

Conference Paper

Modeling of Credit Institution License Withdrawal Based on Panel Data

Domashova D. V.¹, Kripak E. M.², Pisarchik E. E.¹, and Ulanova Zh. U.¹

¹National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe shosse 31, Moscow, 115409, Russia

²Orenburg State University, Orenburg, Russia

Abstract

The research paper involves the construction of panel data binary response models to forecast the probability of a credit institution's license withdrawal based on its financial performance, including the construction of logit and probit models using various sets of source data offers a technique for shaping a general model.

Keywords: credit institutions, panel data, binary response models, license withdrawal

Corresponding Author:

Domashova D. V.

Received: 11 December 2017

Accepted: 20 January 2018

Published: 13 February 2018

Publishing services provided by
Knowledge E

© Domashova D. V. et al. This article is distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use and redistribution provided that the original author and source are credited.

Selection and Peer-review under the responsibility of the FinTech and RegTech: Possibilities, Threats and Risks of Financial Technologies Conference Committee.

1. Introduction

The banking system is crucial for stable and resilient national economy. However, according to the Central Bank [1], more and more credit institutions lose their licenses (Fig. 1).

Therefore, more effective identification of credit institutions, which are at high risk of losing their licenses, and visualization of research findings for prompt decision-making are relevant to the bank monitoring challenges of today. In this research paper, we are modeling the credit institution's license withdrawal using binary response models with panel data that make it possible to consider individual features of economic objects over time.

Forecasting the probability of a credit institution's license withdrawal is the right way to safeguard clients' money, increase their trust in government initiatives, recover national financial system, avoid federal law violations, and mitigate adverse consequences of a credit institution winding up.

OPEN ACCESS

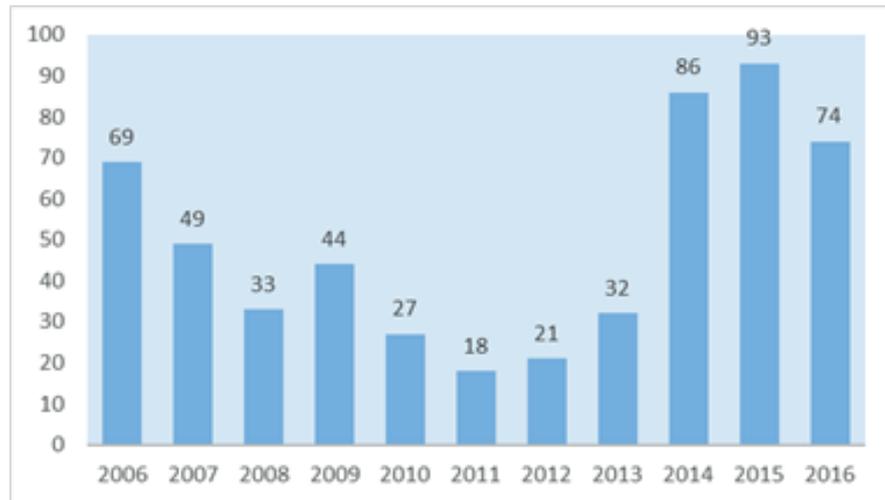


Figure 1: Credit institutions with withdrawn licenses over time.

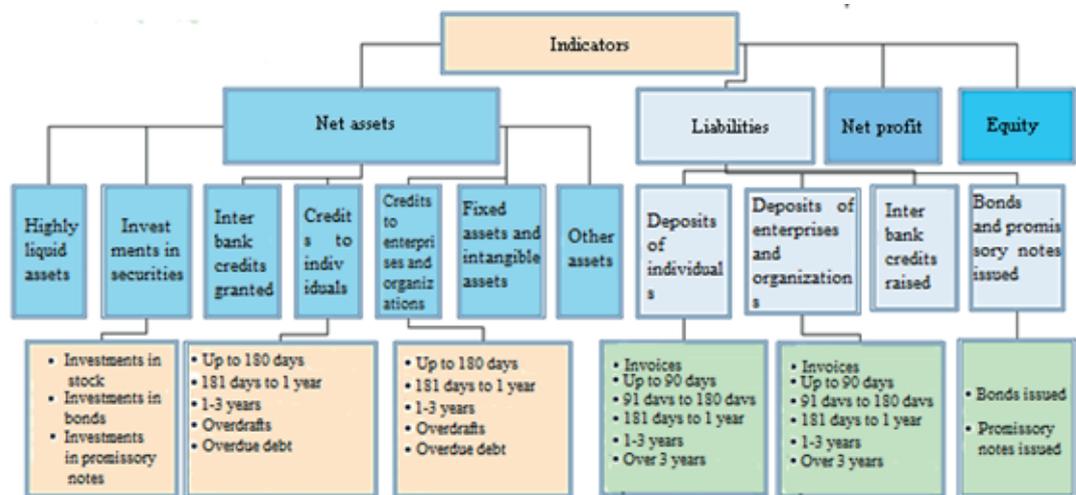


Figure 2: Financial reporting indicators.

