Karas, M., & Srbová, P. (2019). Predicting bankruptcy in construction business: Traditional model validation and formulation of a new model. *Journal of international studies, 12*(1), 283-296. doi:10.14254/2071-8330.2019/12-1/19

Journal of International Studies

Scientific Papers

Predicting bankruptcy in construction business: Traditional model validation and formulation of a new model

Michal Karas
Department of Finances, Faculty of Business and Management,
Brno University of Technology
Czech Republic
karas@fbm.vutbr.cz

Pavla Srbová
Department of Finances, Faculty of Business and Management,
Brno University of Technology
Czech Republic
xpsrbov00@vutbr.cz

Abstract: When predicting bankruptcy of a company basing on its financial statements, the line of business in which the company is operating plays a significant role in terms of prediction accuracy. This accuracy is particularly crucial to banks and businesses which realise sales mostly on credit. The failure to recognise a client's or business partner's financial difficulties or the threat of bankruptcy with sufficient accuracy could lead to significant losses. Bankruptcy prediction models are used for these purposes. Most of the models created have been dedicated to the branch of manufacturing, while the branch of construction is relatively neglected by the mainstream literature. Traditional bankruptcy prediction models cannot be used effectively due to specifics of construction business. The aim of this paper is to test the current accuracies of five selected bankruptcy models in predicting the bankruptcy of construction companies. An additional aim is to create a new model designed specifically for this branch. The research was conducted on the sample of Czech companies. The method of Receiver Operating Characteristic was applied as the measure of accuracy for testing the models. The model created during the course of this research achieved an accuracy higher by 3.6 to 8 percent than the traditional models tested.

Received: October, 2018
1st Revision: January, 2019
Accepted: March, 2019

DOI: 10.14254/2071-8330.2019/12-1/19

Keywords: bankruptcy prediction model, model accuracy, Czech Republic, construction business.

JEL Classification: G33, L74

1. INTRODUCTION

The purpose of a bankruptcy prediction model is to distinguish effectively, on the basis of financial data, between the companies that are financially healthy and companies experiencing financial difficulties. A great deal of attention is paid in literature to the question as to whether previously created models may still be used effectively if they were designed for different economies or industries. In general terms, this issue was the subject of the studies by Platt and Platt (1990), Grice and Dugan (2001), Niemann et al. (2008) and Wu, Gaunt and Gray (2010). Heo and Yang (2014) came to the conclusion that accuracy of such models decreases significantly if they are used in a different environment (or a different industry). Such arguments have motivated efforts to create new bankruptcy models that would better fit the currently prevailing market conditions.

The starting point in creating such a model is to find a limited group of variables that exhibit significant discriminatory power (distinguish between financially healthy companies and companies threatened by bankruptcy). These variables are often called predictors. The result of this approach is called a "reduced form model" in literature and represents a widely used way of creating a model (see Lin, Liang, Chen, 2011; Wang, Lee, 2008; Niemann et al., 2008; Tseng, Hu, 2010; Psillaki, Tsolas, Margaritis, 2009; Cheng, Chen, Fu, 2006). Doubt has been cast on such an approach by Scott (1981) who pointed out that there is a risk connected with every reduction of potential predictors based on their significance for a given case or, rather, its specific conditions or environment. According to this study, such reduction could result in lower robustness of the created model or, in other words, the found (created) group of predictors could be ineffective when applied to different companies, time periods or economic environments, generally under the conditions different from those which were used for deriving the model. Most of the previously created models were derived from data on manufacturing companies (see Grice, Dugan, 2001). According to some authors, these models are ineffective when used on companies from other fields. For example, Thomas, Wong and Zhang (2011) pointed out that there is a necessity for creating models for branches such as construction, as the existing models are inappropriate for this branch. According to Heo, Yang (2014), the specifics of construction companies is in high values of liquidity ratios, high debt and the fact that positive cash flow generated from contracts is concentrated only in their later stages. Sun, Liao, Li (2013) add some more specifics of this sector: The construction industry is a capital-intensive industry that requires long-term project periods and huge investment, and takes a long time to receive returns on the investment. It therefore has a different capital structure as compared to other industries and the same criteria used for other industries cannot be applied to evaluate its financial risk effectively (Sun, Liao, Li, 2013 in: Heo, Yang, 2014). The given opinion is also confirmed by another study (Barrie, Paulson, 1992 in: Tserng et al., 2014) as follows: "due to the distinctive operational behaviours of the construction industry, its financial characteristics also differ from other industries". Sun, Liao, Li (2013) also add that for such reasons *"the same criteria used for other industries cannot be applied to effectively evaluate its financial risk"*. Lee, Choi (2013) compared the accuracy of their model based exclusively on the data from construction companies with a similar model based on the data on companies from different industries. The model designed specifically for construction companies achieved a classification accuracy 6–12% higher as compared to the model created on the data of companies from different industries. The authors believe that the accuracy of the model would be even higher if predictors specific to the construction industry were used.

The aim of this paper is to test a set of five traditional bankruptcy models on the sample of small and medium-sized Czech construction companies. Moreover, the paper also aims to formulate a new model which incorporates a set of significant variables found to fit construction businesses and to compare its accuracy with the accuracies of the tested models.

2. LITERATURE REVIEW

Several studies have been concerned with the classification accuracy of famous bankruptcy models (e.g. Altman Z-score, Ohlson O-score and others), while the possibilities of re-estimating these models on a more up-to-date data set have also been analysed in order to obtain more accurate models.

