BANKA
SLOVENIJE
EVROSISTEM

DELOVNI ZVEZKI BANKE SLOVENIJE/
BANK OF SLOVENIA WORKING PAPERS:

MODELLING CREDIT RISK WITH A TOBIT MODEL OF DAYS PAST DUE

Arjana Brezigar-Masten
Igor Masten
Matjaž Volk

1 / 2018

Title/*Naslov*: Modelling Credit Risk with a Tobit Model of Days Past Due

No./*Številka*: 1/2018

Published by/*Izdajatelj*: BANKA SLOVENIJE
Slovenska 35
1505 Ljubljana
tel.: 01/+386 1 47 19 000
fax: 01/+386 1 25 15 516
e-mail: bsl@bsi.si
http://www.bsi.si

The BANK OF SLOVENIA WORKING PAPERS collection is drawn up and edited by the Bank of Slovenia's Analysis and Research Department (Tel: +386 01 47 19680; Fax: +386 1 4719726; Email: arc@bsi.si).
The views and conclusions expressed in the papers in this publication do not necessarily reflect the official position of the Bank of Slovenia or its bodies.

Zbirko DELOVNI ZVEZKI BANKE SLOVENIJE *pripravlja in ureja Analitsko-raziskovalni center Banke Slovenije (telefon: 01/ 47 19 680, fax: 01/ 47 19 726, e-mail: arc@bsi.si).*
Mnenja in zaključki, objavljeni v prispevkih v tej publikaciji, ne odražajo nujno uradnih stališč Banke Slovenije ali njenih organov.

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani
COBISS.SI-ID=293665536
ISBN 978-961-6960-20-5 (pdf)

Modelling Credit Risk with a Tobit Model of Days Past Due

Arjana Brezigar-Masten[1]

Bank of Slovenia

Igor Masten

University of Ljubljana, Faculty of Economics, and Grant Thornton

Matjaž Volk[1]

Bank of Slovenia

Abstract

In this paper we propose a novel credit risk modelling approach where number of days overdue is modelled instead of a binary indicator of default. Conventional approach transforms number of days overdue to dichotomous variable by applying 90-days overdue threshold and use it as the dependent variable in default probability model. However, a lot of potentially useful information is lost with this transformation. Lower levels of days past due are expected to be good predictors of future defaulters. We show that a dynamic tobit model, where number of days overdue is used as the dependent variable, significantly outperforms other more conventional approaches. It correctly identifies more than 70% of defaulters and issues less than 1% of false alarms. Its superiority is confirmed also by more accurate rating classification, higher rating stability and more timely identification of defaulted borrowers. The implications for banks are clear. By modelling number of days past due they can significantly improve risk identification and reduce procyclicality of IRB capital requirements. Moreover, predictions for number of days overdue can be very well used also for the new stress testing methodology that needs to be aligned with IFRS 9 accounting standards.

JEL-Codes: C24, C25, G21, G32, G33

Keywords: credit risk, probability of default, rating, IRB, tobit

---

Email addresses: arjana.brezigar-masten@bsi.si (Arjana Brezigar-Masten), igor.masten@ef.uni-lj.si (Igor Masten), matjaz.volk@bsi.si (Matjaž Volk)

[1]The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Bank of Slovenia.

Povzetek

V članku razvijemo in predlagamo novo metodologijo za modeliranje kreditnega tveganja, kjer je namesto binarne spremenljivke položaja neplačila, modelirano število dni zamud pri odplačevanju posojila. Konvencionalni pristop pretvori število dni zamud v dihotomno spremenljivko, ki je enaka ena, če komitent pri odplačevanju posojila zamuja več kot 90 dni. Ta je uporabljena kot odvisna spremenljivka v modelu verjetnosti neplačila. S transformacijo števila dni zamud v binarni indikator se izgubijo potencialno koristne informacije za pojasnjevanje položaja neplačila. Pričakovati je, da nižje število dni zamud dobro pojasnjuje prihodnje prehode v položaj neplačila. V članku pokažemo, da ima tobit model, kjer je število dni zamude uporabljeno kot odvisna spremenljivka, precej boljšo napovedno moč kot konvencionalni pristopi. Tobit model pravilno identificira več kot 70% neplačnikov in ima napako druge vrste manjšo od 1%. Superiornost tobit modela je potrjena tudi z bolj natančnim razvrščanjem v bonitetne razrede, večjo stabilnostjo bonitetne lestvice in pravočasnejšo identifikacijo neplačnikov. Rezultati imajo pomembne implikacije za banke, ki lahko z modeliranjem števila dni zamud precej izboljšajo identifikacijo kreditnega tveganja in zmanjšajo procikličnost IRB kapitalskih zahtev. Poleg tega so napovedi števila dni zamud lahko uporabljene tudi v novi metodologiji stres testov, ki mora slediti IFRS 9 računovodskim standardom.

1. Introduction

Credit default models are extensively used by banks and regulators. Within Basel capital regulation, internal rating based (IRB) regulation requires banks to provide their own estimates of probability of default, which is one of the key parameters that determines capital requirements (BCBS, 2001, 2006). Credit default models will play an even more important role with IFRS 9 accounting standards, where banks will need to project transitions between three stages of riskiness and use them in calculation of expected credit losses (see Official Journal of the European Union L323, 2016). Appropriate and timely identification of defaulted borrowers is therefore of key importance.

In this paper we propose novel methodology for modelling credit risk, where we directly model the exact number of days overdue. Credit default is typically modelled using discrete choice methodology as was first proposed by Altman (1968) and further adopted by Löffler and Maurer (2011), Bonfim (2009), Carling et al. (2007) and others. The binary dependent variable is usually defined following BCBS (2006) default definition, which is based on the number of days past due. The default event occurs when borrower is more than 90 days overdue. By transforming number of days overdue into a dichotomous variable, a lot of potentially useful information is lost. In addition, number of days past due is already a risk measure and therefore it seems reasonable to model it directly, without any transformations.

Credit default indicators show a lot of persistence. Once a borrower defaults (becomes more than 90 days overdue), a fast return to performing status is unlikely. Moreover, number of days overdue, once being positive, is expected to increase in time. Lower levels of days past due can therefore be used as early warning signal for potential defaulters in the future and are expected to improve model performance. Standard default prediction models are unable to take this advantage. First, they lose all the detailed information on days past due when the binary indicator of credit default is used. Second, they do not account for strong positive autoregressive dynamics in the number of days overdue. We model days overdue in a dynamic tobit panel data setting, which specifically takes this into account, and compare its performance with classical default probability models.

We evaluate the performance of the models in two ways. First, we look at their ability to discriminate between defaulted and non-defaulted borrowers using standard performance measures calculated from the contingency matrix. Second, we build a rating scale based on each model's score function. In this way the comparison is done as it would appear in reality in banks using the IRB regulatory approach. We test for some desired features of a rating scale like high accuracy in discriminating between defaulted and non-defaulted borrowers, stability over time, and timely identification of potential defaulters. 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 inclusion of the autoregressive component affects the performance in predicting defaulted borrowers. The performance of the models is evaluated using the data of Slovenian non-financial firms.

