ISSN: 2278-3369
International Journal of Advances in Management and Economics
Available online at www.managementjournal.info

RESEARCH ARTICLE

Credit Risk Modeling using Multiple Regressions

Maria Dimitriu[1], Ioana-Aurelia Oprea[2]*, Marian-Albert Scrieciu[1]

[1]Academy of Economic Studies, Bucharest, Romania..

[2]*BCR, Erste Group*

Corresponding Author: E-mail:** ***ioanaaurelia.oprea@bcr.ro

Abstract

In classical theory, the risk is limited to mathematical expectation of losses that can occur when choosing one of the possible variants. For banks, risk is represented as losses arising from the completion of one or another decision. Bank risk is a phenomenon that occurs during the activity of banking operations and that causes negative effects for those activities: deterioration of business or record bank losses affecting functionality. It can be caused by internal or external causes, generated by the competitive environment.

Keywords: *Banking system, Credit risk, Multiple regression.*
JEL Classification: *G21, G32*

Introduction

The experience of developed countries underlines the need to develop portfolios of information on all bank customers and a database for information processing as a main way to strengthen the bank's position in its relations with customers, regardless of their size, and as a weapon the prevention and avoidance of credit risk [1].

By using credit derivatives, banks keep the loan on their balance sheet. Transferring credit risks using derivatives means risks that credit risk transfers with loan sales or securitizations do not have. Banks using these derivatives have to bear associated counterparty, operational, and legal risks [2].

In studies of the economy or companies, using data and statistical methods for analysis of information is inevitable, therefore perception objectively and effectively of economic reality recommend the use of quantitative analysis methods. Methods exploiting the information collected via the econometric models are based on specific concepts of logic and mathematics. Using the scoring, the lender can appreciate quickly, objectively and consistently the previous loans, and can calculate the probability that the loan will be repaid according to the contract. [3].Econometric tools available in the investigation of an economic phenomenon are considering the following: [4].

- Identifying features of the phenomenon studied. The first step to be taken in this regard is the choice of economic theories that will guide the research of this phenomenon: it is defined the quantitative relationships between different statistical variables used to characterize this phenomenon.
- Testing the statistical assumptions on some specific aspects of the phenomenon studied.
- Making predictions for a specific time horizon. Companies are considering these forecasts to anticipate and apply a correction to future developments of the economic phenomenon.

Using multiple regression can determine the impact of several independent variables on certain variables (called dependent variable). The general form of multiple regression equation is:

$Y_t = a_0 + a_1 * X1t + a_2 * X_{2t} + a_3 X_{3t} + ... + a_k * X_{kt} + \varepsilon_t,$

where:

$t = 1, 2, ..., n$ – observations of the sample

Y_t - observation t of the dependent variable

X_j - independent variables, explanatory, $j = 1, 2, 3, ..., k$

X_{jt} - observation t of independent variables X_j

a_0 - constant, free term of equation

$a_{1, ... K}$ - coefficients of independent variables

ε_t - error term of equation.

The coefficient of independent variable reflects how dependent variable Y_t changes when the independent variable, X_j, changes by one unit, while the other independent variables remain

constant. If the dependent variable and independent variables are specified in natural logarithms, the coefficients of independent variables can be interpreted as elasticities. Thus, these coefficients will show the percent change of the dependent variable if the independent variable changes by 1 percent.
For the model determined by multiple linear regression equation to be valid, it must meet the following assumptions:

Hypothesis 1: residual variables are random variables with average zero, namely: E (ε_t) = 0 for t = 1, 2, ..., n

Hypothesis 2: residual variables are not correlated: COV (ε_i, ε_j) = 0, for $i \neq j$

Hypothesis 3: *The* residual variance is unchanging variable, homoscedasticity property: var (ε_t) = σ_ε^2)

Hypothesis 4: residual variables are not correlated with explanatory variables: COV (X, ε_t) = 0

Hypothesis 5: The regression model must be correctly specified: the explanatory variables are properly chosen, the regression formula is correctly specified, and not least, the residual term has the correct form.

