Acta Univ. Sapientiae, Economics and Business, **7** (2019) 65–82

DOI: 10.1515/auseb-2019-0005

Modelling the Financial Failure of Romanian Stock Companies

Gergely FEJÉR-KIRÁLY
Sapientia Hungarian University of Transylvania (Cluj-Napoca, Romania)
Department of Business Sciences,
e-mail: fejerkiralygergely@uni.sapientia.ro

Norbert ÁGOSTON
IPHD in Regional Development, Faculty of Business and Economics,
University of Pécs
e-mail: agostonn@yahoo.com

József VARGA
Institute of Finance and Accounting, Faculty of Economic Science,
Kaposvár University
e-mail: varga.jozsef@ke.hu

Abstract. The aim of the present article is to model and predict the financial failure of non-financial companies listed on the Bucharest Stock Exchange. The prediction models are based on the companies' financial reports. The paper emphasizes the importance of processing outlier data and the significance of categorical independent variables. The paper contributes to bankruptcy and corporate financial failure research by presenting a Romanian situation. Results show that the model that uses 3-year financial data prior to the failure has a better accuracy. Several models have been compared, and it was found that using categorical independent variables as explanatory variables increased the accuracy of the models against those that used numerical regressors.

Keywords: corporate failure prediction, parametric modelling, non-parametric modelling, bankruptcy
JEL Classification: G33 – bankruptcy, liquidation

1. Introduction

The present article deals with corporate failure prediction, which is one of the main topics of modern finance (Fantazzini–Figini, 2009). The financial failure of a company can cause significant losses for the business sector (investors, creditors, employees) and for the society as well. For this reason, there is a high demand for developing reliable failure prediction models which are able to predict difficulties in a timely manner (Horta–Camanho, 2013).

In Romania, the highest number of companies struggling with financial problems was registered in the year 2013: almost 28,000 companies (Guda, 2018). However, this number decreased to 8,300 in 2017. From this point of view, Romania is one of the countries with the highest number in Central Eastern Europe (Sielewicz, 2018). Thus, there is a clear need for research on bankruptcy and financial failure in this area (Karas–Režňáková, 2013). One of the objectives of the present study is to fill the research gap in this field. The article focuses on modelling and predicting financial failure and seeking methodological answers to make more accurate predictions.

At this point, it is important to make a difference between financial failure and bankruptcy. Legally speaking, the financial failure process in Romania either starts with the petition on the creditor's behalf or the procedure is started by the company itself. The company enters in a restructuring process, from where there are two possibilities: if it is a successful reorganization, the company survives and if not, the company files for bankruptcy based on Law no 84/2014 on insolvency prevention procedures and insolvency proceedings in Romania.

The present article is organized as follows. Section 2 briefly summarizes the literature of bankruptcy prediction. Section 3 provides a short description of the data and methods used in financial failure prediction. Section 4 summarizes the results. Section 5 concludes and provides some suggestions for future research on financial failure prediction.

2. Literature Review on Bankruptcy Prediction

Several models have been developed in the past decade in order to predict bankruptcy and financial failure. These models have been continuously improved; however, there is no unified and generally accepted model. This is because there is no generally accepted definition of bankruptcy (Constad–Yazdipour, 2011), and researchers tend to interpret the concept of bankruptcy differently when applying different statistical methods (Bellovary et al., 2007).

The first bankruptcy modelling attempt dates back to the 1930s, when Fitzpatrick compared the financial indicators of financially healthy and unhealthy companies (Fitzpatrick, 1932). In 1966, William Beaver published his research results, where the author used discriminant analysis (DA) to predict financial failure one year before bankruptcy. The model was accurate in 90% of the cases (Beaver, 1966). Altman was the first to adopt multivariate discriminant analysis (MDA) for predicting bankruptcy. Altman's failure prediction model was based on 5 financial indicators and had an accuracy of 95% (Altman, 1968).

Altman's model is still used these days; many articles and research papers use it for comparison and reference (Bellovary et al., 2007).

Since 1980, a new model family has been developed in financial prediction, the logistic regression modelling (LRM), which can be linked to Ohlson. Ohlson compared 105 bankrupt companies with 2,058 healthy companies between 1970 and 1976 (Ohlson, 1980). Another form of the LRM is the probit modelling, firstly used by Zmijewski (1984). Since 1990, a new prediction model family was born with the fast developing of the IT sector: the neural-network-based prediction models (Meesier–Hanesen, 1988). From this time on, research turned to artificial intelligence and non-parametric models such as hybrid-, fuzzy-, or generic algorithmic prediction models (Fejér-Király, 2015).

