one if the customer is allocated to the retail business
EXTENSION	A dummy variable equal to one if the contract is a subsequent contract
Additional information at default	
EAD	The outstanding exposure at default in euros
EADTOVALUE	The outstanding exposure at default divided by the original value of the leased asset
CONTRACTAGE	The term between execution and default of the contract divided by the term to maturity of the contract
SOLD	A dummy variable equal to one if the monitoring of the contract was ceded to another company
RESIDUALTOEAD	The calculated residual value of the leased asset (possible zero) divided by the outstanding exposure at default
Macroeconomic factors	
GDP	Growth rate of the gross domestic product in comparison to the preceding quarter - seasonally adjusted values using Census-X12-Arima
IFO-INDEX	Monthly collected index business climate of the Ifo Institute for Economic Research
IFO-EXPECTATION	Monthly collected index business expectations of the Ifo Institute for Economic Research
ASSETINVESTMENT	Growth rate of the gross fixed asset investments in comparison to the preceding quarter - seasonally adjusted values using Census-X12-Arima
INTERESTRATELEVEL	Monthly average of the yield to maturity for domestic bearer bonds

4 Loss Given Default-Adjusted Workout Processes for Leases

4.1 Introduction

Credit risk modeling is an essential assignment of risk management in financial institutions. One of the major drivers of credit risk is the loss given default (LGD). The knowledge of potential losses is crucial for an efficient allocation of regulatory and economic capital and also for credit risk pricing. Pursuant to Article 107 (1) of the Capital Requirement Regulation (CRR), financial institutions shall apply either the Standardised Approach or the Internal Ratings Based Approach (IRBA) in order to calculate their regulatory capital requirements for credit risk. When implementing the advanced IRBA, it is mandatory to develop internal models for estimating the probability of default (PD), exposure at default (EAD), and LGD. One of the main objectives of the IRBA is to achieve risk-adjusted capital requirements (see Basel Committee on Banking Supervision (2003)). Accurate forecasts of PD, EAD, and LGD may result in competitive advantages for the applying financial institution in general, as is indicated by Gürtler and Hibbeln (2013).

While the procedure for calculating the PD might be almost identical for loans and leases, models for estimating the LGD should consider specific characteristics of leasing contracts. The analysis of the leasing business is particularly important considering that a high amount of externally financed investments and total investments in European economies are financed by leasing, e. g. 50% and 25%, respectively, in 2015 in Germany. In contrast to loans, the collateralization of a lease by its leased asset is obligatory. In particular, being the legal owner of the leased asset, the lessor can retain any recovered value of the leased asset's disposal. Thus, unlike in the case of loans, the lessor has legal access to this additional source of payments in the event that a contract defaults. Eisfeldt and Rampini (2009) argue that the main benefit of leasing is that repossession of a leased asset is easier than foreclosure on the collateral of a secured loan. During the workout process of a defaulted loan, the lender receives payments exclusively from the debtor and the liquidation of collateral. These incomes also occur during the workout process of leases. Consequently, considering the additional incomes from disposing of the leased asset, the cash flows of the leasing workout process consist of two parts. One part comprises the asset-related cash flows, the other part comprises all remaining cash flows. Han and Jang (2013), Töws (2014), and Frontczak and Rostek (2015) argue that the level of LGD crucially depends on the actions taken during the workout process. Hence, the specific features of the workout process of leases should be taken into account when estimating LGD.

In the recent literature various advanced approaches for estimating the LGD have been analyzed. Bastos (2010), Hartmann-Wendels et al. (2014), and Yao et al. (2015) find that complex models are able to generate robust and precise LGD predictions in principle. Nevertheless, either for loans or leases, no single estimation approach has been established yet. Remarkably, the majority of the estimation approaches introduced so far has in common that the LGD is regarded as a holistic measure of risk. With regard to the LGD of loans, such an approach is reasonable. However, according to the specific characteristics of leasing contracts, the LGD of leases typically consists of cash flows from two distinct sources. Thus, a holistic approach to estimate the LGD of leases might be inappropriate.

Therefore, we present a new approach to forecasting leasing LGDs. In our study, we consider the specific characteristics of leases and, consequently, we suggest an economically motivated separation of the LGD into an asset-related part and a miscellaneous part. The required information about the breakdown of the cash flows is compulsory for institutes using the IRBA according to Article 181 CRR. Coming from different payment sources, both parts should be driven by different factors.

While our approach is explicitly designed to estimate leasing LGDs, the basic idea can be adjusted in general to estimate the LGD of other instruments such as collateralized loans and in particular mortgages. The only requirement is that the considered instruments include cash flows obtained during the workout process from distinct payment sources.

In the course of this paper, we describe the development of a multi-step estimation model, which is built upon the economic composition of the LGD of leasing contracts. Estimating the asset-related and miscellaneous parts, we derive an estimation of the overall LGD. Our easily traceable model results in a significant advantage in terms of estimation accuracy.

Moreover, the estimated asset-related and miscellaneous LGD can be used to support decisions concerning the accomplishment of the workout process. In fact, the separation of LGD has extensive practical implications for handling a defaulted contract's workout process. The derived shares of LGD are indicators for the success of both the asset's disposal and the effort of collecting further payments. Consequently, we find that our inferred suggestions for the actions to be taken by the lessor during the workout process lead to significant improvements in the resulting LGD value of the respective contracts.

For our study, we use a real-life dataset provided by a major German lessor. The data is of high quality with regard to details, which is particularly important in our approach. We compare the performance of our procedure to that of traditional holistic methods for LGD estimation carried out, e. g., by ordinary least squares (OLS) regression. In particular, to measure the accuracy and robustness of the models, we use in-sample, out-of-sample, and out-of-time validation. Moreover, considering the economic context and the obtained estimation errors, we discuss theoretical and practical advantages or disadvantages of each step in our approach.

4.2 Related literature

The linear regression has so far been the most frequently used method for estimating the LGD in the existing literature on LGD research. Nevertheless, when regarding the specific features of the LGD distribution, linear regression seems to be at least econometrically inappropriate for the estimation task. Typically, the workout LGD of loans and leases, calculated from discounted cash flows after the default of the customer, is bimodally or even multimodally distributed (compare Laurent and Schmit (2005), Zhang and Thomas (2012), Hartmann-Wendels et al. (2014), and Li et al. (2014)). This unusual shape of the density suggests that LGD estimation requires the use of advanced methods. These methods should be able to approximate the complex relationships between the available information and the LGD as precisely as possible in order to produce accurate and comprehensible estimations.

Against this theoretical and practical background, a number of different methods have already been investigated in the literature. In particular, the relevant studies examine the models' suitability and predictive accuracy for LGD estimation.

Several studies focus on reproducing the LGD's density function in order to extrapolate accurate estimations in this manner. For this purpose, Calabrese and Zenga (2010) use a mixed random variable to model LGD on the unit interval. They apply their concept to a large set of defaulted Italian loans. Altman and Kalotay (2014) adopt a similar approach based on the mixture of Gaussian distributions. They report successful estimations using Moody's Ultimate Recovery Database (MURD). Hartmann-Wendels et al. (2014) also apply an approach based on finite mixture models in order to estimate the LGD of leases. However, out-of-sample, their approach performs poorly. The authors conclude that reproducing the LGD density is only of secondary importance to the estimation accuracy.

Further studies examine the suitability of various parametric and nonparametric methods for LGD estimation. Applying several regression techniques to the data of six different banks, Loterman et al. (2012) conclude that nonlinear methods perform better than linear methods. Qi and Zhao (2011) obtain a similar result. They compare different parametric and nonparametric methods using MURD. The authors argue that nonparametric methods can generate more accurate LGD estimations due to their ability to model nonlinear relationships between the LGD and continuous explanatory variables. In particular, they find regression trees to be a suitable nonparametric method for estimating LGD. Bastos (2010) obtains a similar outcome when he uses regression trees on Portuguese bank loans. Likewise, Hartmann-Wendels et al. (2014) successfully apply model trees to estimate the LGD of German leases.

Recently, some studies have applied ensemble learning techniques to estimate LGD. These are an extension of the analysis of single procedures. Bastos (2013) improves the estimation accuracy significantly by using regression trees in an ensemble approach on MURD. On a set of leases, Töws (2014) finds that random forests achieve higher coefficients of determination than does linear regression.

In addition to single-stage models, some studies implement two-stage models to forecast LGD. Typically, these models split the observations ex ante according to a specific key feature. To predict the LGD of mortgage loans, Leow and Mues (2012) first estimate the probability of mortgage accounts undergoing repossession. Then, they calculate the loss in the event of repossession using a certain haircut value. The latter is the ratio of the forced sale price and the valuation of the repossessed property. Concerning the LGD of leases, Töws (2014) successfully introduces a two-stage approach. He distinguishes between recovered and written off contracts and then estimates the respective LGD.

Although the findings of several studies show that complex models can generate more accurate LGD estimations than standard econometric techniques such as the linear regression, the results of Qi and Zhao (2011) and Hartmann-Wendels et al. (2014) indicate that overfitting is a common concern of complex models. Hence, as overfitting may negatively affect forecasting accuracy, a prerequisite for advanced models to perform well is the existence of a correspondingly large database, both in terms of observations and associated information. The lack of data and issues with controlling overfitting are presumably the reasons why linear regression has been the most frequently used method for estimating the LGD in the literature so far. Moreover, some studies also demonstrate the practical suitability of the linear regression. Zhang and Thomas (2012) apply linear regression and survival analysis to a dataset of defaulted personal loans from the UK. They find that linear regression generates the best LGD estimates in general and outperforms more advanced estimation techniques. Bellotti and Crook (2012) obtain a similar result when estimating the LGD of UK credit cards.

The approach for predicting the LGD of leasing contracts we present in this study differs in crucial aspects from the majority of the advanced estimation methods discussed in the literature. We introduce an estimation approach that is explicitly based on economic considerations. In particular, by applying an economic separation of the LGD, we consider the LGD a composed measure of risk.

4.3 Dataset

The dataset consists of 1,493 defaulted leasing contracts with 907 lessees from a large German leasing company. The key figures of the dataset are shown in Table 4.1. The contracts were executed between 1996 and 2009. Their default occurred between 2002 and 2009. The default status of any contract was determined by the default events outlined in Section 452 of Basel II. These events correspond to Article 178 (1) of the CRR. The associated EAD of the contracts ranges between €216 and €1,620,114 with an average of €53,025 per contract. This corresponds to an average ratio of EAD to the lessor's calculated contract value of 60%. The average LGD amounts to 35%. LGD and its distribution will be discussed in detail in following section.

		LGD			EAD (in €)		
# Contracts	# Lessees	Mean	Median	Std	Mean	Median	Std
1,493	907	0.35	0.30	0.48	53,025	20,955	120,845

Table 4.1: Numbers of contracts and lessees as well as loss given default (LGD) and exposure at default (EAD) key figures of the dataset. Std is the standard deviation.

All contracts defaulted without recovering and were finally written off. The contracts default after an average of 50% of their maturity. That is approximately 2.5 years after the execution of the average contract. The mean workout lasts about two years. The workout of all contracts has been completed. The last of these was completed in 2010. Further data has not been provided.

Our data is extremely valuable with respect to its high level of detail, particularly regarding the workout process. The breakdown of cash inflows and outflows during the workout process is of particular importance for the derivation and economic interpretation of the approach we present in this study. The carefully documented costs concerning the disposal of the leased asset and the collection of overdue payments, allow a precise and economically meaningful separation of asset-related and miscellaneous revenues. In general, Schneider et al. (2010) find that rating and industry of the customer already explain 50% of the variance in LGD. Beside customer specific information our data also comprises details with regard to the contract, the leased asset as well as additional information on the contract's default. These are for instance the distinction between customer types, type of contract, e. g. full pay-out lease contract, asset class, e. g. vehicle, and default reason, e. g. insolvency.

