Siniša Bogdan[*]
Luka Šikić[**]
Suzana Bareša[***]

JEL Classification: C33, C34
Preliminary statement
https://doi.org/10.32910/ep.72.5.1

PREDICTING BANKRUPTCY BASED ON THE FULL POPULATION OF CROATIAN COMPANIES

This paper analyses the bankruptcy prediction based on the population of companies representative of the total business sector in Croatia. The representativity of the sample is achieved through the propensity score matching of the full population of bankrupt and similar non-bankrupt companies. The robust estimation of bankruptcy prediction is carried out through the multiple discriminant analysis (MDA) and logistic regression (logit). The results indicate high classification accuracy of both models, but more favourable performance of the logit estimation. Overall accuracy of the MDA model was 73.7%, while the overall accuracy of the logit model was 76.3%. The results serve as a bankruptcy estimation benchmark for the business sector in Croatia.

Keywords: *Multiple discriminant analysis, MDA, logistic regression, logit, financial ratios*

[*] S. Bogdan, PhD, Assistant Professor, University of Rijeka, Faculty of Tourism and Hospitality Management (e-mail: sinisab@fthm.hr).

[**] L. Šikić, PhD, postdoctoral researcher, University of Zagreb, Institute of Social Sciences Ivo Pilar (e-mail: luka.sikic@pilar.hr).

[***] S. Bareša, PhD, Assistant Professor, University of Rijeka, Faculty of Tourism and Hospitality Management (e-mail: suzanab@fthm.hr). The paper was received on 02.07.2020. It was accepted for publication on 19.10.2020.

1. INTRODUCTION

According to the data retrieved from the Croatian Financial Agency — FINA (December 31, 2019), 17,903 businesses were blocked in Croatia due to the aggregate debt amount of HRK 5.88 billion. Relative to the previous year, the number of blocked business entities in Croatia decreased by 5.6% but still remains high. In the near future, the number of bankruptcies in Croatia is expected to increase due to the COVID 19 pandemic, signifying the need for robust bankruptcy prediction estimations.

Bankruptcy prediction has been in the focus of academic research since the first half of the 20th Century. The focus of early studies (Ramser and Foster (1931), Fitzpatrick (1932), Smith and Winakor (1935) and Wall (1936) was on the analysis of financial ratios in order to find financial indicators able to predict business failure. A greater popularisation of business failure prediction models followed in the 1960s. After Beaver's (1966) paper, which applied the first univariate statistical model for bankruptcy prediction, the seminal study of Altman (1968) was published. Altman applied Multiple Discriminant Analysis (MDA) to analyse bankruptcy prediction. Building on Altman's model, Deakin (1972), Edmister, 1972.; Wilcox, 1973.; Altman and McGough, 1974.; Blum, 1974.; Sinkey, 1975.; Libby, 1975; Ohlson 1980 and others expanded the literature on bankruptcy predictions. Since the 1980s, the application of logit models to predict the bankruptcy has become more common. In 1990s, besides MDA, logit and probit models (Zmijewski 1984), neural networks have been the most widely used technique (Odom and Sharda 1990, Tam and Kiang 1992, Fletcher and Goss 1993). Due to the recent growth in computation possibilities, machine learning models like support vector machines (SVM), Bagging, Boosting and Random Forest have been applied in recent literature, such as Barboza, Kimura and Altman (2017), Zięba, Tomczak and Tomczak (2016), Carmona, Climent and Momparler (2019).

The literature on company failure prediction in Croatia is homogeneous in terms of methodological approach where early bankruptcy prediction literature is used as an empirical reference. However, significant differences in terms of data sample and identification of the failure event among Croatian studies exist so that findings are not directly comparable. The objective of this study is to improve global comparability and representability of domestic bankruptcy literature by using the full population of bankrupt companies in Croatia and legally claimed bankruptcy of the company as a failure event. In this way the analysis of bankruptcy prediction represents the total business sector in Croatia and has aggregate implications. This objective was achieved through the careful[1] construction of the

[1] Legal definition of bankruptcy was used as a failure event and sample bankruptcy data from Croatian Court Registry.

empirical data sample with the propensity score matching and database merging while applying robust[2] methodology for the bankruptcy prediction. The paper contributes to the literature by setting the benchmark bankruptcy prediction reference for Croatian companies due to the sample size and the explicit definition of bankruptcy as a failure event. Furthermore, the analysis improves global comparability of bankruptcy estimation results for Croatian companies.

The remainder of this paper is structured as follows: the next section reviews relevant foreign and domestic literature. Although there are number of papers that analyse bankruptcy in Croatia, this topic has still been insufficiently researched. Section 3 describes empirical data set and financial indicators used in the analysis. Methodology and empirical results are explained in section 4. Concluding remarks, limitations of the study and recommendations for future research are provided in the last section.

2. LITERATURE OVERVIEW

Different techniques have been applied in bankruptcy prediction analysis. Discriminant analysis was widely applied in late 1960s due to the simple estimation procedure and ease of interpretation. The first bankruptcy prediction model on the basis of multivariate discriminant analysis (MDA) was developed by Altman (1968). Altman applied five-factor multivariate discriminant analysis model on a sample of 60 manufacturing companies. His results show 95% prediction accuracy one year before failure, but the model's accuracy drops to 79% two years before failure and 48% three years before failure. Altman's model became a workhorse approach for bankruptcy prediction and was widely extended (Deakin 1972, Edmister 1972, Blum 1974, Moyer 1977, Northon and Smith 1979, Betts and Belhoul 1987, Rose and Giroux 1984, Gardiner, Oswald and Jahera 1996, and many others).

Edward Deakin (1972) used 14 financial ratios (as did Beaver in 1966) on an independent sample consisting of 11 failed and 23 non-failed companies selected at random from the 1963 and 1964 Moody's Industrial Manuals. The accuracy of his model for correctly predicted failed companies was 77% and 82% for correctly predicted non-failed companies one year before failure. In the second year, the model had a higher accuracy of 96% for failed companies and 92% for non-failed companies. Compared to the second and third year models, the fourth and fifth year models show lower accuracy for non-failed and failed companies. Robert O. Edmister (1972) analysed small business failures. He examined 19 ratios and five

[2] Multiple discriminant analysis (MDA) and logistic regression (logit).

hypothesised methods of ratio analysis which had been categorised as significant in previous studies by Altman, Beaver and others. His MDA model achieved a 93% degree of classification precision. Marc Blum (1974) tested the data set of 115 failed and 115 non-failed companies. Companies were matched by industry, sales, employees, and fiscal year, and twelve predictors were used in the study. The predictive accuracy was 93-95% at the first year before failure, 80% at the second year and 70% at the third year.

