Theoretical and Applied Economics
Volume XX (2013), No. 5(582), pp. 7-14

Early warning models of financial distress. Case study of the Romanian firms listed on RASDAQ

Mădălina Ecaterina ANDREICA
The Bucharest University of Economic Studies
madalina.andreica@gmail.com

Abstract. *In this paper we design an early warning model for Romanian distressed firms. The logit model was built based on financial ratios of 66 Romanian firms listed on RASDAQ that were facing financial difficulties in 2011. In addition, we identified the main principal components obtained with minimum loss of information after applying the principal component analysis and proposed a new estimation of the logit model by replacing the initial set of input data matrix with the main principal components of the financial observations. The results indicated an increase with 12 percentage points in the performance of the one year ahead prediction of financial distress of the new warning model.*

Keywords: distress prediction; financial crisis; logit model; principal component analysis.

JEL Codes: C53, G32.
REL Codes: 11B, 17B.

1. Introduction

The early warning models of unproductive firms that are confronted with financial difficulties and higher risk of bankruptcy bring a true decision support in the context of the current financial crisis, not only by offering warning signs of financial distress, but also by helping minimizing the computational time for determining a firms risk of insolvency when requesting funding to a bank.

In this paper we aimed to build an efficient one-year-ahead prediction model of Romanian unproductive firms and to establish the main financial ratios that contribute to rapid screening of the firms with high financial distress risk. In order to do that we collected financial ratios of 66 Romanian firms listed on RASDAQ during 2009-2010, out of which half were facing penalties for late payments in 2011 and had an unproductive behavior.

In addition, we identified the main components obtained after applying the principal component analysis with minimal loss of information and proposed a new estimation of a logit model by replacing the initial set of input data with their main principal components and reached better prediction results.

The issue concerning the early warning models of financial distress or bankruptcy gained great interest since the late 1960s. The first step in this direction was made by Beaver (1966), who proposed a method to classify firms into two groups by using a t-test statistic. Beaver's study was continued by Altman (1968), who proposed a model of multivariate discriminant analysis, while Eisenbeis (1977), Ohlson (1980) and Jones (1987) identified some inconsistencies in the method and proposed new methods of financial distress prediction.

For instance, in 1980 Ohlson proposed a logit model for financial distress prediction and his results showed that size, financial structure, performance and current liquidity were important determinants of bankruptcy. Then in 1984 Zmijewski (1984) proposed the probit model in order to predict the bankruptcy risk of a firm, but the model had less applicability in the literature.

Instead Shumway (2001) proposed the hazard model for predicting bankruptcy, that was defined as a multi-period logit model. The main particularities of the hazard model consist in the fact that the explanatory variables vary during several time periods, leading to more efficient estimators. Moreover, the duration model uses a baseline hazard function that can be estimated directly through macroeconomic variables in order to reflect the economic changes (Nam, Kim, Park, Lee, 2008).

In recent years many types of heuristic algorithms such as neural networks and decision trees have also been applied to the bankruptcy prediction problem and several improvements in the financial distress prediction were noticed. For instance, the studies made by Tam and Kiang (1992), Salchenberger et al. (1992) and Jain and Nag (1998) concluded that neural networks outperform conventional statistical models of financial distress prediction. Soon after that, hybrid Artificial Neural Network methods were proposed in some financial distress prediction studies such as Yim and Mitchell (2005) and Andreica (2009).

Instead, Zheng and Yanhui (2007) and Andreica (2012) used decision tree models for corporate financial distress prediction and presented the advantages of using CHAID decision trees in comparison to a neural network model, which is complicated to build up and to interpret or to a statistic model such as multivariate discriminate regression and logistic regression, where the patterns need to be linearly separable and samples are assumed to follow a multivariate normal distribution. In their study, Zheng and Yanhui (2007) noticed that it is not appropriate to use financial information to predict financial distress ahead of four years and the results showed that decision trees are a valid model to predict listed firms' financial distress in China, with a 80% probability of correct prediction.

