Industry specifics of joint-stock companies in Poland and their bankruptcy prediction

Sergiusz Herman[1]

Abstract

The bankruptcy of companies is a characteristic of every developed market economy. The risk of bankruptcy is an object of interest for a wide group of stakeholders, including owners, employees, managers, creditors and suppliers. Negative consequences of bankruptcy led to many attempts to predict it. The one of direction of research on predicting the bankruptcy of companies is building models depending on characteristics of the industry of researched companies. Due to a difficulty of gathering a research sample that is big enough, Polish researchers rarely try to build models depending on industries. There are only two examples of authors who have compared the choice of classification for industry models and a “general” model (which does not include the characteristics of industry). There are two main aims of research. The first of them is to compare predictive ability of industry and general models error of industry and general models. The other aim was to define determinants of joint-stock company bankruptcy in particular industries. Empirical studies were conducted on 180 joint-stock companies in the Polish capital market. They represent three industries of the economy that is construction, manufacturing and trade. The calculations were performed using the bootstrapping method and the multivariate discriminant analysis.

Keywords: *companies bankruptcy, predicting, discriminant analysis, industry specifics*

JEL Classification: C530, G330

1 Introduction

Apart from, among others, using advanced statistical tools and finding new prognostic variables, the research on bankruptcy prediction also focuses on the development of models including industry characteristics of companies. E. I. Altman, a world-renowned authority on bankruptcy prediction, agrees with this approach. He believes that bankruptcy estimation models for companies should be based on financial data of companies pursuing homogeneous business activity (Altman, 1983). Due to difficulties in collecting a research sample that is large enough, Polish researchers rarely try to build industry-based models. There are only two examples of authors who compared classification adequacy of industry models and a non-industry dependent model (the “general” model) (Hołda, 2006; Juszczyk and Balina, 2014).

According to the author, due to the used split sample/holdout method, method of estimating the prediction error of the models, none of these results can unequivocally and

[1] Corresponding author: Poznań University of Economics and Business, Department of Econometrics, al. Niepodległości 10, 61-875 Poznań, Poland, e-mail: sergiusz.herman@ue.poznan.pl.

certainly answer the question whether industry models have better predictive ability compared to "general" models.

The article aims to estimate and compare predictive ability of industry and general models by using the bootstrapping method. The study enumerates financial ratios which may be deemed as determinants of bankruptcy of joint-stock companies for constructed industry models. It was verified whether determinants differ depending on the industry.

2 Research methodology

To conduct the study, the author gathered a research sample consisting of two groups: bankrupt companies in poor financial condition and healthy or non-bankrupt companies in good financial condition. Bankrupt companies were defined as the ones that were declared bankrupt by an applicable court. The sample was selected basing on information included in the Internetowy Monitor Sądowy i Gospodarczy. Financial data of the following companies was collected:

- 30 joint-stock companies from the construction industry (according to the Polish Classification of Activities – PKD 41.10-43.99z),
- 30 joint-stock companies from the manufacturing industry (PKD 10.11-33.20z),
- 30 joint-stock companies from the wholesale and retail trade (PKD 46.11-47.99z).

Each of these companies was paired with a joint-stock company in good financial condition. The companies were paired basing on: industry, the main scope of activity and asset size. Financial data of bankrupt companies was found in financial statements from a year before filing a bankruptcy petition – years 2000-2013. Financial statements of healthy companies concerned the same periods. Data was found in databases of Notoria Serwis and Bisnode Dun & Bradstreet as well as Monitor Polski B.

In empirical studies, 19 financial ratios, concerning profitability, liquidity, equity and asset structure as well as operating performance, were used (table 1). They were chosen based on literature review – these ratios are most frequently used in bankruptcy estimation models. The choice was also based on the availability of data in companies` financial statements.

The linear discriminant analysis was used to develop bankruptcy prediction models. Altman (1968) was the first person to use it for that purpose. Despite a dynamic development of statistical models, it is still vastly popular among business advisors.

