Discriminant Analysys of Default Risk

Aragon, Aker

CARIFIN

21 October 2004

DISCRIMINANT ANALYSIS ON DEFAULT RISK

By Aker Aragón Paz. October, 2004.

The present work intends to propose a way to get prediction functions on the defaults of companies, based on discriminant scores. The use of the Multivariate Discriminant Analysis (MDA), applied to the quantification of the bankruptcy and default risk, has been replaced in the last few years by other techniques such as the logistical regression, because of the necessary normality and homoskedasticity required when the MDA is applied.

However, the objective of this work is to show and suggest a way to determine reliable equations based on MDA, provided that the attainment of such equations is previously supported by non-parametric techniques in the process of variables selection, and by box-cox transformations in order to get the normality of the indicators. The use of Principal Components Analysis is also proposed, in order to avoid the multicollinearity or interrelation of the explicative variables, generally present in the financial ratios and often not taken into account.

In summary, the application of these techniques shows a very reliable way to get probabilities of default of the companies, based on ratios and other financial indicators.

1. INTRODUCTION

The assessment of liquidity risk, or the likely insufficient available funds to face debts taken, turns out to be an essential issue for any company. At Loan Institutions or companies which base their revenue on credit sales, the liquidity risk depends mainly on the credit risk. Though enough assets available in order to face the liabilities are high, the actual Available funds of such payment rights shall depend upon their customers´ payment ability, that is, **the liquidity risk of a financing entity shall depend, to a great extent, on the credit risk and the latter on the liquidity risk of its portfolio of customers.**

Due to the above mentioned facts, we can see the need for the money-lenders, to **perform a analysis, as accurate and quantified as possible, on the liquidity risk or default risk of their customers.** This is the main objective of the present work.

Traditionally, the analysis of the financial conditions of the companies has focused on the analysis of accounting statements, studying the financial ratios by means of univariant techniques. The usual procedure is to make a separate analysis of financial ratios of the company and then making a global qualitative assessment of the company.

The use of the techniques of multivariate analysis becomes especially important at the analysis of default risk of the companies. This information is mostly numerical, therefore the cognition and application of the statistical techniques allow analyzing the behavior of the variables simultaneously, assessing its complete effect on the subject-matter studied.

It is important to stress that the main limitation faced when making this work, was related to the available sample which only comprised the 52 present customers from the non-Banking Financial Institution “TRANSFIN”. Thus, such an important limitation must be taken into account before making any generalization of the results attained.

Despite the aforementioned, the purpose of the research is proposing the methodology described in this work, including the statistical tools applied, which can be generalized for studies with a wider sample of companies. Out of the 52 companies analyzed, 22 belong in the group of default, which was defined as companies with over-90-day delay in payments. The information used as a basis for the calculation of the variables was the Profit and Loss Statement, Balance Sheet, state of receivables and payables ranked by time, as well as internal records on the historical fulfillment of payments with the Financing Institution.

2. SELECTION OF VARIABLES. BACKGROUND

The use of the discriminant analysis as a way to quantify the default risk, started to be replaced in the 80´s by some techniques such as the logistical regression, mainly due to the quality loss by using models based on financial variables, which frequently fail to fulfill the requirements of normality and constant variance[1]. Despite this marked trend to discard the multivariate discriminant analysis, it was decided to take this technique again in the present work. Such technique was proposed by E. I. Altman for predicting the corporate bankruptcy; taking into account the greatest possible normality of the variables in the application phase, and performing non-parametric tests in the phases of selection of original variables.

Not including the significant variables in the multivariate analysis greatly influences its results. This phase was based on choosing indicators that would allow **explaining the possibility of short-term payment.**

One of the most important studies made on the bankruptcy risk and default risk was by Edward I. Altman in 1968. According to some materials consulted, it was the first research proposing the multivariate discriminant analysis, in order to determinate a predictive function of corporate bankruptcy. It must be considered that outstanding studies on the company failure were previously made, with a single-variant vision (Fitzpatrick, 1932 and Beaver, 1966); the latter was supported by the single-variant discriminant analysis.

In E. Altman's[2] research, as well as in further studies made by this author, the selection of variables was taken from a previous selection. In the study which was a second improved version of the first model[3], the variables were chosen from the following list:
LIST OF VARIABLES USED BY ALTMAN, HALDEMAN AND NARAYANAN TO SELECT THE VARIABLES INCLUDED IN THE DISCRIMINANT FUNCTION Z (THE Z CREDIT RISK MODEL, 1977).

Variable		Population Means		Univariate
No.	Name	Failed	Non-Failed	F-Test
(1)	EBIT/TA	-0.0055	0.1117	54.3
(2)	NATC/TC	-0.0297	0.0742	36.6
(3)	Sales/TA	1.3120	1.6200	3.3
(4)	Sales/TC	2.1070	2.1600	0.0
(5)	EBIT/Sales	0.0020	0.0070	30.2
(6)	NATC/Sales	-0.0153	0.0400	33.1
(7)	Log tang. Assets	1.9854	2.2220	5.5
(8)	Interest coverage	-0.5995	5.3410	26.1
(9)	Log no. (8)	0.9625	1.1620	26.1
(10)	Fixed charge coverage	0.2992	2.1839	15.7
(11)	Earnings/debt	-0.0792	0.1806	32.8
(12)	Earnings 5 yr. Maturities	-0.1491	0.6976	8.8
(13)	Cash/flow fixed charges	0.1513	2.9512	20.9
(14)	Cash flow/TD	-0.0173	0.3136	31.4
(15)	WC/LTD	0.3532	2.4433	6.0
(16)	Current ratio	1.5757	2.6040	38.2
(17)	WC/total assets	0.1498	0.3086	40.6
(18)	WC/cash expenses	0.1640	0.2467	5.2
(19)	Ret.earn/total assets	-0.0006	0.2935	114.6
(20)	Book equity/TC	0.2020	0.5260	64.5
(21)	MV equity/TC	0.3423	0.6022	32.1
(22)	5yr.MV equity/TC	0.4063	0.6210	31.0
(23)	MV equity/total liabilities	0.6113	1.8449	11.6
(24)	Standard error of estimate of EBIT/TA (norm)	1.6870	5.784	33.8
(25)	EBIT drop	-3.2272	3.179	9.9
(26)	Margin drop	-0.2173	0.179	15.6
(27)	Capital lease/assets	0.2514	0.178	4.2
(28)	Sales/fixed assets	3.1723	4.179	3.5

Notation:
EBIT = earnings before interest and taxes
NATC = net available for total capital
TA = total tangible assets
LTD = long term debt
MV = market value of equity
TC = total capital
TD = total debt
WC = working capital
CF = cash flow (before interest #13,after interest #14)

---

[1] ***The Multivariate Discriminant Analysis (MDA) implies obtaining a linear combination of several, independent variables, discriminating between previously defined groups. The procedure consists in finding the coefficients linked to the independent variables, maximizing the differences between the groups of classification, and minimizing the differences inside each of these groups (Maximum value of the quotient Between Groups Variability / Within Groups Variability). Differently from other techniques such the logistical regression, the MDA supports itself from two issues: multivariate normality, necessary due to significant tests used in the process to estimate the discriminant function; and the other one, equal covariance and dispersion matrices for all the groups, required for the classification process.***

[2] ***Altman, E., "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," Journal of Finance, New York, USA. September 1968.***

[3] ***Altman, E., R. Haldeman, and P. Narayanan, "ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations," Journal of Banking and Finance, New York, USA. June 1977.***

As shown, the election of variables was supported by the univariant discriminant power measured by a variance analysis of a two-level factor: bankruptcy and non-bankruptcy. Although this variance analysis was not definitive for the election of variables, it was indeed used as a further criterion, supporting the selection process.