Before any separation or estimation of the LGD, we briefly discuss its calculation. The LGD is defined as the portion of EAD that could not have been recovered in the case of a contract's default. Its counterpart is the recovery rate (RR). The workout RR is the ratio of the amount recovered and EAD, which is equivalent to $1 - \text{LGD}$. In line with Article 5 (2) CRR, we use the term-related refinancing interest rate to discount all incurred cash flows (CF) and workout costs (WC) to the time of default. The EAD is the present value of the defaulted contract's outstanding exposure, calculated as the sum of outstanding payments at the time of default.

The detailed breakdown of incoming and outgoing cash flows during each contract's workout enables us to determine LGD very precisely. Formally, we calculate the LGD as

$$\text{LGD} = 1 - \frac{\text{CF} - \text{WC}}{\text{EAD}} = 1 - \text{RR}. \tag{4.1}$$

Beyond the pure determination of LGD, we calculate component LGDs. The asset-related LGD (ALGD) summarizes all asset-related payments, such as the asset's liquidation proceeds and incurred liquidation costs. We call the remaining part of the LGD miscellaneous LGD (MLGD). The MLGD comprises revenues from capital services, such as interest rates and customer payments, the costs of collateral, such as recovery costs and maintenance costs, and proceeds of collateral, other indirect costs, and other payments. Both component LGDs refer to the overall EAD. However, they differ particularly in terms of the lessor's influence on the respective cash flows. While repossession of the leased asset as well as its disposal is entirely in the responsibility of the lessor, miscellaneous cash flows depend on several factors outside his/her control. For instance, if the defaulted lessee goes through an insolvency proceeding, the insolvency estate is distributed pro rata between all relevant creditors. Basically, the MLGD of a leasing contract is the equivalent of a loan's LGD.

We derive the two component LGDs from Equation (4.1) by identifying the asset proceeds (AP) within the incoming cash flows and the related asset liquidation costs (LC) within the workout costs. This splitting results in

$$\begin{aligned} \mathrm{LGD} &= 1 - \frac{(\mathrm{CF_M} + \mathrm{AP}) - (\mathrm{WC_M} + \mathrm{LC})}{\mathrm{EAD}} \\ &= 1 - \frac{\mathrm{AP} - \mathrm{LC}}{\mathrm{EAD}} - \frac{\mathrm{CF_M} - \mathrm{WC_M}}{\mathrm{EAD}} \\ &= 1 - \mathrm{ARR} - \mathrm{MRR}, \end{aligned} \tag{4.2}$$

with $\mathrm{CF_M}$ and $\mathrm{WC_M}$ denoting the remaining miscellaneous incoming cash flows and workout costs respectively.

Subsequently, we derive the asset-related RR (ARR) and the miscellaneous RR (MRR). As usual, we obtain the LGD as the counterpart of the RR

$$\mathrm{ALGD} = 1 - \mathrm{ARR}, \qquad \mathrm{MLGD} = 1 - \mathrm{MRR}. \tag{4.3}$$

In terms of ALGD and MLGD, the LGD then is calculated as

$$\mathrm{LGD} = \mathrm{ALGD} + \mathrm{MLGD} - 1. \tag{4.4}$$

Descriptive statistics

In contrast to various studies, we do not restrict LGD to the unit interval, such as is done by Chalupka and Kopecsni (2009), Bastos (2010), and Zhang and

Thomas (2012). For leases, LGDs outside the unit interval are frequently observed. Hartmann-Wendels and Honal (2010) argue that LGDs less than 0 may occur in cases where the asset disposal covers more than the amount of EAD. Additionally, incorporating workout costs may cause the LGD to rise beyond 1. Table 4.2 provides a brief overview of the LGDs' distribution parameters. The overall LGD averages near 35%, and we observe an average ALGD of 69% and MLGD of 65%. The standard deviation of ALGD is notably lower than that of MLGD and LGD. Minimum and maximum of ALGD and MLGD are consequently higher than those of the LGD.

Part of LGD	Mean	Std	Min.	P25	Median	P75	Max.
ALGD	0.69	0.41	−0.99	0.41	0.98	1.00	2.03
MLGD	0.65	0.51	−1.04	0.22	0.89	1.01	2.68
LGD	0.35	0.48	−1.36	0.00	0.30	0.76	1.50

Table 4.2: Distribution parameters of the loss given default (LGD). Std is the standard deviation, Min. is the minimum, and Max. is the maximum LGD value. P25 and P75 are the respective quartiles. ALGD is the asset-related LGD and MLGD is the miscellaneous LGD. We derive both partial LGDs from Equation (4.3).

We find that the ratio of asset value at default to EAD is 54% on average. Although, the lower quartile of ALGD is quite high, for more than 10% of the contracts the asset value even exceeds EAD. For these contracts, ARR is higher than 1. While the default value of the leased asset is not an explicit part of the EAD, this value qualifies for incoming cash flow during the workout process in the case of asset disposal. In the same way as any other cash income, the disposed asset value reduces the LGD. Moreover, in contrast to the liquidation of a loan's collateral, the lessor, as the legal owner of the leased asset, can keep any surpluses from disposing of the leased asset even if the resulting ARR exceeds 1.

For a lessor's internal risk management, determination of ALGD is useful. If interpreted as a stand-alone parameter, ALGD is theoretically an upper limit to the LGD. This is true if the MLGD does not exceed a value of 1, which implies the

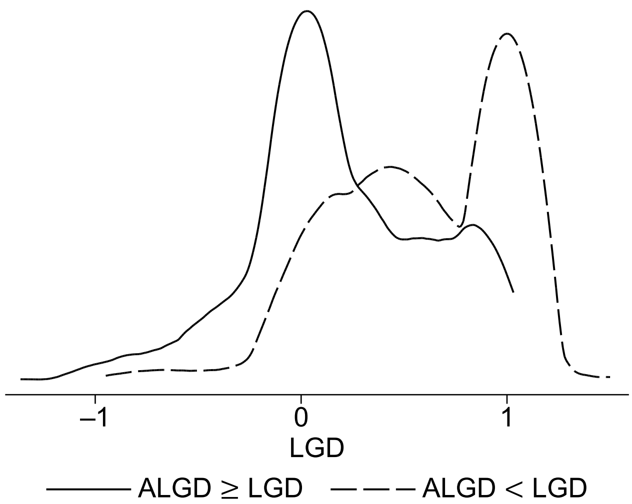


Figure 4.1: Densities of the two classes of the loss given default (LGD), after separating the contracts according to their relationship of LGD to asset-related LGD (ALGD). The full amount of the exposure at default (EAD) is recovered in the case of 0. -1 defines an EAD recovery of 200% while an LGD value of 1 signifies the loss of 100% of the EAD.

success of the workout process. Therefore, depending on the amount of ALGD, the lessor can determine whether the asset sales proceeds already cover the EAD or if further workout actions should be taken to collect overdue payments.

Frontczak and Rostek (2015) argue that knowledge of the effect of disposal efficiency and related costs on the LGD may affect a lender's disposal policy. Consequently, it would be useful for the lessor to know ex ante whether the MLGD will exceed 1. If it does, the lender loses more than the full amount of EAD. Strictly speaking, MLGDs > 1 indicate that the incurred collection costs will exceed the payments collected. In such cases, even if the asset sales proceeds cover only a small portion of EAD, the workout should be restricted to the disposal of the leased asset because collecting further payments is inefficient from an economic standpoint.

Theoretically, it is also possible that the ALGD exceeds 1. Nevertheless, in our data we find that asset sales proceeds exceed the incurred disposal costs in 99% of all cases. This outcome could have been expected, because leasing companies are experts in disposing of their leased assets. Hence, the disposal is economically

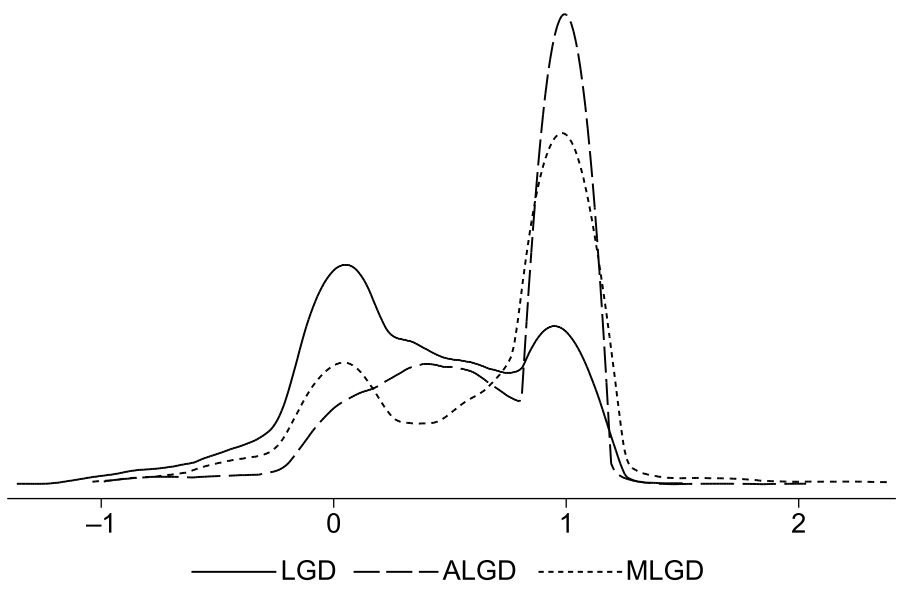


Figure 4.2: Densities of the shares of the loss given default (LGD), asset-related LGD (ALGD), and miscellaneous LGD (MLGD). The full amount of the exposure at default (EAD) is lost in the case of 1. −1 defines an EAD recovery of 200% while an LGD value of 2 signifies the loss of 200% of the EAD.

reasonable in almost any case. Interestingly, for about 35% of the examined contracts, the MLGD exceeds 1. This implies that the ALGD as an upper limit of the LGD holds for only about 65% in practice. Nevertheless, as can be seen in Figure 4.1, this upper limit is an important feature in distinguishing the contracts. More precisely, categorizing the contracts according to this upper limit leads to LGD distributions that are, to a large extent, disjointed. The realized LGDs of the contracts that satisfy ALGD $\geq$ LGD concentrate around 0 with a mean of 0.20. In contrast, for contracts with ALGD exceeding the LGD, LGDs are located, in particular, around 0.5 and 1 with a mean of 0.59.

Figure 4.2 visualizes the density of the calculated overall LGD of the underlying dataset. In addition, ALGD and MLGD densities are plotted. Both overall LGD and MLGD exhibit a pronounced bimodal shape, with concentrations around an LGD level of 0 and 1. The LGD's mean of 0.35 in Table 4.2 indicates that the overall LGD is rather small in most cases. Its median of 0.30 confirms this finding. ALGD and MLGD, however, have more density around 1. From the perspective of a regular lender, a high ALGD is of less concern than a high MLGD. Because the asset's fair value is not included in the EAD, the cash inflow from the asset's disposal has an unexpected reducing effect on the LGD. In contrast, the cash inflows considered by MLGD are fully accounted for in the EAD. A high MLGD reflects a poor outcome from the workout process. However, the ALGD is important to lessors because revenues from disposing of the leased asset in the case of default are a substantial aspect of a lessor's business model.

