 doi:10.20944/preprints201805.0039.v1

Article

Models of forecasting in financial analysis of non-financial corporations

Eva Litavcová [1,] *, Sylvia Jenčová [2] and Róbert Štefko [3,]

[1] Department of Mathematical Methods and Managerial Informatics, Faculty of Management, The University of Presov; Konstantinova 16, 08001 Presov, Slovakia; eva.litavcova@unipo.sk

[2] Department of Finance, Faculty of Management, The University of Presov; Konstantinova 16, 08001 Presov, Slovakia; sylvia.jencova@unipo.sk

[3] Department of Marketing and International Trade, Faculty of Management, The University of Presov; Konstantinova 16, 08001 Presov, Slovakia; Robert.stefko@unipo.sk

* Correspondence: eva.litavcova@unipo.sk; Tel.: +421-948-208-033

Abstract: Corporate Diagnosis is now recognized as an important tool by decision makers to predict and correct burgeoning problems that a corporation may face. Methods based on this model stem from the use of mathematics and are increasingly being applied in the analysis of production processes. The goal of this paper is to use a logistic regression to design a scoring model for non-financial corporations in industry. Based on the data obtained from the Registry of the Slovak Republic for 738 non-financial corporations, according to SK NACE 26, SK NACE 27, the proportional financial metrics, using the logistic regression method, were calculated. By applying these methods, two logistic regression models were found to reliably estimate the probability of bankruptcy for a firm.

Keywords: logit analysis; company decline; model

---

1. Introduction

The prediction of future developments based on a company's financial assessment requires a broad range of mathematical and statistical methods. Many comprehensive evaluation methods of companies excel with regards to transparency but suffer due to their inaccuracy. Therefore, it is necessary for financial analysts to use various financial forecasting methods to assess the financial health of a company at the same time. For example, Kubíčková and Jindřichovská [1], Neumaierová and Neimaier [2, 3], Kameníková [4], Pilch [5], Bondareva [6], Kadarová, Turisová [7], Vochozka [8] and Kabát et al. [9] deal with the problem of forecasting of financial health of companies in their works. The first author of the bankruptcy model based on logistic regression was Ohlson [10]. The application of this model through logistic regression is mentioned by Slavíček and Kuběnka in [11] for the construction sector in the Czech Republic, Marcinkevičius and Kanapickiené [12] for the construction sector of Lithuania, Harumová, and Janisová [13] for small and medium enterprises. These analyses were based on data from enterprises studied by Presov and Kosice regions of Slovak Republic (NUTS III classification), Valecký and Slivková [14], with the scoring model for Czech enterprises, next Jakubík and Teplý [15].

There are currently many predictive models, but unfortunately few of them are applicable to Slovak companies, because they were developed in other countries, other conditions, etc. In the Czech Republic, the popular Czech model, IN95, IN99, IN01, IN05 or the Králiček index [16], is a rather popular model. A Beerman's discriminatory function is suitable for a manufacturing enterprise. According to [16], due to conditions in Slovakia the Altman's model [17] is often used as

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2018 doi:10.20944/preprints201805.0039.v1

a bankruptcy model, and the Goodwill Index as a good model in general. According to Kislingerová and Hnilica [18], "there are countless methods and approaches to assessing a company's creditworthiness and predicting the possibility of bankruptcy. Financial institutions tend to guard their practices as their special and proprietary know-how. In every of these models do financial indicators play a significant role. Sophisticated statistical procedures that utilize historical data to calculate a company's probability of default based on certain values of financial indicators are often difficult to handle.

