

# ***Modeling of Probability of a Default of Mortgage Borrowers Using Econometric Methods***

Alexander Almosov

The Volgograd institute of management – RANEPA branch  
Economics department  
Volgograd, Russia  
[vsamol@yandex.ru](mailto:vsamol@yandex.ru)

Yulia Brekhova

The Volgograd institute of management – RANEPA branch  
Economics department  
Volgograd, Russia  
[brechova7@mail.ru](mailto:brechova7@mail.ru)

Elena Malysheva

The Volgograd institute of management – RANEPA branch  
Economics department  
Volgograd, Russia

Sofya Potomova

The Volgograd institute of management – RANEPA branch  
Economics department  
Volgograd, Russia  
[potomova@inbox.ru](mailto:potomova@inbox.ru)

**Abstract**—The article includes an analysis of probability of a default of loans being part of the mortgage pool. For the purpose of analysis it is proposed to develop an econometric model of probability of a default. A group of loan variables and a group of macroeconomic variables are selected as explanatory variables. As a result, the authors obtain and test the model consisting of 10 variables and demonstrating a high predictive power.

**Keywords**— *credit risk, mortgage, securitization, probability of a default*

## **Introduction**

Development of securitization in Russia has a relatively short history and is still perceived as a financial innovation. Nevertheless, use of this tool can improve competitiveness and sustainability of the Russian financial market. Securitization is based on the following elements:

- coverage, which is initial financial assets (liabilities) forming the basis of a securitization transaction;
- improvement of the loan quality, which is the mechanism aimed at enhancing reliability of securities and protecting investors from exposure to loss or other risks associated with the assets included in the coverage;
- standardization, which is the process of establishing general rules and requirements for lending that ensures the unity of framework contractual terms and conditions;
- liquidity, which is one of the goals of securitization reflecting the degree of transferability of issued securities, i.e., simplicity of their transfer by one investor to another one.

The basis of existence of the very securitization transaction is availability of an adequate, high-quality and homogeneous coverage [1]. The sequence of settlement transactions and features of the mortgage pool set the quality of issued securities and establish the peculiarities of structuring. Residential mortgage loans are more often used among a wide range of assets which can serve as collateral. The reason for this is availability of two principal properties: first, this type of lending is mass, and, second, there is a legally proven mechanism of assignment of debts under mortgage loans and forfeiture of mortgaged immovable property.

The analysis of the mortgage pool regarding probability of a default of loans being part of the mortgage pool allows identifying existence or lack of a threat for investors and timely protecting from exposure to loss. The need for this analysis is confirmed by the mortgage crisis in 2008, one of the main causes of which is considered a large percentage of defaulted borrowers in the coverage of issued bonds. At the moment, the market of residential mortgage loans (hereinafter “RMLs”) in Russia is artificially warmed up with diverse state support programs [2, 3]. These programs are developed to increase the volumes of RMLs and, due to this, support various sectors associated with the construction industry.

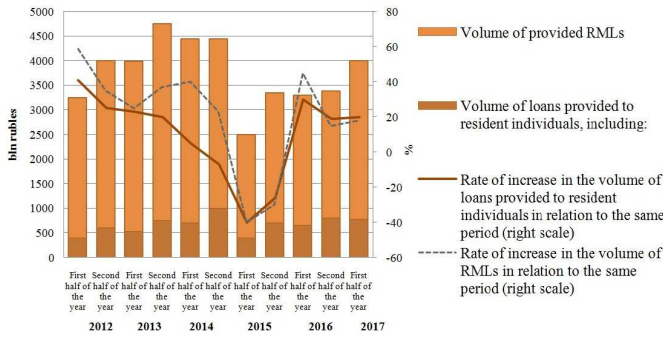


Fig. 1. Dynamics of volumes of loans provided to physical persons including RMLs.

The government plans to increase the volumes of provided RMLs from 4.8 trillion rubles as of October 1, 2017 to 10 trillion rubles by the end of 2020. The annual volume of provided mortgage loans is to grow from 1.9 trillion rubles in 2017 to 3-3.5 trillion rubles in 2020. At the same time, these measures can form a similar mortgage bubble and lead to a new crisis [4]. In order to detect the growing level of the credit risk timely, it is proposed to develop the model for assessment of probability of a default of loans being part of the mortgage pool.

