Simulation of Banking Sector’s Investment Potential

L Kurmanova¹, D Kurmanova²
¹Institute of Economics, Finance and Business, Department of Finance and Taxation, Bashkir State University, Ufa, Russia
²Department of Finance, monetary circulation and economic security, UFA State Aviation Technical University, Ufa, Russia

E-mail: kurmanova_ugaes@mail.ru

Abstract. The article analyzes the factors of the macroeconomic environment that affect the level of the investment potential of the national banking system in modern conditions. Investment potential is a quantitative indicator assessing the availability of certain resources and their sufficiency to meet investment demand and investment activities. It is the investment opportunities of commercial banks that acquire special significance for economic growth, and therefore it is extremely important to develop an up-to-date mechanism for developing and realizing the investment potential of the Russian banking sector. The situation with long-term financing of the real economy is aggravated by the introduction of sanctions, which led to the closure of access to the Russian banks for international capital markets for borrowing "long" money. Issues related to the definition of the investment potential of the banking system and the evaluation of formation factors are widely considered in the scientific literature. However, there is no general definition and unified methodology for assessing the investment potential of the banking sector. All this testifies to the urgency of the scientific problem of developing assessment tools and forecasting investment potential, taking into account all possible factors affecting its magnitude. To determine the points of investment potential growth on the basis of correlation-regression analysis, a three-factor model for forecasting investment activity and expanding the potential of Russian commercial banks was formed.

1. Introduction

Commercial banks are the key link in the financial and credit system that carry out investment lending to the real sector of the economy. The opportunities for activating bank lending to companies are largely determined by the investment potential of the banking system, namely the volume and quality of resources involved in the investment process. Investment potential is an important indicator that gives an idea of the banks' ability to ensure a sustainable expanded reproduction of the resource base while increasing the efficiency of its use to finance the needs of the national economy.

The investment potential of the bank is a set of strategic resources that determine the boundaries of financial opportunities for the functioning of the credit institution under certain conditions [1]. Accordingly, it is possible to consider the investment potential of the banking sector of the country as a set of strategic prerequisites and factors for its sustainable development, including by attracting foreign investments, so it is important to identify the factors shaping the investment potential of the banking sector.
Data from the Bank of Russia show that, in general, the volume of investment is small, if compared with the volumes in developed countries. The volume of investment of commercial banks in Russia is about 1.3 trillion. doll, while in the US this figure is 5.5 trillion. doll. [2]. In the total financing of investments in fixed assets, the share of bank loans varies between 8-10%. While in the economics of developed countries, almost half of investments in fixed assets are carried out through loans from the banking sector. The concept of long-term social and economic development of the Russian Federation for the period until 2020 provides for an increase in the share of banks in sources of financing investments in fixed assets to 20-25% in 2020, which is an extremely difficult task [3].

1.1. Selection of factors for constructing a multifactorial model
To assess the degree of influence of a group of factors, the relationships of the effective feature were investigated – bank loans in investments in fixed capital (Y), as the main indicator of the investment potential of the banking system, and some factor attributes of the main macroeconomic indicators of the Russian Federation, which are given in table 1.

| Table 1. Indicators for calculating the correlation coefficient, billion rubles. |
|---------------------|---------------------|---------------------|
| **Years** | **Bank loans to investments in fixed assets** | **Deposits of individuals in banks** | **GDP** |
| 2009 | 812.6 | 10856.6 | 38807.2 |
| 2010 | 887.4 | 11425.6 | 46308.5 |
| 2011 | 904.7 | 12145.8 | 59698.1 |
| 2012 | 925.6 | 12356.8 | 66926.9 |
| 2013 | 915.5 | 12945.6 | 71016.7 |
| 2014 | 1098.7 | 13699.1 | 79199.7 |
| 2015 | 849.9 | 16347.1 | 83232.6 |
| 2016 | 1152.8 | 18472.0 | 86043.6 |

| **Years** | **Salary of the population** | **Incomes of the population** | **Gross profit of enterprises** |
|-----------|-----------------------------|-------------------------------|-------------------------------|
| 2009 | 20411.6 | 26168.7 | 11921.1 |
| 2010 | 22995.6 | 29481.6 | 15093.7 |
| 2011 | 26226.7 | 33624.0 | 24724.4 |
| 2012 | 29592.3 | 37938.8 | 27503.9 |
| 2013 | 33143.6 | 42491.8 | 27810.8 |
| 2014 | 37387.2 | 47920.7 | 30808.3 |
| 2015 | 37571.5 | 53746.6 | 36388.4 |
| 2016 | 40224.3 | 56871.2 | 36406.1 |

*Developed by the author.