2. Materials and Methods

For the purpose of determining the financial situation of non-financial enterprises, prediction models are often employed: Tafflerov model Zt (1), Springate model SM according to (2), Altman's model of Z-score according to (3), AGR model according to (4).

$$Zt = 0.53 \cdot \frac{EBT}{Current\ liabilities} + 0.13 \cdot \frac{OA}{Liabilities} + 0.18 \cdot \frac{Current\ liabilities}{A} + 0.16 \cdot \frac{S}{A}, \tag{1}$$

$$SM = 1.03 \cdot \frac{NPC}{A} + 3.07 \cdot \frac{EBIT}{A} + 0.66 \cdot \frac{EBT}{Current\ liabilities} + 0.44 \cdot \frac{S}{A}, \tag{2}$$

$$Z = 0.717 \cdot \frac{NWC}{A} + 0.847 \cdot \frac{EAT}{A} + 3.107 \cdot \frac{EBT}{A} + 0.42 \cdot \frac{E}{Liabilities} + 0.998 \cdot \frac{S}{A}, \tag{3}$$

$$\begin{gathered} AGR = \frac{EBITDA}{S} + \frac{EAT}{E} + \frac{EBIT}{A} + \frac{S}{A} + \frac{E}{A} + \frac{EBIT + Depreciation}{Depreciation} + \\ + \frac{FA + (Receivables \cdot 0.7)}{Current\ liabilities} \end{gathered}, \tag{4}$$

At present, the method based on logit analysis by Ohlson [10] and another one based on probit analysis by Zmijewski [19] are still popular methods for predicting business failure. In logistic regression the binomial, not the normal distribution describes the distribution of the errors [20], [21]. Estimates of logit model parameters were obtained using a non-linear estimate of maximum likelihood in logistic regression, according to equation (5). Logistic regression models a relationship between predictor variable and a categorical response variable. Independent variables can be continuous, discrete or categorical. Binary logistic regression is used in case of a binary response. A multiple logistic regression is modeling the probability of the variance of the dependent variable depending on the variations of several independent variables. The resulting model can be used to predict the variance of a dependent variable because the output of this model is a probability estimate and defines the dependence between variables. Vochozka [8] (2011, p. 48) describes the advantages and disadvantages of logistic regression. An important advantage of the model is that the logit-score is between 0 and 1, which instantly indicates the probability of the bankruptcy of a company. The weight of the determined coefficients can be interpreted separately, but only if there are no multiple dependences between variables. The initial logarithmic function suggests that, compared to an average healthy company, an extremely healthy company must achieve the majority of improvements (worsening) of its variables proportionately to the improvement (deterioration) of its financial health assessment score.

One disadvantage to this method is that logit models are extremely sensitive to the problem of collinearity in a multiple regression. Therefore, it is necessary to prevent the inclusion of highly dependent variables into model. Ratio financial indicators sometimes have the same nominator or denominator, which can cause a serious problem of multiple dependences. The conditional

 doi:10.20944/preprints201805.0039.v1

probability of occurrence of an event under the condition of occurrence of the x vector (the vector of independent variables, covariates here), can be written p = P (Y = 1 | x), where Y is a binary variable that acquires two possible outcomes. The logistic function expressing the relation between the probability and the vector of the explanatory variables is non-linear and has the form of an exponential function (5).

$$p=\frac{e^{\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_kx_k}}{1+e^{\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_kx_k}}=\frac{1}{1+e^{-(\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_kx_k)}}=\frac{e^z}{1+e^z}=\frac{1}{1+e^{-z}}, \tag{5}$$

By definition, the odds for an event is *p/(1-p)* such that *p* is probability of the event and it equals e^z. In order to use linear regression, the dependent variable is transformed into a continuous value by calculating the logarithm of the odds ("logit" transformation in, order to obtain values from the interval (-∞; ∞)), with odds and probabilities expressing the same information only in a different form (6).

$$\mathrm{logit}(p)=\ln\left(\frac{P(Y-1|X)}{P(Y-0|X)}\right)=\ln\left(\frac{p}{1-p}\right), \tag{6}$$

By logit transformation we get from non-linear to linear dependence, i.e. the relationship between the logarithm of the odds and the vector of the explanatory variables has a linear character. The equation of the logit model has a form according to (7).

$$logit(p)=\ln\left(\frac{p}{1-p}\right)=\ln\left[e^{(\beta_0+\beta_1x_1+\beta_2x_2+...+\beta_kx_k)}\right]=\beta_0+\beta_1x_1+\beta_2x_2...\beta_kx_k, \tag{7}$$

3. Results

The following text contains a proposal for a predictive model for non-financial corporations in the electronics industry using logistic regression.

