UNIVERSITY OF LJUBLJANA
FACULTY OF ECONOMICS

MATJAŽ VOLK

CREDIT RISK AND THE BUSINESS CYCLE

DOCTORAL DISSERTATION

Ljubljana, 2015

AUTHORSHIP STATEMENT

The undersigned Matjaž Volk, a student at the University of Ljubljana, Faculty of Economics, (hereafter: FELU), declare that I am the author of the doctoral dissertation entitled Credit Risk and the Business Cycle (Kreditno Tveganje in Poslovni Cikel), written under the supervision of prof. dr. Igor Masten.

In accordance with the Copyright and Related Rights Act (Official Gazette of the Republic of Slovenia, Nr. 21/1995 with changes and amendments) I allow the text of my doctoral dissertation to be published on the FELU website.

I further declare

- the text of my doctoral dissertation to be based on the results of my own research;
- the text of my doctoral dissertation to be language-edited and technically in adherence with the FELU's Technical Guidelines for Written Works which means that I
  - cited and / or quoted works and opinions of other authors in my doctoral dissertation in accordance with the FELU's Technical Guidelines for Written Works and
  - obtained (and referred to in my doctoral dissertation) all the necessary permits to use the works of other authors which are entirely (in written or graphical form) used in my text;
- to be aware of the fact that plagiarism (in written or graphical form) is a criminal offence and can be prosecuted in accordance with the Copyright and Related Rights Act (Official Gazette of the Republic of Slovenia, Nr. 55/2008 with changes and amendments);
- to be aware of the consequences a proven plagiarism charge based on the submitted doctoral dissertation could have for my status at the FELU in accordance with the relevant FELU Rules on Doctoral Dissertation.

Date of public defense:
Committee Chair: prof. dr. Matej Marinč
Supervisor: prof. dr. Igor Masten
Member: prof. dr. Timotej Jagrič
Member: prof. dr. Luigi Marattin

Ljubljana, December, 2015. Author's signature: ____________________

CREDIT RISK AND THE BUSINESS CYCLE

Abstract

Credit risk was during the crisis shown to be the most pressing problem for Slovenian banks. This was the main motivation for my doctoral dissertation, which studies some aspects of credit risk. First, I show, which are the main factors that influence the probability of default of Slovenian firms. In general these can be divided into idiosyncratic or firm-specific and aggregate, which influence all firms equally. Second, a more important result of this doctoral dissertation is the underestimation of credit risk by banks in the crisis period. I show that this can be a consequence of incentives to apply discretion in credit risk assessment. Third, I propose novel methodologies for modelling credit risk. As I show, these have high performance in predicting state of non-performing borrowers and transitions to default.

Understanding credit risk drivers of Slovenian non-financial firms was the first goal of my doctoral dissertation. For this reason, I build a default probability model, where the event of default is explained with firms' financial ratios and aggregate macroeconomic effects. The results reveal that several firm-specific variables, like firm size, liquidity, cash flow, indebtedness, etc., influence firms' ability to pay back the loan. Macroeconomic factors also have an important contribution, since they capture dynamics in time and thus explain the movement of average default probability.

The default probability model is used to test the hypothesis of underestimation of credit risk by Slovenian banks during the crisis period. Its classification accuracy is compared to the model, which uses only internal credit ratings as explanatory variables. These are assigned directly by banks and thus represent their own estimate of credit risk. The results show that in the crisis period the classification accuracy of the model based on credit ratings dropped in absolute terms and, more importantly, relative to benchmark default probability model. This is evidence of underestimation of credit risk. Given that credit ratings are closely related to loan loss provision rates, this analysis shows that underestimation of credit risk served to avoid recognising losses in banks' balance sheets. I also show that underestimation of credit risk was more pronounced in banks with higher incentives to apply discretion in credit risk assessment.

The crisis revealed the need for more advanced credit risk modelling techniques. I test the performance of various credit default models in predicting non-performing borrowers and transitions to default. In addition to conventional binary classifiers, which are typically used in practice, I evaluate the performance of two novel methodologies - dynamic and tobit credit risk models. I show that tobit model, where overdue in loan repayment is modelled explicitly, outperforms all the other models. The choice between dynamic and static version of the model depends whether one is interested in predicting state

of non-performing borrowers or new defaulters. For the former, the persistence is of a key importance and therefore the dynamic model is the advantageous methodology. For predicting new defaulters, however, static tobit model is shown to outperform all other models by a large margin.

This dissertation contributes to the science in several ways. First, it shows, by estimating the model of default probability, what are the key credit risk factors of Slovenian firms. Second, to my knowledge, this is the first analysis that explores the information content of credit ratings. Their capacity to predict financial distress of firms is compared to predictive accuracy of classical model of default probability, which gives an answer of underestimation of credit risk. Third, I propose that an important factor behind underestimation of credit risk are incentives for discretionary credit risk assessment. As shown, the underestimation of credit risk was the most pronounced among small domestic banks, who had high proportion of non-performing loans, weak capital adequacy and were owned by financially weak owners. Forth, this is the first study that applies dynamic and tobit methodology to credit risk modelling.

The dissertation has several implications for banking regulation. Underestimation of credit risk is an attempt of banks to temporary avoid recognising losses in their balance sheets. Huge capital shortfall revealed by comprehensive review of Slovenian banking system is to a large part a consequence of underestimation of credit risk in the crisis period. To prevent similar episodes in the future, regulation should monitor credit risk management practices by banks more strictly. This is already possible under current regulatory regime, but regulatory forbearance is often applied in similar crisis situations. Regulators should also monitor incentives for discretionary credit risk assessment, since these can destabilise banking system, which has, as we saw in the crisis period, negative consequences for the whole economy. In addition, increasing capital requirements in times of financial distress can have counterproductive effect, since this even increases incentives to underestimate credit risk.

There are also several important implications from methodological perspective. The proposed model of default probability in already in use at Bank of Slovenia to simulate probability of default under different stress scenarios. The model could also be used to test the adequacy of banks' estimates of credit risk, similarly as I test the predictive capacity of credit ratings in this dissertation. In addition, my results show that the prevailing credit risk methodologies can be significantly improved by including the dynamics and choosing the tobit functional form. This is especially pronounced for conventional default probability model that is typically used by banks and regulators and is shown to have very low classification accuracy.

Keywords: credit risk, credit rating, probability of default, underestimation of credit risk, dynamic model, tobit model

KREDITNO TVEGANJE IN POSLOVNI CIKEL

Povzetek

Kreditno tveganje se je tekom krize pokazalo kot najbolj pereč problem slovenskih bank. To je bila poglavitna motivacija za doktorsko disertacijo, v kateri analiziram določene vidike kreditnega tveganja. Prvič, pokazano je, kateri so ključni dejavniki, ki vplivajo na kreditno tveganje slovenskih podjetij. V splošnem je te moč razdeliti na dejavnike, ki so značilni za posamezno podjetje ter agregatne dejavnike, ki delujejo na vsa podjetja v enaki meri. Drugič, pomembnejši rezultat te disertacije je podcenjevanje kreditnega tveganja s strani bank v času krize. Pokazano je, da je to lahko posledica spodbud za diskrecijsko ocenjevanje kreditnega tveganja. Tretjič, v disertaciji je predlagana nova metodologija za modeliranje kreditnega tveganja, ki precej izboljša pojasnjevalno moč modela pri napovedovanju stanja neplačila.

Prepoznavanje faktorjev, ki vplivajo na kreditno tveganje slovenskih nefinančnih družb, je bil prvi cilj moje doktorske disertacije. S tem namenom sem ocenil model verjetnosti neplačila, kjer je indikator neplačila pojasnjen s finančnimi kazalci podjetij ter agregatnimi makroekonomskimi spremenljivkami. Rezultati kažejo, da na verjetnost neplačila vplivajo številni kazalci podjetij kot npr. velikost podjetja, likvidnost, denarni tok, zadolženost, itd. Makroekonomski faktorji imajo tudi pomemben prispevek, saj zajamejo dinamiko v času in tako pojasnjujejo gibanje povprečne verjetnosti neplačila.

Model verjetnosti neplačila je uporabljen za testiranje hipoteze o podcenjevanju kreditnega tveganja slovenskih bank v času krize. Natančnost napovedi je primerjana z modelom, kjer so kot pojasnjevalne spremenljivke vključene samo bonitetne ocene. Te so določene neposredno s strani bank in predstavljajo njihovo lastno oceno kreditnega tveganja. Rezultati kažejo, da se je v času krize klasifikacijska natančnost modela, ki temelji na bonitetnih ocenah znižala tako absolutno, kot glede na primerjalni model verjetnosti neplačila. To kaže na podcenjevanje kreditnega tveganja. Glede na to, da so bonitetne ocene tesno povezane z oslabitvami in rezervacijami, ta analiza kaže, da so se s podcenjevanjem kreditnega tveganja banke izogibale priznavanju izgub v svojih bilancah. Pokazano je tudi, da je bilo podcenjevanje kreditnega tveganja bolj prisotno pri bankah, ki so imele večje spodbude za diskrecijsko ocenjevanje tveganja.

Kriza je pokazala potrebo po bolj naprednih metodah modeliranja kreditnega tveganja. V disertaciji je testirana napovedna moč različnih modelov kreditnega tveganja za napovedovanje nedonosnih komitentov in prehodov v stanje neplačila. Poleg standardnih binarnih modelov, ki se navadno uporabljajo v praksi, sem ovrednotil tudi napovedno moč dveh novih metodologij - dinamičnih in tobit modelov kreditnega tveganja. Tobit model, kjer je eksplicitno modelirana zamuda pri odplačevanju posojila, kaže najboljše rezultate. Odločitev za statično oz. dinamično specifikacijo je odvisna ali je primarni namen identi-

ficirati vse neplačnike, ali samo tiste, ki na novo preidejo v stanje neplačila. Pri prvem je persistentnost ključnega pomena, zato dinamični model daje boljše rezultate. Pri napovedovanje novih neplačnikov pa je pokazano, da statični tobit model z veliko razliko prekaša vse ostale testirane modelske pristope.

Ta disertacija na več področjih pomembno prispeva k znanosti. Prvič, z oceno modela verjetnosti neplačila pokažem kateri so ključni dejavniki kreditnega tveganja slovenskih podjetij. Drugič, po mojem vedenju je to prva analiza, ki raziskuje informacijsko vrednost bonitetnih ocen. Njihova sposobnost napovedati finančne težave podjetij je primerjana z natančnostjo napovedi klasičnega modela verjetnosti neplačila, kar daje odgovor o podcenjevanju kreditnega tveganja. Tretjič, v disertaciji kot pomemben faktor v ozadju podcenjevanja tveganja izpostavljam spodbude za diskrecijsko ocenjevanje kreditnega tveganja. Kot je pokazano, je bilo podcenjevanje kreditnega tveganja najbolj izrazito pri majhnih domačih bankah, ki so imele visok delež nedonosnih posojil, nizko kapitalsko ustreznost in so bile v lasti finančno šibkih lastnikov. Četrtič, gre za prvo študijo, ki dinamični in tobit modelski pristop aplicira na področje modeliranja kreditnega tveganja.

Disertacija ima številne implikacije za bančno regulativo. Podcenjevanje kreditnega tveganja je poskus bank, da se začasno izognejo priznavanju težav v njihovih bilancah. Dejanska kreditna sposobnost dolžnikov pa je vedno prej ali slej razkrita in izgube morajo biti poknjižene. Velik kapitalski primanjkljaj, ki ga je razkril skrbni pregled slovenskega bančnega sistema v letu 2013, je v veliki meri posledica podcenjevanja tveganja v času krize. Za preprečitev podobnega razvoja dogodkov mora regulativa v prihodnje bolj striktno spremljati prakse bank pri ocenjevanju kreditnega tveganja. To je mogoče že v okviru sedanje ureditve, vendar se regulatorji v času krize pogosto odločijo za regulatorno popuščanje. Analiza kaže tudi pomembnost nadzora nad spodbudami za podcenjevanje tveganja, ker lahko te destabilizirajo bančni sistem, kar ima negativne posledice za celotno gospodarstvo. Poleg tega ima povečevanje kapitalskih zahtev v času krize lahko neželjene učinke, saj to še povečuje spodbude za podcenjevanje kreditnega tveganja.

Disertacije ima pomembne implikacije tudi z metodološkega vidika. Predlagan model verjetnosti neplačila se že uporablja v Banki Slovenije za simuliranje verjetnosti neplačila ob različnih predpostavkah stresnih scenarijev. Model bi lahko bil uporabljen tudi za testiranje ustreznosti ocenjevanja kreditnega tveganja s strani bank, podobno kot je v tej disertaciji testirana natančnost napovedi bonitetnih ocen. Poleg tega moji rezultati kažejo, da je z uporabo dinamičnega in tobit modelskega pristopa mogoče precej nadgraditi prevladujoče modele kreditnega tveganja. To velja zlasti za standardni model verjetnosti neplačila, ki ga navadno uporabljajo banke in regulatorji in se je izkazal kot model z zelo nizko napovedno natančnostjo.

Ključne besede: kreditno tveganje, bonitetne ocene, verjetnost neplačila, podcenjevanje kreditnega tveganja, dinamični model, tobit model

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Introduction

Banking crisis in Slovenia was very severe and it basically still persists since many banks still have large proportion of non-performing loans (more than 90 days overdue). Although not being the only problem in Slovenian banks, large share of NPLs certainly was the prevailing and had many negative consequences for the banks and also for the economy as a whole. First, NPLs put pressure on bank capital through loan loss provisions and write-offs. In financial distress, when raising capital is extremely difficult it leads to deleveraging in order to fulfil the required capital adequacy ratio. This process is additionally intensified with increasing capital requirements due to higher riskiness. Second, large share of bad loans increases the costs of funding. Wholesale investors increase the required premium, whereas depositors typically move their savings to safer banks - flight to quality effect. Share of NPLs also has an important effect on bank's credit rating, which also significantly affects bank's ability to raise funds. At the same time revenues are also severely affected, since bank do not get any interest income from defaulted borrowers.

All the mentioned aspects of non-performing loans have a strong negative effect on banks' ability to provide loans to the economy. In addition, banks also become more risk averse in the crisis and possibly also increase lending interest rates (as was the case in Slovenia), which additionally limits the economy to raise funds in banks. This is especially problematic in such country as Slovenia where banks' loans are the prevailing external finance source for the economy. Any obstacle in the stream of funds from banks to the economy can have a significant impact on macroeconomic activity.

Slovenian banking crisis was mainly credit risk crisis. Once the crisis hit, the share of non-performing loans started rising sharply and for the corporate sector, which represent the large majority of all NPLs, exceeded 25% in 2013. The crisis opened many questions to be explored in order to better understand the situation and to be able to take some measures that could prevent similar developments in the future. For this reason credit risk is the main topic of my dissertation, which consists of three papers. Next paragraphs shortly present the motivation behind these papers.

From a financial stability point of view it is very important to understand the credit risk factors that led to several firm defaults during the crisis. Since the pioneering work of Altman (1968), who uses discriminant analysis technique to model credit risk, a large set of studies find that credit risk is in general driven by idiosyncratic and systemic factors (Bangia, Diebold, Kronimus, Schagen & Schuermann, 2002; Bonfim, 2009; Carling, Jacobson, Lindé & Roszbach, 2007 and Jiménez & Saurina, 2004). Some authors also stress that strong GDP or credit growth in the pre-crisis period may have increased the share of defaulted firms or deteriorate NPL dynamics in the crisis (Festić, Kavkler & Repina, 2011; Foss, Norden & Weber, 2010 and Jiménez & Saurina, 2006). The first goal therefore is, to show empirically, at micro level, what drives credit risk of Slovenian non-financial firms and to choose the best model specification, which is then also used in other two papers.

Credit risk is procyclical and is expected to increase in economic downturns. However, this effect might be amplified when banks ex-ante overestimate borrowers' creditworthiness. Under conditions of fierce competition and high credit growth banks might indeed be willing to assign higher credit ratings to borrowers. Moreover, banks might also have an incentive to apply discretion in credit risk assessment in the crisis period, due to pressing burden of NPLs and difficulties in raising fresh capital. Regulators could in principle require from banks to strictly follow the regulatory standards and thus always reveal the actual state of their credit portfolio. However, in the crisis period regulators often apply regulatory forbearance in order to partially alleviate the pressure on capital through loan loss provisions (Hoffman & Santomero, 1998 and Brown & Dinc, 2011).

The comprehensive review of Slovenian banking system, which was carried out in 2013, provide an unique environment to study the hypotheses of underestimation of credit risk. The results revealed significant capital shortfalls. In particular, estimated capital shortfalls of all examined banks amounted to 214% of existing capital. There were, however, large differences in the capital shortfalls between different groups of banks. The smallest shortfall - 78% of existing capital - was found for foreign owned banks. For domestic banks the required recapitalisation was substantially larger and amounted to 244%. There were important differences also within the group of domestic banks. The recapitalisation requirement for the largest two and majority state owned banks on the market, amounted to 228% of existing capital, while for the small and predominantly privately owned the figure was 274%.

The revealed large capital shortfalls can to a large extend be attributed to underestimation of credit risk. Moreover, the differences in required recapitalisation between groups of banks could be a consequence of different incentives to apply discretion in credit risk assessment. Foreign-owned banks are expected to have less incentives to underestimate credit risk, since in the crisis period they had lower exposure to non-performing loans, higher capitalisation and easier access to funding through internal capital markets. On the other hand, domestic banks were heavily exposed to non-performing loans, reaching 35% in 2013. Moreover, small domestic banks were owned by financially weak private firms, which could not provide fresh capital to the banks to cover losses. This group of banks is thus expected to have the highest incentives to underestimate risk. Discretion in credit risk assessment and incentives behind it is the main topic of the second paper.

The crisis revealed a need for more advanced modelling techniques of credit risk. Researchers, banks and regulators usually apply a model similar to Altman's (1968), where the default occurrence is explained by firm financial ratios like liquidity, indebtedness, cash flow, efficiency, etc. The models are usually estimated using logit or probit estimator. Adding the autoregressive component in the model could substantially improve the performance of the model, since the information about the default status in the previous period is very informative for current state. However, the estimation of dynamic non-linear panel models is not straightforward, since, unlike in the linear case, the endogeneity problem cannot be solved with differencing and applying Arellano and Bond (1991) or Blundell and Bond (1998) estimator. In addition, instead of modelling the credit risk in discrete choice model, the overdue in loan repayment could be modelled explicitly, which would keep all the information content in it. Since loan overdue is censored at zero, the tobit model needs to be used. Similarly as for the discrete choice case, I also estimate the dynamic version of tobit model. The predicting ability of all these novel modelling techniques in relation to classical default probability models is the topic of the third paper. To my knowledge, it is the first attempt to model credit risk using the tobit modelling methodology and applying the dynamic structure of the model.

Research questions

The dissertation answers to several research questions, which can be decomposed into two groups, one dealing with different aspects of modelling credit risk and the other answering the questions about the developments related to credit risk in Slovenian banks during the crisis period.

Research questions related to credit risk modelling:

- *Which are the main credit risk factors of Slovenian non-financial firms?*
- *Do macro effects improve the performance of PD model?*
- *Can the prediction ability of the model be improved by applying different functional form of the model (probit vs. tobit) and/or specifying the dynamic structure of the model?*

Research questions regarding the credit risk assessment by Slovenian banks:

- *Are credit ratings assigned by banks in line with estimated PDs?*
- *Did banks underestimate credit risk in the crisis?*
- *Was the discretion in credit risk assessment more pronounced in banks with higher incentives to underestimate credit risk (higher NPL burden, weaker capital position, limited ability to raise funds)?*
- *Can the differences in underestimation of credit risk between banks rationalise the differences in capital shortfall revealed by the comprehensive review of Slovenian banking system in 2013?*

Scientific contribution

The dissertation has several important contributions to the science, which can be divided into four groups.

First, it helps to explain what are the key credit risk factors of Slovenian non-financial firms. Besides the firm-level variables, which are the main drivers of credit risk variation between firms, special attention is paid to variation in time to find out how sensitive are probabilities of default to macroeconomic fluctuations. Similar analysis has already been done by Kavčič (2005), but in a different manner. Whereas she uses internal credit ratings as the dependent variable, I use overdue in loan repayment to define whether a borrower is in default. This is a much more objective setting since overdue, as opposed to credit rating, is not subject to bank discretion in credit risk assessment. In addition, I also add macroeconomic factors in the model and found, similar as Bonfim (2009), that they improve the performance of the model.

Second, I use internal credit ratings as banks' assessment of credit risk and explore the information content through implied probabilities of default. Credit ratings are very valuable data, since they are assigned directly by banks and thus give an information of banks own assessment of credit risk. There are only a few research papers that use internal credit ratings (like for instance Carling et al., 2007), but, to my knowledge, non of them explores their information content through implied probabilities of default.

Third, it gives an answer about the importance of incentives in credit risk assessment. Banks with higher NPL burden, weaker capital position or limited access to funding could have stronger incentives to apply discretion in credit risk assessment. This information is especially valuable for regulators, which can, based on the results, take some measures to monitor these incentives and reduce underestimation of credit risk. My work is related to the paper by Huizinga and Laeven (2012), but is still different. While Huizinga and Laeven (2012) focus on discretion in banks' valuation of assets, I focus on discretion in credit risk assessment. The empirical test of the hypothesis is novel and is based on predictive capacity of credit ratings in predicting financial distress of banks' clients. I estimate how the predictive ability of credit ratings evolves over time and across groups of banks. The predictive capacity of credit ratings is compared to that of a conventional PD model that uses only publicly available information, such as various financial ratios of banks' clients.

Fourth, I propose a novel methodology for estimating credit default and overdue in loan repayment. There are many important contributions from this perspective. First, I estimate the dynamic credit default model by applying Wooldridge's (2005) method. To my knowledge, it is the first attempt to model credit risk as an autoregressive process. Second, I go a step further and instead of modelling the credit default in discrete choice setting, like it is usually done in the literature (see for instance Bonfim, 2009 and Carling et al., 2007), I model the overdue in loan repayment explicitly by applying the tobit model. This keeps all the information content of overdue and is expected to have better performance. In addition, overdue is already a risk measure and it is therefore reasonable to model it directly. Like in the discrete choice case I also estimate dynamic tobit model, following Wooldridge (2005). This is a novel approach, which enable early detection of possibly problematic borrowers and, given the dynamic setting, is expected to have better predicting ability of non-performing borrowers. For this reasons, it is very useful for banks and regulators in practice. Third, I propose an approach for modelling credit default on quarterly frequency using mixed frequency data.

Data

Data for the analysis are combined from three data sources. First and most important is Credit Register of the Bank of Slovenia, which contains credit-related information about all bank-borrower relationships from 1993 on. These are confidential data so any individual piece of information cannot be revealed. Second, I use balance sheet and income statement data that are for all Slovenian firms collected by the Agency of the Republic of Slovenia for Public Legal Records and Related Services (AJPES) at yearly basis. Third, macroeconomic and other financial series are obtained from Statistical Office of the Republic of Slovenia (SURS) and from the Bank of Slovenia. The analysis is restricted to non-financial firms from 2007 on, since the data about overdue in loan repayment, which is a key variable in all three papers, are not available before 2007.

Structure of the dissertation

The dissertation is structured as follows. Chapter 1 presents the first paper of the doctoral dissertation entitled Estimating Probability of Default and Comparing it to Credit Rating Classification by Banks. Chapter 2 studies the incentives for underestimation of credit risk, which is the main topic of the second paper entitled Discretionary Credit Rating and Bank Stability in a Financial Crisis. The comparison of novel modelling methodologies is presented in Chapter 3, which is like the underlying paper entitled Evaluating the Performance of Dynamic and Tobit Models in Predicting Credit Default. After a concluding remarks, which summarizes the main findings and gives some policy implications, the dissertation concludes with a list of references. The appendix includes two additional sets of results not shown in the main part of the dissertation and extended abstract in Slovene language.

1 Estimating Probability of Default and Comparing it to Credit Rating Classification by Banks[1]

1.1 Introduction

After the start of the crisis in 2007 credit risk has become one of the main issues for analysts and researchers. The deteriorated financial and macroeconomic situation forced many firms into bankruptcy or to a significantly constrained business activity. Banks were to a large extent unprepared to such a large shock in economic activity so they suffered huge credit losses in the following years. Although it is clear that credit risk increases in economic downturn, this effect might be amplified when banks ex-ante overestimate the creditworthiness of borrowers. Under conditions of fierce competition and especially in periods of high credit growth banks might indeed be willing to assign higher credit ratings to obligors, which could cause problems in their portfolios when economic situation worsens.

Knowing why do some firms default while others don't and what are the main factors that drive credit risk is very important for financial stability. Since the pioneering work of Altman (1968), who uses discriminant analysis technique to model credit risk, a large set of studies find that credit risk is in general driven by idiosyncratic and systemic factors (Bangia et al., 2002; Bonfim, 2009; Carling et al., 2007 and Jiménez & Saurina, 2004). The importance of macroeconomic effects is to capture counter-cyclicality and correlation of default probabilities. On the other hand there is also a strong reverse effect of credit risk on macroeconomic activity. In recent study Gilchrist and Zakrajšek (2012) find that a level of credit risk statistically significantly explains the movement of economic activity. They construct a credit spread index (GZ spread), which indicates high counter-cyclicality movement and has high predictive power for variety of economic indicators.

This paper analyses credit risk of Slovenian non-financial firms using an indicator of firm default based on credit overdue. I focus on modelling default probability and use similar approach as those proposed by Bonfim (2009) and Carling et al. (2007). The results obtained suggests that probability of default (PD) can be explained by firm specific characteristics as well as macroeconomic or time effects. While macro variables influence all firms equally, and thus drive average default probability, firm specific variables are crucial to distinguish between firms' creditworthiness. Similar as Bonfim (2009), I find a model that includes time dummies as time effects to perform slightly better than model with macroeconomic variables. This result is expected, since time dummies also capture institutional, regulatory or other systemic changes in time.

The main contribution of this paper is the comparison of the estimated PDs to credit

[1] Volk M. (2012). Estimating Probability of Default and Comparing it to Credit Rating Classification by Banks. *Economic and Business Review, 14*(4), 299-320.

rating classification by banks. I select two models that best fit the data and link the estimated firm-level PDs with all credit grades which are given to borrowers by banks. I find that estimated PDs and credit ratings by banks often exhibit quite different measures of credit risk. The results also suggest that in the crisis banks allow for higher risk borrowers in credit grades A, B and C. This could be due to additional borrower-related information that banks take into consideration in assessing borrowers' riskiness, to the lags in reclassification process or a possible underestimation of systemic risk factors by banks.

The rest of the paper is structured as follows. Next section presents the data. Section 1.3 describes the modelling approach used to estimate the probability of default. Estimation results of various credit risk models and an application of the models in analysis of credit rating classification is presented in Section 1.4. Section 1.5 concludes the paper.

1.2 Data

Three different data sources are combined to construct dataset used in the econometric analysis. First, balance sheet and income statement data for all Slovenian firms are collected by the Agency of the Republic of Slovenia for Public Legal Records and Related Services (AJPES) at yearly basis. The analysis is restricted only to non-financial corporations. Second, data about credit exposures, credit ratings, credit overdue, etc. are gathered in Credit register at Bank of Slovenia. The banks are mandatory to report these data every month, but since firms' balance sheet and income statement data are only available at yearly basis, I use the end-of-year data. Third, to capture the business cycle effects when modelling PD, I use a set of macroeconomic and financial series which are obtained from Statistical Office of the Republic of Slovenia (SURS) and from Bank of Slovenia.

Two different subsets of the data are used for modelling probability of default and in comparison of the estimated PDs to credit rating classification by banks. Hence I present each of them separately.

1.2.1 Data for modelling PD

Under the framework of Basel II the obligor defaults on his credit obligation if (1) he is unlikely to pay the obligation or (2) is passed overdue more than 90 days (BCBS, 2006). Since it is difficult to set the objective criteria for unlikeliness of paying the obligation, I derive the indicator of firm default from credit overdue. Firm i is in default if its principal or interest payments are more than 90 days overdue to at least one bank in year t. The stock of defaulted firms increased significantly in the crisis, from 3.9% in 2007 to 9.9% in 2010.

To model the PD I use yearly data from 2007 to 2010. Since PD is the probability that a firm will default in year t given that it did not default in year $t-1$, all firms who have for the first time taken the loan (in any bank) in year t are excluded from the sample. Firms that were in the state of default for two or more consecutive years are also excluded and only their first migration to the state of default is taken into account. Similar to Bonfim (2009), I keep all the firms that defaulted twice or more in a given sample, but not in two consecutive years.

The firm's financial ratios like measures of liquidity, solvency, indebtedness, cash flow, profitability, etc. are key inputs to PD models. They capture firm specific effects and reflect the riskiness of firms. The sample additionally excludes firms with significant outlier in some of their characteristics so that all the observations in the 1st and the 100th percentile are dropped. Table 1 presents the summary statistics for some financial ratios and the other firm's characteristics for defaulted and non-defaulted firms, which are taken into account in the analysis, for the period 2007-2010. I now turn to the descriptive analysis of the variables.

Table 1: Summary statistics for firms with and without defaults

	Firms with no default		Firms in default	
	Mean	St. dev.	Mean	St. dev.
Total sales (EUR million)	2.2	13.5	1.1	3.7
Firm age (in years)	13.4	6.6	12.0	6.7
Quick ratio	1.3	1.6	0.9	1.1
Debt-to-assets	0.7	0.3	0.9	0.5
Cash flow	0.1	0.2	-0.1	0.5
Asset turnover ratio	1.5	1.9	0.8	1.0
Interest coverage	4.6	11.2	-0.3	7.3
Blocked account (in days)	5.9	32.6	105.8	127.5
Total credit (EUR million)	0.4	1.1	0.7	1.4
No. of bank-borrower relationships	1.4	0.7	1.8	1.0
No. of observations	65557		2887	

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the summary statistics for firms' financial ratios and other variables used for PD modelling. It is calculated for the period 2007-2010. Statistics for Interest coverage are computed on reduced sample of 45236 observations due to the missing values.

Total sales which is a measure of firm size indicates that defaulted firms are on average smaller. Similar result is found by other researchers like Antão and Lacerda (2011), Carling et al. (2007), Kavčič (2005) and Psillaki, Tsolas and Margaritis (2010). Smaller firms are less diversified and rely on less or perhaps on a single project. They are often also more financially constrained comparing to larger firms and may have problems in raising funds in economic downturns (Bernanke, Gertler & Gilchrist, 1996, 1999).

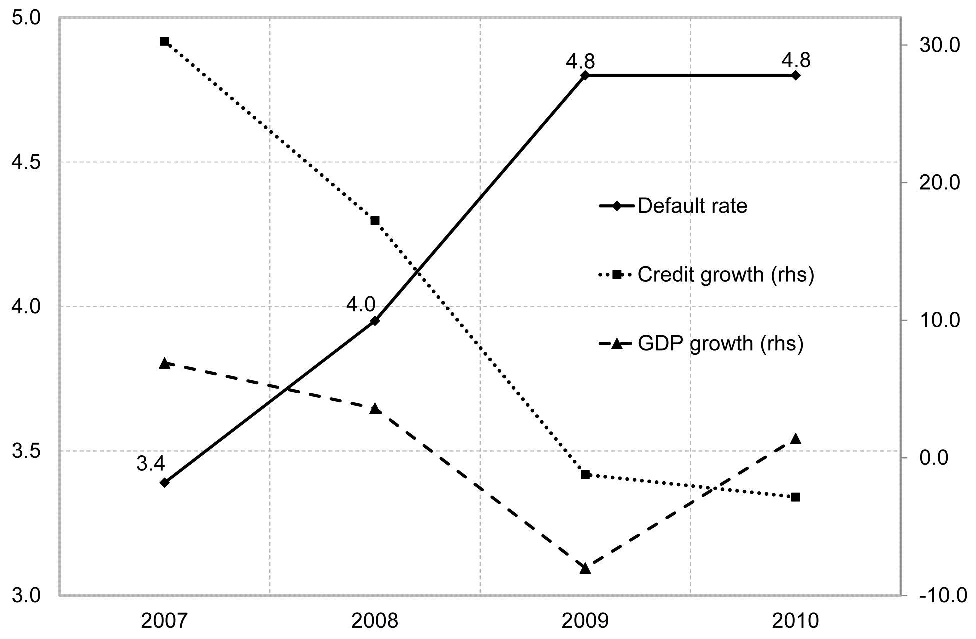
Defaulted firms are on average younger, have lower liquidity, higher leverage, lower cash flow, worse operating performance and have lower interest coverage, comparing to non-defaulted firms. A significantly useful indicator to separate between firms in default and non-default is also a variable which measures a number of days a firm has blocked bank account per year. It shows that in a given sample defaulted firms' bank accounts were on average blocked 106 days per year, whereas accounts for firms with no default were on average blocked only 6 days per year.

