The Prediction Capabilities of Bankruptcy Models in a Different Environment: An example of the Altman Model under the Conditions in the Visegrad Group Countries

*Mária REŽŇÁKOVÁ – Michal KARAS**

Abstract

This paper presents the results of research into the discrimination capability of the Altman bankruptcy model. The authors are contributing in this way to discussion of the possible transferability of models that have been created in a different environment or a different time period. Efforts at model transfer are motivated by an assumption to obtain the same or similar discrimination accuracy for the given model as that declared by its creators. The tests performed have clearly shown that the discrimination accuracy of a model falls significantly when it is used in a different environment. This led in turn to an investigation of ways in which the discrimination capability of a model may be increased by means of the determination of new weightings for model variables and grey-zone boundaries. The accuracy of the original models was not attained, although an increase was seen in the discrimination accuracy of these models.

Keywords: *bankruptcy prediction model transferability, linear discrimination analysis, Wilcoxon test, bootstrap*

JEL Classification: G33, C51

1. Introduction

The first attempts to predict bankruptcy sufficiently in advance reach back to the 60s of the last century. Beaver (1966) demonstrated as the first one that to predict bankruptcy, financial indicators can be used. Altman (1968) continued his work and created the first bankruptcy model. In response to these works, more bankruptcy models were created (see Deakin, 1972; Martin, 1977; Altman, Haldeman and Narayanan, 1977; Altman, 2000; Ohlson, 1980; Taffler, 1982;

* Mária REŽŇÁKOVÁ – Michal KARAS, Brno University of Technology, Faculty of Business and Management, Department of Finance, Kolejní 2906/4, 612 00 Brno, Czech Republic; e-mail: reznakova@fbm.vutbr.cz; karas@fbm.vutbr.cz

Zmijewski, 1984; Tam and Kiang, 1992; Shumway, 2001; Sánchez-Lasheras et al., 2012, and many others). The Altman model is among the most cited and hence the best known model. The original version of the Altman model was intended only for companies listed on the capital market. Later the modification of the model was published for companies not listed in the capital market (see Altman, 2000): the so-called *revised Z-score*, which became very popular even in our conditions. The modification of the model that dates from 1983 enabled its wider use, which was probably contributed to by the simplicity of the formula. The popularity of the model is summarized by Mandru et al. (2010), according to whom the Altman model (see Altman, Haldeman and Narayanan, 1977) is still robust, even though it was developed more than 30 years ago. This view was confirmed also by other studies (see, Li and Ragozar, 2012; Satish and Janakiram, 2011; El Khoury and Al Beaïno, 2014; Al Khatib and Al Bzour, 2011). Conversely, Wu, Gant and Grey (2010), Grice and Dugan (2001), Pitrová (2011) and others have come to the opposite conclusion. The results of these researches show that predictive accuracy of models significantly decreases if the model is used in another industry, in another time and/or in another business environment than that in which the data used to derive the model were obtained. According to Niemann, Schmidt and Neukirchen (2008), the cause can be found in a different structure of values in the financial statements of companies in individual countries. These differences in the structure of the financial statements arise from different values of key macroeconomic indicators, such as interest rates, the level of taxation, the wage levels, the access to the capital market, and so on. The effectiveness of the Altman model in the Czech Republic was investigated by Machek (2014), who came to the conclusion that the Z-score is indeed more effective for Czech companies than, for example, the Taffler model or the Kralicek quick test, but less effective compared with domestic model IN 05. The attention of scientists focused on studying the causes for decreasing discrimination abilities of the Altman model. According to Shumway (2001) and Li (2012), who studied the significance of variables of the Altman Z-score in the US environment, the reason for less predictive accuracy of the Altman model may lie in the different discrimination ability of individual variables occurring in the model.

Our research has followed up the results of the above studies; it was conducted on the data of companies of the countries of the Visegrad Group (the Czech Republic, Slovakia, Poland and Hungary, hereinafter referred to as Visegrad Group – V4). The aim of the research is to test the predictive capability of the original version of the Altman model in the environment different from the environment of its origin, thus exploring its transferability to a different economic environment. By transferability we understand the possibility of the prediction capability of the model to achieve results similar to those originally obtained by the author of the model, i.e. in the capability of differentiating companies in the risk of bankruptcy from financially healthy (active) companies in a country other than the country of origin of the model with the same or similar accuracy. Another goal is to test the prediction capability of the models derived for each country based on the country data, using the same variables as used by Altman in the model for 1983. During this we assume that modified models derived for each country separately will have higher prediction accuracy. The secondary goal is to offer the process of the model modification, which leads to increasing its prediction capability.

The goals above led to the formulation of the following hypotheses:

Hypothesis H_1: *The predictive accuracy of the original version of the model in the V4 countries is statistically different from the accuracy in the original sample.*

Alternative hypothesis $H_{1(1)}$: *The predictive accuracy of the original version of the model in the V4 countries is the same as the accuracy in the original sample*.

If hypothesis H_1 is confirmed, following hypothesis H_2 will be also tested.

Hypothesis H_2: *By deriving the new function using the Altman model and determining the new grey-zone boundaries, statistically higher accuracy of the model can be achieved.*

The alternative hypothesis is $H_{2(1)}$: *By deriving the new function using the variables of the Altman model and determining the new grey-zone boundaries, statistically higher accuracy of the model cannot be achieved.*

If hypothesis H_2 is confirmed, hypothesis H_3 will be tested.