To be precise, the average revenue from disposal of the leased asset amounts to €15,322 per contract. The miscellaneous payments during the workout process add up to €13,616 on average per contract. This allocation of cash inflows emphasizes the importance of both sources of revenues for a leasing company. It confirms that the workout process of defaulted leases is quite different from that of loans. Consequently, for leasing contracts it is essential to consider both ALGD and MLGD when estimating the overall LGD. This entails that especially the prediction of leasing LGDs benefits from a large amount of information. Based on the findings of Schneider et al. (2010) particularly customer specific information are important to explain the MLGD. Additionally, to estimate the ALGD, naturally, information on the leased asset is crucial. We note that the share of revenues from disposing of the leased asset is indeed slightly higher on average than the remaining share. However, in particular for less valuable assets, the traditional payments collection during the workout process are substantial.

In Table 4.2 we observe higher standard deviations of MLGD and LGD compared to ALGD. Thus, the latter is less volatile, Miller (2015) noting much the same. Therefore, ALGD might be easier to estimate. In addition, Table 4.3 displays the key figures of the realized LGD, ALGD, and MLGD values over the default years of the observation period. At this level of aggregation, ALGD is still less volatile than MLGD and LGD in a year by year comparison. Concerning LGD, we observe rather small fluctuations during the period of 2003 to 2008. It

	LGD		ALGD		MLGD	
Year	Mean	Std	Mean	Std	Mean	Std
2002	0.3723	0.5283	0.7386	0.4076	0.6337	0.5652
2003	0.2995	0.4597	0.6978	0.4360	0.6017	0.5344
2004	0.3068	0.4341	0.6637	0.4267	0.6431	0.5302
2005	0.3245	0.4516	0.6640	0.4505	0.6604	0.4603
2006	0.3218	0.4557	0.6506	0.4218	0.6712	0.4750
2007	0.3303	0.4225	0.6503	0.4221	0.6799	0.4555
2008	0.3413	0.4218	0.6167	0.3968	0.7247	0.4062
2009	0.3999	0.4179	0.6654	0.2914	0.7345	0.3826

Table 4.3: Distribution parameters of the loss given default (LGD) for each default year. Std is the standard deviation. ALGD is the asset-related LGD and MLGD is the miscellaneous LGD. We derive both parts of the LGD from Equation (4.3).

is only for the years 2002 and 2009 that the realized LGD is, in comparison, noticeably higher on average. In particular, the higher average LGD in 2009 might be a result of the global financial crisis. Regarding ALGD, the means fluctuate around 65%. However, we find it interesting that ALGD does not increase unusually in 2009. This finding indicates that the fluctuations of ALGD are driven by each year's asset disposals but are not driven by the economy. Apparently, ALGD does not increase during the financial crisis. We attribute this effect to the lessor's excellent knowledge of secondary markets. Obviously, there is a difference in the course of MLGD. It also fluctuates only rarely between the years 2002 and 2007. However, we observe a marked but manageable increase in 2008 and 2009, which might be a result of the financial crisis.

The evolution of the three LGD ratios supports our hypothesis that ALGD might be easier to estimate for the lessor than MLGD or LGD. However, we find no empirical evidence, that the economy, accounting, e. g., for gross domestic product and unemployment rate, has an impact on ALGD. The economy might influence MLGD and LGD slightly. Nevertheless, the potential effect seems to be minor. Moreover, Miller (2015) shows that the LGD estimation at the default of a lease benefits only slightly, if at all, from considering the economic situation.

Consequently, we do not include macroeconomic factors in our approach. In fact, for the estimation of the LGD ratios we focus on contract related factors, such as the type of the leased asset, the customer, and the reason for default.

4.4 Methods

In contrast to recent studies on LGD estimation, we do not focus on the comparison of very complex or even black box methods, such as support vector machines or neural networks. Instead, we develop an economically based and consistent technique for estimating LGDs. Rather than regarding LGD as a holistic measure of risk, we separate the LGD into an asset-related part and a miscellaneous part and, hence, taking into account the specific characteristics of leases. In order to provide evidence that the increase in estimation accuracy does not arise solely from particularly suitable methods but from sophisticated economic consideration, we essentially apply two distinct methods to our proposed multi-step approach. These are OLS and as an advanced estimation method, the tree algorithm random forest (RF). Throughout the study, we set the traditional direct estimation of LGD by OLS and RF as a benchmark to compare the performance of our multi-step estimation model and to measure the improvement.

As OLS is a common estimation method, we will only give a brief overview of the RF model. The RF tree algorithm was constructed by Breiman (2001). It has many similarities to regular regression and classification trees. These trees subsequently divide the initial dataset according to a series of if-then conditions. At every node of the tree, the best split is performed according to an appropriate split criterion, e. g., the greatest expected reduction in standard deviation. Each contract terminates in one leaf of the final tree. Each leaf's estimation value then is the average value of the contracts of the respective leaf. In terms of classification, the contracts' realized class in each leaf determines the leaf's class estimation.

RF differs from regular regression trees in three important ways. First, instead of building only one tree, a series of trees and, thus, a forest is built. Second, each tree is calibrated with a random sample of the dataset. Third, at each node the available set of splitting variables is a random sample of all available variables. The final estimation of a contract is the average of the single tree estimations. For classification, the majority vote determines a contract's class. We use the RF standard parameters suggested by Breiman (2001). For classification, these are $\sqrt{m}$ randomly chosen variables for each split and $m/3$ variables for regression out of a total of m variables. The size of the forest is fixed at 1,000 trees, as proposed by Hastie et al. (2009).

Beside the frequently used OLS, numerous studies have shown that tree-based algorithms are particularly well suited to estimating LGD. While Bastos (2010) and Hartmann-Wendels et al. (2014) find that regression and model trees generate robust and accurate LGD estimations, Töws (2014) reports similar outcomes for RFs explicitly.

4.4.1 Direct estimation

To begin with, we take a look at direct estimation methods. Direct estimation is easy to implement and, therefore, the most elementary and common method for estimating LGD. In this study, direct estimation by OLS and RF serves as a useful benchmark when we compare it to the respective multi-step model by measuring the models' performance improvements. The left-hand side of Figure 4.3 visualizes a simplified direct estimation method. In general, the method uses a set of variables to produce estimations of the LGD.

Using OLS, we model the LGD dependent on the available and relevant variables (VAR) in a linear combination

$$\mathrm{LGD} = \alpha + \sum_{i=1}^{m} \beta_i \cdot \mathrm{VAR}_i + \varepsilon, \tag{4.5}$$

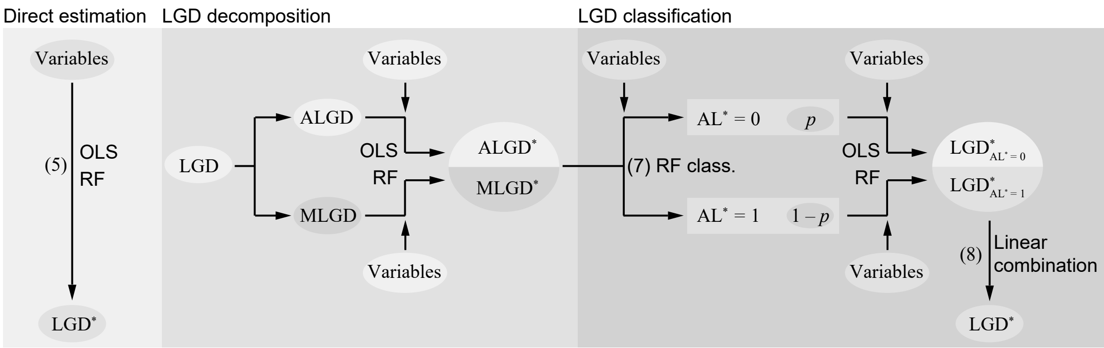


Figure 4.3: Procedure of the developed models. Our approach consists of three consecutive parts. Equations (4.5), (4.7), and (4.8) used at the model's steps are explained in detail in the following paragraph. Direct estimation determines the loss given default (LGD^*) according to Equation (4.5) using the variables available, both in an ordinary least squares (OLS) and in a random forest (RF) regression model. In the LGD decomposition, we divide the realized LGD into an asset-related LGD (ALGD) and a miscellaneous LGD (MLGD). Then again, using the available variables, two OLS or RF models are calibrated to estimate ALGD^* and MLGD^*. Subsequently, the contracts of our dataset are classified into two classes. An RF classification model uses the available variables including ALGD and MLGD and their estimated values to perform the classification of Equation (4.7). It aims to assign $\text{AL}^* = 0$ correctly to contracts with an ALGD exceeding its LGD, and $\text{AL}^* = 1$ in case ALGD falls short of LGD. Based on these two disjoint datasets, we calibrate an OLS or RF model on each to estimate the two $\text{LGD}^*_{\text{AL}^*}$. Using the linear combination of Equation (4.8), we calculate the final LGD estimation by weighting $\text{LGD}^*_{\text{AL}^*}$ with their classification probabilities p and $1-p$.

with α the regression's constant, β_i the slope coefficient of variable VAR_i, ε the residual, and m the number of included variables. For RF regression, we train a forest based on the same information set, estimating the dependent variable directly.

The advantage of direct estimation is the plain analysis of the influence of the independent variables. In the case of OLS, the significance and slope of single influencing factors are fairly easy to measure and have economic interpretations. The importance measure of RF allows for similar conclusions. However, from a methodological perspective, OLS comes with a major disadvantage. It only models linear relationships between dependent and independent variables. This means that it is not possible to consider many latent influences and changes of influences according to independent variable values. Although, OLS has been successfully used for estimating LGD, e. g., by Bellotti and Crook (2012) and Zhang and

Thomas (2012), RF should, theoretically, be much more suited to the estimation task. In particular, the latter can consider nonlinear dependencies between the LGD and its explanatory variables by generating homogeneous subsets of the data. Still, both methods can only process the plain information available.

4.4.2 Loss given default decomposition

From an economic point of view, the LGD is a linear combination of cash flows relative to EAD. With leasing contracts, this relationship plays a particularly important role because, unlike with loans, the cash flows are typically issued from very different sources. Observing the cash flows in detail, we attempt to provide additional information to the estimation of LGD by breaking down the LGD to ALGD and MLGD. Equations (4.2) and (4.3) provide the necessary mathematical steps of this calculation. Figure 4.3 outlines the procedure of the LGD decomposition. Similar to LGD, neither ALGD nor MLGD are available at the time of contract's default. Therefore, the idea is that, instead of estimating LGD directly, we estimate ALGD and MLGD and combine these parameters to form a new LGD estimation. Again, for estimating ALGD and MLGD, we apply OLS and RF models. In principle, any other method could be utilized.

From a mathematical and economic perspective, the separate estimation is reasonable in three ways. First, asset-related cash flows obviously depend on influencing factors different than those applying to miscellaneous cash flows. Indeed, a simple OLS regression finds that both partial LGD ratios are driven by the type of the contract and the asset's acquisition value. In addition, ALGD is significantly influenced by the asset class, term to default, collateral, and assessment basis. MLGD on the other hand is additionally influenced by the relative contract age, type of customer, and the default reason. Second, according to the different density shapes outlined in Figure 4.2, the estimation of the two LGDs might vary in its accuracy. In particular, the markedly lower standard deviation of ALGD compared to MLGD highlighted in Table 4.2, indicates that the estimations of ALGD might be significantly more precise. Third, both of these estimated components of the LGD provide decision support concerning the actions that should be taken during the workout process in order to achieve LGDs which are as low as possible. The last argument is particularly important from an economic point of view.

