• FINANSE I PRAWO FINANSOWE •
• Journal of Finance and Financial Law •

Grudzień/December 2024 • vol. 4(44): 79–93

CONSTRUCTION OF DISCRIMINATION MODELS IN PREDICTION OF BANKRUPTCY OF POLISH NON-PUBLIC ENTERPRISES

Bernard Kokczyński*

https://doi.org/10.18778/2391-6478.4.44.05

CONSTRUCTION OF DISCRIMINATION MODELS IN PREDICTION OF BANKRUPTCY OF POLISH NON-PUBLIC ENTERPRISES

ABSTRACT

The purpose of the article. The aim of this study is to predict bankruptcy among Polish non-financial firms by constructing discriminant models and comparing the outcomes with prognostic models developed by other Polish scholars. Utilizing financial data from 2017–2021 for 416 companies across the trade, production, and service sectors, this research strives to devise the most effective model for classifying entities into two groups.

Methodology. The study employed a discriminant function, a statistical method enabling the classification of objects based on several explanatory variables simultaneously. Two methods for selecting independent variables for the discriminant function were compared using group mean equality tests and Hellwig's method. Additionally, two techniques of winsorization were applied to minimize the impact of outliers on the study results.

Results of the research. The study's findings underscore the importance of operational profitability relative to total assets and the logarithm of total assets as key variables in bankruptcy prediction models. Results confirm the significance of industry specificity on the models' classification accuracy. The use of different methods for selecting independent variables for models and winsorization directly impacts classification efficacy. A comparative analysis with models from selected Polish researchers reveals that the models developed in this study achieved a higher level of effectiveness than existing models in terms of classification accuracy.

* Master's degree, University of Lodz, e-mail: bernard.kokczynski@edu.uni.lodz.pl, https://orcid.org/0000-0002-9379-0376

Keywords: bankruptcy prediction models, winsorization of data, non-financial information, methods of selecting variables for models.

JEL Class: C51, C52, C53, G17, G33.

INTRODUCTION

The article presents the results of predicting bankruptcy among Polish non-public enterprises and compares these outcomes with prognostic models developed by other Polish authors.

The data used in the analysis comes from the EMIS (Emerging Markets Information Service) financial reports from 2017 to 2021, covering 416 enterprises across three sectors: trade, production, and services. Within the scope of the research, 16 discriminant models were constructed. In these models, financial indicators for constructing the discriminant function were selected based on the group mean equality test and Hellwig's variable selection method (Witkowska, 2023: 275–277).

The study differentiated enterprises by industry using a binary-coded non-financial variable. The efficacy of models constructed on raw data was compared with models based on data processed through winsorization, using Tukey's biweight criterion (Pociecha et al., 2014: 67) and the three-sigma rule (Witkowska, 2023: 49) to eliminate the influence of outliers.

The primary goal of the research was to obtain a model with the highest possible classification efficacy for the test sample and to compare the efficiency of the constructed models with those of selected Polish authors who achieved an average classification effectiveness of models in the test sample above 75%. Additionally, it was verified which method of variable selection would prove to be more effective and how the data processing process would impact classification results.

1. DISCRIMINANT BANKRUPTCY PREDICTION MODELS IN THE LITERATURE

The literature on bankruptcy prediction models in Poland is very rich. Therefore, it is necessary to mention the most important works of domestic authors (Ptak-Chmielewska, 2021: 41).

Since the introduction of the Z-score model by Altman (1968: 589–609), which first utilized discriminant analysis for bankruptcy prediction purposes, most scientific research has focused on forecasting financial difficulties of companies. In many cases, researchers consider bankruptcy as a critical threshold intended to distinguish between financially distressed firms and those in good condition. The fundamental issue is defining bankruptcy itself, as in models, bankruptcy is often understood as the inability to continue operational activities – however, it is not related to the legal definition of insolvency. This raises the question of how poorly a company must be managed to be considered bankrupt (Bombiak, 2010: 148; Kokczyński, 2022: 158). In the literature, bankruptcy is defined as a state in which a company is unable to continue operations without external financial support (Pasternak-Malicka et al., 2021: 251) or a state in which the value of a company's liabilities exceeds the value of its assets (Boratyńska, 2014: 21).