Hypothesis H_3: *It is possible to achieve the same or higher predictive accuracy of the model with fewer variables than contained in the original model, with the same or lower share of non-evaluated companies.*

Alternative hypothesis $H_{3(1)}$: *It is not possible to achieve the same or higher predictive accuracy of the model with fewer variables than contained in the original model, with the same or lower share of non-evaluated companies.*

The hypotheses will be verified on the basis of comparing medians of the achieved predictive accuracy of the original version of the Altman model or of the derived models (the so-called modified and reduced models for each country).

2. Sample Studied and Methods Used

The Altman model is based on financial ratio indicators, which are used in the long term to evaluate the stability and financial health of companies (see, for example, Czillingová, Petruška and Tkáč, 2012). In the case of the Altman model, it is a combination of five indicators, which – according to its author – surpassed other alternatives in terms of its predictive accuracy and correlation between indicators. The Revised Z-score, which is the subject of our research (hereinafter referred to as the original version of the model), can be written as follows (see Altman, 2000):

$$\text{Z-score} = 0.717 \cdot X1 + 0.847 \cdot X2 + 3.107 \cdot X3 + 0.42 \cdot X4 + 0.998 \cdot X5 \qquad (1)$$

where

X1 – (current assets – short-term debt)/total assets,
X2 – retained profit/total assets,
X3 – operating profit/total assets,
X4 – book value of equity/total debts, X5 = sales/total assets.

The grey zone of the model is represented by interval <1.23; 2.9>. On the basis of the said model one year before bankruptcy, Altman correctly identified 90.9% bankruptcy companies and 97% active companies (hereinafter the reference values). These reference values will be tested as part of the verification of hypothesis H_1.

The sample studied includes the financial statements of 5 977 companies in the manufacturing industry (NACE rev. 2 main section C), operating in one of the V4 countries (hereinafter the analysed data), of which 4 220 companies are financially healthy (active), and 1 757 companies, which went bankrupt in the following year (bankruptcy). The data were obtained from the Amadeus database provided by the company Bureau Van Dijk. In the bankruptcy companies, data from the statements one year before the bankruptcy were used. The structure of the studied sample is given in Table 1.

T a b l e 1

Structure of the Studied Sample in each Country

Country	**Population**	**Sample studied**				
		Active	Bankruptcy	Total	Share of active	Share in population
Czech Republic (CZ)	172 162	857	379	1 236	69.34%	0.72%
Hungary (HU)	51 161	1 463	1 070	2 533	57.76%	4.95%
Poland (PL)	176 471	1 583	127	1 710	92.57%	0.97%
Slovakia (SK)	69 083	317	181	498	63.65%	0.72%
Total	468 877	4 220	1 757	5 977	70.60%	1.27%

Source: Our processing on the basis of data from the Amadeus database.

In the sample, all companies were included whose data were contained in the database and which went bankrupt in the period 2007 – 2012. These data were then supplemented by 4 220 active companies, selected randomly. The Beaver-Altman approach of the so-called matched-pairs was not deliberately followed; it consists in comparing companies of the same size with each other, because it reduces the size of the sample and hence the degree of freedom (Taffler, 1982).

To assess the relevance of the sample, the minimum size of the sample was determined; for this, the Cochran process was used (see Cochran, 1977). At the 5% level of significance, at the 5% error rate and at the maximum variance of the sample (p = 0.5), the minimum number of companies ranged from 381 to 383, depending on the size of the population. From this point of view, the studied sample is large enough to generalize the results.

To estimate the parameters of the probability distribution of the samples, the non-parametric bootstrap was applied at 1 000 replications. This procedure allowed us to derive descriptive statistics of the model accuracy, and to test the significance of the differences of the measured values against the reference values of the accuracy of the original version of the model.

Let us assume that by testing, we obtained n independent values x_1, x_2, …, x_n, of which we calculated characteristics X of our interest (here, for example, the model accuracy). *The bootstrap sample will be obtained by generating ("selecting by repeating") n random numbers from set* x_1, x_2... x_n; $x^* = (x_1^*, x_2^* \ldots x_n^*)$. *For this sample, we also calculate relevant characteristics* X*. *If we repeat this entire process B-times, we get values* X_1^*, X_2^* ... X_B^*, *which represent the bootstrap-population of characteristics* X* (Menčík, 2001).

For the testing, the Wilcoxon paired test was used (see Wilcoxon, 1945), which can be used to verify the hypothesis that two random variables X and Y are the same in terms of the position (i.e. their medians coincide), or H_0: $z_{0.50} = x_{0.50} - y_{0.50}$ or $z_{0.50} = 0$.

Test statistic for large samples can be written in the following format:

$$Z = \frac{W - \frac{n \cdot (n+1)}{4}}{\sqrt{\frac{n \cdot (n+1) \cdot (2n+1)}{24}}} \tag{2}$$

and

$$W = \sum_{Z_i > 0} R_i^+ \tag{3}$$

where

R_i^+ – the order of quantity $|Z_i|$,

n – number of observations.