2. Shaping and analyzing source data

The research paper takes monthly performance indicators of some 1,000 Russian banks over the period from 2008 to 2015, available on the banki.ru info portal [2], with the banks being ranked by a set of indicators based on the data from statement 101 posted by the Central Bank of Russia on its website, statement 102 (net profit) and 134 (equity). The indicator calculation technique is described on the website banki.ru. For the composition and hierarchy of analyzed indicators, see Fig. 2. The elaborated training sampling contains Russian credit institutions failing to comply with federal laws, covers the period from 2008 through 2013, and specifies the license withdrawal cause and details. The models are tested using the 2014-2015 data.

To take the macroeconomic environment into account, the research involved the following additional indicators: per capita GDP in Russia with account of the purchasing power parity (in current prices, RUB), equity investments in the Russian Federation (million RUB), annual average oil prices (RUB per barrel), and the refinancing rate, as posted by the Federal State Statistics Service [3] and the Central Bank on their websites. The above indicators may indirectly illustrate what is going on in Russia today.

To predict the license withdrawal probability, we use the binary response models with panel data; in particular, logit and probit regression models. The panel data are derived from the same economic units or objects being observed over consecutive periods of time [4]. Binary response models with panel data mean models where a dependent variable is of binary nature, i.e., may return 0 or 1, as in examples (1-2).

$$y_{it}^* = x_{it}\beta + \alpha_i + \varepsilon_{it}, \quad (1)$$

where x_{it} denotes the values of regressors on object i in moment of time t (dimensional vector k), β – regression coefficients, ε_{it} – respective error, α_i – individual effect of the object, y_{it} – dependent (target) variable of object i in moment of time t , returning the following values:

$$y_{it} = \begin{cases} 1, & \text{if } y_{it}^* \geq 0,5, \\ 0, & \text{if } y_{it}^* < 0,5; \end{cases} \quad (2)$$

where errors ε_{it} are independent in terms of i and t and similarly distributed, while values of α_i demonstrate individual differences between objects, with a range of change y_{it}^* being limited by various functions to $[0, 1]$.

Before the research, all indicators were normalized, centered and transformed into comparable values. As a target attribute of regression, we have a binary value 'closed' that equals 1 for banks losing their licenses in the following year and 0 for other banks. In addition, we excluded from the analyzed sampling the credit institutions of definitely good standing, which are considered to be backbone for national economy, such as Sberbank of Russia, Alfa-Bank, Promsvyazbank, etc.

3. Modeling

Initially, we employed StataCorp STATA software to build random-effects logit and probit models using data of the lowest hierarchy level and aggregated indicators. For a random-effects probit regression based on aggregated indicators, see Fig. 3. It is

Random-effects probit regression		Number of obs	=	5765		
Group variable: license_number		Number of groups	=	1122		
Random effects u_i - Gaussian		Obs per group: min	=	1		
		avg	=	5.1		
		max	=	6		
Integration method: mvaghermite		Integration points	=	12		
Log likelihood = -1090.3085		Wald chi2(13)	=	12.26		
		Prob > chi2	=	0.5066		

closed	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
vysokoi_aktivy	-.0305068	.1497007	-0.20	0.839	-.3239147	.2629011
kredity_fiz_licam	-.0661677	.529717	-0.12	0.901	-1.104394	.9720586
vydannye_MBK	-.3054309	.3324888	-0.92	0.358	-.9570968	.3462351
kredity_predp_i_org	.825099	.9671294	0.85	0.394	-1.07044	2.720638
vloj_v_cb	-.2892955	.4021621	-0.72	0.472	-1.077519	.4989277
vloj_v_kapitaly_drug_org	.228704	.492798	0.46	0.643	-.7371623	1.19457
osn_sredstva_i_nemat_aktivy	.458015	.4495831	1.02	0.308	-.4231518	1.339182
prochie_aktivy	.2674539	.1963	1.36	0.173	-.1172871	.6521949
vklady_fiz_lic	-1.061301	1.135575	-0.93	0.350	-3.286988	1.164386
sredstva_predp_i_org	.4047379	.6089804	0.66	0.506	-.7888417	1.598317
privlech_MBK	-.2988513	.4036463	-0.74	0.459	-1.089983	.4922808
vypush_oblig_i_vekselja	-.8666614	.3416899	-2.54	0.011	-1.536361	-.1969615
chist_prib	-.1884837	.3082793	-0.61	0.541	-.7927	.4157326
_cons	-1.739354	.0429364	-40.51	0.000	-1.823507	-1.6552
/lnsig2u	-12.118	19.65023			-50.63173	26.39574
sigma_u	.0023367	.0229587			1.01e-11	539214.2
rho	5.46e-06	.0001073			1.03e-22	1

