View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Electronic archive of Ternopil National Ivan Puluj Technical...

# LOGIT AND PROBIT MODEL FOR PREDICTION THE FINANCIAL HEALTH OF INSURANCE COMPANY

**Viktoriia Kremen[1], Maryna Brychko[2]**
[1] Kyiv National University of Technologies and Design, Kyiv, Ukraine, kremenviktori@gmail.com
[2] Sumy State University, Sumy, Ukraine, m.brychko@uabs.sumdu.edu.ua

**Abstract**
A model for prediction the financial health of insurance company has been developed. The existing leading world and Ukrainian experience of prediction the financial health of financial intermediaries − banks and insurance companies − has been investigated. It was shown that for this need one of the most commonly used methods is the different types of regression equations − linear, logit and probit. As a result of the financial crisis 2015 in Ukraine, some insurance companies have ceased their activities and were deprived of their licenses. Therefore, the financial statements of these companies have been investigated in order to develop a model for prediction their financial health. This analysis was determined 31 indicators of their activities for the groups of liquidity, solvency and financial stability, profitability, business activity, as well as for the group of specific indicators of early warning tests. The selection of normally operated insurance companies and those that were withdrawn in Ukraine was made based on open sources, since an open register of insurance companies were not created by the financial regulator. For the screening of predictors to be included in the model for prediction the financial health of insurance companies, a two-choice $F$ –test for variance, the Farrar-Glauber algorithm and the pair correlation method were applied. Three factors were included in the linear, logit and probit regression equations − gross return on equity, asset turnover, and indicator of changes in equity. The quality of the linear regression equation was checked using the coefficient of determination, $F$ –criterion and $p$ –statistics, logit and probit regression equations − level of significance (maximum likelihood estimator), bankruptcy cases, the residual sum of squares (RSS). Based on the financial statements of Ukrainian insurance companies, linear, logit and probit regression equations have been obtained, which can be used to assess the contributions of various factors to the future financial position of insurance companies, determine the probability of bankruptcy and prospective financial health.

**Keywords:** insurance company, prediction the financial health, logit and probit model, regression.

## INTRODUCTION

Amidst instability in the economic and financial system, the management of financial intermediaries should constantly monitor financial health and assess its propensity for bankruptcy. Financial supervisory authorities are also interested in this, since an operational analysis of the financial position of financial intermediaries and the prevention of their crisis states forms the basis for improving the effectiveness of supervisory activities and achieving its primary objective − to ensure stability in the financial sector, which is not possible without the stability of individual financial intermediaries. This task can be accomplished within the framework of off-site financial supervision, based on the use of models for prediction the financial health and early diagnosis of bankruptcy probability. On the one hand, financial intermediary management cannot neglect the ability to improve financial performance, on the other, their managers are required to identify problems, risks,

threats and crisis phenomena in an institution's activities in a timely manner.

## PREVIOUS RELATED RESEARCH

Nowadays, there is a wide range of methodological tools for determining the bank financial health and stability, such as statistical method, stress testing, rating system, macroprudential, expert, coefficient, factor, integral, structural and functional discriminant analysis [4, p. 117-118]. The same methods can be applied to non-bank financial intermediaries. For timely determination of the financial heath and a thorough understanding of the financial situation, these methods should be combined, since the results obtained and the estimates regarding the financial situation with the financial intermediary make it possible to draw a comprehensive picture.

Daniel Martin developed a group of equations for early warning of bank failure in 1969-1974. For each of these years, he determined the number of bankrupt banks and banks that continued operations [5]. Based on these calculations, he hypothesized that each bank has the same propensity for bankruptcy, which is equal to the ratio between bankrupt and non-bankrupt banks in the respective year. In his research, he examined the impact of 25 factors on banks activities. These factors were grouped as follows: asset risk, liquidity, capital adequacy, income. Building upon the progress already made, Martin constructed logit regression equations for different forecast horizons.

