Article

Bankruptcy Prediction with the Use of Data Envelopment Analysis: An Empirical Study of Slovak Businesses

Róbert Štefko, Jarmila Horváthová [ID] **and Martina Mokrišová** *

Faculty of Management, University of Prešov, Konštantínova 16, 080 01 Prešov, Slovakia;
robert.stefko@unipo.sk (R.Š.); jarmila.horvathova@unipo.sk (J.H.)
* Correspondence: martina.mokrisova@unipo.sk

check for
updates

Abstract: The paper deals with methods of predicting bankruptcy of a business with the aim of choosing a prediction method which will have exact results. Existing bankruptcy prediction models are a suitable tool for predicting the financial difficulties of businesses. However, such tools are based on strictly defined financial indicators. Therefore, the Data Envelopment Analysis (DEA) method has been applied, as it allows for the free choice of financial indicators. The research sample consisted of 343 businesses active in the heating industry in Slovakia. Analysed businesses have a significant relatively stable position in the given industry. The research was based on several studies which also used the DEA method to predict future financial difficulties and bankruptcies of studied businesses. The estimation accuracy of the Additive DEA model (ADD model) was compared with the Logit model to determine the reliability of the DEA method. Also, an optimal cut-off point for the ADD model and Logit model was determined. The main conclusion is that the DEA method is a suitable alternative for predicting the failure of the analysed sample of businesses. In contrast to the Logit model, its results are independent of any assumptions. The paper identified the key indicators of the future success of businesses in the analysed sample. These results can help businesses to improve their financial health and competitiveness.

Keywords: bankruptcy; data envelopment analysis; logit; model

1. Introduction

Determining the probability of bankruptcy is becoming one of the most important risk management tasks. We pay close attention to predictions of bankruptcy due to the fact that it is important from the point of view of creditors, employees, and other entities around the affected company that would feel the effect the bankruptcy brings (Štefko et al. 2012). Bankruptcy prediction methods can, with a certain degree of probability, alert a company to a negative situation. By taking early remedy measures, businesses can prevent future bankruptcy events (Gundová 2015). To date, empirical studies have found that inefficiency, high corporate indebtedness, and solvency problems are a prerequisite for bankruptcy (Altman 1968). According to Achim et al. (2012) the risk of bankruptcy of an enterprise is closely linked to economic and financial risks. The financial risk is determined by the level of indebtedness, while the economic risk depends on the ratio of fixed and variable costs. In general, knowing these risks allows us to quantify the risk of bankruptcy. Although there is no uniform definition of the concept of bankruptcy, it is advisable to follow the definition of Dimitras et al. (1996), according to whom bankruptcy is a situation where an enterprise is unable to repay its creditors' obligations and meet obligations towards shareholders or suppliers, or where bankruptcy proceedings were commenced under applicable law. Ding et al. (2008) described bankruptcy in a similar way: Bankruptcy is a situation where a

firm could not pay lenders, preferred stock shareholders, suppliers, etc., or a bill was overdrawn, or a firm was bankrupt. Trahms et al. (2013) have contributed the most to the current research on finding the causes of bankruptcy. They state that more complex indicators of business performance decline should be identified in any research on bankruptcy. Scarlat and Delcea (2011) introduced a new concept, a bankruptcy syndrome, which they define as a summary of related symptoms that characterise a particular situation that can lead to bankruptcy. Many authors devote their attention to financial symptoms of bankruptcy, which are expressed by financial indicators. These authors include, among others, Beaver (1966), Altman (1968), Altman et al. (1977), Geng et al. (2014), and Ding et al. (2008) who attributed the risk of bankruptcy to reduced company profitability. One of the most common aspects involved in assessing the cause of bankruptcy is either financial or non-financial causes. Chen et al. (2009) determined that non-financial causes of corporate bankruptcy include: customer relationship levels, IT infrastructure, innovation potential, employee professionalism, and other factors. Martin et al. (2012) referred to other parameters of bankruptcy such as operational risk, competitiveness, credibility analysis, training of employees, quality, etc. The loss of competitiveness as a cause of bankruptcy was also examined by Suhányi and Suhányiová (2017). In addition to the previous classifications, research studies often classify the causes of bankruptcy based on the environment they come from through external and internal causes (Blazy et al. 2008). Despite the existence of various other causes of bankruptcy, our empirical study prioritises the financial causes of bankruptcy.

The next part of the text is structured as follows: The second chapter outlines the theoretical basis of the studied problem. This part of the paper lists various methods and models used for predicting the risk of bankruptcy. A special part is devoted to the summary of the theoretical knowledge about the Data Envelopment Analysis (DEA) method. At the end of the theoretical part, a research problem and goal are formulated. The aim of the paper is to find out whether businesses from the analysed sample have been faced the threat of bankruptcy due to financial difficulties. The third chapter describes the data, the analysed sample of companies, and the processing methods. When addressing the research problem, we made use of selected financial indicators, a correlation matrix, the Additive DEA model (ADD model) model, and the Logit model. We formulated the ADD non-oriented model with variable returns to scale that were solved with the use of Efficiency Measurement Systems (EMS) software. The fourth chapter includes results and discussion of the results achieved. This chapter lists and compares the results of bankruptcy prediction with the use of the DEA and Logit models. The final part of the paper is the conclusion in which the essential conclusions resulting from the research problem are addressed.

2. Theoretical Background

At present, we can observe several tendencies regarding the application of prediction models in predicting bankruptcy. Sun et al. (2014) specifically point to three trends: the transition from one-dimensional analysis of variables to multidimensional prediction, a shift from classical statistical methods to machine learning methods based on artificial intelligence, and more intensive involvement of hybrid and ensemble classifiers. Aziz and Dar (2006) divided prediction models into statistical prediction models, models that use artificial intelligence, and theoretical models. Individually, the use of multiple discriminant analysis (MDA) and Logit models dominates the research (Altman and Saunders 1997; In: Csikosova et al. 2019).

Fitzpatrick (1931) was the first to deal with bankruptcy prediction in his study of solvent and insolvent businesses. In the following years, research on this topic has been carried out by Merwin (1942), Chudson (1945), Jackendoff (1962), and Beaver (1966) (In: Delina and Packová 2013). Beaver (1966) demonstrated that financial ratios can be useful in the prediction of an individual firm's failure. He has proven that not all financial indicators can be used to predict business difficulties. However, the use of simple financial indicators was questioned in practice because of their possible mismanagement by managers. Univariate analysis was later followed by authors who used multivariate analysis. In the beginning of multivariate prediction models, discriminant analysis (DA) was applied. In 1968 Altman

developed a multiple discriminant analysis model (MDA) called the Z-Score Model. Since Altman's study, the number and complexity of these models has increased dramatically. DA was explored by Blum (1974), Elam (1975), Altman et al. (1977), Norton and Smith (1979), and Taffler (1983). Altman's original model required the fulfilment of multinormality, homoskedasticity, and linearity assumptions. These prerequisites for financial indicators were often not met. The main drawback of DA, however, is that although it is able to identify businesses that are likely to go bankrupt, it is not able to estimate the likelihood of this situation occurring. Based on these shortcomings, the next step in the theory of bankruptcy prediction was to develop methods and models that would be able to provide such information (Mihalovič 2015). That was the reason why logistic regression began to be preferred, as this method does not have to meet these conditions. Compared to methods based on multi-dimensional discriminant analysis, logistic regression has several advantages. It has a higher predictive ability and its application does not require compliance with assumptions that could limit its usability. The method was first used to predict the bankruptcy of banks by Martin (1977). Ohlson (1980) was the first to use it to assess companies. Ohlson, as a pioneer in the application of Logit analysis, did not agree with the application of discriminant analysis to predict bankruptcy due to its requirement for a variance-covariance matrix (Klieštik et al. 2014). However, even Logit models have their weakness—their sensitivity to remote observation.

