EQUILIBRIUM
Quarterly Journal of Economics and Economic Policy
2017 VOLUME 12 ISSUE 4, December
p-ISSN 1689-765X, e-ISSN 2353-3293
www.economic-policy.pl

ORIGINAL PAPER

Citation: Kovacova, M., & Kliestik, T. (2017). Logit and Probit application for the prediction of bankruptcy in Slovak companies. *Equilibrium. Quarterly Journal of Economics and Economic Policy, 12*(4), 775–791. doi: 10.24136/eq.v12i4.40

Contact to corresponding author: maria.kovacova@fpedas.uniza.sk, University of Zilina, Faculty of Operation and Economics of Transport and Communications, Department of Economics, Univerzitna 1, 010 26 Zilina, Slovakia
Received: 26 January 2017; Revised: 19 May 2017; Accepted: 9 July 2017

Maria Kovacova
University of Zilina, Slovakia

Tomas Kliestik
University of Zilina, Slovakia

Logit and Probit application for the prediction of bankruptcy in Slovak companies

JEL Classification: *G33; C15; C5*

Keywords: *bankruptcy; logit; probit; Slovak companies; financial health*

Abstract
Research background: Prediction of bankruptcy is an issue of interest of various researchers and practitioners since the first study dedicated to this topic was published in 1932. Finding the suitable bankruptcy prediction model is the task for economists and analysts from all over the world. forecasting model using. Despite a large number of various models, which have been created by using different methods with the aim to achieve the best results, it is still challenging to predict bankruptcy risk, as corporations have become more global and more complex.
Purpose of the article: The aim of the presented study is to construct, via an empirical study of relevant literature and application of suitable chosen mathematical statistical methods, models for bankruptcy prediction of Slovak companies and provide the comparison of overall prediction ability of the two developed models.
Methods: The research was conducted on the data set of Slovak corporations covering the period of the year 2015, and two mathematical statistical methods were applied. The methods are logit and probit, which are both symmetric binary choice models, also known as conditional probability models. On the other hand, these methods show some significant differences in process of model formation, as well as in achieved results.

Findings & Value added: Given the fact that mostly discriminant analysis and logistic regression are used for the construction of bankruptcy prediction models, we have focused our attention on the development bankruptcy prediction model in the Slovak Republic via logistic regression and probit. The results of the study suggest that the model based on a logit functions slightly outperforms the classification accuracy of probit model. Differences were obtained also in the detection of the most significant predictors of bankruptcy prediction in these types of models constructed in Slovak companies.

Introduction

Application of bankruptcy prediction models had been widely spread in advanced economies mainly in the western part of the world since the first study in this area carried out by Fitzpatrick (1932, pp. 598–605). Since that time, numerous economists and analysts from all over the world have been trying to find an appropriate company's bankruptcy forecasting model applying different methods with the aim to achieve the best results (Ravi Kumar & Ravi, 2007, pp. 1–28). Later it has become also an issue of growing interest for researchers of capitalist, socialist, and transitional economies as well (Brada, 1993, pp. 82–96). Boratynska (2014, pp. 43–57) emphasizes not only the importance of predicting probability of default of companies, but also the aspects of measurement of costs of corporate bankruptcy.

In Slovakia, bankruptcy issue has come to the attention after the success of Slovak transition in 1995, which initiated an institutional evolution, proving remarkably robust (Schonfelder, 2003, pp. 155–180). During that time few studies dealing with the bankruptcy prediction were published (see: Chrastinova, 1998, pp. 34; Gurcik, 2002, pp. 373–378), but the main attention to this issue aroused after the year 2008, when the global financial crisis appeared (Dixon, 2016, pp. 28–62). Because of the deepening globalization and growing independency across economies, also Slovak companies had to cope with various types of financial difficulties.

Adamko and Svabova (2016, pp.15–20) studied the prediction ability of global Altman´s model on the data set of Slovak companies. Similarly, Delina and Packova (2013, pp. 101–112) validated three selected bankruptcy prediction models: Altman model, Beerman discriminatory function and Index IN05 in condition of Slovakia, and according to gained results they proposed a model for bankruptcy prediction using regression analysis.