When looking for an optimal model, we took into account the profitability indicators (x_2, x_3, x_9), activity (x_1, x_5), debt (x_6, x_8) and liquidity (x_4). There were 8 ratios; x_1 - the time in which the firm pays its trade commitments; x_2 - the return on sales measured by the ratio of the operating result before interest and tax and depreciation and net turnover; x_3 - the return on total invested capital measured by the ratio of the profit before tax and total assets; x_4 - ratio of current assets and short-term liabilities; x_5 share of net turnover and assets; x_6 - foreign capital attributable to total capital, x_8 - total assets per unit of equity; and x_9 - return on equity measured by the economic turnout for the accounting period. Based on a database of 738 non-financial corporations, the parameters of the estimated models are shown in Table 1. Of the original database of 748 corporations, 10 breakpoints were removed during the search for the optimal model. As it turned out, it was a measure with a common characteristic, i.e. they were the enterprises that originated during the analyzed period (the second half of 2016).

Table 1. Estimated parameters of the proposed prediction models, their significance, goodness-of-fit statistics and percentage of correctly estimated corporations. The data consists of healthy and 138 defaulted corporations.

	Metrik	Model 1	Model 2
	intercept	–3237.629 (7038.048)	–1.170^d (0.242)
x_1	DSZ_{os}	-	-
x_2	ROS	-	–0.109 (0.097)
x_3	ROI	-	–1.764^d (0.412)
x_4	L_3	-	–0.446^d (0.092)
x_5	OA	-	0.268^c (0.077)
x_6	CZ	3522.748 (7658.021)	-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2018 doi:10.20944/preprints201805.0039.v1

x_8	FP	-	0.011[b] (0.005)
x_9	ROE	-	−0.049[b] (0.016)
	Hosmer-Lemeshow test	0.000	42.734[d]
	$G^2 = -2$ log likelihood	0.009	531.838[d]
	Cox & Snell R^2	0.619	0.216
	Nagelkerke R^2	1.000	0.349
	0(non d.) - 1 (default), good estimated %	100% - 100%	99% - 31.2%

The designation of parameters a, b, c and d in Table 1 expresses the significance of Wald statistics at $p < 0.1$; $p < 0.05$; $p < 0.01$; $p < 0.001$; the number in brackets indicate a standard S.E. The probability of bankruptcy of a company expressed by model 1 is given in equation (8) and by model 2 in equation (9).

$$p = \frac{e^{-3237.629-3522.748x_6}}{1+e^{-3237.629-3522.748x_6}}, \qquad (8)$$

$$p = \frac{e^{-1.170-0.109x_2-1.764x_3-0.446x_4+0.268x_5+0.011x_8-0.049x_9}}{1+e^{-1.170-0.109x_2-1.764x_3-0.446x_4+0.268x_5+0.011x_8-0.049x_9}}, \qquad (9)$$

Logistic regression has no limitations on the distribution of explanatory variables, but they should not be highly correlated (correlation matrix is shown in Table 2). Collinearity has been checked (Variance inflation factor; Condition index) and has not been demonstrated. The sample is large enough for the Hosmer-Lemeshow test, (Hosmer et al., [22], cited at least 400 measurements), which is suitable for testing the model for fit (fit = model fits to given data) with continuous explanatory variables. These are the so-called ungrouped data for which G^2 and chi-square statistics can misinform about the model's lack of fit. For the Hosmer-Lemeshow test, the values of the explanatory variables are artificially distributed on the basis of estimated probabilities to approximately equally large groups, and the chi-square test is applied to them. (Agresti [23]).

