Equilibrium. Quarterly Journal of Economics and Economic Policy
Volume 14 Issue 2 June 2019
p-ISSN 1689-765X, e-ISSN 2353-3293
www.economic-policy.pl

ORIGINAL ARTICLE

Citation: Svabova, L., & Durica, M. (2019). Being an outlier: a company non-prosperity sign? *Equilibrium. Quarterly Journal of Economics and Economic Policy, 14*(2), 359–375. doi: 10.24136/eq.2019.017

Contact to corresponding author: lucia.svabova@fpedas.uniza.sk; Department of Economics, The Faculty of Operation and Economics of Transport and Communications, University of Zilina, Univerzitna 1, SK-01026 Zilina, Slovak Republic

Received: 13.12.2018; Revised: 19.04.2019; Accepted: 7.05.2019; Published online: 27.06.2019

Lucia Svabova
University of Zilina, Slovak Republic
orcid.org/0000-0002-4722-4103

Marek Durica
University of Zilina, Slovak Republic
orcid.org/0000-0002-2118-166X

Being an outlier: a company non-prosperity sign?

JEL Classification: *C38; G33*

Keywords: *bankruptcy prediction models; financial ratios; failure prediction; financial distress*

Abstract

Research background: The state of financial distress or imminent bankruptcy are very difficult situations that the management of every company wants to avoid. For these reasons, prediction of company bankruptcy or financial distress has been recently in a focus of economists and scientists in many countries over the world.

Purpose of the article: Various financial indicators, mostly financial ratios, are usually used to predict the financial distress. In order to create a strong prediction model and a statistically significant prediction of bankruptcy, it is advisable to use a deep statistical analysis of the data. In this paper, we analysed the real financial ratios of Slovak companies from the year 2017. In the phase of data preparation for further analysis, we checked the existence of outliers and found that there are some companies that are multivariate outliers because are significantly different from other companies in the database. Thus, we deeply focused on these outlying companies and analysed whether to be an outlier is a sign of financial distress.

Methods: We analysed whether there are much more non-prosperous companies in the set of outlier companies and if their financial indicators are significantly different from those of the prosperous companies. For these analyses, we used testing of the statistical hypotheses, such as the test for equality of means and chi-square test.

Findings & Value added: The ratio of non-prosperous companies between the outliers is significantly higher than 50% and the attributes of non-prosperity and being an outlier are dependent. The means of almost all financial ratios of prosperous and non-prosperous companies among outliers are significantly different.

Introduction

Identification of the impending financial problems of the company can be important not only for company owners or managers, but also for business partners, potential or existing creditors or for employees (Kral *et al*., 2018, pp. 282–294). This is why the issue of prediction of financial distress has been important in recent decades and is of importance also nowadays. For early detection of impending problems in the analysed company, prediction models are usually used (Siekelova *et al*., 2017, pp. 3–10). Their task is to evaluate the financial health of the company based on selected financial indicators. Subsequently, the company should identify imminent financial troubles or even bankruptcy in advance (Kral *et al*., 2016, pp. 224–231). In these prediction models, financial indicators of the companies are usually used. The most used ones are financial ratios (Zvarikova *et al*., 2017, pp. 143–155). In order to create a functioning prediction model with high prediction ability, it is important to undergo the data preparation phase and deeper statistical analysis of these prediction variables. For this reason, we deeply focus on the financial ratios of the Slovak companies to check the existence of potential outlying or extreme values. The outliers can significantly influence the results of statistical analyses and tests. Therefore, as a standard practice, it is appropriate to consider exclusion of such extreme values from further analyses not to distort the values of statistical characteristics, test results and resulting prediction models. However, on the other hand, extreme or outlying values of some financial indicators of the company may be important in identifying its financial difficulties.

The main aim of this study is to focus deeply on the Slovak companies being marked as potential multivariate outliers with respect to the values of all financial ratios of all companies in the dataset. The purpose of this study is to make a deeper analysis in order to discover whether there is any connection between being an outlier and being a non-prosperous company.

Multivariate outliers were detected using Mahalanobis distance according to Tabachnick and Fidell (2012). To identify the dependence between the features “to be an outlier” and “to be a non-prosperous company”, standard Pearson chi-square independence test is used (Benhamou & Melot, 2018). The difference between the proportions of non-prosperous companies in a dataset of potential outliers and in a dataset of non-outliers is tested by the test of equality of the proportions (Eberhardt & Flinger, 1977, pp. 151–155). Finally, mean values of financial ratios of prosperous and non-prosperous companies in the set of outlying companies are compared by t-test (Wilks, 1946, pp. 257–281).

