PREDICTING COMPANY PERFORMANCE BY DISCRIMINANT ANALYSIS

Dumitra STANCU[1]
Andrei-Tudor STANCU[2]

ABSTRACT
This paper aims at evaluating the performance through discriminant analysis of 20 companies traded on the Bucharest Stock Exchange (BVB). As these companies are similar in terms of business profile (manufacturing industry), we choose ten financial indicators that relate to stock value (PRICE, BETA, ALPHA, etc.) and book value (Debt / Equity, ROA and ROE) to assess and classify the companies as "good" or "bad". For a sustainable characterization the average value of the financial indicators is estimated between the first quarter of 2005 and third quarter of 2013. The initial grouping is made according to return on assets (ROA) and splits the sample into 10 "good" and 10 "bad" companies. We find that discriminant analysis correctly validates the classification of firms by ROA criterion in 90% of cases (18 of 20 companies). Moreover, our analysis establishes that ROA is of first importance in evaluating company performance as suggested by the F test-statistic and Wilks'Lambda coefficient.

KEYWORDS: *scoring function, multiple discriminant analysis, insolvency, and default risk.*

JEL CLASSIFICATION: *C10, G30, L25*

1. INTRODUCTION

Several academic researchers, as well as financial institutions, have been interested in developing formal methods for predicting the performance of firms by developing a score function. The methodology used is the statistical technique of discriminant analysis that considers several economic and financial aspects of the company to determine its future performance. It starts with a selection of indicators closely related to the financial health of the company and then searches for a linear combination of these indicators that best differentiate between companies more likely to do well and those more likely to do worse in the next period.
This linear combination is used to build an indicator, called Z score, which gives a good approximation of performance status, or risk status, for a given company when compared to its pairs. The Z score for a company *i* can be formalized as:
$Z_i = a_0 + a_1R_1 + a_2R_2 + ... + a_nR_n$

where:
- a_0 is a constant measuring the unexplained part of the score;
- $R_1, R_2 ... R_n$ are financial indicators (e.g. profitability, structure capital, risk, etc.) of the companies considered and are assumed to be independent among themselves;
- $a_1, a_2 ... a_n$ represent the elasticity coefficients of the Z score to unit changes in financial indicators and are estimated through simple least squares estimation.

[1] Universitatea Tehnică de Construcţii - Bucureşti, stancudumitra@yahoo.com
[2] Henley Business School at the University of Reading, stancudoru@yahoo.com.

The main objective in developing a reliable scoring function that discriminates between companies as "good" or "bad" based on financial indicators (profitability, risk, etc.) is to make the score distributions of these two groups most independent from each other. Specifically, we seeks the scoring function that maximizes the distance between the average score of good companies and the average score of bad companies while at the same time minimizes the standard deviation of scores in each group:

$$\max \quad \overline{Z}_B - \overline{Z}_R$$

$$\min \ \sigma(Z_B) \text{ and } \min \ \sigma(Z_R)$$

where:

- $\overline{Z}_B$ and $\overline{Z}_R$ represent the average score of “good” and “bad” companies, respectively,
- $\sigma(Z_B)$ and $\sigma(Z_R)$ are the standard deviations of the “good” and “bad” companies scores.

This way the overlapping area between the two distributions is minimised offering, thus, the lowest degree of ambiguity in the classification of companies (see dashed area in Figure 1).