Somewhat less expectedly firms in default have on average higher amount of total credit. Jiménez and Saurina (2004) indeed show that there is an inverse relationship between the size of the loan and the probability of default since larger loans are more carefully screened. The difference between the two approaches is that their research is done at loan level, whereas this analysis is at firm level, where the default occurs if a firm defaults in any bank in year t.

According to the relationship banking theory banks and borrowers can benefit from a close relationship (Boot, 2000). Especially small banks tend to have comparative advantage in using soft information technologies (Berger & Udell, 2002). Nevertheless, Berger and Black (2011) show that bank will generally choose a hard information technology over a soft information technology if a sufficient hard information is available. The results of Jiménez and Saurina (2004) indicate that when borrower's loans are spread across several banks there is less of an incentive to finance riskier borrowers. Banks are willing to finance higher risk borrowers if they have a close relationship with them. This seems not to hold in the case of Slovenia since firms in default have on average higher number of bank-borrower relationships. One explanation might be that risky firms seek for credit in other banks because current creditors don't want to lend them any more if they are not paying off the loan regularly. The borrower's credit history is in general not available to new creditors, thus they can only assess firms' creditworthiness through their financial ratios.

Jacobson, Lindé & Roszbach (2013) argue that firm-specific variables account for the cross-section of the default distribution, while macroeconomic variables play the role of shifting the mean of the default distribution in each period. Finally, care is taken to include business cycle effects in the model. Figure 1 illustrates the movement of default rate for a given sample against the two indicators of the business cycle. The default rate appears to be highly countercyclical and it seems more tightly related to credit growth than to GDP growth. As shown by Bonfim (2009), Jiménez and Saurina (2006) and others, most of the credit risk is built up during periods of strong credit growth when banks apply looser credit standards. This risk materializes when the economy hits a downturn. With looser credit standards banks attract more risky borrowers which deteriorate their average assets quality. Marcucci and Quagliariello (2009) find that banks with lower asset quality are much more vulnerable in recessions. The increase in default rates due to one percentage point decrease in output gap is almost four times higher for those banks than the effect on banks with better portfolios.

Figure 1: Default rate, credit growth and GDP growth in percentages



Source: Bank of Slovenia; SURS; own calculations.

1.2.2 Data used in the comparison of the estimated PDs to the banks' rating classification

While data for modelling probability of default are at borrower level, analysis of credit rating classification is done at bank-borrower level. Similar as in the PD estimation part, the analysis is restricted to the period 2007-2010.

The credit ratings exhibit banks' assessment of the debtors' ability to discharge the liabilities to the bank. As opposed to credit overdue, credit ratings are more subjective measure of firms' riskiness. Crouhy, Galai and Mark (2001) argue that rating systems are usually based on general considerations and experience and not on mathematical modelling. Although financial health of the firm is a key factor of rating classification, analysts must also take into account managerial and other qualitative information like feature of the industry. As Crouhy et al. (2001) point out, obligor credit ratings exhibit the probability of default by a borrower in repaying its obligation in the normal course of the business.

According to the Regulation on the assessment of credit risk losses of banks and savings banks (hereinafter Regulation) Slovenian banks classify borrowers into five credit grades, from A to E. Each bank uses its own methodology for estimating borrowers' riskiness, which should for reporting purposes be transformed to this five-grade scale. The two main criteria that they consider in classification are the financial health and credit overdue of a firm. The quality of the pledged collateral does not affect the credit rating. All firms for whom there is a substantial likelihood of the loss of part of the financial asset or bank assesses that it will not be paid, are more than 90 days overdue, are insolvent or are in bankruptcy should be classified in either D or E class.

Credit ratings are pro-cyclical (Amato & Furfine, 2004), thus it is expected to find deteriorating rating structure in economic downturn. This is confirmed in Table 2, which shows that there is a decreasing trend of borrowers with grade A, whose share has dropped by 6.5 percentage points since 2008, whereas the share of non-performing firms has increased by two thirds. A similar shift is also noted from credit overdue in Table 3, where the share of borrowers who are more than 90 days overdue increased by 3 percentage points since its lowest value in 2007. Thus the proportion of firms who are more than 90 days overdue is much lower on bank-borrower level than on firm level, where this proportion increased by 6 percentage points (2.5-times) in the same period. Once a firm is in overdue in one bank, there is a substantial likelihood that in the following periods it becomes a delinquent also in other banks to which it has liabilities. Especially in economic downturns when firms struggle to repay the debt, significantly increased credit overdue in one bank clearly indicates that the firm has financial problems and thus exhibits a higher credit risk to all banks to which it has liabilities.

Table 2: Credit rating structure in percentages

Credit rating	2007	2008	2009	2010
A	57.3	57.9	53.7	51.4
B	32.7	31.7	33.0	33.9
C	5.2	6.1	7.5	7.5
D	3.4	3.2	4.4	4.6
E	1.3	1.2	1.5	2.5
No. of bank-borrower relationships	28318	29876	30633	31524

Source: Bank of Slovenia; own calculations.

Note: The statistics are calculated for the firms that were not in default in time $t-1$.

Table 3: Percentages of firms across overdue classes

Credit overdue	2007	2008	2009	2010
0 days	93.1	90.9	89.3	88.1
1-90 days	3.4	5.0	5.1	5.4
more than 90 days	3.4	4.1	5.6	6.5

Source: Bank of Slovenia; own calculations.

Note: The statistics are calculated for the firms that were not in default in time $t-1$.

Table 4 shows the distribution of firms according to their credit rating and credit overdue in particular bank in the period 2007-2010. It shows that these two measures exhibit a quite different assessment of firms' riskiness. Although according to the Regulation credit grade D should include borrowers that are in relatively bad condition or are more than 90 days overdue, around 50% of D borrowers is repaying its obligations without overdue. On the other hand, among borrowers who are more than 90 days overdue, around 5% are classified as A borrowers and 38% are classified in grades A, B or C.

Table 4: Credit ratings versus credit overdue in the period 2007-2010

Overdue in days		Credit Rating A	B	C	D	E	Total
0	Frequency	60200	35767	5786	2372	284	104409
	Row percentage	57.7	34.3	5.5	2.3	0.3	100.0
	Column percentage	96.9	91.6	73.6	50.8	14.6	90.3
1-90	Frequency	1667	2498	914	368	44	5491
	Row percentage	30.4	45.5	16.7	6.7	0.8	100.0
	Column percentage	2.7	6.4	11.6	7.9	2.3	4.8
>90	Frequency	261	768	1159	1928	1619	5735
	Row percentage	4.6	13.4	20.2	33.6	28.2	100.0
	Column percentage	0.4	2.0	14.8	41.3	83.2	5.0
Total	Frequency	62128	39033	7859	4668	1947	115635
	Row percentage	53.7	33.8	6.8	4.0	1.7	100.0
	Column percentage	100.0	100.0	100.0	100.0	100.0	100.0

Source: Bank of Slovenia; own calculations.

Note: The table reports the distribution of firms according to their credit rating and credit overdue in the 2007-2010 period. The statistics are calculated for the firms that were not in default in time $t-1$.

Banks probably use also internal soft information in determining firms' credit rating. Credit overdue is not the only measure for classifying borrowers into credit grades. Close relationship with firms can provide a more detailed information which cannot be inferred from firms' financial accounts but adds valuable information in assessing firms' creditworthiness. For this reason credit overdue and credit ratings exhibit quite different risk structure. However, the proportion of borrowers with more than 90 days overdue in high credit grades still seems quite high.

Table 5 displays credit rating transitions which are computed on one-year horizons. In 2009 when macroeconomic conditions deteriorated significantly, banks downgraded larger share of borrowers than in the pre-crises period. Comparing to 2009 only downgrades from credit grades C and D increased in 2010, whereas those form A and B decreased. There was also larger proportion of credit rating improvements in 2010. Despite the first signs of slowdown in the second half of 2008, banks upgraded higher proportion of borrowers in 2008 than a year before when GDP grew substantially. In the following sections I check what would be the change in rating structure according to the model-estimated PDs.

Table 5: Proportions of increases, decreases and no-changes of credit ratings in percentages

	Rating increased				Rating did not change				Rating decreased			
	2007	2008	2009	2010	2007	2008	2009	2010	2007	2008	2009	2010
A					89.9	91.4	87.1	87.6	10.1	8.6	12.9	12.4
B	9.6	10.1	3.8	7.3	83.0	79.2	81.9	81.2	7.5	10.8	14.3	11.6
C	21.0	21.7	11.0	15.9	68.2	65.7	71.1	63.8	10.8	12.6	17.9	20.3
D	16.5	22.8	10.8	10.1	78.7	69.6	79.2	62.5	4.8	7.6	9.9	27.5
E	8.5	12.1	2.9	3.6	91.5	87.9	97.1	96.4				

Source: Bank of Slovenia; own calculations.

Note: The table reports the percentages of credit rating transitions, calculated on one-year horizons. The statistics are calculated for the firms that were not in default in time $t-1$.

1.3 Empirical model

Credit losses are typically measured with expected loss, which is a product of probability of default, loss given default and exposure at default ($EL = PD*LGD*EAD$). While PD is counter-cyclical, recovery rates are usually pro-cyclical since the value of collateral falls in economic downturn. Bruche and González-Aguado (2010) find that macroeconomic variables are in general significant determinants of default probabilities but not so for recovery rates. They show that although variation in recovery rate distributions over time has an impact on systemic risk, this impact is small relative to the importance of the variation in default probabilities. Hence, I focus on modelling PD, which also enables me to compare estimated PDs with credit rating classification by banks.

Many different approaches for modelling default probability are proposed in the literature. Altman (1968) proposes a model which relies on firm-specific variables, like asset turnover ratio, EBIT/total assets, working capital/total assets, etc. With some modifications this approach is widely used nowadays. Instead of discriminant analysis modelling technique researchers now use logit or probit models. Since the defaults are correlated, aggregate time varying factors (like GDP growth, unemployment rate, etc.) have to be included in the models. These factors are common to all obligors and drive their credit risk into the same direction. In this respect I follow previous work by Bangia et al. (2002), Bonfim (2009), Carling et al. (2007), Jiménez and Saurina (2004) and others. Some authors, such as Festić et al. (2011), Foss et al. (2010) and Jiménez and Saurina (2006) stress another important aspect of macro effects on credit risk, arguing that strong GDP or credit growth before the crisis may have increased the share of defaulted firms or deteriorate NPL dynamics. The reason for this is that banks apply looser credit standards in expansions and thus attract more risky borrowers, which shows up during recessions when default rates rise.

Merton (1974) introduces a structural credit risk model where defaults are endogenously generated within the model. It is assumed that the default happens if the value of assets falls short of the value of liabilities. One of the model's major drawbacks is the availability of market prices for the asset value. Such data are usually not available for small and medium sized enterprises. As shown by Hamerle, Liebig and Rösch (2003), Hamerle, Liebig and Scheule (2004) and Rösch (2003) it is possible to overcome this problem with latent variable approach. They model the default event as a random variable Y_{it} which takes value 1 if firm i defaults in time t and 0 otherwise. The default event happens when borrower's return on assets, R_{it}, falls short of some threshold c_{it}. The probability that a firm i will default in time t, given the survival until time $t-1$ is described by the threshold model:

$$\lambda_{it} = P(R_{it} < c_{it} \mid R_{it-1} \geq c_{it-1}) \tag{1}$$

Equivalently this probability can be described with discrete time hazard rate model, which gives the probability that firm i defaults in time t under the condition that it did not default before time t:

$$\lambda_{it} = P(T_i = t \mid T_i > t - 1) \tag{2}$$

As discussed by Hamerle et al. (2003) it can always be assumed that the default event, Y_{it}, is observable. On the other hand the observability of the return on firm's assets, R_{it}, depends on the available data. If R_{it} is observable then the model is linear. Otherwise a nonlinear model, such as logit or probit, is estimated which treats the return on assets as a latent variable.

I estimate the probability that firm i defaults in year t given that it did not default in previous year $P(T_i = t \mid T_i > t - 1)$ using different specifications of the model:

$$P(Y_{it} = 1 \mid X_{it}, Z_t) = F(\alpha + \beta X_{it} + \gamma Z_t) \tag{3}$$

where Y_{it} is a binary variable, which takes value 1 if firm i defaults in time t and 0 otherwise, α is constant term, X_{it} is a vector of firm specific variables including also time invariant factors like sectoral dummies and Z_t is a vector of time varying explanatory variables, such as time dummies and macroeconomic effects. $F(\cdot)$ is cumulative distribution function which is standard normal distribution function $\Phi(\cdot)$ in the case of probit model and logistic distribution function $\Lambda(\cdot)$ in the case of logit model.

The estimated PDs are used in comparison to credit rating classification by banks. Banks can observe firms' riskiness in time t through monitoring process and can also observe the state of the economy. Moreover, the main criterion that banks consider in classifying borrowers in credit grades is credit overdue, which is available to banks regularly in time

t. This means that in time t banks have a large set of information to decide about firms' creditworthiness. To ensure that I'm using the same set of information as available to banks in time t, all the variables are included in the model at their values in time t, with few exceptions.

To estimate $P(Y_{it} = 1 \mid X_{it}, Z_t)$ I apply random effects probit model. This estimator is most often used in other research and is the underlying model in Basel II risk assessment procedures. Hamerle et al. (2003) show that when only defaults are observable, an appropriate threshold model leads to random effects probit or logit model.

I use the measures of goodness of fit described by BCBS (2005) and Medema, Koning and Lensink (2009). The most often used method for determining the discrimination power of binary models is Receiver Operating Characteristics (ROC) curve. It is obtained by plotting hit rate (HR) against false alarm rate (FAR) for different cut-off points. HR is percentage of defaulters that are correctly classified as defaulters and FAR is percentage of non-defaulters incorrectly classified as defaulters. The area under this curve indicates that the model is noninformative if it is close to 0.5 and the closer it is to 1, the better the discriminating power of the model.

The Brier Score is defined as $BS = \frac{1}{N}\sum_{i=1}^{N}(\widehat{PD}_i - Y_i)^2$ where $\widehat{PD}_i$ is estimated probability of default. As explained by Medema et al. (2009) it can be interpreted as the mean of the sum of squares of the residuals. The better the model, the closer BS is to zero.

Finally, pseudo R^2 is based on log-likelihood values of estimated model (L_1) and a model which contains only constant as explanatory variable (L_0): Pseudo $R^2 = 1 - \frac{1}{1+2(\log L_1 - \log L_0)/N}$. I also use Likelihood Ratio (LR) test, which enables to compare two models of which one is nested into the other. It is defined as $LR = 2[\log L(\widehat{\theta}) - \log L(\widetilde{\theta})]$, where $\log L(\widehat{\theta})$ and $\log L(\widetilde{\theta})$ are log-likelihoods of unrestricted and restricted models, respectively.

1.4 Results

In the first part of this section, I present the estimation results of different credit risk model specifications. Estimated PDs from the two model specifications that best fit the data are then used in the second part in the comparison of estimated PDs to credit rating classification by banks. In the third part I check the robustness of the obtained results by excluding a variable that measures number of days a firm has blocked bank account from the model.

1.4.1 Estimation results

Table 6 shows the results of random effects probit models with various firm specific variables, sector dummies and time effects. The estimated coefficients are divided into firm-specific variables, sectoral dummies and time effects. Robust standard errors are used in all the estimates.

The basic model is given in first column of Table 6. It includes only firm specific variables. All coefficients are different from zero at 1% probability and display the expected sign. Total sales displays a negative coefficient, suggesting that larger firms have lower probability of default. Size of a firm is in many researches found as one of the most important ingredients of credit risk models, since smaller firms are in principle less diversified, have lower net worth and are more financially constrained. Similar result is also found for firm age, which indicates that younger firms who are usually more sensitive to shocks default more often.

Quick ratio, which is an indicator of liquidity, measures the ability of firm to use its quick assets (current assets minus inventories) to meet its current liabilities. As expected, firms with higher liquidity ratios have lower default probabilities. Defaulted firms are generally expected to have more debt in their capital structure. The positive sign on the coefficient for debt-to-assets ratio in the model clearly indicates that firms with higher leverage default more often. Cash flow, which is a ratio between operating cash flow and revenues, displays a negative coefficient. It is expected that stable, mature and profitable firms generate sufficient cash flows to pay off the owners and creditors. Asset turnover ratio measures firm's efficiency in generating sales revenues with assets. The estimated coefficient indicates that firms that are more efficient default less often. Number of days a firm has blocked bank account also seems to offer an important contribution in explaining firm's credit default. The longer the firms have blocked bank account in a given year, the higher the probability of default.

Number of bank-borrower relationships displays highly statistically significant coefficient with positive sign, which is contrary to the findings of Jiménez and Saurina (2004) and indicates that those firms with more credit relationships have on average higher default probability. This result suggests that less creditworthy firms seek for credit in more banks, possibly because current creditors don't want to lend them any more or are only prepared to grant smaller amount of credit due to their riskiness.

I now extend the model with aggregate variables, i.e. the sectoral and time dummies. Many authors like Antão and Lacerda (2011) and Crouhy et al. (2001) suggest taking into account the features of the industry when modelling credit risk. In my sample defaulters and non-defaulters are similarly distributed across sectors, with the highest representativeness of Commerce (28%), Manufacturing (18%), Professional activities (17%) and Construction (11%). By including sectoral dummies in model (2), the dummy variable for manufacturing firms is omitted, so that the coefficients for other sectors indicate the relative riskiness of a particular sector in relation to manufacturing one. Year dummies (omitting the dummy variable for 2007) are capturing the time effects. It is wider

Table 6: Estimated PD models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Firm variables							
Total sales	-0.017***	-0.019***	-0.018***	-0.018***	-0.018***	-0.018***	-0.019***
Firm age	-0.014***	-0.013***	-0.014***	-0.013***	-0.014***	-0.013***	-0.013***
Quick ratio	-0.042***	-0.036***	-0.041***	-0.036***	-0.035***	-0.034***	-0.036***
Debt-to-assets	0.540***	0.560***	0.540***	0.554***	0.542***	0.547***	0.559***
Cash flow	-0.138***	-0.137***	-0.137***	-0.135***	-0.138***	-0.140***	-0.136***
Asset turnover r.	-0.268***	-0.278***	-0.270***	-0.274***	-0.271***	-0.274***	-0.277***
Blocked account	0.007***	0.008***	0.007***	0.008***	0.007***	0.008***	0.008***
No. of relations	0.363***	0.379***	0.369***	0.375***	0.368***	0.371***	0.378***
Sectoral dummies							
Agric., Forestry, Fish. and Mining		0.070	0.067	0.068	0.067	0.063	0.068
Electricity, gas and water supply		-0.404***	-0.391***	-0.395***	-0.391***	-0.393***	-0.399***
Construction		-0.001	-0.000	-0.001	-0.000	0.002	-0.000
Commerce		-0.045	-0.045	-0.045	-0.045	-0.043	-0.044
Tran. and storage		0.071	0.068	0.071	0.067	0.070	0.072
Accomodation and food service		0.168***	0.156***	0.166***	0.157***	0.161***	0.169***
Inf. and commun.		-0.246***	-0.241***	-0.246***	-0.238***	-0.241***	-0.248***
Fin. and insur.		-0.304*	-0.301*	-0.299*	-0.303*	-0.298*	-0.298*
Real estate		0.087	0.077	0.084	0.080	0.083	0.086
Professional act.		-0.169***	-0.164***	-0.168***	-0.164***	-0.164***	-0.169***
Public services		-0.206***	-0.201***	-0.203***	-0.198***	-0.201***	-0.206***
Time effects							
2008		0.212***					
2009		0.173***					
2010		0.209***					
GDP growth			-0.011***				
GDP growth (t-1)						-0.038***	
NFC loan growth				-0.005***			-0.007***
NFC loan g. (t-1)						0.018***	
Interest rate					0.024***		0.046***
Quick r.*GDP g.			0.006***			-0.001	
Constant	-2.426***	-2.635***	-2.403***	-2.414***	-2.401***	-2.654***	-2.381***
Observations	68444	68444	68444	68444	68444	68444	68444
Pseudo R^2	0.094	0.096	0.095	0.095	0.095	0.095	0.096
Log. lik.	-8382.7	-8313.6	-8331.2	-8326.7	-8332.5	-8328.9	-8321.1
LR test	-	138.4	103.1	112.1	100.4	107.6	123.3
AUC	0.888	0.890	0.889	0.889	0.890	0.890	0.889
Brier score	0.030	0.029	0.030	0.029	0.030	0.030	0.029

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the estimated probit models for the period 2007-2010, where the dependent variable is the indicator for firm's default, based on credit overdue. Blocked account is a number of days a firm has blocked bank account, No. of relations measures to how many banks a particular firm is related to. GDP growth is in real terms. NFC credit growth is real growth of loans to non-financial corporations, Interest rate is long-term interest rate on loans to non-financial corporations, AUC is area under ROC curve. Robust standard errors are used. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

category than macroeconomic variables, which will be added in further specifications, since it also captures institutional, regulatory or any other systemic factors that affect all firms. Although some of the sectoral dummies are insignificant, it is clear that there are some differences in credit risk across sectors. Sectors like electricity, gas and water supply, information and communication, professional activities and public services are less risky than manufacturing, whereas only accommodation and food service have on average higher statistically significant default probability. By adding both sectoral and time dummies coefficients of firm specific variables are changed only slightly, which indicates that these two set of aggregate variables are close to independent from firm specific effects. According to likelihood ratio test, sector and time dummies improve the fit considerably comparing to model (1).

Since the default rate is highly related to the business cycle - increasing in economic downturns - a set of macroeconomic and financial variables is included in models (3) to (7). GDP growth as the main indicator of economic activity is added in model (3). The estimated coefficient suggests that higher economic activity lowers the probability of default, because better macroeconomic situation enables a better performance of all firms. The only significant interaction effect between GDP growth and firm specific variables is the one with the quick ratio, which shows how the effect of liquidity changes with one percentage point increase in GDP growth and vice versa. Similar result is also found in model (4) where growth of credit to non-financial corporations is used as an alternative indicator of the business cycle. According to the likelihood ratio test, credit growth actually seems to be a more powerful business cycle variable for explaining default probability than GDP growth. The interest rate on bank loans is also expected to have an important influence on the borrowers' ability to repay loans. As suggested by the coefficient on interest rate in model (5), a higher interest rate leads to a higher probability of default, which also make sense, since it increases borrowers' credit burden.

Among macroeconomic variables, the credit growth seems to have the highest explanatory power in turns of default probabilities. When credit growth and interest rates are put together, as in model (7), it further improves the fit as can be seen from the likelihood ratio test statistic. I also estimate models with different combinations of business cycle indicators, but many of them were found insignificant or with unexpected sign. Short time series does not allow to include many variables that vary in time and are constant for all firms.

Model (6) includes GDP and credit growth lagged one year. Lagged GDP growth exerts a negative effect on probability of default, as in contemporaneous case, although the displayed coefficient is now higher in absolute terms. On the other hand, lagged credit growth displays a positive coefficient, which suggests that high past credit growth increases probability of default, as expected. When economic situation turns around, as it did in

2009-2010, and risk premium starts rising due to the tightening credit standards, these borrowers quickly get into trouble and may default on their credit obligations.

1.4.2 The comparison of estimated PDs to credit rating classification by banks

As the estimated PD exhibit a measure of risk conditional on a large set of available information, it is interesting to compare it to the credit ratings by banks. Credit ratings indeed exhibit the banks' assessment of debtors' ability to repay the debt. For the purpose of comparison, I link firms' probabilities of default with all credit ratings by banks. Since PDs are estimated at firm level, a particular firm represents the same level of risk to all banks that have exposure to this firm.

To select the model specification for this analysis I use root-mean-square error, which is defined as $RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(DR_{Pt} - DR_{At})^2}$, where DR_{Pt} and DR_{At} are predicted and actual default rate in time t, respectively. Table 7 shows that the in-sample predicted default rate from model (2), which includes year dummies as time effects, is the most unbiased. This result might be expected since time dummies do not only capture the macroeconomic dynamics but also other institutional, systemic or regulatory changes. Among models with macroeconomic variables, model (7), which includes credit growth and interest rate as business cycle effects, is the most accurate. Since these two models give the most unbiased in-sample predictions for the default rate and have high overall classification accuracy rate (96.3%) I use them in the comparison to the banks' risk grades.

Table 7: Actual vs. in-sample predicted default rate

	2007	2008	2009	2010	RMSE
Actual default rate	3.39	3.95	4.80	4.80	
In-sample predicted default rate					
Model 1	4.07	3.56	4.58	4.36	0.46
Model 2	3.38	3.92	4.71	4.67	0.08
Model 3	3.89	3.46	4.94	4.31	0.43
Model 4	3.66	3.40	4.90	4.68	0.32
Model 5	3.80	3.65	4.98	4.18	0.41
Model 6	3.80	3.82	4.29	4.73	0.34
Model 7	3.58	3.57	4.59	4.92	0.24

Source: Bank of Slovenia; AJPES; SURS; own calculations.

Note: In-sample predicted default rate is calculated as average of firms' PDs.

Table 8 shows the distribution of firms according to their credit ratings and the level of estimated PD in the period 2007-2010. In all credit grades, except E, the majority of firms have PD between 1 and 5 percent. Although we would expect borrowers in credit grade D to have high PDs on average, 43% have PD below 5%. Among high-risk borrowers with PD above 50%, around 13% are classified as A borrowers and approximately 57% are classified in grades A, B or C. This results are similar as those in Table 4 where instead of PD, the distribution is done according to credit overdue.

Table 8: The distribution of firms among PD buckets and credit grades

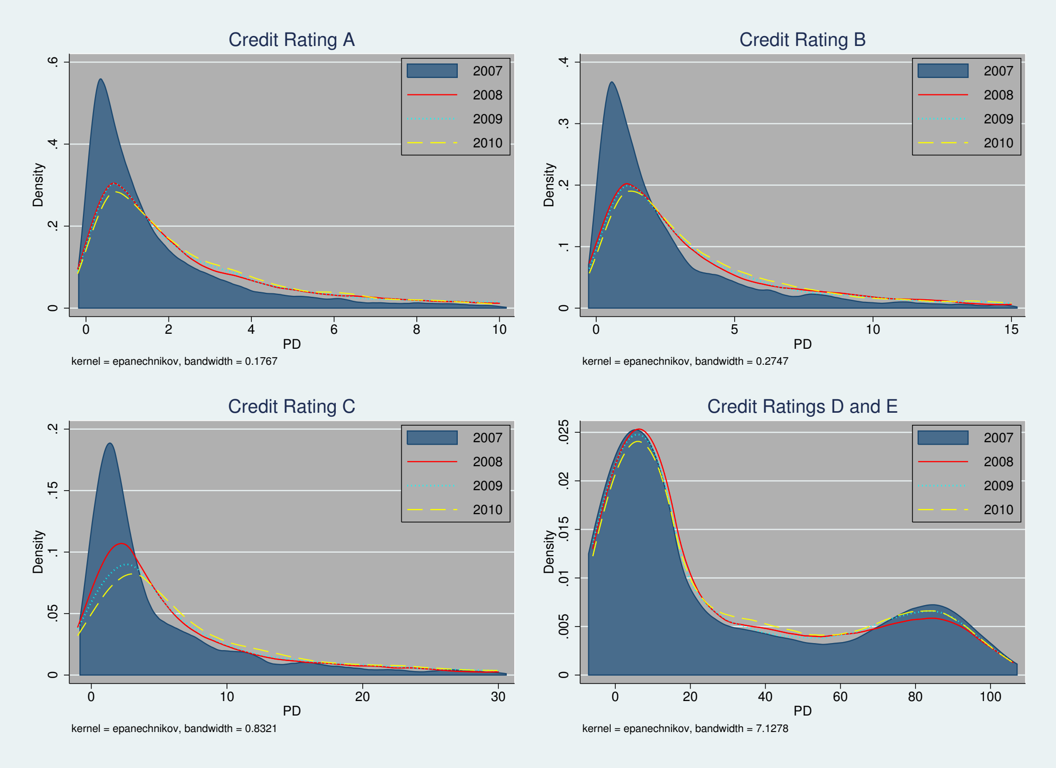
	Credit Rating					
PD	A	B	C	D	E	Total
			Model 2			
$PD \leq 1$	18641	7596	868	399	16	27590
$1 < PD \leq 5$	24621	15815	2296	808	57	43597
$5 < PD \leq 10$	5669	4727	1062	406	58	11922
$10 < PD \leq 25$	2992	2944	933	383	63	7315
$25 < PD \leq 50$	748	838	395	285	68	2334
$PD > 50$	202	368	334	504	178	1586
			Model 7			
$PD \leq 1$	18519	7488	855	396	13	27271
$1 < PD \leq 5$	24776	15899	2317	815	62	43869
$5 < PD \leq 10$	5711	4784	1070	386	53	12004
$10 < PD \leq 25$	2949	2938	924	398	66	7275
$25 < PD \leq 50$	710	795	398	280	70	2253
$PD > 50$	208	384	324	510	176	1602
Total	52873	32288	5888	2785	440	94274

Source: Bank of Slovenia; AJPES; SURS; own calculations.

Note: The table reports the number of firms distributed among PD buckets and credit grades for 2007-2010 period.

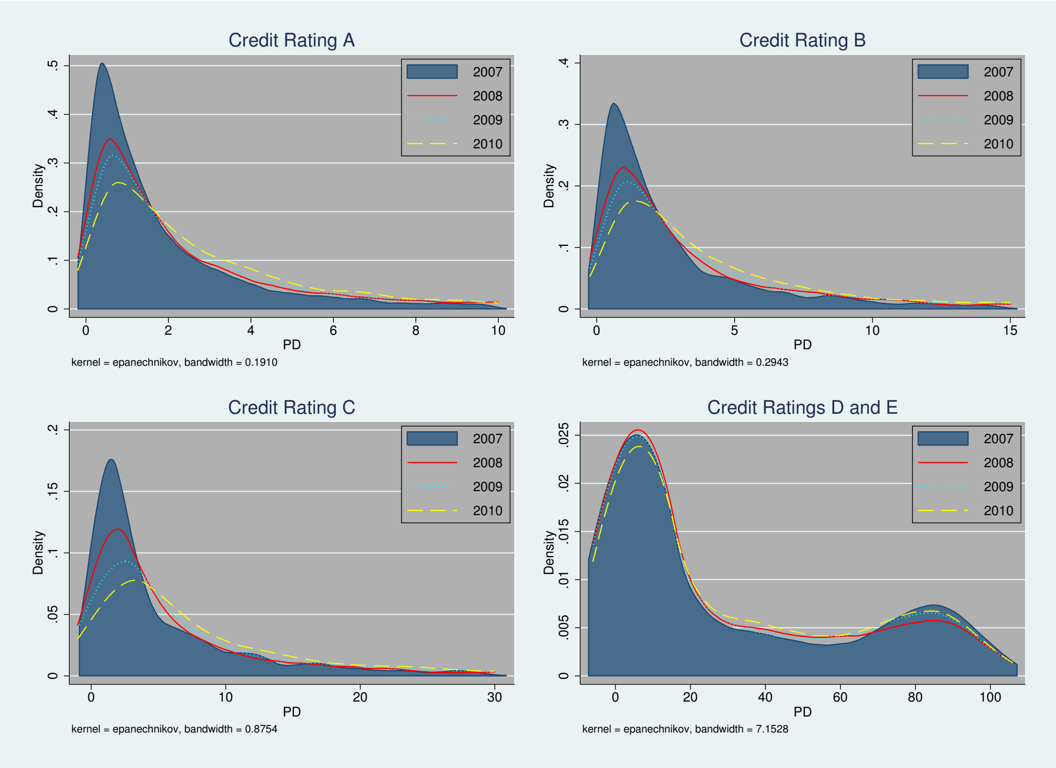
Estimated PDs allow us to test whether banks' rating criteria were constant in time. If banks use unique criteria to assess borrowers riskiness, the risk structure in terms of PDs of firms in each credit grade should be stable in time. Figure 2, Figure 3 and Table 9 indicate that the risk structure was changing in time, particularly in credit grades A, B and C. This can be best seen from the changing shapes in distributions of the estimated PDs in different credit grades. This holds for model (2) estimates (Figure 2) as well as for model (7) estimates (Figure 3), with the largest change in distribution between years 2007 and 2008. In Table 9, average default probability in credit grades A, B and C rose by 1.3, 1.5 and 1.5 percentage point, respectively, as estimated with model (2) and by 0.7, 0.7 and 0.5 percentage point, respectively, as estimated with model (7). This trend continued also in 2009 and 2010 where especially for credit grades A and B model (7) gives more pronounced results. Risk structure deteriorated the most in credit grade C, where the average default probability estimated with models (2) and (7) increased by 4.4 and 4.6 percentage point, respectively, from 2007 to 2010. Somehow surprisingly, in credit grades D and E the average estimated PD actually decreased in 2008. It is possible that this result is driven by small number of borrowers in credit grades D and E.