The study conducted by Hanweck is based on a sample of 177 banks, whose activities were estimated for the period 1973–1975. He identified the following factors for probit regression: net operating income / total assets; own equity / total assets; percentage change in net operating income / total assets; percentage change in total assets; loans / total capital; the size of the bank by volume of assets [4].

In 1996, R.S. Barr and T.S. Siems offered to include in an investigation the factors that not only explains direct banking activities (own equity / loans, inactive loans / assets, net profit / total assets, large deposits / total assets), but also such indicators that describe management quality and economic conditions [2]. The models were built for one- and two-year forecast horizons, with estimates being 92.4% and 94.8%, respectively.

In the same vein, by studying the activities of about 1,500 US banks, J. Tatom developed a logit and probit model based on CAMEL estimates and other indicators for one-year, two-year and one-quarter forecast horizons [10].

There are also some developments in determining the future financial health of insurance companies and in predicting their likelihood of bankruptcy. C. Anghelache and D. Armeanu classified the insurance companies operating in the Romanian insurance market in 2006. According to the results of the study, the following indicators were selected: gross premiums, net mathematical reserves, gross payments, net reserves, net income, authorized capital, reinsurance operations [1]. A broader perspective has been adopted by Z.D. Martinez, J.F. Menendez and J.S. Vargas who used non-parametric methods in order to predict the financial health of insurance companies along with discriminant analysis [6]. These methodological approaches have been used for Non-life insurance companies in Spain.

## RESEARCH METHODOLOGY

The large number of bankrupt financial intermediaries in 2014-2016, and in particular of banking institutions and insurance companies, determines the need to improve the methods of express diagnostics of financial health, which will improve the practice of financial supervision. One of the most common approaches to prediction the financial health and forecasting the likelihood of bankruptcy of financial intermediaries involves the model construction based on economic and mathematical methods as a combination of various factors that affect the activities of the institution. According to this method, when constructing the function, it is a need to determine the weight that characterizes the power of influence of a factor on the financial health of the intermediary as a whole. We use a toolkit of binary choice models, including logit and probit functions, which can take values from 0 to 1. We also use a linear regression for better understanding the impact of factors on bankruptcy probability and the prospective financial position of an intermediary. The analysis can be based on bankruptcy data of insurance companies that have occurred in recent years in Ukraine.

A study sample consists of 12 insurance companies, of which 6 ceased their activities and 6 that functioned normally. An open register of insurance companies operating in the market and their financial and management reporting have not been established in Ukraine. Using open-source information, it is impossible to obtain accurate information about the characteristics of insurance companies and to create a large sample of financial intermediaries.

The scale for a binary variable is determined on the basis of an empirical study of the financial intermediaries activities: $p = 1$, if a financial intermediary bankrupt, $p = 0$, if the intermediary works in the financial market without problems. Independent variables are predictors that reflect the financial position of intermediaries: $x_1$, $x_2$, …, $x_k$. In doing so, the following types of regression equations should be constructed:

– linear:

$$p = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k, \quad (1)$$

– logit:

$$p = \frac{1}{1+e^{-Z}}, \quad (2)$$

where

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k; \quad (3)$$

– probit:

$$p = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-z^2}{2}} dz, \quad (4)$$

where

$$Z = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k. \quad (5)$$

The construction of the linear equation in our study is only due to the fact that we can judge the specific impact of predictors on the probability of bankruptcy of insurance companies or banks. However, given the linearity of the relationship, bankruptcy probabilities $p$ can go beyond this equation $[0; 1]$, making it difficult to predict using a linear model. Therefore, logit and probit models are better suited for this purpose.

The difference between logit regression and probit regression is that the first value of the function $p$ is in the range from 0 to 1, and in the second − as close as possible or equal to 0 or 1. The logit-regression toolkit was developed in the 1950's and 1960's, while the probit-regression toolkit was developed somewhat earlier in the 1930's. Logistic regression is used to construct logit-regression and normal when constructing probit-regression.

The parameters of the regression equations is determined using the statistical analysis software package STATISTICA 10.