In the case of H_0 validity, the test statistic has normal distribution N (0, 1). Hypothesis H_0 is rejected if $|Z| \geq u_{1-\alpha/2}$, where $u_{1-\alpha/2}$ is the quantile of standard normal distribution.

To verify hypothesis H_1, the original version of the model was tested on V4 data. The share of correctly evaluated bankruptcy (or active) companies was studied, as well as the proportion of companies in the grey zone (i.e. non-evaluated) to the number of observations of bankruptcy (or active) companies (always for the given country). The hypothesis is confirmed if the medians of accuracy of the model in all countries studied are statistically different from the reference value in the sample of both active and bankruptcy companies. The calculation was performed in the Statistica programme.

To test hypothesis H_2, coefficients of variables contained in the Altman model for each country were derived, and the so-called modified models created in four variants. The same method of the linear discrimination analysis was used for the derivation as used by Altman for the creation of his model. The resolution ability of the models obtained was tested in the same manner as in the case of verifying hypothesis H_1.

Hypothesis H_3 was tested in the same manner. The assumption included the derivation of the so-called reduced models for each country using the method of backward stepwise discrimination, in which variables with lower significance were eliminated from the model.

3. Results and their Discussion

In the original setting, the Altman model was very successful in recognizing prosperous companies and companies at the risk of bankruptcy, i.e. for the prediction of bankruptcy. The prediction capability of the original version of the model was therefore tested first.

3.1. Results of Testing the Original Version of the Model

The testing of the original version of the model was carried out in three steps. First, the number of companies located in the grey zone was evaluated. The results are shown in Table 2.

T a b l e 2

Percentage of Non-evaluated Companies in the Original Version of the Model

	Median	Min.	Max.	Std. Dev.		Median	Min.	Max.	Std. Dev.
CZ (A)	41.10	36.11	47.01	1.68	CZ (B)	31.78	25.01	40.10	2.41
SK (A)	45.43	35.39	54.19	2.88	SK (B)	39.73	27.91	51.23	3.62
PL (A)	42.23	38.61	46.26	1.21	PL (B)	23.48	13.56	35.66	3.74
HU (A)	49.06	43.49	54.89	1.75	HU (B)	31.84	27.99	36.53	1.45

Source: Our processing on the basis of data from the Amadeus database.

The median of the share of active companies the model failed to evaluate ranged from 41.10% (the Czech Republic) to 49.06% (Hungary). The median of the share of bankruptcy companies the model failed to evaluate ranged from 23.48% (Poland) to 39.73% (Slovakia), while in the grey zone, active companies prevailed over bankruptcy ones. This situation is most noticeable in the sample of Polish companies, where 1.79 times more active Polish companies are found in the grey zone than the bankruptcy companies. This situation is least obvious in the sample of Slovak companies, when there were 1.14 times more active companies in the grey zone. The accuracy with which the original model is able to identify an active or bankruptcy company is shown in the Table 3.

Table 3

Discrimination Accuracy of the Original Version of the Model (in percentage points – pp)

	Median	**Min.**	**Max.**	**Std. Dev.**		**Median**	**Min.**	**Max.**	**Std. Dev.**
CZ (A)	48.51	43.45	54.16	1.70	CZ (B)	46.07	37.56	55.06	2.48
SK (A)	39.31	32.21	48.75	2.79	SK (B)	39.66	29.06	53.07	3.61
PL (A)	46.59	43.33	50.42	1.24	PL (B)	66.96	49.61	80.7	4.07
HU (A)	38.91	33.16	44.15	1.78	HU (B)	43.16	36.98	48.52	1.60

Source: Our processing on the basis of data from the Amadeus database.

Within the V4 countries, it is possible to most accurately recognize active companies in the Czech Republic on the basis of the original version of the model (48.51% of correctly evaluated companies); conversely, the lowest accuracy applies to Hungary (38.91%). In the case of bankruptcy companies, the model is most accurate in Poland (66.96%), and least accurate in Slovakia (39.66%). These values are considerably lower than the values originally obtained by the author for the American environment, which confirms hypothesis H_1. Yet the statistical significance of the difference was tested; the results are shown in Table 4.

Table 4

Results of Hypothesis H_1 Testing

	no.	**W-stat.**	**Z-stat**	**p-val.**		**no.**	**W-stat.**	**Z-stat**	**p-val.**
CZ (A)***	1000	0	27.393	0.000000	CZ (B) ***	1000	0	27.393	0.000000
SK (A) ***	1000	0	27.393	0.000000	SK (B) ***	1000	0	27.393	0.000000
PL (A) ***	1000	0	27.393	0.000000	PL (B) ***	1000	0	27.393	0.000000
HU (A) ***	1000	0	27.393	0.000000	HU (B) ***	1000	0	27.393	0.000000

Note: *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Source: Our processing on the basis of data from the Amadeus database.

According to the Wilcoxon test conclusion, medians of the model accuracy in all studied states are statistically different from the reference value on a sample of both active and bankruptcy companies. This means that null hypothesis H_1 was confirmed.

3.2. Derivation and Testing of Modified Versions of the Model

To verify the validity of hypothesis H_2, a new function was derived (i.e. coefficients of model variables were recalculated) separately for each V4 country; this created a total of four new modified models (see equations 4 to 7). We succeeded in significantly increasing the overall resolution capability of the models – see Tables 6 and 7.