Based on regression model it is advisable to consider the closeness of different variables. The connection between the variable Y (t) and m independent factors can be characterized by the regression function Y (t) = f (x1, x2, ..., xm), which shows what the average value of the variable y will be if the variables x take a specific value. This circumstance makes it possible to use the regression model not only for analysis, but also for forecasting economic phenomena.

The main stages of building the regression model are:
1) Statement of the problem. Construction of a system of indicators (factors). Collection and preliminary analysis of baseline data.
2) Correlation analysis (construction of matrix of coefficients of pair correlation, their analysis and selection of factors participating in the model).
3) Choice of model type and numerical evaluation of its parameters.
4) Check the quality of the model.
5) Evaluation of the individual factors influence on the model basis.
6) Forecasting based on the regression model [4].

The coefficients of linear and multiple correlation, determination of partial correlation coefficients can be used to determine the most significant factors. Selection of factors for the multifactor models construction is made according to the analysis objectives based on qualitative and quantitative analysis using statistical and mathematical criteria [5]. The purpose of the analysis is to determine the factors that have the strongest impact on the formation of the investment potential of the banking sector. As variables, the following factors are selected:

- $X_1$ – Deposits of individuals in banks, bln. rub.
- $X_2$ – GDP, bln. rub.
- $X_3$ – Salary of the population, bln. rub.
- $X_4$ – Total incomes of the population, bln. rub.
- $X_5$ – Gross profit of Russian companies, bln. rub.

Each of the factors has a direct or indirect effect on the factor $Y$ – banks loans to investments in fixed assets. In this case, $n = 8$ (the number of observations), $m = 5$ (the number of variables).

The main issue of the correlation analysis is to identify the connection between random variables by estimating the pair (partial) correlation coefficients, as well as the selection of factors that have the most significant effect on the outcome [6].

The selective pair correlation coefficient is determined by the formula:

$$r_{x,y} = \frac{\sum (y - \bar{y})(x - \bar{x})}{\sqrt{\sum (y - \bar{y})^2 \cdot \sum (x - \bar{x})^2}}$$

(1)

The value of the pair correlation coefficients is in the range from -1 to +1. Its positive value indicates a direct connection, negative one about feedback, in which one variable grows, the other decreases. The closer its value to 1, the closer the connection. The connection is considered strong enough if the correlation coefficient exceeds 0.8 and weak, if less than 0.4. If it is equal to zero, the connection is completely absent.

1.2. Assessment of the factors affecting the investment potential of the banking system
Calculation of the pair correlation coefficients is possible using the tool "Correlation" in the Excel program.

|       | Y   | X1  | X2      | X3      | X4      | X5  |
|-------|-----|-----|---------|---------|---------|-----|
| Y     | 1   |     |         |         |         |     |
| X1    | 0.6230 | 1   |         |         |         |     |
| X2    | 0.6545 | 0.8571 | 1       |         |         |     |
| X3    | 0.7029 | 0.8788 | 0.9867  | 1       |         |     |
| X4    | 0.6503 | 0.9393 | 0.9716  | 0.9859  | 1       |     |
| X5    | 0.5850 | 0.8617 | 0.9891  | 0.9569  | 0.9554 | 1   |

*Developed by the author.