In Romania, bankruptcy research is based on logistic regression. In their article, Brîndescu–Goleț studied 4,327 companies from Timiş County, of which 266 companies were bankrupt. In their research based on logistic regression, they obtained 5 significant variables, and their model had 68% accuracy in predicting the bankrupt company (Brîndescu–Goleț, 2013). Based on Altman's (1966) and Tafflers' (2008) models, Cimpoeru compared 75 healthy and 30 bankrupt companies using logistic regression (Cimpoeru, 2014). In one of his later researches, Goleț used more logistic regression models and tested their differences on a sample of 5,908 companies (of these, 354 were bankrupt) between 2008 and 2012 (Goleț, 2014). The use of neural network modelling in Romania was carried out by Slavici et al. (2015). Their research included 55 samples, 18 financial indicators, and 3 hidden layers. The model had an accuracy of 98% (Slavici et al., 2015). Compared to the international datasets, in Romania, research on bankruptcy is limited due to the low availability of official financial data and fiscal behaviour. Only 1 out of 2 companies struggling with financial difficulties submitted a financial report one year prior to entering the phase of financial failure.

The present article suggests an alternative logistic model based on companies listed on the Bucharest Stock Exchange. According to the international research results, financial data one year prior to financial failure is the most accurate, while financial data from 2 or 3 years before the failure will decrease the accuracy of the model (Beaver, 1966; Altman, 1968). On the other hand, there are studies which confirm that models in which the financial indicators are calculated based on a longer period are more accurate (Berg, 2007). The present study also aims to examine these contrasting findings.

In many studies, one of the problems is that some variables in the data sample are outliers. This has an effect on the accuracy of the models. The present article also wishes to find an answer to this issue. Is it possible to make the logistic regression-based model more accurate in predicting financial failure by transforming these outlier variables? The classification of the logistic regression model is better when the independent numerical variables are categorized (Lázár, 2011). In the present article, these independent variables are transformed into categorical variables; in this way, the accuracy of the logistic regression model can be improved.

3. Data and Methods Used in Predicting Financial Failure

Financial data were collected from the Bucharest Stock Exchange site and from the companies' annual financial reports. The financially failed companies were selected from the National Trade Register Office and were legally deemed as failed companies. 6,142 companies are listed on the Bucharest Stock Exchange, and legal proceedings were launched in the case of 441 companies for financial failure or bankruptcy between 2005 and 2016. From this sample, the research focuses on those companies that financially failed between 2010 and 2016. From this list, financial institutions, banks, and companies without financial data or with 0 sales were eliminated. The remained sample had 65 financially failed companies. The entire sample contains 160 items, out of which 65 are in a state of financial failure and 95 are healthy companies. Taking into consideration all companies listed on the Bucharest Stock Exchange, 5–7% of the companies listed are in a state of financial failure. Compared to this, the sample of the present study includes 60% healthy companies and 40% companies in a state of financial failure. In the literature, healthy companies are usually overrepresented, or a similar number of healthy and unhealthy companies are included in the research (Altman et. al, 1977; Amendola et. al, 2011; Nyitrai, 2015; Virág et al. 2006).

Based on the companies' business activities, the sample used in the present research has a similar distribution to those listed on the Bucharest Stock Exchange. Within the sample, the distribution of industrial and commercial companies is also very similar. *Table 1* shows that industrial companies form the majority, while commercial companies account for the smaller part of the sample.

Table 1. *The Bucharest Stock Exchange (BVB) and data sample distribution*

	BVB	**Data sample**		**Financially failed group**		**Healthy group**	
	distribution	distribution	number	distribution	number	distribution	number
Industry	59.6%	65.0%	104	64.6%	42	65.3%	62
Services	27.6%	24.4%	39	26.1%	17	23.2%	22
Agriculture	4.7%	6.2%	10	6.2%	4	6.3%	6
Commerce	8.1%	4.4%	8	3.1%	3	5.2%	5
Total	**100%**	**100%**	**160**	**100%**	**65**	**100%**	**95**

Source: Bucharest Stock Exchange, own calculation

The data collection was based on the year when financial failure proceedings began. The study used financial reports of the last three years before the financial failure. Therefore, the data come from the period between 2007 and 2015, and the ratio of industrial and commercial companies is similar. The two groups of data are similar in terms of company size and number of employees (between 50 and 249).

The data were analysed using binomial logistic regression. Logistic regression is widely used within the literature, and it was often used at the beginning of research on bankruptcy as well (Bellovary et al., 2007). Moreover, it is the most widely used method in developing credit scorecards (Thomas, 2009). One advantage of this method is that it does not require the normal distribution of independent variables, but it is sensitive if outlier data are present (du Jardin, 2010). A large sample is needed for the success of the procedure; a research needs to have 10 times more observations than dependent variables (Peduzzi et al., 1996). Our 160-item sample fulfils this condition. 16 variables were selected to be used in the model as independent variables, taking into consideration relevant literature and using the companies' yearly financial reports (Bellovary et al., 2007; Virág–Kristóf, 2005). Stock data were examined at their end-of-year closing value because averaging stock value requires more data (reference and base value are also needed) and more work, while it will not improve results significantly (Nyitrai, 2017). The financial indicators and their calculation methods are shown in *Table 2*.

