JEL Classification: C02, G13, G33
Keywords: credit losses, loss given default, recovery rates, workout LGD

The Time Dimension of the Links Between Loss Given Default and the Macroeconomy*

Tomáš KONEČNÝ - European Systemic Risk Board and Czech National Bank, (tomas.konecny @ecb.europa.eu) *corresponding author*

Jakub SEIDLER – ING Bank NV, Prague, Czech Republic, Czech National Bank at the time this paper was written

Aelita BELAYEVA - CERGE-EI, Charles University, Prague

Konstantin BELAYEV - CERGE-EI, Charles University, Prague

Abstract

Most studies focusing on the determinants of loss given default (LGD) have largely ignored possible lagged effects of the macroeconomy on LGD. We fill this gap by employing a wide set of macroeconomic covariates on a retail portfolio that represents 15% of the Czech consumer credit market over the period 2002–2012. We find an important time dimension to the links between LGD and the aggregate economy in the Czech Republic. The model that allows exclusively for contemporaneous effects includes a number of significant macroeconomic variables, some of which have non-intuitive signs. Nonetheless, a more general time structure of the LGD model makes current macroeconomic variables largely irrelevant and highlights the importance of delayed responses of LGD to the macroeconomic environment.

1. Introduction

The amount of research on loss given default (LGD) has increased substantially since the adoption of the Basel II Capital Accord and its successor Basel III. Part of this research interest has focused on the macroeconomic determinants[1] of LGD, as the different characteristics influencing the recovery process can be determined by the current stage of the economy. These efforts have been justified by regulators' need to gauge potential losses in banks' loan portfolios and to forecast the credit losses of the banking sector with respect to macroeconomic developments, as well as by banks' need to estimate their credit losses more accurately.

The present study focuses on the link between the LGD and macroeconomic factors using data on consumer credit (i.e., personal loans), credit cards and overdrafts in a sample representing approximately 15% of the Czech retail consumer credit market. The contribution of the paper is threefold. First, the study explicitly models potential delayed time effects of the macroeconomy on the final LGD value. Changes in macroeconomic conditions proxied by key macroeconomic indicators might have a postponed impact on various dimensions important for the recovery of defaulted loans. In the retail segment, clients' willingness to cooperate probably depends both on their

[1] In the following text we will use the term "macroeconomic" both for variables relating to the overall state of the economy (such as real GDP growth and inflation) and for aggregates in the retail credit segment (such as default rates and loan growth).

current situation (e.g., unemployment status or earnings) and on long-term aspects such as unemployment spells or declining savings. The success of the recovery process might likewise depend on prevailing conditions in collateral markets, including real estate prices, though this effect might be limited in the case of retail loans, which are rarely hedged by real estate. To the best of the authors' knowledge, the only study explicitly considering possible time effects of macroeconomic variables within the context of LGD modeling is Bellotti and Crook (2012). The authors use leads and lags of up to six months and conclude that using macroeconomic variables generates inferior forecasts against benchmarks both without the macroeconomy and with the macroeconomic variables at the time of default. The present study differs in its explicit focus on potential lagged time effects, considering in particular a notably wider set of variables and lags of up to eight quarters within a simple and coherent estimation framework.

The second contribution of the study is the selection of macroeconomic variables in our LGD models via elastic net regularization, which is an extension of the lasso regression (Tibshirani, 1996). The elastic net is a convenient tool for selecting from among many, possibly highly correlated, variables, as is often the case with macroeconomic time series. While the lasso tends to generate parsimonious models with only one variable from a set of highly correlated covariates, the elastic net provides for multicollinearity by choosing a group rather than a single representative candidate. The multicollinearity problem within the context of ratings migration has been addressed by Figlewski et al. (2012). In their study of the U.S. corporate bond market, the authors find that the direction and significance of the links between U.S. corporate defaults and ratings transitions on the one hand, and macroeconomic variables on the other hand, depend heavily on the selection of the remaining (highly correlated) macro covariates.

Finally, the study uses a unique comprehensive dataset on loan losses corresponding to approximately 15% of the retail consumer credit market in the Czech Republic and thus complements the rare evidence on the links between consumer LGD and the macroeconomy (e.g., Caselli et al. 2008; Bellotti and Crook 2012; and Calabrese 2012). Besides discussing the results for a whole sample covering all three portfolios, we analyze each portfolio in more detail to explore possible subsegment-specific macroeconomic effects. Our focus is on the "workout LGD" resulting from the collection of defaulted debt obligations and we leave out the market LGD observed from the market price shortly after the default event. We are aware that the scope of this study might be limited given its reliance on a limited sample of the retail consumer portfolio rather than on the whole retail credit market. However, we believe our study still offers valuable information due to the scarce and comparatively rich data source in the context of evaluating LGD determinants as well as the straightforward transferability of our methodology to other institutions.

Section 2 presents in more detail still scant evidence on the links between LGD and the macroeconomy. The following section describes the data on LGD, client-level information, and macroeconomic variables employed in the study. Section 4 outlines the overall methodology and the mechanics of the elastic net approach. The results are presented and discussed in Section 5. The final section concludes.

2. Empirical Literature

A few studies have explicitly addressed the corporate recovery rate (i.e., 1-LGD) and its relationship to the state of the economic cycle (Altman et al., 2005a, 2005b, Frye 2000, 2005) by employing publicly traded defaulted bonds.[2] Similarly, Acharya et al. (2007) used data on observed prices of defaulted securities in the United States over the period 1982–1999 and concluded that the recovery rates in distressed industries are, on average, lower than those in healthy industrial sectors. Nonetheless, the characteristics of defaulted corporate bonds differ significantly from workout LGDs on defaulted bank loans, as the post-default price is available only for the fraction of the debt that is traded and for which an after-default market exists, i.e., very often for corporate bonds issued by large companies only.

Focusing on bank loans instead of corporate bonds, the empirical results by Dermine and Neto de Carvalho (2006) examine the timing of recoveries of bad and doubtful bank loans and the distribution of cumulative recovery rates. The authors estimate models on a European bank's portfolio of loans to small and medium-sized enterprises and do not find any support for a significant role of macroeconomic variables, including GDP growth, frequency of default in the industry sector, and the interest rate. This fact was explained by the absence of a sufficient recession during the period under consideration.

Much less attention has been devoted to retail portfolios in the current literature. Bellotti and Crook (2012) found a significant relationship between LGD and retail interest rates, the level of unemployment, and earnings growth in their analysis of credit card LGDs at the transaction level. Employing aggregated data, Caselli et al. (2008) examined the sensitivity of LGD to systematic risk and found a relationship between LGD and the development of the macroeconomy. The authors employed separate multivariate models for two customer segments—SMEs and households. For SMEs the best model incorporated the aggregate number of employees and the GDP growth rate; for households the relevant macroeconomic variables included the default rate (approximated by the change in defaulted loans to total loans), the unemployment rate, and household consumption. The authors furthermore demonstrated a positive relationship between LGD and recovery collection length, but did not model the above-mentioned link in their multivariate models. Another study using data on individual loans (including the consumer segment) in Italy is Calabrese (2012). The author estimates a mixed continuous-discrete model of recovery rates and finds systematic links between recovery rates and the macroeconomy proxied by the interest rate, GDP growth, unemployment, and the aggregate default rate. Finally, a top-down stress test on the Finnish economy by Jokivuolle and Virén (2011) included a system of equations on the economy-wide probability of default (PD), LGD, and macroeconomic variables. It estimated a significantly negative relationship between the aggregate LGD and the gross profit rate.

[2] A number of studies on workout LGD focus exclusively on firm-specific factors or details of the recovery process and avoid the potential influence of the macroeconomic environment (e.g., Bastos, 2010; Calabrese and Zenga, 2010; Emery et al., 2004; Grippa et al., 2005; Grunert and Weber, 2009). Witzany et al. (2010) employed survival analysis methods for LGD modeling, leaving macroeconomic factors unexplored.

3. Data

The retail loan and consumer credit market in the Czech economy was significantly underdeveloped at the beginning of the transformation process to market economy compared with western EU countries. However, financial deepening has caused a rapid credit expansion in most European emerging economies, leading to considerable accumulation of both retail and corporate credit in the last twenty years. As such, the ratio of consumer credit to GDP in the Czech Republic has converged to a similar level as in Western European countries (see Table A1 for averages for selected EU countries).

For this study we employ a dataset consisting of over 18,500 defaulted accounts over the period January 2003 to June 2010 provided by a large Czech bank. Using regulatory reporting data, this sample represents more than 15% of the Czech retail consumer credit market and consists of consumer loans (credit provided to an individual—or family—on an unsecured basis, also known as personal loans), overdrafts (a type of revolving loan where a client—a natural person—gets additional credit through a current account up to a limit agreed in advance), and credit cards. It does not include loans for housing purchases or mortgages. The bank switched to the IRB foundation approach at the beginning of 2007, and data for creating the LGD model for this transition were collected from 2002 onwards (five years for the retail portfolio, in accordance with the Basel requirement). Consumer loans account more than half of the sample, overdrafts for roughly 30%, and the credit card sub-portfolio for the rest.

