Macroeconomic Factors as Drivers of LGD Prediction

Belyaev, Konstantin; Belyaeva, Aelita; Konečný, Tomáš; Seidler, Jakub; Vojtek, Martin
2012

WORKING PAPER SERIES 12

2012

Konstantin Belyaev, Aelita Belyaeva, Tomáš Konečný, Jakub Seidler, and Martin Vojtek:

Macroeconomic Factors as Drivers of LGD Prediction: Empirical Evidence from the Czech Republic

CZECH NATIONAL BANK

WORKING PAPER SERIES

Macroeconomic Factors as Drivers of LGD Prediction: Empirical Evidence from the Czech Republic

Konstantin Belyaev
Aelita Belyaeva
Tomáš Konečný
Jakub Seidler
Martin Vojtek

12/2012

Reviewed by: Tigran Poghosyan (International Monetary Fund)
Petr Franěk (Česká Spořitelna, a.s.)
Michal Nyklíček (Czech National Bank)

Project Coordinator: Bořek Vašíček

Macroeconomic Factors as Drivers of LGD Prediction: Empirical Evidence from the Czech Republic

Konstantin Belyaev, Aelita Belyaeva, Tomáš Konečný, Jakub Seidler, and Martin Vojtek*

Abstract

This paper focuses on key macroeconomic driving factors influencing the loss given default (LGD) – an important credit risk parameter determining credit losses of the banking sector. Various econometric approaches are applied on both individual and aggregated data for different bank segments in order to identify the sensitivity of LGD parameters to both the micro characteristics of debtors and aggregated macro-level data. Despite the relatively low importance of macro variables in the model combining micro- and macroeconomic information, our estimates suggest that the macroeconomic environment contributes directly to the variation in LGD. The results from the different approaches confirm a negative link between LGD and consumption growth for the retail portfolio, while in the case of the corporate segment, a negative link between LGD and real GDP growth is identified. Importantly, given that aggregation effects and non-linearities may substantially affect the choice of relevant macroeconomic variables, it is essential to distinguish between models employing purely macroeconomic data and models combining micro- and macro-based information.

JEL Codes: C02, G13, G33.

Keywords: Credit losses, loss given default, recovery rates, work-out LGD.

*Konstantin Belyaev, konstbell@gmail.com; Aelita Belyaeva, aelita.belyaeva@gmail.com; Tomáš Konečný, Czech National Bank, tomas.konecny@cnb.cz; Jakub Seidler, Czech National Bank, jakub.seidler@cnb.cz; Martin Vojtek, Czech National Bank, martin.vojtek@cnb.cz.

The views expressed in this paper are those of the authors and do not necessarily represent those of the Czech National Bank or other financial organizations. The authors acknowledge support from Czech National Bank Research Project C4/2010. The authors thank Petr Franěk, Roman Horváth, Marian Krajč, and Michal Nyklíček for helpful comments. However, all errors and omissions are those of the authors.

Nontechnical Summary

This paper analyzes the determinants of work-out loss given default (LGD) with respect to macroeconomic factors. Various econometric approaches are employed on both individual and aggregated data for different segments (retail and corporate) in order to identify the sensitivity of LGD parameters to both the individual characteristics of debtors and macro-level data.

We use both an individual and aggregated data set on retail and corporate defaulted loans and LGDs. We investigate this subject using two different approaches. Firstly, we estimate a microeconometric model for explaining the LGD of a portfolio. However, in addition to individual client characteristics, macroeconomic variables are used in the search for the proper model specification (i.e., we use a combined model). Secondly, a pure aggregated-data model is estimated using exclusively macroeconomic information. The results from the two approaches are compared to obtain more robust information regarding the link between LGD and macro factors.

The results obtained from the combined model and the aggregated model using time series of macro data lead to notable differences in the suggested macro links obtained from the different estimation frameworks, even though the importance of macro variables in the models employing individual data might be limited. Firstly, our micro and macro perspectives identify jointly only a single macroeconomic factor within each of the two segments concerned. While for retail customers the models indicate consumption-related factors as significant determinants, for corporate clients the major determinant is real GDP growth. Other potential candidates do not robustly enter either the microeconometric model or the macroeconomic model. Still, our study partly conforms to the results from previous literature, as Caselli et al. (2008) also identified consumption growth among the macroeconomic factors influencing the retail LGD, and GDP growth among those influencing the SME LGD for Italian banking data.

The results of our study should help us gain more detailed information about the link between the banking sector LGD and the business cycle of the Czech economy. This might be useful for – among other things – identifying the potential losses in each bank's portfolio based on macroeconomic developments. Nonetheless, it is important to note that the differences in the relevant macroeconomic factors derived from a purely macro-based as compared to combined framework might have substantial implications for the conduct of top-down solvency stress tests performed by regulatory authorities. These rely typically on a battery of so-called satellite models linking macroeconomic developments to the financial sector. If micro-level information, such as client balance sheet data, is missing, estimates obtained exclusively from macro data might paint a rather different picture than more richly specified microeconometric models. As a result, the outcome of this study might have positive implications for the current CNB stress-testing framework, as it provides link between LGD behavior and macroeconomic variables of the Czech economy.

1. Introduction

Over past 20 years, the literature has mostly emphasized the crucial importance of default rate (PD) modeling as one of the important credit risk components indicating credit quality. Less discussion has focused on other credit risk parameters such as loss given default (LGD) and exposure at default (EAD). These two parameters have received more attention only in recent years with the advent of the Basel II Capital Accord adopted in 2006, which enables banks to estimate PD, LGD, and EAD parameters for determining the capital requirements within the Advanced Internal Rating Based (AIRB) approach. The Basel II rules were introduced in the European Union in 2007 with the so-called Capital Requirements Directive (CRD). The CRD strongly emphasizes the importance of LGD and provides incentives for accurate measurement of this parameter.

Since then, the focus on modeling LGD by banks and practitioners has substantially increased, making its estimates – and therefore also banks' capital requirement calculations – more accurate. Research interest has also focused on the macroeconomic determinants of LGD, since its value might vary with the economic cycle, as different characteristics influencing the recovery process might be determined by the current stage of the economy (e.g., the price of the collateral encumbered in the defaulted loan contract).

In addition, regulators' need to possess reliable predictions of potential losses in banks' loan portfolios and to forecast the credit losses of the banking sector with respect to macroeconomic developments has led to increased attention being paid to the estimation of a model linking economic developments with the value of banks' LGD parameters. This model is one of the crucial input parameters into macro stress-testing models, which are used by regulators for assessing the resilience of the banking sector to adverse economic developments. Moreover, banks are also being encouraged to develop stress tests of their credit portfolio performance in terms of macroeconomic shocks and to estimate the so-called downturn LGD – the value of the parameter reflecting losses during downturns in the business cycle. The given reasons imply close investigation not only of the relationship between the LGD outcome and potential predictors characterizing clients, but also of the linkages between LGD performance and the economic cycle.

Nonetheless, recent studies have shown that the development of LGD models suffers from limited and/or incomplete data on the recovery process, including an insufficient time span covering the collection of debts after a default event. As a result, a well established and accepted paradigm for LGD modeling has still not been determined.

The present study focuses on the link between the loss given default (LGD) and key characteristics of debtors and macro-level data. We focus in particular on the relationship between LGD and macroeconomic factors for the Czech Republic. The analysis is performed both on individual data on defaulted clients and on aggregated time series. Importantly, given that modelers' aggregate solvency stress tests often lack sufficiently granular data, we employ both a "bottom up" and a "top down" approach to model selection in order to better understand the gains from the inclusion of more detailed client- and/or facility-level information. Put differently, by performing a separate model selection exercise using combined data ("bottom up") and aggregate-level data ("top down"), we allow for maximum flexibility to exploit the information at each level of aggregation and then compare the final specifications. We do not restrict ourselves to nested models such as Jacobson et al. (2011). Similarly, our analysis does not derive from an integrated micro-macro framework à la De Graeve et al. (2008). One should note, however, that the generality of our setup allows for each of the two approaches as special cases.

The main hypothesis of this study is that LGD for both retail and corporate portfolios is determined to some extent by macroeconomic factors, i.e., that the recovery rate of bank loans may depend on the state of the economic cycle; however, different client segment LGDs might have a different set of explanatory macro indicators influencing LGD. Furthermore, we expect the "bottom up" and the "top down" approaches to deliver mutually consistent results with respect to the choice of relevant macro determinants.

The study uses a unique comprehensive data set on loan losses corresponding to about 15% of the corporate credit market and 15% of the retail credit market in the Czech Republic. We concentrate on the so called work-out LGD resulting from the collection of defaulted debt obligations and 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 focus on an individual institution rather than a whole sector. Nonetheless, we believe it offers valuable information based on a scarce and comparatively rich data source as well as straightforward transferability of our methodology to other institutions.

2. Literature Review

The last decade has seen pronounced growth in research interest in the macroeconomic determinants of LGD risk parameters (see, for example, Altman et al., 2005b, and Düllmann and Trapp, 2004, and references therein). The reason is that the estimation of capital requirements might be more precise if LGD is modeled as a function of macroeconomic factors, which would, in turn, refine the assessment of banks' capital adequacy ratios (CAR). It is apparent that in the event of a severe downturn banks might suffer twin losses – a higher rate of default among their borrowers, and a lower rate of recovery of defaulted loans (recovery rate RR = 1−LGD). Getting long-term average LGDs that do not take into account the possible consequences of a severe downturn can lead to significant capital underestimation (Frye, 2005).

A few studies have explicitly addressed the recovery rate and its relationship to the state of the economic cycle (Altman et al., 2005a, 2005b, 2005c; Frye 2002, 2005) by employing publicly traded defaulted bonds instead of work-out LGDs from defaulted bank loans, whose characteristics are significantly different from those of corporate bonds.[1] A number of studies, on the other hand, have worked with work-out LGDs (e.g., Dermine and Neto de Carvalho, 2006, and Caselli et al., 2008).

The empirical results by Dermine and Neto de Carvalho (2006) relate to the timing of recoveries of bad and doubtful bank loans and to the distribution of cumulative recovery rates. The authors

[1] A number of studies on work-out 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).

estimate models constructed on the basis of individual loan transactions and estimating the economic determinants of the LGD of European bank loans granted to small and medium-sized companies. Their multifactor models include explanatory variables such as loan size, type of guarantee/collateral support, industry sector, default year, and age of the firm. Nevertheless, macroeconomic variables such as GDP growth, frequency of default in the industry sector, and the interest rate were additionally tested with no statistical significance. This fact was explained by the absence of a sufficient recession during the period under consideration. Bellotti and Crook (2009) on the contrary found a significant relationship between LGD and a number of macroeconomic variables in their analysis of credit card LGDs on individual loan transactions.

Caselli et al. (2008) examined the sensitivity of LGD to systematic risk and found a relationship between LGD and the macroeconomic conditions on aggregated data. The authors developed separate multivariate models for two customer segments – SMEs and households. For SMEs the best model incorporates the aggregate number of employees and the GDP growth rate; for households LGD is best explained in relation to the default rate, 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.

A thorough analysis linking LGD with macroeconomic variables has not been performed for the Czech Republic so far. First of all, the insufficient database of work-out LGDs makes it difficult to develop an accurate model linking LGDs to macroeconomic factors in the case of the Czech economy.[2] Secondly, recent literature regarding LGD in the Czech economy has focused more on proper methods within the context of individual defaulted loans and client characteristics. Witzany et al. (2010) explored survival analysis methods for LGD modeling, leaving macroeconomic factors unexplored. Finally, Seidler et al. (2009) showed a positive correlation between the estimated value of market LGD and the aggregated corporate default rate. However, their estimates relate solely to selected publicly listed companies, which implies limited relevance for analysis of the whole corporate sector.