In the area of bankruptcy prediction models, the Authors mostly focus on either checking the suitability of older established models for specified country or companies, or correcting the parameters of these models, or creating new models using various methods. As in this study, we focused on the deeper analysis of data used for bankruptcy prediction modelling, we consider this study to be a pioneer in this area. Its contribution is the innovative approach to the analysis of outliers in relation to the prosperity of the company.

This article is divided into five main parts. The literature review summarizes the current status of publications in the field of bankruptcy prediction models and highlights the place of the study in this field. The next section describes the used data, the set hypotheses, and the methods of their verification. The third part presents the results and their interpretation. The discussion evaluates and summarizes the results. The last part, the conclusion, contains a general summary of the article and its results, research limitations and suggestions for the future direction of research.

Literature review

The creation of bankruptcy prediction models has been the subject of analysis for many authors in different countries over the last years. The first prediction models were created at the end of the 1960s by well-known authors, such as Beaver (1966), Altman (1968), and then in 1980s by Ohlson (1980), Zmijewski (1984), and others. Currently, there are hundreds of prediction models being developed in different countries over the world. Many of them are used in economic practice. The models were created using real data about financial indicators of selected companies by various methods. Some of them are based on historically known discriminant analysis and logistic regression (for example Jing & Fang, 2017, pp. 235–256; Barkar, 2017, pp. 658–672; Szetela *et al*., 2016, pp. 839–856; Lohk & Siimann, 2016, pp. 297–306) or even more modern methods of neural networks (Dima & Vasilache, 2016, pp. 127–143), genetic algorithms, classification trees (Brozyna *et al*., 2016, pp. 93–114), and random forests (Jabeur & Fahmi, 2017, pp. 1173–1186). Several prediction models were created also in Slovakia. In addition to already known models of Chrastinova developed in 1998 and Gurcik developed in 2002, several Slovak authors tried to create a prediction model with the best classification power, such as Gavliak (2006, pp. 65–69); Bielikova *et al*. (2014, pp. 48–56); Harumova and Janisova (2014, pp. 522–539); Mihalovic (2016, pp. 101–118); Kovacova and Kliestik (2017, pp. 775–791); Gavurova *et al*. (2017, pp. 370–

383). Researchers from Slovakia also deal with the application of existing models to predict the financial difficulties of companies in Slovakia (Delina & Packova, 2013, pp. 101–112; Adamko & Svabova, 2016, pp. 64–71; Valaskova *et al*., 2017, pp. 30–38).

Several authors have also dealt with the occurrence of outliers in data used for bankruptcy prediction models in recent years. However, they mostly examined the impact of outliers on the resulting prediction power of the models created. For example, Tsai and Cheng (2012, pp. 333–342) studied bankruptcy prediction performance achieved after removal of different outlier volumes from datasets. Linares-Mustaros *et al*. (2018, pp. 1–10) dealt with problems occurring in financial ratios, such as the occurrence of outliers, in using cluster analysis to classify firms according to their financial structures. Alrawashdeh *et al*. (2018, pp. 284–298) wanted to eliminate the problem of the high sensitivity of the linear discriminant analysis to outliers in data and to improve the classification ability of created models also in bankruptcy prediction. Figini *et al*. (2017, pp. 91–97) in their study described novel approaches to predict default for SMEs by detecting multivariate outliers.

Pawelek *et al*. (2015, pp. 164–173) made an empirical study about the influence of detecting and eliminating outliers on the effectiveness of the bankruptcy prediction logit model for Polish companies. In this study and also in their subsequent studies (Kostrzewska *et al*., 2016, pp. 72–81; Pawelek *et al*., 2017, pp. 29–42) the authors considered both univariate and multivariate methods for detecting outliers in the dataset.

All the authors mentioned, but also other ones, in their studies dealt mostly with the impact of outliers on the resulting bankruptcy prediction model. In our study, we focus on a deeper analysis of the outliers to determine whether being an outlier can be a sign of the corporate non-prosperity. In this respect, therefore, our study is considered innovative in this field.

