ORIGINAL RESEARCH

The recovery rate for retail and commercial customers in Germany: a look at collateral and its adjusted market values

Peter-Hendrik Ingermann[1] · Frederik Hesse[1] ·
Christian Bélorgey[1] · Andreas Pfingsten[1]

Abstract   Based on a unique data set of 909 defaulted retail and commercial (self-employed and SMEs) credit customers in Germany, whose original loans were made by 123 different banks, our article confirms a significant positive influence of collateral, and of amicable agreements between the debtor and the bank (redemption), on the recovery rate [1 − loss given default (LGD)]. In a further analysis of collateral, systematic biases between the realized market price and the expected market values of real estate are revealed, even though the appraisal reports should have already considered all factors influencing the value. Using valuations that were adjusted for these recognized biases, we can increase the explanatory power of the underlying models. Moreover, we compare these models to models that apply, as is common practice in the banking industry, flat haircuts to collateral values and show the superior performance of our proposed approach.

✉ Peter-Hendrik Ingermann
   peter-hendrik.ingermann@wiwi.uni-muenster.de

   Frederik Hesse
   frederik.hesse@wiwi.uni-muenster.de

   Christian Bélorgey
   christian.belorgey@wiwi.uni-muenster.de

   Andreas Pfingsten
   andreas.pfingsten@wiwi.uni-muenster.de

[1]   Finance Center Muenster, University of Muenster, Universitaetsstr. 14-16, 48143 Muenster, Germany

1 Executive summary

Our study investigates the recovery rate [1 − loss given default (LGD)] of bank loans, a key parameter in the context of Basel II, for retail and commercial (self-employed and SMEs) customers in Germany. The recovery rate relates the proceeds and costs of realization to the outstanding amount at the time of default [exposure at default (EAD)]. It is included in the calculation of expected and unexpected losses in the credit portfolio and thereby influences the calculation of the value at risk (VaR) as well as (future) credit terms.

Based on a unique data set of 909 defaulted customers in Germany, whose loans were originally made by 123 different banks, our study shows that the collateralization of a loan is the major driver of the recovery rate for retail and commercial customers in Germany; this confirms earlier research on recovery rates. We also find that an amicable agreement between the debtor and the bank (redemption) has a significant positive influence on the recovery rate. Furthermore, we find systematic biases between the realized and the expected values of collateral in real estate, even though the appraisal reports should have already considered all factors influencing this value. Adjusting the collateralization values for these recognized systematic biases significantly improves the explanatory power of our model for recovery rates. We compare our model to models applying flat haircuts to the collateral values, a fairly common practice in the banking industry. Our study demonstrates that collateral values should not be taken as given. In particular, systematic biases may influence estimated LGDs, and should be corrected by appropriate adjustments. While the results may hold for a greater universe, they are particularly relevant for savings banks and cooperative banks because of the focus of these banks on the customer segment analyzed here.

2 Introduction

The probability of default (PD) and the recovery rate [1 − loss given default (LGD)] are the key parameters for risk-adjusted pricing of loans in the context of Basel II. The recovery rate relates the proceeds and costs from realization to the outstanding amount at the time of default [exposure at default (EAD)]. For internal risk management purposes, the recovery rate, the PD, the EAD, and their correlations are included in the calculation of expected and unexpected losses of the credit portfolio, and thereby influence the calculation of the value at risk (VaR) as well as (future) credit terms. Therefore, a bank's internal risk management has to develop a model that allows an accurate estimate of the recovery rate. A precise estimate, on one hand, can generate a competitive advantage and, on the other, can reduce problems arising from adverse selection due to small differences in loan spreads (e.g., Gürtler and Hibbeln 2013). Moreover, banks that prove to have appropriate internal models

for predicting the recovery rate are allowed to use advanced internal rating based (IRB) approaches. These approaches can result in reduced capital requirements. Hence, a detailed estimate of the recovery rate is beneficial from a regulatory perspective as well.[1]

While several prior studies analyze recovery rates for corporate bonds (e.g., Altman 1989; Renault and Scaillet 2004; Altman et al. 2005), the body of literature concerning bank loan recovery rates has been growing rapidly in recent years. The better data availability means that a large proportion of the literature examines the corporate sector (e.g., Franks et al. 2004; Grunert and Weber 2009). Literature on the recovery rate for retail customers and SMEs is also growing steadily. Our work contributes to this strand of literature in multiple ways.

First, we fill a gap in recovery rate research in Germany. While a few German studies analyze corporate customers (e.g., Grunert 2005; Grunert and Volk 2008; Grunert and Weber 2009; Franks et al. 2004), only one study so far (Gürtler and Hibbeln 2013) has analyzed a sample of loans to retail and commercial customers. However, the focus of Gürtler and Hibbeln (2013) lies in the field of consumer loans, with an average EAD of EUR 9665 in their sample. Although the German banking system is, in terms of total assets, among the largest in the world, a study of SME loans, as well as a study of larger retail loans (not consumer loans), is missing so far. Our study is also linked to the body of literature on the recovery rate of real estate loans (see, e.g., Calem and LaCour-Little 2004; Qi and Yang 2009; Leow and Mues 2012; Tong et al. 2013; Park and Bang 2014). A study for Germany that predominantly analyzes real estate loans is also missing so far. Loans to retail customers and SMEs are of particularly high importance in Germany, since more than 50 % of the total loan volume is granted to this specific customer group (Deutsche Bundesbank 2014). Our results confirm the intuitive expectation of a positive influence of collateral on the recovery rate for retail and commercial customers. A major fraction of the variance of the recovery rates is explained by our models, which is mainly caused by the fact that collateral in real estate (the predominant collateral in our data set) usually carries a high fundamental value. Furthermore, our results demonstrate that in 40 % of the cases an amicable agreement (redemption) between the bank and the debtor can be found. A positive influence on the recovery rate can be shown for cases in which an amicable agreement can be achieved.

Secondly, the existing studies on recovery rates usually use data from just one bank (e.g., Dermine and Neto de Carvalho 2006; Grunert and Weber 2009; Leow and Mues 2012; Zhang and Thomas 2012; Tong et al. 2012, 2013), which raises the issue of the generality of the empirical results. Our data set comes from BAG, a bank that specializes in the acquisition of non-performing loans. Therefore, we can analyze a sample of 909 defaulting loans that were made by 123 different banks, and our findings are not driven by the particular actions of a single bank.

---

[1] A study conducted by the Basel Committee on Banking Supervision (2006) concludes that for almost all banking segments, the capital requirements are reduced when switching from the standard approach to the IRB approach.

Thirdly, other studies have also been able to prove the intuitive view that recovery rates are significantly higher if loans are collateralized (e.g., Thorburn 2000; Grunert and Weber 2009; Qi and Yang 2009; Bastos 2010; Gürtler and Hibbeln 2013). However, most recovery rate studies consider the collateral value as given (exceptions are, for example, Jokivuolle and Peura 2003; Somers and Whittaker 2007; Leow and Mues 2012). We take a closer look at the valuation of collateral itself and, in particular, at the valuation of collateral in real estate. A couple of studies for the German market analyze the proceeds ratio of collateral (realized proceeds divided by the fundamental value of the collateral assumed by the bank) (Franks et al. 2004; Schaaff 2009; Grunert 2010). These studies show that the realized proceeds are, on average, smaller than the original valuation of the collateral. Our results confirm these earlier findings. For the originating bank, the valuation (market value) of the collateral in real estate is based on the appraisal report, which is a proxy for the expected selling price. Since the appraisal report should consider all possible factors that have an influence on the property value, the realized selling price should not differ systematically from the expected value. For the first time in recovery rate research, our method uses an index based on postal codes to measure the attractiveness of a property's location and to arrange for the finest possible discrimination between different locations and property types. We demonstrate that the condition as well as the location of the property, even on average, is not adequately considered in the appraisal reports. To correct for this systematic bias, we develop a model to predict the proceeds ratio to eliminate the bias in the valuation of properties.

Fourthly, banks commonly apply a flat haircut to the collateral value. For collateral in real estate, this haircut ranges between 10 and 40 % in Germany (see, e.g., Grunert and Weber 2009; Bruhn 2009; Pfnür 2011). Hence, we apply flat haircuts to the collateral values to adjust the collateralization ratios and explain the recovery rate again. In comparison to the flat haircuts, we use predicted proceeds ratios to adjust the market value of real estate. We thereby obtain an adjusted collateralization ratio that can also be used to explain the recovery rate. Both adjustments of the collateralization ratios considerably increase the explanatory power of the original model, whereas the haircut models perform worse in terms of explanatory power in comparison to the models using model-adjusted collateralization ratios.

Furthermore, the validity of our results can be demonstrated using various out-of-sample tests. Our findings indicate that the combined analysis of the collateral and borrower levels should be an important part of the analysis of factors influencing the recovery rate in future empirical studies.

Our study provides valuable insights into the practice of banking with regard to the design of adequate models, and provides optimization advice with respect to the collection and processing of data. The results are of particular importance for small and medium-sized banks, such as savings banks and cooperative banks. The customer segment analyzed here is the core clientele for these businesses. And banks, because of to their size, often do not have the appropriate number of cases to perform a comparable analysis.

Our paper is structured as follows. In Sect. 3, we derive our research hypotheses and present the data set. Then, we analyze our hypotheses using bivariate and multivariate analyses in Sect. 4. Afterwards, we check the valuation of collateral in real estate for systematic biases. In Sect. 5, we correct for the recognized biases and use the adjusted collateralization ratios for a new approach to explaining the recovery rate. Section 6 summarizes the results.

3 Empirical analysis of the recovery rate on the borrower level

3.1 Derivation of the research hypotheses

In contrast to the practice for defaulted bonds, it has become common practice to use the so-called workout method for the determination of the recovery rate on bank loans. When a debtor defaults, the individual proceeds (e.g., the realization proceeds from collateral) as well as the costs of the workout-process (e.g., material and personnel costs, and legal expenses) are determined for each time t. The proceeds and expenses are discounted to the time of default to take the present value of money into account.[2] Putting the aggregated net proceeds in relation to the EAD at the time of default (EAD_0) results in the recovery rate:

$$\text{Recovery rate} = \frac{\sum_{t=0}^{T} \dfrac{\text{Proceeds}_t - \text{Expenses}_t}{(1 + r_t)^t}}{EAD_0}. \tag{1}$$

Empirical studies analyzing the recovery rate for bank loans have analyzed a number of potential influencing factors. However, the results regarding most of the factors are ambiguous. Franks et al. (2004), for example, identify a significant, positive influence of macroeconomic factors on the recovery rate, while these factors have no influence according to a study by Dermine and Neto de Carvalho (2006). Yet, nearly all studies (e.g., Carty et al. 1998; Gupton et al. 2000; Thorburn 2000; Bos et al. 2002; Araten et al. 2004; Emery et al. 2004; Franks et al. 2004; Hamilton et al. 2004; Grunert and Weber 2009; Qi and Yang 2009; Bastos 2010; Gürtler and Hibbeln 2013) demonstrate that the recovery rate is higher if loans are collateralized.

We also expect a higher recovery rate for loans with a higher collateralization ratio, since the bank can generate higher proceeds from the collateral provided. As the null hypothesis is to be rejected, we state:

Hypothesis 1 (H1)   The collateralization ratio has no significant effect on the recovery rate.

---

[2] The risk-free interest rate r, the contractually agreed interest rate (the variable interest rate might be weighted so that it is time related), the return on equity, or the effective interest rate in the sense of IAS 39, can all be used as the interest rate r for the discounting factor. No explicit specifications are made by the Basel Committee.

Besides collateral accounted for on a value basis (in EUR), in the banking practice there are additional types of collateral that, although they may result in realization proceeds, are not taken into account because of legal uncertainty or market dependence in general. By analogy to the collateralization ratio, we would expect those to influence the recovery rate positively and formulate as the null hypothesis to be rejected:

Hypothesis 2 (H2)   Additional collateral that is not accounted for on a value basis (in EUR) has no effect on the recovery rate.

Studies by Elsas and Krahnen (2002), Grippa et al. (2005), and Grunert and Weber (2009) have identified the process by which borrowers are put into default as another important factor influencing the recovery rate. Generally, there are three essential processes through which borrowers are put into default: continuing their business activities in line with a restructuring, an amicable settlement (redemption), and the liquidation of the loan and the provided collateral. For economic reasons, continuing business activities is only applicable to sufficiently large debtors, which are usually corporate customers.[3] Active restructuring, that is, the bank arranging an operational and strategic realignment, is rarely possible or economically reasonable for retail and commercial customers.

With retail and commercial customers, the banks prefer direct negotiations in an attempt to achieve an amicable agreement or, if necessary, they settle for receiving only part of the receivables. If an amicable settlement can be achieved, redemption[4] is agreed on.[5] The debtor then makes the previously agreed redemption payment, while the bank releases the provided collateral in return. The bank will, however, only accept a redemption payment if the payment that is offered equals at least the sum of the expected discounted (net) returns from the collateral and the cost of personal arbitration measures against the debtor (lower bound). On the other hand, the debtor will not pay back more than the outstanding liability (upper bound). In the case of redemption, the redemption payment should lie within the interval identified above. The difference between the realized redemption payment and the lower bound constitutes the premium. If negotiations fail or are not possible because

---

[3] In the study by Grunert and Weber (2009), who analyze 120 defaulting corporate customers of a single German bank, 40 % were able to continue their business operations and, therefore, this type of workout had a significant influence on the recovery rate. Even though the study by Elsas and Krahnen (2002) can only be considered to lie partly in the field of recovery rate research, the authors show continuing operations after restructuring for 85 % of 128 German corporate customers.

[4] It should be noted that redemption in this sense is different to the 'statutory right of redemption' that is given by law in some US states. For the influence of the 'statutory right of redemption' on the recovery rate see, e.g., Qi and Yang (2009) and Park and Bang (2014).

[5] Amicable agreements (redemptions) are somewhat different to so-called 'cured' cases that have recently been discussed in the default prediction literature (e.g., Tong et al. 2012; Wolter and Rösch 2014). While 'cured' cases always show a 100 % recovery (see Wolter and Rösch 2014 for a brief introduction to 'cured' events), amicable agreements do not necessarily lead to a full recovery (the average recovery rate for amicable agreements in our sample is 82 %; see Table 1).

Table 1 Descriptive statistics on the borrower level

Variable	Description	Mean	Median	Min.	Max.	σ	N
Metric							
RR	Recovery rate	0.58	0.59	0	1	0.36	499
CR	Collateralization ratio	0.66	0.73	0	1	0.32	499
COL	Number of additional pieces of collateral not recognized on a value basis	1.68	1	0	14	2.18	499
$CV^{real\ estate}$	Collateral value of real estate (in thousand EUR)	160	79	0	2621	289	491
$CV^{deposit}$	Collateral value of deposits (in thousand EUR)	41	20	2	404	63	60
EAD	Exposure at default (in thousand EUR)	286	158	5	4110	413	499
WOP	Length of workout-process (in months)	31.5	30	4	69	12.2	499

Variable	Description	Number of realisations		$\varnothing$ RR		Wilcoxon rank sum test	
		0	1	0	1	Significance	(p-value)
Binary							
RED	= 1, if redemption was settled	304	195	0.42	0.82	***	0.0000
BT	= 1, if customer is commercial	368	131	0.58	0.57		0.6560
INS	= 1, if customer is bankrupt	412	87	0.61	0.40	***	0.0000

***, **, * indicates a 1, 5 and 10 % confidence level using the Wilcoxon rank sum test

of the debtor's lack of financial assets, the bank can only fall back to the third form of the workout-process, the liquidation of the collateral.[6]

In their analysis of 11,649 defaulted loans to retail customers and SMEs in Italy, Grippa et al. (2005) demonstrate that an amicable agreement between the bank and the borrower was achieved in 41 % of the cases. However, there is a specific characteristic of their study: the authors did not derive the percentage of redemptions from the individual debtors' data; the percentage was identified by the participating banks in a survey. To perform a multivariate analysis, the authors combined bank-specific data (here, the estimated fraction of redemptions) with other surveyed quantitative data (individual information regarding the individual loans, collateral, etc.). Thereby, they were able to demonstrate a significant, positive influence of the workout type, redemption, on the recovery rate. We again form the null to be rejected:

Hypothesis 3 (H3)   Amicable agreements with the debtors (redemptions) have no effect on the recovery rate.