Figure 3: Probit regression over the highest hierarchy.

worth noting that meaningful indicators include issued bonds and promissory notes only. According to the Wald statistics, the model is not meaningful in general.

The random-effects probit regression using indicators of the lowest hierarchy level is shown in Fig. 4, with only meaningful indicators being displayed.

The random-effects logit regression using aggregated indicators and those of the lowest hierarchy level returns almost the same results as the respective probit models do. However, neither model provided the required prediction quality.

4. Modeling with account for macroeconomic environment

Adding macroeconomic indicators improved the model adequacy. For example, the chi-square test returned better statistics for logit regression using the indicators of the lowest hierarchy level, as displayed in Fig. 5.

Thus, in addition to the above-listed coefficients, the refinancing rate also turned to be meaningful as it may indicate nationwide volatility.

Random-effects probit regression	Number of obs = 5,765		obs per group: min = 1, avg = 5.1, max = 6		LR chi2(29) ¹ = 66.61 Prob > chi2 ² = 0.0001	
group variable: license number	number of groups = 1,122					
	Coef.	Std. err.	z ³	P> z ⁴	95% Conf. Interval	
Credits to individuals with a maturity date of 1 to 3 years	-1.95	0.69	-2.83	0.01	-3.30	-0.60
Credits to individuals with a maturity date of more than 3 years	1.55	0.78	1.98	0.05	0.16	3.07
Credits to enterprises and organizations with a maturity date of up to 180 days	-0.72	0.33	-2.19	0.03	-1.36	-0.08
Credits to enterprises and organizations with a maturity date of over 3 years	-2.08	1.05	-1.97	0.05	-4.14	-0.01
Credits to enterprises and organizations with a maturity date from 90 to 180 days	0.50	0.19	2.58	0.01	0.12	0.88
Credits to enterprises and organizations with a maturity date from 1 to 3 years	-1.43	0.72	-1.99	0.05	-2.84	-0.02
Bonds issued	-0.90	0.44	-2.06	0.04	-1.76	-0.05
constant	-2.10	0.12	-17.14	0.00	-2.32	-1.85

¹—likelihood ratio chi-square test,
²— probability of receiving this chi-square value,
³— z-value, ⁴— double p-value.

Figure 4: Random-effects probit regression using indicators of the lowest hierarchy.

Random-effects logit regression	Number of obs = 5,765		obs per group: min = 1, avg = 5.1, max = 6		LR chi2(29) = 108.65 Prob > chi2 = 0.0000	
group variable: license number	number of groups = 1,122					
	Coef.	Std. err.	z	P> z	95% Conf. Interval	
refinancing rate	-1.50	0.33	-4.51	0.000	-2.15	-0.85
Credits with a maturity date of 1 to 3 years	-5.35	2.06	-2.59	0.010	-9.39	-1.30
Credits to enterprises and organizations with a maturity date of up to 180 days	-2.07	0.90	-2.30	0.021	-3.83	-0.31
Credits to enterprises and organizations with a maturity date of over 3 years	-2.08	1.05	-1.97	0.049	-4.14	-0.01
Credits to enterprises and organizations with a maturity date from 90 to 180 days	1.26	0.48	2.60	0.009	0.31	2.20
Bonds issued	-2.40	1.19	-2.01	0.044	-4.74	-0.06
constant	-5.56	0.66	-8.43	0.000	-6.86	-4.27

Figure 5: Logit regression over the lowest hierarchy, including macroeconomic indicators.

5. Modeling based on stratified samples

Since the sampling contains much more still licensed banks than those with licenses being withdrawn, it makes sense to elaborate subsamples of source data containing greater number of closed credit institutions, in order to improve the model quality.

Thus, with Statsoft Statistica software, we generated 12 subsamples, each containing 100 records, where credit institutions having their licenses withdrawn appear with a priori probability (p) from 0.9 to 0.1, while still operating credit institutions, from 0.1 to 0.9 respectively. The same models were built based on each of subsamples.

