



Modeling the default probability of Moroccan companies

Siham LOTFI 1^{a*}, Hicham MESK 2^b

^a International university of casablanca, Morocco
siham.lotfi@uic.ac.ma

^b Faculty of Legal and Social Economics of Casablanca, Morocco
Hicham.mesk@univ2c.ma

ARTICLE INFO

ABSTRACT

Keywords

Default risk,
Valuation,
Logistic regression,
Ratios

The financial crisis that has rocked the world in recent years, expressed by the bankruptcy of large international banks (Lehman Brothers in the United States for example), it has led to a questioning of the banking risk management model including credit risk. The objective of our research is to assess the probability of default of firms through a sample that includes 2030 Moroccan companies made up of SMEs and large companies. For this, we used the logistic regression technique which was appreciated in the field of finance mainly in epistemological surveys and credit scoring. This quantitative approach allowed us to measure solvency and at the same time identify healthy borrowers from defaulting borrowers through a number of financial ratios calculated from the balance sheet of the 2015-2017 financial statements.

1. INTRODUCTION

Over the past few decades to the present day, studies devoted to the problem of default risk assessment and forecasting the financial distress of companies are constantly increasing. Among these research works, we cite those of Bardos and Zhu (1997), Chava and Jarrow (2004) and Hillegeist (2004). A large number of these studies are based on techniques of statistical analysis of accounting quantities and financial ratios in order to discriminate between healthy companies from failing companies.

Indeed, the prudential regulations imposed by the Basel Committee are of great importance insofar as they allow banks to foresee the risk of default that may be generated by a company requesting a loan, thus judging the quality of each loan line and optimize the return on funds granted. Thus, mastering the management of financial distress remains an objective sought by banks insofar as they can recover their borrowed capital, decide to refuse to renew or grant new loans.

2. General information on credit risk: concepts and definitions

Granting loans to companies remains a main source of funds for the majority of banking establishments and also for the development of a country's economy, however, these revenues generate risks that must be controlled from the first phase of the request. lending, hence the central theme of the new Basel agreements.

2.1 Some notions of risk

In general, risk corresponds to situations of loss whose probability of occurrence is not zero, so this situation is probable and measurable. (Greenbaum and Thakor, 2007), highlights risk as a fundamental element influencing financial behavior and financial institutions, including banking institutions. They must manage it well to survive in an increasingly uncertain environment. "Risk is endemic to business but central to banking". For Joël BESSIS, all risks are defined as losses associated with adverse developments. The important direct consequence

* Corresponding author
E-mail address: siham.lotfi@uic.ac.ma (S. LOTFI)

is that any measure of risk relies on the assessment of such impairments and their impact on results. According to SAMPSON A, it is about "the tension that lives in bankers is inseparable from their job, they watch over the savings of others and yet they make profits by lending them to others, which inevitably involves risks. . A banker who does not take risks is not one. Nalleau G and Roucham M designate risk as "a commitment bearing uncertainty endowed with a probability of gain and harm, whether this is a degradation or a loss".

2.2. Credit Definitions

Credit is an agreement by which a certain sum is lent against a promise of repayment and against payment of interest (Josette and Max Peyrard, 2001). As for (Ahmed Silem and Jean-Marie Albertini, 1983), define credit as an act of trust resulting in a loan in kind or in cash granted in return for a promise of reimbursement, within a period generally fixed in advance. . On the way to this last definition. (Lukuitshi, 2010), emphasizes that credit is an operation during which a lender can make available to a borrower or debtor, a sum of money subject to a repayment commitment on a date determined in advance.

According to G. Petit Dutaillis (1967), to give credit is to trust, it is to make available a real good, a purchasing power, against the promise that the same good will be returned within a certain period, the most often with remuneration for the service rendered and the danger incurred, danger of partial or total loss inherent in the very nature of this service. Thus, credit can be defined as a loan granted by a banker for remuneration taking into account the duration of the loan and the risk linked to the situation of the borrower.

In summary, credit results from the combination of three determinants: The time or period during which the beneficiary has the loaned funds, the trust placed by the creditor in the debtor, the promise of restitution of the loaned funds.

In general, a credit transaction, considered from the lender's point of view as a risky transaction that requires the intervention of regulations intended to reduce the risk incurred.