In the 1980s, the logit analysis gained popularity in bankruptcy prediction literature due to the robustness gains, less restrictive assumptions and ease of interpretation compared to MDA. The first author who argued for the logit model instead of the MDA was Ohlson (1980). His study uses the sample of 105 bankrupt and 2,058 non-bankrupt companies and applies the logit model with a set of nine accounting ratios. The model achieved a prediction accuracy of 96%. Zavgren (1985) used the logit model to predict "financial health" for a five years prior to the failure. She tested seven factors with a model accuracy of 69% for each tested year before failure. Platt, Platt and Pedersen (1994) examined 124 oil and gas companies in the 1982-1988 period. This model correctly classified 80% of bankrupt companies but with deflated financial indicators the accuracy was improved to 94%. The overall classification accuracy of their model was 95%. The logit methodology is still a commonly used empirical model for bankruptcy prediction (Martin (1977), Wiginton (1980), Laitinen and Laitinen (2000), Jakubík and Teplý (2008), Chi and Tang (2006), but there are also a significant number of papers (Lo (1986), Back et al. (1996), Kim and Gu (2006), Araghi and Makvandi (2013), Mihailovič (2016) that used both, logit and MDA.

Several papers analyse bankruptcy prediction in Croatia based on different samples and with different specifications of the bankruptcy event. Šarlija Penavin and Harc (2009) investigate short-term illiquidity with a five-factor logit model. The results show a prediction accuracy of 68.16% for liquid companies and 74.22% for illiquid companies. Their data sample consists of 60,116 companies for model training and 15,029 for model testing. Pervan, Pervan and Kuvek (2018) apply three separate logit models based on the three failed company statuses: bankruptcy, rescue plan and financial distress. The data sample consists of 244 bankrupt, 808 rescue-plan and 3,200 financial distressed companies one year before failure (t-1); 483 bankrupt, 853 rescue plan and 3242 financial distress companies two years before failure (t-2) and 551 bankrupt, 887 rescue-plan and 3033 financial distress companies threes before failure (t-3). They use five financial explanatory variables and three non-financial ones. The overall error for the one-year lagged bankruptcy prediction model was 20.7%, 11.5% for the rescue-plan companies and financial distress model had the lowest error at 4.7%. The results show that the financial distress model offers the highest level of predictive power and outperforms traditional legal status models in terms of accuracy. They argue in favour of using non-micro companies in the estimation sample due to the improved model prediction power. Kozjak, Šestanj-Perić and Bešvir (2014) tested Altman's z-score model, the Springate model, Kralicek's Quicktest, the FP Rating model, BEX index and Bonitest on Croatian mid-sized companies from the manufacturing sector. The results show that foreign MDA models have higher prediction accuracy than domestically developed models (Bonitest and BEX index). The main drawback of this research was its small sample size. Similar research was conducted by Bogdan, Baresa and Hađina (2019), where they tested Altman's z-score and calibrated the model based on the sample of 52 Croatian companies. As they had to exclude one variable from the original model (due to multicollinearity), the prediction accuracy was lower 0n average for (t-1, t-2 and t-3) by 1.17 percentage points after testing the upper bound and (t-1, t-2 and t-3) by 20.51 percentage points lower after testing the lower bound. The list of remaining bankruptcy prediction studies in Croatia is provided in the Table 1.

Table 1.

OVERVIEW OF BANKRUPTCY PREDICTION AND CLOSELY RELATED STUDIES IN CROATIA

Study	Sector/ Observations	Model/ Factors	Model accuracy		
Novak (2003)	Financial sector/ 39 banks	MDA/4	Distressed banks	–	100.0%
			No distressed banks	–	100.0%
		Logit/4	Distressed banks	–	100.0%
			No distressed banks	–	100.0%
		MDS	N/A		
Novak and Crnković (2007)	General/141 companies	MDA/3	Dubious companies	–	57.1%
			Good companies	–	100.0%
	General/90 companies	MDA/3	Dubious companies	–	72.7%
			Good companies	–	100.0%
	General/90 companies	MDA/6 (after MDS)	Dubious companies	–	88.0%
			Good companies	–	100.0%
	General/141 companies	Logit/3	Dubious companies	–	85.7%
			Good companies	–	95.3%
	General/90 companies	Logit/3	Dubious companies	–	77.3%
			Good companies	–	94.1%
	General/90 companies	Logit/3 (after MDS)	Dubious companies	–	88.0%
			Good companies	–	95.4%

Study	Sector/ Observations	Model/ Factors	Model accuracy		
Sajter (2008)	General/100 companies	MDA/1	Bankrupted companies	–	50.0%
			Healthy companies	–	98.6%
		Logit/2	Bankrupted companies	–	50.0%
			Healthy companies	–	100.0%
		MDS/9	N/A		
Zenzerović (2009)	General/108 companies	MDA/12	Financially unstable companies	–	88.7%
			Financially stable companies	–	98.2%
		MDA/6	Financially unstable companies	–	92.5%
			Financially stable companies	–	98.1%
Pervan, Pervan and Vukoja (2011)	General/156 companies	MDA/3	Bankrupted companies	–	79.5%
			Healthy companies	–	80.8%
		Logit/3	Bankrupted companies	–	85.9%
			Healthy companies	–	80.8%
Šarlija and Jeger (2011)	2008/2009 General SMEs/ *development sample -1987* companies, *validation sample-993* companies	Logit/10	*Development sample*		
			Distressed companies	–	70.7%
			Healthy companies	–	80.1%
			Validation sample		
			Distressed companies	–	63.6%
			Healthy companies	–	78.9%
	2007/2008 General SMEs/ *development sample -1988* companies, *validation sample-998* companies	Logit/9	*Development sample*		
			Distressed companies	–	66.7%
			Healthy companies	–	82.3%
			Validation sample		
			Distressed companies	–	65.4%
			Healthy companies	–	80.4%
	2006/2007 General SMEs/ *development sample -1986 companies, validation sample-995* companies	Logit/8	*Development sample*		
			Distressed companies	–	67.4%
			Healthy companies	–	85.1%
			Validation sample		
			Distressed companies	–	67.4%
			Healthy companies	–	85.1%

Study	Sector/ Observations	Model/ Factors	Model accuracy		
Pervan and Kuvek (2013)	General/ 825 companies/ clients of Croatian commercial bank	Logit/4	Defaulted companies	–	52.0%
			Non-defaulted companies	–	88.4%
		Logit/7 *(Combined- 3 nonfinancial factors)*	Defaulted companies	–	64.6%
			Non-defaulted companies	–	92.4%
Ježovita (2015)	General/343 companies	MDA/5	Unstable companies	–	92.2%
			Stable companies	–	100.0%
	General/*Independent sample* – 344 companies	MDA/5	Unstable companies	–	92.8%
			Stable companies	–	100.0%
	General/343 companies	Logit	Unstable companies	–	97.2%
			Stable companies	–	100.0%

