

COMPARISON OF CREDIT SCORING MODELS ON PROBABILITY OF DEFAULT ESTIMATION FOR US BANKS

Petr Gurný, Martin Gurný*

Abstract:

This paper is devoted to the estimation of the probability of default (PD) as a crucial parameter in risk management, requests for loans, rating estimation, pricing of credit derivatives and many others key financial fields. Particularly, in this paper we will estimate the PD of US banks by means of the statistical models, generally known as credit scoring models. First, in theoretical part, we will briefly introduce the two main categories of credit scoring models, which will be afterwards used in application part – linear discriminant analysis and regression models (logit and probit), including testing the statistical significance of estimated parameters. In the main part of the paper we will work with the sample of almost three hundred US commercial banks which will be separated into two groups (non-default and default) on the basis of historical information. Subsequently, we will stepwise apply the mentioned above scoring models on this sample to derive several models for estimation of PD. Further we will apply these models to the control sample to determine the most appropriate model.

Keywords: probability of default (PD), credit scoring models, linear discriminant analysis, logistic regression, probit regression.

JEL Classification: C51, G01, G21

1. Introduction

Estimating the borrower's risk level, namely the probability of default (PD), by assigning a different PD to each borrower is now widely employed in many banks. The PD indicates that a given counterparty will not be able or willing to meet its obligations. The false estimation of PD leads to unreasonable rating, incorrect pricing of financial instruments and thereby it was one of the causes of the recent global financial crisis. Probability of default is also a crucial parameter used in the calculation of economic capital or regulatory capital under Basel II for a banking institution.

* VŠB – Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava-Poruba, Czech Republic (petr.gurny@vsb.cz).

The paper is based on research activities sponsored through SGS Research Project No. SP2011/166 and ESF Project No.CZ.1.07/2.3.00/20.0296. The support is greatly acknowledged.

From these reasons it is obvious, that estimation of the probability of default of financial subjects is topical theme in the field of the financial research for a long time. There are a lot of models to estimate the probability of default in financial world. Among the most widely used ones belong models, generally known as credit scoring models. These are multivariate models which use the main financial indicators of a company as input, attributing a weight to each of them that reflects its relative importance in forecasting default. We can group these models into three categories: a) discriminant analysis (linear, quadratic); b) regression models (linear, logit, probit); c) inductive models (neural networks, genetic algorithm). See *e.g.* Resti and Sironi (2007), Green (2008) or Engelmann and Rauhmeier (2006) for more details. We will be concerned with linear discriminant analysis and logit and probit model in more detail in the next part of the paper.

This paper is related to a number of other studies focused on the credit scoring models analysis. Although the techniques underlying these models were devised in the 1930s by authors such as Fischer (1936) and Durand (1941), the decisive boost to the development and spread of these models came in the 1960s, with studies by Beaver (1967) and Altman (1968) who, by comparing matched samples of failed *versus* non-failed firms, show that business failures can be predicted by information contained in financial ratios. Other important contributions in that field are: McFadden (1976), Altman *et al.* (1981), Queen and Roll (1987), Shumway (2001), and Balthazar (2006). Within the Czech Republic, the best known models are so called IN models (see *e.g.* Neumaierová and Neumaier (2002)).

The vast majority of already proposed credit scoring models were derived on a sample of non-financial institutions, mainly due to the fact that defaults of financial institutions occur relatively scarcely and not all the data are publicly available. Nevertheless, there were several more or less sufficient attempts to identify the key factors for healthy financial institutions, originating from financial statements, see *e.g.* Peresetsky and Karminsky (2008) and references therein or Gurný and Gurný (2009 and 2010).

The goal of the paper is to estimate, verify and compare one-period prediction models for determining the probability of default for US commercial banks during the financial crisis period and subsequently apply these models to the control sample to determine the most appropriate model for prediction of default.

We proceed as follows. In the theoretical part of the paper we expand in more detail the three types of credit scoring models, namely linear discriminant analysis, logit model and probit model, including testing the statistical significance of estimated parameters. There will be introduced a sample of almost 300 financial institutions, that are to be divided in healthy ones and default ones, including their financial indicators in the application part of the paper. Further we will estimate three above mentioned models and then apply these models to the control sample. In conclusion we will discuss the results.

2. Credit Scoring Models

In this section we will briefly describe the above mentioned credit scoring models and process of selecting variables. Credit scoring models are the most common tools used to estimate the probability of default of the borrower. They work on the principle of assigning weights to financial and economic indicators. Weights express the significance of these indicators in the estimation of the borrower's default.

2.1 Regression models (*logit and probit*)

Logit and probit regression analysis are the multivariate techniques which allow for estimating the probability that an event occurs or not, by predicting a binary dependent outcome from a set of independent variables. The response, y_i , is equal to 0 if default occurs (with probability P_i) and to 1 if default does not occur (with probability $1 - P_i$). In regression models we wish to model the probability P_i that the default will occur by specifying the following model

$$P_i = f(\alpha + \beta'x_i) \quad (1)$$

where x_i are particular financial indicators and α , β are estimated parameters.

There are a lot of ways of specifying P_i , but in this paper we will focus on the logit and probit transformation, thus logit and probit model.

