WHAT DO POST-COMMUNIST COUNTRIES HAVE IN COMMON WHEN PREDICTING FINANCIAL DISTRESS?

Madalina Ecaterina Popescu, Victor Dragotă*

Abstract

Business failure prediction is an important issue in corporate finance. Different prediction models are proposed by financial theory and are often used in practice. Their application is effortless, selecting only few key inputs with the greatest informative power from the large list of possible indicators. Our paper identifies the financial distress predictors for 5 post-communist countries (Bulgaria, Croatia, the Czech Republic, Hungary and Romania) based on information collected from the Amadeus database for the period 2011–2013 using CHAID decision trees and neural networks. We propose a short list of indicators, which can offer a synthetic perspective on corporate distress risk, adapted for these countries. The best prediction models are substantially different from country to country: in the Czech Republic, Hungary and Romania the flow-approach indicators perform better, while in Bulgaria and Croatia – the stock-approach indicators. The results suggest that the extrapolation of such models from one country to another should be made cautiously. One interesting finding is the presence of the ratios *per* employee as predictors of financial distress.

Keywords: financial distress, predictors, prediction models, post-communist countries, CHAID decision trees, neural networks

JEL Classification: G33, L25, C53

1. Introduction

Business failure prediction has been an issue of interest in corporate finance literature over the decades. Especially after the financial crisis of 2007–2009, the topic was even more intensively studied (see, among others, Altman *et al*., 2014). If the financial distress prediction is reliable, managers can effectively initiate remedial measures, whereas investors can adjust their investment strategies. Appropriate prediction models can help firms in preventing financial distress. Financial distress prediction models appear as statistical solutions designed to predict if companies will face some forms of financial distress (*e.g*. loan defaults or non-payment of creditors) or even fail (Keasey and Watson, 1991).

In our paper, we use Chi-square Automatic Interaction Detector classification tree models (hereafter, CHAID) and neural networks (hereafter, NN) for finding the best predictors of failure. CHAID are non-linear architectures that perform well with large data

* Madalina Ecaterina Popescu, Bucharest University of Economic Studies, Faculty of Economic Cybernetics, Statistics and Informatics, and INCSMPS (madalina.andreica@csie.ase.ro);
Victor Dragotă, Bucharest University of Economic Studies, Faculty of Finance, Insurance, Banking and Stock Exchange, and CEFIMO (victor.dragota@fin.ase.ro).
The authors wish to thank Ingrid Mihaela Dragotă, Maria-Isadora Lazăr, Ruxandra Trifan, and Elena Țilică, for their useful comments. The remaining errors are ours.

in a short time being easy to build and to interpret. NN are effective in distress prediction, but they tend to require more time to accomplish iterative processes and imply difficulties in model interpretation.

We collected public financial information for the period 2011–2013 from the Amadeus database, for 5 post-communist countries: Bulgaria, Croatia, the Czech Republic, Hungary and Romania. Having over four decades of similar political backgrounds (between the end of the Second World War and the fall of communist regimes), one might assume similarities among these countries in terms of financial development. However, differences between them can derive from factors like the method of privatization or the speed of the economic reform, or even from some cultural factors.

Financial analysis provides a large number of indicators recommended for different purposes (Jakubík and Teplý, 2011; Scarlat *et al.*, 2012). However, in some instances, they can provide conflicting solutions in the decision-making process. Thus, this large number of indicators can be not only redundant, but dangerous for financial management. For this reason, in this study we propose a short list of indicators, which can offer a synthetic perspective on corporate distress risk adapted for post-communist countries[1].

Insolvency can be approached based on a stock-philosophy or on a flow-philosophy (Ross *et al.*, 2010). Our study identifies, which of these ones predicts better financial distress. For some countries, financial distress is better predicted using the former (Bulgaria and Croatia), while in others the latter performs better (the Czech Republic, Hungary and Romania). The results suggest that the extrapolation of such models from one country to another should be made cautiously: the best prediction models can be substantially different. Moreover, even for the same country, differences appear from one year to the next.

Through this paper we provide a list of indicators suitable for failure prediction for the 5 post-communist countries. Some studies have adopted a similar methodology for financial prediction on newer market economies, but only for the listed companies (Geng *et al.*, 2015; Jaba *et al.*, 2016; Tudor *et al.*, 2015). Our paper conducts an empirical research on a sample of private limited companies. As the countries included in our study have a low level of market capitalization in GDP[2], a study on private limited companies can offer a better understanding of the economic environment. In the same perspective, the countries analysed in our study have bank-oriented systems, and not capital market-ones, so a deeper analysis of unlisted companies can be more useful.

For practitioners, this study can offer some clues regarding the indicators, which can be used as best predictors for failure, adapted for the 5 countries included in our sample. For academics, we suggest a more circumspect attitude in generalizing some patterns

1 With a relatively similar purpose, Welc (2016) proves that indebtedness ratio and current ratio identify the healthy and the bankrupt firms (with a one-year-ahead forecast horizon) for the Polish capital market.

2 According to http://data.worldbank.org/, stock market capitalization of listed domestic companies (% of GDP) was 14.49% in Bulgaria (2011), 36.23% in Croatia (2011), 17.39% in Czech Republic (2008), 14.73% in Hungary (2013) and 7.57% in Romania (2011).

identified for other countries, based only on financial indicators. Other factors seem to matter in failure prediction, too.

The rest of the paper is structured as follows. Section 2 provides the theoretical background. Section 3 is dedicated to the data analysis, where all aspects regarding data collection and matching, financial ratios description, as well as the modelling phase are presented. The main findings are examined in Section 4. Section 5 concludes.

2. Theoretical Background

Academics and practitioners agree that distress prediction is an important issue in finance. The concept of *financial distress* is associated with different situations, in which firms face financial difficulty in paying its debts, but no standard definition is generally accepted (Ross *et al*., 2010; Wruck, 1990).

