PREDICTION OF COMPANY BANKRUPTCY USING STATISTICAL TECHNIQUES – CASE OF CROATIA

Ivica Pervan
Faculty of Economics, University of Split
Matice hrvatske 31, 21000 Split
Phone: ++ <21 430 639>; E-mail: <pervan@efst.hr>

Maja Pervan
Faculty of Economics, University of Split
Matice hrvatske 31, 21000 Split
Phone: ++ <21 430 680>; E-mail: <mpervan@efst.hr>

Bruno Vukoja
Faculty of Economics, University of Split
Matice hrvatske 31, 21000 Split
E-mail: <b1051024@efst.hr>

Abstract

Bankruptcy prediction research in Croatia is still pretty limited and due to data unavailability researches use different samples and statistical techniques that result with diverse findings regarding the variables that most accurately predict business failure. In this research the authors have decided to use a sample of differently sized bankrupted companies from manufacturing and wholesale/retail trade industry. Bankruptcy prediction model was developed on the basis of only publicly available information on bankruptcy and financial statements. The bankruptcy sample data consists of 78 companies that have declared bankruptcy in the Official Gazette during 2010. An equal number of stabile companies sample was randomly selected from the same industry sectors. All financial data were collected from the web site of FINA. The research has shown that publicly available financial statements and calculated financial ratios have informational value since they can be effectively used for prediction of companies' bankruptcy.

Key words: *Bankruptcy, Prediction, Statistical techniques, Croatia*

1. INTRODUCTION

Bankruptcy represents the situation in which company is unable to settle its liabilities (to banks, suppliers, employees, tax authorities, etc) and therefore, according to law, company enters the bankruptcy procedure. The Croatian judicial practice shows that bankruptcy procedures often result with selling of company assets, termination of employees' contracts and company operations. Namely, empirical data in Croatia for the period 2000-2005 indicate that 5,597 bankruptcies were started while only in 7 cases restructuring plan was adopted. In all other cases company assets were sold and companies were liquidated (Sajter, 2008). The reason for such devastating statistics regarding the low rate of recovery of companies in bankruptcy probably lies in the fact that the bankruptcy procedure in Croatia is started very late i.e. when it is too late for restructuring. Bankruptcy procedures are usually started when companies in financial problems have very high liabilities in comparison to their assets. Very often liquidation value of company assets is not sufficient to settle all the liabilities and therefore some creditors suffer losses. In order to avoid such situation companies should take into account financial health of their clients before establishing business relations with them. Significant limitation in development of bankruptcy prediction model may be limited information on bankruptcies and financial statements. Therefore, in this study bankruptcy prediction model is based only on publicly available information for Croatian companies. Bankruptcy prediction research has drawn attention of many researchers and many papers develop different models for business failure prediction. Review of early (Beaver & Altman) and recent research is presented in the second section of the paper. The third section deals with data description, variables and methodology. Empirical findings are presented in the fourth section, which is followed by conclusion.

2. LITERATURE REVIEW

Bankruptcy prediction literature has it roots in the 1930-es, when first authors started to analyze financial profile of companies with risk of business failure (Fitz Patrick, 1932). More sophisticated analysis on business failure has started with Beaver (1967) research. Beaver defines failure as inability of a firm to pay its financial obligations. His research sample consisted of 79 listed failed firms in period 1954-1964. Each failed firm was matched with a non-failed firm from the same industry and with similar asset size. Empirical findings indicated that the best univariate discriminator between failed and non-failed firms was cash flow to total debt, where cash is calculated as net income plus depreciation, depletion and amortization. Classification accuracy of the cash flow to total debt was in the range from 87% (one year before failure) to 78% (five years before failure).

Instead of univariate analysis used by Beaver, Altman (1968) uses multivariate technique, multiple discriminate analysis–MDA in order to develop model for prediction of bankruptcy. His research sample consisted of 33 listed bankruptcy firms and 33 listed nonbankruptcy firms from manufacturing sector. The sample of nonbankruptcy firms was matched by industry, size and year. Altman used 22 financial ratios, but MDA model indentified only five ratios (working capital/assets, retained earnings/assets, EBIT/assets, market value of equity/book value of equity and sales/assets) as significant bankruptcy predictors. Classification accuracy of Altman's model was 95%, while model error was 5%, when testing was done on estimation sample and with data one year before bankruptcy. Classification error has increased to 17% percent with data two years before bankruptcy, indicating classification accuracy of 83%. Deakin (1972) has raised some methodological issues regarding the Altman's 1968 bankruptcy prediction model. Namely, one of basic MDA assumptions is that observations in each group are randomly selected. Altman did not use random selection, but match pair sample approach. Deakin used randomly selected 11 failed and 23 non-failed firms and developed failure prediction model. Classification error of the model was relatively low up to three years before bankruptcy (3-4.5%), but for fourth and fifth year before bankruptcy prediction error has sharply risen (21% and 17%).