Logit model

In logit model we use the so called logistic transformation:

$$P_i = \frac{\exp(\alpha + \beta'x_i)}{1 + \exp(\alpha + \beta'x_i)} = \frac{1}{1 + \exp(-\alpha - \beta'x_i)}. \quad (2)$$

Probit model

In case of probit model we use the cumulative distribution function of normal distribution:

$$P_i = \int_{-\infty}^{\alpha + \beta'x_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt. \quad (3)$$

Due to nonlinear features of these models it is necessary to use maximum likelihood method for parameters estimation. Given P_i and assuming that defaults are independent, we can form the logarithm of likelihood function as follows:

$$\ln L = \sum_{i=1}^n y_i \ln P_i + \sum_{i=1}^n (1 - y_i) \ln(1 - P_i). \quad (4)$$

For selection of the variables we can use the stepwise method and for testing of the models significance the log-likelihood ratio test and Wald test are usually used (see any econometric textbook, e.g. Tabachnick and Fidell (2007) or Hair, Anderson, Tatham and Black (2005)).

2.2 Linear discriminant analysis

The purpose of discriminant analysis is to find the so called discriminant function and to classify objects into one of two or more groups based on a set of features that describe the objects. A basic principal is to maximize the difference between two groups, while the differences among particular members of the same group are minimized. Within credit risk models, one group consists of good borrowers (non-defaulted – group A), while the other includes bad ones (already defaulted – group B). The differences are measured by means of the discriminant variable – score z . For a given borrower i , we calculate the score as follows:

$$z_i = \sum_{j=1}^n \gamma_j x_{j,i}, \quad (5)$$

where x denotes a given feature (usually financial indicator, e.g. obtained from financial statements), γ is its coefficient within the estimated model and n is a number of indicators.

Linear discriminant analysis can be used to produce a direct estimate of the probability of default. It can be shown, see Altman (1968) or Resti and Sironi (2007) for more detail, that the company's probability of default is given by:

$$PD = p(B|x_i) = \frac{1}{1 + \frac{1 - \pi_B}{\pi_B} e^{z_i - \alpha}}, \quad (6)$$

where z_i is quantity defined in (5), π_B represents the prior probability of default and $\alpha = 0.5\gamma'(x_A - x_B)$, where x_A and x_B are vectors containing the mean values of the n independent variables for the group of healthy companies and the group of insolvent ones.

The vector of gamma coefficients in (5) is calculated as follows, for more detail see Resti and Sironi (2007):

$$\gamma = \Sigma^{-1}(x_A - x_B), \quad (7)$$

where Σ is the matrix of variances and covariances between the n independent variables.¹

1 Σ can be estimated as the "average" Σ of the matrixes of variances and covariances for each group of companies, weighted by the number companies n_A , n_B present in each group.

3. Data Description

We will work with the sample of 298 American commercial banks for model estimation in this subsection. As a first step we have to divide this sample of banks into two groups, the non-default banks and default banks. The defaulted banks can be defined in a variety of ways. In this paper these banks are thought the financial institutions which have gone into liquidation or undergone financial restructuring processes (e.g. take-over by another bank or by government). The samples of the financial institutions for both of these groups were chosen randomly pursuant to the publicly available information.² As a second step we have to identify the financial indicators from financial statements, see Karminsky and Peresetsky (2008). Since the empirical research within the prediction models estimation suggests that for the best predictive ability of a model is appropriate to record the financial statements one – two years before the day of default (see e.g. the one-year Z-metrics model, Altman (2010)), we calculated the database of the financial indicators in this study from the annual financial statements of a particular bank in the year $t - 2$, where t is the year of default for particular defaulted bank or year of model estimation for the healthy banks.³ This assumption will lead to the estimation of a one-period (1-2 years) prediction models. It's also necessary to note that all the data were collected during the financial crisis during the years 2007 - 2010. Due to the large sensitivity of the models on the input data, it's necessary to understand the limitations of the estimated models and achieved results only on this phase of the market evolution. We have worked a total of sixteen financial indicators that describe the financial health of the subject (indicators describing the size of the subject, profitability, efficiency, assets quality or capital adequacy) in this paper.⁴ Table 1 shows the means of the chosen indicators for both groups of banks.

Table 1

The Mean Values of Indicators for Non-default and Default Banks in $t - 2$

non-default banks				default banks			
<i>fin. indicator</i>	<i>mean value</i>	<i>fin. indicator</i>	<i>mean value</i>	<i>fin. indicator</i>	<i>mean value</i>	<i>fin. indicator</i>	<i>mean value</i>
x_1 : <i>LTA</i>	15.80	x_9 : <i>PE OI</i>	23.21%	x_1 : <i>LTA</i>	11.91	x_9 : <i>PE OI</i>	26.62%
x_2 : <i>YAEA</i>	5.81%	x_{10} : <i>PL GL</i>	3.71%	x_2 : <i>YAEA</i>	6.65%	x_{10} : <i>PL GL</i>	15.15%
x_3 : <i>CIBL</i>	3.30%	x_{11} : <i>LLR GL</i>	1.96%	x_3 : <i>CIBL</i>	3.69%	x_{11} : <i>LLR GL</i>	3.24%
x_4 : <i>NIM</i>	3.56%	x_{12} : <i>PL EQ LLR</i>	27.42%	x_4 : <i>NIM</i>	3.15%	x_{12} : <i>PL EQ LLR</i>	39.09%
x_5 : <i>ROAA</i>	1.14%	x_{13} : <i>T1</i>	10.70%	x_5 : <i>ROAA</i>	-4.31%	x_{13} : <i>T1</i>	7.84%
x_6 : <i>ROAE</i>	7.62%	x_{14} : <i>EQ TA</i>	10.92%	x_6 : <i>ROAE</i>	-47.53%	x_{14} : <i>EQ TA</i>	5.58%
x_7 : <i>IE II</i>	37.87%	x_{15} : <i>CAR</i>	12.60%	x_7 : <i>IE II</i>	55.43%	x_{15} : <i>CAR</i>	8.35%
x_8 : <i>CIR</i>	86.81%	x_{16} : <i>D EQ</i>	7.86	x_8 : <i>CIR</i>	115.54%	x_{16} : <i>D EQ</i>	15.32

2 www.federalreserve.gov, www.failedbankreporter.com

3 It means that the lag for both groups in the sample is between 12 and 24 months (depending on the particular month of default during the year t).