Research methodology

In our analysis, we focused on the data of Slovak companies. We describe the data file in the next part of the article and in the tables in Annex (Table 2 and Table 3) in more details. Since primary data showed a high number of extreme values, we applied two approaches to mark them for future analysis. Firstly, we focused on the values of all variables (financial ratios) and secondly, we analysed the existence of multidimensional outlying observations. For each individual variable, we marked as potential outliers the values of the variable lying outside the 2.2-multiple of the interquartile range (IQR). Usually, 1.5-multiple of IQR is used, but as shown in Hoaglin & Iglewicz (1987, pp. 1147–1149), this value can sometimes mark as outliers also those values of variables that are not real ones.. Therefore, according to these authors, the use of 2.2-multiple of IQR is preferable. To identify multivariate outliers, the Mahalanobis distance is a suitable metric. The procedure of detecting multidimensional extreme values is processed according to Tabachnick and Fidell (2012). To verify that some measurement is a multivariate outlier, we create a variable P_{MD} defined as:

$$P_{MD} = 1 - CDF_{Chisq}(MD, Df), \tag{1}$$

where CDF_{Chisq} is the cumulative distribution function of the random variable with χ^2 −distribution, Df is the number of financial ratios in the analysis, and MD is the Mahalanobis distance for i-th observation. The P_{MD} variable is used to identify multivariate outliers. If it holds:

$$P_{MD} < 0.001, \tag{2}$$

it indicates that the unit is a multivariate outlier. A value of 0.001 is recommended by Tabachnick and Fidell (2012).

In our analysis we suppose that the fact that a company is a potential outlier may be related to being a non-prosperous entity. We, therefore, focused on the dependence between these two features. Using the statistical procedures, we need to check whether there exists a statistically significant dependence between the fact that the company is a potential outlier or not and the fact that the company is prosperous or not. To identify the dependence, we use the standard Pearson chi-square independence test with the null hypothesis about the independence of the attributes "being an outlier" and "being prosperous". The test variable and critical area of the test is counted according to (Benhamou & Melot, 2018). Rejection of the null hypothesis means that there is a statistically significant association between being a potential outlier and a non-prosperity of a company.

Another point of view we focused on in a dataset of potential outliers, was the fact whether or not there are significantly more non-prosperous companies among outliers than among non-outliers. Thus, we compared the proportion of non-prosperous companies of outliers and non-outliers by the test of equality of proportions in two independent samples. Zero hypothesis of this test is that the proportion of non-prosperous companies is the same, i.e. there are just as much non-prosperous companies among outliers as among non-outliers. The test variable and critical area of this test are counted according to Eberhardt and Flinger (1977, pp. 151–155). Rejecting a zero hypothesis means that, among outliers, there are much more companies that are non-prosperous than prosperous. Therefore, we need to think properly if it is appropriate to exclude these companies from the database in order to avoid loss of information that could be useful in constructing a bankruptcy prediction model.

Finally, we focused on the values of the financial ratios of outlying companies. We compared the mean values for prosperous and non-prosperous companies by using a standard t-test according to Wilks (1946, pp. 257–281). Zero hypothesis is that the mean values of the ratios are the same for prosperous and non-prosperous companies. Rejecting a zero hypothesis indicates that among outliers, the average value of the financial ratio is significantly different for prosperous and non-prosperous companies.

Data

Similarly, like the authors of other studies in Slovakia, we decided to choose the predictors that are the most frequently used in the prediction models worldwide (Valaskova *et al.*, 2017, pp. 30–38). We used financial ratios of real Slovak companies from the Amadeus database from the year 2017. After a thorough check of the data in terms of correctness and completeness, we chose the most commonly used ratios, depicted in Table 1. These variables, which will be the predictors in the prediction model of the financial health of Slovak companies, were subsequently checked for outliers based on the quartile margins and multivariate outliers using (1) and (2). After the checking, the dataset of 62,932 companies was divided into 256 outliers and 62,676 non-outliers.

Then, the prosperity of the company was verified according to the current amendment to Act no. 513/1991 Coll. Commercial Code, where the institute of "the company or the firm in crisis" was established from January 1st, 2016 (Valaskova *et al.*, 2017, pp. 30–38). In Table 2, there are counts and the percentages of companies in the set of potential outliers and in the set of non-outliers divided into prosperous and non-prosperous companies. As we can see from Table 2, there is a much larger proportion of non-prosperous companies among outliers (nearly 40%), while only 15% of non-prosperous companies are non-outliers.