For example, the study by Begley, Ming and Watts (1996) studied the classification errors of the Altman (1968) and Ohlson (1980) models. The authors re-estimate the models' coefficients using the COMPUSTAT data on companies listed on the NYSE and AMEX during the period 1980–1989. They focus only on large companies (asset value over 10 mil. USD). Their final sample consists of 165 companies which declared bankruptcy and 3,300 randomly selected non-bankrupt firms. By analysing the dataset from the 1980s, the authors of this study come to the conclusion that the combined error rate increased if the original coefficients are applied. The possible cause of this is that in both the Altman and Ohlson models *"a leverage ratio plays an important role, while during the 1980s there was an increasing acceptance of relatively high corporate debt levels. As a result, a given level of debt in the 1980s may not be associated with the same likelihood of bankruptcy as it was in the pre-1980 period."* (see Begley, Ming, Watts, 1996). These authors come to the conclusion that the Altman and Ohlson models do not perform well in a more recent time period (1980s data), even after re-estimating the model, as the combined error was not reduced.

Grice and Dugan (2003) analysed the stability of the coefficients of the Ohlson and Zmijewski models. This was performed by re-estimating the coefficients of the models and comparing them with the original values. The authors of this study used a Compustat database of North American and Canadian companies. Their hold-out sample contained 1,024 companies (183 distressed, 841 non-distressed) for the Zmijewski model and 1,043 (154 distressed, 889 non-distressed) for the Ohlson model. They come to the conclusion that the coefficients of the Zmijewski model are not stable over time periods and are not sensitive to industry classification. In the case of the Ohlson model, the result regarding sensitivity to time periods was similar, moreover the authors also confirmed the sensitivity to industry classification. Grice and Dugan (2003) also come to the conclusion that the relationship between financial ratios and financial distress changes over time. The original overall accuracy of the Zmijewski and Ohlson models for the original sample (originally reported by the authors of the models) were 98.2 and 96.4 %. The results of testing the model accuracies for a hold-out sample (see Grice, Dugan, 2003) indicate significantly lower results of 81.3 % for the Zmijewski model and 39.8 % for the Ohlson model. After re-estimating the coefficient, the accuracies range from 85.7 to 86.1 % in the case of the Zmijewski model and 88.1 to 88.7 % in the case of the Ohlson model.

The research by Singh and Mishra (2016) focused on a similar question. Specifically, one of the aims of their research was to compare the accuracies of the Altman (1968 version), Zmijewski and Ohlson models in two ways. Firstly, when the original coefficients are applied, and secondly, when the re-estimated coefficients are applied. Their study was based on a research sample of 208 Indian manufacturing firms (130 distressed and non-distressed used as an estimation sample, 78 used as a hold-out sample). The distressed firms in the sample registered were financially sick during the period 2006 to 2014. Singh and Mishra (2016) confirmed the results of Grice and Ingram (2001) regarding the instability of the given models' coefficients, claiming these coefficients unstable and sensitive to time periods, though in this case for Indian manufacturing firms. The predictive accuracy of the Altman model (in the case of the hold-out sample) was 61.538 %, while the accuracy rose to 88.462 % after re-estimation of the coefficients. The corresponding numbers were 79.487 % when using the original coefficients and 76.923 % after the coefficients were re-estimated in the case of the Zmijewski model. The overall accuracy was 64.103 % on the estimation sample and 89.744 % after re-estimation in the case of the Ohlson model. Re-estimation was performed using the original methodology.

It is, however, also possible to reach an opposite conclusion. Altman et al. (2017) carried out an extensive study on the classification performance of the Altman Z-score model in predicting bankruptcy. The performances of the models were analysed for firms from 28 European and 3 non-European (China, Columbia and the United States) countries. The time period under analysis was from 2002 to 2010. The sample consists of 2,602,563 non-failed and 38,215 failed firms. The authors come to the conclusion that the Z-score with the original coefficients performs very well in the international context. The topic of re-estimating the model coefficients was also addressed, though the conclusion of the study was opposite or, rather, does not support the results of the previously mentioned studies (e.g. Grice and Ingram, 2001), as the *"re-estimation of the coefficients using MDA only marginally improved the classification performance, or, put differently, showing that the original coefficients are extremely robust across countries and over time*" (see Altman et al., 2017).

3. SAMPLE AND METHODS USED

The research sample consists of 4,420 small and medium-sized companies (4,243 non-failed, 177 failed) which operated in the construction business in the Czech Republic. In terms of the population, with 172,283 SMEs operating in the construction business in the Czech Republic in 2016 (see mpo.cz), the research sample covers 2.56 % of the population. We focus on construction companies for several reasons. Firstly, we agree that the construction branch is usually neglected by the main stream of literature on bankruptcy prediction. Secondly, we agree with Heo, Yang (2014) and Thomas Ng, Wong and Zhang (2011) who pointed out that the existing models are unsuitable for predicting bankruptcy in the construction industry. During the course of this research, we test a set of five traditional models – these models were created for different environments (countries or branches) and by using different methods. The bankrupt firms in the sample declared bankruptcy during 2011 and 2015, while we analysed the last five years of the bankrupt companies, i.e. the period under investigation is from 2006 to 2015. We focus on a five-year period prior to bankruptcy as, according to the literature (e.g. Beaver et al., 2005), the financial ratios have a predictive power up to five years prior to bankruptcy. The traditional models under investigation were the following: the revised Z-score, Springate's model, the Zmijewski model, Taffler's model and the IN05 model. A short description of the models follows.