The dynamic tobit model, where we model the number of days overdue, outperforms all other models. It correctly identifies more than 70% of defaulters and issues less than 1% of false alarms. An important advantage of the tobit model is also that its prediction, the number of days past due, enables to form different classes of overdue. One can for instance predict defaulters using any overdue threshold, not only 90 days as it is standard in binary models. We show that the dynamic tobit model has high classification accuracy across different classes of overdue, from 30 to 360 days. The superiority of the dynamic tobit model is confirmed also in the rating scale analysis. Compared to other methods, the dynamic tobit model classifies more new defaulters to the worst (non-defaulted) rating class and less to other classes. Importantly, this does not come with a cost of a high false positive rate. Each rating class contains exactly the same number of firms and it is the superior ability of the dynamic tobit model to place borrowers that are close to default to the worst rating class. Moreover, the rating scale based on the dynamic tobit model is also more stable in time. Higher stability with at the same time higher accuracy clearly reveals that the dynamic tobit model has significantly better discrimination ability and more timely identification of potential defaulters.

Our paper is related to recent studies performed by Jones et al. (2015) and Bauer and Agarwal (2014). Jones et al. (2015) 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 and support vector machines. They find that newer classifiers significantly outperform all other modelling approaches. Bauer and Agarwal (2014) compare three prevailing bankruptcy approaches: (i) models based on accounting information, (ii) contingent claims based models, and (iii) hazard models, which is a mix of previous two. They show that the hazard models dominate the other two with the highest value for area under ROC curve. Our paper does not only evaluate the performance of different modelling approaches, but goes a step further and provides new pieces of evidence. We show that instead of binary classifiers, one should directly model the number of days past due. This approach contains a lot of detailed information, that is otherwise lost, and results in higher accuracy and more timely identification of defaulters. In addition, to the best of our knowledge, this is the first work where tobit and dynamic methodologies are applied to default prediction. As we show, a dynamic specification significantly improves the performance of the models in comparison to the static version.

The findings of our paper have important implications for banks and banking regulation. We show that the performance of default prediction models can be significantly improved if days overdue are modelled instead of the binary indicator of default. Importantly, this approach is simple to implement, both from the modelling perspective and also final rating classification that can be used to calculate the IRB capital requirements. In addition to the superior accuracy, banks should aim for this approach also because it results in higher rating stability and more timely identification of defaulted borrowers. This is expected to reduce the procyclicality of the IRB capital requirements. Furthermore, predictions of days past due can be very useful also for the new stress testing procedures that need to be aligned with the IFRS 9 accounting standards (see EBA, 2017). IFRS 9 defines three riskiness stages and in stress testing banks need to model the transition probabilities between these stages and use them to calculate expected credit losses under different macroeconomic scenarios. The main criterion for classification in stages is the number of days overdue (Official Journal of the European Union L323, 2016). Having the prediction for days overdue, it is fairly straightforward to calculate transitions between classes.

The rest of the paper is structured as follows. Section 2 provides the descriptive analysis of the dynamics of different credit risk measures. In Section 3, we present the methodology for estimating and evaluating different credit default models. Estimation and evaluation results are presented in Section 4. Finally, Section 5 concludes the paper.

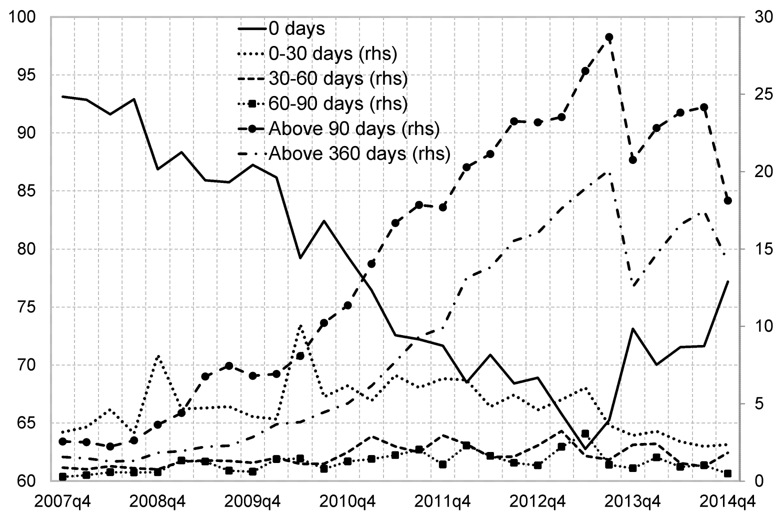
2. The dynamics of credit default measures

Basel regulation (BCBS, 2006) requires banks to classify a loan as non-performing if one of the following criteria is met: (1) the borrower is more than 90 days overdue in loan repayment or (2) the bank determines that the borrower is unlikely to pay its obligations to the bank in full amount.[2] Whereas the unlikeliness to pay can be subject to some discretion by banks, days past due are an objective measure of borrower's riskiness and are thus a key indicator to determine the default status. However, current regulation uses the information on days past due only to classify borrowers as either defaulted or non-defaulted. In the following paragraphs we demonstrate how the identification of riskiness can be improved by using the full information content of days past due.

The key data source for our analysis is the Credit registry of the Bank of Slovenia, which is exceptionally rich database containing also information that is not publicly available. The variable we are most interested in is the number of days overdue in loan repayment. It is first available in the fourth quarter of 2007, 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 a large sample of more than 1 million observations represented by a triple firm-bank-time.

Figure 1 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 rose by more than 25 percentage points until the third quarter of 2013. In 2013q4 it dropped by 8 percentage points as a result of a transfer of bad loans from two largest banks to the Bank Assets Management Company (BAMC). Thus, the decline in share of loans in overdue should not be understood as autonomous improvement of banks' credit portfolios, but rather as an institutional measure that reduced the pressing burden of non-performing loans. The 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 distress.

Figure 1: 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

[2]Similarly also holds for the EBA definition of non-performing exposure, which harmonises the definition of default across the EU (see EBA, 2016).

the 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 1 reveals that overdue is a highly autoregressive process. It can be best seen from parallel movement of loans with more than 90 days overdue and loans with more than 360 days overdue. Once the number of days overdue bridges a certain threshold, it is expected to increase in time and reach a higher number of days past due. In 83% of cases when the number of days overdue changed between two consecutive quarters, this change was positive.[3] This dynamic is, however, very heterogeneous across different classes of overdue. As can be seen in Table 1, the number of days 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 the overdue, the more likely it is, that it will further increase. This is to be expected, since once number of days overdue exceeds a certain threshold, it is not very likely that a firm will ever be able to repay the debt.

Table 1: Share of increases and decreases of overdue over different classes, in %

Overdue	One quarter horizon		One year horizon	
class	% of increases	% of decreases	% of increases	% 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 in number of days overdue over different classes of overdue and two horizons.

Looking at changes in one year horizon in Table 1, the dynamic is similar, but percentage of overdue decreases prevail only until overdue is below 10 days. In addition, with exception of last three classes, the increases of days overdue are more frequent on yearly basis than quarterly. This means that also borrowers with fewer days past due can be more problematic in the long run. Although they were mostly able to repay their debt in the short run, they might not be able to do so in the long run. Overall, Table 1 clearly reveals that the number of days overdue has a strong positive autoregressive component, especially when it is higher than 30 days. This information could be very informative in predicting credit default. However, existing default prediction models fail to account for this advantage. By transforming days past due to dichotomous variable, all such information is lost.

Default rate and its projection, probability of default, is typically of a main interest to banks, since it is one of the key factors that determines projected expected losses and capital

[3]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.

requirements for IRB banks. In addition, PD is also an important factor in loan approval and pricing. Table 2 shows the default rate over different classes of days 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 number of days overdue in previous period and it further increases when calculated on a one year horizon. Lower levels of days overdue can thus be used as an early warning signal for potential defaulters in future periods. The classical PD model, where the transition to default is typically explained with borrower-specific factors, is unable to 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 per year, which disable updating the estimated probabilities of default on the same frequency as an information on number of days overdue is refreshed.