Further, we will verify three hypotheses:

- The first hypothesis says that mean values of financial ratios of outliers and of non-outliers are significantly different for non-prosperous companies. This will be verified by a standard t-test for equality of means of two independent samples (Ahmad *et al*., 2018, p. 3060).
- The second hypothesis says that the proportion of non-prosperous companies among outliers is higher than among non-outliers. This will be verified by the test of equality of proportions in two independent samples.
- The third hypothesis says that the fact the company belongs to the set of outliers is dependent on the fact that the company is non-prosperous. This will be verified by the test of independence in the contingency table.

The weakness of this approach is the fact that in the case of using another way to identify outliers the results might be different. However, the strength of the study is that we propose three different points of view to verify the relationship between the features of the companies "to be an outlier" and "to be a non-prosperous company".

Results

Differences between outliers and non-outliers companies

Table 3 shows the basic statistical characteristics of all financial ratios mentioned in the previous section. The characteristics are presented separately for prosperous and non-prosperous companies. Moreover, each group is divided into companies that were marked as outliers and for those that were not. As was supposed, the means are different for outliers and non-outliers. For prosperous companies, the values of financial ratios of outliers are on average much higher than for non-outliers, which could be the reason for consideration to exclude them from the dataset. For non-prosperous companies, the situation is similar as in the case of prosperous companies. The means of profitability ratios, debt and capital structure ratios and ratios of activity are higher for outlying companies than for non-outlying ones. In the case of liquidity ratios, the situation is opposite. Except for L4, liquidity of outliers is lower than of non-outliers. Comparing the variability of outliers and non-outliers, we can see that the variability of outliers is much higher than of non-outliers. This also could be the reason for consideration of their exclusion from the database of Slovak companies. We have to consider this step in the process of the formation of the bankruptcy prediction model.

As it is visible in Table 3, there are also differences between means of financial ratios of prosperous and non-prosperous companies among outliers. Therefore, we performed a test of differences of these mean values to analyse, to check whether they are statistically significant. Due to the huge ranges of samples, we can use a standard independent two-sample t-test (Wilks, 1946, pp. 257–281). The test results are summarized in Table 4. The equivalence of variances was confirmed by statistical test only for the ratios R3 and A2. The p-values of the tests show that using the significance level 0.05, the null hypotheses about the equivalence of mean values were rejected only for ratios R1, R2, L1, Z4, A1. All other ratios do not have significantly different mean values for prosperous and for non-prosperous companies in the set of outliers. This could mean that these five ratios could be considered as suitable predictors of a company non-prosperity. However, the current ratio *L*3 is often used for this identification. Thus, we can say that even for this reason it could be considered to exclude these outlying companies from the database, which will be further used to derive the prediction model for identification of the company failure.

Relationship between being an outlier and being a non-prosperous entity

Now we focus on the set of outlying companies and analyse whether there exists some relationship between the fact, that the company is an outlier and the fact that the company is a non-prosperous enterprise. The count and percentage of the prosperous and non-prosperous companies between outlying and non-outlying companies are given in contingency Table 5. It shows that there is a larger proportion of non-prosperous companies among outliers than among non-outliers. This indicates a statistically significant association between “being an outlying company” and “being a non-prosperous company”. This hypothesis is tested by the Pearson chi-square independence test (Bengamou & Melot, 2018). The results of this test and other tests of independence are in Table 6. As the p-value (Asymp. Sig.) of all tests is less than any commonly used significance level, we reject the null hypothesis about independence between the variables. In addition, we can claim that there exists a significant association between the non-prosperity of company and the fact whether or not the company is an outlier. The intensity of this association is measured by Phi and Cramer's V coefficients (Table 7). According to these results, the association between the features "to be an outlier" and "to be a non-prosperous" is weak but nevertheless is statistically significant.

Proportions of non-prosperous companies in the sets of outliers and non-outliers

As already mentioned, almost 40 % of companies among the outliers and 15 % among non-outliers are non-prosperous. Testing the equality of the proportions (Eberhardt & Flinger, 1977, pp. 151–155) we determine whether this difference is statistically significant. The result of the test is in Table 8. According to the value of significance level, the null hypothesis about an equal proportion of non-prosperous companies between outliers and non-outliers is rejected. It can be concluded that the proportion of non-prosperous companies is significantly higher in the set of outlying companies than in the non-outlying one.