Hypothesis 6: Explanatory variables are linearly independent

Hypothesis 7: *The* residual variable is distributed as a normal distribution: $\varepsilon \sim N(0, \sigma_\varepsilon^2)$

Defining Variables for Multiple Regression Model

The main activity of commercial banks is lending activity, so a special importance is given to credit risk management. [5].

This study has the starting point the economic and financial analysis of 337 companies for 3 consecutive years, taking into account 12 quantitative factors. The econometric model proposed considers that the probability of default of loans by a bank is directly dependent on the 12 factors analyzed. Each customer's probability of default is determined by the category of risk to which it belongs. So we divided the customers into five classes of risk. Depending on the values of the indicators analyzed each firm is classified into the following classes of risk, thus giving a value for the probability of default:

- Risk class a with a probability of default 0.3%
- Risk class b with a probability of default a 0.5%
- Risk class c with a probability of default 1.5%
- Risk class d with a probability of default 5%
- Risk class e with probability of default 10%

In this study the dependent variable is the probability of default (PD) and independent variables are: evolution of total turnover (ca), commercial return or profit margin (margin), return on equity (ROE) , return on investment (ROI), net cash flow (cf), intensity of investments (imob), investment ratio (inv), equity ratio (KPR), quick liquidity (lich), overall net indebtedness (indat), average accounts receivable (cl) and average accounts payable (fz).

The 12 endogenous variables were calculated by following formula:
Turnover rate increase = $(CA_1-CA_0)/CA_0$ where,
CA_1 – turnover for year 1
CA_0 – turnover for year 0
Profit margin = Profit / Turnover

$$ROE = \frac{net\ profit}{equity}$$

ROI = Net profit / investment
CF = Net Profit + Depreciation - fixed assets expenses - working capital increase
Share of current assets to total assets = current assets / total assets
Investment rate increase = (I1-I0) / I0, where
I1 – investment for year 1
I_0 – investment for year 0
Degree of financing total assets from equity = total assets / equity

$$Quick\ ratio = \frac{(current\ assets - inventories)}{current\ liabilitie\ s}$$

$$Degree\ of\ debt = \frac{total\ debts}{net\ asset\ value}$$

The average receipt for customer = Number of days * Customer / Sales

The average payment to suppliers = Number of days * Supplier / Cost of goods sold

Therefore the model equation is as follows:

PD = $a_0 + a_1$*ca + a_2*margin + a_3*roe + a_4*roi + a_5*cf + a_6*imo + a_7*inv + a_8*kpr + a_9*lich + a_{10}*indat + a_{11}*cl + a_{12}*fz

It is worth mentioning that this case study involves the definition of three different models, one for each year, and at the end of the analysis the best model is recommended.

Determination of the Optimal Regression Model

First we made a descriptive analysis of data, analysis that helped us to identify if data series are stationary. To check the stationarity of variables we used the Phillip Perron test and noticed that all independent variables are stationary, hence their fluctuations around a trend occurring in parallel with the abscissa, and the probability shown by Phillip Perron test is equal to 0. As example we test for the probability of default.

Phillips-Perron Unit Root Test on PD

Null Hypothesis: PD has a unit root
Exogenous: Constant
Bandwidth: 5 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-18.87996	0.0000
Test critical values:	1% level	-3.449620	
	5% level	-2.869927	
	10% level	-2.571307	

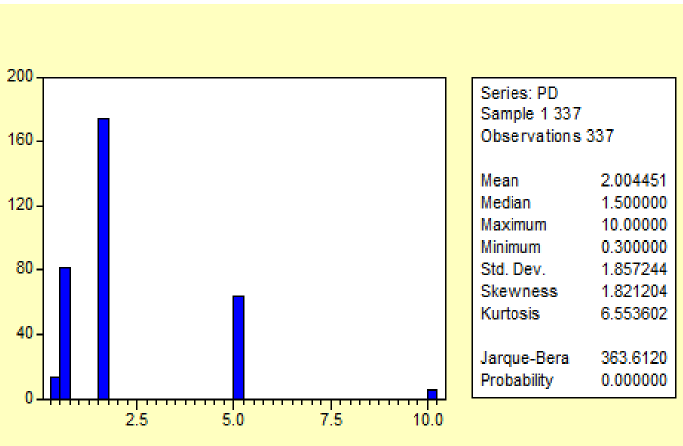
*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	3.444792
HAC corrected variance (Bartlett kernel)	3.411605

Source: own calculations

Fig. 1: Test PD Phillips-perron

Finally Jarque-Bera test confirms that the variable is normally distributed, as associated test probability 0.00. The assumption of normality of the variable can be checked and rejected also by specific tests Eviews: *simple hypothesis test* and *quantile-quantile graph*.