On the other hand, Rybarova, *et. al.* (2016, pp. 298–306) applied in their analysis the Altman Z–score bankruptcy model only on the key sector of Slovakia, which is construction industry. Selection of one sector, in this case the Slovak logistic sector, was proposed also by Brozyna, *et. al.* (2016, pp. 93–114). They proposed four bankruptcy prediction models based on discriminant analysis, logit, decision trees and k-nearest neighbours' method and validated prediction power of these models in comparison with Poland logistic sector.

Bankruptcy prediction is focusing not only on companies, but the subject of interest can be city or other municipal entities. Alexy (2015, pp.111–117) highlighted the importance of studying financial health of cities. Furthermore, the modelling of default probability of cities in Slovakia through logit model was identified by Kacer and Alexy (2015, pp. 484–491).

Despite the fact that one can find studies focusing on bankruptcy prediction in Slovakia, there is still a lack of models developed on the basis of the Slovak environment. Similarly, Mihalovic (2016, pp. 101–118) emphasizes the reasons for development of such models and proposed multiple discriminant analysis and logit models for bankruptcy prediction.

Despite a large number of various models, it is still challenging to predict bankruptcy risk as corporations have become more global and more complex. According to the above mentioned, the primary focus in this study is on the creation of bankruptcy prediction models which will be based on two various statistical methods applied on Slovak companies. These methods include both logistic regression as well as probit regression, given the fact that mostly discriminant analysis and logistic regression are used for the construction of bankruptcy prediction models (Spuchlakova & Michalikova-Frajtova, 2016, pp. 2093–2099). Under creation of these models the most significant financial ratios best distinguishing among groups of default and no default companies may be detected. Furthermore, the main objective of this study is to compare the performance of the two proposed bankruptcy prediction models on a sample of selected companies operating in Slovak economic environment. To achieve these efforts, two scientific questions were build:

- Are variables included in the created bankruptcy prediction models statistically significant?
- Are created bankruptcy prediction models statistically significant?

Although in Slovakia some bankruptcy prediction models have been constructed, there is no generally accepted model which can be used not only by researchers, but also by practitioners and analysts to predict financial health of the Slovak Republic. So the aim of this study is to find out and propose such bankruptcy prediction models which will set a basis for different groups of users and will be generally accepted as delivering high prediction accuracy.

Due to the above mentioned reasons, the composition of the article is the following: the introduction part, stressing the significance of bankruptcy prediction according to provided literature review, followed by the methodology part, describing the data set and research methodology used. The next part displays result of provided research resulting in discussion part and conclusion of the presented study.

Research methodology

Methodology part of the study describes theoretical basis of models employed, data uses, sample design and variable selection procedure. To construct bankruptcy prediction model in this study, two mathematical statistical methods were used, namely logistic regression and probit. In spite of the fact that these methods are both symmetric binary choice models, they show some significant differences in the process of model formation, as well as in achieved results.

The data for the study were obtained from annual financial reports of Slovak companies (Register of financial statements, Ministry of Finance of the Slovak Republic) covering the year 2015. Firstly, there is a need to stress terminological differences between bankruptcy and insolvency. (Boratynska, 2016, pp. 107–129) Currently, the Slovak legal system considers company as default according to three criterions:

- the total amount of payable and not payable liabilities is higher that the value of company´s assets,
- company has at least two liabilities 30 days after due date from different creditors,
- the value of financial independence indicator is less than 0.04.

Additionally to those criteria, we have detected other relevant characteristics which are considered significant according to the Slovak environment. (see Svabova & Kral, 2016, pp. 1759–1768; Svabova & Durica, 2016, pp. 2–11) Considering these specifications, we have specified three criteria for the subsequent classification of the company as default or no default. Thus, the company is included in the default group of sample if it satisfies these conditions:

- negative value of earnings after taxes,
- the value of current ratio indicator is less than 1.
- the value of financial independence indicator is less than 0.04.

So the final sampling was done by applying the above mentioned criteria, three criteria given by the Slovak legal system and three criteria given by the specifics of the Slovak environment. Furthermore, the application of those criteria on the results of financial analysis of set of companies and removal of detected outliers led to the designation of basic data set from which data of companies serving as inputs for models construction were chosen (Table 1).

The final sample consisted of 500 default and 500 no default companies following the suggestion of Agrawal and Maheshwari (2016, pp. 268–284). The selection was done randomly from basic data set, while no specifics, such as industry in which companies are doing their business, size or the legal form of the companies were not taken into considerations.