3.1. Revised Z-score Model (see Altman, 2000)

The revised Z-score represents the original Z-score model (see Altman, 1968) adapted for non-listed companies (see Altman, 1983). The formula of the model is as follows (see Altman and Sabato, 2013):

$$Z = 0.717 * NWC/TA + 0.847 * RE/TA + 3.107 * EBIT/TA + 0.420 * BVE/TA + 0.998 * S/TA$$

where: *NWC – net working capital (= current assets-current liabilities), TA – total assets, RE – retained earnings, EBIT – earnings before interest and taxes, BVE – book value of equity, S – sales.*

The grey zone interval is (1.23; 2.9). For $Z < 1.23$ the company is classified by the model as threatened by bankruptcy; for $Z > 2.9$ the company is classified as not threatened by bankruptcy, i.e. financially healthy. Altman and Sabato (2013) tested the model on a sample of US SMEs over the period from 1994 to 2002. The resulting overall accuracy of the model was 68 %, while type I error (the percentage of bankrupt firms classified as non-bankrupt) was 25.81 %.

3.2. Springate's Model

This model was derived in 1978 by using the method of discrimination analysis (see Imanzadeh et al., 2011). The model is inspired by Altman's model, but adjusted to the conditions of the Canadian market.

The accuracy of the model, at the time it was derived, was 92.5 percent. The model could be described by the following formula:

$S = 1.3 * NWC/TA + 3.07 * EBIT/TA + 0.66 * EBT/CL + 0.4 * S/TA$

where: *NWC – net working capital (= current assets-current liabilities), TA – total assets, EBIT – earnings before interest and taxes, EBT – earnings before taxes, CL – current liabilities, S – sales.*

The model interpretation is as follows: if $S < 0.862$ the given company is threatened by bankruptcy.

3.3. The Zmijewski Model

Mark Zmijewski published his model in 1984. He derived his model by using probit analysis (see Zmijewski, 1984). The model could be described by the following formula:

$Z = – 4.3 – 4.5 * EAT/TA + 5.7 * TL/TA + 0.004 * CA/CL$

where: *EAT – earnings after taxes, TA – total assets, TL – total liabilities, CL – current liabilities, CA – current assets.*

The model provides results in the form of a probability of bankruptcy (P). This probability is given by the formula: $P = 1 / (1+\exp(-Z))$. For $P > 0.5$, the company is considered threatened by bankruptcy.

3.4. Taffler's Model

The model was published in 1977. Its construction is based on Altman's model and the model is also based on the method of discrimination analysis (see Taffler, 1982). The model could be described by the following formula:

$T = 0.53 * EBT/CL + 0.13 * CA/CL + 0.18 * CL/TA + 0.16 * S/TA$

where: *EBT – earnings before taxes, CL – current liabilities, CA – current assets, TA – total assets, S – sales.*

The grey zone interval is (0.2; 0.3). For $T < 0.2$, the company is classified by the model as threatened by bankruptcy, for $T > 0.3$ it is classified as not threatened by bankruptcy, i.e. financially healthy.

3.5. The IN05 Model

The IN05 is the only one of the tested models developed specifically for Czech companies (see Neumaier and Neumaierová, 2005). The formula of the model is as follows:

$IN05 = 0.13 * TA/TL + 0.04 * EBIT/IE + 3.97 * EBIT/TA + 0.21 * OR/TA + 0.09 * CA/CL$

where: *TL – total liabilities, OR – operating revenue, CA – current assets, CL – current liabilities.*

The grey zone interval is (0.9; 1.6). For $IN05 < 0.9$, the company is classified by the model as threatened by bankruptcy; for $IN05 > 1.6$ it is classified as not threatened by bankruptcy, i.e. financially healthy. For $0.9 < IN05 < 1.6$ the predicted fate of the analysed company is not clear (the so-called grey zone). At the time at which the model was created, its authors summarised its prediction ability as follows (Neumaier and Neumaierová, 2005, p. 146): "*If the index value for a given company falls beneath the lower limit, there is a 9 % probability that the company is headed for bankruptcy and a probability of 76 % that it will not create value. A company in the grey zone has a practically 50 % probability of bankruptcy and a 70 % probability of creating value. A company above the upper limit will have a 92 % probability of not going bankrupt and a 95 % probability of creating value*".

4. RESULTS OF EVALUATING MODEL ACCURACY

The following table shows descriptive statistics of the analysed sample. The values are shown for the period of a year prior to bankruptcy (referred to as the T+1 period). The character of the subsample is differentiated by the abbreviations "A" for active (non-defaulted) companies and "B" for companies prior to bankruptcy (see column "Bankrupt").