Symbol	Description	Symbol	Description
ROA	Return on Assets	WPP	Cash ratio
ROE	Return on Equity	ZO	Total Debts to Assets
ZB	Gross Profitability	ZD	Long-Term Debt to Assets
ZS	Gross Profit Margin	KW	Equity to Assets Ratio
MZ	EBT to Sales Revenue	KWZ	Equity to Debt Ratio
MZ2	Net Profit Margin	RN	Accounts Receivable Turnover
MZO	Operating Margin	RZ	Inventory Turnover
KP	Working Capital to Total Assets ratio	RZob	Payables Turnover
WBP	Current liquidity ratio	Rakt	Total Assets Turnover
WSP	Quick liquidity ratio		

Table 1. Financial indicators underlying the research.

The quality of a developed classifier is defined by its ability to predict adherence of objects to defined populations. Such quality may be measured by a classifier true error rate. The world literature mentions empirical studies which aimed to compare various methods of estimating a prediction error (Wehberg and Schumacher, 2004; Braga-Neto and Dougherty, 2004; Molinaro et al., 2005; Kim, 2009). These studies concern the most popular classifiers in medicine. All of those studies bring about one common conclusion that prediction error estimators developed through the holdout method (split sample) are the most volatile of all analyzed estimators. Authors emphasize that the holdout method can be used only for large sets of data which makes it possible to define large independent training and test groups (Ripley, 1996). It is very difficult to meet this condition in the case of studies concerning the Polish capital market.

According to author`s previous studies, when predicting bankruptcy of Polish join-stock companies, prediction error estimators developed through the bootstrapping method have had the most valuable features (Herman, 2016). These methods are based on generation of *B* samples of $x^{*1}, x^{*2}, x^{*3}, ..., x^{*B}$ bootstrap type in such a way that each of them is developed by *n*-fold simple random sampling with replacement from the available set $\{x_1, x_2, x_3, ..., x_n\}$. Next, these samples are used as training samples. Objects which were not drawn in next iterations constitute the test sample. Based on the samples, a true prediction error is estimated. It is often defined as a prediction error of a model developed basing on the entire *n* set and tested on a large and independent sample. In the study, the +632 estimator, proposed in 1997 by Efron and Tibshirani (1997), was used.

3 Analysis of industry differences of financial ratios

Assumptions for constructing a discriminant function concern normal distribution and equal variance in studied groups. Firstly, the hypothesis on population observation of normal distribution was verified with the use of the Kolmogorov-Smirnov test. It turned out that only three ratios (KP, ZO and KW) in the case of healthy companies and one ratio (WBP) for bankrupts have a normal distribution. Next, with the use of Levene`s test, it was assessed whether the variance of ratios was the same in both tested populations. Results of only three efficiency ratios (RN, RZ and Rakt) do not constitute the basis to reject the null hypothesis stating that populations have equal variances. According to the literature, results achieved with the use of this method are not deteriorated when assumptions concerning the linear discriminant analysis are not met (Hand, 1981; Hadasik, 1998). Therefore, despite not meeting the assumptions, the linear discriminant analysis was used in the study.

The Mann-Whitney nonparametric U test was used to verify whether values of ratios for healthy and bankrupt companies in particular industries come from the same distribution. Only in the case of four ratios in the construction industry (ZD, RN, RZ, Rakt), four ratios in the manufacturing industry (ROE, ZD, RN, RZ) and three ratios in trade (ZD, RN, RZ) with a statistical significance of 5%, the distribution of values does not differ in the studied populations. The value of statistics of manufacturing sector is much higher than other researched.

The last step of the initial analysis of financial data used later in the study was to compare distribution of values of company financial ratios between particular industries. For this purpose, the Mann-Whitney U test was used. Healthy and bankrupt companies were studied separately. In the case of healthy companies (5% of statistical significance), distribution of their values differ between industries for 9 financial ratios (WSP, WPP, ZO, KW, KWZ, RN, RZ, RZob, Rakt). When analyzing bankrupt companies, distribution of values of their financial ratios differ more frequently. Only in the case of three financial ratios (ROE, ZO and ZD, with 5% of statistical significance), distribution of values do not differ between industries. The same conclusions were reached when verifying the hypothesis on equal medians of studied financial ratios. A test based on chi-squared distribution was used for that purpose. Differences may be an indicator of a need to construct other models for predicting bankruptcy of joint-stock companies for particular industries.