3.1 Definition of LGD

The Basel Accord defines a default event as a realization of one or more of the following circumstances: the credit obligor i) is unlikely to pay, ii) is more than 90 days past due, or iii) is declared bankrupt. For retail clients only, the second and/or third condition has to be met. The default rate is defined as the ratio of the number of clients that defaulted during a given time period to the number of all observed clients at the beginning of the period concerned. As we focus only on standard bank loans excluding marketable instruments, our measure of LGD is derived as the ratio of losses to exposure at default (EAD). In order to measure LGD in this way, recovery cash flows from defaulted loans, as well as the costs of the bank's workout process, must be observed. The loss experienced by a bank is understood to mean the economic loss, i.e., the loss adjusted for discount effects, funding costs, and direct and indirect costs associated with collection of the instrument (BCBS, 2006). As the final amount collected from a defaulted loan can exceed the EAD, the real range of observed LGDs may vary from negative numbers to positive numbers higher than one. We truncate the LGD at zero and one in order to make it comparable with the market LGD (obtained after the sale of defaulted market instruments).[3] In order to unify the calculation

[3] The choice to truncate LGD by zero and one is driven by a presumption that LGD behaviour out of range [0,1] cannot be explained just by consumer's characteristics (covariates) used in the regression procedure, but would require incorporating a bank's collection policy in terms of applied interests, fees and costs. Due to commercial confidentiality of data, the exact figures of due interests, fees and costs cannot be provided. Nonetheless, while the issue of truncation might be generally relevant, it is likely to matter more for the

procedure for either closed loans or still open loans, we have chosen time horizon of one year for observed LGD to be recorded. Belotti and Crook (2012) are also have used period of one year in the collection process for LGD calculation.

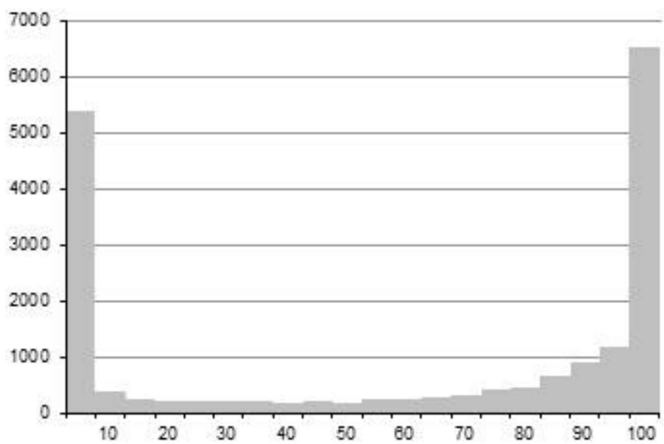
Table 1 LGD Summary Statistics for the Whole Sample and Sub-portfolios, 2003q1-2010q2

Variable	*Obs*	*% of total*	*Mean*	*Median*	*Std. Dev.*
Total	18 698	100	0.58	0.79	0.40
Consumer loans	10 287	55.0	0.59	0.78	0.42
Overdrafts	5 627	30.1	0.46	0.46	0.30
Credit cards	2 784	14.9	0.77	0.90	0.30

Table 1 presents summary statistics of the LGD values for the whole sample and individual sub-portfolios. The whole sample average LGD of 0.58 is very close to the value of 0.54 reported for the SME and household portfolio by Caselli et al. (2008) and to the value of 0.62 reported for the Italian resident portfolio by Calabrese (2012). The value is nonetheless larger than the LGDs reported for corporates (e.g., Acharya et al., 2007). The highest average LGD is reported for the credit card portfolio (0.77) and the lowest is reported for overdrafts (0.46).

Figure 1 presents the histogram of the LGD for the retail lending segment.

Figure 1 Distribution of LGD in the Sample, Counts on the y-axis



Notes: axis x: LGD (%); axis y: number of observations.

3.2 Transaction-level Data

We used historical observations from transactional systems on a monthly basis. Although our original data covers the period from January 2003 through June 2012, in order to avoid possible estimation bias due to defaults with unfinished workout

---

evaluation of LGD determinants at the client and bank level, rather than at the macro-level, which is the main focus of the present study.

processes, we exclude cases after June 2010. Apart from the information on LGDs, the data contain the standard socio-demographic characteristics of retail clients usually collected by banks when a credit account is opened. These include the client's age, gender, number of children, education, family, employment status, and phone ownership. All retail loans were provided in Czech crowns (CZK).

The retail portfolio offers information on the duration of the client's relationship with the bank in years and the exposure at default. Both variables are transformed logarithms. These client-related variables were used in prior studies, for example in Bellotti and Crook (2012).

Table 2 contains the full list of transaction-level (henceforth microeconomic) variables employed.

Table 2 Summary Statistics of Micro Variables Employed, Total Sample 2003q1-2010q2

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
LGD	18 698	0.6	0.4	0	1
Relationship with bank*	18 698	2.9	0.7	1.1	4.6
Exposure at default*	18 698	9.9	1	4.7	13.6
Children	18 698	0.3	0.7	0	13
Education	18 698	0.5	0.6	0	2
Employment	18 698	1.6	0.3	0	2
Age	18 698	37.9	11.7	18.4	71.2
Female	18 698	0.3	0.5	0	1
Phone	18 698	0.5	0.5	0	1
Family status	18 698	0.6	0.3	0	1

Notes: The variables have been transformed into logarithm for the reasons of confidentiality as required by the data owner. Education =0 if Primary, =1 if Secondary, =2 if Graduate; Employment =0 if Unemployed, =1 if Student, pensioner, or house wife, =2 if employee or entrepreneur; Family status =1 if married or registered partnership.

3.3 Macroeconomic Data

The macroeconomic data we use reflect the choices made in the existing literature on LGD modeling. These data will serve as the starting point for our estimations, and variable selection techniques will be applied for the ultimate model specifications.

Table 3 Studies on LGD Employing Macro Variables

Author(s) and year of publication	*Sample coverage*	*Macro variables employed by the study*	*Did the study use lags or leads of macro variables?*
Acharya et al. (2007)	Corporate bonds	Default rate, volume of defaulted bonds, GDP growth, S&P stock index	
Altman et al. (2005a)	Corporate bonds	Default rate, Δdefault rate, total volume of (high yield) bonds, volume of defaulted bonds, GDP growth (y-o-y), ΔGDP growth (y-o-y), GDP indicator, S&P 500 index, ΔS&P 500 index, U.S. Treasury rate	no
Altman et al. (2005b)	Corporate bonds	GDP growth (y-o-y), ΔGDP growth (y-o-y)	no
Belotti and Crook (2012)	Credit cards	Retail interest rates, unemployment level, earnings growth	yes 0-6months leads and lags

Bruche and Gonzalez-Aguado (2008)		GDP growth (y-o-y), unemployment, investment growth (y-o-y), S&P 500 return	no
Calabrese (2012)	Corporate loans	Interest rate on delayed payment, GDP growth (y-o-y), unemployment, default rate	no
Caselli et al. (2008)	Households and SMEs	ΔDefault-to-loan ratio (ΔD/L, y-o-y), volume of bank loans, GDP growth (y-o-y), employment, Δunemployment, household consumption, total gross investment, total production, gross annual available income	no
Dermine and Neto de Carvalho (2006)	SMEs	GDP growth (y-o-y), industry default rates	no
Jokivuolle and Viren (2011)	Corporate bank loans	Output gap, gross profit rate, interest rate, indebtness, stock market index and a housing price index, real house prices, real stock index, real interest rate	no

Table 3 lists the empirical studies on LGD that have considered macroeconomic effects, their portfolio of interest, a full list of the macroeconomic variables employed, and an indication of whether they allowed for time effects. We focus in particular on those candidate variables which are potentially relevant to the retail segment and either take their exact counterparts or construct the closest possible approximation. Given the boundedness of our loss-given-default measure, we furthermore transform level variables with a clear growing trend into yearly growth rates.

The indicators of retail credit risk we use include the amount of retail loans defaulting, the 3-month retail default rate, and the first difference of the retail default rate (e.g., Calabrese, 2012; Caselli et al., 2008; Dermine and Neto de Carvalho, 2006). In addition, the ratio of the amount of retail loans classified as non-performing to the total retail credit portfolio (the NPL ratio) proxies for materialized credit risk in banks' balance sheets. The retail credit growth variable aims to approximate demand for retail credit (Caselli et al., 2008). Labor market developments are captured by the unemployment rate, change in the unemployment rate, and total employment (Caselli et al., 2008). Instead of the earnings index, as employed by Bellotti and Crook (2012),[4] we use real wage growth to approximate the evolution of household earnings. The indicators referring to aggregate supply and demand are real GDP growth and change in real GDP growth (Altman et al., 2005a; Altman and Brady, 2002), growth in real industrial production, real consumption growth, and real investment growth (Caselli et al., 2008). Apart from the real 3-month Pribor as a measure of the monetary policy rate, we use average retail credit rates and the retail credit spread, defined as the difference between the average rates on retail loans and deposits (e.g., Bellotti and Crook, 2012). The inflation in the economy is approximated by CPI index growth, while real estate price index growth reflects the situation on the real estate market. Growth in the PX-50 stock market index tracks developments in stock markets (e.g., Jokivuolle and Virén, 2011).