In particular, the random-effects logit model over the highest hierarchy, with macroeconomic environment involved, did not return meaningful coefficients against

Random-effects logit regression	Number of obs = 522		obs per group: min = 1, avg = 5.3, max = 6		Wald chi2(29) = 21.94 Prob > chi2 = 0.9874	
	number of groups = 98					
group variable: license number	Coef.	Std. err.	z	P> z	95% Conf. Interval	
Credits to enterprises and organizations with a maturity date from 181 days to 1 year	188.53	83.12	2.27	0.023	25.61	351.45
Credits to enterprises and organizations – overdrafts	48.17	17.86	2.70	0.007	13.17	83.17
Investments in bonds	387.29	131.92	2.94	0.003	128.74	645.85
Investments in promissory notes	-30.35	14.83	-2.05	0.003	128.74	645.85

Figure 6: Logit regression over the lowest hierarchy, including macroeconomic indicators, for the sampling with $p = 0,8$.

Random-effects probit regression	Number of obs = 522		obs per group: min = 1, avg = 5.3, max = 6		Wald chi2(29) = 21.94 Prob > chi2 = 0.9874	
	number of groups = 98					
group variable: license number	Coef.	Std. err.	z	P> z	95% Conf. Interval	
Credits to enterprises and organizations with a maturity date from 181 days to 1 year	99.54	38.34	2.60	0.009	24.41	174.68
Credits to enterprises and organizations – overdrafts	26.01	9.23	2.82	0.005	7.93	44.10
Investments in bonds	199.88	65.57	3.05	0.002	71.37	328.39
Investments in promissory notes	-16.45	7.31	-2.25	0.025	-30.78	-2.12
constant	-36.10	18.04	-2.00	0.045	-71.46	-0.74

Figure 7: Probit regression over the highest hierarchy, including macroeconomic indicators, for the sample with $p = 0,8$.

the sampling with license withdrawal probability of 0.8. The respective probit regression returned almost similar results. The random-effects logit model over the lowest hierarchy, with macroeconomic environment involved, against the sampling with license withdrawal probability of 0.8 is displayed in Fig. 6.

Similar probit regression is displayed in Fig. 7.

As can be seen from the table, the model coefficients are a bit more meaningful than those in the respective logit model.

6. General model building

The adequacy of each model will be assessed against the share of open and closed credit institutions being forecasted correctly. The bar chart in Fig. 8 shows how the prediction accuracy for open and closed banks depends on the respective a priori license withdrawal probabilities applied to samplings.

Here the conclusion to come: When the number of credit institutions with withdrawn licenses increases in the shaped sample, the models better predict banks to be closed, if compared with the models built on samples with lower license withdrawal probability.

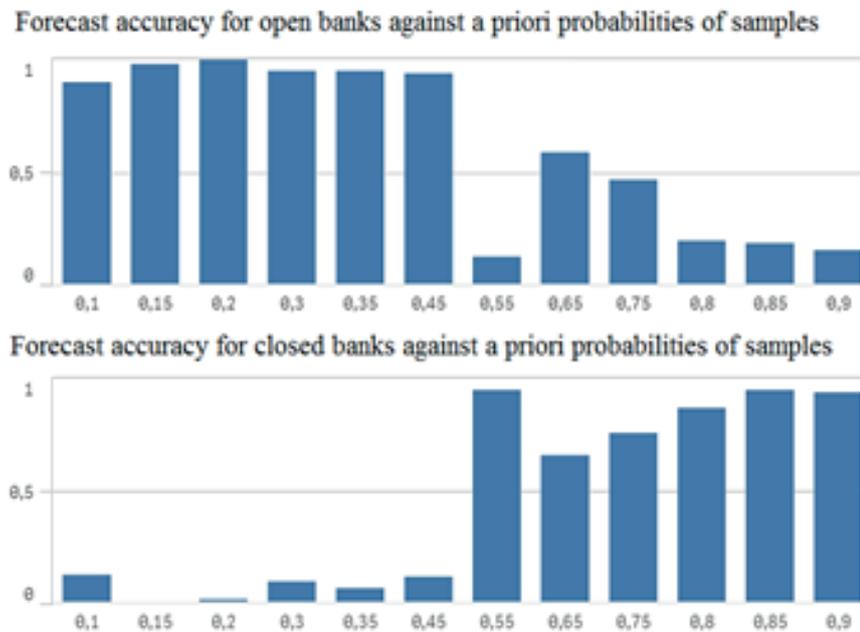


Figure 8: Bar chart showing how the prediction accuracy for open and closed banks depends on a priori license withdrawal probability.

