

# Methodology of model structure choice in logistic modelling

Polyakov L. Konstantin

candidate of technical sciences, associate professor  
National Research University “Higher School of Economics”  
Polyakov.kl@hse.ru

Chuvakov Dmitry

Master of Arts  
National Research University “Higher School of Economics”  
Dmchuvakov@gmail.com

**Abstract.** The methodology, used in this research, allow to identify the structure of non-linear connections in the models of statistical interconnections. The empirical base of research is the reasons for revoking of licenses of Russian banks from 01.2013 to 12.2015 in two groups — due to economic reasons and laundering of money. Models of binary choice for the forecast of revoking licenses were built with the help of the fractional polynomial regression. Models were built for above-mentioned groups. Each model contains financial and social variables. The latter have been got with the help of networks theory from banks relationship analysis.

**Keywords:** bank, license revocation, bank default probability models, binary choice models, multivariable fractional polynomial regression, social capital, network theory.

## 1 Introduction

The aim of this paper is to develop the methodology of modeling non-linear statistical interactions between various statistical indicators of economic activity. Authors model the probability of revoking banking license on the basis of factors, characterizing social capital of the bank. The methodology is based on the conception of multivariable fractional polynomial regression. In addition, this paper offers the original methodology of evaluation of social capital.

Social capital can be considered as an asset, as it can generate profit for an organization. It influence its sustainability and success. The main problem is to identify the functional form of impact of social capital on the financial result. Any organization has many partners, build many formal and informal relations. In fact, the firm is the part of a huge net, with the help of witch the firm can find many new partners. This is all part of social capital.

Bank, like other organizations, is the part of a net, inside which it interacts with partners, use it to solve problems, achieve the objectives, and gain additional benefits. Such nets can be analyzed with the help of network theory, analytical apparatus of which can be used to measure quality and quantity of connections, and the position of the organization with respect to others.

## 2 Literature review

In the article «Early warning of bank failure. A logit regression approach» by Daniel Martin, 1977, author defined 25 financial indices. There were about 5700 banks in the sample, all were in the reserve system of the USA for the period 1970-1976. Daniel Martin showed, that in the period of general instability in the banking sector or crisis financial indices were good indicators of risks. Also, author showed, that logit model determined fail banks in most cases.

In the article, «A factor-analytic approach “To bank condition”» by Robert Craig West, 1985, information from auditing organization was used in addition to financial indices. Author showed, that this information helps to determine problem banks more accurately.

In the article, «Models of reasons for revocation of licenses of Russian banks» by A. A. Peresetskiy, all banks were divided into groups: functioning, the license was withdrawn due to money laundering, the license was withdrawn due to the violation of the Federal Law and unauthenticity, financial insolvency and voluntary license. 5 macroeconomic and 14 financial variables were used. It was shown that in the model of binary choice, significant variables are different in different groups.

In the article, «The Value of Human and Social Investments for the Business Performance of Start-ups» by Niels Bosma, Mirjamvan Praag, Roy Thurik and Gerrit de Wit, 2004, authors considered the impact of social capital on the success of start-ups in the Netherlands in 1994-1997. Spouse support, use of commercial relations and connection with other entrepreneurs through informal channels were used as proxies for social capital. Authors got significant associations between the first 2 indices and the income, number of employees and the life time of the firm.

In the article, «Cents and Sociability: Household Income and Social Capital in Rural Tanzania» by John F. Padgett and Cristopher K. Ansell, 1993, the role of civil and social norms, and voluntary participation in different groups were taken as proxies for social capital. The research showed the positive connection between the social capital and the income of households.

In the article, «Robust action and the rise of the Medici, 1400-1434» by John F. Padgett and Cristopher K. Ansell, 1993, author showed that the rise use of net can lead to the increase of the position of an individual. Medici did not create net, but managed to use it for getting huge political power among rich families of Florence.

### 3 Data description

Monthly banking reporting of 704 banks were got from open sources for the period 2013-2015, 36 periods, financial variables were taken from this base. For the period 151 licenses were revoked, including 51 because of laundering of money. Hence, 2 samples would be considered separately (with laundering of money and without).

At the base of articles devoted to financial analysis, 10 financial variables were chosen: asset ratio, coverage ratio, loan activity ratio, customer base utilization ratio, customer base ratio, external financing level, capital adequacy ratio, current liquidity ratio, long-term liquidity ratio, working assets security ratio.