Some researchers conceptualize bankruptcy as a stage following the declaration of insolvency, when the debtor's assets are insufficient not only to satisfy creditors' claims but also to cover the costs of the bankruptcy procedures themselves. They define bankruptcy as the ultimate state, when the debtor's financial capabilities are so limited that they prevent any restructuring or negotiation activities with creditors (Kopczyński, 2022: 13). In this work, companies that have filed for bankruptcy in court are considered bankrupt. However, it should be added that Polish authors approach this issue differently. In the work of Gajdka and Stos, bankrupt companies are considered those that have started the liquidation process due to financial situation, entered into a court agreement with creditors, or reached a settlement with the bank under the act on financial restructuring of enterprises and banks (Gajdka & Stos, 1996: 143). In the work of Mączyńska and Zawadzki, enterprises at risk of bankruptcy are considered to be economic entities where symptoms such as negative equity, losses, and loss of liquidity were observed (Mączyńska & Zawadzki, 2006: 12). Meanwhile, in the collective work of Pociecha et al., bankrupt companies are considered those that have declared bankruptcy (Pociecha et al., 2014: 59), and Kopczyński also adopted this way of defining bankrupts (Kopczyński, 2022: 15). The adopted definition of a bankrupt determines which entities will be included in the research group as enterprises at risk of bankruptcy. Different approaches to the definition can lead to the selection of various data sets, which directly affects the representativeness of the research sample and the generalization of research findings. For example, if only companies that have formally filed for bankruptcy are considered bankrupt, the model may be less effective in identifying companies at earlier stages of bankruptcy risk. The literature emphasizes the low classification effectiveness of predictive models based on data describing the financial situation of bankrupts in the years preceding the declaration of bankruptcy by more than two (Pociecha et al., 2014: 61).

The bankruptcy of a company signifies a disruption in its operational continuity and has a significant impact on all stakeholders, including creditors, owners, and suppliers. In extreme cases, the accumulation of so-called bad debts can not only lead to the bankruptcy of individual enterprises but also trigger a cascading effect of bankruptcies among financially interlinked companies, generating another wave of uncollectible obligations and causing a so-called domino effect (Janus et al., 2022: 72). On the other hand, bankruptcies serve as a form of catharsis for the economy, necessitating the cleansing of the market by eliminating insolvent units that cannot meet the rules and requirements of market efficiency (Mączyńska, 2013: 4). Consequently, there is a need to develop reliable models for predicting financial distress that can timely diagnose entities with financial difficulties. Such models are a crucial informational tool for investors, shareholders, company management, and financial institutions like banks (Shi & Li, 2019: 116).

The diverse conditions under which enterprises operate in various regions are the reason why the issue of bankruptcy prediction cannot be generalized and requires an individual research approach based on the use of empirical data related to a specific economy or group of economies with similar operating conditions (Jaki & Ćwięk, 2021: 3).

Among the most frequently cited works by Polish authors in the field of bankruptcy prediction there are: Gajdka and Stos (1996), Hadasik (1998), Wierzba (2000), Appenzeller and Szarzec (2004), Mączyńska & Zawadzki (2006), Hołda (2006), Wojna (2007), Hamrol and Chodakowski (2008), Pociecha (2011), Pociecha et al. (2014), Kopczyński (2016). It is worth noting that Polish authors also explore alternatives to standard financial indicators, including works such as: Korol (2010), Ptak-Chmielewska and Matuszyk (2017), Ptak-Chmielewska (2021).

2. DATA AND METHODS

2.1. Data

The database contains observations on 416 non-financial enterprises not listed on the capital market. Half of the observations consist of enterprises that have filed for bankruptcy in court (208 observations), while the other half are enterprises able to continue their business operations. For simplification in the study, these enterprises will be referred to as non-bankrupts. The data comes from the EMIS – Emerging Markets Information Service.

The entities included in the study can be divided into three sectors, which are presented in Table 1. The database gathers both financial and non-financial information about the entities under examination. The data is complete, as missing information was supplemented based on reports from the Ministry of Justice's website, from which financial reports of the companies under study for the year and two years before bankruptcy (financial reports from 2017–2021) were downloaded. Reports from the year in which an entity filed for bankruptcy were not included in the database.

Table 1. Number of enterprises included in the research set by industry

Category	Trade	Production	Services
Bankrupts	72	68	68
Non-bankrupts	72	68	68

Source: own study based on enterprise data obtained from the EMIS website.