The gain of information by estimating ALGD and MLGD may be used in different ways to enhance the accuracy of LGD estimation. Theoretically, the LGD can be calculated reversely by using Equation (4.4)

$$\mathrm{LGD}^* = \alpha \cdot \mathrm{ALGD}^* + \beta \cdot \mathrm{MLGD}^* - \varepsilon, \tag{4.6}$$

where LGD^*, ALGD^*, and MLGD^* are the estimated LGD, ALGD, and MLGD, respectively. α and β are slope coefficients and ε is the constant. In the theoretical calculation, these three parameters are set to 1. However, for practical purposes it might be suitable to set up an OLS regression to find the optimal values for these parameters. Nevertheless, one major disadvantage of this procedure is that the full estimation error of both estimated ALGD and MLGD enters the estimated LGD. Consequently, we do not pursue this approach any further.

4.4.3 Loss given default classification

Instead of deriving an LGD estimation from ALGD^* and MLGD^*, we use the estimated values to classify the contracts into two classes. In Section 4.3, we show that the ALGD is a theoretical upper boundary to the LGD. By generating a dummy variable

$$\mathrm{AL} = \begin{cases} 0 & \text{if ALGD} \geq \text{LGD} \\ 1 & \text{if ALGD} < \text{LGD}, \end{cases} \tag{4.7}$$

we identify contracts, which realize an LGD exceeding their ALGD. According to Figure 4.1 this categorization leads to a largely disjointed separation of the contracts in terms of the LGD distributions. Moreover, the two resulting distributions of the LGD feature less distinctive bimodal shapes than the LGD distribution of all contracts, illustrated in Figure 4.2. Consequently, we expect that estimating LGD separately in each class is easier than estimating LGD without this separation. On account of this, we calibrate an RF classification model with AL as the dependent variable to predict whether a contract's LGD is expected to be below or above its ALGD. This model uses the relevant and available information at contract's default. Expanding this information set, we additionally use ALGD and MLGD determined according to Equation (4.3) to calibrate the classification model. For predictive classification, we consequently use the respective estimates of ALGD and MLGD from Section 4.4.2, as is indicated by the right-hand side of Figure 4.3.

Theoretically, it is possible to classify the contracts directly by using only the estimates of ALGD and MLGD. However, in this case, the estimation error of these estimates would directly impact the classification accuracy negatively. Therefore, we do not rely on these two ratios but rather calibrate a classification model using a set of information.

For each contract, we obtain the classification probability p of the respective contract in class 0, and its estimated class AL*. Based on the contracts of these two classes, we calibrate two separate LGD regression models. In the estimation step, every contract receives exactly two LGD estimations, one from each of the two models calibrated. Finally, we calculate the estimated LGD in a linear combination

$$\mathrm{LGD}^* = p \cdot \mathrm{LGD}^*_{\mathrm{AL}^*=0} + (1-p) \cdot \mathrm{LGD}^*_{\mathrm{AL}^*=1}, \tag{4.8}$$

using the classification probability p to weight the single LGD estimates.

The additional classification step enriches the overall LGD estimation by interpreting ALGD as an upper limit to the LGD. Economically, the classification of a contract indicates, which actions the lessor should take during its workout process. In particular, if the LGD is likely to exceed its ALGD, the lessor should consider restricting the workout process to the disposal of the leased asset, because, in this case, the miscellaneous workout costs are expected to exceed the miscellaneous cash inflows. Considering that MLGD > 1 for about 35% of the contracts of the studied lessor, the proper implementation of the workout process such as we suggest, would lower its realized LGD. In case all workout decisions are followed as proposed by our model, this reduction would amount to nearly 10% reducing the mean realized LGD to 0.32. Considering the portfolio EAD (see Table 4.1), this reduction of the LGD would lead to lower losses to the leasing company of about €2,250,000.

4.4.4 Validation techniques

In order to validate the estimation accuracy and to verify the robustness of the methods used, we apply three fundamentally different validation techniques outlined in Figure 4.4. Beside common in-sample and out-of-sample validation, we also use out-of-time validation. The last simulates an estimation scenario that is as close to reality as possible. In the course of the study, we are estimating different parameters, such as LGD, ALGD, and MLGD. Furthermore, we perform a classification to predict whether the ALGD is greater or less than the LGD. Since the following validation techniques apply to all of these parameters, we will use a uniform synonym and call them dependent variables.

For the in-sample model calibration, all observations and available information at the time of contracts' default are used. The estimation of the dependent variable is then carried out on the same data. Consequently, the estimation accuracy is expected to be relatively high. On the one hand, this effect is based on the

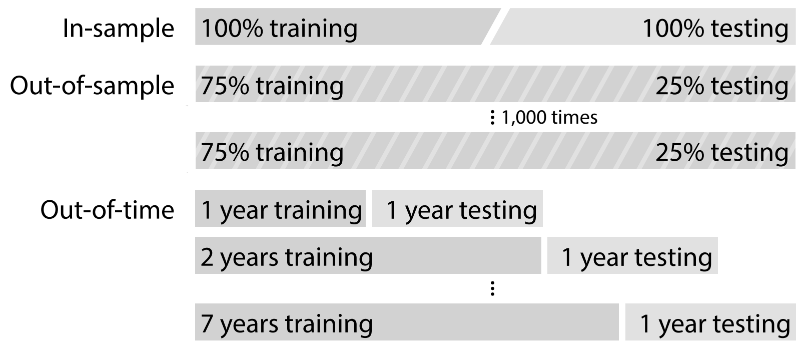


Figure 4.4: Each validation approach divides the total dataset into $x\%$ training set to calibrate the estimation model, and $(1-x)\%$ validation set. Out-of-sample, we divide the data randomly 1,000 times. In the case of out-of-time validation, the data is divided by the contracts' year of default. The first model is then calibrated on contracts that default in the first year and validated on contracts that default in the following year.

particularly large dataset used for the model's calibration. On the other hand, when estimating the dependent variable, each combination of information that occurs in the validation set is already known to the model. One problem, however, is that a high in-sample estimation accuracy frequently results from the overfitting of the model to the underlying data. In fact, in reality, most validation sets consist of unknown observations and combinations of information.

Therefore, it is reasonable, and for the estimation of LGD, it is required by the regulator, that the estimation model be calibrated on a sample of the data. Article 179 (1)(d) CRR states that this sample shall be sufficient to provide the performing institution with confidence in the accuracy and robustness of its estimates.

For out-of-sample validation, these samples can be implemented by k-fold cross-validation. While earlier studies on LGD estimation used this method frequently, Kohavi (1995) employs different validation methods, such as cross-validation, leave-one-out, and random subsampling. The last divides the data into training and validation sets and is run l times. In their recent study, Hartmann-Wendels et al. (2014) use random subsampling to validate their regression results. Dividing the data into 75% training and 25% validation sets, they repeat the procedure 25 times. Yao et al. (2015) perform a similar out-of-sample validation using 70% and

30% randomly chosen observations as training and validation sets respectively. Their procedure is repeated 100 times.

We produce randomly drawn subsamples of 75% for the training set without returning the observations. The remaining 25% form the validation set. On each training set, an estimation model is calibrated. Subsequently, we estimate the dependent variable for the corresponding validation set. We perform this step 1,000 times and average the resulting performance measures. The estimation error obtained out-of-sample is usually greater than that of in-sample validation. However, the error reflects a much more realistic allocation of the model's predictive accuracy.

The final step in validating the predictive accuracy of estimation models is out-of-time validation. In the recent literature on LGD, models are rarely validated out-of-time due to special requirements to the underlying data. In particular, a comprehensive dataset and time information are necessary. For out-of-time validation, Gupton and Stein (2005) propose a growing window, subsequently using observations prior to a fixed year for the training set. The following year serves as the validation set. Most recently, Altman and Kalotay (2014) conduct an out-of-sample, out-of-time simulation experiment. They calibrate a model on observations prior to 2002 and randomly draw 100 observations from the period between 2002 and 2011 for the validation set. This step is repeated 50,000 times.

However, we are convinced that an adequate and realistic out-of-time estimation model should be built upon the available historical data and should forecast the forthcoming period. Considering a period of several years for the validation set, as proposed by Altman and Kalotay (2014), might dilute specific characteristics of single years. Hence, the estimation model would produce out-of-time results that might be too optimistic.

Employing the method of Gupton and Stein (2005), we divide our dataset according to the contracts' time of default. To calibrate a solid first model, built

	2002	2003	2004	2005	2006	2007	2008	2009
# Contracts	575	191	116	116	144	122	127	102

Table 4.4: Year of default and frequency of contracts.

upon a sufficiently large dataset, we use the contracts that defaulted in 2002. Table 4.4 shows that a total of 575 contracts defaulted in the first year of the observation period.[5] The trained model is then validated on those 191 contracts which defaulted in the following year – in this case, 2003. Calibrating the next model, we expand the training set by one year. By doing so, the first two years are used for the model's calibration. Subsequently, further models are build by expanding the training set. Validation is always performed on the contracts of the year following the training period. Consequently, the final model is based on the contracts that have defaulted between 2002 and 2008. This model's predictive accuracy is validated by contracts that defaulted in 2009. Finally, we weight the outcomes of each year with the relevant number of observations.

4.4.5 Performance measurements

In order to compare the results of our different estimation models, we use four performance measurements. These are: mean absolute error (MAE); mean squared error (MSE); normalized area under the regression error characteristic curve (NA-REC); and Theil inequality coefficient (TIC). Each of these performance measurements focuses on the evaluation of specific aspects of the estimation.

MAE and MSE are common measures to evaluate the performance of estimation methods. With LGD and LGD* denoting the realized and estimated LGD, respectively, and n being the number of observations, we calculate MAE and MSE

[5]The large amount of defaults in 2002 arises from an inaccuracy in the default date provided. Some of these contracts may have defaulted before 2002 but were uniformly assigned to this specific first default year. However, this inaccuracy has no impact on the out-of-time validation.

according to the following definition

$$\mathrm{MAE} = \frac{1}{n}\sum_{j=1}^{n} |\mathrm{LGD}_j - \mathrm{LGD}_j^*|, \tag{4.9}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{j=1}^{n} \left(\mathrm{LGD}_j - \mathrm{LGD}_j^*\right)^2. \tag{4.10}$$

MSE punishes larger deviations between predicted and realized values harder. In general, a low parameter outcome is preferable for both measurements.

NAREC can be used to evaluate the performance of regression models in total. This measure is based on the regression error characteristic (REC) curve developed by Bi and Bennett (2003) as a generalization of receiver operating characteristic curves for regression problems. The REC curve draws the error tolerance δ against the models accuracy $\mathrm{acc}(\delta)$. The latter computes as

$$\mathrm{acc}(\delta) = \frac{\#\left\{\mathrm{LGD}^*\colon |\mathrm{LGD}_j^* - \mathrm{LGD}_j| = \epsilon_j \leq \delta,\ j = 1, \ldots, n\right\}}{n}. \tag{4.11}$$

It specifies the percentage of observations whose estimates do not exceed the error tolerance. NAREC is defined as the area under the REC curve (AUC) normalized to the interval $[0, 1]$

$$\mathrm{AUC} = \int_0^\infty \epsilon p(\epsilon)\,\mathrm{d}\epsilon, \tag{4.12}$$

with ϵ the error of the model regarded as random variable and $p(\epsilon)$ the corresponding probability density function. A higher outcome of NAREC implies more accurate estimations produced by the estimation model in total.

TIC was introduced by Theil (1966) and sets the mean squared error in relation to the sum of the quadratic realized and estimated LGD. It aims to quantify the goodness of fit and robustness of a model. We use

$$\mathrm{TIC} = \frac{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{LGD}_i - \mathrm{LGD}_i^*)^2}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\mathrm{LGD}_i^*)^2} + \sqrt{\frac{1}{n}\sum_{i=1}^{n}\mathrm{LGD}_i^2}}, \tag{4.13}$$

to calculate the TIC. A low parameter outcome is preferable.