4 For more detail description of the financial indicators see Appendix A.

4. Application and Results

In this part of the paper we will estimate three types of credit scoring models – logit model, probit model and LDA model (model based on linear discriminant analysis), including testing the statistical significance of estimated parameters. Logit and probit model will be estimated using statistical software STATA. After comparison of these estimated models we will apply them to the control sample to determine the most appropriate model.

4.1 Logit model

The estimated logit model is shown in Table 2.⁵

Table 2

The Estimated Logit Model

begin with empty model						
p = 0.0000	<	0.0500	adding	x10_PL_GL		
p = 0.0000	<	0.0500	adding	x1_LTA		
p = 0.0170	<	0.0500	adding	x16_D_EQ		
p = 0.0181	<	0.0500	adding	x12_PL_EQ_LLR		
p = 0.0279	<	0.0500	adding	x6_ROAE		
p = 0.8739	>=	0.1000	removing	x16_D_EQ		
p = 0.5968	>=	0.1000	removing	x12_PL_EQ_LLR		
Logistic regression				Number of obs	=	298
				LR chi2(3)	=	395.95
Log likelihood = -7.6143581				Prob > chi2	=	0.0000
				Pseudo R2	=	0.9630
Stat	us	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
x10_PL_GL		87.69232	35.89412	2.44	0.015	17.34115 158.0435
x1_LTA		-4.477891	1.786542	-2.51	0.012	-7.979449 -.9763325
x6_ROAE		-36.10652	15.96435	-2.26	0.024	-67.39607 -4.816973
_cons		52.08043	21.22815	2.45	0.014	10.47401 93.68685

This model according to (2) looks as follows:

$$P_i = \frac{1}{1 + \exp(-52.08 + 4.48x_{1,i} + 36.11x_{6,i} - 87.69x_{10,i})},$$

where x_1 , x_6 and x_{10} denotes LTA (*Logarithm of total assets*), ROAE (*Return on average equity*) and PL GL (*Problem loans on Gross loans*), respectively.⁶

The likelihood-ratio test

$H_0: \hat{\beta}_j = 0$ where $j = 1, 2, 3$;

5 This table is the outcome of the software STATA[™] 10.0 MP.

6 These indicators were selected on the basis of stepwise method (forward selection followed by backward elimination) with significance levels 5% for adding variable into the model and 10% for removing variable from the model.

$H_A: \hat{\beta}_1 \neq 0 \vee \hat{\beta}_2 \neq 0 \vee \hat{\beta}_3 \neq 0$ (at least one of the parameters is simultaneously nonzero).

This test is a part of the Table 2. Considering $0.0000 < 0.05$, we reject null hypothesis and accept alternative hypothesis which means that at least one of the parameters of the model is different from zero, so the model is statistically significant at the 5% significance level.

The Wald test

Hypotheses are the same as in likelihood-ratio test. The values of Wald test are, in case we test all the parameters simultaneously, shown in Table 3.

Table 3

Wald Test of Logit Model

(1)	x10_PL_GL = 0	
(2)	x1_LTA = 0	
(3)	x6_RCAE = 0	
	chi 2(3) =	6.99
	Pr ob > chi 2 =	0.0421

Considering $0.0421 < 0.05$, we reject null hypothesis and accept alternative hypothesis again which means that at least one of the parameters of the model is different from zero, so the model is statistically significant at the 5% significance level.

The Table 4 shows the estimated mean values of PD for both groups of sample and Figure 1 shows the PD for every particular bank from the sample, where on x-axis is a ordinal number of the particular bank and on y-axis is the probability of default.

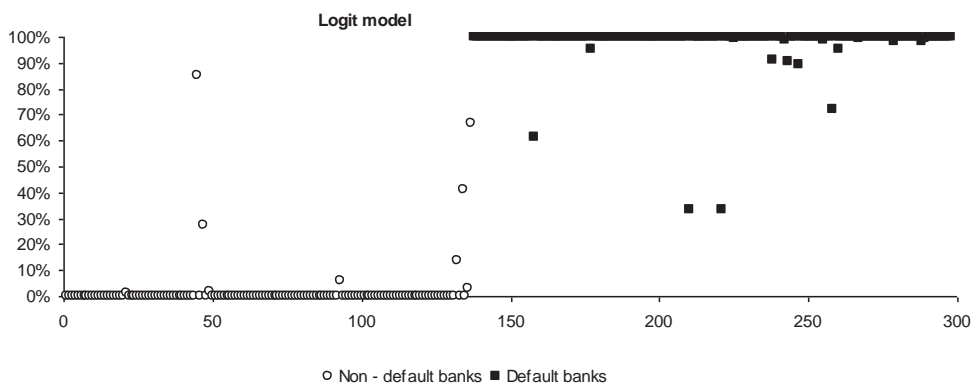
Table 4

The Estimated Mean Values of PDs for Logit Model

non-default banks	mean value	default banks	mean value
<i>PD</i>	1.80%	<i>PD</i>	98.47%

Figure 1

The Estimated PDs for Logit Model



4.2 Probit model

The estimated logit model is shown in Table 5.