Table 2. *Financial indicators used in the calculation*

Financial indicators	Calculation formula
1. Return on Assets (ROA)	Net income/Assets
2. Return on Sales (ROS)	Net income/Total sales
3. Return on operating results	Operating results/Total sales
4. Liquidity ratio	Current assets/Current liabilities
5. Net working capital	(Current assets – Current liabilities)/ Current liabilities
6. Working capital	Current assets – Current liabilities
7. Equity ratio	Equity/Assets
8. Covered long-term assets	(Equity + Long-term liabilities)/Long-term assets
9. Debt ratio	Total liabilities/Assets
10. Asset Turnover	Total sales/Assets
11. Days Sales Outstanding (DSO)	Total sales/Accounts receivable
12. Inventory ratio	Inventory/Current assets
13. Current liabilities ratio	Current liabilities/Total liabilities
14. Change in total sales	Total sales_1/Total sales_0
15. Employees	Average employees
16. Company foundation year	Years since foundation of company

It must be noted that it is necessary to prepare the data before modelling, on the one hand, because logistic regression can lead to distorted results if outlier data is still present and, on the other hand, as we also need to consider multicollinearity that might distort the results. To identify multicollinearity, we used the variance inflation factor, the general formula of which can be seen below:

$$\mathrm{VIF_j} = \frac{1}{\mathrm{T_j}}, \tag{1}$$

where T_j is the tolerance indicator, and $T_j = 1 - R^2_j$, which shows how many unexplained variables are left after *j*-th independent variable. Therefore, in formula (1), if the *j*-th independent variable is linearly independent from other variables, then the value of the indicator is one, while in the case of an extreme multicollinearity the value of the indicator is infinite. The general threshold value is 5; if the indicator is greater or equal to 5, then there is a strong multicollinearity (Kovács, 2008).

The analysis revealed that there is multicollinearity between profitability indicators such as return on sales, return on operating results, equity, or debt ratio. For this reason, only one pair of indicators was used at a time for modelling, making it easier to identify the best performing indicators to be used in the final model, using a single iteration.

There is no consensus in the literature about identifying and handling outlier data. However, most seem to agree about the fact that outlier data can have a significant impact on parametric as well as non-parametric tests (Zimmerman, 1994). Winsorization is often used in bankruptcy research in order to handle outliers (Wu et al., 2010; Mansi et al., 2010; Araujo et al., 2012), and therefore the present study also made use of this method. Outliers were examined with the help of interquartile range, and it showed that almost all indicators had outliers, in some case even extreme outliers. Reducing the size of an already small sample was not an option, and so we did not exclude outliers. Outlier values were substituted by one and a half times the value of the interquartile range. Statistical analyses also include identifying outliers which are three times greater than the average range of dispersion (Li–Sun, 2011).

Using random sampling, the whole sample was divided into 80% test and 20% control groups in order to examine the predicting ability of the models. It is important to note that the number of failed companies within the test sample needs to have at least 50 items in order to perform multivariate statistical analysis (Engelman et al., 2003). Our sample fulfils this criterion. To create the final models, the Enter method was applied using the results of several trial tests. For the graphic evaluation of the models' performance and the calibration plot, the ROC (Receiver Operating Curve) was used. The ROC shows the accuracy of classification within the model compared to the real classification, including all cut-off values. Generally, the horizontal axis presents the probability of false alarm, while the vertical axis presents the ratio of correctly classified failed companies. There is a 45-degree line, which illustrates the chance diagonal and the level at which the model is not acceptable. The accuracy of the model can be measured with the help of the area below the ROC curve. The higher the value of the AUC, the better the model. While a value of 0.5 AUC refers to chance, a value of 1 indicates a perfect model.

4. Results

The study aims to model the 3 years before the companies' financial failure, where the point of reference is the state of financial failure (marked with a "t"). All three models showed a significant correlation between the dependent and independent variables at a level of 0.05 based on the Chi-square test. The Hosmer–Lemeshow tests showed that the models fit the data. The indicators included in the final models are shown in *Table 3* below.

Table 3. *Financial indicators present in the final model (years prior financial indicators – year t-n)*

Financial indicators	year t-1	year t-2	year t-3
ROA	*	*	*
DSO	*	*	*
Equity ratio	*	*	
Working capital	*	*	
Changes in total sales	*		
Debt ratio			*

** means that the financial indicator is present in the final model* *Source: own calculations*

As far as the final models are concerned, there is a relatively consistent result regarding the significance of the indicators. Return on assets proved to be one of the most stable indicators, which had a significant impact on all three models. This was hardly a surprising result as Bellovary et al. (2007) concluded in their study that the ROA indicator is the most commonly used indicator in the history of bankruptcy research. The DSO is also present in all three models as a significant independent variable. Equity ratio had a significant impact on the first and second year before the state of financial failure. Debt ratio was significant only in the third year before the state of financial failure. Changes in the total sales indicator were only significant in the first year before the state of financial failure. Only one liquidity indicator, the working capital, contributed to the prediction of financial failures in models *t-1* and *t-2*. The ROS seemed to be significant in all cases when the ROA indicator was not included in the model. This is because of the strong connection between the two indicators, but it is worth mentioning that using ROS resulted in less accurate models, and therefore it was left out of the final models. Return on operating results was similarly in a strong correlation with both ROS and ROA. Results showed that most of the indicators used in modelling had no effects on the financial failure.