In addition to financial data from reports, the EMIS database also contains data on calculated financial indicators. However, it was decided to forego using this information due to its largely incomplete nature, as well as significant discrepancies in the calculation methods for the indicators among different entities. Therefore, to minimize additional factors that could disrupt the models, a decision was made to independently calculate all considered financial indicators, selected from six groups: liquidity, indebtedness, profitability, operational efficiency, dynamics, and size and structure, based on the literature (Pociecha et al., 2014: 64–67; Mączyńska & Zawadzki, 2006: 23–24; Hamrol & Chodakowski, 2008: 21–24), totaling 57 financial indicators. Table 2 presents these indicators used in the study. Additionally, a non-financial variable coded as binary, B01 with a value of 1 for the industry and B02 with a value of 1 for services, was implemented in the study.

Table 2. Financial indicators

Indica-tors	Formula
Liquidity	
P01	Current assets / Short-term liabilities
P02	(Current assets - Inventories) / Short-term liabilities
P04	(Current assets - Short-term liabilities) / Total assets
Indebtedness	
Z01	(Long-term liabilities + Short-term liabilities) / Total assets
Z02	(Long-term liabilities + Short-term liabilities) / Equity
Z04	Equity / Total assets
Z07	(Net income + Depreciation) / (Long-term liabilities + Short-term liabilities)
Z10	(Equity + Long-term liabilities) / Fixed assets
Z12	Current assets / (Long-term liabilities + Short-term liabilities)
Profitability	
R01	Operating income + Depreciation

R02	(Operating income + Depreciation) / Total assets
R03	(100 · Gross profit) / Net sales revenue
R05	(100 · Net income) / Equity
R07	Operating income / Total assets
R09	Gross profit of t–1 and t–2 / Total assets
R10	Net profit /((Current assets t–1 + Current assets t–2) / 2) · 100
R12	Net profit / ((Current assets of t–1 + Current assets of t–2) / 2) · 100
R13	(Operating income – Depreciation) / Total assets
R14	(Net income / Current assets) ·100
R16	Net sales revenue / Total assets
Operational Efficiency	
S03	Inventories / Operating costs
S04	Inventories / Net sales revenue
S05	Short-term receivables / Net sales revenue
S06	Operating costs / Short-term liabilities
S07	Net sales revenue / Short-term receivables
S09	(Inventories ·360) / Operating revenue
S14	(((Inventories of t–1 + Inventories of t–2) / 2) · 360) / Operating revenue
S15	(((Receivables of t–1 + Receivables of t–2) / 2) · 360) / Operating revenue
S16	(((Short-term liabilities of t–1 + Short-term liabilities of t–2) / 2) · 360) / Operating revenue
S17	S14 + S15
S18	S17 – S16
S19	(((Short-term liabilities of t–1 + Short-term liabilities of t–2) / 2) / Operating costs · 360)
Dynamics	
D01	Revenue of t–1 / Revenue of t–2
D02	Equity of t–1 / Equity of t–2
Size and Structure	
W01	Fixed assets / Current assets
W02	Log (Fixed assets / Current assets)
W03	Log (Fixed assets + Current assets)

Source: own study based on: Pociecha et al. (2014: 64–67); Mączyńska and Zawadzki (2006: 23–24); Hamrol and Chodakowski (2008: 21–24).

The collected data was subjected to winsorization, a statistical estimation process that involves modifying outlier values to reduce their impact on the analysis results. In this approach, variable values exceeding predefined threshold limits are replaced with those thresholds, thereby making the estimator resistant to the effects of large residuals. This process divides units into a group of data used unchanged and a group of outlier observations, which are modified and included in the sample in an altered form, enabling the estimation of parameters based on such a transformed dataset (Dehnel, 2017: 61–62).

The winsorization process was conducted using various methodologies to assess their impact on the models' classification efficiency. In the first approach, the three-sigma rule was applied to modify the data (Witkowska, 2023: 49), while the second method was based on the application of the Tukey's biweight criterion (Pociecha et al., 2014: 67–68). Both methods of winsorization were applied separately for each set of companies, and it should be emphasized that the discriminant models were also constructed based on raw data.

2.2. Methods

For the construction of bankruptcy prediction models, a linear discriminant function was used, which is a statistical method that allows for the classification of objects based on multiple explanatory variables simultaneously according to a specific criterion (Tłuczak, 2013: 424). The form of the linear discriminant function is as follows:

$$\mathrm{Z} = \mathrm{a}_0 + \mathrm{a}_1\mathrm{X}_1 + \mathrm{a}_2\mathrm{X}_2 + \cdots + \mathrm{a}_\mathrm{k}\mathrm{X}_\mathrm{k}$$

where:
Z – represents the dependent variable,
a_i – are discriminatory coefficients są współczynnikami dyskryminacyjnymi,
a_0 – is the constant,
X_i – denotes the explanatory variables.