Table 2: 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.

Our analysis so far reveals three potential upgrades of the current prevailing credit risk modelling techniques. First, the number of days overdue by itself is already a risk measure and thus it seems natural to model it directly. Valuable information is lost, when it is transformed to a dichotomous variable and modelled with a discrete choice model. Number of days past due, even if it is low, signals financial problems of a firm and it is therefore important to monitor the whole spectrum of delays in loan repayment. Second, the autoregressive component seems to be an important factor in modelling credit risk. As shown, the number of days overdue positively correlated in time. Using this information in a dynamic model could significantly improve data fit compared to a static model. Third, credit risk should be monitored and modelled on a higher frequency. The yearly frequency for modelling the probability of default, that is typically used in the literature and also proposed by BCBS (2001) to IRB banks, is a long period within which significant changes in riskiness of firms can occur. In the extreme, the number of days overdue might increase from 0 to over 360 days. A standard PD model estimated at yearly frequency is not able to capture such changes in risk profiles, since its information set is not updated within a year. It is expected that a higher-frequency dynamic model would allow for a more timely identification of financial distress.

3. Methodology

This section presents the methodology for estimating credit default models. Our choice of methodology is guided by two modelling features. First regarding the choice of the dependent variable, we compare the performance of a probit model, where the default is modelled as a binary variable, and a model where the number of days overdue in loan repayment is modelled explicitly, without any transformations. Since the number of days overdue is censored at zero, standard OLS estimator would result in biased estimates. Therefore, we apply a tobit estimator, which captures this source of non-linearity. Second, we investigate how a dynamic specification of the model affect the precision in modeling the probability of default. To this end, we estimate both a probit and a tobit model including the autoregressive term, and compare these models to their static counterparts.

Overall, we estimate and compare the respective performance of four models, which can be divided into two 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 a binary indicator of default: a *static probit* and *dynamic probit*. In the second group we model the number of days overdue in loan repayment by means of a censored regression: *static tobit* and *dynamic tobit*. 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 inclusion of autoregressive component affect the performance of the model.

To the best of our knowledge, this is the first attempt to model credit default in a dynamic panel setting. There are some analyses, like Costeiu and Neagu (2013), where past information is included in the model, but not explicitly in the form of a lagged dependent variable.

The key issue in estimating dynamic panel data models is the correlation between unobserved heterogeneity and past values of the dependent variable. In linear models, this problem can be easily solved via appropriate transformations of variables, such as first differencing, which eliminates the unobserved firm-specific effects. Although the transformed error term is correlated with transformed lagged dependent variable, instrumental variables can be used to achieve a consistent estimator.[4] In non-linear models, however, there is in general no transformation that would eliminate the unobserved effects. Suppose we are interested in modelling the process:

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

where y^*_{it} is latent index of either probit or tobit functional form, 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}. By definition, the unobserved individual-specific effect η_i is correlated with lagged dependent variable y_{it-1}. One option is to assume that initial values of dependent variable y_{i0} 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 a serious bias.

Another way of dealing with this bias is to use the fixed effect approach. Honoré and Kyriazidou (2000) and Arellano and Carrasco (2003) propose an estimation method for fixed effects logit model, which solves the initial condition problem by eliminating the unobserved heterogeneity. However, these models 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 and the resulting sample could be biased. In addition, there is no such solution for the tobit model.

[4]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 random effects solutions are much more common and attractive in practice. Wooldridge (2005) solves this issue by specifying the functional form for unobserved heterogeneity:[5]

$$\eta_i = \xi_0 + \xi_1 y_{i0} + x_i' \xi_2 + \psi_i \tag{2}$$

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 lagged dependent variable y_{it-1} is captured by equation (2), which gives another unobserved individual effect ψ_i that is not correlated with y_{it-1} and its 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 (1) with functional form (2), which results in:

$$y_{it}^* = \alpha y_{it-1} + x_{it}' \beta + \xi_0 + \xi_1 y_{i0} + x_i' \xi_2 + \psi_i + \varepsilon_{it}. \tag{3}$$

The main advantage of this methodology is that it is computationally very simple and can be implemented using standard random effects estimator. Additionally, the same methodology can be used for estimating the dynamic probit and the dynamic tobit model. Since we are interested in comparing the performance of different functional forms of credit default models, it is important that it is not affected by different methodology for estimating the probit and tobit model. A strong support for using this estimator is also a study by Akay (2012), who finds that it performs especially well in the panels that are longer than 8 periods. Our dataset consists of 29 periods.

3.1. Model specification

In order to estimate the credit default models, we link the Credit registry 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) on yearly basis. To do so, we aggregate Credit registry data to firm-time level by taking the highest number of days overdue a particular firm has to any bank in quarter t. Note that our final dataset is of a mixed frequency. Whereas Credit registry data are on quarterly frequency, balance sheet and income statement data vary only yearly. For this reason, we select model specification that explicitly accounts for this feature.

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

$$y_{it} = f(y_{it-1},\ x_{it-1}^q,\ d_j x_{it-1}^y,\ \eta_i), \quad i = 1, ..., N,\ t = 1, ..., T_i,\ j = 1, ..., 4 \tag{4}$$

where y_{it} is the dependent variable, which can take two different forms. First, for static and dynamic probit models we define the dependent variable as an indicator, which equals one if firm i is more than 90 days overdue in quarter t. Second, for static and dynamic tobit models, we keep overdue as it is, without any transformations and therefore model the exact number of days past due. y_{it-1} is a lagged value of the dependent variable, i.e. a lagged default indicator in dynamic probit case and a lagged number of days overdue in the dynamic tobit case.

Due to mixed frequency data, the distinction needs to be made between regressors that are available quarterly (x_{it-1}^q) and those that vary only yearly (x_{it-1}^y). Since the latter can have different effect across quarters, we multiply them with d_j, which is simply a quarter-specific

[5] 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.

dummy variable. In this way we get a quarter-specific effect of yearly varying regressors on our dependent variables, which are observed quarterly. All regressors are included with one period lag.[6] 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) and Brezigar-Masten et al. (2015), who model the probability of default as a function of firm size, age, liquidity, indebtedness, cash flow, efficiency, the number of days with blocked account, and the number of relations a particular borrower has with banks.[7] The last two listed variables are observed quarterly, while 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. It is important to note that these variables were selected as the best predictors in a binary choice model of default in Volk (2012) and Brezigar-Masten et al. (2015), which in our empirical application favors the probit model. Model selection in the tobit model could have resulted in a different set of predictors and a better empirical fit of the model. However, in our analysis we want to evaluate the isolated effects of the choice of the dependent variable and dynamic specification, which is most effectively achieved by keeping the set of predictors unchanged across the competing models.

Term η_i in equation (4) captures the functional form for unobserved heterogeneity. As can be seen in equation (2), Wooldridge (2005) proposes to include an initial value of the dependent variable and the realizations of other regressors in each time period. This procedure would in our case lead to approx. 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 estimate 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, we use the following functional form for individual specific effects in dynamic probit and tobit model:

$$\eta_i = \xi_0 + \xi_1 y_{i0} + x'_{i0}\xi_2 + \bar{x}'_i\xi_3 + z'_i\xi_4 \tag{5}$$

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 number of days past due in dynamic tobit case. The majority of initial values are taken from 2007q4 when our dataset starts. However, for those firms that enter subsequently, their first observation is taken as an inital value.[8] x_{i0} is a vector of inital values for all the regressors, whereas $\bar{x}_i$ are within means of the regressors, defined as $\frac{1}{T_i}\sum_{t=0}^{T_i} x_{it}$ [9]. 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 models. There are mainly two reasons

[6]For variables that are observed at yearly frequency this means including their values from previous year.