Another outstanding research consulted, was that of Moody's Investors Service[4], which published in May 2000 a deep research comprising the greatest sampling chosen so far in this kind of works: 1,621 corporations that ran into default and 23,089 which did not. For this study, the authors started from a first high number of variables, analyzing for each of them its univariant discriminant power through logistical regressions. The variables chosen in this research were as follows:

VARIABLES INCLUDED IN THE MODEL FOR PREDICTION OF DEFAULT OF MOODY´S INVESTORS SERVICE (Moody's Default Model, 2000).

Ratios	Calculation
Assets	Assets / CPI of the year
Inventory/COGS	Inventory/COGS
Liabilities/Assets	Liabilities /Assets
Net Income/Assets	Net Income /Assets
Net Income Growth	(current Net Income/ Assets)-(prior Net Income/Assets)
Quick Ratio [= (curr. assets - invent)/curr. Liab.]	(Current Assets - Inventory)/Current Liabilities
Retained Earnings/Assets	Retained Earnings/Assets
Debt Service Coverage Ratio	EBIT/interest
Cash/Assets	Cash & Equivalents/Assets
Sales Growth	(current Sales/prior Sales)-1

In the researches consulted, the selection of variables is performed from a considerable number of financial ratios, and many times variance analysis tests are performed so as to support the selection process. For the present work, a high number of variables used in previous studies in addition to those already commented was compiled. But a special care was taken in the analysis, as it is an unusual subject-matter in Cuba. For this, variables considered as most suitable for the Cuban conditions were added, and others which evidently made no sense in this environment were excluded. This way, a previous selection of 52 indicators was attained. Different from many of the researches previously made, in this work the application of Mann Whitney's[5] statistical technique was considered as univariant analysis supporting the process of variable selection, aimed at determining the capacity of every variable to distinguish between the two groups of corporations (Payment and Default Groups), by means of a non-parametric statistical test.

This way, the risks implying the use of parametric methods, like the variance analysis (ANOVA) without verifying the issues of normality are avoided. The variance analysis may lead to totally mistaken judgments when dealing with financial ratios, which in general, are not normally distributed.

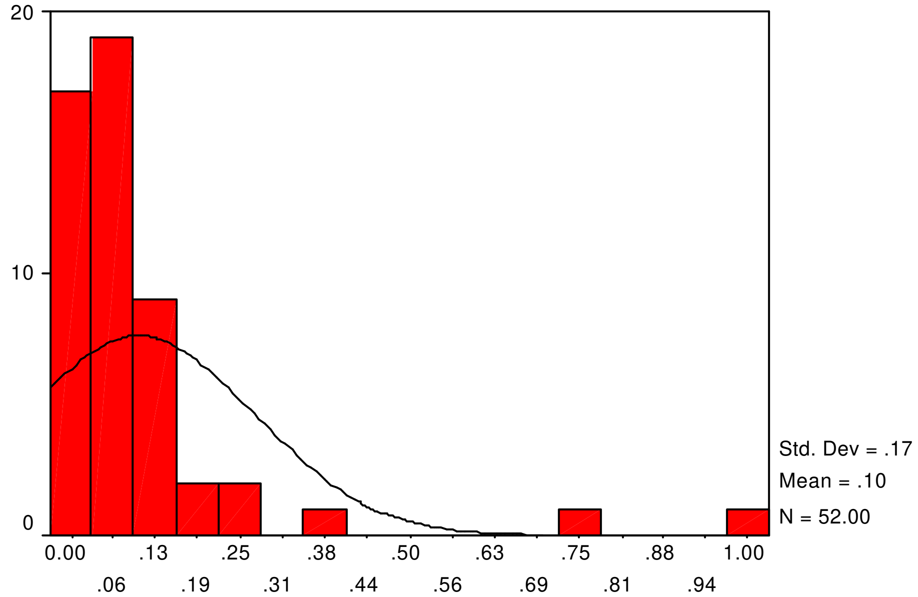
[4] Moody's, E. Falkensten, A. Boral and L. Carty, "RiskCalc ™ Private Model: Moody´s Default Model for Private Firms", Global Credit Research, New York, USA. May 2000.

[5] The goal of Mann Whitney's test is proving the existence or non-existence of significant differences between the median of two independent populations. For this, it is based on a statistical test calculated on the sampling values taken to an ordinal scale, excluding the effect of asymmetry in the "Non Normal" distribution.

When applying the Kolmogorov-Smirnov test, only 12 out of the 52 previously selected variables have a normal distribution, if an error probability Type I of 5% is taken for the decision criterion. However, with a more conservative decision criterion (a = 10%), only 10 out of the 52 variables have a normal distribution.

Such “non normal” feature, is usual in many financial variables like the Cash & Equivalents / Current Liabilities ratio which shows the cash available funds to face short-term debts. It is quite logical that most corporations try to keep the lowest quick assets, as the proper theory on working with “naught idle liquidity” is generally followed in finance. Notwithstanding, cash & equivalents must not be negative, so the central trend shall have a value next to zero, having some cases with very positive, extreme values. This way, the variable shall have a negative asymmetry, with an O minimum value and extreme values quite to the right-hand side, being the median a measurement of central trend more suitable than the average of the observations.

Histogram of the variable (Cash and Equivalents/Current liabilities)



When applying a variance analysis to the variable Cash and Equivalents / Current Liabilities, the result shows that there are no significant differences between the two groups:
Ho: $\mu_1=\mu_2$

Variance Analysis between two groups for the variable Cash and Equivalents / Current Liabilities

ANOVA	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.742E-02	1	3.742E-02	1.299	.260
Within Groups	1.441	50	2.881E-02		
Total	1.478	51			

That is to say, this univariant analysis of variance, points out that there are not perceptible differences between the mean of the groups of corporations which pay and those which do not. (Probability of error over .260 very high if the hypothesis on equality of mean is rejected). This could suggest the exclusion of this variable from the study.