[4] The authors used an economy-wide earnings index, including bonuses as a ratio of the retail price index.

All growth variables are expressed in year-on-year terms, while level variables are logged.[5] We generate leads and lags of the macroeconomic variables of up to eight periods. Some aggregates relating to the retail loan segment are available only from the beginning of 2002 or later. These include the amount of retail loans and defaults, retail credit rates and spreads, and the default rate and the NPL ratio. As a consequence, we had to shift the initial sample date in order to obtain the desired number of lags. The most restricted sample, allowing for eight lags of all the macroeconomic variables considered, was limited to defaults that occurred after 2005q1.[6] Instead of interpolating the quarterly macroeconomic covariates to monthly frequency, we mapped the same quarterly values to all defaults in the relevant quarter. The sources of macroeconomic data are the Czech National Bank (CNB) and the Czech Statistical Office (CZSO). A statistical summary of the macroeconomic variables in the current study is presented in Table 4. These will be accounted for within the elastic net framework, which is described in more detail in the following section.[7]

Table 4 Summary Statistics of Macroeconomic Variables, Whole Sample 2003q1-2010q2

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Defaulted loans growth (y-o-y)	18 698	0.3	0.2	-0.1	0.6
Default rate (3 months)	18 698	1.7	0.6	0.6	3.2
ΔDefault rate (3 months)	18 698	0.0	0.2	-0.3	0.3
NPL ratio	18 698	8.0	0.8	6.4	9.4
Retail loans growth (y-o-y)	18 698	20.3	7.0	9.3	31.4
Unemployment rate	18 698	7.0	1.2	4.3	8.5
ΔUnemployment rate	18 698	0.2	0.5	-0.7	1.3
Total employment (in thds)*	18 698	3875.9	88.0	3780.1	4058.2
Real wage growth (y-o-y)	18 698	3.4	1.2	1.5	6.1
Real GDP growth (y-o-y)	18 698	1.3	4.9	-5.5	7.6
ΔReal GDP growth (y-o-y)	18 698	-0.1	1.6	-3.8	2.3
Real ind. production growth (y-o-y)	18 698	0.5	2.7	-6.5	6.0
Real consumption growth (y-o-y)	18 698	2.2	2.1	-1.1	5.9
Real investment growth (y-o-y)	18 698	-4.8	13.9	-23.0	25.1
Real Pribor3m	18 698	0.9	0.8	-1.2	2.9
Rate on retail loans	18 698	14.0	0.5	13.0	15.0
Retail rate spread	18 698	12.6	0.6	11.5	13.7
CPI(y-o-y)	18 698	2.0	1.9	-0.3	7.4
Property price index (y-o-y)	18 698	1.0	0.2	0.8	1.4
Stock market index (PX-50) (y-o-y)	18 698	19.5	33.6	-52.7	67.2

Notes: *The variable is expressed in absolute value.

4. Methodology

We estimate the generalized-linear model (GLM) regressions with logit transformation of the LGD parameter (Hastie and Tibshirani, 1990). The GLM framework is a standard benchmark that has been used in a number of existing studies

[5] Interest rates are not considered as level variables.
[6] The estimation results presented in the following sections use the sample starting in 2003q1, i.e., the sample only includes contemporaneous values of the yearly growth rates of total and defaulted retail loans, retail rates and spreads, and default rates and the NPL ratio. Our results from the most restricted sample (2005q1–2010q2) are qualitatively very similar to the reported output. Similarly, earlier cutoff dates before 2003q1, which effectively dropped some of the above-mentioned variables, left our results practically unchanged. These results are available upon request.
[7] Table A2 in the Appendix presents the correlation table of the macroeconomic variables employed. The shaded areas highlight correlation coefficients in excess of 0.6, indicating a high degree of collinearity.

on LGD determinants and forecasting (e.g., Altman et al., 2005b; Caselli et al., 2008; Grunert and Weber, 2009). In fact, other parametric approaches, such as inverse-Gaussian, Tobit, and beta transformation, do not seem to outperform the standard approaches (Qui and Zhao, 2011; Bellotti and Crook, 2012).

We run our estimations on the dataset described in Section 2, which aimed to match or at least approximate all the macro variables used in existing studies using macroeconomic variables (see Table 3). For the results presented in later sections we select variables (and their respective time effects where applicable) using elastic net regularization (Friedman et al., 2010). Elastic net regularization is an extension of the lasso regression developed by Tibshirani (1996). The lasso regression reduces the variability of the estimates through coefficient shrinkage. Given that some coefficients are set to zero, the procedure generates a sparser model and can serve as a variable selection tool at the cost of only moderate bias.[8] While the advantage of the lasso regression over other selection procedures, such as stepwise regression, is its lack of path dependency (some of the coefficients set to zero might change again over the iteration process), the lasso applies only to cases of more observations than variables. Furthermore, in the case of multicollinearity typical of macroeconomic time series data, the procedure tends to select only one variable in the group. Elastic net regularization aims to overcome the above-mentioned limitations. The elastic net derives from the lasso regression, yet it allows for cases of more variables than observations and selects groups of highly correlated variables instead of a single group representative as is typical of the lasso (Zhou and Hastie, 2005). Empirical studies suggest that the elastic net tends to outperform the lasso in settings with highly correlated groups of variables (e.g., Zhou and Hastie, 2005; Friedman et al., 2010). The elastic net regularization problem for generalized linear models can be defined as

$$\min_{\beta} |\boldsymbol{y} - \boldsymbol{X}\beta|^2 + \lambda P_\alpha(\beta), \tag{1}$$

where $\boldsymbol{y} = (y_1, \dots, y_n)^T$ is a vector of responses from the data sample of N, $\boldsymbol{X}$ is a matrix of candidate explanatory variables and β are the estimated p regression parameters. The penalty term $P_\alpha(\beta)$ corresponds to $P_\alpha(\beta) = \frac{(1-\alpha)}{2}\|\beta\|_2^2 + \alpha\|\beta\|_1 = \sum_{j=1}^{p}\left(\frac{(1-\alpha)}{2}\beta_j^2 + \alpha|\beta_j|\right)$ for $0 < \alpha < 1$ and $\lambda \geq 0$.

The elastic net converges to the lasso as $\alpha \to 1$ and to the ridge regression as $\alpha \to 0$. Solution of the problem for given values of lambda and alpha are obtained via maximizing the λ-penalized log likelihood of (1).

We obtain our final model specification by the following procedure. First, we employ the default ten-fold cross validation and calculate the cross-validated mean squared error (CVMSE) for all values of lambda and alpha.[9] Then we locate a value of lambda over all alphas with the minimum CVMSE. Finally, to obtain a more parsimonious specification, we select a model corresponding to a lambda value within one standard error of the minimum CVMSE.[10]

[8] The ridge regression—an alternative shrinkage method—shrinks yet keeps all the coefficients in the model.
[9] We considered a sequence of alpha values ranging from 0 to 1 with increments of 0.1.
[10] More details on the variable selection algorithm for the glmnet package in R can be found at http://cran.r-project.org/web/packages/glmnet/glmnet.pdf.

Our benchmark specification contains exclusively client-level information. In the following models, we first estimate the specification using client-level information and contemporaneous macroeconomic variables, that is, we do not allow for time effects. Finally, we present a full model including client-level and macroeconomic data with time effects (lagged and lead values).