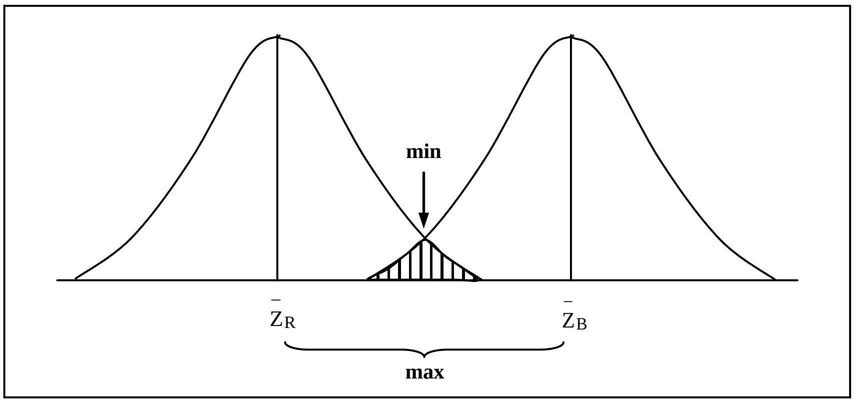

Figure 1. Frequency distribution of scores on “good” and “bad” companies

2. LITERATURE REVIEW

In an influential study, Robu et al. (2012) present an excellent classification of performance prediction models and determine four main categories (Robu et al., 2012) A first category concerns models developed according to **the accounting system of the country**, Anglo-Saxon models, continental models and national models. The second category involves models that depend on **the degree of development of the country**, developed versus developing countries models. The third group of models has a focus on the **information set employed to create a score function**. This group contains models based on financial ratios, on cash–flow, on variations in outcome and non-financial models. Finally, the last group is formed according to **the type of analysis used to develop the model or function score**: discriminant analysis, logit and probit regression analysis, adopting heuristic algorithms as neural networks or decision trees.

Jaba et al. (2007) provide a methodology to classify the eight regions and 42 counties of Romania using a set of variables that characterize the economic and social development in each of these areas. Their framework can be used by the government when making its investment policy decision to reduce the development gaps between these regions and counties.

Mironiuc et al. (2010) study the influence in variance of several economic and financial indicators of a company on its financial autonomy (by the total equity / total debt ratio as proxy). To do so, they perform a score function to classify 80 companies traded on the BVB in companies with high, medium, low and no financial autonomy. They find that companies with no financial autonomy (i.e. high level of debt) have the highest level of financial profitability during the time period 2006-2008.

Jaba and Robu (2009a, 2009b) perform a discriminatory function to classify 60 Romanian companies in three categories according to their economic and financial results (i.e. leader, middle and least). The authors performed two functions of discrimination and group the Romanian companies by their economic and financial indicators but also by region, the activity and number of employees. It follows thus one function for each group according to the classification of companies in Leader, Middle and Least.

Armeanu et al. (2012) use an Altman type score on a sample of 60 companies listed on the Bucharest Stock Exchange to determine the relevant score thresholds specific to the Romanian stock market. They focus on seven financial indicators to capture several operating aspects of the company (i.e. total assets, net turnover, operating profit, operating cash flow, net profit, total debt and average market capitalization). Their multivariate analysis yields the following three areas of risk: companies with a Z score lower than -2.34 belong to the safe area and are associated with very low probability of default; companies with a Z score between -2.34 and -0.102 are included in the uncertain area; and companies with a Z score higher than -0.102 are regarded as having increased default risk and are included in the risky area.

3. DATA, METHODOLOGY AND IMPLEMENTATION

We use discriminant analysis to predict whether a firm can be classified as "good" or "bad". Our sample consists of 20 companies traded on the Bucharest Stock Exchange (BVB) that are similar in terms of business profile (manufacturing industry). For each company, we obtain quarterly data on ten financial indicators between the Q1 2005 and Q3 2013 from Thompson Reuters. Where missing, we complete the time series with data from Bloomberg, BSE and KTD.
The average values of these financial indicators are presented in Table 1. For a sustainable characterization, we estimate the average value of these financial indicators registered between first quarter 2005 and third quarter 2013. Since our time sample spams over the 2007-2009 subprime crisis and over the Euro-zone debt crisis starting in 2009, the stock returns have often been negative for our 20 BVB companies selected. In order to harmonize the scale of the variables we apply a simple standardization on the averages for PRICE, MKT_CAP, and FF_MKT, expressing their size by the ratio of the average to standard deviation of these variables.

Our main aim is to classify the companies according to 10 relevant financial indicators: Price, Beta, Alpha, Market Capitalization, Free Float Market Capitalisation, Price Earnings Ratio, Price Book Ratio, Debt to Equity, ROE and ROA. We start by defining a state variable that takes a value of 1 when the company is perceived as being “good” and a value of 0 when the company is perceived as being “bad”. The first 10 companies with highest average return on assets (ROA) for the period 2005-2006 are assigned to the “good” group leaving the rest of the companies to be included in the “bad” group. Next, this classification will be validated by discriminant analysis.