However, if a test of non-parametric hypothesis is made instead of an ANOVA, to verify if there are differences between the median of both groups, we have got:

Mann-Whitney Test
H_o: $Me_1 = Me_2$

Mann-Whitney Test for the variable Cash and Equivalents / Current Liabilities

Ranks

	DEFAULT	N	Mean Rank	Sum of Ranks
Cash and Equivalents / Current Liabilities	0	30	31.95	958.50
	1	22	19.07	419.50
	Total	52		

Test Statistics[a]

	Cash and Equivalents / Current Liabilities
Mann-Whitney U	166.50
Wilcoxon W	419.500
Z	-3.028
Asymp. Sig. (2-tailed)	.002

[a] Grouping Variable: DEFAULT

Once the Mann-Whitney Test is applied, significant differences are noted between both groups for the variable Cash and Equivalents / Current Liabilities (small Probability of error of 0.002 if the null hypothesis is rejected), which corroborates the inadequacy of using parametric tests when normality is absent.

One variant which would allow applying parametric tests could be the transformation of the 42 ratios which are not normally distributed. Obviously, this would hinder the selection process of variables, above all, if taken into account that such transformations are only done in this phase so as to apply univariant tests, serving as a support for the election of the future variables to be included in the study, among the wide group of previously selected variables. Therefore, it is quite suitable to apply the non-parametric Mann-Whitney test on the previously selected variables, instead of other parametric techniques usually used over this selection phase.

The results of applying this test, making inferences on the mean of two independent groups, are shown below. The variables in red have got Type-I error probability critical values (Over 10 %). These results, and some theoretical criteria and knowledge on the variables which are supposed to predict the customers´ default, were taken into account.

MANN WHITNEY TEST AND VARIABLES SELECTED FOR THE MULTUVARIATE ANALYSIS

Concept	*No*	*Variable*	*Mann-Whitney U*	*p-value (2-tailed)*	*Selected Variables*
Liquidity	*1*	*(Cash and Equivalents / Current liabilities)*	*169.5*	*0.00246*	*x*
Liquidity	*2*	*Quick ratio ((Cash & Equivalents + Accounts & Bills Receivable) / Current liabilities)*	*214.0*	*0.01866*	
Liquidity	*3*	*Quick ratio 2 ((Cash & Equivalents + Accounts & Bills Receivable) / Current liabilities – Accounts Receivable over 90 days) / Current liabilities)*	*166.0*	*0.00154*	*x*
Liquidity	*4*	*Current assets / Total assets*	*275.0*	*0.20125*	
Liquidity	*5*	*Current ratio (Current assets / Current liabilities)*	*177.0*	*0.00198*	*x*
Liquidity	*6*	*Working capital / Total assets*	*168.0*	*0.00111*	
Liquidez	*7*	*Short-term liquidity((Cash & Equivalents + Accounts & Bills Receivable) / Short-term debts)*	*202.0*	*0.01059*	
Receivable & Payable	*8*	*Accounts & Bills receivable / Accounts & Bills payable*	*254.0*	*0.08168*	
Receivable & Payable	*9*	*Accounts payable over 90 days / Accounts payable*	*186.0*	*0.00098*	*x*
Receivable & Payable	*10*	*Accounts receivable over 90 days / Accounts receivable*	*217.0*	*0.02255*	
Receivable & Payable	*11*	*Accounts receivable over 90 days / Current assets*	*244.0*	*0.06108*	
Receivable & Payable	*12*	*Accounts payable over 90 days / Current liabilities*	*244.0*	*0.04350*	

Solvency	*13*	*Long-term solvency (Permanent resources / Fixed assets)*	*254.0*	*0.07538*	
Leverage	*14*	*Leverage (Total liabilities / Total assets)*	*162.0*	*0.00075*	*x*
Leverage	*15*	*Debt Quality (Current liabilities / Total liabilities)*	*336.0*	*0.63011*	
Activity	*16*	*Sales / Total assets*	*237.0*	*0.04750*	
Activity	*17*	*Sales / Current Assets*	*248.0*	*0.06671*	
Activity	*18*	*Sales / Working capital*	*272.0*	*0.13841*	
Activity	*19*	*Receivable days (((Accounts & Bills receivable)*360 days) / Sales)*	*246.0*	*0.06671*	*x*
Activity	*20*	*Payment days (((Accounts & Bills payable)*360days)/Cost of sales)*	*190.0*	*0.00487*	*x*
Activity	*21*	*Receivable days / Payment days*	*283.0*	*0.21462*	
Profitability	*22*	*Sales margin (Net Profit / Sales)*	*240.0*	*0.03979*	
Profitability	*23*	*Earnings before taxes Total Assets)*	*187.0*	*0.00239*	*x*
Profitability	*24*	*Net Profit / Total Assets)*	*193.0*	*0.00386*	
Profitability	*25*	*Fundamental Activity Economic Profitability (Operating Profits/ Total Assets)*	*206.0*	*0.00724*	
Profitability	*26*	*Operating Profit / Total Assets*	*222.0*	*0.01688*	
Fulfilment	*27*	*Fulfilment of payments*	*35.0*	*0.00000*	*x*
Others	*28*	*Cash & Equivalents / Net Sales*	*317.5*	*0.66337*	
Others	*29*	*Fixed assets / liabilities*	*237.0*	*0.04749*	
Others	*30*	*Liabilities / Net Sales*	*144.0*	*0.00021*	*x*
Others	*31*	*(Cash & Equivalents + Accounts & Bills Receivable) / Net Sales*	*245.0*	*0.04961*	
Size	*32*	*Total assets*	*230.0*	*0.05643*	
Cash Flow	*33*	*Net Profit – Growth of Accounts & Bills Receivable + Growth of Accounts & Bills Payable / Current liabilities*	*352.0*	*0.95569*	
Cash Flow	*34*	*Net Profit – Growth of Accounts & Bills Receivable + Growth of Accounts & Bills Payable / Total liabilities*	*346.0*	*0.95569*	
Cash Flow	*35*	*Net Profit – Growth of Accounts & Bills Receivable + Growth of Accounts & Bills Payable / Total Assets*	*347.0*	*0.94094*	
Cash Flow	*36*	*Net Profit – Growth of Accounts & Bills Receivable + Growth of Accounts & Bills Payable / Sales*	*324.0*	*0.61701*	
Cash Flow	*37*	*(Net Profit – Growth of Accounts & Bills Receivable) / Current liabilities)*	*227.0*	*0.02624*	
Cash Flow	*38*	*(Net Profit – Growth of Accounts & Bills Receivable) / Total liabilities)*	*233.0*	*0.03635*	
Cash Flow	*39*	*(Net Profit – Growth of Accounts & Bills Receivable) / Total Assets)*	*226.0*	*0.02624*	*x*
Cash Flow	*40*	*(Net Profit – Growth of Accounts & Bills Receivable) Sales)*	*250.0*	*0.07240*	
Cash Flow	*41*	*(Net Profit – Growth of Accounts & Bills Receivable) / Current Liabilities*	*135.0*	*0.00016*	
Trend	*42*	*Sales Growth*	*326.0*	*0.73884*	
Trend	*43*	*Growth of Accounts Payable over 90 days*	*172.0*	*0.00046*	
Trend	*44*	*Growth of Accounts Receivable over 90 days*	*333.0*	*0.75173*	
Trend	*45*	*Growth of Accounts & Bills Payable*	*246.0*	*0.07848*	
Trend	*46*	*Growth of Accounts & Bills Receivable*	*318.0*	*0.50491*	
Trend	*47*	*Growth of Accounts payable over 90 days / Accounts Payable*	*201.0*	*0.00223*	
Trend	*48*	*Growth of Accounts Receivable over 90 days / Accounts Receivable*	*323.0*	*0.62850*	
Repayment	*49*	*Maximum monthly amount to refund / Cash & Equivalents*	*252.0*	*0.13351*	
Repayment	*50*	*Maximum monthly amount to refund / Current Assets*	*320.0*	*0.76696*	
Repayment	*51*	*Maximum monthly amount to refund / (Current Assets – Accounts Receivable over 90 days)*	*290.0*	*0.41509*	
Repayment	*52*	*(Sales – Growth of Accounts & Bills Receivable) / Maximum monthly amount to refund*	*192.0*	*0.01003*	