Figure 2: Kernel densities of PDs estimated with Model 2, by credit rating



Source: Bank of Slovenia; AJPES; SURS; own calculations.

Figure 3: Kernel densities of PDs estimated with Model 7, by credit rating



Source: Bank of Slovenia; AJPES; SURS; own calculations.

Table 9: Summary statistics for PDs estimated with Model 2 and Model 7, by credit rating

	Model 2						Model 7					
	Mean	Std.	P50	P90	Skew.	Kurt.	Mean	Std.	P50	P90	Skew.	Kurt.
Credit Rating A												
2007	2.6	5.8	1.1	5.8	7.8	91.1	2.8	5.9	1.2	6.2	7.4	83.3
2008	3.9	6.7	1.8	9.0	5.0	39.8	3.5	6.3	1.6	8.1	5.3	44.5
2009	4.0	7.3	1.8	9.0	5.3	41.4	3.9	7.2	1.8	8.7	5.4	42.6
2010	3.9	6.6	2.0	8.5	5.4	46.2	4.2	6.8	2.1	9.1	5.2	42.7
Credit Rating B												
2007	4.1	8.9	1.6	8.7	5.7	44.0	4.4	9.1	1.8	9.3	5.6	41.6
2008	5.6	8.9	2.6	12.8	4.2	26.4	5.1	8.5	2.3	11.7	4.4	29.2
2009	6.1	10.4	2.7	14.1	4.2	25.8	5.9	10.3	2.6	13.7	4.3	26.6
2010	5.9	9.7	2.9	13.3	4.5	28.6	6.2	9.9	3.1	14.1	4.3	27.1
Credit Rating C												
2007	8.7	15.9	2.7	22.6	3.4	15.4	9.0	16.2	2.9	23.7	3.3	14.8
2008	10.2	15.6	4.1	26.6	2.7	10.9	9.5	15.1	3.7	24.8	2.9	11.6
2009	12.7	19.1	5.0	38.4	2.4	8.4	12.4	18.9	4.8	37.4	2.4	8.6
2010	13.1	18.8	5.7	37.6	2.4	8.7	13.6	19.0	6.1	38.8	2.4	8.5
Credit Rating D												
2007	22.2	29.9	5.8	79.1	1.3	3.2	22.7	30.1	6.3	80.0	1.3	3.1
2008	16.9	22.8	6.4	55.9	1.7	4.7	16.0	22.2	5.7	53.6	1.7	5.0
2009	21.2	27.6	6.2	72.5	1.3	3.4	20.9	27.4	6.0	71.6	1.4	3.5
2010	23.2	28.0	8.9	74.2	1.2	3.1	23.8	28.2	9.4	75.2	1.2	3.0
Credit Rating E												
2007	42.9	36.0	34.1	92.7	0.2	1.4	43.5	36.1	35.2	93.1	0.2	1.4
2008	30.9	27.9	18.5	75.1	0.7	2.2	29.4	27.4	17.0	73.1	0.8	2.3
2009	47.0	35.1	40.3	91.6	0.1	1.3	46.5	35.0	39.6	91.2	0.1	1.3
2010	42.5	33.1	33.5	92.6	0.4	1.7	43.2	33.1	34.4	92.7	0.3	1.7

Source: Bank of Slovenia; AJPES; SURS; own calculations.

Notes: P50 and P90 are 50^{th} and 90^{th} percentile, respectively. Std., Skew. and Kurt. are abbreviations for standard deviation, skewness and kurtosis.

To get a more clear insight in comparing risk evaluations I check what would be the model-predicted rating structure if banks would keep constant rating criteria in time. To be able to do this I need to predict credit ratings by setting threshold PDs between each credit grade. Since there is a lot of overlapping in default probability between credit ratings, perfect discrimination is not possible. Hence, I set the cut-off PDs so as to ensure that the predicted rating structure in a particular date is equal to actual one. I use as a point of reference first 2007 and then 2008. Thus for 2007, I classify the top 56.2% in terms of PDs of firms as A borrowers, next 34.1% as B and so on. In this way, rating structure does not change, but the actual and predicted structure of borrowers in each credit grade is quite different. I repeat this in predicting credit ratings based on rating structure in 2008.

Table 10 shows the actual and predicted rating structures based on estimates with models (2) and (7). I focus on the crises years 2009 and 2010. Based on the estimated default probabilities with model (2), the proportion of A borrowers should have been lower for 15.8 percentage points in 2009 and 17.4 percentage points in 2010 if banks would apply the same rating criteria as in 2007. Similar results are also found with model (7), although with slightly better predicted rating structure in 2009. Using thresholds from 2008, predicted rating structures based on model (2) are almost equal to actual ones. On the other hand, based on model (7), the proportion of A borrowers should have been 6.6 percentage points lower in 2010.

Table 10: Actual vs. predicted rating structure in percentages

			Model 2				Model 7			
	Actual		Cut-off 2007		Cut-off 2008		Cut-off 2007		Cut-off 2008	
	2009	2010	2009	2010	2009	2010	2009	2010	2009	2010
A	55.4	54.3	39.5	36.9	56.4	54.7	43.9	37.4	53.8	47.6
B	34.3	35.9	45.3	48.4	33.3	35.4	42.8	48.0	35.2	40.5
C	7.1	6.4	8.8	8.8	6.6	6.5	7.6	8.7	7.1	7.9
D	3.0	3.2	5.6	5.3	3.0	2.7	5.0	5.3	3.1	3.1
E	0.3	0.3	0.8	0.7	0.8	0.7	0.7	0.7	0.8	0.9

Source: Bank of Slovenia; AJPES; SURS; own calculations.

Notes: The table reports the actual and predicted rating structure for the firms included in the model.

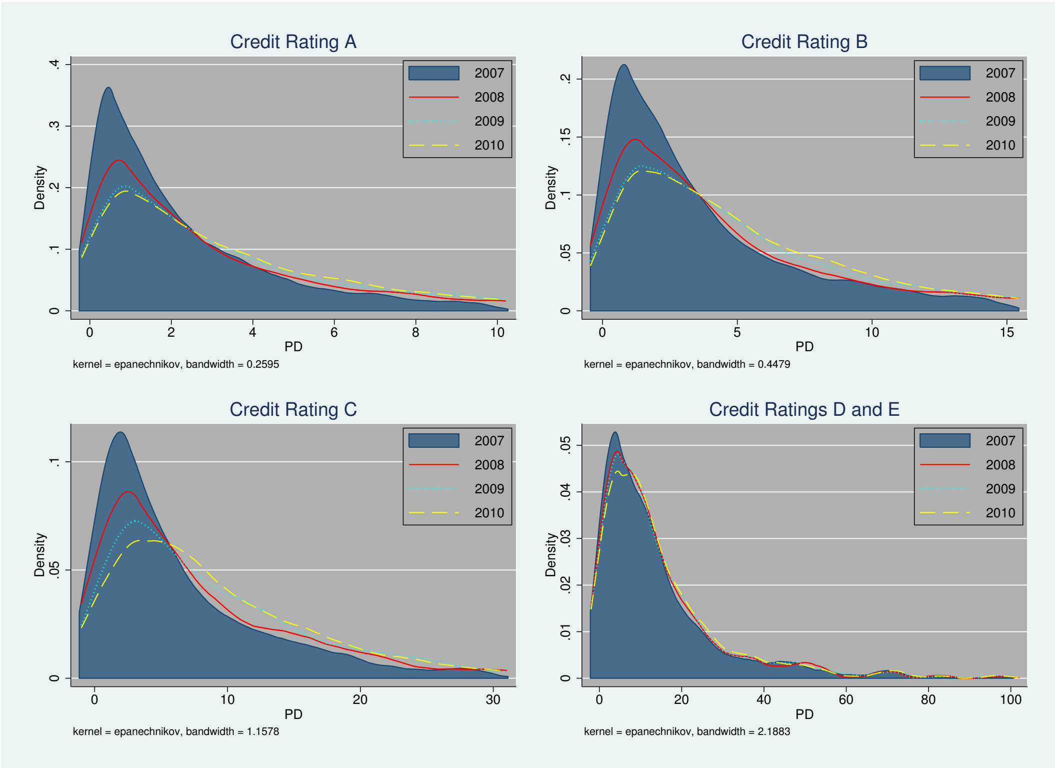
Besides the indication, that the risk assessment strategy by banks might have significantly changed over time, one could interpret these results in two ways. On the one hand, the banks risk classification might underestimate the underlying risk structure, with the risk grades attributed being too high. This could be due to a possible underestimation of the underlying risk. In particular, systemic risk factors might be more accurately captured in the model, which includes the macroeconomic factors, that drive average default probability over the business cycle. On the other hand, this could also be due to banks taking into account additional borrower-related information, e.g. the information gathered through bank-borrower relation, or to the lags in reclassification process.

1.4.3 Robustness check

To test the validity of the obtained results I exclude the variable that measures the number of days a firm has blocked bank account from the model. This variable could be a source of endogeneity bias since both, the dependent variable, which is based on credit overdue, and number of days with blocked account are measures of default and both are dated in time t in the model.

I reestimate model (2) by excluding the Blocked account. All the estimated coefficients are highly statistically significant and display the expected sign. Discriminating power of the model, measured with area under ROC curve, is slightly lower and is equal to 0.83. As before, I use the estimated PDs in the comparison to credit rating classification by banks. As shown in Figure 4, the results are similar as before and indicate that in the crisis the risk structure of borrowers in credit grades A, B and C deteriorated. The largest change in the distribution is between years 2007 and 2008, when average default probability in credit grades A, B and C rose by 0.8, 1.0 and 0.9 percentage point, respectively. Although there are some differences in the shapes of distributions of estimated PDs, the results seems to be robust also when Blocked account is excluded from the model.

Figure 4: Kernel densities of PDs estimated with Model 2, excluding number of days with blocked account



Source: Bank of Slovenia; AJPES; SURS; own calculations.

1.5 Conclusion

This paper uses the data on the characteristics of non-financial firms, which have credit obligations to Slovenian banks. In the first part, I estimate several credit risk models, which suggest that probability of default can be explained with firm specific factors as well as macroeconomic or time effects. I find the model that includes year dummies as time effects to perform slightly better than models with macroeconomic variables. This result is expected, since time dummies are much broader category that also capture institutional, regulatory or other systemic changes in time.

Estimated PDs from the two models that fit the best are used in the comparison to credit rating classification. I link estimated firms' PDs with all their relations to banks and analyse banks' classification of borrowers into credit grades. I find that PDs and credit ratings often imply a quite different measure of debtors' creditworthiness. Similar result is also found by using credit overdue instead of estimated PDs. By looking at PD densities for each credit rating, I find that in the crisis banks allow for higher risk borrowers in credit grades A, B and C. This could be due to banks taking into account additional borrower-related information, to the lags in reclassification process or a possible underestimation of systemic risk factors by the banks.

One of the shortcomings of the estimated models is short time series. Problematic can be especially coefficients of macro variables, which are based on only four observations. Nevertheless, a supportive argument for the validity of the estimated PDs is that they remain very similar when time dummies are used instead. Also the coefficients on firm-specific determinants of PD all exhibit expected signs, are highly statistically significant and are very stable across different model specifications.

2 Discretionary Credit Rating and Bank Stability in a Financial Crisis[2]

2.1 Introduction

Existing accounting standards and banking regulation induce procyclicality in the loan-loss provisions through capital adequacy requirements.[3] In economic downturns the incidence of corporate default increases and the value of banks' assets decreases. Resulting higher loan-loss provisions negatively reflect in the profit and loss account and, consequently, in bank capital, which creates an incentive for banks to relax the standards of credit risk assessment and valuation of assets in times of economic downturn.[4] In a financial crisis raising additional capital to meet with minimum capital requirements is particularly difficult, which only increases the incentives for discretion in credit risk assessment and leads to systematic underestimation of credit risk of loan portfolios.

To overcome the problems with underestimation of credit risk bank regulators could in principle apply discipline on banks to comply with regulatory standards. In a financial crisis, however, financial regulators often apply forbearance in order to partly alleviate the problems with procyclicality of capital requirements and prevent significant disruptions in the banking system (Hoffman & Santomero, 1998).

This paper studies how the current regulatory framework amplifies incentives for discretionary credit risk assessment in times of financial distress by studying the case of Slovenia in the Great recession after 2008. In 2013, ten Slovenian banks, accounting for approximately 70% of total bank assets, went through a comprehensive review, consisting of asset quality review (AQR henceforth) and stress tests, performed by independent external examiners using uniform methodology. The results, announced publicly in December 2013, revealed significant capital shortfalls.[5] In particular, estimated capital shortfalls of all examined banks amounted to 214% of existing capital, with some pronounced differences across banks.[6] Namely, there were stark differences in the capital shortfalls

[2] Brezigar-Masten A., Masten I. & Volk M. (2015). Discretionary Credit Rating and Bank Stability in a Financial Crisis. *Eastern European Economics, 53*(5), 377-402.
First person plural is used for narration.

[3] This feature is present both in banks using the internal ratings based or the standardized approach under Basel II.

[4] Such incentives might be amplified if the regulators make the minimum capital requirements stricter in times of economic downturn. Such was the case of the European Banking Authority, who with the aim of boosting confidence in the European banking system in 2011 set the provision that a minimum of 9% of risk-weighted assets should be held in the form of Core Tier 1 capital.

[5] For details see the Full report on the comprehensive review of the banking system, Bank of Slovenia (2013b).

[6] Such capital shortages indicate potential problems with insolvency. In fact, for two small private domestically owned banks the central bank initiated insolvency procedures already before the results of the comprehensive review were published. These two banks were subsequently excluded from the stress test part of the comprehensive review.

between domestic and foreign owned banks. For the former the recapitalisation requirement amounted to 244% of existing capital, while for the latter this figure was "only" 78%. Also within the group of domestic banks we can find important differences. The recapitalisation requirement for the largest two and majority state owned banks on the market, holding 36% of total assets, amounted to 228% of existing capital, while for the small and predominantly privately owned the figure was 274%.

Our working hypothesis is that the incentive to underestimate credit risk is an important factor behind the estimated differences in capital shortfalls. We evaluate this hypothesis in a rather unique historical episode. The Slovenian banking system went through a comprehensive review a year before similar comprehensive reviews became standard practice in the Single Supervisory Mechanism of the Euro area. The comprehensive review revealed large capital shortfalls, which can be almost entirely attributed to underestimation of credit risk in classic loan portfolios of non-financial enterprises. This provides us with a unique environment to study the incentives to underestimate credit risk.

The empirical test of our hypothesis is based on predictive capacity of credit ratings - assigned by banks to their non-financial corporate clients - in predicting financial distress of banks' clients. Given that banks need to provide on average more loan loss provisions for lower rated borrowers, the credit ratings clearly reflect banks incentives to underestimate risk. We estimate how the predictive ability of credit ratings evolves over the period 2007-2012 and across groups of banks. The predictive capacity of credit ratings is compared to that of a conventional econometric (logit) model that uses only publicly available information, such as various financial ratios of banks' clients. Such a model can be used to predict financial distress by an econometrician free of any incentives to underestimate risk and as such serves as our benchmark of comparison.

Our results indicate that the precision of bank ratings in predicting financial distress deteriorated during the crisis, both in absolute terms and, more importantly, against the predictive capacity of the econometric model using publicly available information only. While the loss of predictive capacity of financial ratios in general can be rationalised by a potential structural break induced by the crisis, our finding that predictive capacity of credit ratings on average deteriorated even relative to the econometric model, reflects the fact that, given mounting non-performing loans, banks had an incentive to underestimate credit risk and inflate their balance sheets by keeping lower levels of loan-loss provisions. The classification accuracy of credit ratings assigned by foreign-owned banks outperforms the predictive capacity of credit ratings assigned by domestic banks by a large margin. Within the group of domestic banks we also observe differences between large and state owned, and small banks. The latter group reveals the worst predictive capacity of credit ratings. These conclusions appear to be robust to the effects of the timing of public release of corporate balance-sheet data and prediction horizon.

The results on predictive capacity of credit ratings across groups of banks - domestic/ foreign, private/state owned or large/small - align with the results of estimated capital shortfalls in the 2013 comprehensive review in Slovenia. Banks with the largest capital shortfalls were the ones with the least reliable credit ratings as predictors of default, and vice versa. Predictive capacity of credit ratings can be significantly affected by the incentives to underestimate risk in relation to the ability to raise capital and funding in times of financial distress. Small banks with financially weak owners faced the most pronounced difficulties in raising capital and thus had stronger incentives to apply discretion in risk assessment. Foreign-owned banks, at the other end of the spectrum, had access to the internal capital markets of international banking groups they belong to, and thus had access to more stable sources of funding (Navaretti, Calzolari, Pozzolo & Levi, 2010; de Haas & van Lelyveld, 2010). Smaller incentives to under estimated risk by foreign banks can stem also from superior managerial and organizational capacity to absorb losses. Large domestic banks fall in between as they enjoyed the bail-out guarantee by the government.

Our work is related to the paper by Huizinga and Laeven (2012), who for the case of the US mortgage crisis report significant discrepancies between market value and banks' valuation of real-estate related securities. These differences are attributed to the use of discretion over classification of mortgage-backed securities with the aim of inflating banks' books. Moreover, Huizinga and Laeven (2012) notice that the over-valuation of distressed assets is more pronounced if banks are bigger and exhibit higher exposure to these assets.

While addressing a similar topic, our paper is different. Our focus is not on discretion in banks' valuation of assets, but on discretion in credit risk assessment. Our approach is novel in the literature in the sense that we use credit ratings and explore the information content through implied probabilities of default. In addition, our analysis is applied to classic loan portfolios, which constitute the majority of bank assets in continental Europe. Such assets are not traded. For this reason we cannot rely on market information to infer about their implicit riskiness. Instead, we use econometric analysis.

We study the case of a Euro area member country. On one hand, this is instructive as we study the behaviour of banks that operated in the same market and were exposed to same systematic risks and the same regulatory environment. This means that we can abstract from cross-country differences in factors that are outside of banks' control. On the other hand, results of our analysis are relevant also for the advanced banking systems. A recent paper by Mariathasan and Merrouche (2014) shows that the reported riskiness of banks decreased upon the adoption of the IRB. They find this effect to be especially pronounced among weakly capitalised banks, which have higher incentives to under-report actual riskiness. A similar conclusion is also suggested by Blum (2008), who analyses the effectiveness of regulatory risk-sensitive capital requirements in an adverse selection model. The study by Brown and Dinc (2011) provides empirical evidence that the likelihood of regulatory forbearance is indeed higher when the banking system is weak. These studies suggest that the incentives to underestimate risk exist in banking systems of a wide variety of countries. The methodology we propose to study the incentives to underestimate credit risk is thus applicable quite generally.

Last but not least, the Slovenian case is instructive for the comprehensive review of the Euro-area banking system prior to the launch of the Single Supervisory Mechanism, conducted in 2014, a year later than the comprehensive review in Slovenia. In this review the results also revealed that capital shortages in smaller banks exceeded those in large banking groups by a significant margin. Based on our analysis, it can be argued that the differences in estimated capital shortfalls can be, at least to some extent, attributed to systematic underestimation of credit risk due to incentives to do so in time of financial distress.

The rest of the paper is organised as follows. Section 2.2 provides stylised facts incentives for discretionary risk assessment in the Slovenian case. Section 2.3 presents our modelling approach for testing for discretion in credit risk assessment. Section 2.4 presents the main results. Section 2.5 contains the robustness checks, while Section 2.6 concludes and discusses policy implications.

2.2 Incentives for discretionary risk assessment

To highlight the incentives for discretionary credit risk assessment embedded in the regulatory system, this section provides basic empirical evidence about key developments in the Slovenian banking system in the period 2006-2012. We look at how the credit rating structure of banks' portfolios evolved during the Great recession and what were the corresponding dynamics of loan-loss provisions. In addition to looking at the banking system as a whole, we provide descriptive statistics across groups of banks, divided according to size and ownership.

The source of credit ratings data is the Credit Registry of the Bank of Slovenia, which contains bank-borrower information. These credit ratings represent banks' subjective assessment of firms' creditworthiness. Each bank has its own methodology for estimating borrowers' riskiness, which should at the end be transformed into five-grade rating scale (from A to E) set by the Bank of Slovenia in the Regulation. Banks classify borrowers into credit grades based on the assessment of their financial position, the ability to provide sufficient cash flow to regularly fulfil the obligations to banks, and information on the borrowers' potential arrears on loan repayments. The latter piece of information is regularly available to the banks and in practice carries significant role in determining the credit rating. The credit ratings are independent of the pledged collateral and thus give the assessment of the quality of the borrower and not necessarily of the quality of bank's claims to this borrower.

We combine the Credit Registry data with the balance sheet and income statement data for all Slovenian firms, collected by the Agency of the Republic of Slovenia for Public Legal Records and Related Services (AJPES) at yearly basis.

In the fourth quarter of 2008 Slovenian economy entered a deep recession. From the peak in the third quarter of 2008 to the end of 2012 the cumulative loss of real output exceeded 9%. This protracted economic slump reflected also in the quality of banks assets and corresponding credit rating structure of banks' borrowers. Table 11 shows how the economic and financial crisis resulted in a deteriorated structure of credit ratings of borrowers. The share of A-rated borrowers had dropped by 14.4 percentage points between 2006 and 2012. On the other hand, the share of the worst performing borrowers rated D or E increased by 6.8 percentage points. The deterioration is even more significant if we look at bottom panel of Table 11 that reports the rating structure weighted by the banks' exposure to borrowers. According to this measure the share of A-rated borrowers decreased by almost 32 percentage points, while it increased by more than 10 points for C-rated, 6.6 points for D-rated and 13.2 percentage points for E-rated clients respectively. The difference between the upper and bottom panel of of Table 11 reveals that credit risk was concentrated in large credit exposures.

Table 11: Credit rating structure over the business cycle

Rating	2006	2007	2008	2009	2010	2011	2012
% of borrowers							
A	54.0	53.1	53.5	48.9	46.3	41.7	39.6
B	30.6	32.0	31.0	32.0	32.5	34.0	34.3
C	5.1	5.2	6.1	7.5	7.4	8.0	9.0
D	4.7	4.5	4.6	6.3	5.6	5.5	5.6
E	5.6	5.2	4.8	5.3	8.2	10.8	11.5
% of borrowers weighted by loan size							
A	71.7	70.0	67.8	57.4	52.0	44.8	39.8
B	22.1	24.8	26.5	29.4	25.7	25.4	23.0
C	3.2	3.0	3.6	8.2	11.8	13.5	14.4
D	1.5	1.0	1.1	3.4	7.0	7.8	8.1
E	1.5	1.2	1.0	1.6	3.5	8.5	14.7

Source: Bank of Slovenia; own calculations.

Notes: The table reports the percentage of firms and banks' exposure (in terms of classified claims) to the firms in each grade over time.

Deterioration in the quality of assets as reflected in the credit rating structure led to an increase in the amount of loan-loss provisions banks needed to take onto their balance sheets. In principle, for each individual firm loan-loss provisions need not increase automatically in response to a worse credit rating of the borrower, but the relation is nevertheless positive. Each bank uses its own methodology for determining loan-loss provisions, both for collective as well as individual provisioning. The former is in general based on the credit ratings, although the banks may also use differently formed groups of financial assets. For each of the credit ratings A, B and C the banks calculate the (past) incurred loss for the borrowers that migrated to ratings D or E and thus determine their internally required coverage of loans with loan-loss provisions for each of these three rating classes. The banks are required to regularly update the migration matrices, meaning that their required coverage ratio can change in time.

Collateral plays an important role in provisioning. Banks can apply a lower coverage ratio for the borrowers that pledged the best-quality collateral, but only for the part of the claims that is secured with this collateral. According to the Regulation, individual provisioning is used for the borrowers for which there exists an objective evidence of a possible loss. This could be either significant financial difficulties of the debtor, default on the obligations to the bank, information about potential bankruptcy, financial reorganization, decrease of the estimated future cash flows or other changes that could represent a loss for the bank. When the bank finds that such objective evidence exists, it assesses the value of collateral and expected cash flow and thus determines the individual provisions for such borrower. In general, all the borrowers that are either more than 90 days overdue or are rated D or E are assessed individually.

Table 12 reports the average coverage of outstanding corporate loans with the loan-loss provisions (end-of-period stocks) banks held across the credit rating classes in the period 2006-2012. It is evident that banks on average needed to provide more for expected loan losses for firms with lower credit ratings. Even though the exact rate of the loan-loss provisions depends also on the potential collateral used to secure a loan, we see that on average loan-loss provisions significantly increase with deteriorating credit-rating structure of banks' portfolios. In the period under investigation this structure exhibited a significant deterioration and an increase in the loan-loss provisions banks took on their books. Consequently increased also the pressure on bank capital, which created amplified incentives for the banks to underestimate credit risk. This corroborates the findings of the comprehensive review of the banking system of 2013 that the loan-loss provision, even though significant, were insufficient to cover the estimated potential future losses.

Table 13 provides further evidence on the link between credit ratings and loan-loss provisions by focusing on a subset of observations. Many firms in our dataset are clients of more than one bank and quite a significant number of these are assigned different credit ratings by different banks. The total number of firms in our sample is 41.964, which yields 174.048 firm-year observations. Out of these, 45.161 pertain to clients of more than one bank. And within this sub-sample 24.365 firm-year observations are such that firms are assigned different credit rating by different banks.

Banks are required to assign credit ratings independently of the existing or potential pledged collateral, which implies that the differences in credit ratings we observe are the

Table 12: Loan-loss provisions across credit ratings

Credit rating	Loan-loss provisions in percent of outstanding loans
A	0.6
B	3.2
C	13.6
D	41.4
E	80.6

Source: Bank of Slovenia; own calculations.

Notes: The table reports the average coverage of classified claims by the stock of loan-loss provisions, calculated for the 2006-2012 period.

result of banks' specificities in credit risk assessment and are not conditioned by firm characteristics. Table 13 shows that banks applied on average considerably higher rates of loan-loss provisions for worse credit ratings assigned to the same firms. For instance, the difference in the realized coverage ratio for the firms who have at one bank rating A and at the other bank rating B is on average 3 percentage points. In general, the differences in the ratios of loan-loss provisions increase monotonically with the difference in the assigned credit ratings.

The results in Table 13 might be plagued by the value of collateral held by different banks for the same client. In this respect two comments are in order. First, it is true that banks with better access to the collateral of a given firm have smaller incentives to underestimate credit risk. What we observe in the data, however, is that on average this is not the case. As Table 13 suggests, a downgrade of a client on average led to an increase in loan-loss provisions even after a potentialy neutralizing effect of collateral.

The second comment is about potential biases in valuation of collateral in addition to under-estimation of credit risk. Namely, in addition to having the incentive to overestimate credit ratings, banks have an incentive to overvalue the collateral (see Huizinga & Laeven, 2012). In particular, banks with weakly collateralised loans would have an incentive to over value the collateral and decrease the amount of required loan-loss provisions without the need to assign overly optimistic credit ratings. This effect, however, would bias the differences in the ratios of loan-loss provisions downward, which would imply that the true differences in the ratios of loan-loss provisions among credit-rating classes would only be larger than those reported in Tables 12 and 13.

Overall, the information in Table 13 is in line with that in Table 12. It confirms that in times of a financial crisis and deteriorating credit rating structure of bank portfolios, banks can reduce their provisioning costs by underestimating credit risk. Empirical evidence of underestimation of credit risk for the case of Slovenia is found by Volk (2012). He notices that even though the banks downgraded a considerable share of borrowers, the average implied probability of default of rating classes A, B and C increased. This indicates that the risk assessment strategy by the banks changed significantly in the crisis.

Table 13: Average loan-loss ratios for the sub-group of firms with different ratings across banks

Pairs of ratings	Average coverage given higher rating	lower rating	Difference
A-B	0.7	3.7	3.0
A-C	1.1	11.9	10.8
A-D	2.2	35.1	32.9
A-E	12.3	81.8	69.5
B-C	4.0	12.9	8.9
B-D	4.2	32.4	28.2
B-E	4.0	73.4	69.4
C-D	18.8	35.3	16.5
C-E	22.0	64.1	42.1
D-E	48.2	75.2	27.0

Source: Bank of Slovenia; own calculations.

Notes: The table reports the average coverage of classified claims by the stock of loan-loss provisions for the same firms with different ratings assigned by different banks. For all pairs of credit ratings the average coverage ratio is reported separately for the higher and lower assigned rating. The statistics are given in percentages (difference in percentage points) and are calculated for the 2006-2012 period.

The deterioration of banks' portfolios in 2006 - 2012 was very heterogeneous. To demonstrate this we divide 25 banks under analysis into three groups. The first division is according to residence of owners into foreign and domestic. Foreign-owned banks are 11 in number. The second division is that of domestic banks into large and small.[7] The group of large domestic banks consists of three banks. Two, the NLB bank and the NKBM bank, were the largest on the market, holding more than 35% of total bank assets. From the point of view of the ability to raise capital in times of financial distress these two banks can be deemed too big to fail and enjoying an implicit bailout guarantee by the government. The implicit state guarantee assumption rests also on the ownership structure of the largest two banks. They both had the government or government-controlled enterprises as the largest or even majority owners. The classification of these banks as large should thus be understood also in the functional sense due to direct presence of the state ownership. In addition, state ownership could be reflected in business strategy of these banks as it provides a channel for political intervention into bank management and consequently allocation of loans based on other than purely financial criteria. Evidences of political influence on loan allocation and interest rates charged by state-owned banks are provided by Dinc (2005), Khwaja and Mian (2005) and Sapienza (2004). The third

[7] All foreign-owned banks operating in Slovenia in period 2006-2012 can be classified as small.

bank in the group of large domestic banks, the SID bank, is similarly included not merely because of its size, but because it is a 100% state-owned bank.[8] The remaining 11 domestic banks were classified as small in terms of size. These banks were also predominantly privately owned.

Table 14 reports the share of non-performing loans (defined as loans with more than 90 days overdue - upper panel) and the coverage of NPLs with loan-loss provisions (lower panel). Non-performing loans increased rapidly after the onset of the crisis and virtually exploded in domestic banks after 2009, exceeding 25% in 2012. These levels of NPL ratios are very high by international standards. The increasing dynamics in foreign-owned banks was significantly less pronounced. Similarly, it can be argued that a better performance of foreign-owned banks was not due to miss-classification of NPLs as performing. The AQR providers reported that the rate of miss-classification of NPLs as performing loans was on average roughly 4% for SMEs (ranging from 0% to 13%), roughly 13% for large corporates (ranging from 0% to 21%) and roughly 10% for real estate developers (ranging from 0% to 19%). Our econometric analysis of classification accuracy of firms in default below suggests that foreign-owned banks were at the lower end of these ranges, which implies that miss-classification was mostly concentrated among domestically-owned banks. These banks had on average considerably higher NPL ratios and thus also higher incentives to miss-classify loans.

The bottom panel of Table 14 reports the corresponding coverage ratios of NPLs with loan-loss provisions. Relative to the levels before the crisis we see that only large domestic banks on average kept the ratio at the same level and, provided unchanged level of collateralization, took on their books the full account of increasing burden of NPLs. Foreign and especially small domestic banks, on the other hand, decreased the coverage ratio quite significantly and thus did not let the required provisions on expected losses from the NPLs to pass in full onto their profit and loss accounts.

Descriptive statistics presented thus far allow to make the following summary observations. The banking regulation in place embeds significant procyclicality in loan-loss provisions. During the Great recession the quality of bank assets in Slovenia deteriorated significantly. This was reflected in the credit rating structure presented in Table 11. The table shows that banks downgraded their borrowers, but the results of the comprehensive review from 2013 suggest that downgrading did not reflect fully the increase in credit risk. Table 14 finally indicates that the incentives to objectively asses the creditworthiness of assets differed significantly across banks.