Source: Authors

The local bankruptcy literature overview provided in Table 1 defines the failure event differently. Novak (2003) analyses banks with business difficulties. Novak and Crnković (2007) classify companies as good and bad, i.e. those with which the bank had problems when collecting receivables. Zenezerović (2009) defines financially unstable companies as those who went bankrupt or disclosed losses exceeding their equity in their financial statements. In their research Šarlija and Jeger (2011) use distressed companies, defined as companies that cannot pay a single obligation continuously over the period longer than 90 days in one year. Pervan and Kuvek (2013) use defaulted and non-defaulted companies in order to estimate default prediction models with financial and non-financial predictor variables. Ježovita (2015) uses several financial criteria to classify a company as stable or unstable. These papers use the sample of companies with poor financial performance rather than bankrupt companies. Poor financial performers “rush” towards bankruptcy, but often never declare bankruptcy and eventually recover to continue their business. This paper uses open bankruptcy proceedings at the Court Registry as an indicator of distress and is more similar to Sajter (2008), Pervan, Pervan and Vukoja (2011) and Pervan and Kuvek (2018).

3. DATA AND VARIABLES

This paper uses two sources in the construction of the empirical dataset. Firstly, the Open Data Portal of the Court Registry is used to retrieve information on the entire population of 161,028 companies in Croatia as of March 1, 2020. The focus of the analysis is on profit-oriented, private and public companies, so non-government organizations and budgetary users are excluded from the sample. This sample consists of 3,992 companies with bankruptcy proceedings pending and represents the complete population of bankrupt companies in Croatia. It should be noted that the bankruptcy proceedings is only the first of six steps leading towards the full deletion of the company from the Court Registry, so the number of companies with debt-servicing difficulties in Croatia is significantly larger, amounting to 9,720.

This paper defines bankruptcy as the first official signal that the company is operating unsustainably and only companies with open bankruptcy proceedings are preserved in the empirical sample. The variables obtained for bankrupt companies are the company identification number (OIB) and the date of bankruptcy that was sampled for the period 2015-2019. Secondly, for each of the bankrupt companies the annual financial statements are collected from the Croatian Financial Agency (FINA), the official public database of business entities. The financial statements are sampled for the period 2014-2018 and merged with bankruptcy data in a way that the submitted financial statement precedes the company´s bankruptcy date by one year. Merging the data from the Court Registry and FINA was done by OIB code, and this made it possible to add a variety of financial statement variables to every bankrupt company.

This results in the largest available data set and covers the full-time distribution of declared bankruptcies in the positive business cycle that started after the 7 year long recession in Croatia. The number of bankruptcies depends on macroeconomic conditions (Bhattacharjee et al. 2004, Bruneau, Bandt and Amri 2008) and the choice of time period in this paper aims to reduce the business cycle bias that might arise from the inclusion of positive and negative periods of business cycle in the same sample. The common robustness checks in the related literature (Ward 1994, Poddig 1995, Šarlija and Jeger 2011) include sampling of financial statements preceding the bankruptcy date for two or more years but this paper doesn't find this approach to be appropriate due to the several reasons. The inclusion of more bankruptcy preceding years would in the case of the present analysis imply: 1) a significant reduction of the empirical sample size, 2) the distribution shrinkage of bankruptcy dates across the business cycles and 3) would produce incomparable results for different periods preceding bankruptcy.

To achieve a mixed sample of bankrupt and non-bankrupt[3] companies, the propensity score matching (PSM) model introduced by Rosenbaum and Rubin (1983) is used. PSM is the method for estimating the effect of receiving treatment when a random assignment of treatments to subjects is not feasible. This method pairs treatment and control units with similar values on the propensity score and other covariates. The PSM matching model is run on a sample of 634 companies that filed for bankruptcy one year[4] before the last financial statement was submitted to FINA and a complete dataset of non-bankrupt Croatian companies. The matching was executed using the nearest neighbour method and by controlling for the total assets, number of employees and sector of activity (National Classification of Activities-NKD2007). The procedure resulted in an equal number of bankrupt and similar non-bankrupt companies according to the specified control variables. Companies with missing accounting data were excluded from the sample. The initially matched sample of 1,268 companies was further cleaned of outliers and finally included 508 (46.2%) non-bankrupt and 591 (53.8%) bankrupt companies. The data sample of the business entities according to the sectoral classification is shown in table 2.

[3] As no financial indicators have been calculated to confirm that the companies are balance-sheet healthy, the term non-bankrupt is used to denote successful business in our data sample.

[4] In order to accommodate changing economic conditions and the loss of a bankruptcy model's predictive power over time (Šarlija and Jeger (2011), Zavgren (1985), one year prior to bankruptcy was used in the analysis.

Table 2.

OVERVIEW OF SAMPLE CLASSIFICATION SECTIONS ACCORDING TO THE NATIONAL CLASSIFICATION OF ACTIVITIES –NKD 2007

Code	Sector/Industry	Percentage
A	Agriculture, forestry and fishing	3.7%
B	Mining and quarrying	0.5%
C	Manufacturing	16.4%
D	Electricity, gas, steam and air conditioning supply	0.9%
E	Water supply, sewerage, waste management and remediation activities	0.8%
F	Construction	18.1%
G	Wholesale and retail trade; repair of motor vehicles and motorcycles	24.7%
H	Transportation and storage	3.5%
I	Accommodation and food service activities	7.3%
J	Information and communication	1.8%
K	Financial and insurance activities	0.3%
L	Real estate activities	6.2%
M	Professional, scientific and technical activities	9.2%
N	Administrative and support service activities	3.2%
P	Education	0.4%
Q	Human health and social work activities	0.6%
R	Arts, entertainment and recreation	1.5%
S	Other service activities	1.1%

Source: Authors' own calculations

The accounting information from financial statements was used to construct financial indicators. The seventeen financial ratios that cover main categories of financial ratios: liquidity, profitability, leverage, efficiency, and solvency were chosen as predictor variables and were calculated one year before bankruptcy. These indicators have been widely used in the related literature. Low liquidity serves as a signal of financial distress and increases default risk. Solvency indicators together with liquidity indicators are common factors in credit risk assessment. Profitability ratios are strong predictors of distress as they show how effectively a company is using its assets to generate earnings. According to Brindescu-Olariu (2016), the high profitability ratios are expected to have a positive impact on cash-flows and thus reduce payment difficulties and failure. Leverage ratios are also important indicators of financial distress since company with higher debt leverage are more likely to face financial difficulties when their debt falls due. Financial efficiency indicators describe how efficiently a business uses its assets to create a return or income whereby the higher ratio is expected to reduce probability of bankruptcy. The results in this analysis show that the effect of independent variables on bankruptcy outcome (Table 3) have signs as implied by theory.

Table 3.