After deriving the models, it was necessary to analyse the interval of the values of the function, in which the models achieve the greatest error rate. The error here means the designation of the bankruptcy company as active (empirical error II) and vice versa, i.e. the designation of the active company as a bankruptcy company (empirical error I). Intervals in which errors occur are very scattered, but concentrated in a relatively narrow interval. For example, under the model derived for Poland, errors occur in the interval from –72.638 to 204.044, but 50% of the values were identified in the range from 4.095 (bottom quartile) to 5.599 (upper quartile).

Quantile boundaries of error values were thus used to determine the grey-zone boundaries. For this, different combinations of order statistics had to be explored. The criterion for determining the grey-zone boundaries is achieving the highest predictive accuracy of the model in both active and bankruptcy companies while minimizing the share of non-evaluated companies. The best results were achieved in setting the boundaries of the grey zone to the value of the lower quartile and the error median. Final forms of the modified models can be written in the following format (grey-zone intervals for each model variant are shown after the equation):

$$\text{Model CZ} = -0.0218 \cdot X1 + 0.0750 \cdot X2 + 0.0327 \cdot X3 + 0.0669 \cdot X4 + 0.0159 \cdot X5; \quad \langle 0.065, 0.115 \rangle \tag{4}$$

$$\text{Model SK} = 0.1128 \cdot X1 - 0.01155 \cdot X2 + 1.5202 \cdot X3 - 0.0065 \cdot X4 - 0.0286 \cdot X5; \quad \langle -0.04, 0.044 \rangle \tag{5}$$

$$\text{Model PL} = 0.629 \cdot X1 + 0.744 \cdot X2 + 6.77 \cdot X3 + 0.0043 \cdot X4 - 0.152 \cdot X5; \quad \langle -0.043, 0.423 \rangle \tag{6}$$

$$\text{Model HU} = -0.042 \cdot X1 + 0.046 \cdot X2 + 0.001 \cdot X3 + 0.003 \cdot X4 - 0.034 \cdot X5; \quad \langle -0.083, -0.055 \rangle \tag{7}$$

Modified models were tested on the same sample as the original version of the model; non-parametric bootstrap process was also used again. First of all the share of non-evaluated companies was studied.

T a b l e 5

Percentage of Non-evaluated Companies in Modified Model Versions

	Median	Min.	Max.	Std. Dev.		Median	Min.	Max.	Std. Dev.
CZ (A)	25.60	20.16	29.94	1.47	CZ (B)	17.89	12.15	24.48	1.96
SK (A)	26.73	19.16	36.59	2.49	SK (B)	24.14	14.67	35.86	3.04
PL (A)	25.53	20.88	29.19	1.09	PL (B)	15.70	5.97	27.13	3.28
HU (A)	33.33	28.17	38.30	1.66	HU (B)	24.15	19.98	27.84	1.16

Source: Our processing on the basis of data from the Amadeus database.

In all modified models, reduction of the share of non-evaluated companies was achieved in the sample of both active and bankruptcy companies. For active companies, this difference was biggest in the sample of Slovak companies (18.70 pp), while the smallest difference was reached in the sample of Czech companies (15.5 pp). Within the bankruptcy companies, this difference was biggest in the sample of Slovak companies (15.59 pp), and smallest in the sample of Hungarian companies (7.69 pp).

T a b l e 6

Discrimination Accuracy of Modified Models (in pp)

	Median	Min.	Max.	Std. Dev.		Median	Min.	Max.	Std. Dev.
CZ (A)	50.79	45.62	56.55	1.69	CZ (B)	72.28	64.99	79.66	2.25
SK (A)	53.66	45.70	62.83	2.78	SK (B)	53.67	41.18	64.89	3.59
PL (A)	50.71	46.69	54.48	1.24	PL (B)	72.59	57.66	84.68	3.90
HU (A)	55.93	50.71	61.72	1.73	HU (B)	27.23	23.36	31.26	1.21

Source: Our processing on the basis of data from the Amadeus database.

Median accuracy of modified models in active enterprises ranges from 50.71% in Poland up to 55.93% in Hungary. By recalculation, these values increased most in Hungarian companies (17.02 pp), and least in Czech companies (2.28 pp). The values of median accuracy for bankrupt companies range from 27.23% in Hungary up to 72.59% in Poland. By recalculation, these values increased most in Czech companies (26.21 pp), while in Hungarian companies the accuracy decreased by 15.93 pp. The difference between the median accuracy of the original version of the model and modified models was tested using the Wilcoxon test. The results are shown in the Table 7.

T a b l e 7

Results of Hypothesis H_2 Testing

	No.	W-stat.	Z-stat	p-val.		No.	W-stat.	Z-stat	p-val.
CZ (A)***	1000	45 049	22.462	0.000000	CZ (B)***	1000	0	27.393	0.000000
SK (A)***	1000	0	27.393	0.000000	SK (B)***	1000	20	27.391	0.000000
PL (A)***	1000	158	27.376	0.000000	PL (B)***	1000	40 309.5	22.981	0.000000
HU (A)***	1000	0	27.393	0.000000	HU (B)***	1000	0	27.393	0.000000

Source: Our processing on the basis of data from the Amadeus database.