Apparently, financial distress can be easily connected with a high leverage. However, a high leverage is not mandatory associated with a high probability of financial distress (Harris and Raviv, 1991). Moreover, some empirical evidences suggest that leverage is inversely correlated with the probability of failure (Castanias, 1983). In addition, some theories (*e.g*. Ross, 1977) state that a high leverage can signal a good performance. The relation between leverage and financial distress is more complicated by the cumulative action of different impact factors, like the company's profitability, size, life cycle, growth opportunities, *etc*. (Harris and Raviv, 1991; Myers, 2001; Buus, 2015).

Financial indicators used in financial distress analysis can be integrated in a stock or in a flow-approach. The first one concentrates on assets: if their market value is higher than the level of debts, the company is still solvable because it can liquidate some of its assets to cover its debt service[3]. This approach is not too much concerned on company's future cash flows. In some cases, it can suggest an over-optimistic figure of the company; for instance, one firm can be vulnerable to a generalized fall of assets' prices, as in the case of the crisis of 2007–2009. Solvency and liquidity ratios are relevant for this analysis.

In the flow-approach, the concern is if the company's current obligations can be paid through its cash flows. It is assumed that the past performance indicators of the company are good proxies for the future. Different ratios, like the rates of return, margin rates, efficiency rates and debt ratios are typically used in this analysis.

It can be argued that these approaches are over-simplistic because of their restrictive assumptions. In fact, forecasting different indicators (including here the probability of default) is a common practice in finance (Ekinci, 2016). For instance, Monte Carlo simulations can be used in forecasting the probability of an investment project to generate a positive net present value (*e.g*. Dragotă and Dragotă, 2009). However, their application over a great number of companies can become expensive. Also, monitoring the complex processes involved in these cases can be challenge.

3 In this approach, often, the assets' market values are considered to be equal to their book values and to their liquidation values. Both assumptions are simplifications.

Different statistical techniques have been used to predict business failure. They tend to fall in three main groups: Multivariate Discriminant Analysis (MDA) (Altman, 1968), dichotomous econometric models (Ohlson, 1980; Shumway, 2001) and, in recent years, machine learning methods based on NN and decision trees.

NN imply a nonlinear mathematical approach inspired by the biological networks of the human nervous system. Some empirical findings suggest that NN often outperforms conventional statistical models such as MDA or logit models in predicting financial distress (Jain and Nag, 1998). In contrast, Zheng and Yanhui (2007) use decision trees for financial distress prediction and argue in favour of the CHAID models, which are less difficult to build and to interpret compared to NN or to statistical models, where the patterns need to be linearly separable and samples are assumed to follow a multivariate normal distribution.

More recent studies have emphasized the need of improved data mining techniques and cross-validation methods to predict business failure. For instance, Geng *et al.* (2015) used 3 different time windows and repeated random sub-sampling validation to build financial distress warning models. Their main findings suggest that NN performs better in predicting financial distress for listed companies. Different studies on distress prediction refer to the countries included in our sample. A summary of empirical findings is presented in Table 1.

A classification tree is a predictive model built in the process of learning from instances, which can be viewed as a tree. Specifically, in the tree, each branch is a classification question and the leaves are partitions of the dataset with their classification. Because of their ability to easily generate consistent rules for segmentation of the original database, decision trees can become an efficient method in predicting financial distress. In the category of decision tree algorithms, CHAID has the advantage of generating non-binary trees (Tudor *et al*., 2015).

In our case, the input attributes of the CHAID model consist of the 24 initial financial ratios, while the target attribute is the binary variable indicating, whether the firm is distressed or not. For each input attribute the algorithm uses a Pearson chi-square test to find the pair of values that is least significantly different with respect to the target attribute. For each selected pair, CHAID checks then if p-value obtained is greater than a merge threshold of 0.05. If the answer is positive, it merges the values and searches for an additional potential predictor. Moreover, a Bonferroni correction was used to account for multiple testing using SPSS software (Hsu, 1996). The process of building and testing the CHAID models is briefly described in Figure 1.

NN models are an alternative to statistical techniques for distress prediction. Unlike parametric statistical models, they do not require a specification for the functional relationship between variables. They are adaptive and respond to structural changes in the data generating process in manners that parametric models cannot. The purpose of the model is to capture the causal relationships between dependent and independent variables in the data set under consideration. NN have the ability to construct nonlinear models by scanning the data for patterns.

Table 1 | Brief Selection of Empirical Research on the 5 Post-Communist Country Panel