2.3. Credit risk: Literature review

By definition, credit risk is the risk that a borrower does not repay all or part of his credit on the due dates set out in the contract signed between him and the bank. In the context of risk management, credit risk is considered a major risk in a bank. Indeed, many authors have attempted to define the notion of credit risk, for Faye (1993), this risk is defined as being the risk of losing all or part of the receivables in the event that the borrower does not have more at the end of the term the will or the possibility of honoring its commitments. For Wonou (2006), credit risk can be defined as the probability that loans granted to one or more customers will not be reimbursed. Credit risk can be defined as the risk of loss linked to the default of a borrower on a commitment to repay debts that he has contracted.

The Basel Committee defines credit or counterparty risk as the risk of non-reimbursement associated with a loan granted by a bank. Indeed, it corresponds to the uncertainty related to the ability of a borrower to settle his debts and to respect his contractual commitments.

For Desmicht (2007), several works of research have treated the credit risk of companies with a term that is different from one study to another, it is generally called risk of default, risk of cessation of payment or risk of default. Thus he defines credit risk as the risk of loss in the event of default by the borrower. This is the risk of non-payment or risk of default.

3. Empirical study of credit risk

Among the methods used in predictive analysis, we find the technique of logistic regression which has been appreciated in the field of finance mainly in epistemological surveys and credit scoring.

3.1. Determination of the representative sample

Our sample constituted covers a three-year period from 2015 to 2017, and includes 2030 companies, made up of SMEs and large companies (GE) with both financial and descriptive information.

For the elaboration of our development and test sample, we opted for a simple stratified sampling.

Table 1: Breakdown of the study sample

LABELS LINES	NUMBER OF DEFAULTS	DEVELOPMENT SAMPLE	% OF DEVELOPMENT SAMPLE	OF VALIDATION SAMPLE	% OF THE VALIDATION SAMPLE
0	1821	1275	90%	546	90%
1	209	146	10%	63	10%
TOTAL	2030	1421	100%	609	100%

Namely that modality 1 represents the failing company and 0 the healthy compan

3.2. Presentation of the explanatory variables for predicting failure

Among the financial determinants of failure, we find in particular the ratios relating to liquidity, solvency, size and indebtedness. In our study, we retained 17 financial ratios as explanatory variables of the target variable which is credit default. The literature on financial economics has shown the primary role ordinal played by these financial variables in the treatment of the situation and the financial health of the company through the financial analysis or other evaluation method used by financial analysts, investors or rating agencies.

Based on the empirical literature review of various works in the field of predicting corporate financial difficulties, the explanatory variables used are generally financial ratios.

Below are the different ratios considered as explanatory variables used in our model.

Table 2: The financial ratios used in our study

TYPE OF RATIOS	VARIABLES	FINANCIAL RATIOS	FORMULA
<i>LIQUIDITY RATIOS</i>	R1L	Stock rotation	(Stock/Turnover) *365
	R2L	Rotation of trade receivables	(Customer receivables/Turnover) *365
	R3L	Cash Ratio	Current Assets/Liabilities Cash
	R4L	Turnover/Fixed Assets	
	R5L	Equity/Medium and Long Term Debt	Equity/financing debt

TYPE OF RATIOS	VARIABLES	FINANCIAL RATIOS	FORMULA
<i>SIZE RATIOS</i>	R1T	Age	Rating Date- Creation Date
	R2T	Turnover	Log Turnover

TYPE OF RATIOS	VARIABLES	FINANCIAL RATIOS	FORMULA
<i>ACTIVITY RATIOS</i>	R1A	Asset Turnover	Turnover / Total assets

TYPE OF RATIOS	VARIABLES	FINANCIAL RATIOS	FORMULA
SOLVENCY RATIOS	R1S	Gearing	Medium and long-term debt/Equity
	R2S	% Equity in the Structure	Equity/ Debt + Equity

TYPE OF RATIOS	VARIABLES	FINANCIAL RATIOS	FORMULA
DEBT RATIOS	R1E	Debt	(Assets-Equity) /Assets
	R2E	Coverage of financial costs	Financial Charges/Added Value
	R3E	turnover in working capital requirement	working capital requirement/ Turnover *365

3.3. Univariate analysis of variables

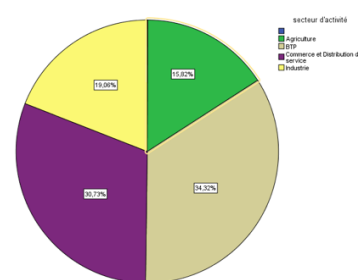
The initial analysis of the descriptive statistics is useful for understanding the variability of the elements used in the study and detecting the errors that may occur in the estimates. The table below presents the main descriptive statistics of the variables used in the analysis.