The selection of variables for discriminant models was based on the test of group mean equality using the SPSS software. This test, conducted through the ANOVA analysis of variance, assesses whether the average values of the variables under study statistically differ between the specified groups (IBM, 2023), in this case, between bankrupt companies and those that remained on the market. The selection criteria for variables into the model were the F statistic values and the significance level p-value, with statistically significant differences considered at a p-value less than 0.05. Variables that met this criterion and showed high F values, indicating strong differences between groups, were selected for further analysis. Although in the study, the lowest possible Wilks' lambda value is desirable (Wojnar & Kasprzyk, 2011: 412–413), this value was not a selection criterion for variables, as the values for most included variables were similar.

After selecting the variables, the construction of linear discriminant functions for various combinations of explanatory variables was initiated. A stepwise forward method was adopted, which allowed for the identification of factors with the most significant impact on classification. As a result of numerous experiments, models that demonstrated the highest classification efficiency were presented. These models were constructed based on variables: R02, W03, and R09, as well as on W03 and R10 variables.

The second method applied for the selection of diagnostic variables was the Hellwig method, aimed at selecting a set of variables that best characterize the phenomenon of bankruptcy while avoiding informational redundancy (Witkowska, 2023: 275–276). As a result of implementing the Hellwig method, variables such as S18, R09, P02, R13, R14, R16, S04, S19, and R10 were selected for the central variable group, and Z02, Z03, R01, R02, R05, S03, S07, W01, W02, D01, Z04, W03, D02, S09 were classified into the isolated variable group. Based on the selected variables, two discriminant functions were constructed, separately for the set of central and isolated variables. All constructed models were presented in Tables 3 and 4.

Table 3. Statistical significance of estimated discriminant functions with variables selected using the progressive stepwise method

Discriminant function equation	**Eigen-value**	**Cano-nical corre-lation**	**λ**	**χ**	**df**	**p-value**
Raw data						
F01 = 0.217 · R02 + 1.026 · W03 + 0.032 · R09 – 3.322	0.150	0.361	0.870	46.468	3	0.000
F02 = 1.043 · W03 + 0.003 · R10 – 3.357	0.159	0.360	0.863	49.115	2	0.000
Winsorized data – 3 sigma rule						
F03 = 0.226 · R02 + 1.051 · W03 + 0.023 · R09 – 3.399	0.154	0.366	0.866	47.754	3	0.000
F04 = 1.063 · W03 + 0.003 · R10 – 3.415	0.163	0.374	0.860	50.279	2	0.000
Winsorized data – Tukey's biweight criterion						
F05 = 0.245 · R02 + 1.238 · W03 – 0.030 · R09 – 3.952	0.219	0.424	0.821	65.752	3	0.000
F06 = 1.210 · W03 + 0.002 · R10 – 3.842	0.226	0.429	0.816	67.736	2	0.000
Raw data						
F07 = 0.217 · R02 + 1.028·W03 + 0.034 · R09 – 0.093 · B01 + 0.088 · B02 – 3.326	0.151	0.362	0.869	46.561	5	0.000

F08 = 1.045 · W03 + 0.003 · R10 – 0.090 · B01 + 0.133 · B02 – 3.378	0.160	0.372	0.862	49.344	4	0.001
Winsorized data – 3 sigma rule						
F09 = 0.226 · R02 + 1.054 · W03 + 0.024 · R09 – 0.102 · B01 + 0.083 · B02 – 3.400	0.155	0.367	0.866	47.860	5	0.000
F10 = 1.066 · W03 + 0.003 · R10 – 0.098 · B01 + 0.131 · B02 – 3.435	0.164	0.376	0.859	50.529	4	0.000
Winsorized data – Tukey's biweight criterion						
F11 = 0.249 · R02 + 1.249 · W03 – 0.034 · R09 – 0.214 · B01 + 0.066 · B02 – 3.941	0.222	0.426	0.818	66.397	5	0.000
F12 = 1.219 · W03 + 0.002 · R10 – 0.191 · B01 + 0.115 · B02 – 3.845	0.229	0.432	0.814	68.495	4	0.000

Source: own study based on enterprise data obtained from the EMIS website and analysis in SPSS.