Source: own calculations

Fig. 2: Histogram for probability of default

If variables were not stationary it was necessary to logarithms as: **l_variable $_i$ = log (variable $_i$)** The next step was to determine the coefficients of multiple regression model using the least square method that shows the following values for the variables.

Dependent Variable: PD
Method: Least Squares
Date: 04/10/11 Time: 00:07
Sample: 1 337
Included observations: 337

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CA	0.001154	0.000744	1.551000	0.1219
MARJA	0.003238	0.001987	1.629710	0.1041
ROE	-0.014866	0.002668	-5.571055	0.0000
ROI	-0.030644	0.012140	-2.524275	0.0121
CF	-5.40E-06	6.81E-06	-0.793530	0.4280
IMOB	0.001496	0.004565	0.327758	0.7433
INV	0.000446	0.000481	0.928838	0.3537
KPR	-0.018978	0.004087	-4.643343	0.0000
LICH	-0.001937	0.001830	-1.058410	0.2907
INDAT	0.001230	0.000563	2.184005	0.0297
CL	0.001382	0.001822	0.758683	0.4486
FZ	6.15E-05	0.001205	0.051004	0.9594
C	3.055120	0.297092	10.28340	0.0000

R-squared	0.340480	Mean dependent var	2.004451
Adjusted R-squared	0.316054	S.D. dependent var	1.857244
S.E. of regression	1.535960	Akaike info criterion	3.734000
Sum squared resid	764.3722	Schwarz criterion	3.881363
Log likelihood	-616.1791	F-statistic	13.93889
Durbin-Watson stat	2.169237	Prob(F-statistic)	0.000000

Source: own calculations

Fig. 3: Estimation parameters for 2008

Thus the first equation using Eviews software for 2008, has the form:

PD = 0.00115354791*CA + 0.0032383296*MARGIN - 0.01486619484*ROE-0.03064377156*ROI- 5.401135453e-06*CF + 0.001496178241*IMOB + 0.0004463429102*INV - 0.01897753683*KPR - 0.001936911715*LICH + 0.00122992785*INDAT + 0.001381941557*CL + 6.148389557e-05*FZ + 3.055119835 (1)

As previously mentioned the proposed model to be used in lending decisions will be chosen as the best model of the three analyzed.

Regression equation for Model 2 is as follows:

PD = -0.001838769405*CA-0.05770658157*MARGIN - 0.0008878188804*ROE- 0.009553228523*ROI+ 1.39035012e-05*CF+ 0.008825282231*IMOB- 0.0001422659158*INV- 0.03281036515*KPR+ 0.0003608745752*LICH + 0.008793179356*INDAT - 0.0002003597285*CL + 0.0009775815716*FZ + 2.750100729 (2)

Analyzed indicators have values displayed in the table below:

For Model 3 has produced the following values:

With form equation:

PD = -0.005379910878*CA - 0.05169358257*MARGIN -0.004833026573*ROE - 0.02375204639*ROI - 1.046135381e-05*CF + 0.01552840358*IMOB - 0.0001930041028*INV - 0.0417891334*KPR + 0.0008371068961*LICH + 0.0003458981285*INDAT + 0.00274293772*CL + 0.002757301955*FZ + 3.074346603 (3)

Dependent Variable: PD
Method: Least Squares
Date: 04/10/11 Time: 00:19
Sample: 1 337
Included observations: 336
Excluded observations: 1