5. Results

5.1 Whole Portfolio

We begin the analysis with an examination of the marginal contributions relating to each of the macroeconomic variables with respect to the loss given default (LGD). Table 5 presents the condensed output from the OLS regressions of the logit LGD on a particular contemporaneous macro variable (i.e., no time effects, 'no lags' column), and on its lags from t to t-8. Values listed in the second column represent the sum of beta coefficients and thus provide initial information on the long-term correlation between a macro variable of concern and the logit LGD.[11]

Table 5 Marginal Contributions of Macroeconomic Variables, Whole Sample 2003q1-2010q2

	(1) no lags			*(2) lags t to (t-8)*		
	Coefficient		*SE*	*Coefficient*		*SE*
Defaulted loans growth (y-o-y)	0.25		0.17	-1.10		1.10
Default rate (3 months)	0.05	***	0.02	0.16	**	0.06
ΔDefault rate (3 months)	0.10		0.18	0.00	**	0.31
NPL ratio	-0.02		0.03	-0.12	*	0.07
Retail loans growth (y-o-y)	-0.53		0.55	-0.60		1.10
Unemployment rate	0.04	***	0.01	0.06	***	0.02
ΔUnemployment rate	0.16	***	0.06	0.30	**	0.11
Total employment (in thds)	-3.11	***	1.04	2.36		1.60
Real wage growth (y-o-y)	-0.03	*	0.02	-0.01	***	0.05
Real GDP growth (y-o-y)	-0.02	***	0.00	-0.05	***	0.02
ΔReal GDP growth (y-o-y)	0.03	*	0.02	-0.07	*	0.04
Real ind. production growth (y-o-y)	0.06		0.72	0.02		0.13
Real consumption growth (y-o-y)	-0.03	*	0.02	-0.02		0.05
Real investment growth (y-o-y)	-0.01	***	0.00	-0.01	*	-0.01
Real Pribor3m	0.07	***	0.02	0.05		0.10
Rate on retail loans	0.04		0.04	0.03		0.15
Retail rate spread	0.04		0.04	0.02		0.12
CPI(y-o-y)	-0.03	***	0.01	0.02		0.02
Property price index (y-o-y)	-0.51	***	0.14	0.38		0.47
Stock market index (PX-50) (y-o-y)	0.18	*	0.1	-0.05	**	0.02

Notes: *, **, *** stand for 10 %, 5 %, and 1 % significance levels respectively. The signs and standard errors in the specification with lags relate to the sum of coefficients on lagged values.

The preliminary evidence is consistent with the expected significant link between LGD and the macroeconomic environment. Most variables are significant and have the expected signs. The only counterintuitive and statistically significant signs

[11] Furthermore, a preliminary analysis from simple regressions of average LGD on individual lagged values of macroeconomic variables (see Figure A1 in the Appendix) reveals a temporal pattern of the correlations between periods t and t-8. The temporal patterns, as well as beta coefficients in the OLS regressions in Table 5 indicate smooth correction-style dynamics between the LGD and macro variables. Individual beta coefficients from the OLS regressions in Column (2) can be provided upon request.

are the coefficients for the yearly absolute change in the GDP growth rate and the growth rate of the stock market index in the specification with contemporaneous values (Column 1). Other cases with counterintuitive signs, such as the positive correlation of LGD with growth of industrial production or with growth in real estate prices in Column (2), are imprecise with large standard errors and hence not statistically significant.

The marginal contributions for the individual sub-portfolios in Table 6 provide a few more detailed insights. First of all, the correlations between LGD and macroeconomic covariates tend to be weaker for overdrafts and credit cards than for consumer loans, where the coefficients on most macro variables are statistically significant. Secondly, while there are fewer statistically significant coefficients in the overdraft portfolio as compared to consumer loans, in both cases the coefficients remain intuitive. This is less so for the credit card portfolio—for example, the positive correlation between LGD and total employment, and the negative correlation between the LGD and the unemployment rate as well as the absolute change in the unemployment rate in Column (6) run strongly against ex ante expectations. Finally, the credit card portfolio provides rather ambiguous information on the direction of the links between the macroeconomic variables and LGD, as in most cases the coefficient sign varies depending on whether or not time effects are allowed for.

Table 6 Marginal Contributions of Macroeconomic Variables by Subportfolios, 2003q1-2010q2

	Consumer loans				*Overdrafts*				*Credit cards*			
	(1) *no lags*		*(2)* *lags t to (t-8)*		*(3)* *no lags*		*(4)* *lags t to (t-8)*		*(5)* *no lags*		*(6)* *lags t to (t-8)*	
	Coefficient	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**
Defaulted loans growth (y-o-y)	0.33	0.27	-0.21	0.19	0.05	0,19	-0.10	0.08	0.54	0.40	0.81	0.96
Default rate (3 months)	0.09 ***	0.00	0.11 **	0.04	0.00	0,94	0.02	0.04	0.04	0.04	-0.13 **	0.05
ΔDefault rate (3 months)												
NPL ratio	0.06 ***	0.01	0.02	0.05	0.03	0.14	0.01	0.04	0.01	0.56	-0.09 **	0.03
Retail loans growth (y-o-y)	-1.29 ***	0.00	-3.64 **	1.23	-0.70 ***	0.14	-1.93 *	1.04	-0.53 ***	0.05	-1.23	0.85
Unemployment rate	0.06 ***	0.00	-0.01	0.02	0.02	0.15	-0.05	0.29	0.00	0.95	-0.08 *	0.04
ΔUnemployment rate	0.20 ***	0.01	0.46 ***	0.07	0.04	0.46	0.17	0.10	0.07 ***	0.01	-0.24 **	0.098
Total employment (in thds)	-5.31 ***	0.00	-4.58 ***	1.13	-1.74 ***	0.08	-0.66	1.10	-0.95 ***	0.13	4.69 **	1.37
Real wage growth (y-o-y)	-0.04	0.14	-0.13 ***	0.03	-0.02	0.20	-0.11 **	0.04	-0.01	0.05	0.09	0.09
Real GDP growth (y-o-y)	-0.02 ***	0.00	-0.09 ***	0.02	-0.01	0.11	-0.04 *	0.02	-0.01 **	0.00	0.02	0.02
ΔReal GDP growth (y-o-y)	0.03	0.16	-0.05	0.07	0.02	0.14	-0.03	0.58	-0.01	0.05	-0.11 **	0.03
Real ind. production growth (y-o-y)	0.02	0.03	-0.21	0.17	0.00	0.04	0.01	0.30	0.00	0.12	0.06	0.07
Real consumption growth (y-o-y)	-0.04 **	0.02	-0.18 ***	0.02	-0.04 ***	0.01	-0.12 ***	0.03	-0.01	0.03	-0.01	0.03
Real investment growth (y-o-y)	-0.01 ***	0.00	-0.02 ***	0.00	0.00	0.08	-0.01 ***	0.00	0.00	0.01	0.01	0.01
Real Pribor3m	0.06 ***	0.00	-0.07	0.11	0.02	0.28	-0.04	0.10	0.03	0.06	-0.02	0.28
Rate on retail loans	0.11 ***	0.00	0.37 **	0.12	0.05	0.06	0.20 *	0.10	0.05	0.06	-0.04	0.05
Retail rate spread	0.10 ***	0.00	0.30 ***	0.07	0.05	0.07	0.15 **	0.06	0.03	0.15	-0.07	0.04
CPI(y-o-y)	-0.05 ***	0.00	0.00	0.02	-0.02	0.13	0.01	0.02	0.00	0.80	0.05	0.03
Property price index (y-o-y)	-0.68 ***	0.00	-0.55 *	0.28	-0.40 ***	0.00	0.01	0.44	-0.08	0.12	0.99	0.01
Stock market index (PX-50) (y-o-y)	0.19	0.16	-0.07 *	0.04	0.13	0.30	0.35	0.34	-0.13	0.65	0.01	0.02

Notes: *, **, *** stand for 10 %, 5 %, and 1 % significance levels respectively. The signs and standard errors in the specification with lags relate to the sum of coefficients on lagged values.

Table 7 GLM Estimates with Logit LGD, Whole Sample 2003q1-2010q2

	Client specific		***All variables, no lags***		***All variables, lags included***	
Explanatory variable logit LDG	*(1)*		*(2)*		*(3)*	
	Coefficient	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**
Client- specific factors						
Exposure at default	1.448***	0.059	1.036***	0.058	1.044***	0.057
Relationship with bank	-0.033***	0.004	-0.082***	0.004	-0.080***	0.004
Age	0.397**	0.192	0.815***	0.183	0.774***	0.181
Children	-1.046***	0.079	-0.059	0.087		
Phone	-0.498***	0.117	0.388***	0.113		
Employment	0.206***	0.059				
Education	-0.893***	0.097	-0.631***	0.093	-0.607***	0.093
Female	0.362***	0.123	-0.000	0.116		
Macroeconomic variables, current values						
Real GDP growth (y-o-y)			0.023	0.094		
∆Real GDP growth (y-o-y)			0.284***	0.080		
Real Consumption Growth (y-o-y)			-0.398***	0.090	-0.232***	0.070
Real Investment Growth (y-o-y)			-0.056***	0.019		
Real Pribor3m			0.395**	0.165		
Inflation rate (y-o-y)			-0.229***	0.064		
Property prices (y-o-y)			-1.690**	0.745		
Default rate			-0.774***	0.126		
Retail loan growth (y-o-y)			-0.093***	0.027	-0.054**	0.021
Macroeconomic variables, lagged and lead values						
Real GDP growth (y-o-y) (t-1)					-0.123*	0.063
Real GDP growth (y-o-y) (t-2)					-0.143**	0.063
Real investment growth (y-o-y) (t-2)					-0.035***	0.013
Unemployment rate (t-8)					0.315	0.225
Real wage growth (y-o-y) (t-3)					-0.097	0.109
Real wage growth (y-o-y) (t-4)					-0.444***	0.106
Real wage growth (y-o-y) (t-5)					-0.161*	0.088
Constant	-13.778***	0.909	-5.001***	1.235	-7.704***	1.563
Alpha	0.9		0.9		0.3	
Observations	18698		18698		18698	
Adjusted R-squared	0.056		0.152		0.152	
AIC	130270.953		128276.671		128264.055	

Notes: *, **, *** stand for 10 %, 5 %, and 1 % significance levels respectively. Robust standard errors reported. The values of Alpha stand for the elastic net weighting parameter with lowest residual mean-squared error.