Our study also relates to the literature on the feedback between the real sector and the financial sector combining detailed information from credit registers with macro variables. For example, De Graeve et al. (2008) develop an integrated micro-macro framework for testing the relationship between the aggregate default rate and output, inflation, and the interest rate. The authors estimate firm-level default rates using a large firm-level database from the Credit Bureau register augmented by selected macroeconomic variables. Similarly, Jacobson et al. (2005) examine the relation between macroeconomic fluctuations and defaults using Swedish corporate data. In particular, the authors' evaluations indicate that the combined micro-macro framework is superior both to models that exclude macro information and to best-fitting standard time-series models. Besides focusing on LGD instead of default rates as in Jacobson et al. (2005), our framework is more flexible by allowing for variable selection at both the micro and macro level.[3] This possible

[2] Kocenda and Vojtek (2011) faced a similar problem related to the availability of comprehensive retail default data. The data were ultimately provided by an anonymous wholesale bank operating on the Czech market.

[3] Several studies have focused on joint modeling of default and recovery rates. Other credit risk measures (apart from traditional default rates and LGD) employed within stress-testing frameworks include the write-off to loan ratio (Hoggarth et al., 2005), the non-performing loans ratio (e.g., Jimenez and Saurina, 2006; Quagliariello, 2007), the ratio of net loan losses to total loans (Pesola, 2007), and loan loss provisions (Quagliariello, 2007).

improvement might also have positive implications for the current CNB stress-testing framework, which has so far employed only simplifying assumptions concerning the behavior of LGD and its link to macroeconomic variables (GDP and property prices; see Geršl and Seidler, 2012).

3. Data Employed

3.1 Micro-level Data and Definitions of Basic Risk Parameters

While all financial institutions following the Internal-Rating Based Approach (IRBA) need to interpret and quantify the risk parameters entering their capital calculations, a slight difference in parameter definitions might exist. Therefore, we first provide our definitions and calculations of the basic risk parameters, such as default, the default rate, LGD, and EAD, which are further used in this paper.

The Basel Accord defines a default event as a realization of one or more of the following circumstances: the credit obligor is i) unlikely to pay, ii) is more than 90 days past due, or iii) is declared bankrupt. We employ the first default condition in the following way: a default event occurs if a client has had at least one of his credit accounts restructured during a calendar year. In such case the client becomes labeled as "unlikely to pay." While any of the three conditions applies for corporate clients, 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):

$$LGD = LOSS/EAD. \tag{1}$$

In order to measure LGD in this way, recovery cash flows from defaulted loans as well as the costs of the bank's work-out process must be observed. Loss experienced by a bank is understood to mean economic loss, i.e., 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. Usually, the LGD experienced is truncated at zero and one in order to make it comparable with the market LGD (obtained after the sale of defaulted market instruments). We follow this approach as well.

The data correspond to about 15% of the corporate credit market and 15% of the retail credit market in Czech Republic. The sample of defaulted accounts historically collected in the retail portfolio consists of 34,078 cases. The sample of defaulted clients gathered from the corporate credit portfolio is 3,193 cases.

The results are focused mainly on the elasticities between LGD and macroeconomic variables, leaving the levels and evolution of fitted LGD values undiscovered.

For the retail portfolio we used historical observations from transactional systems on a monthly basis. The time span of the retail data covers the period from January 2002 until June 2012. Apart from the information on LGDs, the data contain the standard social-demographic characteristics of retail clients usually collected by banks when a credit account is opened. The data set on retail customers likewise includes client behavioral characteristics relating to different credit products such as overdrafts, consumer loans, and credit cards. However, our data set does not contain mortgages, which are not the subject of our study.

The retail portfolio offers information on contract length and the balance and interest rate of the credit account each month from the beginning of the account's existence until June 2012. In addition, we calculate the default date for each credit account. This is used for sample construction when we calculate the duration of the work-out process.

The data set relating to corporate clients covers the time span from 1993 until 2012. However, it suffers from some imperfections; for example, the information is collected only at the moment of default for each client. Dynamic changes in the balances of client credit accounts are therefore not available. We observe the client's segment (corporate or SMEs), the default date, exposure at default, the client's credit rating before the default event, the lender's fiscal situation, and some other variables relating to credit collateralization.

3.2 Macroeconomic Data

The macroeconomic data we use reflect the choices made in the existing literature as well as (given the absence of a generally accepted structural framework for the determinants of LGD behavior) other macroeconomic variables available from the core DSGE "g3" forecasting model of the Czech National Bank (CNB). The availability of the above-mentioned variables facilitates their use within the aggregate solvency stress-testing framework of the CNB.

We present the data as a normalized time series (Figures 1 and 2).

Table 1 presents the macroeconomic variables employed by selected studies on LGD. Table 2 contains the full list of macroeconomic variables used in both the aggregated macro model and the combined model. The main sources of macroeconomic data are the Czech National Bank (CNB) and the Czech Statistical Office (CZSO).

Table 1: Macroeconomic Variables Used in the Literature on LGD Modeling

Altman et al. (2005b)	GDP growth
Calabrese (2010)	GDP growth, interest rate, unemployment rate, default rate
Caselli et al. (2008)	GDP growth, employment, delta of unemployment rate, household consumption, total gross investment, total production, total income, delta of default-to-loan ratio, total bank loans
Crook and Belotti (2009)	interest rate, unemployment rate
Dermine and Neto de Carvalho (2006)	GDP growth, industry default rates
Jokivuolle and Viré (2011)	output gap, interest rate, indebtness, gross profit
Qui and Yang (2009)	GDP growth

Table 2: Macroeconomic Variables Used in the Combined and Aggregated Models

Real Consumption Growth (QoQ)	Export Price Inflation (Home Curr.) (QoQ)
Real Consumption Growth (YoY)	Export Price Inflation (Home Curr.) (YoY)
MP Inflation (QoQ)	Foreign Inflation (Foreign Curr.) (QoQ)
MP Inflation (YoY)	Foreign Inflation (Foreign Curr.) (YoY)
CPI Inflation (QoQ)	Nominal Depreciation (QoQ)
CPI Inflation (YoY)	Nominal Depreciation (YoY)
Real Government Cons. growth (QoQ)	Nom. Wage Growth (QoQ)
Real Government Cons. growth (YoY)	Nom. Wage Growth (YoY)
Real GDP Growth (QoQ)	Real Export Growth (QoQ)
Real GDP Growth (YoY)	Real Export Growth (YoY)
Nominal Government Cons. Growth	Interest Rate
Nominal Government Cons. Growth	CZK/EUR
Real Investment Growth (QoQ)	1-Year Interbank Rate % pa
Real Investment Growth (YoY)	Euro 1-Year Interbank Rate % pa
Real Import Growth (QoQ)	Euro 3-Month Interbank Rate % pa
Real Import Growth (YoY)	Euro Zone Real GDP % pa y-o-y
Foreign Demand Growth (QoQ)	Unemployment Rate % of Labour Force
Foreign Demand Growth (YoY)	ILO Unemployment Rate
Consumption Price Inflation (QoQ)	Total loans
Consumption Price Inflation (YoY)	Total corporate loans
Government Cons. Inflation (QoQ)	Total household loans
Government Cons. Inflation (YoY)	Total consumption loans
GDP Deflator (QoQ)	Non-performing loans/Total loans
GDP Deflator (YoY)	Non-performing loans/Corporate loans
Investment Price Inflation (QoQ)	Non-performing loans/Household loans
Investment Price Inflation (YoY)	Non-performing loans/Cons. loans
Import Price Inflation (QoQ)	Property price index
Import Price Inflation (YoY)	Adjusted operating profit

Source: CNB

4. Methodology

The paper investigates data at both the micro and macro level. As Han and Jang (2013) note, different studies suggest different factors and there is no consensus on these factors except collateral. This inconsistency can be attributed to "the differences in loan portfolios among banks, lending and debt collection procedures among countries, LGD measurement methods and/or sample periods." Given the above-mentioned uncertainties about the selection of proper micro and macro determinants, we decided to allow for maximum flexibility by employing a backward-selection algorithm for both the combined and aggregated (time series) models.

Our micro approach to LGD modeling (i.e., the combined model) exploits individual client information on socio-demographic as well as behavioral characteristics related to credit accounts. Specifically, in the retail portfolio the unit of observation is the credit facility, i.e., our sample might include several credit products granted to the same client. For the corporate portfolio, on the other hand, we use granularity on the client level. The micro data are supplemented with selected macroeconomic aggregates at quarterly frequency within a duration and/or hierarchical/multi-level model framework, depending on the segment and the corresponding data limitations.

The alternative macro perspective (i.e., the aggregated model) investigates the sensitivity of aggregated LGDs for different bank portfolios to the set of macroeconomic indicators using a general-to-specific model selection algorithm within the ARMAX (AutoRegressive Moving Average with eXogenous inputs) class of time-series models.

The aim of both modeling approaches – combined and aggregated – is to highlight the importance of macroeconomic characteristics in the estimation of LGD by using aggregate and individual data sets and employing different types of modeling techniques.

4.1 Methodology for Individual-level Data – Combined Model

The combined approach involves models based on individual data on credit losses which incorporate macroeconomic factors to control for changes in the macroeconomic environment. Similarly to the aggregated data analysis, the sensitivity of LGD is investigated across different banks' portfolios. The client-level data for the corporate segment include, for example, the number of employees in the SME and the length of cooperation with the bank, and for retail portfolios include, for example, age and education.

Past studies have used both (semi-)parametric (Generalized Additive Models, duration models, etc.) and non-parametric methods (such as regression trees and neural networks). Bastos (2010) argues for regression trees and against the alternative of fractional response regressions as a specific form of generalized additive models. More generally, Han and Jang (2013) argue that non-parametric methods tend to perform better than parametric methods for the purposes of LGD modeling. Nonetheless, for the sake of simplicity in comparing the macro and combined models we prefer the parametric framework.[4]

For the retail segment we explore two different approaches to LGD modeling: the survival analysis model and the generalized linear regression model (GLM). The two methods are adopted in order to exploit the richer data available for the survival framework and to compare the results with the GLM output, where the information available on clients/facilities is more limited. Given that the survival analysis allows for different durations of the recovery process while the GLM does not, the two perspectives might give us at least some clue as to what factors might affect the outcome (LGD) of the collection process as opposed to its actual length.

4.2 Survival Analysis for Retail Portfolio

One of the specific issues of LGD modeling is that the LGD outcome is evident not immediately after default but after a certain period, when the debt is to be collected in the so-called work-out process. The success and length of the work-out process depend on a number of factors known and unknown to the observer. We take the intrinsic feature of collection duration into account by adopting the survival modeling framework. Survival (or duration) analysis has been used in other studies on LGD, for instance, Witzany et al. (2010) and Zhang and Thomas (2012).

To estimate the final work-out LGD we need to wait until the collection process has been completed. The final LGD for clients who defaulted in 2012 might be available after approximately 1 year in the retail segment and 2–3 years in the corporate segment. This implies that we can work only with defaults registered before 2009. As a result, the data sample is reduced substantially. An alternative approach might be to keep additional files which defaulted after 2009 and which have a fully completed collection process. Nonetheless, given that the file duration is shorter than the average collection time, our sample would be overweighted with files closed relatively fast and the above-mentioned approach might lead to underestimation of the LGD outcome. Similarly, with the inclusion of all defaulted clients who defaulted just recently, the sample would contain many observations with 100% LGD, since the recovery process has not started yet and the resulting LGD prediction might thus be overestimated. Setting the right balance between these approaches and choosing the appropriate sample are challenging yet still unresolved issues.