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CA	-0.001839	0.000514	-3.574843	0.0004
MARJA	-0.057707	0.014038	-4.110787	0.0001
ROE	-0.000888	0.000866	-1.024777	0.3062
ROI	-0.009553	0.008109	-1.178031	0.2397
CF	1.39E-05	1.02E-05	1.362160	0.1741
IMOB	0.008825	0.005806	1.520040	0.1295
INV	-0.000142	0.000656	-0.217034	0.8283
KPR	-0.032810	0.005744	-5.711913	0.0000
LICH	0.000361	0.002512	0.143661	0.8859
INDAT	0.008793	0.002359	3.727100	0.0002
CL	-0.000200	0.002416	-0.082931	0.9340
FZ	0.000978	0.001434	0.681714	0.4959
C	2.750101	0.364019	7.554833	0.0000

R-squared	0.311878	Mean dependent var	2.374107
Adjusted R-squared	0.286313	S.D. dependent var	2.285271
S.E. of regression	1.930596	Akaike info criterion	4.191457
Sum squared resid	1203.886	Schwarz criterion	4.339143
Log likelihood	-691.1648	F-statistic	12.19946
Durbin-Watson stat	1.939870	Prob(F-statistic)	0.000000

Fig. 4: Estimation parameters for 2009

Dependent Variable: PD
Method: Least Squares
Date: 04/10/11 Time: 00:28
Sample: 1 337
Included observations: 337

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CA	-0.005380	0.002050	-2.624536	0.0091
MARJA	-0.051694	0.012351	-4.185431	0.0000
ROE	-0.004833	0.001765	-2.738203	0.0065
ROI	-0.023752	0.014775	-1.607588	0.1089
CF	-1.05E-05	1.08E-05	-0.964641	0.3354
IMOB	0.015528	0.005158	3.010331	0.0028
INV	-0.000193	0.000707	-0.273109	0.7849
KPR	-0.041789	0.006225	-6.713513	0.0000
LICH	0.000837	0.002386	0.350813	0.7260
INDAT	0.000346	0.000425	0.813855	0.4163
CL	0.002743	0.001860	1.474826	0.1412
FZ	0.002757	0.001238	2.227117	0.0266
C	3.074347	0.359246	8.557777	0.0000

R-squared	0.403210	Mean dependent var	2.676261
Adjusted R-squared	0.381107	S.D. dependent var	2.516073
S.E. of regression	1.979389	Akaike info criterion	4.241265
Sum squared resid	1269.426	Schwarz criterion	4.388627
Log likelihood	-701.6532	F-statistic	18.24206
Durbin-Watson stat	2.051530	Prob(F-statistic)	0.000000

Source: own calculations

Fig. 5: Estimation parameters for 2010

As a criterion for choosing between the three competing models we used Akaike and Schwartz tests. According to them the best performing model is one that has the minimum value for one of the two indicators, because these two indicators decrease their values with decreasing adjustment errors. At the same time the quality of estimation is improved by increasing the size of data series used to estimate parameters. [6]

Analyzing the output of each regression we concluded that the optimal model is:

PD = 0.00115354791*CA + 0.0032383296*MARGIN - 0.01486619484*ROE - 0.03064377156*ROI - 5.401135453e-06*CF + 0.001496178241*IMOB + 0.0004463429102*INV - 0.01897753683*KPR - 0.001936911715*LICH + 0.00122992785*INDAT + 0.001381941557*CL + 6.148389557e-05*FZ + 3.055119835 (4)

For this equation, Figure 6, the test has the lowest Akaike value, respectively 3.734000 compared to 4.191457 for the corresponding Model 2, and 4.241265 for the Model 3. At the same time Schwartz criterion displays the minimum value also for Model 1: 3.881363, compared to 4.339143 for Model 2 and respectively 4.388627 for Model 3.