In the case of the binomial logistic regression variable, weight is shown by variable coefficient. *Table 4* takes the *t-1 model* as example, and by examining the B coefficient it can be seen that an increase in any financial indicator decreases the chance of financial failure, while the other indicators remain unchanged. The exception to this rule is the working capital indicator, the increase of which also increases the chance of financial failure. The likelihood ratio, or Exp(B) shows how an increase of regressors by one unit increases the chance of financial failure. A likelihood ratio greater than 1 increases, while a likelihood ratio smaller than 1 decreases the chance of financial failure. Looking at *Table 4*, it can be observed that a one-unit increase in the working capital increases the likelihood of financial failure by one unit. The ROA indicator has the most powerful effect on the likelihood of financial failure – a unit increase in ROA decreases the likelihood of financial failure with almost 100%, leaving other indicators unchanged. Equity ratio and change in total sales work in the same way though their impact is less powerful. The DSO has the least effect on financial failure, where a one-unit increase, ceteris paribus, increases the likelihood of financial failure by 0.825 times, meaning a 17.5% decrease. Due to the limited space, the parameters and formulas for models *t-2* and *t-3* can be found in the *Annex*.

Table 4. *The parameters of one year prior to failure (t-1) model*

Financial ratio	B	S.E.	Wald	Sig.	Exp(B)
ROA (x_1)	-11.137	2.982	13.949	0.000	0.000
Equity ratio (x_2)	-4.421	1.053	17.629	0.000	0.012
DSO (x_3)	-0.193	0.084	5.285	0.022	0.825
Changes in total sales (x_4)	-1.899	0.872	4.741	0.029	0.150
Working capital (x_5)	0.000	0.000	11.360	0.001	1.000
Constant	3.356	1.023	10.765	0.001	28.666

Source: own research and calculations with SPSS

Formula for model *t-1*:

$$P(\text{failed}) = \frac{1}{1+e^{-(3.356-11.137\cdot \text{ROA}-4.421\cdot \text{equity ratio}-0.193\cdot \text{DSO}-1.899\cdot \text{change in total sales})}} \qquad (2)$$

The models' ability to predict financial failure can be tested on the control groups, in the case of which all three models show a lower ranking accuracy

(*Table 5*). In model *t-1*, it is lower by 1.5%, in model *t-2*, by 6.2%, and in model *t-3* by 4.6% compared to the test group, which is normal. For this reason, the models prove to be adequate in predicting financial failure. At the same time, it can be observed that in spite of the large proportion of the test group, we cannot speak of overtesting, which in turn reinforces the predicting ability of the models.

Table 5. *The main characteristics of the three models*

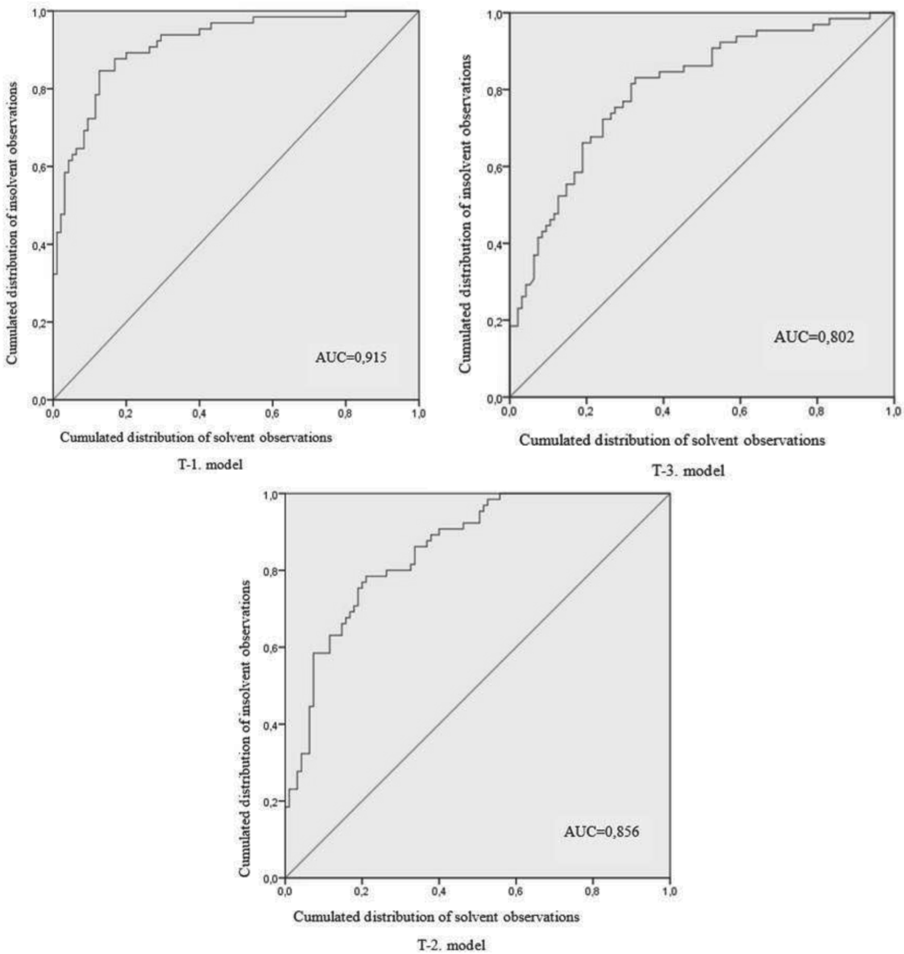
cut-off value = 0.50	**t-1. model**	**t-2. model**	**t-3. model**
No. of significant indicators	5	4	3
Hosmer–Lemeshow Sig.	0.935	0.876	0.858
Nagelkerke R^2	0.647	0.505	0.396
Test group accuracy	82.8%	78.1%	73.4%
Control group accuracy	81.3%	71.9%	68.8%