In order to use as much information as possible and avoid unnecessary sample restrictions we employ survival analysis, which preserves practically all the defaulted files in our sample. The survival methodology deals with objects which have duration and analyzes the length of time that passes from the beginning of some event (state) either until it ends or until the observation period is completed. We assume that a random variable T capturing the time of leaving a given state during a given observation period has a continuous probability distribution $f(t)$. The cumulative probability is expressed as

[4] Other parametric approaches, such as inverse-Gaussian or beta transformation, in fact do not seem to outperform standard approaches (Qui and Zhao, 2011). Preceding studies resorting to linear regression include Altman et al. (2005b), Caselli et al. (2008), and Grunert and Weber (2009).

$$F(t) = \int_0^t f(s)ds = P(T \le t), \tag{2}$$

which means the probability of leaving the state before time *t*. The probability of not leaving the state (or the event of interest) before time *t* corresponds to

$$S(t) = 1 - F(t) = P(T > t), \tag{3}$$

which represents the so-called survival function. The survival methodology builds upon the concept of the hazard function *H(t)*, which is defined as the probability of leaving a given state at duration (time) *t* conditional upon staying there up to that point in time:

$$h(t) = \lim_{\partial t \to 0} \frac{P(t \le T < t + \partial t / T \ge t)}{\partial t} = \lim_{\partial t \to 0} \frac{F(t + \partial t) - F(t)}{\partial t \, S(t)} = \frac{f(t)}{S(t)}. \tag{4}$$

The observation period for the retail portfolio starts in January 2002 and ends in June 2012 and time *t* is measured in months. Unfortunately, limited data availability prevents duration analysis for the corporate portfolio, as changes in client recoveries over time are missing. Consequently, it is not possible to construct our duration variable.

The starting point of the collection process corresponds to the moment of default, which in general differs across facilities. The collection process lasts until the facility is repaid. More specifically, we assume that the duration starts at the moment of default and finishes once the process of collection of the client's debt obligations has been effectively closed. Such a situation occurs, for example, when the client's repayments reach 90% of the total debt amount (i.e., LGD is less than 10%). For some defaulted facilities – especially in the most recent years – the collection process has not been finished yet, due to the relatively short time spell after default.[5]

For each facility, we record the number of months the facility stays in the particular state (defined as the payoff of the past-due receivable being less than 90%) before it exits this state. Secondly, we follow whether each observed facility was censored or not. If the whole repayment history of the facility did not reach 90% of the debt amount, it is labeled as censored. Otherwise, the facility is labeled as uncensored.[6] Censored observations are addressed directly in the likelihood function of the duration model (Green, 2000).

The probability that the facility is not recovered until time *t* is expressed by the survival function *S(t)*. Conversely, *F(t)* means the probability that the facility is paid by time *t*. Furthermore, in our notation the hazard function *H(t)* determines the instantaneous probability of 90% of the debt being paid off at moment *t* given that until this moment the pay-offs have not reached 90% of the outstanding debt.

In order to incorporate macroeconomic variables into the survival model, we chose the semi-parametric form of the hazard function proposed by Cox (1972). The model presents the hazard function in the following form:

[5] The survival analysis for these censored observations adjusts the estimation procedure accordingly.
[6] Our censoring is induced by the specifics of the credit process for retail products: a product in default with a past-due history has to be repaid and then annulled. Facilities without full repayment during the observation period are considered censored given that the final collection remains unknown.

$$h(t) = exp\,(-\boldsymbol{\beta}'\boldsymbol{x})h_0(t), \tag{5}$$

where h_0 is the baseline function, expressing the individual heterogeneity of each observation. Vector $\boldsymbol{x}$ includes among the explanatory variables the macroeconomic indicators taken for the analysis. The Cox procedure first provides a partial likelihood estimation of unknown parameters β without requiring estimation of the baseline function h_0 and second constructs non-parametric estimates of the baseline hazard using the estimated $exp(-\beta'x)$ (for more details see Pudney, 1989, or Witzany et al., 2010).

4.3 Multi-level Regressions

As an alternative we employ the classical and widely used generalized linear regression model (GLM), where the dependent variable is the (possibly transformed) observed LGD for each facility (for the retail segment) or client (for the corporate segment). The linear regression approach has to deal with effectively closed files only, as it does not account for the time dimension of the collection process. For the retail portfolio, we determined the file as being effectively closed if its work-out period is at least one year after default. For the corporate portfolio, due to the dramatically different scale of loans and legislative procedures for legal entities, the work-out period was prolonged to three years after the default event.

Despite the disadvantage of working only with effectively closed files, the linear regression approach has its advantages. Specifically, multi-level generalization applied to linear regression takes into account different levels of aggregation in the data structure and estimates the standard errors of the regression coefficients in an efficient manner.[7]

As the data set employed for the linear regression model has a different structure than the data set used for the survival analysis, we did not use information about the repayment cash flow over a given (case-dependent) time period as we did in the case of the survival analysis. We used information on collected recoveries only for a specific time point differing for each segment – three years after default for corporate clients and one year after default for retail clients – when we observed most of the defaulted loans being closed. As a result (and unlike in the survival analysis), we had to substantially reduce the data set of defaulted clients. The predictors used for LGD modeling were observed at the time of default.

In order to model LGD we considered different generalized linear regressions of the form:

$$F(LGD_{it}) = \alpha \mathrm{X}_t + \delta \mathrm{Y}_{it} + \varepsilon_{it}, \tag{6}$$

where F is a suitable transformation function chosen in order to deal with the response variable in the range 0 to 1.

Besides the ordinarily used linear regression, we tried different specifications, for instance Fractional Logit Transformation (applied in Dermine and de Carvalho, 2006) and Logit transformation (considered in Bellotti and Crook , 2009). We did not find any notable advantages

[7] For the purposes of brevity we did not perform a multi-level survival analysis (Hox, 2010). This might be the subject of future research.

for the above-mentioned transformations and for the final model we used multi-level linear regression.[8]

The set of explanatory variables X_t contains the macroeconomic factors observed at time t and common to each defaulted client in the cohort t. Y_{it} contains the client's characteristics at time t, and LGD_{it} represents the individual loss rate of a defaulted client i at time t with a different post-default history related to the collection process.

Our assumption is the following: since we have information on each defaulted client at time t and each time cohort obtains a different set of defaulted clients, we adopt the following hierarchical structure for our combined model:

Level I: $LGD_{it} = LGD_{0t} + \beta_t Y_{it} + \varepsilon_{it}$

Level II: $LGD_{0t} = LGD_0 + \alpha_t X_{it} + \alpha_{t-1} X_{it-1} + \ldots + \alpha_{t-p} X_{it-p} + u_{0t}$.

The Level I model expresses the i-th individual's LGD_{it} at time t as a function of the mean LGD specific for each time cohort t, and of an individual set of characteristics Y_{it}. The factor loadings (β_t) can vary by time cohort t. By explicitly modeling in a cohort-specific way, we allowed the mean LGD to vary over time. ε_{it} represents an idiosyncratic individual-level error.

We extended this generalization by assuming that the time-specific mean LGD_{0t} can be explained by a set of macro-level variables X_t. This extension is represented by the Level II equation, where each time cohort LGD_{0t} is described as a function of the average LGD outcome for the whole population (the so-called grand mean) and for a set of macroeconomic factors (with factor loadings α), and u_{0t} represents a time-specific error term. The macro level allows for more lags to capture potential delays in the responses to aggregate fluctuations not captured by Level I variables (see, for example, Jokivuolle and Virén, 2011). Heterogeneity in firms' responsiveness might be caused, for example, by time-to-build effects (Kydland and Prescott, 1982) or industry sensitivity to external finance (Braun and Larrain, 2005), which could translate to a firm's performance at a more general level, i.e., including its default status and subsequent recovery propensity.

The combined model has the following specification:

Combined: $LGD_{it} = LGD_0 + \alpha_t X_{it} + \alpha_{t-1} X_{it-1} + \ldots + \alpha_{t-p} X_{it-p} + \beta_t Y_{it} + \varepsilon_{it} + u_{0t}$. (7)

This model includes a set of estimates for the coefficients of the macro factors common to the whole population, which is the primary focus of our research.[9] Furthermore, it includes a set of coefficients for clients' characteristics and the error term, which helps to account for individual- and group-level variation in estimating group-level regression coefficients. The introduction of this error-term specification controls for heteroskedasticity of errors in the group-structured data and leads to more efficient estimation of unknown coefficients.

The outcome of the multi-level regression allows variance estimation of the combined model as the sum of the variances of the individual-specific errors (σ^2) and the macro-level errors (τ^2). The

[8] The results from the alternative model specifications are available upon request.

[9] Note that the selection procedure might select a lag structure with gaps.

percentage of the observed variation in the dependent variable LGD attributable to macro-level characteristics can be estimated as:

$$\rho = \frac{\tau^2}{\tau^2 + \sigma^2}, \tag{8}$$

where σ^2 and τ^2 were defined above and ρ is referred to as the infraclass correlation coefficient, which can be used as a measure of the magnitude of the influence of the macroeconomic environment on the recovery process.

4.4 Model for Aggregated Data

With respect to the model for aggregated data, we are interested in the links between the average LGD levels for both the corporate and retail segments and the macroeconomic variables at quarterly frequency. The starting point of our analysis is a simple dynamic regression framework with segment-level LGDs as the left-hand side variable and proxies for the real economy as the explanatory variables. As the available sample length – totaling 37 observations at most – is rather limited, our task is to propose a robust and well-specified parsimonious model with reasonable pseudo out-of-sample forecast performance. That is, we will not address potential shifts in LGD-macro relationships following the outbreak of the financial crisis in September 2007. Nonetheless, the framework will allow for possible intercept shifts.

Similarly to the combined model, the resulting specification builds upon the backward-selection method. In particular, it relies upon an automated general-to-specific model-selection algorithm (Gets) as discussed by Doornik (2009). The Gets algorithm is an iterative search procedure allowing for tree search and maintaining model congruency throughout the selection process. The advantage of the Gets approach is that it is not path-dependent like the forward method and a number of other backward-selection methods. At the same time, the algorithm can handle the case of more variables than observations (Hendry et al., 2008). This is not the case with other non-path-dependent approaches such as Bayesian model averaging.

One should note that using the Gets algorithm nonetheless does not imply we resort to mechanistic model building. Our variable pre-selection reflects the variables that have already been employed in the empirical literature and have received extensive justification therein. Furthermore, given that the theoretical relationship between LGD and the macroeconomic environment provides only a partial explanation and remains to be further explored, we admit our limited knowledge and allow for an extended set of variables that, we believe, might relate to the LGD data-generating process and provide a possible alternative perspective (Hendry and Morgan, 1995). These variables include a range of related indicators capturing aggregate supply and demand. Furthermore, we consider proxies for institutional development that might be closely linked to the LGD work-out process (such as government efficiency and rule of law). In this context, our case study might contribute to the discussion on the relevance of a wider range of macroeconomic determinants useful for modeling LGD behavior.

Equation (9) presents the model specification:

$$\Delta LGD_t = \alpha_1 \Delta LGD_{t-1} + \alpha_2 \Delta LGD_{t-2} + \ldots + \alpha_p \Delta LGD_{t-p} + \boldsymbol{\beta_1 Z_t} + \ldots \boldsymbol{\beta_{t-r} Z_{t-r}} + \varphi_b d_{b,t} + \varphi_s d_{s,t} + \varphi_{stp} d_{stp,t} + \varepsilon_t,$$

$t = 1, \ldots, T \quad \varepsilon_t \sim NI(0,\sigma),$ (9)

where ΔLGD_t represents the first difference of our aggregate measure of segment LGD,[10] $\boldsymbol{Z_t}$ stands for an m x 1 vector of exogenous macroeconomic variables,[11] $d_{b,t}$ is a b x 1 vector of permanent blip variables (...,0,0,1,0,0,...) to be estimated by the impulse indicator saturation procedure (Hendry et al., 2008),[12] $d_{s,t}$ is a c x 1 vector of transitory dummy variables (...,0,0,0,1,-1,...), and $d_{stp,t}$ is a vector of step dummies (…,0,0,1,1,1,..) or (…1,1,1,0,0,…).[13] The dummies are included to account for possible misspecification due to omitted structural breaks (e.g., legislative changes, which might influence the work-out process, etc). In particular, the permanent blip dummies should account for possible discrete shifts in the LGD parameter, the transitory dummies for potential one-off discrepancies in, for example, LGD measurement, and the step dummies for trending periods of LGD not accounted for by the macroeconomic determinants under consideration. p and r represent the number of lags for the autoregressive LGD term and exogenous variables, respectively.