Dependent Variable: PD
Method: Least Squares
Date: 04/10/11 Time: 00:07
Sample: 1 337
Included observations: 337

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CA	0.001154	0.000744	1.551000	0.1219
MARJA	0.003238	0.001987	1.629710	0.1041
ROE	-0.014866	0.002668	-5.571055	0.0000
ROI	-0.030644	0.012140	-2.524275	0.0121
CF	-5.40E-06	6.81E-06	-0.793530	0.4280
IMOB	0.001496	0.004565	0.327758	0.7433
INV	0.000446	0.000481	0.928838	0.3537
KPR	-0.018978	0.004087	-4.643343	0.0000
LICH	-0.001937	0.001830	-1.058410	0.2907
INDAT	0.001230	0.000563	2.184005	0.0297
CL	0.001382	0.001822	0.758683	0.4486
FZ	6.15E-05	0.001205	0.051004	0.9594
C	3.055120	0.297092	10.28340	0.0000

R-squared	0.340480	Mean dependent var	2.004451
Adjusted R-squared	0.316054	S.D. dependent var	1.857244
S.E. of regression	1.535960	Akaike info criterion	3.734000
Sum squared resid	764.3722	Schwarz criterion	3.881363
Log likelihood	-616.1791	F-statistic	13.93889
Durbin-Watson stat	2.169237	Prob(F-statistic)	0.000000

Source: Own calculations

Fig. 6: Estimated parameters for 2008

The model shows satisfactory statistical results. *F statistic* value and its associated probability value 0.000, suggesting the model is correct. Dependence of exogenous variable regression with regression factors is given by the coefficient of determination Adjusted R-squared, noted R2. Mathematically it is calculated as:

$R^2 = \frac{SPR}{SPT}' = 1 - \frac{SPE}{SPT}$, where

- $SPT = \sum_{t=1}^{n} (y_t - \bar{y})^2$ quantifies dispersion of endogenous variable below the action of endogenous factors in the model, and factors unregistered

- $SPR = \sum_{t=1}^{n} (\hat{y}_t - \bar{y})^2$ measures the influence of exogenous variables in the total endogenous variable of series size.

- $SPE = \sum_{t=1}^{n} (\varepsilon_t)^2$ is the sum of squared errors adjustment and measure the influence of unregistered factors of multiple regression model. [4]

The value of coefficient of determination must belong to interval [6] and increases with increasing number of endogenous variables used to define the regression model. If the indicator is different from 0, then the endogenous variable is explained, largely, by endogenous variables.Value of the coefficient R-squared indicates that over 34% of the variability of the probability of default is explained by the evolution of total turnover (ca), commercial return or profit margin (margin), return on equity (ROE) , return on investment (ROI), net cash flow (cf), intensity of investments (imob), investment ratio (inv), equity ratio (KPR), quick liquidity (lich), overall net indebtedness (indat), average accounts receivable (cl) and average accounts payable (fz); the rest is the contribution of factors that are not included in the analysis. In this category we include staff productivity, employment, the average stationary stocks and not in the least the represented qualitative factors: credit history, quality of ownership, quality management, securities received and market coverage.

Durbin Watson statistic (DW) is a statistical test which tests the serial correlation of errors. If the errors are not correlated, the value of DW will be around 2. Value 2.17 of Durbin-Watson test in Fig. 7, suggests the autocorrelation of the first order residues, which has a negative impact on model validation, even if the value of F statistics and the associated probability value is 0.000 suggesting the correct model specification.

Dependent Variable: PD
Method: Least Squares
Date: 04/10/11 Time: 00:07
Sample: 1 337
Included observations: 337

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CA	0.001154	0.000744	1.551000	0.1219
MARJA	0.003238	0.001987	1.629710	0.1041
ROE	-0.014866	0.002668	-5.571055	0.0000
ROI	-0.030644	0.012140	-2.524275	0.0121
CF	-5.40E-06	6.81E-06	-0.793530	0.4280
IMOB	0.001496	0.004565	0.327758	0.7433
INV	0.000446	0.000481	0.928838	0.3537
KPR	-0.018978	0.004087	-4.643343	0.0000
LICH	-0.001937	0.001830	-1.058410	0.2907
INDAT	0.001230	0.000563	2.184005	0.0297
CL	0.001382	0.001822	0.758683	0.4486
FZ	6.15E-05	0.001205	0.051004	0.9594
C	3.055120	0.297092	10.28340	0.0000

R-squared	0.340480	Mean dependent var	2.004451
Adjusted R-squared	0.316054	S.D. dependent var	1.857244
S.E. of regression	1.535960	Akaike info criterion	3.734000
Sum squared resid	764.3722	Schwarz criterion	3.881363
Log likelihood	-616.1791	F-statistic	13.93889
Durbin-Watson stat	2.169237	Prob(F-statistic)	0.000000

Source: Own calculations

Fig. 7: Autocorrelation of residues

Hypothesis Testing for Multiple Regression Model

So far, the results of econometric regression analysis show that the model can be validated. To make sure that the estimated parameters are effective, homebound and linear, the multifactorial model implies complying with the assumptions outlined above, where we have defined multiple regression, so that the next step is hypothesis testing.