For the construction of logistic regression models, the stepwise method was used, testing the significance of adding or eliminating the variable at each step. The test is based on testing the significance of a change in G^2 statistics by adding or excluding a variable. Model 1 was found by a stepwise procedure entering all nine predictors as covariates. The resulting model (Model 1) contains only one explanatory variable x_6 - total indebtedness, where even functional dependence was found. The resulting model 1 with one predictor fits perfectly, $p = 1.000$ for the Hosmer-Lemeshow' test. The probabilities produced by the model reached only two values, 0.000 and 1.000. According to Bewick, Cheek, Ball [24], a small insignificant Wald statistic can be generated for data that produce large coefficient estimates. Therefore the explanatory variable may be incorrectly assumed to be unimportant in the model, which was the case for Model 1. (Wald's Z is 0.212; $p = 0.646$, the percentage match of estimated bankruptcies to actual is 100%). If total debt x_6 after deletion of deviant values is functionally dependent on the default variable, the question arises as to how far default would depend on the other variables considered unless the variable x_6 is not inserted into the model.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 2 May 2018 doi:10.20944/preprints201805.0039.v1

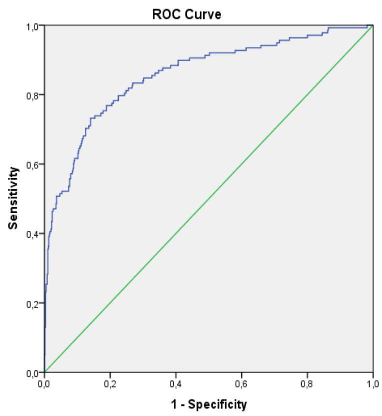
Table 1. Pearson correlations of of variables entering the models, N=738 corporations. The designation r [a, b, c, d,] represents significance at $p<0{,}1$; $p<0{,}05$; $p<0{,}01$; $p<0{,}001$

	x_1	x_2	x_3	x_4	x_5	x_6	x_8	x_9
x_1	1	−.572[d]	−.031	−.019	−.095[b]	.032	.038	.005
x_2		1	.105[c]	.024	.065[a]	−.078[b]	.007	−.004
x_3			1	.042	.138[d]	−.629[d]	−.016	−.473[d]
x_4				1	−.027	−.101[c]	−.018	.014
x_5					1	.161[d]	−.066[a]	−.268[d]
x_6						1	.014	−.014
x_8							1	−.438[d]
x_9								1

After removing the 100% fitting variable x_6, a model 2 with predictors x_2, x_3, x_4, x_5, x_8 and x_9 was designed for a set of 738 non-financial corporations. The Wald test showed a significant contribution of these predictors to the model, except for x2. After removing the insignificant x_2 predictor, we obtain the previously unpublished model, which is very similar in all characteristics to the above model 2. Odds ratio (e^B) means for example, if explanatory variable x_5 is increased by one unit, the odds ratio is $e^{0.268} = 1.308$ and the estimated odds of default of firm are multiplied by 1.308 for given values of other explanatory variables being fixed. The Hosmer-Lemeshow test with $p<0.001$ indicates lack of fit of Model 2 despite the significant coefficients of the five predictors of this model. The number of companies in default differs significantly from the number predicted by this model. The Nagelkerke R^2 statistics (adjusted version of the Cox and Snell R^2) equals 0.349 and indicates 34.9% usefulness of predictors to predict default of firm. But Nagelkerke R^2 do not measure goodness of fit of the model only indicate usefulness.

The discrimination of a model – that is, according to Bewick et al. [24], how well the model distinguishes non defaulting firms from defaulting – can be judged by the area under the receiver operating characteristic curve (AUROC, or ROC). The value of the ROC is the probability that a defaulting firm had a higher predicted probability than did a firm which is not in default. The ROC for the Model 2 gave a value of 0.857 (S.E.=0.020; p=0.000; 95% C.I.=0.818-0.895), indicating that the model discriminates well. For example, Valecký and Slivková [14] report the value of AUC ROC 0.862 as highly reliable.