For the purpose of this study, the procedure of variable selection includes variables significant in previous studies (Kliestik & Majerova, 2015, pp. 537–543; Zvarikova *et. al.*, 2017, pp. 145–157). According to this criterion, the initial set of variables is drawn from 14 explanatory variables ($x_1 \ldots x_{14}$) in 4 categories (see Table 2.), which served as a basis for construction of bankruptcy prediction models.

Based on given specifications logistic regression and probit were applied to classify the observation (company) into one of the predetermined group. In this type of models, the dependent variable y may obtain only two values. In this study y is a dummy variable representing the occurrence of an event (default of the company or no) expressed by value 0 (no default) and 1 (default). The goal is to quantify the relationship between the individual characteristics (explanatory variables) and the probability of default.

Fundamentals of logistic regression were applied according to Meloun and Militky (2012). The procedure is given by the logit transformation of dependent variable resulting in obtaining the probability of the default of the company P_1 towards the probability of no default of the company $P_0=1-P_1$ through the probability ratio P_1/P_0, where P_1 is computed by the cumulative logistic function:

$$P_1 = \frac{1}{1+e^{-(Z)}} = \frac{e^Z}{e^Z+1} \tag{1}$$

where

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \tag{2}$$

Following Hebak, *et. al.* (2015, pp. 877) the logit can be defined as:

$$\log it(P_1) = \ln\left(\frac{P_1}{1-P_1}\right) = f(x,\beta) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \tag{3}$$

where β are values of coefficients $\beta_0, \beta_1, \beta_2, \ldots \beta_n$ estimated from the data set of companies by maximizing the log-likelihood function. At the centre of the logistic regression is the task estimating the odds ratio $\frac{P_1}{1-P_1}$ and *ln* of this relationship indicates logit transformation. Additionally, based on assumed probability, the company is classified as default or no default, using a cut-off score (usually 0.5), attempting to minimize the type I and type II errors. The type I error arises when the default company is classified as no default, and the type II error arises when the no default company is classified as default.

After a logistic regression model has been fitted, a global test of goodness of fit of the resulting model should be performed (Archer & Lemeshow, 2006, pp. 97–105). To answer the question "How well does my model fit the data?" is widely used the Hosmer-Lemeshow (HL) test for logistic regression (Hosmer, *et. al.*, 1997, pp. 965–980). According to the given p-value of this test (higher better), we suggest to reject or accept the model. According to Hu *et. al.* (2006, pp. 1383–1395) various *R square* statistics have been proposed for logistic regression to quantify the extent to which the binary response can be predicted by a given logistic regression model and covariates. The Nagelkerke's R Square, Cox & Snell R Square and -2 Log likelihood can provide assessing the goodness of fit of the logistic regression model. These statistics show the power of explanation of the model. Cox & Snell R Square is the ratio of the likelihoods reflecting the improvement of the full model over the intercept model (the smaller the ratio, the greater the improvement). Furthermore, Nagelkerke's R Square adjusts Cox & Snell's so that the range of possible values is in interval $\langle 0,1 \rangle$ while considering smaller as greater.

The probability of the observed results given the parameter estimates is known as the Likelihood. Since the likelihood is a small number less than 1, it is customary to use -2 times the log likelihood (-2LL) as an estimate of how well the model fits the data. A good model is one that results in a high likelihood of the observed results.

Significance of explanatory variables and appropriate coefficients is provided by Wald test (see Bewick, *et. al.*, 2005, pp. 112–118), which tests the null hypothesis that the constant equals 0. This hypothesis is rejected if the p-value is smaller than the critical p-value of .05. Hence, we conclude that the constant is not 0. Logit models are often compared to probit models. Probit regression is a specialized regression model of binomial response variables and is also used to analyse the relationship between dependent and explanatory variables. Although these methods are similar in their application, the process of model creation differs. Supposing that a binary dependent variable, *y*, takes only values 0 and 1 (same as in logit), the probit model is given by:

$$P_1 = 1 - \Phi(-x, \beta) = \Phi(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n) \tag{4}$$

where Φ is the cumulative distribution function of the standard normal distribution:

$$\Phi(x, \beta) = \int_{-\infty}^{x,\beta} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \tag{5}$$

