ORIGINAL PAPER

DOI: 10.26794/2587-5671-2024-28-3-94-108
UDC 336(045)
JEL G32, C51, C53

CC BY 4.0

Simulation of the Bankruptcy Event of Companies Associated with a Business Group

V.V. Lopatenko[a], **A.M. Karminsky**[b]
[a, b] National Research University Higher School of Economics, Moscow, Russia;
[a] PJSC Sberbank, Moscow, Russia

ABSTRACT

The **purpose** of the study is to determine the influence of a business group on the assessment of the borrower's creditworthiness, as well as to identify the most significant credit risk factors. Despite the fact that creditworthiness assessment is widely disseminated in both domestic and foreign literature, the impact of the consolidated group in the context of this problem is practically not mentioned. The authors use a statistical modeling **method** using logistic regression. The variable models are based on the annual financial statements of both individual companies and business groups. To select factors and build a model, approaches used in statistics and machine learning were used to obtain unbiased and effective estimates, independent of the sample generating these estimates. **Analyzed data** of 8691 companies providing annual financial statements in accordance with Russian accounting standards from 2015 to 2021. The total sample size was 22 201 observations. The number of bankruptcy events in the sample is 238 observations. Variables calculated from consolidated financial statements in accordance with international standards were used as information about the group. Various views on the concepts of "business group" and "holding" in the domestic literature are considered and systematized. Features of the behavior of companies united in groups are given. Variables associated with the business group that are significant in assessing the probability of bankruptcy of individual companies have been identified. Various specific aspects of the activities of companies associated with the group are mentioned. A statistical model is constructed to confirm a number of hypotheses, which is subject to verification and analysis. The bankruptcy event is used to determine the significant deterioration of a company's creditworthiness. It is **concluded** that the use of group reporting data can improve the quality of model prediction for companies associated with a business group.

Keywords: credit risk; bankruptcy; bankruptcy risk factors; business group; holding; IFRS; logistic regression; machine learning

For citation: **Lopatenko V.V., Karminsky A.M. Simulation of the bankruptcy event of companies associated with a business group. *Finance: Theory and Practice*. 2024;28(3):94-108. (In Russ.). DOI: 10.26794/2587-5671-2024-28-3-94-108**

© Lopatenko V.V., Karminsky A.M., 2024

INTRODUCTION

Different Views on the Definition

There are special categories of companies whose creditworthiness assessment, in isolation from external factors, does not always give a reliable assessment of the current economic condition. Companies affiliated with the group can be identified in one of these categories. There is no single view on the definition of a group of companies in Russian legislation or literature. For example, according to International Financial Reporting Standards (IFRS) 10 "Consolidated Financial Statements" (hereinafter — IFRS 10), the group of companies is the parent company and all its subsidiaries. The parent company is understood as an enterprise that controls one or more subsidiaries through:

- rights that provide an opportunity to manage the significant activities of a subsidiary;
- a majority share in a subsidiary that provides voting rights to the extent sufficient to determine operational and financial policies;
- opportunities to use their powers in relation to the company in order to influence the income of the investor — the parent company.

On the other hand, there are no concepts of "group" and "consolidated reporting" in Russian accounting standards. In accordance with the requirements of the Federal Law of 27.07.2010 № 208-FZ (edit. from 26.07.2019, amend. from 07.04.2020) "On consolidated financial statements", consolidated statements are prepared in accordance with IFRS standards and are mandatory only for a narrow circle of legal entities — banks, insurance and public companies.

Nevertheless, the concept of "banking group" is legislated for banks, which means an association of legal entities (hereinafter — legal entities), in which one or more legal entities are under the control or significant influence of one credit institution.[1] At the same time, the definitions of "control" and "influence" refer to them in accordance with IFRS.

The concept of "holding" has become a little more widespread, which is often used in a synonymous sense with the concept of "group of companies". Despite the fact that it is also not reflected in Russian legislation, there are many works defining the holding as a group of persons in which the holding company has the right to manage the activities of other members of the holding due to the prevailing participation in their authorized capital or otherwise [1–3]. Anyway, in the literature you can find many interpretations of the meaning of this entity, but they all boil down to several common features inherent in all of them:

1) availability of parent (independent) and subsidiary (dependent) companies;
2) presence of control or influence by the parent company over subsidiaries;
3) involvement in a common activity is also highlighted.

Note that these criteria lead us to a very close definition, already given by the IFRS 10 standard.

Summarizing the above, we can distinguish the criteria specific to both the concept of "group of companies" and the concept of "holding" (*Table 1*).

CHARACTERISTICS OF COMPANIES IN THE FRAMEWORK OF BUSINESS GROUPS

After formalizing the definition of companies affiliated with the group, it is necessary to identify behavioral factors specific to them. Among the most characteristic are the following:

1. Companies affiliated with the group can use internal resources to overcome difficulties in gaining access to external sources of financing [4]. Such an exchange of

---
[1] Federal Law from 02.12.1990 No. 395-1 (ed. from 29.12.2022) "Banks and banking activities".