3.2 Data and descriptive statistics

Our data set was provided by Bankaktiengesellschaft Hamm (BAG). BAG was founded to resolve the troubled Hammer Bank Spadaka eG, and therefore formed Germany's first bad bank. In the years following its foundation, BAG regularly took over sub-performing loans from troubled cooperative banks, because it had the necessary know-how. At the beginning of 2005, BAG started to offer its services— the acquisition and management of sub-performing loans—to the whole German banking sector. As is common in the context of the acquisition of sub-performing loans, BAG will take over the whole settlement and liquidation risk, and will cope with the risk via a sufficiently low purchase price.

The full sample contains information on 909 debtors that were transferred to BAG between December 2005 and August 2010. The following analysis of the recovery rate for debtors (borrower level) in Sect. 3 only includes the 499 debtors whose workout-processes had been completed.[7] The workout-process is considered to be completed when collateral has been processed entirely, or when an amicable agreement has been reached.

The analysis of individual collateral (collateral level) in Sect. 4 includes all 909 debtors and considers all collateral for which the workout-process is completed.

---

[6] When liquidating, the bank tries to sell the collateral that has been provided for as much as possible, and may initiate arbitration measures against the individual debtor (e.g., the attachment of his wages or salary).

[7] The authors are aware of the potential selection bias. Cases with long workout-processes are underrepresented in this sample since only cases where the workout-process is completed are considered. The average recovery rate could be overestimated, because defaults with a relatively long workout-process usually show lower recovery rates (see Gürtler and Hibbeln 2013). However, we find in our subsequent analysis no significant influence of the duration of the workout-process on the recovery rate. Moreover, the average duration of a workout-process in our sample is relatively long (31.5 months), with a maximum of 69 months. Consequently, we consider this potential selection bias to be negligible.

Fig. 1 Inclusion of cases in the borrower and collateral levels
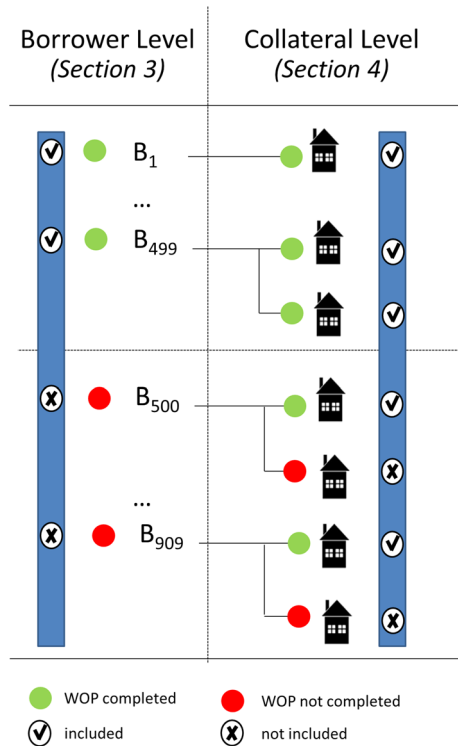
Whether a debtor or a particular item of collateral was included in our analysis at the borrower or the collateral level is illustrated in Fig. 1.

The data set also contains encoded information about the bank that originally granted the loan. As a result, we are able to conclude that the debtors were transferred to BAG by 123 different banks that are predominantly part of the cooperative banking sector.[8] For this reason, the data set to a large extent reflects the customer structure of cooperative banks, which is typically composed of retail and commercial (self-employed and SMEs) customers, whereas larger corporations (corporate customers) are usually provided with financial services in cooperation with or exclusively by the banks' central institutions (i.e., "Landesbanken" for savings banks and cooperative central banks for cooperative banks).

Since our data set is obtained from BAG—a bank that specializes in the acquisition of sub-performing loans—the question arises whether the data set is subject to a further sample selection bias in that only a fraction of sub-performing loans are transferred to BAG. It may be that banks do not offer unsecured or smaller sub-performing loans to BAG, and directly depreciate these loans instead. On the contrary, the opposite may happen and only unsecured loans are transferred to BAG.

---

[8] For the 123 banks, we are unaware of the exact outstanding loan volume (and the number of debtors) that were transferred to BAG, as our sample only includes those debtors whose workout-processes had already been entirely completed or at least had been completed for one collateral.

7 % of the loans in our sample are completely unsecured loans. This fraction seems very low, but since most of our loans are real estate loans, and such loans are typically collateralized, this fraction seems plausible. König and Beimborn (2008) conducted a survey among the 1020 largest German banks. The authors report an average SME loan size on the books of EUR 470,000. This number is comparable to the average EAD for SMEs of EUR 490,000 (median EUR 329,000) in our sample (see Table 1). Dr. Klein & Co. AG (2012) state that the average loan size for properties in Germany for retail customers is about EUR 160,000. This is below the average EAD for retail customers in our sample, which is EUR 213,000, but is above the median of EUR 130,000. Another possibility is that banks only transfer difficult loans that are hard to manage to BAG. This should lead to lower recovery rates compared to the German average. Hesse et al. (2012) report in their survey of cooperative banks and savings banks in Germany an average recovery rate for savings banks of 54 % and for cooperative banks of 52 %, both with a median of 55 %. These recovery rates are just slightly below the ones in our sample: we find an average of 58 % and a median of 59 %. Our data set is, therefore, in terms of EAD and recovery rate, comparable to the German average. Nevertheless, the selection biases mentioned above cannot be ruled out with absolute certainty. Yet the pieces of evidence mentioned indicate that any remaining bias should not be so severe that our analysis of the valuation of collateral is significantly affected.

Another concern is related to the relatively small sample size. But studies for the German market usually consist, because of limited data availability, of even fewer observations (Franks et al. 2004 have 276 observations, and Grunert 2005; Grunert and Volk 2008; Grunert and Weber 2009 all use the same data set consisting of 120 observations). One exception in terms of data points is the study by Gürtler and Hibbeln (2013), in which 69,985 relatively small sub-performing consumer loans are analyzed. But, since there is not yet a similar study that analyzes larger loans to individuals and SMEs for Germany, we consider our sample size of 499 (respectively, 909) debtors to be appropriate in comparison to other German studies. In addition, a problem for a majority of the existing recovery rate studies is that their data stems from just a single bank (e.g., Dermine and Neto de Carvalho 2006; Grunert and Weber 2009; Leow and Mues 2012; Zhang and Thomas 2012; Tong et al. 2012, 2013), which also raises the issue of the generality of the empirical results. Our data stem, in contrast, from 123 different banks, which makes it possible for us to rule out the peculiarities of a specific bank's loans.

For the analysis, we first determine the recovery rate (RR_j) for every debtor $j \in (1, 2, \ldots, 499)$. In our data set, the proceeds include the returns from the liquidation of any possible collateral as well as the payments made by the debtor. Costs include all external costs (e.g., costs for arbitration measures, the real estate agent's commission fees, etc.).[9] Determining the EAD appears to be critical, since the exact time of the default event as well as the criteria for the default are not known.[10]

---

[9] As for many other studies in this field (e.g., Grippa et al. 2005; Grunert and Weber 2009), we have no information regarding the internal costs.

[10] See Grunert and Volk (2008) regarding the influence of the definition of a default event when determining the recovery rate.

Therefore, we use the definition of a default event specified under the Basel Accord. Following Basel Committee on Banking Supervision (2006), a loan is considered to be in default if the bank is selling the loan obligation and thereby accepting an economic loss. Due to the reduction in the selling price, selling a receivable to BAG always satisfies this definition.[11] Hence, we take the debtor's existing liabilities at the time of transfer to BAG as the exposure at default (EAD). Using this definition, the difficulty of employing inconsistent definitions of a default event among the 123 included banks can be avoided.

For the EAD, we only have aggregated information; we are not, therefore, able to identify the types of loans (i.e., credit line, housing loan, etc.) that were originally granted. While other studies analyzing the retail and/or commercial customers segment determine the recovery rate for every loan individually,[12] we follow the approach used in studies on corporate customers (e.g., Grunert and Weber 2009) and look at the debtor as a unit by summing up all the loans and collateral for each customer. Focusing on the debtor as a unit instead of looking at each loan individually matches with the viewpoint of banking practice in Germany. First, the liens created under the general terms of business in Germany can be used for collateralizing all the loans of one debtor. Second, an amicable settlement covers the total commitment instead of individual loans and pieces of collateral.

A number of theoretical works (see e.g., Frye 2000; Pykhtin 2003; Altman et al. 2005; Bade et al. 2011; Rösch and Scheule 2014) suggest that default and recovery processes are negatively correlated. Empirical studies confirm this negative relationship between recovery rates and default probabilities (see e.g., Bade et al. 2011). One channel for this correlation may be the collateral value. In the case of an economic downturn, more borrowers default and the collateral value drops at the same time (Frye 2000). Heckman (1979) demonstrates that isolated models for two variables lead to biased parameter estimates if the two variables are correlated, and one variable can be observed only if the other variable exceeds a particular threshold. The higher the correlation between both variables the larger is the bias of the parameter estimates (Bade et al. 2011). As realized recovery rates can only be observed if a default event occurred, our study is also affected by this potential bias. A solution to this problem would be the application of an appropriate selection model (e.g., a Heckman selection model). Selection models correct the parameter estimates of a recovery rate regression for the probability of a default in the first place. However, since our data set solely contains finally defaulted borrowers, we do not see any way for us to apply a selection model here. Furthermore, the data set

---

[11] Furthermore, the origin of our data is the reason why we cannot observe potential recoveries of non-defaulted loans. If a loan is sold to BAG the originating bank is finally admitting that the borrower has defaulted. Following this reasoning, our data set solely contains finally defaulted customers, and no 'cured' cases. This sample selection issue is a problem for default prediction studies (see, for a brief discussion, Wolter and Rösch 2014), but the focus of our study lies on finally defaulted loans, their recovery rates, and the valuation of their collateral.

[12] This approach is especially reasonable when dealing with certain types of loans, such as credit card debt (e.g., Bellotti and Crook 2012), consumer credit (e.g., Gürtler and Hibbeln 2013) or vehicle financing (e.g., Appasamy et al. 2008). Particularly for automotive financing as asset financing, there is a close relationship between loan and collateral.

does not give any information about characteristics at the time when the loan was originally granted, such as the debtor's creditworthiness (i.e., credit rating), the interest rate, or the original loan amount. Therefore, all our following results have to be considered under the shortcoming of potentially biased parameter estimates.

Not only the EAD but also the proceeds and the expenses were only available in an aggregated form and without explicit information regarding their timing. However, Franks et al. (2004) and Schaaff (2009) demonstrate in their research only a marginal difference between the results if the present value is used and if not. In combination with the historically low level of interest rates during the time of observation, we consider the impact of not discounting to be negligible—partly because of the absence of discounting, it is, therefore, possible for the recovery rate to take on a value greater than 1. On the other hand, the recovery rate can take a negative value if the costs exceed the proceeds. In the same way as most empirical studies (e.g., Friedman and Sandow 2003; Grippa et al. 2005; Khieu et al. 2012; Loterman et al. 2012), we limit the recovery rate to the interval [0, 1]. This is done by setting all negative values to zero and all values above the unit interval to one. This data censoring influences the distribution of the recovery rate and can cause some problems. The issue of censored data was first mentioned by Tobin (1958) and is further discussed among others, by Greene (2012). Calabrese and Zenga (2010) address this issue in the context of recovery rates. In a robustness check, we excluded all cases that were censored (see Table 16 in the Appendix). Our results stay qualitatively the same. Therefore, we consider the potential problems caused by this censoring to be negligible.

Figure 2 shows, on the left-hand side (a), the frequency distributions of the unbounded recovery rate. There are peaks at 0 and 100 %; this bimodality can be found in almost all empirical studies (e.g., Asarnow and Edwards 1995; Hurt and Felsovalyi 1998; Araten et al. 2004; Grippa et al. 2005; Dermine and Neto de Carvalho 2006; Caselli et al. 2008). Some recoveries are below 0 % if the costs exceed the proceeds, and some are above 100 % if the net proceeds exceed the EAD. This finding is common if workout recovery rates are calculated. Araten et al. (2004), for example, report recovery rates between −10 and 173 %. The maximum observed recovery rate in our sample is 191 % and the minimum is −34 %. In total,
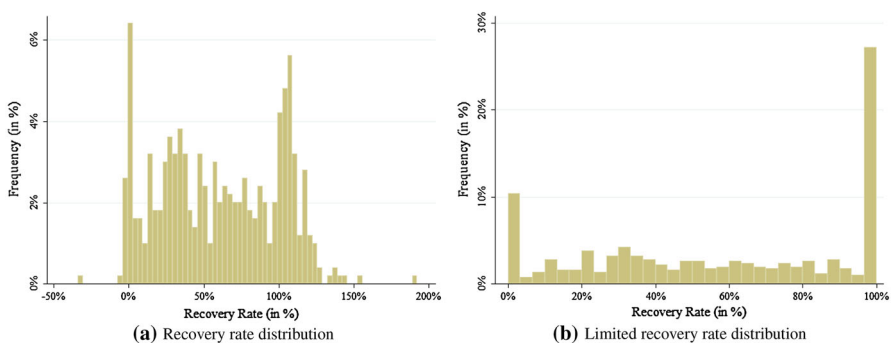

(a) Recovery rate distribution   **(b)** Limited recovery rate distribution

Fig. 2 Histograms of realized recovery rates

Table 2 Default years and recovery rates

Default year	Full sample		Retail customers			Commercial customers		
	N	$\varnothing$ RR	N	Fraction (%)	$\varnothing$ RR	N	Fraction (%)	$\varnothing$ RR
2005	143	0.71	102	71.3	0.72	41	28.7	0.67
2006	35	0.45	16	45.7	0.40	19	54.3	0.48
2007	244	0.52	192	78.7	0.51	52	21.3	0.53
2008	64	0.59	46	71.9	0.60	18	28.1	0.55
2009	12	0.57	11	91.7	0.58	1	8.3	0.46
2010	1	0.85	1	100.0	0.85	0	0.0	–

15 cases are below 0 % and 120 are above 100 %. On the right-hand side (b) of Fig. 2, the distribution of the censored recovery rate limited to the unit interval is displayed. The bimodality mentioned before is now even more pronounced. Table 1 contains the descriptive statistics of the variables on the borrower level and shows an average recovery rate of 58 % (median of 59 %).

The 499 customers were, as mentioned earlier, transferred to BAG between December 2005 and August 2010. Table 2 clearly shows that the number of transfers per year, as well as the corresponding recovery rates, differ over the observation time. Most customers were transferred in 2005 (143) and in 2007 (244). The average recovery rate per year was the highest in 2010 at 85 % (for just a single observation) and the lowest was in 2006 at 45 %. Table 2 also presents the distribution of the transferred cases per customer type.