Table 1. Q1 2005 to Q3 2013 standardized data and average of financial indicator by company

	Firm	PRICE	BETA	ALPHA	MKT_CAP	FF_MKT	PER	PBR	D_Eq	ROE	ROA	State	ROA_2005_6
1	CARBOCHIM	1.69	1.00	1.17	10.23	2.14	0.78	1.76	1.77	0.08	0.07	1	0.238
2	AEROSTAR	2.25	-0.30	10.23	11.54	2.92	3.05	2.27	3.23	0.15	0.10	1	0.142
3	PRODPLAST	2.08	0.73	-1.19	10.43	2.83	1.42	1.30	1.19	0.06	0.05	1	0.085
4	TURBOMECANICA	1.03	1.59	-2.10	10.90	3.30	0.80	1.09	1.57	-0.09	-0.03	1	0.074
5	SC TRANSILVANIA	1.76	1.68	-1.45	10.10	1.35	0.83	0.46	1.51	0.08	0.05	1	0.071
6	GR. IND.ELECONT.	1.73	3.38	-2.69	8.91	0.82	1.78	2.14	2.97	-0.02	-0.01	1	0.071
7	ARTROM	0.70	1.25	5.43	12.01	2.65	0.82	0.98	2.47	0.06	0.05	1	0.069
8	VRANCART	2.68	2.26	-0.15	10.97	2.75	2.67	2.80	1.65	0.06	0.05	1	0.063
9	COMELF	1.65	0.77	0.85	10.89	2.52	1.07	1.71	5.72	0.09	0.04	1	0.058
10	MECANICA	1.39	0.83	-3.78	10.37	2.27	1.25	1.69	2.49	0.02	0.03	1	0.037
11	VOESTALPINE VAE	4.00	2.12	3.35	11.14	1.18	1.51	3.80	1.64	0.07	0.07	0	0.036
12	COMPA	1.67	1.73	7.62	11.70	4.45	0.71	1.60	2.00	0.04	0.03	0	0.035
13	C_NIA ENGPETROL	1.40	1.01	1.55	8.24	0.61	0.82	1.39	2.12	0.02	0.01	0	0.029
14	ZIMTUB	2.27	-0.49	4.81	9.21	0.44	1.22	2.72	2.77	0.01	0.02	0	0.020
15	UAMT	3.01	0.48	9.04	9.76	1.61	1.75	3.18	7.67	0.01	0.01	0	0.005
16	TITAN	4.08	1.13	-4.96	11.80	1.17	1.28	1.65	1.92	0.03	0.03	0	-0.037
17	UCM	1.52	0.33	-8.49	9.72	0.95	0.28	0.61	0.57	-1.50	-0.22	0	-0.053
18	ARMATURA	1.53	0.81	-1.46	9.02	1.24	0.89	0.78	0.55	-0.52	-0.07	0	-0.072
19	MEFIN	2.49	0.52	-0.12	8.94	0.34	0.53	1.80	1.18	-0.04	-0.04	0	-0.110
20	ELECTROPUTERE	0.85	0.54	7.60	10.26	1.94	6.71	0.63	0.62	-0.59	-0.17	0	-0.212

Discriminant analysis has the same assumptions as linear regression analysis (e.g. normality, stationarity, etc.) but discriminant analysis is more robust to these assumptions. However, discriminant analysis is sensitive to outliers of the independent variables. (UCLA academic technology services, SPSS – Discriminant Analyses, chapter 6, http://www.cs.uu.nl/docs/vakken/arm/SPSS/spss6.pdf; Burns & Burns, 2008)

Jaba et al. (2007), Jaba and Robu (2009a, 2009b), Mironiuc et al. (2010) and Robu et al. (2012) use the SPSS statistical software, which has the advantage of holding an in-built discriminant analysis application. We follow their example and execute the following steps in SPSS. We start by specifying the state variable as a grouping variable and defined the lowest and highest value of the groups. In our case, we introduced only two values, 0 and 1. We then select our ten financial indicators (PRICE, BETA, MKT_CAP, etc.) as independent variables to be entered together in the analysis.

4. EMPIRICAL RESULTS

This section presents the discriminant analysis results for our sample of 20 companies traded on the BVB. We first start by looking at the univariate ANOVA statistics performed for each independent variable. Results are reported in Table 2.

Here, FF_MKT, ROE and ROA perform best and are significant at 10%. PRICE, BETA, and MKT_CAP indicators also seem to provide some differentiation, although not significant, between the two groups. We will consider these 6 variables further in the estimation of the score function. We do not consider the variables ALPHA, PER, PBR and D_Eq further since these do not provide a useful discrimination between companies.

