

THE IMPACT OF EXTERNAL FACTORS ON THE DECISION TO GRANT A LOAN. A CASE STUDY ON ROMANIAN BANKS

Lucian Găban¹

ABSTRACT: The objective of this study is to identify external factors having a significant impact on the probability of granting a loan, and to construct a credit scoring model for small and medium enterprises. Applying the logit analysis, we have found three external predictors which characterize the firm's financial condition. These predictors are as follows: EBITDA, ROI, and CR, and they influence the decision to grant a loan. The significant non-linear effects improve the quality and prediction power of the model. These external factors predict 85 percent, a high percentage of correctly classified observations.

KEYWORDS: the probability of granting a loan, credit scoring, logit model, non-linear effects

JEL CODES: G21, G23

Introduction

There are a lot of micro entrepreneurs who want and need financing support for their businesses. One of the factors with relevant impact on increasing the level of credit consumer has been the increase in the use of credit scoring. In the U.S. microlending field, the credit scoring plays an important role. Contributing to the reduction of the cost of making loan decisions, credit scoring has made small business loans profitable. The credit scoring in financial credit institutions and banks is to help more precisely evaluate a credit. The applications of the scoring methodologies have been widely extended to include different areas, particularly in financial institutions, to predict their client's behaviour. The recent world financial crisis has mobilized even more the financial institutions regarding estimating and predicting the credit risk.

According to Berger et al. (2007), credit scoring is a means of estimating the desirability of a loan customer based on his/her basic characteristics and past experiences with credits. The statistical procedure is based on quantitative information that is collected on the factors related to the business of customer when applying for a credit. Once the factors which relate to strong or weak payments are identified, one can build a statistical model that determines the weight of each factor. The factors which have a relevant impact on the repayment behavior are weighted more heavily in score. Based on this analysis, one can develop an evaluation system that gives a score that describes the data accumulated on loan applicants, and a probability of granting a loan can be estimated.

The credit risk remains one of the major threats which financial institutions face, and it is essential to model the credit risk of financial institutions. The problems resulting from the non-performing loans can cause financial distress in banks, which in turn creates economic and social turmoil. Banks need to manage the credit risk inherent in the entire portfolio, as well as the risk in individual credits. In a lot of researches on credit risk modeling there are used statistical and econometric techniques, for instance multivariate discriminant analysis, logit and probit models and these are among the most widespread for credit risk modeling (see Altman and Saunders (1998), Gordy (2000), and Lin and Chang (2006). Developing credit scoring models for financial

institutions, especially for banks, is of major interest to analysts, practitioners, and supervising authorities. Such models assist users to make quick decisions and this is extremely desirable when banks are dealing with large number of clients. Identification and determination of risks in different parts of financial activities have a major impact on stability and survival of financial institutes. In the paper Samad (2012), the author using a deep technical analysis, determines the significant internal factors in predicting bank failures, applying the probit model. In order to give facilities to firms, the bank should study external factors and analyze the financial statements, in other words a financial analysis should be made. In this paper we study the impact of external factors on the probability of granting a loan.

In accordance with the rules of lending, as specified in Basel II (2006), the objective of this study is to identify external factors having a significant impact on the probability of granting a loan, and to construct a credit scoring model for small and medium enterprises. Basel II encourages the implementation of internal own models, for measuring their financial risks. The credit scoring model is a function of the quality of a bank's loans, which depends of the financial state of of the firms that are credited. According to the declared financial statements, the crediting bank analyses two groups of indicators: liquidity and profitability indicators. Which are the external factors that influence the decision to grant a loan? What is the explanatory power of the identified factors? The identification of external predictors having a direct impact on the probability of granting a loan is an important contribution of this study in the banking literature. This paper proposes a study based on a random sample of small size of small and medium enterprises that want loans, compared with other works (such as Lennox (1999), Samad (2012), and Danila (2012)) which use large samples. After careful consideration of the nature of the problem and of the purpose of this study, based on the following relevant papers: McFadden (1973), Ohlson (1980), Zmijewski (1984), Lennox (1998), and Greene (1998), we have decided to choose the logistic analysis as an appropriate statistical technique.