As to collateral, BAG only considers certain collateral on a value basis (in EUR); this includes money in accounts (e.g., building loan contracts, savings deposits, endowment life insurances, etc.) and collateral in real estate. The collateral value of money in accounts ($CV^{deposit}$) equals the corresponding balance or the surrender value at the time of the transfer.

The value of collateral in real estate ($CV^{real\ estate}$) results from a sequential calculation. Possible existing preloads (PL) are subtracted from the most recent valuation of the property (MV).[13] Moreover, the collateral value is limited to the sum of the property's recorded land charge (LC).[14] The collateral value of real estate i is, therefore, given by

$$CV_i^{real\ estate} = MAX(MIN(MV_i - PL_i; LC_i); 0) = MIN(MAX(MV_i - PL_i; 0); LC_i).$$
(2)

---

[13] The most recent valuation is the value resulting from the most recent appraisal report at the time of the transfer and is materially equivalent to the fair value used in international accounting. For the different types of appraisal reports, see Sect. 4.2.

[14] The German Civil Code (BGB) defines two types of collateral in real estate. The 'mortgage' (Hypothek) is used for accessory charges on property and the 'land charge' (Grundschuld) is used for non-accessory charges on property. In contrast to the 'mortgage', a 'land charge' continues to exist even if the debt is extinguished. 'Land charges' are common in German banking practice and 'mortgages' are only rarely used. See European Mortgage Federation (2007) for further information on the 'mortgage' and the 'land charge' in Germany.

Other studies do not provide many details on how they determine the collateral value, but it can be assumed that the $CV^{real\ estate}$ determined by the bank is applied. We, on the other hand, explicitly combine information from different sources for each collateral individually to determine the $CV^{real\ estate}$ using formula (2). The total collateral value for every debtor j is computed by summing up m collateral values in real estate and n collateral values of money in accounts.

Depending on the quality and quantity of the available data, the collateral can be displayed in various ways in an analysis. Dermine and Neto de Carvalho (2006) incorporate the collateral using binary variables for every type of collateral (collateral in real estate, guarantees, etc.). In contrast, more recent studies (Caselli et al. 2008; Grunert and Weber 2009; Qi and Yang 2009; Leow and Mues 2012; Park and Bang 2014) take collateral into account using a collateralization ratio or its inverse, the (current) loan-to-value ratio, which both represent the relation between the bank's valuation of the collateral and the EAD. We follow this approach and employ the collateralization ratio (CR) by dividing the total collateral value by the EAD:

$$CR_j = \frac{\sum_{i=1}^{m_j} CV_{i,j}^{real\ estate} + \sum_{i=1}^{n_j} CV_{i,j}^{deposit}}{EAD_j}. \tag{3}$$

A collateralization ratio above 1 reflects over-collateralization. We limit the collateralization ratio—in line with the banking practice—to a maximum of 1.[15]

On average, the collateralization ratio is 66 %, with only a small percentage (approx. 6–7 %) of debtors not providing any collateral at all (CR = 0; see Fig. 3). Our data set thus shows a significantly lower percentage of unsecured liabilities

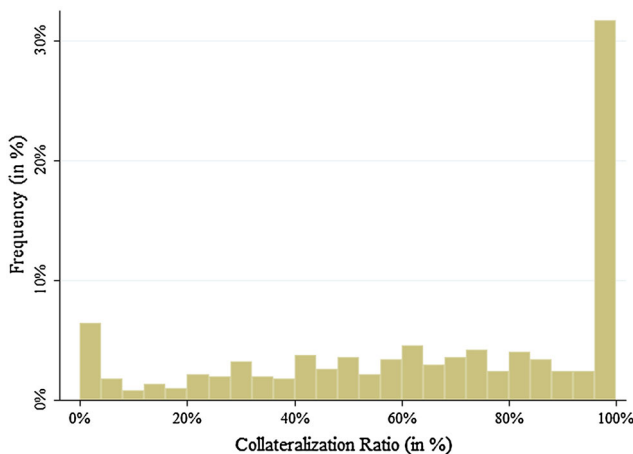
Fig. 3 Frequency distribution of collateralization ratio

---

[15] This limitation is also used by Grunert and Weber (2009). The issue discussed earlier caused by censoring of the distribution is also valid for the collateralization ratio.

compared to other studies.[16] This could be caused, as discussed earlier, by the origin of the data set. It is possible that unsecured liabilities are directly depreciated by the originating bank and rarely offered for purchase to BAG.

The average collateral values—EUR 160,000 for collateral in the form of real estate and EUR 41,000 for collateral in money in accounts—lead to the conclusion that the collateralization ratio is mainly determined by the collateral in real estate. The collateralization in our data set is comparable to that in the study of Qi and Yang (2009), who even show a collateralization ratio of 100 %. The 240,000 mortgages of US retail customers included in their study may explain the higher collateralization ratio. Interestingly, there is a larger difference in the collateralization rate (100 vs. 66 %) than in the realized recovery ratio (70 vs. 58 %).

To assess whether additional collateral not recognized on a value basis (in EUR) has an effect on the recovery rate (H2), we determine the variable COL as the number of additional pieces of collateral (e.g., individual assignments of claims, global or overall assignments, assignments of loan receivables, guarantees or transfers of ownership of vehicles, machinery and warehouses) that are not recognized on a value basis (in EUR). On average, 1.8 additional pieces of collateral (median: 1) are pledged.

Since there is no restructuring for the debtors in our data set, we only differentiate between an amicable agreement and a liquidation to assess Hypothesis H3. The binary variable redemption (RED) takes a value of 1 if an amicable agreement between BAG and the debtor was reached. The portion of redemptions in our data set amounts to 40 % (195 out of 499 cases). This proportion is consistent with the survey results of Hesse et al. (2012) for cooperative banks and savings banks in Germany regarding problem loans. According to their 171 responses, the average proportion of redemptions equals 34 %. Grippa et al. (2005) report a fraction of 41 % of such private agreements between debtors and banks for an Italian sample and find for these cases a considerably higher recovery rate of 68 % compared to their overall average of 37 %.

To check for scale efficiencies via the credit volume, we utilize the logarithm of the EAD (EAD^{LN}) which is EUR 286,000 in our sample. Additionally, we consider the duration of the workout-process in months (WOP). BAG's purchase of a certain liability marks the starting point, and the completion of the workout-process marks the end of the workout-process for a debtor. On average the workout-process lasts 31.5 months, which is considerably lower than the average duration of 54 months reported by Calabrese and Zenga (2010) for Italy. Khieu et al. (2012), on the other hand, state an average duration of the workout time in their US sample of just 14 months. As to the debtor, we differentiate between 368 retail and 131 commercial (self-employed and SMEs) customers. The binary variable borrower type (BT) takes a value of 1 if the debtor is a commercial customer. The variable INS shows the status of bankruptcy and takes a value of 1 if the debtor is declared bankrupt before or during the workout-process, which is the case for 87 observations. The annual

---

[16] 36 % of the 374 loans to defaulting small- and medium-sized companies in Portugal analyzed by Dermine and Neto de Carvalho (2006) were unsecured. Nevertheless, a recovery rate of 71 % was achieved. However, there is no information regarding the collateralization ratio. A high portion of unsecured securities also becomes evident in a study by Jiménez and Saurina (2002) on the Spanish market.

Table 3 Comparison of retail and commercial customers

Borrower type	Quantity	∅ CR	∅ COL	Fraction of RED (%)	∅ EAD (in thousand EUR)	∅ WOP	∅ RR
Retail customer	368	0.70	1.20	39	213	30.72	0.58
Commercial customer	131	0.58	3.00	38	490	33.60	0.57

growth rate of the real gross domestic product (GDP) and the unemployment rate (UER) at the time of default (i.e., transfer to BAG) take the role of macroeconomic control variables.[17]

To analyze the extent to which there are differences between retail and commercial customers, important variables are presented by borrower type in Table 3. It becomes obvious that loans to commercial customers exhibit, on average, a lower collateralization ratio (58 vs. 70 %). But commercial customers pledge, on average, in contrast to the results for the collateralization ratio, a higher number of additional pieces of collateral (3.0 vs. 1.2). One reason may be that commercial customers are usually in possession of more specific assets that can be used as collateral (e.g., special purpose machines). By looking at the customer structure, the average EAD (EUR 213,000 for retail customers and EUR 490,000 for commercial customers) demonstrates again that our study is quite different from studies of classical corporate customers, e.g., of Grunert and Weber (2009) who found a mean value of EUR 5.6 million.

3.3 Bivariate analysis

In the same way as in Grunert and Weber (2009), the Bravais and Pearson correlation coefficient is used for the bivariate analysis of the metric variables (see Table 12 in the Appendix). There is a strong significant positive correlation of 0.7 between the collateralization ratio and the recovery rate (Hypothesis H1). Thereby, Hypothesis H1 of no association can be rejected in the bivariate context. The correlation coefficient for additional collateral (COL) equals 0.03, which leads to the conclusion that Hypothesis H2 cannot be rejected. Furthermore, there are no other remarkable correlations between the control variables.

The binary analysis of the influence of redemption may lead to the conclusion that Hypothesis H3 can be rejected. The average recovery rate of redemptions (0.82) is considerably higher than the rate of liquidation (0.42; see Table 1). Interestingly, we find a significant positive correlation (0.31) between redemption and the collateralization ratio (see Table 12 in the Appendix). In the following multivariate analysis, we are going to assess whether redemption can influence the recovery rate beyond the collateralization ratio.

---

[17] According to Basel Committee on Banking Supervision (2006), banks have to consider economic upswings when estimating the recovery rate. These requirements are satisfied, following Basel Committee on Banking Supervision (2005), using the variation of both the GDP and the unemployment rate.

Table 4 Fractional Logit regression with the recovery rate as dependent variable

Variable	Model (1) Recovery rate std. coefficient (standard error)	Model (2a) Recovery rate std. coefficient (standard error)	Model (2b) Recovery rate std. coefficient (standard error)
H1			
CR	0.6291*** (0.212)	0.6040*** (0.297)	0.7150*** (0.428)
H2			
COL	0.0368 (0.036)	0.0427 (0.045)	0.0523 (0.045)
H3			
RED	0.3813*** (0.134)	0.4286*** (0.147)	0.2982*** (0.268)
Control variables			
EAD^{LN}	0.0277 (0.070)	0.1002** (0.080)	−0.1494** (0.108)
WOP	0.0062 (0.005)	−0.0292 (0.007)	0.0130 (0.011)
BT	0.0691** (0.134)		
INS	−0.0689** (0.161)	−0.0757* (0.202)	−0.0019 (0.198)
GDP	−0.0359 (0.038)	−0.0386 (0.362)	0.0289 (0.061)
UER	0.0828 (0.073)	0.0624 (0.073)	0.1495* (0.106)
Constant	−4.0712*** (1.007)	−5.3583*** (1.224)	−1.0187 (1.485)
Observations	499	368	131
AIC	0.7994	0.7989	0.8838
BIC	−2893.48	−2013.36	−563.08
Mc Fadden R^2	0.3304	0.3411	0.3348
R^2	0.6324	0.6350	0.6925
Adj. R^2	0.6256	0.6268	0.6724
Wald test (p-value)	467.95 (0.000)	390.85 (0.000)	273.84 (0.000)

***, **, * indicates a 1, 5 and 10 % confidence level. The robust standard errors that are clustered for the originating bank are given in brackets

3.4 Multivariate analysis

In this section, we analyze the influence of the nine previously introduced variables on the recovery rate of debtor j. To do so, we use a Fractional Logit Model, similar to the one used in the studies by Grippa et al. (2005), Dermine and Neto de Carvalho (2006), Bastos (2010) and Khieu et al. (2012), among others:[18]

$$\mathrm{RR}_{j,t} = \beta_0 + \beta_1 \cdot \mathrm{CR}_j + \beta_2 \cdot \mathrm{COL}_j + \beta_3 \cdot \mathrm{RED}_j + \beta_4 \cdot \mathrm{EAD}_j^{LN}$$
$$+ \beta_5 \cdot \mathrm{WOP}_j + \beta_6 \cdot \mathrm{BT}_j + \beta_7 \cdot \mathrm{INS}_j + \beta_8 \cdot \mathrm{GDP}_t + \beta_9 \cdot \mathrm{UER}_t + \epsilon_j. \quad (4)$$

Using the Fractional Logit Model with robust standard errors that are clustered for the originating bank leads to the estimation results listed in Table 4.

---

[18] The Fractional Logit Model is considered to be a Generalized Linear Model and is specifically designed for a dependent variable in the interval [0;1] when there is a positive probability that the values 0 and 1 can be realized.

Model (1) shows the estimation results including the borrower type using the binary variable BT. We have already illustrated (see Table 3) that there is a big difference between the two types of customers, for example, with respect to the collateralization ratio. Therefore, we split the sample into the two customer types [Model (2a) = retail customers and Model (2b) = commercial customers]. This sample split allows for an individual analysis of the independent variables. By doing so, we are able to demonstrate differences in the influence of the other independent variables.

Regardless of the customer type it can be shown that the collateralization ratio has a strong positive influence on the recovery rate ($p < 0.01$). Therefore, Hypothesis 1 (H1) can be rejected in the multivariate context. The outstanding importance of the collateralization ratio is not surprising since it is mainly determined by collateral in the form of real estate, to which a high fundamental value has been attributed. Our result is in line with the empirical literature on recovery rates (e.g., Grunert and Weber 2009; Qi and Yang 2009; Bastos 2010; Gürtler and Hibbeln 2013). The straightforward explanation for this finding is that either the direct realization of the collateral results in direct proceeds or, in the case of an amicable agreement, the bank would only reach this agreement if the payment offered equals at least the sum of the expected discounted (net) returns from the collateral and the personal arbitration measures against the debtor.

As to Hypothesis H2, there is no significant influence in the multivariate analysis of additional collateral (COL) on the recovery rate—neither for retail nor for commercial customers, i.e., this hypothesis can not be rejected. This finding seems surprising at first glance. A possible explanation for this finding is that additional pieces of collateral are not recognized on a value basis (in EUR). These additional pieces of collateral are usually very specific assets that are difficult to sell (e.g., special purpose machines) and, therefore, often did not yield any cash inflows. Their specificity is why they were not or even could not be valued in the first place.

Different from H2, we have been able to reject Hypothesis H3. There is a significantly positive coefficient for redemption, which shows a relation with the recovery rate beyond the collateralization ratio. To identify the differences with respect to the different workout-processes in more detail, we form the variable DIF_j for each debtor, which is the difference between the realized recovery rate (RR_j) and the collateralization ratio (CR_j). If DIF_j is positive, a premium can be realized. Figure 4 shows the distribution of DIF depending on the different workout-processes.

