FINANCIAL DISTRESS CRITERIA DEFINED BY CLUSTERING OF LONGITUDINAL DATA

Maria Stachova – Lukas Sobisek

Abstract

Financial distress is a situation in which a company cannot pay or has a difficulty to pay off its financial obligations. The crucial step in each analysis of companies' financial status is to define the criteria that would describe the financial difficulties of enterprises with high accuracy. In our analyses of financial distress of Slovak companies, three criteria are considered: (1) the equity, (2) the earnings after taxes and (3) the current ratio value. Our main goal is to investigate if it is possible to identify homogeneous clusters regarding the companies' financial health by using the financial longitudinal data. The *k*-means partitioning of companies is based on time trajectories of these criteria. The clustering is made by using statistical system R and the package "kml". The results show, that recognition of patterns in panel data set can be helpful in the process of financial distress analysis, but it is necessary to add an expert point of view as well.

Key words: financial distress, longitudinal data clustering, *k*-means partitioning

JEL Code: C38, G33

Introduction

There exist many valid alternatives how to define financially distressed (or healthy) companies. Generally, we can say, that financial distress is the situation in which company is not able to pay off its financial obligation. Authors in (Baixauli & Modica-Milo, 2010) use four financial indicators of profitability and an expert opinion to define financial distress. Abid & Zouari (2000) use a modified (option pricing) Black-Scholes model. Li & Liu (2009) state, that company is in financial distress if their economic results for two consecutive years are negative. In the paper (Stachova & al., 2015), it is defining financial distress as a situation in which a company went bankrupt, has some ongoing liquidation or has some overdue obligations (on an obligation).

Our analysis is based on an approach to financial distress as defined in (Boďa & Úradníček, 2016). "An enterprise was considered financially distressed if

a) its Equity was negative,
b) its EAT (Earnings after taxes) was negative,
c) its Current ratio attained a value lower than 1.

All the three conditions have to be satisfied in order for an enterprise to be considered financially distressed."

The condition of the negative value of Equity comes from the legislative of the Slovak Republic, and it can indicate the high level of indebtedness of a company. The second condition, negative Earnings after taxes, can signalise unprofitable business operations. Because all of the enterprises used in this analysis operate in a homogeneous environment, there is a reason to monitor this criterion. The last indicator (Current ratio) is connected with undercapitalised state of an enterprise. This state is described in (Boďa & Úradníček, 2016) as "a situation in which long-term assets are covered not only by long-term liabilities but also by a fraction of current liabilities. Such a situation is not sustainable in the long term and it also implies an accounting form of insolvency. By definition, an undercapitalised enterprise is unable to liquidate all its current assets to pay off or settle all its current liabilities. Too low a level of liquidity is an impediment to honouring short-term payables and implementing investment strategies. The significance of liquidity for Slovak enterprises is discussed e.g. by Strachotová (2012, pp. 3-4) who assesses the liquidity of the whole entrepreneurial sector in Slovakia as well."

The models for classifying and predicting whether a company is a potential candidate for being financially distressed has become a subject of many studies since well-known Altman's Z-score (Altman, 1968) and its revision (Altman, 1983). Many similar studies and approaches are based on static classification models constructed using various statistical methods, -e.g. discriminant analysis, logistic regression, decision trees (Boďa & Úradníček, 2016; Balcean & Ooghe, 2006; Brezigar-Masten, 2012). We believe, that time dynamic is being incorporated into these well-known static models can improve their predictive accuracy. This idea is supported by studies presented in (Kráľ & al., 2014; Stachová & al. 2015).

We assume that the power of static financial distress predictive models might be enhanced by adding information about negative dynamics of the financial indicators to their static cut-off values. Thus, the aim of our contribution is to investigate if it is possible to identify homogeneous clusters regarding the companies' financial distress by using the financial longitudinal data collected over four consecutive years. In order to satisfy this goal, we use the *K*-means partitioning of companies based on their time trajectories of values of three criteria mentioned above.

The paper is divided as follows. In section 1 we shortly present the data set and the methodology; Section 2 consists of the results achieved in our analysis and in Section 3 we conclude and discuss our achievements.

1 Data and Methodology

In the paper, we use data set that consists of 3 numeric financial distress indicators from 2900 companies that are recognized as sectors of Manufacturing, Construction, and Wholesale and retail trade, repair of motor vehicles and motorcycles, in accordance with SK NACE classification. This database is purchased from the Slovak corporate analytical agency CRIF – Slovak Credit Bureau, s.r.o. (http://www.crif.sk) and covers the data from year 2010 to 2013. The selected companies belong to the higher risk section of economic activities in order of the number of bankruptcy declarations. Descriptive statistics of yearly percentage change of indicators can be found in Table 1.

Tab. 1: Descriptive Statistics of Mean Annualized Percentage Change of Indicators

Financial Indicator	Mean	Standard deviation	Median
Equity	1.24	34.5	0.75
EAT	-17.86	35.55	-24.5
Current ratio value	1.47	22.45	0

Source: the authors.

We expect, that the significant changes in the values of each of these three criteria can indicate the changes in financial health of a monitored company especially if they decrease rapidly over the time.