To measure the performance of classification methods, we use the classification error. It is the ratio of misclassified cases to all cases, which is the relative frequency of misclassification. We calculate the classification error (CE) as

$$\mathrm{CE} = \frac{1}{k}\sum_{i=1}^{k} I(\mathrm{AL}_i^* \neq \mathrm{AL}_i) = \frac{\text{Misclassified cases}}{\text{All classified cases}}, \tag{4.14}$$

where AL_i is the realized class of case i defined in Equation (4.7), AL^* is the estimated class, I is the indicator function, and k is the number of classified cases.

4.5 Results

In order to reliably compare the results of the applied models, we use a consistent set of explanatory variables listed in Table 4.5. The performance results for the three different validation techniques are contained in Tables 4.6–4.8. The remainder of the results section is arranged as follows: We discuss the LGD estimation results and models' performance of the in-sample, out-of-sample, and out-of-time validation in Sections 4.5.1–4.5.3. Section 4.5.4 gives a detailed explanation of the single steps of our model's performance. Finally, Section 4.5.5 interprets the results and puts them into perspective by comparing these with the results of the relevant recent literature.

Variable	Description
Asset class	Class of asset, e. g., vehicle, information technology equipment, or machinery
Acquisition value	Acquisition value of the leased asset in €
Assessment basis	Lessor's calculated contract value
Calc. residual value	Residual asset value calculated at the execution of the contract
Contract type	Type of contract, e. g., full payout lease or sale and lease back
Collateral	Dummy for provided collateral beside the leased asset
Retail	Dummy to distinguish between retail and non-retail customers
Term to default	Term between execution and default of the contract in years
Rel. contract age	Ratio of term between execution and default to maturity of the contract
Default reason	Default trigger, e. g., 90 days past due or insolvency
EAD	Outstanding exposure at the default of the contract
ALGD/ALGD*	Asset-related loss give default respectively estimated asset-related loss given default
MLGD/MLGD*	Miscellaneous loss give default respectively estimated miscellaneous loss given default

Table 4.5: Consistent set of variables used for regression and classification. Asset-related loss given default (ALGD) and miscellaneous loss given default (MLGD) and their estimates are additionally used in the classification step of our model in Figure 4.3.

4.5.1 In-sample validation

The in-sample results that we present in Table 4.6 show that RF produces more accurate estimates than OLS throughout the performance measurements. Irrespective of whether the direct or the multi-step approach is used, the RF model strictly outperforms both OLS models. In particular, the exceptionally low errors in terms of MAE and MSE indicate a close adaptation of RF to the training data.

Method	MAE	MSE	NAREC	TIC
Direct estimation				
OLS	0.3436	0.1821	0.6593	0.1835
RF	0.1484	0.0362	0.8325	0.0340
Multi-step estimation				
OLS	<u>0.2757</u>	<u>0.1184</u>	<u>0.7252</u>	<u>0.1159</u>
RF	<u>0.0671</u>	<u>0.0095</u>	<u>0.8849</u>	<u>0.0084</u>

Table 4.6: In-sample loss given default (LGD) estimation results. The methods used are ordinary least squares (OLS) and random forest (RF) regression. The determined performance measurements are mean absolute error (MAE), mean squared error (MSE), normalized regression error characteristic curve area (NAREC), and Theil inequality coefficient (TIC). These are calculated according to Equations (4.9), (4.10), (4.11), and (4.13) respectively. Comparing the direct estimation with the multi-step estimation approach, we underline the better results per method used.

In general, the multi-step approach is beneficial for both methods. We observe that the multi-step model has a distinct advantage over direct estimation. According to the additional information presented in Section 4.5.4, we attribute this outcome to the almost perfect classification. Upon closer inspection, we further note that the classification error of RF is half that of OLS. Consequently, the reduction in estimation error from the direct approach to the multi-step approach is even larger for RF than for OLS.

Additionally, the estimation results of ALGD and MLGD, presented in Table 4.9, show that estimating ALGD is easier than estimating MLGD. This confirms our expectations outlined in Section 4.3. Again, we note that RF has a markedly higher estimation accuracy than OLS. Consequently, the advantage of RF when forecasting ALGD and MLGD improves the classification accuracy.

4.5.2 Out-of-sample validation

The out-of-sample results that we report in Table 4.7 mostly confirm the findings of the in-sample validation. However, the performance gaps between the models are now less pronounced. In particular, the benefit of RF turns out to be less distinctive.

Method	MAE	MSE	NAREC	TIC
Direct estimation				
OLS	0.3505	0.1894	0.6538	0.1907
RF	0.3272	<u>0.1722</u>	0.6768	0.1725
Multi-step estimation				
OLS	<u>0.3387</u>	<u>0.1777</u>	<u>0.6655</u>	<u>0.1782</u>
RF	<u>0.3233</u>	0.1772	<u>0.6813</u>	<u>0.1708</u>

Table 4.7: Out-of-sample loss given default (LGD) estimation results. The methods used are ordinary least squares (OLS) and random forest (RF) regression. The determined performance measurements are mean absolute error (MAE), mean squared error (MSE), normalized regression error characteristic curve area (NAREC), and Theil inequality coefficient (TIC). These are calculated according to Equations (4.9), (4.10), (4.11), and (4.13) respectively. Comparing the direct to the multi-step estimation approach, we underline the better results per method used.

Considering the out-of-sample outcomes more closely, we again note that the RF models strictly outperform both OLS models. Moreover, in line with the in-sample results, we find that the multi-step approach is beneficial for both methods out-of-sample. With OLS, the multi-step approach notably outperforms direct estimation for all applied performance measurements. With RF, the multi-step approach is also advantageous in terms of MAE, NAREC, and TIC, but not concerning MSE.

Although, out-of-sample each multi-step approach has a notable advantage over the respective direct estimation model, this advantage is not, in absolute terms, as significant as in-sample. The reason for this is that the classification error in Table 4.10 increases similarly for RF and OLS in the out-of-sample validation compared with in-sample. Surprisingly, the classification error does indeed increase slightly more for RF. Consequently, in contrast to the in-sample results, out-of-sample the improvement of using the multi-step approach is more distinctive for OLS than for RF.

Typically, an increased classification error particularly affects the MSE. Compared to MAE, NAREC, and TIC, the MSE penalizes large deviations of estimates from their realized value stronger. Consequently, for the multi-step approach even a small number of falsely classified observations might increase the MSE significantly. That might be the case, even if the estimates are more accurate in general compared to direct estimation. Interestingly, we observe this effect only for RF but not for OLS. Apparently, OLS corrects for the bias of incorrectly classified contracts by estimating conditional expectations.

Again, as in-sample, we note more accurate estimations of ALGD than of MLGD. Moreover, Table 4.9 shows that RF once more produces lower errors than OLS. However, in line with the estimation of the overall LGD, out-of-sample the advantage of RF over OLS is less pronounced than in-sample.

4.5.3 Out-of-time validation

Our results of the most realistic scenario, the out-of-time validation, are presented in Table 4.8. Concerning the multi-step approach, we find that the outcomes confirm our in-sample and out-of-sample findings. However, the results differ concerning the direct estimation models.

Method	MAE	MSE	NAREC	TIC
Direct estimation				
OLS	0.3451	0.1876	0.6632	0.1959
RF	0.3457	<u>0.1830</u>	0.6605	0.2003
Multi-step estimation				
OLS	<u>0.3372</u>	<u>0.1778</u>	<u>0.6694</u>	<u>0.1897</u>
RF	<u>0.3412</u>	0.1858	<u>0.6611</u>	<u>0.1963</u>

Table 4.8: Out-of-time loss given default (LGD) estimation results. The methods used are ordinary least squares (OLS) and random forest (RF) regression. The determined performance measurements are mean absolute error (MAE), mean squared error (MSE), normalized regression error characteristic curve area (NAREC), and Theil inequality coefficient (TIC). These are calculated according to Equations (4.9), (4.10), (4.11), and (4.13) respectively. Comparing the direct to the multi-step estimation approach, we underline the better results per method used.

Considering these, we find RF to be no longer strictly advantageous. Instead, we rather observe better outcomes for OLS in terms of MAE, NAREC, and TIC. We attribute this finding to the overly good adaptation of RF to the training data, which becomes obvious when we look at the in-sample accuracy discussed in Section 4.5.1. Therefore, RF seems to experience difficulties with validation sets that differ significantly from the training sets. Related literature frequently observes relatively poor out-of-sample estimates and, in particular, poor out-of-time estimates of complex models with an excellent in-sample performance compared to OLS. Hartmann-Wendels et al. (2014) address this phenomenon concerning finite mixture models and Töws (2014) observes similar results in particular for RF.

A closer examination of the out-of-time outcomes shows that the multi-step approach still has a general advantage over direct estimation. Being most accurate out-of-time, the multi-step approach with OLS clearly outperforms the respective direct estimation, independently of the applied performance measurement. However, the advantage of the OLS multi-step model over its direct estimation is slightly smaller than out-of-sample. This result could have been expected because classification is even more difficult out-of-time. The reason for this is that validation data might differ significantly from training data. This effect is documented by the somewhat increased classification error in Table 4.10.

Analyzing the outcomes of the multi-step approach and direct estimation with RF, we observe many similarities to the out-of-sample approach. As in the out-of-sample validation, the multi-step approach with RF outperforms the respective direct model in terms of MAE, NAREC, and TIC. However, the multi-step approach produces a higher MSE. With respect to the classification errors in Table 4.10, we attribute the latter again to a small number of falsely classified observations. As mentioned above, such false classifications result in a rather large deviation of predicted LGD from realized LGD. By improving the estimation accuracy using the multi-step approach, RF achieves better MAE and MSE values than direct OLS. Nevertheless, when we compare the results of both multi-step models, OLS remains advantageous throughout. Moreover, in line with the out-of-sample findings, the benefit of using the multi-step approach is again more distinctive for OLS than for RF.

The estimation of ALGD and MLGD conforms to our expectations. We still note that ALGD estimation is more accurate than that of MLGD. On closer inspection of the results presented in Table 4.9, we see that RF again produces at least slightly lower out-of-time errors than OLS when estimating MLGD. However, in contrast to the in-sample and out-of-sample findings, OLS becomes somewhat advantageous when estimating ALGD. We attribute these comparatively inaccurate estimates of RF to its general difficulties in forecasting unseen observations from future periods. These results might also play a part in the advantage of the multi-step approach over the direct model being smaller when we use RF instead of OLS.

4.5.4 Further estimation and classification

Calculating separate LGD ratios and classifying the contracts increases the complexity of the estimation process from a methodological perspective. Recent studies show that greater complexity might negatively influence estimation accuracy (see, e. g., Qi and Zhao (2011)). However, in previous sections we have shown that our multi-step approach is clearly advantageous compared to direct estimation. We attribute this to the fact that our approach is based on economic considerations. Nevertheless, the accuracy of the final LGD estimation in our multi-step approach crucially depends on the estimation and classification accuracy in each step.

	In-sample		Out-of-sample		Out-of-time	
Method	MAE	MSE	MAE	MSE	MAE	MSE
ALGD estimation						
OLS	0.2884	0.1385	0.2945	0.1455	0.3097	0.1583
RF	0.1256	0.0298	0.2739	0.1385	0.3101	0.1588
MLGD estimation						
OLS	0.3758	0.2227	0.3829	0.2310	0.3836	0.2211
RF	0.1629	0.0452	0.3557	0.2103	0.3780	0.2186

Table 4.9: Asset-related loss given default (ALGD) and miscellaneous loss given default (MLGD) estimation results. The realized ALGD and MLGD are calculated according to Equation (4.3). The methods used are ordinary least squares (OLS) and random forest (RF) regression. The determined performance measurements are mean absolute error (MAE) and mean squared error (MSE). These are calculated according to Equations (4.9) and (4.10) respectively. The table summarizes the results of the three validation techniques: in-sample, out-of-sample, and out-of-time.