The base (1.6 million of agents and 60 million of records), which contains data from news and informational web sites for the year 2013, was used to get 33000 agents, which are banks or their partners. Then net was build. After that, 4 variables were got. They characterize the position of the bank in the net and its connections. These variables were taken as proxies for social capital of banks.

The first variable is degree centrality (Deg\_c) – sum of all connections of the point. More complex is betweenness centrality (Bet\_cen) – sum of ratios, how often the point is on the shortest pass between 2 points to the amount of all shortest pass between 2 points. The third variable is closeness centrality – the average way between the current point and all other points in the net. The formula: nominator – the amount of all points minus 1, denominator – sum of all shortest passes to other points. The last variable was introduced to detect the intensity of use the net (Am\_Inter). It is the sum of all straight interactions of the point and its neighbors.

### 4 Building models

Fractional polynomial regression was used to identify non-linear connections in logit models of panel data. The method is based on the automatic choice of complex non-linear dependence from the determined class, according to the data properties. Then, results are analyzed with the help of contribution functions, which characterize the direction of change of probability with respect to change in the variable.

6 models were built in total, 3 for each sample, using 3 pairs of social variables (Deg\_c, Am\_Inter), (Bet\_cen, Am\_Inter), (Clos\_c, Am\_Inter), and all financial measures. 4 models have significant social variables.

For the sample without laundering of money, one model was got (Picture 1, Model 1).

At the graph 1 we may see that betweenness centrality, which shows the convenience of use by partners the ability of the bank, as a «staging post» at the connection with other agents, has the point of changing the direction of influence, if the variable is small, the probability of the revoke a license increases. This situation can relate to the movement of capital, the more connections, the more flows go through the bank, as a result bank gain additional benefits.

At the graph 2 we can see unusual interconnection; the probability rises with respect to this variable (Am\_Inter). Also, we can see the switching at small values of variables. We can say, that if the activity of the bank increases, the probability of the

loss of license rises. This can be explained by the situation, when credit activity increases, bank give more money, risks increase.

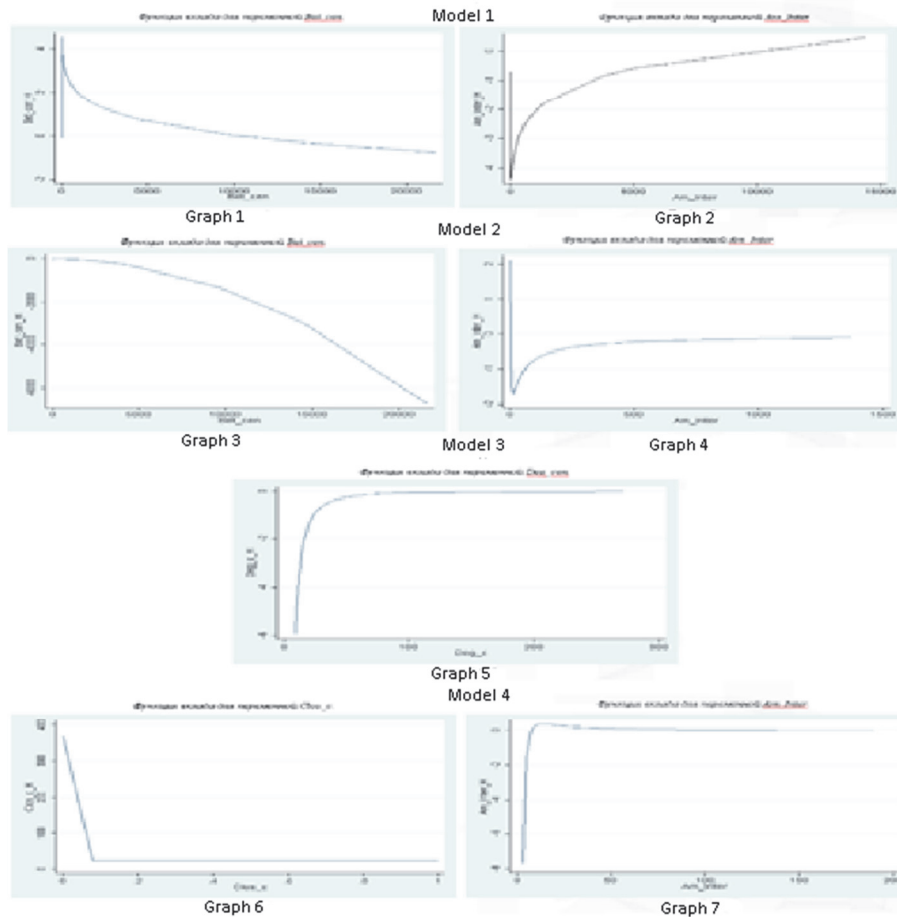


Fig. 1. Picture 1

For the sample with laundering of money, three models were got.