DESCRIPTION OF INDEPENDENT VARIABLES

Variable	**Description**	**Financial indicator**	**Effect on bankruptcy outcome**
X1	cash/short term liabilities	liquidity	positive
X2	(cash + receivables)/short-term liabilities	liquidity	positive
X3	current assets/current liabilities	liquidity	positive
X4	current assets/total liabilities	liquidity	positive
X6	total debt/total assets	leverage	negative
X7	EBIT/Total assets	profitability	positive
X8	net income/total assets	profitability	positive
X10	working capital/total assets	liquidity	positive
X11	retained earnings/total assets	profitability	positive
X12	long term debt/total assets	solvency	negative
X13	net worth/total assets	solvency	positive
X14	operating income/total assets	profitability	positive
X15	cash flow from operations/total assets	efficiency	positive
X16	cash flow from operations/total outstanding debt	efficiency	positive
X19	cash/total assets	liquidity	positive
X20	short-term liabilities/total assets	liquidity	negative
X21	net worth/total liabilities	leverage	positive

Source: Authors' own calculations

The multicollinearity of the predictors was tested using Pearson correlation and a variance inflation factor–VIF, and shows high correlation. The VIF indicates the presence of multicollinearity, as shown in Table 5, so the following six variables were excluded from the sample: x1, x2, x6, x8, x11, x13 and x20.

Table 4.

CORRELATION MATRIX

	x1	***x2***	***x3***	***x4***	***x6***	***x7***	***x8***	***x10***	***x11***	***x12***	***x13***	***x14***	***x15***	***x16***	***x19***	***x20***	***x21***
x1	1																
x2	0.89	1.00															
x3	0.63	0.72	1.00														
x4	0.61	0.66	0.66	1.00													
x6	-0.09	-0.14	-0.15	-0.15	1.00												
x7	0.04	0.07	0.06	0.07	-0.37	1.00											
x8	0.04	0.06	0.06	0.06	-0.39	0.90	1.00										
x10	0.10	0.16	0.17	0.16	-0.92	0.36	0.37	1.00									
x11	0.06	0.11	0.11	0.11	-0.88	0.07	0.02	0.80	1.00								
x12	-0.01	-0.03	-0.01	-0.13	0.44	-0.10	-0.13	-0.12	-0.43	1.00							
x13	0.09	0.14	0.15	0.15	-1.00	0.37	0.39	0.92	0.88	-0.44	1.00						
x14	0.00	0.01	-0.03	0.05	-0.05	0.05	-0.02	0.05	0.09	-0.13	0.05	1.00					
x15	0.04	0.03	0.00	0.03	-0.15	0.57	0.56	0.14	-0.09	-0.03	0.15	0.07	1.00				
x16	0.17	0.17	0.17	0.28	-0.06	0.07	0.07	0.04	0.03	-0.05	0.06	0.03	0.26	1.00			
x19	0.44	0.37	0.22	0.32	-0.04	0.10	0.08	0.09	0.00	-0.03	0.04	0.27	0.15	0.16	1.00		
x20	-0.09	-0.15	-0.16	-0.11	0.93	-0.37	-0.37	-0.98	-0.80	0.08	-0.93	0.00	-0.16	-0.05	-0.03	1.00	
x21	0.18	0.21	0.24	0.35	-0.10	0.03	0.03	0.06	0.05	-0.07	0.10	-0.03	0.01	0.23	0.04	-0.08	1.00

Source: Authors' own calculations

The estimation with the reduced number of variables shows that the Pearson correlation matrix[5] does not have coefficients higher than 0.7, and the highest VIF (table 5) was 2.15, so the multicollinearity issue is avoided.

[5] Due to space limitation table is available at the request.

Table 5.

VARIANCE INFLATION FACTOR

Collinearity Statistics		
	Tolerance	**VIF**
x3	0.547	1.829
x4	0.465	2.152
x7	0.589	1.697
x10	0.832	1.202
x12	0.940	1.064
x14	0.901	1.110
x15	0.614	1.628
x16	0.824	1.214
x19	0.808	1.238
x21	0.851	1.175

Source: Authors' own calculations

4. METHODOLOGY AND RESULTS

This paper uses multivariate discriminant analysis (MDA) and logistic regression (LR) as the two most common and widely employed methods in predicting bankruptcy.

4.1. Multivariate discriminant analysis

The aim of multivariate discriminant analysis is to classify observations (company) by a set of independent variables $X = (x_1, x_2, x_n)$ into one of two or more categories (bankruptcy and non-bankruptcy). If each observation's discriminant score Z_i is a linear function of X_i, it is possible to write a discriminant function that linearly separates the observations as:

$$Z_i = \beta_0 + \beta_i X_{i1} + \beta_{i2} X_{i2} + \ldots + \beta_n X_{im}. \quad (1)$$

The discrimination boundary Z^* is defined by the set of points where

$$\beta_0 + \beta_i X_{i1} + \beta_{i2} X_{i2} + \ldots + \beta_n X_{im} = Z'. \tag{1.1}$$

The discriminant function maps observations of two different categories from n-dimensional attribute space into low dimensional space that maximises the separation of the groups. The classification procedure uses linear combinations of independent variables to estimate coefficients in a way that the ratio of the squared distance between the means of the two groups ($\bar{Z}_1$ *and* $\bar{Z}_2$) to the variance of Z is maximised. If the means of the vector X for the two groups are μ_1 and μ_2, and their respective covariance matrices θ_1 and θ_2, the means of the linear function Z in the two groups are $b\mu_1$ and $b\mu_2$ where b is a vector of discriminant coefficients $b_i s$. Under the assumption of covariance equality $\theta_1 = \theta_2 = \theta$, the variance of the discriminant function Z is, so the:

$$\varphi = \frac{\left[b\left(\mu_1 + \mu_2\right)\right]^2}{b\theta b} = \frac{\textit{between group variance}}{\textit{within group variance}} \tag{1.2}$$

needs to be maximised. Taking the first derivative with respect to b and setting the φ to zero yields:

$$b = \theta^{-1}\left(\mu_1 - \mu_2\right). \tag{1.3}$$

The parameters θ^{-1}, μ_1, μ_2 can be estimated by the sample mean $\bar{X}_1, \bar{X}_2$ and sample variance S^{-1} so that the means of the discriminant functions can be expressed as:

$$\bar{Z}_1 = b'\bar{X}_1 = \left(\bar{X}_1 - \bar{X}_2\right)' S^{-1}\bar{X}_1 \tag{1.4}$$

$$\bar{Z}_2 = b'\bar{X}_2 = \left(\bar{X}_1 - \bar{X}_2\right)' S^{-1}\bar{X}_2. \tag{1.5}$$