The first research on SME business failure was done by Edminister (1972) who also used MDA statistical technique to discriminate among loss and nonloss SME borrowers. Empirical analysis has resulted with MDA model with seven variables financial ratios. Classification accuracy of Edminister's model was 93%, while model error was 7%. Research also revealed that models predictive power depends on ratios calculation approach. Namely, dividing ratios with industry averages has shown to be useful technique. Significant methodological changes in bankruptcy prediction research were introduced by Ohlson paper (1980). Firstly, he used logit model instead of MDA, since MDA has the following main characteristics: requirement for normality of predictors, requirement for the same variance-covariance matrices for both groups and MDA score has little intuitive interpretation. The logit model on the other hand does not have assumptions on a priori probabilities and distribution of predictors. Ohlson did not use match pair sample, but he used 105 listed bankruptcy firms and randomly chosen 2,058 listed nonbankruptcy firms. Of nine ratios included into analysis only the following four appeared to be statistically significant: size, total liabilities to total assets, net income to total assets and working capital to total assets. Classification accuracy of Ohlson's model was 96.3%.

All previously mentioned research was done on cross sectional data. However, some researchers focus on selected industries. For example, bankruptcy prediction study for retail sector was done by Bhargava et al. (1998). The study compared bankruptcy predictive power of Altman's Z-Score and two single performance measures (cash flow and inventory turnover), which are popularly believed to be particularly relevant in the retail sector. Logit model for the retail sector indicated that Z-Score had better bankruptcy predictive power in comparison with cash flow and inventory turnover. Also, ROA outperformed cash flow and inventory turnover, while cash flow outperformed inventory turnover in terms of bankruptcy predictive power. Research for 42 insolvent small Portuguese firms from the footwear manufacturing industry used MDA and logistic regression (Pindado and Rodrigues, 2004). Sample of 42 insolvent firms (with loan default) was matched with sample of 42 solvent firms from the same industry. Empirical findings indicate that in both models (MDA and logistic regression) the following two variables are significant for insolvency prediction: accumulated profit/total assets and interest charge/total income. Classification accuracy for MDA model was 89.58%, while for logistic regression it was 91.67% when estimation sample was used. But testing the prediction models on the data for holdout sample indicated decline in predictive accuracy (77.95% for MDA and 75.98% for logistic regression).

Bankruptcy prediction research has become very popular among academics and many studies are conducted all over the world. Analysis of corporate bankruptcy during the Korean economic crisis (1997-1998) was conducted by Nam & Jinn (2000). They used logit model with data for 46 listed bankrupted companies and 46 matched (by industry and size) nonbankrupted listed companies. Authors used 33 financial ratios, while only the following 3 ratios have been significant in prediction of bankruptcy: financial expenses/sales, (net income + depreciation + financial expenses)/(total borrowings + bonds payable + financial expenses) and receivables turnover. Classification accuracy for estimation sample was 77.2%, while for holdout sample it was little lower (76.2%). Analysis of business failure of Turkish public and private companies was done by Vuran (2009), who used data for 78 failed companies in period 1999-2007. Failed companies sample included major types of failure: default on loan obligation, reorganizing debt structure and bankruptcy. Failed firms were randomly matched with data for 91 nonfailed firms (from the same industry and year). As statistical methods author uses MDA and logistic regression. MDA analysis resulted with finding that the following variables discriminate failed and nonfailed firms:

- total debt/total assets, cash flow from operations/interest expense and net profit/total assets (one year before bankruptcy)
- short term debt/total assets and sales/total assets (two years before bankruptcy)

Classification accuracy of model with data one year before failure was 84.4%, while model with data for two years before failure had classification accuracy of 80.3%. Logistic regression resulted with finding that the same variables like in MDA models were significant for failure prediction. Classification accuracy of the logistic regression model with data one year before failure was 84.4%, while the model with data for two years before failure had classification accuracy of 82%.