Foreign-owned banks had the smallest burden of non-performing loans. Foreign-owned

[8] SID bank was established with a special purpose of securing international trade deals and enjoys an explicit government guarantee for its liabilities. During the crisis it served as a vehicle to stimulate corporate lending through state guarantee schemes and for disbursement of loans of international financial institutions.

Table 14: Share of NPLs and coverage ratio for three groups of banks

	2007	2008	2009	2010	2011	2012
			Share of NPLs			
Large domestic banks	2.8	3.3	6.6	16.5	26.7	33.2
Small domestic banks	2.3	3.5	7.1	11.7	19.2	26.4
Foreign banks	2.1	3.9	6.6	8.7	9.0	11.9
			Coverage ratio			
Large domestic banks	44.5	37.1	40.9	35.6	42.3	46.1
Small domestic banks	68.7	49.7	40.1	33.1	36.0	34.9
Foreign banks	52.2	29.7	26.1	33.7	34.0	36.0

Source: Bank of Slovenia; own calculations.

Notes: The table reports the share of NPLs and coverage of NPLs by the stock of loan loss provisions (in percent). Non-performing loans are defined as classified claims more than 90 days overdue.

banks are also part of international banking groups who have easier access to wholesale finance and can provide funds to daughter affiliates through the internal capital markets. Moreover, it is possible that these banks employed better risk management techniques and corporate governance mechanisms. A plausible explanation of a smaller NPL burden of foreign banks is the "cherry picking" behavior under which foreign banks choose to serve less risky firms (see Claessens & van Horen, 2014 and references therein).[9] Such a behavior is especially pronounced in emerging markets (Sengupta, 2007).[10] A smaller NPL burden and easier access to sources of capital offered foreign-owned banks capacity for easier absorption of credit losses. Both factors imply less incentives to apply discretion in credit risk assessment.

Just the opposite was the case of domestic banks, which had a considerably higher NPL burden. Small domestic banks, moreover, had very limited possibilities to raise additional capital.[11] Our prior is that both features translate into significant incentives to underestimate credit risk under stress.

Large domestic banks, even though heavily burdened with non-performing loans, enjoyed

[9] This appeared to have happened despite the fast expansion of lending activity by foreign-owned banks who actually led the pace of credit expansion in Slovenia prior to the crisis. In the period 2003-2008 the total amount of loans outstanding of foreign-owned banks expanded by 372%, while those of domestically-owned small and large banks grew by 274% and 207% respectively.

[10] While better risk management and corporate governance and "cherry picking" represent plausible explanations for a lower share of the NPL burned reported in Table 14 it is important to note that the comprehensive review revealed a significant 78% capital shortfall for foreign-owned banks also. This demonstrated that, on average, risk management and corporate governance was not at a high level in absolute terms.

[11] The Financial Stability Review of the Bank of Slovenia (Bank of Slovenia, 2013a) documented that throughout the crisis smaller domestic banks had on average higher capital requirements for credit risk on overdue and high-risk exposures, with foreign-owned banks on the opposite side of the spectrum. Similar developments were observed regarding capital adequacy. During the crisis it increased the most for the group of foreign-owned banks, while domestic banks, especially small, experienced significant difficulties in raising additional capital.

an implicit bail-out guarantee by the government. The effect of size and state ownership on the incentives to underestimate credit risk is not clear a priory. Higher capacity to raise additional capital through the government offers the management to take full account of deteriorating credit portfolio and hence reduces the incentives to underestimate risk. Bail-out guarantee, on the other hand, leads to moral hazard and thus higher incentives to underestimate risk. For a particular Slovenian situation our prior is that the first factor prevails and hence we conclude that large domestic banks had smaller incentives to underestimate credit risk than small banks. Nevertheless, we could think of it as an empirical issue. But given that the two effects work on the incentives in opposite directions we can say that if our methodology proposed below confirms our prior that bigger size led to smaller incentives to underestimate risk, it will be an overestimation of true incentives to underestimate risk.[12]

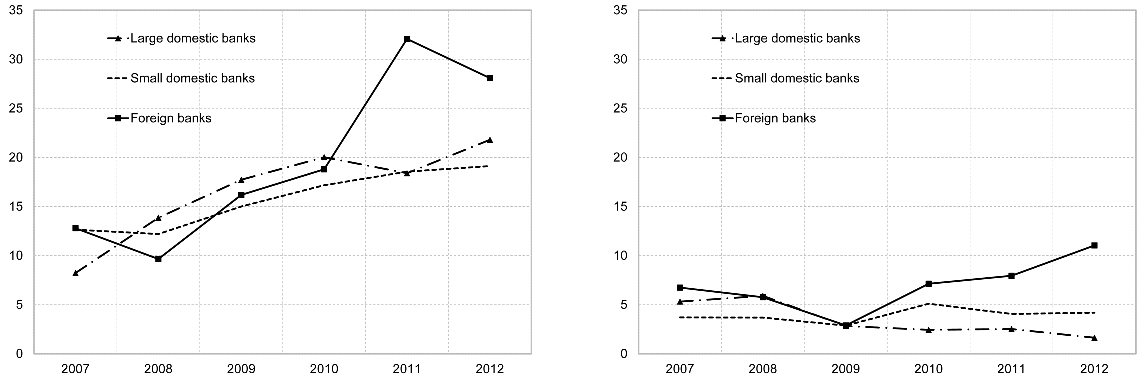
Our ranking of the groups of banks to underestimate risk can be supported with simple descriptive statistics on the evolution of credit rating changes. Figure 5 reports the share of rating changes per year in the period 2007 - 2012 divided into rating downgrades (left panel) and rating upgrades (right panel). It clearly emerges from the figure that as the crisis unfolded the frequency of rating downgrades increased. From roughly 10% ratings revised downwards on average in 2007 the share increased to above 20% on average in 2012. The share of rating upgrades hovered around 5% on average. Across groups of banks we again observe a large degree of heterogeneity. At the onset of the crisis in 2009 and 2010 the pace of rating downgrades was led by large domestic banks, but closely followed by foreign banks. In the last phase of the crisis (2011 - 2012) foreign-owned banks dramatically increased the pace of downgrades to 30%, while in domestic banks it levelled off at 20%. It is true that foreign-owned banks led also the pace of rating upgrades, while domestic banks lowered this rates to levels below 5%, however, the rates of upgrades did not exceed one third of the rates of downgrades. This implies that on net foreign banks led a more restrictive ratings policy and most actively downgraded the quality of their portfolios in face of deteriorating economic conditions. Large domestic banks can be ranked second in this respect, and they even led the pace at the beginning of the crisis.

Small domestic banks introduced the smallest changes to ratings structure of their portfolios. Such a ranking of bank groups is in line with our prior ranking of incentives to underestimate credit risk and consequently inflate bank books.

[12] The incentives to underestimate risk cannot be linked to difference in capital buffers formed before the crisis. In 2008 the average capital adequacy ratio of foreign-owned banks was at 10.6% below average.

Figure 5: Share of rating changes (in %): rating downgrades (left panel) and upgrades (right panel)

(a) Rating downgrades *(b) Rating upgrades*



Source: Bank of Slovenia; own calculations.

2.3 Predictive power of credit ratings and a test of discretion

Our approach to testing the potential bias in credit risk assessment in times of an economic downturn is the following. We focus on loans to non-financial corporations as this segment of bank portfolios held 80% of overall value of loans more than 90 days overdue (our measure of default) in the period 2007 - 2012. For these corporations we have access to data on their credit ratings assigned by corresponding banks. In the process of credit risk assessment banks dispose with publicly available information on firms' balance sheet and income statement and other information collected by the banks such as the information on overdue payments on bank loans. Such information is systematically recorded and can be in principle used also by an econometrician in modelling default. In addition, banks can keep regular contacts with the borrowers to obtain other information that is not systematically recorded and apply expert judgement in assessing creditworthiness and assignment of credit ratings. It is thus sensible to assume that credit ratings are formed using more information about firms' creditworthiness relative to the information used by an econometrician. Indeed, the additional information used in assigning credit ratings could also be a strategic decision to apply discretion in order to inflate the value of the bank's portfolio in a crisis.

The idea of the test for discretion is rather simple. We have two information sets, one in the form of pure financial information and the other in the form of credit ratings. We take both sets and include them separately into two econometric models. The model using pure financial information is denoted below as the *balance-sheet model.* The model that uses the information embedded in credit ratings is denoted as the *credit-ratings model.* Both models are logistic regressions. With both we test for respective classification accuracy of defaulted firms. For the purpose of our analysis it is important to emphasize that the balance-sheet model is free from incentives to underestimate risk.

Credit ratings should embed superior information and expert knowledge in assessing creditworthiness. In addition, the potential superiority of expert information should become more pronounced in times of financial distress, when rapidly changing economic conditions call for prompt collection of new information from the market. The econometrician using the balance sheet model cannot adapt to the situation as it disposes only with information that is reported by private enterprises for each fiscal year. Under fair assessment of credit risk we should thus observe that in the a financial crisis the information advantage of credit rating results in a relative improvement of default prediction of the credit-ratings model. Conversely, if in such a setting we observe a reduction in explanatory power of credit ratings for the probability of default relative to a model using financial information (ratios) only, this is an indication of the incentives to underestimate risk. In other words, if in times of financial distress banks have incentives to underestimate risk, we should observe a reduction in explanatory power of credit ratings model relative to the balance sheet model. Moreover, we should observe a more pronounced reduction for those banks whose incentives to underestimate risk are larger.

In sum, we test for the presence of discretionary risk assessment with the aim to underestimate risk by comparing the explanatory power of models using (1) credit ratings and (2) pure financial information both along the time dimension and along the cross section of banks. If the hypothesis of increased incentives to underestimate risk is empirically relevant, we expect to observe a deterioration of explanatory power of credit ratings in time as the financial crisis unfolds, and in the cross-section for banks with weaker capital structures and corporate governance, higher exposures to credit risk, and with limited access to the market for funds.

Such a testing approach is in line with Krahnen and Weber (2001) who note that an important requirement for the risk rating system to function properly is that it takes into account possible incentive problems. In relation to this, Kirstein (2002) demonstrated theoretically that even if assumed that banks have better knowledge of the customers than rating agencies, external ratings are better able to implement the goals of the Basel Committee than internal ratings. He argues this is due to lack of the banks' incentive to truthfully assess firms' creditworthiness. Consequently, banks' credit ratings need not be more reliable indicators of financial distress.

Before moving to the presentation of bankruptcy prediction models three remarks are in order. Firstly, it should be noted that a reduction in explanatory power of credit ratings can also be a consequence of standard financial information (like financial ratios) becoming less reliable in a crisis, potentially due structural breaks, and hence be in general less reliable indicators of financial distress. Because this can happen in periods when incentives for discretionary risk assessment increase, the change in explanatory power of credit ratings would not be a reliable gauge of discretionary risk assessment. Our approach, however, does not suffer from this problem because we test our hypothesis through differences between the model with credit ratings and the model with pure financial information. Because the latter in principle enter the information set of both models, they should be similarly affected by a potential structural breaks in explanatory power of financial ratios.

Secondly, with exception of one, all banks in our sample used a standardized approach to determining capital requirements for credit risk. Our testing approach, however, is applicable also to banks using the internal rating based (IRB) system. Under IRB banks use a fully parametric model to determine the probability of default as one of the key ingredients to calculating capital requirements.[13] Based on such a model a rating scale is determined and borrowers sorted accordingly. The information entering the model is of two types: purely financial and information based on bank's expert judgement of various non-measurable determinants of borrower's creditworthiness. The latter set of information is in principle subject to discretionary assessment in times of financial distress (Kirstein, 2002). In such a case, a test of the presence of under-estimation of risk would be based on the test for systematic differences (both across time and banks) between the probabilities of default given by the bank's model and a bankruptcy prediction model free of subjective information. Similarly, the test could be based on the predictive power of credit ratings of the banks operating an IRB system. Both approaches are fully consistent with the approach we use.

Last but not least, our measure of comparison of the models is classification accuracy of firms in default. Classification accuracy of healthy firms is not of primary interest. Our focus is on incentives to under-estimate credit risk through avoiding the regulatory requirement to form more loan-loss provisions for firms with worse credit rating. This means that we primarily study a "one-sided incentive", which corresponds to evaluating the classification accuracy of defaults.

2.3.1 The balance sheet model

We model corporate default in a fairly standard way in the literature, paved by Altman (1968). The balance sheet model uses financial ratios as predictors of default. Default is defined from the information on number of days overdue on loan payments, which is a common indicator of default in the literature (see Bonfim, 2009, and Carling et al., 2007, among others). Such an indicator is also in line with the Basel recommendations (BCBS,

[13] The second important quantity is the estimate of the loss given default, which depends on the valuation of underlying collateral. Valuation of collateral is another area where banks can apply discretion to inflate their books. Given that the focus of our analysis is the informativeness of credit rating we focus our discussion on discretion in modeling probability of default.

2006). We define the event of default as:

$$Y_{it} = \begin{cases} 1 & \text{if firm } i \text{ is more than 90 days overdue to at least one bank in time } t \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The probability that the binary dependent variable Y_{it} equals one given the covariates is modelled using the following specification:

$$P(Y_{it} = 1|X_{it-1}) = \Lambda(\alpha + \beta X_{it-1}) = \frac{e^{\alpha+\beta X_{it-1}}}{1 + e^{\alpha+\beta X_{it-1}}} \tag{5}$$

where α and β are parameters to be estimated and X_{it-1} is a vector of firm specific variables measuring size (log of sales), firm life-cycle (age), liquidity (quick ratio, cash-flow ratio), number of days with blocked account, asset turnover, financial structure (debt-to-assets ratio) and position on the financial market (number of relations with banks). The information in the balance-sheet model is also available to banks in the process of assessing firms' creditworthiness. All explanatory variables enter the model lagged one year.

The model is estimated for the period 2007-2012. Given that our aim is to compare the classification accuracy between the balance-sheet model and the credit-ratings model during the crisis, the models are estimated for each year separately.

2.3.2 The credit ratings model

To be able to compare the prediction accuracy of the balance sheet model with the banks' accuracy in firms' credit risk assessment we propose the following credit-ratings model. Since each bank assesses firms' riskiness with its own methodology and same firms can thus have different credit ratings across banks, we define the default event at bank-borrower level as

$$Y_{ijt} = \begin{cases} 1 & \text{if firm } i \text{ is more than 90 days overdue to bank } j \text{ in time } t \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

and estimate the logit model:

$$P(Y_{ijt} = 1|R_{ijt-1}) = \Lambda(\gamma + \delta R_{ijt-1}) = \frac{e^{\gamma+\delta R_{ijt-1}}}{1 + e^{\gamma+\delta R_{ijt-1}}} \tag{7}$$

where γ and δ are parameters to be estimated and R_{ijt-1} is a set of four dummy variables for each of the credit ratings from A to D, indicating firm i's credit rating, given by bank j in time $t-1$. The credit rating E is accounted for by the constant. Similarly to the balance sheet model we estimate the model for the 2007-2012 period year by year.

2.4 Results

Tables 15 and 16 present the estimated coefficients of the balance sheet and credit ratings model respectively. Classical standard errors are used in both models. In the period under analysis, the variables included in the balance-sheet model are consistently statistically significant. The estimates of the credit rating model in Table 16 show that all credit rating dummies enter statistically different from zero.[14] The base rating is E - the worst rating - and in line with our expectations we can observe the coefficients of other dummy variables monotonically decrease with increasing rating. Through time the constant increases quite significantly, corresponding to an increase in probability of default in line with an increase in the share of non-performing loans in the banking system. Note that other coefficients that measure differential effects relative to the credit rating E do not exhibit similar changes, which reflects the fact that the probability of default increased consistently across all credit ratings.

Table 15: The balance sheet model - Estimated coefficients for each year separately

	2007	2008	2009	2010	2011	2012
log(Total sales)$_{it-1}$	-0.292***	-0.170***	-0.143***	-0.105***	-0.122***	-0.107***
Age$_{it-1}$	-0.024***	-0.040***	-0.051***	-0.046***	-0.041***	-0.043***
Quick ratio$_{it-1}$	-0.131***	-0.090***	-0.158***	-0.215***	-0.221***	-0.156***
Debt-to-assets$_{it-1}$	0.016**	0.006	0.069*	0.378***	-0.018	0.037
Cash flow ratio$_{it-1}$	-0.300***	-0.224***	-0.133**	-0.272***	-0.433***	-0.317***
Asset turnover r.$_{it-1}$	-0.459***	-0.643***	-0.450***	-0.723***	-0.598***	-0.470***
Blocked account$_{it-1}$	0.011***	0.011***	0.012***	0.012***	0.014***	0.014***
No. of relations$_{it-1}$	0.484***	0.429***	0.420***	0.418***	0.468***	0.540***
Constant	-0.911**	-2.247***	-1.862***	-1.399**	-0.486	-0.744
No. of observations	15638	15970	17546	17985	18164	18218

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the logit estimates for each year from 2007 to 2012, where the dependent variable is equal 1 if firm i is more than 90 days overdue to at least one bank in year t and zero otherwise. Blocked account is a number of days a firm has blocked bank account, No. of relations measures to how many banks a particular firm is related to. Sectoral dummies are included to control for the specificity of each sector. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 17 contains a comparison of the classification accuracy of the two models. While overall classification accuracy decreases slightly from 2007 to 2012 (and less so for the balance sheet model), the classification accuracy of defaulted firms exhibits more pronounced dynamics that is depicted in Figure 6.[15] What we observe is that before the crisis (default in 2007 based on 2006 data) both models had a very similar classification

[14] We also estimate the credit rating model using clustering standard errors across firms. This only marginally increases the standard errors and all the coefficients are still significant at less than 1% probability of type I error. The results are shown in Appendix A.

[15] To distinguish between defaulters and non-defaulters we use standard 0.5 cut-off of predicted probability.

Table 16: The credit ratings model - Estimated coefficients for each year separately

	2007	2008	2009	2010	2011	2012
Credit rating A_{ijt-1}	-6.184***	-5.556***	-5.173***	-5.795***	-6.102***	-5.833***
Credit rating B_{ijt-1}	-4.776***	-4.218***	-4.204***	-4.668***	-4.944***	-4.673***
Credit rating C_{ijt-1}	-3.423***	-2.954***	-3.043***	-3.340***	-3.233***	-3.058***
Credit rating D_{ijt-1}	-1.900***	-1.898***	-1.918***	-2.240***	-2.411***	-2.112***
Constant	1.290***	1.177***	1.480***	1.946***	2.168***	1.946***
No. of observations	21200	21480	23906	24926	25203	25595

Source: Bank of Slovenia; own calculations.

Notes: The table reports the logit estimates for each year from 2007 to 2012, where the dependent variable is equal 1 if firm i is more than 90 days overdue to bank j in year t and zero otherwise. Credit rating A to D are dummy variables for each of the credit ratings, which are assigned to the firms by the corresponding banks. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

precision. The balance sheet model outscored the credit ratings model only by 2 percentage points. In the initial years of the crisis, 2008 and 2009, the classification accuracy of both models drops.[16] Such a result is expected. The beginning of the crisis represents also a turnaround in defaults. Bankruptcy prediction models use $t-1$-dated information. This means that predicting default in the first year of the crisis involves using only information from before the crisis, when balance sheets of firms appeared healthy. It is important to note, however, that the deterioration in classification precision is higher for the credit ratings model. Credit ratings should in principle reflect information superior to pure $t-1$-dated information of the balance-sheet model and thus suffer less from the problem of time delays in availability of information. Banks can learn about the crisis before its effects are recorded in end-of-the year balance sheet and income statement data that the balance-sheet model uses. This information advantage could serve to adjust the ratings in a timely manner so as to reflect the increase in the incidence of firm default. It would be thus sensible to expect that the credit ratings would suffer less in terms of loss of defaults classification precision. This is not what we observe in our estimation results, which leads us to conclude that the lack of adjustment of credit ratings in face of financial crisis was used to inflate banks' balance sheets.

A diverging performance of the models continues to the end of the period under investigation. We can note first that the turnaround in classification precision of the balance sheet model occurs one year before the turnaround of the credit ratings models. This represents another piece of evidence that the banks were slower to incorporate new overwhelming evidence of deteriorating financial health of enterprises than a pure mechanical econometric procedure would do. The last column of Table 17 shows that because of this in 2010 the difference in classification accuracy of defaulted firms grows to 19 percentage points, almost tenfold of the pre-crisis difference.

[16] Slovenia slid into a recession in the fourth quarter of 2008.

Table 17: Classification accuracy through the business cycle

	The balance sheet model		The credit ratings model		Difference	
	Overall	Defaulters	Overall	Defaulters	Overall	Defaulters
2007	96.2	21.3	96.9	19.3	-0.7	2.0
2008	95.2	17.2	96.0	13.9	-0.8	3.3
2009	93.8	17.0	94.5	10.2	-0.7	6.8
2010	93.8	28.0	93.8	9.2	0.0	18.8
2011	93.8	33.2	93.5	16.0	0.3	17.2
2012	93.3	35.1	92.9	19.9	0.4	15.2

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the overall classification accuracy and correctly classified defaulters in percentages as predicted with the balance sheet model and the credit ratings model estimated for each year in the sample. The difference between both models is given in percentage points.

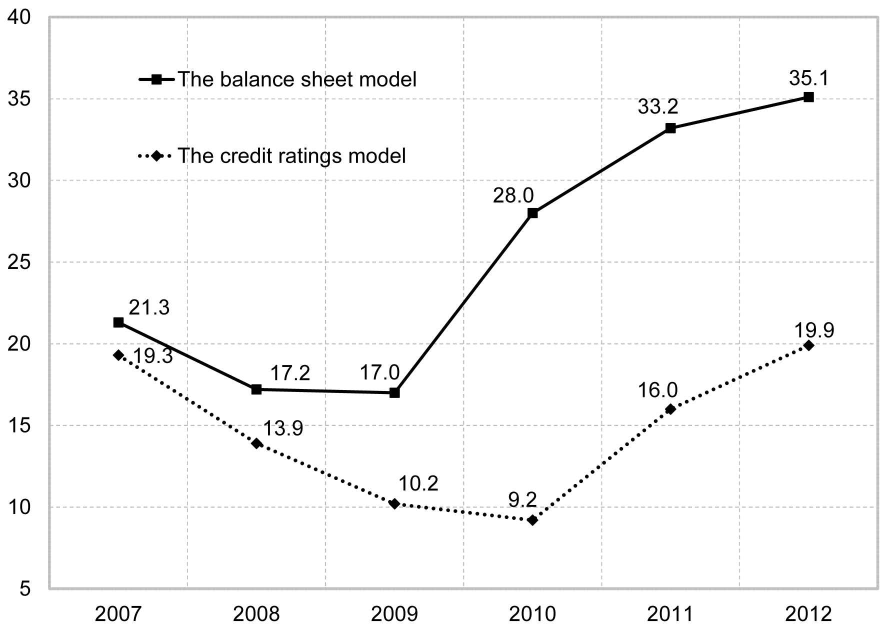
In 2011 and 2012 the classification accuracy of the credit rating model picks up and closes some of the gap to the balance sheet model, but remains at more than 15 percentage points, which is seven times higher than before the crisis. Moreover, for the credit rating model the classification accuracy in 2012 returns to the pre-crisis level. For the balance sheet model, however, we can see that it is considerably higher than before the crisis, 35.1% relative to 21.3%.[17] Overall, this comparison provides time-series evidence of a potential problem with discretion in credit risk assessment. Banks could in principle incorporate information about mounting financial difficulties of their clients much faster and in a forward-looking manner than a purely econometric backward-looking procedure that uses only published information from the previous period. In the data we observe just the opposite.

We now turn our attention to classification accuracy of defaulters across groups of banks. The corresponding results are presented in Figure 7. The results for the credit rating model estimated on the data for banking system as a whole (solid line) and the balance-sheet model (dashed line with circle markers) are the same as in Figure 6. The remaining results are for the credit ratings model estimated on observations corresponding to each of the banking groups: foreign-owned banks, large domestic banks and small domestic banks.

What we can observe are large differences across groups of banks. Classification accuracy of foreign-owned banks stands out as the most precise. With 40% accuracy before the crisis (2007) it outperforms all other models, declines significantly in the first three years of the crisis, but remains quite comparable to the balance-sheet model, and results again

[17] To some extent, this can explained by the fact that with the crisis the number of firms in default increased considerably. In estimation of a bankruptcy prediction model this implies that a higher probability mass is accounted for by the default cases, which facilitates a more precise discrimination between healthy firms and firms in default. See Brezigar-Masten and Masten (2012) for a discussion.

Figure 6: Correctly classified defaulters (in %)



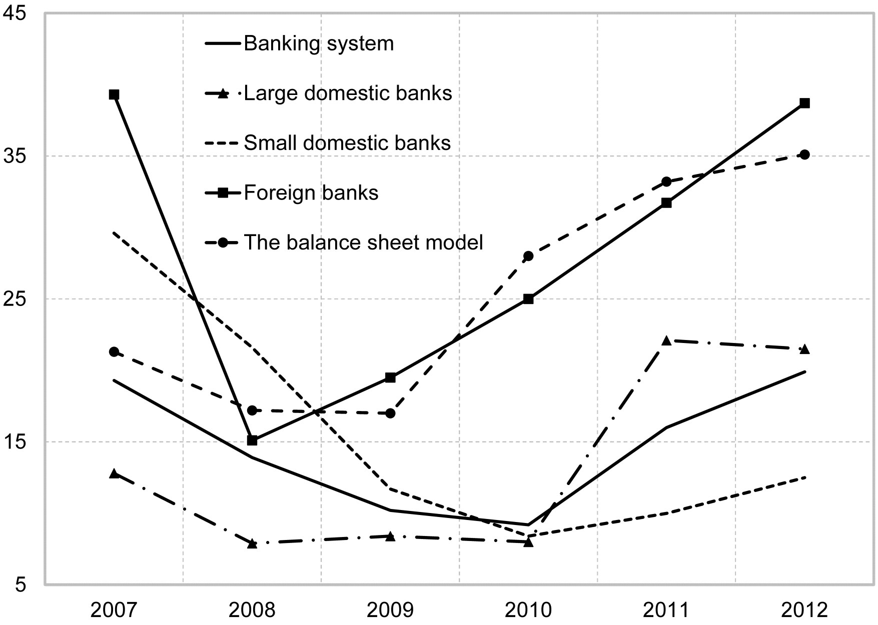
Source: Bank of Slovenia; AJPES; own calculations.

to be the best performing model in the final year under analysis. Moreover, evolution of the classification accuracy of foreign-owned banks is very similar to the evolution of the balance-sheet model. The latter is a purely econometric procedure free of incentives to underestimate credit risk. The experience of the classification accuracy of credit ratings of small banks is at the opposite end of the spectrum. While they performed quite well before the crisis, their classification precision of defaulters steadily decreases through to 2010 to less than 10% accuracy and remains at these low levels thereafter. The classification accuracy of the credit-rating model of large domestic banks stands in between. It is the least precise before the crisis and the initial two years, but it picks up quite significantly at the end of the period, reaching levels of precision above 20%, which is double the pre-crisis rate.[18]

These results go hand in hand with the evidence on the incentives for discretionary risk assessment presented in Section 2.2. Mounting burden of non-performing loans in the banking system as a whole led to an average increase in incentives to under-estimate risk, assign higher credit ratings on average and consequently make smaller loan-loss provisions. In such a case, we would expect to find that with the financial crisis unfolding credit ratings on average lose the explanatory power for default. The empirical evidence in Figure 6 is consistent with this view.

[18] While the worst classification accuracy of large domestic banks before the crisis is consistent with the moral hazard explanations of incentives to underestimate risk, this no longer prevails in the crisis. The fact that the classification accuracy of large banks outperforms those of the small banks indicates a stronger effect of smaller incentives to underestimate risk due easier access to capital and funding.

Figure 7: Correctly classified defaulters across groups of banks (in %)



Source: Bank of Slovenia; AJPES; own calculations.

Section 2.2 presents also evidence that the incentives for under-estimation of risk differed significantly across banks. Foreign-owned banks experienced smaller problems with the NPLs and had smaller difficulties with maintaining capital adequacy and funding because of a more stable access to finance through internal capital market of the banking groups they belong to (Navaretti et al., 2010). Domestically-owned banks were more heavily exposed to NPLs and had weaker capital adequacy ratios. Among them, large banks enjoyed a strong implicit state bail-out guarantee, which also materialized in capital injections into two largest banks in 2011 and 2012. Small domestic banks, on the other hand, experienced significant problems with raising additional capital and with access to wholesale funding. As we noticed above, this group of banks did not make a similar adjustment of credit risk assessment standard towards more stringent policy we observe for foreign-owned banks and large domestic banks. In sum, these observations suggest that it was the group of foreign-owned banks with smallest incentives to under-estimate risk in order to artificially protect their balance sheets. Small domestic banks were on the other end of the spectrum. The results on classification accuracy in Figure 7 are in line with these observations. Credit-ratings of foreign-owned banks appear considerably more reliable determinants of default than those of domestic banks. In the latter group it is the group of small banks whose credit ratings' classification accuracy deteriorated most significantly during the financial crisis and even remained well below the classification accuracy of large domestic banks.

The best classification accuracy of credit rating of foreign-owned banks could be in principle attributed also to better corporate governance and better risk management practices.

Note, however, that the pre-crisis levels of classification accuracy of small domestic and privately owned banks, that did not have any direct access to technology of large multinational banking groups, was quite comparable. About 5 percentage points lower than foreign banks, but at the same time roughly 15 percentage points higher than the classification accuracy of large domestic banks. Because of state ownership the weak initial performance of the latter group could be attributed also to political interference and moral hazard. Relative positions in classifications accuracy before the crisis can thus be rationalized with factors related to corporate governance. In addition, corporate governance and risk management technology can play an important role in in shaping the incentives to underestimate risk in times of financial distress. However, in the period under analysis no major ownership changes occurred in the Slovenian banking system. For this reason we consider the potential effect of corporate governance heterogeneity as fixed in time. Consequently, the dynamics in classification accuracy in the crisis can be explained with the differences in the incentives to underestimate credit risk described above.

2.5 Robustness checks

The analysis so far has been conducted on yearly frequency in which we use only end of year data to predict default one year ahead. Our data on credit ratings are on quarterly, which enables us to conduct two robustness checks. Both check robustness of our results with respect to real-time availability of the information banks have at their disposal when setting credit ratings.[19]

2.5.1 Controlling for rating changes

It is sensible to expect that ratings that change as banks acquire new information about specific clients will have a better explanatory power of firms' default. Namely, banks' cannot instantaneously and simultaneously review all the clients in the portfolio, because of insufficient capacity to do so. Instead, priority is given to subsets of firms. In a crisis these are foremost firms in distress. Not controlling for rating changes thus potentially biases our previous analysis at the expense of the credit rating model, thereby showing it less reliable in predicting default.

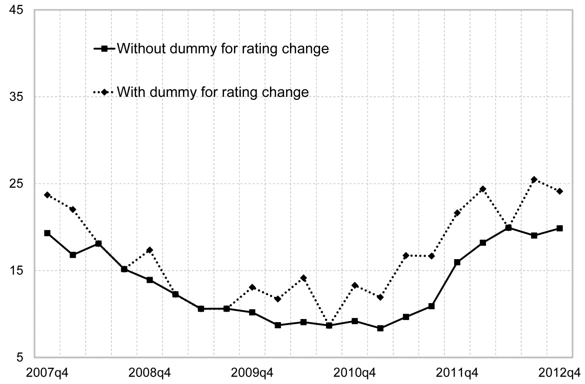
Data on credit ratings are on quarterly frequency, which enables us to trace the timing of rating changes and hence the time when a bank updated the information set. Accounting for updating of the information sets is imperfect since we cannot identify cases where information set was updated, but rating did not change. Nevertheless, given that empirical analysis uses data mostly from the Great recession in which rating downgrades heavily

[19] We checked also robustness of our results with respect to the forecast horizon. In particular, horizons between 1 and 4 quarters were considered. Our conclusions obtained in the previous section are robust to different forecast horizons. Results are shown in Appendix B.

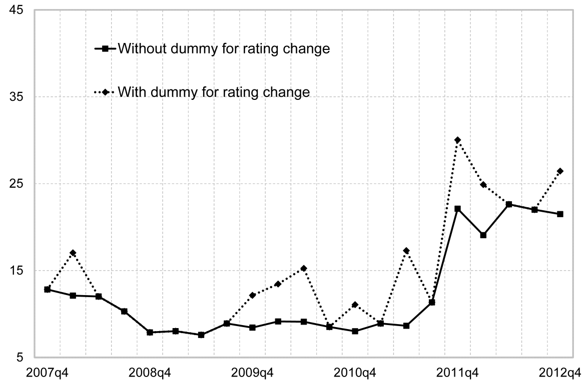
dominated the process (see Figure 5), controlling for rating changes could improve the performance of the credit ratings model.