This paper is organized as follows: Section 2 gives a survey of the scientific literature. The theoretical developments and statistical methodology are discussed in Section 3. Data and empirical results are provided in Section 4. Section 5 gives concluding remarks.

A Survey of the Scientific Literature

In the last period a significant number of international journal articles have addressed different credit scoring techniques in different fields. In the following, we shall present the most relevant papers of scientific literature that were the basis for the justification and development of this paper.

Altman and Sabato (2005) developed a distress prediction model specifically for the small and medium sized enterprises, and they analyse its effectiveness compared to a generic model. This paper analyses a set of financial ratios, linked to US SMEs, and finds out which the most predictive variables affecting the credit worthiness are. The variables from our study are from these categories: liquidity, and profitability. Sabato (2008) emphasizes the importance of dividing the lifecycle of every model into three phases: assessment, implementation, and validation. In this paper, based on the random sample of small size, we have developed an early warning model corresponding to the first phase of lifecycle of every model. The idea of the construction of early warning models is found in Platt & Platt (2006).

According to Abdou and Pointon (2011) the credit scoring is "one of the most important kits, to classify a bank's customers, as a part of the credit evaluation process to reduce the current and the expected risk of a customer being a bad credit". Credit scoring models are some of the most successful applications in finance and banking. The authors point out the advantages of credit

scoring systems compared to subjective systems. A credit scoring system is a consistently process applied to all credit decisions.

Medina et al. (2010), focuses on the internal models for the assessment of credit risk IRB – Internal Ratings-Based approach. “These models are based on the bank's own estimates and their objective is to calculate the unexpected loss from credit exposures. The amount of this loss depends on a set of factors: probability of default, loss in the event of default, exposure at the time of the default or severity, maturity and granularity”. The probability of granting a loan is a component of credit risk, estimated in the initial phase of analysis of a potential loan, helping to strengthen the banking prudence.

For the first time, Ohlson (1980) applied the conditional logit model to the predictions problems. The empirical results of study identify two factors, that measure performance and current liquidity, as being statistically significant in affecting the probability of failure. The same factors affect the probability of granting a loan.

Zmijewski in Zmijewski (1984) was the pioneer in applying probit analysis to predict default, but, until now, logit analysis has given better results in this field.

The conditional logit analysis is mathematically founded by McFadden (1973). The impact of conditional logit analysis to economic analysis can be indicated by many consumer choice problems to which these models have been applied. An additional argument in choosing logit models is the interpretive simplicity, outlined by Ohlson (1980).

The smooth functions produced by local scoring procedures, developed by Hastie and Tibshirani (1995), can be used to suggest covariate transformations. GAM provides a method for identifying non-linear covariate effects in exponential family models. Several studies emphasize that the relationship between explanatory variables, in the logit models, are often nonlinear. Dakovic et al. (2010) have used GAM (Generalized Addaptive Models) to reveal non-linear relations to be used in the generalized linear model. The mixed logit model with heterogeneous parameters and nonlinear relationships for covariates, as in Kukuk and Ronnberg (2013), help to improve the quality of models. The significant non-linear effects, see Lennox (1998), improve the model's explanatory power, as outlined in my paper.

Under normal conditions, the data would be drawn randomly from a population and used to fit the model, which could then be used to make predictions. In reality, the conditions necessary to obtain an accuracy model are not always fulfilled. Greene (1998) made a thorough analysis of the impact process of obtaining data on the accuracy of the model. In this respect, in this paper, we have used the technical means necessary to obtain a random sample of small size.

Theoretical Developments and Statistical Methodology

A wide range of statistical techniques are used in building the scoring models, such as classical techniques: regression analysis, discriminant analysis, probit analysis, logistic regression, and linear programming. Several other statistical techniques have been tested to improve the prediction accuracy of credit scoring models, such as: Cox's proportional hazard model, decision trees, neural networks, k -nearest-neighbour, but the empirical results have never shown really significant benefits. Most of these statistical models are applicable to build an efficient and effective credit scoring system that can be effectively used for predictive purposes. According to Berger and Scott Frame (2005), "Credit scoring is a statistical approach to predict the probability that a credit applicant will default or become delinquent". A special place is granted to evolutionary computing techniques, necessary for an integrating approach regarding: classification, variable selection, and parameter optimization, specific to credit scoring as in Marques et al. (2013).