Figure 1. ROC curve for model 2; (AUC ROC = 0.857, S.E. = 0.020, p = 0.000, 95% CI = 0.818-0.895)



4. Discussion

By applying logistic regression, two logistic regression models were found to estimate the company's probability of bankruptcy. The first model is a deterministic model in which it is possible to estimate bankruptcy using a single explanatory variable, i.e. the indebtedness index x_6 - CZ. After the omission of this predictor, a second stochastic model was developed that explains the bankruptcy of the company using six predictors: profitability is defined by (x_2 - ROS, x_3 - ROI, x_9 - ROE), activity (x_5 - OA) (x_4 - L3). It should be noted that by selecting from the given 8 predictors, none of the resulting stochastic models lead to good fits, although AUC ROC often exhibits high reliability.

5. Conclusions

For further research and modeling by logistic regression, it is necessary to take into account other variables. As well as qualitative variables, other variables that should be taken into account include the region or size of the business entity. The size of a business is an important factor in predicting the failure of a newly established companies, because small businesses are more susceptible to bankruptcy than major corporations. Large companies are predicted to have lower failure rates due to their size because they can make larger transactions on more advantageous terms, etc.

6. Patents

Author Contributions: Conceptualization, Sylvia Jenčová and Eva Litavcová; Methodology, Eva Litavcová and Sylvia Jenčová.; Software, Eva Litavcová; Validation, Sylvia Jenčová and Róbert Štefko; Formal Analysis, Eva Litavcová; Investigation, Róbert Štefko; Resources, Eva Litavcová; Data Curation, Sylvia Jenčová; Writing-Original Draft Preparation, Eva Litavcová and Sylvia Jenčová; Writing-Review & Editing, Róbert Štefko; Visualization, Sylvia Jenčová; Supervision, Sylvia Jenčová; Project Administration, Eva Litavcová; Funding Acquisition, Róbert Štefko.

Funding: This research was funded by the grant No. 1/0470/18 and by the grant No. 1/0945/17 of the Grant Agency VEGA.

References

1. Kubíčková, D., Jindřichovská, I. *Finanční analýza a hodnocení výkonnosti firmy.* Praha: C.H.BECK, Czech, 2015. ISBN 978-807400-538-1.
2. Neumaierová, I., Neumaier, I. Index IN05. In ČERVINEK, P. ed. *Evropské finanční systémy*, Brno: MU, Czech. pp.143-148. 21st June 2005 -23rd June 2005. Available online: http://is.muni.cz/do/1456/sborniky/2005/evropske-financni-systemy-2005.pdf (accessed on 15 February 2017).
3. Neumaierová, I., Neumaier, I. Proč se ujal index IN a nikoli pyramidový systém ukazatelů INFA. *Ekonomika a management.* **2008**, *Volume 4*, pp. 1-10. Available online: https://www.vse.cz/eam/51 (accessed on 21 February 2017).
4. Kameníková, K. Limitation of models used for predicting the financial development of firms in the Slovak Republic. *Acta Montanistica Slovaca*. **2005**, Volume 10 (3), pp. 337-343,
5. Pilch, C. K modelom hodnotenia finančného zdravia podniku. *Finančné trhy.* **2008**, *Volume* 5, pp.1-7. Available online: http://www.derivat.sk/index.php?PageID=1420 (accessed on 17 October 2017).
6. Bondareva, I. Analysis of explanatory models of the predictive ability of the financial condition of the company in Slovakia. *Manažment podnikania a vecí verejných.* Bratislava: SAM, 2011, pp. 59-64.
7. Kadarová, J., Turisová, R. Finančné modely predikcie finančných problémov v priemyselných podnikoch. In: Modelování, simulace a optimalizace podnikových procesú v praxi, 29. March 2011; Tuček, D., Eds. Praha: ČSOP, Czech, 2011. pp. 167-173.
8. Vochozka, M. *Metody komplexního hodnocení podniku.* Praha: Grada Publishing, Czech, 2011. ISBN 978-80-247-3647-1.
9. Kabát, L., Sobeková Majková, M., Vincúrová, Z. *Hodnotenie podniku a analýza jeho finančného zdravia.* Bratislava: Iura Edition, Slovakia, 2013. ISBN 978-80-8078-608-3.
10. Ohlson, J.A. Financial Ratios and Probabilistic Prediction of Bankruptcy. *Journal of Accouting Research.* **1980**, *Volume 18 (1)*, pp. 109-131.