Figure 8: Effect of rating changes - Correctly classified defaulters 1-year ahead across groups of banks (in %)

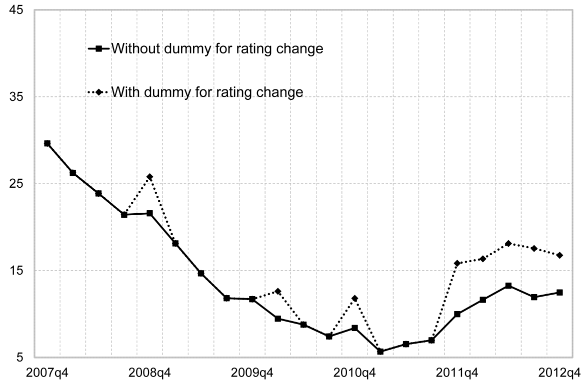
(a) Banking system



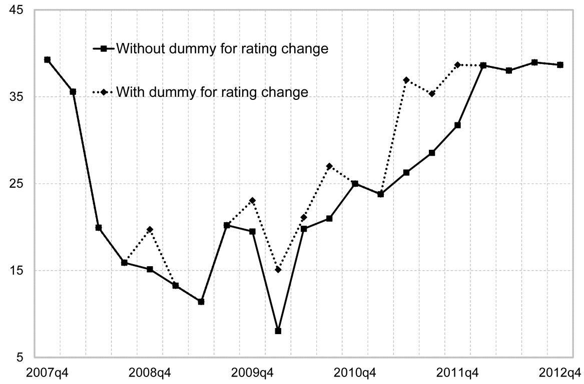
(b) Large domestic banks



(c) Small domestic banks



(d) Foreign banks



Source: Bank of Slovenia; own calculations.

Classification accuracy of credit rating models, estimated for each quarter, with and without a dummy variable for rating change is shown in Figure 8. For the whole banking system, controlling for rating changes improves the share of correctly classified defaulters in 14 out of 21 quarters. In the cases where its impact is positive, it contributes on average 4.8 percentage points to the classification accuracy. Rating change thus improves the performance of the credit rating model. However, the credit rating model still hits considerably lower share of defaulters than the balance sheet model, especially in the crisis period.

Controlling for rating changes also does not change our conclusions about the different behaviour of the three groups of banks. From Figure 8 we can still conclude that it is the ratings of foreign banks that are throughout the crisis the most precise in classifying firms in default. Conclusions about large domestic and small domestic banks are also fully consistent with the evidence presented in Figure 7. This leads us to conclude that our basic conclusions are robust to the timing of rating changes.

2.5.2 Controlling for public release of corporate balance sheet data

In this section we investigate whether the timing of public release of balance sheet and income statement data has an effect on the dynamics of rating changes. Namely, firms are required to report their balance sheet data for the past fiscal year to the Agency of public and other legal records until the end of March of the current year. The data become publicly available during the second quarter of current year. We could thus assume that banks are informed about the financial state of each firm in year t in the second quarter of year $t+1$. The balance sheet model in the previous sections, however, assumes that this information is available already at the end of each year t. In this sense the balance sheet model has a potential information advantage over the credit-rating model, which uses the information on credit ratings in real time.

Table 18 reports the percentage of credit rating changes over individual quarters (averaged across years). If the newly available balance sheet data would be the main driver of rating changes, we expect that the large majority of changes would happen in the second and/or third quarter. Results in the table indicate that this might be the case on average, but not systematically so. On average, the frequency of rating changes is the lowest in the first quarter, but this observation does not hold uniformly across bank groups as the minimum for large domestic banks is in the third quarter. The frequency of rating changes is on average the largest in the second quarter, but again not uniformly across bank groups. Finally, one can note that the differences in rating changes across quarters are not stark. The largest is the difference between the third and the first quarter for small domestic banks of about 15 percentage points. The remaining differences are considerably smaller.

Table 18: Credit rating changes over quarters (in %)

	Q1	Q2	Q3	Q4
Large domestic banks	25.0	28.7	22.5	23.8
Small domestic banks	18.5	25.5	34.2	21.8
Foreign banks	18.0	28.4	25.8	27.8
Overall	21.6	27.7	26.8	23.9

Source: Bank of Slovenia; own calculations.

Notes: The table reports the percentage of credit rating changes in each quarter. The statistics are calculated for the period 2007q1-2012q4.

Thus, the timing of the release of the balance sheet and the income statement data might not be a decisive element in determining the relative performance of the credit ratings model. There are two reasons for such a finding. The information on delays in loan repayment and information on blocked accounts are an important determinant of the credit rating and available in real time. Moreover, banks' reliance on the release of official balance sheet and income statement data is crucial for smaller firms. Larger firms, to which banks have larger exposures, are monitored on a more regular basis.

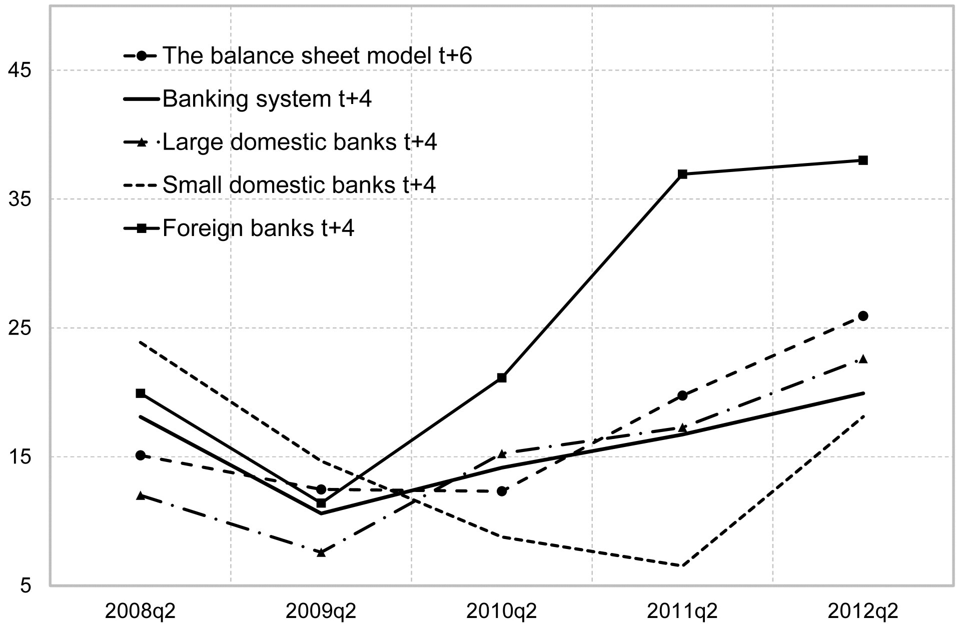
To investigate the effects of the potential information advantage of the balance sheet model formally, we compare the classification accuracy of competing models by suitably adjusting the prediction horizons. Namely, we position the models in a pseudo-real time context into the second quarter of each year, when corporate balance sheet data become publicly available. Exploiting the availability of credit ratings at quarterly frequency we can estimate the credit ratings model for the second quarter of each year and consider predicting default one year ahead. From the point of view of the balance-sheet model the corresponding prediction horizon is 6 quarters because this model uses the data for the end of the previous year, which become available with a 2-quarter delay, i.e. in the second quarter of the current year. Different forecast horizons between models are in terms of notation. From the point of view of balance sheet data availability in real time, the forecast horizons are aligned.

Constructing the comparison this way puts the models at par as regards the availability of balance sheet and income statement data. It favours, however, the credit ratings model from two other important points of view. First, the balance sheet model uses the information on incidence and the duration of blocked transactions accounts of firms. This information is available in virtually real time to banks and, as explained above, it is an important determinant of credit rating. So it is in principle contemporaneously embedded in the credit rating model. However, because we compare the 4-quarter ahead prediction of the credit ratings model with the 6-quarter ahead prediction of the balance-sheet model, the latter essentially uses the information on blocked accounts that is 2 quarters old.

The second aspect is the information on overdue loan repayments. This information is not used in the balance sheet model, but is readily available to banks in real time when they monitor their clients. Consider, for instance, a firm that in the second quarter accumulates significant days overdue, but not more than 90, which is our measure of default. This is a clear indication of a high probability that the firm will be in default in the near future. Clearly, the information on cumulating overdue is a signal to the bank to downgrade the firm, which should thus be a good predictor of default. This is another feature that puts the credit ratings model in a more favourable position from the point of view of available information.

In sum, by adjusting the forecast horizons as explained above we neutralize the information advantage of the balance-sheet model due to availability of balance sheet data in real time. At the same time, however, we additionally penalize the balance-sheet model from the point of availability of data on blocked transaction accounts and payments overdue. The robustness check we consider is thus rather extreme.

Figure 9: Effect of public release of data - Correctly classified defaulters one year ahead (Balance sheet model t+6, Credit rating models t+4) (in %)



Source: Bank of Slovenia; AJPES; own calculations.

Figure 9 presents the results of the robustness check. As in the previous subsection, the credit rating model is augmented with the dummy variable for the rating change. Our conclusions from the previous section appear to be robust. The balance-sheet model exhibited a less precise classification accuracy than the credit ratings model before the crisis. Even if put into an informational disadvantage, however, results end up as more precise than the average and comparing to domestic banks in particular. As before, the credit ratings model estimated for the group of foreign banks outperforms all other models, especially in the crisis period. On the other hand, the classification accuracy of small domestic banks dropped considerably during the crisis, becoming the least precise. From the best position before the crisis, it declined both relative to other two banking groups and, more importantly, to the balance sheet model. The patter for large domestic banks is also fully consistent with our basic analysis. From the worst position before the crisis, classification accuracy improves in the crisis significantly more that it does for small banks. This is another indication that our basic conclusions that the largest incentives to underestimate risk could be found with the small domestic banks are robust.

2.6 Conclusion

In this paper we study the discretion in credit risk assessment for the case of Slovenian banking system during the Great recession. The Slovenian case is instructive as 10 major banks of the system in the second half of 2013 went through a comprehensive review similar to comprehensive reviews major banks in the Euro area went through before the establishment of the Single Supervisory Mechanism in November 2014. The comprehensive review applied common methodology to all banks involved and estimated significant shortages of capital in banks that reported sufficient capital adequacy ratios just a quarter before. Moreover, the review revealed stark differences across groups of banks that differ primarily with respect to ownership (domestic - foreign, state - private) and size. Our analysis addressed the question whether these differences can be explained with the incentives of banks to apply discretion in credit risk assessment, whereby over-estimation of credit ratings helped the banks to conceal some of the problems with deteriorating quality of their portfolios. This allowed them to temporarily avoid taking additional loan-loss provisions and hence inflate their balance sheets.

Our empirical analysis, by using data on firms' credit ratings, shows that discretion in credit risk assessment is a plausible explanation of differences in the required recapitalisation revealed by the comprehensive review. Banks that needed higher relative recapitalizations resulted to be the ones with the highest incentives to over-estimate ratings and whose ratings, as a results, provided to be the least reliable indicators of the incidence of financial distress of borrowers. These conclusions remain valid also after considering two robustness checks about the information structure in the credit rating process.

Moreover, the analysis provides a plausible explanation for the results of the comprehensive review of banks that entered the Single Supervisory Mechanism in November 2014. The results released by the European Central Bank in October 2014 indicate a clear divide between the estimated shortages of bank capital across the size distribution of banks. Capital shortfall for banks with total assets above the average on average amounts to 0.66% of their capital. For banks below average size the number is almost ten times bigger, 6.46%. By taking the median as the divisor between small and large banks, the respective numbers are 1.70% and 8.18%. Across the quartiles of the bank size distribution the estimated capital shortfalls are 5.96%, 9.05% and 6.30% for lower three quartiles respectively, and only 0.64% for the top quartile. These results are clearly in line with the conclusions of our analysis. Smaller banks likely have more limited options to raise additional capital either from the market or through government intervention and thus have bigger incentive to underestimate credit risk and inflate their books.

Our empirical findings have a number of important implications for banking regulation. Discretion in credit risk assessment is nothing but an attempt to temporarily sweep the problems under the rug. The fact is that the true creditworthiness of borrowers is always eventually revealed, credit risk realized and losses incurred. These losses are higher the longer the under-estimation of risk postpones solving the problems. As it turned out for the case of Slovenia, the estimated direct fiscal costs of bailing out these banks exceeded 10% of GDP.

In this respect, for future prevention and better management of such episodes it is important for the regulation to respond to the problem of incentives to under-estimate credit risk in times of financial crises and economic downturns in general. Clearly, discretion can be in principle mitigated by stricter control over credit risk assessment. Stricter control is possible already in the current system, but regulatory forbearance is often applied in similar crisis situations. Stricter regulatory control should thus take the form of standardized and externally controlled credit rating procedures. Currently, banks using both the basic and advanced rating approaches under Basel Accord regulation develop internal methodologies that need to be approved by the regulators. The application of these methodologies is, however, still subject to discretion. Discretion can only be avoided if risk assessment is subject to simultaneous external evaluation or even externally determined.

A more important result of our analysis is the importance of monitoring the incentives for discretion in credit risk assessment. As we show, the firms' ratings that are regularly reported to the central bank can be tested for their precision in predicting distress. In times of financial crisis significant differences across time and banks emerge that, if persistent, may lead to a significant destabilization of the banking system. Indeed, smaller banks and banks with weaker position on the market for funds may represent a disproportional risk to the system as a whole. The current IFRS provisioning model, based on incurred losses, led to delays in loss recognition and to significant pro-cyclicality in loan-loss provisions during the financial crisis. The International Accounting Standards Board (IASB) intends to introduce a new impairment model where losses will be recognised in more forward-looking manner. According to their proposal from March 2013 there would no longer be a threshold before credit losses would start to be recognized. Instead, expected credit losses would be recognized from the point at which financial instruments are originated or purchased. The amount of expected credit losses would be regularly updated to reflect changes in the credit quality. In this way, the credit losses would in principle not be delayed until the default event, but would at least partly be recognized in earlier stages. These provisions, however, assume away the problems with discretion in valuation of assets and credit risk assessment. Despite being forward-looking in nature, such a provision could be distorted by the banks incentives to over-value assets and under-value risk. Indeed, if such new regulation would result to be the most binding at times of extreme financial distress, its major expected effect could be undone by amplified incentives to conceal the true state of banks' portfolios.

In addition, policy measures increasing capital requirements in times of financial distress, increase the incentives to under-estimate risk and thus may undo the expected effect of strengthening confidence in the banking system. An example of such a measure in the Great recession is the measure by the European Banking Authority that required banks to hold at least 9% Core Tier 1 capital adequacy ratio by mid 2011, which was in the middle of still intense financial turmoil in the Euro area. From the point of view of our analysis, the timing of this policy measure could have amplified the incentives to under-estimate credit risk. The new Basel III and CRD IV capital regulation introduce a countercyclical capital buffer that could somewhat alleviate this problem. In the periods of excessive credit growth and possible build-up of system-wide risk, banks will be required to build a capital buffer (of up to 2.5% of RWA) in the form of Common Equity Tier 1 capital. When the crisis hits the buffer could be released and banks would thus have additional capital at hand, increasing their loss absorption capacity and possibly decreasing the incentives to underestimate credit risk.

3 Evaluating the Performance of Dynamic and Tobit Models in Predicting Credit Default[20]

3.1 Introduction

Credit default models are extensively used by banks and regulators. IRB regulation requires from banks to provide their own estimates of probability of default, which is one of the key parameters that determines capital requirements (BCBS, 2001, 2006). Identifying non-performing borrowers also enables banks and regulators to project expected losses and to assess potential capital needs to cover these losses. In addition, default probability models can also be used for stress testing purposes to simulate the effect of different scenarios.

In this paper we propose and test the performance of two novel methodologies for modelling credit risk. Credit default is typically modelled using discrete choice methodology as was first proposed by Altman (1968). The binary dependent variable is usually defined following BCBS (2006) default definition, which is based on number of days past due. The default event occurs when borrower is more than 90 days overdue. By transforming an overdue into a dichotomous variable, a lot of potentially useful information is lost. In addition, overdue is already a risk measure and therefore it seems reasonable to model it directly, without any transformations. Since it is censored at value zero, we apply tobit modelling approach. Our first set of tests is aimed to evaluate and compare the performance of classical binary probit model versus tobit model.

Credit default indicators show a lot of persistence. Once a borrower defaults (becomes more than 90 days overdue), it is not very likely that he will become performing again. Moreover, an overdue, once being positive, is expected to increase in time. Estimating default probability model, which includes autoregressive dynamics can thus significantly improve predicting performance. Our second proposed novelty is thus to estimate dynamic probit and dynamic tobit model using Wooldridge (2005) methodology and compare their performance with static version of the models.

We evaluate the performance of the models by looking at their ability to discriminate between performing and non-performing borrowers. Conventional default probability models, however, usually follow the discrete time hazard rate modelling approach, which gives the probability that borrower defaults in current period under the condition the default event did not occur before (see for instance Bonfim, 2009 and Carling et al., 2007). As described by Hamerle et al. (2003) this in an underlying methodology of IRB regulation. We therefore also estimate classical default probability model, where only transitions to default are taken into account, and compare its performance in predicting new defaulters

[20] The paper is written in co-authorship with Arjana Brezigar-Masten and Igor Masten. First person plural is used for narration.

with other proposed models. Our goal is not to find the best performing model specification, but rather to use the same explanatory variables in all the estimates and see how different functional form (probit vs. tobit) and different information set (static vs. dynamic) affects the performance in explaining state of default and transition to default. The performance of the models is evaluated using the data of Slovenian non-financial firms.

We find that tobit modelling methodology outperforms all other models. In predicting non-performing borrowers, where persistence is of a key importance, dynamic tobit correctly identifies more than 70% of defaulters and issues less than 1% of false alarms. High performance - 66% true positive rate - is also achieved by dynamic probit model, which outperforms the static version by more than 30 percentage points. An important advantage of tobit model, however, is that its prediction is number of days past due, which enables to form different classes of overdue. One can for instance predict defaulters using any overdue threshold, not only 90 days as is standard in binary models. We show that dynamic tobit has high classification accuracy across different classes of overdue, from 30 to 360 days. For predicting new defaulters, however, we find that the static tobit model is the advantageous modelling approach. It correctly identifies more than 50% of new defaulters and outperforms all the other models by a large margin. It also issues more false alarms comparing to other methodologies, but given the gain in identifying defaulters, this loss is relatively small and acceptable. This is especially true if one is more concerned in missing defaulters (type I error) than issuing false alarms (type II error), like is typically assumed in early warning literature (see Alessi & Detken, 2011 and Sarlin, 2013). Even though classical binary default probability model is estimated explicitly on transitions to default, it is able to correctly identify only 5% of new defaulters. Three sets of robustness checks confirm the validity of our results.

Our paper is related to a recent study performed by Jones, Johnstone and Wilson (2015). They test the performance of various binary classifiers in predicting credit rating changes. In addition to conventional techniques such as probit/logit, they also evaluate the performance of more advanced approaches like non-linear classifiers, neural networks, support vector machines and others. They find that newer classifiers significantly outperform all other modelling approaches. Although the goal of our paper is very similar, it provides two new pieces of evidence. First, we show the performance of the models can be significantly improved if, instead of conventional binary model, tobit modelling methodology is applied. Second, we provide evidence that the dynamic specification of the model significantly improves the performance in predicting non-performing borrowers. To our knowledge both, tobit and dynamic methodologies, have not yet been applied to credit risk modelling. Moreover, we propose an approach for modelling credit default on quarterly frequency using mixed frequency data. This enables to monitor the changes in credit portfolio on higher frequency and also more accurately since the information set is updated each quarter.

The findings of this paper have important implications for banks and banking regulation. We show that the conventional default probability model that is typically used by IRB banks achieves very low classification accuracy. This poses a question whether this modelling approach, which at the end determines banks capitalisation, is an appropriate methodology. A simple upgrade of the model with dummies indicating overdue in previous period significantly improves the classification accuracy. The performance can be further improved by using the tobit modelling approach. Although the prediction of the tobit model, which is days past due, cannot be directly used in IRB formula for capital requirements, this approach is far more accurate in identifying new defaulters, and therefore it seems reasonable to use it in practice.

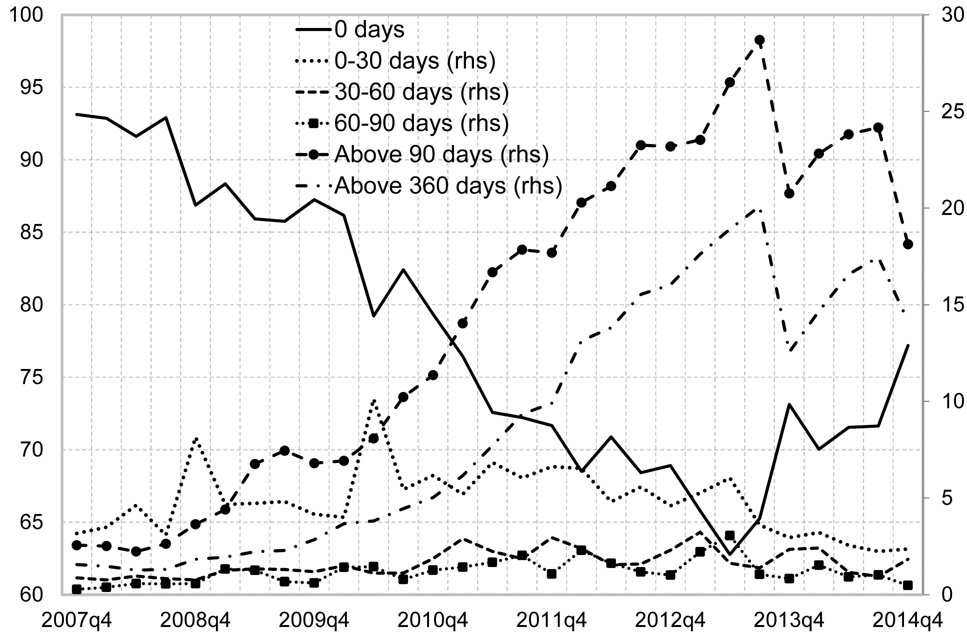
The rest of the paper is structured as follows. Section 3.2 provides descriptive analysis of the dynamics of different credit risk measures. In Section 3.3 we present the methodology for estimating and evaluating different credit default models. Estimation and evaluation results are presented in Section 3.4. Section 3.5 presents three sets of robustness checks, while Section 3.6 concludes the paper and discusses implications.

3.2 The dynamics of credit default measures

The key data source for our analysis is Credit register of Bank of Slovenia, which is exceptionally rich database with many information that are not publicly available. The variable we are most interested in is overdue in loan repayment, which signals financial problems of firms and is also a key credit risk measure under Basel regulation (see BCBS, 2006). It is first available in 2007q4, which limits our analysis to 29 quarterly cross sections from 2007q4 to 2014q4. Restricting the analysis to non-financial firms, which were during the crisis shown to be the most problematic segment, results in large sample of more than 1 million observations represented by a triple firm-bank-time.

Figure 10 shows the evolution of loans broken down to different classes of days of overdue in loan repayment. It can be seen that after the start of the crisis in 2008q4, the share of non-performing loans started rising rapidly and reached very high levels. The share of loans with more than 90 days overdue, which is a standard measure of non-performing loans (BCBS, 2006), rose by more than 25 percentage points until the third quarter of 2013. In 2013q4 it dropped by 8 percentage points, which is the result of transfer of bad loans from two largest banks to Bank Assets Management Company (BAMC). It should thus not be understand as natural improvement of banks' credit portfolio, but rather as an institutional measure that reduced the pressing burden of non-performing loans. Second tranche of transfer was carried out at the end of 2014. Contrary to non-performing loans, the share of loans with 0 days overdue dropped considerably in times of financial stress.

Figure 10: Share of loans across different classes of overdue (in %)



Source: Bank of Slovenia; own calculations.

Other classes between 0 and 90 days overdue represent only a small share of total loans, since these are in many cases only transition classes to higher days past due. The only exception is class between 0 and 30 days, which represents around 3 to 10 percentage share of total loans. There are many borrowers who occasionally have small delays in loan repayment, but whose overdue does not necessarily increase from one period to another.

Figure 10 reveals that overdue is highly autoregressive process. It can be best seen by increasing share of loans with overdue above 360 days. Once an overdue bridges a certain threshold, it is expected to increase in time and reach higher number of days past due. Since these borrowers are financially very weak and are not able to pay back their debt to banks, they are sooner or latter expected to bankrupt. In 83% of cases when an overdue changed between two consecutive quarters, this change was positive. This finding is partly the result of the fact that overdue is censored at zero, which means that by the nature of the variable the increases could be much more frequent. However, even when we look only at the cases when overdue>0, we get a similar result: 80% increases and only 20% decreases. This dynamic is, however, very heterogeneous across different classes of overdue. As can be seen in Table 19, an overdue is more likely to decrease between two consecutive quarters when it is lower than 30 days. This is the result of already mentioned occasional delayers who are in majority of cases able to repay the debt and their overdue thus typically returns to zero in the next quarter. In other classes positive dynamic prevails and the higher is the overdue, more likely it is, that it will further increase. This is to be expected, since once an overdue exceeds a certain threshold, it is not very likely that a firm will ever be able to repay the debt.

Table 19: Share of increases and decreases of overdue over different classes, in %

Overdue class	One quarter horizon % of increases	One quarter horizon % of decreases	One year horizon % of increases	One year horizon % of decreases
0 days	4.4	-	8.7	-
0-5 days	27.4	57.7	34.4	56.5
5-10 days	36.2	58.7	43.3	52.9
10-20 days	41.0	52.9	48.3	47.5
20-30 days	46.6	47.6	50.1	45.0
30-60 days	53.4	43.5	57.5	40.2
60-90 days	62.9	35.5	66.0	32.6
90-180 days	75.6	23.5	74.1	25.2
180-360 days	88.8	10.9	84.3	15.4
>360 days	95.3	4.6	91.5	8.4

Source: Bank of Slovenia; own calculations.

Note: The table reports the percentage of increases and decreases of overdue over different classes of overdue and two horizons.

Looking at changes in one year period in Table 19 reveals similar dynamic, but decreases prevail only until overdue is below 10 days. In addition, with exception of last three classes, the increases of overdue are more frequent on yearly basis than quarterly. This means that also borrowers with fewer days past due can be more problematic on a long run. Although they were in majority of cases able to repay their debt on a short run, this signals that they might not be able to do so on a long run. Overall, Table 19 clearly reveals that overdue has strong positive autoregressive component, especially when it is higher than 30 days.

Default rate and its projection, probability of default, is typically of a main interest in banks, since it is one of the key factors that determines projected expected losses and capital requirements for IRB banks. In addition, PD is also an important factor in loan approval and pricing. Table 20 shows the default rate over different classes of overdue. It is calculated as a share of borrowers that had been performing in time $t-1$ and became more than 90 days overdue in time t. As expected, the share of transitions to non-performing status is higher, the higher was the overdue in previous period and it further increases when calculated on one year horizon. Lower levels of overdue can thus be used as an early warning signal for potential defaulters in future periods. Classical PD model, where the transition to default is typically explained with borrower-specific factors, is unable to fully capture this information. It only captures some part of it when problems in loan repayment are reflected also in firm financial ratios. These, however, are usually available only once a year, which disable updating the estimated probabilities of default on the same frequency as overdue is refreshed.

Our analysis thus far reveals three potential upgrades of current prevailing credit risk

Table 20: Default rate over different overdue classes, in %

Overdue class	One quarter horizon	One year horizon
0 days	0.3	3.7
0-5 days	6.1	15.7
5-10 days	12.3	25.1
10-20 days	16.2	31.2
20-30 days	23.3	35.5
30-60 days	40.3	49.1
60-90 days	59.6	64.1

Source: Bank of Slovenia; own calculations.

Note: The table reports the default rate - share of borrowers that were less than 90 days overdue in time $t-1$ and became more than 90 days overdue in time t - over different classes of overdue and two horizons.

modelling techniques. First, overdue by itself is already a risk measure and thus it seems natural to model it directly. A lot of useful and valuable information is lost, when it is transformed to dichotomous variable and estimated with discrete choice model. An overdue, even if it is low, signals financial problems of a firm and it is thus important to monitor the whole spectrum of delays in loan repayment. Second, autoregressive component seems to be an important factor in modelling credit risk. As shown, an overdue is expected to increase in time, whereas default status shows a lot of persistence. Past information on days past due can also significantly contribute to explaining transition to default. It is therefore sensible to estimate dynamic credit risk model and see if it adds valuable information comparing to static one. Third, credit risk should be monitored on higher frequency. One year horizon for modelling probability of default that is typically used in the literature and also proposed by BCBS (2001) to IRB banks, is a very long period, since a lot can change over such a long horizon. In extreme case, an overdue may increase from 0 to over 360 days. Standard PD model, which is usually estimated using firm financial ratios is not able to capture such severe deterioration, since its information set is not updated during the year.

3.3 Methodology

This section presents the methodology for estimating and comparing credit default models. We are interested in three sorts of comparison. First, does it matter if we change the functional form of the model? More specifically, we compare the performance of probit model, where the default is modelled as a binary variable, and model where overdue in loan repayment is modelled explicitly, without any transformations. Since overdue is censored at zero, standard OLS estimator would result in biased estimates. We therefore apply tobit estimator, which captures this source of non-linearity. Second, does the dynamic specification of the model improve performance? We estimate both probit and tobit model including autoregressive term and compare the resulting performance with static specification of the models. Third, we compare the performance of the models in explaining the state of default and transition to default. For the latter, the accuracy of classical PD model is compared with the dynamic version of PD model and with prediction ability of aforementioned models.

Overall, we estimate and compare performance of six models, which can be divided into three groups. They differ in the definition of the dependent variable and in functional form of the model. The first group includes the models where the dependent variable is state of default: *static probit* and *dynamic probit*. In the second group we model overdue in loan repayment and apply censored regression: *static tobit* and *dynamic tobit*. Lastly, the transition to default is modelled with *static PD model* and *dynamic PD model*. Our goal is not to find the best performing model specification, but rather to use the same explanatory variables in all the estimates and see how different functional form (probit vs. tobit) and different information set (static vs. dynamic) affects the performance in explaining state of default and transition to default.

To our knowledge this is the first attempt to model credit default using the dynamic and tobit methodology. There are some analysis, like for instance Costeiu and Neagu (2013), where past information are included in the model, but not explicitly as lagged dependent variable. Hence, we first present some theory and solutions on how to estimate dynamic non-linear panel data models. Next, we present the specification of all the models and describe how we evaluate their performance.

3.3.1 Dynamic non-linear panel data models

The key issue in estimating dynamic panel data models is the initial conditions problem, which is the result of correlation between unobserved heterogeneity and past values of the dependent variable. In linear models this problem can be easily solved with appropriate transformation, like first differencing, which eliminates the unobserved effects. Although the transformed error term is correlated with transformed lagged dependent variable, instrumental variables can be used to achieve a consistent estimator. Anderson and Hsiao (1982) propose using y_{it-2} as an instrument in first-differenced equation. Arellano and Bond (1991) upgrade this approach by using a GMM-type of model with all possible instruments in each time period, whereas Blundell and Bond (1998) propose a system estimator, where also level equation with instruments in differences is estimated.

The problem with initial conditions is even more complicated in non-linear models. There are no transformations that would eliminate the unobserved effects. Suppose we are interested in modelling the process:

$$y^*_{it} = \alpha y_{it-1} + x'_{it}\beta + \eta_i + \varepsilon_{it} \tag{8}$$

where y^*_{it} is latent index, y_{it-1} is first lag of the dependent variable, x_{it} is a vector of strictly exogenous variables, η_i is unobserved individual effect and ε_{it} is error term, which is assumed to be distributed with mean 0 and variance σ^2_ε. As described by Akay (2012) the type of the model depends on how the dependent variable is observed. If y_{it} is observed as an indicator

$$y_{it} = \begin{cases} 1 & \text{if} \quad y^*_{it} > 0 \\ 0 & \text{if} \quad y^*_{it} \leq 0 \end{cases} \tag{9}$$

the model to be estimated is dynamic probit or logit model. If, on the other hand, y_{it} is observed as the variable that is censored at zero

$$y_{it} = \begin{cases} y^*_{it} & \text{if} \quad y^*_{it} > 0 \\ 0 & \text{if} \quad y^*_{it} \leq 0 \end{cases} \tag{10}$$

this leads to tobit model specification. Referring to our case, binary credit default models - state probit and transition probit - fit into equation 9, whereas overdue is censored at zero and can thus be represented with equation 10.