The analysis of the pair correlation coefficients matrix given in table 2 shows that the dependent variable $Y$ has an appreciable direct connection with the factor $X1 (r_{y1} = 0.6230)$, an appreciable direct connection with the factor $X2 (r_{y2} = 0.6545)$, a close direct connection with factor $X3 (r_{y3} = 0.7029)$,
an appreciable direct connection with factor X4 ($r_{yx4} = 0.6545$), an appreciable direct connection with factor X5 ($r_{yx5} = 0.5850$). Rate the statistical significance of the correlation coefficients using Student's t-test.

Figure out the calculated values of t-statistics:

$$t = |r| \sqrt{\frac{n-2}{1-r^2}}$$

- $t_{x1} = |0.6230| \times \sqrt{\frac{8-2}{1-0.6230^2}} = 2.107 < 2.365$
- $t_{x2} = |0.6545| \times \sqrt{\frac{8-2}{1-0.6545^2}} = 2.490 > 2.365$
- $t_{x3} = |0.7029| \times \sqrt{\frac{8-2}{1-0.7029^2}} = 2.615 > 2.365$
- $t_{x4} = |0.6503| \times \sqrt{\frac{8-2}{1-0.6503^2}} = 2.265 < 2.365$
- $t_{x5} = |0.6850| \times \sqrt{\frac{8-2}{1-0.6850^2}} = 2.488 > 2.365$

The table value t-test at a significance level of 5% and degrees of freedom (8 - 1 - 1 = 6) is 2.365 (found using the function TINV).

Correlation analysis showed that $X_2$, $X_3$, $X_5$ variables out of all can be in the model as significant variables. The original data after correlation analysis are presented in table 3.

**Table 3.** The original data for building multiple regression models, billion rubles.

| Years | Bank loans to investments in fixed assets | GDP | Salary of the population | Gross profit of enterprises |
|-------|------------------------------------------|-----|--------------------------|-----------------------------|
| 2009  | 812.6                                    | 38807.2 | 20411.6                  | 11921.1                     |
| 2010  | 887.4                                    | 46308.5 | 22995.6                  | 15093.7                     |
| 2011  | 904.7                                    | 59698.1 | 26226.7                  | 24724.4                     |
| 2012  | 925.6                                    | 66926.9 | 29592.3                  | 27503.9                     |
| 2013  | 915.5                                    | 71016.7 | 33143.6                  | 27810.8                     |
| 2014  | 1098.7                                   | 79199.7 | 37387.2                  | 30808.3                     |
| 2015  | 849.9                                    | 83232.6 | 37571.5                  | 36388.4                     |
| 2016  | 1152.8                                   | 86043.6 | 40224.3                  | 36406.1                     |

*Developed by the author

The linear model of multiple regression will look like:

$$Y = a_0 + a_2X_2 + a_3X_3 + a_4X_5$$  \hspace{1cm} (2)

The regression parameters can be estimated using the Regression tool in Excel program. The regression equation for the dependence of the variable $Y$ on $X_2$, $X_3$, $X_5$, obtained with EXCEL, looks like this:

$$Y(x) = 350.733 + 0.0206 \times X_2 + 0.00456 \times X_3 + 0.0348 \times X_5$$  \hspace{1cm} (3)

The quality of the model is estimated in a standard way for mathematical models: by adequacy and accuracy based on the analysis of the remainders of regression. The calculated values are obtained by substituting in the model the actual values of all the included factors.
The analysis of the remainders gives an idea of how well the model is chosen and how correctly the coefficient estimation method is chosen. According to general assumptions of regression analysis, the remainders should behave as independent (in effect, almost independent), equally distributed random variables. In the classical methods of regression analysis, the normal distribution of remainders is also assumed [7]. The remainders are presented in table 4.

**Table 4.** Remainder’s deduction.

| Observations | Predicted Y, % | Remainders |
|--------------|---------------|------------|
| 1            | 828.4         | -15.8      |
| 2            | 884.3         | 3.1        |
| 3            | 839.7         | 65.0       |
| 4            | 907.2         | 18.4       |
| 5            | 997.0         | -81.5      |
| 6            | 1080.6        | 18.1       |
| 7            | 970.3         | -120.4     |
| 8            | 1039.7        | 113.1      |

*Developed by the author

Calculation of the main indicators for the remainders will be found using the "Descriptive statistics" tool in the Excel program.