Source: own research and calculations with SPSS

If we go back in time from the moment of financial failure, we can find less significant indicators, weaker explanatory power, lower accuracy, and the AUC is also becoming smaller. While in the first year before the state of financial failure there are five significant indicators, with an explanatory power of 64.2%, in the third year, there are only three significant indicators with an explanatory power of 39.6%. This means that on a short term the results are more accurate, and the further we go back in time, the more the accuracy of models and predictions are decreased, which is perfectly realistic and supports the reliability of the models (*Table 5*).

Examining the performance of the models based on ROC, we can find that in the case of model *t-1* there is a significant deviation from the 45-degree line, which shows an excellent model. Model *t-2* proves to be good, while the AUC value of the *t-3* model can also be considered as good based on *Figure 1* (0.802).

Calibration of the cut-off value revealed that in the case of model *t-1* the optimal cut-off point is 0.41, in model *t-2* 0.39, and in model *t-3* 0.30. Going further back in time from the state of financial failure, the calibration of the cut-off value seemed to be less advantageous: in the case of model *t-1*, it brought about an accuracy increase of 3.1% in the test and control groups, while in model *t-3* this increase was only of 1.6% in the test group, but it had no effect on the control group.



Source: own research and calculations with SPSS

Figure 1. *The ROC-curve of the three models*

Multi-Year Model vs Model t-1

Besides creating the 3 models, one of the aims of the present paper is to assess the reliability and forecasting ability of the model which includes all financial indicators three years prior to the financial failure. After several testing procedures, the best model contains 6 independent variables. According to the Hosmer–Lemeshow test, this model fits our data, and according to the Nagelkerke R^2 the selected independent variables have a confidence level of 72.4% in predicting financial failure. Its accuracy proved to be the best for the test group and the control group alike. Compared to the *t-1* model, it has a better classifying ability for the test group, but it performs worse in the case of the control group. The comparison is illustrated in *Table 6* below.

Table 6. *Key features of the multi-year model vs model t-1*

cut-off value = 0.50	Multi-year model	t-1. model
No. of significant indicators	6	5
Hosmer–Lemeshow Sig.	0.825	0.935
Nagelkerke R^2	0.724	0.647
Test group accuracy	89.1%	82.8%
Control group accuracy	75.0%	81.3%
AUC	0.921	0.915

Source: own research and calculations with SPSS

The Importance of Managing Outliers

The paper also seeks to find and answer to the question whether winsorizing outliers improves the models' accuracy, classifying ability, and explanatory power. In order to run this test, the whole sample of 160 items was used without dividing it into test and control groups. Managing outliers was done according to the procedure mentioned in the theoretical part of the paper; outlier values were substituted by one and a half times the value of the interquartile range. Managing outliers was performed on all three models created previously, and the cut-off value was set at 0.50.

Looking at the raw data with outlier values in the first year before the financial failure, we can see that the number of significant independent variables decreased to four because the ROS indicator was not significant. Based on the Nagelkerke R^2 test, the explanatory power of the model was at a confidence level of 56.6% with the outlier data and 63.2% after managing outliers. According to the Hosmer–Lemeshow test, both models fit de data. Classification accuracy was better after managing outliers, and based on the area under the ROC curve the data without outliers resulted in an excellent model, whereas with the outlier data it is an acceptable model. *Table 7* summarizes the arguments above.