Andrews and Hosmer-Lemeshow tests provide evaluation of goodness-of-fit of the proposed model. Additionally, there are several likelihood-based statistics. Along with Log likelihood, Avg. log likelihood and Restr. log likelihood is recommended to assess according to McFadden R-squared, which is the likelihood ratio index, and it is an analogy to the R-squared reported in linear regression models. The discriminant ability of logistic regression model, as well as probit model, can be designed by ROC Curve (Received Operation Characteristic Curve). The ROC curve is a graphical technique allowing for visual analyses of the trade-offs between the sensitivity and the specificity of a test with regard to the various cut-offs that may be used. (see Fawcett, 2006, pp. 861–874) The curve is obtained by calculating the sensitivity and specificity of the test at every possible cut-off point, and plotting sensitivity (the proportion of true positive results) against 1-specificity (the proportion of false positive results). The curve may be used to select optimal cut-off values for a test result, to assess the diagnostic accuracy of a test, and to compare the usefulness of different tests.

Results

During the research process presented in this study, two models were constructed. One was developed on logistic regression and another one through probit regression. Firstly, we assessed our results separately for each model. According to the provided backward stepwise conditional method of logistic regression, logit function coefficients variables were estimated. (see Table 3)

The significance of individual explanatory variable on dependent variable is performed by Wald´s test statistic and given that, the final logit function involves eight variables and constant, which are statistically significant. The resulting logit function providing the probability of default of the company is:

$$P_1 = \frac{1}{1+e^{-(248{,}882-57{,}992*R3+15{,}337*L3-15{,}250*L4+2{,}686*Z1-260{,}981*Z2-7{,}6*Z3-316{,}811*Z5+11{,}137*A2)}} \tag{6}$$

Hosmer-Lemeshow tests signalize good conformity of the final model with given data. The P-value is according to Table 4 0.181, which is consistent with findings of Karan, *et. al*. (2009, pp. 9–26).

Following the suggestions of Menard (2000, pp. 17–24), the overall explanatory power of estimated model is provided in Table 5. Assessing through Nagelkerke´s R Square statistics the model explains 93,8% variability of binary dependent variable. This is confirmed also by the relatively high value of -2 Log likelihood statistics providing the residual deviance of the model with value 169.365.

In addition to logistic regression, the probit regression model was estimated to compare gained results. (see Table 6) In contrast with logit model, final probit models includes all 14 explanatory variables. Furthermore, variables R1, R2, L1, L2, Z3, Z4 and A1 are not statistically significant according to the p-value of z-statistics. However, developed probit model is statistically significant according to the value of McFadden R-squared statistic 87.59% indicating a good fit of the model. (following Hwang, *et. al.*, 2010, pp. 120–137)

Given that, the resulting probit function take the following form:

$$\begin{aligned} P_1 = \Phi 131.8074 - 10.09076*R1 + 7.299365*R2 - 23.84973*R3 + \\ + 0.233393*L1 - 0.316407*L2 + 7.095480*L3 - 6.464321*L4 + \\ + 1.892825*Z1 - 138.0107*Z2 - 3.126644*Z3 + 0.14*Z4 - \\ - 168.0274*Z5 + 0.879841*A1 + 4.962289*A2 \end{aligned} \tag{7}$$

The overall characteristics of probit model is similarly to logit evaluated by Hosmer-Lemeshow test supplemented by Andrews test proving the overall significance of estimated probit function. (see Table 7)

Discussion

In order to assess the overall performance of constructed models (logit and probit models), classification accuracy matrix and ROC curve were provided. There is a need to highlight the fact that overall classification accuracy of proposed models is assessed on the sample of testing data proved by a data sample of training data. The raining data sample, equal to the training data sample, consists of 500 default and 500 no default companies. Table 8 summarizes all classification results of two estimated models proving results of Jones *et. al.* (2015, pp. 72–85) that classification accuracy of logit and probit function is quite similar. In the case of dataset consisting of Slovak companies, the overall prediction accuracy is high (logit 97% and probit 97.3%), which was confirmed by testing the prediction ability of these models resulting in more than 86.5% accuracy of both constructed models.