Table 5

The Estimated Probit Model

		begin with empty model				
p =	0.0000	<	0.0500	adding	x1_LTA	
p =	0.0000	<	0.0500	adding	x5_ROAA	
p =	0.0001	<	0.0500	adding	x9_PE_OI	
p =	0.0003	<	0.0500	adding	x6_ROAE	
p =	0.3117	>=	0.1000	removing	x5_ROAA	
p =	0.0273	<	0.0500	adding	x14_EQ_TA	
Probit regression				Number of obs	= 298	
				LR chi 2(4)	= 394.11	
				Prob > chi 2	= 0.0000	
				Pseudo R2	= 0.9585	
Log likelihood =		-8.5362047				
Stat us	Coef .	Std. Err .	z	P> z	[95% Conf. Interval]	
x1_LTA	-1.702366	.4842775	-3.52	0.000	-2.651532	-.7531992
x6_ROAE	-18.26809	6.159848	-2.97	0.003	-30.34117	-6.195008
x9_PE_OI	-17.04526	6.168434	-2.76	0.006	-29.13517	-4.955347
x14_EQ_TA	-82.14669	37.22711	-2.21	0.027	-155.1105	-9.182897
_cons	33.75012	9.742617	3.46	0.001	14.65494	52.8453

This model according to (3) looks as follows:

$$f(\alpha + \beta'x_i) = 33.75 - 1.7x_{1,i} - 18.27x_{6,i} - 17.05x_{9,i} - 82.15x_{14,i},$$

$$P_i = \Phi(33.75 - 1.7x_{1,i} - 18.27x_{6,i} - 17.05x_{9,i} - 82.15x_{14,i}),$$

where x_1, x_6, x_9 and x_{14} denotes LTA (*Logarithm of total assets*), ROAE (*Return on average equity*), PE OI (*Personnel expenses on Operating income*) and EQ TA (*Shareholders' equity on Total assets*), respectively.

The likelihood-ratio test

$$H_0: \hat{\beta}_j = 0 \quad \text{where } j = 1, 2, 3, 4;$$

$$H_A: \hat{\beta}_1 \neq 0 \vee \dots \vee \hat{\beta}_4 \neq 0 \quad (\text{at least one of the parameters is simultaneously nonzero}).$$

This test is again a part of the Table 5. Considering $0.0000 < 0.05$, we reject null hypothesis and accept alternative hypothesis which means that at least one of the parameters of the model is different from zero, so the model is statistically significant at the 5% significance level.

The Wald test

Hypotheses are the same again as in likelihood-ratio test. The values of Wald test are, in case we test all the parameters simultaneously, shown in Table 6.

Table 6

Wald Test of Probit Model

(1)	$x1_LTA = 0$	
(2)	$x6_ROAE = 0$	
(3)	$x9_PE_CI = 0$	
(4)	$x14_EQ_TA = 0$	
	chi 2(4) =	12.88
	Prob > chi 2 =	0.0119

Considering $0.0119 < 0.05$, we reject null hypothesis and accept alternative hypothesis again which means that at least one of the parameters of the model is different from zero, so the model is statistically significant at the 5% significance level.

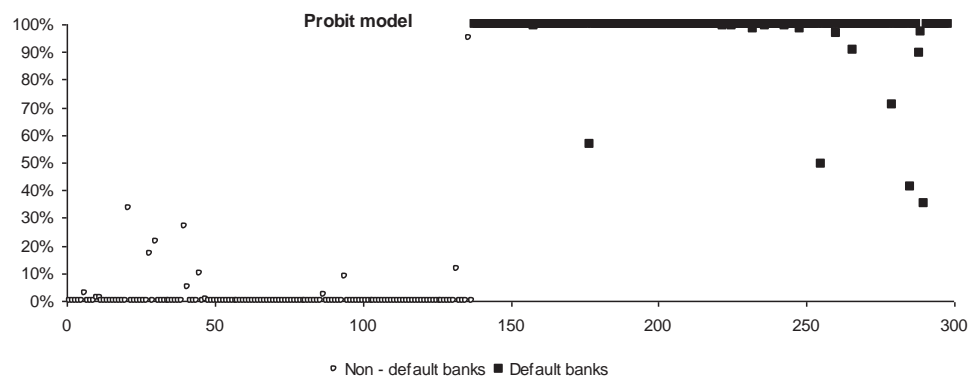
The results are in Table 7 and Figure 2 by analogy.

Table 7:

The Estimated Mean Values of PDs for Probit Model

non-default banks	mean value	default banks	mean value
<i>PD</i>	1.74%	<i>PD</i>	98.27%

Figure 2

The Estimated PDs for Probit Model**4.3 LDA model**

This model is based on the linear discriminant analysis and according to (5) and (7) we get the general discriminant function for estimating z -score:

$$\hat{z}_i = 1.14x_{1,i} + 24.91x_{5,i} - 5.76x_{7,i} - 9.26x_{10,i} + 52.57x_{14,i},$$

where x_1 , x_5 , x_7 , x_{10} and x_{14} denotes LTA (*Logarithm of total assets*), ROAA (*Return on average assets*), IE II (*Interest expense on Interest income*), PL GL (*Problem loans on Gross loans*) and EQ TA (*Shareholders' equity on Total assets*), respectively.

t-test

$$H_0: \hat{\beta}_i = 0 ;$$

$$H_A: \hat{\beta}_i \neq 0 \text{ where } i = 1, 2, 3, 4, 5.$$

Statistics of *t*-test for all parameters are shown in Table 8.

Table 8:

t-test of LDA Model

	beta	SEbeta	t^{calc}	value P
x₁: LTA	1.14	0.3306	3.4487	0.0006
x₅: ROAA	24.91	11.4084	2.1835	0.0298
x₇: IE II	-5.76	2.8446	-2.0249	0.0438
x₁₀: PL GL	-9.26	3.8457	-2.4079	0.0167
x₁₄: EQ TA	52.57	22.2995	2.3575	0.0191

We can see that value P of all parameters is lower than the chosen 5% significance level. Thus, we reject null hypothesis and accept alternative hypothesis which means that all calculated parameters are in the critical area and therefore are statistically significant at the 5% significance level.