**Materials and Methods (Model)**

The basis of the analysis is the method of econometric modeling, which allowed identifying certain indices of the quality of mortgage loans and their impact on the level of credit risk of the entire pool of securitized assets. The graphic method and the method of scientific abstraction were applied for visual demonstration of the identified interrelations.

**Results and Discussion**

VTB 24 bank is of most interest for the analysis of the mortgage pool, as one of the main players in the Russian market of mortgage loans, which is a member of VTB Group, the largest participant of the securitization market. At the moment, in the market of securities there are 9 issues for a total amount of over 200 billion rubles. The following issue was selected for the analysis: nonconvertible certified interest-bearing bearer bonds, subject to mandatory deposit, 2-IP series, two grades (hereinafter “issue”). It is characterized by the possibility of early redemption, emitted in the number of 4 million pieces of par value of 1,000 rubles each, with the redemption date of November 26, 2043. This issue was registered on September 1, 2011, includes 1599 loans accumulated for the period from 2010 to 2011.

This issue has the longest circulation in the market among similar issues of VTB 24. The issue was registered on September 1, 2011, the mortgage pool as of the end of 2017 comprises 1599 loans for an amount of 1,474,887,325.5 rubles. The data on the mortgage pool are published on a monthly basis, which allows establishing a 2-month time lag for the analysis. Looking ahead, let us specify that the further analysis showed a dependence in terms of which the increasing lag leads to both the decreasing quality and adequacy of the model to the empirical data. It appears logical

that financial variables taken into account directly before a default (or successful continuation of loan disbursement) will allow predicting a default (continuation of disbursement) more accurately.

Before constructing the econometric model, let us present the construction stages in the form of algorithm.

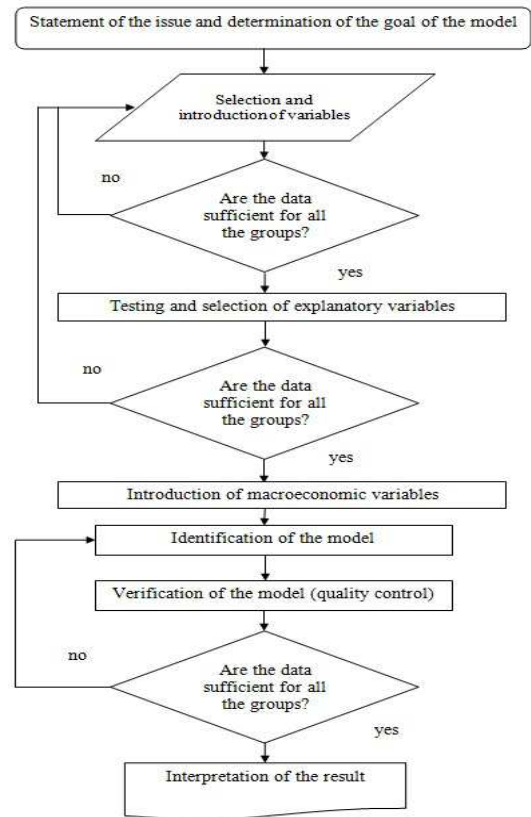


Fig. 2. Algorithm for constructing the econometric model of probability of a borrower's default.

First of all, a potential set of explanatory variables was formed from the presented mortgage pool to construct the model of probability of a borrower's default. Explanatory variables were selected on the basis of the economic meaning of the available data. The following indicators were selected initially:

1. Loan amount –  $x_1$ ;
2. Federal district –  $x_2$ ;
3. Area of mortgage item –  $x_3$ ;
4. Market value of mortgage –  $x_4$ ;
5. Number of years before discharge of the obligation –  $x_5$ ;
6. Full term of loan –  $x_6$ ;
7. Amount of payment –  $x_7$ ;
8. Volume of default on the principal of the debt –  $x_8$ ;
9. Volume of default on the interest of the debt –  $x_9$ ;
10. Loan/Mortgage Indicator –  $x_{10}$ .