The output from whole sample regressions of logit LGD on client-specific and macroeconomic variables in Table 7 provides a somewhat more complex picture.[12] Column (1) presents the benchmark specification. The results highlight intuitive attributes that should, on average, increase the chances of a successful workout

[12] We use variables identified by the elastic net regularization procedure and estimate the GLM with a logit link function for the LGD variable (see Section 4). For interested readers, estimates for the combined portfolio of overdrafts and consumer loans (excluding credit cards) can be found in the Appendix in Table A3.

process. The factors contributing to lower LGD include a longer relationship with the bank and smaller exposure at default. The positive relationship between LGD and exposure at default is consistent with Calabrese (2012), Dermine and Neto de Carvalho (2006), and Grippa et al. (2005). On the other hand, Asarnow and Edwards (1995) do not find a significant link between LGD and exposure at default. The output in Column (1) furthermore suggests that younger, better educated male clients with children and phone ownership represent the most promising outcome of the workout process. The only counterintuitive result is the contribution of employment or entrepreneurship status to higher LGD, i.e., a lower recovery rate.

Column (2) represents an extended model with contemporaneous macroeconomic variables. The coefficient sign and size of the client's most robust characteristics, i.e., exposure at default, length of relationship with the bank, and education, do not depart substantially from the benchmark specification in Column (1). Importantly, once the macroeconomic dimension is included, the remaining client-level information that was significant in the benchmark specification does not pass the elastic net shrinkage procedure.

These less-stable client factors include gender, number of children, phone ownership, and employment status with a counterintuitive positive sign in the benchmark regression.

The elastic net estimates in Column (2) put forward several macroeconomic factors consistent with previous findings. The negative association of LGD with real consumption growth is in line with the findings on households and SMEs by Caselli et al. (2008). A similar result holds for real investment growth. A negative and significant relationship between investment and LGD was reported for the retail segment by Caselli et al. (2008). Bruche and Gonzalez-Aguado (2008), on the other hand, do not find a significant link. Most of the other coefficient estimates likewise provide an intuitive interpretation. This concerns the positive sign on the interest rate and real estate price growth (e.g., Jokivuolle and Virén, 2011) and the inflation rate. Unlike Caselli et al. (2008), who used the total volume of retail loans, we find that the negative sign of the coefficient on retail loan growth is highly significant.[13]

Inspection of the estimates of macro variables in Column (2) nonetheless indicates that some coefficients have counterintuitive yet significant signs. This concerns the positive sign (significant at a 1% level) of the change in the real GDP growth, and the negative sign of the default rate.

Change in real GDP was used in the study of corporate bonds by Altman et al. (2005a). The authors found that while real GDP growth in their multivariate specification of the recovery rate model was insignificant and had the incorrect sign, change in real GDP growth improved the model performance markedly.[14] In the present specification, the estimates do not conform to the results of Altman et al. (2005a). The signs on change in real GDP growth in Column (2) are in line with the variable's marginal contribution in the contemporaneous specification in Table 5 (Column 1). The more reasonable sign estimates of the marginal contribution using

[13] Altman et al. (2005a) likewise find a significantly negative sign for corporate bonds.

[14] Similarly, Calabrese (2012) reports a positive link between the LGD and real GDP growth. However, Caselli et al. (2008) and Acharya et al. (2007) found a negative link, though in the latter case only partly significant.

lagged effects (Column 2) nonetheless suggest that the model specification not allowing for time effects might be too restrictive.

The significant positive relationship between the aggregate default rate and LGD, as suggested by existing studies, is not supported by the counterintuitive results in Column (2) (e.g., Altman et al., 2005b; Frye, 2005; Düllmann and Trapp, 2005, for corporate bonds; Altman et al., 2005c, for corporate loans; Caselli et al., 2008, for SMEs and households). While the unconditional contemporaneous correlation between the default rate and LGD is positive (see Table 5), the conditional correlation becomes negative once other variables have been controlled for. The above-mentioned result is partly consistent with the ambiguous results by Caselli et al. (2008), who report an insignificant link for a pooled portfolio of SMEs and households, as well as by Calabrese (2012), who finds a negative link between the default rate and extreme values of the LGD.

The adjusted R-squared of 0.06 for the benchmark client-specific model in Column (1) improves notably to approximately 0.15 for specifications including macroeconomic determinants. This finding is consistent with Altman (2005a), who reports a significant increase in the explained variation once change in GDP growth is included in the estimations. The same study, however, did not observe an increase in R-squared when real GDP growth was included instead of change in real GDP growth. The results from specifications with contemporaneous macroeconomic variables thus partly conform to the extant studies, but should be evaluated carefully given the counterintuitive signs on key macro correlates such as change in real GDP growth or the default rate.

The results listed in Column (3) point to the importance of the time dimension for the evaluation of links between LGD and the macroeconomy. In particular, while only two contemporaneous macroeconomic indicators from Column (2) passed the elastic net regularization procedure (real consumption growth and retail loan growth), all remaining variables enter the models with lags. Furthermore, the overall effect of each variable preserves an intuitive sign and is significant in at least one specification. In Column (3), LGD is negatively linked to real GDP in the last two quarters and real investment growth in period t-2. In the case of real GDP growth, the negative coefficient sign conforms to the results reported by Bruche and Aguado (2010) and Calabrese (2012). For both real GDP and real investment growth, the coefficient furthermore exceeds the marginal contributions from Table 5.

A number of existing studies have found a significant contribution of contemporaneous or recent unemployment dynamics for LGD modeling (e.g., Acharya et al., 2007; Bellotti and Crook, 2012; Bruche and Aguado, 2008; and Caselli et al., 2008).[15] The model with time effects in Column (3) points to unemployment links over a longer time horizon of eight quarters, but contrary to the abovementioned studies statistically insignificant. The final variable that passed the elastic net procedure in the specification with time effects in Column (3) is real wage growth. Unlike the estimates for real GDP growth and real investment growth, where the difference from the contemporaneous model was not substantial in terms of the size of the lag, real wage growth relates to LGD with a lag of about one year. Furthermore, wage effects did not

[15] Interestingly, some studies report rather counterintuitive results for the unemployment rate (Calabrese, 2012).

appear in the specification without time effects, which therefore provided imprecise estimates of macroeconomic links to LGD. The general evidence for a negative link between wages and LGD is consistent with the findings of Bellotti and Crook (2012), who found a significant link between the recovery rate on credit cards and earnings growth (though only at the 10% level).

As a part of our exercise we allowed for both lead and lagged interlinkages between the macroeconomy and LGD. While all the coefficient estimates kept their signs, and remained of roughly similar size, in some cases they became smaller or even non-significant. Most importantly, the elastic net procedure did not select any new factor that is consistently and significantly related to our LGD measure, in line with the study by Bellotti and Crook (2012).[16]

One should note that the specification allowing for time effects yielded a sparser model in terms of client-level information. The only dimensions that remained robust were the length of the relationship with the bank, the size of exposure at default, and education. This relates to the explained variation of the flexible specification with delayed responses of LGD. Given that the value added in terms of the adjusted R-squared is virtually zero (a stagnant 0.152 in Columns 2 and 3), allowing for lagged responses of LGD should provide a better understanding of the interactions with the macroeconomy over time, rather than a major improvement in the forecasting of LGD performance.

The results presented in Columns (2) and (3) of Table 7 represent a key deliverable of our study. We distinguish between models with and without time effects, employ variable selection techniques, and see how they differ in terms of the final specifications and performance. We show that the model in Column (2), which selected exclusively from macroeconomic variables with no time effects, and the model in Column (3), which considered both contemporaneous values and time effects, tend to differ to a substantial degree. This outcome points strongly in favor of the inclusion of time effects in LGD estimations. In the following section we check how this pattern is robust at the level of the individual portfolios.

5.2 Consumer Loans and Overdrafts

The logit estimates for the individual portfolios provide more detailed information on the drivers of the results from the whole sample. Table 8 provides estimates for the consumer loan and overdraft portfolios, while credit cards have been relegated to the Appendix (Table A4). The main reason in the case of credit cards was convergence issues arising in the estimations using the elastic net algorithm, as the large variance of the minimum cross-validated mean squared error (CVMSE) implied that only a constant should be selected for the lambda value within one standard error of the minimum CVMSE. As a result, we treat the results for credit cards with some caution and report the output for a lambda value corresponding to the minimum CVMSE. It should nonetheless be noted that the results for credit cards using lambda for the minimum RMSE are consistent with the output for consumer loans and overdrafts. Specifically, the overall pattern is qualitatively identical for the whole sample and the individual sub-portfolios; while the coefficients on client-level and

[16] The results of the estimations allowing for lagged and lead effects can be provided by the authors upon request.

contemporaneous macroeconomic covariates tend to be significant if no time effects have been allowed for, they become largely redundant once lagged macroeconomic effects enter the model. The only client-level factors that pass the elastic-net regularization procedure once time effects are included are the size of exposure at default and the length of the relationship with the bank. This, again, is fully consistent with the whole sample.