We first analyze the estimations of ALGD and MLGD, outlined in Table 4.9. Moreover, comparing the results of ALGD with those of direct LGD estimation, shown in Tables 4.6–4.8, we find that the estimates of ALGD are significantly more accurate than those of LGD. Additionally, in particular, in-sample and out-of-

Method	In-sample	Out-of-sample	Out-of-time
OLS	0.0248	0.2191	0.2484
RF	0.0100	0.2215	0.2364

Table 4.10: Classification results of classifying according to Equation (4.7). The classification error is calculated according to Equation (4.14). We use random forest classification in each case. The incorporated estimates of asset-related loss given default (ALGD) and miscellaneous loss given default (MLGD) from step one of our approach are estimated by ordinary least squares (OLS) and random forest (RF) regression. The table summarizes the results of the three validation techniques: in-sample, out-of-sample, and out-of-time.

sample, the results of MLGD are only slightly worse than their LGD counterparts. This small difference seems to play a part in our multi-step approach having an advantage over direct estimation. As regards direct estimation, we further observe that RF outperforms OLS both in-sample and out-of-sample regarding ALGD and MLGD. This effect is particularly evident in-sample. Out-of-sample the advantage of RF over OLS is less pronounced because the level of the estimation error increases significantly for RF, but remains stable for OLS. Consequently, RF benefits more from using the multi-step approach in-sample than OLS, whereas out-of-sample the opposite holds. Moreover, out-of-time, the above-mentioned difficulties of RF in forecasting unseen observations result in slightly more accurate MLGD estimations than OLS, but worse ALGD predictions.

The second crucial aspect of generating accurate LGD estimations with our multi-step approach is the classification. After estimating ALGD and MLGD with OLS or RF in the first step, classification is performed by random forest classification. Because the estimates of ALGD and MLGD are used to classify the contracts, we report the classification results labeled OLS and RF in Table 4.10. We find that the classification error varies significantly according to the validation technique. As expected, the classification is very precise in-sample but at about 20 times this rate out-of-sample and out-of-time. Nevertheless, classification remains sufficiently accurate as the multi-step approach still yields more accurate LGD predictions than direct estimation. However, the advantage is not as pronounced as in-sample. In general, it should be noted that despite the advantage of our multi-step approach over direct estimation, the concrete accuracy of the final LGD estimation depends on the applied method. For instance, if OLS generates better direct estimates than RF, its multi-step approach also produces more accurate results than that of RF.

Classification accuracy is important not only from a methodological perspective but also from an economic point of view. Based on the outcome of the classification the lessor might decide to restrict the workout process to the disposal of the leased asset. However, a false restriction results in waiving additional payment collection during the workout process and negatively affects the realized LGD. While classification is almost perfect in-sample, we see in Table 4.10 that out-of-sample and out-of-time classification is less reliable. Here, the classification errors are about 22% and 24%, respectively. False classification typically arises from classification probabilities near 50%, indicating that classification is ambiguous. These are probably cases in which a lessor does not restrict the workout process to the asset's disposal, although our classification would suggest this. To address such ambiguous cases, we present in Table 4.11 the classification results for contracts with a classification probability of below 25% or above 75%. As expected, we note that the classification error decreases consistently. In particular, regarding OLS we note significantly lower classification errors of about 13% out-of-sample and 17% out-of-time. Consequently, for these contracts the classification is clearly more reliable and seems to be suitable for practical use.

Finally, in order to make the comparison process most transparent, we want to elaborate on the direct regression of LGD with OLS. The direct in-sample estimation achieves an adjusted R^2 of 0.15, which is comparable to Bellotti and Crook (2012) with 0.13. The main drivers are the class of the leased asset, the term to default, and the type of the contract. Out-of-sample, pseudo R^2 amounts

Method	In-sample	Out-of-sample	Out-of-time
OLS	0.0000	0.1336	0.1726
RF	0.0000	0.1745	0.2089

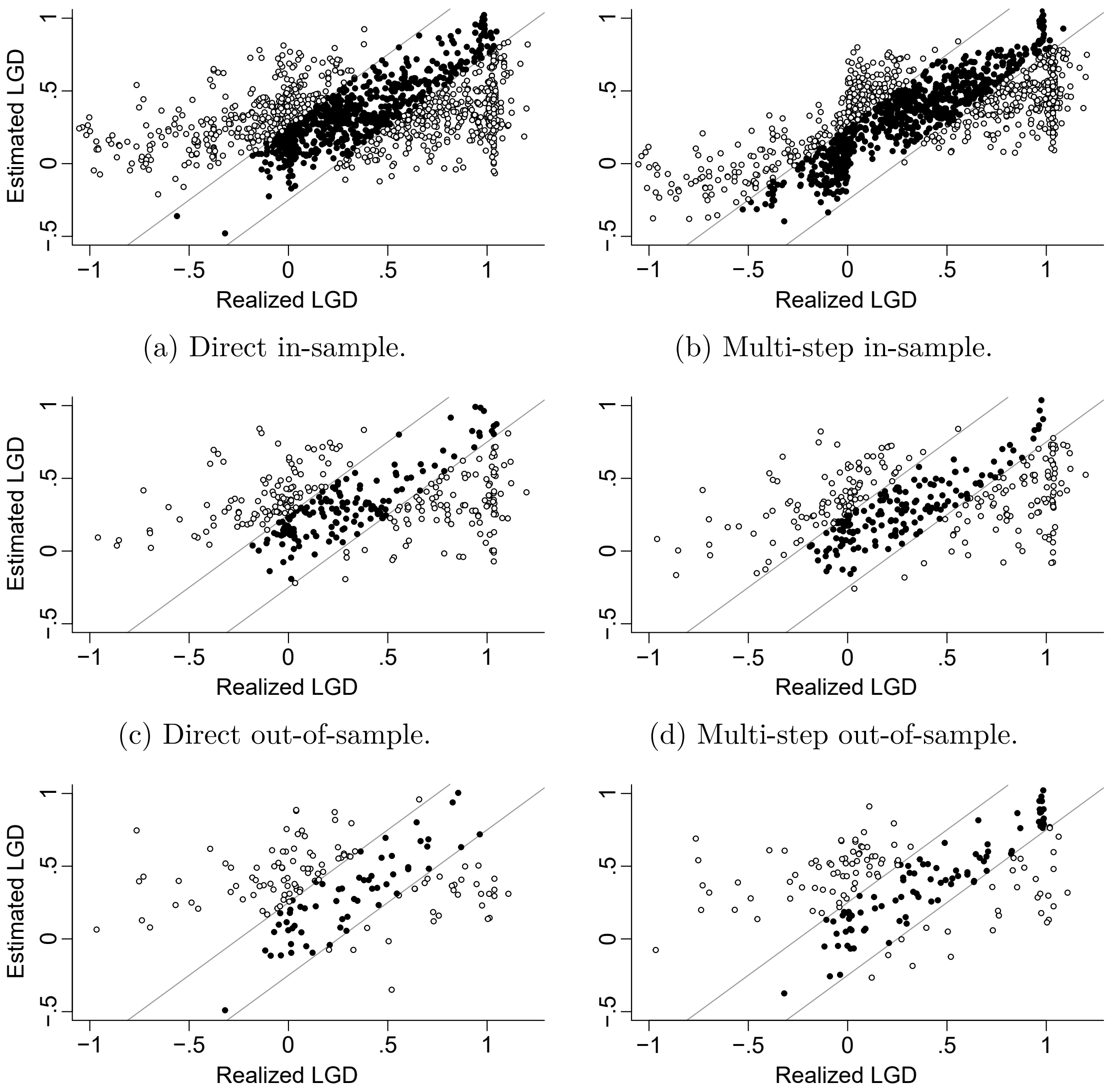
Table 4.11: Classification results of classifying according to Equation (4.7), considering exclusively classification probabilities below 25% or above 75%. The classification error is calculated according to Equation (4.14). We use random forest classification in each case. The incorporated estimates of asset-related loss given default (ALGD) and miscellaneous loss given default (MLGD) from step one of our approach are estimated by ordinary least squares (OLS) and random forest (RF) regression. The table summarizes the results of the three validation techniques: in-sample, out-of-sample, and out-of-time.

to 0.15 on average. In this model, the LGD is additionally significantly influenced by the lessee's default reason. Standard errors are marginally higher compared to the in-sample regression. In the out-of-time regression pseudo R^2 increases to 0.17 on average. The estimation is additionally positively driven in the case of retail customers. Again, standard errors increase slightly.

4.5.5 Interpretation

The previously discussed results clearly show the benefit of our multi-step approach compared to direct LGD estimation in terms of the applied performance measurements. While the chosen measurements are convenient for a precise comparison of the models, the scatter plots in Figure 4.5 provide an additional visual proof of our findings. In particular, the figures allow a more detailed analysis than the aggregated measures MAE, MSE, or NAREC. In these figures, the perfect estimation of LGD would be located on the diagonal through the plot's origin. We draw two diagonal lines to frame a 0.5-wide interval around the perfect estimation. The interval contains all estimates that are close to the realized LGD. These are displayed as solid points.

The scatter plots in Figure 4.5 refer to OLS, but the outcomes are similar concerning RF. According to Figure 4.5a, in-sample direct estimation produces a large number of accurate estimates. Nevertheless, the multi-step estimation in



(a) Direct in-sample. (b) Multi-step in-sample.

(c) Direct out-of-sample. (d) Multi-step out-of-sample.

(e) Direct out-of-time. (f) Multi-step out-of-time.

Figure 4.5: Visual comparison of realized and estimated loss given default (LGD). Figures a, c, and e display direct ordinary least squares (OLS) estimations in the in-sample, out-of-sample, and out-of-time validation respectively. Out-of-sample we randomly choose and display one run out of 1,000. Out-of-time we plot estimates and realizations of LGD of one year. Figures b, d, and f show the counterparts of the multi-step approach. The simple diagonal lines frame a 0.5-wide interval to highlight estimates close to their realized value. Additionally, these points are solid, whereas points outside the interval and, thus, far from their realized LGD are hollow.

Figure 4.5b generates a larger number of accurate estimates on the whole range of realized LGD values. In particular, due to a downward shift in estimation, it is visibly better than direct estimation for realized LGDs smaller than 0. Regarding the out-of-sample estimates in Figures 4.5c and 4.5d, we also note that the multi-step approach is again more accurate than direct estimation. To be precise, we observe a significantly higher concentration of estimates within the drawn interval for the multi-step model than in the case of direct estimation. This observation is particularly true for realized LGDs larger than about 0.3. Moreover, again the multi-step estimates generally tend to converge toward their realized value. The outcomes of the out-of-time validation in Figures 4.5e and 4.5f show a similar picture to that of in-sample and out-of-sample. Compared to direct estimation the predictions of the multi-step approach move closer to the diagonal lines from outside the interval. The increased number of estimates within the drawn interval indicates notably precise predictions for realized LGDs larger than 0.5.

For all three validation techniques the scatter plots in Figure 4.5 confirm that the estimates of the multi-step approach tend to converge toward their realized value. Therefore, the results of our performance measurements should not be affected by outliers.