Another method used in the area of bankruptcy prediction is DEA (Horváthová and Mokrišová 2018). Compared to statistical methods, DEA is a relatively new, non-parametric method, which represents one of the main possible approaches to assessing the financial health of a business and its risk of bankruptcy (Štefko et al. 2018). This method was first applied in Charnes et al. (1978). It is based on the idea mentioned in the article "Measuring efficiency of decision making units", published by Farrell in 1957. His work was based on the works of Debreu (1951) and Koopmans (1951). Farrell (1957) proposed a new approach to efficiency measuring based on a linear convex envelopment curve and the use of distance measurement functions between the enterprise of interest and the projected point on the efficiency frontier. In this way, he proposed a new level of efficiency based on the calculation of two components of the overall business efficiency: technical efficiency and resource allocation efficiency. Farrell's approach measures the ability of the business to transform inputs into outputs. Therefore, it is also called the input-oriented approach. Charnes et al. (1978) have applied a multiplicative input-output model to measure business efficiency. The approach of these authors represents a two-stage efficiency calculation. The first step is to identify the production frontier, while businesses that lie on this line are among the best businesses. In the second step, the efficiency score is calculated for the analysed enterprises and their distance from the production frontier is determined. From the point of view of their input, DEA models can be divided into DEA CCR (Charnes et al. 1978) and DEA BCC (Banker et al. 1984). This method was further developed by Färe et al. (1985). The DEA method was also used by the following authors: Tone (2001); Wang et al. (2007); Kao and Hwang (2008); Sadjadi and Omrani (2008); Zhu (2015); Oanh and Ngoc (2016); Ghomi et al. (2019); Dumitrescu et al. (2020); and many others.

The first idea to use the DEA method to predict bankruptcy was recorded by Simak (1997), who was the first to compare its results with the results of Altman's Z-score. Other authors dealing with the DEA bankruptcy prediction included Cielen et al. (2004). The authors used the DEA radial model to predict bankruptcy and compared the results with DA results. In the same year Paradi et al. (2004) applied an additive and radial model along with the peeling technique. The model achieved 100% success in predicting the bankruptcy of businesses. In 2009, Premachandra et al. used an ADD model and compared its results with the results of logistic regression. The result of this research was a satisfactory level of correct prediction of business bankruptcy. The prediction rate for financially sound businesses was less accurate. Sueyoshi and Goto (2009) applied an ADD model to create a line under which businesses go bankrupt. The results were then compared with the DEA-DA approach. In 2011, Premachandra et al. combined the radial and ADD model and created the DEA ranking index. Shetty et al. (2012) applied the DEA model in 2012 to determine the bankruptcy likelihood for their

analysed business sample. The result of their study was the designing of indicators that should be applied as predictors of bankruptcy.

Other methods that are suitable for the application in the field of bankruptcy prediction include neural networks. In this context, it is possible to mention the authors Odom and Sharda (1990), who developed a neural network to investigate business bankruptcy using selected financial indicators. Gherghina (2015) made a significant contribution to the application of neural networks in this area. The neural network in the field of bankruptcy prediction was also applied by Altman et al. (1994). Other methods include decision trees (Breiman et al. 1984; Frydman et al. 1985). However, in conclusion, it should be noted that the most commonly used methods today are discriminant analysis and logistic regression.

In line with the above-mentioned text, we identified the following research problem: Is the DEA method a suitable alternative in predicting failure of businesses from the analysed sample? In relation to the research problem, the aim of the paper was formulated: To predict business failure with the use of the ADD model and to compare its results with the results of the Logit model. The aim was also to analyse classification and estimation accuracy of the ADD DEA model and to compare it with the classification and estimation accuracy of the Logit model.

3. Methodology and Data

DEA models are designed to assess the technical efficiency of production units based on the size of inputs and outputs. There are two possible approaches to creating DEA models: multiplicative and dual. The dual model is an additional task to the multiplicative one. A significant problem of the DEA analysis are production externalities (negative outputs) and desirable inputs. Generally, in DEA models, the basic prerequisite is data positivity. However, situations in which negative inputs and outputs occur are not uncommon. In the case of the sample of companies analysed, negative outputs occurred in the case of profitability. The ways to deal with this problem are different. Some software programmes attach zero weight to negative inputs and outputs. Another frequently used option is to treat the negative outputs as inputs (thus minimising them) and the desired inputs as outputs (thus maximising them). However, this procedure is not universally applicable. One of the simpler options is to use an additive model in which the positive and negative inputs and outputs are evaluated separately (Premachandra et al. 2009; Mendelová and Stachová 2016).

The ADD model is one of the non-oriented models. This model was formulated by Charnes et al. (1985). A Decision Making Unit (DMU) was introduced as a unit for which efficiency was solved and which describes any entity for which the process of transforming inputs into outputs is in progress. Determining DMU efficiency with an additive model for variable returns to scale means solving the following linear programming model:

$$
\begin{aligned}
&\max_{\lambda, s^x s^y} A_o = \left(e_m^T s^x + e_s^T s^y \right) \\
&\text{s.t. } \textstyle\sum_{j=1}^{n} x_j \lambda_j + s^x = x_o, \quad s^x \geq 0, \\
&\quad\;\; \textstyle\sum_{j=1}^{n} y_j \lambda_j - s^y = y_o, \quad s^y \geq 0, \\
&\quad\;\; \textstyle\sum_{1}^{n} \lambda_j = 1, \qquad \lambda_j \geq 0,
\end{aligned}
\tag{1}
$$

where e_m, e_s, are unit vectors of appropriate length and s^x, s^y are additional variables-slacks. DMU$_o$ o = {1, . . . , n}, is efficient when $s^x = 0$, $s^y = 0$, in other words, when the objective function and all slacks equal zero. Otherwise, the DMU$_o$ is inefficient.