If $\bar{Z}_1$ is greater than $\bar{Z}_2$, Z_0 will be closer to $\bar{Z}_1$ than to $\bar{Z}_2$ if the condition $\left|Z_0 - \bar{Z}_1\right| > \left|Z_0 - \bar{Z}_2\right|$ is satisfied. This implies that $Z_0 > \frac{1}{2}\left(\bar{Z}_1 + \bar{Z}_2\right)$ so that the optimal cut-off point Z^* defines the average of the two means. The square of the difference between the means is called Mahalanobis (generalised) distance and can be written as:

$$D^2 = \left(\bar{Z}_1 - \bar{Z}_2\right)^2 = \left(\bar{X}_1 - \bar{X}_2\right)' S^{-1}\left(\bar{X}_1 - \bar{X}_2\right). \tag{1.6}$$

Under the assumption that independent variables come from normally distributed populations with means μ_1, μ_2 and have equal covariance matrix θ, the F ratio can be used to test statistically significant differences between the groups:

$$F = \frac{n_1 n_2 (n_1 + n_2 - k - 1)}{(n_1 + n_2)(n_1 + n_2 - 2)k} D^2 \qquad (1.7)$$

where k is the number of independent variables.

The stepwise MDA was estimated using the predictors: x19, x10, x14, x3, x12, x7 and x21.The specification check of the equality of the group means is entered in the Table 6 test and shows that all variables differ, since sig.<1%.

Table 6.

TESTS OF EQUALITY OF GROUP MEANS

	Wilks' Lambda	**F**	**df1**	**df2**	**Sig.**
x19	0.908	110.550	1	1097	0.000
x10	0.938	72.386	1	1097	0.000
x14	0.930	81.957	1	1097	0.000
x3	0.929	83.837	1	1097	0.000
x12	0.968	36.241	1	1097	0.000
x7	0.963	42.450	1	1097	0.000
x21	0.980	22.618	1	1097	0.000

Source: Authors' own calculations

Another measure of variable's potential is Wilks' Lambda, an index of the discriminating power which ranges from 1 (no discriminatory power) to 0 (perfect discriminatory power). Smaller values indicate higher discriminatory power between groups. Table 6 suggests the following variable ranking: x19, x3, x14, x10, x7, x12 and x21.

According to the Box's M in the Table 7, the variance-covariance matrices are not equal. However, the Box's M is very sensitive to large data files Feldesman (2002), meaning that when there is a large number of cases, this assumption is very often violated. According to Landau and Everitt (2004, p. 317), "in practice, Box's test is not of a great use since even if it suggests a departure for the equality hypothesis, the linear discriminant may still be preferable over a quadratic function. " According to Salkind (2010, p. 371), "when sample sizes are large across groups, the significance test of discriminant function is usually robust with respect to the violation of the homogeneity assumption." The same issue is present in Pervan, Pervan and Vukoja (2011), Situm (2015), Memić (2015), Svabova, Durica and Podhorska (2018), Kliestik, Vrbka and Rowland (2018) and many other authors. It is assumed that this violation will not have a large impact on developing the model, but the results should be interpreted with caution.

Table 7.

BOX'S M TEST RESULTS

Box's M		8602.959
F	Approx.	305.165
	df1	28
	df2	4002696.213
	Sig.	0.000

Source: Authors' own calculations

The canonical correlation is calculated as the Pearson correlation between the discriminant function scores and group variables (0 and 1). The canonical correlation is 0.489 and 23.9% of the variance in the discriminant function scores can be explained by group differences. Wilks' lambda quality of the discriminant power and significance of the function for the proposed model is 0.761 with Chi-square 298.86 and significance lower than 1%, so the corresponding function explains the group membership well. According to the classification test results (Table 8.), the model accuracy is 92.9% for non-bankrupt companies, 51.4% for bankrupt companies and its overall accuracy is 73.7%.

Table 8.

CLASSIFICATION RESULTS

			Predicted Group Membership		
		fail	0.0	1.0	Total
Original	Count	0.0	549	42	591
		1.0	247	261	508
	%	0.0	92.9	7.1	100.0
		1.0	48.6	51.4	100.0
Overall percentage					73.7

Source: Authors' own calculations

The final empirical specification has good accuracy but somewhat lower than similar papers that apply MDA on the global sample of companies like: Altman (1968) whose model achieves 95% accuracy, Deakin (1972) with 77% accuracy of prediction for bankrupt companies and 82% accuracy for healthy companies and Blum (1974) whose model produces overall accuracy of 87%. The MDA estimation results in this analysis are close to Libby (1975) that reports accuracy of 74%. It is important to note that all of these papers use a smaller data sample than the present analysis. Results in the present analysis are similar to Sajter (2008) although he uses only one control variable. Pervan, Pervan and Vukoja (2011) use three control variables but achieve a lower accuracy prediction of bankrupt companies (79.5%) and higher accuracy of healthy companies (80.8%).

4.2. LOGIT

The second model used in the analysis is logistic regression (logit). The logit is a special form of a regression model that can predict and explain a binary categorical outcome conditional on a set of metric and non-metric independent variables. The logit model does not require the normality of distribution assumption for independent variables or equality of covariance matrices of two groups and provides meaningful interpretation of probability that an observation will fall within a given group. The model assumes the existence of a latent dependent variable:

$$Z_i^{'} = \sum_{i=0}^{n} b_j X_{ij} + \varepsilon_i = b' X_i + \varepsilon_i \tag{2}$$

Where $Z_i^{'}$ is the outcome variable, b_j are the coefficients of independent variables X_j that need to be estimated, X_{ij} is the i-th observation's value of the j-th independent variable and ε_i is the residual of the model, assumed to be random and identically distributed with zero mean.

Since Z' is not observable, a logit binary classification problem can be related to an observable dummy variable Y so that $Y = 1$ if the company is bankrupt ($Z_i^{'} > 0$) and $Y = 0$ for the non-bankrupt company. It is now possible to write:

$$P(Y_i = 1) = P(\varepsilon_i > -\beta' X_i) = 1 - F(-\beta' X_i) \tag{2.1}$$

where F is the cumulative distribution function for ε. As the observed values of Y are realizations of a binomial process with probabilities specified in (2) and varying from trial to trial depending on X_i, the likelihood function of observing Y_i can be specified as:

$$LF = \prod_{Y=0} F\left(-\beta' X_i\right) \prod_{Y=1} \left[1 - F\left(-\beta' X_i\right)\right] \tag{2.2}$$

where the functional form of LF depends on the cumulative distribution of ε from Equation (2) and describes the relationship between the response and control variables. The logistic function is chosen to be:

$$P = F(Z) = {}^{1}\!\big/_{1+e^{-Z_\theta}} \quad \text{where} \quad 0 < P_\theta < 1 \; and \; Z_\theta = \beta' X_i \tag{2.3}$$