The distribution of DIF is reminiscent of the normal distribution; however, there are too many observations with a negative difference (p-value 0.000 in the Shapiro–Wilk test). Splitting the sample into the different workout-processes reveals the main differences. The collateralization ratio is realized (DIF = 0) for redemption in the majority of the cases, which is intuitive considering that redemption is only optional for BAG. From a relative viewpoint, there are significantly more cases with a premium since the related curve lies above the curve for the entire population for values of the x-coordinate greater than 0 %. Reciprocally, there are, from a relative viewpoint, more cases with a discount for liquidation. Altogether, our illustration regarding the premium clarifies the positive influence of the redemption workout-process.
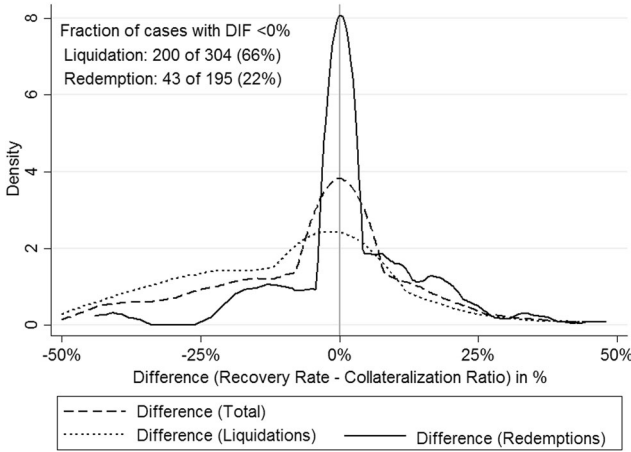
Fig. 4 Distribution of difference (DIF) dependent on the workout type
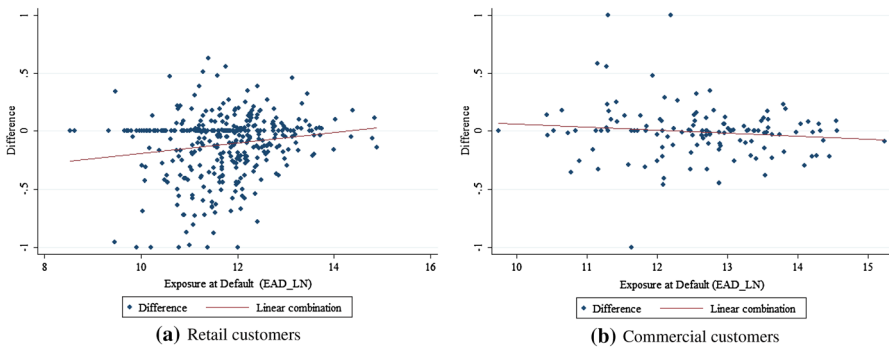
Fig. 5 Scatter plots of difference (DIF) dependent on the exposure at default (EAD^{LN})

Our results indicate an ambiguous influence of the outstanding loan amount (EAD^{LN}). There is a positive, significant coefficient for retail customers, while the coefficient is negative for commercial customers. Banks may intensify their effort in the workout-process if the EAD is higher and, as a consequence, if higher losses are possible. This could lead to a higher recovery rate for larger EAD. On the other hand, it could be that higher outstanding amounts are more difficult for the borrower to pay back, which could lead to a lower recovery rate. The empirical evidence for the influence of the EAD on the recovery rate is mixed. Bellotti and Crook (2012) find a significant negative influence of a higher EAD on the recovery rate, while Tong et al. (2013) find the opposite effect. Another possible explanation for our finding is offered by the premium (DIF$_j$ is positive). For retail customers, there is a positive (but weak) correlation between the outstanding credit volume and the premium (see Fig. 5a), expressed in a decreasing liquidation risk with increasing volume. The inverse relation exists for commercial customers (see Fig. 5b).

Table 5 Collateralization ratio, redemptions and differences dependent on the status of bankruptcy as well as the borrower type

Borrower type	Status of bankruptcy	$\varnothing$ CR	Fraction of RED (%)	DIF-points (%)
Commercial customer	0	0.58	42.72	0.0033
	1	0.56	21.43	−0.0468
Retail customer	0	0.72	45.95	−0.0895
	1	0.59	5.08	−0.2488

The status of bankruptcy (INS) is accompanied by a significantly lower recovery rate for retail customers ($p < 0.1$). This finding is in line with the results of Han and Jang (2013). Since the proceeds from bankrupt debtors only come from the collateral provided, with no further payments to be expected, the relation appears coherent. Moreover, an amicable agreement is less likely, because there are no other assets or sources of financing available for redemption due to the debtor's bankruptcy. Therefore, the fraction of redemptions for bankrupt retail customers is significantly lower (5 %) than that for non-bankrupt debtors (46 %; see Table 5).

Calabrese and Zenga (2010) and Gürtler and Hibbeln (2013) argue that defaulted borrowers with a long workout-process yield lower recovery rates. This could be driven by the additional costs that arise over time. Caselli et al. (2008), Qi and Yang (2009) and Khieu et al. (2012) find, in line with this argument, a negative influence of the length of the workout-process on the recovery rate. We cannot confirm their findings, since we find no significant coefficient of the duration of the workout-process on the recovery rate.

Some empirical studies find a significant positive relation between real GDP growth and the recovery rate (e.g., Khieu et al. 2012; Han and Jang 2013; Park and Bang 2014). An explanation for this finding could be that when the economy is in a poor state defaulters are less able to repay their debt, which negatively influences the recovery rate, and vice versa. But a few studies, on the other hand, find no significant relation (see e.g., Dermine and Neto de Carvalho 2006) or, especially for the case of Germany, the opposite effect (see Franks et al. 2004 as well as Grunert and Weber 2009 who show this result for the regional GDP growth). However, our results show no significant relation between real GDP growth and the recovery rate.

We find a weak significant positive relation between the unemployment rate and the recovery rate for commercial customers. This finding is in contrast to the findings of Bellotti and Crook (2012) and Park and Bang (2014). Bellotti and Crook (2012) argue that an increase in the level of unemployment means that most people find themselves in circumstances where they cannot repay their loans, which lowers the average recovery rate. Grunert and Weber (2009) find no significant influence on the recovery rate for their German sample. Our counterintuitive finding of a positive relationship between the level of unemployment and the recovery rate could be driven by our sample period. The unemployment rate is quite high in 2005 and 2006, at 11.7 and 10.8 %, and drops to about 8 % between 2008 and 2010. The high unemployment rate in 2005, which was associated with an above-average recovery rate of 71 % for this year, is probably a particular driver for our finding. We

excluded all cases that were transferred to BAG in 2005 in a robustness test. All coefficients for real GDP growth turn positive, and a positive significant influence ($p < 0.01$) for commercial customers [Model (2a)] can be seen. Moreover, two of the three coefficients for the unemployment rate turn negative, and the earlier positive impact for commercial customers loses its significance. All other results stay qualitatively the same.

As to the macroeconomic factors, the variables recommended by the Basel Committee—gross domestic product (GDP) and unemployment rate (UER)—can only partly be confirmed to have an influence on the recovery rate for commercial customers. To overcome the concern that our macroeconomic variables do not pick up all the macroeconomic effects, we use year dummies for the respective default year in a robustness test (see Table 14 in the Appendix). All our results stay qualitatively the same. Therefore, we consider our macroeconomic variable setup to be suitable.

We split our sample in a further robustness check among the two different workout types liquidation and redemption. Table 17 in the Appendix shows the results of the regressions [Models (13a) and (13b) for the liquidations, again divided into the two customer types, and Models (14a) and (14b) for the redemptions]. Sign and statistical significance of the collateralization ratio are confirmed in all models. It is noticeable that the status of bankruptcy (INS) has a significant negative influence on the recovery rate for both customer types only in the case of an amicable agreement. An insolvency in conjunction with an amicable agreement is only observed for nine customers (three retail and six commercial customers) in our sample. The reliability of this particular result is, therefore, debatable. Nevertheless, all other results are qualitatively unchanged. We, therefore, consider the joint analyses of the two workout types to be appropriate.

In conclusion, the collateralization ratio (H1) and an amicable agreement (redemption) (H3) have a positive influence on the recovery rate. Our model is able to explain 63 and 67 %, respectively, of the variance using the adjusted R^2 as a measure. This high explanatory power for recovery rates of loans that are predominantly collateralized by real estate is typical of empirical studies. Qi and Yang (2009), for example, report an adjusted R^2 of 61 % and Park and Bang (2014) of as much as 74 %. The high explanatory power is mainly driven by the high fundamental value of collateral in real estate. If the collateralization ratio is excluded from the analysis, the adjusted R^2 drops to 20 % for commercial customers and to 39 % for retail customers. Qi and Yang (2009) achieve a similar result with a reduction of the adjusted R^2 from 61 to 15 % if the collateralization by real estate is not included. Because of the major influence of the collateralization ratio on the recovery rate, we are going to analyze the essential driver—the valuation of the real estate—in the following section in more detail.

4 Empirical analysis of the influence of the proceeds ratio on the collateral level

4.1 Derivation of our research hypothesis

Since empirical studies have identified a significant positive influence of collateral on the recovery rate, there has been a second strand of literature, in which the collateral provided, instead of the individual debtors or loans, is studied. Studies in this field analyze the fundamental value of different forms of collateral, such as collateral in real estate, guarantees, etc. The fundamental value is measured by the proceeds ratio, which relates the realized proceeds to the original fundamental value assumed by the bank. Grunert (2010) analyzes potential factors influencing the proceeds ratio, based on a German data set consisting of 104 pieces of collateral (of which 62 are collateral in real estate). The average proceeds ratio is 61 %, and the author concludes that the proceeds ratio fluctuates depending on the type of collateral. Franks et al. (2004) analyze the fundamental value of collateral for 790 pieces of collateral in Germany, France, and Great Britain. The German subsample consists of 120 pieces of collateral with an average proceeds ratio of 73 % (median 77 %). The authors likewise conclude that the average proceeds ratio fluctuates depending on the type of collateral. Nevertheless, an above-average fundamental value of collateral in real estate is pointed out. Their result coincides with practical experience: according to a survey carried out by Grunert (2009), banking professionals have identified collateral in real estate to have the highest fundamental value.

Following this result, Schaaff (2009) focuses on an isolated analysis of different factors influencing the proceeds ratio of collateral in real estate. She analyzes 1120 pieces of collateral in real estate that were transferred to BAG between 1999 and 2005. The resulting average proceeds ratio amounts to 61 %, with a median of only 28 %. The strong discrepancy can be traced back to the fact that for 41 % of the assets there was no cash inflow. This specific feature could be caused by the data set itself. Even the author herself notes that the representativeness of the data set has to be reviewed critically. The collateral used, as well as its most recent valuations, originate from banks undergoing restructuring that have sought help from the central organization of the cooperative banking group in Germany (Bundesverband der Deutschen Volksbanken und Raiffeisenbanken) because of their economic situation. The author identifies the macroeconomic environment, the type of workout-process for the property, the size, and the location as some important factors influencing the proceeds ratio.

Previous studies have shown average liquidation proceeds below the value originally set by the bank, i.e., a proceeds ratio below 1. Therefore, the collateralization ratio used to analyze the recovery rate is too high in most empirical studies (exceptions are e.g., Jokivuolle and Peura 2003; Somers and Whittaker 2007; Leow and Mues 2012) and this influences the results of the recovery rate analysis. As previously mentioned, the collateralization ratio in our sample is dominated by collateral in real estate. The collateral value of real estate ($CV^{real\ estate}$) stems from a sequential calculation, as shown in formula (2), with the

property's latest valuation in the most recent appraisal report as the starting point. According to §194 of the German Federal Building Code (BauGB), the valuation by the appraiser (determination of the (expected) market value) should reflect the price that can be achieved in an arm's length transaction on an active market, considering the specific qualities of the property such as its condition and location. The realization of the (expected) market value is the price actually achieved on the market. Following Kleiber and Simon (2007), this market price reflects the result of negotiations between the seller and the buyer, who each have different opinions regarding the object's value, depending on their subjective and personal beliefs.

If it is assumed that the appraisal reports on the (expected) market value of real estate consider all factors that can possibly influence its value, according to §194 of the German Federal Building Code (BauGB), there should be no systematic deviation from the realized prices.

Hypothesis 4 (H4)   The proceeds ratio of real estate is not systematically biased.

4.2 Data, descriptive statistics and bivariate analysis

Our data set includes 1236 properties, relating to 909 debtors, that were sold between May 2006 and January 2011. 470 of these properties were sold through foreclosure auctions, and 248 by over-the-counter trades. Redemption took place in 376 of the cases.[19] The remaining 142 objects are classified as 'miscellaneous'. This 'miscellaneous' category comprises properties for which the realized market price could not be found because, for example, the intention of liquidating was not pursued any further due to extensive rights of third parties. In cases of redemption, there is no objective market price since there is no market-oriented realization. If the debtor's EAD is smaller than the property's market value, and if the liability is completely repaid with the redemption payment, this yields a proceeds ratio of less than 1. However, this cannot necessarily be traced back to the property alone. If a debtor is in possession of more than one property, a package price is paid. Splitting the price would always be subject to a criticism that an arbitrary figure had been chosen, as the debtor's willingness to pay for the individual properties is unknown. Therefore, the property in a redemption case will not be considered in the following analysis. Hence, in these cases as well as for the 142 'miscellaneous' cases, the realized market price is missing. For the explanatory model, we therefore only consider 718 objects for which a realized market price is known, here from foreclosure auctions and over-the-counter trades.

To measure the proceeds ratio (PR) for property i, we divide the realized market price of the property (MP) by the expected market value of the property (MV):

$$PR_i = \frac{MP_i}{MV_i}. \tag{5}$$

Table 6 shows descriptive statistics for the variables on the collateral level.[20]

---

[19] This matches approximately 30 % of the real estate. The previously mentioned portion of 40 % is measured at the borrower level (195 out of 499 debtors).

[20] The correlation between the variables is shown in Table 13 in the Appendix.

Table 6 Descriptive statistics on the collateral level

Variable	Description	Mean	Median	Min.	Max.	σ	N
PR	Proceeds ratio	0.78	0.74	0	3.13	0.37	718
MV	Market value of real estate (in thousand EUR)	200	122	3	4441	326	718
Location	Index value for a property's location	0.94	0.86	0.40	2.82	0.35	718

Variable	Description	Number of realizations		$\varnothing$ PR of realizations		Wilcoxon rank sum test	
		0	1	0	1	Significance	(p-value)
Report							
BANK	= 1, if report from bank	518	200	0.78	0.77		0.3800
COURT	= 1, if report from court	580	138	0.83	0.57	***	0.0000
BAG	= 1, if report from BAG	338	380	0.69	0.86	***	0.0000
Condition							
CONDITION$^{\text{WITHOUT}}$	= 1, if condition cannot be ascertained	438	280	0.77	0.79		0.5226
CONDITION$^{\text{NORMAL}}$	= 1, if backlog of repair ratio at most 25 %	450	268	0.74	0.84	***	0.0000
CONDITION$^{\text{BAD}}$	= 1, if backlog of repair ratio more than 25 %	548	170	0.81	0.67	***	0.0000
Type of property							
APH	= 1, if apartment house	618	100	0.78	0.79		0.8410
COMPROP	= 1, if commercial property	442	276	0.77	0.80		0.5763
SFH	= 1, if single-family home	519	199	0.78	0.78		0.8558
CONDO	= 1, if condominium	575	143	0.79	0.74		0.2900
FO	= 1, if sold through foreclosure auction	248	470	0.92	0.71	***	0.0000

***, **, * indicates a 1 , 5 and 10 % confidence level using the Wilcoxon rank sum test

The proceeds ratio is, on average, 78 %, with a median of 74 %. Because of different definitions of the proceeds ratio and differences in the groups of the originating banks, the proceeds ratio is only comparable to a limited extent with the results of other studies. Schaaff (2009) reports an average proceeds ratio of 62 % (median of 28 %) and Grunert (2010) of 61 % (median of 53 %).[21] The average market value of the properties is EUR 200,000 with a median of EUR 122,000. This value is higher than in the study by Schaaff (2009), whose figures show EUR 108,000 on average and a median of EUR 40,000.