Autocorrelation residues (errors)

The value of *Durbin Watson* test (DW) different from 2 shows that it is possible residue autocorrelation, which must be confirmed by the Breusch-Godfrey test Figure 8. Statistical value of 1.69 and R-squared of 3.51 suggests rejection of the null hypothesis, i.e. the lack of residual values correlation.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.696204	Probability	0.185013
Obs*R-squared	3.513424	Probability	0.172611

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 04/10/11 Time: 01:41
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CA	1.05E-05	0.000742	0.014111	0.9888
MARJA	0.000295	0.001992	0.147915	0.8825
ROE	5.25E-05	0.002673	0.019653	0.9843
ROI	-0.000221	0.012151	-0.018191	0.9855
CF	9.76E-07	6.82E-06	0.143212	0.8862
IMOB	-0.000142	0.004562	-0.031213	0.9751
INV	-4.70E-05	0.000483	-0.097280	0.9226
KPR	-0.000150	0.004083	-0.036755	0.9707
LICH	0.000121	0.001843	0.065782	0.9476
INDAT	9.33E-05	0.000565	0.165196	0.8689
CL	-7.78E-05	0.001818	-0.042772	0.9659
FZ	3.70E-05	0.001203	0.030785	0.9755
C	-5.52E-06	0.297584	-1.85E-05	1.0000
RESID(-1)	-0.093303	0.056534	-1.650388	0.0998
RESID(-2)	-0.053328	0.056848	-0.938072	0.3489

R-squared	0.010426	Mean dependent var	-4.82E-16
Adjusted R-squared	-0.032599	S.D. dependent var	1.508283
S.E. of regression	1.532670	Akaike info criterion	3.735389
Sum squared resid	756.4032	Schwarz criterion	3.905423
Log likelihood	-614.4131	F-statistic	0.242315
Durbin-Watson stat	1.999304	Prob(F-statistic)	0.998008

Source: own calculations

Fig. 8: *Breusch-godfrey* test

Heteroscedasticity

According to this hypothesis, variant of residues must be constant; otherwise the estimators are no longer effective. As above, we proceed to make a statistical test. F-statistic value of 3.56 and R-squared and their associated probability of 0.00 accept null hypothesis, so there is no heteroscedasticity.

Normality of Residues

The verification for normality residues is done by residues histogram and *Jarque-Bera* test, Figure 10. The test measures the difference between the coefficient of asymmetry and kurtosis for the distribution analyzed with the normal distribution. In our study, coefficients of asymmetry *(skewness)* and flatness (kurtosis) are significantly different from 0, respectively 3, and the histogram is not symmetric. Kurtosis indicator has a value greater than 3, so we have a leptokurtosis distribution. [6].

White Heteroskedasticity Test:

F-statistic	3.566977	Probability	0.000000
Obs*R-squared	72.55827	Probability	0.000001

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/10/11 Time: 01:57
Sample: 1 337
Included observations: 337