F-test

$$H_0: \hat{\beta}_i = 0 \text{ where } i = 1, 2, 3, 4, 5;$$

$$H_A: \hat{\beta}_1 \neq 0 \vee \dots \vee \hat{\beta}_5 \neq 0 \text{ (at least one of the parameters is simultaneously nonzero).}$$

Statistic of *F*-test is shown in Table 9.

Table 9:

F-test of LDA Model

Regression statistic	
R	0.885653
R²	0.784381
SE	17.117242
T	298

ANOVA

	df	SS	MS	F^{calc}	significance F
regression	5	58.057407	11.611481	227.699869	1.13581E-98
residues	293	14.941440	0.050994		
total	298	74.016778			

We can see that significance F is lower than the chosen 5% significance level. Thus, we reject null hypothesis and accept alternative hypothesis which means that estimated model is statistically significant at 5% significance level.

The Table 10 shows the estimated mean values of the z-score and PD for both group of sample⁷, Figure 3 then shows the PD for every particular bank from the sample again.

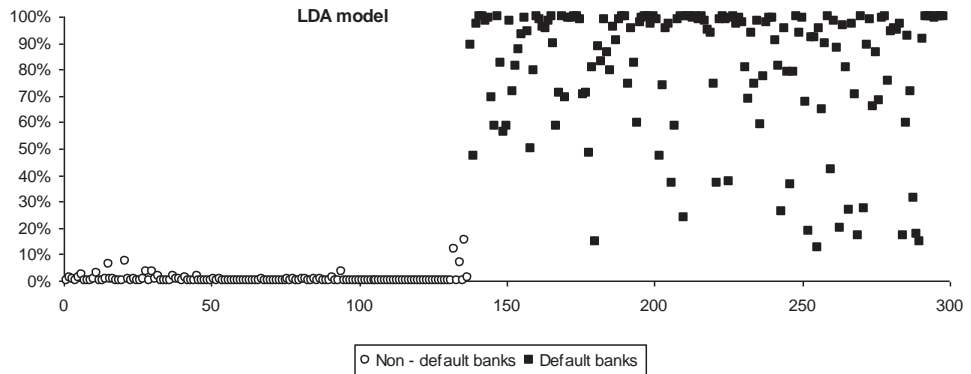
Table 10

The Estimated Mean Values of z-scores and PDs for LDA Model

non-default banks	mean value	default banks	mean value
z	21.50	z	10.83
PD	0.67%	PD	81.65%

Figure 3

The Estimated PDs for LDA Model



4.4 Comparison of the estimated models

Comparison of the estimated models is done in Table 11.

Table 11

Comparison of the Estimated Models

	Logit model		Probit model		LDA model	
number of indicators	3		4		5	
indicators	LTA, ROAE, PL GL		LTA, ROAE, PE OI, EQ TA		LTA, ROAA, IE II, PL GL, EQ TA	
pseudo R² / R²	0.9630		0.9585		0.7844	
E (PD)⁸	1.80%	98.47%	1.74%	98.27%	0.67%	81.65%

⁷ We can directly estimate the PD for every particular bank from the sample according to (11).

⁸ The first value indicates the mean value of PD for non-default banks, the second for default banks.

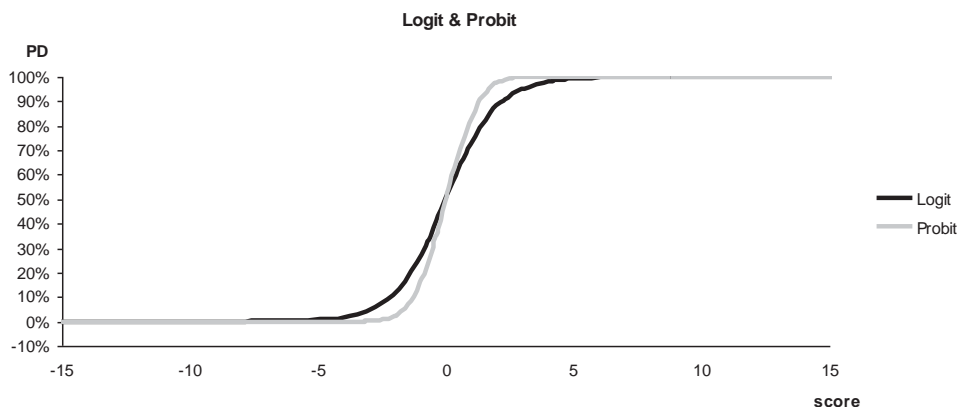
We can see in this table that logit model contains at least indicators (3). Logit and probit model have similar explanatory power (96.30% for logit model and 95.85% for probit model), although probit model has one extra indicator. From mean values of PDs and from graphical results it is clear that both models separate groups of the banks very good (see mean values 1.80% and 98.47% for logit model and 1.74% and 98.27% for probit model, respectively).

Compared to these two model LDA model contains five indicators, but its explanatory power is lower (78.44%). We can say the first two models are better for prediction of default (under the assumptions given above). LDA model predicts default slightly better at non-default banks, but much worse at default banks (mean value of PD is only 81.65%).

We can also use the graphical dependency between z -score, $(\alpha + \beta'x_i)$ and PD to demonstrate the differences between logit and probit model. We can observe the results on the Figure 4.