Variables not making significant differences between the so-called "Payment" and "Default" groups of corporations for a 10% significant level.

It is important to point out that the shortness of available information (Sample size of only 52 observations), avoids the use of more variables than those chosen. In that case a sample of at least 100 corporations would be necessary. Despite the theoretical validity for the selection of a bigger group of indicators, the fact of considering so many variables for a relatively small sample, would lead to create an ·"over-adjusted" prediction model, which would just show the specific reality of a chosen sample. That is, the model obtained could not characterize a new case properly (A new customer or a customer included in the sample in a new period of time).

The selection of 11 variables was based on the fact that it is the maximum number, not exceeding the recommended ratio of about 5 observations for every independent variable[6].

Within the 11 variables selected, those measuring liquidity are more important. However, it is important to underline that a great part of the indicators chosen, takes directly or indirectly into account, the behaviour of the receivables and payables and their ages. Such aspects are not generally included in the calculation of default predicting variables.

Besides that, the ratios taking the Fixed Assets into account were not included, because of special problems faced by many Cuban entities on their valuation. Also, the trend indicators showed, in general, a smaller predictive capacity. Such aspect agrees with the results attained by previous studies on the default risk[77].

3. QUANTIFICATION OF THE DEFAULT RISK. DETERMINATION OF THE DISCRIMINANT FUNCTION

3.1. DESIGN

In order to attain the final objective of the present research (Finding a mathematical function allowing to predict a customer's future default, and quantifying the probability of its happening), the fulfilment of the facts of normality, homoskedasticity and non-multicollinearity is previously required. All these restrictions would be broken if the discriminant analysis were made, taking directly the 11 variables chosen, because of the great correlation among many of these variables and the non-normal distribution of most of the financial indicators.

Due to the lack of normality and to different covariance matrices, many times the statistical signification of the results is little reliable, when a Discriminant Analysis is applied. Also the multicollinearity, given by the high relationship of the variables, is especially critical in the stage-process of the discriminant analysis, since a variable can be completely excluded, if an indicator highly relate to that variable is chosen on a previous step. So, the measurement of the actual contribution from each variable to the predictive capacity of the discriminant function is difficult.

In short, breaking these three requirements (Normality, Homoskedasticity and Non Multicollinearity), along with an improper number of predicting variables, rouses a high predisposition to the distortion of the results from the discriminant analysis, and many times that leads to getting seemingly significant discriminant functions, but with a high bias actually.

Therefore, the first step to take, must be attaining the greatest possible normality of the 11 first variants, and then through an Principal Components Analysis (PCA) from a smaller number of indicators, creating new correlated synthetic variables, which follow in the highest degree a normal distribution, and fulfil the fact of homoskedasticity. From these new synthetic indicators obtained through the PCA, then the Discriminant Analysis can be performed. By means of such analysis a reliable discriminant function shall be attained. This will classify a corporation as to its future possibility of payment.

3.2. ANALYSIS ON THE "NORMALIZATION" OF THE VARIABLES

In order that the greatest possible number of variables have a normal distribution, the transformation is recommended by means of the Box-Cox method[8]. This procedure allows

***

[6] According to Hair, J. F. Jr.; Anderson, R. E.; Tatham, R. L.; Black, W.C. "Análisis Multivariante"(Multivariate Analysis), Fifth Edition. Prentice Hall Iberia; Madrid, Spain, 1999; less than five observations for each independent variable are not recommendable.

[7] Fundamentally in "RiskCalc™ Private Model: Moody's Default Model for Private Firms", the least predictive capacity of the trend indicators is shown in section "Growth vs. Levels".

[8] One of the Box-Cox transformation families used the most is $(X+C)^p$. The transformation is focused then on determining the p constant, which can be calculated by iterative processes so as to meet some optimality criterion, consisting in some cases, in maximizing the correlation coefficient between the distribution of the variable transformed and Normal theoretical distribution.

making transformations type $(X+C)^{\wedge p}$, in which C is a constant that makes the independent variable X positive and the p value is determined from an iterative procedure which leads to maximize the normality of the indicators; for p=0, the Box-Cox transformation becomes logarithmic. The box-cox coefficients for each of the 11 variables were attained through the statistical pack MINITAB.

From the results of the Box-Cox analysis, the value obtained for every variable, indicates the p value that must be chosen to make the transformation, always taking into account that, theoretically speaking it is comprehensible and logical. For instance, in the case of Current Ratio, the value calculated by the Box-Cox procedure was 0.113. But determining $(X+C)^{\wedge 0.113}$ as a more suitable transformation lacked a practical sense. So p=0 was selected. It is within the 95% confidence interval, and indicates a natural logarithmic transformation for the original variable: LOG (X+C).

The table below show the results of applying the Kolmogorov Smirnov test to each of the 11 original variables, as well as the value selected to perform the transformation of every indicator.