The specific characteristics of our data set mean that we can distinguish between three types of experts providing appraisal reports on the market value of a property. If there is a report prepared by the originating bank (BANK) or a judicial report (COURT) available at the time of transfer of the loan, the most recent of these is considered to be the relevant report. If there is no such appraisal, or if the appraisal is outdated, a local expert is assigned by BAG to prepare the appraisal report (BAG). In conclusion, there is only one relevant report on the market value. Its source is shown using three binary variables (BANK, COURT and BAG) that take on the value of 1 if the report stems from the respective source. As to the authors of the appraisal reports, it becomes obvious that the reports prepared by BAG show a significantly higher proceeds ratio (86 %) than the reports for the comparable groups. While there is no difference for the proceeds ratio derived from appraisal reports prepared by the originating banks (77 %), the average proceeds ratio for sales on the basis of judicial reports is only 57 %, which is significantly lower. One possible explanation for this might be that, in comparison to judicial reports, BAG and originating banks have to bear the consequences themselves if their estimate is wrong, and therefore make a more conservative estimate.

Looking back at the legal definition for the determination of the (expected) market value (§194, BauGB), it becomes obvious that a property's condition as well as its location have to be considered. To include the property's condition, we analyzed, based on the appraisal reports, the backlog of repairs and put this in relation to the market value of a property in perfect condition (backlog of repair ratio). If the ratio lies between 0 and 25 %, we assume a normal condition. The binary variable Condition^{NORMAL} then takes a value of 1. A ratio above 25 % is displayed with the binary variable Condition^{BAD}. In 40 % of the cases, we were not able to identify the condition explicitly from the information provided by the data set. Not including these cases would lead to the loss of a large proportion of the observations. For that reason, the binary variable Condition^{WITHOUT} represents properties with unknown condition.[22] As to Condition, it can be demonstrated that real estate in worse condition has a significantly lower proceeds ratio (67 %) than the comparable group (81 %), although the condition as well as the location should have already been considered by the experts. The fact that property in a bad condition only attracts a very limited range of prospective buyers could offer one

---

[21] For example, at the time of the study by Schaaff (2009), BAG mostly took over loans from banks that were subject to restructuring.

[22] We excluded in a robustness test (not reported) all properties with unknown condition from our analyses. All results stay qualitatively the same.

explanation. The very limited range of potential buyers, as well as the demand of the buyers for a risk discount if they fear additional, as yet undiscovered, defects, could explain the negative influence of the variable Condition. There is no significant difference in the proceeds ratio for property in unknown condition (79 %) and a comparable group consisting of properties in good and bad condition (77 %), which may indicate that the group with unknown condition is in fact a mixture of properties in good and bad condition and does not yield a selection bias.

The location is the second factor that is explicitly mentioned for determining a property's market value. The attractiveness of a location has to be measured on a very fine regional scale. In cities with a high population density, the attractiveness of the location varies greatly between the various neighborhoods (e.g., for the wealthy area of Duesseldorf). When measuring particular neighborhoods, the attractiveness of a location should be expressed by an aggregate ratio that expresses as many advantages and disadvantages (e.g., transport links, nuisance, etc.) as possible at the same time.

To fulfill these requirements, we make use of data provided by F+B GmbH. F+B GmbH identifies an objective average selling price per square meter for different types of property (single-family home, condominium, apartment house, commercial property) on the postal code level for Germany.[23] Since we cannot compare the absolute average purchase prices of the different types of property, we construct an index for each type of property. To do so, we first form the average purchase prices for all postal codes on an annual basis. After this, we reference the given value for the corresponding postal code to the previously determined average. Thereby, we provide an index value for the relative attractiveness of a location. An index value above 1 indicates prices above the German average for a certain property type in a specific postal code area. We capture the index value for every property i with the metric variable LOCATION. The year in which the workout-process started is considered to be the relevant year, to display the expert's knowledge optimally.

The index value for the variable location is on average 0.94, and indicates that the attractiveness of the location of the properties included in the data is slightly below the national German average. At the same time the maximum value (2.82) and the minimum value (0.40) indicate strong differences in terms of attractiveness. Moreover, a statistically significant, positive relation between a property's location and the proceeds ratio can be demonstrated (see Table 13 in the Appendix). Although the location should have been considered in the experts' appraisal reports, there still seems to be a significant influence on the proceeds ratio at least in the bivariate context. This contradicts Hypothesis H4.

The four previously mentioned types of property are displayed with the binary variables SFH (single-family house), CONDO (condominium), APH (apartment) and COMPROP (commercial property). The results show no significant difference in the proceeds ratios for the different types of property.

---

[23] F+B GmbH quarterly evaluates all purchase offers in Germany for single-family houses, apartments, condominiums, and commercial properties based on the purchase offers on 100 websites on real estate. These offers are then validated by the realized purchase prices. The final data include the average realized market price on a postal code level for the different types of property (F+B GmbH 2013).

Furthermore, we control for the type of workout-process. If there is an amicable agreement between BAG and the debtor, all collateral from the debtor is released. If there is no amicable agreement, the collateral is liquidated. First, collateral in real estate is offered by a regional real estate agent mandated by BAG. If a buyer can be found, the property is sold in an over-the-counter trade OTC. If the debtor is not willing to cooperate or another creditor denies his consent and an over-the-counter trade is not possible, BAG initiates a foreclosure auction. The disposal of the property in a foreclosure auction is displayed with the binary variable FO. At 92 %, the proceeds ratio for over-the-counter trades is significantly higher than the proceeds ratio for foreclosure auctions (71 %).

Additionally, we control for changes in the real gross domestic product (GDP) and the unemployment rate (UER) at the time of the transferal to BAG.

4.3 Multivariate analysis

In the multivariate analysis, the proceeds ratio (PR_i) of property i serves as the dependent variable. The base categories are the appraisal reports issued by BAG (BAG), commercial property (COMPROP), property with unknown condition (Condition^{WITHOUT}) as well as over-the-counter trades (OTC). The complete model is as follows.

$$PR_{i,t} = \beta_0 + \beta_1 \cdot BANK_i + \beta_2 \cdot COURT_i + \beta_3 \cdot MV_i^{LN} + \beta_4 \cdot Condition_i^{NORMAL}$$
$$+ \beta_5 \cdot Condition_i^{BAD} + \beta_6 \cdot LOCATION_i + \beta_7 \cdot APH_i + \beta_8 \cdot CONDO_i$$
$$+ \beta_9 \cdot SFH_i + \beta_{10} \cdot FO_i + \beta_{11} \cdot GDP_t + \beta_{12} \cdot UER_t + \epsilon_i.$$

$$(6)$$

In contrast to the recovery rate, the proceeds ratio is not limited to the interval [0,1]. However, the proceeds ratio only takes positive values. Hence, we employ a Tobit model using robust standard errors that was proposed by Tobin (1958) and is designed for the case of a non-negative-dependent variable.

The estimation results can be found in Table 7, where we display the average marginal effects at the means of the covariates of the Tobit regression. The results from the bivariate analysis for the authors of the appraisal reports can be confirmed: appraisal reports prepared by the loan issuing bank or the court are associated with a significantly lower proceeds ratio than reports by BAG.[24] As mentioned before, a possible explanation for this finding could be the fact that in comparison to reports prepared by the court, BAG and the originating banks have to bear the consequences themselves if their estimate is wrong, and therefore make more conservative

---

[24] Because of potential differences in the effects of the independent variables depending on the authors of the appraisal reports, which might not be obvious in a bivariate analysis, we performed, by analogy to the customer groups, a sample split as a robustness test. In other words, the overall sample was split into three subsamples. There was no change in the essential results, especially regarding the property's location and condition. In particular, keeping in mind future analyses and also for the sake of clarity, we consider it more reasonable to display the authors of the appraisal reports with binary variables.

Table 7 Results of Tobit estimation with proceeds ratio as dependent variable

Variable	Model (3) Proceeds ratio marginal effect (standard error)
BANK	−0.1300*** (0.033)
COURT	−0.2537*** (0.031)
MV^{LN}	−0.0433** (0.017)
Condition^{NORMAL}	−0.0285 (0.031)
Condition^{BAD}	−0.1319*** (0.033)
LOCATION	0.2420*** (0.043)
APH	−0.0442 (0.043)
CONDO	−0.0902** (0.039)
SFH	0.0002 (0.032)
FO	−0.1691*** (0.026)
GDP	−0.0153** (0.007)
UER	0.0266** (0.011)
Constant	1.1019*** (0.220)
Observations	718
Pseudo R^2	0.3068

***, **, * indicates a 1, 5 and 10 % confidence level. The robust standard errors are given in brackets

estimates; the estimates of BAG seem to be even more conservative than the estimates of the originating banks.

Interestingly, our results still indicate that properties with high values, as measured by taking the market value's logarithm, are linked to significantly lower proceeds ratios. This is consistent with the results of Grunert (2010).[25] This result seems especially reasonable for real estate, bearing in mind that with an increasing market value, the number of potential buyers decreases.

Regarding the property's condition, properties in bad condition show a significantly lower proceeds ratio. See Sect. 4.2 for a discussion of this finding in the bivariate context. Our results are in line with the educated guess by Schaaff (2009), who was not able to analyze the influence of condition in more detail because of missing data.

Besides a property's condition, its location also has a significant influence on the proceeds ratio, as already demonstrated in the bivariate analysis. Objects located in an area that is attractive according to the index show a significantly higher proceeds ratio. In this way, our results indicate that the measurement of the attractiveness of a location on a very fine regional scale and a distinction between the types of properties are important features when controlling for a property's location. Our approach is more precise in catching regional particularities and object specifics,

---

[25] We have to point out that the analysis by Grunert (2010) is not only based on collateral in real estate. Furthermore, in his study, a piece of collateral is assumed to be 'small' if its value is less than EUR 660,000.

and is in contrast to the studies by, for example, Qi and Yang (2009) and Leow and Mues (2012) who use broader house price indices in the context of recovery rate literature. A study explicitly analyzing the proceeds ratio using an index controlling for location has not yet been performed.

When analyzing the properties' market values, it was mentioned that the negative influence of the logarithmized market value on the proceeds ratio could be caused by the limited class of potential buyers since an increasing market value is often linked to an increasing specificity. On the contrary, the property type 'condominium' is linked to a significantly lower proceeds ratio in comparison to commercial properties. In comparison to other types of properties, especially commercial properties, this type of property is characterized by a very low specificity. However, realizing more than one condominium of a single debtor usually results in a package price that is lower than the sum of the individual prices. This could offer an explanation for the negative influence.

Reviewing the control variables, a highly significant influence of the workout-process can be demonstrated. In the multivariate context, the proceeds ratio for property sold through foreclosures is significantly lower than the proceeds ratio achieved by over-the-counter sales. Nevertheless, it remains questionable whether the workout-process can really be considered to be an influencing factor or whether the workout-process is only a result of the various characteristics of the property. In conclusion, the influence of the workout-process could also be caused by the fact that many characteristics, such as the year of construction, cannot be analyzed but are partly captured by the type of workout-process.

There are some empirical analyses from the US discussing a possible discount for foreclosure auctions (e.g., Shilling et al. 1990; Carroll et al. 1997; Clauretie and Daneshvary 2009). These studies usually try to explain the selling price or the premium instead of the proceeds ratio, using hedonistic regressions. Nevertheless, there are certain similarities with regard to the content of these studies. Shilling et al. (1990) identify a significant discount of up to 24 %, and Carroll et al. (1997) a discount between 12 and 14 %, for foreclosure auctions in comparison to over-the-counter trades. More recent studies, which are testing for more discriminating factors, assume that the differences are determined by the property's characteristics instead of the type of workout-process. The discount for foreclosure auctions equals only 10 % in the study by Clauretie and Daneshvary (2009) and 3–7 % in the study by Campbell et al. (2011).

We find significant influences of our macroeconomic control variables on the proceeds ratio. There is a significantly positive relation between the unemployment rate (UER) and the proceeds ratio, and a negative relation between real GDP growth and the proceeds ratio.[26] The direction of the macroeconomic influences is surprising at first glance. But the results regarding the unemployment rate (UER) are consistent with the results published by Schaaff (2009) and partially so with those

---

[26] Again, to overcome the concern that our macroeconomic variables do not pick up all macroeconomic effects, we perform a robustness test using year dummies for the respective default year (see Table 15 in the Appendix). All our results stay qualitatively the same. Therefore, we consider our macroeconomic variables to be suitable on the collateral level as well.

published by Grunert (2010), both for German data sets. The negative influence of real GDP growth on the proceeds ratio could be caused by the observation period that includes the financial crisis. To cross-check the consistency of our findings, we calculated the correlation of the percentage change of the house price index reported by the German Federal Statistical Office (Statistisches Bundesamt) in comparison to the preceding year, as well as the percentage change of the residential property index (likewise reported by the German Federal Statistical Office) with the annual real GDP growth for the period 2005 through 2011. We find for both indices a negative, though not significant, correlation with real GDP growth, which supports our findings. Schaaff (2009) and Grunert (2010) find no significant influence of GDP growth on the proceeds ratio. Likely this variable sometimes picks up artifacts.

In conclusion, we are able to identify various factors that have a significant relationship with the proceeds ratio. However, considering the point in time, we have to differentiate between two kinds of factors: there are factors already known at the time when the appraisal report is prepared, such as the condition and the location, and there are control variables which we only get to know after the workout-process is completed, such as the type of liquidation. Since the factors that are known at the time of preparing the appraisal report significantly influence the proceeds ratio even when we control for other variables, there is an indication of a systematic bias in appraisal reports. Therefore, Hypothesis H4 has to be rejected.

However, the question arises as to how this may affect the explanation of the recovery rate. To give an answer, in the following section (Sect. 5.1) we will develop a prediction model based on our results to identify an individual correction factor δ_i for each individual property i. This equals the predicted proceeds ratio $\hat{PR}_i$. Multiplying this factor by the market value results in the adjusted market value and, therefore, in a new adjusted collateralization ratio (Sect. 5.2). In consequence, we are able to analyze the influence of the adjustment at the borrower level (Sect. 5.3).

5 Influence of the adjusted market values on the recovery rate

5.1 Prediction model

In the context of our prediction model, we have to differentiate between two groups of properties: one group is for calibrating the model (group^{CAL}), and the other group is for predicting the proceeds ratio (group^{PRED}). If the observations for the two groups are identical, it is a mere in-sample analysis, whereas we have an out-of-sample analysis if there is a difference between group^{CAL} and group^{PRED}. In the following, we will describe the composition of both groups in more detail.

For the context of our analysis, group^{PRED} always consists of all 711 properties assigned to the 499 debtors for whom the workout-process was completed (see Fig. 6). To calibrate the model, we use three differently compounded groups. group$_A^{CAL}$ consists of the 718 properties that were analyzed in the previous section. group$_B^{CAL}$ consists of the 324 properties for which the workout-process on

Fig. 6 Borrower and collateral
levels
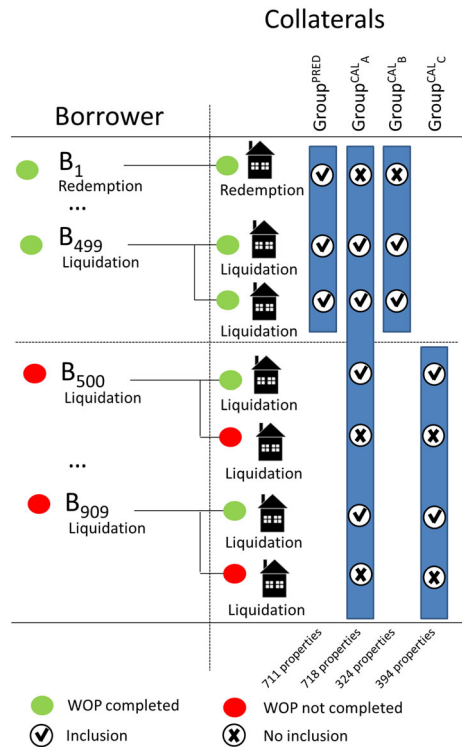
the borrower level was completed, while $\text{group}_C^{\text{CAL}}$ is made up of the 394 properties with incomplete workout-processes on the borrower level.