Since our paper does not address the efficiency of the analysed sample, but rather covers bankruptcy, the input vectors x_o, were replaced by output vectors y_o. The efficiency condition in this case served as a condition for the assumed bankruptcy of the company. In our research, we used 9 financial indicators. We selected this group of indicators in such a way that it contains indicators from all areas of financial health evaluation (liquidity, profitability, activity, indebtedness) and there is not a strong correlation between indicators. As output variables, we applied indicator

LLTA—long-term liabilities/total assets used as a leverage measure which indicates long-term financial obligation, and indicator CLTA—current liabilities/total assets which indicates a lack of cash flow to fund business operations. As input variables, we applied 7 indicators: TRTA—total revenue/total assets, CR (Current ratio)—(financial assets + short-term receivables)/current liabilities, WCTA—working capital/total assets, CATA—current assets/total assets, EBTA—earnings before interest and taxes/total assets, EBIE—earnings before interest and taxes/interest expense, and ETD—equity/total debt. For the creation of the ADD model, we used the Efficiency Measurement System (EMS) software. We divided the results of the DEA model into 6 zones (businesses in financial distress—3 zones and financially healthy businesses—3 zones) according to Mendelová and Bieliková (2017).

3.1. Logit Model

The Logit regression model was applied to compare the results obtained with the DEA model. The Logit model is a widespread model that has been used by several authors to predict the default/no default probability of a company (Premachandra et al. 2009; Kováčová and Klieštik 2017; Mendelová and Stachová 2016). This model is a type of multivariate statistical model. It captures the relationship between the dependent variable Y and the independent variable X.

Logistic regression works very similar to linear regression, but with a binomial response variable (Sperandei 2014). The dependent variable y_i can only take two values: $y_i = 1$ if the probability of bankruptcy occurs and $y_i = 0$ if the probability of bankruptcy does not occur. Therefore, we can assume that probability $y_i = 1$ is given by P_i; probability $y_i = 0$ is given by $1 - P_i$. By using logistic transformation, we could specify the probability P_i using the following model: $P_i = f(\alpha + \beta x_i)$, where x_i are the chosen financial indicators while α and β are estimated parameters. P_i is then calculated using the logistic function:

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

A logistic regression models the chance of an outcome based on individual characteristics (Sperandei 2014). According to Kováčová and Klieštik (2017), the Logit can be defined as:

$$Logit = \ln\left(\frac{P_i}{1 - P_i}\right) = f(\alpha + \beta x_i). \tag{3}$$

The above represents the logarithm of the odds ratio of the two possible alternatives (P_i, $1 - P_i$). It is called the Logit. The goal of logistic regression is to calculate the odds ratio ($\frac{P_i}{1-P_i}$); ln in this relationship represents the Logit transformation.

For the creation of Logit model, it was necessary to divide businesses into bankrupt and non-bankrupt. When choosing the appropriate conditions for bankruptcy occurrence evaluation, we studied the papers of various authors. Some of them assume that a company goes bankrupt if it does not make a profit (Beaver 1966; Altman 1968; Altman et al. 1977; Geng et al. 2014) or reaches negative cash flow (Ding et al. 2008). Based on the bankruptcy definition stated in the Introduction, we chose the value of indebtedness as the bankruptcy condition. We then detected 50 bankrupt businesses.

When creating the Logit model, we started with the same 9 indicators which we used for the ADD model. However, there was a strong correlation between ETD and the indicator of indebtedness which we used as a bankruptcy criterion. Therefore, we did not use the indicator ETD in the Logit model. We also excluded the indicator CATA because when applying it, the Logit model did not process any coefficients. We assumed that the indicators CATA and WCTA are related indicators which evaluate the same financial area of evaluation and express the same reality. For the creation of the Logit model, we used software Statistica 13.1.

Using the results of the Logit model, it is possible to determine whether a company is about to go bankrupt or not. This classification may use a cut-off score (usually 0.5), with businesses above this value facing a probability of going bankrupt and businesses below this value facing lower (or no) probability of going bankrupt. Two types of misclassification can occur when evaluating

business failure. The type I error (false negative rate) arises when a bankrupt company is classified as non-bankrupt, and the type II error (false positive rate) arises when the non-bankrupt company is classified as bankrupt (Kováčová and Klieštik 2017).

The prediction ability of the Logit model can be verified by using the Area Under Curve (AUC) method, which measures the area under the Receiver Operating Characteristic curve (ROC curve). This analysis represents a statistical procedure for evaluating correct and false positives as well as correct and false negatives. ROC curve analysis describes the relationship of sensitivity and specificity at different discriminatory levels. AUC measures overall performance of the model. It can take on any value between 0 and 1. The closer AUC is to 1, the better is the overall performance of the model (Park et al. 2004).

One of the important tests that can be mentioned in verifying the Logit model is the Wald test, which confirms the significance of variables in the model. Based on the results of Likelihood ratio test, the model includes those variables, which increases its maximum credibility. This test is suitable not only to assess the significance of the model, but also to assess the contribution of individual predictors to the model. The higher the Chi-square test statistic, the better the model reflects the situation of a business. In addition to the above tests, the results of the Hosmer–Lemeshow test should be mentioned. This test indicates the compliance of the model with the applied data. Nagelkerke's R Square explains the percentage of variance, while we could also find out how successful the model is in explaining the "variability" of a dependent variable.

3.2. Description of the Sample of Companies

The input database of this empirical study was created from data obtained for 497 companies operating in Slovakia in the heat supply industry. The database of the data from financial statements of these companies for the year 2016 was obtained from the Slovak analytical agency CRIF—Slovak Credit Bureau, s.r.o. According to SK NACE Rev. 2, the sample of enterprises analysed falls under section D: "Supply of electricity, gas, steam and cold air". Sources and distribution of heat of these businesses were built along with the development of urban agglomerations. Their systems allow the effective use of various sources of energy generated in a city, including renewable sources, waste heat, and so on. These systems are an energy infrastructure integrator which can efficiently link production and consumption and enable the storing of energy (in the form of heat) at the time of its surplus. As part of independent heat production, today about 54% of the heat is produced in combined production (Janiš 2018). The European Commission's winter energy package sets new targets for energy efficiency. These goals and new trends in energy bring new opportunities and challenges for the heating industry. These facts are a precondition for the occurrence of risk factors which affect the performance and competitiveness of analysed businesses from outside. A more detailed analysis of this sample excluded 154 companies due to a negative value of equity or deficiencies in the database. The resulting analysed sample consisted of 343 companies. In terms of each business's legal status, 15% of the companies are joint stock companies and the remaining 85% are limited liability companies. The results of the financial analysis show that the analysed companies do not have a liquidity problem. The average value of current ratio found is 3.92. However, we also obtained a median of current ratio of 0.951. This was also reflected in the negative value of net working capital. The analysed sample of companies reported a high creditors payment period, which results in a negative value of cash conversion cycle. The assets of these companies change on average once a year. The average value of the return on assets is 5%. The capital structure of these companies is 35:65 in favour of equity. The performance of companies active in the heating industry was not found to reach the required value to avoid bankruptcy.

The Table 1 shows the descriptive statistics of indicators, which represent indicators applied in the DEA model. The values of the indicators are divided into two groups. The first group consists of bankrupt businesses and the second group consists of non-bankrupt businesses. Of these values, the negative values for WCTA, EBTA, ETD, and EBIE should be pointed out, as these negative values are one of the signs of bankruptcy. Analysed businesses also have high indebtedness.