The data were selected in a few stages. The first stage of data selection is elimination of obvious errors (in input or measurement): observations with negative values for variables were excluded and normalization of data was performed. For preliminary generation and cleaning of explanatory variables an analysis of variance test (ANOVA-test) was carried out in relation to the model of probability of a default of borrowers. In order to apply this test, the average values of financial indicators are formed by groups of solvent and defaulted borrowers. The results of the analysis show that the difference in the default level of borrowers with different variables is significant (substantial). This is evidenced by the value of F: 4659.099 > 1.721, as well as by the low probability of acceptance of the null hypothesis (p-value = 0.01). Consequently, the null hypothesis is rejected and the alternative hypothesis that the borrower's default level depends on the listed variables is accepted.

The next stage was examination of pair correlations of the selected variables. In the article the correlation level equal to 0.3 is considered as the minimum admissible level of coherence among explanatory variables, which is acceptable for such economic studies. The matrix of correlations of preliminarily selected explanatory variables is presented in Table 1.

TABLE I. CORRELATION MATRIX

Variable	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
$x_1$	1									
$x_2$	0.1	1								
$x_3$	0.4	0.0	1							
$x_4$	0.7	0.1	0.6	1						
$x_5$	0.0	0.0	0.0	0.0	1					
$x_6$	0.0	0.0	0.0	0.0	0.9	1				
$x_7$	0.7	0.0	0.2	0.5	0.0	0.0	1			
$x_8$	0.7	0.0	0.2	0.5	0.4	0.4	0.7	1		
$x_9$	0.4	0.0	0.1	0.3	0.7	0.7	0.3	0.8	1	
$x_{10}$	0.3	0.0	0.2	0.2	0.1	0.1	0.2	0.2	0.2	1

As a result, we rejected using the following variables:  $x_3$  – area of item,  $x_{10}$  – loan/mortgage,  $x_4$  – market value of mortgage,  $x_9$  – default on the interest of the debt,  $x_6$  – full term of loan. In order to consider the environment of borrowers, it is necessary to introduce consideration of macroeconomic factors in the model [5]. Meanwhile, not only absolute values of macrovariables (for example, the cost of housing in the federal district) are considered, but also relative features which can show their unexpected and abrupt changes affecting the borrower's financial status. Therefore, the range of explanatory variables used further considers the rates of

growth of macroeconomic variables [6]. The macroeconomic variables which were viewed as candidates for inclusion in the model of probability of a mortgage borrower's default are as follows:

1. Real disposable money income of the population –  $x_{11}$ ;
2. Unemployment rate –  $x_{12}$ ;
3. Price indices in the housing market –  $x_{13}$ ;
4. Key rate of the Central Bank of the Russian Federation –  $x_{14}$ .

As a result, the model was as follows:

$$P_{def} = f(x_{1,lag2}; (x_1, lag2)^2; x_{2,lag2}; (x_2, lag2)^2; x_{5,lag2}; (x_5, lag2)^2; x_{7,lag2}; (x_7, lag2)^2; x_{8,lag2}; (x_8, lag2)^2; x_{11,lag2}; (x_{11, lag2})^2; x_{12,lag2}; (x_{12, lag2})^2; x_{13,lag2}; (x_{13, lag2})^2; x_{14,lag2}; (x_{14, lag2})^2; C). \tag{1}$$

Before starting the direct interpretation of the results, it is necessary to generalize coefficients of variables for 100 models constructed for all the subsamples (the final specification at the stage of construction of the model was obtained by one of 100, the results of modeling of which appeared to be the closest to the empirical data). As a result, the model was as follows:

$$P_{def} = f(2,238 \cdot x_{1,lag2} - 2,451 \cdot (x_1, lag2)^2 + 1,582 \cdot x_{2,lag2}; - 3,063 \cdot (x_2, lag2)^2 + 1,946 \cdot x_{5,lag2} - 0,392 \cdot (x_5, lag2)^2 + 0,879 \cdot x_{7,lag2} + 1,628 \cdot (x_7, lag2)^2 + 1,694 \cdot x_{8,lag2} - 6,276 \cdot (x_8, lag2)^2 - 1,378 \cdot x_{11,lag2} - 0,812 \cdot (x_{11, lag2})^2 + 0,932 \cdot x_{12,lag2} - 4,947 \cdot (x_{12, lag2})^2 - 1,204 \cdot x_{13,lag2} + 2,706 \cdot (x_{13, lag2})^2 + 0,994 \cdot x_{14,lag2} - 3,256 \cdot (x_{14, lag2})^2 - 7,156). \tag{2}$$