MDA model is also used by Lugovskaja (2009) in paper that predicts bankruptcy of Russian SMEs. Research sample consisted of 260 bankrupted and 260 randomly chosen nonbankrupted SMEs. The first MDA model resulted with finding that the following six variables were significant for bankruptcy prediction: cash/current liabilities, current assets/current liabilities, (cash + short term debtors)/current liabilities, current liabilities/total assets, ROA and cash/ total assets. The second MDA model besides financial ratios included non-financial variables (size and age), which both appear to be significant. Model with financial ratios only had classification accuracy for estimation sample of 76.2%, while for holdout sample it was 68.1%. Model with financial ratios and mentioned non-financial variables had higher classification accuracy (77.9% for estimation sample and 79% for holdout sample).

In Croatia, according to the available literature only a few authors have analyzed company failure prediction. Their research was based on different definition of failure variable, different sampling approaches, different approach of models efficiency measurement and therefore findings can not be directly compared. Sajter (2008) developed bankruptcy prediction model on the sample of 72 healthy and 18 bankrupted companies from the four selected counties in the eastern Croatia. With usage of MDA, only the ratio of working capital/total assets was significant, while model total accuracy was 88% (only 50% accuracy for bankrupted companies). The logit model besides ratio of working capital/total assets, as statistically significant variable has identified ratio of long lived assets/(equity+long term debt). The logit model total accuracy was 90%, although like in the case of MDA model predictive power for bankrupted companies was only 50%. Šarlija et al. (2009) developed model for prediction of insolvency on the sample of 4,213 insolvent firms and 55,903 solvent firms. Insolvency prediction model was based on 14 financial ratios and 2 non-financial ratios, while its predictive accuracy was 68.16% for solvent firms and 74.22% for insolvent firms, when tested on holdout sample. Zenzerović (2009) used stratified sample (by size and industry) of 55 stabile and 55 unstable Croatian companies. Unstable companies' definition includes companies which went into bankruptcy procedure or have loss over equity. Usage of MDA resulted with final model in which 5 variables were significant discriminators, while total model's accuracy was 95.3%, while testing was done on the estimation sample.

3. DATA AND METHODOLOGY

As mentioned earlier, the purpose of this research is to develop model that would be useful for bankruptcy prediction using only publicly available information. Therefore, firstly we had to identify companies that went into bankruptcy, which was done be empirical insight on the web site of Official Gazette[1] that publishes such data. Namely, according to the regulation in Croatia when company enters into the official bankruptcy procedure such information must be published via Official Gazette. Analysis was done for the period January-June 2010 discovering that 78 companies from manufacturing and wholesale/trade sector went into bankruptcy. Selected industry sectors together account for about 47% of the Croatian registered legal entities. Also, the sample of financial healthy companies had to be defined and we decided to use an equal number of randomly selected companies. After the samples were defined the next step was to obtain publicly available financial information for all 156 companies that were taken into research. Financial information was obtained from the web site of Croatian Financial Agency, since from 2009 all companies in Croatia have to send their financial statements to the Agency[2]. A very important element of analysis is selection of financial ratios that should explain probability of bankruptcy. In selection of ratios we decided to use all major groups of ratios: liquidity, activity, financial structure, profitability and cash flow. On the basis

[1] http://narodne-novine.nn.hr/
[2] http://rgfi.fina.hr/JavnaObjava-web/izbornik.do

of financial statements from 2007 we calculated the initial set of 15 financial ratios that were often used in bankruptcy literature. As statistical methodology for bankruptcy prediction discriminant analysis-DA and logistic regression-LR are used. DA has assumptions for ratios normality and equal dispersion of covariance matrices for groups defined by dependent variable. Many studies report that the mentioned two assumptions are violated and therefore results might be questionable. Although the evidence on the issue of sensitivity of DA in case of violated assumptions is mixed, the results must be interpreted with potential impacts of violated assumptions. Especially in the case of small samples and unequal covariance matrices statistical significance of the estimation can be adversely affected (Hair, et. al, 2010). When DA assumptions are violated a good alternative is LR. Namely, LR does not have requirements for data normality of equal dispersion of covariance matrices of groups. With fewer assumptions LR can be used more efficiently than DA, when assumptions of data normality and equal dispersion of covariance matrices is not met.