Table 4. Statistical significance of estimated discriminant functions with variables selected using the Hellwig method

Discriminant function equation	**Eigen-value**	**Cano-nical correla-tion**	**λ**	**χ**	**df**	**p-value**
F13 = 0 · S18 + 0.271 · R09 – 0.001 · P02 + 0.028 · R13 + 0·R14 – 0.018 · R16 – 0.181 · S04 + 0·S19 + 0.002 · R10 + 0.243	0.050	0.218	0.952	16.101	9	0.065
F14 = –0.002 · Z02 + 0.078 · Z03 + 0 · R01 + 0.129 · R02 + 0·R05 – 0.061 · S03 + 0 · S07 – 0.051 · W01 + 0.358 · W02 + 0 · D01 + 0.034 · Z04 + 0.779 · W03 – 0.001 · D02 + 0 · S09 – 2.133	0.253	0.449	0.798	73.755	14	0.000
F15 = 0 · S18 + 0.276 · R09 – 0.001 · P02 + 0.028 · R13 + 0 · R14 – 0.018 · R16 – 0.187 · S04 + 0 · S19 + 0.002 · R10 – 0.129 · B01 + 0.160 · B02 + 0.237	0.051	0.220	0.952	16.264	11	0.132
F16 = –0.001 · Z02 + 0.077 · Z03 + 0 · R01 + 0.130 · R02 + 0 · R05 – 0.065 · S03 + 0 · S07 – 0.050 · W01 + 0.357 · W02 + 0 · D01 + 0.035 · Z04 + 0.783 · W03 – 0.001 · D02 + 0 · S09 – 0.182 · B01 – 0.003 · B02 – 2.085	0.255	0.451	0.797	73.997	16	0.000

Source: own study based on enterprise data obtained from the EMIS website and analysis in SPSS.

The verification of classification accuracy should be performed based on a test set of observations, containing objects that did not participate in the process of estimating parameters of the discriminant function. The process of selecting companies for the sample should be random to avoid errors stemming from a subjective selection (Witkowska, 2023: 318–319).

In the study, observations for the test sample were selected through randomization, using a random number generator in Excel. The effectiveness of the models was verified based on an 80% training sample and a 20% test sample ratio. To maintain the representativeness of industry participation in the study for the entire population, the drawing was conducted separately within each industry: trade, production, and services. This approach allowed for maintaining symmetry between the number of available observations and industry participation. The results of the drawing from each industry were then summed up, creating a balanced test sample that reflects the industry structure of the available database. This should translate into greater reliability and accuracy in verifying the effectiveness of the models.

Such a method of sample selection, unfortunately, still carries the choice-based sample bias. This means a situation in which units are selected for the sample based on prior information regarding the dependent variable. For instance, initially, data concerning a group of bankrupt companies are collected. The probability that units will be included in such a sample depends precisely on the characteristics of the dependent variable. The sample is constructed, for example, by including all the insolvent units, while the rest are selected using a specific matching scheme (Gruszczyński, 2017: 24).

3. RESULTS AND DISCUSSION

Table 5 presents the classification effectiveness of the 16 constructed discriminant models.

The least accurately classifying models, F13 and F15, were constructed using the Hellwig method based on central variables. These models are characterized by an average classification effectiveness of 67.50%. It is important to emphasize their low effectiveness in classifying bankruptcies at 42.50% while simultaneously achieving a very high effectiveness in classifying non-bankruptcies at 92.50%. The low effectiveness in recognizing bankruptcies may indicate a suboptimal selection of variables.

F02, F04, F06, F11, F12, and F16 models stand out in terms of classification effectiveness. F02, F04, and F06 models are based on financial variables and were constructed using a progressive stepwise method based on the test of group means equality. Models F11 and F12 were also constructed using this method but incorporate non-financial information about industries. The independent variables in F16 model were selected using the Hellwig method, consisting of variables from the set of isolated features. The analysis of these models shows the impact of individual variables on classification effectiveness. Variables such as W03, the logarithm of total assets, and R02, an indicator of operational profitability relative to total assets, appear in the most effective models, emphasizing their significance in assessing the financial condition of companies. Additionally, the use of a non-financial variable that differentiates observations by industry in some models, such as F11, F12, and F16, shows that complementing financial analysis with non-financial information can contribute to increased prediction accuracy.