4 The study of predictive ability of industry models

For constructing industry classifiers and assessing their predictability, the following assumptions were formulated:

a) models are constructed separately for 60 companies for each analyzed industry,
b) models are constructed basing on linear discriminant function and financial indicators presented in table 1,
c) before the learning process, variables strongly correlated with others (correlation coefficient higher than 0.90) are removed,
d) prediction error is assessed with the use of the bootstrapping method, the number of bootstrap samples B=50,
e) method of choosing model variables – the choice of 5 variables with the highest absolute value of the t-statistic for the test comparing an average value of financial ratios in the studied populations.

The true prediction error of industry-based classifiers was assessed with the use of the above assumptions. The results are presented in table 2.

	Construction	**Manufacturing**	**Trade**
The number of firms classified	1064	1064	1064
The number of firms correctly classified	821	950	749
non-bankrupt	395	503	394
bankrupt	426	447	355
The number of firms incorrectly classified	243	114	315
non-bankrupt	137	29	138
bankrupt	106	85	177
Prediction error	21.86%	10.45%	27.67 %

Table 2. Evaluation of classification effectiveness of industry models.

According to the analysis of data included in the table, a classifier constructed for the manufacturing industry has the lowest prediction error. Compared to the construction industry and trade, it is lower by more than 10 and 17 percentage points respectively. Prediction errors for all three industries were estimated based on the test sample which included 1064 objects. It is a result of the assumption on the bootstrap sampling and makes it possible to come up with an average prediction error for analyzed industries. The average prediction error for industry-based models is about 20.00%.

The study aims to compare the prediction ability of industry models and the general model which was constructed basing on all 180 joint-stock companies. Therefore, its prediction error needs to be estimated. In order to do that, the same b)-e) assumptions were used. The prediction error was estimated with the use of two different methods:

1) I method – all 180 companies were used at the same time,
2) II method – 100 subsamples with 60 objects were drawn from the population of 180 companies.

The results of the first method of assessing the model predictability are presented in table 3.

General model	
The number of firms classified	3278
The number of firms correctly classified	2623
non-bankrupt	1337
bankrupt	1286
The number of firms incorrectly classified	655
non-bankrupt	302
bankrupt	353
Prediction error	19.79%

Table 3. Evaluation of classification effectiveness of general model.

The estimated prediction error is lower than the average prediction error for industry models (which is 20.00%). In addition, a detailed analysis of the results of the classification of individual objects showed that the general model classifies accurately the companies from researched sectors, compared to the industry models.

The figure no. 1 shows the prediction error of general models for 100 subsequent subsamples with 60 companies (II method of assessing the prediction error) compared with the average prediction error of industry-based models.