Table 1

Business Group Criteria

Criterion	Compatibility with the concept of "holding"	Compatibility with the concept of "group of companies"
Presence of parent and subsidiary companies	Yes	Yes
Control by the parent	Yes	Yes
Control criteria: dominant share in the capital, executive functions	Yes	Yes

Source: Compiled by the authors.

resources within the group often goes beyond the financial resources of the group and is an important factor determining the result of the firm's activities [5].

2. Groups can be considered as a set of implicit agreements through which companies affiliated with the group support each other as needed. According to these mutual agreements, groups are effectively involved in the risk distribution process, thereby preventing the default of individual participants subject to temporary liquidity shortages [6].

Based on the above, group affiliation and the model of behavior of the company within the group play an important role in assessing the credit risk of companies belonging to the group. However, it also depends on a number of factors related, for example, to the role of the company in the group, its share in the financial result of the business, the ownership structure.

Another important characteristic of such companies is that owners can control a large number of companies, limiting their exposure to risk through limited liability — as parent companies may not be responsible for the obligations of their subsidiaries and may decide not to support the subsidiary in a difficult situation where it is expensive for the entire group of companies (the so-called selective default option). Parent companies adhere to strategic behavior that they seek to deliberately use limited liability to protect themselves from the obligations of their subsidiaries. Therefore, it is unlikely that these companies will be able to help out their subsidiaries that are in a difficult financial situation. A similar business model is discussed in some articles. According to it, the beneficiary creates several firms controlled only by him, and by distributing assets, profits and losses avoids taxes and leads to the bankruptcy of firms that are debt centers [7]. Following these arguments, the connection with the group should not, in fact, have a meaning to predict default.

On the other hand, bankruptcy courts can be raised to the level of the parent company, in which case the latter may be held responsible for the debts of subsidiaries. Moreover, the default of a subsidiary may expose the maternity to additional non-trivial costs associated with reputation risks/restriction of access to capital, and thus generate a series of defaults within the group. As a result of these costs and the probability of escalation, the parent company can still decide to support the subsidiary, and on this side, affiliation with the group will already matter in the context of the probability of the company defaulting [8].

DESCRIPTION OF THE APPROACH TO BANKRUPTCY

To simulate the bankruptcy event of a company associated with the group, a model was built on combined financial data: both the group and the company itself. Consolidated statements in accordance with the IFRS standard were used as the group's statements, the statements prepared according to Russian accounting standards were used for the statements of individual companies.

Companies associated with the group have been identified as companies over which the relevant IFRS issuer is able to have significant influence. It manifests itself in the presence of one of the following connections between companies:

- the presence of a controlling stake in the controlling company in the dependent;
- the ability to manage a significant activity of a dependent company;
- the situation in which the controlling and dependent company has a general director;
- a situation in which the CEO of the controlling company is the founder of the dependent.

The simulated event was defined as the arbitration court's acceptance of an application for bankruptcy proceedings against the company within a year after the reporting date. For this purpose, data on arbitration proceedings were used. The initiation of the bankruptcy procedure meant the court's acceptance of an application for bankruptcy of the debtor, and if the application was returned to the plaintiff, it was assumed that the event did not happen. Observations on which bankruptcy proceedings have already been initiated on the reporting date were also excluded from the sample.

Risk factors calculated using the relevant company reporting indicators were used for modeling. All factors can be divided into several groups, depending on what aspect of the borrower's economic activity they characterize:

- Creditworthiness — factors that have in the calculation of articles related to borrowed obligations;
- Size — factors based on the company's revenue or assets;
- Profitability — factors in which different types of profit are correlated with other reporting items;
- Liquidity — factors that characterize the company's ability to have prompt access to funds;
- Financial stability — factors operating with equity and balance sheet items;
- Activity and turnover are factors that demonstrate the financial result of the company in dynamics.

A full list of factors is presented in *Table 2*.

For a full list of indicators, an independent analysis of factors was carried out to assess the ranking capacity and stability of the basic model based on one factor. On the basis of this analysis, a list of factors that meet the criteria for stability and predictive power was selected, on the basis of which correlations were further evaluated, features were selected and the final model was built.

The Gini index was used to assess the discriminatory ability of the model,[2] describing how correctly the model organizes observations from best to worst.

The following describes in detail the step-by-step process of building a model.

Formation of a Sample for Modeling

The sample was based on financial reporting companies in accordance with international standards. All such companies with at least one report published since 2016 have been added to the sample. Methodologically, it was accepted that these companies are the parent company of the group of companies, while the corresponding reporting indicators represent the financial results of the group.

Further, companies related to them were found, publishing the results of financial activities in accordance with Russian accounting standards, while companies

[2] Gini Coefficient. From Economics to Machine Learning. Habr. Blog of Open Data Science. URL: https://habr.com/ru/company/ods/blog/350440/ (accessed on 10.12.2022).