Table 8 GLM estimates with Logit LGD, Consumer Loans and Overdrafts, 2003q1-2010q2

	Consumer loans						***Overdrafts***					
	Client specific		*All variables, no lags*		*All variables, lags included*		*Client specific*		*All variables, no lags*		*All variables, lags included*	
Explanatory variable logit LDG	**(1)**		**(2)**		**(3)**		**(4)**		**(5)**		**(6)**	
	Coefficient	***SE***	***Coefficient***	***SE***	***Coefficient***	***SE***	***Coefficient***	***SE***	***Coefficient***	***SE***	***Coefficient***	***SE***
Client- specific factors												
Exposure at default	1.986***	0.078	1.198***	0.088	1.180***	0.087	1.536***	0.163	1.130***	0.158	1.102***	0.157
Relationship with bank			-0.020***	0.008	-0.026***	0.008	-0.049***	0.005	-0.108***	0.006	-0.111***	0.006
Age	0.774***	0.255										
Children	-0.823***	0.105					-1.163***	0.141	-0.180	0.152		
Phone	-0.387**	0.153					-0.792***	0.248				
Employment	-0.679***	0.130					0.310***	0.119				
Education	0.654***	0.164	-0.399***	0.123			-1.549***	0.197	-1.131***	0.188	-1.112***	0.187
Macroeconomic variables, current values												
Real GDP growth (y-o-y)									-0.964***	0.148	-0.241	0.187
ΔReal GDP growth (y-o-y)			0.389***	0.063					0.188**	0.081		
Real Consumption Growth (y-o-y)			-0.393***	0.098								
Real Investment Growth (y-o-y)			-0.039***	0.012								
Inflation rate (y-o-y)			-0.449***	0.077	0.022	0.096			-0.548***	0.088		
Property prices (y-o-y)			-2.076*	1.082					-0.607	1.459		
Default rate									-0.784***	0.251		
Retail loan growth (y-o-y)			-0.101***	0.021					-0.095***	0.034		
Macroeconomic variables, lagged and lead values												
Real GDP growth (y-o-y) (t-1)					-0.131	0.083						
Real GDP growth (y-o-y) (t-2)					0.008	0.088					-0.197**	0.097
ΔReal GDP growth (y-o-y) (t-2)											-0.126	0.167
ΔReal GDP growth (y-o-y) (t-3)											-0.089	0.167
Real investment growth (y-o-y) (t-1)					-0.001	0.021						
Real investment growth (y-o-y) (t-2)					-0.057***	0.018						

Real investment growth (y-o-y) (t-4)					-0.012	0.022						
Real consumption growth (y-o-y) (t-1)											-0.111	0.170
ΔUnemployment rate (t-1)					0.418	0.493						
ΔUnemployment rate (t-2)					1.378***	0.441					0.488	0.541
Property prices (y-o-y) (t-1)					-0.725	1.045						
Real wage growth (y-o-y) (t-3)											-0.325*	0.193
Real wage growth (y-o-y) (t-4)											-0.409*	0.235
Real wage growth (y-o-y) (t-5)					-0.644***	0.136					-0.181	0.197
Constant	-21.667***	1.314	44.793	66.878	105.835	84.422	-12.366***	1.509	-0.684	2.194	-3.745**	1.670
Alpha	0.90		0.90		0.30		0.9		0.9		0.5	
Observations	10287		10287		10287		5627		5627		5627	
Adjusted R-squared	0.070		0.186		0.187		0.052		0.160		0.163	
AIC	71266.880		69902.510		69898.895		40075.436		39408.654		39393.744	

Notes: *, **, *** stand for 10 %, 5 %, and 1 % significance levels respectively. Robust standard errors reported. The values of Alpha stand for the elastic net weighting parameter with lowest residual mean-squared error.

The benchmark specifications in Columns (1) and (4), which considered exclusively client-level information, largely conformed to expectations, but delivered strongly counterintuitive results for Education in the former case and Employment in the latter. These irregularities nonetheless proved irrelevant in the extended specifications and confirmed the prominence of socio-economic status as proxied by education of clients in affecting the success of the work-out process especially for overdrafts. In Columns (2) and (5) of Table 8, which summarize the estimates accounting for contemporaneous macro links, real consumption growth and investment growth proved to be statistically significant only in the case of consumption loans, which ultimately drove the results for the whole sample. The counterintuitive estimates for the default rate and absolute change in the GDP growth rate in Columns (2) and (5) replicate the output of the whole sample (Column (2), Table 7). Turning to the specifications with time effects in Columns (3) and (6), one can observe that the statistically significant macroeconomic effects are relevant to either consumer loans or overdrafts only, the only exception being real wage growth. Nonetheless, even in the case of real wage growth the actual statistically significant lags between the two sub-portfolios differ. It is thus relatively straightforward to assign the drivers of the results for the whole sample to a particular sub-portfolio. While real investment growth, unemployment rate-related factors, and real wage growth seem to be key macroeconomic correlates for the consumer loan LGD, for overdrafts the relevant macroeconomic time effects are real GDP and real wage growth.

The emergence of the above macroeconomic factors in the specification allowing for time effects is perhaps not too surprising given that real economic activity, real wages and unemployment status could substantially affect ability of defaulted clients to recover their debt. On the other hand, a partial absence of the same correlates in less flexible specifications in Table 8 (e.g., Columns (2) and (5)) is relatively striking and points to the importance of model flexibility when modelling LGD with macroeconomic covariates. In fact, a lack of strong priors on the most appropriate macroeconomic correlate in the present study stands in contrast to ad hoc specifications in existing studies that have found significant macroeconomic correlations across a wide range of variables (see Table 3).

6. Conclusion

Our approach allowed for investigation of the relevance of delayed macroeconomic effects to the realized LGD on a sample representing 15% of the Czech retail credit market. One should be careful when considering possible links between the macroeconomy and LGD, given that macroeconomic time series data tend to be heavily correlated and identification of size effects and their direction can suffer from multicollinearity issues. By opting for explicit treatment of model selection through elastic net regularization, we highlighted the importance of lagged macro effects other than the commonly employed contemporaneous effects. Furthermore, the socio-demographic background of the clients in our portfolio plays a lesser role than that indicated by exclusively client-level data or the model with contemporaneous values. Our results should thus have profound implications for the modeling of retail LGD for regulatory purposes, given that overly focusing on contemporaneous links between the macroeconomy and LGD might lead to misperception of the risks in banks' portfolios.

This conclusion is important both from the perspective of commercial banks developing credit risk models for estimating losses on their retail portfolios and to regulators assessing the resilience of the banking sector to adverse economic developments. This is particularly relevant for stress-testing exercises and the need to correctly link the sensitivity of the LGD parameter to macroeconomic developments. In this respect, our results show that the lagged effect of the macroeconomic environment must be taken into account to obtain reliable predictions of potential losses in banks' loan portfolios and correctly assess the robustness of the banking sector in the event of adverse macroeconomic developments.

REFERENCES

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

Altman E, Brady B, Resti A, Sironi A (2005a): The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications, *Journal of Business*, 78(6):2203–2228.

Altman E, Brady B, Resti A, Sironi A (2005b): The PD/LGD link: Empirical evidence from the bond market. In: Altman E, Resti A, Sironi A (eds) *Recovery risk*. Riskbooks, London, 217–233.

Asarnow E, Edwards D (1995): Measuring loss on default bank loans: A 24-year study. *Journal of Commercial Lending*, 77:11–23.

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

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

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

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

Calabrese, R (2012): *Predicting bank loan recovery rates with a mixed continuous-discrete model*, Applied Stochastic Models in Business and Industry.

Caselli S, Gatti S, Querci F (2007): The Sensitivity of the Loss Given Default Rate to Systematic Risk: New Empirical Evidence. *Journal of Financial Services Research*, 34:1–34.

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

Düllmann K, Trapp M (2004): Systematic risk in recovery rates—an empirical analysis of US corporate credit exposures. *Deutsche Bundesbank Discussion Paper*, Issue 2.

Emery K, Cantor R, Arner R (2004): *Recovery rates on North American syndicated bank loans, 1989–2003*, Moody's special comment.

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

Friedman J, Tibshirani R, Hastie T (2010): Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1):1–22.

Frye J (2000): Depressing recoveries. *Federal Reserve Bank of Chicago Working Paper*, Emerging Issues Series, 1–17 October.