Figure 4

Comparison of Logit and Probit Model



From this figure we can see that in this case logit and probit models give qualitatively similar results and the main difference between logit and probit model is that logistic has slightly flatter tails. Minor differences are seen also in the values of z -score close to zero.

4.5 Application of the estimated model to the control sample

For verification of the success rate of the estimated models we applied these models on the control sample of 100 American commercial banks. We again randomly chosen 50 non-default and 50 default banks and for the verification of the predicting abilities of

the estimated models we applied them on the data one-period ($t - 2$ annual reports data) before decisive day in year t . The results are presented on the estimated mean values of the PD (Table 12) for every model and on the graphical expressions of the particular PD for every bank from control sample (Figure 5), where first 50 banks are non-defaulted banks.

Table 12

The Estimated Mean Values of PDs of Banks from Control Sample

estimated model	non - default banks	default banks
	$E(PD)$	$E(PD)$
Logit model	14.47%	90.54%
Probit model	26.37%	75.38%
LDA model	7.39%	34.64%

Figure 5

Estimated PD for Control Sample (logit, probit, LDA model, respectively)

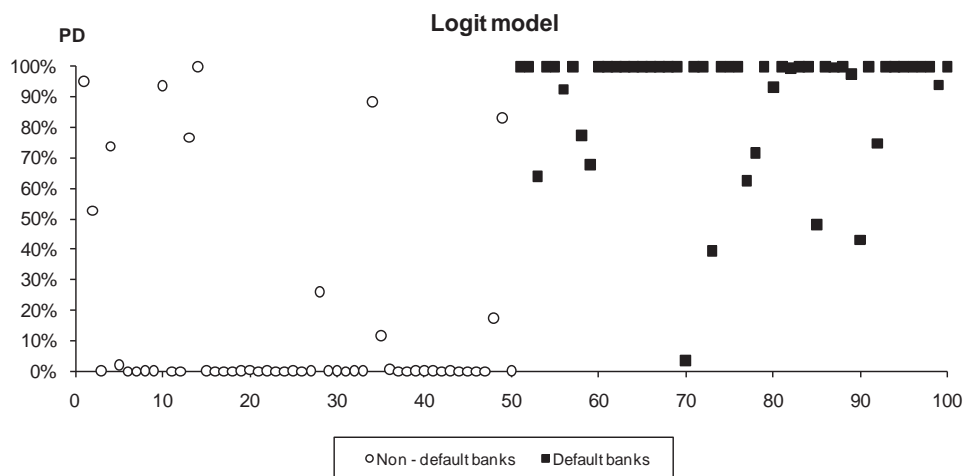
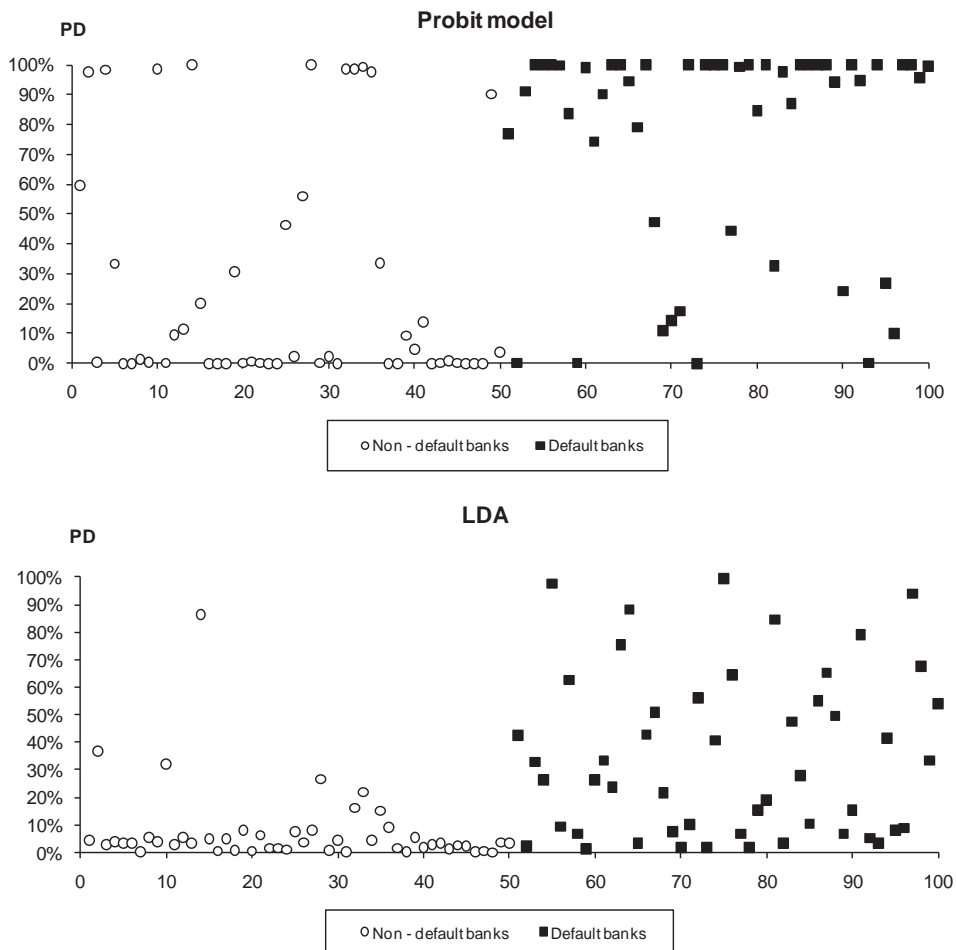


Figure 5 - Continuation



On the next tables (Tables 13-15) we can observe the success rate of the estimated models for various levels of the cut values.