[7]We also ran a stepwise selection procedure, which resulted in a model with very similar performance. The results are available upon request.

[8]For robustness, 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). The results are in line with the findings presented in section 4.

[9]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.

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 5 for dynamic models, with the only difference that we exclude initial values of the dependent variable.

3.2. Model comparison

The comparison of competing models proceeds in two ways. First, we calculate several performance measures from the contingency matrix presented in Table 3. The columns represent the actual observed state, whereas the rows are predicted state by the model. The prediction accuracy measures that we use are shown under Table 3. The measure that we put considerable weight on 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 at the cost of issuing too many false alarms. For this reason, we show also other measures that will help us assess model classification precision. Accuracy, as an overall classification accuracy measure, is also important, but is largely driven by the classification of non-defaulters, which represent the largest share of observations in our data.

Table 3: 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)

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

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

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

Second, we build a rating scale for each model. Rating classification is a standard procedure in the IRB regulatory approach. Even though a bank assesses the probability of default for each borrower, it needs to aggregate them to rating classes and apply the same PD, typically the realised default rate, to all the borrowers within the class. We build a rating scale mainly for two reasons. First, to compare the model performance as would appear in real application in banks. Second, to demonstrate how tobit model can be applied to rating classification and used in the IRB approach. Even though the prediction of the tobit model is the number of days past due, it can be used to define the rating scale and calculate the PD for the IRB purposes in exactly the same way as standard binary models are.

The evaluation exercise is based on in-sample predictions. We are interested if different functional forms, tobit vs. probit, lead to more accurate identification of defaulted borrowers and more stable rating classification. These relations are not expected to change over time, since the information set for all the models is exactly the same.[10]

[10]As robustness check we evaluate also the out-of-sample performance. We recursively estimate the model using the data until each quarter and predict one quarter ahead. We use simplified pooled OLS methodology due to a large computational burden of random effects estimator. As expected, results are in line with presented in-sample predictions and are available upon request.

4. Results

In this section we present the results. First, we present the estimated coefficients for all four models. We then turn to performance of the models and show that dynamic tobit model outperforms all other models and results in the most timely identification of new defaulters and the most stable rating classification.

Table 4 presents the estimated coefficients of all models. In addition to the variables that are shown in the table, the models include also controls for unobserved heterogeneity as presented in section 3.1. 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 dependent variables display, as expected, positive and highly statistically significant effect. Estimated coefficient for the lagged indicator of default shows that the default status, 0 or 1, is highly persistent. Being zero in the previous quarter, it very likely stays zero also in the 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 lagged dependent variable is also found in the dynamic tobit model, which shows that number of days overdue is expected to increase in time. All these results are in line with the findings presented in section 2.

Results in Table 4 also reveal 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 the 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.

We now turn to classification accuracy of models. Table 5 presents the classification accuracy measures of probit and tobit models in predicting defaulted and non-defaulted borrowers. Defaulted borrowers are the ones with more than 90 days overdue in loan repayment. The dynamic tobit model shows the best performance. It correctly identifies more that 71% of defaulters and has the highest overall accuracy of 97.3%. A high true positive rate, 68.8%, is also achieved by the static tobit model, but this comes with more than 5% of false alarms. When the dynamic tobit specification is chosen instead, the proportion of false alarms drops below 1%. Note that the dynamic specification significantly improves also the performance of the probit model. A true positive rate increase by more than 30 percentage points compared to the static version of the probit model, but is, nevertheless, still about 5 percentage points lower than in the dynamic tobit model.

Overall, these results indicate that a dynamic model specification significantly increases the classification accuracy in predicting default. Both dynamic model specifications outperform their static counterparts. The dynamic tobit appears the most accurate according to virtually all measures, but the dynamic probit does not appear significantly worse. It should be noted, however, that the basis for comparison for both is predicting the indicator of 90 days past due. The potential advantages of the tobit specification, however, become more pronounced in application where the whole distribution of days past due is of importance, such as the rating scale assignment through the cycle, which is what we consider in subsection 4.1.

Table 6 shows the classification accuracy of dynamic probit and dynamic tobit model in predicting corporate default where we let the autoregressive process to proceed four quarters ahead. These are in-sample predictions where 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 the dynamic tobit is the superior model also on a longer horizon. Its true positive rate is expectedly decreasing on a longer forecast horizon, but it stays above the performance of the probit model. The dynamic probit achieves a slightly higher

Table 4: Estimated coefficients

	Static probit	Dynamic probit	Static tobit	Dynamic tobit
Dependent variable	$I(>90)_{it}$	$I(>90)_{it}$	Overdue$_{it}$	Overdue$_{it}$
Dependent variable$_{it-1}$		2.096***		1.067***
log(Total sales)$_{it-1}$	-0.182***	-0.056***	-81.820***	1.369*
Age$_{it-1}$	0.267***	0.133***	60.193***	11.249***
Quick ratio$_{it-1}$	-0.023***	-0.014***	-1.368***	-1.249***
Debt-to-assets$_{it-1}$	0.005*	0.002	2.954***	0.802***
Cash flow ratio$_{it-1}$	-0.011	-0.018	-8.654***	-7.447***
Asset turnover ratio$_{it-1}$	-0.263***	-0.149***	-26.766***	-22.156***
No. of days with blocked account$_{it-1}$	0.017***	0.010***	2.805***	1.053***
No. of relations$_{it-1}$	0.345***	0.198***	69.505***	29.457***
d2*log(Total sales)$_{it-1}$	0.026***	0.019***	2.980***	-0.449
d2*Age$_{it-1}$	-0.003	-0.002	-0.040	-0.524***
d2*Quick ratio$_{it-1}$	0.019***	0.011*	0.123	0.269
d2*Debt-to-assets$_{it-1}$	0.006*	0.006	0.750	-0.072
d2*Cash flow ratio$_{it-1}$	-0.050***	-0.027	-1.355	1.759
d2*Asset turnover ratio$_{it-1}$	-0.007	-0.010	5.392**	7.079***
d3*log(Total sales)$_{it-1}$	0.053***	0.037***	2.529***	-0.793*
d3*Age$_{it-1}$	-0.005***	-0.003	0.743**	-0.238
d3*Quick ratio$_{it-1}$	0.019***	0.008	1.381***	1.294***
d3*Debt-to-assets$_{it-1}$	0.008**	0.008*	1.332**	-0.113
d3*Cash flow ratio$_{it-1}$	-0.054***	-0.021	-10.257***	0.531
d3*Asset turnover ratio$_{it-1}$	0.002	0.001	1.306	5.449***
d4*log(Total sales)$_{it-1}$	0.057***	0.022***	5.843***	-0.651
d4*Age$_{it-1}$	-0.007***	-0.003	0.099	-0.783***
d4*Quick ratio$_{it-1}$	0.019***	0.007	1.360***	1.256***
d4*Debt-to-assets$_{it-1}$	0.007**	0.005	1.634***	-0.226
d4*Cash flow ratio$_{it-1}$	-0.056***	-0.020	-12.949***	1.365
d4*Asset turnover ratio$_{it-1}$	0.028**	0.031**	4.097*	9.355***
Constant	-10.629***	-7.002***	-824.385***	-190.530***
Observations	517964	517964	517964	517964

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. The dependent variable in both tobit models is the exact number of days past due. No. of days with blocked account 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. In addition to the variables that are shown in the table, the models also include controls for unobserved heterogeneity as described in section 3.1. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

overall accuracy, but this is only due to better prediction of non-defaulters. A false positive rate still stays very low for both models.