Comparing gained results with prediction accuracy of other models constructed in condition of Slovakia, it can be summarized that the accuracy of logit and probit models overdo prediction ability of multiple discriminant analysis (approximately 62%) and logistic regression (approximately 73%) provided by Mihalovic (2016, pp. 101–118). On the other hand, he suggested the use of other relevant mathematical statistical prediction techniques including artificial intelligence expert system. Furthermore, this is proved by Mendelova and Bielikova (2017, pp. 26–44) applying DEA analysis on the set of Slovak companies. The prediction accuracy of their model was lower (78,5%) than the prediction accuracy of models designed by us. The need for development of relevant bankruptcy prediction models based on the environment of Slovakia is proved by Delina and Packova (2013, pp. 101–112). Considering national environment and specific of individual economy is highlighted also by Szetela *et. al.* (2016, pp. 839–856) as well as Antonowicz (2014, pp. 35–45). Additionally, ROC curves providing graphic illustration of trade-offs between the sensitivity and the specificity of the classification table providing prediction accuracy of proposed models were constructed. Graphical presentation of four ROC curves constructed for each data set (training and test) of both models (logit and probit) are shown in Figure 1.

According to the graphic illustration, it is clear that the area under the ROC curve is higher for test data than for training data representing a metric for classification accuracy for various cut-off points. The following Table 9 provides the evidence of these results. According to obtained results in the case of logit model applied on test data set, there is 86.7% probability of correct classification and probit model presents 86.6% probability.

In spite of that numerous bankruptcy prediction models have been created worldwide the originality and novelty of proposed models lie in combination of popular statistical methods while taking into account specific conditions of Slovak environment. Given the high prediction accuracy of proposed models, they have a potential to become generally accepted in the Slovak Republic.

Conclusions

Although the issue of bankruptcy prediction is widely spread worldwide, up till now there has been no generally accepted bankruptcy prediction model considering the specifics of Slovak national environment and economics. Therefore, the goal of the presented study was to construct models for bankruptcy prediction of Slovak companies. Thus, two prediction models based on logit regression and probit regression were projected to fill this gap. The proposed bankruptcy prediction models were developed using a data set of Slovak companies covering the period of the year 2015 and models have been evaluated by their classification accuracy and Receiver Operating Characteristic curves. The selection of input variables resulted in collection of the most relevant explanatory variables following by detection of outliers for starting the model creation.

According to provided logistic regression, one rentability, two liquidity, four debt and capital structure and one activity variables are statistically significant providing the best distribution between the group of default and no default companies. Additionally, the final model is also statistically significant providing high classification accuracy, 97.0% for training data and 86.7% for test data. In the case of probit model, we were aiming to study if there were any relevant differences in the obtained results between models, since those methods are both symmetric binary choice models. The results did not prove any significant dissimilarities as probit model obtained 97.3% prediction accuracy for training data and 86.6% prediction accuracy for test data.

Although the probit model is statistically significant, it included variables which are not all significant, excluding two rentability, two liquidity, two debt and capital structure, and one activity variable. In summary, given the fact that mostly discriminant analysis and logistic regression are used for the construction of bankruptcy prediction models, this study aims to overcome these standards. The proposed models can serve as a basis for further research, due to their quite high accuracy prediction.

On the other side, the results could differ based on the provided data set. In addition, it can be assumed that the proposed models should be tested in following years to find out possibilities for construction of the overall bankruptcy prediction model generally accepted in the condition of Slovakia.

References

Adamko, P., & Svabova, L. (2016). Prediction of the risk of bankruptcy of Slovak companies. In *Proceedings of 8th international scientific conference managing and modelling of financial risks.* Ostrava.

Alexy, M. (2015). Financial health of cities in Slovakia. In *Proceedings of 17th International scientific conference Finance and Risk*, Bratislava.

Agrawal, K., & Maheshwari, Y. (2016). Predicting financial distress: revisiting the option-based model. *South Asian Journal of Global Business Research*, *5*(2), doi: 10.1108/SAJGBR-04-2015-0030

Antonowicz, P. (2014). The multi-dimensional structural analysis of bankruptcy of enterprises in Poland in 2013 – results of empirical studies. *Journal of International Studies*, *7*(1) doi: 10.14254/2071-8330.2014/7-1/3.

Archer, K. J., & Lemeshow, S. (2006). Goodness-of-fit test for a logistic regression model fitted using survey sample data. *Stata Journal*, *6*(1).

Bewick, V., Cheek, L., & Ball, J. (2005). Statistics review 14: logistic regression. *Critical Care*, *9*(1).

Boratyńska, K. (2014). The theoretical aspects of measuring the costs of corporate bankruptcy. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *9*(3). doi: 10.12775/EQUIL.2014.017.