In estimating these models one needs to deal with unobserved individual-specific effect η_i, which is correlated with initial values y_{i0}, unless the start of the observed panel data set coincides with the start of the stohastic process. In this case initial values are non-stohastic constants and there is no need to deal with the initial conditions problem. In practice, however, we usually observe data after the start of the stohastic process and the conditional distribution of initial values must be specified. One option is to assume that initial values are not affected by past developments, i.e. to treat them as exogenous variables independent of all other regressors including unobserved individual effects. As described by Akay (2012) this is a very naive assumption, which typically leads to serious bias.

Another way of dealing with initial values is to use the fixed effect approach. Although explicit modelling of individual effects seems attractive, the results can be biased due to incidental parameters problem (Neyman & Scott, 1948). Honoré and Kyriazidou (2000) and Arellano and Carrasco (2003) propose a method for fixed effects logit model, which solves the initial condition problem by eliminating the unobserved heterogeneity. These models, however, can only be estimated for individuals that in the observed period switch between both observed states. If the states are persistent, like in our case, the number of observations would be considerably reduced.

The random effects solutions are much more common and attractive in practical applications.[21] Wooldridge (2005) proposes to use the density $(\eta_i | y_{i0}, x_{it})$ that specifies the functional form of unobserved heterogeneity:

$$\eta_i = \xi_0 + \xi_1 y_{i0} + x_i' \xi_2 + \psi_i \tag{11}$$

where x_i is $(x_{i1}, x_{i2} ... x_{iT})$. The basic logic of this procedure is that correlation between unobserved heterogeneity η_i and initial value y_{i0} is captured by equation 11, which gives another unobserved individual effect ψ_i that is not correlated with initial value y_{i0}. This follows the logic of Chamberlain (1984) who proposes to model conditional expectation of the unobserved effect as a linear function of the exogenous variables and initial conditions. All that needs to be done is to replace η_i in equation 8 with functional form 11, which results in:

$$y_{it}^* = \alpha y_{it-1} + x_{it}' \beta + \xi_0 + \xi_1 y_{i0} + x_i' \xi_2 + \psi_i + \varepsilon_{it}. \tag{12}$$

The main advantage of this methodology is that it is computationally very simple and can be implemented using standard random effects software. Additionally, the same methodology can be used for estimating dynamic probit and dynamic tobit model. Since we are interested in comparing the performance of different functional forms of credit default models, it is very important that it is not affected by different methodology for estimating probit and tobit model. A strong support for using this estimator in our analysis is also study by Akay (2012), who finds that it performs especially well in panels that are longer than 5-8 periods, which is also the case in our models.

3.3.2 Model specification

In order to estimate the credit default models we link Credit register data with firm balance sheet and income statement data, which are for all Slovenian firms collected by the Agency of the Republic of Slovenia for Public Legal Records and Related Services (AJPES) at yearly basis. To do so, we aggregate Credit register data to firm-time level by taking the highest overdue a particular firm has to any bank in quarter t. Note that our final dataset is of a mixed frequency. Whereas Credit register data are on quarterly basis, balance sheet and income statement data vary only yearly. As is presented below, we select a model specification that takes this into account.

General specification of our models can be characterised with the following non-linear

[21] Another random effects estimator is suggested by Heckman (1981a,b) who proposes approximating the conditional distribution of initial values using reduced form equation, estimated on the pre-sample information. As discussed by Akay (2012), the main problem with this method is that it requires simultaneous estimation of reduced form and structural model, which is computationally very difficult. In addition, it is not that often applied in empirical work.

function:

$$y_{it} = f(y_{it-1},\ x^q_{it-1},\ d_j x^y_{it-1},\ \eta_i), \quad i = 1, ..., N,\ t = 1, ..., T_i,\ j = 1, ..., 4 \qquad (13)$$

where y_{it} is the dependent variable, which is defined as presented in Table 21. In both, probit and PD models, we apply the 90-days threshold, which is very common in the literature (see for instance Bonfim, 2009) and also in line with the recommendations of Basel Committee (BCBS, 2006). For static and dynamic probit we define the default indicator that is equal one if firm i is more than 90 days overdue in quarter t. Similarly also for the PD model where the indicator is equal one if firm became a defaulter in time t, but had still been performing in $t-1$. For the tobit models, we keep overdue as it is, without any transformations and thus use all the information content in it. y_{it-1} is lagged value of the dependent variable, i.e. lagged default indicator in dynamic probit case and lagged overdue in dynamic tobit case. In PD model lagged dependent variable cannot be included explicitly since we are modelling the transition to default and thus it is equal to zero for all the firms. Similarly as Costeiu and Neagu (2013), we introduce the dynamics in the PD model by including dummies for different classes of overdue in previous period.

Table 21: Dependent variables in the models

Model	Definition of the dependent variable
Static & dynamic probit	state of default: $I(>90)_{it}$
Static & dynamic tobit	overdue$_{it}$
Static & dynamic PD	transition to default: $I(>90)_{it}/(\leq 90)_{it-1}$

Note: The table reports the dependent variables for probit, tobit and PD models.

Due to mixed frequency data, the distinction needs to be made between regressors that are available quarterly (x^q_{it-1}) and those that vary only yearly (x^y_{it-1}). Since the latter can have different effect across quarters, we multiply them with d_j, which are simply the dummy variables for each quarter. In this way we get a quarter-specific effect of yearly varying regressors on our dependent variables, which are observed quarterly. All the regressors are included with one period lag.[22] There are mainly two reasons for this. First, given current information, this will enable us to predict credit default at least one period ahead. Second, by including past values of regressors we avoid possible simultaneous causality problems.

In selecting the explanatory factors we follow the model specification by Volk (2012), who models the probability of default as a function of firm size, age, liquidity, indebtedness, cash flow, efficiency, number of days with blocked account and number of relations a particular borrower has with banks. The last two variables are observed quarterly, while

[22] For variables that are observed at yearly frequency this means including its values form previous year not previous quarter, since this would result in contemporaneous values for quarters 2, 3 and 4.

others that are calculated on a basis of firm balance sheet and income statement data, are available only once per year. Hence, we interact them with quarterly dummies.

η_i term in equation 13 captures the functional form for unobserved heterogeneity. As can be seen in equation 11, Wooldridge's (2005) original proposal is to include initial value of the dependent variable and the realizations of other regressors in each time period. This procedure would in our case lead to approximately 100 additional parameters to estimate. Given that we work with a large panel of data, this might not be so problematic. However, increasing the number of parameters to be estimated significantly extends the optimization procedure when the dataset is large and given that the model is already complex, this might also lead to problems with convergence. To avoid these problems we rely on evidence provided by Rabe-Hesketh and Skrondal (2013) who show that including only within means and initial values of each regressor does not lead to any bias comparing to Wooldridge's (2005) original specification. Therefore, our functional form for individual specific effects in dynamic probit and tobit model is the following:

$$\eta_i = \xi_0 + \xi_1 y_{i0} + x'_{i0}\xi_2 + \bar{x}'_i\xi_3 + z'_i\xi_4 \tag{14}$$

where y_{i0} is initial value of the dependent variable for each firm, which is the initial value of default indicator in case of dynamic probit model and the initial overdue in dynamic tobit case. The majority of initial values is taken from 2007q4 when our dataset starts. However, for those that enter subsequently, their first observation is taken as an inital value. x_{i0} is a vector of inital values for all the regressors, whereas $\bar{x}_i$ are within means of the regressors, defined as $\frac{1}{T_i}\sum_{t=0}^{T_i} x_{it}$ [23]. As explained by Wooldridge (2005), functional form for individual specific effects may include also other time invariant regressors. We add z_i, which is a set of industry dummies that controls for specificity of each industry.

We control for unobserved heterogeneity also in static and PD models. There are mainly two reasons for this. First, we capture the correlation between error term and firm specific effect and thus achieve consistent estimates (Chamberlain, 1984). Second, in this way the dynamic models do not have any advantage in terms of performance stemming from this additional terms. We use the same functional form as presented in equation 14 for dynamic models, with the only difference that we exclude initial values of the dependent variable. The same approach is used also for the dynamic PD model, which does not explicitly include lagged dependent variables and is thus not subject to initial conditions problem presented in section 3.3.1.

[23] For yearly varying regressors the mean is calculated by taking into account only one observation per year. In this way we avoid possible miscalculations for those firms that enter the dataset in the middle of the year.

3.3.3 Model evaluation

Basic goal of this paper is to compare the performance of different functional forms and specifications of presented credit default models. We do this by looking at several measures that can be calculated from the contingency matrix presented in Table 22. The columns represent the actual observed state, whereas the rows are predicted state by the model. For the latter we take the in-sample fit that is actually the prediction one quarter ahead. The prediction accuracy measures that we use are shown under the Table 22. The most important measure is the true positive rate, which shows the share of correctly predicted defaults. Banks and regulators are mostly concerned in identifying problematic loans, but of course, not on the cost of issuing too many false alarms.[24] For this reason, we show also other measures that will help us to assess model performance. Accuracy, as an overall classification accuracy measure, is also important, but is largely driven by the classification of non-defaulters, which represent a large majority in our data.

We use several criteria that places the observations in the contingency matrix. First, we compare probit and tobit models in terms of their ability to predict non-performing borrowers - more than 90 days past due. Second, the main advantage of tobit model is that its outcome is the whole distribution of overdue, which enables to test the performance also on other overdue classes, like 30, 60, 90, 180 and 360 days past due. Lastly, we compare the models' ability to predict the transition to default - $\leq$90 days overdue in $t-1$, >90 days overdue in time t. In all the cases the predicted indicator is equal one if the predicted probability of state or transition probit models bridges the 0.5 cut-off, whereas for the tobit models it is equal one if its predicted overdue is above a certain threshold, like 90 days.

Table 22: Contingency matrix

	Actual ($I_{it} = 1$)	Actual ($I_{it} = 0$)
Predicted ($P_{it} = 1$)	True positive (TP)	False positive (FP)
Predicted ($P_{it} = 0$)	False negative (FN)	True negative (TN)

$$\text{True positive rate} = \frac{TP}{TP+FN} \qquad \text{True negative rate} = \frac{TN}{FP+TN}$$

$$\text{False positive rate} = \frac{FP}{FP+TN} \qquad \text{False negative rate} = \frac{FN}{TP+FN}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$$

[24] An alternative way of defining this is to use the loss function proposed by Alessi and Detken (2011) and Sarlin (2013), where different weights are placed on type I and type II error.

3.4 Results

Table 23 presents the estimated coefficients of all the models. In addition to the variables that are shown in the table, all the models also include controls for unobserved heterogeneity as presented in section 3.3.2. Most of the coefficients for these controls are statistically significant, which indicates that it is indeed important to control for these effects in order to achieve consistent estimates.

Lagged default indicator in dynamic probit model has, as expected, highly statistically significant positive effect on current value of indicator. This indicates that the default status, 0 or 1, is highly persistent. Being zero in previous quarter, it is very likely it stays zero also in current period. On the other hand, once a firm is more than 90 days overdue it is not likely to become performing in the next quarter. Similarly, the positive effect of the dependent variable is also found in dynamic tobit model, which shows that the overdue is expected to increase in time. Past information on overdue is also included in dynamic PD model in the form of dummies for different classes of days past due (dummy for 0 days past due is excluded). It can be seen that higher overdue in previous quarter adds more to the default probability. All these results are in line with the findings presented in section 3.2.

Table 23 also reveals the importance of using the model specification that takes into account the mixed frequency structure of the data. Most of the interaction terms between quarterly dummies and firm specific variables are statistically significant, especially so for static version of the models. This indicates that the effect of yearly-observed variables on default probability or days past due is indeed heterogeneous across quarters. It is expected that the shorter the information lag, the more informative are the variables about credit default indicators. It is exactly what we find in our estimates. The majority of statistically significant coefficients can be found for the first quarter (the terms that are not pre-multiplied with quarterly dummy in Table 23), where the information lag to the observed firm-specific variables is only one quarter.

We now turn our attention to prediction accuracy of the models. Table 24 presents the classification accuracy of probit and tobit models in predicting non-performing borrowers. It can be seen that the dynamic specification of the models significantly improves the performance, especially for the probit model where the true positive rate increases by more than 30 percentage points comparing to static version of the model. Tobit model has even better performance. Static tobit achieves more than 33 percentage points higher true positive rate than static probit model, whereas dynamic tobit adds additional 3 percentage points to the classification accuracy of defaulters. Importantly, this high prediction accuracy of defaulters is not on a cost of issuing too many false alarms. Tobit model has slightly higher false positive rate, but these are still very low values, especially in the case

Table 23: Estimated coefficients

	Static probit	Dynamic probit	Static tobit	Dynamic tobit	Static PD	Dynamic PD
Dependent variable	$I(>90)_{it}$	$I(>90)_{it}$	Overdue$_{it}$	Overdue$_{it}$	$I(>90)_{it}/(\leq 90)_{it-1}$	$I(>90)_{it}/(\leq 90)_{it-1}$
Dependent var.$_{it-1}$		2.096***		1.067***		
log(Total sales)$_{it-1}$	-0.182***	-0.056***	-81.820***	1.369*	0.033**	0.004
Age$_{it-1}$	0.267***	0.133***	60.193***	11.249***	0.107***	0.038***
Quick ratio$_{it-1}$	-0.023***	-0.014***	-1.368***	-1.249***	-0.018***	-0.006
Debt-to-assets$_{it-1}$	0.005*	0.002	2.954***	0.802***	-0.003	-0.002
Cash flow ratio$_{it-1}$	-0.011	-0.018	-8.654***	-7.447***	-0.045**	-0.025
Asset t. ratio$_{it-1}$	-0.263***	-0.149***	-26.766***	-22.156***	-0.209***	-0.076***
No. of days bl. ac.$_{it-1}$	0.017***	0.010***	2.805***	1.053***	0.012***	0.006***
No. of relations$_{it-1}$	0.345***	0.198***	69.505***	29.457***	0.241***	0.050***
d2*log(Total sales)$_{it-1}$	0.026***	0.019***	2.980***	-0.449	0.026***	-0.001
d2*Age$_{it-1}$	-0.003	-0.002	-0.040	-0.524***	-0.002	0.003
d2*Quick ratio$_{it-1}$	0.019***	0.011*	0.123	0.269	0.016**	0.007
d2*Debt-to-assets$_{it-1}$	0.006*	0.006	0.750	-0.072	0.008	0.003
d2*Cash flow ratio$_{it-1}$	-0.050***	-0.027	-1.355	1.759	0.003	-0.001
d2*Asset t. ratio$_{it-1}$	-0.007	-0.010	5.392**	7.079***	-0.023	-0.005
d3*log(Total sales)$_{it-1}$	0.053***	0.037***	2.529***	-0.793*	0.045***	0.032***
d3*Age$_{it-1}$	-0.005***	-0.003	0.743**	-0.238	-0.004*	-0.001
d3*Quick ratio$_{it-1}$	0.019***	0.008	1.381***	1.294***	-0.000	-0.001
d3*Debt-to-assets$_{it-1}$	0.008**	0.008*	1.332**	-0.113	0.009	0.004
d3*Cash flow ratio$_{it-1}$	-0.054***	-0.021	-10.257***	0.531	-0.011	-0.008
d3*Asset t. ratio$_{it-1}$	0.002	0.001	1.306	5.449***	-0.005	-0.023
d4*log(Total sales)$_{it-1}$	0.057***	0.022***	5.843***	-0.651	0.029***	0.026***
d4*Age$_{it-1}$	-0.007***	-0.003	0.099	-0.783***	-0.003	-0.000
d4*Quick ratio$_{it-1}$	0.019***	0.007	1.360***	1.256***	-0.020*	-0.015
d4*Debt-to-assets$_{it-1}$	0.007**	0.005	1.634***	-0.226	0.009	0.009
d4*Cash flow ratio$_{it-1}$	-0.056***	-0.020	-12.949***	1.365	-0.029	-0.033
d4*Asset t. ratio$_{it-1}$	0.028**	0.031**	4.097*	9.355***	0.038**	0.012
Overdue 0-5$_{it-1}$						1.095***
Overdue 5-10$_{it-1}$						1.352***
Overdue 10-20$_{it-1}$						1.533***
Overdue 20-30$_{it-1}$						1.840***
Overdue 30-60$_{it-1}$						2.290***
Overdue 60-90$_{it-1}$						2.716***
Constant	-10.629***	-7.002***	-824.385***	-190.530***	-2.164***	-2.546***
Observations	517964	517964	517964	517964	487969	487969

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the coefficients for all the estimated models. The dependent variable for static and dynamic probit is an indicator $I(>90)_{it}$ that is equal one if firm i is more than 90 days past due in time t and zero otherwise. For both PD models, the dependent variable is defined as transition to default ($\leq$90 days overdue in time $t-1$, >90 days overdue in time t). No. of days bl. ac. measures number of days a firm has blocked account. No. of relations is number of relationships between each firm and banks. d2 to d4 are dummy variables from second to fourth quarter. Overdue 0-5 to Overdue 60-90 are dummy variables for number of days a firm is past due. In addition to the variables that are shown in the table, the models also include controls for unobserved heterogeneity as described in section 3.3.2. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

of dynamic model. Comparing to the gain in true positive rate, the loss in terms of false alarms is relatively minor.

Table 24: Performance of probit and tobit model in predicting performing and non-performing borrowers

	Probit		Tobit	
	Static	Dynamic	Static	Dynamic
True positive rate	0.356	0.663	0.688	0.714
True negative rate	0.990	0.993	0.947	0.991
False positive rate	0.010	0.007	0.053	0.009
False negative rate	0.644	0.337	0.312	0.286
Accuracy	0.949	0.972	0.930	0.973

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the classification performance of probit and tobit models in predicting performing and non-performing borrowers (more than 90 days past due). See section 3.3.3 for the description of classification accuracy measures.

Table 25 shows the classification accuracy of dynamic probit and dynamic tobit model in predicting non-performing firms where we let the autoregressive process to proceed four quarters ahead. These are still the in-sample predictions, with the only difference that instead of actually observed values of lagged dependent variable its predictions are taken, which are obtained by recursively running the predictions four times. The results show that tobit is the superior model also on a longer horizon. Its true positive rate is expectedly decreasing on longer forecast horizon, but it stays above the performance of the probit model. Dynamic probit achieves slightly higher overall accuracy, but this is only due to better prediction of non-defaulters. False positive rate still stays very low for both models.

Table 25: Performance of dynamic probit and tobit model in predicting performing and non-performing borrowers from one to four quarters ahead

	Dynamic probit				Dynamic tobit			
	1q	2q	3q	4q	1q	2q	3q	4q
True positive rate	0.663	0.571	0.505	0.454	0.714	0.603	0.544	0.508
True negative rate	0.993	0.992	0.993	0.993	0.991	0.987	0.986	0.986
False positive rate	0.007	0.008	0.007	0.007	0.009	0.013	0.014	0.014
False negative rate	0.337	0.429	0.495	0.546	0.286	0.397	0.456	0.492
Accuracy	0.972	0.964	0.959	0.954	0.973	0.962	0.956	0.952

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the classification performance of dynamic probit and dynamic tobit model in predicting performing and non-performing borrowers (more than 90 days past due) one to four quarters ahead. See section 3.3.3 for the description of classification accuracy measures.

Tobit model enables to form the predictions for different overdue classes. Table 26 shows the classification accuracy results of static and dynamic tobit model for five thresholds of days past due. All the classes are defined in the same way: when overdue or prediction is above a certain threshold the indicator is equal one, otherwise it is zero. It can be seen that in the case of static model, true positive rate is decreasing with higher overdue threshold. static model correctly classifies 85% of firms with overdue above 30 days, but only 39% of firms with overdue higher than 360 days. The performance of the dynamic model is much more stable and its true positive rate is fluctuating around 75%. Static model outperforms the dynamic one in terms of true positive rate for 30 and 60 days class. It, however, also has significantly higher false positive rate, which is for the 30-days class equal to 20%, comparing to only 3% of the dynamic model. The results presented in Table 26 thus reveal, that both, static and dynamic tobit model, are quite successful in classifying borrowers to different classes of overdue, but the dynamic version of the model is shown to be the superior one.

Table 26: Classification accuracy of static and dynamic tobit model across different groups of overdue

	Static tobit					Dynamic tobit				
Overdue threshold	30	60	90	180	360	30	60	90	180	360
True positive rate	0.854	0.752	0.688	0.561	0.390	0.746	0.720	0.714	0.733	0.774
True negative rate	0.800	0.918	0.947	0.973	0.987	0.972	0.986	0.991	0.997	0.998
False positive rate	0.200	0.082	0.053	0.027	0.013	0.028	0.014	0.009	0.003	0.002
False negative rate	0.146	0.248	0.312	0.439	0.610	0.254	0.280	0.286	0.267	0.226
Accuracy	0.805	0.906	0.930	0.953	0.967	0.952	0.967	0.973	0.984	0.991

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the performance of static and dynamic tobit model in classifying borrowers into different groups of days past due. In all the cases an indicator is equal one if overdue is above certain threshold (30, 60, 90, 180 or 360 days past due) and zero if it is equal or below that threshold. See section 3.3.3 for the description of classification accuracy measures.

Banks and regulators are mostly concerned about predicting new non-performing borrowers. PD as a measure of likelihood that a borrower will default on a certain horizon is also a key credit risk parameter under the IRB capital regulation. Table 27 presents the performance of the models in predicting the transition to default ($\leq$90 days overdue in time $t-1$, >90 days overdue in time t). Classical static PD model that is most frequently used in practice and where only firm specific variables are used as regressors, is shown to have very low performance. It correctly identifies only 5% of new defaults. Extending the model with dummies for different classes of overdue in $t-1$ significantly improves the performance to 27% true positive rate. This is to be expected since, as we already presented in section 3.2, past information on days past due is very informative about current default status. The higher the overdue in previous quarter, more likely it is that the firm defaults in current period. A minor change in the model can thus lead to much more accurate estimates of the default probability.

The performance in predicting new defaulters can also be calculated for the models where

Table 27: Models' performance in predicting transition to default

	Static PD	Dynamic PD	Static probit	Dynamic probit	Static tobit	Dynamic tobit
True positive rate	0.047	0.274	0.150	0.045	0.501	0.139
True negative rate	0.997	0.997	0.991	0.997	0.950	0.995
False positive rate	0.003	0.003	0.009	0.003	0.050	0.005
False negative rate	0.953	0.726	0.850	0.955	0.499	0.861
Accuracy	0.983	0.986	0.979	0.983	0.943	0.983

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the performance of the estimated models in predicting the transition to default ($\leq$90 days overdue in time $t-1$, >90 days overdue in time t). See section 3.3.3 for the description of classification accuracy measures.

either state of default or overdue is used as the dependent variable. It is interesting to find that any other version of the model outperforms the classical PD model. The only exception is the dynamic probit model, where the autoregressive term leads to persistence of states and there is thus not a lot of switching between performing and non-performing states.[25] As expected, the dynamic tobit with lagged information about days past due is better able to capture the transition to default, but however, is still performing worse than some other models. Given also low false positive rate, it seems that the autoregressive component is not strong enough to lead to a sufficient increase of overdue between two consecutive periods. This is the result of empirical findings presented in Table 19, where we show that overdue quite frequently also decreases between two consecutive periods, especially when it is still low.

The best performing model for predicting transition to default is found to be the static tobit. It achieves 50% classification accuracy of new defaulters and outperforms all other models by a large margin. It also has the highest false positive rate (5%), which also explains lower overall accuracy. This measure, however, is typically not of a primary interest in evaluating the performance of default probability models and comparing to the gain in correctly identifying defaulters, the loss of over-signalling is relatively small. We formally compare the performance of the two best performing models, dynamic PD and static tobit, using the methodology proposed by Alessi and Detken (2011). Applying equal weights on type 1 and type 2 error results in a loss of 0.365 for dynamic PD and only 0.275 for static tobit. Given that regulators and banks are typically more concerned about missing the defaulters than issuing false alarms, which would be reflected in higher weight on type 1 error, places the static tobit model to even more superior position.

Let us summarize our main results. We find two strong peace of evidence that the tobit modelling technique of credit risk is the advantageous one. Dynamic tobit model is shown

[25] As shown in Table 24, dynamic probit achieves a high accuracy in predicting non-performing borrowers, where the persistence of both states is of a key importance.

to achieve the best classification accuracy of non-performing borrowers, whereas static tobit outperforms all the other models in predicting new defaulters. The advantage of tobit model is also that it enables classifying borrowers to different groups of overdue and thus get the whole spectrum of riskiness of credit portfolio. We also show that the static PD model, that is widely used by banks and regulators, actually has the worst classifying performance. Given the evidence in our paper, it thus seems reasonable to upgrade credit risk modelling techniques, since these lead to much more accurate predictions. We now check the robustness of our results.

3.5 Robustness checks

This section presents three sets of robustness checks. First, we show the out-of-sample performance results. Second, we extend the horizon in PD models from one quarter to one year. Third, we show the dynamic model predictions on a sub-sample of firms that are present at the beginning of the sample.

3.5.1 Out-of-sample performance

The classification accuracy results presented thus far are in-sample predictions. Models are typically used to forecast credit default on a certain horizon. We therefore check the validity of our results by also predicting out-of-sample. We do this by recursively estimating the models and predicting the state of default or overdue one quarter ahead. For instance, we estimate the models until 2010q4 and forecast 2011q1. We start the estimating process in 2008q4, such that we get the estimates for all the coefficients, including the interactions between quarterly dummies and firm specific variables. The applied estimating methodology, however, needs to be simplified due to a large computational burden. Using random effects estimator, it took the computer approximately 12 days to estimate all the models presented in Table 23. Given that now all the estimates would need to be replicated 24-times, it would take a very long time to estimate all the models. We therefore use pooled estimators, which proceed much faster. The only difference comparing to random effects estimator is that the pooled version is less efficient, since it does not take into account the autoregressive structure of the variance-covariance matrix. Since we use an alternative methodology, the prediction accuracy of this procedure should not be directly compared to the results presented in previous section.

Table 28 presents the out-of-sample performance in predicting non-performing borrowers. Even though the estimation methodology is now different, the prediction accuracy is similar as presented in Table 24 for in-sample predictions. Similarly, we also find that the dynamic version of the models outperform the static ones. Dynamic probit achieves the highest classification accuracy of defaulters (78%) with low false positive rate below 1%. This model, however, is not able to break down firms to different overdue classes. Table

29 displays these results for static and dynamic tobit. Similar as before, we find that the predictions of the dynamic model are much more stable and accurate.

Table 28: Out-of-sample performance of probit and tobit model in predicting performing and non-performing borrowers

	Probit		Tobit	
	Static	Dynamic	Static	Dynamic
True positive rate	0.396	0.783	0.710	0.721
True negative rate	0.990	0.992	0.951	0.992
False positive rate	0.010	0.008	0.049	0.008
False negative rate	0.604	0.217	0.290	0.279
Accuracy	0.949	0.978	0.935	0.974

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the out-of-sample classification performance of probit and tobit models in predicting performing and non-performing borrowers (more than 90 days past due). See section 3.3.3 for the description of classification accuracy measures.

Table 29: Out-of-sample classification accuracy of static and dynamic tobit model across different groups of overdue

	Static tobit					Dynamic tobit				
Overdue threshold	30	60	90	180	360	30	60	90	180	360
True positive rate	0.887	0.777	0.710	0.592	0.427	0.751	0.725	0.721	0.739	0.780
True negative rate	0.776	0.923	0.951	0.975	0.988	0.975	0.988	0.992	0.997	0.998
False positive rate	0.224	0.077	0.049	0.025	0.012	0.025	0.012	0.008	0.003	0.002
False negative rate	0.113	0.223	0.290	0.408	0.573	0.249	0.275	0.279	0.261	0.220
Accuracy	0.786	0.911	0.935	0.955	0.969	0.955	0.968	0.974	0.983	0.991

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the out-of-sample performance of static and dynamic tobit model in classifying borrowers into different groups of days past due. In all the cases an indicator is equal one if overdue is above certain threshold (30, 60, 90, 180 or 360 days past due) and zero if it is equal or below that threshold. See section 3.3.3 for the description of classification accuracy measures.

We now turn to out-of-sample prediction accuracy of new defaulters. The results presented in Table 30 reveal that the prevailing modelling methodology, static PD model, performs very badly with true positive rate below 1%. The model is basically uninformative in identifying transitions to default. The performance can be significantly improved by moving to dynamic PD model and even more by using static tobit, which correctly classifies 53% of new defaulters. Overall, we can conclude that the out-of-sample prediction results are totally in line with the results obtained using in-sample predictions.

3.5.2 Yearly horizon of default probability

Low prediction accuracy of static PD model could be the result of modelling the default probability on quarterly horizon. The underlying default rate is a very volatile series

Table 30: Out-of-sample performance in predicting transition to default

	Static PD	Dynamic PD	Static probit	Dynamic probit	Static tobit	Dynamic tobit
True positive rate	0.008	0.233	0.157	0.005	0.528	0.114
True negative rate	0.999	0.997	0.991	1.000	0.955	0.997
False positive rate	0.001	0.003	0.009	0.000	0.045	0.003
False negative rate	0.992	0.767	0.843	0.995	0.472	0.886
Accuracy	0.984	0.986	0.978	0.985	0.948	0.983

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the out-of-sample performance of the estimated models in predicting the transition to default ($\leq$90 days overdue in time $t-1$, >90 days overdue in time t). See section 3.3.3 for the description of classification accuracy measures.

and the model may not be able to sufficiently capture all these dynamics. In addition, probabilities of default are usually estimated on a one year horizon as is also suggested by BCBS (2001) to IRB banks. Hence, to check the robustness of presented results, we re-estimate all our models using only end-of-year data. Models' specification is similar as before, with the only difference that the interactions between quarterly dummies and firm specific variables are now dropped and instead of one quarter lags, yearly lags of the dependent variables are used. Similarly, the dependent variable for PD models is now defined as transitions to default on one year horizon.

The results of yearly estimates are presented in Table 31. As expected, the prediction accuracy of static PD model is now improved. However, with 16% true positive rate it is still among the worst performing. We again find that static tobit outperforms all the other models' predictions by a large margin. It correctly classifies 56% transitions to default, which is even slightly improved comparing to quarterly estimates. Similar as we find before, static tobit issues more false alarms, but the evaluation of a loss function is still considerably in favor of this modelling approach.

Table 31: Models' performance in predicting yearly transition to default

	Static PD	Dynamic PD	Static probit	Dynamic probit	Static tobit	Dynamic tobit
True positive rate	0.159	0.235	0.248	0.145	0.559	0.354
True negative rate	0.994	0.993	0.988	0.994	0.955	0.982
False positive rate	0.006	0.007	0.012	0.006	0.045	0.018
False negative rate	0.841	0.765	0.752	0.855	0.441	0.646
Accuracy	0.963	0.966	0.961	0.963	0.941	0.959

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the performance of the estimated models in predicting the yearly transition to default ($\leq$90 days overdue in time $t-1$, >90 days overdue in time t). See section 3.3.3 for the description of classification accuracy measures.

3.5.3 Dynamic model estimates on a sub-sample of firms

Wooldridge's (2005) methodology, which we use to estimate the dynamic models, requires that the estimates are performed on a balanced panel data.[26] Due to the nature of the modelling problem, our panel is unbalanced. Firms that become overdue on their credit obligation sooner or later bankrupt and disappear from the sample. In addition, new firms are entering into the sample. In applied work researchers usually estimate the dynamic models only on balanced part of the dataset (see for instance O'Neill & Hanrahan, 2012). In our case, however, this would lead to serious sample selection bias. We would be mostly left with the firms that defaulted during the last periods of the sample. We rely on the evidence provided by Akay (2009), who shows by simulations that using Wooldridge's (2005) methodology on unbalanced panel does not lead to any serious bias. In addition, we also estimate our models on a sub-sample of firms that are represented at the beginning of the sample. We therefore exclude all the firms that subsequently enter the dataset. In this way we achieve that the initial values for all the firms are taken from the same time period (2007q4).

Table 32 presents the performance results of dynamic probit and tobit models estimated on a sample of firms present in 2007q4. As it can be seen, this only marginally changes the classification accuracy results. We can still find that the dynamic models outperform the static ones and that the dynamic tobit is the advantageous methodology for identifying non-performing borrowers. Overall, the results are in line with the findings presented in section 3.4.