Table 13

Classification Table for Estimated Logit Model

cut value	non-default banks			default banks			overall percentage
	observed	predicted	correct	observed	predicted	correct	
50%	50	42	84.0%	50	46	92.0%	88.0%
30% (70%)	50	42	84.0%	50	43	86.0%	85.0%
10% (90%)	50	39	78.0%	50	40	80.0%	79.0%
5% (95%)	50	39	78.0%	50	37	74.0%	76.0%
1% (99%)	50	38	76.0%	50	36	72.0%	74.0%
						mean	80.4%

Table 14

Classification Table for Estimated Probit Model

cut value	non-default banks			default banks			overall percentage
	observed	predicted	correct	observed	predicted	correct	
50%	50	38	76.0%	50	37	74.0%	75.0%
30% (70%)	50	34	68.0%	50	37	74.0%	71.0%
10% (90%)	50	31	62.0%	50	31	62.0%	62.0%
5% (95%)	50	29	58.0%	50	26	52.0%	55.0%
1% (99%)	50	24	48.0%	50	24	48.0%	48.0%
						mean	62.2%

Table 15

Classification Table for Estimated LDA Model

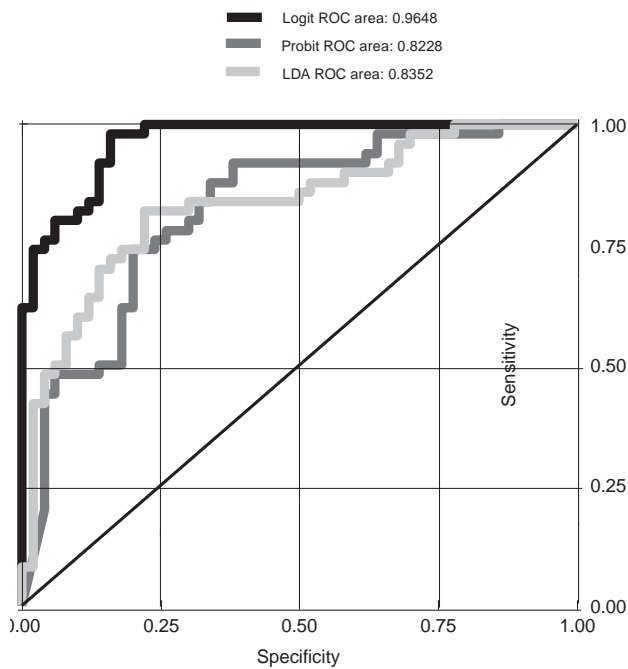
cut value	non-default banks			default banks			overall percentage
	observed	predicted	correct	observed	predicted	correct	
50%	50	49	98.0%	50	15	30.0%	64.0%
30% (70%)	50	47	94.0%	50	7	14.0%	54.0%
10% (90%)	50	43	86.0%	50	3	6.0%	46.0%
5% (95%)	50	35	70.0%	50	2	4.0%	37.0%
1% (99%)	50	11	22.0%	50	1	2.0%	12.0%
						mean	42.6%

The very powerful statistical tool for measure of the model's quality prediction is also ROC analysis, because it does not assume the equality of classification error costs what is the case with bankruptcy prediction models.⁹ ROC analysis for all three models (within application on control sample) is observable in the next Figure 6.

⁹ It means ROC analysis does not assume the equality between Type I error (models classify the bank as non-defaulted while in reality it went default) and Type II error (error of classifying the non-defaulted bank as defaulted).

Figure 6

ROC Analyses for Application of Estimated Models on the Control Sample



In the Figure 6 x -axis represents the relation of default incorrectly classified with total default group - specificity (probability that the models will correctly classify non-defaulted banks), while y -axis represents sensitivity (probability that bankruptcy prediction models will correctly classify defaulted). The more the ROC curve is closer to y -axis, the better the estimated prediction model is. Or in other words, the model has higher discriminant power if its sensitivity and specificity are higher with respect to other model sensitivity and specificity. From Figure 6 it is clear that logit model has the best predictive ability. We can express this also by quantification and statistical verification of the area under the ROC curve, see Table 16.¹⁰

Table 16

Statistical Description of the Area under the Curve for Estimated Models

	Obs	ROC Area	St d. Err.	—Asymptotic Normal— [95% Conf. Interval]
Logit	100	0.9648	0.0147	0.93596 0.99364
Probit	100	0.8228	0.0422	0.74016 0.90544
LDA	100	0.8352	0.0408	0.75530 0.91510

¹⁰ For more detailed description of ROC curve see Hanley and McNeil (1982) or any statistics textbook.

From the results it is clear that **logit model** shows the best results in application to the control sample and so we can say this model is the most appropriate model for prediction of the banks default (there are surprisingly relatively big differences between logit and probit model). On the other hand the LDA model appears as quite inappropriate for the prediction of the banks failure. However, it is still need to keep in mind limitations of the estimated logit model, as we mentioned above (one-period model for utilization primarily in the time of market depression).

5. Conclusion

The main purpose of the estimation of probability of default consists in its usage in the risk management, valuation of the credit derivatives, estimation of the creditworthiness of the borrowers and estimations of the banks' capital adequacy. Incorrect estimation of the PD can lead to the false valuation of the risk and consequently to the financial problems of the particular company.

In the paper we have described the possibility of PD's estimation of US commercial banks by means of the three types of credit scoring models (logit model, probit model and linear discriminant analysis). We have derived the three models for estimation of PD from sample of 298 American commercial banks, including testing the statistical significance of estimated parameters. These models were afterwards applied to the control sample of next 100 American commercial banks to determine the most appropriate model.