Table 2

Long List of Factors

Factor code	Description
prof_btax_rev	Ratio of pre-tax profit to revenue
prof_btax_bal	Ratio of pre-tax profit to balance
prof_btax_turn	Ratio of pre-tax profit to assets
prof_net_rev	Ratio of net profit to revenue
prof_net_bal	Ratio of net profit to balance sheet
prof_net_turn	Ratio of net profit to assets
liq_abs	Complete liquidity
liq_inst	Instantaneous liquidity
liq_cur	Current liquidity
liq_short	Short liquidity
liq_mid	Average liquidity
debt_bal	Ratio of company debt to balance sheet
cur_asset_bal	Ratio of current assets to balance sheet
bal_long_liab	Ratio of company balance to long-term liabilities
notrnvr_asset_long_liab	Ratio of company assets to liabilities
profit_debt	Ratio of net profit to debt
profit_net_debt	Ratio of net profit to net debt
profit_liab	Ratio of net profit to liabilities
profit_net_liab	Ratio of net profit to net liabilities
profit_btax_debt	Ratio of pre-tax profit to debt
profit_btax_net_debt	Ratio of pre-tax profit to net debt
profit_btax_liab	Ratio of pre-tax profit to corporate liabilities
profit_btax_net_liab	Ratio of pre-tax profit to net liabilities
debt_rev	Ratio of company debt to revenue
net_debt_rev	Ratio of net debt to revenue
liab_rev	Ratio of company's liabilities to revenue
net_liab_rev	Ratio of Company Net Liabilities to Revenue
debt_due_turn	Accounts receivable turnover
acc_due_turn	Accounts payable turnover
cash_rev	Ratio of cash to revenue

Source: Compiled by the authors.

Table 3

Introductory Sample Passage for Modeling

Field name	Observation 1	Observation 2	Observation 3	Observation 4
inn_num	5 003 050 143	5 030 076 824	5 911 063 420	5 444 100 990
report_dt	2021-01-01	2022-01-01	2018-01-01	2018-01-01
prof_btax_rev	768	−836	0.016457	0.03806
prof_btax_bal	0.001463	−0.01739	0.118958	0.127227
prof_btax_turn	0.108367	−0.05188	0.10019	0.184778
prof_net_rev	614.25	−836	0.010861	0.02963
prof_net_bal	0.00117	−0.01739	0.078506	0.099045
prof_net_turn	0.108367	−0.05188	0.10019	0.184778
liq_abs	187	873	1094	4679
liq_inst	187	873	1094	3729
liq_cur	22 673	16 115	28 436	17 632
liq_short	22 673	16 115	26 372	16 086
liq_mid	22 673	16 115	28 425	17 632
debt_bal	0.050783	0	0	0
cur_asset_bal	0.010801	0.335205	0.783577	0.536025
bal_long_liab	19.69147	48 075	36 290	32 894
notrnvr_asset_long_liab	19.47879	31 960	7854	15 262
profit_debt	0.023048	−836	2849	3258
profit_net_debt	819	−836	2849	3258
profit_liab	0.023048	−836	2849	3258
profit_net_liab	819	−836	2849	3258
profit_btax_debt	0.028817	−836	4317	4185
profit_btax_net_debt	1024	−836	4317	4185
profit_btax_liab	0.028817	−836	4317	4185
profit_btax_net_liab	1024	−836	4317	4185
debt_rev	26 651	0	0	0
net_debt_rev	26 604.25	−873	−0.00417	−0.03391
liab_rev	26 651	0	0	0
net_liab_rev	26 604.25	−873	−0.00417	−0.03391
debt_due_turn	5621.5	15 242	0.096361	0.103741
acc_due_turn	113	14 400	0.069124	0.046109
cash_rev	46.75	873	0.00417	0.033913
report_type	RSBU	RSBU	RSBU	RSBU
bank_flag	0	0	0	0
cg_inn_num	7 751 188 020	7 718 560 636	6 607 000 556	4 205 003 440
prof_btax_rev_cg	97.97468	0.107876	0.201286	0.068
prof_btax_bal_cg	0.085656	0.087471	0.059612	0.100137
prof_btax_turn_cg	2.378571	0.231876	0.123405	0.168145
prof_net_rev_cg	84.3038	0.106889	0.145785	0.05442

Table 3 (continued)

Field name	Observation 1	Observation 2	Observation 3	Observation 4
prof_net_bal_cg	0.073704	0.086671	0.043175	0.08014
prof_net_turn_cg	2.378571	0.231876	0.123405	0.168145
liq_abs_cg	0.348837	0.214353	2.142049	1.944146
liq_inst_cg	0.348837	0.214353	2.142049	1.944146
liq_cur_cg	65.11628	1.434122	4.79237	4.480129
liq_short_cg	39.95349	0.385852	2.777121	3.20981
liq_mid_cg	42.16279	0.879116	4.490951	4.137845
debt_bal_cg	0.016478	0.431269	0.271597	0.370153
cur_asset_bal_cg	0.030987	0.37378	0.349867	0.476613
bal_long_liab_cg	6.829491	4.288029	3.076	2.984087
notrnvr_asset_long_ liab_cg	6.617867	2.685251	1.999809	1.561832
profit_debt_cg	4.472801	0.200966	0.158968	0.216505
profit_net_debt_cg	2.22E + 09	5.63E + 09	4.28E + 09	7.01E + 08
profit_liab_cg	0.501733	0.175503	0.108453	0.18152
profit_net_liab_cg	2.22E + 09	5.63E + 09	4.28E + 09	7.01E + 08
profit_btax_debt_cg	5.19812	0.202823	0.219488	0.270529
profit_btax_net_debt_ cg	2.58E + 09	5.68E + 09	5.9E + 09	8.76E + 08
profit_btax_liab_cg	0.583095	0.177125	0.149741	0.226814
profit_btax_net_liab_ cg	2.58E + 09	5.68E + 09	5.9E + 09	8.76E + 08
debt_rev_cg	18.8481	0.531874	0.917069	0.251358
net_debt_rev_cg	18.65823	0.462974	0.389038	0.11091
liab_rev_cg	168.0253	0.609041	1.344226	0.299803
net_liab_rev_cg	167.8354	0.540141	0.816195	0.159355
debt_due_turn_cg	21.55696	0.055125	0.15655	0.091433
acc_due_turn_cg	5.936709	0.115802	0.061109	0.089157
cash_rev_cg	0.189873	0.0689	0.528031	0.140448
bank_flag_cg	0	0	0	0
target	0	0	0	0