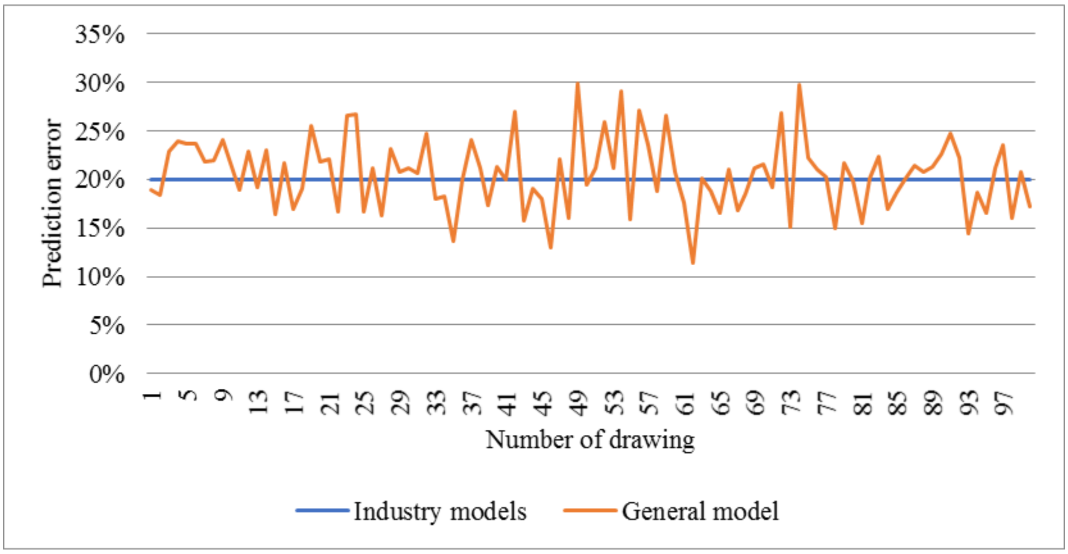

Fig. 1. Prediction error of general model for 100 another sub-samples against average prediction error of industry models.

With the use of the one sample *t*-test, the author verified the hypothesis that an average prediction error for 100 subsamples is the same as the error assessed for industry models (20.00%). The assumptions have been fulfilled in this test. The test statistic value of 1.546 indicated that, with the significance level of 5%, there is no basis to reject the null hypothesis.

The results indicate that industry-based models do not have a lower average prediction error which means that they do not have a higher predictive ability compared to the general model. If the financial ratios of companies have a desirable statistical properties - high discriminative power, as in the case for manufacturing sector, both general and industry models predicts bankruptcy correctly.

5 Determinants of joint-stock company bankruptcy in chosen industries

The previous part included a comparison of predictability of industry models and the general model which does not take into account industry characteristics. According to the approach of Efron (1983), the true prediction error of industry classifiers (models) constructed on the available sample was estimated. The form of classifiers is presented here.

The first model was constructed basing on the sample including construction companies. Two versions of the equation – non-standardized (W) and standardized (SW) – are presented below.

$$D(W) = -0{,}35 \cdot WBP + 2{,}40 \cdot MZO + 3{,}00 \cdot KP + 0{,}40 \cdot KWZ + 0{,}84 \cdot KW - 0{,}42 \qquad (1)$$

$$D(SW) = -0{,}30 \cdot WBP + 0{,}29 \cdot MZO + 0{,}66 \cdot KP + 0{,}46 \cdot KWZ + 0{,}23 \cdot KW \qquad (2)$$

The model was constructed in such a way that positive values of the discriminant function are an indicator of good condition of a company and qualify it as a healthy company. Whereas, negative values suggest that a company has a poor financial situation and is classified as a company at risk of bankruptcy. Additionally, the higher the positive values of financial ratios, the lower risk of bankruptcy. So, the increase of profitability (MZO ratio) and liquidity (KP ratio) decreases the risk of bankruptcy of analyzed companies. The situation is similar for equity and asset structure ratios. The growth of equity share – the growth of KW and KWZ ratios – leads to the lower risk of bankruptcy. In the case of WBP ratio, a negative value may raise some doubts. However, according to the literature, liquidity ratios have a certain value range deemed as correct. Too large deviations, both up and down, have a negative effect on company condition.

The second model of joint-stock company bankruptcy prediction was estimated based on financial data of entities from the manufacturing industry. The relevant equations are presented below.

$$D(W) = 1{,}29 \cdot ZB + 4{,}17 \cdot ZS - 0{,}43 \cdot ZO + 0{,}70 \cdot MZO + 0{,}32 \cdot KWZ + 0{,}34 \tag{3}$$

$$D(SW) = 0{,}22 \cdot ZB + 0{,}49 \cdot ZS - 0{,}19 \cdot ZO + 0{,}15 \cdot MZO + 0{,}46 \cdot KWZ \tag{4}$$

Also in this case, the profitability growth (ZB, ZS and MZO ratios) and a higher equity share in the asset and equity structure (KWZ) decreases the probability of bankruptcy. The situation is different in the case of debt – the higher the debt ratio (ZO ratio), the higher the risk of bankruptcy.