Therefore, we have a (partial) out-of-sample analysis for the subsequent observation at the borrower level for all three calibrating groups because the calibrating group and the estimation group always differ. The strongest out-of-sample test is that for $\text{group}_C^{\text{CAL}}$ because there is no overlap between the analyzed borrowers and the collateral used for calibration.

In the context of the calibration of the three prediction models, we use the proceeds ratio (PR_i) of property i as the value to be explained and the factors known at the time of creating the appraisal report as the explanatory variables (i.e., we exclude FO, GDP and UER). The explanatory variables are the author of the appraisal reports, the condition and the type of the property and its location, and the (expected) market value of the property. The base categories are the binary variables BAG, $\text{Condition}^{\text{WITHOUT}}$, and COMPROP. The complete model is the following.

$$\text{PR}_i = \beta_0 + \beta_1 \cdot \text{BANK}_i + \beta_2 \cdot \text{COURT}_i + \beta_3 \cdot \text{MV}_i^{\text{LN}} + \beta_4 \cdot \text{Condition}_i^{\text{NORMAL}}$$
$$+ \beta_5 \cdot \text{Condition}_i^{\text{BAD}} + \beta_6 \cdot \text{LOCATION}_i + \beta_7 \cdot \text{APH}_i + \beta_8 \cdot \text{CONDO}_i$$
$$+ \beta_9 \cdot \text{SFH}_i + \epsilon_i. \tag{7}$$

The Tobit model using robust standard errors is once more applied for the estimation. Since not all the included variables are of statistical significance, but are

Table 8 Tobit regression with proceeds ratio as dependent variable

Variable	Model (4) Proceeds ratio marginal effect (standard error)	Model (5) Proceeds ratio marginal effect (standard error)	Model (6) Proceeds ratio marginal effect (standard error)
BANK	−0.1042*** (0.0327)		−0.1680*** (0.0437)
COURT	−0.2749*** (0.0280)	−0.2455*** (0.0347)	−0.2666*** (0.0404)
MV^{LN}	−0.0398** (0.0170)	−0.0560* (0.0294)	
Condition^{NORMAL}			
Condition^{BAD}	−0.1416*** (0.0285)	−0.1647*** (0.0406)	−0.1138*** (0.0419)
Location	0.2562*** (0.0417)	0.3103*** (0.0723)	0.1453*** (0.0432)
APH			−0.1001** (0.0464)
CONDO	−0.0986*** (0.0326)	−0.1048** (0.0505)	−0.0807** (0.0372)
SFH			
Constant	1.1347*** (0.1997)	1.2458*** (0.3351)	0.7881*** (0.0523)
Observations	718	324	394
Pseudo R^2	0.2135	0.2877	0.1900

In Model (4), (5), and (6), the estimation of the coefficients is based on group$_A^{CAL}$, group$_B^{CAL}$, and group$_C^{CAL}$, respectively. ***, **, * indicates a 1, 5 and 10 % confidence level. The robust standard errors are given in brackets

correlated with the remaining factors of influence and could, therefore, influence the quality of the estimation, we use a backward elimination for the estimation models.[27]

The average marginal effects at the means of the Tobit regression are displayed in Table 8 for the three calibration groups. On the one hand, it becomes obvious that the majority of the variables for all three models (e.g., Condition^{BAD}, Location) have statistical significance. On the other hand, there are considerable differences between other variables depending on the model. For example, the preparation of the appraisal report by the bank is not included in Model (5) due to its lack of statistical significance. Because of these differences, it seems interesting to analyze how adjusting the market value influences the recovery rate.

5.2 Adjusting the market values

In the next step, as previously explained, we use the three calibrated estimation models (A, B and C) to estimate individual correction factors $\delta_{i,k}$ with $k \in (A, B, C)$ for all 711 properties. We analyze whether and how the result on the collateral level influences the recovery rate, i.e., the borrower level. Thus, we

---

[27] A partial F-test, testing how the coefficient of determination differs when a variable is excluded, given that the variable is not statistically significant, is the basis for excluding variables from the model. We ran the regressions without a backward elimination as well. Their results show nearly the same coefficients and significance level as our reported values for the backward elimination.

employ the individual correction parameters to determine the adjusted collateralization value for real estate. Equation (2) changes to

$$\text{CV}_{i,k}^{\text{real estate}} = \text{MAX}(\text{MIN}(\delta_{i,k} \cdot \text{MV}_i - \text{PL}_i; \text{LC}_i); 0). \tag{8}$$

Therefore, the adjusted collateralization ratio for every debtor j and the calibration group k is now defined as

$$\text{CR}_{j,k} = \frac{\sum_{i=1}^{m_j} \text{CV}_{i,j,k}^{\text{real estate}} + \sum_{i=1}^{n_j} \text{CV}_{i,j}^{\text{deposit}}}{\text{EAD}_j}. \tag{9}$$

Next, we form the corresponding collateralization ratios for all 499 debtors whose workout-processes were completed, and get three adjusted collateralization ratios CR_A, CR_B and CR_C, derived from the different calibration groups k. Again, we limit the collateralization ratio to the interval [0,1]. In German banking practice the correction of collateral values is not uncommon. But in contrast to our approach, which adjusts the collateral value depending on the associated attributes of the collateral, banking practice commonly applies a flat haircut to the collateral value. For collateral in real estate in Germany this haircut ranges between 10 and 40 % (e.g., Grunert and Weber 2009; Bruhn 2009; Pfnür 2011). Unlike this flat discount, our individual correction parameter can also take on values above 1 and, therefore, cause discounts and premiums. In the following, we compare our model-adjusted collateralization ratios with two flat haircuts of 20 % (CR_{80}) and 30 % (CR_{70}) as well as with our original results using the original collateralization ratio. Table 9 shows descriptive statistics for the original and the resulting adjusted collateralization ratios. Figure 7 illustrates the corresponding frequency distributions. The average original collateralization ratio drops in all three model-adjusted cases from 66 to 55 %. The collateralization ratio drops in the case of a 30 % flat haircut to as little as 50 % on average, whereas the resulting average of 56 % in the case of a 20 % haircut is highly comparable to the model-adjusted cases. If the sample is split between retail and commercial customers, the main results stay about the same. As discussed earlier, the collateralization ratios are considerably lower for commercial customers than for retail customers.

Table 9 Descriptive statistics of adjusted collateralization ratios

Variable	Full sample				Retail customers				Commercial customers			
	Mean	Median	σ	N	Mean	Median	σ	N	Mean	Median	σ	N
CR	0.66	0.73	0.32	499	0.70	0.77	0.31	368	0.58	0.59	0.34	131
CR_A	0.55	0.55	0.33	499	0.57	0.59	0.33	368	0.48	0.46	0.34	131
CR_B	0.55	0.55	0.33	499	0.57	0.60	0.33	368	0.49	0.46	0.34	131
CR_C	0.55	0.55	0.33	499	0.57	0.57	0.33	368	0.49	0.47	0.34	131
CR_{70}	0.50	0.47	0.32	499	0.52	0.50	0.31	368	0.43	0.40	0.32	131
CR_{80}	0.56	0.56	0.32	499	0.59	0.60	0.31	368	0.48	0.47	0.34	131
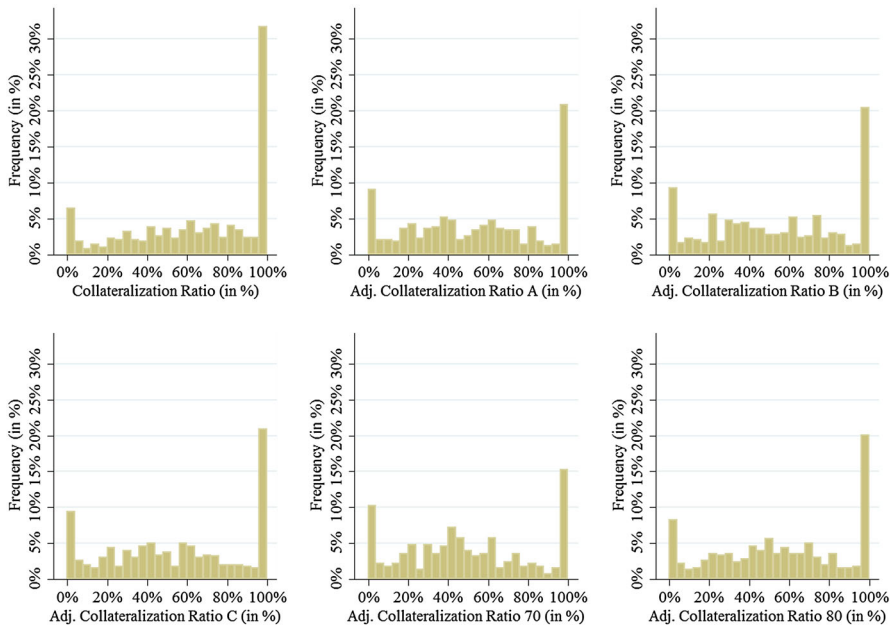
Fig. 7 Frequency distributions of adjusted collateralization ratios

5.3 Regression results

Using the Fractional Logit Model to explain the recovery rate at the borrower level, with robust standard errors that are clustered for the originating bank, leads to the results displayed in Table 11. Model (2) repeats the results of the regressions of the unadjusted collateralization ratio for retail customers [Model (2a)] and commercial customers [Model (2b)] already discussed in detail in Sect. 2.4. Models (7)–(11) show the adjusted collateralization ratios (CR_A, CR_B, CR_C, CR_{70} and CR_{80}) resulting from the different calibration groups as well as the flat haircuts explained in the previous section.

The results show that considering adjustments results in a significant improvement of the model's explanatory power. The value for the adjusted R^2 of Models (7)–(11) is higher than that of Model (2). For both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), lower values mean higher

Table 10 Degrees of preference according to Raftery (1995)

Absolute difference	Degree of preference
0–2	Weak
2–6	Positive
6–10	Strong
> 10	Very strong

explanatory power, therefore, confirming the necessity of an adjusted collateraliza-
tion ratio for a precise estimate of the recovery rate.

Another advantage of using the BIC to discriminate between models is the
possibility of evaluating the strength of preference for one model over another.
According to Raftery (1995), there are four categories oriented to the absolute
difference between two BIC values (see Table 10). In comparison to Model [2a(b)],
Models [7a(b)] to [9a(b)] result in absolute differences significantly above 10(6)
and, therefore, show a very strong (strong) preference for explanatory models using
the adjusted collateralization ratio. Interestingly, this result is very robust as it is
valid for all three calibration groups, and the maximum difference of the BIC over
Models (7) to (9) is 3.67, which is, according to Table 10, a positive though neither
strong nor even very strong preference for Model (7a) over Model (9a). The inferior
values are reasonable because in Model (9a) there is no relation between the
analyzed debtors and the collateral used for calibration (strongest out-of-sample
test). However, we are able to conclude that employing the correction parameters
significantly improves the explanatory power, and this result can be confirmed for
all three calibration groups.

For the haircut models (10) to (11), we find a positive preference [for Model (10a)
even a strong preference] in comparison to the unadjusted models (2). If we compare the
correction parameter models (7) to (9) with the flat haircut models (10) to (11), we find a
very strong preference for retail customers and a positive preference for commercial
customers for Models (7) to (9). The considerably higher adjusted R^2 confirms this
argument. Therefore, we can conclude that our models that use the correction
parameters outperform the unadjusted models and the flat haircut models.

So far, we have discussed the explanatory power of the various models. Now we
are going to analyze the significant changes in the influence of the different
variables when employing the adjusted collateralization ratios. While the standard-
ized coefficients of the collateralization ratio increase for Models (7) to (11), the
coefficients for redemption decrease for Models (7) to (9) and increase for Models
(10) to (11). Apparently, a part of the distorted illustration of collateral was captured
in the redemption variable.

Furthermore, we are able to demonstrate that some of the control variables
(EAD^{LN} and INS) for retail customers lose their previous significance, whereas
there is no change for commercial customers. The results for the macroeconomic
factors stay qualitatively the same for all models.[28]

6 Conclusions

Previously, academic research as well as banking practice have focused on models
and methods to estimate the probability of default (PD). Well-founded results for
the recovery rate are still rare, although their number has increased in recent years.

---

[28] Again, we use in a robustness test year dummies for the respective default year (see Table 14 in
the Appendix). All our results stay qualitatively the same. Therefore, we consider our
macroeconomic variable setup to be suitable.

Table 11 Fractional Logit Regression with recovery rate as dependent variable for the unadjusted collateralization ratio [Models (2)] as well as the five adjusted collateralization ratios [Models (7)–(11)]. Models with 'a' are for retail customers and models with 'b' are for commercial customers

Variable	Model (2a)	Model (2b)	Model (7a)	Model (7b)	Model (8a)	Model (8b)	Model (9a)	Model (9b)	Model (10a)	Model (10b)	Model (11a)	Model (11b)
Recovery rate std. coefficient (standard error)												
H1												
CR	0.6040*** (0.297)	0.7150*** (0.428)										
CR_A			0.7535*** (0.318)	0.8451*** (0.305)								
CR_B					0.7478*** (0.299)	0.8720*** (0.342)						
CR_C							0.7215*** (0.299)	0.8388*** (0.306)				
CR_{70}									0.6492*** (0.284)	0.8084*** (0.360)		
CR_{80}											0.6380*** (0.289)	0.8041*** (0.374)
H2												
COL	0.0427 (0.045)	0.0523 (0.045)	0.0610 (0.050)	0.0589 (0.034)	0.0534 (0.048)	0.0526 (0.042)	0.0723* (0.050)	0.0622 (0.039)	0.0422 (0.045)	0.0499 (0.046)	0.0443 (0.045)	0.0481 (0.046)
H3												
RED	0.4286*** (0.147)	0.2982*** (0.268)	0.3532*** (0.141)	0.2784*** (0.279)	0.3521*** (0.152)	0.2488*** (0.259)	0.3733*** (0.154)	0.2844*** (0.272)	0.4658*** (0.164)	0.3289*** (0.307)	0.4368*** (0.154)	0.3126*** (0.300)
Control variables												
EAD^{LN}	0.1002** (0.080)	−0.1494** (0.108)	0.0662 (0.080)	−0.1839*** (0.095)	0.0767* (0.078)	−0.1522*** (0.099)	0.0378 (0.079)	−0.2091*** (0.094)	0.0687 (0.102)	−0.1889*** (0.106)	0.0972** (0.089)	−0.1789*** (0.110)
WOP	−0.0292 (0.007)	0.0130 (0.011)	−0.0671 (0.006)	−0.0082 (0.011)	−0.0685* (0.006)	−0.0421 (0.008)	−0.0628 (0.006)	−0.0090 (0.011)	0.0005 (0.006)	0.0042 (0.011)	−0.0150 (0.007)	−0.0134 (0.009)
INS	−0.0757* (0.202)	−0.0019 (0.198)	−0.0448 (0.175)	0.0044 (0.198)	−0.0513 (0.174)	0.0159 (0.188)	−0.0403 (0.172)	0.0008 (0.194)	−0.0806** (0.171)	−0.0269 (0.211)	−0.0819** (0.184)	−0.0156 (0.200)