Source: Compiled by the authors.

in bankruptcy proceedings were excluded. A target event trigger indicator has been defined for each company. *Table 3* shows an introductory fragment of the sample for modeling.

The characteristics of the sample collected are presented in *Table 4*.

Selection of the Training Sample

In machine learning, the so-called deferred or control sample is usually used to assess the quality of the model. The idea is to get scores on one sample, called a training sample, and test on another, thereby confirming or disproving the hypothesis of the presence of

Table 4

Modeling Sample Properties

Property	Value
Number of different groups of companies	509
Number of associated companies	8 691
Number of related companies for which the target event occurred	236
Relative frequency of bankruptcy, %	1.07%

Source: Compiled by the authors.

the generalizing ability of the resulting model. Along with the described approach, a method of splitting called stratified cross-validation into k-blocks is often also used to select model parameters, when one of the samples, usually training, is divided k times into training and control samples, so that the control samples between the splits do not intersect. At the same time, k models are built on each of the samples, which are verified on each of the control samples. Next, the best one is selected from the obtained models, which is then checked on the control sample obtained from the initial cleavage [9]. This approach avoids retraining the model and building shifted estimates.

The sample obtained by separating 25% of the sample from the original sample was used as the control sample. When selecting, stratification by the target variable was used.

Transformation of Factors

Before the selection of risk factors, each factor was transformed using the *WOE* algorithm (Weight of Evidence) transformations [10, 11]. According to this algorithm, each factor is converted into a categorical variable so that a hypothesis of a statistically significant difference between the averages for the samples presented by each of the categories is executed for neighboring categories. Next, a WOE is calculated for each category — the value corresponding to this group and the next formula:

$$WOE_i = ln\left(\frac{Ngood_i / Ngood_{all}}{Nbad_i / Nbad_{all}}\right), \tag{1}$$

where $WOEi$ — value of the *WOE* indicator for the factor group with the ordinal number *i*; $Ngoodi$ — the number of observations for which the bankruptcy event was not realized in the factor group with serial number *i*; $Ngoodall$ — the total number of observations for which the bankruptcy event was not realized; $Nbadi$ — the number of observations for which the bankruptcy event was realized in the factor group with serial number *i*; $Nbadall$ — the total number of observations for which the bankruptcy event was not realized.

Factor Selection

To highlight the most stable and ranking factors in the model, the above-described cross-validation mechanism for 3 blocks was used. Previously, a training sample was allocated from the entire sample for development, which is 75% of the original sample size. Further, according to the algorithm, it was sequentially divided into 3 disjoint samples for training, which make up 75% of the original sample for training, and 3 disjoint control samples, which make up, respectively, 25% of the original sample for development, which do not overlap each other. Thus, the final check was performed on 3 independent samples with an empty intersection with each of the other 3. At the stage of independent selection of factors for each factor for each of the splits obtained using the stratified cross-validation algorithm into 3 segments, the following requirements were put forward:

- discriminating ability of at least 5% of Gini on the learning sample;
- discriminating power of at least 5% Gini in the control sample;
- statistical significance of constructed single-factor regression at the level of 99% on the training sample;
- statistical significance of the constructed single-factor regression at 99% on the control sample;
- re-learning: the absolute rating of the factor is not more than 10 points or 20% in relative terms.

The Gini index was used to measure the ranking level, it is measured on a scale from 0 to 100%. With regard to binary classification problems, it represents the degree of stratification of two classes by any feature. Using the probability predicted by the model as a feature, by calculating the Gini index, you can understand what proportion of the sample the model ranks correctly. The higher the value of this indicator, the higher the quality of model prediction. The value of this coefficient reaches a maximum at a value equal to one, when at a certain value of the predicted probability is reached, all observations belonging to the same class are less than/not greater than this value, and all observations belonging to the second class are not more than/less than this value. The value of the Gini coefficient will be zero if both classes in equal shares are present for any of the given values in the resulting split.

To measure the statistical significance of the coefficient obtained by constructing one-factor regression, the Wald test [12] was used, the mechanism of application of which consists in the calculation of statistics:

$$Z_w = \frac{\hat{\beta}}{E_{st}\left(\hat{\beta}\right)}, \tag{2}$$

where $\hat{\beta}$ — the resulting value of the coefficient before the variable; $E_{st}\left(\hat{\beta}\right)$ — standard regression coefficient error.

For the obtained statistical value, the corresponding p-value value was used, which was further compared with the threshold value of 0.01, in excess of which the factor was cut off from further analysis.