Therefore, on the basis of the obtained results dependence of the borrower's default on the loan amount can be observed. This dependence as we expected is nonlinear and a big amount of loan increases probability of the borrower's default. Meanwhile, even in terms of the average level of the loan/mortgage indicator in most cases the borrower with the biggest amount of loan was recognized defaulted.

The indicator of number of years before discharge of the obligation appeared to be connected with the indicator of "default on the principal on the debt". We identified the expected and generally accepted regularity which consists in the highest probability of recognition by the borrower of the loan as default in case of payment of less than 50% of the mortgage loan value, i.e., in case when the borrowers are in a difficult financial situation, they are more likely to refuse to continue paying the loan if the outstanding amount exceeds 50% of the loan amount, and vice versa. During study of the pool of loans selected by us it was discovered that only 30% of the defaulted borrowers paid more than 50% of the principal of the debt. This confirms the stated assumption about the inverse dependence of the term of loan and the level of its credit risk.

The amount of payment and the key rate of the Central Bank of the Russian Federation demonstrated a weak dependence, and only in 2017 the key rate appeared to be more significant. The reason for this was the decrease in the key rate and further appeal of borrowers to lenders for

reduction of the interest rate of the mortgage loan subsequent to the change in the key rate of the Bank of Russia [7].

The rest of the macroeconomic variables (real disposable money income of the population; unemployment rate; price indices in the housing market) showed the expected significance and inclusion of them in the model definitely led to the higher predictive ability of the model.

### Conclusion

The presented model includes 10 variables, among them 4 macrovariables, which allows achieving the highest predictive power of the model. In practice, in order to enhance the accuracy of assessment and identification of borrowers requiring greater attention, the alternative models can be used simultaneously, for example, on the basis of developed ratings or neural networks. The advantage of this particular model is simplicity of its use and availability of the data, which is a considerable advantage for private investors or for application of this analysis as a rapid test in case of large amounts of data.

### References

- [1] Tano Santos, comments by Andrés Velasco. «Credit booms: implications for the public and the private sector». 2015. 1-15 pages. Available at: <http://www.bis.org/publ/work481.pdf>
- [2] The federal law of 11.11.2003 N 152-FZ (an edition of 01.07.2017) "About mortgage securities" ["Ob ipotechnykh tsennykh bumagakh"]
- [3] Basel III Document Revisions to the securitisation framework // Bank for International Settlements, 11 December 2014, pp. 1-43
- [4] Lavrova M.I., Hudoroshkova M.A. Enhancement of assessment of a subject to pledge for the purpose of mortgage lending [Sovershenstvovanie otsenki ob'ekta zaloga v tselyakh ipotechnogo kreditovaniya] // Young scientist [Molodoi uchenyig]. 2016. vol. 10-5 (114). pp. 33-36.
- [5] Medvedeva I.A. Territorial inequality of incomes of the population in Russia [Territorial'noe neravenstvo dokhodov naseleniya v Rossii] // Science Krasnoyarya [Nauka Krasnoyar'ya] 2017. T. 6. vol. 3-3. pp. 153-155.
- [6] Encyclopedia of the Russian securitization of 2017 [Entsiklopediya rossiiskoi sek'yuritizatsii 2017]. St. Petersburg. – 2017. – P. 207.
- [7] Alpatskaya I.E., Alpatsky D.G. Tendencies of development of the Russian market of mortgage securities in the conditions of crisis//Economy and an entrepreneurship. 4-1 (69-1). 2016. vol.4-1 (69-1). 2016. pp. 731-734