An important issue in application of both methods can be the problem of multicollinearity of independent variables. Since some of financial ratios use the same variables in the calculation (assets, net income, liabilities, etc) there is real possibility of multicollinearity problem in the estimated model. The problem of multicollinearity in the estimated model causes inefficiently estimated parameters and high errors, which in turn results with many insignificant variables and high explanatory power of the estimated model. In order to control this problem we have decided to use two approaches. The first one is the usage of matrix of Pearson Correlation coefficients, where correlation higher than 0.8 indicates multicollinearity problem. The second test for multicollinearity can be done by Variance Inflation Factors – VIFs, where linear regression of one discriminating variable was run, while all other variables were used as explanatory variables. This auxiliary regression model resulting with VIFs less than 5 indicates that the estimated model is free of multicollinearity. Matrix of Pearson Correlation coefficients revealed that in our analysis many financial ratios were highly correlated, since they had coefficients higher than 0.8. After elimination of highly correlated independent variables only the following six financial ratios were selected:

Table 1: Description of the selected financial ratios

Financial ratio	Description
Current liquidity-CL	Current assets/Current liabilities
Net working capital-NWC	(Current assets-Current liabilities)/Current assets
Ratio of liquid assets-RLA	Current assets/Total assets
Leverage-LEV	Total debt/Total assets
Assets turnover-ATURN	Sales/Total assets
EBIT	EBIT/Total assets

Regression model was run with CL as dependent and all other variables as independent, resulted with all VIFs less than 5, also confirming that the estimated model is free of multicollinearity.

4. EMPIRICAL FINDINGS

In the estimation part of research we firstly decided to use the DA stepwise procedure. But before running DA it was necessary to test for independent variables normality. In order to test for variables normality we used Kolmogorov-Smirnov test that is supported by SPSS statistical package used for the analysis. The test resulted with a finding that only RLA variable was normally distributed while all other variables did not have normal distribution. An important step in DA analysis is a test of group means equality, which revealed that RLA and ATURN did not have significant mean differences, while CL, NWC, LEV and EBIT had significant differences between healthy and bankrupted firms.

Table 2: Tests of Equality of Group Means

Variable	Wilks' Lambda	F	df1	df2	Sig.
CL	0.905	16.175	1	154	0.000
NWC	0.878	21.463	1	154	0.000
RLA	0.986	2.178	1	154	0.142
LEV	0.817	34.509	1	154	0.000
ATURN	0.991	1.369	1	154	0.244
EBIT	0.893	18.458	1	154	0.000

Box M test with significance less than 5% revealed that groups did not have equal covariance matrices, indicating that second DA assumption is violated. Canonical correlation was 0.531, which means that model explains 53.1% of the variation in the grouping variable. Wilks Lambda was 0.718 (Chi-Square 50.53), with significance less than 1% indicating significance of estimated DA model.

Table 3: Canonical Discriminant Function Coefficients

Variable	Unstandardized coefficient
CL	0.165
LEV	-0.681
EBIT	3.389
Constant	0.236

DA model unstandardized coefficients from the table 3 reveal that LEV has negative sign meaning that increase of this variable reduces discriminate score and increased probability of being bankrupted. Contrary to such finding, CL and EBIT have positive sign, meaning that increase of these variables causes increase of discriminate score and decreases probability of being bankrupted. Here we can see that discrimination among bankrupted and healthy firms can be done by financial ratios from three groups (liquidity, leverage and profitability). After estimation phase of research follows the classification. Here we used cross-validated procedure of classification (i.e. jack knife classification), which is more honest than classical classification. The data from the table 4 reveal that DA model accuracy is 80.8% for healthy firms, 79.5% for bankrupted firms and 80.1% in total.

Table 4: Classification Results

			Predicted group		
			Healthy firm	Bankruptcy firm	Total
Cross-validated	Count	Healthy firm	63	15	78
		Bankruptcy firm	16	62	78
	%	Healthy firm	80.8	19.2	100,0
		Bankruptcy firm	20.5	79.5	100,0

LR model was run with usage of same six independent variables as reported in the Table 1. The forward stepwise procedure was done in 3 steps. In the final third step LR models Chi-square was 7.31, with significance 0.7% indicating that the overall fitting of the estimated model is good. Another approach of measuring the model fitting is Nagelkerke R Square. In this case Nagelkerke R Square is 66.9% indicating moderate strong relationship between the used financial ratios and bankruptcy prediction.

Table 5: Variables in the LR Equation

Variable	B	S.E.	Wald	df	Sig.
RLA	-1.172	0.385	9.262	1	0.002
LEV	6.510	1.458	19.926	1	0.001
EBIT	-13.555	5.803	5.456	1	0.019
Constant	-4.644	1.236	14.111	1	0.001

Variables that entered in the final LR model were RLA, LEV and EBIT. The results of LR model are similar to that of DA model since ratios from liquidity, leverage and profitability are shown to be significant for prediction of bankruptcy. The only difference is that in DA model CL was indicator of liquidity, while in LR model liquidity significant variable is RLA. LEV variable has the positive sign indicating that increase of this variable increases probability of being bankrupted. Contrary to that RLA and EBIT variable have negative sign indicating that the increase of these variables decreases the probability of being bankrupted. The LR models classification accuracy for healthy firms was 80.8%, for bankrupted firms 85.9%, while total accuracy was 83.3%.