The tobit model is suitable to form predictions over different overdue classes. Table 7 shows the classification accuracy results of the static and dynamic tobit model for five thresholds of days past due. All the classes are defined in the same way: when the number of days overdue or its prediction exceeds a certain threshold, the indicator is equal one, otherwise it is zero. The results confirm the superiority of the dynamic tobit model. A true positive rate achieved by the

Table 5: Performance of probit and tobit model in predicting defaulted and non-defaulted 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 3.2 for the description of classification accuracy measures.

Table 6: Performance of dynamic probit and tobit model in predicting defaulted and non-defaulted 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 the dynamic probit and the dynamic tobit model in predicting performing and non-performing borrowers (more than 90 days past due) one (1q) to four (4q) quarters ahead. See section 3.2 for the description of classification accuracy measures.

static tobit is decreasing with a higher overdue threshold. The 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. However, it also has significantly higher false positive rate, which for the 30-days class equals to 20%, comparing to only 3% of the dynamic model. Overall accuracy of the dynamic tobit is higher across all the classes.

The results presented in Table 7 have important implications for the stress testing procedures that need to be aligned with the new IFRS 9 accounting standards (see EBA, 2017). Under IFRS 9 banks need to classify loans into three stages, where number of days past due is the key criterion for this classification. At origination, a loan is classified to Stage 1, which is the class with a stable risk profile. If the credit risk increases significantly, the loan needs to be assigned to Stage 2, and finally to Stage 3 when it defaults. Among other criteria, the regulation states that the credit risk increases significantly when contractual payments are more than 30 days past due and the borrower defaults when overdue exceeds 90 days (see Official Journal of the European Union L323, 2016).[11] Given that the tobit model has very high prediction accuracy across all

[11]There are several other criteria that determine the significant increase in credit risk but the 30-days overdue

Table 7: 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 3.2 for the description of classification accuracy measures.

these classes (see Table 7) and that its prediction is the exact number of days past due, it could be very well used for stress testing purposes based on the IFRS 9 accounting standards.

Let us summarize our main findings up to this point. We show three pieces of evidence that confirm the superiority of the dynamic tobit model. It has the highest prediction accuracy for identifying the defaulted borrowers and importantly, with a low false positive rate. An advantage of the tobit model is also that it enables to form different groups based on number of days past due and it shows very good performance across all these groups. We now turn to the rating classification, where we compare model performance as would appear in reality in banks using the IRB approach and demonstrate how the tobit model can be used for the IRB rating classification in exactly the same way as binary models are. As we show below, the dynamic tobit model superiority is confirmed also in this application.

4.1. Rating dynamics through the cycle

We build a 10-grade rating scale with 9 classes for non-defaulted borrowers and one for defaulters. Standard 90-days past due threshold is applied to determine the default status. We do the classification based on each model's score function by splitting its values to nine equally-sized classes. For instance, 11.1% of borrowers with lowest values of score function are classified in grade one and so on. This classification applies only to borrowers that were not in default already in previous period. By previous period we mean either previous quarter or previous year, since we compare the model performance over two horizons of default probability. Borrowers that were in default already in previous period are assigned to class 10.

Once the rating scale is defined, we attach default probabilities to each class. These are calculated as through-the-cycle (TTC) default rate[12] and are reported in first four columns of Table 8, across two horizons. As expected, yearly default rates are higher since more borrowers pass to default status on longer horizon.

Perfect rating classification would assign all the firms that are expected to default in the following period to the lowest rating class. The dynamic tobit model is closest to this perfection.

threshold is the most important one. Similarly for the definition of defaulters (Stage 3) where it is assumed that the default does not occur later than when a borrower is 90 days past due.

[12]For instance, TTC default rate for first rating class is calculated by counting the number of transitions (on quarterly or yearly horizon) to default in total period 2007q4-2014q4 and dividing this by total number of observations in first rating class. Note that all classes contain the same number of observations, i.e. 11.1% of non-defaulters.

As can be seen from Table 8, the dynamic tobit has the highest TTC default rate in the last (non-defaulted) rating class and the lowest in all other classes. This holds for both, quarterly and yearly horizon. Another piece of evidence of this superiority is presented in Figure 2, where we display share of new defaulters that were at the time of default classified in the last (non-defaulted) rating class. Ideally, this would be equal to 100%. We again find that the dynamic model is the most successful in identifying new defaulters since it assigns the largest proportion of new defaulters to the last rating class and on average the lowest proportion to other classes (see Figure A1 and A2 in the Appendix). Its superiority is especially pronounced on a shorter - quarterly - horizon. This was expected, since detailed information on the number of days past due in previous quarter is very informative for the default probability in current period. Standard binary models are unable to use this advantage.

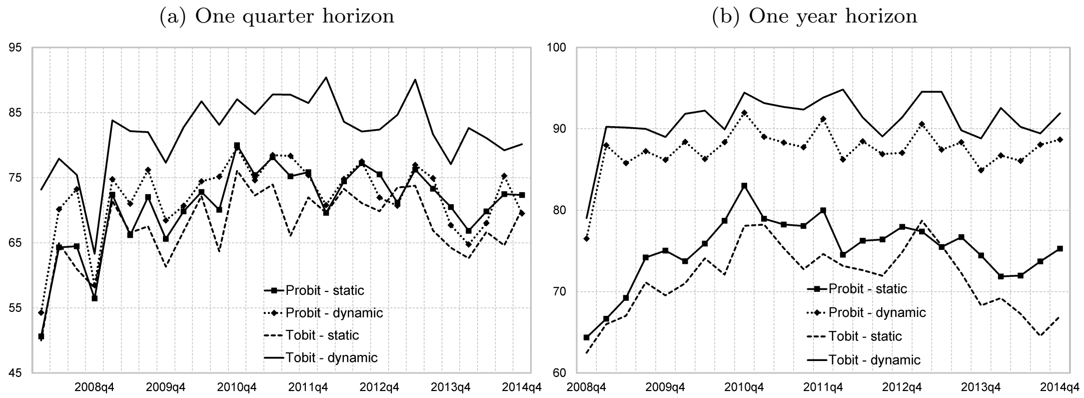
Table 8: Through the cycle default rate and standard deviation of number of firms, across ratings

	TTC Default rate in %				St.dev of number of firms			
	Probit		Tobit		Probit		Tobit	
	Static	Dynamic	Static	Dynamic	Static	Dynamic	Static	Dynamic
	One quarter horizon							
Rating 1	0.08	0.10	0.11	0.06	2096	2039	1795	1301
Rating 2	0.14	0.13	0.15	0.05	1515	1482	1299	1116
Rating 3	0.23	0.20	0.21	0.10	1069	1029	825	697
Rating 4	0.26	0.21	0.24	0.08	918	863	683	493
Rating 5	0.31	0.31	0.39	0.16	911	878	708	578
Rating 6	0.45	0.39	0.48	0.20	1037	1026	870	800
Rating 7	0.70	0.67	0.80	0.41	1273	1271	1080	990
Rating 8	1.63	1.60	1.86	1.27	1435	1396	1170	893
Rating 9	9.43	9.60	8.97	10.88	346	304	266	145
	One year horizon							
Rating 1	0.19	0.12	0.24	0.07	1889	1858	1623	1192
Rating 2	0.29	0.14	0.30	0.06	1345	1311	1142	974
Rating 3	0.36	0.18	0.38	0.06	933	894	729	600
Rating 4	0.48	0.21	0.49	0.11	769	721	560	409
Rating 5	0.58	0.29	0.72	0.15	763	744	610	490
Rating 6	0.85	0.45	1.01	0.19	921	929	769	725
Rating 7	1.48	0.74	1.70	0.50	1175	1182	990	869
Rating 8	3.84	1.98	4.40	1.75	1264	1223	1011	714
Rating 9	24.80	28.74	23.61	29.97	203	186	147	172

Source: Bank of Slovenia, own calculations.
Notes: The table reports through the cycle (TTC) default rate and standard deviation of number of firms in each rating class. Both statistics are calculated for the period 2007q4-2014q4.