Observing the predetermined event (default), we construct the dependent variable. The dependent variable D_n is a binary variable: $D_n = 1$ for firms that have accessed credits, representing

the default event; $D_n = 0$ for firms that have not accessed credits. Since the dependent variable is binary, the binary outcome models are used for identifying the probability of success (granting credit), which is modeled to depend on regressors. The model used, explains an unobserved continuous random variable y^* , further on called latent variable. What is observable is the binary variable y , which may take the value 1 or 0, as y^* exceeds a threshold. Let y^* be a unobserved variable, and regression model for y^* is the index function model (see Cameron and Trivedi (2005))

$$y^* = X\beta + u \tag{1}$$

where the regressor vector X is a $K \times 1$ column vector, the parameter vector β is a $K \times 1$ column vector, and the error vector u is a $K \times 1$ column vector. Then $X\beta = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$. The parameter vector β from model (1) cannot be estimated, because y^* is not observed. We have

$$y = \begin{cases} 1 & \text{if } y^* > 0.5 \\ 0 & \text{if } y^* \leq 0.5 \end{cases} \tag{2}$$

where the threshold is 0.5. From (1) and (2) we have

$$\Pr(y = 1 / X) = F(X\beta) \tag{3}$$

where $F(X\beta)$ is the cumulative distribution function (c.d.f.) of $-u$. If u is standard normally distributed is obtained the probit model, and the logit model if u is logistically distributed. Logistic regression has been extensively used in credit scoring applications like in: Ohlson (1980), Kukuk and Ronnberg (2013, Danila (2012). Henley and Hand (1996) provide a comparison between decision trees and other techniques, such as logistic regression and K -nearest neighbour, in terms of average bad risk rate. However the bad risk rates were clearly similar for the different scoring techniques.

Let a vector of data denoted as $X_i = (X_{i1}, \dots)$, $i = 1, \dots, N$ from N observations. From the model (3), the conditional probability is given by

$$p_i = \Pr[y_i = 1 / X_i] = F(X_i \beta) \tag{4}$$

The maximum likelihood estimator (MLE) is the estimator for binary models. Given a sample (y_i, X_i) , $i = 1, \dots, N$ of N independent observations, we obtain a ML estimation $\hat{\beta}$. The Wald test and the likelihood-ratio (LR) test are used to test the significance of coefficients $\hat{\beta}$. The estimated coefficients help us to weight the impact of each of the independent variables on the estimated probability of default. The change on the probability $\Pr(y = 1)$, using the following formula

$$\frac{\partial \Pr\left(y_i = \frac{1}{X_i}\right)}{\partial x_{ij}} = F'(X_i \beta) \beta_j$$

represents marginal effects of X_i . A measure of goodness of fit in the nonlinear models is pseudo- R^2 . This measure is computable for the binary models, and is proposed by McFadden. The guide, Cameron and Trivedi (2009), explains how to perform a good regression analysis.

Data and Empirical Results

The objective of a credit scoring model is to assign a loan to customers regarding a verdict if to access a credit or not. Therefore, scoring problems are related to classification analysis. The selection of the variables used in building scoring models depends on the quality of the data and the availability of those who supply them. There is no default number of variables that should be used in building scoring models. The data may affect the quality of the model. By improving the data base, the predictive classification power of a model should be more realistic, and the standard tests of statistical significance should be more compelling.

Data

The study utilizes a sample containing financial data for 60 small and medium sized enterprises from Romania, which are randomly chosen. These form a random sample from the companies of various fields that have been analyzed by the bank in order to access a credit. Taking into account just the economic-financial data, we shall develop a warning model which estimates the likelihood of granting a loan. The analyzed financial statements taken into account by the credit bank refer to the latest activity time of the company: one, two or three years.