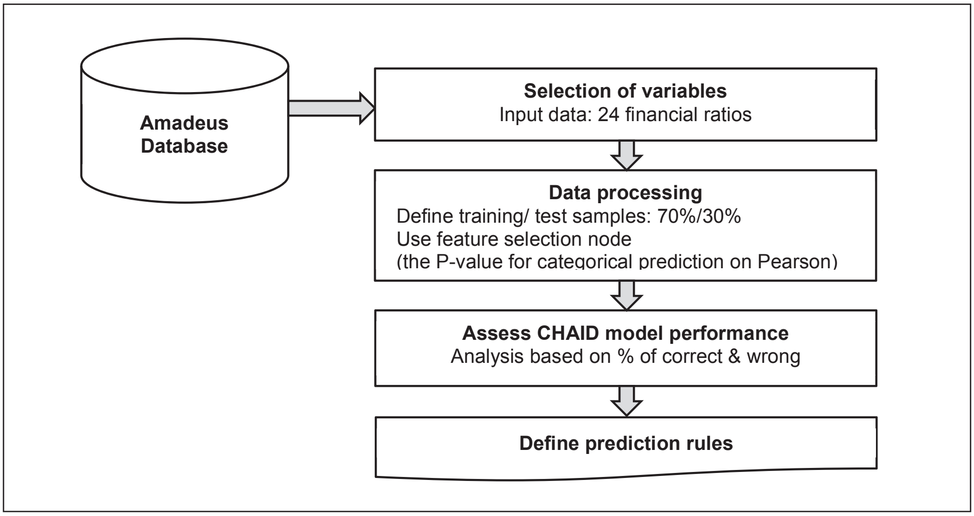
Study	Country	Best predictors of financial distress	Main results
Virág and Hajdu (1996)	HU	Quick liquidity ratio, Cash flow / Total debts and Current assets / Total assets	Proposed bankruptcy models using MDA and logistic regression.
Anghel (2002)	RO	Net profit / Total income, Cash flow / Total assets, Total Debt / Total assets and (Obligations / Sales) · 360 (in days)	Estimated a MDA model score for the period 1994–1998
Virág and Kristóf (2005)	HU	The NN was built based on all available indicators.	NN outperform MDA and logistic regression models.
Šarlija and Jeger (2011)	HR	ROE, Operating revenues / operating expenses, LT Assets / (Equity + LT Liabilities), ST Liabilities / Total Assets ratio and Equity / sales	Estimated logistic models based on financial data of privately-owned SME in HR from 2006–2009.
Altman *et al.* (2014)	29 EU countries (including BG, RO, HU, HR and CZ) and 3 non-EU countries	X1 = Working capital / Total assets; X2 = Retained Earnings / Total assets; X3 = Earnings before interest and taxes / Total assets; X4 = Book value of equity / Book value of total liabilities	The initial Altman's Z'-Score gives good results for HR and CZ, while logit models for BG and RO. For HU, the model can be improved with additional variables.
Reznakova and Karas (2015)	The Visegrad Group (CZ, Slovakia, Poland and HU)	For CZ: Retained profit / total assets and Book value of equity / total debts For HU: Retained profit / total assets and Sales / total assets	Analysed the Visegrad Group using Amadeus database. The original version of the Altman model worked with a statistically lower accuracy.
Tudor *et al.* (2015)	RO	ROA and Growth rate on net profit	Developed a support system to predict distress on a sample of listed firms. CHAID outperformed NN and statistical models.

Note: BG – Bulgaria, HR – Croatia, CZ – Czech Republic, HU – Hungary, RO – Romania, EU – European Union. LT – long term, ST – short term.

Source: Authors' prelucration

The structure of NN is made up by neurons that are hierarchically organized in layers having the first layer as input, the last one as output and, between them, one or more hidden layers. The input values (in our case, all the 24 financial indicators) are then converted to an output value (in this case, a binary variable indicating whether the firm is distressed or not) through a transfer function. Next, when the output value exceeds a specific threshold level, the neuron is activated and the output is fed towards the next layer. Every neuron is connected to the neurons of the next layer through specific weights that represent the connection's strength between two neurons of two succeeding layers. The network starts with arbitrary weights and changes them through an activation function during the training stage, in order to minimize the error function. The basic idea consists in an iterative training process of generating weights as the network gets fed with data.

Figure 1 | The Process of Building CHAID Prediction Models



Source: Authors' own computation

3. Data Preparation and Modelling Phase

3.1 Data collection and matching procedure

We collected public financial information for the period 2011–2013 (due to data availability) using Amadeus (Bureau van Dijk) database. The initial sample consisted in 20,119 firms corresponding to a panel of five countries: Bulgaria, Croatia, the Czech Republic, Hungary and Romania (see Table 2).

From the initial sample, 4% (768 companies) were declared inactive in 2014. According to the information provided by the Amadeus database, we considered a firm to be distressed if it became either inactive, dissolved, in liquidation or bankrupt in 2014[4]. Since modelling financial distress implies building prediction models based on balanced samples of both distressed and healthy firms having similar observable characteristics, we applied a matching procedure. First, we assigned each company to one of two distinct groups, labelled as "healthy" or "distressed" (see columns 2 and 3 in Table 2).

4 This definition has some limits. As one anonymous reviewer suggested, in some cases, companies can effectively defer bankruptcy for several years after the start of the insolvency litigation. In these cases, a closer look to the financial situations of these companies or to other sources (*e.g.* media) can provide an identification of some of these troubled companies. However, this is difficult when available information is scarce (*e.g.* an external financial analysis) or when analysis is made on large volumes of data.

Table 2 | Sampling and Matching Distressed and Healthy Firms

Country	Initial sample of healthy firms	Initial sample of distressed firms	Matched distress firms	Matched healthy firms	Total matched firms
Bulgaria	2,675	130	130	130	260
Croatia*	102	118	118	102	220
Czech Republic	5,968	124	124	124	248
Hungary	3,396	126	126	126	252
Romania	7,194	286	173	173	346
Total	19,351	768	671	655	1,326

Note: *for Croatia no matching technique was required since the initial samples were already limited to a reduced number of companies balanced between distressed and non-distressed firms.

Source: Own calculation, based on Amadeus database

We applied a matching procedure in order to obtain two identically sized groups of either distressed or non-distressed firms for each country. For that, a 1:1 nearest neighbour matching (hereafter, NNM) algorithm without replications was implemented using STATA 14 software and applied for each country and each year considered in the analysis. First, we estimated a probit model having as dependent variable a binary variable that takes value 1 in case of distressed and 0 otherwise, and as independent variables the following three elements: level of turnover, total assets, and economic sector described by the four-digit NACE[5] code of each firm. Because of the high degree of heterogeneity in data and of the limited available information we had to accept as limits to our study the lack of more specific data regarding the typology of firms' activity and relative size which could matter in explaining the capital structure (Harris and Raviv, 1991; Myers, 2001; Buus, 2015).

The probit model allowed us to estimate each firm's probability of being distressed through a propensity score. Next, the 1:1 NNM procedure paired each distressed firm with the closest non-distressed firm for assuring a minimum average within-pair difference in propensity scores. The "no replications" condition suggests that once a match was done, the selected non-distressed firm was no longer eligible for further matching.

The matching results (see Table 2) follow the same structure of matched distressed with healthy firms for each of the years 2011–2013. Altogether 1326 firms resulted, equally divided between distressed and healthy.

5 NACE is the statistical classification of economic activities in the European Community. In this study the revised NACE Rev. 2 classification was used, as provided by the Amadeus database.