As a result, the factors that meet the obtained criteria for each of the five splits were selected.

Study of the Combined Influence of Factors

At this stage, the mutual influence of the resulting list of factors was investigated. Similar to the stage of independent selection of factors, stratified cross-validation into 3 non-overlapping segments was used to highlight the model list of factors. The variable exclusion procedure based on the calculated values of the Pearson correlation coefficient was first applied to the list of factors obtained at the previous stage, then the step-by-step regression algorithm was applied. The list of factors used to construct the outcome model was determined based on factors present in each of the three lists obtained after applying the step-by-step regression algorithm.

A correlation analysis was subsequently carried out [13]. If the correlation factor for a pair of variables exceeded 0.7, the factor less than the Gini index was excluded from further consideration.

Subsequently, a step-by-step regression algorithm was applied, which consists of the sequential inclusion of the most statistically significant factors at each step with the subsequent exclusion of the minor factors at every step. The statistical significance is determined on the basis of the p-value obtained by testing the zero hypothesis of the significance of zero of the linear regression factor before the relevant variable.

Model structure

After studying the combined influence of factors and identifying a list of factors for

Table 5

Model Properties

Factor	Coefficient value	Standard error	Z-value	p-value
Free indicator	−4.5339	0.087	−51.888	0.000
Accounts receivable turnover, WOE	−0.6841	0.237	−2.882	0.004
Ratio of pre-tax profit to balance sheet, WOE	−0.5421	0.142	−3.826	0.000
Ratio of group debt to balance sheet, WOE	−1.2964	0.299	−4.338	0.000
Ratio of group pre-tax profit to its revenue, WOE	−0.8544	0.251	−3.410	0.001
Ratio of pre-tax profit to net liabilities, WOE	−0.5720	0.120	−4.768	0.000
Ratio of net liabilities to net profit, WOE	−0.8827	0.129	−6.825	0.000
Ratio of non-current assets to long-term liabilities, WOE	−1.9480	0.504	−3.864	0.000

Source: Compiled by the authors.

modeling, a logistic regression model was developed. In the Logit model, the probability of the event is defined as:

$$p = \frac{1}{1+e^{-Z}}, \quad (3)$$

where

$$Z = \sum\nolimits_i X_i \beta_i + \alpha, \quad (4)$$

where X_i — independent factors, β_i — corresponding regression coefficients, α — free indicator.

Table 5 shows the characteristics of the built-in model, namely:

- the value of the regression coefficients before the relevant factors;
- the standard errors of regression factors used to verify the hypothesis of the equality of zero values of the coefficients;
- z-statistics corresponding to the zero-hypothesis, calculated in accordance with the Wald test;
- p-value relevant to the statistics.

In addition to the factors built on the financial statements of companies, the model included indicators calculated on the accounts of the controlling company. At the same time, one of them, namely the ratio of the debt of the group to the balance sheet, has the second most absolute value of the coefficient.

All coefficients are significant at 99%. The value of q-squares statistics was 248.44.

The Gini index score in the training sample for the built-in model was 59.08%. To assess the contribution of group factors, they were removed from the model, and then the regression factors were reassessed. The Gini

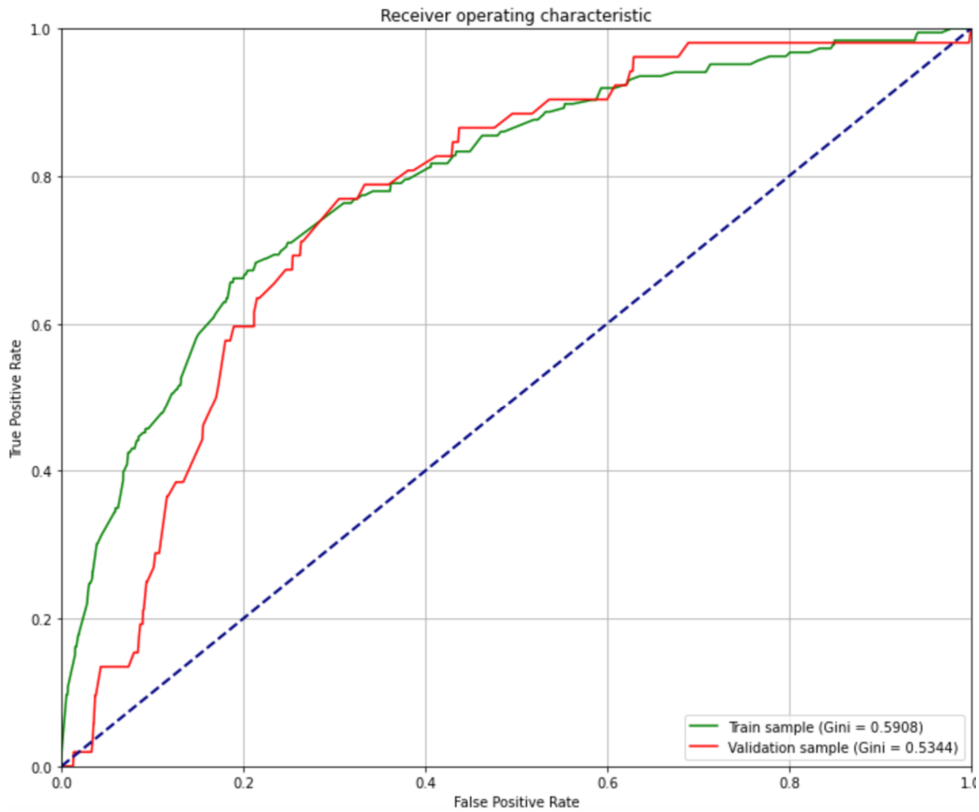

Fig. 1. **Model Discriminatory Ability**

Source: Compiled by the authors.