Table 3: Descriptive statistics

VARIABLES	MEDIUM	MEDIANE	MODE	ECART TYPE	MIN	MAX
AGE	3,073	3,000	3,000	,8215	1,000	4,000
ENDETTEMENT	5,50106	5,00000	3,000	2,875742	1,000	10,000
CA	5,50669	6,00000	10,000	2,876959	1,000	10,000
TB	5,50598	6,00000	10,000	2,874021	1,000	10,000
ROTATION STOCK	4,60662	5,00000	1,000	2,732060	1,000	9,000
ROTATION CREANCES CLIENTS	5,50457	6,00000	4,000 ^a	2,874514	1,000	10,000
AC/PC	5,50739	6,00000	7,000	2,874753	1,000	10,000
CA/ACTIF IMMOBILISE	5,50739	6,00000	10,000	2,876712	1,000	10,000
TRESORERIE NET	5,50457	6,00000	10,000	2,876228	1,000	10,000
FP/STRUCTURE	5,20901	6,00000	8,000	2,477920	1,000	8,000
BFR/CA	5,50457	6,00000	2,000 ^a	2,876228	1,000	10,000
FP/DETTES	5,51513	6,00000	10,000	2,884749	1,000	10,000

All the companies operate in various sectors of activity, we find 30.73% of the companies carry out their activities in the trade and distribution of services 34.32% in the construction sector, 20% in industry and 15, 82% in agriculture. The following graph illustrates the distribution of companies by sector of activity.

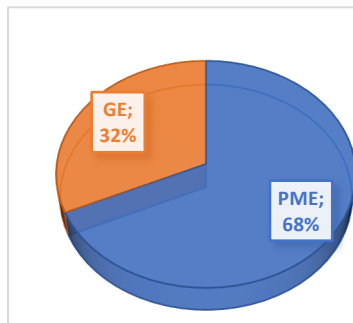
Graph 1: Breakdown of companies in the sample by sector of activity



Source : spss

For the legal form variable, our sample is made up of 54% SAs and 46% SARLs.

The circular below represents the business segment variable, our sample consists of 68% represented by the category of small and medium-sized enterprises SME and 32% by the category of large enterprises.

Graph 2: Breakdown of SMEs and large enterprises (LE) in the sample

Source : SPSS

3.4. Bivariate analysis of variables

To test the dependence between the explained variable and the explanatory variables, we used the chi-square test, this

test of independence of the chi-square is used to assess the existence or not of a relationship between two characters within a population, when these characters are qualitative or when one character is quantitative and the other is qualitative, or even when the two characters are quantitative but the values have been grouped together. The null hypothesis is:

H0i: there is independence between the explained variable Y and the explanatory variable Xi.

The chi-square test, represented in the table below, shows that, for all the variables retained, the hypothesis of independence is rejected at the 5% significance level except for the variables Age, Rotation Stock, current assets/current liabilities, turnover/fixed assets, ROE, so we can say that there is a dependence between the explained variable and the explanatory variables chosen.