Table 1

Descriptive statistics of the variables under analysis

	Bankrupt	Valid N	Mean	Median	Minimum	Maximum	Std. Dev.
BVE/TL	A	4174	3.06315	0.620806	-78.81	1949.00	33.6327
	B	155	0.8837	0.019829	-1.00	77.000	8.1004
CA/CL	A	4166	5.72174	1.568201	-59.85	3526.00	71.7540
	B	154	2.0076	0.895561	0.00	78.000	8.3716
CL/TA	A	4180	0.50959	0.461252	-2.96	8.31	0.4253
	B	156	65.0784	0.875958	0.00	9421.000	754.9961
EAT/TA	A	4180	0.06892	0.042944	-10.34	2.96	0.2513
	B	156	0.7595	-0.001517	-4.40	136.500	10.9533
EBIT/IE	A	4184	56.71145	9.000000	-9852.00	23472.00	621.2808
	B	177	-63.4502	7.718391	-9607.00	180.343	728.7063
EBIT/TA	A	4180	0.09803	0.062618	-1.65	2.97	0.2012
	B	156	-0.0836	0.000000	-4.31	3.300	0.5830
EBT/CL	A	4166	0.55945	0.115441	-51.00	293.00	5.5319
	B	154	0.4173	-0.000375	-2.25	77.000	6.2238
NWC/TA	A	4180	0.14867	0.086932	-6.18	1.00	0.2480
	B	156	-60.3987	0.015307	-9421.00	5.641	754.2838
OR/TA	A	4180	2.69027	2.130641	-33.69	303.90	5.4913
	B	156	1.9115	1.452503	-0.02	16.683	2.2538
RE/TA	A	4180	0.24644	0.227210	-9.38	10.68	0.4932
	B	156	-66.9375	0.000000	-9606.50	0.781	770.8610
S/TA	A	4180	2.56704	2.065575	-33.55	303.90	5.2736
	B	156	1.8154	1.352583	-0.02	18.133	2.2669
TA/TL	A	4174	4.06661	1.621280	-77.81	1950.00	33.6349
	B	155	1.8802	1.019829	0.00	78.000	8.1010
TL/TA	A	4180	0.62480	0.612234	-2.96	9.06	0.4600
	B	156	66.7065	0.978730	0.00	9421.000	755.9135

Source: Own calculation based on data from the Amadeus Database

The accuracies of the models were evaluated in two ways. First, as a percentage of correctly classified bankrupt and non-bankrupt companies, with respect to the original setting of the cut-off score (or generally grey zone borders). Second, by using ROC curves and the corresponding Area Under Curve (AUC) value, regardless of the setting of the cut-off score.

The first evaluated model is Altman's model.

Table 2

Results of testing Altman's model

Model	Category	T+1	T+2	T+3	T+4	T+5
Altman	Non-failed (%)	63.13	56.56	53.42	53.97	53.2
	Failed (%)	46.45	41.4	29.53	18.84	49.23
	Grey zone (non-failed) (%)	29.23	33.66	34.93	34.44	35.79
	Grey zone (failed) (%)	35.48	43.31	47.65	42.75	30

Source: Own calculation based on data from the Amadeus Database

The percentage of correctly classified non-failed companies in the period t+1 is 63.13 %, while the corresponding number for failed companies is just 46.45 %. For the more distant periods prior to bankruptcy the numbers are even lower, just 53.2 % for the t+5 period in the case of non-failed companies, while the number is just 49.23 % in the case of failed companies for the same period. A large proportion of the analysed companies end up in the grey zone; on average 33.6 % of non-failed companies and 39.8 % of failed companies.

The next model under analysis is Springate's model.

Table 3

Results of testing Springate's model

Model	Category	T+1	T+2	T+3	T+4	T+5
Springate	Non-failed	78.3	73.97	68.09	66.44	70.47
	Failed	62.34	61.15	51.01	39.86	35.38

Source: Own calculation based on data from the Amadeus Database

The accuracies are significantly higher in the case of Springate's model; for the period t+1 the percentage of correctly classified non-failed companies attained the value of 78.3 %, while in the case of failed companies the number is again higher than in the case of Altman's model, specifically 62.34 %. The possible reason for this is that Springate's model does not apply the grey zone interval for evaluation of the results. On average, the model leads to correct classification in the case of 71.4 % of non-failed companies and 49.9 % of failed companies in the period up to five years prior to bankruptcy. So far, the analysed models were derived using the method of discrimination analysis.

The third model under investigation is Zmijewski's model which was derived using the probit method which applies a probabilistic approach.

Table 4

Results of testing Zmijewski's model

Model	Category	T+1	T+2	T+3	T+4	T+5
Zmijewski	Non-failed	68.58	64.56	63.23	64.51	66.59
	Failed	85.71	80.25	74.5	64.49	63.08

Source: Own calculation based on data from the Amadeus Database

The percentage of correctly classified failed companies is, in contrast to the previous model, higher than the percentage of correctly classified non-failed companies. Namely, for the t+1 period, the model leads to correct classification in the case of 85.71 % of failed companies and 68.58 % in the case of non-failed companies. On average, the model leads to correct classification in the case of 65.4 % of non-failed companies and 73.6 % of failed companies in the period up to five years prior to bankruptcy.

The testing of Taffler's model follows; the model was derived using linear discrimination analysis and applies the grey zone concept.