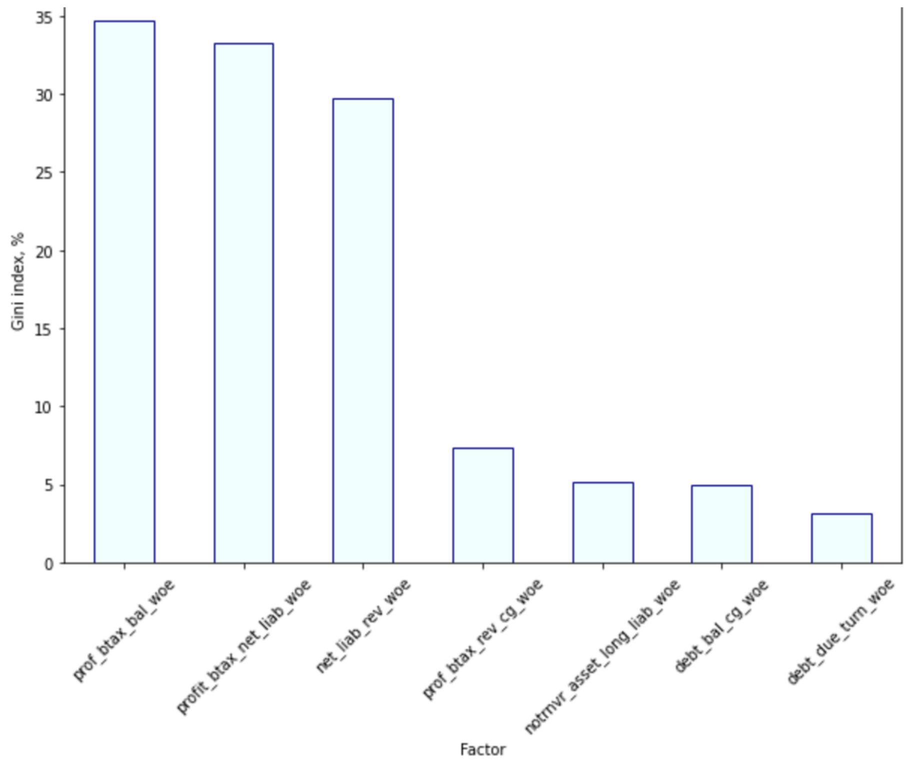

Fig. 2. **Discriminating Ability of Factors**

Source: Compiled by the authors.

Table 6

Table of *WOE* Values for the Factor "Ratio of Profit Before Taxes to Balance Sheet"

WOE value	Relative frequency, %
−1.288797	3.38
−0.700082	1.89
0.380075	0.84
1.369820	0.33

Source: Compiled by the authors.

Index declined to 56.27%, corresponding to a 5% deterioration in discrimination.

Validation and Analysis of Model

The model's structure was evaluated for suitability, and its created model's conformance with economic reasoning was examined using the validation sample. The sample was structured in such a way that companies associated with the same group were included either in the training or in the control sample, since this approach excludes the existence of dependent observations. The validation sample had the following characteristics:

- the sample amounted to 4 709 observations;
- the number of observations for which the bankruptcy event occurred was 52 observations.

The discrimination of the model in the validation sample was 53.44%. Similarly, the training sample excluded group factors and reassessed the model coefficients, which also led to a decline in the Gini index to 51.73%. *Figure 1* shows a comparison of the discrimination of the model in the educational and control samples (*Fig. 1*).

The discriminatory capacity of individual factors are presented in *Fig. 2*.

The highest predictability are the factors "Ratio of pre-tax profit to the balance sheet", "Ratio of pre-tax profit to net liabilities", " Ratio of net liability to revenue". It is important to note that each of these factors, independently of the others, identifies a third of all observations for which the target event occurred. From the point of view of the stability of the model, this is a positive feature: for example, in a situation where the model is applied to one of the factors with a high noise, the model will prove to be more stable compared to a model of similar discriminatory capacity built on a single dominant factor. Such a situation may occur in a number of cases:

- manual errors in reporting to the system;
- technical infrastructure errors;
- manipulation of individual financial statements.

The first factor is the ratio of the income tax to the balance sheet. As the value of this factor increases, the relative frequency of occurrence of the simulated event increases consistently from 0.32% to 3.32%. The resulting *WOE* values and their corresponding relative frequencies are presented in *Table 6*.

It is worth noting that the maximum value of all pairs of correlation coefficients does not exceed 0.7. All factors are significant at a 99% level. Additionally, a f-test was performed for the significance of the regression equation [14], the resulting statistical value of 29.86 corresponds to the p-value << 0,001, on the basis of which a conclusion was made about the importance of a regression equation.