Kolmogorov Smirnov Test and p value used to transform the variables

	VARIABLES	*N*	*Kolmogorov-Smirnov Z*	*p-value*	*Distribution*	***P Value** in the transformation[a)]*
X1	*Cash & Equivalents / Current liabilities*	*52*	*2.00032*	*0.00067*	*Non Normal*	*0.00*
X2	*Quick ratio ((Cash & Equivalents + Accounts & Bills Receivable – Accounts Receivable over 90 days) / Current liabilities)*	*52*	*1.96837*	*0.00086*	*Non Normal*	*0.25*
X3	*Current Ratio (Current assets / Current liabilities)*	*52*	*1.95939*	*0.00088*	*Non Normal*	*0.00*
X4	*Accounts payable over 90 days / Accounts payable*	*52*	*0.71946*	*0.67864*	*Normal*	*1.00*
X5	*Leverage (Total liabilities / Total assets)*	*52*	*0.88253*	*0.41730*	*Normal*	*1.00*
X6	*Receivable days (((Accounts & Bills receivable)*360 days) / Sales)*	*52*	*1.79674*	*0.00314*	*Non Normal*	*0.00*
X7	*Payment days (((Accounts & Bills payable)*360days)/Cost of Sales)*	*52*	*1.84431*	*0.00222*	*Non Normal*	*0.00*
X8	*Earnings before taxes / Total Assets*	*52*	*1.56404*	*0.01501*	*Non Normal*	*0.50*
X9	*Fulfilment of payments*	*52*	*2.31487*	*0.00004*	*Non Normal*	*1.00*
X10	*Liabilities / Net Sales*	*52*	*2.01543*	*0.00059*	*Non Normal*	*0.00*
X11	*(Net Profit – Growth of Accounts & Bills Receivable) / Total Assets*	*52*	*0.86774*	*0.43877*	*Normal*	*1.00*

a) ***The p value is used to make the type $(X+C)^{\wedge p}$ transformation. For p=o, the natural logarithmic transformation: LOG(X+C) is performed; for p=10, the exponential transformation : $e^{\wedge X}$ is made; for p=1, no transformation is made at all.***

After performing the corresponding transformations to every variable, the Kolmogorov-Smirnov test was applied again, and its result was that out of the 11 variables transformed, only one does not follow a normal distribution. That is the case of the “Fulfilment” indicator, because of the maximum and minimum value agreeing with the most probable values from the variable. That is why its opposed behaviour to a normal distribution. However, including this indicator is important because of its high univariant predictive power.

3.3 FINAL PREPARATION FOR THE DISCRIMINANT ANALYSIS: PRINCIPAL COMPONENTS

After improving the normality of the set of independent variables, the Principal Components Analysis (PCA) becomes obvious as a previous step to the discriminant analysis. The justification for its use, is firstly due to the fact that the multicollinearity does not hinder the application of this technique, but quite the contrary: the presence of correlation between the variables is necessary, for the further attainment of non-correlated synthetic variables (Factors), which contain, in turn, the greatest possible part of the explicative power from the initially chosen, independent variables. Another factor for the application of PCA, is the fact that the presence of normality in this method is not so decisive as in the Discriminant Analysis, because as a rule, like in our case, a statistical significance test for the factor coefficients is not used.

The validity or non-validity of applying the PCA, was analyzed through the Bartlett's Test of Sphericity, and Kaiser-Meyer-Olkin Measure of sampling adequacy:

Index of Sampling Adequacy (KMO) and Barlett's Test of Sphericity

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.705
Bartlett's Test of Sphericity	Approx. Chi-Square	325.039
	df	55
	Sig.	.000

The KMO measure of sampling adequacy, shows a 0.7 value. Although it is not optimal, it is higher than the acceptable minimum value (0.5). On its side, the Barlett' s Test of Sphericity, clearly shows that the variables are correlated (p value=0.000 quite lower than the 5% significance level).

On the other hand, it is important to verify the validity of including in the PCA the 11 variables selected, for which their communalities are analyzed.

Contribution from the variables chosen to the model of PCA determined.

Communalities		
Variables (Orden Descendente)		***Extraction***
Receivable days		*0.916532329*
Earnings before taxes / Total Assets		*0.897204641*
Quick ratio ((Cash & Equivalents + Accounts & Bills Receivable – Accounts Receivable over 90 days) / Current liabilities)		*0.888306384*
(Net Profit – Growth of Accounts & Bills Receivable) / Total Assets		*0.867347829*
Current Ratio		*0.8457532*
Payable days		*0.835913161*
Accounts payable over 90 days / Accounts Payable		*0.822367807*
Liabilities / Net Sales		*0.803917894*
Fulfilment		*0.69975296*
Leverage (Total liabilities / Total Assets)		*0.618169367*
Cash & Equivalents / Current liabilities		*0.607404121*

All the variables have got communalities justifying their inclusion in the model (Variables lacking enough explanation are those with values lower than 0.5). This indicates that the variance rate contributed by every indicator to the final solution is significant. After the application of the PCA was regarded as valid, the Factoring was carried out, aimed at forming the new synthetic variables or components, attained by diagonalizing the Correlation matrix. These components obtained by this method, are 11 suitable orthogonal vectors, associated to the suitable values (j) correlated among themselves:

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.741	43.102	43.102	4.741	43.102	43.102
2	1.560	14.185	57.287	1.560	14.185	57.287
3	1.395	12.684	69.971	1.395	12.684	69.971
4	1.106	10.053	80.024	1.106	10.053	80.024
5	.616	5.596	85.620			
6	.527	4.793	90.414			
7	.349	3.175	93.588			
8	.255	2.320	95.908			
9	.193	1.752	97.660			
10	.174	1.586	99.245			
11	8.300E-02	.755	100.000			

Extraction Method: Principal Component Analysis.

The number of factors was determined by means of a combination of several criteria, rendering greater importance to the ones below:
Criterion on variance percentage: considering that the factors extracted contain at least 75% from the overall variance. In this case, the first four factors mange to contain 80.02% of the variance.
Criterion on latent root: considering the factors which have latent roots or auto-values higher than 1. The last factor extracted has got a suitable value higher than 1 (1.106).
A priori Criterion: It is the most "reasonable" criterion, because it is based on the fact that the factors attained, agree with what the logic indicates as to some theoretical justification. It is related to the practical interpretation on the independent variables after making the orthogonal rotation.

At the rotation phase, the use of the Varimax method is proposed, for this criterion focuses on simplifying the columns of the factor matrix. And it is most recommended when the aim of the Principal Components Analysis is reducing the number of variables for using the new non-correlated synthetic indicators, in a discriminant or multiple regression analysis. Through the Varimax rotation, keeping the total variance explained, the factorial charges of the variables are polarized into the four components. Thus, the following polarization of the variables in each factor is obtained:

Significant variables included in the new Synthetic Variables (Factors)

Factor 1: LIQUIDITY	***Factorial Charge***
Quick ratio ((Cash & Equivalents + Accounts & Bills Receivable – Accounts Receivable over 90 days) / Current liabilities)	*0.895169*
Current Ratio	*0.836406*
Cash & Equivalents / Current liabilities	*0.684522*
Leverage (Total liabilities / Total Assets)	*-0.639667*
Payable days	*-0.708725*
Liabilities/Net Sales	*-0.781526*
Factor 2: FULFILMENT	***Factorial Charge***
Accounts Payable over 90 days / Accounts Payable	*0.901607*
Fulfilment	*-0.737011*
Factor 3: PROFITABILITY	***Factorial Charge***
Earnings before taxes / Total Assets	*0.913105*
(Net Profits – Growth of Accounts & Bills Receivable) / Total Assets	*0.892839*
Factor 4: PAYMENT / RECEIVABLE DAYS	***Factorial Charge***
Receivable days	*0.937971*
Payment Days	*0.535763*

To facilitate the interpretation, the making-up of the factors after the rotation process, has been shown only with the variables having factorial charges higher than 0.5. From this value, the correlation between the variable and the factor attained, which is nothing but the expression of these factorial charges, is esteemed significant.