Table 2. Univariate ANOVA statistics for each financial indicator

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
PRICE	.891	2.208	1	18	**.155**
BETA	.917	1.624	1	18	**.219**
ALPHA	.983	.313	1	18	.583
MKT_CAP	.902	1.947	1	18	**.180**
FF_MKT	.793	4.704	1	18	**.044**
PER	.998	.036	1	18	.851
PBV	.988	.224	1	18	.642
D_Eq	.989	.201	1	18	.660
ROE	.841	3.405	1	18	**.082**
ROA	.781	5.055	1	18	**.037**

Table 3 presents the summary of the canonical discriminant function. A higher eigenvalue for our discriminant function translates into a larger proportion of variance that is explained and, thus, into a stronger function of separating the companies into the two groups chosen. This is apparent from the canonical correlation which takes the value of 77.2%. The proportion of the explained variance in the state variable amounts to $(77.2\%)^2$ or approximately 60%. The proportion of the total variance not explained is expressed by the Wilks'Lambda coefficient. The Wilks'Lambda coefficient is significant value at 5% in our case, indicating that the two groups, “good” and “bad”, seem to differentiate quite well.

Using the function coefficients, we first determine the threshold score value that will be used for the new classification of companies in the “good” and “bad” groups. The threshold is simply estimated by taking the average values of the financial indicators across companies:

$$Z = -1.164 - 1.339 * PRICE + 0.579 * BETA + 0.300 * MKT\_CAP - 0.102 * FF\_MKT -$$
$$- \quad 2.202 * ROE + 24.445 * ROA$$

$$\bar{\mathbf{Z}} = -1.164 - 1.339 * 1.9894 - 0.579 * 1.0691 + 0.3 * 10.3067 - 0.102 * 1.8735 -$$
$$- \quad 2.202 * -0.983 + 24.445 * 0.0036 = \mathbf{-0.00443}$$

Table 3. Summary of Canonical Discriminant Functions

Unstandardized Function Coefficients

	Function
	1
PRICE	-1.339
BETA	.579
MKT_CAP	.300
FF_MKT	-.102
ROE	-2.202
ROA	24.445
(Constant)	-1.164

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	1.472[a]	100.0	100.0	.772

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.405	13.575	6	.035

The average score amounts to – 0.004 and separates firms with higher individual score as "good" and those with lower individual score as "bad". The distributions of the discriminant function scores are illustrated in Figure 2. As is apparent from the graphs, the discriminant function separated the two groups well. The "good" companies were all correctly identified and only 8 of the 10 "bad" companies were correctly identified since two companies are to the right of the - 0.004 threshold. In large, we can deduce that the ROA variable provides a good characterization of company performance.

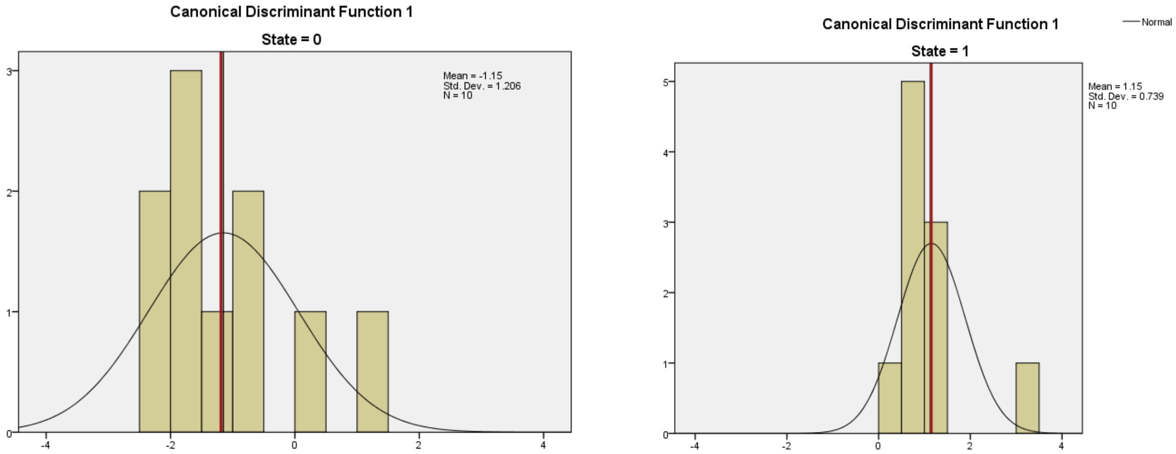

Figure 2. Distribution of "bad" and "good" companies

6. CONCLUSIONS:

In conclusion, the ROA indicator was validated by discriminant analysis for the classification of companies into "good" and "bad". We find that 90% of companies (18 of 20) are correctly classified. The variable ROA has the highest value of the test statistic F and the lowest Wilks'Lambda coefficient.