One might be concerned that the superiority of the dynamic tobit model comes with a cost of a large number of false alarms. For instance, the model could assign a large proportion of firms to the lowest rating grade and this could be the reason why we find majority of new defaulters in rating grade 9. This is not the case because each rating class contains exactly the same number of observations (over the whole cycle), i.e., 11.1% of non-defaulters and it is the ability of the dynamic tobit model to place in rating 9 exactly those firms that are close to default. As we show next, contrary to the above concern, rating scale based on the dynamic tobit model is also the most stable one.

Figure 2: Share of new defaulters classified in rating 9 over two horizons for calculating the default



Source: Bank of Slovenia, authors' calculations.

A desired feature of the rating system is stability of rating assignments through the business cycle. Firm financial conditions change over time and, depending on the properties of the rating model, this typically leads to changes in obligors' credit ratings. However, it is desired that number of downgrades and upgrades through the business cycle is limited. This is especially important if the rating scale is used to calculate IRB capital requirements. Large number of downgrades in recession could induce strong procyclicality of capital requirements since there would be a sudden shift in calculated risk weights. It is therefore important that the model timely assigns appropriate rating that through time changes as least as possible.

We measure rating stability with the standard deviation of the number of firms in each rating class over time. This statistic is reported in last four columns in Table 8. The dynamic tobit model results in the most stable rating scale. It has the lowest standard deviation of number of firms across all ratings. The only exception is rating 9 on a one year horizon. Additional support to this result is shown in Figure A3 in Appendix where we display differences in number of firms between two consecutive quarters (calculated as four-quarter moving average) for each credit rating. As can be seen, the dynamic tobit model results in the most stable rating scale with on average the lowest movement of firms across rating grades. It displays the lowest number of downgrades from the first rating class at the beginning of the crisis and the lowest number of downgrades to the last rating class in recovery period. In general, similar observations apply to other rating classes. Rating transitions are on average the lowest for the dynamic tobit model.

Overall, given that the dynamic tobit model is also the most accurate methodology and with the highest default rate in the last rating class, this clearly reveals that it has significantly better discrimination ability and more timely identification of potential defaulters.

5. Conclusion

In this paper we introduce two novelties in empirical modelling of the probability of default in bank portfolios. First, we propose and evaluate the performance of credit risk model where the number of days overdue is used as the dependent variable. Overdue in loan repayment is already a risk measure and therefore it seems reasonable to estimate it directly, using the tobit model methodology. Second, the state of default and number of days overdue are highly autoregressive processes. Days past due is expected to increase in time, whereas state of default shows a lot of persistence. Estimating a dynamic model, where lagged dependent variable is included among regressors, can significantly improve the performance of the model. Since in our empirical application all competing models use the same predictors, the differences in classification accuracy can be fully attributed to different functional forms (probit vs. tobit) and/or additional information that enter the model in the form of lagged dependent variable.

We show that the dynamic tobit model outperforms all other models. 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 the performance of the dynamic tobit model is very high and stable across different overdue classes, from 30 to 360 days. This information is important also in light of the stress testing procedures that need to be aligned with the new IFRS 9 accounting standards. In this respect, banks need to classify loans to three stages. Key indicator that determines the classification across stages is the number of days overdue. Predictions for the number of days past due can therefore be used to count transitions between stages and use them in the calculation of expected credit losses.

Dynamic tobit model superiority is confirmed also by rating scale analysis. We build 10-grade rating scale based on each model's score function. The dynamic tobit model assigns more new defaulters to last rating class and less to other classes, comparing to other methods. Its rating scale is also more stable in time and leads to more timely identification of defaulted borrowers. These findings have important implications especially for banks using the IRB regulatory approach to calculate capital requirements. By modelling days overdue instead of a binary indicator of default, banks could benefit a lot by increasing accuracy in risk identification and by reducing procyclicality of capital requirements. The latter is the result of a more stable rating classification with more timely identification of defaulted borrowers.

Overall, our results show that a small change in modelling approach, where binary dependent variable is replaced by exact number of days overdue, leads to significant improvements in risk identification. Given also other benefits like higher rating stability and potential use of predictions of the number of days past due for stress testing purposes, banks should put some effort to implement this approach in practice. Importantly, the implementation should be fairly straightforward. The only novelty that banks would need to take care of is to change the dependent variable in the model and to apply tobit modelling approach. Rating classification, on the other hand, works in exactly the same way as with binary classifiers, which implies that our novel methodological approach can be easily integrated into conventional risk management procedures in banks.

References

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

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

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

[4] 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.

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

[6] Bauer J. & Agarwal V. (2014). Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance, 40*, 432-442.

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

[8] BCBS (2006). *International Convergence of Capital Measurements and Capital Standards: A Revised Framework Comprehensive Version.*

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

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

[11] Brezigar-Masten A., Masten I. and Volk M. (2015). Discretionary Credit Rating and Bank Stability in a Financial Crisis. *Eastern European Economics, 53*, 377-402.

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

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

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

[15] EBA (2016). *Final Report - Guidelines on the application of the definition of default under Article 178 of Regulation (EU) No 575/2013.*

[16] EBA (2017). *2018 EU-Wide Stress Test, DRAFT methodological note.*

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

[18] 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.

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

[20] 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.

[21] Löffler G. & Maurer A. (2011). Incorporating the dynamics of leverage into default prediction. *Journal of Banking & Finance, 35*, 3351-3361.

[22] Official Journal of the European Union L323 (2016). Commission regulation (EU) 2016/2067 of 22 November 2016 amending Regulation (EC) No 1126/2008 adopting certain international accounting standards in accordance with Regulation (EC) No 1606/2002 of the European Parliament and of the Council as regards International Financial Reporting Standard 9.

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

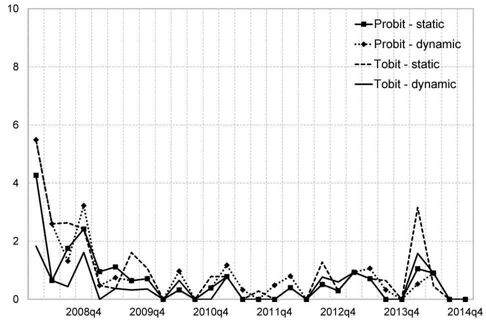
[24] Volk M. (2012). Estimating Probability of Default and Comparing it to Credit Rating Classification by Banks. *Economic and Business Review, 14*(4), 299-320.

[25] 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.