Table 1. Descriptive statistics for bankrupt and non-bankrupt businesses.

Financial Indicators	TRTA	CR	WCTA	CATA	EBTA	EBIE	ETD	LLTA	CLTA
Bankrupt businesses (50)									
Mean	0.336	3.925	−0.239	0.185	−0.026	−1.1120	−0.217	0.499	0.424
Median	0.205	0.473	−0.067	0.138	−0.000	0.014	−0.200	0.615	0.227
Standard deviation	0.875	12.190	0.590	0.173	0.129	5.244	0.150	0.384	0.600
Skewness	6.752	4.433	−3.273	2.239	1.181	−4.877	−0.909	−0.150	3.450
Non-bankrupt businesses (293)									
Mean	0.854	3.644	−0.036	0.300	0.070	830.989	0.838	0.330	0.332
Median	0.286	0.813	−0.018	0.192	0.055	2.243	0.241	0.300	0.240
Standard deviation	1.460	12.43	0.270	0.270	0.160	11,531.54	3.5800	0.300	0.300
Skewness	3.461	6.075	−0.541	1.416	6.669	16.550	13.889	0.167	0.893

4. Results and Discussion

During the research process presented in the paper, we created two models. One was developed based on DEA as an ADD model. As the outcome of the ADD model is an extensive set of data, we present selected data as an example in Table 2. In the case of businesses that have a score of zero and all their slacks equal to zero, it can be said that they are likely to go bankrupt. From the Table 3, it can be seen that such a situation applies to e.g., company number 122 and company number 126.

Table 2. Selected results of the Additive DEA model (ADD model).

DMU	Score	TRTA	CR	WCTA	CATA	EBTA
TP121	2.48	0	0	0	0.52	0.15
TP122	0	0	0.01	−0.27	0.09	0
TP123	2.21	0	0	0	0	0.16
TP124	2.54	0.43	0	0	0	0.19
TP125	2.62	0.47	0	0	0	0.27
TP126	0	0.05	0	−0.45	0.45	0.01
TP127	1.77	0	0	0	0	−0.17
TP128	1.2	0.6	0	0	0	0
TP129	0.1	0.45	0	0	0	−0.18

Table 3. Results of a multi-zones ADD model.

DEA Zones	1	2	3	4	5	6
Number of businesses	17	15	23	56	81	151

The classification of businesses into individual DEA zones is shown in Table 3. In case of three financial distress zones, the risk of going bankrupt affects 55 companies. With regard to the first zone of the DEA model, it turned out that 17 companies face risk of bankruptcy. With regard to the second zone, 15 companies face such risk, and in case of the third zone, the risk is being faced by another 23 companies.

To determine a model's estimation accuracy, a cut-off point of 0.5 is usually used. However, this cut-off is not appropriate for every given model. Therefore, we were looking for the optimal cut off corresponding to a point in which the sum of sensitivity and specificity is the highest. The optimal cut-off was found at the level of 0.63. The results are shown in Table 4. In this case, 82% estimation accuracy for bankrupt businesses and 56% estimation accuracy for non-bankrupt businesses could be achieved. We can consider this to be an adequate estimation accuracy rate.

Table 4. Classification accuracy of the ADD model corresponding to a cut-off point of 0.63.

	Predicted: Yes	**Predicted: No**	**% Correct**	**Error %**
Observed: yes	41	9	82 188	18 (I)
Observed: no	131	163	56	44 (II)

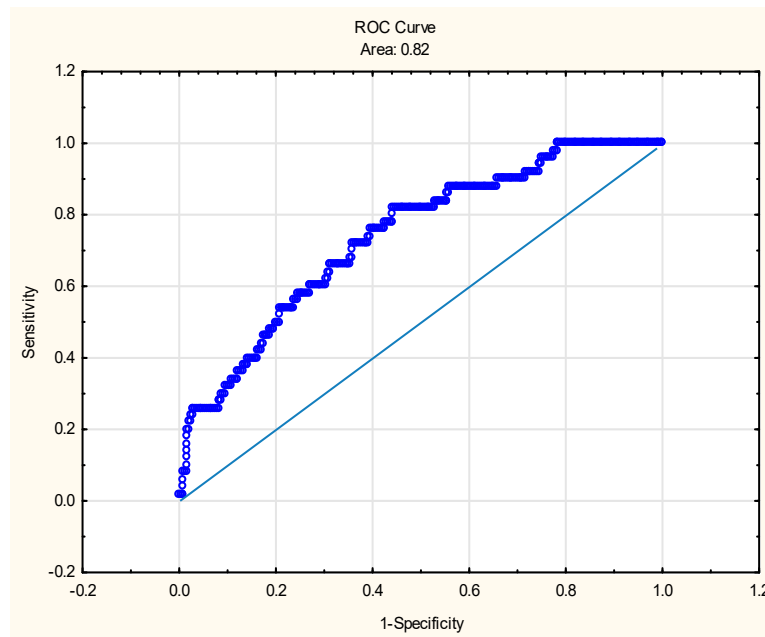We also constructed a ROC curve for the ADD model (see Figure 1), where the AUC is at the level of 0.82.

Figure 1. ROC curve for ADD model.

We then formulated Logit model to identify businesses that are likely to go bankrupt. We selected 7 financial indicators for the Logit model. They are indicators TRTA, CR, WCTA, EBTA, LLTA, EBIE and CLTA. Table 5 shows the results of Logit model.

Table 5. Logit function coefficients.

Effect	**Bankrupt-Parameter Estimates**						
	Column	**Estimate**	**Standard Error**	**Wald Stat.**	**Lower CL 95.0%**	**Upper CL 95.0%**	**P**
Intercept	1	−2.15434	0.443236	23.62436	−3.0231	−1.28562	0.000001
TRTA	2	−0.48536	0.415575	1.36406	−1.2999	0.32915	0.242835
CR	3	0.01264	0.011864	1.13604	−0.0106	0.03590	0.286491
WCTA	4	−1.58510	1.136295	1.94594	−3.8122	0.64200	0.163025
EBTA	5	−8.52758	1.987190	18.41503	−12.4224	−4.63276	0.000018
LLTA	6	1.31126	0.591269	4.91819	0.1524	2.47012	0.026575
EBIE	7	−0.00026	0.002247	0.01353	−0.0047	0.00414	0.907416
CLTA	8	0.19154	1.156441	0.02743	−2.0750	2.45812	0.868449
Scale		1.00000	0.000000		1.0000	1.00000	

Based on the Wald confidence intervals, it can be stated, with 95% confidence, that the coefficients of the variables EBTA, LLTA are within the specified limits of the interval and none of the intervals contains a value of 0, which would exclude the variable from the model. Since no Wald statistics

parameter is equal to zero, it can be assumed that all explanatory variables can be included in the model. It follows that the tested variables are suitable for the Logit model. At the same time, Wald statistics determines which of the independent variables is more important than the others. A statistically significant relationship determining probability of bankruptcy was confirmed for the indicators EBTA and LLTA. The above results show that the probability of bankruptcy is determined by profit and indebtedness. The resulting Logit function providing the probability of business bankruptcy is:

$$P_1 = \frac{1}{1 + e^{-(-2.150.49 \times TRTA + 0.01 \times CR - 1.59 \times WCTA 8.53 \times EBTA + 1.31 \times LLTA - 0.0003 \times EBIE + 0.19 \times CLTA)}}, \quad (4)$$