The test confirmed that all medians of accuracy of modified models were statistically different at the level of significance of 1% from medians of accuracy achieved when the original version of the model was used. The modification of the model decreased the median of the share of non-evaluated companies as opposed to the original version of the model (see Tables 2 and 5), while the median of predictive accuracy of the models in all studied countries increased (see Tables 3 and 6). Although the aforementioned differences are statistically significant, hypothesis H_2 can be considered as confirmed only for the Czech Republic, Poland and Slovakia. The hypothesis was not confirmed for the Hungarian bankruptcy companies. This finding is consistent also with the result of the tests of overall discrimination capability of modified models – see Table below 8.

Table 8

Overall Discrimination Capability of Modified Models

Model	Wilk's lambda	F-stat.	p-val.	Model	Wilk's lambda	F-stat.	p-val.
CZ***	0.95492	11.613	<0.0000	PL***	0.75193	112.43	<0.0000
SK***	0.93748	6.5626	<0.0000	HU***	0.97893	9.1610	<0.0000

Source: Our processing on the basis of data from the Amadeus database.

All models are statistically significant at the level of 1%. The highest discrimination capability was achieved by the model for the data of Polish companies (Model PL), the lowest for Slovak companies.

One of the reasons for the lower discrimination capability of models can be the method used to derive the model (see e.g. Karas and Režňáková, 2014). To derive the weights of the variables, the same method was used as used by Altman to derive his model: the method of linear discrimination analysis, which assumed multivariate normal distribution of data; this – however – is a very rare phenomenon with financial ratio indicators. Frequent disproportionality between the numerator and denominator of the ratio indicators can be considered as the cause (Whittington, 1980). Although this method shows the drawbacks mentioned, we did not want to influence the results by the selection of the method.

3.3. Results of Testing Statistical Significance of Model Variables

Although there was an increase of discrimination capability of modified models, their prediction capability did not reach the level declared by the author of the model. The cause of the lower discrimination capability of the model can be the variables used. Although the variables were used, characterized by a high discrimination capability in the original environment, in a different environment they may loose this property. The reason is the multi-collinearity of variables.

As a basis, we use the Cochran thesis (Cochran, 1964) – that positive correlation between individual pairs of indicators reduces the discrimination capability of the model based on the discrimination analysis method. To evaluate the degree (severity) of the multi-collinearity, the so-called tolerance is used, which expresses unique contribution of the variable to the overall explanatory capability of the model (Craney and Surles, 2002). For the above reasons, the statistical significance of the contribution of individual variables to the discrimination capability of the model in individual countries and their tolerances were tested in the next step. The results are shown in the Table 9.

T a b l e 9

Significance of Model Variables According to Countries

	Wilk's Lam.	Part. Lam.	F to remove	p-val.	Toler.		Wilk's Lam.	Part. Lam.	F to remove	p-val.	Toler.
X1 (CZ)***	0.971	0.9830	20.99	0.000005	0.033	X1 (PL)***	0.777	0.968	55.90	0.000000	0.668
X2 (CZ)***	0.975	0.9800	25.30	0.000001	0.027	X2 (PL)***	0.779	0.965	61.60	0.000000	0.653
X3 (CZ)***	0.961	0.9940	7.15	0.007593	0.482	X3 (PL)***	0.792	0.950	89.64	0.000000	0.024
X4 (CZ)***	0.973	0.9820	22.67	0.000002	0.998	X4 (PL)	0.752	1.000	0.01	0.909240	0.956
X5 (CZ)***	0.967	0.9880	14.92	0.000118	0.165	X5 (PL)***	0.790	0.952	86.21	0.000000	0.024
X1 (SK)*	0.945	0.9920	3.79	0.052088	0.237	X1 (HU)**	0.981	0.998	4.64	0.031393	0.012
X2 (SK)	0.938	0.9990	0.42	0.517474	0.240	X2 (HU)***	0.983	0.996	8.45	0.003686	0.009
X3 (SK)***	0.975	0.9610	19.78	0.000011	0.978	X3 (HU)	0.980	0.999	1.79	0.181416	1.000
X4 (SK)	0.938	0.9990	0.36	0.546642	0.990	X4 (HU)**	0.982	0.997	6.06	0.013902	0.999
X5 (SK)	0.938	1.0000	0.21	0.650693	0.986	X5 (HU)***	0.995	0.984	34.56	0.000000	0.203

Note: *significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Source: Our processing on the basis of data from the Amadeus database.

The variable 'ratio of net working capital and total assets (X1)' is statistically significant at the significance level of 1% in Model PL and Model CZ, in which tolerance has the second lowest value, at the significance level of 5% in Model HU, and at the significance level of 10% in Model SK. The variable 'ratio of retained profit to total assets (X2)' is statistically significant at the level of 1% in Model CZ, in which it is the most significant variable, as well as in Model HU and Model PL. The variable 'return on assets (X3)' is statistically significant at the significance level of 1% in all models except Model HU; it represents the most significant variable in Model PL. The variable 'ratio of the book value of equity and total external sources (X4)' is statistically significant at the level of 1% only in Model CZ, where – in addition – this indicator reaches the highest value of tolerance; i.e. it most contributes to the differentiation of active companies from bankruptcy companies. In Model HU, the given indicator is significant at the level of 5%; in other models it is not significant at any standard level of significance. The variable 'ratio of sales and assets (X5)' is statistically significant at the level of 1% in all models except Model SK.