P_θ represents the conditional probability of any X_i and β so with the condition that the cumulative distribution of the residual structure is logistic, the logit model is defined as:

$$F\left(-\beta' X_i\right) = {}^{1}\!\big/_{1+\exp\left(\beta' X_i\right)} \quad \text{and}$$

$$1 - \left(-\beta' X_i\right) = 1 - {}^{1}\!\big/_{1+\exp\left(\beta' X_i\right)} = {}^{\exp\left(\beta' X_i\right)}\!\big/_{1+\exp\left(\beta' X_i\right)} \tag{2.4}$$

The logit model as specified in (2.4) and (2.5) assumes that P_θ represents the probability of bankruptcy for the i-th company with the characteristic X_i and can be described as the logistic cumulative distribution function that asymptotically approaches zero and one. The estimation of the logit coefficients Z is best done using the maximum likelihood method (MLE) in (2.2), through multiplication of products of all P_i for bankrupt companies times the product of all instances $1 - P_i$ for all non-bankrupt companies. Therefore, higher failure probabilities for bankrupt companies and lower failure probabilities for non-bankrupt companies represent higher points on the likelihood function. The likelihood function can now be specified as:

$$\begin{aligned}
Lf &= \prod_{i=1}^{n} P\left(Y_i = 0\right) \prod_{i=1}^{n} P\left(Y_i = 1\right) \\
&= \prod_{i=1}^{n} P\left(\varepsilon_i < -\beta' X_i\right) \prod_{i=1}^{n} P\left(\varepsilon_i > -\beta' X_i\right) \\
&= \prod_{y=o}^{n} F\left(-\beta' X_i\right) \prod_{y=1}^{n} \left[1 - F\left(-\beta' X_i\right)\right] \\
&= \prod_{y=o}^{n} \left({}^{1}\!\big/_{1+\exp\left(\beta' X_i\right)} \right)^{1-y_i} \prod_{y=1}^{n} \left({}^{\exp\left(\beta' X_i\right)}\!\big/_{1+\exp\left(\beta' X_i\right)} \right)^{y_i} \\
&= \frac{\exp(\beta') \sum_{i=1}^{n} X_i Z_i}{\prod_{y=o}^{n} \left[1 + \exp\left(\beta' X_i\right)\right]}
\end{aligned} \tag{2.5}$$

The coefficients of interest b and β can be obtained by finding the global maximum of the logarithm of the likelihood function. In the process of differentiation of the likelihood function, the Newton-Rapson iterative method is recommended due to the non-linearity of the partial derivatives in β. MLE estimates can be considered consistent and efficient for a large number of observations, so probabilities P can be achieved through the estimated parameters. The probabilities of being in the group k will lie between 0 and 1, irrespective of Z. Then, observations can be classified into the group to which they have the highest predicted probability of belonging.

The forward stepwise LR logit method is employed to estimate the model with the best fit and accuracy precision. Variables that enter the model are described in Table 9.

Table 9.

VARIABLES IN THE EQUATION

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
x3	0.263	0.059	20.154	1	0.000	1.300
x7	3.942	0.673	34.270	1	0.000	51.517
x10	0.655	0.161	16.560	1	0.000	1.925
x12	-0.791	0.202	15.291	1	0.000	0.453
x14	0.538	0.114	22.114	1	0.000	1.713
x19	4.156	1.072	15.043	1	0.000	63.847
Constant	-0.540	0.137	15.607	1	0.000	0.583

Source: Authors' estimation

All six explanatory variables in the final logit function and constant are statistically significant. As expected, all coefficients have a positive sign, except x12, which is negative. The overall fit of the logit model is assessed by using the Nagelkerke R Square. According to the results from Table 10, the model explains 46.5% of the variability of the dependent variable.

Table 10.

LOGISTIC REGRESSION MODEL SUMMARY

-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1046.815[a]	0.348	0.465
a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.		

Source: Authors' estimation

According to the results of Hosmer and Lemeshow Test, it can be concluded that the data fit the model well, which is confirmed by p > 5%.

Table 11.

HOSMER AND LEMESHOW TEST

Chi-square	df	Sig.
9.699	8	0.287

Source: Authors' estimation

The classification table summarises the results of the prediction. The next table correctly classifies 86.1% of bankrupt companies and 64.8% of non-bankrupt companies. The overall classification percentage is 76.3%.

Table 12.

CLASSIFICATION TABLE

Observed		Predicted		
		fail		Percentage Correct
		0.0	1.0	
fail	0.0	509	82	86.1
	1.0	179	329	64.8
Overall Percentage				76.3

Source: Authors' estimation

Relative to MDA model, the logistic regression is performing slightly better (overall accuracy 73.7%). Both methodologies use same predictor variables except that the logit model uses one more control variable, x21. Compared to the similar literature that uses logit model on the global sample of companies, the model classification accuracy in this paper is somewhat lower than Ohlson (1980) who achieves 96% accuracy but higher than Zavgren (1985) who reports 69% accuracy. Luoma and Laitinen (1991) achieve a lower accuracy for bankrupt companies (73.5%) but higher accuracy for predicting non-bankrupt companies (70.6%). There is only a small number of papers that use logit model in the analysis of bankruptcy prediction in the local literature. Among those, Pervan, Pervan and Vukoja (2011) achieve 85.9% accuracy of classification for bankrupt companies and slightly higher result for healthy companies on the sample of 156 companies. Pervan and Kuvek (2013) use logit model on the sample of 825 companies (127 unable to repay debts and 698 non defaulted companies) with several specifications. In the first specification with 4 control variables authors achieve 52% accuracy for defaulted and 88.4% accuracy for non-defaulted companies. In the second specification, they use 7 control variables and improve overall accuracy for defaulted companies to 64.6% and non-defaulted companies to 92.4%. Even though Pervan and Kuvek (2013) do not use legal bankruptcy as a failure event, the results are comparable with this paper. Other papers that apply the logit model on the sample of Croatian companies use various indicators of financial difficulties as a failure (bankruptcy) proxy and shouldn't be compared to results in this analysis directly.

5. CONCLUSION

Bankruptcy prediction literature is a well-established field of economic research, but this topic is still insufficiently explored in Croatian literature, particularly in terms of data representativity and comparability of results. The data sample in this analysis includes 1,099 companies created through the propensity score matching of legally declared bankrupt and similar non-bankrupt companies. The paper uses MDA and logit models to estimate the impact of a set of financial ratios on the incidence of bankruptcy. After conducting MDA stepwise procedure, seven variables that include liquidity, solvency and profitability indicators entered the final model. The overall accuracy of MDA model was 73.7%; it correctly classified 51.4% of non-bankrupt companies and correctly classified 92.9% of bankrupt companies. The logit model achieved a slightly higher overall accuracy of 76.3%, correctly classifying 64.8% of non-bankrupt companies and 86.1% bankrupt companies.