Table 7. *Results before and after managing outliers in year t-1*

Year t-1	No. of significant indicators	Match (%)			Nagelkerke R^2	Hosmer–Lemeshow test (Sig.)	AUC
		Healthy	Financially failed	Total			
Without outliers	5	87.4	75.4	82.5	63.2%	0.601	0.913
With outliers	4	86.3	70.8	80.0	56.6%	0.436	0.887

Source: own research and calculations with SPSS

Examining the second year before the financial failure based on the existing model, both sets of data contain four significant variables, and they fit based on the Hosmer–Lemeshow test. Results in *Table 8* show that the explanatory value of raw data containing outliers is 4.3% less compared to the database without outliers. Classification accuracy is higher in the sample managed for outliers, but the raw sample yielded better results in terms of identifying healthy companies. Based on their position under the ROC curve, both samples of data can be considered good.

Table 8. *Data with and without outliers in year t-2*

Year t-1	No. of significant indicators	Match (%)			Nagelkerke R^2	Hosmer–Lemeshow test (Sig.)	AUC
		Healthy	Financially failed	Total			
Without outliers	4	84.2	64.6	76.3	47.2%	0.106	0.863
With outliers	4	86.3	60.0	75.6	42,9%	0.505	0.842

Source: own research and calculations with SPSS

In the third year before the financial failure, out of three independent variables in the raw data, the ROA and the DSO were only acceptable on a 10% significance level, but they remained in the model. We can see in *Table 9* that both samples fit well according to the Hosmer–Lemeshow test. There is a significant difference in their explanatory power in favour of the sample without outliers based on the R^2 test. The total classification accuracy of the sample without outliers is higher, while the classification accuracy of raw data is more precise. Based on the AUC, the data without outliers results in a good model, while raw data yield a mediocre/average model.

Table 9. *Data with and without outliers in year t-3*

Year t-1	No. of significant indicators	Match (%)			Nagelkerke R^2	Hosmer–Lemeshow test (Sig.)	AUC
		Healthy	Financially failed	Total			
Without outliers	3	81.1	60.0	72.5	35.9%	0.760	0.809
With outliers	3	86.3	46.2	70.0	18.1%	0.669	0.783

Source: own research and calculations with SPSS

The Importance of Categorical Independent Variables

Another interesting aspect was the possibility of obtaining better classification results by using categorical independent variables instead of numerical variables. In order to achieve this, we transformed the sample containing outliers to categorical values. Each financial index with a value in the first quartile was assigned into Category 1, Category 2 was made up of values between the first and second quartile, the median and upper quartile were assigned into Category 3, and values in the upper quartile were placed into Category 4. For the modelling process, the entire 160-item sample was used, taking into consideration the variables of the previous models using a cut-off value of 0.50. Testing the first year before the financial failure and using categorical variables, we revealed that only three were significant: ROA, equity ratio, and DSO. The Hosmer–Lemeshow test confirms that the models fit well; the explanatory power of the categorical data sample is 2.1% higher compared to the numerical data sample. Classification accuracy was higher when using categorical variables in both classes, and total classification accuracy is the highest among all models in the current study.

In the second year before the financial failure (*t-2*), the categorical variables allowed for a higher total classification accuracy and a higher explanatory power. Unfortunately, the Hosmer–Lemeshow test revealed that the model does not fit the data well, wherefore it is unacceptable.

In the third year before the financial failure, the categorical values perform better again, the model fits the data, and the explaining power is 9% higher compared to the numerical data. Total classification accuracy using categorical values at 75% is good and higher than its numerical counterpart. The summary of the models' results is shown in *Table 10*.

Table 10. *Models using categorical vs numerical variables*

Year t-1	No. of significant indicators	Match (%)			Nagelkerke R^2	Hosmer–Lemeshow test (Sig.)
		Healthy	Financially failed	Total		
Numerical (year t-1)	5	87.4	75.4	82.5	63.2%	0.601
Categorical (year t-1)	3	88.4	80.0	85.0	65.3%	0.221
Numerical (year t-2)	4	84.2	64.6	76.3	47.2%	0.106
Categorical (year t-2)	3	83.2	75.4	80.0	51.5%	0.027
Numerical (year t-3)	3	81.1	60.0	72.5	35.9%	0.760
Categorical (year t-3)	3	80.0	67.7	75.0	44.6%	0.697

Source: own research and calculations with SPSS

5. Conclusions

The final conclusion of the present paper is that it is possible to predict the financial failure of non-financial companies listed on the Bucharest Stock Exchange. The most accurate predictions were calculated using short-term, one-year data. The study confirms the validity of the general approach on financial failure prediction models, more specifically that they work as a short-term, early warning/forecasting system. It is worth mentioning that models built on data three years prior to the financial failure were also acceptable, and this way it can be stated that the phenomenon of financial failure is not a single event but rather the end result of a financially problematic period.