**Table 5.** Descriptive statistics.

| Reminders | Average       | Standard deviation | Minimum | Maximum | Sum       | Score | Level of reliability (95.0%) |
|-----------|---------------|--------------------|---------|---------|-----------|-------|-------------------------------|
|           | 4.50958E-12   | 562.3821135        | -13.09562 | 11.77897 | 5.41149E-11 | 12    | 1518.924726                   |

1. Table 5 shows that the mathematical expectation of the residual component is practically 0, i.e. $M(e) = 4.50958E-12$; Check by the Student’s test:

$$t_{calc} = \frac{|\mu(\omega)|}{S_e \sqrt{n}} = 2.82E - 1 \xi,$$

where $S_e = \sqrt{\frac{\sum (e_i - \overline{e})^2}{N-1}} = 98.84$ is the standard deviation. Since $t_{calc} < t_{tabl}(\alpha=0.05, k=7)=2.228$, the average value of the residual component is assumed to be 0.

2. The levels of the residual component are not correlated to each other, since the calculated value of the Durbin-Watson criterion $d = \sum (e_i - e_{i-1})^2 + \sum e_i^2 = 76961.0 : 39081.4 = 1.96$ is more than the upper boundary of the tabulated $d_2$ (d1 = 1.06; d2 = 1.34). Calculations are carried out according to table 6.

Table values are obtained at the significance level $\alpha = 0.05$, the number of explanatory variables $k = 3$ and the number of levels of the residual component $n = 8$.

3. The levels of the residual component are distributed according to the normal law and the requirement is fulfilled: $RS_{calc}=3.34e[R_{calc}=2.8; R_{calc}=3.91]$. The calculated value of the RS-test is determined by the formula:
All the properties of the residual component are satisfied and, consequently, the model is adequate. The coefficient of multiple correlation (correlation index), squared ($R^2$), is called the coefficient of determination.

$$R^2 = 1 - \frac{\sum e(t)^2}{\sum (y_i - \bar{y})^2}$$

(5)

This coefficient shows the variation of the effective trait share under the influence of the studied factors, that is, determines how much of the variation of the trait $Y$ is taken into account in the model and is due to the influence of factors on it. So, $R^2 = 0.6074$. Consequently, about 60.74% of the variation of the dependent variable is taken into account in the model and is due to the influence of the included factor. This is a fairly high value, characterizing the high quality of the model.

### Table 6. Calculation of the Durbin-Watson coefficient.

| №  | $Y_i$  | Predicted $Y_i$ | $E_i^2$ | $(E_i-E_{i-1})^2$ |
|----|--------|----------------|--------|------------------|
| 1  | 812.6  | 828.4          | 250.4  | -                |
| 2  | 887.4  | 884.3          | 9.5    | 357.2            |
| 3  | 904.7  | 839.7          | 4227.4 | 3836.9          |
| 4  | 925.6  | 907.2          | 338.0  | 2174.7          |
| 5  | 915.5  | 997.0          | 6641.0 | 9975.4          |
| 6  | 1098.7 | 1080.6         | 327.5  | 9918.2          |
| 7  | 849.9  | 970.3          | 14491.3| 19176.1         |
| 8  | 1152.8 | 1039.7         | 12796.3| 31522.5         |
| Sum| 7547.2 | 7547.2         | 39081.4| 76961.0         |

*Developed by the author.

We will verify the significance of the regression equation on the basis of a calculation of Fisher's $F$-test:

$$F = \frac{R^2 / k}{(1 - R^2) / (n - k - 1)} = 101.00$$

The table value of the $F$-criterion for the confidence probability 0.95 $v_1 = k = 1$ и $v_2 = n-k-1 = 8-1-1 = 6$ and $v_1 = v_2 = n-k-1 = 8-1-1 = 6$ is 4.965 (found using the FINV function). Since $F_{calc} > F_{table}$, the regression equation should also be considered adequate and meaningful. The significance of the coefficient (except free) of the regression equation $a_1$ is estimated using Student's $t$-test. Estimated values of $t$-statistics: $t_{a2} = 4.7835$; $t_{a3} = 0.8872$; $t_{a5} = 7.5341$. The tabular value of the $t$-test at a significance level of 5% and degrees of freedom (8-1-1 = 6) is 2.365. Since $|t_{calc}| > t_{tabl}$, the coefficients $a_2$, $a_3$ are significant.