Boratyńska, K. (2016). Corporate bankruptcy and survival on the market: lessons from evolutionary economics. *Oeconomia Copernicana*, *7*(1). doi: 12775/OeC.2016.008.

Brada, J. C. (1993). The comparative economics of bankruptcy – dealing with loss-making firms in capitalist, socialist, and transitional economies. *Eastern European Economics*, *31*(4).

Brozyna, J., Mentel, G., & Pisula, T. (2016). Statistical methods of the bankruptcy prediction in the logistics sector in Poland and Slovakia. *Transformations in Business & Economics*, *15*(1).

Delina, R., & Packova, M. (2013). Prediction bankruptcy models validation in Slovak business environment. *E & M Ekonomie a management*, *16*(3).

Dixon, Ch. (2016). Why the global financial crisis had so little impact on the banking systems of emergent East Asia. *Journal of Self-Governance and Management Economics*, *4*(2).

Chrastinova, Z. (1998). *Metódy hodnotenia ekonomickej bonity a predikcie finančnej situácie poľnohospodárskych podnikov*. Bratislava: VÚEPP. 34.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8). doi: 10.1016/j.patrec.2005.10.010.

Fitzpatrick, P. (1932). A comparison of ratios of successful industrial enterprises with those of failed firms. *Certified Public Accountant*, *2*.

Gurcik, L. (2002). G-index – the financial situation prognosis method of agricultural enterprises. *Agricultural Economics*, *48*(8).

Hebak, J. (2015). *Statisticke mysleni a nastroje analyzy dat*. Prague: Informatorium.

Hosmer, D. W., Hosmer, T., LeCessie, S., & Lemeshow, S. (1997). A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in Medicine, 16*(9). doi: 10.1002/(SICI)1097-0258(19970515)16:9%3C965::AID-SIM509%3E3.0.CO;2-O.

Hu, B., Palta, M., & Shao, J. (2006). Properties of R-2 statistics for logistic regression. *Statistics in Medicine*, *25*(8).

Hwang, R. C., Chung, H. M., & Chu, C. K. (2010). Predicting issuer credit ratings using a semiparametric method. *Journal of Empirical Finance*, *17*(1). doi: 10.1016/j.jempfin.2009.07.007.

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

Kacer, M., & Alexy, M. (2015). Modelling of the default of cities in Slovakia. In *Proceedings of 10th international scientific conference on financial management of firms and financial institutions*. Ostrava.

Karan, M. B., Arslan, O., & Alatli, M. (2009). Detection of factors leading to business failure for petroleum filling stations in Turkey. *Iktisat Isletme ve Finans*, *24*.

Kliestik, T., & Majerova, J. (2015). Selected issues of selection of significant variables in the prediction models. In *Proceedings of 10th international scientific conference financial management of firms and financial institutions.* Ostrava.

Meloun, M., & Militky, J. (2012). *Interaktivni statisticka analyza dat*. Prague: Karolinum.

Menard, S. (2000). Coefficients of determination for multiple logistic regression analysis. *American Statistician*, *54*(1). doi: 10.2307/ 2685605.

Mendelova, V., & Bielikova, T. (2017). Diagnosing of the corporate financial health using DEA: an application to companies in the Slovak Republic. *Politicka Ekonomie*, *65*(1).

Mihalovic, M. (2016). Performance comparison of multiple discriminant analysis and Logit models in bankruptcy prediction. *Economics and Sociology*, *9*(4). doi: 10.14254/2071-789X.2016/9-4/6.

Ravi Kumar, P., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review European. *Journal of Operational Research*, *180*(1).

Rowland, Z., Dvorakova L., & Rousek P. (2016). Determination of bankruptcy model of transport companies in the Czech Republic by using multiple discriminant analysis. *Ekonomicko-manazerske spektrum*, *10*(1).

Rybarova, D., Braunova, M., & Jantosova, L. (2016). Analysis of the construction industry in the Slovak Republic by bankruptcy model. *Procedia Social and Behavioral Sciences, 230*.

Schonfelder, B. (2003). Debt collection and bankruptcies in Slovakia: a study of institutional development. *Post-Communist Economies*, *15*(2).

Spuchlakova, E., & Michalikova-Frajtova, K. (2016). The selected global bankruptcy model. In *Proceedings of international scientific conference globalization and its socio-economic consequences 2016*. Rajecke Teplice.