All distressed firms with available data were included in the analysis. In building the group of healthy firms, the matching procedure assumed random selection of these firms. So, even though the final size of our data samples is modest, it is representative for the considered statistical populations.

Building the prediction models for each of the five countries implied applying the same methodology on five distinct datasets. To increase the comparability power of our results, we expressed all financial datasets in the same currency.

Then, we generated several sub-samples, resulting distinct "in-sample" and "out-of-sample" datasets that were later used for testing the prediction ability of the models. Finally, in the modelling phase, we applied both CHAID and NN, using the financial indicators as input data.

3.2 Financial ratios

For the selected companies, we collected the financial ratios used in the modelling phase. The selection of the set of financial ratios was restricted to the financial data provided by the Amadeus database. As effect, we did not include in our analysis some other useful indicators (*e.g*. see Harris and Raviv, 1991). Moreover, out of the 32 financial ratios provided, only 23 were effectively used because of missing values and multicollinearity issues[6]. We have also used, as proxy for the company size, the natural logarithm of Total assets.

Finally, we used 24 financial ratios grouped in two categories reflecting either a stock or a flow-approach (see Table 3). Amadeus database provides some financial indicators, in which the denominator is the number of employees, less used in financial analysis. Thus, we were able to check if these indicators are good predictors of default.

Some useful indicators based on market values (like market value of equity / book value of debt used in Altman Z's score) were not used because of data unavailability. This unavailability is explainable as long as we included in our database unlisted companies.

Table 3 presents the financial indicators used in the models.

6 For instance, variables such as: Gross margin, Enterprise value / EBITDA, Gearing or R&D expenses / Operating revenue were not included in the analysis due to data unavailability, while variables such as Return on Capital Employed (ROCE) using P/L before tax or ROCE using Net income were excluded due to multicollinearity issues.

Table 3 | Financial Indicators Used in the Models

Approach	Indicators	Code	Definition(*)	Unit
Stock-approach	Shareholders liquidity ratio	SHLQ	Shareholders funds / Non current liabilities	x
	Solvency ratio (Asset based)	SOLR	(Shareholders funds / Total assets) · 100	%
	Solvency ratio (Liability based)	SOLL	(Shareholders funds / (Non current liabilities + Current liabilities)) · 100	%
	Current ratio	CURR	Current assets / Current liabilities	x
	Liquidity ratio	LIQR	(Current assets –Stocks) / Current liabilities	x
	Working capital *per* employee	WCPE	Working capital / Employees	Th. €
	Total assets	TA	Ln (Total Assets)	x
	Total assets *per* employee	TAPE	Total assets / Employees	Th. €
	Shareholders' funds *per* employee	SFPE	Shareholders´ funds / Employees	Th. €
Flow-approach	ROE using P/L before tax	RSHF	(Profit before taxation / Shareholders funds) · 100	%
	ROE using net income	ROE	(Net income / Shareholder funds) · 100	%
	ROA using P/L before tax	RTAS	(Profit before tax / Total assets) · 100	%
	Profit margin	PRMA	(Profit before tax / Operating revenue) · 100	%
	EBITDA margin	ETMA	(EBITDA / Operating revenue) · 100	%
	EBIT margin	EBMA	(EBIT / Operating revenue) · 100	%
	Cash flow / Operating revenue	CFOP	(Cash flow / Operating revenue) · 100	%
	Net assets turnover	NAT	Operating revenue / (Shareholders funds + Non current liabilities)	x
	Stock (inventory) turnover	STOT	Operating revenue / Stocks	x
	Collection paid	COLL	(Debtors / Operating revenue) · 360	days
	Profit *per* employee	PPE	Profit before tax / Employees	Th. €
	Operating revenue *per* employee	TPE	Operating revenue / Employees	Th. €
	Average cost of employee	ACE	Cost of employees / Employees	Th. €
	Costs of employees / Operating revenue	SCT	(Cost of employees / Operating revenue) · 100	%
	Interest coverage	IC	Operating profit / Interest paid	x

Notes: (*) The definitions and the codes are provided by the Amadeus database.
Source: author's prelucration based on Amadeus database

3.3 Modelling phase

In the third step, we used the database presented in Sections 3.1 and 3.2 for the modelling phase. Our database includes both healthy and distressed companies, respectively *N companies · t years · m financial ratios*.

We prepared the sample in order to define both the in-sample and the out-of-sample datasets required to compute the prediction ability of the models. We sub-divided the initial country specific distressed and non-distressed matched samples of firms in 70% observations for the learning phase and 30% for the testing one. The selection of the sub-samples was made randomly from the initial dataset and the same 70/30 structure was kept for all country-year sub-sets (having altogether 15 sub-sets due to the 5 country-panel structure and the 3 year-time-frame).

When building up the classification trees based on the in-sample dataset, we tested each financial indicator's ability to predict business failure considering, one by one, each indicator to be the explanatory variable in the model. As a result, we built up 24 distinct CHAID trees by considering each time a distinct financial ratio as predictor. For each resulted tree model, both the in-sample and the out-of-sample prediction accuracy were computed in order to test the models' efficiency. Thus, we determined a top 5 best predictors for each post-communist country for the interval 2011–2013 using CHAID trees (see Table 4).

The design of the neural network was determined not only by the availability of data inputs but also by the desired classification output, which took value 1 in case of distressed and 0 otherwise. The supervised *feedforward backpropagation* algorithm was selected to minimize the errors (Smith and Gupta, 2002).

For each year and country dataset, NN had a multilayer structure consisting of an input layer, an output layer and one hidden layer. We chose the solution with only a single layer based on several previous empirical studies (Jain and Nag, 1998; Kumar and Tan, 2004; Olson *et al.,* 2012) that evidence that a network with multiple hidden layers cannot improve the accuracy of training.

In our NN the input layer had a predetermined number of nodes, corresponding to the 24 financial indicators, while the output layer had just a single node, since the output consisted in a prediction of just two values (0 for healthy firms and 1 for distressed ones). The number of nodes used in the hidden layer (ranging between 1 and 10) was optimally determined each time through an automatic architecture selection.