It is important to note the assumptions of the estimated models and hence the possibilities and limitations of their utilization. All three prediction models were estimated from the dataset obtained in the time of financial crisis and thus their utilization is also restricted to this particular phase of market evolution. Another limitation comes from the lag between calculation of relevant indicators and the date of banks bankruptcy. From this perspective, the estimated models are one-period prediction models, with a rather short time prediction of default (1-2 years). Taking into account these limitations, from the obtained results we can say that the estimated logit model is the most appropriate model for prediction of banks' default.

Question, what lag should be taken into account when compiling a database of financial indicators, is next key issue in prediction models estimation. Unfortunately, there is a lack of empirical information about a lag between the date of the calculation of financial indicators with the most explanatory power for detection of higher default probability and the date of bankruptcy for the financial institutions. For the reasons mentioned in Section 3 we calculated with approximately 1-2 years lag within models estimation in the paper (and the results seem to be quite acceptable). However, it could be convenient to test also different lag within following research. Another interesting possibility is to try to calculate the values of the financial indicators as a weighted average of their values over the last few years and subsequently to compare the prediction power of the models with results presented in this paper. However, these possibilities again face a lack of input information, particularly for defaulted banks.

Appendix

Indicator	Indicator	Indicator's group
LTA	Logarithm of total assets	Size
YAEA	Interest Income / Average Interest Earning Assets (%)	Profitability
CIBL	Interest Expense / Average Interest Bearing Liabilities (%)	Profitability
NIM	Net Interest Margin	Profitability
ROAA	Return on Average Assets (%)	Profitability
ROAE	Return on Average Equity (%)	Profitability
IE II	Interest Expense / Interest Income (%)	Profitability
CIR	Cost to Income Ratio (%)	Efficiency
PE OI	Personnel Expenses / Operating Income (%)	Efficiency
PL GL	Problem Loans / Gross Loans (%)	Assets Quality
LLR GL	Loan Loss Reserve / Gross Loans (%)	Assets Quality
PL EQ LLR	Problem Loans / (Shareholders' Equity + Loan Loss Reserve) (%)	Assets Quality
T1	Tier 1 ratio (%)	Capital adequacy
EQ TA	Shareholders' Equity / Total Assets (%)	Capital adequacy
CAR	Capital Adequacy (%)	Capital adequacy
D EQ	Total Deposits / Shareholders' Equity	Capital adequacy

References

- Altman, E. I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy." *Journal of Finance*, September, pp. 589-609.
- Altman, E. I. et al. (1981), "Application of Classification Techniques in Business." In *Banking and Finance*. JAI Press, Greenwich.
- Altman, E. I. (2010), *The Z-Metrics Methodology For Estimating Company Credit Ratings And Default Risk Probabilities*. RiskMetrics Group.
- Balthazar, L. (2006), *From Basel 1 to Basel 3: The Integration of State-of-the-Art Risk Modeling in Banking Regulation*. New York: Palgrave Macmillan.
- Beaver, W. (1967), "Financial Ratios as Predictors of Failures. Empirical Research in Accounting: Selected Studies – 1966." In *Supplement to Journal of Accounting Research* 4, pp. 71-111.
- Durand, B. (1941), "Risk Elements in Consumer Installments Financing." Working Paper, NBER.
- Engelmann, B., Rauhmeier, R. (2006), *The Basel II Risk Parameters*. Springer Verlag.
- Fisher, R. (1936), "The Use of Multiple Measurements in Taxonomic Problems." *Annals of Eugenics* 7, pp. 179-188.
- Green, W. (2008), *Econometric Analysis*. 6th ed. Prentice Hall.
- Gurný, P., Gurný, M. (2009), "Estimation of PD of Financial Institutions within Linear Discriminant Analysis." *Mathematical Methods in Economics*.
- Gurný, P., Gurný, M. (2010), "Comparison of the Credit Scoring Models on PD Estimation of US Banks." *Mathematical Methods in Economics*.

- Gurný, P., Gurný, M.** (2010), "Logit and Probit Model within Estimation of US Banks PD." *Proceedings of the 47th Meeting of the Euro Working Group of Financial Modelling*, pp. 73-80.
- Hair, J. F., Anderson, R. E., Tatham, R. L., Black, W. C.** (2005), *Multivariate Data Analysis*. 6th ed. Prentice Hall.
- Hanley, J. A., McNeil, B. J.** (1982), "The Meaning and Use of the Area under a Receiver Operating Characteristics (ROC) Curve." *Radiology*, pp. 561-557.
- Hosmer, D. W.; Lemeshow, S.** (2000), *Applied Logistic Regression*. 2nd ed. New York: Wiley.
- McFadden, D.** (1976), "A Comment on Discriminant Analysis versus Logit Analysis." *Annals of Economic and Social Measurement*, pp. 511–523.
- Neumaierová, I., Neumaier, I.** (2002), *Výkonnost a tržní hodnota firmy*. Praha: Grada Publishing.
- Peretsky, A., Karminsky, A.** (2008), "Models for Moody's Bank Rating." BOFIT Discussion Papers 17.
- Queen, M., Roll, R.** (1987), "Firm Mortality: Using Market Indicators to Predict Survival." *Financial Analysts Journal* 3, pp. 9–26.
- Resti, A., Sironi, A.** (2007), *Risk Management and Shareholders' Value in Banking*. Chichester: Wiley.
- Shumway, T.** (2001), "Forecasting Bankruptcy more Accurately: A Simple Hazard Rate Model." *Journal of Business* 74, pp. 101–124.
- Tabachnick, B. G., Fidell, L. S.** (2007), *Using Multivariate Statistics*. 5th ed. Pearson Education, Inc.