In addition, the performance of the model was studied only on observations in which the controlling company itself went into bankruptcy. It is noteworthy that the discriminatory capacity of the model using group data, compared with the model that did not use group indicators, increased from 47.82% to 60.33% of Gini. *Fig. 3* shows comparative curves.

Speaking of the economic interpretation of this phenomenon, it can be said that the factors built on the accounts of the controlling company also indirectly model the bankruptcy event for that company. The bankruptcy of the

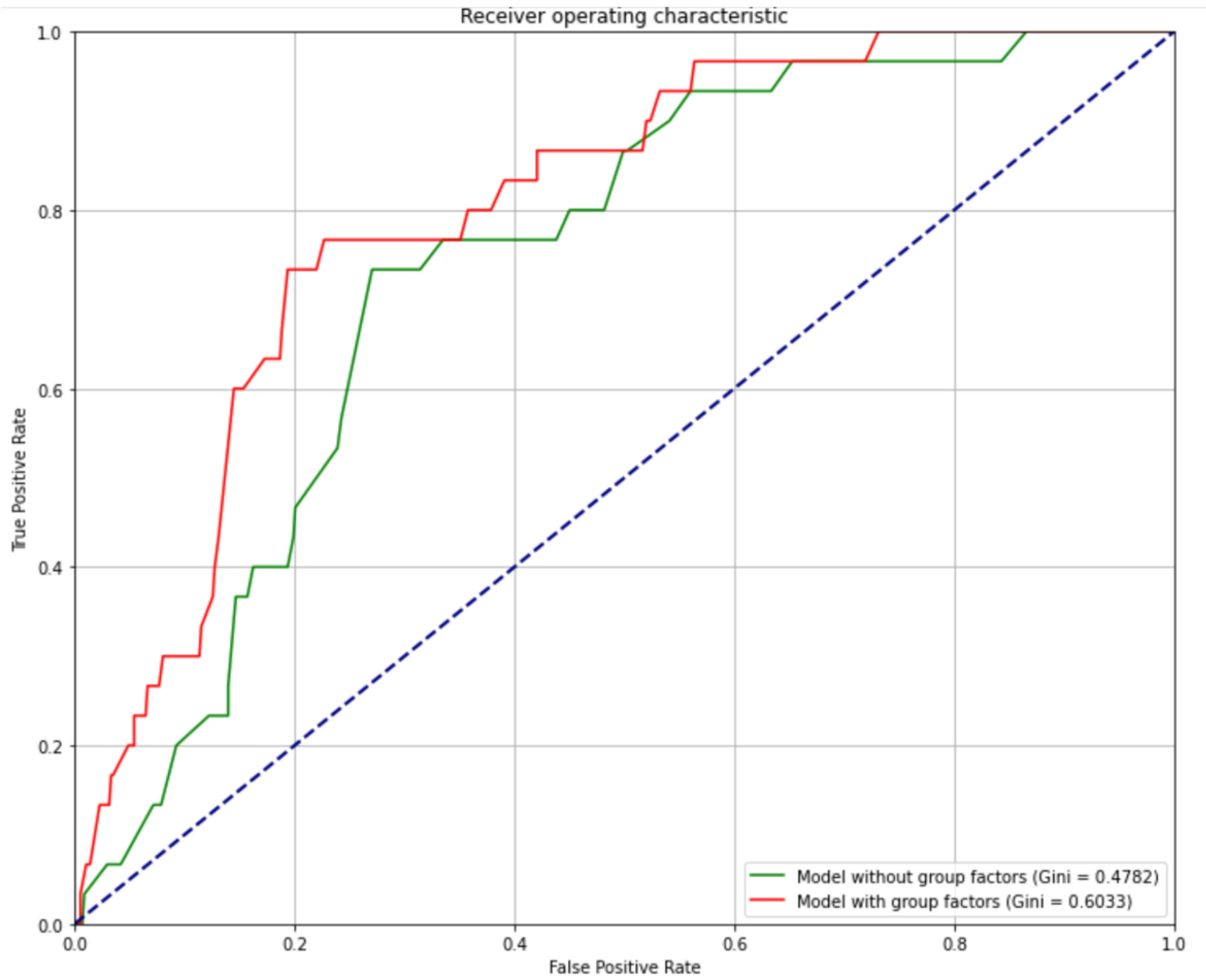

Fig. 3. **Comparison of the Discriminatory Ability of Models in the Bankruptcy of the Parent Company**

Source: Compiled by the authors.

parent company of the group often entails the effect of infection — the spread of the event to other companies. In a difficult financial situation, the controlling company, having access to the capital markets of dependent companies, uses their resources to save its own position, which is highly likely to result in the bankruptcy.

A comparison of discriminatory curves clearly shows that the second model is more comprehensive, and in the entire area of definition. In other words, no matter what threshold we choose for the model to work, the second model will always identify more "bad" companies, with always a lower percentage of false operations. This attribute indicates an increase in the completeness of the model, which may be important in a number of situations where the first and second types of errors are not equal.