## RESEARCH RESULTS AND DISCUSSION

For this study, the selection of liquidated and properly functioning insurance companies was made on the basis of data of the National Commission, which carries out state regulation in the sphere of financial services markets.

Based on the financial statements of insurance companies for 2015 [9], the study sample was formed based on array of metrics that are reflected in the statement of financial position and income and expenses of insurance companies. To consider whether to include insurance companies in the financial forecasting model, 31 indicators were identified:

$x_1$ – working capital / (short-term liabilities + insurance reserves);

$x_2$ – (cash + short-term financial investments) / short-term liabilities;

$x_3$ – (working capital − long-term accounts receivable) / (short-term accounts payable + insurance reserves);

$x_4$ – (cash + short-term financial investments + short-term receivables) / (short-term accounts payable + insurance reserves);

$x_5$ – cash / short-term liabilities;

$x_6$ – value of balance sheet profit / volume of capital;

$x_7$ – net profit / own equity;

$x_8$ – income from investment of insurance reserves / (short-term financial investments + long-term financial investments + securities);

$x_9$ – most liquid assets / short-term liabilities;

$x_{10}$ – (working capital − deferred costs) / short-term liabilities;

$x_{11}$ – income of the insurance company / assets;

$x_{12}$ – total income / own equity;

$x_{13}$ – total income / current assets;

$x_{14}$ – own equity / total capital;
$x_{15}$ – total capital / own equity;
$x_{16}$ – (current assets − current liabilities) / own equity;
$x_{17}$ – own equity / borrowed capital;
$x_{18}$ – borrowed capital / total capital;
$x_{19}$ – net income / assets;
$x_{20}$ – net income / own equity;
$x_{21}$ – accounts receivable / total capital;
$x_{22}$ – highly liquid assets / liabilities;
$x_{23}$ – net premiums / total capital;
$x_{24}$ – liabilities / total capital;
$x_{25}$ – net profit / total capital;
$x_{26}$ – payouts and expenses / net earned premiums;
$x_{27}$ – total capital at the end of the period / total capital at the beginning of the period;
$x_{28}$ – the amount of net premiums at the end of the period / the amount of net premiums at the beginning of the period;
$x_{29}$ – net premiums / gross premiums;
$x_{30}$ – net insurance reserves / total capital;
$x_{31}$ – profit from financial activities / financial investments.

The indicators selected in this research belong to different groups - liquidity, solvency and financial stability, profitability, business activity, as well as to a group of specific indicators of early warning tests of insurance company instability. The choice of predictors to be included in the model is followed by the use of a two-choice $F$ –test for variance, the Farrar-Globe algorithm, and the pair correlation method.

To test the adequacy of the obtained models, special indicators should be calculated: for linear regression – $R^2$, $F$ – criterion, $p$ – statistics, for logit and probit equations: $\chi^2$, level of significance (maximum likelihood estimator), bankruptcy cases, the residual sum of squares (RSS). Equation testing should also be based on determining theoretical values of a function and comparing them with empirical values.

In order to select the predictors of the insurance companies activity for inclusion in the model, we tested their significance using a two-choice $F$ –test for variances (Table 1). Given the levels of freedom with a probability of 0.95, the tabular value of the $F$ –criterion is 0.198.

Comparing the calculated values with the critical ones, a number of indicators from the totality of the insurance companies indicators were removed due to their statistical insignificance: $x_1$, $x_2$, $x_3$, $x_4$, $x_5$, $x_9$, $x_{10}$, $x_{12}$, $x_{13}$, $x_{14}$, $x_{15}$, $x_{16}$, $x_{17}$, $x_{18}$, $x_{21}$, $x_{22}$, $x_{24}$, $x_{26}$, $x_{28}$, $x_{29}$, $x_{30}$ and $x_{31}$.