Two companies, COMPA and ENGPETROL, were first classified as "bad". However, these companies achieve a score of 1.29 and 0.25, respectively. Since both scores are higher than the threshold, these companies are included in the "good" group indicating that their financial position has improved over time.

Reclassification of 20 companies

New order	Firm	Old order	Previous state	Current state	Scor
1	ARTROM	**7**	1	1	3.04
2	CARBOCHIM	**1**	1	1	1.47
3	SC TRANSILVANIA	**5**	1	1	1.34
4	COMPA	**12**	0	1**	1.29
5	MECANICA	**10**	1	1	1.11
6	COMELF	**9**	1	1	0.97
7	AEROSTAR	**2**	1	1	0.97
8	GR. IND.ELECONT.	**6**	1	1	0.79
9	TURBOMECANICA	**4**	1	1	0.73
10	VRANCART	**8**	1	1	0.58
11	PRODPLAST	**3**	0	1	0.45
12	C_NIA ENGPETROL	**13**	0	1**	0.25
13	VOESTALPINE VAE	**11**	0	0	-0.62
14	ARMATURA	**18**	0	0	-0.73
15	ZIMTUB	**14**	0	0	-1.31
16	TITAN	**16**	0	0	-1.80
17	UAMT	**15**	0	0	-1.92
18	ELECTROPUTERE	**20**	0	0	-1.92
19	UCM	**17**	0	0	-2.37
20	MEFIN	**19**	0	0	-2.41

The five firms with the lowest score can potentially be considered as having a high probability of being insolvent or bankrupt. At the end of 2013, the financial health of four of the five companies was still in difficulty. One of these is insolvent and the others have low or even negative shareholder's equity.

However, one should not inflate the informative power if the score function. A company can be regarded as an economic and social system that operates in a complex environment with many more variables to determine its health or weakness. The score is a simple tool for "early detection" of default risk and investment opportunities, but this information should be used with caution. In order to reach a proper decision regarding the financial soundness of a company, the analysis should be supplemented by observing the evolution of the score over several years for the company in comparison with its peers and by traditional methods of financial analysis.

REFERENCES:

Afifi A., Clark V. (2003), SPSS Textbook Examples Computer Aided Multivariate Analysis, 3rd Edition, chapter 11, SPSS - Discriminant Analyses, *UCLA academic technology services,* http://www.cs.uu.nl/docs/vakken/arm/SPSS/spss6.pdf

Armeanu, Ş. D. et al. (2012), Utilizarea tehnicilor de analiză cantitativă a datelor pentru estimarea riscului de faliment al corporațiilor, *Economie teoretică şi aplicată*, Volumul XIX (2012), No. 1(566), pp. 86-102 http://store.ectap.ro/articole/681_ro.pdf

Burns R., Burns R. (2008)**,** Business Research Methods and Statistics using SPSS**,** chapter 25, Discriminant Analysis, *SAGE Publications Ltd.*, California, USA
http://www.uk.sagepub.com/burns/website%20material/Chapter%2025%20-%20Discriminant%20Analysis.pdf

Jaba E. et al. (2007), Discriminant analysis in the study of romanian regional economic development, *Analele ştiințifice ale universității „Alexandru Ioan Cuza" din Iaşi, Tomul LIV, Ştiințe Economice*, http://anale.feaa.uaic.ro/anale/en/Arhiva%202007%20-%20Jaba/184

Jaba E., Robu I.-B. (2009), Utilizarea analizei discriminant pentru obtinerea probelor de audit (I şi II), *Revista de audit financiar nr.11 şi 12 (2009),* https://www.researchgate.net/publication/235996415_Utilizarea_analizei_discriminant_pentru_obtinerea_probelor_de_audit_(I) şi https://www.researchgate.net/publication/235996402_Utilizarea_analizei_discriminant_pentru_obtinerea_probelor_de_audit_(II)

Mironiuc M. et al. (2010), The Discriminant Analysis: an Exploratory Study Concerning the Degree of Financial Autonomy of Companies in the Context of the Romanian Business Environment , *Studies and Scientific. Economic Edition, nr 15, 2010*, Bacău, http://sceco.ub.ro/index.php/SCECO/article/viewFile/99/99

Robu M. A., Mironiuc M., Robu I.-B. (2012), Un model practic pentru testarea ipotezei de "going-concern" în cadrul misiunii de audit financiar pentru firmele româneşti cotate, *Revista de audit financiar nr. 2 (2012),* http://www.cafr.ro/uploads/AF%202%202012%20-%20Site-2faa.pdf