Table 4: Results of the Chi 2 test

VARIABLES	CHI (VALUE)	2	CHI2 (SIG%)	SIGNIFICANCE	PREDICTIVE POWER	CORRECTED SIGNIFICANCE	CORRECTED PREDICTIVE POWER
AGE	7,2207		,065	>5%	Non	>5%	Non
DEBT	40,349		,000	<5%	Oui	<5%	Oui
TURNOVER	42,091		,000	<5%	Oui	<5%	Oui
BALANCE SHEET TOTAL	64,865		,000	<5%	Oui	<5%	Oui
ROTATION STOCK	12,249		,140	>5%	Non	>5%	Non
TRADE RECEIVABLES	54,273		,000	<5%	Oui	<5%	Oui
ROTATION CURRENT ASSETS/CURRENT LIABILITIES	11,060		,272	>5%	Non	>5%	Non
TURNOVER/FIXED ASSETS	17,937		,036	>5%	Non	>5%	Non
NET CASH	55,740		,000	<5%	Oui	<5%	Oui
EQUITY/STRUCTURE	17,863		,013	<5%	Oui	<5%	Oui
WORKING CAPITAL REQUIREMENT / TURNOVER	21,261		,012	<5%	Oui	<5%	Oui
FP/DETTES	29,206		,001	<5%	Oui	<5%	Oui
FRAIS FINANCIERS/CA	41,656		,000	<5%	Oui	<5%	Oui
ROE	16,005		,067	>5%	Non	>5%	Non
CASH RATIO	27,055		,001	<5%	Oui	<5%	Oui
ASSET TURNOVER	52,470		,000	<5%	Oui	<5%	Oui
GEARING	340,2		,000	<5%	Oui	<5%	Oui

3.5. Correlation analysis and selection of explanatory variables

The indicator used to measure the degree of correlation between the variables is the Spearman correlation coefficient which is a correlation indicator by ranks, it is a

coefficient which expresses the relationship between two variables, either numerical or alphanumeric.

The value of p which is the correlation coefficient is measured as follows:

- If the result < 0.3% : Means a weak correlation

- If the result between 0.3% and 0.5% : Means a medium correlation

- If the result >0.5% : Means a strong correlation

In our case study, we used the Spearman test, we kept only the variables with a value less than <0.3.

After determining the correlation effects in particular in a multivariate framework, we retain the following explanatory variables which will be used to present our model :

- The explained variable : the default of the company
- The explanatory variables : Debt, Turnover, Rotation of customer receivables, Equity in structure, Cash ratio, Financial costs, WCR/CA, Asset Turnover, Equity/ML debt.

Regarding the correlated ratios, we tried several tests under SPSS in order to eliminate redundancies and we retained 9 ratios in our prediction model which represent the best explanatory power.

We can thus notice that variables from five categories of financial ratios (Debt, Size, Liquidity, Solvency, Activity) are significant.

4. Development of a default prediction model based on the logistic regression method under SPSS

At this level, we will present the logistic regression model with the interpretation of the results obtained.

4.1. Presentation of the model

We have a final sample of 1421 observations. Within the framework of compliance with the time horizon and in accordance with the regulations of the Basel Committee (Basel Committee on Banking Supervision, 1999), the duration used in our default prediction model is 2 years, which largely respects the duration required, which is 12 months in the credit risk prediction models used by banks subject to the international standards of the Basel agreements.

We define the multiple linear regression model as any linear regression model with at least two explanatory variables.

The goal of the regression is to establish the law $y = f(x)$. Once this law has been estimated, we will try to predict a value of y for a given value of x .

The model we want to estimate

$$\text{Log}(\pi/1-\pi) =$$

With π : the probability of being “Healthy” $1-\pi$: the probability of being “Failing”.

The estimation of the model is made by the SPSS software such as the explained variable which is a dichotomous variable and can only take two modalities, respectively (0 and 1).

$Y_i = 0$ If the company is healthy

$Y_i = 1$ If the company is failing

After several tests using the logistic regression method on SPSS, the default prediction model retained 9 financial ratios and 3 signal variables.

4.2. Logistic regression results

The table below indicates the weight of each variable of the study in the prediction of default risk, we note that the variables with the most important predictive power are indebtedness, legal form, size of the company , liquidity and structure.

Table 5: calculation of the marginal effects of the study variables

CATEGORY VARIABLES RATIOS	OF OR EXPLANATORY VARIABLES	B-FACTOR	MARGINAL CONTRIBUTION	WEIGHT
SIGNAL VARIABLES	Code_forme juridique	-,025	0,02491566	20%
	Code_secteur d'activité	,002	0,001550511	1%
	Code_région	-,004	0,004034625	3%
SEIZE	CA	,017	0,017125153	14%
DEBT	Endettement	,003	0,003408292	3%
LIQUIDITY	Rotation Créances Clients	,013	0,012879715	10%
STRUCTURE	Fp dans structure	-,003	0,003355382	3%
LIQUIDITY	BFRCA	,005	0,004759032	4%
DEBT	Fd/dettes ML	,024	0,024194076	20%
DEBT	Frais financiers/CA	,004	0,003803625	3%
PROFITABILITY	Cash ratio	-,007	0,007199971	6%
LIQUIDITY	Asset turnover	-,015	0,015498618	13%
TOTAL				100%