The said analysis clearly revealed the causes of lower discrimination capabilities of modified models than that the author declares for the original version of the model.

3.4. Derivation and Testing of Reduced Models

To verify hypothesis H_3, the so-called reduced models were derived using backward stepwise discrimination, in which the variables of lower significance are omitted from the model (see equations 8 to 11). This process was based on the conclusion that of the two models, the one is preferred, which achieves the same performance with a smaller number of explanatory variables (Greene, 2012).

$$\text{Model CZ red.} = 0.0079 \cdot X2 + 0.0692 \cdot X4; \qquad <0.0287,\ 0.0654> \tag{8}$$

$$\text{Model SK red.} = 1.6274 \cdot X3; \qquad <-0.1650,\ -0.0851> \tag{9}$$

$$\text{Model PL red.} = 0.6294 \cdot X1 + 0.7436 \cdot X2 + 6.7841 \cdot X3 - 0.1523 \cdot X5; \qquad <0.0515,\ 0.5437> \tag{10}$$

$$\text{Model HU red.} = 0.007 \cdot X2 - 0.016 \cdot X5; \qquad <-0.020, -0.036> \tag{11}$$

The overall discrimination capability of the derived reduced models is shown in the Table 10.

T a b l e 10

Overall Discrimination Capability of Reduced Models

Model	Wilk's lambda	F-stat.	p-val.
CZ***	0.97127	18.233	<0.0000
SK***	0.95393	23.953	<0.0000
PL***	0.75194	140.62	<0.0000
HU***	0.98572	18.327	<0.0000

Source: Our processing on the basis of data from the Amadeus database.

All reduced models as a whole are statistically significant at the level of 1%. Nevertheless, the overall discrimination capability according to Wilk's lambda is very low for all models except Model PL. In reverse discrimination, only variables X2 and X4 were left in the Model CZ red., variables X2 and X5 in the Model HU red., and one variable, X4, in the Model SK red. Even though all variables of reduced models are statistically significant at the significance level of 1%, the variables of the model for Poland achieve significantly higher significance in comparison with variables of reduced models for the SR and CR.

For all reduced models, the error rate was also analysed, and the grey zone derived in a similar manner as in the previous case (see equations 4 – 7). The share of non-evaluated companies in the total number of companies (valid observations) was explored first, see the Table 11.

T a b l e 11

Percentage of Non-evaluated Companies in Reduced Model Versions

	Median	Min.	Max.	Std. Dev.		Median	Min.	Max.	Std. Dev.
CZ (A)	24.96	20.33	29.99	1.40	CZ (B)	10.77	6.14	15.03	1.58
SK (A)	26.21	17.07	34.92	2.55	SK (B)	28.64	17.22	39.67	3.46
PL (A)	25.49	21.89	29.83	1.12	PL (B)	13.39	5.08	22.41	2.92
HU (A)	40.15	33.25	45.65	1.77	HU (B)	24.21	20.36	27.59	1.19

Source: Our processing on the basis of data from the Amadeus database.

By comparing medians of the shares of non-evaluated companies among reduced (equations 8 – 11) and modified models (equations 4 to 7), it can be ascertained that in models for the CR and Poland, the share of non-evaluated companies was reduced in the sample of both active and bankruptcy companies. In the case of Slovakia, there has been a decline in the share of non-evaluated enterprises in active companies only; in terms of bankruptcy companies, the share increased by 4.51 pp. In the case of Hungary, there has been a decline in the sample of both active and bankruptcy companies – by 6.82 pp and 0.06 pp. The prediction capability of the model, i.e. the accuracy of the reduced models on the bootstrap sample, is shown in the Table 12.

T a b l e 12

Discrimination Accuracy of Reduced Models (in pp)

	Median	Min.	Max.	Std. Dev.		Median	Min.	Max.	Std. Dev.
CZ (A)	50.34	44.65	55.39	1.68	CZ (B)	83.15	76.04	88.75	1.92
PL (A)	50.92	46.79	55.22	1.28	PL (B)	81.08	68.70	91.53	3.41
SK (A)	52.90	42.12	62.01	2.86	SK (B)	46.97	34.68	60	3.71
HU (A)	45.49	39.01	50.98	1.79	HU (B)	27.57	23.16	31.58	1.24

Source: Our processing on the basis of data from the Amadeus database.

The comparison of medians of accuracy of reduced models with accuracies of modified models has shown that the reduction of the number of indicators led to increased accuracy both on the sample of active (by 0.22 pp) and bankruptcy companies (by 8.49 pp) only in the model for Polish companies. In Model CZ red., where the reduction of the number of indicators was more noticeable, there was an increase of accuracy in the sample of bankruptcy companies only (by 10.87 pp), while in the sample of active companies the accuracy slightly decreased (by 0.45 pp). In SK model red., in which the reduction of the number of indicators was most significant, the predictive accuracy in the sample of active enterprises decreased by 0.76 pp, and in the sample of bankruptcy companies by 6.70 pp. In HU model reduced, the accuracy in active companies decreased by 10.44 pp; conversely, the accuracy increased by 0.34 pp in the sample of bankruptcy companies. The difference of the median accuracy of the models was also tested, see the Table 13.