To assess the estimation accuracy of the model, we constructed a Receiver Operating Characteristic (ROC) curve (see Figure 2). In our case, the AUC accounts for 79.65%, which we evaluated positively. Therefore, we can state that our model has good estimation accuracy.
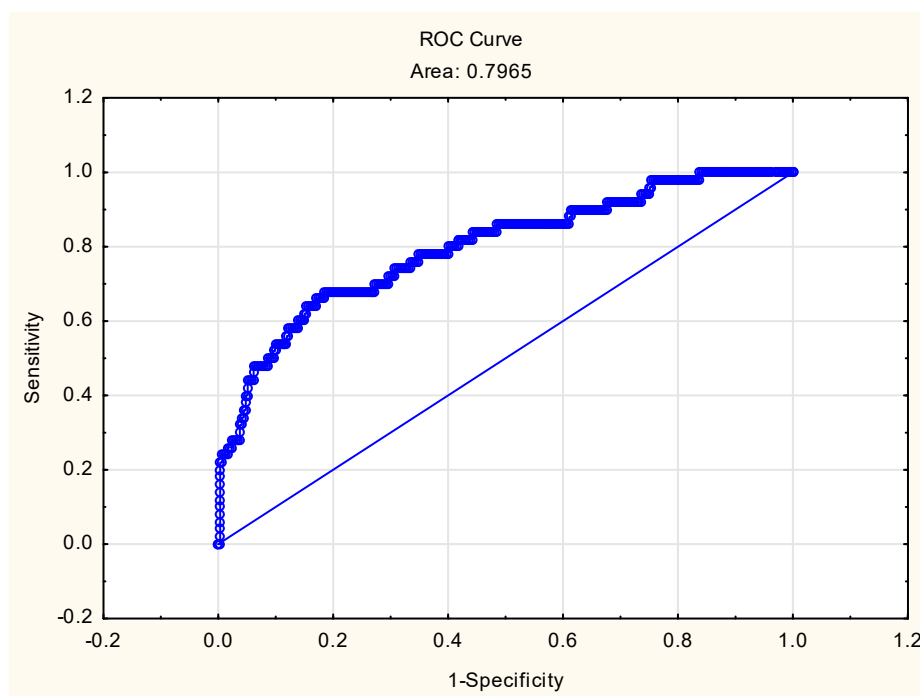

Figure 2. Receiver Operating Characteristic (ROC) curve for the Logit model.

The Hosmer–Lemeshow test signalized good conformity of the final model with given data. The *p*-value of the test was 0.71. This value is higher than the significance level, so we accepted the null hypothesis—the distribution of predicted and achieved results is the same across all groups of businesses. According to Nagelkerke's R Square, the model explains 26.73% variability of the binary dependent variable. Total estimation accuracy of the Logit model was found to be 86.6%, for non-bankrupt businesses the result was 96%, and estimation accuracy for bankrupt businesses was 30% (see Table 6). The error type I was 70% and error type II was 4%. The model was found to have higher classification accuracy for businesses that are financially sound.

Table 6. Classification accuracy of the Logit model corresponding to the same cut-off as the ADD model.

	Classification of Cases		
	Predicted: Yes	**Predicted: No**	**Percent Correct**
Observed: yes	15	35	30%
Observed: no	11	282	96%

A comparison of the classification ability of the models is given in Table 7 (corresponding to a cut-off point of 0.63). Several researchers compared the results of the ADD model and the Logit model. We already described the research of Premachandra et al. (2009) in the Introduction. The comparison of these two models was also performed by the Slovak researchers Mendelová and Stachová (2016, p. 103) who based on their research concluded that, in general, they cannot say that one method is better than the other one, because the accuracy and suitability of each method depends on the particular data used, its size, and its proportions. The results of the ADD model and Logit model were also compared by Araghi and Makvandi (2012). They found out that DEA is an effective tool for predicting business bankruptcy, but it is not as efficient as the Logit model—DEA achieved a weak performance in identifying bankrupt and non-bankrupt companies.

Table 7. Comparison of the estimation accuracy of the ADD and Logit models corresponding to optimal cut-off for the DEA model.

Model	Error Type I	Error Type II	Overall Estimation Accuracy	Sensitivity	Specificity
Logit	70%	4%	87%	30%	96%
DEA	18%	44%	59%	82%	56%

The ADD model was found to have a lower classification accuracy for non-bankrupt businesses corresponding to a cut-off of 0.63 and the Logit model has lower classification accuracy for bankrupt businesses at this cut-off (see Table 7). Therefore, it is necessary to state the optimal cut off for the Logit model, which is 0.16. At this point, the Logit model achieves a higher classification accuracy for bankrupt businesses.

Figure 3 shows the estimation accuracy for bankrupt and non-bankrupt businesses, as well as the percentage of total correct predictions for the DEA model. In the case of non-bankrupt businesses, the highest estimation accuracy was found at a cut-off of 1. Then the estimation accuracy decreases. In the case of bankrupt businesses, the estimation accuracy gradually increases up to 100% at a cut-off point of 0.5 and lower. The overall estimation accuracy decreases with a decreasing cut-off.
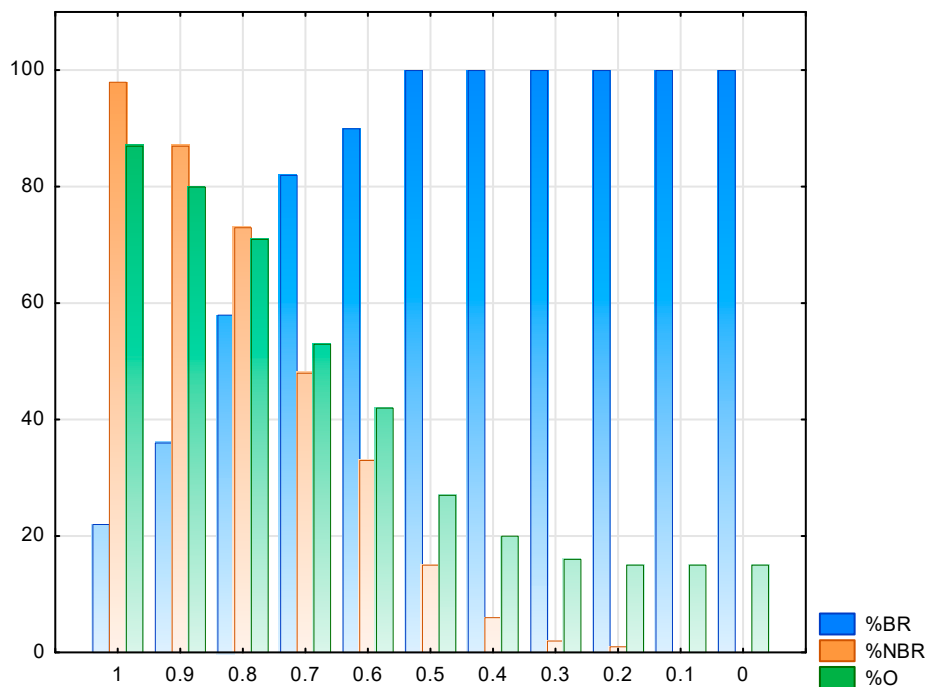

Figure 3. Percentage of correct predictions using the DEA model. Legend: BR—Estimation Accuracy for Bankrupt Businesses. NBR—Estimation Accuracy for Non-Bankrupt Businesses, O—Overall Estimation Accuracy.