Table 5

Results of testing Taffler's model

Model	Category	T+1	T+2	T+3	T+4	T+5
Taffler	Non-failed	94.22	91.7	90.03	89.59	91.33
	Failed	9.74	12.1	6.71	5.07	5.38
	Grey zone (non-failed)	2.28	3.49	3.42	3.59	3.48
	Grey zone (failed)	9.09	8.28	7.38	4.35	3.85

Source: Own calculation based on data from the Amadeus Database

Analysis of the percentage of correctly classified companies reveals a quite evident disproportion in the figures. Application of the model leads to a very high percentage of correctly classified observations in the case of non-failed companies (namely 94.22 % for the t+1 period), while the corresponding number was only 9.74 % (for the t+1 period) in the case of failed companies. The situation is rather analogous on average, with the model leading to correct classification in the case of 91.3 % of non-failed companies and 7.8 percent of failed companies in the period up to five years prior to bankruptcy. As the grey zone interval is relatively short, the percentage of observations which ended up in the grey zone is relatively low, i.e. 3.3 % of non-failed companies and 6.6 % of failed companies.

The last model subjected to analysis was the Czech model IN05.

Table 6

Results of testing the IN05 model

Model	Category	T+1	T+2	T+3	T+4	T+5
IN05	Non-failed	50.36	43.29	37.91	37.02	39.34
	Failed	68.83	62.42	54.36	47.83	38.46
	Grey zone (non-failed)	32.14	34.36	34.38	33.89	36.02
	Grey zone (failed)	21.43	25.48	33.56	28.26	39.23

Source: Own calculation based on data from the Amadeus Database

The IN05 model was derived specifically for Czech businesses. The percentages of correctly classified companies are rather comparable to those which resulted from testing Altman's model. However, the percentage, in the case of failed companies, is higher, being 68.83 % for the t+1 period (it was 46.45 % in the case of Altman's model). The equivalent number in the case of non-failed companies is 50.36 % which is, in contrast, lower (it was 63.13 % for Altman's model). When comparing the percentage of failed companies in the grey zone interval, the number is seen to be rather lower, specifically 29.6 %, in the case of the IN05 model (39.8 % in the case of Altman's model).

The above results of testing model accuracies were obtained with respect to the original cut-off score values. The results showed that the accuracies are lower as compared to the original values at the time at which the models were derived. A possible cause of this is a shift of the cut-off score or grey zone interval. Therefore, the next step of testing the models is to apply ROC curves.

Table 7

Results of testing the models

Tested model	AUC	Std. Error[a]	Asymp. Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
Altman	0.807	0.021	0.000000	0.767	0.848
Springate	0.772	0.021	0.000000	0.731	0.814
Taffler	0.693	0.023	0.000000	0.648	0.739
In05	0.809	0.019	0.000000	0.772	0.847
Zmijewski	0.839	0.017	0.000000	0.806	0.871

Notes: a. Under the nonparametric assumption, b. Null hypothesis: true area = 0.5. *Source*: Own calculation based on data from the Amadeus Database

According to the Area Under Curve (AUC), all models provide the user with a result better than a random choice would, as all the AUC values are higher than 0.5, while all the found values of AUC are significant at the 1% level. The highest score was achieved by the application of Zmijewski's model (AUC of 0.839), followed by the IN05 model (with an AUC of 0.809), with Altman's model attaining a highly comparable value of 0.807. In contrast, Springate's model was related with the second lowest score (AUC of 0.772), while the lowest score was attained by application of Taffler's model. The highest AUC value of the analysed model (the value of Zmijewski's model) will further serve as a reference value for the purposes of testing the newly created model.

5. CREATING A NEW MODEL

In this phase, we focused on whether better results could be achieved if a new model, specifically derived for construction companies, is created. A list of 35 potential predictors was drawn up on the basis of a review of the literature.

Table 8

The list of potential predictors

No.	Variable	No.	Variable
1	cash flow/sales	19	net income/operating revenue*
2	cash flow/total assets*	20	net income/total assets
3	cash flow/total liabilities*	21	operating revenue/current assets*
4	current assets/total liabilities	22	operating revenue/current liabilities*
5	current assets/current liabilities	23	operating revenue/fixed assets
6	current assets/total assets*	24	operating revenue/total assets
7	current liabilities/sales	25	operating revenue/total liabilities
8	current assets/sales	26	profit margin (3-year average)
9	EBIT/interest paid	27	retained earnings/total assets
10	EBIT/total assets	28	sales/total assets*
11	EBITDA/interest paid*	29	shareholder funds/total liabilities
12	EBITDA/total liabilities*	30	tangible fixed assets/total assets
13	EBT/current liabilities*	31	total assets/total liabilities
14	EBT/operating revenue*	32	total liabilities/EBITDA
15	intangible fixed assets/total assets	33	total liabilities/total assets*
16	net income/capital*	34	working capital/total assets
17	net income/current assets*	35	working capital/sales*
18	net income/fixed assets*		

Source: Beaver, 1966; Altman, 1968; Deakin, 1972; Ohlson, 1980; Ding et al., 2008; Wang, Lee, 2008; Niemann et al., 2008; Beaver et al., 2005; Tseng, Hu, 2010; Psillaki, Tsolas, Margaritis, 2009-
Note: *variables removed from the initial sample due to strong correlation

As the initial set of financial ratios was gathered from different sources, it was necessary to check for similarly defined variables which would exhibit a strong correlation. It was found by correlation analysis that 16 of the 35 analysed ratios need to be excluded from the sample due to strong correlation (Spearman's rank coefficient higher than 0.9). Stepwise discrimination analysis was applied (both forward selection and backward elimination) to find a significant set of variables for deriving the model. The backward elimination method leads to a more significant model. The Wilk's lambda of the resultant model attained a value of 0.62387 (which is equivalent to $F(4; 266) = 40.093$ for which the p-value is $p < 0.001$). The variables of the model and their contribution to the discriminant power of the model are listed below. All the variables are significant at the 1% level.