Svabova, L., & Durica, M. (2016). Korelačná analýza prediktorov použitých v bankrotných predikčných modeloch na Slovensku. *Ekonomicko-manazerske spektrum*, *10*(1).

Svabova, L., & Kral, P. (2016). Selection of predictors in bankruptcy prediction models for Slovak companies. In *Proceedings of 10th international days of statistics and economics.* Prague.

Szetela, B., Mentel, G., & Brozyna, J. (2016). In search of insolvency among European countries. *Economic Research – Ekonomska Istrazivanja*, *29*(1). doi: 10.1080/1331677X.2016.1237301.

Zvarikova, K., Spuchlakova, E., & Sopkova, G. (2017). International comparison of the relevant variables in the chosen bankruptcy models used in the risk management. *Oeconomia Copernicana*, *8*(1). doi: 10.24136/oc.v8i1.10.

Acknowledgements

This work was supported by the Slovak Research and Development Agency under Grant number APVV-14-0841: Comprehensive Prediction Model of the Financial Health of Slovak Companies.

Annex

Table 1. Data set for models construction

	No default	Default	Total	% of Default
Basic data	7867	1342	9209	17,06%
Training data	500	500	1000	50%
Test data	500	500	500	50%

Table 2. Set of variables for models construction

Label	Category	Name	Label	Category	Name
R1	Rentability	Net return on assets	**Z1**	Debt and capital structure	Retained Earnings to Total Assets ratio
R2		Gross return on assets	**Z2**		Total debt to Total Assets ratio
R3		Net return on total income	**Z3**		Current debt to Total Assets ratio
L1	Liquidity	Cash ratio	**Z4**		Loan to assets ratio
L2		Quick ratio	**Z5**		Equity to assets ratio
L3		Current ratio	**A1**	Activity	Assets to Total incomes ratio
L4		Net working capital ratio	**A2**		Current Assets to Total incomes ratio

Table 3. Estimated logit function coefficients

Variables in the Equation							
		B	**S.E.**	**Wald**	**df**	**Sig.**	**Exp(B)**
Step 7[a]	**R3**	-57.992	7.636	57.684	1	.000	.000
	L3	15.337	5.371	8.153	1	.004	4578905.876
	L4	-15.250	5.575	7.482	1	.006	.000
	Z1	2.686	1.178	5.195	1	.023	14.667
	Z2	-260.981	39.061	44.641	1	.000	.000
	Z3	-7.600	2.954	6.618	1	.010	.001
	Z5	-316.811	45.974	47.488	1	.000	.000
	A2	11.137	1.750	40.520	1	.000	68681.681
	Constant	248.882	38.190	42.470	1	.000	1.22426373044781E+108

a. Variable(s) entered on step 1: R1, R2, R3, L1, L2, L3, L4, Z1, Z2, Z3, Z4, Z5, A1, A2.

Table 4. Hosmer-Lemeshow test

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
7	11.377	8	.181

Table 5. Logistic regression model summary

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
7	169.365[a]	.704	.938

a. Estimation terminated at iteration number 12 because parameter estimates changed by less than ,001.

Table 6. Estimated probit function coefficients

Dependent Variable: Neprosperuje
Method: ML- Binary Probit (Newton-Raphson / Marquardt steps)
Sample: 1 1000
Included observations: 1000
Convergence achieved after 12 iterations
Coefficient covariance computed using observed Hessian

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	131.8074	19.42462	6.785586	0.0000
R1	-10.09076	7.258377	-1.390223	0.1645
R2	7.299365	7.107661	1.026971	0.3044
R3	-23.84973	6.614007	-3.605943	0.0003
L1	0.233393	0.567420	0.411324	0.6808
L2	-0.316407	0.550507	-0.574755	0.5655
L3	7.095480	2.939340	2.413971	0.0158
L4	-6.464321	3.148683	-2.053024	0.0401
Z1	1.892825	0.864191	2.190284	0.0285
Z2	-138.0107	19.75534	-6.985995	0.0000
Z3	-3.126644	1.782101	-1.754471	0.0793
Z4	0.140000	0.682199	0.205219	0.8374
Z5	-168.0274	23.11148	-7.270300	0.0000
A1	0.879841	0.900051	0.977545	0.3283
A2	4.962289	1.284828	3.862219	0.0001