5. Results – Retail Portfolio

5.1 Combined Model – Results of Survival Analysis for Retail Portfolio

As described above, the advantage of survival analysis is that it keeps all defaulted credit files, closed and unclosed, allowing for more precise LGD prediction. There are several thousand defaulted retail transactions in our data set of retail clients. This data set covers 123 months of repayment history of defaulted facilities. In this set, 40% of the lifetimes of individual repayments are uncensored and the rest are censored. This implies that no data will be lost from the collected observations and all the information will be used in the estimation procedure: the coefficients of regression will be determined through uncensored observations maximizing the partial maximum likelihood, while an individual baseline hazard function characterizing the individual heterogeneity of each repayment regime is estimated by adding information from the censored observations and maximizing the full maximum likelihood function.

Using the balance amount known at the end of each month, we constructed a pair of output variables for each facility included in the modeling sample: (T_k, d_k), where T_k is the number of months for which a defaulted loan stays in the collection process and d_k takes one of two possible values – 0 if the duration in the collection process was censored, and 1 if the collection process is censored and the final outcome is not known.

To build the model we used the internal explanatory variables collected by the bank. First of all we used a set of individual characteristics of retail clients that proved to affect the duration of the

[10] The aggregate LGD was calculated as the average weighted by the exposure at default.

[11] For the purposes of our study we will not examine the order of integration of the modeled variables unless the estimated system is not stable (Lütkepohl, 2006).

[12] Impulse indicator saturation is a procedure for estimating permanent blip dummies from a complete (“saturated”) set of *N* impulse indicators, i.e., one for every observation. The procedure selects subsets across combinations of indicators, each search path leading to a specific model, followed by searches across the union of these.

[13] The relevant *b* permanent blip variables and *s* mean-shift variables will be determined endogenously.

state of interest. We also included aggregated information about the internal default rate over the retail portfolio due to its potential and plausible link to the collection process. After a backward stepwise procedure, eight internal variables were kept. Our study focuses on macroeconomic drivers of LGD and does not present the influence of client characteristics. The second set of variables is a set of macro indicators that are publicly collected. The results are available in Table 3.

Table 3: Results for the Combined Model Using Survival Analysis of the Retail Portfolio[14]

Parameters	Coefficient	Std.Error	Wald Z	p-value
Unemployment_rate_region_Q	-0,016	0,003	27,4	0
PRIBOR_Year_Q	0,19	0,015	156,7	0
NPL_household_Q	-0,061	0,009	40,8	0
Real_Consumption_Growth_Q	0,046	0,004	167,1	0
Real_GDP_Growth_Q	0,008	0,002	11,4	0,001
Nom_Wage_Growth_Q	0,01	0,003	10,5	0,001

Source: Authors' calculations

Two of the six macroeconomic variables – the unemployment rate and real consumption growth – are closely related to the variables identified by Caselli et al. (2008) in their study on Italian retail data. This result supports a more pronounced influence of these variables for the explanation of LGD fluctuations at the aggregate level. However, the other four macro indicators proved to be relevant only for the Czech retail segment.

The marginal effects of the macro variables introduced have the expected direction: the coefficients on the unemployment rate and the ratio of non-performing loans to GDP are negative, which means that an increase in these macroeconomic factors implies a lower probability of paying off at least 90% of the debt obligation at time *t*. On the other hand, the positive coefficient values for the 1-year Pribor, real consumption growth, real GDP growth, and nominal wage growth indicate a higher probability of paying off at least 90% of the debt obligation at time *t* following a rise in any of the above-mentioned factors, in line with intuition.

The negative link between LGD and the 1-year Pribor might be related to the effect mentioned by Dell'Ariccia and Marquez (2006), who suggest that lower interest rates reduce financing costs and might therefore motivate banks to perform less thorough checks of the credit quality of their debtors. Lower interest rates might thus lead to lending to lower credit quality clients, leading to lower recovery in the event of default and consequently higher LGD. Although Geršl et al. (2012) do not find evidence of excessive risk-taking by banks in the Czech Republic in a low-interest rate environment, their methodology might not capture our results for the retail portfolio, as they monitor risky lending to legal persons only, since their study is based on the Central Credit Register operated by the Czech National Bank.

5.2 Combined Model – Results of Multi-level Linear Regression for Retail Portfolio

To build the multi-level linear model, we constructed a data set of LGD outcomes for retail clients collected between 2002 and 2010. Likewise, we collected macroeconomic variables at quarterly

[14] The individual variables included in the regression are not reported in this paper for confidentiality reasons.

frequency and individual client characteristics at the moment of default. The macroeconomic variables observed for each defaulted case at the moment of default played a role in the explanation of recovery after one year after default; a natural time lag was therefore present between the dependent variable (LGD) and the macroeconomic variables. Additionally, the potential relevance of macroeconomic variables with longer lags was considered in the model-building procedure, but they were not included in the final model since their incorporation caused insignificancy of some important individual variables that were obviously linked with LGD.

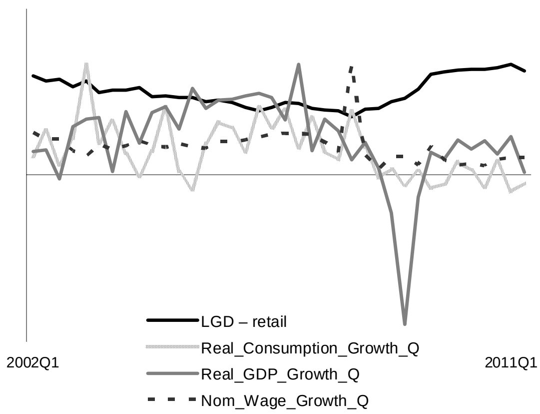
The results of the multi-level linear regression model, where we kept the input of lagged macroeconomic variables as parsimonious as possible, are presented in Table 4. The final specification includes the real GDP growth rate, nominal wage growth, and real consumption growth observed at the moment of default as significant macroeconomic indicators determining the LGD outcome of each individual case. These results differ from the results obtained by Caselli et al. (2007) and also slightly depart from the output of our duration model. Conditional on the results of the multi-level model, both micro models seem to support simultaneously the finding that the LGD for the retail portfolio is negatively related to GDP growth, nominal wage growth, and consumption growth.

Table 4: Results for the Combined Model Using the Multi-level Model for the Retail Portfolio[15]

Parameter	Coefficient	Std. Error	Wald Z	p-value
Real_GDP_Growth_Q	-0.016	0.005		0.007
Nom_Wage_Growth_Q	-0.014	0.007		0.069
Real_Consumption_Growth_Q	-0.014	0.008		0.084
Covariance Parameters:				
Residual	0.155	0.001	110.1	0
Intercept [subject=quarter] Variance	0.012	0.004	3.2	0.002

Source: Authors' calculations

Figure 1: Normalized Time Series



Analysis of the covariance parameters indicates that the Level I (i.e., ε_{it}) residual variance component totals σ^2 (individual-specific errors) = 0.15, while the Level II variance (i.e., u_{0t}) accounting for between-group variation of LGD corresponds to τ^2 (macro-level error) = 0.01, The two variance components can be used to partition the variance across levels according to equation (8). The intraclass correlation coefficient for the retail portfolio is equal to 0.012/(0.012 + 0.155) = 0.0719, which represents 7.2% of the total variation.

As a measure of model performance we used the ordinal power indicator, which is an extension of the Gini coefficient devoted to cases of a non-binary dependent variable. The survival regression has an ordinal power of 34.9%. The multi-level regression has a power of 44% given the lower number of macroeconomic variables included in the final model.

[15] The individual variables included in the regression are not reported in this paper for confidentiality reasons.

5.3 Aggregated Model – Results for Retail Portfolio

The main purpose of the study is to evaluate the links between macroeconomic aggregates and LGD and to compare the relative performance of models using aggregate and micro-level data. This subsection estimates the potential links between macro covariates and LGD at the aggregate level. We start with the estimation using the Gets selection procedure described above. For exposition purposes, we present an additional model using macro covariates from the combined model's multi-level linear regression and discuss the corresponding model performance.

The results can be found in Table 5, which presents the model parameter estimates for the retail segment over the period 2004q4–2011q2.[16] The LGD variable was transformed into first differences to dispose of breaks observed in the original data. The remaining irregularities were identified by the Gets procedure in combination with the dummy saturation procedure (Doornik, 2009) and are located in the first half of the sample, spanning from 2002q2 to 2011q2. Note that the ultimate model selected by the Gets procedure contains variables above the 5% significance cut-off threshold (3M Pribor(t-1)) to preserve model congruency. The results obtained from the trimmed sample starting in 2004q4 provide both qualitatively and quantitatively very similar results.[17]

Table 5: Output for the Retail Portfolio on Aggregated Data

Aggregated model - macro variables using gets algorithm		Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant		0,042	0,013	3,190	0,004	0,316
3M Pribor (t-1)		-0,013	0,007	-1,850	0,078	0,135
3M Pribor (t-6)		-0,019	0,006	-3,090	0,005	0,303
Nominal cons. growth YoY(t)		-0,018	0,004	-5,000	0,000	0,532
Nominal cons. growth YoY(t-1)		0,014	0,004	3,820	0,001	0,399
sigma	0,026		RSS	0,014		
R^2	0,566		F(7,20) =	7.17 [0.001]**		
Adj.R^2	0,487		log-L	63,438		
N	27		no. of pars	5		
mean(Y)	0,007		se(Y)	0,036		
Aggregated model - macro variables from combined model		**Coefficient**	**Std.Error**	**t-value**	**t-prob**	**Part.R^2**
Real cons. Growth QoQ(t)		0,000	0,002	-0,046	0,964	0,000
Real GDP growth QoQ(t)		0,003	0,001	1,940	0,066	0,146
Nominal wage growth QoQ(t)		0,000	0,002	-0,245	0,809	0,003
step dummy 2008q2		-0,083	0,020	-4,150	0,000	0,439
step dummy 2009q4		0,059	0,012	4,890	0,000	0,521
sigma	0,026		RSS	0,014		
R^2	0,580		F(5,22) =	6.08 [0.001]**		
Adj.R^2	0,485		log-L	63,410		
N	27		no. of pars	5		
mean(Y)	0,007		se(Y)	0,036		

Source: Authors' calculations

[16] The results refer to aggregate LGD calculated as a simple average. The output for the weighted average remains practically unchanged and can be provided upon request.

[17] The results from the complete sample including all the dummies picked by the selection procedure are listed in Table A3 in the Appendix.

The aggregate LGD in the retail segment is correlated in particular with the real consumption growth rates (y-o-y) and 3M interest rates on the Czech interbank market.[18] Our results thus partly conform to those of Caselli et al. (2008), who found household consumption to matter for the household segment. On the other hand, the effect of y-o-y nominal consumption growth on LGD is rather short-lived, as the cumulative effect of a change in nominal consumption remains statistically not different from zero. The signs of the 3M Pribor rates are negative, implying a negative cumulative impact of a rise in the 3M Pribor on aggregate LGD. This is in line with the positive sign of the 1Y Pribor (ignoring the different maturity), suggesting a higher probability of paying off at least 90% of the debt obligation at time t, as was found in the survival analysis model (Table 1).

Unlike in the combined model, however, the Gets algorithm in the aggregate-level ARMAX estimation framework did not select real GDP, nominal wage growth, the unemployment rate, or the household NPL ratio as statistically relevant correlates with the aggregate LGD. Furthermore, and perhaps not surprisingly, when the macro-level variables selected within frameworks based upon microeconomic data are put into the ARMAX framework, they tend to perform rather poorly. This inconsistency of different explanatory variables based on micro- and aggregated-level (i.e., macro) data is common and might lead to different results (see, for example, Altissimo et al., 2007, or Horváth et al., 2009, for the case of aggregation bias of inflation time series). The only variable avoiding low statistical significance and negligible partial R^2 is real GDP growth, which, however, has the opposite sign to its multi-level linear regression model counterpart.[19]

The lower part of Table 5 presents the parameter estimates for the aggregated model employing macroeconomic variables identified by the combined model. The presented output represents the best model (in terms of model congruency and model fit) that includes the macroeconomic variables from the combined model estimated by multi-level regression. The first observation is that the model employing macro variables from the combined setup, even after allowing for a richer lag structure and/or the dummy saturation procedure, performed notably worse. Comparing the present output with the aggregated Gets model in terms of the variation explained by macroeconomic factors, the adjusted R^2 of the two models reveal relatively small differences in adjusted R^2 for the trimmed sample as well as in the residual sum of squares. Nonetheless, the results for the aggregate model using the variables from the combined framework are driven mainly by adjustments captured by two step dummies (1,1,1,0,0,…) over the period 2008q2–2009q4, rather than by the macro variables themselves.