Modelling was based on examining data from three consecutive years before the financial failure occurred, and a model containing data from all three years was also created. The results of the study show that the multi-year model proved to be better than our best model built on data from one year before the financial failure (*model t-1*). This led us to the conclusion that the most reliable models need to be based on more data than those available in the year before the failure.

Calibration of the cut-off value only improved accuracy by approximately 3 percentile points; furthermore, in the case of the test sample, there was no increase in classification accuracy in the long term. Fine-tuning the cut-off value did not yield outstanding results; however, the improved measuring, those couple of per cents could be of key importance when making decisions.

The classification accuracy of the logistic regression improved after winsorizing outliers, transformations based on the one and a half times the interquartile range did not cause a significant loss of data, and in each case better models were obtained compared to the scenario of using raw data with outliers. Therefore, it can be stated that outliers outside the one and a half times the interquartile range do not contain vital information; rather they have a distorting effect. Further studies could examine the effect of more strict or lenient winsorizing on the results.

Compared to the models using numerical data, a higher classification accuracy was achieved by grouping independent variables into categories. This process is also a good outlier management tool. Thus, a further question is to be answered: does it make sense to place numerical variables in more categories than the existing four in order to achieve better accuracy?

The results of the present study have their limitations, but at the same time they might point out further research directions. The limitations of the study are that it is heavily reliant on the public data available on the Bucharest Stock Exchange website; therefore, the conclusions can be formulated only for a part of the listed companies. The viability of this small-scale study could be confirmed by a similar study carried out on the national level.

A significant distortion comes from the fact that the database is from the years 2007–2015, and we cannot overlook the effect of this. The global economic crisis of 2008 is also a distorting factor because in addition to normal economic operations some companies failed due to the domino effect, which caused economic difficulties. Taking the years affected by the crisis out of the data range and rebuilding the models is the object of a future research.

Out of more than a dozen indicators, only 3–5 proved to be of significant influence on the financial failure – among these are the ROA and DSO, which proved to be very stable.

They are included in each of the models as a significant explanatory variable. ROA had a substantial impact on the development of financial failure, just as we expected. Only one liquidity indicator proved to be significant, the working capital, which contributed to the estimation of financial failure in models *t-1* and *t-2*. The debt ratio turned out to be significant only in the third year before the financial failure (model *t-3*) and showed a strong causality of financial failure. This last result is believed to indicate that the debt ratio is an early warning sign of financial failure, but this needs to be confirmed by further studies.

With a single exception, only static indicators were used, and for this reason in the future it might be useful to create a model with dynamic indicators, where we could identify the rank between dynamic and static indicators.

Appendix

Appendix 1. *The parameters and formulas for the t-2 model*

Financial ratio	B	S.E.	Wald	Sig.	Exp(B)
ROA	-10.408	2.825	13.572	0.000	0.000
Equity ratio	-4.444	1.009	19.412	0.000	0.012
DSO	-0.245	0.076	10.447	0.001	0.782
Working capital	0.000	0.000	8.760	0.003	1.000
Constant	2.563	0.743	11.893	0.001	12.980

Source: own research and calculations with SPSS

Formula for model *t-2*:

$$P(\text{failed}) = \frac{1}{1+e^{-(2.563-10.408\cdot\text{ROA}-4.444\cdot\text{equity ratio}-0.245\cdot\text{DSO})}}$$

Appendix 2. *The parameters and formulas for the t-3 model*

Financial ratio	B	S.E.	Wald	Sig.	Exp(B)
ROA	-10.210	3.584	8.117	0.004	0.000
Debt ratio	4.002	0.904	19.593	0.000	54.681
DSO	-0.263	0.073	13.028	0.000	0.768
Constant	-0.867	0.450	3.713	0.054	0.420

Source: own research and calculations with SPSS

Formula for model *t-3*:

$$P(\text{failed}) = \frac{1}{1+e^{-(-0.867-10.210\cdot \text{ROA}+4.002\cdot \text{debt ratio}-0.263\cdot \text{DSO})}}$$

References

Agarwal, V.; Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance* 32: 1541–1551.

Altman, E. I. (1968). Financial ratios, discriminant analysis and prediction of corporate bankruptcy. *The Journal of Finance* 23(4): 589–609.

Altman, E. I.; Haldeman R. G.; Narayanan P. (1977). ZETA Analysis: A new model for bankruptcy classification. *Journal of Banking and Finance* 1(1): 29–54.

Amendola, A.; Restaino, M.; Sensini, L. (2011). Variable selection in default risk models. *The Journal of Risk Model Validation* 5(1[Spring]): 3–19.

Araujo, A.; Ferreira, R.; Funchal, B. (2012). The Brazilian bankruptcy law experience. *Journal of Corporate Finance* 18: 994–1004.

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

Bellovary, J.; Giacomino, D.; Akers, M. (2007). A review of bankruptcy prediction studies: 1930–present. *Journal of Finance Education* 33: 1–42.