Discussion

Companies may be multivariate outliers with respect to other companies because they are in financial distress and therefore the value of their financial ratios differ extremely from those of other companies. It is, therefore, necessary to make deep preparation of the database with respect to these identified facts so that we do not lose the information that might be needed to create the bankruptcy prediction model. Similar studies of other authors focus mainly on these aspects of database preparation for creating bankruptcy prediction models. They focused in their study mainly on the presence of outliers in the database and their impact on the resulting model and its prediction or classification ability.

For example, Pawelek *et al.* (2015) in their study focused on the influence of outliers on the effectiveness of the logit model constructed with the database including or omitting the detected (one-dimensional and also multivariate) outliers. Their model was created for Polish companies covering the year 2009. Consequently, the difference between the distributions of the variables with and without the outliers was tested. Because of a small number of bankrupt companies, the authors tried to identify them as one-dimensional outliers comparing with the healthy companies. Finally, they identified seven financial indicators that contained a higher number of bankrupt companies as outliers. Regarding the created logistic prediction models, the authors compared the prediction power of models created with and without outliers. However, the resulting model did not reach the prediction power of 50%, so they do not consider it suitable for predicting the financial difficulties of Polish companies (Pawelek *et al*., 2015).

Kostrzewska *et al*. (2016) in their study used the results of previous studies that presented an improvement in the prediction ability of the model created after exclusion of outliers from the database. Therefore, they used six methods of one-dimensional and multivariate outliers, which were excluded from the database of companies. Then, the authors compared the prediction power of models created from the original database and also from a database in which outliers were removed by some of the methods used. The results in this study confirmed the improvement of the prediction ability of the generated models after excluding the outliers from the database (Kostrzewska *et al*., 2016).

Thsai and Cheng (2012) focused their paper on the effect of removing different outlier volumes on bankruptcy prediction for different datasets. Detection of outliers was distribution-based, distance-based and density-based. An interesting feature of this study is the use of a database of excluded outlying businesses as a test sample in prediction models created. The study showed that when larger numbers of outliers are removed, the prediction models perform better and prediction accuracy of models excluding outliers is always higher than that of models trained and tested without outlier removal (Thais & Cheng, 2012).

In the study of Linares-Mustaros *et al*. (2018), the authors used cluster analysis to create covariates, which are then used to create a model for predicting bankruptcy. In doing so, they also focused on the detection of outliers in the data used. However, they state that when working with financial ratios, either outliers need to be removed from the sample prior to performing the cluster analysis or the number of clusters has been increased (Linares-Mustaros *et al*., 2018).

It can be seen that other authors do not analyse if the fact that an enterprise is an outlier proves that it is non-prosperous, as was presented in this study. Therefore, we do consider the study to be a pioneer in this area.

Conclusions

In this paper, we focused on the analysis of a database of Slovak companies and their financial ratios. In the database preparation phase, we checked the dataset for the existence of potential outlying values, not only one-dimensional but also as multivariate outliers according to Tabachnick and Fidell (2012). On one hand, it could be appropriate to exclude these companies from the database because the values of outliers could cause changes in the results of statistical tests and procedures in the subsequent creation of the corporate bankruptcy prediction model. Moreover, looking at the mean values of the financial ratios of the outliers, we found that it was problematic to identify prosperous and non-prosperous companies as they did not differ significantly in the mean values of financial ratios. For this reason, it could be indeed appropriate to exclude these companies from the database.

However, on the other hand, we have found that the fact that the company is an outlier is somehow interconnected with the fact that it is in a state of financial distress. As it was shown in this paper, an association between being an outlier and being a non-prosperous company exists. Despite the fact, that this association is not very strong, it is statistically significant. The results also indicate that among outliers there is a significantly higher proportion of non-prosperous companies than the prosperous ones. All the findings obtained in this study should be taken into account when developing a database for further creation of bankruptcy prediction model with a strong prediction ability.

The limitation of this study is that both multivariate and one-dimensional outliers were investigated in the database in one way only. Therefore, in the future research it would be appropriate to apply different methods of identifying outliers in data files and to compare the results obtained. It would also be useful to create a prediction model and compare its prediction ability if the database contains outliers and the changes in prediction ability after the removal outliers.