Figure 4 illustrates the estimation accuracy for bankrupt and non-bankrupt businesses, as well as the percentage of total correct predictions for the Logit model.
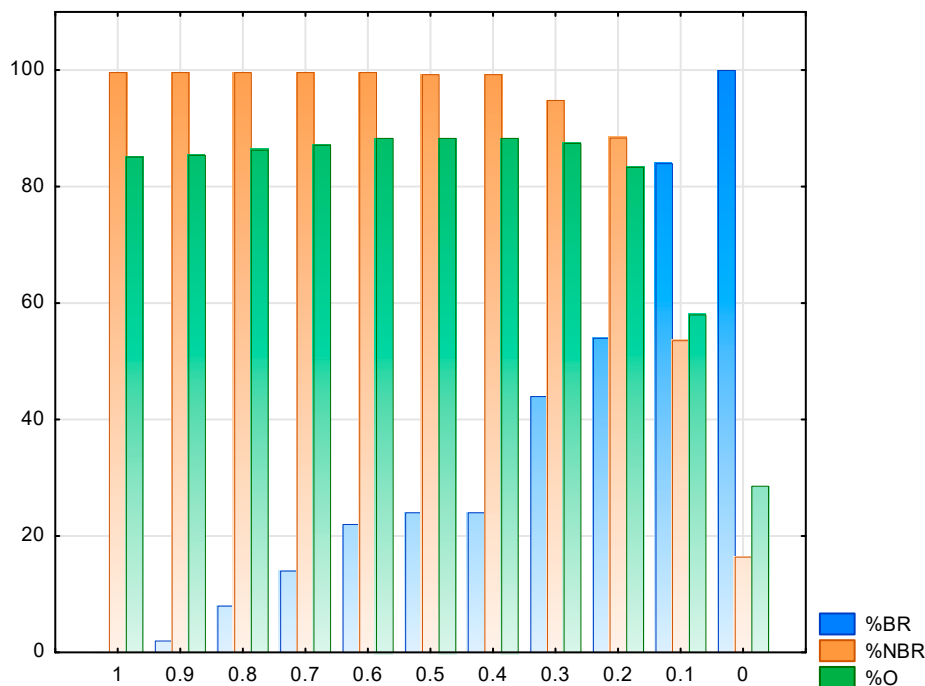

Figure 4. Percentage of correct predictions using the Logit model.

Figure 4 shows that estimation accuracy in the case of the Logit model is different in comparison with the results of the DEA model. In the case of non-bankrupt businesses, the highest estimation accuracy is from a cut-off point of 1 to cut-off point of 0.4; from a cut-off of 0.3, the estimation accuracy decreases rapidly. In the case of bankrupt businesses, the estimation accuracy gradually increases up to 100% at a cut-off point of 0. The overall estimation accuracy slightly increases up to a cut-off of 0.3 and then rapidly decreases. At the end of this discussion, it is necessary to point out the fact that each of the models has its optimal cut-off and the estimation accuracy of the models is given by the selected value of each cut-off.

5. Conclusions

In this paper, we created specific bankruptcy prediction models for the analysed sample of businesses with the use of the DEA method and Logit model. Inspired by authors who dealt with the causes of bankruptcy and based on the correlation matrix, we selected financial indicators as inputs and outputs for the constructed models. Based on the scientific literature, we also identified the bankruptcy condition of indebtedness, which was necessary to classify businesses into bankrupt and non-bankrupt businesses. This condition was applied in the case of the Logit model and together with the profit, it was verified by this model as a symptom of bankruptcy. Based on the results outlined in the Results and Discussion section, we can say that the ADD model achieved an estimation accuracy for bankrupt businesses of 82%. A similar estimation accuracy rate for businesses threatened with bankruptcy was presented in the work of Premachandra et al. (2009) of 84.89%; Mendelová and Stachová's (2016) accuracy rate was 10–42.86% and Cielen et al.'s (2004) accuracy rate was 74.4–75.7%. We obtained an estimation accuracy of the ADD model for non-bankrupt businesses of 56%, and an overall estimation accuracy of DEA model of 59%. We can compare this result with the outcomes of the above-mentioned authors: Premachandra et al. (2009) obtained 75–77%, Mendelová and Stachová (2016) obtained 88–95%, and Cielen et al. (2004) obtained 85.1–86.4%. The error type I for the ADD model was 18% and the error type II was 44%.

Several studies (Premachandra et al. 2011; Paradi et al. 2014) confirm that the traditional cut-off point of 0.5 may not be appropriate for assessing bankruptcy models' estimation accuracy. This was also confirmed in our research. The optimal cut off for DEA model was found to be 0.63, and we compared the results of the DEA model with the results of theLogit model at this cut-off. Overall estimation accuracy and estimation accuracy for non-bankrupt businesses was higher in the case of the Logit model. On the other hand, DEA has a higher estimation accuracy for bankrupt businesses. Also error type I was lower in the case of the DEA model. The optimal cut-off for Logit model was different, corresponding to 0.16. It should be noted that estimation accuracy of the models depends on their cut-off values. Both models have different optimal cut-offs, so the results cannot be clearly compared. DEA identifies fewer businesses at risk of bankruptcy, but at a higher probability of achieving bankruptcy. Logit identifies more businesses at risk of bankruptcy, but with a lower probability of the identified businesses achieving bankruptcy. This fact may also speak in favour of the application of DEA model in predicting the financial distress of businesses. In this paper, the optimal cut-off was set as the value at which the sum of sensitivity and specificity is the highest. Another way of determining an optimal cut-off is to calculate an index based on two DEA models, one representing the financial health frontier and the other representing the financial distress frontier.

Results of the constructed models can be a starting point to improve financial health, prosperity, and competitiveness of analysed businesses. Based on the achieved results, we can conclude that within our research sample, DEA identifies bankrupt businesses at a higher probability of bankruptcy than the Logit model. The DEA method does not take into account initial bankruptcy conditions but its results are based on the achieved values of financial indicators, so they are independent of any assumptions. A significant benefit of this method is that it allows us to accept the specifics of companies and industry. In contrast to the Logit model, it offers us goal values of indicators, which the Logit model does not offer. Based on the above-mentioned factors, we can conclude that the DEA method is a suitable alternative for predicting the failure of businesses from the analysed sample. In order to increase estimation accuracy of the DEA model and decrease type I and II errors for this model, in our further research we will focus on selecting explanatory variables from a wider range of financial and even non-financial indicators.