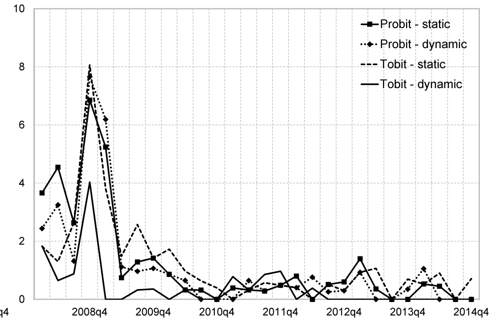
Appendix: Additional figures across credit ratings

Figure A1: Share of new defaulters across ratings - one quarter horizon of default

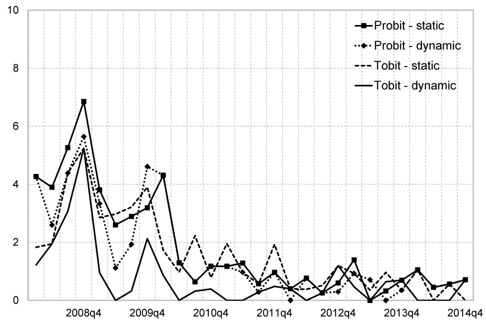
(a) Rating 1



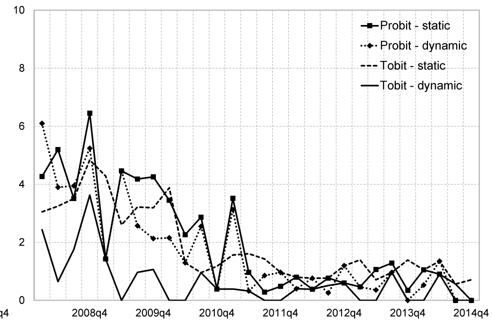
(b) Rating 2



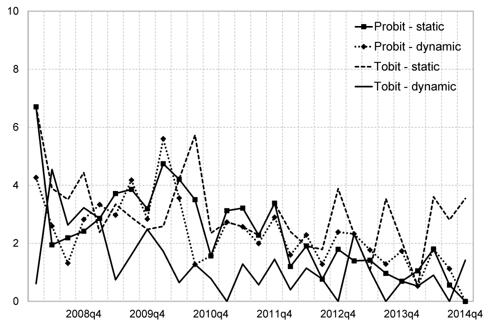
(c) Rating 3



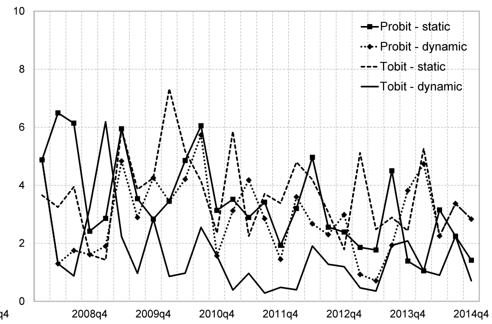
(d) Rating 4



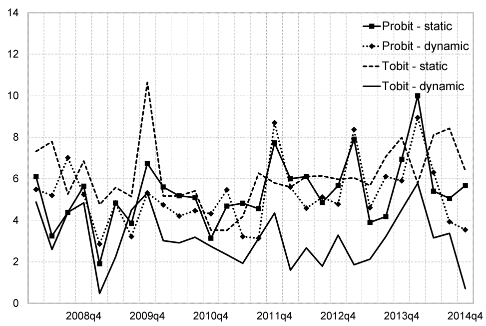
(e) Rating 5



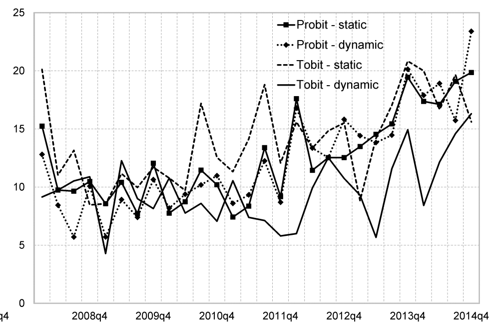
(f) Rating 6



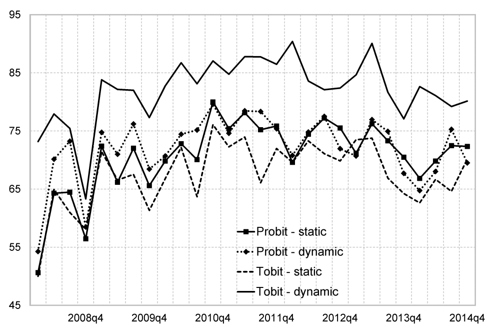
(g) Rating 7



(h) Rating 8



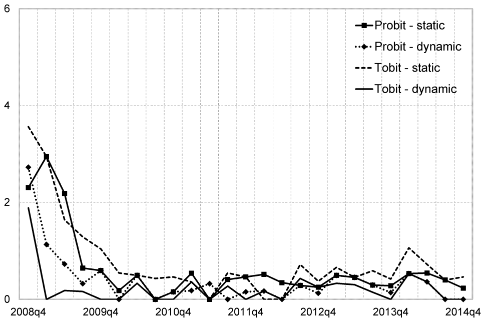
(i) Rating 9



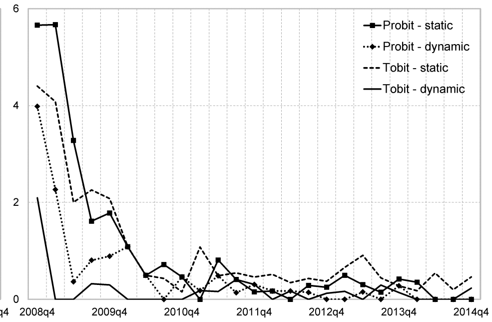
Source: Bank of Slovenia, authors' calculations.