We adopted the following model:

$$\begin{aligned} \text{ZScore} = & 0.2 * \text{legal form} + 0.01 * \text{sector of activity} + 0.03 * \\ & \text{Region} + 0.14 * \text{Turnover} + 0.03 * \text{Debt} + 0.1 * \text{Rotation} \\ & \text{Customer Receivables} + 0.03 * \% \text{equity in} \\ & \text{structure} + 0.04 * \text{working capital requirement} / \text{Turnover} \\ & + 0.2 * \text{working capital/medium and long term} \\ & \text{debt} + 0.03 * \text{Financial costs/CA} \\ & + 0.06 * \text{Cash ratio} + 0.13 * \text{Asset turnover} \end{aligned}$$

4.3. Discussion of results

According to the table of the marginal effects of the different variables used in our study, we see that the category of the ratio which has a significant weight on the default variable is the liquidity ratio with a weight of 27% followed by the ratio of debt (26%), the legal form variable with a percentage of 20%, the size ratio (14%), the profitability ratio with an effect of 6%, the region variable (3%) and finally the sector of activity (1%).

The coefficients obtained thus show the signs of the partial effects of each explanatory variable on the probability of default.

For the liquidity ratio, it includes the following financial variables: Customer Receivables Rotation, WCR/CA and Asset turnover.

- With regard to the variable Customer Receivables Rotation, it positively influences the risk of default, which means that the more the company grants its customers loans with a long duration, the more it finds itself in a situation of illiquidity, which leads to an increase in the credit risk of this company vis-à-vis its bank.

- BFR/CA has a positive effect on the forecast of credit risk, the more the company observes a change in its BFR/CA, the more a reduction in the risk of default is expected.

- The Asset turnover ratio is negatively linked with the probability of default, which means that the greater the turnover of the company's assets, the lower the risk of default.

In our study, the debt ratio includes the following financial variables : Debt, Equity/Medium and Long-term debt, Financial costs/Turnover.

- For the indebtedness variable, we note that it is positively correlated with the default risk forecast, so the more the company is indebted, the greater the probability of falling into a default situation.

- For the Equity/Medium and Long-term debt variable, it has a positive effect on credit risk, which means that the greater the share of debt in the company's equity, the greater the risk of default. .

- A change in financial costs/turnover leads to a strong chance of being risky. Financial costs show a company's indebtedness, the higher the ratio, the greater the risk of default.

- With regard to the size ratio, we have the turnover variable which has a negative effect on the prediction of default, the more the turnover of a company increases the more the probability of being in default decreases.

- As for the profitability ratio, we find the cash ratio which acts negatively on the forecast of credit risk, which supposes that the more the company is profitable the more it does not have payment difficulties.

5. Evaluation of the discriminating power of the model on the modeling sample

In order to test the discriminating power of our estimation model, we adopted the methods most used in research work, the Fisher test and the rate of good classification.

5.1. Evaluation of the significance of the model by Fisher's test

The principle of this test is based on the frequencies obtained in each cell of the crosstab (categorical variables do not have a mean or variance as a reference).

In our model, the Fisher test is presented in the table below :

Table 6 : ANOVA

ANOVA ^A						
MODEL		sum of squares	Ddl	F	Sig.	
1	Regression	9,099	12	0,758	8,825	,000 ^b
	Residues	84,887	988	0,086		
	Total	93,986	1000			

According to the results of the Fisher statistic which presents a significance, we can conclude that the model is able to discriminate the target variable which is the defect.

5.2. Nickname R2

The pseudo-R2 statistic is therefore well between 0 and 1. The two extreme cases are:

- If the constrained model is correct, = and the statistic is 0 (the variations of Y_i are not explained by the X_i).
- If the unconstrained model fits the data perfectly, = 0 and the pseudo-R2 is 1.