Frye J (2005): The effects of systematic credit risk: A false sense of security. In: Altman E, Resti A, Sironi A (eds) *Recovery risk*. Riskbooks, London, 187–200.

Grippa P, Iannotti S, Leandri F (2005): Recovery rates in the banking: stylised facts emerging from Italian experience. In: Altman E, Resti A, Sironi A (eds) *Recovery risk*, Riskbooks, London.

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

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

Hastie TJ, Tibshirani RJ (1990): Generalized Additive Models. Chapman & Hall/CRC.

Jokivuolle E, Virén M (2013): Cyclical default and recovery in stress testing loan losses. *Journal of Financial Stability*, 9(1):139–149.

Qi M, Zhao X (2011): Comparison of modeling methods for Loss Given Default. *Journal of Banking and Finance*, 35:2842–2855.

Tibshirani R (1996): Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society*, 58:267–288.

Van der Kooij A (2007): Prediction accuracy and stability of regression with optimal scaling transformations. *Technical report*, Dept. Data Theory, Leiden Univ.

Witzany J, Rychnovský M, Charamza P (2010): Survival Analysis in LGD Modeling. *Working Papers IES 2010/02*, Charles University Prague, Faculty of Social Sciences, Institute of Economic Studies.

Zhou H, Hastie T (2005): Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society*, 67:301–320.

APPENDIX

Table A1 Consumer Credit in the Czech Republic and Selected EU Countries

Consumer Credit to GDP Ratio (%)											
	CZ	EU25	EA12	SK	PL	DE	AT	BG	FR	GR	ES
2004	2.3	7.1	6.5	2.7	4.7	7.9	10.3	6.5	7.9	9.0	7.1
2008	5.2	7.3	7.3	3.5	8.4	7.2	9.8	16.6	8.7	13.5	10.4
2013	6.1	6.8	6.7	7.4	9.3	7.1	8.0	16.4	8.1	17.6	6.7
Average Consumer Annual Credit Growth (%)											
	CZ	EU25	EA12	SK	PL	DE	AT	BG	FR	GR	ES
2004–2014	16.8	1.4	1.7	15.6	13.8	0.1	0.5	19.2	1.4	9.8	1.9
2004–2008	30.6	4.5	5.3	12.0	22.1	-0.2	4.6	38.9	4.0	18.6	13.3
2009–2013	3.1	-1.7	-2.0	19.2	5.5	0.3	-3.6	-0.5	-1.3	1.0	-9.5

Notes: Average growth rates computed from EUR-denominated values of the outstanding credit at the end of the particular year.
Source: ECB, Eurostat.

Table A2 Correlation plot of macroeconomic variables and LGD, whole sample 2003q1-2010q2

	Defaulted loans growth (y-o-y)	*Default rate (3 months)*	*ΔDefault rate (3 months)*	*NPL ratio*	*Cons. loans growth (y-o-y)*	*Unempl. rate*	*ΔUnempl. rate*	*Total empl.*	*Real wage growth (y-o-y)*	*Real GDP growth (y-o-y)*
Defaulted loans growth (y-o-y)	1									
Default rate (3 months)	0.51	1								
ΔDefault rate (3 months)	0.04	-0.30	1							
NPL ratio	0.51	0.51	-0.35	1						
Retail loans growth (y-o-y)	-0.52	-0.07	-0.35	-0.44	1					
Unemployment rate	0.40	0.79	-0.60	0.63	-0.03	1				
ΔUnemployment rate	0.24	0.34	0.48	-0.07	-0.35	0.09	1			
Total employment	-0.59	-0.66	0.17	-0.70	0.61	-0.75	-0.44	1		
Real wage growth (y-o-y)	-0.46	-0.24	-0.42	-0.21	0.51	0.03	-0.66	0.36	1	
Real GDP growth (y-o-y)	-0.34	-0.12	-0.68	0.03	0.67	0.13	-0.83	0.44	0.68	1
ΔReal GDP growth (y-o-y)	0.25	0.32	-0.30	0.42	-0.22	0.51	-0.30	-0.45	0.16	0.20
Real consumption growth (y-o-y)	-0.57	-0.32	-0.50	-0.37	0.78	-0.14	-0.59	0.61	0.72	0.79
Real investment growth (y-o-y)	-0.27	-0.26	-0.53	0.18	0.36	-0.04	-0.85	0.39	0.67	0.84
Real Pribor3m	0.02	0.23	0.09	0.14	-0.40	0.27	0.28	-0.49	0.18	-0.42
Rate on retail loans	0.44	0.92	-0.34	0.70	-0.20	0.81	0.24	-0.74	-0.24	-0.10
Retail rate spread	0.45	0.88	-0.42	0.77	-0.20	0.88	0.18	-0.78	-0.16	-0.03
CPI(y-o-y)	-0.25	-0.33	0.22	-0.45	0.39	-0.67	-0.25	0.74	-0.10	0.23
Property price index (y-o-y)	-0.54	-0.73	0.06	-0.49	0.43	-0.70	-0.59	0.84	0.58	0.51
Stock market index (PX-50) (y-o-y)	0.23	0.56	-0.58	0.15	0.31	0.70	-0.32	-0.19	0.51	0.52
Real ind. production growth (y-o-y)	0.33	0.18	-0.37	0.31	-0.03	0.51	-0.42	-0.28	0.30	0.37

	ΔReal GDP growth (y-o-y)	Real cons. growth (y-o-y)	Real inv. growth (y-o-y)	Real Pribor3m	Rate on retail loans	Retail rate spread	CPI (y-o-y)	Prop. pr. Index (y-o-y)	Stock market index (PX-50) (y-o-y)	Real ind. prod. growth (y-o-y)
Defaulted loans growth (y-o-y)										
Default rate (3 months)										
ΔDefault rate (3 months)										
NPL ratio										
Retail loans growth (y-o-y)										
Unemployment rate										
ΔUnemployment rate										
Total employment										
Real wage growth (y-o-y)										
Real GDP growth (y-o-y)										
ΔReal GDP growth (y-o-y)	1.00									
Real consumption growth (y-o-y)	-0.20	1.00								
Real investment growth (y-o-y)	0.06	0.69	1							
Real Pribor3m	0.23	-0.26	-0.28	1						
Rate on retail loans	0.34	-0.33	-0.14	0.31	1					
Retail rate spread	0.39	-0.29	-0.06	0.32	0.98	1				
CPI(y-o-y)	-0.36	0.29	0.20	-0.70	-0.43	-0.54	1			
Property price index (y-o-y)	-0.32	0.64	0.64	-0.37	-0.74	-0.73	0.57	1		
Stock market index (PX-50) (y-o-y)	0.63	0.21	0.33	0.24	0.36	0.42	-0.34	-0.12	1	
Real ind. production growth (y-o-y)	0.48	0.03	0.46	-0.03	0.23	0.33	-0.38	-0.06	0.54	1

Table A3 GLM Estimates with Logit LGD, Sample Excl. Credit Cards 2003q1-2010q2

	Client specific		***All variables, no lags***		***All variables, lags included***	
Explanatory variable logit LDG	*(1)*		*(2)*		*(3)*	
	Coefficient	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**
Client- specific factors						
Exposure at default	1.640***	0.064	1.044***	0.064	1.054***	0.063
Relationship with bank	-0.032***	0.004	-0.090***	0.005	-0.085***	0.004
Age	0.327	0.217	0.832***	0.207		
Children	-1.039***	0.085	-0.102	0.093		
Phone	-0.527***	0.132	0.357***	0.125		
Employment	0.186**	0.075	-0.096	0.069		
Education	-0.999***	0.109	-0.734***	0.104	-0.707***	0.103
Female	0.364***	0.139	0.077	0.131		
Macroeconomic variables, current values						
∆Real GDP growth (y-o-y)			0.386***	0.051		
Real Consumption Growth (YoY)			-0.555***	0.082	-0.314***	0.078
Real Pribor3m			0.154	0.140		
Inflation rate (y-o-y)			-0.392***	0.085		
Property prices (y-o-y)			-0.061***	1.215		
Default rate			-0.491**	0.191		
Retail loan growth (y-o-y)			-0.077***	1.909	-3.282	2.273
NPL ratio			-0.232	0.150		
Retail rate spread			-0.796***	0.248		
Macroeconomic variables, lagged and lead values						
Real GDP growth (y-o-y) (t-1)					-0.177**	0.072
Real GDP growth (y-o-y) (t-2)					-0.257***	0.073
Real investment growth (y-o-y) (t-2)					-0.067***	0.015
Unemployment rate (t-8)					1.123***	0.258
Real wage growth (y-o-y) (t-3)					-0.271**	0.125
Real wage growth (y-o-y) (t-4)					-0.516***	0.121
Real wage growth (y-o-y) (t-5)					-0.135	0.099
Constant	-15.597***	1.024	121.929	84.326	-9.593***	1.691
Alpha	0,2		0,4		0,7	
Observations	15914		15914		15914	
Adjusted R-squared	0.065		0.178		0.178	
AIC	111591.197		109556.178		109563.544	