As for adequate models, it makes sense to set the task of estimating their accuracy. The accuracy of the model is characterized by the magnitude of the deviation of the model output from the actual value of the simulated variable. As a statistical indicator of accuracy, we use the indicator of the average relative error of approximation, which is calculated by the formula:

$$\bar{e}_{rel} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{Y - \hat{Y}}{Y} \right| \cdot 100\% \approx 11.6$$

The obtained value of the average relative error indicates a good level of accuracy of the constructed model, since less than 15%.

### 2. Conclusion

The resulting regression equation for the dependence of the variable $Y$ on $X_2$, $X_3$, $X_5$ has the following form:

$$Y(x) = 350.733 + 0.0206 \times X_2 + 0.00456 \times X_3 + 0.0348 \times X_5$$
The coefficient $a_2 = 0.0206$ shows that with the growth of GDP, the capacity of banks to issue investment loans significantly increases. The coefficient $a_3 = 0.00456$ indicates that this factor, connected with the growth of the population's labor payment (factor $X_3$), as a source of savings, acquires significant weight in case of general macroeconomic instability, inaccessibility of the euromarkets and in conditions of general lack of liquidity. In the conditions of narrowing of the sources of raising funds, the stability of this factor may be significant for the functioning of banks. The increase in the gross profit of Russian companies (factor $X_5$) leading to an increase in balances on bank accounts and the coefficient $a_5 = 0.0348$ shows that this factor has a positive impact on the growth of the investment potential of the banking system due to the fact that the access of banks to domestic sources of resources, as client capital.

After checking all the main criteria, it can be concluded that the constructed model approximates correctly the initial data (the relative error is 11.6%), therefore, it can be used for further analysis and forecasting of the investment potential of the banking sector.

Thus, the investment potential model confirms a significant dependence of the development dynamics of the banking system on the influence of external macroeconomic factors that can both limit and expand these objective opportunities. The investment potential of the banking sector is able to develop only in conditions of the country's economic growth and the creation of favorable macroeconomic and institutional conditions. In an uncertain economic situation, both enterprises and banks can not build their investment strategies. An essential condition for influence is the behavioral tendencies of the population associated with the credibility of credit institutions and the degree of satisfied consumer demand for banking services. In the conditions of narrowing of foreign capital markets, the growth of the investment potential of banks can rely mainly on internal sources, which requires the development of new financial instruments and technologies to attract savings and savings of economic entities.

3. References
[1] Ajupov A A, Kurmanova L R, Abelguzin N R, Kurmanova D A 2016 A comprehensive assessment of the financial security of banking system // International Business Management 10(23) 5620-5624
[2] Gerasimenko K V 2016 Investment potential of the Russian financial market: sources, structure, dynamics Fundamental research 2 (part 2) 352-356
[3] The concept of long-term social and economic development of the Russian Federation for the period up to 2020: approved by the decree of the Government of the Russian Federation No. 1662-r dated 17.11.2008
[4] Ajupov A A, Kurmanova L R, Khabibullin R G, Kurmanova D A 2017 Evaluation of financial stability of Russian companies Journal of Engineering and Applied Sciences 12(19) 4905-4907
[5] Safiullin M R, Elstina L A, Shakirova A I 2012 Evaluation of business and economic activity as a short-term forecasting tool Herald of the Russian Academy of Sciences 4 290-294
[6] Yusupov K N, Yangirov A V, Akhunov R R, Toktamyshova Y S 2014 The region reproductive potential structure pyramid Life Science Journal 11 403-405
[7] Omelchenko A N, Khrustalev O E 2010 Methods of increasing the credit and investment potential of the banking system of Russia Finance and credit 17 (40)21-30