The last model was constructed basing on the sample of companies from the wholesale and retail trade. The model equation can be found below.

$$D(W) = 1{,}38 \cdot ROA + 2{,}41 \cdot MZ2 + 0{,}36 \cdot WBP - 0{,}11 \cdot WSP + 0{,}58 \cdot KWZ - 0{,}61 \tag{5}$$

$$D(SW) = 0{,}26 \cdot ROA + 0{,}38 \cdot MZ2 + 0{,}31 \cdot WBP - 0{,}08 \cdot WSP + 0{,}51 \cdot KWZ \tag{6}$$

Basing on the dependencies, the growth of profitability, shown by ROA and MZ2 ratios, and of equity share in the asset and equity structure decreases the risk of bankruptcy. In the case of liquidity ratios, the growth of WBP ratio is perceived as positive; whereas, it is negative for the WSP ratio.

When comparing models estimated on the basis of company data from particular industries, it can be noted that they were constructed with the use of different set of financial ratios. The KWZ (debt-to-equity) ratio is the only common feature of all three models. The adopted method of choosing variables proved that different financial ratios are useful in predicting company bankruptcy for certain industries.

The form of industry models indicates that certain industries have different determinants of joint-stock company bankruptcy in Poland. It proves that bankruptcy prediction models should be industry-based, despite the fact that it does not improve their average predictability.

Conclusion

Results presented in the article indicate that compared to general models, industry-based models do not have a lower average prediction error, which means higher predictability. Another goal of this study was to define determinants of joint-stock company bankruptcy in particular industries. The form of industry models as well as variables used for the construction of models estimating true prediction errors were taken into account. As a result, determinants of bankruptcy of joint-stock companies in Poland differ in particular industries. Therefore, it is worth to construct industry-based bankruptcy prediction models, despite not always resulting in models with higher predictability.

The results might have been affected by the fact that analyzed joint-stock companies often conducted various business activities (various sections in the classification of business activities in Poland – PKD) even though they belonged to the same industry. In further studies, it is worth to conduct analysis based on companies conducting more uniform business operations but having different legal forms.

The study may be further analyzed by the use of different methodology. In further studies, other methods for predicting company bankruptcy could be used; for instance, soft computing methods that have been popular recently.

References

Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance, 23*(4), 589-609.

Altman, E. I. (1983). *Corporate financial distress: a complete guide to predicting, avoiding, and dealing with bankruptcy*. New York: Wiley.

Braga-Neto, U. M., & Dougherty, E. R. (2004). Is cross-validation valid for small-sample microarray classification? *Bioinformatics, 20*(3), 374-380.

Efron, B. (1983). Estimating the Error Rate of a Prediction Rule: Improvement on Cross-Validation. *Journal of the American Statistical Association, 78*(382), 316-331.

Efron, B., & Tibshirani, R. (1997). Improvements on Cross-Validation: The .632 Bootstrap Method. *Journal of the American Statistical Association, 92*(438), 548-560

Hadasik, D. (1998). *Upadłość przedsiębiorstw w Polsce i metody jej prognozowania*. Poznań: Akademia Ekonomiczna w Poznaniu.

Hand, D. J. (1981). *Discrimination and classification*. Chichester: Wiley.

Herman, S. (2016). Analiza porównawcza wybranych metod szacowania błędu predykcji klasyfikatora. *Przegląd Statystyczny, 62*(4)*, 449-463.*

Hołda, A. (2006). *Zasada kontynuacji działalności i prognozowanie upadłości w polskich realiach gospodarczych*. Kraków: Wydawnictwo Akademii Ekonomicznej.

Juszczyk, S., Balina, R. (2014). Prognozowanie zagrożenia bankructwem przedsiębiorstw w wybranych branżach, *Ekonomista*, *1*, 67-95.

Kim, J. (2009). Estimating classification error rate: Repeated cross-validation, repeated hold-out and bootstrap. *Computational Statistics & Data Analysis, 53*(11), 3735-3745.

Molinaro, A. M., Simon, R., & Pfeiffer, R. M. (2005). Prediction error estimation: a comparison of resampling methods. *Bioinformatics, 21*(15), 3301-3307.

Ripley, B. D. (1996). *Pattern recognition and neural networks*. Cambridge: Cambridge University Press.

Wehberg, S., & Schumacher, M. (2004). A Comparison of Nonparametric Error Rate Estimation Methods in Classification Problems. *Biometrical Journal, 46*(1), 35-47.