In finance applications, a rank from only three variables to around twenty variables has been used in building the scoring models. The model used for credit scoring are based on the idea of delimitation between two groups in a data sample. But what determines the categorisation of a applicant? In order to assess the discriminatory power of each variable, according to Fernandes (2005), three of the variables have been selected. Result of correlation of selected variables is provided in Table 1. There are three variables which are used as explanatory variables. They are:

- Earnings before interest, taxes, depreciation and amortization (EBITDA);
- Return on investment (ROI);
- Current ratio (CR).

EBITDA is an accounting measure calculated using a company's net earnings, before interest expenses, taxes, depreciation and amortization are subtracted. ROI, the most common profitability ratio, evaluates the performance of a business by dividing net profit by net worth. CR is a popular financial ratio used to measure company's liquidity, by deriving the proportion of current assets available to cover current liabilities. Further on we analyze the impact of variables: EBITDA, ROI and CR, on the likelihood of granting a loan. Table 1 shows no strong correlation among the three variables.

Table 1

Correlation Matrix			
	EBITDA	ROI	CR
EBITDA	1.0000		
RO	0.4532	1.0000	
CR	0.1374	0.4717	1.0000

They can be used as explanatory variables in credit scoring model.

The model

It is not necessary to use the robust sandwich form for the variance-covariance matrix of the estimator VCE of the ML estimation, because data are independent and c.d.f. is correctly specified, according to Cameron and Trivedi (2005). It is observed that the value of the coefficients and standard error, presented in Table 2 and 3, no significantly differs. In Table 2 we have the logit model without robust option, and in Table 3 the model with robust option.

Table 2

Logit model without robust option

Number of obs = 60	Prob > chi2 = 0.0000
LR chi2(3) = 26.25	Pseudo R2 = 0.3156

	Dn	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
EBITDA	.0024077	.0016987	1.42	0.156		-.0009217 .005737
ROI	-.0609919	.0442086	-1.38	0.168		-.1476391 .0256554
CR	.0605788	.0183942	3.29	0.001		.0245269 .0966307
_cons	-5.348501	1.589868	-3.36	0.001		-8.464585 -2.232418

Table 3

Logit model with robust option

Number of obs = 60	Prob > chi2 = 0.0000
Wald chi2(3) = 23.40	Pseudo R2 = 0.3156

	Dn	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
EBITDA	.0024077	.0014584	1.65	0.099		-.0004508 .0052661
ROI	-.0609919	.0387372	-1.57	0.115		-.1369153 .0149316
CR	.0605788	.0135059	4.49	0.000		.0341077 .0870499
_cons	-5.348501	1.167233	-4.58	0.000		-7.636236 -3.060767

The impact of the robust option is insignificant. The p-value for EBITDA and ROI is greater than 0.05 level. In order to identify nonlinear relationships in model, we adopt the approach found in Dakovic et al. (2010) and review the linearity assumptions using generalized additive (GA). The (GA) approach indicates nonlinear relationships for the variables EBITDA and ROI, as shown in the "Gain" column of Table 4. The values in the "Gain" column suggested me to use a polynomial of degree 2 to model the influence of EBITDA and ROI on the probability of granting a loan.

Table 4

Generalized Additive Model with family binom, link logit

	Dn	df	Lin. Coef.	Std. Err.	z	Gain
EBITDA	3.713		.0335077	.0006557	51.100	86.244
ROI	1.953		-1.310497	.0164341	-79.743	-13.705
CR	2.100		.5207825	.0035592	146.322	2.539
_cons	1		3.26306	.147225	22.164	.