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.281415	1.362488	3.142350	0.0018
CA	0.002627	0.005953	0.441206	0.6594
CA^2	-1.12E-06	7.61E-06	-0.147867	0.8825
MARJA	-0.066748	0.038811	-1.719794	0.0865
MARJA^2	8.47E-05	5.62E-05	1.508980	0.1323
ROE	-0.032834	0.008395	-3.911281	0.0001
ROE^2	0.000113	4.39E-05	2.564908	0.0108
ROI	-0.039508	0.066877	-0.590758	0.5551
ROI^2	0.002595	0.001413	1.836350	0.0673
CF	-4.74E-05	5.92E-05	-0.800601	0.4240
CF^2	2.27E-10	3.05E-10	0.744304	0.4573
IMOB	0.024289	0.047323	0.513255	0.6081
IMOB^2	-0.000128	0.000488	-0.262250	0.7933
INV	-0.006685	0.003594	-1.860356	0.0638
INV^2	3.79E-06	2.90E-06	1.307884	0.1919
KPR	-0.087370	0.023451	-3.725696	0.0002
KPR^2	0.000485	0.000117	4.127209	0.0000
LICH	0.000170	0.011187	0.015225	0.9879
LICH^2	1.15E-06	1.79E-05	0.064115	0.9489
INDAT	0.006951	0.006186	1.123705	0.2620
INDAT^2	-9.04E-06	5.52E-06	-1.637964	0.1024
CL	0.007443	0.010663	0.698063	0.4857
CL^2	-2.01E-06	2.64E-05	-0.076187	0.9393
FZ	-0.004474	0.007827	-0.571617	0.5680
FZ^2	6.13E-06	1.64E-05	0.373710	0.7089

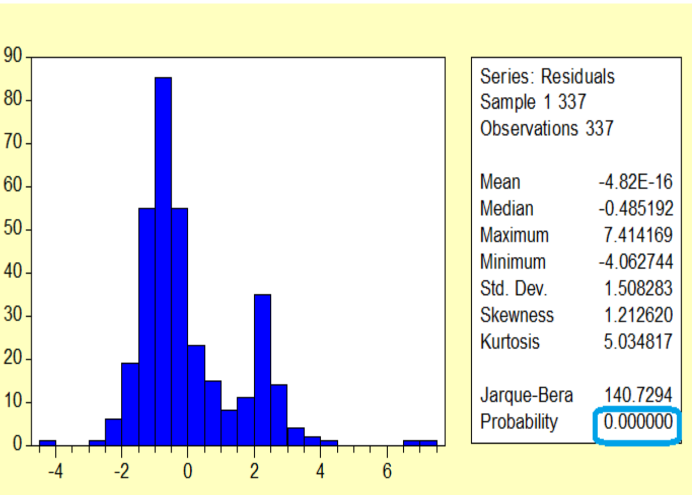
R-squared	0.215306	Mean dependent var	2.268167
Adjusted R-squared	0.154945	S.D. dependent var	4.562808
S.E. of regression	4.194446	Akaike info criterion	5.776688
Sum squared resid	5489.133	Schwarz criterion	6.060077
Log likelihood	-948.3719	F-statistic	3.566977
Durbin-Watson stat	2.081696	Prob(F-statistic)	0.000000

Source: own calculations

Fig. 9: White test

The value 3.56 for F-statistic and R-squared and also their associated probability of 0.00 accept null hypothesis, so there is no heteroscedasticity

These results induce the failure of assumption of normality of residues. Also, *Jarque-Bera* test suggests that errors are not normally distributed, since the probability is 0.00.



Source: own calculations

Fig. 10: *Jarque-bera* test

Test parameters can be achieved by Wald test, Fig. 11. The test shows that the probability of the coefficient to be 0 is very small.

Wald Test:
Equation: EQ1

Test Statistic	Value	df	Probability
F-statistic	13.93889	(12, 324)	0.0000
Chi-square	167.2667	12	0.0000

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(1)	0.001154	0.000744
C(2)	0.003238	0.001987
C(3)	-0.014866	0.002668
C(4)	-0.030644	0.012140
C(5)	-5.40E-06	6.81E-06
C(6)	0.001496	0.004565
C(7)	0.000446	0.000481
C(8)	-0.018978	0.004087
C(9)	-0.001937	0.001830
C(10)	0.001230	0.000563
C(11)	0.001382	0.001822
C(12)	6.15E-05	0.001205

Restrictions are linear in coefficients.