Our results in Section 4.5 and the scatter plots in Figure 4.5 clearly show that the proposed multi-step approach outperforms direct estimation of the LGD. To evaluate the results in the context of related literature, we summarize the results of several studies in Table 4.12. Yao et al. (2015) argue that it is hard to compare empirical results when using different data and information sets. Nevertheless, they compare absolute values of R^2 from selected literature on LGD prediction performance. Instead of using absolute values we propose examining the improvement of the estimation accuracy of a model compared to OLS. Almost all related studies use the latter as a benchmark. For comparison, we focus on MAE, MSE, and root mean squared error (RMSE). More precisely, in Table 4.12 we present the maximal improvement of a study's best model compared to OLS regarding the respective performance measurement for out-of-sample and out-of-time validation.

Studies	Data	Best technique	ΔMAE	ΔMSE	ΔRMSE
Out-of-sample validation					
Bastos (2010)	SME loans	RT			6.9[6]
Loterman et al. (2012)	Four types of loans	SVM and NN	5.0		
Zhang and Thomas (2012)	Personal loans	OLS	3.7		
Bastos (2013)	Loans and bonds	RT ensemble	28.0	25.0	
Hartmann-Wendels et al. (2014)	Leases	Model tree	5.5		0.8
This study	**Leases**	**Multi-step approach**	**7.8**	**6.5**	
Out-of-time validation					
Bastos (2010)	SME loans	RT			6.7[6]
This study	**Leases**	**Multi-step approach**	**2.3**	**6.2**	

Table 4.12: Performance improvements in loss given default (LGD) estimation literature. The table reports the maximal percentage improvement of a study's best model compared to direct ordinary least squares (OLS) regression. The error measurements are: mean absolute error (MAE); mean squared error (MSE); and root MSE (RMSE). The techniques are: regression tree (RT); support vector machine (SVM); and neural network (NN).

Across the performance measurements, the authors primarily achieve improvements in the range from 2% to 10%. One major exception is Bastos (2013) with improvements around 25%. This exceptional improvement in estimation accuracy might be due to specific characteristics of the data used. The most frequently reported performance measurement in the literature is MAE. Out-of-sample our multi-step approach clearly achieves the highest MAE improvement among the studies considered, except for Bastos (2013). Concerning MSE or RMSE, Bastos (2010), in particular, reports a promising increase in estimation accuracy using regression trees. However, instead of OLS, he uses the historical average as benchmark. This outcome should be treated with caution. According to the results of Hartmann-Wendels et al. (2014), OLS performs at least 3 percentage points better than the historical average in terms of RMSE. Yashkir and Yashkir (2013) find a similar deviation. Therefore, comparing the regression tree results of Bastos (2010) with OLS, the improvement would probably not exceed 4%. Consequently, our multi-step approach also obtains competitive results in terms of MSE and

[6] Bastos (2010) uses the historical average as benchmark. Hence, the value reported is the performance increase of RT compared to the historical average.

RMSE.

Currently, several studies report out-of-sample errors, but out-of-time results are still scarce. Hence, concerning the latter, we can hardly evaluate our multi-step approach. Nevertheless, when we consider the benchmark used by Bastos (2010), our multi-step approach seems to generate good out-of-time results. Moreover, we emphasize that our multi-step approach can indeed perform better than direct OLS out-of-time. As our results of the direct estimation with RF indicate, it is not common for complex models that perform well out-of-sample also to produce stable and accurate out-of-time estimates.

In addition to the comparison with direct estimation methods, we compare our approach to an existing multi-step model. We apply the methodology introduced by Bellotti and Crook (2012) to our dataset. In the first step, they use a logistic regression to classify the data according to the contracts' LGD into three classes: (1) LGD < 0, (2) $0 <$ LGD < 1, and (3) $1 <$ LGD. The second step calibrates an OLS linear regression model to data of the second class. The LGD estimate then is set to 0 for contracts of class (1), to LGD* from the OLS prediction model for class (2), and to 1 for class (3) contracts. Gürtler and Hibbeln (2013) pick up this method and improve it by accounting for pit falls of LGD estimation and additional explanatory factors, such as the contracts' recovery from default.

Using the same explanatory factors established on our dataset within this study, we find that the estimation approach of Bellotti and Crook (2012) is inferior throughout our three validation methods. For most performance measurements the results are even worse compared to OLS direct estimation.

While the proposed multi-step approach, in general, is able to generate more precise estimates than direct OLS, according to Table 4.12 the improvement of advanced estimation methods compared to direct OLS appears to be rather small in general. However, in order to evaluate the economic benefit of the increased estimation accuracy by implementing advanced estimation approaches it is essential to consider the value of the underlying portfolio. Referring to the mean EAD of the employed dataset presented in Table 4.1 and the out-of-time results shown in Table 4.8, our multi-step approach with OLS predicts the loss of a contract on average by €420 more accurate than direct OLS. Consequently, particularly by taking into account that the value of the entire portfolio considerably exceeds the exposure of the studied subportfolio including exclusively defaulted contracts, the economic benefit by the improved estimation accuracy is crucial. Obviously, the difference between the proposed multi-step approach and other advanced estimation methods is not as pronounced, however the gap is still economically relevant. Moreover, unlike other advanced estimation methods, our multi-step model additionally provides decision support concerning the actions that should be taken during the workout process.

4.6 Conclusion

The development of an appropriate and dynamic model for estimating LGD requires the consideration of mathematical aspects and economic factors. For defaulted leasing contracts, we argue that detailed consideration of the revenues during the workout process is a key driver for improving LGD forecasting accuracy. Instead of the traditional holistic consideration of LGD, we separate LGD into asset-related and miscellaneous parts. This separation is economically reasonable because, typically, cash flows have different sources. To account for the different revenues at the time of contracts' default, we estimate an ALGD and a MLGD.

Leasing companies are experts in evaluating and disposing of their leased assets. The estimation of the ALGD takes this expertise into account. Moreover, together with the estimated MLGD, it provides decision support for actions to be taken during the workout process. We show that ALGD is theoretically an upper boundary to the LGD. Likewise, the estimation of MLGD yields economic value. Its value indicates whether the effort of collecting overdue payments during the workout process will be rewarding or rather unprofitable considering incurred workout costs. Consequently, we present a guideline for organizing the workout process, i. e., if workout costs are expected to exceed collected payments, the workout process should be restricted to the disposal of the leased asset.

This finding is particularly interesting because cash flows from the asset's disposal are positive in 99% of the cases, net of disposal costs. However, for 35% of the contracts, the collection of miscellaneous payments turns out to generate losses due to the incurred costs. We find that following our suggestion to restrict the workout process to the asset's disposal would, in general, significantly reduce the average LGD. With our data, the reduction of the average LGD amounts to 10% or €2,250,000 in absolute losses.

Based on the sophisticated economic separation of the LGD, we introduce a new multi-step LGD estimation approach. We apply our approach to a real-life dataset of a German leasing company and perform in-sample, out-of-sample, and out-of-time validation. While the approach supports the workout process, we find that the separation of LGD is very beneficial for its estimation accuracy. We apply OLS and RF regression to our approach. With both methods, we note a significant increase in estimation accuracy compared to the benchmarking results of the respective direct estimation. The proposed multi-step approach is more complex than direct estimation. However, the increase in complexity does not lead to overfitting, which is a common concern of advanced LGD estimation models. Nonetheless, the interpretability of the variables' influence might suffer slightly. However, to put it in Bastos (2013) words, it is often the case that simplicity has to be sacrificed in order to achieve a higher degree of precision.

5 Summary and conclusions

This thesis includes three essays dealing with the modeling and estimation of the LGD for leasing contracts. The first essay focuses on the methodological aspects of estimating the LGD and compares different estimation approaches in order to investigate which methods are particularly suitable to predict the LGD of leases. Motivated by insights gained in the first study, the second essay analyzes the factors driving the LGD of leasing contracts and examines whether and to what extent the key drivers are different for the individual leasing companies. Based on the findings of the first two studies, the third essay introduces an advanced approach for estimating the LGD of leases that explicitly considers the specific characteristics of the leasing business.

Basically, there are many different methods that can be used to estimate the LGD. By comparing the outcomes of several estimation approaches, the first essay demonstrates that the quality of LGD estimates varies significantly depending on the used method. Moreover, it turns out to be indispensable to consider and interpret the results of different validation techniques in order to reliably evaluate which methods are particularly suitable to predict the LGD of leasing contracts. Specifically, it is exposed that in-sample results might be significantly misleading when estimating out-of-sample LGDs, which are crucial for proper risk management and are required for regulatory purposes.

Among the analyzed estimation methods, FMMs appear to be a less suited concept to predict the LGD of leases. Although the FMMs are able to reproduce the multimodality of the LGD density more properly than other applied approaches, in particular the out-of-sample estimations generated by the FMMs are quite poor. This observation basically supports the thesis that reproducing the LGD density is only of minor importance for precisely predicting the LGD. In fact, by generally providing the most accurate out-of-sample results of the approaches employed, a nonparametric method, namely the model tree M5', turns out to be particularly suitable to predict the LGD of leasing contracts. Unlike parametric methods, nonparametric methods make no assumptions concerning the distribution of the underlying data. This feature is obviously highly beneficial when estimating the LGD. Nevertheless, the results stress that the applicability of an estimation method depends on the lessor's specific characteristics and should be considered in each individual case. For example, if a lessor's database only contains a small number of observations it appears to be practical to predict the LGD by OLS linear regression.

Actually, the findings of the first study also reveal that the quality of LGD predictions depends not exclusively on the applied estimation method. The results illustrate that the prediction accuracy differs fundamentally between the studied leasing companies and is in particular determined by the used set of information. On this account, the second essay employs data from two lessors to analyze which determinants are driving the LGD of leases. Consistent with the findings of the first study, it turns out that the factors identified as drivers of the LGD are at least to some extent different for the considered leasing companies. In particular, the differences noted among the lessors refer to both the set of idiosyncratic factors driving the LGD and the determined relationship between the LGD and macroeconomic factors. The confirmation that the factors influencing the LGD vary depending on the analyzed company is especially important with regard to the heterogeneous results concerning the drivers of the LGD in the literature. Nevertheless, although it is crucial to analyze the potential drivers of the LGD individually for each lessor, the results affirm that the LGD of leasing contracts generally depends on factors that are related to the leased asset, such as, in particular, the type of the leased asset. Yet, it is also shown that, e. g., the contract structure has an impact on the LGD of leases. Moreover, the results expose that the estimation accuracy might be improved if in addition to idiosyncratic factors also macroeconomic factors are considered. Although the relationship between LGD and macroeconomic conditions depends on the characteristics of the leasing company and in particular its organization of the workout process, especially the estimates carried out at contract's execution benefit from including macroeconomic factors.

In summary, the results of the first two studies corroborate that the quality of LGD predictions generally depends on both the applied estimation method and the considered set of explanatory factors. However, due to the different nature of leasing companies, it is worthwhile to select the explanatory factors and also the estimation method individually for each lessor according to its specific characteristics. Nonetheless, the results of the second study likewise point out that across different lessors the LGD of leasing contracts generally depends on factors that are related to the leased asset. Hence, to further improve the prediction of leasing LGDs, it appears to be logical to develop advanced LGD estimation approaches which explicitly consider the peculiarities of the leasing business.