Table 11 continued

Variable	Recovery rate std. coefficient (standard error)											
	Model (2a)	Model (2b)	Model (7a)	Model (7b)	Model (8a)	Model (8b)	Model (9a)	Model (9b)	Model (10a)	Model (10b)	Model (11a)	Model (11b)
GDP	−0.0386	0.0289	−0.0030	0.0649	0.0208	0.0559	−0.0240	0.0486	−0.0538	0.0398	−0.0539	0.0382
	(0.362)	(0.061)	(0.041)	(0.068)	(0.040)	(0.066)	(0.043)	(0.064)	(0.043)	(0.068)	(0.043)	(0.067)
UER	0.0624	0.1495*	0.0284	0.0719**	0.0141	0.0721**	0.0439	0.0777**	0.0358	0.0738**	0.0293	0.0779**
	(0.073)	(0.106)	(0.069)	(0.097)	(0.065)	(0.091)	(0.077)	(0.097)	(0.086)	(0.104)	(0.078)	(0.099)
Constant	−5.3583***	−1.0187	−4.3194***	−0.2612	−4.1562***	−0.8298	−4.0371***	0.1489	−4.4999***	−0.0622	−5.0829***	−0.4242
	(1.224)	(1.485)	(1.069)	(1.383)	(1.071)	(1.392)	(1.018)	(1.388)	(1.182)	(1.538)	(1.215)	(−0.424)
Observations	368	131	368	131	368	131	368	131	368	131	368	131
AIC	0.7989	0.8838	0.7422	0.8287	0.7465	0.8178	0.7522	0.8323	0.7821	0.8564	0.7862	0.8512
BIC	−2013.36	−563.08	−2034.20	−570.30	−2032.65	−571.71	−2030.53	−569.81	−2019.52	−566.66	−2018.03	−567.34
Mc Fadden R^2	0.3411	0.3348	0.3908	0.3839	0.3871	0.3936	0.3821	0.3806	0.3558	0.3592	0.3522	0.3638
R^2	0.6350	0.6925	0.7143	0.7640	0.7076	0.7832	0.7010	0.7619	0.6426	0.7204	0.6457	0.7332
Adj. R^2	0.6268	0.6724	0.7079	0.7485	0.7011	0.7690	0.6943	0.7463	0.6347	0.7021	0.6378	0.7157
Wald test (p-value)	390.85	273.84	519.76	521.05	589.71	421.57	474.49	490.92	560.60	334.45	452.80	285.46
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

***, **, * indicates a 1, 5 and 10 % confidence level. The robust standard errors that are clustered for the originating bank are given in brackets

Banks seeking the transition to advanced IRB approaches have to demonstrate that they have the appropriate internal models to estimate the recovery rate. To keep the effort of collecting and processing data as small as possible, results from academic studies could provide essential insights for identifying the important factors influencing the recovery rate. However, the greater part of the empirical research is based on US and UK banking data. These results cannot be, or can only to a limited extent, transferred to Germany because of differences in the collateral and in the legal systems.

This article provides empirical evidence for the drivers of the recovery rate for retail and commercial customers in Germany, based on a data set of 909 defaulting customers from exactly this market segment whose original loans were with 123 different banks. We are able to confirm a positive relationship between the collateralization ratio as well as the type of workout-process and the recovery rate. Furthermore, we demonstrate the important role of redemption in the processing of problematic loans, because an amicable agreement with an already defaulting customer was achieved in 40 % of the cases. As to the two customer groups (retail and commercial), it becomes obvious that a separate analysis is necessary, because there are diametrically opposite effects for some of the control variables (e.g., the EAD). In total, our model is able to explain 63 and 67 %, respectively, of the variance of the customers' recovery rates measured with the adjusted R^2, although it seems likely that this result is merely based on the collateral in real estate.

Because of the major influence of collateral based on real estate, we analyzed the valuation of the properties included in the data set in more detail. The proceeds ratio, that is, the relation between the realized market price and the expected market value, is only 78 % on average. This result is surprising because the valuation in the form of the (expected) market value should already have included all the factors influencing the value. A multivariate analysis of the proceeds ratio shows that, among others, the condition of the property is not considered appropriately in the value derived from the appraisal reports. The same is the case for the attractiveness of the property's location. To approximate a location's attractiveness, we were able to use a postal code-based index depending on the property type. Subsequently, we developed a prediction model for the proceeds ratio to eliminate the identified systematic bias in the market value. In this way, the present article goes beyond previous studies.

In a next step, we use the predicted proceeds ratio to adjust the market value of real estate. We thereby obtain an adjusted collateralization ratio and then create additional adjusted collateralization ratios by applying, in accordance with common banking practice, flat haircuts of 20 and 30 % to the collateral in real estate. The new analysis of the recovery rate provides several new insights. First, the models' explanatory power can be increased by adjusting by the recognized biases and using the adjusted collateralization ratio. This is confirmed by a (partially) out-of-sample test. Therefore, we can conclude that our models that use the correction parameter outperform the unadjusted models and the flat haircut models. Second, some of the control variables for retail customers lose their previously significant influence. The combined analysis at the borrower and collateral levels should, therefore, be a

crucial part of the analysis of factors influencing the recovery rate in future empirical studies.

Our results make an important contribution to further academic research. On the one hand, the analysis should be applied separately for every customer group because single variables could function in diametrically opposite ways. On the other hand, factors influencing the recovery rate that were identified in other studies (in particular for retail customers), could be caused by the biased valuation of the collateral. Future studies should, therefore, not consider the value of the collateral as a given. Additionally, the banks should review their current practice of flat haircuts and should instead apply adjustments of the collateral value depending on the associated attributes of the collateral as presented in our study. Moreover, the final model is able to explain a large fraction of the recovery rate's variance (adjusted R^2 equals approximately 70–77 %) and is able to identify the collateralization ratio as the essential driver of the recovery rate. Therefore, the banking practice should concentrate on the collateralization ratio and its unbiased form. To allow a systematic analysis of the collateralization ratio, the collection and processing of data should focus on the creation of databases for real estate collateral upon which there have been defaults.

Appendix

Table 12 Correlation matrix for the variables on the borrower level

Variable	EAD^{LN}	CR	COL	RED	BT	WOP	INS	GDP	UER	RR
EAD^{LN}	1.00									
CR	−0.20*	1.00								
COL	0.31*	−0.09	1.00							
RED	−0.03	0.31*	0.06	1.00						
BT	0.35*	−0.16*	0.36*	−0.01	1.00					
WOP	0.18*	0.04	0.10*	−0.10*	0.10*	1.00				
INS	0.05	−0.12*	0.03	−0.27*	0.06	0.14*	1.00			
GDP	−0.01	−0.09	−0.05	−0.18*	0.03	0.17*	0.09	1.00		
UER	−0.13*	−0.08	0.08	0.07	0.09*	0.35*	−0.14*	−0.24*	1.00	
RR	−0.07	0.70*	0.03	0.55*	-0.02	0.01	−0.23*	−0.17*	0.16*	1.00

* indicates a 5 % confidence level

Table 13 Correlation matrix for the variables on the collateral level

Variable	BANK	BAG	COURT	MV^{LN}	$Condition^{WITHOUT}$	$Condition^{NORMAL}$	$Condition^{BAD}$	LOCATION
BANK	1.00							
BAG	−0.66*	1.00						
COURT	−0.30*	−0.52*	1.00					
MV^{LN}	−0.12*	0.18*	−0.10*	1.00				
$Condition^{WITHOUT}$	0.21*	−0.33*	0.18*	−0.17*	1.00			
$Condition^{NORMAL}$	−0.16*	0.32*	−0.22*	0.27*	−0.62*	1.00		
$Condition^{BAD}$	−0.06	0.01	0.05	−0.11*	−0.45*	−0.43*	1.00	
LOCATION	0.02	0.05	−0.09*	0.31*	−0.02	0.08*	−0.07	1.00
APH	0.08*	−0.06	−0.02	0.06	0.26*	−0.18*	−0.10*	0.03
COMPROP	0.00	0.09*	−0.12*	0.22*	−0.18*	0.09*	0.10*	−0.06
CONDO	0.01	−0.06	0.07	−0.30*	−0.01	0.08*	−0.07	0.11*
SFH	−0.07	0.00	0.08*	−0.02	0.00	−0.03	0.04	−0.05
FO	−0.05	−0.11*	0.19*	−0.09*	0.02	−0.10*	0.09*	0.01
GDP	−0.00	−0.01	0.01	0.01	−0.09*	−0.05	0.17*	0.01
UER	0.28*	−0.24*	−0.02	0.12*	0.03	−0.02	−0.01	0.30*
PR	−0.01	0.23*	−0.28*	0.06	0.02	0.12*	−0.17*	0.23*

Variable	APH	COMPROP	CONDO	SFH	FO	GDP	UER	PR
BANK								
BAG								
COURT								
MV^{LN}								
$Condition^{WITHOUT}$								
$Condition^{NORMAL}$								

Table 13 continued

Variable	APH	COMPROP	CONDO	SFH	FO	GDP	UER	PR
Condition[BAD]								
LOCATION								
APH	1.00							
COMPROP	−0.32*	1.00						
CONDO	−0.20*	−0.39*	1.00					
SFH	−0.25*	−0.49*	−0.31*	1.00				
FO	−0.03	−0.07	0.05	0.05	1.00			
GDP	−0.05	0.04	−0.03	0.02	0.13*	1.00		
UER	0.11*	0.01	−0.06	−0.04	0.03	0.06	1.00	
PR	0.01	0.04	−0.05	0.00	−0.27*	0.00	0.06	1.00

* Indicates a 5 % confidence level

Table 14 Fractional logit regression with recovery rate as dependent variable for the unadjusted collateralization ratio [Models (1)–(2)] as well as the five adjusted collateralization ratios [Models (7)–(11)] using default year dummies

Variable	Recovery rate std. coefficient (standard error)					
	Model (1)	Model (2a)	Model (2b)	Model (7a)	Model (7b)	Model (8a)
H1						
CR	0.6216*** (0.215)	0.5959*** (0.300)	0.7095*** (0.430)			
CR$_A$				0.7472*** (0.324)	0.8423*** (0.311)	
CR$_B$						0.7410*** (0.305)
CR$_C$						
CR$_{70}$						
CR$_{80}$						
H2						
COL	0.0321 (0.036)	0.0414 (0.047)	0.0353 (0.039)	0.0586 (0.051)	0.0382 (0.031)	0.0506 (0.049)
H3						
RED	0.3845*** (0.135)	0.4284*** (0.151)	0.3125*** (0.256)	0.3562*** (0.152)	0.2968*** (0.259)	0.3563*** (0.162)
Control variables						
EAD^{LN}	0.0342 (0.070)	0.1066** (0.083)	−0.1483** (0.110)	0.0696 (0.083)	−0.1832*** (0.098)	0.0802* (0.081)
WOP	0.0117 (0.005)	−0.0215 (0.006)	0.0093 (0.011)	−0.0650 (0.006)	−0.0150 (0.011)	−0.0667*** (0.006)
BT	0.0693** (0.132)					
INS	−0.0709*** (0.162)	−0.0828** (0.205)	0.0220 (0.226)	−0.0450 (0.182)	0.0335 (0.219)	−0.0508 (0.181)
Default year 2006	−0.0612* (0.259)	−0.0488 (0.415)	−0.0713 (0.239)	−0.0341 (0.479)	−0.0855 (0.274)	−0.0294 (0.471)
Default year 2007	−0.1192* (0.223)	−0.1067* (0.220)	−0.1353 (0.321)	−0.0696 (0.196)	−0.1554** (0.277)	−0.0264 (0.190)
Default year 2008	−0.0707 (0.343)	−0.0381 (0.348)	−0.1331 (0.473)	−0.0561 (0.308)	−0.1464* (0.441)	−0.0311 (0.300)
Default year 2009	−0.0374 (0.381)	−0.0338 (0.446)	−0.1059*** (0.675)	−0.0407 (0.489)	−0.1246*** (0.566)	−0.0328 (0.456)
Default year 2010	0.0230** (0.371)	0.0356*** (0.346)		0.0004 (0.346)		0.0057 (0.326)

Table 14 continued

Variable	Recovery rate std. coefficient (standard error)					
	Model (1)	Model (2a)	Model (2b)	Model (7a)	Model (7b)	Model (8a)
Constant	−2.9488*** (0.956)	−4.5354*** (1.024)	1.2843 (1.821)	−3.3965*** (1.085)	2.3082 (1.479)	−3.7079*** (1)
Observations	499	368	131	368	131	368
AIC	0.8059	0.8073	0.8950	0.7527	0.8381	0.7569
BIC	−2881.81	−2002.45	−558.73	−2022.53	−566.19	−2021.00
Mc Fadden R^2	0.3317	0.3432	0.3384	0.3912	0.3891	0.3875
R^2	0.6346	0.6388	0.7002	0.7145	0.7745	0.7083
Adj. R^2	0.6233	0.6276	0.6752	0.7057	0.7557	0.6993
Wald test (p-value)	461.99 (0.000)	371.87 (0.000)	266.20 (0.000)	508.64 (0.000)	461.04 (0.000)	580.03 (0.000)

Variable	Recovery rate std. coefficient (standard error)						
	Model (8b)	Model (9a)	Model (9b)	Model (10a)	Model (10b)	Model (11a)	Model (11b)
H1							
CR							
CR_A							
CR_B	0.8729*** (0.342)						
CR_C		0.7136*** (0.305)	0.8331*** (0.313)				
CR_{70}				0.6380*** (0.288)	0.8015*** (0.358)		
CR_{80}						0.6271*** (0.293)	0.7983*** (0.368)
H2							
COL	0.0306 (0.032)	0.0691* (0.050)	0.0434 (0.031)	0.0367 (0.045)	0.0333 (0.039)	0.0404 (0.046)	0.0304 (0.039)

Table 14 continued

Variable	Recovery rate std. coefficient (standard error)						
	Model (8b)	Model (9a)	Model (9b)	Model (10a)	Model (10b)	Model (11a)	Model (11b)
H3							
RED	0.2632*** (0.243)	0.3770*** (0.163)	0.3028*** (0.256)	0.4726*** (0.168)	0.3482*** (0.287)	0.4408*** (0.159)	0.3302*** (0.280)
Control variables							
EAD^{LN}	−0.1543** (0.103)	0.0426 (0.082)	−0.2070*** (0.097)	0.0787 (0.104)	−0.1852*** (0.108)	0.1046** (0.093)	−0.1766*** (0.111)
WOP	−0.0488 (0.009)	−0.0598 (0.006)	−0.0146 (0.012)	0.0052 (0.006)	−0.0015 (0.012)	−0.0087 (0.007)	−0.0184 (0.010)
BT							
INS	0.0438 (0.210)	−0.0407 (0.179)	0.0276 (0.215)	−0.0800** (0.176)	0.0017 (0.234)	−0.0847** (0.187)	0.0126 (0.226)
Default year 2006	−0.0697 (0.297)	−0.0507 (0.486)	−0.0947 (0.234)	−0.0891 (0.487)	−0.1056* (0.296)	−0.0685 (0.458)	−0.1005* (0.298)
Default year 2007	−0.1653** (0.263)	−0.1146* (0.225)	−0.1694** (0.268)	−0.1142 (0.266)	−0.1589* (0.304)	−0.1026 (0.239)	−0.1720** (0.288)
Default year 2008	−0.1353 (0.437)	−0.0873 (0.337)	−0.1645* (0.441)	−0.0774 (0.373)	−0.1618* (0.479)	−0.0545 (0.347)	−0.1652* (0.472)
Default year 2009	−0.1202*** (0.565)	−0.0532 (0.512)	−0.1174*** (0.550)	−0.0447 (0.455)	−0.1163*** (0.640)	−0.0369 (0.444)	−0.1175*** (0.632)
Default year 2010		−0.0021 (0.369)		0.0116 (0.402)		0.0228** (0.396)	
Constant	1.7888 (1.558)	−2.6207** (1.069)	2.8703* (1.501)	−3.4590** (1.368)	2.4738 (1.658)	−4.2164*** (1.151)	2.2842 (1.714)
Observations	131	368	131	368	131	368	131
AIC	0.8279	0.7624	0.8425	0.7904	0.8657	0.7950	0.8610
BIC	−567.51	−2018.98	−565.61	−2008.68	−562.57	−2006.98	−563.18