References

Adamko, P., & Svabova, L. (2016). Prediction of the risk of bankruptcy of Slovak companies. In M. Culik (Ed.). *Proceedings of the 8th international scientific conference managing and modelling of financial risks*. Ostrava.

Ahmad, I., Olah, J., Popp, J., & Mate, D. (2018). Does business group affiliation matter for superior performance? Evidence from Pakistan. *Sustainability, 10*(9). doi: 10.3390/su10093060.

Alrawashdeh, M. J., Radwan, T. R., & Abunawas, K. A. (2018). Performance of linear discriminant analysis using different robust methods. *European Journal of Pure and Applied Mathematics*, *11*(1). doi: 10.29020/nybg.ejpam.v11i1.3176.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*. doi: 10.1111/j.1540-6261.1968.tb00843.x.

Barkar, A. A. (2017). Development of instruments of financial diagnostics of crisis. *St. Peterburg University Journal of Economic Studies, 33*(4). doi: 10.21638/11701/spbu05.2017.408.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, *4*. doi: 10.2307/2490171.

Bengamou, E., & Melot, V. (2018). Seven proofs of the Pearson chi-squared independence test and its graphical interpretation. *SSRN*, doi: 10.2139/ssrn.3239 829.

Bielikova, T., Banyiova, T., & Piterkova, A. (2014). Prediction techniques of agriculture enterprises failure. *Procedia Economics and Finance, 12*. doi: 10.1016/ s2212-5671(14)00319-0.

Brozyna, J., Mentel, G., & Pisula, T. (2016). Statistical methods of the bankruptcy prediction in the logistics sector in Poland and Slovakia. *Transformations in Business & Economics*, *15*(1).

Delina, R., & Packova, M (2013). Prediction bankruptcy models validation in Slovak business environment. *E + M Ekonomie a management*, *16*(3).

Dima, A. M., & Vasilache, S. (2016). Credit risk modeling for companies default prediction using neural networks. *Romanian Journal of Economic Forecasting*, *19*(3).

Eberhardt, R., & Flinger, A. (1977). Comparison of two tests for equality of two proportions. *American Statistician*, *31*(4). doi: 10.2307/2683536.

Figini, S., Bonelli, F., & Giovannini, E. (2017). Solvency prediction for small and medium enterprises in banking. *Decision Support Systems, 102*. doi: 10.1016/j. dss.2017.08.001.

Gavliak, R. (2006). Bankruptcy prediction in Slovak companies using linear probability models. *Forum Statisticum Slovacum*, *2*(5).

Gavurova, B., Janke, F., Packova, M., & Pridavok, M. (2017). Analysis of impact of using the trend variables on bankruptcy prediction models performance. *Ekonomicky casopis*, *65*(4).

Harumova, A, & Janisova, M. (2014). Rating Slovak enterprises by scoring functions. *Ekonomicky casopis*, *62*(5).

Hoaglin, D. C., & Iglewicz, B. (1987). Fine-tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association, 82*. doi: 10.2307/228 9392.

Jabeur, S. B., & Fahmi, Y. (2017). Forecasting financial distress for French firms: a comparative study. *Empirical Economics*, *54*(3). doi: 10.1007/s00181-017-1246-1.

Jing, Z., & Fang, Y. (2017). Predicting US bank failures: a comparison of logit and data mining models. *Journal of Forecasting*, *37*(2). doi: 10.1002/for.2487.

Kostrzewska, J., Kostrzewski, M., Pawelek, B., & Galuszka, K. (2016). The classical and Bayesian logistic regression in the research on the financial standing of enterprises after bankruptcy in Poland. In M. Papiez & S. Smiech (Eds.). *Proceedings of 10th professor Aleksander Zelias international conference on modelling and forecasting of socio-economic phenomena*. Cracow: Foundation of the Cracow University of Economics.

Kovacova, M., & Kliestik, T. (2017). Logit and Probit application for the prediction of bankruptcy in Slovak companies. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *12*(4). doi: 10.24136/eq.v12i4.40.

Kral, P., Fleischer, M., Stachova, M., Nedelova, G., & Sobisek, L. (2016). Corporate finance distress prediction of Slovak companies: Z-score model vs. alternatives. In M. Boda & V. Mendelova (Eds.). *Proceedings of the applications of mathematics and statistics in economics AMSE 2016*. Banska Bystrica: Financ.