The four factors estimated clearly show four basic aspects which must be analyzed when assessing the short-term payment capacity of a corporation.
Component 1: It is made up of financial ratios describing as a whole, the quick assets so as to face debts. Thus, this factor was defined as "***LIQUIDITY***".
Component 2 It shows the "***PROFITABILITY***", as it is made up of the Economic Profitability and one variant of this indicator, including the increase of accounts receivable.
Component 3: This synthetic variable is greatly important in the model, as it measures the "***FULFILMENT***", being represented by the importance of accounts payable over 90 days within

the whole accounts payable, and by the corporate historical fulfilment concerning its payments on due date.

Component 4: The fourth factor represents the "***PAYMENT / RECEIVABLE DAYS***", made up by the Payment and Receivable days.

This last component contains the only variant which was not polarized after making the orthogonal rotation. The presence, with significant factorial charges, of the indicator Payment Days in two factors, is understood to the effect that is not only important to assess this variable along with the receivable days as part of a corporation financial maturity period, but it is also by itself, an indirect reflection of insufficient available funds to face debts taken. Generally, the entities with an excessive payment days, is not due to credits from sellers but to delay in payments as liquid assets are not enough.

In short, to determine the four synthetic variables, the original variables must be transformed firstly through the box-cox coefficients, then standardized (Deducting the estimated mean and dividing by the standard deviation), and afterwards multiplying the values attained by the coefficients shown in the table below:

Coefficients for attaining the Factorial Scores

	Standardized coefficients of the Factors			
Variables	*1*	*2*	*3*	*4*
Cash & Equivalents / Current liabilities	*0.215844*	*0.050874*	*0.206657*	*-0.145869*
Quick ratio ((Cash & Equivalents + Accounts & Bills Receivable – Accounts Receivable over 90 days) / Current liabilities)	*0.283130*	*-0.067857*	*-0.009418*	*0.176448*
Current Ratio (Current assets / Current liabilities)	*0.284607*	*-0.025092*	*0.039301*	*0.286934*
Accounts payable over 90 days / Accounts payable	*0.152027*	*0.078542*	*0.692793*	*-0.077120*
Leverage (Total liabilities / Total assets)	*-0.130086*	*-0.064191*	*0.120881*	*-0.067540*
*Receivable days (((Accounts & Bills receivable)*360 days) / Sales)*	*0.090098*	*-0.023351*	*-0.006305*	*0.619824*
*Payment days (((Accounts & Bills payable)*360days)/Cost of Sales)*	*-0.229180*	*0.231483*	*-0.054230*	*0.328094*
Earnings before taxes / Total Assets	*-0.042929*	*0.533187*	*0.074592*	*0.105348*
Fulfilment of payments	*-0.064259*	*-0.007070*	*-0.489933*	*-0.021027*
Liabilities / Net Sales	*-0.216750*	*0.008518*	*-0.080477*	*0.175522*
(Net Profit – Growth of Accounts & Bills Receivable) / Total Assets	*-0.110587*	*0.515241*	*0.015934*	*-0.058496*

4. DETERMINATION OF DISCRIMINANT FUNCTION

The discriminant function is based from the four synthetic indicators attained before.

4.1 Assumptions for the Discriminant Analysis

The point that the variables are factors attained through PCA, guarantees their non-multicollinearity. The normality and homoskedasticity, so important in the discriminant analysis so as to get stable and reliable results, are verified through Kolmogorov-Smirnov and M-Box tests:

One-Sample Kolmogorov-Smirnov Test

		REGR factor score 1 for analysis 1	REGR factor score 2 for analysis 1	REGR factor score 3 for analysis 1	REGR factor score 4 for analysis 1
N		52	52	52	52
Normal Parameters [a,b]	Mean	-2.5611E-09	5.1939E-10	4.2626E-09	-5.7312E-10
	Std. Deviation	1.0000000	1.0000000	1.0000000	1.0000000
Most Extreme Differences	Absolute	.054	.185	.092	.081
	Positive	.052	.185	.092	.081
	Negative	-.054	-.140	-.077	-.053
Kolmogorov-Smirnov Z		.389	1.337	.665	.586
Asymp. Sig. (2-tailed)		.998	.056	.769	.882

a. Test distribution is Normal.

b. Calculated from data.

M Box Test to check up the homoskedasticity in both groups.

Test Results

Box's M		17.166
F	Approx.	1.562
	df1	10
	df2	9625.704
	Sig.	.111

Tests null hypothesis of equal population covariance matrices.

For the 95% confidence level, the four synthetic variables have no significant evidence on not following a normal distribution, so the normality is accepted. The same happens regarding the constant variance, because after applying the M-Box test, the p-value (.111) is higher than the 0.05 signification level.

4.2 VALIDITY OF THE DISCRIMINANT FUNCTION

The percentage from the overall variation, explained by the differences between the groups is given by a high canonical correlation[9] of 0.79, which linked to the rejection of the null hypothesis from Wilks´ Lambda test[10] with a p-value virtually null, clearly points out the validity of a discriminant function calculation for the "Default" and "Payment" groups.

4.3 ESTIMATE OF THE DISCRIMINANT MODEL

The calculation of the discriminant function is made following the Fisher's procedure, consisting in finding a linear combination of the predictive variables, whose coefficients are calculated so as to maximize the variance between groups and minimize the variance within groups. The structure matrix obtained after applying the Fisher's procedure, shows the synthetic variable "FULFILMENT" having the biggest correlation coefficient with the discriminant function calculated:

Structure Matrix (Matrix of correlations with the discriminant function)

Structure Matrix

	Function
	1
FULFILMENT	.603
LIQUIDITY	-.340
PROFITABILITY	-.183
RECEIVABLE/PAYMENT DAYS	.152

Pooled within-groups correlations between discriminating
Variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

Logically speaking, these correlations agree with the coefficients from the function obtained, which gives a bigger weighting to the variables FULFILMENT and LIQUIDITY, followed by PROFITABILITY and finally by the RECEIVABLE/PAYMENT DAYS:

9 The Canonical Correlation is given by the expression: CC = $\sqrt{\frac{\lambda_i}{1+\lambda_i}}$; λ = $\sum_{g=1}^{q} n_g \left(\bar{d}_g^i - \bar{d}^i\right)^2$ in which $\bar{d}_g^i$ stand for the mean scores from the discriminant function -*i in the q groups and $\bar{d}^i$ is the overall mean score. As can be seen, the CC takes values between 0 and 1, measuring in relative terms the discriminant power from a discriminant function, obtaining the percentage of the overall variation in the function analyzed.