Using the Farrar-Glouber algorithm, the variables were standardized and a correlation matrix was made. Then, having defined the criterion $\chi^2$, the multicollinearity in the data set were defined. This conclusion was made by comparison defined the criterion $\chi^2$ (134.66) and its table value (23.27). If $\chi^2 > \chi_{tabl}$

there is multicollinearity in the data set. By constructing a matrix of pair correlation coefficients (Table 2), the tight relationship with factors such like $x_6$ and $x_7$; $x_6$ and $x_{25}$; $x_7$ and $x_{25}$; $x_8$ and $x_{20}$; $x_{11}$, $x_{19}$, $x_{20}$, and $x_{23}$; $x_{19}$, $x_{20}$ and $x_{23}$; $x_{20}$ and $x_{23}$ was found. To get rid of multicollinearity, the factors $x_7$, $x_8$, $x_{19}$, $x_{20}$, $x_{23}$ and $x_{25}$ were removed.

Table 1. $F$ –criterion values for indicators of insurance companies performance

| **Indicator** | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ |
|---|---|---|---|---|---|---|---|
| $F$ –criterion | 0.000 | 0.012 | 0.000 | 0,000 | 0,037 | 9,218 | 39,901 |
| Indicator | $x_8$ | $x_9$ | $x_{10}$ | $x_{11}$ | $x_{12}$ | $x_{13}$ | $x_{14}$ |
| $F$ –criterion | 54.423 | 0.012 | 0.000 | 0,434 | 0,022 | 0,120 | 0,054 |
| Indicator | $x_{15}$ | $x_{16}$ | $x_{17}$ | $x_{18}$ | $x_{19}$ | $x_{20}$ | $x_{21}$ |
| $F$ –criterion | 0.007 | 0.059 | 0.026 | 0,063 | 0,730 | 10,953 | 0,001 |
| Indicator | $x_{22}$ | $x_{23}$ | $x_{24}$ | $x_{25}$ | $x_{26}$ | $x_{27}$ | $x_{28}$ |
| $F$ –criterion | 0.000 | 6.659 | 0.007 | 34,022 | 0,000 | 1,945 | 0,001 |
| Indicator | $x_{29}$ | | $x_{30}$ | | $x_{31}$ | | |
| $F$ –criterion | 0.000 | | 0.094 | | 0.000 | | |

Source: calculated on the basis of the Agency for Stock Market Infrastructure Development of Ukraine [9], annual reports and financial statements of insurance companies

Table 2. Matrix of pair correlation coefficients

| | $x_6$ | $x_7$ | $x_8$ | $x_{11}$ | $x_{19}$ | $x_{20}$ | $x_{23}$ | $x_{25}$ | $x_{27}$ |
|---|---|---|---|---|---|---|---|---|---|
| $x_6$ | 1.000 | | | | | | | | |
| $x_7$ | 0.995 | 1.000 | | | | | | | |
| $x_8$ | 0.241 | 0.283 | 1.000 | | | | | | |
| $x_{11}$ | -0.243 | -0.193 | 0.405 | 1.000 | | | | | |
| $x_{19}$ | -0.086 | -0.034 | 0.504 | 0.828 | 1.000 | | | | |
| $x_{20}$ | -0.030 | 0.025 | 0.675 | 0.782 | 0.951 | 1.000 | | | |
| $x_{23}$ | -0.287 | -0.228 | 0.560 | 0.749 | 0.930 | 0.941 | 1.000 | | |
| $x_{25}$ | 0.804 | 0.750 | -0.013 | -0.329 | -0.268 | -0.256 | -0.481 | 1.000 | |
| $x_{27}$ | 0.466 | 0.493 | 0.419 | 0.287 | 0.425 | 0.542 | 0.283 | 0.230 | 1.000 |

Source: calculated on the basis of the Agency for Stock Market Infrastructure Development of Ukraine [9], annual reports and financial statements of insurance companies

Taking into account previously removed factors, the final list of predictors of insurance companies' activity for constructing regression functions (gross return on

capital ($x_6$), turnover of assets ($x_{11}$) and indicator of changes in capital ($x_{27}$)) were obtained. The obtained parameters of the equations for the regression equations for prediction the financial health of Ukrainian insurance companies are given in Table 3. To evaluate the adequacy of the linear equation, the standard error and $t$ –statistics for all regression equation coefficients were calculated. The multiple regression coefficient is 0.85; the determination coefficient is 0.72. The value of $F$ –criterion is 6.72 with a theoretical value of 4.07, which indicates the adequacy of the obtained regression equation. The $p$ –statistic values for the parameters $\beta_6$ and $\beta_{27}$ correspond to a significance level of 0.05.