Table 5. Classification efficiency in the test set

Effectiveness of classification							
Model	**Bank-rupts**	**Non-bank-rupts**	**Total**	**Model**	**Bank-rupts**	**Non-bank-rupts**	**Total**
F01	70,00%	75,00%	72,50%	F09	67,50%	75,00%	71,25%
F02	70,00%	80,00%	75,00%	F10	65,00%	77,50%	71,25%
F03	70,00%	75,00%	72,50%	F11	72,50%	77,50%	75,00%
F04	70,00%	80,00%	75,00%	F12	65,00%	87,50%	76,25%
F05	70,00%	75,00%	72,50%	F13	42,50%	92,50%	67,50%
F06	70,00%	82,50%	76,25%	F14	72,50%	67,50%	70,00%
F07	67,50%	75,00%	71,25%	F15	42,50%	92,50%	67,50%
F08	65,00%	77,50%	71,25%	F16	75,00%	72,50%	73,75%

Source: own study based on enterprise data obtained from the EMIS website and analysis in SPSS.

The study confirms that both data processing and the methodology of variable selection are crucial for the effectiveness of discriminant models. Models based on raw data, while effective, seem to be slightly less precise compared to those utilizing processed data. Among the two applied winsorization methods, better results were obtained using the method based on the Tukey's biweight criterion. The introduction of a non-financial variable into the models affected their classification effectiveness, but it did not always translate into a clear improvement compared to models based solely on financial variables.

As part of the conducted research, a comparison of the obtained classification results was also made with discriminant models developed by selected Polish researchers. These models, presented in Table 6, were estimated using the data employed in this study. It should be emphasized, however, that the estimation of models by the Polish authors was based on the use of variables indicated by these researchers, and not on an exact replication of the discriminant functions they proposed. The presented models were based solely on variables available in the collected database, which allowed for their direct comparison with the discriminant models developed as part of the research.

Table 6. Classification efficiency in the test set

Models		**Effectiveness of classification**		
Authors	**Variables**	**Bankrupts**	**Non-bankrupts**	**Total**
Wierzba's	Z04, R12, R13, Z12	20,00%	92,50%	56,25%
Hadasik's	P01, P02, P04, Z01, S04, S05	30,00%	92,50%	61,25%
Pogodzińska and Sojak's	P02, R03	12,50%	95,00%	53,75%
Pociecha's (D_1)	S06, R02, Z10	52,50%	95,00%	73,75%
Pociecha's (D_2)	Z07, Z10, S06	10,00%	97,50%	53,75%
Mączyńska and Zawadzki's (G)	P01, R07, Z04, Z07	25,00%	92,50%	58,75%

Source: own study based on enterprise data obtained from the EMIS website and analysis in SPSS and subject literature: Pociecha et al. (2014: 109); Mączyńska and Zawadzki (2006: 21); Hamrol and Chodakowski (2008: 21–23).

This analysis aimed to verify the effectiveness of our solutions in the context of existing concepts. Among the models of Polish researchers, Pociecha's model in version D1 showed the highest classification effectiveness, achieving an average effectiveness of 73.75%. A common feature of this model and the developed F06 and F16 models is the R02 variable, which refers to operational profitability relative to total assets. This variable, as it turns out, has a significant impact on classification effectiveness, as confirmed in both Pociecha's model and F06 and F16 models.

CONCLUSIONS

The study focused on analyzing the bankruptcy prediction of Polish non-publicly traded non-financial companies. The method of variable selection and data processing techniques proved significant, where the application of the Tukey's biweight criterion for data winsorization contributed to the improvement of the models' classification effectiveness. The use of the non-financial B01 and B02 variable to distinguish companies based on their industry had an impact on the accuracy of model classification, though it did not always lead to a clear improvement in classification accuracy. Comparing the accuracy of classification results of constructed discriminant models with the works of Polish authors revealed significant classification effectiveness, especially for F06 and F16 models, which showed a better ability for precise identification of bankruptcies. This seems to be caused by continuous changes occurring in socio-economic phenomena that may influence the rapid obsolescence of models (Witkowska, 2023: 233). Among the financial variables, the key indicator proved to be R02 variable, relating to operational profitability relative to total assets, indicating that operational profitability can be a strong predictor of bankruptcy risk. Additionally, W03 variable, the logarithm of total assets, suggests that larger companies have a lower probability of bankruptcy, which may reflect the scale effect and greater financial stability of large entities.

The obtained results, however, cautiously challenge the rationale behind constructing models, emphasizing that key variables such as R02 and W03 could themselves serve as criteria for assessing the financial condition of companies. This conclusion may seem contradictory to the literature on the subject (Mączyńska & Zawadzki, 2006: 22–23; Gajdka & Stos, 1996: 147), yet, it is worth noting that there are some authors who have arrived at similar conclusions in their works (Kopczyński, 2016: 385–391; Pociecha et al., 2014: 117–118). Nonetheless, the necessity to include statistically significant financial variables and non-financial information, as well as the careful selection of their processing methods, remains crucial for developing effective predictive models capable of identifying companies deemed bankrupt.