Figure A2: Share of new defaulters across ratings - one year horizon of default

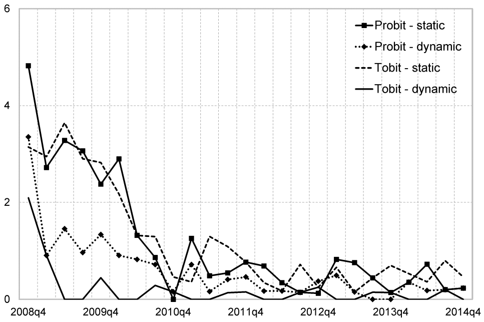
(a) Rating 1



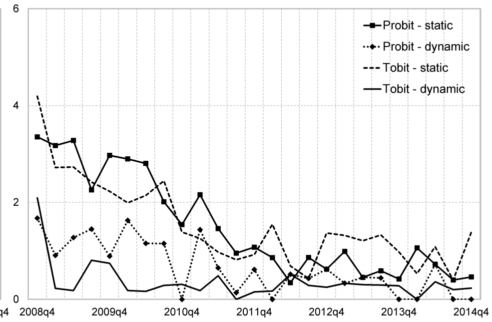
(b) Rating 2



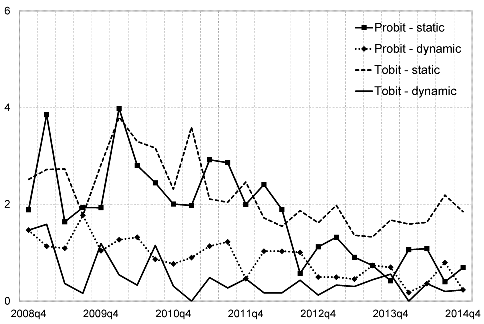
(c) Rating 3



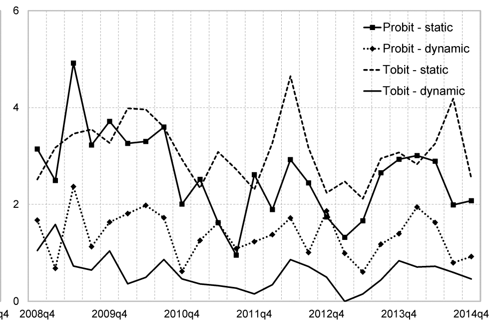
(d) Rating 4



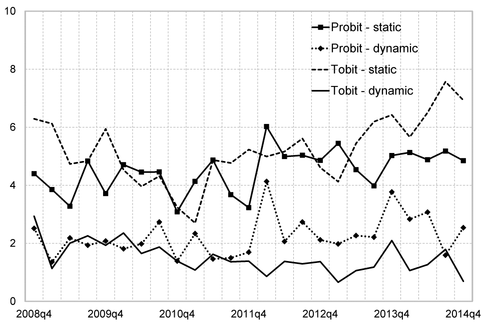
(e) Rating 5



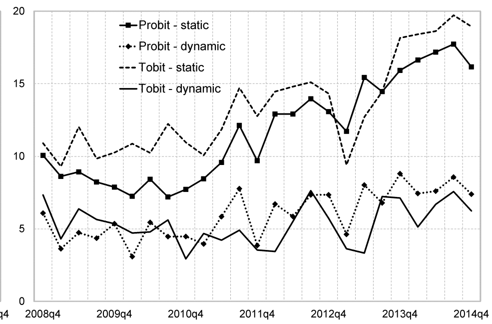
(f) Rating 6



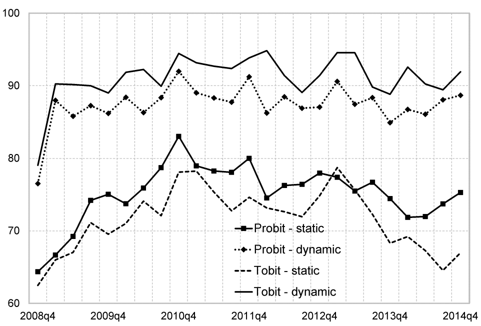
(g) Rating 7



(h) Rating 8



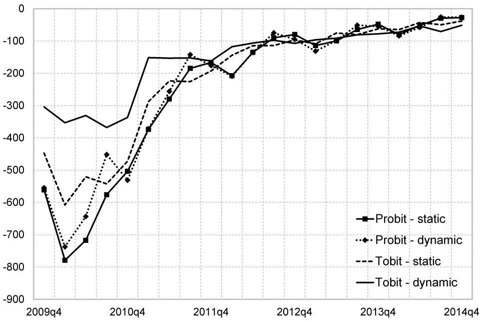
(i) Rating 9



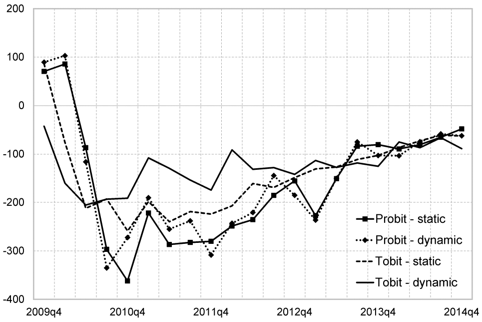
Source: Bank of Slovenia, authors' calculations.

Figure A3: Differences in number of firms across ratings - moving average for one year horizon of default

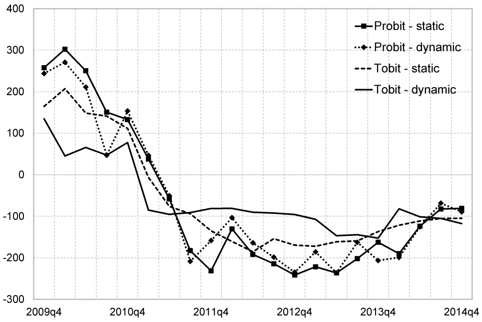
(a) Rating 1



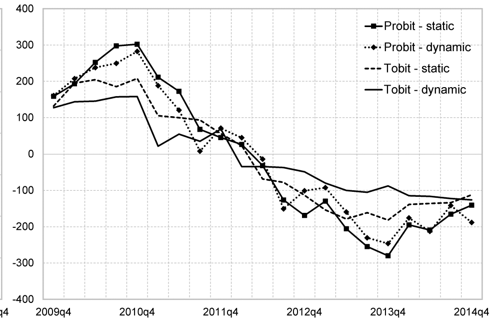
(b) Rating 2



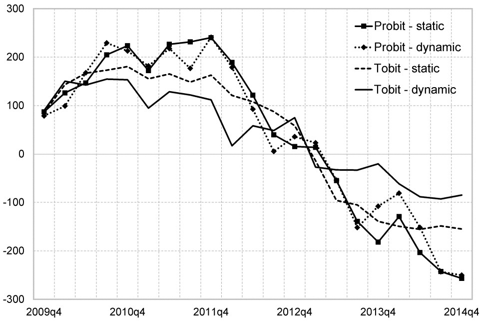
(c) Rating 3



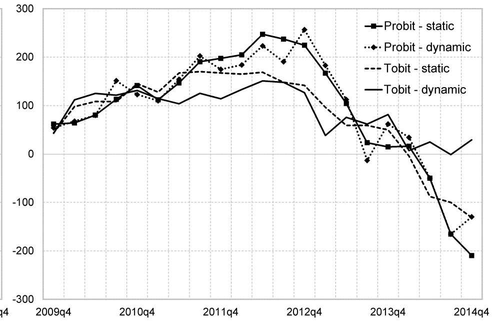
(d) Rating 4



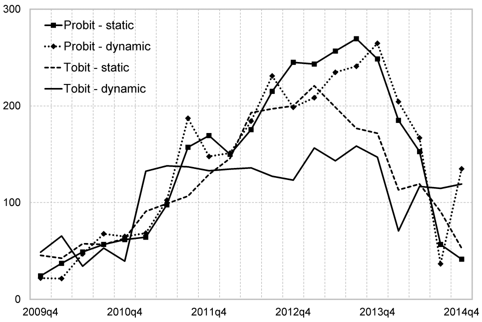
(e) Rating 5



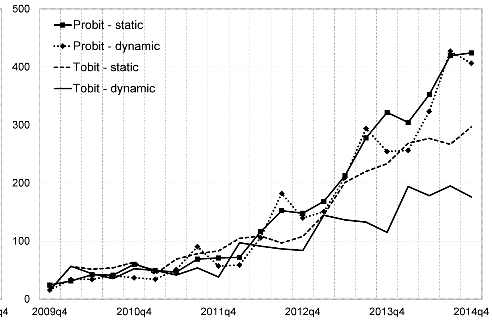
(f) Rating 6



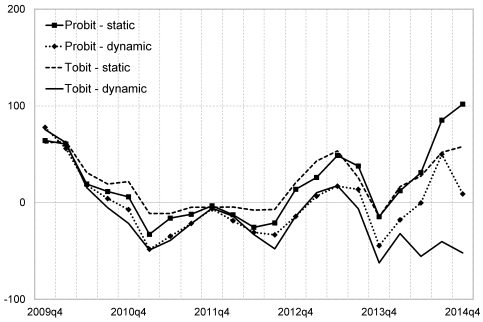
(g) Rating 7



(h) Rating 8



(i) Rating 9



Source: Bank of Slovenia, authors' calculations.

INDEX