Kral, P., Janoskova, K., & Kliestik, T. (2018). Key determinants of the public transport user's satisfaction. *Administratie si Management Public*, *31*. doi: 10.24818/amp/2018.31-03.

Linares-Mustaros, S., Coenders, G., & Vives-Mestres, M. (2017). Financial performance and distress profiles. From classification according to financial ratios to compositional classification. *Advances in Accounting*, *40*. doi: 10.1016/j.adiac.2017.10.003.

Lohk, P., & Siimann, P. (2016). Predicting the risk of encountering financial difficulties by the example of Estonian municipalities. In *Proceedings of the 5th international conference on accounting, auditing, and taxation (ICAAT 2016)*. doi: 10.2991/icaat-16.2016.30.

Mihalovic, M. (2016). Performance comparison of multiple discriminant analysis and logit models in bankruptcy prediction. *Economics & Sociology*, *9*(4). doi: 10.14254/2071-789x.2016/9-4/6.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*. doi: 10.2307/2490395.

Pawelek, B., Galuszka, K., Kostrzewska, J., & Kostrzewski, M. (2015). Classification methods in the research on the financial standing of construction enterprises after bankruptcy in Poland. In F. Palumbo, A. Montanari, M. Vichi (Eds.). *Data science: innovative developments in data analysis and clustering*. Springer International Pulishing. doi: 10.1007/978-3-319-55723-6_3.

Pawelek, B., Kostrzewska, J., & Lipieta, A. (2015). The problem of outliers in the research on the financial standing of construction enterprises in Poland. In M. Papiez & S. Smiech (Eds.). *Proceedings of 9th professor Aleksander Zelias international conference on modelling and forecasting of socio-economic phenomena*. Cracow: Foundation of the Cracow University of Economics.

Siekelova, A., Spuchlakova, E., & Zvarikova, K. (2017). Selected predictive models of financial health based on the multiple discriminant analysis method and their application in selected enterprises. *Podnikova ekonomika a manazment*, *2*.

Szetela, B., Mentel, G., & Brozyna, J. (2016). In search of insolvency among European countries. *Economic Research - Ekonomska Istraživanja*, *29*(1). doi: 10.1080/1331677x.2016.1237301.

Tabachnick, B. G., & Fidell, L. S., (2012). *Using multivariate statistics*. Boston, USA: Pearson Education.

Tsai, C. F., & Cheng, K. C. (2012). Simple instance selection for bankruptcy prediction. *Knowledge-based Systems*, *27*. doi: 10.1016/j.knosys.2011.09.017.

Valaskova, K., Svabova, L., & Durica, M. (2017). Verification of prediction models in conditions of the Slovak agricultural resort. *Economics, Management, Innovation*, *9*(3).

Wilks, S. S. (1946). Sample criteria for testing equality of means, equality of variances, and equality of covariances in a normal multivariate distribution. *Annals of Mathematical Statistics*, *17*(3). doi: 10.1214/aoms/1177730940.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, *22*. doi: 10.2307/2490859.

Zvarikova, K., Spuchlakova, E., & Sopkova, G., (2017). International comparison of the relevant variables in the chosen bankruptcy models used in the risk management. *Oeconomia Copernicana*, *8*(1). pp. 143-155. doi: 10.24136/oc.v8i1.10.

Acknowledgments

The paper is an output of the science project VEGA 1/0428/17 Creation of new paradigms of financial management at the threshold of the 21st century in conditions of the Slovak republic.

Annex

Table 1. Financial ratios used in analysis

Ratio name	Group	Method for calculation	Ratio name	Group	Method for calculation
$\boldsymbol{R1}$	ratios of profitability	ROA EAT (net return / assets)	$\boldsymbol{Z1}$	ratios of debt and capital structure	coverage ratio of total assets / retained earnings
$\boldsymbol{R2}$		ROA EBT (gross return / assets)	$\boldsymbol{Z2}$		total debts ratio
$\boldsymbol{R3}$		net return / total incomes	$\boldsymbol{Z3}$		current debts ratio
$\boldsymbol{L1}$	ratios of liquidity	cash ratio	$\boldsymbol{Z4}$		loan / assets ratio
$\boldsymbol{L2}$		quick ratio	$\boldsymbol{Z5}$		equity / debt
$\boldsymbol{L3}$		current ratio	$\boldsymbol{A1}$	ratios of activity	total incomes / assets
$\boldsymbol{L4}$		net working capital ratio	$\boldsymbol{A2}$		total incomes / current assets