REFERENCES

Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23(4), 589–609. https://doi.org/10.2307/2978933

Appenzeller, D. & Szarzec, K. (2004). Prognozowanie zagrożenia upadłością polskich spółek publicznych. *Rynek Terminowy*, *1*, 120–128.

Bombiak, E. (2010). Modele dyskryminacyjne jako metoda oceny sytuacji finansowej przedsiębiorstwa. *Zeszyty Naukowe Akademii Podlaskiej. Administracja i Zarządzanie*, *13*(86), 141–152.

Boratyńska, K. (2014). Mechanizm bankructwa wybranego przedsiębiorstwa spożywczego. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, *361*, 20–28. https://doi.org/10.15611/pn.2014.361.02

Dehnel, G. (2017). Winsoryzacja w ocenie małych przedsiębiorstw. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, *468*, 60–68. oai:dbc.wroc.pl:37242

Gajdka, J. & Stos, D. (1996). Wykorzystanie analizy dyskryminacyjnej do badania podatności przedsiębiorstwa na bankructwo. W: J. Duraj (red.), *Przedsiębiorstwo na rynku kapitałowym* (ss. 138–148). Wydawnictwo UŁ.

Gruszczyński, M. (2017). Błędy doboru próby w badaniach bankructw przedsiębiorstw. *Kwartalnik Nauk o Przedsiębiorstwie*, *44*(3), 22–29. https://doi.org/10.5604/01.3001.0010.4688

Hadasik, D. (1998). *Upadłość przedsiębiorstw w Polsce i metody jej prognozowania*. Wydawnictwo AE w Poznaniu.

Hamrol, M. & Chodakowski, J. (2008). Prognozowanie zagrożenia finansowego. Wartość predykcyjna polskich modeli analizy dyskryminacyjnej. *Badania Operacyjne i Decyzyjne*, *3*, 17–32.

Hołda, A. (2006). *Zasada kontynuacji działalności i prognozowanie upadłości w polskich realiach gospodarczych*. Wydawnictwo AE w Krakowie.

IBM. (2023). *Testy równości średnich grupowych*. https://www.ibm.com/docs/pl/spss-statistics/29.0.0?topic=predictors-tests-equality-group-means

Jaki, A. & Ćwięk, W. (2021). Bankruptcy Prediction Models Based on Value Measures. *Journal of Risk and Financial Management*, *14*(6), 1–14. https://doi.org/10.3390/jrfm14010006

Janus, A., Jóźwicki, R., Kobus-Ostrawska, D. & Nyk, M. (2022). *Przedsiębiorstwo w warunkach współczesnej gospodarki rynkowej. Finansowanie, pomiar, efektywność*. Wydawnictwo UŁ.

Kokczyński, B. (2022). Ewolucja statystycznych modeli predykcji bankructwa. W: A. Adamczyk (red.), *Finanse – kierunki i wymiary zmian* (ss. 147–162). Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.

Kopczyński, P. (2016). *Prognozowanie upadłości spółek giełdowych*. Wydawnictwo UŁ.

Kopczyński, P. (2022). *Prognozowanie upadłości przedsiębiorstw*. Wydawnictwo UŁ.

Korol, T. (2010). *Early Warning Systems of Enterprises to the Risk of Bankruptcy*. Wolters Kluwer.

Mączyńska, E. (2013). Bankructwa i upadłości przedsiębiorstw — podstawowe problemy i obszary badawcze. *Biuletyn Polskiego Towarzystwa Ekonomicznego*, *1*(60), 3–6.

Mączyńska, E. & Zawadzki, M. (2006). Dyskryminacyjne modele predykcji upadłości przedsiębiorstw. *Ekonomista*, *2*, 205–235.

Pasternak-Malicka, M., Ostrowska-Dankiewicz, A. & Dankiewicz, R. (2021). Bankruptcy – an assessment of the phenomenon in the small and medium-sized enterprise sector – case of Poland. *Polish Journal of Management Studies*, *24*(1), 250–267. https://doi.org/10.17512/pjms.2021.24.1.15

Pociecha, J. (2011). Modele prognozowania bankructwa w systemie wczesnego ostrzegania przedsiębiorstw. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, *165*, 124–134.