A limitation of our research is the sample of businesses used, which consisted of a limited number of companies and insufficient data; therefore, we will improve our sample and data to overcome these shortcomings in the future. However, it is important to note that the research sample consisted of real businesses and it took into account all businesses in the Slovak heat sector. Therefore, we can say that results will be beneficial for that industry.

References

Achim, Monica Violeta, Codruta Mare, and Sorin Nicolae Borlea. 2012. A statistical model of financial risk bankruptcy applied for Romanian manufacturing industry. *Procedia Economics and Finance* 3: 132–7. [CrossRef]

Altman, Edward I. 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance* 23: 589–609. [CrossRef]

Altman, Edward I., and Anthony Saunders. 1997. Credit risk measurement: Development over the last 20 years. *Journal of Banking & Finance* 21: 1721–42. [CrossRef]

Altman, Edward I., Robert G. Haldeman, and Paul Narayanan. 1977. ZETA ANALYSIS, a new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance* 1: 29–54. [CrossRef]

Altman, Edward I., Giancarlo Marco, and Franco Varetto. 1994. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance* 18: 505–29. [CrossRef]

Araghi, Khalili, and Sara Makvandi. 2012. Evaluating Predictive power of Data Envelopment Analysis Technique Compared with Logit and Probit Models in Predicting Corporate Bankruptcy. *Australian Journal of Business and Management Research* 2: 38–46.

Aziz, Adnan M., and Humayon A. Dar. 2006. Predicting corporate bankruptcy: Where we stand? *Corporate Governance* 6: 18–33. [CrossRef]

Banker, Rajiv D., Abraham Charnes, and William W. Cooper. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30: 1078–92. [CrossRef]

Beaver, William H. 1966. Financial ratios as predictors of failure. *Journal of Accounting Research* 4: 71–111. [CrossRef]

Blazy, Régis, Bertrand Chopard, and Agnès Fimayer. 2008. Bankruptcy law: A mechanism of governance for financially distressed firms. *European Journal of Law and Economics* 25: 253–67. [CrossRef]

Blum, Marc. 1974. Failing company discriminant analysis. *Journal of Accounting Research* 12: 1–25. [CrossRef]

Breiman, Leo, Jerome Friedman, Richard Olshen, and Charles J. Stone. 1984. *Classification and Regression Trees*. New York: Taylor & Francis.

Charnes, Abraham, William W. Cooper, and Edwardo L. Rhodes. 1978. Measuring the efficiency of decision making units. *European Journal of Operation Research* 2: 429–44. [CrossRef]

Charnes, Abraham, William W. Cooper, Boaz Golany, Lawrence M. Seiford, and Joel D. Stutz. 1985. Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions. *Journal of Econometrics* 30: 91–127. [CrossRef]

Chen, Hsueh-Ju, Shaio Yan Huang, and Chin-Shien Lin. 2009. Alternative diagnosis of corporate bankruptcy: A neuro fuzzy approach. *Expert Systems with Applications* 36: 7710–20. [CrossRef]

Cielen, Anja, Ludo Peeters, and Koen Vanhoof. 2004. Bankruptcy prediction using a data envelopment analysis. *European Journal of Operational Research* 154: 526–32. [CrossRef]

Csikosova, Adriana, Maria Janoskova, and Katarina Culkova. 2019. Limitation of Financial Health Prediction in Companies from Post-Communist Countries. *Journal of Risk and Financial Management* 12: 15. [CrossRef]

Debreu, Gerard. 1951. The coefficient of resource utilization. *Econometrica* 19: 273–92. [CrossRef]

Delina, Radoslav, and Miroslava Packová. 2013. Validácia predikčných bankrotových modelov v podmienkach SR. *Ekonomie a Management* 16: 101–12.

Dimitras, Augustinos I., Stelios H. Zanakis, and Constantin Zopounidis. 1996. A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research* 90: 487–513. [CrossRef]

Ding, Yongsheng, Xinping Song, and Yueming Zen. 2008. Forecasting Financial Condition of Chinese Listed Companies Based on Support Vector Machine. *Expert Systems with Applications* 34: 3081–9. [CrossRef]

Dumitrescu, Dalina, Ionela Costică, Liliana Nicoleta Simionescu, and Stefan Cristian Gherghina. 2020. DEA Approach Towards Exploring the Sustainability of Funding in Higher Education. Empirical Evidence from Romanian Public Universities. *The Amfiteatru Economic journal* 22: 593–607. [CrossRef]

Elam, Rick. 1975. The effect of lease data on the predictive ability of financial ratios. *The Accounting Review* 50: 25–43.

Färe, Rolf, Shawna Grosskopf, and C.A. Knox Lovell. 1985. *Measurement of Efficiency of Production*. Boston: Kluwer-Nijhoff Publishing Co.

Farrell, Michael James. 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A* 120: 253–90. [CrossRef]

Fitzpatrick, Paul Joseph. 1931. A comparison of the ratios of successful industrial enterprises with those of failed companies. *Certified Public Accountant* 6: 727–31.

Frydman, Halina, Edward I. Altman, and Duen-Li Kao. 1985. Introducing recursive partitioning for financial classification: The case of financial distress. *Journal of Finance* 40: 269–91. [CrossRef]

Geng, Ruibin, Indranil Bose, and Xi Chen. 2014. Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European Journal of Operational Research* 241: 236–47. [CrossRef]

Gherghina, Stefan Cristian. 2015. An Artificial Intelligence Approach towards Investigating Corporate Bankruptcy. *Review of European Studies* 7: 5–22. [CrossRef]

Ghomi, Hamidreza, Morteza Rahmani, and Morteza Khakzar Bafruei. 2019. A two-stage method for assessing the efficiency of the three-stage series network data envelopment analysis model with two feedback. *Decision Science Letters* 8: 339–352. [CrossRef]

Gundová, Petra. 2015. Verification of the selected prediction methods in Slovak companies. *Acta Academica Karviniensia* 14: 26–38. [CrossRef]

Horváthová, Jarmila, and Martina Mokrišová. 2018. Risk of Bankruptcy, its Determinants and Models. *Risks* 6: 117. [CrossRef]

Janiš, Stanislav. 2018. Slovak Heating Companies are Preparing for the Fourth Generation of Heating. In *Slovak Enetgy Annual 2018*. Bratislava: RE-PUBLIC, pp. 76–79.

Kao, Chiang, and Shiuh-Nan Hwang. 2008. Efficiency decomposition in twostage data envelopment analysis: An application to non-life insurance companies in Taiwan. *European Journal of Operational Research* 185: 418–29. [CrossRef]

Klieštik, Tomáš, Alexander N. Lyakin, and Katarína Valašková. 2014. Stochastic calculus and modelling in economics and finance. Paper Presented at 2nd International Conference on Economics and Social Science ICESS 2014, Shenzhen, China, July 29–30; pp. 161–7.