Table 9

The list of potential predictors

Variable	Wilk's lambda	Parc. lambda	F to rem	p-value	Toler.	1-toler. (R^2)
Net income/total assets***	0.692	0.902	28.925	0.000	0.105	0.895
EBIT/total assets***	0.65	0.959	11.341	0.001	0.104	0.896
Retained earnings/total assets***	0.675	0.924	21.757	0.000	0.917	0.083
Current liabilities/sales***	0.665	0.938	17.666	0.000	0.917	0.083

Note: ***significant at the 1% level.
Source: Own calculation based on data from the Amadeus Database

The final model consists of four variables – the return on assets based on EAT (net income/total assets), followed by the return on assets based on EBIT (EBIT/total assets), the past profitability of assets (retained earnings/total assets) and the liabilities turnover based on sales (current liabilities/sales). The discrimination function of the model could be described by the following formula:

$$M = 20.8 * EAT/TA - 12.054 * EBIT/TA + 3.116 * RE/TA - 2.399 * CL/S$$

where: *EAT – earnings after taxes, TA – total assets, EBIT – earnings before interest and taxes, RE – retained earnings, CL – current liabilities, CA – current assets, S – sales.*

The model's interpretation is as follows: for $M > -0.6$ the company is evaluated as threatened by failure (bankruptcy), otherwise it is evaluated as not threatened by failure (bankruptcy). The presented model was tested both on the learning and the test sample for the period t+1, the results are listed below.

Table 10

The percentage of correctly classified companies – the created model

Category/sample	Learning sample (%)	Test sample (%)
Non-failed	84.62	77.28
Failed	82.3	85.71

Source: Own calculation based on data from the Amadeus Database

When comparing the results with the results of testing the Zmijewski model, which achieved the highest accuracies of the analysed models, we can come to the conclusion that the created model exhibits a comparable percentage on the sample of failed companies (82.3 on the test sample and 85.71 on the test sample, while the Zmijewski model recorded 85.71). In the case of non-failed companies, the results are more favourable for the created model as the percentages are higher (train sample 84.62 and test sample 77.28, while the Zmijewski model recorded 85.71). The next step is to compare the accuracies in terms of AUC.

6. DISCUSSION

When viewing bankruptcy prediction models in the context of the environment specifics (current market or branch conditions) it is possible to deduce that these models are environment or branch specific which results in their lower prediction accuracies when applied under alternative conditions (see, for example, Wu, Gaunt and Gray, 2010 or Heo, Yang, 2014). As a result, much scientific effort is expended in deriving new models which would better fit current conditions. Nevertheless, there are branches which are relatively neglected by the literature, for example construction, as has been pointed out by Thomas, Wong and Zhang (2011). The aim of the paper was to contribute a new bankruptcy model to the current literature which would better fit the conditions of construction businesses. The presented model was derived using linear discrimination analysis which is, according to Aziz, Dar (2006), the most frequently applied method.

There are four variables in the model. Three of them are measures of profitability, either current (net income/total assets or EBIT/total assets) or past (retained earnings/total assets). Profitability ratios play a vital role in bankruptcy prediction in general. An explanation can be found in Altman (1968, p. 595): *"Since a firm's ultimate existence is based on the earning power of its assets, this ratio appears to be particularly appropriate for studies dealing with corporate failure. Furthermore, insolvency in a bankruptcy sense occurs when the total liabilities exceed a fair valuation of the firm's assets with value determined by the earning power of the assets"*.

The most frequently used profitability indicator is the return on assets (based on EBIT). This indicator is often mentioned as the most significant predictor of most Altman models (for example, Altman, 1968, 1973, 1977, 1983) or other authors' models, e.g. Li, Sun (2009), Mileris, Boguslaukas (2011) or Psillaki, Tsolas, Margaritis (2009). In his paper, Shumway (2001) suggests that tests of the significance of bankruptcy predictors are often biased as the methods used do not consider the time factor. In this context he re-evaluated several typically applied predictors and found that only two of them remain significant, of which one was the return on assets (based on EBIT).

Moreover, speaking about the profitability of construction businesses in the Czech Republic, Spička (2013) concluded that typical bankruptcy manifestations in construction companies in the Czech Republic included high indebtedness due to current liabilities, low labour productivity and a negative return on assets. Working capital management is particularly problematic of SME financing, in which payment discipline plays a vital role, and this problem often turns into a direct threat of bankruptcy (Ključnikov, Kozubíková, Sopková, 2017). The factor of current liabilities is incorporated in the created model in the form of a ratio of current liabilities and sales; moreover, the return on assets is also a significant part of the model.

The profitability of assets is also represented in the model by the return on assets (based on net income); the difference between these two versions of return on assets is mainly in including financial cost or not. The version of return on assets based on net income is again part of many models – of the analysed models it is part of the Zmijewski model, though it is also part of many others, e.g. Beaver (1966), Deakin (1972), Ohlson (1980), Cheng, Chen, Fu (2006), Grunert et al. (2004), Lin (2009) and Wang, Lee (2008).