A hyperbolic tangent activation function for the hidden layer was used when building NN model, along with a Softmax activation function for the output layer and a cross-entropy error function. Following Kumar and Tan (2004) the training stage was set to last for 2,000 iterations and the network was trained in order to learn how to classify firms as distressed and non-distressed. Once again, the same sub-sampling structure was used by considering a 70% of observations for the learning phase and a 30% for the testing phase.

4. Results of the Analysis

Based on the results of the prediction models, we performed a comparative analysis, in order to check the consistency of each predictor of financial distress in time. The top five most appropriate indicators of financial distress for each country and year that resulted based on CHAID trees is presented in Table 4. The top five was defined in terms of models' prediction ability scores in the test phase, followed by the learning phase prediction accuracy.

In general, we can notice a three-year time frame similarity for Croatia and a rather two-year long similarity for Romania. Regarding the most effective financial distress predictor for 2013, with the exception of Bulgaria and Croatia, for which the same predictor – *TA* turned out to be the most effective, heterogeneity between countries can be noted. For instance, *SFPE* turned out to be best predictor for Romania, *PRMA* for Hungary, whereas *WCPE* – for the Czech Republic.

Although both stock and flow indicators are present in each country's top five best predictors of financial distress, some particularities can be noticed. For instance, the stock-approach indicators are better predictors of financial distress for Bulgaria and Croatia, and the flow-approach one for the Czech Republic, Hungary and Romania. These different patterns can be explained through a greater variability in time of flow indicators in some countries; as a result, they become irrelevant for future companies' performance and for failure prediction. Stock indicators are in general more stable in time.

From the stock-approach indicators, the most stable predictors of distress on a three-year time frame tend to be solvency ratio (*SOLR*) for Bulgaria, as well as *TA*, *TAPE* and *SFPE* for Croatia. Similarly, *SFPE* is also found to be rather consistent to predict distress in Romania and Bulgaria, but on a rather two-year time frame, while no indicator was found to be stable for the Czech Republic.

From the flow-approach perspective, *PRMA* and *PPE* are the most stable predictors for Hungary, while *NAT* is the most stable for Croatian firms' distress prediction. *PPE* was also found to be rather relevant for Romania, especially for the interval 2012–2013, while *IC* turned out to predict well for the Czech Republic on a two-year time frame.

As a particularity, for the Czech Republic and Bulgaria, when building the prediction models based on financial data for 2012, less than five CHAID trees registered prediction accuracy in the testing phase higher than the cut-off value of 50%. Thus, a top of only 2 (and not 5) best predictors was determined.

NN models were built by considering all 24 financial indicators as input data. For each yearly and country dataset, an optimal number of hidden layers was computed and then used in the network. According to the highest final weights of each predictor in the neural network models, a top 5 most relevant predictors resulted (see Table 5).

We can notice a general three-year time frame similarity for Bulgaria based on the *NAT* indicator, as well as for Romania, when considering *EBMA*. Moreover, it can be noticed a rather two-year time frame similarity for Croatia on *ETMA, SOLR* and *TA* and also for the Czech Republic on *SFPE* and *IC*, as well as for Bulgaria on the *COLL* and *STOT* indicators.

Table 4 | Top 5 Best Predictors for the Post-Communist Country Panel and the Interval 2011–2013 Using CHAID

Country	Top predictors	Prediction ability (%)		Top predictors	Prediction ability (%)		Top predictors	Prediction ability (%)	
BG	**2013**	**learn phase**	**test phase**	**2012**	**learn phase**	**test phase**	**2011**	**learn phase**	**test phase**
	TA	53.3	82.1	SFPE	58.8	60.3	SOLL	61.0	64.4
	IC	63.3	69.2	SOLR	56.0	53.8	SOLR	59.9	56.4
	RSHF	60.0	69.2	COLL	53.0	53.2	RTAS	56.6	55.1
	WCPE	55.9	59.2	PPE	57.2	52.6	RSHF	55.5	53.8
	SOLR	58.9	52.6	–	–	–	SFPE	61.0	52.6
HR	**2013**	**learn phase**	**test phase**	**2012**	**learn phase**	**test phase**	**2011**	**learn phase**	**test phase**
	TA	85.7	100.0	TA	84.4	100.0	TA	81.8	98.5
	TAPE	92.2	95.5	NAT	84.4	87.9	TAPE	90.9	92.4
	NAT	87.0	83.3	TAPE	87.7	86.4	NAT	85.7	86.4
	ETMA	72.1	77.3	SFPE	77.8	77.3	SFPE	77.1	86.4
	SFPE	79.1	74.2	CFOP	70.6	75.8	TPE	53.9	74.2
CZ	**2013**	**learn phase**	**test phase**	**2012**	**learn phase**	**test phase**	**2011**	**learn phase**	**test phase**
	WCPE	57.4	70.1	IC	55.9	51.3	ETMA	62.1	55.4
	TA	63.8	68.8	CURR	54.0	50.0	STOT	56.8	54.7
	IC	64.4	65.6	–	–	–	LIQR	59.8	54.1
	ROE	61.0	63.6	–	–	–	SOLR	54.6	54.1
	RSHF	61.0	62.3	–	–	–	CFOP	63.2	52.7
HU	**2013**	**learn phase**	**test phase**	**2012**	**learn phase**	**test phase**	**2011**	**learn phase**	**test phase**
	PRMA	61.9	69.7	EBMA	63.6	69.7	RTAS	60.8	68.4
	SFPE	56.6	69.7	CFOP	64.8	68.4	PPE	59.4	64.5
	ROE	60.8	68.4	PPE	61.7	63.2	PRMA	57.4	64.5
	PPE	62.3	67.1	ROE	60.8	61.8	RSHF	60.8	63.2
	RSHF	61.9	67.1	PRMA	60.8	61.8	SOLR	58.0	61.8
RO	**2013**	**learn phase**	**test phase**	**2012**	**learn phase**	**test phase**	**2011**	**learn phase**	**test phase**
	SFPE	52.1	78.2	SFPE	55.6	74.5	ETMA	57.0	67.3
	PPE	67.4	77	ACE	56.2	72.5	PRMA	61.2	65.4
	IC	68.7	75.4	PPE	59.8	72.3	SOLR	62.8	64.4
	RTAS	67.4	75	RSHF	57.9	62.5	CFOP	58.7	64.4
	RSHF	65.3	73.1	SHLQ	61.0	60.9	RTAS	63.6	57.7

Note: BG – Bulgaria, HR – Croatia, CZ – Czech Republic, HU – Hungary, RO – Romania. The significance of the predictors is provided in Table 3.