Pociecha, J., Pawełek, B., Baryła, M. & Augustyn, S. (2014). *Statystyczne metody prognozowania bankructwa w zmieniającej się koniunkturze gospodarczej*. Fundacja Uniwersytetu Ekonomicznego w Krakowie.

Ptak-Chmielewska, A. (2021). Bankruptcy prediction of small- and medium-sized enterprises in Poland based on the LDA and SVM methods. *Statistics in Transition New Series*, 22(1), 179–195. https://doi.org/10.21307/stattrans-2021-010

Ptak-Chmielewska, A. & Matuszyk, A. (2017). The importance of financial and non-financial ratios in SMEs bankruptcy prediction. *Bank i Kredyt*, *49*(1), 45–62.

Shi, Y., & Li, X. (2019). An overview of bankruptcy prediction models for corporate firms: A systematic literature review. *Intangible Capital*, *15*(2), 114–127. https://doi.org/10.3926/ic.1354

Tłuczak, A. (2013). Zastosowanie dyskryminacyjnych modeli przewidywania bankructwa do oceny ryzyka upadłości przedsiębiorstw. *Zeszyty Naukowe Wyższej Szkoły Bankowej we Wrocławiu*, *2*(34), 423–434.

Wierzba, D. (2000). Wczesne wykrywanie przedsiębiorstw zagrożonych upadłością na podstawie wskaźników finansowych – teoria i badania empiryczne. *Zeszyty Naukowe Wyższej Szkoły Ekonomiczno-Informacyjnej w Warszawie*, *9*, 79–105.

Witkowska, D. (2023). *Wybrane metody ilościowe w finansach*. Wydawnictwo Uniwersytetu Łódzkiego.

Wojna, A. (2007). *Predykcja ekonometryczna oraz modelowanie stochastyczne*. Wydawnictwo Uczelniane Politechniki Koszalińskiej.

Wojnar, J. & Kasprzyk, B. (2011). Analiza dyskryminacji dobrobytu ekonomicznego gospodarstw domowych w województwie podkarpackim. *Metody Ilościowe w Badaniach Ekonomicznych*, *12*(2), 409–417.

KONSTRUKCJA MODELI DYSKRYMINACYJNYCH W PRZEWIDYWANIU BANKRUCTWA POLSKICH PRZEDSIĘBIORSTW NIEPUBLICZNYCH

Cel artykułu. Celem badania jest prognozowanie bankructwa polskich przedsiębiorstw niefinansowych poprzez konstrukcję modeli dyskryminacyjnych oraz porównanie wyników z modelami prognostycznymi opracowanymi przez innych polskich autorów. Wykorzystując dane finansowe z lat 2017–2021 dla 416 firm z sektorów handlowego, produkcyjnego i usługowego, badanie dąży do konstrukcji najbardziej efektywnego modelu w klasyfikacji podmiotów na dwie grupy.

Metoda badawcza. W badaniu wykorzystano funkcję dyskryminacyjną, która jest statystyczną metodą umożliwiającą klasyfikację obiektów na podstawie wielu zmiennych objaśniających jednocześnie. Porównano dwie metody doboru zmiennych niezależnych do funkcji dyskryminacyjnej za pomocą testów równości średnich grupowych oraz metody Hellwiga. Ponadto wykorzystano dwie techniki winsoryzacji w celu zmarginalizowania wpływu obserwacji odstających na wyniki badania.

Wyniki badań. Wyniki badania podkreślają znaczenie rentowności operacyjnej w stosunku do aktywów ogółem oraz logarytmu sumy aktywów jako kluczowych zmiennych w modelach prognozowania upadłości. Wyniki potwierdzają istotność wpływu specyfiki branżowej na skuteczność klasyfikacyjną modeli. Stosowanie różnych metod doboru zmiennych niezależnych do modeli oraz winsoryzacji ma bezpośrednie implikacje dla efektywności klasyfikacyjnej. Analiza porównawcza z modelami wybranych polskich badaczy ujawnia, że modele opracowane w tym badaniu, uzyskały wyższy poziom skuteczności niż istniejące modele pod względem dokładności klasyfikacji.

Słowa kluczowe: modele predykcji bankructwa, winsoryzacja danych, informacje niefinansowe, metody doboru zmiennych do modeli.

JEL Class: C51, C52, C53, G17, G33.

Zakończenie recenzji/ End of review: 04.11.2024
Przyjęto/Accepted: 03.12.2024
Opublikowano/Published: 23.12.2024