Table 3. Parameters of regression equations for prediction the financial health of insurance companies in Ukraine

| Parameters | Regression | | |
|---|---|---|---|
| | linear | logit | probit |
| $\beta_0$ | 1.300 | 49.932 | 10.455 |
| $\beta_6$ | 0.966 | 12.200 | 4.504 |
| $\beta_{11}$ | -0.674 | -65.464 | -13.293 |
| $\beta_{27}$ | -0.305 | -7.521 | -2.295 |

Source: calculated on the basis of Forinsurer data [3] and National Commission for Regulation of Financial Services Markets in Ukraine [7]

An analysis of the impact of the predictors according to the values of linear regression coefficients on the performance trait showed that an increase in net return on equity and a decrease in asset turnover and insurance risk reduced the likelihood of bankruptcy of insurance companies.

The obtained logit-regression equation for f prediction the financial health of insurance companies in Ukraine has a value of $\chi^2$ statistics equal to 16.63. This is indeed a vindication of the fact that selected predictors affect the probability of bankruptcy of Ukrainian insurance companies. The percentage of cases of foreseen insurance companies bankruptcies as well as those that continued their functioning is 100%. The sum of squares of residuals is also insignificant – 0.000805313, which indicates the normal quality of the obtained probit-regression.

The obtained probit-regression equation for forecasting the financial position of Ukrainian insurance companies has a value of $\chi^2$ statistics equal to 16.55, which confirms the hypothesis about the impact of selected predictors on the financial health of Ukrainian insurance companies. The percentage of cases of foreseen insurance companies bankruptcies as well as those that continued their operations is 100%. The sum of the residual squares is also insignificant (0.043650596) and is a testament to the quality of the obtained probit-regression equation.

Based on the obtained parameters of equations for prediction the financial health of insurance companies in Ukraine, the regression equations are as follows:

linear regression

$$p = 1.300 + 0.966x_6 - 0.674x_{11} - 0.305x_{27}, \quad (6)$$

logit-regression

$$p = \frac{1}{1+e^{-(49.932+12.200x_6-65.464x_{11}-7.521x_{27})}}, \quad (7)$$

And also probit-regression

$$p = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(10.455+4.504x_6-13.293x_{11}-2.295x_{27})^2}{2}} dz \quad (8)$$

Also, the theoretical values of the $p$ function were calculated. The results of this assessment presented in the Table 4.

Table 4. Empirical and theoretical values of the functions of regression equations for prediction the financial health of insurance companies in Ukraine

| Name of insurance company | Empirical value | Theoretical values | | |
|---|---|---|---|---|
| | | linear | logit | probit |
| PJSC NASK “ORANTA” | 0.000000 | 0.380410 | 0.000661 | 0.018899 |
| PJSC IC “Providna” | 0.000000 | -0.522770 | 0.000000 | 0.000000 |
| PJSC SG “TAS” | 0.000000 | 0.175730 | 0.000000 | 0.000000 |
| PJSC “UPSK” | 0.000000 | 0.428840 | 0.000003 | 0.001319 |
| PJSC USC “Knyazha” | 0.000000 | 0.167690 | 0.000000 | 0.000000 |
| PJSC IC “UNIQA” | 0.000000 | 0.220740 | 0.000000 | 0.000000 |
| PJSC IC “Vector Invest” | 1.000000 | 0.857960 | 1.000000 | 0.999750 |
| PJSC IC “Vector Invest Life” | 1.000000 | 0.913990 | 1.000000 | 1.000000 |
| PJSC SC “Standard Re” | 1.000000 | 0.843110 | 0.999860 | 0.977322 |
| PJSC IC“Zlagoda” | 1.000000 | 0.947820 | 1.000000 | 1.000000 |
| PJSC SC “Atlanta” | 1.000000 | 0.776828 | 1.000000 | 0.999942 |
| PJSC USC "Garant-Life" | 1.000000 | 0.807170 | 1.000000 | 0.999996 |