Koopmans, Tjalling Charles. 1951. Analysis of production as an efficient combination of activities. In *Activity Analysis of Production and Allocation*. New York: J. Wiley and Sons, London: Chapman and Hall, pp. 33–97.

Kováčová, Mária, and Tomáš Klieštik. 2017. Logit and Probit application for the prediction of bankruptcy in Slovak companies. *Equilibrium. Quarterly Journal of Economics and Economic Policy* 12: 775–91. [CrossRef]

Martin, Daniel. 1977. Early warning of bank failure. A logit regression approach. *Journal of Banking and Finance* 1: 249–76. [CrossRef]

Martin, Aaron Hillsboro, V. Aswathy, Sundarraman Balaji, T. Muthu Lakshmi, and V. Prasanna Venkatesan. 2012. An Analysis on Qualitative Bankruptcy Prediction Using Fuzzy ID3 and Ant Colony Optimization Algorithm. Paper Presented at International Conference on Pattern Recognition, Informatics and Medical Engineering PRIME 2012, Periyar University, Salem, Tamilnadu, India, March 21–23.

Mendelová, Viera, and Tatiana Bieliková. 2017. Diagnostikovanie finančného zdravia podnikov pomocou metódy DEA: Aplikácia na podniky v Slovenskej republike. *Politická Ekonomie* 65: 26–44. [CrossRef]

Mendelová, Viera, and Mária Stachová. 2016. Comparing DEA and logistic regression in corporate financial distress prediction. Paper Presented at International Scientific Conference FERNSTAT 2016, Banská Bystrica, Slovakia, September 22–23; pp. 95–104.

Mihalovič, Matúš. 2015. The Assessment of Corporate Financial Performance via Discriminant Analysis. *Acta Oeconomica Cassoviensia: Scientific Journal* 8: 57–69. [CrossRef]

Norton, Curtis. L., and Ralph. E. Smith. 1979. A comparison of general price level and historical cost financial statements in the prediction of bankruptcy. *The Accounting Review* 54: 72–87.

Oanh, Nguyen Hoang, and Nguyen Hong Ngoc. 2016. An Application of the Data Envelopment Analysis Method to Evaluate the Performance of Academic Departments within a Higher Education Institution. *Journal of Economics & Development* 18: 71–87. [CrossRef]

Odom, Marcus D., and Ramesh Sharda. 1990. A neural network model for bankruptcy prediction. Paper Presented at 1990 IJCNN International Joint Conference on Neural Networks, San Diego, CA, USA, June 17–21; vol. 2, pp. 163–68. [CrossRef]

Ohlson, James A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18: 109–31. [CrossRef]

Paradi, Joseph C., Mette Asmild, and Paul C. Simak. 2004. Using DEA and worst practice DEA in credit risk evaluation. *Journal of Productivity Analysis* 21: 153–65. [CrossRef]

Paradi, Joseph C., D'Andre Wilson, and Xiaopeng Yang. 2014. Data Envelopment Analysis of Corporate Failure for Non-Manufacturing Firms Using a Slack-Based Measure. *Journal of Service Science and Management* 7: 277–90. [CrossRef]

Park, Seong Ho, Jin Mo Goo, and Chan-Hee Jo. 2004. Receiver Operating Characteristic (ROC) Curve: Practical Review for Radiologists. *Korean Journal of Radiology* 5: 11–8. [CrossRef] [PubMed]

Premachandra, Inguruwatt M., Gurmeet Singh Bhabra, and Toshiyuki Sueyoshi. 2009. DEA as a tool for bankruptcy assessment: A comparative study with logistic regression technique. *European Journal of Operational Research* 193: 412–24. [CrossRef]

Premachandra, Inguruwatt M., Yao Chen, and John Watson. 2011. DEA as a Tool for Predicting Corporate Failure and Success: A Case of Bankruptcy Assessment. *Omega* 3: 620–6. [CrossRef]

Sadjadi, Seyed Jafar, and Hashem Omrani. 2008. Data envelopment analysis with uncertain data: An application for Iranian electricity distribution companies. *Energy Policy* 36: 4247–54. [CrossRef]

Scarlat, Emil, and Camelia Delcea. 2011. Complete analysis of bankruptcy syndrome using grey systems theory. *Grey Systems: Theory and Application* 1: 19–32. [CrossRef]

Shetty, Udaya, TPM Pakkala, and Thathaiah Mallikarjunappa. 2012. A modified directional distance formulation of DEA to assess bankruptcy: An application to IT/ITES companies in India. *Expert Systems with Applications* 9: 1988–97. [CrossRef]

Simak, Paul C. 1997. DEA Based Analysis of Coporate Failure. Master's thesis, Faculty of Applied Sciences and Engineering, University of Toronto, Toronto, ON, Canada.

Sperandei, Sandro. 2014. Understanding logistic regression analysis. *Biochemia Medica* 24: 12–18. [CrossRef]

Štefko, Róbert, Beata Slusarczyk, Sebastian Kot, and Cezary Kolmasiak. 2012. Transformation on steel products distribution in Poland and Slovakia. *Metalurgija* 51: 133–6.

Štefko, Róbert, Beáta Gavurová, and Kristína Kočišová. 2018. Healthcare efficiency assessment using DEA analysis in the Slovak Republic. *Health Economics Review* 8: 1–12. [CrossRef]

Sueyoshi, Toshiyuki, and Mika Goto. 2009. Methodological comparison between DEA and DEA-DA from the perspective of bankruptcy assessment. *European Journal of Operational Research* 188: 561–75. [CrossRef]

Suhányi, Ladislav, and Alžbeta Suhányiová. 2017. Competitiveness of Slovak spa facilities. Paper Presented at 4th International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM 2017, Albena, Bulgaria, August 24–30; vol. 4, pp. 153–60.

Sun, Jie, Hui Li, Qing-Hua Huang, and Kai-Yu He. 2014. Predicting Financial Distress and Corporate Failure: A Review from the State-of-the-art Definitions, Modeling, Sampling, and Featuring Approaches. *Knowledge-Based Systems* 57: 41–56. [CrossRef]

Taffler, Richard J. 1983. The assessment of company solvency and performance using a statistical model. *Accounting and Business Research* 13: 295–308. [CrossRef]

Tone, Kaoru. 2001. A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research* 130: 498–509. [CrossRef]

Trahms, Cheryl A., Hermann Achidi Ndofor, and David G. Sirmon. 2013. Organizational decline and turnaround—A review and agenda for future research. *Journal of Management* 39: 1277–307. [CrossRef]

Wang, Ying-Ming, Kwai-Sang Chin, and Jian-Bo Yang. 2007. Measuring the performances of decision-making units using geometric average efficiency. *Journal of the Operational Research Society* 58: 929–37. [CrossRef]

Zhu, Joe, ed. 2015. DEA Based Benchmarking Models. In *Data Envelopment Analysis. International Series in Operations Research & Management Science*. Boston: Springer, vol. 221. [CrossRef]