 doi:10.20944/preprints201805.0039.v1

11. Slavíček O., Kuběnka, M. Bankruptcy prediction models based on the logistic regression for companies in the Czech Republic. In 8th International Scientific Conference Managing and Modelling of Financial Risks. Ostrava: VŠB-TU of Ostrava, Czech, 2016. pp. 924-931. Available online: https://www.ekf.vsb.cz/export/sites/ekf/rmfr/cs/sbornik/Soubory/Part_IIIB.pdf (accessed on 15 November 2017).
12. Marcinkevičius R., Kanapickienė, R. Bankruptcy Prediction in the Sector of Construction in Lithuania. In: 19th International Scientific Conference Economics and Management, ICEM 2014, Riga, Latvia. Available online: https://doi.org/10.1016/j.sbspro.2014.11.239 (accessed on 15 January 2018).
13. Harumová, A., Janisová, M. Rating Slovak Enterprises by Scoring Functions. *Journal of Economics.* **2014**, *Volume 62 (5)*, p. 522 – 539
14. Valecký, J., Slivková, E. Mikroekonomický scoringový model úpadku českých podniků. *Ekonomická revue - Central European Review of Economic.* **2012**, *Volume 15*, p. 15–26, DOI:10.7327/cerei.2012.03.02
15. Jakubík, P., Teplý, P. The JT Index as an Indicator of Financial Stability of Corporate Sector. *Prague economic pape.r* **2011**, *Volume 20 (2)*, p. 157-176. DOI 10.18267/j.pep.394
16. Kraliček, P. *Základy finančního hospodaření.* Praha: Linde, Czech, 1993.
17. Altman, E.I. Predicting financial distress of companies. Revisiting the Z-score and Zeta model.
18. Kislingerová, E., Hnilica, J. *Finanční analýza - krok za krokem.* Praha: C. H. Beck, Czech, 2005. ISBN 80-7179-321-3.
19. Zmijewski, M. E. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accouting Research.* **1984**, *Volume 22*, pp. 59-86. DOI: 10.2307/2490859
20. Hosmer, D. W., Lemeshow, S. *Applied logistic regresion.* John Wiley, Sons. New York, USA, 1989.
21. Theil, H. *Principles of Econometrics*. New York: Wiley, USA, 1971.
22. Hosmer, D.W., S. Lemeshow, Sturdivant, R. X. Applied Logistic Regression, 3rd ed. John Wiley & Sons, Inc., 2013.
23. Agresti, A. (2015). *Foundations of Linear and Generalized Linear Models*. Wiley Interscience, John Wiley & Sons, Inc., Hoboken, New Jersey, USA.
24. Bewick V., Cheek, L., Ball, J. Statistics review 14: Logistic regression. *Critical Care* 2005, *Volume 9 (1)*, p. 112-118. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1065119/ (accessed on 20 December 2017).
25. Rajnoha R., Štefko, R., Merková, M., Dobrovič, J. Business Intelligence as a Key Information and Knowledge tool for Strategic Business Performance Management. *E & M Ekonomie a Management.* **2016**, *Volume 19 (1)*, pp. 183-203. DOI 10.15240/tul/001/2016-1-013.