Table 2. Cross-table of variables “being an outlying company” and “being a non-prosperous company”

	Prosperous	Non-prosperous	Total
Non-outlier	53 189 84.86%	9 487 15.14%	62 676 100.00%
Outlier	154 60.16%	102 39.84%	256 100.00%
Total	53 343 84.76%	9 589 15.24%	62 932 100.00%

Table 3. Statistical characteristics of prosperous and non-prosperous sets of outliers and non-outliers

Variable		**Outliers**		**Non-outliers**	
		Mean	**Std Dev**	**Mean**	**Std Dev**
Prosperous	R1	38.06	299.07	0.086	43678.00
	R2	45.24	315.95	0.131	41275.00
	R3	-71937.00	929 240.49	-14.8	512.91
	L1	454.00	1 068.33	43710.00	21490.00
	L2	2 662.19	24 412.92	30042.00	24.73
	L3	2 766.02	24 405.80	5.221	24.95
	L4	-232.52	1 921.27	0.09	31107.00
	Z1	-383.39	2 478.74	-0.21	16558.00
	Z2	288.56	2 042.15	0.73	35125.00
	Z3	233.21	1 921.30	0.62	31107.00
	Z4	42736.00	14062.00	0.08	0.31
	Z5	3 961.32	24 950.8	43529.00	32.19
	A1	1476223.00	6 127 126.36	446.73	15 105.44
	A2	138776.00	723 068.2	147.37	3 783.91

Table 4. Continued

	Variable	Outliers		Non-outliers	
		Mean	Std Dev	Mean	Std Dev
Non-prosperous	R1	-424.78	2 085.86	-1.1	6.816
	R2	-408.91	2 054.66	-0.99	24624.00
	R3	-19754.00	89 257.82	-163.79	3 282.03
	L1	0.04	0.11	0.12	0.17
	L2	0.09	0.2	0.31	0.27
	L3	0.12	0.26	0.41	0.3
	L4	-1 515.5	6 570.94	-3.43	18.35
	Z1	-1 473.2	5 876.81	-3.42	18.82
	Z2	1 561.51	6 575.34	12510.00	18.94
	Z3	1 516.21	6 571.04	43620.00	18.39
	Z4	27.81	123.61	0.24	43497.00
	Z5	-0.79	0.38	-0.37	0.31
	A1	431871.00	1416639.00	1 828.65	33 744.57
	A2	115523.00	380 569.38	366.97	5 487.10

Table 5. t-test for equality of means between prosperous and non-prosperous companies among outliers

Variable	Equal variances	T	Sig. (2-tailed)
R1	not assumed	2.226	0.028
R2	not assumed	2.215	0.029
R3	assumed	-0.565	0.573
L1	not assumed	5.273	0.000
L2	not assumed	1.353	0.178
L3	not assumed	1.406	0.162
L4	not assumed	1.918	0.058
Z1	not assumed	1.771	0.079
Z2	not assumed	-1.896	0.061
Z3	not assumed	-1.918	0.058
Z4	not assumed	-2.174	0.032
Z5	not assumed	1.971	0.051
A1	not assumed	2.035	0.043
A2	assumed	0.298	0.766

Table 6. Contingency table of prosperous and non-prosperous companies among outliers and non-outliers

outlier * prosperity Cross-table					
			prosperous	non-prosperous	Total
outlier	No	Count	53189	9487	62676
		% within outlier	84.86%	15.14%	100.00%
	Yes	Count	154	102	256
		% within outlier	60.16%	39.84%	100.00%
Total		Count	53343	9589	62932
		% within outlier	84.76%	15.24%	100.00%

Table 7. Test of independence of variables "being an outlier" and "being a non-prosperous company"

Chi-Square Tests	Value	Asymp. Sig.
Pearson Chi-Square	120.506	0.000
Continuity Correction	118.600	0.000
Likelihood Ratio	90.966	0.000

Table 8. Correlation measures of variables "being an outlier" and "being a non-prosperous company"

Correlation Measures	Value	Approx. Sig.
Phi	0.044	0.000
Cramer's V	0.044	0.000

Table 9. Test of equality of proportions of non-prosperous companies among outliers and non-outliers

Test	Value	Asymp. Sig.
Proportions equality test	10.978	0.000