Table 7: Model summary

Summary of models										
Model	R	R-square	Adjusted square	R-	Standard error of the estimate	Edit statistics				
						Variation of R-square	Variation of F	ddl1	ddl2	sig. Variation of F
1	,844 ^a	,713	,632		,093	,713	8,825	12	988	,000

The table above presents a summary of the logistic regression model used, in particular the coefficient of determination R^2 estimated at 71%, which concludes that the model has good discriminating power.

Quality of classification rules: Cross-validation method

The measurement of the quality of a discrimination is done from the percentages of well classified (or badly classified) in each class, and of the overall percentage of well classified.

This measure can also, in some applications, involve misclassification costs.

A ranking table can calculate a percentage of correctly classified on the training sample, which will give an optimistic idea of the quality of discrimination.

The table below shows the results of the predictions of the Logit model developed. We notice that 89.7% of the observations are well classified in the sample.

According to the results, the established model achieved an overall good classification rate of 89.5%, that is to say that out of 100 borrowing companies 89 will be classified correctly and 11 will be badly classified by our model, which is satisfactory as a result.

6. CONCLUSION

Credit risk remains the primary concern for banks and other financial institutions who have put a lot of effort into

trying to find the most effective ways to control or mitigate credit risk.

Mastery of credit risk management remains an objective sought by banks in order to prevent financial difficulties and identify healthy borrowers and defaulting borrowers.

In addition to the statistical approach of logistic regression, new management techniques are emerging so that banks can adapt to these significant changes while remaining efficient, such as models based on artificial intelligence (Neural networks, expert systems, etc.).

Table 8: Good ranking table of companies in the learning sample

CLASSIFICATION TABLE A				
OBSERVED	Forecasts			
	FAILURE code		Percent	
	0	1		Correct
Failcod	0	893	3	99,7
F	1	100	5	4,8
Percent global				89,5
A. LA VALEUR DE COUPE EST ,500				

Bibliographic references

- Asma GUIZANI. (2014) Treatment of refused files in the process of granting credit to individuals.
- Altman, E.I. 1980. "Commercial bank lending: Process, credit scoring and costs of errors in lending", Journal of Financial and Quantitative Analysis 15, p. 813-832.

- Altman, E.I., and Saunders, A. 1998. "Credit risk measurement: Developments over the last 20 years", *Journal of Banking and Finance* 21, p. 1721-1742.
- Glossary of Descriptive Statistics, 2010 "The Pearson Chi-square test of independence". P: 1
- Hand, D. J., Henley, W. E. (1997) "Statistical Classification Methods in Consumer Credit Scoring: A Review". *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160 (3): 523-541.
- Hicham ZMARROU. (2006) the risk management system & internal control within credit institutions.
- Jeanne Lazarus. 2012, "Predicting credit default: the ambition of scoring" *Raisons politiques* 2012/4 (n° 48), p. 103-118. Cairn.info
- Joël BESSIS, Risk Management and Asset-Liability Management of Banks, Edition Dalloz, Paris, 1995, p.15
- NAULLEAU Gérard and ROUACH Michel (1998), management and financial control, *Revue Bancaire*, page 30
- Rim BOUSSAADA. (2012), The impact of banking governance and the banking relationship on credit risk: The case of Tunisian banks.
- SAMPSON A, Banks in a dangerous world, R.Laffont, 1982, p.38
- SAHOSSI & Ibrahima DIARRA, 2009 "Credit risk analysis for construction companies: case of BNDA".
- SAUNDERS A. (1999), "Credit risk measurement – New approaches to value at risk and other paradigms", New York, Wiley, 226 p
- Selma HAJ KHLIFA (2016), "Basel system and access of SMEs to bank financing. » " Doctoral Thesis, ISCAE 2016.
- Treacy W.F. and CAREY M. (2000), "Credit Risk Rating Systems at Large US Banks", *Journal of Banking & Finance*, vol. 24, n° 1/2, January 2000, pp. 167-201

- Matoussi, H., Karâa, A. and Krichène, A. (2004) "Usefulness of financial information in the credit evaluation process: An exploratory study on the informational contribution of the Tunisian SCE to the credit granting process » the fourth international days of research in management sciences, the Tunisian Association of Management Sciences and the Konrad Adenauer Foundation 11,12;13 March 2004.
- Mamoughli, C (1984), "The forecast of the risk of failure of Tunisian companies" Thesis of Doctorate of third cycle, university of Paris Dauphine 1984. »