As noticing from the literature review, the early warning models for distress companies have been intensively studied starting with 1960 and still play a huge role especially during financial crisis. In this context, early warning models could be of great help to prevent financial difficulties or bankruptcy of a firm and to protect the companies that are offering financing to any potential financial distressed firm.

2. Metodologies

In order to build predictive models of financial distress in Romania, some logit econometric models were estimated and then a principal component analysis was applied.

According to Shumway (2001) *the logit model* is a single-period classification model which uses maximum likelihood estimation to provide the conditional probability of a firm belonging to a certain category given the values of the independent variables for that firm. The logit model describes the relationship between a dichotomous variable Y, that takes values 1 or 0 for "distress" and "non-distress", respectively, and k explanatory variables $x_1, x_2, ..., x_k$, representing seven financial ratios. The logit regression model is defined as follows:

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \sum_{j=1}^{k} \beta_j x_{i,j})}} \tag{1}$$

When applying the logit-transformation to the above equation, we get a linear relationship between logit(p_i) and the explanatory variables:

$$\text{logit}(p_i) = \log\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \sum_{j=1}^{k} \beta_i x_{i,j} \tag{2}$$

This equation is also called the logit form of the model, where logit(p_i) is the log odds of distress for the given values $x_{i,1}$, $x_{i,2}$, …, $x_{i,k}$ of the explanatory variables.

Regarding the threshold value to which we compare the predicted probabilities of a firm to become productive or unproductive we chose the level of 0.5. Thus, if the estimated probability is less than 0.5, the observation is classified as a productive firm, while the probability values above 0.5 indicate an unproductive firm.

The Principal component analysis (PCA) is a way of identifying patterns in data where graphical representation is not available, by reducing the number of dimensions, without much loss of information. PCA involves a mathematical procedure that reduces the dimensionality of the initial data space by transforming a number of possibly correlated variables into a smaller number of uncorrelated variables called *principal components*. These components are synthetic variables of maximum variance, computed as a linear combination of the original variables. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible.

3. Building the early warning model

For this study, public financial information for the years 2009-2010 was collected for 66 Romanian listed companies on RASDAQ, corresponding to the same III-R market category with the purpose to predict financial distressed firms one year ahead. In order to have two equal sample groups of both financial distressed and non-distressed firms in the year 2011, 33 "distressed" firms with losses and outstanding payments were selected in the study together with 33 "non-distressed" companies, that were matched by assets size and activity field.

The selection of the main set of financial ratios for each company was conditioned by those variables that appeared in most empirical work, but also restricted to the availability of the financial data provided by the Bucharest Stock Exchange Market. Seven financial ratios were calculated for this study in order to reflect the company's profitability, asset utilization and growth ability. They are presented in Table 1.

Table 1

Financial indicators

CATEGORY	CODE	FINANCIAL RATIOS	DEFINITION
Profitability	**I1**	Profit Margin	Net Profit or Loss / Turnover $\times$ 100
	I2	Return on Assets	Net Profit or Loss / Total Assets $\times$ 100
	I3	Profit per employee	Net Profit or Loss / number of employees
	I4	Operating Revenue per employee	Ln(Operating revenue / number of employees)
Asset utilization	**I5**	Working capital per employee	Working capital / number of employees
	I6	Total Assets per employee	Ln(Total Assets / number employees)
Growth ability	**I7**	Growth rate on total assets	(Total Assets$_1$ – Total Assets$_0$) / Total Assets$_0$

The initial sample of 66 firms was divided into two groups: 76% of the initial observations were used in the learning process (in-sample), while the rest of 24% were used for prediction (out-of-sample). The prediction efficiency of the model was based on the 24% of the observations used for prediction and was calculated by comparing the forecasting results with the actual values.

After estimating the econometric model based on the financial indicators registered in 2010, it resulted the following unifactorial model, with the profit margin indicator (I1) as the explanatory variable and the following logit equation:

$$p_i = P(y_i = 1) = \frac{1}{1 + e^{-(-0,372 - 0,0648\ \times\ I_1)}}$$

Although the McFadden R-squared value is only 32.5%, the result of the *Expectation Predicted Test* suggests that the model brings a total gain of 34% in comparison to the simple constant model and the level of 0.2 of *the goodness of fit Test* (H-L Statistics) confirms the validity of the model. Moreover, the Akaike and Schwartz information criterion values are small and the coefficient sign of variable I1 is according to the economic theory, stating that a reduced level of the profit margin implies higher risks of financial distress.