Source: Own calculation, based on Amadeus database

Table 5 | Top 5 Best Predictors for the Post-Communist Country Panel and the Interval 2011–2013 Using NN

<table>
<tr><th>Country</th><th>Top predictors</th><th colspan="2">Prediction ability (%)</th><th>Top predictors</th><th colspan="2">Prediction ability (%)</th><th>Top predictors</th><th colspan="2">Prediction ability (%)</th></tr>
<tr><td rowspan="6">BG</td><td>2013</td><td>learn phase</td><td>test phase</td><td>2012</td><td>learn phase</td><td>test phase</td><td>2011</td><td>learn phase</td><td>test phase</td></tr>
<tr><td>RSHF</td><td rowspan="5">88.6</td><td rowspan="5">100.0</td><td>STOT</td><td rowspan="5">66.7</td><td rowspan="5">64.0</td><td>STOT</td><td rowspan="5">81.0</td><td rowspan="5">77.3</td></tr>
<tr><td>NAT</td><td>NAT</td><td>PPE</td></tr>
<tr><td>COLL</td><td>CURR</td><td>TA</td></tr>
<tr><td>PRMA</td><td>COLL</td><td>TAPE</td></tr>
<tr><td>SHLQ</td><td>ROE</td><td>NAT</td></tr>
<tr><td rowspan="6">HR</td><td>2013</td><td>learn phase</td><td>test phase</td><td>2012</td><td>learn phase</td><td>test phase</td><td>2011</td><td>learn phase</td><td>test phase</td></tr>
<tr><td>RSHF</td><td rowspan="5">97.0</td><td rowspan="5">100.0</td><td>TA</td><td rowspan="5">88.1</td><td rowspan="5">70.6</td><td>ETMA</td><td rowspan="5">83.7</td><td rowspan="5">76.9</td></tr>
<tr><td>SOLR</td><td>CFOP</td><td>RTAS</td></tr>
<tr><td>ROE</td><td>SFPE</td><td>TAPE</td></tr>
<tr><td>NAT</td><td>SOLR</td><td>TA</td></tr>
<tr><td>TAPE</td><td>ETMA</td><td>NAT</td></tr>
<tr><td rowspan="6">CZ</td><td>2013</td><td>learn phase</td><td>test phase</td><td>2012</td><td>learn phase</td><td>test phase</td><td>2011</td><td>learn phase</td><td>test phase</td></tr>
<tr><td>TAPE</td><td rowspan="5">100.0</td><td rowspan="5">100.0</td><td>IC</td><td rowspan="5">79.5</td><td rowspan="5">71.4</td><td>LIQR</td><td rowspan="5">75.8</td><td rowspan="5">83.3</td></tr>
<tr><td>SFPE</td><td>TA</td><td>TAPE</td></tr>
<tr><td>NAT</td><td>ACE</td><td>EBMA</td></tr>
<tr><td>COLL</td><td>STOT</td><td>IC</td></tr>
<tr><td>PPE</td><td>SFPE</td><td>PRMA</td></tr>
<tr><td rowspan="6">HU</td><td>2013</td><td>learn phase</td><td>test phase</td><td>2012</td><td>learn phase</td><td>test phase</td><td>2011</td><td>learn phase</td><td>test phase</td></tr>
<tr><td>SHLQ</td><td rowspan="5">75.0</td><td rowspan="5">71.0</td><td>ETMA</td><td rowspan="5">100.0</td><td rowspan="5">83.0</td><td>ACE</td><td rowspan="5">85.0</td><td rowspan="5">85.0</td></tr>
<tr><td>SFPE</td><td>NAT</td><td>TPE</td></tr>
<tr><td>TPE</td><td>TA</td><td>CURR</td></tr>
<tr><td>PPE</td><td>RSHF</td><td>SHLQ</td></tr>
<tr><td>COLL</td><td>EBMA</td><td>COLL</td></tr>
<tr><td rowspan="6">RO</td><td>2013</td><td>learn phase</td><td>test phase</td><td>2012</td><td>learn phase</td><td>test phase</td><td>2011</td><td>learn phase</td><td>test phase</td></tr>
<tr><td>NAT</td><td rowspan="5">75.0</td><td rowspan="5">77.0</td><td>TAPE</td><td rowspan="5">71.8</td><td rowspan="5">74.4</td><td>RTAS</td><td rowspan="5">72.7</td><td rowspan="5">71.7</td></tr>
<tr><td>RTAS</td><td>SOLL</td><td>CFOP</td></tr>
<tr><td>PPE</td><td>CFOP</td><td>EBMA</td></tr>
<tr><td>ROE</td><td>EBMA</td><td>ETMA</td></tr>
<tr><td>EBMA</td><td>PPE</td><td>COLL</td></tr>
</table>

Note: BG – Bulgaria, HR – Croatia, CZ – Czech Republic, HU – Hungary, RO – Romania. The significance of the predictors is provided in Table 3.

Source: Own calculation, based on Amadeus database

For Romania, *PPE* and *CFOP* are among the most stable distress predictors, which belong to the flow-approach category. For Hungary, no indicator was found to last for more than one year as best predictor of financial distress. However, SHLQ, TPE and COLL are common for 2011 and 2013. This result can suggest a possible unusual situation for 2012.