10 The statistical Wilks´ Lambda (1/(1+λ)), takes values between 0 and 1, so the closer to 0, is bigger the discriminant power.

Canonical Discriminant Function Coefficients

	Function
	1
LIQUIDITY	-.815
PROFITABILITY	-.464
FULFILMENT	1.247
RECEIVABLE/PAYMENT DAYS	.389
(Constant)	.000

Unstandardized coefficients

The discriminant function obtained is represented as follows:

D = -0.815*F1 – 0.464*F2 + 1.247*F3 + 0.389*F4

Synthetic Variables (Factors) :
F1: LIQUIDITY
F2 : PROFITABILITY
F3 : FULFILMENT
F4 : RECEIVABLE/PAYMENT DAYS

The coefficients of Fisher's linear discriminant function for each group are shown below:

	DEFAULT	
	0	1
LIQUIDITY	.876	-1.194
PROFITABILITY	.499	-.681
FULFILMENT	-1.340	1.827
RECEIVABLE/PAYMENT DAYS	-.418	.571
(Constant)	-1.271	-1.767

Fisher's linear discriminant functions

These two functions, which do contain in this case, a significant value for the constant term, can be used in the classification process. If the function belonging to Group 0 ("Payment") is called DO and the group 1 ("Default") D1, the probability of belonging to either of these two groups is calculated as follows:

$P(g=0/X) = (e^{\wedge(D0)}) / (e^{\wedge(D0)} + e^{\wedge(D1)})$

$P(g=1/X) = (e^{\wedge(D1)}) / (e^{\wedge(D0)} + e^{\wedge(D1)})$

The calculation of the probability of belonging to one of the groups shall also enable to classify the corporation. There are only two possible groups. So when calculating such probability P (gi), if its value is higher than 0.5, the corporation shall be classified in this group. Otherwise, it shall belong in the other group with 1-P(gi) probability.

As it can be seen, both in the discriminant function coefficients for both groups and in Fisher's linear discriminant function for each separate group, the components related to liquidity and to payment fulfilment, have got a weighting significantly higher than the other two factors. Therefore, a "step by step" algorithm was also used, which would allow to value the statistical signification of including every variable in the model.

To determine which variables go in and out in each step, the Wilks´ Lambda criterion was used. Such criterion uses the mentioned indicator to measure the gained or lost power when introducing or withdrawing every variable. After applying this "step by step" algorithm, all the variables contributed a significant discriminating power to the discriminant function, so it is not advisable to do without any of them, not even the "RECEIVABLE/PAYMENT DAYS" factor, despite its lower weighting in the function.

4.4 VALIDATION OF THE RESULTS

Due to the already-mentioned small sampling size the most advisable validation is performing, in addition to a simple classification of the elements in the sample from the function determined (Simple Validation), a Cross Validation, consisting in classifying every observation from the discriminant function obtained with the remaining sample.

Classification Results[b,c]

		IMPAGO	Predicted Group Membership 0	1	Total
Original	Count	0	29	1	30
		1	1	21	22
	%	0	96.7	3.3	100.0
		1	4.5	95.5	100.0
Cross-validated[a]	Count	0	29	1	30
		1	2	20	22
	%	0	96.7	3.3	100.0
		1	9.1	90.9	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 96.2% of original grouped cases correctly classified.

c. 94.2% of cross-validated grouped cases correctly classified.

Just as shown by both validation methods, with the model obtained the percentage of success is quite high, either through a simple classification of the cases, or through a more reliable cross validation, with which a high 94.2% of correctly classified elements is attained.

Before concluding the validation, it is important to underline that it was also based on a qualitative appraisal of different parameters, not only from the discriminant analysis, but from the previously applied Principal Components Analysis, as well.

5. PRACTICAL APPLICATION OF THE RESULTS OBTAINED THROUGH THE DISCRIMINANT ANALYSIS

The most obvious application of the discriminant function consists in using it to classify a new case, that is, a new customer or what's most common, a usual customer who submits his financial statements, over a new period of time, to request a credit. So, by only introducing the 11 original variables in an Excel sheet, the discriminant function and the probability of belonging in the "Default" group is obtained by means of the procedure below:

1. **Transformation of the original variables through the box-cox coefficients determined (X_t= X+C)$^{\wedge p}$.**
2. **Standardization of the variables transformed ((X_{ti}-μ)/s)), taking the corporations in the sample for the estimate of the mean and the standard deviation.**
3. **Multiplying the factorial coefficients obtained through the Principal Components Analysis by the standardized transformed variables, determining this way the values for the four synthetic variables.**
4. **Multiplying the discriminating coefficients by the corresponding factors, in order to determine the discriminant function.**
5. **Multiplying the discriminating coefficients of every classification function per groups (Payment or Default), to determine then the probability of belonging in each group.**

To facilitate the interpretation of this discriminant function found, it was additionally decided to organize hierarchically the calculated discriminant function, that is, the relative place of the new observation is determined on the basis of the scores used as sample. Taking this hierarchy or ranking to a 0-10 score, a 0 value means that the factorial score is the worst compared to the sample of corporations, while a 10 score points out that the factor has an optimal value compared to the sample. This new scale is calculated bearing in mind the value of the normal accumulative distribution of "D" score by means of: 10-(NORMAL.DISTRIB(D)*10). The transformation into a 0 to 10 scale, is also calculated for each of the eleven original variables which are initially introduced.

The procedure of relative ranking is performed with the discriminant function in view of facilitating the meaning the score attained. But, in this case, the probability calculated offers excellent information on the payment capability from the corporation rated. Here below, an example of a corporate classification showing a very low probability of default, supported by great indicators for each of the four synthetic variables. This report not only assesses the corporation probability of default, but also the factors influencing the value of such probability.