The conditioning information from the individual level thus plays an essential part in the choice of macroeconomic aggregates for modeling the aggregate LGD. In this respect combined models, even though potentially better at explaining the total variation in LGD, might not bring additional clarity into the relationship between the aggregate LGD and macroeconomic factors when confronted with time-series models working exclusively with aggregate data. This result is partly at odds with Jacobson et al. (2011), who state that the combined micro-macro framework is

[18] We did not find the autoregressive and moving average components of the model to matter in our LGD model.
[19] All the specification tests for the two aggregate models were satisfied for both the trimmed and full sample , maintaining model congruency. The only exception is the autoregressive test at the 5% level for Model 2 over 2002q2–2011q2, which might be a result of the model's lack of dynamic structure in combination with higher volatility in the sample over 2002–2004.

superior both to models that exclude macro information and to the best-fitting standard time-series models.

6. Results – Corporate Portfolio[20]

6.1 Combined Model – Results of Multi-level Linear Regression for Corporate Portfolio

The modeling sample for the corporate portfolio consists of defaults observed from 2002 until 2009. The dependent variable is the LGD rate calculated for each client. The indicator of quarter is used as the grouping variable. Inside each quarter we considered clients that defaulted in this quarter together with client and credit product characteristics. These individual-level characteristics were used in the Level I equation. The Level II regression employed macroeconomic factors kept unchanged for each quarter and common to each individual case used in each quarter – see equation (6). The time lag of the macroeconomic factors used as explanatory variables for the individually observed LGDs for each corporate client was naturally incorporated into the model: the macroeconomic factors were observed at the time of default and were used to explain differences in recoveries after the work-out process has been completed. Later lags did not prove significant in the model.

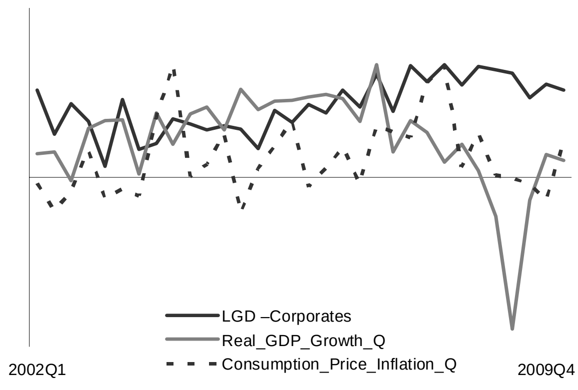
When building the model, we also tested for statistical differences between the models for different subportfolios of the corporate portfolio – the corporate sub-portfolio (CORP) and the SME portfolio. We were not able to estimate a separate model for corporate clients due to the very limited number of observations. Nevertheless, except for the corporate dummy in the pooled regression, which did not prove to be significant, all the remaining variables – both internal and macroeconomic – remained statistically significant.

Table 6: Output of the Multi-level Model for the Corporate Portfolio

Parameter	Coefficient	Std. Error	t-value	Wald Z	p-value
Real_GDP_Growth_Q	-0.012	0.006	-1.852		0.080
Consumption_Price_Inflation_Q	0.020	0.010	2.04		0.052
Covariance Parameters					
Residual	0.142	0.004		###	0.000
Intercept [subject=quarter] Variance	0.020	0.007		2.709	0.007

Source: Authors' calculations

Figure 2: Normalized Time Series



The results for the corporate segment of the multi-level linear regression include real GDP growth and consumption price inflation. This result is different from that for the retail portfolio in our

[20] As mentioned in the methodology section, the survival method is not performed for the corporate portfolio due to data limitations: dynamic changes in client recoveries for historical observations were not collected. As a result, it is not possible to construct the duration variable.

study, indicating that for retail recoveries different macro drivers might be important as compared to the corporate portfolio.

Selection among the macroeconomic variables eventually keeps the most significant factors (real GDP growth and consumption price inflation) in the final model.[21] Real GDP growth appears in both portfolios as a significant control for the macroeconomic situation. The study of the SME sub-segment by Caselli et al. (2007) also confirms the importance of GDP growth.

Analysis of the covariance parameters indicates that the Level I residual variance component induced by the individual variables was $\sigma^2 = 0.14$. The Level II variance quantifies the between-group variation of the LGD means ($\tau^2 = 0.02$), which makes up 12.3% of the total variation.

The ordinal power for the linear regression with macroeconomic variables is equal to 46.7%, while the multi-level regression ordinal power is 55.9%. The corporate models perform better than the retail models, as the individual explanatory characteristics included information about realized collaterals, which are important drivers of the collection process.

6.2 Aggregated Model – Results for Corporate Portfolio

Similarly to the case of the retail portfolio, we estimate (using the Gets selection procedure described in the Methodology section) a time series ARMAX model with aggregate data and compare it with the aggregated model employing macroeconomic variables identified by the combined model. We also discuss the corresponding model performance.

Table 7 presents the model parameter estimates for the corporate segment over the period 2004q4–2011q2.[22] The LGD variable was again transformed into first differences to dispose of breaks observed in the original data, where the remaining irregularities were identified by the Gets procedure in combination with the dummy saturation procedure. The model-selected dummies are spread more equally over the sample period; nonetheless, the major cluster of dummies is again located in the first half of the sample (2003q1 to 2010q4). The results obtained from the trimmed sample starting in 2004q4 provide, similarly to the retail segment, both qualitatively and quantitatively very similar results to the multi-level model combining both micro and macro-level data.[23]

[21] Other macroeconomic indicators which seemed relevant in the classical (i.e., non-hierarchical) regression were excluded due to their lack of significance within the multi-level estimation framework. This was induced by the fact that linear regression does not account for the hierarchical structure of the data used for the modeling. Details of the benchmark linear regression can be found in Table A2 in the Appendix.

[22] Similarly to the case of the retail portfolio, the results refer to aggregate LGD calculated as a simple average. The output for the weighted average preserves the qualitative conclusions and can be provided upon request.

[23] The results from the complete sample including all the dummies picked by the selection procedure are listed in Table A4 in the Appendix.

Table 7: Output for the Corporate Portfolio on Aggregated Data

Aggregated model - macro variables using gets algorithm		Coefficient	Std.Error	t-value	t-prob	Part.R^2
Real GDP growth YoY (t-2)		0,017	0,010	1,710	0,106	0,146
Real GDP growth YoY (t-3)		-0,018	0,010	-1,800	0,089	0,161
Investment price inflation QoQ(t)		0,010	0,002	4,530	0,000	0,547
blip dummy 2006q4		0,214	0,082	2,630	0,018	0,289
blip dummy 2005q3		-0,141	0,078	-1,810	0,088	0,161
sigma	0,074		RSS	0,092		
R^2	0,668		F(,) =			
Adj. R^2	0,581		log-L	34,576		
N	25		no. of pars	5		
mean(Y)	0,012		se(Y)	0,099		
Aggregated model - macro variables from combined model		**Coefficient**	**Std.Error**	**t-value**	**t-prob**	**Part.R^2**
Real GDP growth QoQ(t)		0,005	0,004	1,220	0,239	0,081
Consumption price inflation		-0,006	0,007	-0,906	0,378	0,046
sigma	0,094		RSS	0,151		
R^2	0,356		F(7,25) =			
Adj. R^2	0,129		log-L	26,799		
N	25,000		no. of pars	3,000		
mean(Y)	0,011		se(Y)	0,101		

Source: Authors' calculations

The aggregate LGD work-out in the corporate segment is correlated in particular with the lags of the real GDP growth rate (y-o-y) and investment price inflation (q-o-q).[24] The corporate segment thus reports an explanatory variable that resembles the output from the multi-level linear regression, namely, the q-o-q real GDP growth rate, even though the marginal cumulative impact is not statistically different from zero, as opposed to the negative sign in the multi-level equation. On the other hand, the Gets procedure did not select q-o-q consumption price inflation and instead opted for inflation in q-o-q investment prices. However, the qualitative difference between these two inflation definitions is minor. Furthermore, we experimented with consumption price inflation in the model and obtained impaired model performance in terms of both fit and model congruency. Nevertheless, the sign on consumption price inflation is positive, as is the sign on consumer price inflation in the multi-level model.

The lower part of Table 7 presents the model using the macro-level variables selected within the multi-level/combined framework. Similarly to the case of the retail portfolio, the variables tend to perform poorly within the ARMAX framework. None of the macro covariates preserves statistical significance and maintains a negligible partial R^2.[25]

Comparing the models in terms of the variation explained by macroeconomic factors, there is a large difference in the adjusted R^2 for the trimmed sample as well as in the residual sum of squares. While the partial R^2 measure of the listed dummies is non-negligible, the macroeconomic covariates retain substantial explanatory power and our results are thus not a mere relict of

[24] We did not find the autoregressive and moving average components of the model to matter in our LGD model.

[25] All the specification tests for the two aggregate models were satisfied for both the trimmed and full sample , maintaining model congruency. The only exception is the test of normality, which was rejected at the 5% level for the Gets model over 2003q1–2010q4, which, as for the retail model, might be the result of the model's lack of dynamic structure in combination with higher volatility in the sample over 2002–2004.

possible model overfitting. On the other hand, the partial R^2s for the explanatory variables in the aggregated model using combined inputs are relatively modest.[26] The relevance of the conditioning information from the individual level thus confirms to our previous discussion on the retail segment.

7. Conclusions

This paper analyzes the determinants of work-out loss given default (LGD) with respect to macroeconomic factors of the Czech economy. We approached the topic from two different perspectives. Firstly, we specify a microeconometric ("combined") model for explaining the LGD of a portfolio. However, in addition to individual client characteristics, macroeconomic variables are employed in the final model specification. Secondly, pure time-series models using aggregated macro data are estimated using a general-to-specific approach, and the results are confronted with those obtained from the combined approach.

The comparison of the models described above is aimed at revealing several aspects of their performance, such as the information area, technical aspects, and the persistence of the outcome.

Regarding the first aspect, the information area used in the combined models – survival and multi-level – is much wider than in the case of the aggregated model. Combined models not only take into account factors influencing the common trend in LGD performance over time, but also rely on individual-level information, which tends to enrich the selection of important macroeconomic variables. Moreover, even among the combined methods different information areas were explored: in the case of survival analysis the whole set of defaulted cases were taken into account, so that even unfinished recovery processes affected the output.

Regarding the technical aspects employed in the paper, the methods applied were aimed at exploring an appropriate econometric technique to the full extent in order to disclose as much as possible about our subject of interest. As stated in the survival analysis, the dynamic structure of the recovery process affects the output. On the other hand, another combined method – multi-level regression – has the advantage of proper treatment of the error structure. As a result of the application of survival analysis, a wider spectrum of relevant macroeconomic variables was selected than in previous studies (for instance, Bellotti et al., 2009, applied simple linear regression and found only one relevant macroeconomic factor – the unemployment rate for the retail portfolio). The next logical step for future research is to explore the advantages of both methods and analyze the multi-level survival approach, which will definitely retain the advantages of both the pure survival method and pure multi-level regression. The most advanced technique was also explored in the case of the aggregated models. Our approach controlled for a possible time trend or possible structural breaks present in the dynamics of the aggregated values by considering the first differences of the quarterly observed aggregated LGDs and applying a sophisticated analytical procedure (the Gets approach), which, to the best of our knowledge, has not previously been applied in the context of our study (for instance, Caselli et al., 2007, studied

[26] In the same manner as for the retail segment, this model represents the best model that includes the macroeconomic variables from multi-level regression using micro data in terms of model congruency and model fit. Models with a richer lag structure and/or without the dummy saturation procedure performed notably worse.

the influence of macroeconomic factors without treatment of the possible unit root present in the time series).