The highest final weights of NN suggest the flow-approach indicators as the most relevant predictors of distress for each year and country analysed, with the exception of the Czech Republic, where the proportion of both stock and flow-approach indicators is rather balanced. Also, for the Czech Republic, the presence as predictors of distress of some indicators *per* employee (*TAPE*, *SPFE*, *PPE*) can be noticed.

The *NAT* turned out to be among the most effective financial distress predictors for 2013 in all countries, excepting Hungary. Other relevant predictors were the rates of return (*ROA* and *ROE*), the margin rates (*EBMA* and *PRMA*), as well as *TAPE* and *SFPE*.

Since NN use all 24 input data in the model, the selection of the top 5 most valuable predictors of the NN should be interpreted cautiously. The predictors' importance is limited, as they cannot stand alone in the NN model.

The differences in the top 5 best predictors resulted from either NN or CHAID models are not a huge surprise. They can be explained by the distinct patterns of the two methods. On the one hand, CHAID models assume computing a set of multiple statistical tests for each input attribute in order to find the pair of values that is least significantly different with respect to the target attribute. The resulted model is described through a tree diagram with easily intuitive classification rules. On the other hand, NN uses all 24 financial indicators as input data and builds nonlinear models by identifying data patterns. More important, we do not need to specify the functional relationship between variables and the final weights can indicate which of the input data are the most relevant in predicting business failure.

When comparing the final tops of best predictors resulted from both CHAID and NN models for each of the 5 countries analysed on a 3 year time frame, we notice rather significant differences in the main findings. As a result, models for one country cannot be used for other countries. These findings are not a surprise, as the literature review in the field brings empirical evidence of a large variability of best financial predictors of distress (not just between countries and periods, but also for the same country case when distinct datasets and methods are used).

Even though the two types of models tend to highlight a diversity of best predictors of distress, some similarities draw a short list of indicators of business failure. Moreover, the two methodological approaches assure a wider perspective for analysing the ability to predict business failure through financial indicators. For instance, the most stable financial distress predictors confirmed through both types of prediction models based on a top 5 best predictors on the 3 year time frame were the following:

- Bulgaria: *RSHF* for 2013 and *COLL* for 2012;
- Croatia: *NAT* for 2011 and 2013; *TAPE* for 2011 and 2013; *TA* for 2011–2012*; SFPE* for 2012 and *CFOP* for 2012;
- Czech Republic: *IC* for 2012 and *LIQR* for 2011;
- Hungary: *SFPE* and *PPE* for 2013;
- Romania: *PPE* for 2012–2013; *ROA* for 2011 and 2013; *ETMA* and *CFOP* for 2011.

However, additional country specific information is distinctly given based on the two types of prediction models considered in the empirical analysis.

For Bulgaria, the most effective predictors are: *TA*, solvency ratios, efficiency rates (*NAT, COLL* and *STOT*), rates of return (*ROE*) and debts ratios (*IC*)).

Business failure is best predicted in the Croatian case by *TAPE*, *SFPE*, *TA*, efficiency rates (*NAT*), rates of return (*ROE*) and margin ratios (*ETMA*).

In the case of Hungary, the best predictors are: *SFPE* and *TA*, efficiency rates (*PPE, NAT* and *TPE*), rates of return (*ROE*) and margin ratios (*PRMA, EBMA* and *ETMA*).

Financial distress in the Czech Republic is described best through debts ratios (*IC*), liquidity ratios like *WCPE*, size (*TA*), efficiency rates (*NAT, COLL* and *PPE*), and rates of return (*ROE*).

For Romania, several financial *per* employee indicators turned out to best predict business failure, such as: *SFPE* for 2012–2013 (according to CHAID) and *PPE*. Also, business failure is predicted by some rates of return (*RTAS* and *ROE*) and margin rates, such as: *EBMA* and *ETMA*.

5. Conclusions

Through this paper we identified a short list of best predictors of distress for each of 5 post-communist countries: Bulgaria, Croatia, the Czech Republic, Hungary and Romania. However, it can be noticed that these indicators should be interpreted cautiously as they are not persistent as predictors in time. For instance, for Romania, with the exception of *EBIT margin* (and only when using neural networks) no indicator was present as a predictor of failure in all of the analysed years, no matter of the technique used.

Regarding the indicators' ability to predict financial distress for several consecutive years, the results confirmed through both methods that *Total Assets* and *Net Assets Turnover* are proper predictors for Croatia, *Interest Coverage* for the Czech Republic and *Profit per Employee* for Romania.

We bring empirical evidence on the role of both stock and flow-approach indicators providing a list of indicators suitable for failure prediction. Some country specific patterns were identified even though the results differ substantially in time and from country to country. The invariant application of such models should therefore be made cautiously. From a practical perspective, the use of a larger number of indicators and a careful interpretation of their informational content is recommended. Also, one relevant finding is the importance of ratios per employee for predicting financial distress.

Based on CHAID models (since neural networks brings further insights through the use of a more complex structure) we conclude that in countries, such as the Czech Republic, Hungary and Romania, financial distress can better be predicted through flow-approach indicators, while stock-approach indicators work better in prediction for Bulgaria and Croatia. This result can be a clue that some common patterns exist for some countries, explainable through an economic or even socio-cultural context, but supplementary research is needed.

As an element of novelty, we conducted the analysis on a sample of private limited companies, less studied than the listed ones, in the case of analysed countries. Therefore, our results reflect better these countries' economic environment and financial development. Further on, this study offers several clues to practitioners regarding the indicators that can be used as best predictors for failure, while from an academic perspective, we suggest a more circumspect attitude in generalizing some patterns identified for other countries.

As a limit, our results rely entirely on the available data samples and cannot reflect a more general image of the business failure phenomenon. Moreover, we have to be aware that panel-country datasets normally face high heterogeneity and perhaps computational differences might also occur due to distinct accounting procedures.

Other (non-financial) factors can count in failure prediction, as well. Our research on financial distress was limited to the use of financial indicators disclosed in the accounting statements and ignored the role of the non-financial indicators. In a future study we plan to also use some non-financial indicators to improve the prediction accuracy of the models.