Table 32: Performance of the dynamic models estimated on a sub-sample of firms

	Dynamic probit	Dynamic tobit 30	60	90	180	360
True positive rate	0.678	0.766	0.738	0.731	0.746	0.781
True negative rate	0.993	0.969	0.985	0.991	0.996	0.998
False positive rate	0.007	0.031	0.015	0.009	0.004	0.002
False negative rate	0.322	0.234	0.262	0.269	0.254	0.219
Accuracy	0.972	0.951	0.967	0.974	0.983	0.991

Source: Bank of Slovenia; AJPES; own calculations.

Notes: The table reports the performance of dynamic probit and tobit model estimated on a sample of firms present at the beginning of the sample (2007q4). Dynamic probit performance is shown for the 90 days overdue threshold, whereas dynamic tobit results are for different thresholds from 30 to 360 days. See section 3.3.3 for the description of classification accuracy measures.

[26] Honoré (2002) shows that the initial conditions problem is especially problematic in unbalanced panels.

3.6 Conclusion

In this paper we evaluate the performance of several credit default models, which are compared in their ability to correctly predict non-performing borrowers and transitions to default. In addition to conventional static binary models, we also evaluate the performance of two novel methodologies that, to our knowledge, have not yet been applied in modelling credit risk. Overdue in loan repayment is already a risk measure and therefore it seems reasonable to estimate it directly, using the tobit model methodology. In addition, state of default and overdue are highly autoregressive processes. Overdue is expected to increase in time, whereas state of default shows a lot of persistence. Estimating the dynamic probit and tobit model, where lagged dependent variable is included among regressors, can significantly improve the performance of the model. Same inputs are used in all the models, which means that the differences in classification accuracy can be fully attributed to different functional forms (probit vs. tobit) and additional information that enter the model in the form of lagged dependent variable.

We show that tobit modelling methodology outperforms all other methodologies. Dynamic tobit model is shown to achieve the highest classification accuracy in predicting non-performing borrowers. It correctly identifies more than 70% of defaulters and issues less than 1% of false alarms. In addition, its prediction is number of days past due, which enables to form different classes of overdue. This is a very valuable information, since it gives direct and easily interpretable information on expected portfolio riskiness. We show that tobit model performance is very high and stable across different overdue classes, from 30 to 360 days. High performance (66% true positive rate) is also achieved by dynamic probit, which outperforms the static version by more than 30 percentage points. This shows that the dynamic modelling methodology can significantly improve the performance of credit default models.

Tobit model also has the highest prediction ability for explaining transitions to default. In classifying firms into performing and non-performing class the dynamic structure of the model plays a crucial role. When a certain overdue threshold is bridged, it is not very likely that the borrower will become performing again. In explaining transitions to default, however, this autoregressive process is too slow to sufficiently capture the increase of overdue from one quarter to another. We therefore find the static tobit model to perform the best in terms of true positive rate. It correctly classifies 50% of new defaulters and outperforms all other models by a large margin. It also issues more false alarms, but as we show, the evaluation of loss function, which takes into account type I and type II error, is much in favor of this model. On the other hand, conventional PD model, that is typically used by banks and regulators, has very low performance. It is able to correctly identify only 5% of transitions to default. A number of robustness checks confirm the validity of our results.

The findings in this paper have several important implications for banks, banking regulation and credit risk modelling practitioners. We show that the prevailing credit risk modelling methodology, which is based on binary classifiers, can be significantly improved by including the dynamics and choosing the tobit functional form of the model. In addition, we propose a model specification to estimate credit risk on a quarterly basis, which enables much more frequent and accurate monitoring of expected changes in credit portfolio.

A more important finding of our empirical analysis is very low prediction performance of conventional static PD model. This type of model is usually used by banks to assess riskiness of their portfolio and to determine one of the crucial parameters for calculating capital requirements under IRB regulation - the probability of default. A simple upgrade of the model with dummies indicating overdue in previous period significantly improves the performance. Even higher prediction ability is achieved by static tobit model. Although the prediction of this model is not in the form of default probability and can thus not be directly used in IRB formula, it seems very useful for identifying new defaulters more accurately. This is important information for banks and regulators, since knowing which borrowers are expected to default in next period they are able to assess in advance the required loan loss provisions and capital to cover the losses. In addition, IRB regulation (BCBS, 2001) requires from banks to form classes of default probability and apply the same PD to all the firms within the class. Tobit predictions enable to form similar riskiness classes based on days past due. Combining this predictions with the information about default rate for each overdue class, one can, similarly as under IRB regulation, also attach default probability to each class.

Conclusion

In this chapter I present the main findings of my doctoral dissertation. It consists of two parts. First, I summarize the main results by answering the research questions that were highlighted in the Introduction. Second, the dissertation has several important implications, which can be divided into regulatory and methodological implications. Within this section I also comment on some possible limitations of my research.

Main findings

This doctoral dissertation helps to understand the developments related to credit risk, which was shown to be the most pressing problem for Slovenian banks during the crisis period, and proposes novel methodologies for modelling credit risk. I first show what are the main credit risk drivers of Slovenian firms. These can be decomposed into firm-specific and macroeconomic or time effects. One of the main empirical finding of this dissertation is the underestimation of credit risk by Slovenian banks in the financial crisis. As shown, this can be linked to incentives to apply discretion in credit risk assessment and can rationalise the results of comprehensive review of Slovenian banking system, which was carried out in 2013. Last but not least, I show that performance of the models in predicting credit default can be significantly improved by applying tobit modelling methodology and specifying the dynamic structure of the model. Next paragraphs elaborate on these findings by answering the research questions, which were presented in the Introduction and are the basis of the dissertation.

1. Which are the main factors that drive credit risk of Slovenian non-financial firms?

The results of the first paper show that probability of default can be explained by firm specific characteristics as well as macroeconomic or time effects. While macro variables influence all firms equally, and thus drive average default probability, firm specific variables are crucial to distinguish between firms' creditworthiness. According to the results, firms are expected to default more often if they are smaller, younger, have lower liquidity, lower cash flow, higher leverage and worse operating performance. In addition to standard financial ratios, which are the main input to default probability models, I also found an important contribution by two additional variables - number of days a firm has blocked bank account and number of bank-borrower relationships. They are both highly statistically significant and have positive effect on probability of default.

2. Do macro effects improve the performance of PD model?

Default rate is highly related to the business cycle, which suggests that macroeconomic or time effects should have an important contribution in modelling probability of default. Estimation results indeed confirm this hypothesis. I tested the performance of three macroeconomic variables. GDP growth as the main indicator of business activity was found, as expected, to have a negative effect on default probability. Similarly, the negative effect was also found for credit growth. On the other hand, higher interest rate increases the credit burden and thus have positive effect on probability of default. Among the included variables, the credit growth was found to have the highest explanatory power in terms of default probabilities. When credit growth and interest rate are put together in the model, it further improves the fit as indicated by the likelihood ratio test. The fit is additionally improved when time dummies are included instead of business cycle variables. Such result is expected, since time dummies include all variation in time, which cannot be entirely captured by macroeconomic variables.

The estimated coefficients of the macroeconomic variables should be interpreted with caution since the estimates are based on only four time observations (2007-2010). However, in terms of further analysis, where I compare the estimated PDs with credit rating classification by banks, the model with macroeconomic variables leads to the same conclusions as can be obtained using the model with time dummies.

3. Can the prediction ability of the model be improved by applying different functional form of the model (probit vs. tobit) and/or specifying the dynamic structure of the model?

The results of the third paper reveal that tobit modelling methodology outperforms all other methodologies. Dynamic tobit model is shown to achieve the highest classification accuracy in predicting non-performing borrowers. It is able to correctly identify more than 70% of defaulters with very low false alarm rate. High performance (66% true positive rate) is also achieved by dynamic probit model, which outperforms the static version by more than 30 percentage points. This shows that the dynamic modelling methodology can significantly improve the performance of the models. Another important advantage of tobit model is that its prediction is number of days past due, which enables to form any classes of overdue, not just 90-days threshold as is usually the case in binary models. As I show, dynamic tobit model has high and stable predicting performance across all the overdue classes from 30 to 360 days.

I further evaluate the performance of the models in relation to classical default probability model, where instead of state of default only transitions to default are taken into account. As I show, static tobit model correctly classifies more than 50% of new defaulters and outperforms all other modelling methodologies by a large margin. It also has the highest false positive rate (5%), but comparing to the gain in identifying new defaulters, this loss still acceptable, especially if one is more concerned in missing defaulters than issuing false alarms. On the other hand, conventional PD model, which is typically used by banks and regulators, is shown to have very low performance. It is able to correctly identify only 5% of new defaulters.

4. Are credit ratings assigned by banks in line with estimated PDs?

PDs are estimated at firm level and thus a particular firm represents the same level of risk to all banks that have exposure to this firm. The comparison of estimated PDs to credit ratings assigned by banks reveals that these two measures are quite different. Although we could expect borrowers in credit grade D to have high PDs on average, 43% had PD below 5% in the period 2007-2010. Among high-risk borrowers with PD above 50%, around 13% were classified as A borrowers and approximately 57% were classified in grades A, B or C. One could argue that this is the result of misspecification of the PD model, which leads to biased PDs. However, similar result can be obtained if, instead of estimated PDs, the distribution is made according to the number of days past due, which is used to determine the default indicator for the model (90 days threshold). This confirms that credit ratings were indeed not in line with a more objective measure of risk. A possible explanation for this difference could be that banks also use other soft information, which cannot be captured by the model. Close relationship with firms can provide a more detailed information, which cannot be inferred from firms' financial accounts but adds valuable information in assessing firms' creditworthiness. The result is, nevertheless, a first indication of banks' misclassification of borrowers in terms of credit risk.

5. Did banks underestimate credit risk in the crisis?

The results of the first paper indicate that in the crisis banks allowed for higher risk borrowers in credit grades A, B and C. As estimated by the model with time dummies as time effects, average default probability in credit grades A, B and C rose by 1.3, 1.5 and 1.5 percentage point, respectively, from 2007 to 2008. This trend continued also in 2009 and 2010. To get a more clear insight in comparing risk evaluations I also checked what would be the model-predicted rating structure if banks would have kept constant rating criteria in time. The results show that in the crisis years 2009 and 2010 the credit rating structure should have been considerably worse than the one reported by the banks.

Second paper extends these findings by showing that the ability of credit ratings to predict default deteriorated during the Great recession both in absolute terms and relative to the benchmark econometric model that uses publicly available data only. Incentives for discretionary credit risk assessment were found as an important factor behind these results. Loan-loss provisions put strong pressure on capital in the crisis period, which creates an incentive for banks to underestimate credit risk. Given that credit ratings are closely related to loan-loss provisions, this analysis indicates that under-estimation of credit risk served to inflate banks' books.

6. Was the discretion in credit risk assessment more pronounced in banks with higher incentives to underestimate credit risk?

Different groups of banks had different incentives to underestimate risk. Small banks with financially weak owners faced the most pronounced difficulties in raising capital and thus had stronger incentives to apply discretion in risk assessment. Foreign-owned banks, at the other end of the spectrum, had access to more stable sources of funding through internal capital markets. They also had better capital adequacy position and lower proportion of non-performing loans. Smaller incentives to underestimate risk by foreign banks can stem also from superior managerial and organizational capacity to absorb losses. Large domestic banks fall in between as they enjoyed the bail-out guarantee by the government.

The results of predictive capacity of credit ratings give the same ranking of groups of banks as indicated above according to their incentives to underestimate risk. The classification accuracy of credit ratings assigned by foreign-owned banks outperforms the predictive capacity of credit ratings assigned by domestic banks by a large margin. Within the group of domestic banks one also observe differences between large and state owned, and small banks. The latter group reveals the worst predictive capacity of credit ratings. The results thus confirm that discretion in credit risk assessment was indeed more pronounced in banks with higher incentives to underestimate risk.

7. Can the differences in underestimation of credit risk between banks rationalise the differences in capital shortfall revealed by the comprehensive review of Slovenian banking system in 2013?

The results of the comprehensive review, carried out in 2013, revealed significant capital shortfalls, which for all examined banks amounted to 214% of existing capital. There were, however, stark differences in the capital shortfalls between domestic and foreign owned banks. For the former the recapitalisation requirement amounted to 244% of existing capital, while for the latter this figure was only 78%. Also within the group of domestic banks one can find important differences. The recapitalisation requirement for the largest two and majority state owned banks on the market, amounted to 228% of existing capital, while for the small and predominantly privately owned the figure was 274%. The same ranking of groups of banks is also found by the analysis of predictive capacity of credit ratings. This shows that banks that were more involved in underestimation of credit risk were also more penalized by comprehensive assessment in terms of capital shortfall, since true creditworthiness of borrowers was revealed at that time and losses had to be incurred.

Regulatory implications

This doctoral dissertation has several important implications for banking regulation. With underestimation of credit risk, banks temporary avoid problems of worsening balance sheet. However, true creditworthiness is sooner or later revealed and losses need to be incurred. Theses losses are higher, the longer the banks avoid admitting problems in their balance sheets. Such was the case in Slovenia where comprehensive review, carried out in 2013, revealed huge losses and large amount of recapitalisation was required.

For future prevention of similar episodes it is thus important to monitor credit risk management practices by banks more strictly. Stricter control is already possible under current system, but regulatory forbearance is often applied in similar crisis situations. Under current regulatory paradigm, regulators need to approve the internal methodologies developed by banks to estimate risk. However, the application of these methodologies is still in hands of banks and is thus subject to discretion. Future regulation should thus go in the way to be involved also in applying credit risk methodologies in practice or to require external credit risk evaluation, which would challenge credit risk assessment by banks.

A more important result of my work is the need to monitor the incentives for discretion in credit risk assessment. Higher proportion of non-performing loans, weaker capital adequacy position and limited access to fresh funding increase banks' incentives to underestimate credit risk. In times of financial crisis significant differences across time and banks may emerge that, if persistent, may lead to a significant destabilization of the banking system. It is thus important for the banking regulation to respond to the problems of incentives to underestimate credit risk.

Current IFRS provisioning model, based on incurred losses, leads to significant procyclicality of loan loss provisions, which only increases the incentives to apply discretion in credit risk assessment. The International Accounting Standards Board intends to introduce a new provisioning model, where losses will be recognised in more forward-looking manner. In this way, the credit losses would in principle not be delayed until the default event, but would at least partly be recognized in earlier stages. This, however, assumes away the problems with discretion in valuation of assets and credit risk assessment. Despite being forward-looking in nature, such a provision could be distorted by the banks incentives to underestimate risk. In addition, a large part of provisions will still need to be recognised in times of financial distress. In this way, the expected effect of new provisioning model could be undone by amplified incentives to underestimate risk.

Stricter capital regulation can amplify incentives to apply discretion in credit risk assessment. Such was the case of European Banking Authority who in 2011 required from banks to hold at least 9% Core Tier 1 capital adequacy. The main purpose of this measure was to increase confidence in the banking system. However, it could have counterproductive effect, since higher capital requirements also increase incentives to underestimate risk. The instrument that could somewhat alleviate this problem is countercyclical capital buffer, which will increase banks' loss absorption capacity in times of financial distress and thus possibly decrease incentives to underestimate risk.

Methodological implications

The proposed model of default probability can be used by regulators for practical purposes. The model is already in use at Bank of Slovenia to estimate probability of default under different stress testing scenarios. The model could also be used to challenge banks own estimates of credit risk, similarly as adequacy of credit ratings is tested in this work. In addition to the model, the analysis also reveals which are the main credit risk drivers of Slovenian non-financial firms. This information is valuable for supervisors who can focus on the factors, which were found to have the most significant effect on default probability.

The results of the third paper have several important implications for banks, banking regulation and credit risk modelling practitioners. I show that the prevailing credit risk modelling methodology, which is based on binary classifiers, can be significantly improved by including the dynamics and choosing the tobit functional form of the model. Whereas the dynamic tobit specification is the advantageous modelling approach for predicting non-performing borrowers, static tobit shows the highest performance in classifying transitions to default. An important advantage of tobit model is that one is completely flexible in choosing the threshold number of days past due that determines the default. Moreover, I also propose a model specification to estimate credit risk on a quarterly basis, which enables much more frequent and accurate monitoring of expected changes in credit portfolio.

A more important finding of evaluation of model performance in predicting transitions to default is a very low performance of conventional default probability model that is typically used by banks and regulators. This poses a question whether this modelling technique, which at the end influences the capitalisation of IRB banks, is an appropriate methodology. A simple upgrade of the model with dummies indicating overdue class in previous period significantly improves the prediction ability of the model. The performance can be further improved by using the tobit modelling approach. Although the prediction of tobit model, which is number of days past due, cannot be directly used in IRB formula for capital requirements, this approach is much more accurate in identifying new defaulters and thus it seems reasonable to use it in practice.

References

1. Akay A. (2009). The Wooldridge Method for the Initial Values Problem Is Simple: What About Performance? Discussion Paper No. 3943, Institute for the Study of Labor (IZA), Bonn.

2. Akay A. (2012). Finite-sample comparison of alternative methods for estimating dynamic panel data models. *Journal of Applied Econometrics, 27*, 1189-1204.

3. Alessi L. & Detken C. (2011). Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity. *European Journal of Political Economy, 27*, 520-533.

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

5. Amato J.D. & Furfine C.H. (2004). Are Credit Ratings Pro-Cyclical? *Journal of Banking & Finance, 28*, 2641-2677.

6. Anderson T.W. & Hsiao C. (1982). Formulation and estimation of dynamic models using panel data. *Journal of Econometrics, 18*, 67-82.

7. Antão P. & Lacerda A. (2011). Capital requirements under the credit risk-based framework. *Journal of Banking & Finance, 35*, 1380-1390.

8. Arellano M. & Bond S.R. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*, 277-297.

9. Arellano M. & Carrasco R. (2003). Binary choice panel data models with predetermined variables. *Journal of Econometrics, 115*, 125-157.

10. Bangia A., Diebold F.X., Kronimus A., Schagen C. & Schuermann T. (2002). Ratings migration and the business cycle, with application to credit portfolio stress testing. *Journal of Banking & Finance, 26*, 454-474.

11. Bank of Slovenia (2013a). *Financial Stability Review.* Available at `https://www.bsi.si/library/includes/datoteka.asp?DatotekaId=5263`.

12. Bank of Slovenia (2013b). *Full report on the comprehensive review of the banking system.* Available at `https://www.bsi.si/library/includes/datoteka.asp?DatotekaId=5457`.

13. BCBS (2001). *The Internal-Rating Based Approach. Supporting Document to the New Basel Capital Accord.*

14. BCBS (2005). Studies on the validation of Internal Rating Systems. Working Paper No. 14, BIS.

15. BCBS (2006). *International convergence of capital measurements and capital standards: A revised framework comprehensive version.*

16. Berger A.N. & Black L.K. (2011). Bank size, lending technologies, and small business finance. *Journal of Banking & Finance, 35*, 724-735.

17. Berger A.N. & Udell G.F. (2002). Small business credit availability and relationship lending: The importance of bank organisational structure. *The Economic Journal, 112*, F32-F53.

18. Bernanke B., Gertler M. & Gilchrist S. (1996). The Financial Accelerator and the Flight to Quality. *The Review of Economics and Statistics, 78*(1), 1-15.

19. Bernanke B., Gertler M. & Gilchrist S. (1999). The Financial Accelerator in Quantitative Business Cycle Framework. In: *Handbook of Macroeconomics*, North-Holland, Amsterdam, 1341-1393.

20. Blum J.M. (2008). Why 'Basel II' may need a leverage ratio restriction. *Journal of Banking & Finance, 32*, 1699-1707.

21. Blundell R. & Bond S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115-143.

22. Bonfim D. (2009). Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking & Finance, 33*, 281-299.

23. Boot A.W.A. (2000). Relationship Banking. What Do We Know? *Journal of Financial Intermediation, 9*, 7-25.

24. Brezigar-Masten A. & Masten I. (2012). CART-based selection of bankruptcy predictors for the logit model. *Expert Systems with Applications, 39*(11), 10153-10159.

25. Brown C.O. & Dinc I.S. (2011). Too Many to Fail? Evidence of Regulatory Forbearance When the Banking Sector Is Weak. *The Review of Financial Studies, 24*(4), 1378-1405.

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

27. Carling K., Jacobson T., Lindé J. & Roszbach K. (2007). Corporate credit risk modeling and the macroeconomy. *Journal of Banking & Finance, 31*, 845-868.

28. Chamberlain G. (1984). Panel data. In *Handbook of Econometrics, Vol. 2*, Griliches Z., Intriligator M. North-Holland, Amsterdam, 1247-1318.

29. Claessens S. & van Horen N. (2014). Foreign banks: Trends and Impact. *Journal of Money, Credit and Banking, 46*(s1), 295-326.

30. Costeiu A. & Neagu F. (2013). Bridging the banking sector with the real economy. A financial stability perspective. ECB Working Paper Series, No. 1592.

31. Crouhy M., Galai D. & Mark R. (2001). Prototype risk rating system. *Journal of Banking & Finance, 25*, 47-95.

32. de Haas R. & van Lelyveld I. (2010). Internal capital markets and lending by multinational bank subsidiaries. *Journal of Financial Intermediation, 19*, 1-25.

33. Dinc I.S. (2005). Politicians and banks: Political influences on government-owned banks in emerging markets. *Journal of Financial Economics, 77*, 453-479.

34. Festić M., Kavkler A. & Repina S. (2011). The macroeconomic sources of systemic risk in the banking sectors of five new EU member states. *Journal of Banking & Finance, 35*, 310-322.

35. Foss D., Norden L. & Weber M. (2010). Loan growth and riskiness of banks. *Journal of Banking & Finance, 34*, 2929-2940.

36. Gilchrist S. & Zakrajšek E. (2012). Credit Spreads and Business Cycle Fluctuations. *American Economic Review*, *102*(4), 1692-1720.

37. Hamerle A., Liebig T. & Rösch D. (2003). Credit risk factor modeling and the Basel II IRB approach. Deutsche Bundesbank Discussion Paper, No. 02/2003.

38. Hamerle A., Liebig T. & Scheule H. (2004). Forecasting credit portfolio. Deutsche Bundesbank Discussion Paper, No. 01/2004.

39. Heckman J.J. (1981a). Heterogeneity and state dependence. In *Studies in Labor Markets*, Rosen S. University of Chicago Press, 91-139.

40. Heckman J.J. (1981b). The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In *Structural Analysis of Discrete Data with Econometric Applications*, Manski C., McFadden D, MIT Press, 114-178.

41. Hoffman P. & Santomero A. (1998). Problem Bank Resolution: Evaluating the Options. Working Paper, Wharton Financial Institutions Center.

42. Honoré B. (2002). Nonlinear models with panel data. *Portuguese Economic Journal, 1*, 163-179.

43. Honoré B. & Kyriazidou E. (2000). Panel data discrete choice models with lagged dependent variables. *Econometrica, 68*, 839-874.

44. Huizinga L. & Laeven L. (2012). Bank valuation and accounting discretion during a financial crisis. *Journal of Financial Economics, 106*, 614-634.

45. Jacobson T., Lindé J. & Roszbach K. (2013). Firm Default and Aggregate Fluctuations. *Journal of the European Economic Association, 11*(4), 945-972.

46. Jiménez G. & Saurina J. (2004). Collateral, type of lender and relationship banking as determinants of credit risk. *Journal of Banking & Finance, 28*, 2191-2212.

47. Jiménez G. & Saurina J. (2006). Credit cycles, credit risk, and prudential regulation. *International Journal of Central Banking, 2*(2), 65-98.

48. Jones S., Johnstone D. & Wilson R. (2015). An empirical evaluation of the performance of binary classifiers in the prediction of credit rating changes. *Journal of Banking & Finance, 56*, 72-85.

49. Kavčič M. (2005). Ocenjevanje in analiza tveganj v bančnem sektorju. Prikazi in analize XIII/1, Banka Slovenije.

50. Khwaja A.I. & Mian A. (2005). Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market. *The Quarterly Journal of Economics, 120*(4), 1371-1411.

51. Kirstein R. (2002). The new Basel Accord, internal ratings and the incentives of banks. *International Review of Law and Economics, 21*, 393-412.

52. Krahnen J.P. & Weber M. (2001). Generally accepted rating principles: A primer. *Journal of Banking & Finance, 25*, 3-23.

53. Marcucci J. & Quagliariello M. (2009). Asymmetric effects of the business cycle on bank credit risk. *Journal of Banking & Finance, 33*, 1624-1635.

54. Mariathasan M. & Merrouche O. (2014). The manipulation of basel risk-weights. *Journal of Financial Intermediation, 23*, 300-321.

55. Medema L., Koning R.H. & Lensink R. (2009). A practical approach to validating a PD model. *Journal of Banking & Finance, 33*, 701-708.

56. Merton R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance, 29*(2), 449-470.

57. Navaretti G.B., Calzolari G., Pozzolo A.F. & Levi M. (2010). Multinational banking in Europe - financial stability and regulatory implications: lessons from the financial crisis. *Economic Policy, 25*, 703-753.

58. Neyman J. & Scott E. (1948). Consistent estimates based on partially consistent observations. *Econometrica, 16*, 1-32.

59. O'Neill S. & Hanrahan K. (2012). Decoupling of agricultural support payments: the impact on land market participation decisions. *European Review of Agricultural Economics, 39*(4), 639-659.

60. Psillaki M., Tsolas I.E. & Margaritis D. (2010). Evaluation of credit risk based on firm performance. *European Journal of Operational Research, 201*, 873-881.

61. Rabe-Hesketh S. & Skrondal A. (2013). Avoiding biased versions of Wooldridge's simple solution to the initial conditions problem. *Economics Letters, 120*, 346-349.

62. Regulation on the assessment of credit risk losses of banks and savings banks. *Uradni List RS [Official Gazette of the RS]*, No. 28/07 with amendments 102/2008, 3/2009, 29/12 and 12/13. Available at `http://www.bsi.si/library/includes/datoteka.asp?DatotekaId=5030`.

63. Rösch D. (2003). Correlations and business cycles of credit risk: evidence from bankruptcies in Germany. *Financial Markets and Portfolio Management, 17*(3), 309-331.

64. Sapienza P. (2004). The effects of government ownership on bank lending. *Journal of Financial Economics, 72*, 357-384.

65. Sarlin P. (2013). On policymakers' loss functions and the evaluation of early warning systems. *Economics Letters, 119*(1), 1-7.

66. Sengupta R. (2007). Foreign entry and bank competition. *Journal of Financial Economics, 84*(2). 502-528.

67. Volk M. (2012). Estimating Probability of Default and Comparing it to Credit Rating Classification by Banks. *Economic and Business Review, 14*(4), 299-320.

68. Wooldridge J.M. (2005). Simple solution to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics, 20*, 39-54.

APPENDICES

LIST OF APPENDICES

Appendix A: The credit ratings model - the effect of clustering standard errors

Table A1: The credit ratings model - Estimates without clustering

	2007	2008	2009	2010	2011	2012
Credit rating A_{ijt-1}	-6.184***	-5.556***	-5.173***	-5.795***	-6.102***	-5.833***
	(0.205)	(0.196)	(0.201)	(0.236)	(0.189)	(0.150)
Credit rating B_{ijt-1}	-4.776***	-4.218***	-4.204***	-4.668***	-4.944***	-4.673***
	(0.190)	(0.187)	(0.199)	(0.232)	(0.184)	(0.142)
Credit rating C_{ijt-1}	-3.423***	-2.954***	-3.043***	-3.340***	-3.233***	-3.058***
	(0.204)	(0.199)	(0.206)	(0.236)	(0.187)	(0.146)
Credit rating D_{ijt-1}	-1.900***	-1.898***	-1.918***	-2.240***	-2.411***	-2.112***
	(0.195)	(0.199)	(0.211)	(0.238)	(0.191)	(0.151)
Constant	1.290***	1.177***	1.480***	1.946***	2.168***	1.946***
	(0.176)	(0.179)	(0.193)	(0.228)	(0.178)	(0.136)
Observations	21200	21480	23906	24926	25203	25595

Source: Bank of Slovenia; own calculations.

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: The credit ratings model - Estimates with clustering across firms

	2007	2008	2009	2010	2011	2012
Credit rating A_{ijt-1}	-6.184***	-5.556***	-5.173***	-5.795***	-6.102***	-5.833***
	(0.207)	(0.198)	(0.211)	(0.234)	(0.197)	(0.154)
Credit rating B_{ijt-1}	-4.776***	-4.218***	-4.204***	-4.668***	-4.944***	-4.673***
	(0.195)	(0.191)	(0.208)	(0.229)	(0.190)	(0.145)
Credit rating C_{ijt-1}	-3.423***	-2.954***	-3.043***	-3.340***	-3.233***	-3.058***
	(0.206)	(0.201)	(0.214)	(0.233)	(0.192)	(0.147)
Credit rating D_{ijt-1}	-1.900***	-1.898***	-1.918***	-2.240***	-2.411***	-2.112***
	(0.198)	(0.197)	(0.217)	(0.233)	(0.190)	(0.153)
Constant	1.290***	1.177***	1.480***	1.946***	2.168***	1.946***
	(0.180)	(0.181)	(0.201)	(0.224)	(0.184)	(0.137)
Observations	21200	21480	23906	24926	25203	25595

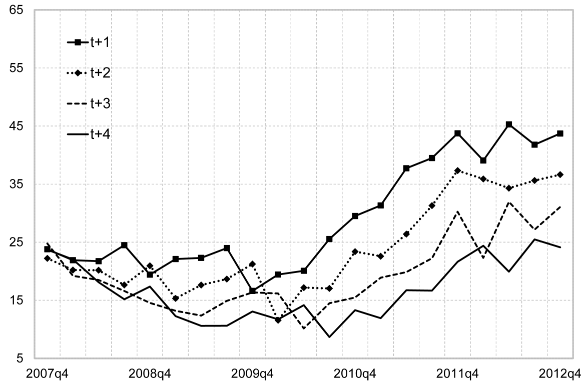
Source: Bank of Slovenia; own calculations.

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

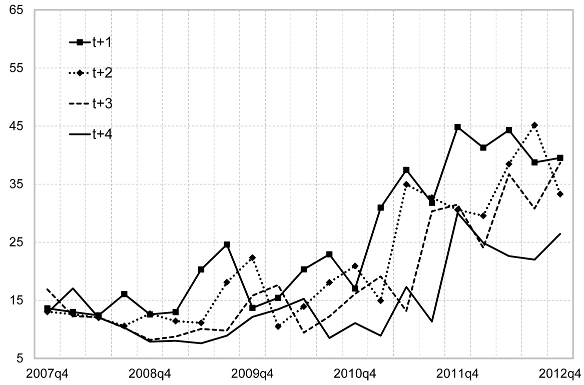
Appendix B: Classification accuracy across different forecast horizons

Figure B1: Correctly classified defaulters across groups of banks and different forecast horizons (in %)

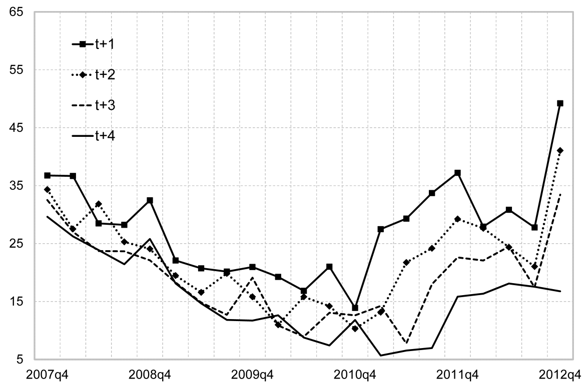
(a) Banking system



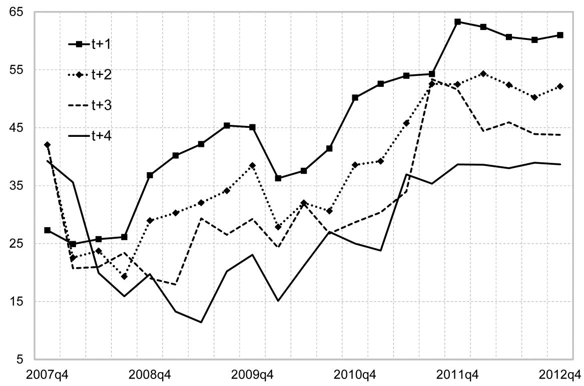
(b) Large domestic banks



(c) Small domestic banks



(d) Foreign banks



Source: Bank of Slovenia; own calculations.