At the Picture 1, model 2, contribution functions for  $Bet_{cen}$  and  $Am_{cen}$  are presented. First function (Graph 3), like in the previous case, decreases, but at all values of the variable, and the concavity is different. Hence, at rather small values, the fall of probability is not so fast, as in the previous case. For the second function (Graph 4), the situation is different. There is switching at small values the probability of revoking rises, however after some point, there is about no relationship. This can be explained by the situation, when economic risks stop to have high influence.

Model 3 has only one significant social variable – degree centrality ( $Deg_c$ ). Graph 5 shows the positive connection between the variable and the probability at small values, and about no relations at big values. This situation is explained by the fact that

if the bank is created for laundering of money, it has not many connections, as the most of all operations are done for some small number of agents.

The model 4, Graph 6 shows that at small values of closeness centrality, the probability declines, and then when value increases, probability does not change. This situation can be interpreted by the following way: banks, that launder money, works almost always only with its closed circle of agents.

The behavior of contribution function of Am\_Inter from graph 7, reminds the situation from model 2. At small values of variable, there is rise, then switching and fall, after that function becomes about horizontal.

## 5 Conclusion

It was demonstrated the effectiveness of above described methodology of model structure choice. It was done on the basis of the modelling of the probability of license revocation for Russian banks. The methodology allowed detecting a non-linear influence of several parameters on the probability of license revocation. In this research, some financial indicators and variables that characterizes social capital were used as independent variables. The latter were got with the help of network theory. There was used the model of binary choice for panel data with application of multi-variable fractional polynomial regression. There were discovered statistically significant non-linear interconnections of social variables and the probability of licenses revocation in 4 out of 6 models. This effect was discovered in one model for banks that lost license only due to economic reasons, and in three models for banks that lost licenses because of laundering of money.

## References

1. Bat Batjargal, Internet entrepreneurship: Social capital, human capital, and performance of Internet ventures in China // *Research Policy* 36 (2007) 605–618
2. Charles Perreault, Gabrielle A. Brenner, Teresa V. Menzies, Louis Jacques Filion, Charles Ramangalahy, Social Capital and Business Performance: Ethnic Enterprises in Canada // *International Journal of Business and Globalisation*, Vol. 1, no 2, 2007, 145-160
3. Daniel Martin, Early warning of bank failure. A logit regression approach. // *Journal of Banking and Finance* 1 (1977) 249-276
4. Deepa Narayan, Lant Pritchett, Cents and Sociability: Household Income and Social Capital in Rural Tanzania // *Economic Development and Cultural Change* Vol. 47, No. 4 (July 1999), 871-897
5. Ed A.W. Boxman, Paul M. De Graaf and Hendrik D. Flap, The impact of social and human capital on the income attainment of Dutch managers // *Social Networks* 13 (1991) 51-73
6. John F. Padgett and Cristopher K. Ansell, Robust action and the rise of the Medici, 1400-1434 // *American Journal of Sociology* Vol. 98, No. 6 (May, 1993), 1259-1319
7. Niels Bosma, Mirjam van Praag, Roy Thurik, Gerrit de Wit, The Value of Human and Social Capital Investments for the Business Performance of Start-ups // *Tinbergen Institute Discussion Paper*, No. 02-027/3 (2002)

8. Patrick Royston and Douglas G. Altman, Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling // *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 43, No. 3(1994), pp. 429-467.
9. Robert Craig West, A factor-analytic approach "To bank condition" // *Journal of Banking and Finance* 9 (1985) 253-266.
10. Thomson, James B. "Predicting Bank Failures in the 1980s,"// *Federal Reserve Bank of Cleveland, Economic Review*, vol. 27, no. 1, pp. 9-20, 03.01.1991.
11. Winnie Yip, S.V. Subramanian, Andrew D. Mitchell, Dominic T.S. Lee, Jian Wang, Ichiro Kawachi, Does social capital enhance health and well-being? Evidence from rural China // *Social Science & Medicine* 64 (2007) 35-49.
12. A. A. Peresetsky, Models of reasons for revocation of licenses of Russian banks // *Applied Econometrics* №30 (2) 2013 49-64.