References

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance,* 23(4), 589–609, https://doi.org/10.2307/2978933

Altman, E. I., Iwanicz-Drozdowska, M., Laitine, E. K., Suvas, A. (2014). *Distressed Firm and Bankruptcy Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model.* [Retrieved 2016-07-03] Available at: http://doi.org/10.2139/ssrn.2536340

Anghel, I. (2002). *Falimentul (Bankruptcy).* Bucharest: Editura Economica. ISBN 973-590-678-3.

Buus, T. (2015). A General Free Cash Flow Theory of Capital Structure. *Journal of Business Economics and Management,* 16(3), 675–695, https://doi.org/10.3846/16111699.2013.770787

Dragotă, V., Dragotă, M. I. (2009). Models and Indicators for Risk Valuation of Investment Projects. *Economic Computation and Economic Cybernetics Studies and Research,* 43(3), 69–75.

Ekinci, A. (2016). Rethinking Credit Risk under the Malinvestment Concept: The Case of Germany, Spain and Italy. *European Financial and Accounting Journal,* 11(1), 39–64, https://doi.org/10.18267/j.efaj.152

Geng, R., Bose, I., Chen, X. (2015). Prediction of Financial Distress: An Empirical Study of Listed Chinese Companies Using Data Mining. *European Journal of Operational Research,* 241(1), 236–247, https://doi.org/10.1016/j.ejor.2014.08.016

Harris, M., Raviv, A. (1991). The Theory of Capital Structure. *Journal of Finance,* 46(1), 297–355, https://doi.org/10.1111/j.1540-6261.1991.tb03753.x

Hsu, J. C. (1996). *Multiple Comparisons: Theory and Methods.* London, UK: Chapman and Hall. ISBN 978-0412982811.

Jaba, E., Robu, I. B., Istrate, C., Balan, C. B., Roman, M. (2016). Statistical Assessment of the Value Relevance of Financial Information Reported by Romanian Listed Companies. *Romanian Journal of Economic Forecasting,* 19(2), 27–42.

Jain, B. A., Nag, B. N. (1998). A Neural Network Model to Predict Long-Run Operating Performance of New Ventures. *Annals of Operations Research,* 78, 83–110, https://doi.org/10.1023/A:1018910402737

Jakubík, P., Teplý, P. (2011). The JT Index as an Indicator of Financial Stability of Corporate Sector. *Prague Economic Papers,* 20(2), 157–176, https://doi.org/10.18267/j.pep.394

Keasey, K., Watson, R. (1991). Financial Distress Prediction Models: A Review of Their Usefulness. *British Journal of Management,* 2(2), 89–102, https://doi.org/10.1111/j.1467-8551.1991.tb00019.x

Kumar, K., Tan, C. (2004). *Artificial Intelligence in Financial Distress Prediction.* [Retrieved 2016-09-10] Available at: www.niitcrcs.com/iccs/papers/2005_37.pdf

Myers, S. (2001). Capital Structure. *Journal of Economic Perspectives,* 15(2), 81–102, https://doi.org/10.1257/jep.15.2.81

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research,* 18(1), 109–131, https://doi.org/10.2307/2490395

Olson, D. L., Delen, D., Meng, Y. (2012). Comparative Analysis of Data Mining Methods for Bankruptcy Prediction. *Decision Support Systems,* 52(2), 464–473, https://doi.org/10.1016/j.dss.2011.10.007

Reznakova, M., Karas, M. (2015). The Prediction Capabilities of Bankruptcy Models in a Different Environment: An Example of the Altman Model under the Conditions in the Visegrad Group Countries. *Ekonomický časopis,* 63, 617–633.

Ross, S. (1977). The Determination of Financial Structure: The Incentive-Signalling Approach. *The Bell Journal of Economics,* 8(1), 23–40, https://doi.org/10.2307/3003485

Ross, S., Westerfield, R., Jaffe, J. (2010). *Corporate Finance, 9th Edition.* Irwin: The McGraw-Hill. ISBN 978-0073382333.

Šarlija, N., Jeger, M. (2011). Comparing Financial Distress Prediction Models before and during Recession. *Croatian Operational Research Review (CRORR),* 2, 133–142.

Scarlat, E., Chiriță, N., Bradea, I. A. (2012). Indicators and Metrics Used in the Enterprise Risk Management (ERM). *Economic Computation and Economic Cybernetics Studies and Research Journal,* 46(4), 5–18.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business,* 74(1), 101–124, https://doi.org/10.2139/ssrn.171436

Smith, K., Gupta, J. (2002). *Neural Networks in Business: Techniques and Applications.* Hershey: Idea Group Publishing. ISBN 1-930708-31-9.

Tudor, L., Popescu, M. E., Andreica, M. (2015). A Decision Support System to Predict Financial Distress. The Case of Romania. *Romanian Journal of Economic Forecasting,* 18(4), 170–179.

Virág, M., Hajdu, O. (1996). Pénzügyi mutatószámokon alapuló csõdmodell-számítások (Financial Ratio Based Bankruptcy Model Calculations). *Bankszemle,* 15(5), 42–53.

Virág, M., Kristóf, T. (2005). Neural Networks in Bankruptcy Prediction – A Comparative Study on the Basis of the First Hungarian Bankruptcy Model. *Acta Oeconomica,* 55(4), 403–425, https://doi.org/10.1556/aoecon.55.2005.4.2

Welc, J. (2016). Empirical Safety Thresholds for Liquidity and Indebtedness Ratios on the Polish Capital Market. *European Financial and Accounting Journal,* 11(3), 37–50, https://doi.org/10.18267/j.efaj.161

Wruck, K. (1990). Financial Distress: Reorganization and Organizational Efficiency. *Journal of Financial Economics,* 27(2), 419–444, https://doi.org/10.1016/0304-405x(90)90063-6

Zheng, Q., Yanhui, J. (2007). Financial Distress Prediction on Decision Tree Models. *IEEE,* 16, 80–120, https://doi.org/10.1109/soli.2007.4383925, http://data.worldbank.org