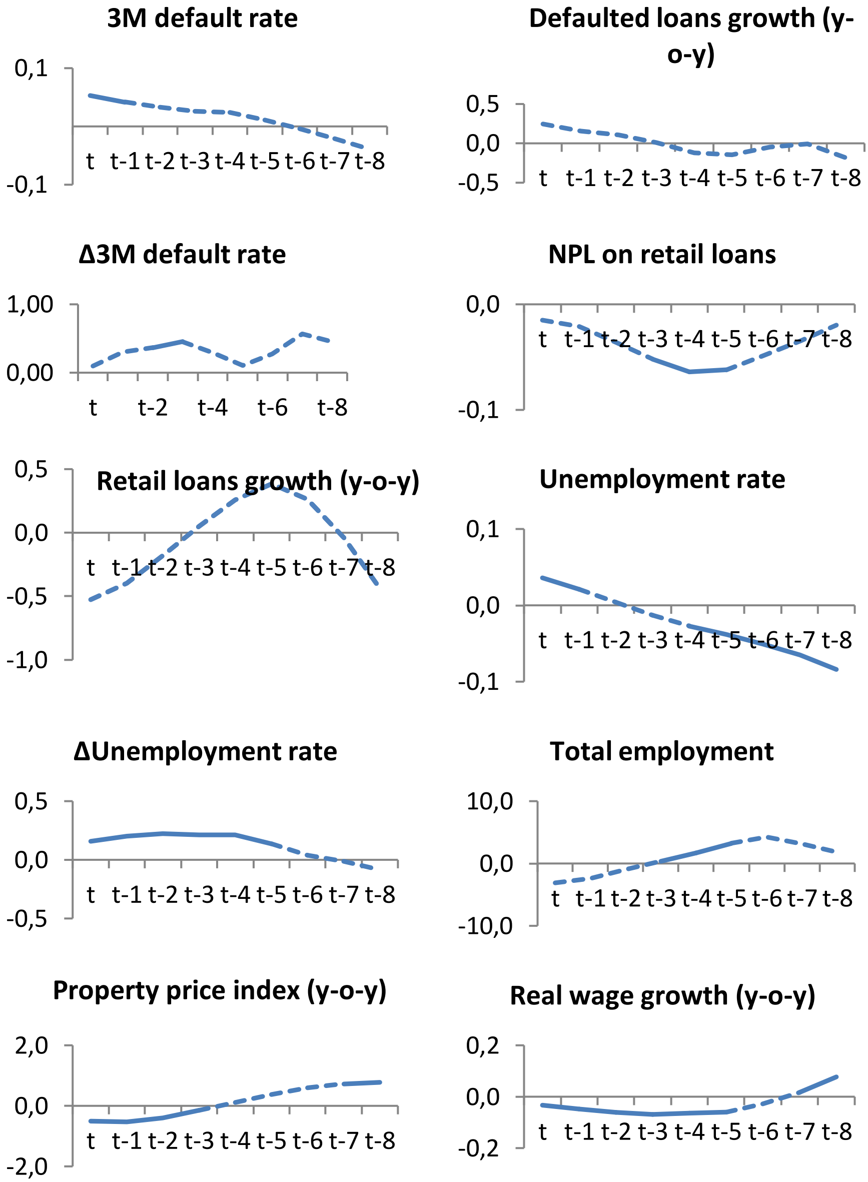
Notes: *, **, *** stand for 10 %, 5 %, and 1 % significance levels respectively. Robust standard errors reported. The values of Alpha stand for the elastic net weighting parameter with lowest residual mean-squared error.

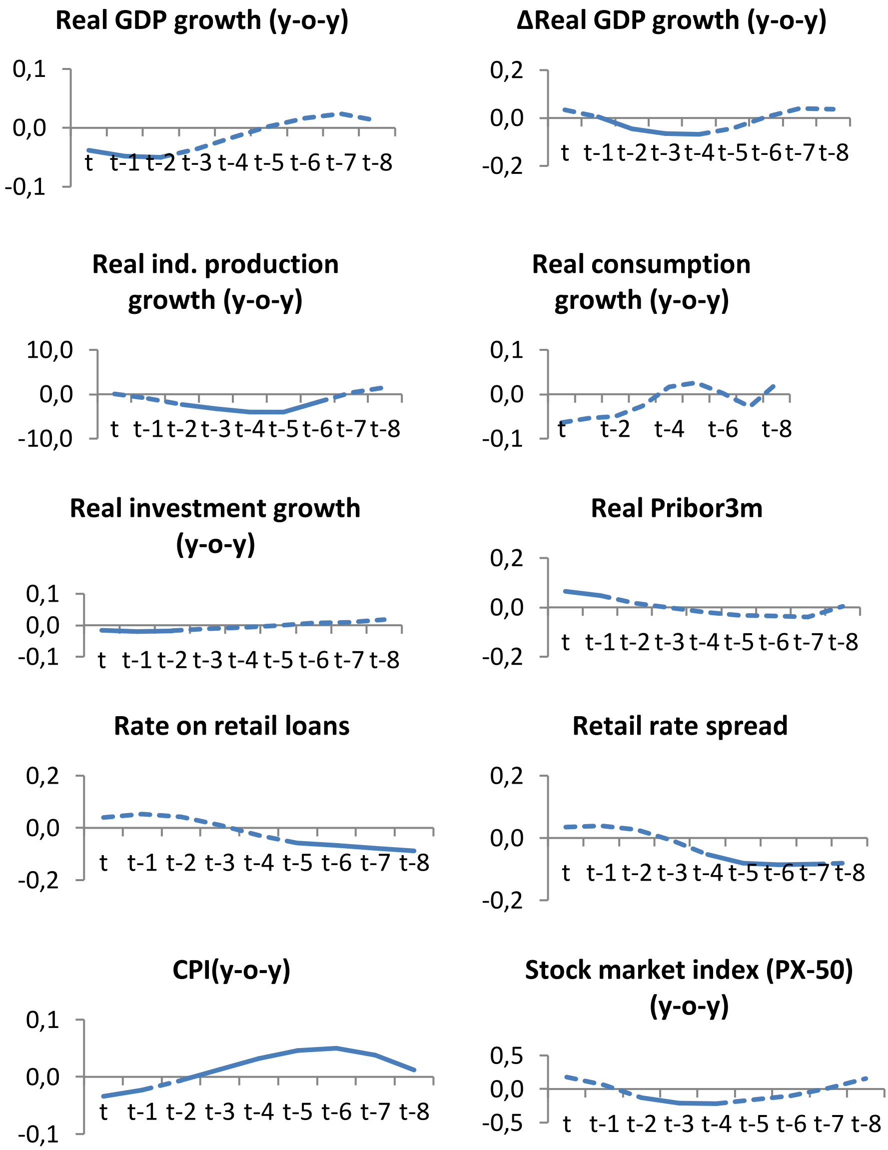
Table A4 GLM Estimates with Logit LGD, Credit Cards, 2003q1-2010q2

	Client specific		***All variables, no lags***		***All variables, lags included***	
Explanatory variable logit LDG	*(1)*		*(2)*		*(3)*	
	Coefficient	**SE**	**Coefficient**	**SE**	**Coefficient**	**SE**
Client- specific factors						
Exposure at default	0.074	0.163	0.125	0.163	0.11	0.161
Relationship with bank	-0.003	0.008	-0.014*	0.009	-0.01	0.008
Age	0.838**	0.369	0.749**	0.368	0.623*	0.349
Children	-0.261	0.191	0.296	0.220		
Phone	-0.039	0.222	0.345	0.234	0.355	0.233
Macroeconomic variables, current values						
Real Investment Growth (y-o-y)			-0.014	0.047		
Total employment (in thds)			2.33	2.734		
Property prices (y-o-y)			6.393*	3.619		
Default rate			-0.866	0.843		
Retail loan growth (y-o-y)			-8.335	5.636		
Macroeconomic variables, lagged and lead values						
Real Pribor3m (t-1)					0.147	0.161
Real Pribor3m (t-2)					0.188	0.151
Total employment (in thds) (t-8)					1.515**	0.75
Property prices (y-o-y) (t-6)					6.249***	1.163
Constant	-1.894	2.125	-238.201	231.413	-134.564**	61.819
Alpha	0.8		0.9		0.9	
Observations	2784		2784		2784	
Adjusted R-squared	0.003		0.023		0.025	
AIC	17484.893		17438.26		17424.22	

Notes: *, **, *** stand for 10 %, 5 %, and 1 % significance levels respectively. Robust standard errors reported. The values of Alpha stand for the elastic net weighting parameter with lowest residual mean-squared error.

Figure A1 Beta Coefficients from Regressions of Average LGD on Individual Lagged Values of Macroeconomic Variables (2003q1/2005q1–2010q2)





Notes: Solid lines represent correlations significant at the 5 % level.