The paper contributes to the literature by construction of the empirical data sample where company failure is identified by the explicit and legally defined bankruptcy event. Furthermore, the full population of failed companies is matched to similar non-bankrupt companies. In this way the analysis provides the benchmark bankruptcy estimation for Croatian companies on the representative sample. The paper uses discriminant analysis that is based on the assumption that the group covariance matrices are equal but the empirical specification in this analysis does not accommodate the respective assumption completely, so the results should be interpreted accordingly, keeping in mind that this is a common problem with large samples. Future research should strive to accommodate this assumption, possibly by applying other bankruptcy prediction methods and to provide additional robustness to the results.

REFERENCES

1. Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4): 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x
2. Altman, E. I., & McGough, T. P. (1974). Evaluation of a Company as a Going Concern. *Journal of Accountancy*, (December): 50–57.
3. Araghi, M. K., & Makvandi, S. (2013). Comparing logit, probit and multiple discriminant analysis models in predicting bankruptcy of companies. *Asian Journal of Finance & Accounting*, 5(1), 48. https://doi.org/10.5296/ajfa.v5i1.2977
4. Back, B., Laitinen, T., & Sere, K. (1996). Neural networks and genetic algorithms for bankruptcy predictions. *Expert Systems with Applications*, 11(4), 407-413. https://doi.org/10.1016/S0957-4174(96)00055-3
5. Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. Expert Systems with Applications, 83, 405-417. https://doi.org/10.1016/j.eswa.2017.04.006
6. Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, 71-111. https://doi.org/10.2307/2490171
7. Betts, J., & Belhoul, D. (1987). The effectiveness of incorporating stability measures in company failure models. *Journal of Business Finance & Accounting*, 14(3), 323-334. https://doi.org/10.1111/j.1468-5957.1987.tb00098.x
8. Bhattacharjee, A., Higson, C., Holly, S., & Kattuman, P. (2009). Macroeconomic instability and business exit: Determinants of failures and acquisitions of UK firms. Economica, 76(301), 108-131.

9. Blum, M. (1974). Failing company discriminant analysis. *Journal of Accounting Research*, 12(1): 1–25. https://doi.org/10.2307/2490525
10. Bogdan, S., Bareša, S., & Hađina, V. (2019). Testing the applicability of the Altman's Z-score model for predicting bankruptcy in the Republic of Croatia. *Notitia-časopis za održivi razvoj*, 5, 31-46. https://doi.org/10.32676/n.5.1.4
11. Bruneau, C., De Bandt, O., & El Amri, W. (2012). Macroeconomic fluctuations and corporate financial fragility. Journal of Financial Stability, 8(4), 219-235.
12. Carmona, P., Climent, F., & Momparler, A. (2019). Predicting failure in the US banking sector: An extreme gradient boosting approach. *International Review of Economics & Finance*, 61, 304-323. https://doi.org/10.1016/j.iref.2018.03.008
13. Chi, L. C., & Tang, T. C. (2006). Bankruptcy prediction: Application of logit analysis in export credit risks. *Australian Journal of Management*, 31(1), 17-27. https://doi.org/10.1177%2F031289620603100102
14. Croatian Financial Agency — FINA. https://www.fina.hr/-/informacija-o-neizvrsenim-osnovama-za-placanje-poslovnih-subjekata-za-prosinac-2019- (accessed 16.02.2020)
15. Deakin, E. B. (1972). A discriminant analysis of predictors of business failure. *Journal of accounting research*, 10(1): 167–179. https://doi.org/10.2307/2490225
16. Edmister, R. O. (1972). An empirical test of financial ratio analysis for small business failure prediction. *Journal of Financial and Quantitative analysis*, 7(2): 1477–1493. https://doi.org/10.2307/2329929
17. Feldesman, M. R. (2002). Classification trees as an alternative to linear discriminant analysis. *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, 119(3), 257-275. https://doi.org/10.1002/ajpa.10102
18. Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: an application using bankruptcy data. *Information & Management*, 24(3), 159-167. https://doi.org/10.1016/0378-7206(93)90064-Z
19. Fitzpatrick, P. J. (1932). *A comparison of the ratios of successful industrial enterprises with those of failed companies*. The Certified Public Accountant. October, November, December, 598-605.
20. Gardiner, L. R., Oswald, S. L., & Jahera Jr, J. S. (1996). Prediction of hospital failure: A post-PPS analysis. *Journal of Healthcare Management*, 41(4), 441.
21. Jakubík, P., & Teplý, P. (2008). The prediction of corporate bankruptcy and Czech economy's financial stability through logit analysis (No. 19/2008). IES Working Paper.

22. Ježovita, A. (2015). Designing the model for evaluating business quality in Croatia. Management: journal of contemporary management issues, 20(1), 101-129.
23. Kim, H., & Gu, Z. (2006). Predicting restaurant bankruptcy: A logit model in comparison with a discriminant model. *Journal of Hospitality & Tourism Research*, 30(4), 474-493. https://doi.org/10.1177/1096348006290114
24. Kliestik, T., Vrbka, J., & Rowland, Z. (2018). Bankruptcy prediction in Visegrad group countries using multiple discriminant analysis. Equilibrium. *Quarterly Journal of Economics and Economic Policy*, 13(3), 569-593. http://dx.doi.org/10.24136/eq.2018.028
25. Kozjak, S. K., Šestanj-Perić, T., & Bešvir, B. (2014, January). Assessment of Bankruptcy Prediction Models Applicability in Croatia. In 7th International Conference" An Enterprise Odyssey: Leadership, Innovation and Development for Responsible Economy".
26. Landau, S. and Everitt S. (2004). A handbook of statistical analyses using SPSS. London: CRC Press Company.
27. Laitinen, E. K., & Laitinen, T. (2000). Bankruptcy prediction: Application of the Taylor's expansion in logistic regression. *International review of financial analysis*, 9(4), 327-349. http://dx.doi.org/10.1016/S1057-5219(00)00039-9
28. Libby, R. (1975). Accounting ratios and the prediction of failure: Some behavioral evidence. *Journal of Accounting Research*, 13(1): 150–161. https://doi.org/10.2307/2490653
29. Lo, A. W. (1986). Logit versus discriminant analysis: A specification test and application to corporate bankruptcies. *Journal of econometrics*, 31(2), 151-178. https://doi.org/10.1016/0304-4076(86)90046-1
30. Luoma, M., & Laitinen, E. K. (1991). Survival analysis as a tool for company failure prediction. *Omega*, 19(6), 673-678. https://doi.org/10.1016/0305-0483(91)90015-L
31. Martin, D. (1977). Early warning of bank failure: A logit regression approach. *Journal of banking & finance*, 1(3), 249-276. https://doi.org/10.1016/0378-4266(77)90022-X
32. Mihalovič, M. (2016). Performance comparison of multiple discriminant analysis and logit models in bankruptcy prediction. *Economics & Sociology*, 9(4), 101.
33. Memić, D. (2015). Assessing credit default using logistic regression and multiple discriminant analysis: Empirical evidence from Bosnia and Herzegovina. Interdisciplinary Description of Complex Systems: INDECS, 13(1), 128-153.