The out-of-sample forecasting performance of the logit model built based on financial ratios of the year 2010 in order to predict financial distress of the Romanian firms one year ahead is only 69%, according to Table 2.

Table 2

The forecasting performance of the logit model

	IN SAMPLE			OUT-OF-SAMPLE		
	Healthy	**Unhealthy**	**TOTAL**	**Unhealthy**	**Healthy**	**TOTAL**
Total	25	25	50	8	8	16
incorect	8	0	8	5	0	5
corect	17	25	42	3	8	11
% incorect	32.0	0.0	16.0	62.5	0.0	31.3
% corect	68.0	100.0	**84.0**	37.5	100.0	**68.8**

4. Improving the efficiency of the early warning model

A principal component analysis was then performed in order to reduce the dimensionality of the original data space with minimum loss of information and in order to see to what extent are the variables used in the analysis relevant for obtaining classification rules for the two types of firms. The eigenvalues obtained from PCA application are shown in Table 3.

Table 3

Total Variance Explained

Compo-nent	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	*2.548*	36.400	36.400	2.548	36.400	36.400	2.056	29.371	29.371
2	*1.640*	23.426	59.827	1.640	23.426	59.827	1.998	28.547	57.918
3	*1.306*	18.663	78.490	1.306	18.663	78.490	1.440	20.572	78.490
4	.748	10.679	89.169						
5	.344	4.910	94.079						
6	.260	3.709	97.788						
7	.155	2.212	100.000						

Extraction Method: Principal Component Analysis.

The first three principal components that have values higher than 1 are the following: $\lambda_1 = 2.55$, $\lambda_2 = 1.64$ and $\lambda_3 = 1.31$, with a minimum loss of information of approximately 21.5%.

The first principal component is highly correlated to I1 and I7, being a profitability indicator. The second principal component is mostly correlated to I2, I3 and I5 and offers information about firms assets and working capital, while the third principal component describes the economic activity of a firm according to its employees, being strongly correlated to I4 and I6 (Table 4).

Table 4

Rotated component matrix(a)

	Component		
	1	2	3
I1	.874	.049	.014
I2	.565	.709	-.096
I3	.269	.782	.117
I4	.545	-.173	.758
I5	-.151	.902	-.072
I6	-.265	.098	.914
I7	.714	.170	-.039

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 5 iterations.

The graphical representation of the 66 firms on a three-dimensional principal component space is shown in Figure 1, where one can notice that, in general, the two types of firms tend to form two distinct groups.

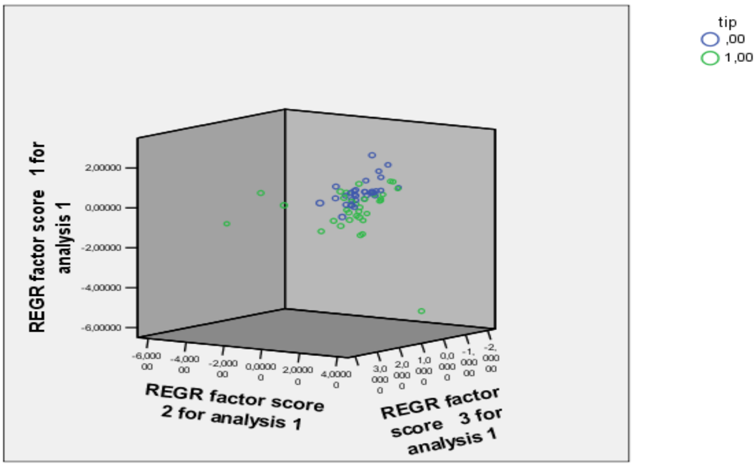

Figure 1. *The firms on a three-dimensional space*

The PCA allowed us to reduce the dimensionality of the initial data space with minimum loss of information so that the two types of firms could be easily identified.