The last factor included in the model is the ratio of retained earnings and total assets. This ratio represents a frequently used predictor of bankruptcy and was, for example, part of Altman's models (Altman, 1968; Altman, 2000; Altman, Sabato, 2013). Or other authors' models as Fulmer H-score (Fulmer et al., 1984; Kalupa, 2001) or specifically for construction business a CART based model (Karas, Režňáková, 2017). According to Altman (1968) this ratio measures cumulative or rather past profitability and implicitly the age of the company. He also added that past profitability is more important in terms of the risk of bankruptcy than the current asset profitability (measured by EBIT/total assets).

7. CONCLUSION

Whether previously created bankruptcy models could still be used for effective bankruptcy prediction under current market conditions or whether they are applicable to an alternative branch is a frequently discussed topic. According to Lee, Choi (2013), based on the example of construction companies, a specially designed model could achieve a 6–12 % higher classification accuracy compared to a model created on data on companies from different industries. The authors believe that the accuracy of the model would be even higher if predictors specific to the construction industry were used. Inspired by this idea, a new bankruptcy model was created, while its predictors were selected to fit the specifics of the construction business. We found that the created model achieved a higher accuracy than the models analysed, as the analysed models were not created exclusively on data on construction companies. We can say that an increase of accuracy could be obtained by creating a branch-specific model. The resultant difference in accuracy (when compared to the results of Zmijewski's model and the results of out-of-sample testing of the created model) is between 3.6 and 8 % in terms of AUC.

ACKNOWLEDGEMENT

The authors are grateful to the Internal Grant Agency of University No.: FP-S-18-5234 "Predictive models in finance: analysis of factors and predictions of bankruptcy, company performance and value" for financial support to carry out this research.

REFERENCES

Altman, E.I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial Distress Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model. *Journal of International Financial Management & Accounting*, 28, 131-171. doi:10.1111/jifm.12053

Altman, E. I, Haldeman, R. G., & Naraynan, P. (1977). ZETA Analysis. A new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance*, 1, 22-54. doi:10.1016/0378-4266(77)90017-6.

Altman, E. I. (1973). Predicting Railroad Bankruptcies in America. *Bell Journal of Economics*, 4(1), 184-211. doi:10.2307/3003144.

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance,* 23(4), 589-609. doi:10.1111/j.1540-6261.1968.tb00843.x.

Altman, E. I. (1983). *Corporate financial distress: A complete guide to predicting, avoiding and dealing with bankruptcy*. New York: John Wiley and Sons.

Altman, E. I. (2000). Predicting financial distress of companies: Revisiting the Z-score and Zeta® models. *Handbook of Research Methods and Applications in Empirical Finance*, 5. doi:10.4337/9780857936097.00027.

Altman, E. I., & Sabato, G. (2013). Modeling credit risk for SMEs: Evidence from US market. *Managing and measuring risk*, 251-279. doi: 10.1142/9789814417501_0009.

Aziz, M., & Dar, H. (2006). Predicting corporate bankruptcy: where we stand? *Corporate Governance*, 6, 18-33. doi:10.1108/14720700610649436.

Barrie, D. S., & Paulson, B. C. (1992). *Professional construction management: Including CM, design-construct, and general contracting*. Hightstown: McGraw-Hill.

Beaver, W. H. (1966). Financial Ratios as predictors of Failure. *Journal of Accounting Research,* 4, 71-111. doi:10.2307/2490171.

Beaver, W. H., McNichols, M. F., & Rhie, J. W. (2005). Have Financial Statements Become Less Informative? Evidence from the Ability of Financial Ratios to Predict Bankruptcy. *Review of Accounting Studies*, 10, 93–122. doi:10.1007/s11142-004-6341-9.

Begley, J., Ming, J. & Watts, S. (1996). Bankruptcy Classification Errors in the 1980s: An Empirical Analysis of Altman's and Ohlson's Models. Review of Accounting Studies, 1, 267-284. doi:10.1007/BF00570833

Cheng, C. B., Chen, C., L., & Fu, C. J. (2006). Financial Distress Prediction by a Radial Basis Function Network with Logit Analysis Learning. *Computers and mathematics with applications*, 51, 579-588. doi:10.1016/j.camwa.2005.07.016.

Deakin, E. B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10(1), 167-179. doi:10.2307/2490225.

Ding, Y., Song, X., & Zen, Y. (2008). Forecasting financial condition of Chinese listed companies based on support vector machine. *Expert Systems with Applications,* 34, 3081–3089. doi:10.1016/j.eswa.2007.06.037.

Fulmer J. G, Moon J. E., Gavin T. A., & Erwin M. J. (1984). A bankruptcy classification model for small firms. *Journal of Commercial Bank Lending*, 66(11), 25–37.

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, 151-166. doi:10.1023/A:1017973604789.

Grice, J. S., & Dugan, M. T. (2003). Re-estimations of the Zmijewski and Ohlson bankruptcy Prediction models. *Advances in Accounting, 20, 77-93.* doi: 10.1016/S0882-6110(03)20004-3

Grunert, J., Norden, L., & Weber, M. (2004). The Role of Non-financial Factors in Internal Credit Ratings. *Journal of Banking & Finance,* 29(2), 509–531. doi:10.1016/j.jbankfin.2004.05.017.

Heo, J., & Yang, J. Y. (2014). AdaBoost based bankruptcy forecasting of Korean construction companies. *Applied soft computing*, 24, 494-499. doi:10.1016/j.asoc.2014.08.009.