The third essay addresses this issue and uses data from a large German lessor to evolve such an advanced model for estimating the LGD of leases that expressly takes into account specific characteristics of the leasing business. Motivated by the economic consideration that the revenues received during the workout process of a defaulted leasing contract come from two distinct payment sources, the LGD is separated into an asset-related and a miscellaneous part. Based upon this economic separation of the LGD into ALGD and MLGD, a multi-step LGD estimation model for leases is developed and its performance is evaluated using in-sample, out-of-sample and out-of-time testing. The results confirm the hypothesis that the prediction of leasing LGDs can be improved by the development of advanced estimation models that explicitly consider the peculiarities of the leasing business. Compared to the results of established estimation approaches, the predictions generated by the introduced multi-step estimation model are considerably more accurate. In particular, there is no evidence that the proposed and economically motivated multi-step estimation model is liable to overfitting, which is a common concern of advanced LGD estimation approaches. Moreover, also unlike other advanced estimation approaches, the developed multi-step LGD estimation model provides valuable interim results that can be used as a decision support for actions to be taken during the workout process of a defaulted leasing contract.

Bibliography

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

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

Altman, E.I., Ramayanam, S., 2007. Default and Returns in the High-Yield Bond Market 2006 in Review and Outlook. Technical Report. Citigroup Global Markets.

Araten, M., Jacobs Jr., M., Varshney, P., 2004. Measuring LGD on Commercial Loans: An 18-Year Internal Study. The RMA Journal 4, 96–103.

Basel Committee on Banking Supervision, 2003. Overview of The New Basel Capital Accord. Bank for International Settlements.

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

Bastos, J.A., 2013. Ensemble Predictions of Recovery Rates. Journal of Financial Services Research 46, 177–193.

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

Bi, J., Bennett, K.P., 2003. Regression Error Characteristic Curves, in: Proceedings of the 20th International Conference on Machine Learning.

Borio, C., Furfine, C., Lowe, P., 2001. Procyclicality of the financial system and financial stability: issues and policy options . BIS papers , 1–57.

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32.

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

Calabrese, R., 2014. Downturn Loss Given Default: Mixture distribution estimation. European Journal of Operational Research 237, 271–277.

Calabrese, R., Zenga, M., 2010. Bank loan recovery rates: Measuring and non-parametric density estimation. Journal of Banking & Finance 34, 903–911.

Caselli, S., Gatti, S., Querci, F., 2008. The Sensitivity of the Loss Given Default Rate to Systemic Risk: New Empirical Evidence on Bank Loans. Journal of Financial Services Research 34, 1–34.

Chalupka, R., Kopecsni, J., 2009. Modeling Bank Loan LGD of Corporate and SME Segments: A Case Study. Czech Journal of Economics and Finance 59, 360–382.

De Laurentis, G., Mattei, J., 2009. Lessors' recovery risk management capability. Managerial Finance 35, 860–873.

De Laurentis, G., Riani, M., 2005. Estimating LGD in the Leasing Industry: Empirical Evidence from a Multivariate Model, in: Altman, E., Resti, A., Sironi, A. (Eds.), Recovery Risk: The Next Challenge in Credit Risk Management. Risk Books. Chapter 8, 143–164.

Dermine, J., Neto de Carvalho, C., 2005. How to Measure Revcoveries and Provisions on Bank Lending: Methodology and Empirical Evidence, in: Altman, E., Resti, A. (Eds.), Recovery Risk: The Next Challenge in Credit Risk Management. Risk Books. Chapter 6, 101–119.

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

Dwyer, D., Korablev, I., 2009. Moody's KMV Losscalc®. Technical Report. Moody's KMV Company.

Eisfeldt, A.L., Rampini, A.A., 2009. Leasing, Ability to Repossess, and Debt Capacity. Review of Financial Studies 22, 1621–1657.

Elbracht, H.C., 2011. Statistische Methoden zur Quantifizierung und Schätzung des Loss Given Default. Ph.D. thesis. University of Cologne. 2011.

Emery, K., 2008. Strong Loan Issuance in Recent Years Signals Low Recovery Prospects for Loans and Bonds of Defaulted U.S. Corporate Issuers. Report Number 109457, Moody's Investors Service.

Frühwirth-Schnatter, S., 2006. Finite Mixture and Markov Switching Models. Springer Series in Statistics, Springer, New York.

Friedman, C., Sandow, S., 2005. Estimating Conditional Probability Distributions of Recovery Rates: A Utility-Based Approach, in: Altman, E., Resti, A. (Eds.), Recovery Risk: The Next Challenge in Credit Risk Management. Risk Books. Chapter 19, 347–359.

Frontczak, R., Rostek, S., 2015. Modeling loss given default with stochastic collateral. Economic Modelling 44, 162–170.

Frye, J., 2005. A False Sense of Security, in: Altman, E., Resti, A., Sironi, A. (Eds.), Recovery Risk: The Next Challenge in Credit Risk Management. Risk Books. Chapter 10, 187–200.

Gadd, A., Wold, H., 1964. The janus coefficient: a measure for the accuracy of prediction.

Gibilaro, L., Mattarocci, G., 2007. The selection of the discount rate in estimating loss given default. Global Journal of Business Research 1, 15–33.

Gibilaro, L., Mattarocci, G., 2011. The Impact Of Discount Rate Choice in Estimating The Workout LGD. Journal of Applied Business Research 27, 139–148.

Gray, J.B., Fan, G., 2008. Classification tree analysis using TARGET. Comput. Stat. Data Anal. 52, 1362–1372.

Grippa, P., Iannotti, S., Leandri, F., 2005. Recovery Rates in the Banking Industry: Stylised Facts Emerging from the Italian Experience, in: Altman, E., Resti, A. (Eds.), Recovery Risk: The Next Challenge in Credit Risk Management. Risk Books. Chapter 7, 121–141.

Grün, B., Leisch, F., 2007. Fitting finite mixtures of generalized linear regressions in R. Computational Statistics & Data Analysis 51, 5247–5252.

Gürtler, M., Hibbeln, M., 2013. Improvements in loss given default forecasts for bank loans. Journal of Banking & Finance 37, 2354–2366.

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

Gupton, G.M., Stein, R.M., 2005. LossCalc v2: Dynamic Prediction of LGD. Technical Report. Moody's KMV Company.

Han, C., Jang, Y., 2013. Effects of debt colletion practices on loss goven default. Journal of Banking & Finance 37, 21–31.

Hartmann-Wendels, T., Honal, M., 2010. Do Economic Downturns Have an Impact on the Loss Given Default of Mobile Lease Contracts? An Empirical Study for the German Leasing Market. Kredit und Kapital 43, 65–96.

Hartmann-Wendels, T., Miller, P., Töws, E., 2014. Loss given default for leasing: Parametric and nonparametric estimations. Journal of Banking & Finance 40, 364 – 375.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning. Springer New York.

Hlawatsch, S., Ostrowski, S., 2011. Simulation and estimation of loss given default. The Journal of Credit Risk 7, 39–73.

Hurt, L., Felsovalyi, A., 1998. Measuring Loss on Latin American Defaulted Bank Loans: A 27-Year Study of 27 Countries. Journal of Lending & Credit Risk Management 81, 41–46.

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

Johnston Ross, E., Shibut, L., 2015. What Drives Loss Given Default? Evidence From Commercial Real Estate Loans at Failed Banks. Working paper.

Khieu, H.D., Mullineaux, D.J., Yi, H., 2012. The determinants of bank loan recovery rates. Journal of Banking & Finance 36, 923–933.

Kohavi, R., 1995. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection, in: Joint Conference on Artificial Intelligence (IJCAI), Morgan Kaufmann. 1137–1143.

Lasfer, M.A., Levis, M., 1998. The determinants of the leasing decision of small and large companies. European Financial Management 4, 159–184.

Laurent, M., Schmit, M., 2005. Estimating "Distressed" LGD on Defaulted Exposures: A Portfolio Model Applied to Leasing Contracts, in: Altman, E., Resti, A., Sironi, A. (Eds.), Recovery Risk: The Next Challenge in Credit Risk Management. Risk Books. Chapter 17, 307–322.

Leow, M., Mues, C., 2012. Predicting loss given default (LGD) for residential mortgage loans: A two-stage model and empirical evidence for UK bank data. International Journal of Forecasting 28, 183–195.

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

Li, P., Qi, M., Zhang, X., Zhao, X., 2014. Further Investigation of Parametric Loss Given Default Modeling. Working paper.

Loterman, G., Brown, I., Martens, D., Mues, C., Baesens, B., 2012. Benchmarking regression algorithms for loss given default modeling. International Journal of Forecasting 28, 161–170.

Miller, P., 2015. Does the Economic Situation Affect the Loss Given Default of Leases? Working paper.

Miller, P., Töws, E., 2014. Berücksichtigung nichtlinearer Einflüsse bei der Schätzung von IRBA-Kennzahlen. Finanzierung Leasing Factoring 2, 56–60.

Mérõ, K., Zsámboki, B., Horváth, E., 2002. Studies on the procyclical behaviour of banks. MNB Occasional Papers. Magyar Nemzeti Bank (the central bank of Hungary).

Qi, M., Zhao, X., 2011. Comparison of modeling methods for Loss Given Default. Journal of Banking & Finance 35, 2842–2855.

Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann.

Renault, O., Scaillet, O., 2004. On the way to recovery: A nonparametric bias free estimation of recovery rate densities. Journal of Banking & Finance 28, 2915–2931.

Royston, P., Sauerbrei, W., 2007. Multivariable modeling with cubic regression splines: A principled approach. Stata Journal 7, 45–70.

Saunders, A., Allen, L., 2002. Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms. New York: John Wiley and Sons.

Schmit, M., 2004. Credit risk in the leasing industry. Journal of Banking & Finance 28, 811–833.

Schmit, M., Degouys, C., Delzelle, D., Stuyck, J., Wautelet, F., 2003. Credit Risk in the Leasing Business – A case study of low probability of default.

Schmit, M., Stuyck, J., 2002. Recovery Rates in the Leasing Industry. Working Paper presented at Leaseurope's Annual Working Meeting, Salzburg.

Schneider, P., Sögner, L., Veža, T., 2010. The Economic Role of Jumps and Recovery Rates in the Market for Corporate Default Risk. Journal of Financial and Quantitative Analysis 45, 1517–1547.

Theil, H., 1966. Applied economic forecasting. Studies in mathematical and managerial economics 4, 28.

Thorburn, K.S., 2000. Bankruptcy auctions: costs, debt recovery, and firm survival. Journal of Financial Economics 58, 337 – 368.

Töws, E., 2014. The impact of debtor recovery on loss given default. Working paper.

Wang, Y., Witten, I.H., 1997. Inducing Model Trees for Continuous Classes. In Proceedings.

Witten, I.H., Frank, E., Hall, M.A., 2011. Data Mining: Practical Machine Learning Tools and Techniques. Elsevier Science.

Yao, X., Crook, J., Andreeva, G., 2015. Support vector regression for loss given default modelling. European Journal of Operational Research 240, 528–538.

Yashkir, O., Yashkir, Y., 2013. Loss given default modeling: a comparative analysis. Journal of Risk Model Validation 7, 25–59.

Zhang, J., Thomas, L.C., 2012. Comparisons of linear regression and survival analysis using single and mixture distributions approaches in modelling LGD. International Journal of Forecasting 28, 204–215.

Curriculum Vitae

Personal Details

Patrick Michel Miller
Im Kalkfeld 5
51379 Leverkusen

0160 4186 991
P-Miller@gmx.de

Place of birth: Leverkusen
Day of birth: 6th March 1985

Education

10/2011 – present	PhD student, Department of Bank Management, *University of Cologne*
10/2004 – 12/2010	Diplom Business Mathematics, *University of Cologne* Thesis: Exakte und heuristische Verfahren zur Lösung von Range-Assignment Problemen

Professional Activities

10/2015 – present	Risk controller, *Deutsche Postbank AG, Bonn*
04/2012 – 09/2015	Research assistant, Department of Bank Management, *University of Cologne*

Leverkusen, January 17, 2017