In addition to that, we found that when replacing the financial indicators of the prediction model with the main principal components of the initial data matrix, the prediction performance of the new logit model built with the first two principal components improved and reached an accuracy of prediction of 81.3%. The new econometric model is described by the equation:

$$p_i = P(y_i = 1) = \frac{1}{1 + e^{-(0{,}354 - 3{,}043 \times fact_1 - 1{,}394 \times fact_2)}}$$

The R^2 McFadden value reaches 39.7%. Moreover, the new model brings a total gain of 28% in comparison to the simple constant model and the level of 0.21 of *the goodness of fit Test* (H-L Statistics) confirms the validity of the model

The out-of-sample forecasting performance of the new logit model built based on the first two principal components of the initial financial data of the year 2010 in order to predict financial distress of the Romanian firms one year ahead is improving by reaching 81.3%.

Table 5

The forecasting performance of the new logit model built on principal components

	IN SAMPLE			OUT-OF-SAMPLE		
	Healthy	Unhealthy	TOTAL	Unhealthy	Healthy	TOTAL
Total	25	25	50	8	8	16
incorect	9	8	17	0	3	3
corect	16	17	33	8	5	13
% incorect	36.0	32.0	34.0	0.0	37.5	18.8
% corect	64.0	68.0	66.0	100.0	62.5	81.3

5. Conclusions

In this paper we design an early warning model for Romanian distressed firms. The logit model was built based on financial ratios of 66 Romanian firms listed on RASDAQ that were facing financial difficulties in 2011. In addition, we identified the main principal components obtained with minimum loss of information after applying the principal component analysis and proposed a new estimation of the logit model by replacing the initial set of input data matrix with the main principal components of the financial observations.

The results indicated an increase with 12 percentage points in the performance of the one year ahead prediction of financial distress of the new warning model.

References

Altman, E.I. (1968). “Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”, *Journal of Finance*, 23, pp. 589-609

Andreica, M.E. (2012). “Financial Distress Prediction of the Romanian Companies Using CHAID Models”, *Metalurgia International*, nr. 12, pp. 196-200, ISSN 1582-2214

Andreica, M.E. (2009). “Early warning models for financial distress: The case of the Romanian companies”, *Metalurgia International,* vol. 14, special Issue nr. 11, pp. 174-178

Andreica, M.E. (2009). “Predicting Romanian Financial Distressed Companies”, *Advances in Economic and Financial Research - DOFIN Working Paper Series 37*, Bucharest University of Economics, Center for Advanced Research in Finance and Banking - CARFIB, http://econpapers.repec.org

Beaver, W. (1966). “Financial ratios as predictors of failure”, *Journal of Accounting Research* (Supplement), 4, pp. 71-102

Eisenbeis, R. (1977). “Pitfalls in the application of discriminant analysis in business, finance and economics”, *Journal of Finance*, 32, pp. 875-900

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, pp. 83-110

Jones, F.L. (1987). “Current techniques in bankruptcy prediction”, *Journal of Accounting Literature*, 6, pp. 131-164

Nam, C.W., Kim, T.S., Park, N.J., Lee, H.K. (2008). “Bankruptcy prediction using a Discrete-Time Duration Model Incoprorating Temporal and Macroeconomic dependencies”, *Journal of Forecasting*, 27, pp. 493-506

Ohlson, J.A. (1980). “Financial ratios and the probabilistic prediction of bankruptcy”, *Journal of Accounting Research*, 18, pp. 109-131

Salchenberger, L.M., Cinar, E.M., Lash, N.A. (1992). “Neural networks: A new tool for predicting thrift failures”, *Decision Sciences,* 23, pp. 899-916

Shumway, T. (2001). “Forecasting bankruptcy more accurately: A simple hazard model”, *Journal of Business*, 74 (1), pp. 101-124

Yim, J., Mitchell, H. (2005). “A comparison of corporate distress prediction models in Brazil: hybrid neural networks, logit models and discriminant analysis”, *Nova Economia Belo Horizonte,* 15, 73-93

Zheng, Q., Yanhui, J. (2007). “Financial Distress Prediction on Decision Tree Models”, *IEEE*