Table 14 continued

Variable	Recovery rate std. coefficient (standard error)							
	Model (8b)	Model (9a)	Model (9b)	Model (10a)	Model (10b)	Model (11a)	Model (11b)
Mc Fadden R^2	0.3981	0.3827	0.3852	0.3581	0.3645	0.3540	0.3687
R^2	0.7927	0.7016	0.7712	0.6466	0.7326	0.6489	0.7441
Adj. R^2	0.7754	0.6924	0.7521	0.6357	0.7103	0.6381	0.7228
Wald test (p-value)	406.69 (0.000)	460.33 (0.000)	436.77 (0.000)	534.38 (0.000)	308.28 (0.000)	420.37 (0.000)	286.39 (0.000)

Models with 'a' are for retail customers and models with 'b' are for commercial customers

***, **, * indicates a 1, 5 and 10 % confidence level. The robust standard errors that are clustered for the originating bank are given in brackets

Table 15 Results of Tobit estimation with proceeds ratio as dependent variable and default year dummies

Variable	Model (3) Proceeds ratio marginal effect (standard error)
BANK	−0.1314*** (0.033)
COURT	−0.2642*** (0.030)
MV^{LN}	−0.0423** (0.017)
Condition^{NORMAL}	−0.0257 (0.031)
Condition^{BAD}	−0.1263*** (0.033)
LOCATION	0.2350*** (0.044)
APH	−0.0411 (0.043)
CONDO	−0.0880** (0.039)
SFH	0.003 (0.032)
FO	−0.1685*** (0.026)
Default year 2006	−0.0719 (0.080)
Default year 2007	−0.1206*** (0.039)
Default year 2008	−0.1338*** (0.047)
Default year 2009	0.0019 (0.069)
Default year 2010	0.0043 (0.113)
Constant	1.3995*** (0.213)
Observations	718
Pseudo R^2	0.3129

***, **, * indicates a 1, 5 and 10 % confidence level. The robust standard errors are given in brackets

Table 16 Fractional logit regression with the recovery rate as dependent variable (without censored cases)

Variable	Recovery rate std. coefficient (standard error)	
	Model (12a)	Model (12b)
H1		
CR	0.4621*** (0.26)	0.5148*** (0.421)
H2		
COL	0.0408 (0.044)	0.0349 (0.035)
Control variables		
RED	0.217*** (0.135)	0.1908*** (0.233)
EAD^{LN}	0.1535*** (0.072)	−0.1157* (0.113)
WOP	0.0314 (0.005)	0.024 (0.009)
INS	−0.0472 (0.169)	−0.0006 (0.169)
GDP	−0.0169 (0.04)	0.0273 (0.037)
UER	0.0352 (0.071)	0.0906 (0.096)

Table 16 continued

Variable	Recovery rate std. coefficient (standard error)	
	Model (12a)	Model (12b)
Constant	−6.3662*** (1.054)	−0.5831 (1.556)
Observations	265	99
AIC	0.9410	1.0288
BIC	−1365.40	−395.7815
Mc Fadden R^2	0.1841	0.2002
R^2	0.5035	0.6042
Adj. R^2	0.4879	0.5690
Wald test (p-value)	322.32 (0.000)	209.62 (0.000)

***, **, * Indicates a 1%, 5% and 10% confidence level. The robust standard errors that are clustered for the originating bank are given in brackets

Table 17 Fractional Logit Regression with the recovery rate as dependent variable (liquidation vs. redemption)

Variable	Recovery rate std. coefficient (standard error)			
	Model (13a)	Model (13b)	Model (14a)	Model (14b)
H1				
CR	0.5403*** (0.354)	0.7535*** (0.475)	0.6258*** (0.425)	0.6717*** (0.521)
H2				
COL	0.0205 (0.048)	0.0883 (0.051)	0.1267 (0.095)	0.0686 (0.046)
Control variables				
EAD^{LN}	0.1561*** (0.090)	−0.0536 (0.138)	−0.0757 (0.107)	−0.2507** (0.192)
WOP	0.0010 (0.008)	0.0673 (0.016)	−0.1085 (0.012)	−0.0342 (0.014)
INS	−0.0781 (0.206)	0.0164 (0.263)	−0.0990* (0.685)	−0.1206** (0.332)
GDP	−0.0041 (0.049)	−0.0401 (0.104)	−0.0491 (0.061)	0.0066 (0.084)
UER	0.0592 (0.071)	−0.0215 (0.184)	0.0893 (0.103)	0.3003*** (0.127)
Constant	−6.6852*** (1.338)	−1.173 (2.247)	−0.8574 (1.247)	0.7023 (2.956)
Observations	223	81	145	50
AIC	0.9650	0.9786	0.6009	0.9665
BIC	−1089.97	−300.99	−652.40	−154.60

Table 17 continued

Variable	Recovery rate std. coefficient (standard error)			
	Model (13a)	Model (13b)	Model (14a)	Model (14b)
Mc Fadden R^2	0.1679	0.3017	0.3003	0.3112
R^2	0.4023	0.6807	0.6003	0.6794
Adj. R^2	0.3829	0.6501	0.5799	0.6260
Wald test (p-value)	153.92 (0.000)	206.68 (0.000)	192.50 (0.000)	166.58 (0.000)

***, **, * Indicates a 1, 5 and 10 % confidence level. The robust standard errors that are clustered for the originating bank are given in brackets

References

Altman, E. 1989. Measuring corporate bond mortality and performance. *The Journal of Finance* 28: 909–922.

Altman, E., B. Brady, A. Resti, and A. Sironi. 2005. The link between default and recovery rates: theory, empirical evidence, and implications. *Journal of Business* 78: 2203–2227.

Appasamy, B., U. Dörr, H. Ebel, and E.A. Stützle. 2008. LGD-Schätzung im Retailgeschäft am Beispiel Automobilfinanzierung. *Zeitschrift für das gesamte Kreditwesen* 61: 206–209.

Araten, M., M. Jacobs, and P. Varshney. 2004. Measuring LGD on commercial loans: an 18-year internal study. *The Risk Management Journal* 86: 96–103.

Asarnow, E., and D. Edwards. 1995. Measuring loss on defaulted bank loans: a 24-year study. *The Journal of Commercial Lending* 77: 11–23.

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

Basel Committee on Banking Supervision. 2005. Guidance on paragraph 468 of the framework document. Basel.

Basel Committee on Banking Supervision. 2006. International convergence of capital measurement and capital standards. Basel.

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

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

Bos, R., K. Kelhoffer, and D. Keisman. 2002. *Ultimate recovery in an era of record defaults*. New York: Standard & Poor's Risk Solutions.

Bruhn, S. 2009. Immobilienfinanzierung: das Beratungs- und Leistungsprogramm der Kreditinstitute. In *Kapitalanlage mit Immobilien*, ed. M. Brunner, 205–219. Wiesbaden: Gabler.

Calabrese, R., and M. Zenga. 2010. Bank loan recovery rates: measuring and nonparametric density estimation. *Journal of Banking and Finance* 34: 903–911.

Calem, P.S., and M. LaCour-Little. 2004. Risk-based capital requirements for mortgage loans. *Journal of Banking and Finance* 28: 647–672.

Campbell, J.Y., S. Giglio, and P. Pathak. 2011. Forced sales and house prices. *American Economic Review* 101: 2108–2131.

Carroll, T.M., T.M. Clauretie, and H.R. Neill. 1997. Effect of foreclosure status on residential selling price: comment. *Journal of Real Estate Research* 13: 95–102.

Carty, L.V., D.T. Hamilton, S.C. Keenan, A. Moss, M. Mulvaney, T. Marshella, and M.G. Subhas. 1998. *Bankrupt bank loan recoveries*. New York: Moody's Investors Service.

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

Clauretie, T.M., and N. Daneshvary. 2009. Estimating the house foreclosure discount corrected for spatial price interdependence and endogeneity of marketing time. *Real Estate Economics* 37: 43–67.

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

Deutsche Bundesbank. 2014. Bankenstatistik—Oktober 2014. Frankfurt.

Dr. Klein & Co AG. 2012. Trendindikator für Baufinanzierung—Juni 2012. http://www.drklein.de/dtb-juni-2012.html.

Elsas, R., and J.P. Krahnen. 2002. Collateral, relationship lending, and financial distress: an empirical study on financial contracting. Center for Financial Studies Working Paper No. 2002 / 17, Frankfurt.

Emery, K., R. Cantor, and R. Arner. 2004. *Recovery rates on North American syndicated bank loans, 1989–2003*. New York: Moody's Investors Service.

European Mortgage Federation. 2007. Study on the efficiency of the mortgage collateral in the European Union. Brussels.

F+B GmbH. 2013. F+B-Marktpreise (Immobilienpreise)—Methodensteckbrief. http://www.f-und-b.de/pdf/geschaeftsbereich/gb1/F+B_114_Marktpreise_Methode.pdf.

Franks, J., A. de Servigny, and S. Davydenko. 2004. A comparative analysis of the recovery process and recovery rates for private companies in the U.K., France and Germany. Standard & Poor's Risk Solutions Report, London.

Friedman, C., and S. Sandow. 2003. Ultimate recoveries. *Risk* 16: 69–73.

Frye, J. 2000. Collateral damage. *Risk* 13: 91–94.

Greene, W.H. 2012. *Econometric analysis*. Essex: Pearson.

Grippa, P., S. Iannotti, and F. Leandri. 2005. Recovery rates in the banking industry: stylised facts emerging from the Italian experience. In *Recovery risk: the next challenge in credit risk management*, ed. E.I. Altman, A. Resti, and A. Sironi, 121–141. London: Risk Books.

Grunert, J. 2005. Empirische Evidenz zur Prognose der Ausfallwahrscheinlichkeit und der Recovery Rate von Bankkrediten an deutsche Unternehmen. Dissertation, University of Mannheim.

Grunert, J. 2009. Kreditsicherheiten: eine Umfrage zu deren Häufigkeit, den erzielten Verwertungserlösen und der Bewertungsfrequenz. *Zeitschrift für Bankrecht und Bankwirtschaft* 2: 126–135.

Grunert, J. 2010. Verwertungserlöse von Kreditsicherheiten: Eine empirische Analyse notleidender Unternehmenskredite. *Zeitschrift für Betriebswirtschaft* 80: 1305–1323.

Grunert, J., and A. Volk. 2008. Die Bedeutung der Ausfalldefinition bei der Berechnung der Recovery Rate von Unternehmenskrediten. *Finanz-Betrieb* 5: 317–326.

Grunert, J., and M. Weber. 2009. Recovery rate of commercial lending: empirical evidence for German companies. *Journal of Banking and Finance* 33: 505–513.

Gupton, G.M., D. Gates, and L.V. Carty. 2000. *Bank loan loss given default*. New York: Moody's Investors Service.

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

Hamilton, D.T., P. Varma, S. Ou, and R. Cantor. 2004. Default & recovery rates of corporate bond issuers—a statistical review of Moody's ratings performance 1970–2001. New York: Moody's Investors Service.

Han, C., and Y. Jang. 2013. Effects of debt collection practices on loss given default. *Journal of Banking and Finance* 37: 21–31.

Heckman, J.J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153–161.

Hesse, F., C. Kuklick, and A. Pfingsten. 2012. Sanierung und Abwicklung in Sparkassen und Genossenschaftsbanken. *Zeitschrift für das gesamte Kreditwesen* 65: 746–750.

Hurt, L., and A. Felsovalyi. 1998. Measuring loss on Latin American defaulted bank loans: a 27-year study of 27 countries. *The Journal of Lending and Credit Risk Management* 80: 41–46.

Jiménez, G., and J. Saurina. 2002. Loan characteristics and credit risk. Working Paper, Bank of Spain.

Jokivuolle, E., and S. Peura. 2003. Incorporating collateral value uncertainty in loss given default estimates and loan-to-value ratios. *European Financial Management* 9: 299–314.

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

Kleiber, W., and J. Simon. 2007. *Verkehrswertermittlung von Grundstücken—Kommentar und Handbuch zur Ermittlung von Verkehrs-, Versicherungs- und Beleihungswerten unter Berücksichtigung von WertV und BelWertV*, 5th ed. Cologne: Bundesanzeiger Verlagsgesellschaft.

König, W., and D. Beimborn. 2008. Sourcing-Trends im KMU-Kreditgeschäft der deutschen Banken. In *Outsourcing in Banken*, ed. B. Kaib, 183–210. Wiesbaden: Gabler.

Leow, M., and C. Mues. 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.

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

Park, Y.W., and D.W. Bang. 2014. Loss given default of residential mortgages in a low LTV regime: role of foreclosure auction process and housing market cycles. *Journal of Banking and Finance* 39: 192–210.

Pfnür, A. 2011. *Modernes Immobilienmanagement*, 3rd ed. Heidelberg: Springer.

Pykhtin, M. 2003. Unexpected recovery risk. *Risk* 16: 74–78.

Qi, M., and X. Yang. 2009. Loss given default of high loan-to-value residential mortgages. *Journal of Banking and Finance* 33: 788–799.

Raftery, A. 1995. Bayesian model selection in social research. *Sociological Methodology* 25: 111–163.

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

Rösch, D., and H. Scheule. 2014. Forecasting probabilities of default and loss rates given default in the presence of selection. *Journal of Operational Research Society* 65: 393–407.

Schaaff, C. 2009. Verwertungsquoten von Grundpfandrechten. Wiesbaden: Gabler.

Shilling, J.D., J.D. Benjamin, and C.F. Sirmans. 1990. Estimating net realizable value for distressed real estate. *Journal of Real Estate Research* 5: 129–139.

Somers, M., and J. Whittaker. 2007. Quantile regression for modelling distributions of profit and loss. *European Journal of Operational Research* 183: 1477–1487.

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

Tobin, J. 1958. Estimation of relationships for limited dependent variables. *Econometrica* 26: 24–36.

Tong, E.N.C., C. Mues, and L.C. Thomas. 2012. Mixture cure models in credit scoring: if and when borrowers default. *European Journal of Operational Research* 218: 132–139.

Tong, E.N.C., C. Mues, and L.C. Thomas. 2013. A zero-adjusted gamma model for mortgage loan loss given default. *International Journal of Forecasting* 29: 548–562.

Wolter, M., and D. Rösch. 2014. Cure events in default prediction. *European Journal of Operational Research* 238: 846–857.

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