Imanzadeh, P., Maran-Jouri, M., & Sepehri, P. (2011). A Study of the Application of Springate and Zmijewski Bankruptcy Prediction Models in Firms Accepted in Tehran Stock Exchange. *Australian Journal of Basic and Applied Sciences*, 5(11), 1546-1550.

Kalupa, L. (2001). Model H-score w badaniu kondycji małych i srednich przedsiębiorstw. Ruch prawniczy, ekonomiczny i socjologiczny. 63(4), 207–220. ISSN 0035-9629.

Karas, M., & Režňáková, M. (2017). Predicting the Bankruptcy of Construction Companies: A CART- Based Model. *Engineering Economics,* 28(2), 145-154. doi:10.5755/j01.ee.28.2.16353.

Ključnikov, A., Kozubíková, L., & Sopková, G. (2017). The Payment Discipline of Small and Medium-sized Enterprises. *Journal of Competitiveness,* 9(2), 45-61. doi:10.7441/joc.2017.02.04.

Lee, S., & Choi, W. S. (2013). A multi-industry bankruptcy prediction model using back-propagation neural network and multivariate discriminant analysis. *Expert Systems with Applications*, 40(8), 2941–2946. doi:10.1016/j.eswa.2012.12.009.

Lin, S. L. (2009). A new two-stage hybrid approach of credit risk in banking industry. *Expert Systems with Applications*, 36(4), 8333-8341. doi:10.1016/j.eswa.2008.10.015.

Li, H., & Sun, J. (2009). Predicting business failure using multiple case-based reasoning combine with support vector machine. *Expert Systems with Applications*, 36(6), 10085-10096. doi:10.1016/j.eswa.2009.01.013.

Lin, F., Liang, & D., Chen, E. (2011). Financial ratio selection for business crisis prediction. *Expert Systems with Applications,* 38, 15094-15102. doi:10.1016/j.eswa.2011.05.035.

Milleris, R., Boguslaukas, V. (2011). Credit Risk Estimation Model Development Process: Main Steps and Model Improvement. *Inzinerine Ekonomika-Engineering Economics*, 22(2), 126-133. ISSN 1392-2785.

Neumaier, I., & Neumaierová, I. (2005). Index IN 05. In: *Sborník příspěvků mezinárodní vědecké konference „Evropské finanční systémy“.* Brno: Ekonomicko-správní fakulta Masarykovy university v Brně, 143-148. ISBN 80-210-3753-9.

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, 434–446. doi:10.1016/j.jbankfin.2007.05.015.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1). doi:10.2307/2490395.

Platt, D. H., & Platt, M. B. (1990). Development of a Class of Stable Predictive Variables: The Case of Bankruptcy Prediction. *Journal of Business Finance & Accounting,* 17(1), 31-51. doi:10.1111/j.1468-5957.1990.tb00548.x.

Psillaki, M. Tsolas, I. T., & Margaritis, M. (2010). Evaluation of credit risk based on firm performance. *European Journal of Operational Research,* 201, 873-881. doi:10.1016/j.ejor.2009.03.032.

Scott, J. (1981). The probability of bankruptcy: a comparison of empirical predictions and theoretical models. *Journal of Banking & Finance*, 5, 317-44. doi:10.1016/0378-4266(81)90029-7.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, 74(1), 101-24.doi:10.1086/209665.

Singh, B. P., & Mishra, A. K. (2016). Re-estimation and comparisons of alternative accounting based bankruptcy prediction models for Indian companies. *Financial Innovation,* 2(6), 1-28. Doi:10.1186/s40854-016-0026-9

Sun, J. Liao, B., & Li, H. (2013). AdaBoost and bagging ensemble approaches with neuralnetwork as base learner for financial distress prediction of chinese construction and real estate companies. *Recent Patents on Computer Science*, 6(1), 47–59. doi:10.2174/2213275911306010007.

Taffler, R. J. (1982). Forecasting company failure in the UK using discriminant analysis and financial ratio data. *Journal of the Royal Statistical Society*, 145, 342–358. doi:10.2307/2981867.

Thomas Ng, S. T., Wong, J. M. W., & Zhang, J. (2011). Applying Z-score model to distinguish insolvent construction companies in China. *Habitat International*, 35(1), 599–607. doi:10.1016/j.habitatint.2011.03.008.

Tseng, F. M., & Hu, Y. C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, 37, 1846–1853. doi:10.1016/j.eswa.2009.07.081.

Tserng, H. P., Chen, P. Huang, W. Lei, M. C., & Tran, Q. H. (2014). Prediction of default probability for construction firms using the logit model. *Journal of Civil Engineering and Management*, 20(2), 247–255. doi:10.3846/13923730.2013.801886

Mpo.cz (2017). *Report on the development of small and medium-sized enterprises and its support in 2016*. Retrieved August 8, 2018, from http:// https://www.mpo.cz/cz/podnikani/male-a-stredni-podnikani/studie-a-strategicke-dokumenty/zprava-o-vyvoji-maleho-a-stredniho-podnikani-a-jeho-podpore-v-roce-2016--232792/

Wang, Y. J., & Lee, H. S. (2008). A clustering method to identify representative financial ratios. *Information Sciences*, 178, 1087-1097. doi:10.1016/j.ins.2007.09.016.

Wu, Y., Gaunt, C., & Gray, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics*, 6, 34-45. doi:10.1016/j.jcae.2010.04.002.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22, 59-82. doi:10.2307/2490859.