CONCLUSION

On the basis of the analysis, it was concluded that the use in the model of information from the reporting of the controlling company increases its discriminatory ability to predict the event of bankruptcy. The proposed approach can be used by commercial banks for use in rating models for companies associated with the group for which the bank does not have reliable information about company defaults. It has also been shown that a built-in model is more complete when the bankruptcy event is established by the parent company. By setting different thresholds, you can adjust the degree of conservation of the model, while adjusting the ratio between the first and second kind of errors.

The model can also be improved by clarifying the list of factors for each company, depending on its role in the group. At the moment, this remains an unresolved problem due to the existence of intra-group transactions between companies. Without additional data on cash flows between companies within the group, it is impossible to calculate the contribution of the individual company to the financial result of the group and its role in it. On the other hand, such data may allow the sample to be segmented or include additional factors in the model, depending on whether the company, for example, is the major holding of assets in the group or the major profit-generating company, but for the time being this remains a subject for future research.

ACKNOWLEDGEMENTS

This work/article is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University). National Research University Higher School of Economics, Moscow, Russia.

REFERENCES

1. Bereznev S.V., Baryshev M.A. The concept of "holding" in Russia: Determination of nature and content. *Mir ekonomiki i upravleniya = World of Economics and Management*. 2012;12(1):109–114. URL: https://cyberleninka.ru/article/n/ponyatie-holding-v-rossii-opredelenie-suschnosti-i-soderzhaniya (accessed on 11.12.2022). (In Russ.).
2. Korneeva T.A., Shatunova G.A. Corporate unions and the consolidated financial reporting. *Audit i finansovyi analiz = Audit and Financial Analysis*. 2006;(3):5–9. URL: https://auditfin.com/fin/2006/3/Korneeva/Korneeva%20.pdf (accessed on 15.12.2022). (In Russ.).
3. Belykh V.S., Bersunkaev G.E. Holdings and holding legislation in Russia. *Biznes, menedzhment i pravo = Business, Management and Law*. 2007;(1):86–88. (In Russ.).
4. Claessens S., Fan J.P.H., Lang L.H.P. The benefits and costs of group affiliation: Evidence from East Asia. *Emerging Markets Review*. 2006;7(1):1–26. DOI: 10.1016/j.ememar.2005.08.001
5. Chang S.J., Hong J. Economic performance of group-affiliated companies in Korea: Intragroup resource sharing and internal business transactions. *Academy of Management Journal*. 2000;43(3):429–448. DOI: 10.2307/1556403
6. Beaver W.H., Cascino S., Correia M., McNichols M.F. Group affiliation and default prediction. *Management Science*. 2019;65(8):3559–3584. DOI: 10.1287/mnsc.2018.3128
7. Mkrtchyan G.M. Some issues of bringing the controlling persons of the debtor to subsidiary liability. *Severo-Kavkazskii yuridicheskii vestnik = North Caucasus Legal Vestnik*. 2019;(3):122–129. (In Russ.). DOI: 10.22394/2074–7306–2019–1–3–122–129
8. Khanna T., Yafeh T. Business groups and risk sharing around the world. *Journal of Business*. 2005;78(1):301–340. DOI: 10.1086/426527
9. Ghojogh B., Crowley M. The theory behind overfitting, cross validation, regularization, bagging, and boosting: Tutorial. arXiv:1905.12787. 2019. DOI: 10.48550/arXiv.1905.12787
10. Morgunov A.V. Modeling the probability of default of the investment projects. *Korporativnye finansy = Journal of Corporate Finance Research*. 2016;10(1):23–45. (In Russ.). DOI: 10.17323/j.jcfr.2073–0438.10.1.2016.23–45
11. Siddiqi N. Credit risk scorecards: Developing and implementing intelligent credit scoring. Hoboken, NJ: John Wiley & Sons, Inc.; 2006. 208 p.

12. Polterovich V.M., ed. Strategy for modernizing the Russian economy. St. Petersburg: Aletheia; 2010. 424 p. (In Russ.).
13. Magnus Ya.R., Katyshev P.K., Peresetskii A.A. Econometrics. Initial course. 6th ed. Moscow: Delo; 2004. 576 p. (In Russ.).
14. Karminsky A.M., Fal'ko S.G., eds. Controlling in a bank. Moscow: Infra-M; 2013. 288 p. (In Russ).
15. Karminsky A.M., Kostrov A.V. Modeling the default probabilities of Russian banks: Extended opportunities. *Zhurnal Novoi ekonomicheskoi assotsiatsii = Journal of the New Economic Association*. 2013;(1):64–86. (In Russ.).

ABOUT THE AUTHORS

Valentin V. Lopatenko — PhD student at the School of Economics, National Research University Higher School of Economics, Moscow, Russia; Leading Data Science Expert, Sberbank PJSC, Moscow, Russia
https://orcid.org/0000-0003-0256-0792
Corresponding author:
vlnlopatenko@gmail.com

Alexander M. Karminsky — Dr. Sci. (Econ.), Research Prof., School of Finance, National Research University Higher School of Economics, Moscow, Russia
https://orcid.org/0000-0001-8943-4611
karminsky@mail.ru

Conflicts of Interest Statement: The authors have no conflicts of interest to declare.

The article was submitted on 10.04.2023; revised on 14.05.2023 and accepted for publication on 25.05.2023.
The authors read and approved the final version of the manuscript.

Translated by V. Timonina