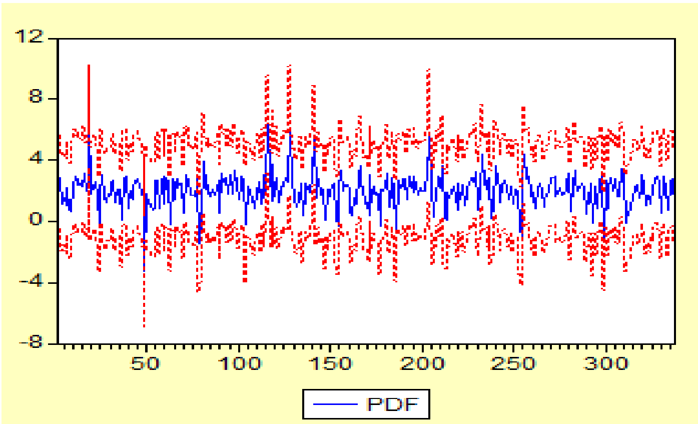
Source: own calculations

Fig. 11: Wald test

The results highlight the acceptance of the null hypothesis of linear relationship between the regression model parameters.

Predictions

To test the quality of the model again, we wanted to do an estimate based on recorded data. We see that the forecast model fails to mimic almost all series of actual values, which is another added advantage to validate the model.



Source: own calculations

Fig. 12: Predictions

In the current context of crisis, access to any resource, including the financial, is more difficult, and more expensive. Therefore, the eligibility of any customer of the bank is carefully assessed and determining the risk profile and a proper management of credit risk is absolutely necessary.

Conclusions

The main problem in creating an external credit assessment institution is to build a statistical model to quantify the probability of default in accordance with the requirements of Basel II. In terms of methodology, credit risk modeling includes:[7].

- Comprehensive evaluation of characteristics of the borrower and facility that he wishes to access;
- Meaningful differentiation of risk, namely granularity grading scale;
- Reasonable accuracy and consistency over time of estimates of quantitative credit risk

After analysis and testing, we concluded that there is a strong correlation between the probability of default and endogenous variables: evolution of total turnover (ca), commercial return or profit margin (margin), return on equity (ROE) , return on investment (ROI), net cash flow (cf), intensity of investments (imob), investment ratio (inv), equity ratio (KPR), quick liquidity (lich), overall net indebtedness (indat), average accounts receivable (cl) and average accounts payable (fz). Probability of Default - PD is determined, largely, by the evolution of the 12 indicators as well as other company specific factors. It is important to note the superiority of the factorial model credit risk management. This is confirmed firstly by the high value associated coefficient of determination, *R-squared = 34.05%,* and the tests based on information criteria.

This is consistent with modern theories that minimize credit risk, specific for loan portfolio and can be achieved through diversification. Because the model assumptions were checked and met, we validate the proposed model for analysis because of good results of statistical tests performed. The scoring system developed with multiple regression is an important element for the commercial banks for credit risk management. This tool for credit risk management provides information on the quality of bank debtor company and thus the bank can make a correct decision as regards the amount, the guarantees to be made, and not least: interest and cost of the loan. Over time it has been shown that the onset of financial crises was based on an inadequate management of credit risk, therefore should be noted that credit risk is the most important factor that influencing the performance of a banking company. Credit risk management should not be restricted only to solve problems arising after its manifestation but rather to prevent such crises.

References

1. Dimitriu M, Oprea IA (2009) Credit risk management in financial crisis, Review of International Comparative Management, 2:987-994.
2. Minton BA, Williamson RW (2005) How much do banks use credit derivatives to reduce risk?, National Bureau of Economic Research, Working Paper 11579, p. 12.
3. Dimitriu M, Oprea IA (2010) Modeling Credit Scoring, Metalurgia International, Special Issue nr. 5:62-67.
4. Tudorel A, Bourbonnais R (2008) Econometrie, Ed. Economica, Bucharest, 10-54.
5. Van Greuning H, Sonja Brajovic Bratanavic (2009) Analyzing banking risk –a framework for assessing corporate governance and financial risk management, Editura The World Bank, Editia a III-a, 87.
6. Codirlaşu A, Chidesciuc NA (2008) Econometrie aplicată utilizând EViews 5.1, Note de curs, Bucharest, p. 30 – 48.
7. CFO Research Services Boston (2010) The Impact of Basel III in Europe. A supplement to Capital Management in Banking: Senior Executives on Capital, Risk, and Strategy, Ernst & Young Bank Executive Series.