Total gain (nonlinearity chisquare) = 75.078 (4.766 df), P = 0.0000

These nonlinearities are modeled parametrically by incorporating a quadratic term for EBITDA and ROI. We improve the model testing the presence of interaction effects between

variables. Two variables are created: ROI2=ROI*ROI and EBITDA2=EBITDA*EBITDA, and they are included in the model. We obtain:

Table 5

The presence of interaction effect

Number of obs = 60	Prob > chi2 = 0.0007
Wald chi2(4) = 19.35	Pseudo R2 = 0.4419

	Dn	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]
EBITDA		.0235526	.0078459	3.00	0.003	.0081749 .0389302
CR		.0813545	.0193404	4.21	0.000	.043448 .119261
EBITDA2		-.0000266	8.95e-06	-2.97	0.003	-.0000441 -9.03e-06
ROI2		.0079695	.0024876	3.20	0.001	.0030938 .0128451
_cons		-9.027901	2.177734	-4.15	0.000	-13.29618 -4.75962

All regressors are statistically significantly different from zero at the 0.05 level. The null hypothesis that the coefficients of EBITDA, CR, EBITDA2, ROI2 except constant are zero is rejected at the 0.05 level. This is confirmed by Wald test. For the fitted logit model the McFadden R^2 is 0.4419. The Table 5 shows a good model about the impact of EBITDA, CR and ROI on Pr(Dn = 1). However some measures of overall goodness of fit are necessary. The Hosmer-Lemeshow specification test, according to Hosmer et al. (2013), evaluates the goodness of fit by comparing the sample frequency of the dependent variable with the fitted probability of granting a loan within subgroups of observations. We perform this test, setting the number of groups to four. The results of the test are presented in the following table. The outcome of Table 6 indicates a good specification, because the p-value is 0,6852.

Table 6

The goodness-of-fit Hosmer-Lemeshow test with 4 groups

Logistic model for Dn, goodness-of-fit test

number of observations	=	60
number of groups	=	4
Hosmer-Lemeshow chi2(2)	=	0.76
Prob > chi2	=	0.6852

To check if the same outcome occurs, we have used a larger number of groups (ten groups) to perform the test. Again the test indicates a good specification, this time with a greater p-value (0.9807). The preceding measure is based on the fitted probability of having access to credit. One measure of goodness of fit is the percentage of correctly classified observations. For the fitted logit model, we obtain the Table 7.

Table 7

The percentage of correctly classified observations

Logistic model for Dn

----- True -----			
Classified	D	~D	Total
+	25	4	29
-	5	26	31

Total	30	30	60
Classified + if predicted $\Pr(D) \geq .5$			
True D defined as $D_n \neq 0$			
Correctly classified	85.00%		

The Table 7 compares fitted with actual values. The percentage of correctly specified values is 85.00. In this table, 4 observations are misclassified as 1 (access to credit) when the correct classification is 0 (no access to credit), and 5 values are misclassified as 0 (no access to credit) when the correct value is 1 (access to credit). The remaining observations (25 + 26) are correctly classified.

Marginal effects and prediction

More discussion about the interpretation of coefficients, marginal effects and predictions are offered in the following. The logit model, presented in Table 5, is linear in CR and non-linear in EBITDA si ROI and provide information in terms of satisfying the sign of coefficient and the significance of the coefficient of variables. The interpretation of coefficients of variables and marginal effects are more difficult to interpret. The sign for the coefficient CR is positive as it is expected in the model and it is statistically significant. The sign and values for the coefficients EBITDA, EBITDA2 and ROI2 are necessary for the calculation of marginal effects. The coefficients of all variables are statistically significant. The model shows the significant impact of EBITDA, CR and ROI on $\Pr(D_n = 1)$.

Based on McFadden R^2 , which provides the explanatory power of the variables, the model provides a significant explanatory power (44.19 percent). The remaining 56 percent is attributed to other factors with direct impact on the credit decision: the market, the shareholders, the executive management and the business.

The marginal effect at the mean (MEM) measures the effect on $\Pr(D_n = 1)$, when the regressor changes, evaluated at $X = \bar{X}$. We obtain the Table 8.