Table 1 shows the overall prediction accuracy for the respective samples with the specified a priori license withdrawal probability.

TABLE 1: Sample prediction accuracy.

License withdrawal probability	0.1	0.15	0.2	0.3	0.35	0.45	0.55	0.65	0.75	0.8	0.85	0.9
Overall accuracy	0.8	0.86	0.87	0.84	0.84	0.82	0.21	0.59	0.49	0.27	0.26	0.24

Thus, modeling powered by stratified samples returned prediction models of higher quality.

The final decision-making may leverage a so-called general model, which accumulates the results obtained in models built.

General model construction comprises the following stages (Fig. 9):

- Generate stratified samples with different license withdrawal probability
- Build sets of models on each sample (logit regression over the lowest hierarchy, including macroeconomic indicators; logit regression over aggregated indicators, including macroeconomic indicators; probit regression over the lowest hierarchy, including macroeconomic indicators; probit regression over aggregated indicators, including macroeconomic indicators)

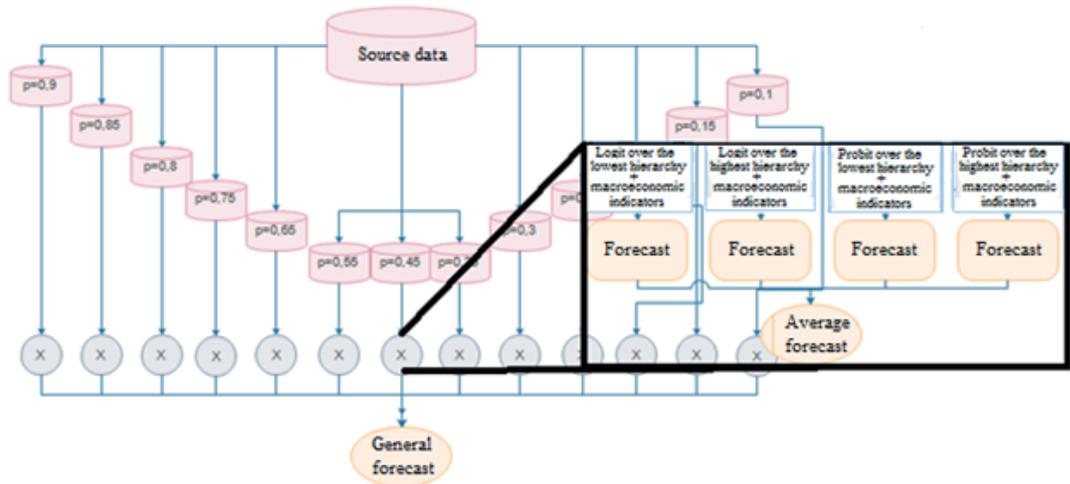


Figure 9: General model construction.

- Create a general model by simple voting

TABLE 2: Forecast accuracy of the general model.

		Actual data		
		closed	open	
Forecast data	closed	54	251	
	open	46	514	
	total	100	765	865
accuracy		0.54	0.67	0.66

The general model construction algorithms are still to be discussed. The general model results can be adjusted by distributing weight between the models, setting thresholds for function activation and/or varying the sample sets, with a proper understanding of which error (of the first or second type) is considered to be less admissible.

7. Conclusion

The research has resulted in an advanced toolbox that can be used by both bank supervisory authorities to forecast credit institution license withdrawal and other banks to assess their competitive landscape.

As part of the research, we have built the panel data binary response models and elaborated the system for aligning the forecast results obtained from the models built. The license withdrawal probability for each sample goes to the final forecast with a

certain weight, thus shaping a general integrated indicator. As an option to improve the model quality, you can either use a weighted voting method, or elaborate another technique to aggregate the forecasts from the models built – a way that may require further research.

Acknowledgements

This work was supported by Competitiveness Growth Program of the Federal Autonomous Educational Institution of Higher Education National Research Nuclear University MEPhI (Moscow Engineering Physics Institute).

References

- [1] Central Bank of the Russian Federation URL: <https://www.cbr.ru/>
- [2] Banki.ru info portal URL: <http://www.banki.ru/>
- [3] Federal State Statistics Service of the Russian Federation URL: <http://www.gks.ru/>
- [4] J.R. Magnus, Econometrics. Introduction Course: Manual. 6th edition, revised and added // P.K. Katyshev, A.A. Peresetsky. – Moscow: Delo Publishing House, 2004 – 576 pages