T a b l e 13

Results of Hypothesis H_3 Testing

	No.	W-stat.	Z-stat	p-val.		No.	W-stat.	Z-stat	p-val.
CZ (A)***	999	197 989	5.6744	0.000000	CZ (B)***	1000	0	27.3930	0.000000
SK (A)***	999	197 650	5.7116	0.000000	SK (B)***	1000	15 059.5	25.7445	0.000000
PL (A)***	1000	213 002	4.0773	0.000050	PL (B)***	1000	4 968	26.8492	0.000000
HU (A)***	1000	0	27.39297	0.000000	HU (B)***	1000	19 7539	5.7699	0.000000

Source: Our processing on the basis of data from the Amadeus database.

According to the conclusion of the Wilcoxon test, the medians of accuracy between modified and reduced models both in active and bankruptcy companies are statistically different at the 1% level of significance. Hypothesis H_3 is therefore entirely confirmed in the sample of Polish companies, when reducing the number of indicators caused the increase of the accuracy of the model in the sample of both active and bankruptcy companies, while reducing the number of non-evaluated bankruptcy companies and with an unchanged number of non-evaluated active companies.

In the sample of Czech companies, this hypothesis was confirmed only partially, because although there was reduction of share of non-evaluated enterprises, the increasing of the accuracy was achieved only in the sample of bankruptcy companies.

The hypothesis was not confirmed in the sample of Slovak companies, as the reduction of the number of variables caused the reduction of accuracy in the sample of both active and bankruptcy companies and at the same time increasing the number of non-evaluated bankruptcy companies. The hypothesis was not quite confirmed in the sample of Hungarian companies: although there was a slight increase in accuracy in the sample of bankruptcy companies, the share of non-evaluated companies increased at the same time.

The significance of variables of the Altman model for bankruptcy prediction in American conditions was tested by Li (2012). Even though it is a test of the original model using market data, conclusions are worth mentioning. He found that statistically significant are only two indicators: the ratio net working capital to assets (X1), and the ratio of the market value of equity and total liabilities (X4). The importance of the ratio net working capital to assets (X1) was confirmed also by our research: this variable is statistically significant in all V4 countries. On the other hand, it occurs only in the Model PL red. – it was dropped from the other models due to redundancy, i.e. the information contained in this indicator was substituted by other indicators.

Shumway (2001) found that in addition to the already mentioned indicator X4, the return on assets indicator (EBIT/assets, X3) was the statistically significant indicator of the Altman model. According to our conclusions, this indicator is statistically significant at the significance level of 1% in all countries except Hungary. In the sample of Slovak companies, this indicator is the most important one.

Conclusion

Absence of a sufficient number of observations concerning bankruptcy companies tends to favouring the models created in different environments or even in another period against the creation of one's own models. But prediction capabilities of the models in another environment are not considered. The issue of transferability of models has been investigated on the example of the Altman model in a number of studies, and was therefore used also in our research.

Testing the accuracy of the Altman model in the sample of data of manufacturing industry companies of V4 countries showed that the original model worked with a statistically lower accuracy and with a high share of non-evaluated companies.

It was also found that by revaluating the weights of the model coefficients and grey-zone boundaries while maintaining the variables of the model, the discrimination capability of the model – i.e. the capability to correctly identify bankruptcy and prosperous companies – can be increased. This hypothesis was not confirmed only for bankruptcy companies of Hungary.

The model in general can be regarded as an optimum combination of variables with suitably set coefficients. In the next step, the effectiveness of the combination of the variables for the given environment was therefore investigated using the reverse discrimination method, during which the model was first compiled of all statistically significant variables, and then the insignificant variables were gradually eliminated (insignificant in the sense that the discrimination accuracy of the model did not drop by omitting the variable).

In the sample of the data examined it was proven that the Altman model variables have very different discrimination capability in different countries, and are therefore not transferable among environments. The above shows that for the particular environment it is necessary to find its own optimal combination of indicators and create original models. We consider this conclusion as very important also in the context with the rating evaluation of borrowers from individual countries.

Although the sample of companies studied does not constitute a population, using the bootstrap method allowed estimating the properties of the population. At the same time, this process enabled us to test whether the recalculation of model coefficients leads to a statistically significant increase of its accuracy.

References

AL KHATIB, K. – AL BZOUR, A. E. (2011): Predicting Corporate Bankruptcy of Jordanian Listed Companies: Using Altman and Kida Models. International Journal of Business and Management, *6*, No. 3, pp. 208 – 215.

ALTMAN, E. I. (1968): Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance, *23*, No. 4, pp. 589 – 609.

ALTMAN, E. I. (2000): Predicting Financial Distress of Companies: Revisiting the Z-score and Zeta® Models. [Online.] Available at: <http://pages.stern.nyu.edu/~ealtman/PredFnclDistr.pdf>.

ALTMAN, E. I. – HALDEMAN, R. G. – NARAYANAN, P. (1977): ZETA Analysis. A New Model to Identify Bankruptcy Risk of Corporations. Journal of Banking and Finance, *1*, No. 1, pp. 22 – 54.

BEAVER, W. H. (1966): Financial Ratios as Predictors of Failure. Journal of Accounting Research, *4*, No. 1, pp. 71 – 111.

COCHRAN, W. G. (1964): Sampling Techniques. 3rd ed. New York: John Wiley & Sons.