Table 8

Marginal effects at mean after logit

$y = \Pr(D_n) \text{ (predict)} = .70234761$

variable	dy/dx	Std. Err.	z	P> z	[95% C.I.]	X
EBITDA	.0049238	.00175	2.82	0.005	.001502 .008346	168.317
CR	.0170076	.00368	4.62	0.000	.009794 .024222	88.9383
EBITDA2	-5.55e-06	.00000	-2.81	0.005	-9.4e-06 -1.7e-06	88588.3
ROI2	.0016661	.00052	3.23	0.001	.000656 .002676	130.527

The impact of current ratio on $\Pr(D_n = 1)$ is 1.7 percent. According to Cameron and Trivedi (2009) some additional calculations are needed to determine marginal effects of EBITDA and ROI. The regressors EBITDA and ROI appear in model as polynomials. The average marginal effect (AME) for ROI can be computed by calculating $2 * _b[ROI2] * ROI$, where $_b[ROI2]$ is the coefficient of ROI2, for each observation and averaging. The average marginal effect for return on investment is 3.1 percent. The AME for EBITDA can be computed by calculating $_b[EBITDA] + 2 * _b[EBITDA2] * EBITDA$, where $_b[EBITDA]$ is the coefficient of EBITDA and $_b[EBITDA2]$ is

the coefficient for EBITDA2, for each observations and averaging. The AME for EBITDA is 0.0064, representing the impact on $\Pr(Dn = 1)$ for a growth of 1000 monetary units. The sign of marginal effect for each of the three variables is positive, they are expected. The marginal effect of model's predictive power on $\Pr(Dn = 1)$ is 70.23 percent. We first summarize, in the table 9, the fitted probability from the model. The prediction is stored as the variable Plogit. Table 9 shows the mean and standard deviation, for Dn and Plogit variables.

Table 9

The mean and standard deviation for Dn and Plogit variables

Variable	Obs	Mean	Std. Dev.	Min	Max
Dn	60	.5	.5042195	0	1
Plogit	60	.5	.3544271	.0038891	1

The results identify the factors that are significantly related to the financial distress, encountered by small and medium enterprises in Romania, in order to access credit.

Conclusions

Considering a lot of relevant papers of scientific literature, we have developed a logit model which estimates the likelihood of granting a loan. This paper proposes a study based on a random sample of small size of small and medium enterprises from Romania that want loans. The model provides a high prediction power. The variables of model - EBITDA, ROI, and CR – correctly predicts 25 out of total 29, i.e. 85 percent. From the three predictors, the return on investment has the highest marginal impact on $\Pr(Dn = 1)$, its marginal impact is 3.1 percent. The impact of current ratio on $\Pr(Dn = 1)$ is 1.7 percent. The average marginal effect for EBITDA is 0.0064, and represent the impact on $\Pr(Dn = 1)$ for a growth of 1000 monetary units. The marginal effect of model's predictive power on $\Pr(Dn = 1)$ is 70.23 percent. It should be emphasised that relying only on quantitative information (economic and financial data) it is possible to design a scoring system with good predictability power. The variables of the model provide a significant explanatory power (44.19 percent). The identification of significant non-linear effects improves the model's explanatory power, and hence the quality of the model. The probability of granting a loan estimated in the assessment phase of analysis of a potential loan, helps to strengthen the banking prudence. The results obtained in this study are comparable to those obtained in: Samad (2012), Kukuk and Ronnberg (2013), Danila (2012), and Altman and Sabato (2005).

Consequently, this study identifies key determinants of the financial state of firm which influence the decision to grant a loan. The random sample is heterogeneous in relation to the activities of firms. This study requires the application separately for each field.

References

1. Abdou H., Pointon J., 2011. Credit Scoring, Statistical Techniques and Evaluation Criteria: A Review of the Literature, *Intelligent Systems in Accounting, Finance & Management*, 18 (2-3), pp. 59-88.
2. Altman E. I., Saunders A., 1998. Credit risk measurement: Developments over the last 20 years, *Journal of Banking and Finance*, 21, pp. 1721–1742.
3. Altman E. I., Sabato G., 2005. Modeling Credit Risk for SMEs: Evidence from US Market, Available at SSRN, <http://ssrn.com/abstract=872336>.
4. Basel II, 2006. International Convergence of Capital Measurement and Capital Standards: A Revised Framework – Comprehensive Version.