5. CONCLUSION

Bankruptcy prediction has been an interesting topic among academics for many years, while research in Croatia is pretty limited. The main objective of the study was to explore whether only publicly available information on bankruptcy and financial statements can be useful for the development of bankruptcy prediction model. Research was conducted on the sample of 78 bankrupted companies and 78 healthy companies from Croatian manufacturing and trade/wholesale industries. As statistical methods DA and LR are used. Use of DA has shown to be problematic since two major assumptions were violated: data normality and equality of covariance matrices. The estimated model has shown that current liability, leverage and ratio of EBIT to total can be effectively used for bankruptcy prediction. DA models accuracy in prediction of bankrupted companies is 79.5%. LR model is more robust since it does not have the previously mentioned assumptions. LR model resulted with similar findings since final model entered leverage and ratio of EBIT to total assets, while significant liquidity measure was ratio of liquid to total assets. LR model had higher classification accuracy (85.9%) in comparison with DA. The research has shown that publicly available financial statements and calculated financial ratios have informational value since they can be effectively used for prediction of companies' bankruptcy. As a limitation of the research we must point out the accuracy of the estimated DA model since two major assumptions were violated. Also, the future analysis might be done with the data for longer period, which would reveal a time dimension of financial ratios relevance in the bankruptcy prediction**.**

LITERATURE:

Altman, E. I. (1968), "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *Journal of Finance*, Vol. 23, No. 4, pp. 889-609.

Beaver, W. (1967), "Financial Ratios as Predictor of Failure", *Empirical Research in Accounting, Empirical Studies, Journal of Accounting Research*, Supplement to Vol. 4, pp. 71-111.

Bhargava, M., Dubelaar, C. and Scott, T. (1998), "Predicting bankruptcy in the retail sector: an examination of the validity of key measures of performance", *Journal of Retailing and Consumer Services*, vol. 5, No. 2, pp. 105-117.

Deakin, E. B. (1972), "A Discriminant Analysis of Predictors of Business Failure“, *Journal of Accounting Research*, Vol. 10, No. 1, pp.167-179.

Edminster, R. O. (1972), "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction", *Journal of Financial and Quantitative Analysis*, March, pp. 1477-1493.

Fitz Patrick, P., J. (1932), "A Comparison of Ratios of Successful Industrial Enterprises with those of failed Firms", *Certified Public Accountant*, October, pp. 598-605; November, pp. 656-662; December, pp. 727-731.

Hair, J., F., Black, W., C., Babin, B., J., Anderson, R., E., (2010): *Multivariate Data Analysis*, Pearson Prentice Hall.

Lugovskaja, L. (2009), "Predicting default of Russian SMEs on the basis of financial and non-financial variables", *Journal of Financial Services Marketing*, Vol. 14, No. 4, pp. 301-313.

Ohlson, J. A. (1980), "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research,* Vol. 18, No. 1, pp.109-131.

Nam, J-H. and Jinn, T. (2000), "Bankruptcy prediction: Evidence from Korean Listed Companies During the IMF Crisis", *Journal of International Financial Management and Accounting*, Vol. 11, No. 3, pp. 178-197.

Pindado, J. and Rodrigues, L., F. (2004), "Parsimonious Models of Financial insolvency in Small Companies", *Small Business Economics*, No. 22, pp. 51-66.

Sajter, D. (2008), *Ekonomski aspekti stečajeva i restrukturiranja u stečaju*, Doktorska disertacija, Ekonomski fakultet Osijek, Osijek.

Šarlija, N., Penavin, S. and Harc, M. (2009): "Predviđanje nelikvidnosti poduzeća u Hrvatskoj", *Zbornik Ekonomskog fakulteta u Zagrebu*, No. 2, pp. 21-36.

Vuran, B. (2009), "Prediction of business failure: a comparison of discriminat and logistic regression analyses", *Istanbul University Journal of the School of Business Admistration*, Vol. 38, No. 1, pp. 47-65.

Zenzerović, R. (2009), "Business financial problem prediction-Croatian experience", *Ekonomska istraživanja*, Vol. 22, No. 4, pp. 1-16.