McFadden R-squared	0.875985	Mean dependent var	0.500000
S.D. dependent var	0.500250	S.E. of regression	0.154637
Akaike info criterion	0.201922	Sum squared resid	23.55405
Schwarz criterion	0.275538	Log likelihood	-85.96081
Hannah-Quinn criter.	0.229901	Deviance	171.9216
Restr. Deviance	1386.294	Restr. Log likelihood	-693.1472
LR statistic	1214.373	Avg. Log likelihood	-0.085961
Prob(LR statictic)	0.000000		
Obs with Dep=0	500	Total obs	1000
Obs with Dep=1	500		

Table 7. Goodness-of-Fit Evaluation

Goodness-of-Fit Evaluation for Binary Specification
Andrews and Hosmer-Lemeshow Tests
Equation: UNTITLED
Grouping based upon predicted risk (randomize ties)

	Quantile of Risk		Dep=0		Dep=1		Total	H-L
	Low	**High**	**Actual**	**Expect**	**Actual**	**Expect**	**Obs**	**Value**
1	0.0000	0.0000	100	100.000	0	0.00000	100	NA
2	0.0000	2.E-80	100	100.000	0	3.3E-80	100	3.3E-80
3	3.E-77	1.E-17	100	100.000	0	4.6E-17	100	4.6E-17
4	1.E-17	0.0004	100	99.9975	0	0.00255	100	0.00255
5	0.0005	0.5459	87	81.4053	13	18.5947	100	2.06780
6	0.5574	0.9441	8	18.7824	92	81.2176	100	7.62134
7	0.9447	0.9970	4	1.66358	96	98.3364	100	3.33689
8	0.9970	1.0000	1	0.06702	99	99.9330	100	12.9958
9	1.0000	1.0000	0	0.00011	100	99.9999	100	0.00011
10	1.0000	1.0000	0	7.5E-10	100	100.000	100	7.5E-10
	Total		500	501.916	500	498.084	1000	NA

Andrew Statistic	74.8493	Prob. Chi-Sq(10)	0.0000

Table 8. Classification results of logit and probit estimated models

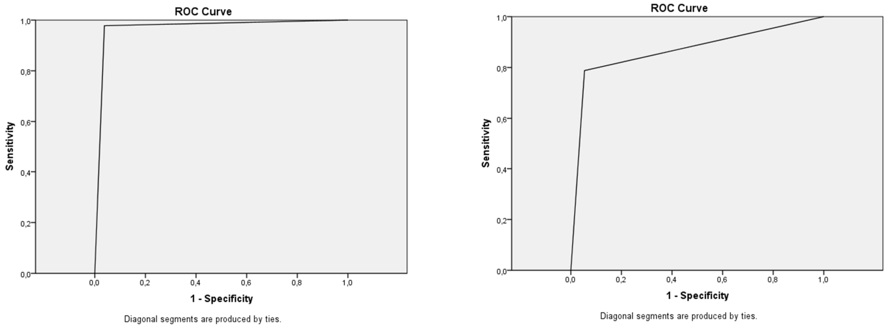
Classification Results (Logistic regression)				
		Predicted (default)		**Percentage correct**
	Observed	**0 (no default)**	**1(default)**	
Training data	**0 (no default)**	481	19	96.2
	1(default)	11	489	97.8
Overall Percentage				97.0
Test data	**0 (no default)**	473	27	94.6
	1(default)	106	394	78.8
Overall Percentage				86.7
Classification Results (Probit regression)				
		Predicted (default)		**Percentage correct**
	Observed	**0 (no default)**	**1(default)**	
Training data	**0 (no default)**	484	16	96.8
	1(default)	11	489	97.8
Overall Percentage				97.3
Test data	**0 (no default)**	474	26	94.8
	1(default)	108	392	78.4
Overall Percentage				86.6

Table 9. Classification results of logit and probit estimated models

	AUC	Sensitivity	False negative rate	Specificity	False positive rate
Logit training	.970	.9776	.0224	.9626	.0374
Logit test	.867	.8169	.1831	.9359	.0641
Probit training	.973	.9778	.0222	.9683	.0317
Probit test	.866	.8144	.1856	.9378	.0622

Figure 1. ROC curves for estimated models

a) Logit model training and test



b) Probit model training and test.