5. Berger A. N., Scott Frame W., 2005. Small Business Credit Scoring and Credit Availability, Working Paper 2005, p. 10.
6. Berger A., Barrera M., 2007. ACCION New Mexico, L. Parsons. ACCION USA, Joyce Klein. The Aspen Institute/FIELD, Credit Scoring for Microenterprise Lenders, Available at fieldus.org/projects/pdf/CreditScoring.pdf.
7. Cameron A. C., Trivedi P. K., 2005. *Microeconometrics: Methods and Applications*, Cambridge: Cambridge University Press, p. 1058.
8. Cameron A. C., Trivedi P. K., 2009. *Microeconometrics Using Stata*. A Stata Press Publication. StataCorp LP, College Station, Texas, p. 692.
9. Dakovic R., Czado C., Berg, D., 2010. Bankruptcy prediction in Norway: a comparison study, *Appl. Econ. Lett.* 17(17), pp. 1739-1746.
10. Danila O. A., 2012. Credit Risk Assessment under Basel Accords, *Theoretical and Applied Economics*, Vol. XIX, No. 3(568), pp. 77-90.
11. Fernandes J. E., 2005. Corporate Credit Risk Modeling: Quantitative Rating System And Probability Of Default Estimation, www.ratingexpress.net/content, SSRN working paper.
12. Gordy M. B., 2000. A comparative anatomy of credit risk models, *Journal of Banking and Finance*, 24, pp. 119–149.
13. Greene W., 1998. Sample selection in credit-scoring models, *Japan and the World Economy* 10, pp. 299-316.
14. Hastie T. J., Tibshirani R. J., 1995. *Generalized additive models*, 1st edn. Chapman & Hall, London.
15. Henley W. E. and Hand D. J., 1996. A k-nearest-neighbour classifier for assessing consumer credit risk, *The Statistician* 45 (1), pp. 77-95.
16. Hostmer D. W., Lemeshov S., Sturdivant R. X., 2013. *Applied logistic regression*, Third Edition, Jhon Wiley & Sons Inc., New York, p. 508.
17. Kukuk M., Ronnberg M., 2013. Corporate credit default models: a mixed logit approach, *Rev Quant Finan Acc* 40, pp. 467-483.
18. Lennox C., 1999. Identifying failing companies: a reevaluation of the logit, probit and DA approaches, *J Econ Bus* 51(4), pp. 347-364
19. Lin S. L., Chang Y. T., 2006. The credit-risk warning system of financial distress in Taiwan, *International Journal of Business and Strategy*, 7 (2), pp. 36–72.
20. Marques A. I., Garcia V., Sanchez J. S., 2013. A literature review on the application of evolutionary computing to credit scoring, *Journal of the Operational Research Society* 64, pp. 1384-1399.
21. McFadden D., 1974. Conditional Logit Analysis of Qualitative Choise Behavior, *Frontiers in Econometrics*, edited by P. Zarembka, New York: Academic Press, pp. 105-142.
22. Medina R. S., Alfonso M. D. O., Cueto M. J. V., 2010. Internal models (IRB) in Basel II: an approach to determining the probability of default, *Banks and Bank Systems*, Vol. 5, Issue 2, pp. 222-229.
23. Ohlson J. A., 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research*, Vol. 18/1, pp. 109-131.
24. Platt H. D., Platt M. B., 2006. Understanding Differances Between Financial Distress and Bankruptcy, *Review of Applied Economics*, Vol. 2, No. 2, pp. 141-157.
25. Sabato G., 2008. Credit Risk Scoring Models. Electronic copy available at: <http://ssrn.com/abstract=1546347>.
26. Samad A., 2012. Credit Risk Determinants of Bank Failure: Evidence from US Bank Failure, *International Business Research*, Vol. 5, No. 9, pp. 11-15.

27. Zmijewski M.E., 1984. Methodological issues related to the estimation of financial distress prediction models, *Journal of Accounting Research, Supplement to Vol 22*, pp. 59-82.