The last aspect of the methods studied in this paper was concerned with the persistence of the output born by their application. As mentioned earlier, the survival approach retained the most significant macroeconomic variables. However, multi-level regression makes the selection of relevant macro factors more scrupulous: for instance, of the six variables kept by the survival approach in the retail portfolio it retained only the three most important ones. The most persistent macroeconomic variables were selected by the aggregated method: nominal consumption growth in the case of the retail portfolio and real GDP growth in the case of the corporate portfolio.

The choice of method crucially depends on the availability of information and the appropriate technique and the degree of generality required for the results obtained.

Despite the relatively low importance of macro variables in the model combining micro- and macroeconomic information, our estimates suggest that the macroeconomic environment contributes directly to the variation in LGD. Furthermore, the results obtained from both the combined and aggregated models point to notable differences in the suggested macro links obtained from the different estimation frameworks. Firstly, our micro and macro perspectives identify jointly only a single macroeconomic domain within each of the two segments concerned. While for retail customers the models indicated consumption-related factors as significant determinants, for corporate clients the major driving factor is real GDP growth. Other potential candidates do not enter robustly either the combined model or the aggregated model. Still, our study partly -confirms to the results from previous literature, as Caselli et al. (2008) also identified consumption among the macroeconomic factors influencing retail LGD, and GDP growth among those influencing the SME LGD for Italian banking data.

The results of our study should help us gain more detailed information about the link between the banking sector LGD and the business cycle of the Czech economy. Nonetheless, it is important to note that the differences in the relevant macroeconomic factors derived from a purely macro-based as compared to combined framework might have substantial implications, not least for the conduct of top-down solvency stress tests performed by regulatory authorities. These rely typically on a battery of so-called satellite models linking macroeconomic developments to the financial sector. If micro-level information, such as client balance sheet data, is missing, estimates obtained exclusively from macro data might paint a rather different picture than more richly specified microeconometric models. Even in cases where macro-based and combined satellite models provide qualitatively similar predictions (e.g., given that macro variables from the aggregate model might approximate the confounded interaction of macro and micro variables in the combined model), the impact of a given macroeconomic stress scenario will become critically dependent on the consistency of the macro variables in the two models.

References

ALTISSIMO, F., B. MOJON, AND P. ZAFFARONI (2007): "Fast Micro and Slow Macro: Can Aggregation Explain the Persistence of Inflation?" European Central Bank Working Paper No. 729.

ALTMAN E., B. BRADY, A. RESTI, AND A. SIRONI (2005a): "The Link between Default and Recovery Rates: Theory, Empirical Evidence, and Implications." *Journal of Business* 78(6), pp. 2203–2228.

ALTMAN, E., B. BRADY, A. RESTI, AND A. SIRONI (2005b): "The PD/LGD Link: Empirical Evidence from the Bond Market." In: Altman E, Resti A, Sironi A (eds) *Recovery Risk*, Riskbooks, London, pp. 217–233.

ALTMAN, E., A. RESTI, AND A. SIRONI (2002): "The Link Between Default and Recovery Rates: Effects on the Procyclicality of Regulatory Capital Ratios." BIS Working Papers, July, Issue 113.

ALTMAN, E., A. RESTI, AND A. SIRONI (2005c): "The PD/LGD Link: Implications for Credit Risk Modelling." In: Altman E, Resti A, Sironi A (eds) *Recovery Risk*, Riskbooks, London, pp. 253–266.

BCBS (2006): "Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework – Comprehensive Version." Basel Committee on Banking Supervision, BIS, June 2006.

BASTOS, J. A. (2010): "Forecasting Bank Loans Loss-Given-Default." *Journal of Banking & Finance* 34(10), pp. 2510–2517.

BELLOTTI, T. AND J. CROOK (2009): "Calculating LGD for Credit Cards." QFRMC Conference on Risk Management in the Personal Financial Services Sector, January 2009.

BRAUN, M. AND B. LARRAIN (2005): "Finance and the Business Cycle: International, Inter-Industry Evidence." *Journal of Finance* 60(3), pp. 1097–1128.

CALABRESE, R. AND M. ZENGA (2010): "Bank Loan Recovery Rates: Measuring and Nonparametric Density Estimation." *Journal of Banking & Finance* 34, pp. 903–911.

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, pp. 1–34.

CHALUPKA, R. AND J. KOPECSNI (2009): "Modeling Bank Loan LGD of Corporate and SME Segments: A Case Study." *Czech Journal of Economics and Finance* 59(4), pp. 360–382.

COX, D. (1972): "Analysis of Survival Data." *Journal of the Royal Statistical Society*, Series B, 34, pp. 89–110.

DE GRAEVE, F., T. KICK, AND M. KOETTER (2008): "Monetary Policy and Financial (In)stability: An Integrated Micro–Macro Approach." *Journal of Financial Stability* 4(3), pp. 205–231.

DELL'ARICCIA, G. AND R. MARQUEZ (2006): "Lending Booms and Lending Standards." *Journal of Finance* 61(5), pp. 2511–2546.

DERMINE, J. AND C. NETO DE CARVALHO (2006): "Bank Loan Losses-Given-Default: A Case Study." *Journal of Banking and Finance* 30(4), pp. 1219–1243.

DOORNIK, J. A. (2009): *Autometrics*, in: Castle J L, Shephard N (eds) *The Methodology and Practice of Econometrics: Festschrift in Honour of David F. Hendry*, Oxford: Oxford University Press.

DÜLLMANN, K. AND M. TRAPP (2004): "Systematic Risk in Recovery Rates – An Empirical Analysis of US Corporate Credit Exposures." Deutsche Bundesbank Discussion Paper, Series 2, Banking and Financial Supervision, No 02/2004.

EMERY, K., R. CANTOR, AND R. ARNER (2004): "Recovery Rates on North American Syndicated Bank Loans, 1989–2003." Moody's special comment.

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, pp. 187–200.

GERŠL, A. AND J. SEIDLER (2012): "How to Improve the Quality of Stress Tests Through Backtesting." *Czech Journal of Economics and Finance* 62(4), pp. 325–346.

GERŠL, A., P. JAKUBÍK, D. KOWALCZYK, S. ONGENA, AND J. L. PEYDRÓ-ALCALDE (2012): "Monetary Conditions and Banks' Behaviour in the Czech Republic." CNB Working Paper No 2/2012.

GREENE, H. W. (2000): *Econometric Analysis (International Edition)*, Pearson US Imports & PHIPEs, fourth edition.

GRIPPA, P., S. IANNOTTI, AND F. LEANDRI (2005): "Recovery Rates in the Banking Industry: Stylised Facts Emerging from the Italian Experience." In: Altman E, Resti A, Sironi A (eds) *Recovery Risk*, Riskbooks, London.

GRUNERT, J. AND M. WEBER (2009): "Recovery Rates of Commercial Lending: Empirical Evidence for German Companies." *Journal of Banking & Finance* 33, pp. 505–513.

HAN, C. AND Y. JANG (2013): "Effects of Debt Collection Practices on Loss Given Default." *Journal of Banking & Finance* 37, pp. 21–31.

HENDRY, D. F., S. JOHANSEN, AND C. SANTOS (2008): "Automated Selection of Indicators in a Fully Saturated Regression." *Computational Statistics* 33, pp. 317–335.

HENDRY, D. F. AND M. S. MORGAN (1995): *The Foundations of Econometric Analysis*, Cambridge.

HOGGARTH, G., S. SORENSEN, AND L. ZICCHINO (2005): "Stress Tests of UK Banks Using a VAR Approach." Bank of England Working Paper No 282.

HORVÁTH, R., F. CORICELLI, AND J. BABECKÝ J (2009): "Assessing Inflation Persistence: Micro Evidence on an Inflation Targeting Economy." *Czech Journal of Economics and Finance* 59(2), pp. 102–127.

HOX, J. (2010): *Multilevel Analysis: Techniques and Applications*, 2nd ed., Routledge, New York.

JACOBSON, T., R. KINDELL, J. LINDÉ, AND K. ROSZBACH (2011): "Firm Default and Aggregate Fluctuations." International Finance Discussion Papers No 1029, Board of Governors of the Federal Reserve System.

JIMENEZ, G. AND J. SAURINA (2006): "Credit Cycles, Credit Risk, and Prudential Regulation." *International Journal of Central Banking* 2(2), pp. 65–98.

JOKIVUOLLE, E. AND M. VIRÉN (2011): "Cyclical Default and Recovery in Stress Testing Loan Losses." *Journal of Financial Stability*, available online.

KYDLAND, F. E. AND E. C. PRESCOTT (1982): "Time to Build and Aggregate Fluctuations." *Econometrica* 50(6), pp. 1345–1370.

KOCENDA, E. AND M. VOJTEK (2011): "Default Predictors in Retail Credit Scoring: Evidence from Czech Banking Data." William Davidson Institute Working Papers Series wp1015, William Davidson Institute at the University of Michigan.

LÜTKEPOHL, H. (2006): *New Introduction to Multiple Time Series Analysis*, 2nd ed., New York: Springer.

PESOLA, J. (2007) "Financial Fragility, Macroeconomic Shocks and Banks' Loan Losses: Evidence from Europe." Bank of Finland Research Discussion Papers No 15.

PODPIERA, R. (2004): "Does Compliance with Basel Core Principles Bring Any Measurable Benefits?" IMF Working Paper WP/04/204.

PUDNEY, S. (1989): *Modelling Individual Choice*, Basil Blackwell.

QI, M. AND X. ZHAO (2011): "Comparison of Modeling Methods for Loss Given Default." Journal of Banking & Finance 35, pp. 2842–2855.

QUAGLIARIELLO, M. (2007): "Banks' Riskiness over the Business Cycle: A Panel Analysis on Italian Intermediaries." *Applied Financial Economics* 17(2), pp. 119–138.

SEIDLER, J., R. HORVATH, AND P. JAKUBÍK (2009): "Estimating Expected Loss Given Default in an Emerging Market: The Case of Czech Republic." *Journal of Financial Transformation* 27, pp. 103–107.

WITZANY, J., M. RYCHNOVSKÝ, AND P. CHARAMZA (2010): "Survival Analysis in LGD Modeling." IES Working Paper 2010/02, Institute of Economic Studies, Faculty of Social Sciences, Charles University Prague.

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(1), pp. 204–215.