Source: calculated on the basis of Forinsurer data [3] and National Commission for Regulation of Financial Services Markets in Ukraine [7]

Thus, for linear regression, $p$ values for some insurance companies go beyond the range $[0; 1]$, while for logit and probit regressions, theoretical $p$ values are completely consistent with empirical values. Therefore, logit and probit regressions can be used for prediction the financial health of insurance companies.

## CONCLUSION

Developing models for prediction the financial health of insurance companies could be viewed as one of the ways to improve the financial management of insurance companies and manage their risks. Other institutions and individuals, such as financial supervisory authorities, potential customers, business partners and other stakeholders interested in understanding the level of stability and reliability of an insurance company, could also use these models.

The first stage of models for prediction the financial health of insurance company's development involves the selection of insurance companies that operate stably and ceased their operations for a certain period of time. Usually, such selection is made in the years corresponding to the financial crisis − when some of the insurance companies operating in the market go bankrupt, lose their licenses and cease their business, while others − overcomes the crisis and adapts its business model to the new conditions of functioning. A primary goal of the next stage is to calculate as many indicators as possible to show the financial position of the insurance companies and form a short list of indicators that will be included in the model, based on the use of such tools as a two-choice $F$ −test for variance, Farrar-Glouber algorithm, pair correlation. The third stage is to determine and analyze the parameters of the linear, logit and probit regression equations. The last, but not least stage is to check the quality of the obtained equations: linear − $R^2$, $F$ –criterion, $p$ –statistics; logit and probit − $\chi^2$, level of significance (maximum likelihood estimator), bankruptcy cases, the residual sum of squares (RSS).

Based on the financial statements of Ukrainian insurance companies, linear, logit and probit regression equations have been obtained, which can be used to assess the impact of factors on the future financial position and determine the probability of bankruptcy of insurance companies and prospective financial health.

## REFERENCE

1. Anghelache, C., and Armeanu, D. Application of Discriminant Analysis on Romanian Insurance Market. *Theoretical and Applied Economics*, 2008. 11 (528). Pp. 51−62.

2. Barr, Richard S., and Thomas F. Siems. Bank failure prediction using DEA to measure management quality. *Interfaces in computer science and operations research*. Springer, Boston, MA, 1997. Pp. 341-365.

3. Forinsurer [Online]. – Available at: https://forinsurer.com/ratings/Nonlife

4. Hanweck, Gerald A. *Predicting bank failure*. No. 19. Board of Governors of the Federal Reserve System (US), 1977.

5. Martin, D. Early warning of bank failure: A logit regression approach. *Journal of banking & finance*, 1977. 1.3. Pp. 249-276.

6. Martinez Z. D., Menendez J. F., Vargas J. S. Algorithm versus Discriminant Analysis. An Application to the Prediction of Insolvency in Spanish Non-life Insurance Companies. Documentos

de trabajo de la Facultad de Ciencias Económicas y Empresariales [Online]. – Available at: http://eprints.ucm.es/6836/1/0412.pdf

7. National Commission for Regulation of Financial Services Markets of Ukraine [Online]. – Available at: https://nfp.gov.ua/ua/Ohliad-rynkiv.html

8. Prymak, Ju.R. Suchasni ukrajinsjki ta mizhnarodni metody analizu finansovoji stijkosti bankivsjkoji ustanovy. *Mizhnarodnyj naukovyj zhurnal,* 2016. 9. Pp. 115−122

9. Stock Market Infrastructure Development Agency of Ukraine [Online]. – Available at: https://smida.gov.ua/db/emitent.

10. Tatom, J., and Houston, R. Predicting failure in the commercial banking industry. *Networks Financial Institute Working Paper*, 2011. Working Paper. 27.