34. Moyer, R. C. (1977). Forecasting financial failure: a re-examination. Financial Management (pre-1986), 6(1), 11.
35. Norton, C. L., & Smith, R. E. (1979). A comparison of general price level and historical cost financial statements in the prediction of bankruptcy. *Accounting Review*, 72-87.
36. Novak, B. (2003). Predviđanje poslovnih teškoća banaka u republici hrvatskoj na osnovi javno dostupnih financijskih pokazatelje. *Ekonomski pregled*, 54(11-12), 904-924.
37. Novak, B., & Crnković, I. (2007). Classification of bank debtor distress based on official financial statements. *Ekonomski pregled*, 58(1-2), 41-71.
38. Odom, M. D., & Sharda, R. (1990, June). A neural network model for bankruptcy prediction. In 1990 IJCNN International Joint Conference on neural networks (pp. 163-168). IEEE.
39. 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
40. Olariu, D. B. (2016). Profitability ratio as a tool for bankruptcy prediction. SEA–Practical Application of Science, 4(11), 369-372.
41. Pervan, I., & Kuvek, T. (2013). The relative importance of financial ratios and nonfinancial variables in predicting of insolvency. *Croatian Operational research review*, 4(1), 187-197.
42. Pervan, I., Pervan, M., & Kuvek, T. (2018). Firm Failure Prediction: Financial Distress Model vs Traditional Models. *Croatian Operational Research Review*, 9(2), 269-279.
43. Pervan, I., Pervan, M., & Vukoja, B. (2011). Prediction of company bankruptcy using statistical techniques–Case of Croatia. *Croatian Operational Research Review*, 2(1), 158-167.
44. Platt, H. D., Platt, M. B., & Pedersen, J. G. (1994). Bankruptcy discrimination with real variables. *Journal of Business Finance & Accounting*, 21(4), 491-510. https://doi.org/10.1111/j.1468-5957.1994.tb00332.x
45. Poddig, T. (1995). Bankruptcy prediction: A comparison with discriminant analysis. Neural Networks in the Capital Markets. Wiley, 311-323.
46. Ramser, J. R., Foster, L. O. (1931). A demonstration of ratio analysis. Bureau of Business Research, Bulletin 4, Urbana, University of Illinois.
47. Rose, P. S., & Giroux, G. A. (1984). Predicting corporate bankruptcy: an analytical and empirical evaluation. *Review of Financial Economics*, 19(2), 1.

48. Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55. https://doi.org/10.2307/2335942
49. Sajter, D. (2008). Ekonomski aspekti stečaja i restrukturiranja u stečaju. Ekonomski fakultet Osijek, raspoloživo na: https://bib. irb. hr/datoteka/377526. Sajter_Disertacija.pdf (02.07. 2017.).
50. Salkind, N. J. (Ed.). (2010). Encyclopedia of research design (Vol. 1). Sage.
51. Sinkey, J. F. (1975). A multivariate statistical analysis of the characteristics of problem banks. *The Journal of Finance*, 30(1): 21–36. https://doi.org/10.1111/j.1540-6261.1975.tb03158.x
52. Situm, M. (2015). The relevance of trend variables for the prediction of corporate crises and insolvencies. *Zagreb International Review of Economics and Business*, 18(1), 17-49. https://doi.org/10.1515/zireb-2015-0002
53. Smith, R. F., & Winakor, A. H. (1935). *Changes in the financial structure of unsuccessful industrial corporations*. University of Illinois.
54. Svabova, L., Durica, M., & Podhorska, I. (2018). Prediction of default of small companies in the Slovak Republic. *Economics and Culture*, 15(1), 88-95.
55. Šarlija, N., & Jeger, M. (2011). Comparing financial distress prediction models before and during recession. *Croatian Operational Research Review*, 2(1), 133-142.
56. Šarlija, N., Penavin, S., & Harc, M. (2009). Predviđanje nelikvidnosti poduzeća u Hrvatskoj. *Zbornik Ekonomskog fakulteta u Zagrebu*, 7(2), 21-36.
57. Tam, K. Y., & Kiang, M. Y. (1992). Managerial applications of neural networks: the case of bank failure predictions. *Management science*, 38(7), 926-947. http://dx.doi.org/10.1287/mnsc.38.7.926
58. Wall, A. (1936). How to evaluate financial statements. New York: Harper
59. Ward, T. J. (1994). An empirical study of the incremental predictive ability of Beaver's naive operating flow measure using four-state ordinal models of financial distress. Journal of Business Finance & Accounting, 21(4), 547-561.
60. Wilcox, J. W. (1973). A prediction of business failure using accounting data. *Journal of Accounting Research*, 11: 163–179. http://dx.doi.org/10.2307/2490035
61. Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. *Journal of Financial and Quantitative Analysis*, 15(3), 757-770. https://doi.org/10.2307/2330408
62. Zavgren, C. V. (1985). Assessing the vulnerability to failure of American industrial firms: a logistic analysis. *Journal of Business Finance & Accounting*, 12(1), 19-45. https://doi.org/10.1111/j.1468-5957.1985.tb00077.x

63. Zenzerović, R. (2009). Business- Financial Problems Prediction-Croatian Experience. *Economic research-Ekonomska istraživanja*, 22(4), 1-15.
64. Zięba, M., Tomczak, S. K., & Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, 58, 93-101.
65. Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting research*, 59-82. https://doi.org/10.2307/2490859

PREDVIĐANJE STEČAJA NA TEMELJU UKUPNE POPULACIJE HRVATSKIH PODUZEĆA

Sažetak

Ovaj rad analizira predviđanje bankrota na temelju pune populacije poduzeća koje predstavljaju ukupni poslovni sektor u Hrvatskoj. Reprezentativnost uzorka postigla se tehnikom uparivanja (propensity score matching) ukupne populacije bankrotiranih i sličnih poduzeća koje nisu u stečaju. Robusna procjena predviđanja bankrota provela se temeljem višestruke diskriminacijske analize (MDA) i logističke regresije (logit). Rezultati su ukazali na visoku točnost klasifikacije oba modela, s naglaskom na povoljniju procjenu uporabom logit metode. Ukupna točnost MDA modela bila je 73,7%, dok je ukupna točnost logit modela bila 76,3%. Rezultati služe kao referentna točka za procjenu bankrota za poslovni sektor u Hrvatskoj.

Ključne riječi: Multipla diskriminantna analiza, MDA, logistička regresija, logit, financijski pokazatelji