Appendix

Table A1: Summary Statistics of the Variables Selected in the Combined and Aggregated Models

Retail portfolio

Variable	Obs	Mean	Std. Dev.	Min	Max
Delta retail discounted LGD	37	0.001	0.038	-0.080	0.107
Unemployment_rate_region_Q	37	8.596	1.331	5.667	10.034
PRIBOR_Year_Q	37	2.739	0.779	1.777	4.320
NPL_household_Q	37	4.538	1.472	2.870	8.231
Real_Consumption_Growth_Q	37	2.822	3.378	-1.860	12.597
Real_GDP_Growth_Q	37	3.325	4.091	-13.646	10.173
Nom_Wage_Growth_Q	37	5.393	3.243	1.183	20.950
3M Pribor	37	0.828	1.006	-1.225	2.872
Nom. cons. growth YoY	37	4.665	2.997	-0.901	10.797

Corporate portfolio

Variable	Obs	Mean	Std, Dev,	Min	Max
Delta corporate discounted LGD	32	0.008	0.146	-0.332	0.442
Real_GDP_Growth_Q	32	3.523	4.313	-13.646	10.173
Consumption_Price_Inflation_Q	32	1.967	3.010	-2.884	9.247
Real_GDP_Growth_YoY	32	3.490	3.436	-4.792	7.447
Investment price inflation QoQ	32	0.895	7.465	-10.707	16.082

Source: ARAD CNB, private database

Table A2: Outcome of the Linear Regression Model for the Corporate Portfolio

	Coefficient	Std.Error	Wald Z	p-value
Real_GDP_Growth_Q	-0,010	0,002	-6,0	0
Real_Investment_Growth_Q	0,002	0,000	6,0	0
Consumption_Price_Inflation_Q	0,012	0,003	4,5	0
Household_consumption_expend_perc	-0,034	0,009	-3,8	0
rule_of_law	0,965	0,177	5,5	0

Table A3: Retail Portfolio, Dependent Variable Delta Discounted LGD, 2002(2)–2011(2)

Aggregated model - macro variables using gets algorithm		Coefficient	Std.Error	t-value	t-prob	Part.R^2
Constant		0,049	0,013	3,740	0,001	0,342
3M Pribor (t-1)		-0,022	0,006	-3,610	0,001	0,326
3M Pribor (t-6)		-0,023	0,006	-3,730	0,001	0,341
Nominal cons. growth YoY(t)		-0,019	0,004	-5,150	0,000	0,496
Nominal cons. growth YoY(t-1)		0,014	0,004	3,680	0,001	0,334
blip dummy 2003q1		0,097	0,030	3,220	0,003	0,277
blip dummy 2003q3		0,078	0,030	2,560	0,016	0,195
blip dummy 2003q4		0,073	0,031	2,380	0,025	0,173
blip dummy 2004q1		0,073	0,031	2,400	0,023	0,176
blip dummy 2004q3		0,090	0,032	2,820	0,009	0,227
sigma	0,028		RSS	0,020		
R^2	0,609		F(10,27) =	4.68 [0.001]**		
Adj.R^2	0,479		log-L	86,284		
N	37		no. of pars	10		
mean(Y)	0,001		se(Y)	0,038		

Aggregated model - macro variables from combined model		Coefficient	Std.Error	t-value	t-prob	Part.R^2
Real cons. Growth QoQ(t)		0,002	0,002	0,792	0,434	0,019
Real GDP growth QoQ(t)		0,002	0,002	1,050	0,300	0,034
Nominal wage growth QoQ(t)		-0,001	0,002	-0,473	0,639	0,007
step dummy 2008q2		-0,079	0,021	-3,860	0,001	0,317
step dummy 2009q4		0,057	0,014	3,980	0,000	0,331
sigma	0,031		RSS	0,031		
R^2	0,407		F(5,32) =	6.08 [0.001]**		
Adj.R^2	0,315		log-L	78,562		
N	37		no. of pars	5		
mean(Y)	0,001		se(Y)	0,038		

Table A4: Corporate Portfolio, Dependent Variable Delta Discounted LGD, 2003(1)–2010(4)

Aggregated model - macro variables using gets algorithm		Coefficient	Std.Error	t-value	t-prob	Part.R^2
Real GDP growth YoY (t-2)		0,021	0,010	2,100	0,047	0,155
Real GDP growth YoY (t-3)		-0,021	0,010	-2,130	0,044	0,159
Investment price inflation		0,009	0,002	4,390	0,000	0,445
blip dummy 2006q4		0,195	0,080	2,440	0,023	0,198
blip dummy 2005q3		-0,138	0,077	-1,790	0,087	0,117
step dummy 2003q2		-0,812	0,110	-7,410	0,000	0,696
step dummy 2003q3		0,878	0,107	8,230	0,000	0,738
step dummy 2003q4		-0,412	0,075	-5,490	0,000	0,556
sigma	0,073		RSS	0,128		
R^2	0,807		F(8,24) =	12,27[0.000]**		
Adj. R^2	0,742		log-L	42,885		
N	32		no. of pars	8		
mean(Y)	0,008		se(Y)	0,146		

Aggregated model - macro variables from combined model		Coefficient	Std.Error	t-value	t-prob	Part.R^2
Real GDP growth QoQ(t)		0,004	0,004	1,020	0,319	0,040
Consumption price inflation		0,000	0,006	0,065	0,949	0,000
Constant		0,342	0,102	3,720	0,001	0,356
step dummy 2002q3		-0,317	0,140	-3,510	0,002	0,331
step dummy 2003q2		0,739	0,141	5,980	0,000	0,589
step dummy 2003q3		-0,755	0,104	-6,050	0,000	0,594
sigma	0,087		RSS	0,191		
R^2	0,713		F(6,26) =	10.33 [0.000]**		
Adj. R^2	0,644		log-L	36,555		
N	32		no. of pars	6		
mean(Y)	0,008		se(Y)	0,146		

CNB WORKING PAPER SERIES

12/2012	Konstantin Belyaev Aelita Belyaeva Tomáš Konečný Jakub Seidler Martin Vojtek	*Macroeconomic factors as drivers of LGD prediction: Empirical evidence from the Czech Republic*
11/2012	Adam Geršl Petr Jakubík Tomáš Konečný Jakub Seidler	*Dynamic stress testing: The framework for testing banking sector resilience used by the Czech National Bank*
10/2012	Tomáš Havránek Marek Rusnák	*Transmission lags of monetary policy: A meta-analysis*
9/2012	Volha Audzei František Brázdik	*Monetary policy and exchange rate dynamics: The exchange rate as a shock absorber*
8/2012	Alexis Derviz Jakub Seidler	*Coordination incentives in cross-border macroprudential regulation*
7/2012	Peter Claeys Bořek Vašíček	*Measuring sovereign bond spillover in Europe and the impact of rating news*
6/2012	Michal Franta Jan Libich Petr Stehlík	*Tracking monetary-fiscal interactions across time and space*
5/2012	Roman Horváth Jakub Seidler Laurent Weill	*Bank capital and liquidity creation: Granger causality evidence*
4/2012	Jaromír Baxa Miroslav Plašil Bořek Vašíček	*Changes in inflation dynamics under inflation targeting? Evidence from Central European countries*
3/2012	Soňa Benecká Tomáš Holub Narcisa Liliana Kadlčáková Ivana Kubicová	*Does central bank financial strength matter for inflation? An empirical analysis*
2/2012	Adam Geršl Petr Jakubík Dorota Kowalczyk Steven Ongena José-Luis Peydró Alcalde	*Monetary conditions and banks' behaviour in the Czech Republic*
1/2012	Jan Babecký Kamil Dybczak	*Real wage flexibility in the European Union: New evidence from the labour cost data*
15/2011	Jan Babecký Kamil Galuščák Lubomír Lízal	*Firm-level labour demand: Adjustment in good times and during the crisis*
14/2011	Vlastimil Čadek Helena Rottová Branislav Saxa	*Hedging behaviour of Czech exporting firms*
13/2011	Michal Franta Roman Horváth Marek Rusnák	*Evaluating changes in the monetary transmission mechanism in the Czech Republic*

12/2011	Jakub Ryšánek Jaromír Tonner Osvald Vašíček	*Monetary policy implications of financial frictions in the Czech Republic*
11/2011	Zlatuše Komárková Adam Geršl Luboš Komárek	*Models for stress testing Czech banks´ liquidity risk*
10/2011	Michal Franta Jozef Baruník Roman Horváth Kateřina Šmídková	*Are Bayesian fan charts useful for central banks? Uncertainty, forecasting, and financial stability stress tests*
9/2011	Kamil Galuščák Lubomír Lízal	*The impact of capital measurement error correction on firm-level production function estimation*
8/2011	Jan Babecký Tomáš Havránek Jakub Matějů Marek Rusnák Kateřina Šmídková Bořek Vašíček	*Early warning indicators of economic crises: Evidence from a panel of 40 developed countries*
7/2011	Tomáš Havránek Zuzana Iršová	*Determinants of horizontal spillovers from FDI: Evidence from a large meta-analysis*
6/2011	Roman Horváth Jakub Matějů	*How are inflation targets set?*
5/2011	Bořek Vašíček	*Is monetary policy in the new EU member states asymmetric?*
4/2011	Alexis Derviz	*Financial frictions, bubbles, and macroprudential policies*
3/2011	Jaromír Baxa Roman Horváth Bořek Vašíček	*Time-varying monetary-policy rules and financial stress: Does financial instability matter for monetary policy?*
2/2011	Marek Rusnák Tomáš Havránek Roman Horváth	*How to solve the price puzzle? A meta-analysis*
1/2011	Jan Babecký Aleš Bulíř Kateřina Šmídková	*Sustainable real exchange rates in the new EU member states: What did the Great Recession change?*
15/2010	Ke Pang Pierre L. Siklos	*Financial frictions and credit spreads*
14/2010	Filip Novotný Marie Raková	*Assessment of consensus forecasts accuracy: The Czech National Bank perspective*
13/2010	Jan Filáček Branislav Saxa	*Central bank forecasts as a coordination device*
12/2010	Kateřina Arnoštová David Havrlant Luboš Růžička Peter Tóth	*Short-term forecasting of Czech quarterly GDP using monthly indicators*
11/2010	Roman Horváth Kateřina Šmídková Jan Zápal	*Central banks´ voting records and future policy*
10/2010	Alena Bičáková Zuzana Prelcová Renata Pašaličová	*Who borrows and who may not repay?*

9/2010	Luboš Komárek Jan Babecký Zlatuše Komárková	*Financial integration at times of financial instability*
8/2010	Kamil Dybczak Peter Tóth David Voňka	*Effects of price shocks to consumer demand. Estimating the QUAIDS demand system on Czech Household Budget Survey data*
7/2010	Jan Babecký Philip Du Caju Theodora Kosma Martina Lawless Julián Messina Tairi Rõõm	*The margins of labour cost adjustment: Survey evidence from European firms*
6/2010	Tomáš Havránek Roman Horváth Jakub Matějů	*Do financial variables help predict macroeconomic environment? The case of the Czech Republic*
5/2010	Roman Horváth Luboš Komárek Filip Rozsypal	*Does money help predict inflation? An empirical assessment for Central Europe*
4/2010	Oxana Babecká Kucharčuková Jan Babecký Martin Raiser	*A gravity approach to modelling international trade in South-Eastern Europe and the Commonwealth of Independent States: The role of geography, policy and institutions*
3/2010	Tomáš Havránek Zuzana Iršová	*Which foreigners are worth wooing? A Meta-analysis of vertical spillovers from FDI*
2/2010	Jaromír Baxa Roman Horváth Bořek Vašíček	*How does monetary policy change? Evidence on inflation targeting countries*
1/2010	Adam Geršl Petr Jakubík	*Relationship lending in the Czech Republic*
15/2009	David N. DeJong Roman Liesenfeld Guilherme V. Moura Jean-Francois Richard Hariharan Dharmarajan	*Efficient likelihood evaluation of state-space representations*
14/2009	Charles W. Calomiris	*Banking crises and the rules of the game*
13/2009	Jakub Seidler Petr Jakubík	*The Merton approach to estimating loss given default: Application to the Czech Republic*
12/2009	Michal Hlaváček Luboš Komárek	*Housing price bubbles and their determinants in the Czech Republic and its regions*
11/2009	Kamil Dybczak Kamil Galuščák	*Changes in the Czech wage structure: Does immigration matter?*
10/2009	Jiří Böhm Petr Král Branislav Saxa	*Percepion is always right: The CNB´s monetary policy in the media*
9/2009	Alexis Derviz Marie Raková	*Funding costs and loan pricing by multinational bank affiliates*
8/2009	Roman Horváth Anca Maria Podpiera	*Heterogeneity in bank pricing policies: The Czech evidence*
	David Kocourek	*The impact of early retirement incentives on labour market*