CRANEY, T. A. – SURLES, J. G. (2002): Model-Dependent Variance Inflation Factor Cutoff Values. Quality Engineering, *14*, No. 3, pp. 391 – 403.

CZILLINGOVÁ, J. – PETRUŠKA, I. – TKÁČ, M. (2012): Financial and Economic Analysis of Steel Industry by Multivariate Analysis. Ekonomický časopis/Journal of Economics, *60*, No. 4, pp. 388 – 405.

DEAKIN, E. B. (1972): A Discriminant Analysis of Predictors of Business Failure. Journal of Accounting Research, *10*, No. 1, pp. 167 – 179.

EL KHOURY, R. – AL BEAÏNO, R. (2014): Classifying Manufacturing Firms in Lebanon: An Application of Altman’s Model. Procedia – Social and Behavioral Sciences, *109*, No. 1, pp. 11 – 18.

GREENE, W. H. (2012): Econometric Analysis. 7th ed. Boston: Pearson, 1228 pp.

GRICE, J. S. – DUGAN, M. T. (2001): The Limitations of Bankruptcy Prediction Models: Some Cautions for the Researchers. Review of Quantitative Finance and Accounting, *17*, No. 2, pp. 151 – 166.

KARAS, M. – REŽŇÁKOVÁ, M. (2014): A Parametric or Nonparametric Approach for Creating a New Bankruptcy Prediction Model: The Evidence from the Czech Republic. International Journal of Mathematical Models and Methods in Applied Sciences, *8*, No. 1, pp. 214 – 223.

LI, J. (2012): Prediction of Corporate Bankruptcy from 2008 Through 2011. Journal of Accounting and Finance, *12*, No.1, pp. 31 – 41.

LI, J. – RAGOZAR, R. (2012): Application of the Z -Score Model with Consideration of Total Assets Volatility in Predicting Corporate Financial Failures from 2000 – 2010. Journal of Accounting and Finance, *12*, No. 2, pp. 11 – 19.

MACHEK, O. (2014): Long-term Predictive Ability of Bankruptcy Models in the Czech Republic: Evidence from 2007 – 2012. Central European Business Review, *3*, No. 2, pp. 14 – 17.

MANDRU, L. – KHASHMAN, A. – CARSTEA, – C. DAVID, N. – PATRASCU, L. (2010): The Diagnosis of Bankruptcy Risk Using Score Function. [Proceedings of the 9th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, pp. 83 – 87.]

MARTIN, D. (1977): Early Warning of Bank Failure: A Logit Regression Approach. Journal of Banking & Finance, *1*, No. 3, pp. 249 – 276.

MENČÍK, J. (2001): Simulace Monte Carlo, metoda bootstrap a spolehlivost výsledků. In: Spolehlivost konstrukcí. [2nd National conference.] Ostrava: Dům techniky, pp. 47 – 50. ISBN 80-02-01410-3.

NIEMANN, M. – SCHMIDT, J. H. – NEUKIRCHEN, M. (2008): Improving Performance of Corporate Rating Prediction Models by Reducing Financial Ratio Heterogeneity. Journal of Banking & Finance, *32*, No. 3, pp. 434 – 446.

OHLSON, J. A. (1980): Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research, *18*, No. 1, pp. 109 – 131.

PITROVÁ, K. (2011): Possibilities of the Altman Zeta Model Application to Czech Firms. Ekonomika a Management, 2011, No. 3, pp. 66 – 76.

SÁNCHEZ-LASHERAS, F. – DE ANDRÉS – J. LORCA, P. – DE COS JUEZ, F. J. (2012): A Hybrid Device for the Solution of Sampling Bias Problems in the Forecasting of Firms' Bankruptcy. Expert Systems with Applications, *39*, No. 1, pp. 7512 – 7523.

SATISH, Y. M. – JANAKIRAM, B. (2011): Turnaround Strategy Using Altman Model as a Tool in Solar Water Heater Industry in Karnataka. International Journal of Business and Management, *6*, No. 1, pp. 199 – 206.

SHUMWAY, T. (2001): Forecasting Bankruptcy More Accurately: A Simple Hazard Model, Journal of Business, *74,* No. 1, pp. 101 – 124.

TAFFLER, R. J. (1982): Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data. Journal of the Royal Statistical Society, *145*, No. 3, pp. 342 – 358.

TAM, K. – KIANG, M. (1992): Managerial Applications of Neural Networks: The Case of Bank Failure Prediction. Management Science, *38*, No. 7, pp. 926 – 947.

WHITTINGTON, G (1980): Some Basic Properties of Accounting Ratios. Journal of Business Finance and Accounting, *7*, No. 2, pp. 219 – 232.

WILCOXON, F. (1945): Individual Comparisons by Ranking Methods. Biometrics Bulletin, *1*, No. 6, pp. 80 – 83.

WU, Y. – GAUNT, C. – GRAY, S. (2010): A Comparison of Alternative Bankruptcy Prediction Models. Journal of Contemporary Accounting & Economics, *6*, No. 1, pp. 34 – 45.

ZMIJEWSKI, M. E. (1984): Methodological Issues Related to the Estimation of Financial Distress Prediction Models. Journal of Accounting Research, *22*, No. 1, pp. 59 – 82.