DEFAULT RISK REPORT

Variables:		
Cash & Equivalents / Current liabilities	0.3853	(9.68 pts)
Quick ratio ((Cash & Equivalents + Accounts & Bills Receivable – Receivables over 90 days) / Current liabilities)	0.9404	(8.11 pts)
Current Ratio (Current assets / Current liabilities)	1.0466	(5.05 pts)
Accounts payable over 90 days / Accounts payable	0.0016	(9.46 pts)
Leverage (Total liabilities / Total assets)	0.8085	(3.68 pts)
Receivable days (((Accounts & Bills receivable)*360 days) / Sales)	37.4185	(6.86 pts)
Payment days (((Accounts & Bills payable)*360days)/Cost of Sales)	28.6909	(9.09 pts)
Earnings before taxes / Total Assets	0.2687	(8.11 pts)
Fulfilment of payments	1.0000	(10 pts)
Liabilities / Net Sales	0.2855	(8.08 pts)
(Net Profit – Growth of Accounts & Bills Receivable) / Total Assets	0.3477	(7.76 pts)
Default Probability:	**0.0028**	

Estimated score by Risk Factors:	Liquidity	Profitability	Fulfilment	Receivable/Payment Days	Weighted Average:
(Base: 10 puntos)	7.46	5.82	8.17	8.20	9.06

	Values from the Discriminant Function			Probability of belonging into a group	
	D	D0	D1	P(g=0/x)	P(g=1/x)
Values from the Discriminant Model:	-2.1209	1.0086	-4.8746	0.9972	0.0028

	F1: Liquidity	F2: Profitability	F3: Fulfilment	F4: Receivable / Payment Days
Factorial scores:	0.6612	0.2068	-0.9058	-0.9160

From the calculation of the default probability, several applications to the cash management and financial risk management are logically derived. Any projected cash flow can be performed, considering the probable inputs and outputs, that is to say influenced by the default probability of indebted corporations.

Also, being able to quantify the customer's payment possibilities, allows defining differed price policies in relation to the default risk, as well as defining the groups of customers where new financial resources must not be invested.

Other several applications derive from the quantification of the credit and liquidity risk, but always taking into account the qualitative assessments on the subject-matter dealt with, since any purely mathematical result unsupported by common sense, may lead to the totally erroneous decision making.

6. CONCLUSIONS

Main Points:

- The procedure followed to obtain the discriminant function, offers a simple and reliable model for predicting the short-term default, enabling the quantification of the default risk.
- The variables selected in this research greatly characterize the financial situation of the companies regarding their liquidity risk. Within the variables stands out the ratio Accounts payable over 90 days / accounts payable, little used when assessing the financial situation of companies.
- In the selection process of variables for the multivariate analysis, the application of non-parametric tests for comparing independent samples is quite important, because of the non normal distribution followed by most financial variables.
- Many of the variables which statistically showed a high discriminating default power include the receivables or payables in their calculation. This shows the importance of this aspect.
- The Principal Components Analysis, based on the normally distributed variables through the Box-Cox transformations, enables to obtain new non-correlated synthetic variables, with normality and constant variance.
- The Multivariate Discriminant Analysis is a very reliable procedure for the default prediction, provided it is based on non-correlated variables and with normal distribution, such as those attained through the Principal Components Analysis.

Main Limitations of the Procedure proposed:

- In the variable selection process, variables containing information on the cash flow produced (From the Cash Flow Statement), must be regarded. They were not taken into account in the present research, as reliable information was not available when the research was made.
- The transformations on the original variables may lead to indefinite expressions when working with extreme values (like logarithms of negative numbers, negative roots, etc.) Therefore, the chosen sample must be as representative as possible, and the coefficients taken in such a way that they ensure a high percentage of the population represented.

REFERENCES

Altman, Edward I. *"Predicting Financial Distress of Companies: Revisiting The Z-Score And Zeta® Models". Stern School of Business, New York University; Nueva York, EUA, 2000.*
Altman, Edward I. *"Revisiting Credit Scoring Models In A Basel 2 Environment". Credit Rating: Methodologies, Rationale and Default Risk, London Risk Books; Londres, Inglaterra, 2002.*
Back, Barbro; Laitinen Teija; Sere Kaisa; van Wezel, Michiel. *"Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis, and Genetic Algorithms". Turku Centre for Computer Science (TUCS), Computational Intelligence for Business; Turku, Finlandia, 1996.*
Calves Pina, Silvia. M. *"Alternativas Para Valorar Instituciones Financieras" (Alternatives for Assessing Financial Institutions). Facultad de Economía, Universidad de la Habana, Cuba, 2003.*
Cortijo Bon, Francisco J. *"Transformación en Componentes Principales (Transformation into Principal Components)". (En Línea) <http://www-etsi2.ugr.es/depar/ccia/rf/www/tema5_00-01_www/node12.html>. Departamento de Ciencias de la Computación e Inteligencia Artificial; Universidad de Granada, España, 2001. (Fecha de Consulta: Ene-Mar/2003).*
Cuadras, Carles M. *"Métodos de Análisis Multivariante". EUB S.L.; Barcelona, España, 1996.*

Damodar N, Gujarati. *"Econometría" (Econometrics), Segunda Edición. Ciudad de La Habana, 1999.*

Falkenstein, Eric. *"Rating Methodology for Private Companies: Moody's Default Model". Moody's Investors Service; Nueva York, EUA, 2000.*

Hair, J. F. Jr.; Anderson, R. E.; Tatham, R. L.; Black, W.C. *"Análisis Multivariante" (Multivariate Análisis), Quinta Edición. Prentice Hall Iberia; Madrid, España, 1999.*

Hyde, Scott. *"Likelihood Based Inference On The Box-Cox Family Of Transformations". Department of Mathematical Sciences, Montana State University. Montana, EUA, 1999.*

Levin, Irwin P. *"Relating Statistics and Experimental Design". Thousand Oaks, CA: Sage Publications. Quantitative Applications in the Social Sciences series #125; California, EUA, 1999.*

Salmi, Timo; Virtanen, Ilkka; Yli-Olli, Paavo. *"On the Classification of Financial Ratios: A Factor and Transformation Analysis of Accrual, Cash Flow, And Marked-Based Ratios". School of Business Studies, University of Vaasa, Finlandia, 1990.*

Salvador Figueras, M. *"Análisis Discriminante" (Discriminant Análisis), (En línea) 5campus.org, Estadística <http://www.5campus.com/leccion/discri>; Universidad de Zaragoza, España, 2000. (Fecha de Consulta: Ene-Mar/2003).*

Salvador Figueras, M. *"Introducción al Análisis Multivariante" (Introduction to the Multivariate Análisis), (En línea) 5campus.org, Estadística <http://www.5campus.com/leccion/anamul>; Universidad de Zaragoza, España, 2000. (Fecha de Consulta: Ene-Mar/2003).*

Siegel, Sidney. *"Diseño Experimental No Paramétrico " (Non-Parametric Experimental Design). Edición Revolución Instituto cubano del libro. La Habana, 1972.*

Siegel, Sidney. *"Nonparametric Statistics for the Behavioural Sciences". McGraw-Hill; Nueva York, EUA, 1956.*

Tomàs, Jaume; Amat, Oriol; Esteve, Mercè. *"Cómo Analizan las Entidades Financieras a sus Clientes". (How the Financial Entibies analyze their Customers) Ediciones Gestión 2000, S.A; Barcelona, España, 2002.*