Berg, D. (2007). Bankruptcy prediction by generalized additive models. *Applied Stochastic Models in Business and Industry* 23: 129–143.

Brîndescu, D.; Goleț, I. (2013). Prediction of corporate bankruptcy in Romania through the use of logistic regression. *The Annals of the University of Oradea, Economic Sciences* XXII(July): 976–986.

Cimpoeru, S. (2014). Scoring functions and bankruptcy prediction models – Case study for Romanian companies. *Procedia Economics and Finance* 10: 217–226.

Constand, L. R.; Yazdipour, R. (2011). Firm failure prediction models: A critique and a review of recent developments. In: *Advances in Entrepreneurial Finance:*

With applications from behavioral finance and economics. Springer Science and Business Media. 185–204.

du Jardin, Philippe (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variables election techniques on model accuracy. *Neurocomputing* 73(10–12): 2047–2060.

Engelman, B.; Hayden, E.; Tasche, D. (2003). Measuring the discriminant power of rating systems. *Discussion Paper, Series 2: Banking and Financial Supervision*. Deutsche Bundesbank, Frankfurt.

Fantazzini, D.; Figini, S. (2009). Random survival forests models for SME credit risk measurement. *Methodology and Computing in Applied Probability* 11(1): 29–45.

Fejér-Király, G. (2015). Bankruptcy prediction: A survey on evolution, critiques, and solutions. *Acta Universitatis Sapientiae, Economics and Business* 3: 93–108.

Fitzpatrick, P. (1932). *A comparison of the ratios of successful industrial enterprises with those of failed companies*. Washington: The Accountants' Publishing Company.

Goleț, I. (2014). Symmetric and asymmetric binary choice models for corporate bankruptcy. *Procedia – Social and Behavioral Sciences* 124: 282–291.

Guda, I. (2018). *Studiul insolvențelor România 2018*. Coface – February 2018.

Horta, I. M.; Camanho, A. S. (2013). Company failure prediction in the construction industry. *Expert Systems with Applications* 40(16): 6253–6257.

Karas, M.; Režňáková, M. 2013. Bankruptcy prediction model of industrial enterprises in the Czech Republic. *International Journal of Mathematical Models and Methods in Applied Sciences* 5(7): 519–531.

Kovács, P. (2008). A multikollinearitás vizsgálata lineáris regressziós modellekben. *Statisztikai Szemle* 86(1): 38–64.

Lázár, E. (2001). Ároptimalizálás – Ökonometriai modellezésen és az empirikus keresleti függvény meghatározásán alapuló árazási technikák. *Budapest Management Review/Vezetéstudomány* 42(6): 56–68.

Li, H.; Sun, J. (2011). On performance of case-based reasoning in Chinese business failure prediction from sensitivity, specificity, positive and negative values. *Applied Soft Computing* 11(1): 460–467.

Mansi, S.; Maxwell, W.; Zhang A. (2010). Bankruptcy prediction models and the cost of debt. *The Journal of Fixed Income* Spring 21(4): 25–42.

Messier, Jr., W.; Hansen, J. (1988). Inducing rules for expert system development: An example using default and bankruptcy data. *Journal Management Science* 34(12): 1403–1415.

Nyitrai, T. (2015). Hazai vállalkozások csődjének előrejelzése a csődeseményt megelőző egy, két, illetve három évvel korábbi pénzügyi beszámolók adatai alapján. *Budapest Management Review/Vezetéstudomány* 46(5): 55–65.

(2017). Stock és flow típusú számviteli adatok alkalmazása a csődelőrejelző modellekben. *Vezetéstudomány/Budapest Management Review* XLVIII(8–9): 68–77.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18(1): 109–131.

Peduzzi, P.; Concato, J.; Kemper, E.; Holford, T. R.; Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology* 49(12): 1373–1379.

Sielewicz, G. (2018). *Central & Eastern Europe insolvencies: Good times come to an end*. Coface Economic Publications, September.

Slavici, T.; Maris, S.; Pirtea, M. (2015). Usage of artificial neural networks for optimal bankruptcy forecasting. Case Study: European small manufacturing enterprises. *Quality & Quantity* 50(1): 385–398.

Thomas, Lyn C. (2009). *Consumer credit models: Pricing, profit and portfolios.* Oxford: *Oxford University Press.*

Virág, M.; Kristóf, T. (2005). Az első hazai csődmodell újraszámítása neurális hálók segítségével. *Közgazdasági Szemle* 52(February): 144–162.

(2006). Iparági rátákon alapuló csődelőrejelzés sokváltozós statisztikai módszerekkel. *Vezetéstudomány* 37(1): 25–35.

Wu, Y.; Gaunt, C.; Gray, S. (2010). A comparison of alternative bankruptcy prediction models. *Journal of Contemporary Accounting & Economics* 6: 34–45.

Zimmerman, D. (1994). A note on the influence of outliers on parametric and nonparametric tests. *The Journal of General Psychology* 121(4): 391–401.

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