DALJŠI POVZETEK DOKTORSKE DISERTACIJE V SLOVENSKEM JEZIKU

Slovenska bančna kriza je povečini izhajala iz kreditnega tveganja bank. Delež nedonosnih posojil (kjer komitenti pri odplačevanju zamujajo več kot 90 dni) je z nastopom krize začel hitro naraščati. V sektorju nefinančnih družb, ki predstavlja večino vseh slabih posojil, je v letu 2013 presegel 25%. Kriza je odprla številna zanimiva raziskovalna vprašanja z vidika boljšega razumevanja razmer kot tudi zmožnosti sprejemanja ukrepov, ki bi v prihodnje lahko preprečili podoben razvoj dogodkov. Kreditno tveganje je tako osrednja tema moje doktorske disertacije, ki je sestavljena iz treh člankov. V nadaljevanju je na kratko predstavljena motivacija v ozadju vsakega od člankov.

Z vidika finančne stabilnosti je zelo pomembno razumevanje faktorjev kreditnega tveganja, ki so tekom krize vodili do precejšnjega naraščanja deleža slabih posojil. Večina raziskovalcev pri modeliranju kreditnega tveganja uporablja podoben pristop, kot ga je razvil Altman (1968), in ugotavlja da se faktorji kreditnega tveganja v splošnem delijo na idiosinkratične in sistemske (Bangia et al., 2002; Bonfim, 2009; Carling et al., 2007 in Jiménez & Saurina, 2004). Nekateri avtorji tudi poudarjajo, da pregrevanje gospodarstva v obdobju pred krizo lahko vodi do pospešenega naraščanja slabih posojil v času krize (Festić et al., 2011; Foss et al., 2010 in Jiménez & Saurina, 2006). Prvi cilj doktorske naloge je s pomočjo mikro podatkov empirično raziskati dejavnike, ki vplivajo na kreditno tveganje oz. verjetnost neplačila slovenskih nefinančnih družb ter izbrati najboljšo specifikacijo modela, ki je uporabljena tudi v preostalih dveh člankih.

Kreditno tveganje je prociklično in je pričakovati, da poraste v času poslabšanih ekonomskih razmer. Ta proces je lahko dodatno okrepljen, če banke v naprej precenjujejo kreditno sposobnost dolžnikov. V razmerah, ko je pritisk konkurence večji in zlasti v času visoke kreditne rasti, so banke bolj popustljive pri ocenjevanju kreditne sposobnosti posojilojemalcev in tako bolj nagnjene k dodeljevanju višjih bonitetnih ocen. Dodatno imajo banke tudi spodbude za podcenjevanje tveganja v času krize, ko breme slabih posojil precej najeda kapital bank, dostop do tega pa je precej omejen. Regulatorji bi tudi v času krize lahko zahtevali od bank, da dosledno sledijo regulatornim standardom in tako vedno razkrijejo dejansko stanje v svojih portfeljih. Pokazano pa je bilo, da regulatroji v času krize preko stroškov oslabitev in rezervacij pogosto popuščajo bankam z namenom, da vsaj delno omilijo pritisk na kapital (Hoffman & Santomero, 1998 in Brown & Dinc, 2011).

Skrbni pregled slovenskega bančnega sistema, ki je bil izveden v letu 2013 (Banka Slovenije, 2013b), predstavlja edinstvene okoliščine za raziskovanje hipoteze o podcenjevanju kreditnega tveganja. Rezultati so razkrili visok kapitalski primanjkljaj bank, ki je skupaj za vse udeležene banke znašal 214% obstoječega kapitala bank. Pri tem pa obstajajo velike razlike med različnimi skupinami bank. Ugotovljen kapitalski primanjkljaj je v veliki meri posledica podcenjevanja kreditnega tveganja. Razlike v potrebni dokapitalizaciji med skupinami bank so lahko posledica različnih spodbud za podcenjevanje tveganja. Za banke v večinski tuji lasti je pričakovati, da imajo manjše spodbude za podcenjevanje tveganja, saj so imele tekom krize nižjo izpostavljenost do slabih posojil, višjo kapitalsko ustreznost ter lažji dostop do financiranja preko internega kapitalskega trga. Na drugi strani, so bile domače banke precej izpostavljene do nedonosnih posojil, katerih delež je v letu 2013 dosegel 35%. Poleg tega so bile majhne domače banke v lasti finančno šibkih lastnikov, ki niso bili sposobni zagotoviti svežega kapitala za pokritje izgub. Za skupino majhnih domačih bank je zato pričakovati, da je tekom krize imela največje spodbude za podcenjevanje tveganja. Diskrecija pri ocenjevanju kreditnega tveganja je tema drugega članka, ki sestavlja to disertacijo.

Kriza je pokazala potrebo po bolj naprednih tehnikah za modeliranje kreditnega tveganja. Raziskovalci, banke in regulatorji v večini primerov uporabljajo modele, kjer je položaj neplačila pojasnjen s finančnimi kazalci podjetij, kot so mere likvidnosti, zadolženosti, učinkovitosti, itd. Modeli so navadno ocenjeni z logit ali probit cenilko. Z vključitvijo avtoregresijske komponente v model, bi se predikcijska natančnost modela lahko precej izboljšala, saj je informacija o položaju neplačila v preteklem obdobju zelo informativna za tekoče stanje neplačila. Ocena dinamičnega nelinearnega modela pa je vse prej kot enostavna, saj za razliko od linearnega modela, problema endogenosti ni mogoče rešiti z diferenciranjem in aplikacijo ene od standardnih cenilk za dinamične panele, kot sta Arellano in Bond (1991) ter Blundell in Bond (1998). Nadaljnji korak pri modeliranju kreditnega tveganja je preskok iz binarnega modela verjetnosti neplačila na eksplicitno modeliranje zamud pri odplačevanju posojila. Podobno kot pri diskretnem modelu sem ocenil tudi dinamično specifikacijo tobit modela. Vsi ti novi modelski pristopi so tema tretjega članka. Po mojem vedenju je to prvi poskus modeliranja kreditnega tveganja z uporabo tobit metodologije ocenjevanja in aplikacijo dinamične strukture modela.

Raziskovalna vprašanja

Disertacija odgovarja na številna raziskovalna vprašanja, ki jih je moč razdeliti v dve skupini. Medtem ko se prva nanaša na različne vidike modeliranja kreditnega tveganja, je osrednja tematika druge skupine razvoj kreditnega tveganja v slovenskih bankah tekom krize.

- *Kateri so ključni faktorji, ki vplivajo na kreditno tveganje slovenksih nefinančnih družb?*
- *Ali makro učinki pomembno prispevajo k pojasnjevanju verjetnosti neplačila?*
- *Ali drugačna funkcijska oblika (tobit proti probit) in/ali dinamična specifikacija*

izboljša natančnost napovedi modela kreditnega tveganja?

- *So bonitetne ocene, ki jih ocenjujejo banke, skladne z ocenjenimi verjetnostmi neplačila?*
- *So banke podcenjevale kreditno tvegaje v obdobju krize?*
- *Je bila diskrecija pri ocenjevanju kreditnega tveganja bolj izrazita pri bankah, ki so imele večje spodbude za podcenjevanje tveganja?*
- *Lahko razlike v podcenjevanju tveganja med skupinami bank razložijo razlike v kapitalskem primanjkljaju, ki je bil ugotovljen v skrbnem pregledu slovenskega bančnega sistema v letu 2013?*

Podatki

Za odgovore na zastavljena raziskovalna vprašanja so bili uporabljeni trije viri podatkov. Prvi in najbolj pomemben vir je kreditni register Banke Slovenije, ki vsebuje informacije o kreditnih poslih na nivoju komitent-banka od leta 1993 dalje. Gre za zaupne podatke, zato individualne informacije ne morejo biti razkrite. Drugi vir so bilance stanja in izkazi uspeha slovenskih podjetij, ki jih na letni ravni zbira Agencija Republike Slovenije za javnopravne evidence in storitve (AJPES). Poleg individualnih podatkov so uporabljene tudi razne makroekonomske in finančne serije, ki so pridobljene iz Statističnega urada Republike Slovenije ter Banke Slovenije. Analiza je omejena na obdobje od leta 2007 dalje, ker pred tem letom ni na voljo podatkov o zamudah pri odplačevanju posojil, ki je ključna spremenljivka v celotni analizi.

Ugotovitve

Doktorska disertacija prispeva k razumevanju razvoja kreditnega tveganja v bankah, ki se je tekom krize pokazalo kot najbolj pereč problem za slovenske banke. Najprej so predstavljeni glavni faktorji, ki vplivajo na kreditno tveganje slovenskih podjetij. Ti so lahko v splošnem razdeljeni na tiste, ki so značilni za posamezno podjetje, ter makroekonomske oz. časovne učinke. Ena od poglavitnih empiričnih ugotovitev te disertacije je podcenjevanje kreditnega tveganja slovenskih bank v času krize. Pokazano je, da je to moč pojasniti s spodbudami za diskrecijsko ocenjevanje kreditnega tveganja. Razlike med skupinami bank pa dajejo smiselno razlago za razlike v potrebnih dokapitalizacijah, ki so bile ugotovljene v okviru skrbnega pregleda slovenskega bančnega sistema v letu 2013. Kot zadnje, pokazano je, da je z uporabo tobit metodologije in dinamične specifikacije modela, mogoče precej izboljšati uspešnost modelov kreditnega tveganja za predikcijo neplačila posojila. V naslednjih odstavkih so z odgovori na raziskovalna vprašanja bolj obširno predstavljene ključne ugotovitve.

1. *Kateri so ključni faktorji, ki vplivajo na kreditno tveganje slovenskih nefinančnih družb?*

Rezultati prvega članka kažejo, da na verjetnost neplačila vplivajo spremenljivke, ki so značilne za posamezno podjetje, kakor tudi makroekonomski oz. časovni učinki. Medtem ko makro spremenljivke vplivajo na vsa podjetja v enaki meri ter tako pojasnjujejo povprečno verjetnost neplačila, so spremenljivke na ravni podjetja ključne za pojasnjevanje razlik v tveganosti med podjetji. Skladno z rezultati je pričakovati, da ima podjetje večjo verjetnost neplačila, če je majhno, mlajše, ima nizko likvidnost, šibek denarni tok, visoko zadolženost in slabšo učinkovitost poslovanja. Poleg standardnih finančnih kazalcev, ki so ključne vhodne spremenljivke v PD model, je bil ugotovljen pomemben prispevek še dveh dodatnih spremenljivk, in sicer števila dni, ko ima podjetje blokiran transakcijski račun ter števila relacij med podjetjem in bankami. Obe spremenljivki sta visoko statistično značilni in imata pozitiven učinek na verjetnost neplačila.

2. Ali makro učinki pomembno prispevajo k pojasnjevanju verjetnosti neplačila?

Stopnja neplačila je precej povezana z gibanjem poslovnega cikla, na podlagi česar je pričakovati, da imajo makroekonomski oz. časovni učinki pomemben prispevek k pojasnjevanju verjetnosti neplačila. Rezultati potrjujejo to hipotezo. Testirana je bila pojasnjevalna moč treh makroekonomskih spremenljivk. Stopnja rasti BDP-ja, kot glavna mera ekonomske dinamike, ima pričakovano negativen učinek na verjetnost neplačila. Podobno je bil negativen učinek ugotovljen tudi za kreditno rast. Na drugi strani, višja obrestna mera povečuje breme odplačevanja posojil, zato ima pozitiven učinek na verjetnost neplačila. Za vključene spremenljivke je bilo ugotovljeno, da ima stopnja rasti posojil največjo pojasnjevalno moč za verjetnost neplačila. Testirane so bile tudi različne kombinacije makroekonomskih spremenljivk. Učinek je največji, ko sta v model skupaj vključeni kreditna rast ter obrestna mera. Pojasnjevalna moč modela se še izboljša, ko so namesto makro spremenljivk, v model vključene časovne slamnate spremenljivke. To je pričakovano, saj slamnate spremenljivke zajamejo vso časovno dinamiko, ki ne more biti v celoti zajeta z makroekonomskimi spremenljivkami.

Pri interpretaciji koeficientov makroekonomskih spremenljivk je potrebno biti pazljiv, saj ocene temeljijo na samo štirih opazovanjih (2007-2010). Ne glede na to pa pri nadaljnji analizi, ko so ocenjene verjetnosti neplačila primerjane z bonitetnimi ocenami bank, model z makroekonomskimi spremenljivkami vodi do enakih zaključkov kot model, ki namesto teh vključuje časovne slamnate spremenljivke.

3. Ali drugačna funkcijska oblika (tobit proti probit) in/ali dinamična specifikacija izboljša natančnost napovedi modela kreditnega tveganja?

Rezultati tretjega članka kažejo, da tobit modelska metodologija prekaša vse ostale testirane metodologije. Dinamični tobit model doseže najvišjo natančnost napovedi pri napovedovanju nedonosnih komitentov. Pravilno identificira več kot 70% slabih komitentov in ima ob tem zelo majhno napako druge vrste. Visoko natančnost (66% pravilnih pozitivnih napovedi) doseže tudi dinamični probit model, ki prekaša statično verzijo modela za več kot 30 odstotnih točk. To kaže, da dinamična specifikacija modela precej izboljša uspešnost napovedovanja modela. Predikcija tobit modela je število dni zamude pri odplačevanju posojila, kar je dodatna pomembna prednost tega modela, saj omogoča oblikovanje poljubnih razredov zamud in ne samo na podlagi 90-dnevne meje, ki je navadno uporabljena v binarnih modelih. Kot je pokazano v disertaciji ima dinamični tobit model visoko in stabilno predikcijsko natančnost v vseh razredih zamud od 30 do 360 dni.

Dodatno sem predikcijsko natančnost modelov ovrednotil tudi v primerjavi s klasičnim modelom verjetnosti neplačila, kjer odvisna spremenljivka namesto stanja neplačila upošteva samo prehode v položaj neplačila. Kot je pokazano, statični tobit model pravilno napove več kot 50% novih neplačnikov in z veliko razliko prekaša vse ostale metodologije. Ta model ima tudi najvišjo napako druge vrste (5%), vendar je ta v primerjavi s precej višjo sposobnostjo identificiranja neplačnikov sprejemljiva. To velja zlasti v primeru, ko se daje večja utež na napako prve vrste v primerjavi z napako druge vrste. Na drugi strani pa standardni PD model, ki se večinoma uporablja v bankah, kaže zelo nizko uspešnost, saj je sposoben identificirati zgolj 5% novih neplačnikov.

4. So bonitetne ocene, ki jih ocenjujejo banke, skladne z ocenjenimi verjetnostmi neplačila?

Verjetnost neplačila je ocenjena na nivoju podjetja, tako da določeno podjetje predstavlja enako tveganje do vseh bank, ki imajo izpostavljenost do njega. Primerjava ocenjenih verjetnosti neplačila in bonitetnih ocen kaže, da se ti dve meri kreditnega tveganja precej razlikujeta. Pričakovati bi bilo, da bi imeli komitenti z bonitetno oceno D v povprečju visoke vrednosti verjetnosti neplačila. Rezultati pa kažejo, da jih je imelo v obdobju 2007-2010 več kot 40% verjetnost neplačila manjšo od 5%. Na drugi strani pa je bilo med najbolj tveganimi komitenti z verjetnostjo neplačila nad 50%, 13% takšnih z bonitetno oceno A, 57% pa je jih je bilo razvrščenih v razrede A, B ali C. Takšno neskladje med ocenjeno verjetnostjo neplačila in bonitetnimi ocenami bi lahko bilo posledica neustrezne specifikacije PD modela. Argument proti temu je, da dobim zelo podoben rezultat, če namesto ocenjenih verjetnosti neplačila, primerjavo naredim na število dni zamude, ki je uporabljena za določitev indikatorja neplačila v modelu (zamuda nad 90 dni). To potrjuje, da bonitetne ocene niso bile skladne z bolj objektivnimi merami kreditnega tveganja. Možna razlaga teh razlik je, da banke bonitete določajo tudi na podlagi mehkih informacij, ki jih ni mogoče vključiti v model. Tesna in dolgoročna povezava banke s podjetjem lahko banki zagotavlja pomemben vir informacij, ki ji pomaga pri določitvi kreditne sposobnosti tega podjetja. Ne glede na to pa opisani rezultati predstavljajo prvo indikacijo neustreznega ocenjevanja kreditnega tveganja s strani bank.

5. So banke podcenjevale kreditno tvegaje v obdobju krize?

Rezultati prvega članka kažejo, da so banke v času krize popustile bonitetne standarde v razredih A, B in C. Ocenjeno z modelom, ki vključuje časovne slamnate spremenljivke, se je povprečna verjetnost neplačila med letoma 2007 in 2008 v bonitetnih razredih A, B in C povečala za 1.3, 1.5 ter 1.5 odstotnih točk. Podoben trend se je nadaljeval tudi v letih 2009 in 2010. Na podlagi ocenjenih verjetnosti neplačila sem tudi preveril, kakšna bi bila bonitetna struktura, če bi banke v času ohranile nespremenjene bonitetne standarde. Rezultati kažejo, da bi morala biti v dveh letih krize (2009 in 2010), za kateri je narejena analiza, bonitetna struktura precej slabša od te, ki so jo poročale banke.

Podcenjevanje kreditnega tveganja je obširneje in bolj detajlno analizirano v drugem članku. Rezultati kažejo, da se je sposobnost bonitetnih ocen za napovedovanje stanja neplačila v času krize poslabšala tako absolutno, kot tudi v primerjavi z ekonometričnim modelom, ki temelji samo na javno dostopnih informacijah. Pomemben dejavnik v ozadju tega rezultata so spodbude za podcenjevanje kreditnega tveganja. Oslabitve in rezervacije v času krize predstavljajo velik pritisk na kapital bank, zaradi česar imajo te spodbudo za podcenjevanje tveganja. Glede na to, da so bonitetne ocene tesno povezane z oslabitvami in rezervacijami, analiza kaže, da je podcenjevanje kreditnega tveganja služilo prikazovanju manjših izgub s področja kreditnega tveganja od dejanskih.

6. Je bila diskrecija pri ocenjevanju kreditnega tveganja bolj izrazita pri bankah, ki so imele večje spodbude za podcenjevanje tveganja?

Skupine bank se med seboj razlikujejo po različnih spodbudah za podcenjevanje tveganja. Majhne domače banke, ki so pretežno v lasti finančno šibkih domačih lastnikov, so imele največ težav pri pridobivanju svežega kapitala in zato največje spodbude za diskrecijsko ocenjevanje kreditnega tveganja. Na drugi strani imajo banke v večinski tuji lasti bolj stabilen dostop do financiranja preko internih kapitalskih trgov (Navaretti et al. 2010; de Haas in van Lelyveld, 2010). Ta skupina bank je imela tekom krize višjo kapitalsko ustreznost in nižji delež nedonosnih posojil. Nižje spodbude za podcenjevanje tveganj bank v večinski tuji lasti izhajajo tudi iz boljšega upravljanja s tveganji, ki se lahko kaže v boljši izbiri komitentov in večji sposobnosti absorbiranja izgub. Skupina velikih domačih bank je glede na spodbude za podcenjevanje tveganja uvrščena med obe prej omenjeni skupini, saj uživa garancijo s strani države.

Rezultati natančnosti napovedi bonitetnih ocen za obdobje 2007-2012 kažejo enak vrstni red skupin bank, kot je bil opredeljen zgoraj, glede na njihove spodbude za podcenjevanje kreditnega tveganja. Natančnost napovedi bonitetnih ocen bank v večinski tuji lasti je bila precej višja od bank v domači lasti. Razlike je moč zaznati tudi znotraj skupine domačih bank. Bonitetne ocene majhnih domačih bank so imele najnižjo napovedno moč za pojasnjevanje stanja neplačila. Rezultati torej potrjujejo, da je bilo diskrecijsko ocenjevanje kreditnega tveganja dejansko bolj izrazito pri bankah, ki so imele večje spodbude za podcenjevanje tveganja.

7. Lahko razlike v podcenjevanju tveganja med skupinami bank razložijo razlike v kapitalskem primanjkljaju, ki je bil ugotovljen v skrbnem pregledu slovenskega bančnega sistema v letu 2013?

Rezultati skrbnega pregleda so razkrili visok kapitalski primanjkljaj, ki je za vse banke skupaj znašal 214% obstoječega kapitala, ki so ga imele banke v svojih bilancah od koncu leta 2012. Ugotovljene so bile precejšnje razlike med, na eni strani, domačimi bankami, ter tujimi bankami na drugi strani. Pri domačih bankah je potrebna dokapitalizacija znašala 244% obstoječega kapitala, medtem ko je bila pri bankah v večinski tuji lasti zgolj 78%. Tudi znotraj skupine domačih bank obstajajo pomembne razlike. Pri velikih domačih bankah, ki so v večinski državni lasti, je bil ugotovljen kapitalski primanjkljaj v višini 228% obstoječega kapitala, medtem ko je bil pri majhnih domačih bankah 274%. Enak vrstni red skupin bank je bil ugotovljen tudi z analizo natančnosti napovedi bonitetnih ocen. To kaže, da so bile banke, ki so v večji meri podcenjevale kreditno tveganje, bolj prizadete v skrbnem pregledu v smislu večje potrebne dokapitalizacije, saj je bila takrat razkrita dejanska kreditna sposobnost komitentov in izgube so morale biti poknjižene.

Prispevki k znanosti

Disertacija ima pomemben prispevek k znanosti na vsaj štirih v nadaljevanju omenjenih področjih.

Prvič, pomaga razložiti ključne dejavnike, ki vplivajo na kreditno tveganje slovenskih podjetij. Poleg spremenljivk na nivoju posameznega podjetja, ki pomembno ločujejo podjetja po njihovem kreditnem tveganju, je posebna pozornost namenjena tudi variaciji v času oz. odvisnosti verjetnosti neplačila od makroekonomskih nihanj. Podobno analizo za slovenska podjetja je naredila že Kavčič (2005), vendar na drugačen način. Bistvena razlika je v določitvi odvisne spremenljivke, ki vstopa v model. V njeni raziskavi so kot mera tveganosti uporabljene interne bonitetne ocene bank. Ker so te precej podvržene subjektivni presoji banke in so zato lahko pristrana mera tveganja, sem v moji raziskavi indikator neplačila določil na podlagi zamud pri odplačevanju posojila, ki je objektivna mera kreditnega tveganja. Poleg tega so v moji analizi dodani tudi časovni oz. makroekonomski učinki, za katere sem podobno kot Bonfim (2009) ugotovil, da izboljšajo pojasnjevalno moč modela.

Drugič, v raziskavi obširno uporabljam interne bonitetne ocene bank in analiziram njihovo informacijsko vrednost v smislu verjetnosti neplačila. Bonitetne ocene so z vidika analize zelo uporabni podatki, saj so določene neposredno s strani bank in tako odražajo njihovo lastno oceno kreditnega tveganja. Obstaja zelo malo člankov, ki uporabljajo bonitetne ocene (eden takšnih je Carling et al., 2007), noben od teh pa ne testira njihove informacijske vrednosti, kar je eden od poudarkov v moji disertaciji.

Tretjič, disertacija daje odgovor o pomembnosti in posledicah spodbud pri ocenjevanju kreditnega tveganja. Banke z višjim deležem slabih posojil, šibko kapitalsko ustreznostjo in omejenim dostopom do virov financiranja imajo višje spodbude za diskrecijsko ocenjevanje kreditnega tveganja. Ta informacija je zelo koristna za regulatorje, ki lahko na osnovi rezultatov sprejmejo določene ukrepe za spremljanje teh spodbud in tako zmanjšajo podcenjevanje kreditnega tveganja. Moje delo na tem področju je sorodno z delom Huizinga in Laeven (2012). Kljub temu pa se analizi razlikujeta. Medtem ko se Huizinga in Laeven (2012) v svojem članku osredotočata na diskrecijo bank pri vrednotenju sredstev, je v moji analizi poudarek na diskreciji pri ocenjevanju kreditnega tveganja. Empirični test, ki testira hipotezo o podcenjevanju tveganja v obdobju krize, je nov in temelji na predikcijski natančnosti bonitetnih ocen za pojasnjevanje stanja neplačila komitentov. Ta je primerjana tako v času, glede na ekonometrični model, ki uporablja samo javno dostopne informacije in tako ni podvržen diskreciji pri ocenjevanju tveganja, kakor tudi med tremi skupinami bank.

Četrtič, v disertaciji predlagam novo metodologijo za ocenjevanje verjetnosti neplačila in zamud pri odplačevanju posojila. Iz tega stališča ima disertacija kar nekaj pomembnih prispevkov. Prvič, predstavljena je dinamična ocena modela kreditnega tveganja, ki za ocenjevanje uporablja Wooldridgevo (2005) metodo. Po mojem vedenju, je to prvi poskus modeliranja verjetnosti neplačila kot avtoregresijskega procesa. Drugič, pri modeliranju grem še korak dlje in namesto ocenjevanja verjetnosti neplačila v binarnem modelu, kot je navadno narejeno v literaturi (Bonfim, 2009 in Carling et al., 2007), modeliram eksplicitno zamude pri odplačevanju posojila z uporabo tobit modela. To ohrani vso informacijsko vrednost zamud in je pričakovati, da ima večjo pojasnjevalno moč. Poleg tega je zamuda kot taka že mera tveganosti in zato jo je smiselno modelirati direktno, brez transformacij. Podobno kot pri diskretnem modelu, je po Wooldridgevi (2005) metodi ocenjen tudi dinamični tobit model. To je nov pristop, ki omogoča zgodnjo detekcijo potencialno problematičnih komitentov in ima, z vključitvijo avtoregresijske komponente v model, boljšo sposobnost napovedovanja. Ta model je zato zelo uporaben za banke in regulatorje. Tretjič, predlagana je rešitev za modeliranje kreditnega tveganja na četrtletni ravni z uporabo podatkov z različno frekvenco.

Implikacije

Doktorska disertacija ima številne pomembne implikacije za bančno regulativo. S podcenjevanjem kreditnega tveganja se banke začasno izogibajo problemu slabšanja kakovosti aktive. Dejanska kreditna sposobnost dolžnikov pa je prej ali slej razkrita in izgube morajo biti priznane. Dlje kot se banke izogibajo priznavanju izgub, večja je ta na koncu. Takšna je bila tudi situacija v Sloveniji, kjer je skrbni pregled aktive leta 2013 razkril visoke izgube banke, čemur je sledila več-milijardna dokapitalizacija.

Da se v prihodnje to ne bi več ponovilo, je potreben bolj strikten nadzor nad bančnim ocenjevanjem kreditnega tveganja. Ta je bil možen že v okviru sedanje ureditve, vendar se regulatorni organi v kriznih razmerah pogosto odločajo za popuščanje regulatornih standardov, s čimer bankam začasno olajšajo poslovanje. Interna metodologija bank za ocenjevanje tveganj mora biti v okviru sedanje ureditve potrjena s strani regulatorja. Aplikacija te regulative v praksi pa je še vedno v rokah bank samih in tako podvržena njihovi subjektivni presoji. Prihodnja regulativa bi zato morala iti v smeri, da bi bila vključena tudi v procesu samega ocenjevanja tveganja. K bolj doslednemu spremljanju tveganja bi pripomoglo tudi zunanje ocenjevanje tveganja, ki bi testiralo ustreznost ocen s strani bank.

Eden od pomembnejših rezultatov te doktorske disertacije je potreba po nadzoru nad spodbudami za diskrecijsko ocenjevanje tveganja. Višji delež nedonosnih posojil, nižja kapitalska ustreznost in omejen dostop do virov financiranja lahko povečujejo spodbude bank za podcenjevanje tveganja. V času finančne krize se lahko pojavijo pomembne razlike tako v času kot med skupinami bank, ki lahko destabilizirajo bančni sistem. S tega vidika je pomembno, da se bančna regulativa odzove na problem spodbud pri ocenjevanju kreditnega tveganja.

Skladno s trenutnim modelom oslabitev in rezervacij v okviru mednarodnih računovodskih standardov se izgube iz kreditnega tveganja poknjižijo, ko te dejansko nastanejo oz., ko obstaja nepristranski dokaz o oslabitvi finančnega sredstva. To vodi do precejšnje procikličnosti oslabitev in rezervacij in tako samo še povečuje spodbude za podcenjevanje kreditnega tveganja v času ekonomske krize. Mednarodni odbor za računovodske standarde (IASB) namerava uvesti nov model oslabljevanja, kjer bodo izgube vsaj deloma lahko priznane že v naprej. Kljub temu, da je to pozitiven premik, pa zanemarja problem diskrecijskega vrednotenja sredstev in ocenjevanja kreditnega tveganja. Čeprav bodo lahko banke izgube poknjižile že v času, ko je pritisk na kapital manjši, bo ocenjevanje izgub še vedno v rokah bank in njihove subjektivne presoje. Poleg tega, bodo banke še vedno morale velik del izgub poknjižiti v času gospodarske krize. Na ta način bodo pričakovani učinki nove regulative lahko izničeni, ker ne odpravljajo problema spodbud za podcenjevanje tveganja.

Strožja kapitalska regulativa, v smislu višjih kapitalskih zahtev, lahko povečuje spodbude za diskrecijsko ocenjevanje tveganja. Takšen je bil primer Evropske Bančne Agencije (EBA), ki je leta 2011 od bank zahtevala, da imajo vsaj 9% kapitalsko ustreznost osnovnega temeljnega kapitala. Poglavitni namen tega ukrepa je bil povečati zaupanje v evropski bančni sistem. Ta ukrep je lahko imel negativne posledice, saj višje kapitalske zahteve povečujejo spodbude za podcenjevanje tveganja. Instrument, ki bi lahko deloma olajšal ta problem, je proticiklični kapitalski blažilnik, ki bo v času krize povečal sposobnost bank za absorbiranje izgub in tako potencialno zmanjšal spodbude za podcenjevanje tveganja.

Disertacija ima pomembne implikacije tudi z vidika metodologije ocenjevanja kreditnega tveganja. Predlagan model verjetnosti neplačila lahko regulatorji uporabljajo za praktične namene. Model se že uporablja v Banki Slovenije za ocenjevanje verjetnosti neplačila ob različnih predpostavkah stresnih scenarijev. Model bi lahko bil uporabljen tudi za preverjanje bančnih ocen kreditnega tveganja, podobno kot je v tej disertaciji testirana ustreznost bonitetnih ocen. Poleg samega modela, analiza razkriva tudi kateri so ključni faktorji, ki vplivajo na kreditno tveganje podjetij. To informacija je uporabna za nadzornike, ki se pri spremljanju kreditnega tveganja lahko osredotočajo na tiste faktorje, za katere je analiza pokazala, da najbolj vplivajo na verjetnost neplačila.

Rezultati tretjega članka imajo številne pomembne implikacije za banke, bančno regulativo in raziskovalce, ki se ukvarjajo z modeliranjem kreditnega tveganja. Kot je pokazano v disertaciji, je moč prevladujočo metodologijo, ki temelji na binarnih modelih, precej izboljšati s prehodom na dinamično specifikacijo in izborom tobit funkcijske oblike modela. Medtem ko dinamični tobit model doseže najvišjo natančnost napovedi pri identifikaciji nedonosnih komitentov, je statična različica modela najbolj uspešna pri predikciji prehodov v stanje neplačila. Pomembna prednost tobit modela je fleksibilnost pri izboru mejnega števila dni zamude, ki definira položaj neplačila. Poleg tega je v disertaciji predstavljana tudi metodologija za ocenjevanje kreditnega tveganja na četrtletni ravni, kar omogoča bolj pogosto in natančno spremljanje sprememb kreditnega portfelja.

Pomembnejša ugotovitev primerjave natančnosti napovedi različnih modelskih metodologij je zelo nizka uspešnost standardnega modela verjetnosti neplačila, ki ga navadno uporabljajo banke in regulatorji. To postavlja vprašanje, če je takšen modelski pristop, ki na koncu vpliva na kapitaliziranost bank, ustrezen. Enostavna nadgradnja modela s slamnatimi spremenljivkami za posamezne razrede zamude precej izboljša natančnost napovedi modela. Uspešnost pri identifikaciji novih neplačnikov je mogoče še nadalje izboljšati z uporabno tobit modelskega pristopa. Čeprav predikcija tobit modela, ki je število dni zamude, ne more biti direktno uporabljena v IRB formuli za izračun kapitalske ustreznosti, ta pristop precej izboljša identifikacijo novih neplačnikov in ga je zato smiselno uporabljati v praksi.