7/2009	Filip Pertold	*participation: Evidence from a parametric change in the Czech Republic*
6/2009	Nauro F. Campos Roman Horváth	*Reform redux: Measurement, determinants and reversals*
5/2009	Kamil Galuščák Mary Keeney Daphne Nicolitsas Frank Smets Pawel Strzelecki Matija Vodopivec	*The determination of wages of newly hired employees: Survey evidence on internal versus external factors*
4/2009	Jan Babecký Philip Du Caju Theodora Kosma Martina Lawless Julián Messina Tairi Rõõm	*Downward nominal and real wage rigidity: Survey evidence from European firms*
3/2009	Jiri Podpiera Laurent Weill	*Measuring excessive risk-taking in banking*
2/2009	Michal Andrle Tibor Hlédik Ondra Kameník Jan Vlček	*Implementing the new structural model of the Czech National Bank*
1/2009	Kamil Dybczak Jan Babecký	*The impact of population ageing on the Czech economy*
14/2008	Gabriel Fagan Vitor Gaspar	*Macroeconomic adjustment to monetary union*
13/2008	Giuseppe Bertola Anna Lo Prete	*Openness, financial markets, and policies: Cross-country and dynamic patterns*
12/2008	Jan Babecký Kamil Dybczak Kamil Galuščák	*Survey on wage and price formation of Czech firms*
11/2008	Dana Hájková	*The measurement of capital services in the Czech Republic*
10/2008	Michal Franta	*Time aggregation bias in discrete time models of aggregate duration data*
9/2008	Petr Jakubík Christian Schmieder	*Stress testing credit risk: Is the Czech Republic different from Germany?*
8/2008	Sofia Bauducco Aleš Bulíř Martin Čihák	*Monetary policy rules with financial instability*
7/2008	Jan Brůha Jiří Podpiera	*The origins of global imbalances*
6/2008	Jiří Podpiera Marie Raková	*The price effects of an emerging retail market*
5/2008	Kamil Dybczak David Voňka Nico van der Windt	*The effect of oil price shocks on the Czech economy*
4/2008	Magdalena M. Borys Roman Horváth	*The effects of monetary policy in the Czech Republic: An empirical study*
3/2008	Martin Cincibuch	*Central bank losses and economic convergence*

	Tomáš Holub Jaromír Hurník	
2/2008	Jiří Podpiera	*Policy rate decisions and unbiased parameter estimation in conventionally estimated monetary policy rules*
1/2008	Balázs Égert Doubravko Mihaljek	*Determinants of house prices in Central and Eastern Europe*
17/2007	Pedro Portugal	*U.S. unemployment duration: Has long become longer or short become shorter?*
16/2007	Yuliya Rychalovská	*Welfare-based optimal monetary policy in a two-sector small open economy*
15/2007	Juraj Antal František Brázdik	*The effects of anticipated future change in the monetary policy regime*
14/2007	Aleš Bulíř Kateřina Šmídková Viktor Kotlán David Navrátil	*Inflation targeting and communication: Should the public read inflation reports or tea leaves?*
13/2007	Martin Cinncibuch Martina Horníková	*Measuring the financial markets' perception of EMU enlargement: The role of ambiguity aversion*
12/2007	Oxana Babetskaia-Kukharchuk	*Transmission of exchange rate shocks into domestic inflation: The case of the Czech Republic*
11/2007	Jan Filáček	*Why and how to assess inflation target fulfilment*
10/2007	Michal Franta Branislav Saxa Kateřina Šmídková	*Inflation persistence in new EU member states: Is it different than in the Euro area members?*
9/2007	Kamil Galuščák Jan Pavel	*Unemployment and inactivity traps in the Czech Republic: Incentive effects of policies*
8/2007	Adam Geršl Ieva Rubene Tina Zumer	*Foreign direct investment and productivity spillovers: Updated evidence from Central and Eastern Europe*
7/2007	Ian Babetskii Luboš Komárek Zlatuše Komárková	*Financial integration of stock markets among new EU member states and the euro area*
6/2007	Anca Pruteanu-Podpiera Laurent Weill Franziska Schobert	*Market power and efficiency in the Czech banking sector*
5/2007	Jiří Podpiera Laurent Weill	*Bad luck or bad management? Emerging banking market experience*
4/2007	Roman Horváth	*The time-varying policy neutral rate in real time: A predictor for future inflation?*
3/2007	Jan Brůha Jiří Podpiera Stanislav Polák	*The convergence of a transition economy: The case of the Czech Republic*
2/2007	Ian Babetskii Nauro F. Campos	*Does reform work? An econometric examination of the reform-growth puzzle*
1/2007	Ian Babetskii Fabrizio Coricelli Roman Horváth	*Measuring and explaining inflation persistence: Disaggregate evidence on the Czech Republic*
13/2006	Frederic S. Mishkin	*Does inflation targeting make a difference?*

	Klaus Schmidt-Hebbel	
12/2006	Richard Disney Sarah Bridges John Gathergood	*Housing wealth and household indebtedness: Is there a household 'financial accelerator'?*
11/2006	Michel Juillard Ondřej Kameník Michael Kumhof Douglas Laxton	*Measures of potential output from an estimated DSGE model of the United States*
10/2006	Jiří Podpiera Marie Raková	*Degree of competition and export-production relative prices when the exchange rate changes: Evidence from a panel of Czech exporting companies*
9/2006	Alexis Derviz Jiří Podpiera	*Cross-border lending contagion in multinational banks*
8/2006	Aleš Bulíř Jaromír Hurník	*The Maastricht inflation criterion: "Saints" and "Sinners"*
7/2006	Alena Bičáková Jiří Slačálek Michal Slavík	*Fiscal implications of personal tax adjustments in the Czech Republic*
6/2006	Martin Fukač Adrian Pagan	*Issues in adopting DSGE models for use in the policy process*
5/2006	Martin Fukač	*New Keynesian model dynamics under heterogeneous expectations and adaptive learning*
4/2006	Kamil Dybczak Vladislav Flek Dana Hájková Jaromír Hurník	*Supply-side performance and structure in the Czech Republic (1995–2005)*
3/2006	Aleš Krejdl	*Fiscal sustainability – definition, indicators and assessment of Czech public finance sustainability*
2/2006	Kamil Dybczak	*Generational accounts in the Czech Republic*
1/2006	Ian Babetskii	*Aggregate wage flexibility in selected new EU member states*

14/2005	Stephen G. Cecchetti	*The brave new world of central banking: The policy challenges posed by asset price booms and busts*
13/2005	Robert F. Engle Jose Gonzalo Rangel	*The spline GARCH model for unconditional volatility and its global macroeconomic causes*
12/2005	Jaromír Beneš Tibor Hlédik Michael Kumhof David Vávra	*An economy in transition and DSGE: What the Czech national bank's new projection model needs*
11/2005	Marek Hlaváček Michael Koňák Josef Čada	*The application of structured feedforward neural networks to the modelling of daily series of currency in circulation*
10/2005	Ondřej Kameník	*Solving SDGE models: A new algorithm for the sylvester equation*
9/2005	Roman Šustek	*Plant-level nonconvexities and the monetary transmission mechanism*
8/2005	Roman Horváth	*Exchange rate variability, pressures and optimum currency area criteria: Implications for the central and eastern european countries*

7/2005	Balázs Égert Luboš Komárek	*Foreign exchange interventions and interest rate policy in the Czech Republic: Hand in glove?*
6/2005	Anca Podpiera Jiří Podpiera	*Deteriorating cost efficiency in commercial banks signals an increasing risk of failure*
5/2005	Luboš Komárek Martin Melecký	*The behavioural equilibrium exchange rate of the Czech koruna*
4/2005	Kateřina Arnoštová Jaromír Hurník	*The monetary transmission mechanism in the Czech Republic (evidence from VAR analysis)*
3/2005	Vladimír Benáček Jiří Podpiera Ladislav Prokop	*Determining factors of Czech foreign trade: A cross-section time series perspective*
2/2005	Kamil Galuščák Daniel Münich	*Structural and cyclical unemployment: What can we derive from the matching function?*
1/2005	Ivan Babouček Martin Jančar	*Effects of macroeconomic shocks to the quality of the aggregate loan portfolio*
10/2004	Aleš Bulíř Kateřina Šmídková	*Exchange rates in the new EU accession countries: What have we learned from the forerunners*
9/2004	Martin Cincibuch Jiří Podpiera	*Beyond Balassa-Samuelson: Real appreciation in tradables in transition countries*
8/2004	Jaromír Beneš David Vávra	*Eigenvalue decomposition of time series with application to the Czech business cycle*
7/2004	Vladislav Flek, ed.	*Anatomy of the Czech labour market: From over-employment to under-employment in ten years?*
6/2004	Narcisa Kadlčáková Joerg Keplinger	*Credit risk and bank lending in the Czech Republic*
5/2004	Petr Král	*Identification and measurement of relationships concerning inflow of FDI: The case of the Czech Republic*
4/2004	Jiří Podpiera	*Consumers, consumer prices and the Czech business cycle identification*
3/2004	Anca Pruteanu	*The role of banks in the Czech monetary policy transmission mechanism*
2/2004	Ian Babetskii	*EU enlargement and endogeneity of some OCA criteria: Evidence from the CEECs*
1/2004	Alexis Derviz Jiří Podpiera	*Predicting bank CAMELS and S&P ratings: The case of the Czech Republic*

CNB Research and Policy Notes

1/2012	Róbert Ambriško Vítězslav Augusta Dana Hájková Petr Král Pavla Netušilová Milan Říkovský Pavel Soukup	*Fiscal discretion in the Czech Republic in 2001-2011: Has it been stabilizing?*

3/2011	František Brázdik Michal Hlaváček Aleš Maršál	*Survey of research on financial sector modeling within DSGE models: What central banks can learn from it*
2/2011	Adam Geršl Jakub Seidler	*Credit growth and capital buffers: Empirical evidence from Central and Eastern European countries*
1/2011	Jiří Böhm Jan Filáček Ivana Kubicová Romana Zamazalová	*Price-level targeting – A real alternative to inflation targeting?*
1/2008	Nicos Christodoulakis	*Ten years of EMU: Convergence, divergence and new policy priorities*
2/2007	Carl E. Walsh	*Inflation targeting and the role of real objectives*
1/2007	Vojtěch Benda Luboš Růžička	*Short-term forecasting methods based on the LEI approach: The case of the Czech Republic*
2/2006	Garry J. Schinasi	*Private finance and public policy*
1/2006	Ondřej Schneider	*The EU budget dispute – A blessing in disguise?*
5/2005	Jan Stráský	*Optimal forward-looking policy rules in the quarterly projection model of the Czech National Bank*
4/2005	Vít Bárta	*Fulfilment of the Maastricht inflation criterion by the Czech Republic: Potential costs and policy options*
3/2005	Helena Sůvová Eva Kozelková David Zeman Jaroslava Bauerová	*Eligibility of external credit assessment institutions*
2/2005	Martin Čihák Jaroslav Heřmánek	*Stress testing the Czech banking system: Where are we? Where are we going?*
1/2005	David Navrátil Viktor Kotlán	*The CNB's policy decisions – Are they priced in by the markets?*
4/2004	Aleš Bulíř	*External and fiscal sustainability of the Czech economy: A quick look through the IMF's night-vision goggles*
3/2004	Martin Čihák	*Designing stress tests for the Czech banking system*
2/2004	Martin Čihák	*Stress testing: A review of key concepts*
1/2004	Tomáš Holub	*Foreign exchange interventions under inflation targeting: The Czech experience*

CNB ECONOMIC RESEARCH BULLETIN	
November 2012	*Financial stability and monetary policy*
April 2012	*Macroeconomic forecasting: Methods, accuracy and coordination*
November 2011	*Macro-financial linkages: Theory and applications*
April 2011	*Monetary policy analysis in a central bank*
November 2010	*Wage adjustment in Europe*
May 2010	*Ten years of economic research in the CNB*
November 2009	*Financial and global stability issues*
May 2009	*Evaluation of the fulfilment of the CNB's inflation targets 1998–2007*
December 2008	*Inflation targeting and DSGE models*
April 2008	*Ten years of inflation targeting*
December 2007	*Fiscal policy and its sustainability*
August 2007	*Financial stability in a transforming economy*
November 2006	*ERM II and euro adoption*
August 2006	*Research priorities and central banks*
November 2005	*Financial stability*
May 2005	*Potential output*
October 2004	*Fiscal issues*
May 2004	*Inflation targeting*
December 2003	*Equilibrium exchange rate*

Czech National Bank
Economic Research Department
Na Příkopě 28, 115 03 Praha 1
Czech Republic
phone: +420 2 244 12 321
fax: +420 2 244 14 278
http://www.cnb.cz
e-mail: research@cnb.cz
ISSN 1803-7070