In order to fulfill the goal of our paper we use the *K*-means partitioning to cluster the time trajectories of selected outcomes which is implemented in the R (R Core Team 2013) package “kml” (Genolini & al., 2015) using the same named function *kml*(). We applied the “kml” clustering to the values from year 2010 to 2013 of Equity (VI), Earnings after taxes (EAT) and Current ratio (L_III) separately, trying to find companies that are not labeled as financially distressed from a static expert’s point of view, but that could be labeled as companies at risk according their past trends in values of VI, EAT and L_III.

The *K*-means clustering applied in the “kml” package is modified for the longitudinal data. This algorithm is based on the original *K*-means clustering (MacQueen, 1967). This method minimizes the utility function iteratively for the time *t*, *N* objects according to an assumption of *C* cluster. The utility function can be expressed as follows:

$$\min \sum_{i=1}^{N} \sum_{c=1}^{C} u_{ict} d_{ict}^{2}, \tag{1}$$

where u_{ict} is a degree of appropriateness of the *i*-th object into the *c*-th cluster in the time *t* with conditions:

$$\sum_{c=1}^{C} u_{ict} = 1, \forall i, t, \tag{2}$$

$$\forall u_{ict} : u_{ict} = \begin{cases} 1 & \lVert \mathbf{x}_{it} - \mathbf{h}_{ct} \rVert = \arg\min_{i} \lVert \mathbf{x}_{it} - \mathbf{h}_{ct} \rVert \\ 0 & \text{elsewhere} \end{cases}.$$

We used the euclidean distance $d_{ict} = \lVert \mathbf{x}_{it} - \mathbf{h}_{ct} \rVert$ between *i*-th vector of objects $\mathbf{x}_{it} = (x_{i1t}, \ldots, x_{ijt}, \ldots, x_{iJt})'$ and *c*-th centroid $\mathbf{h}_{ct} = (h_{c1t}, \ldots, h_{cjt}, \ldots, h_{cJt})'$ in the time *t*. We applied the algorithm to the standardised values of variables.

2 Results

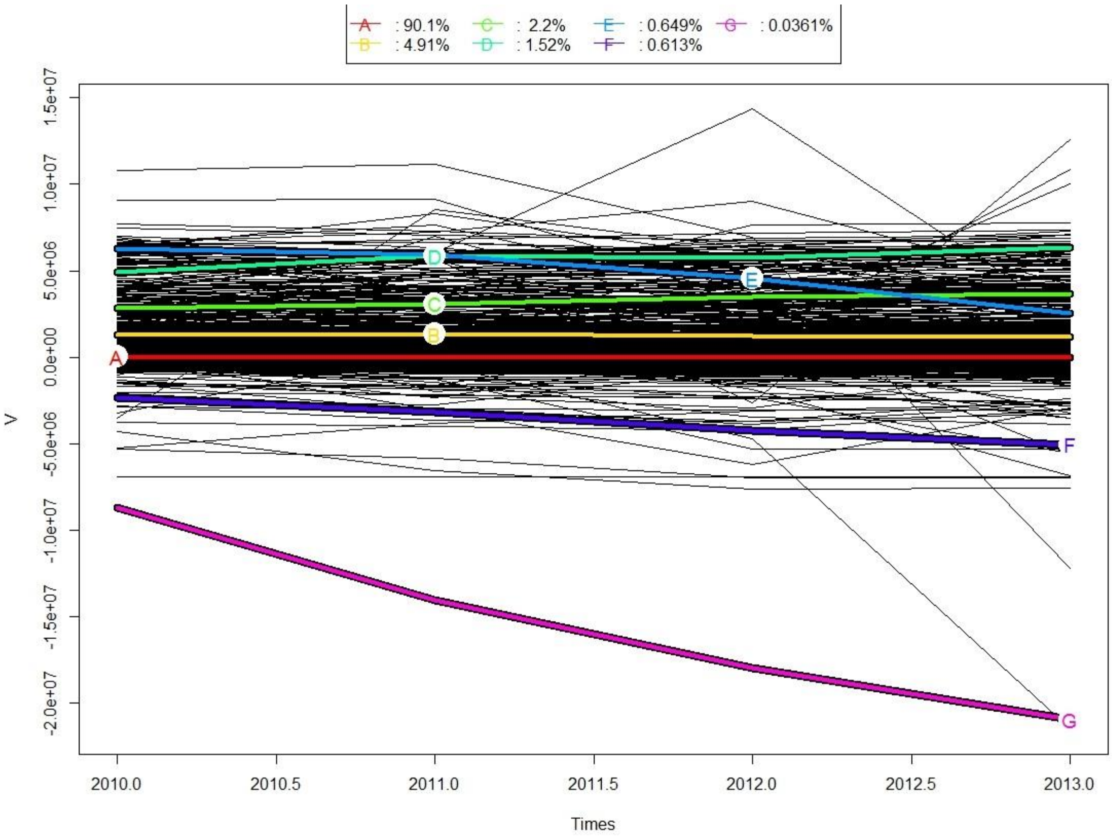
If we follow the static expert’s definition of financial distress as in (Boďa & Úradníček, 2016) and apply it to our data from year 2013, we will find the 901 companies being in financial distress.

In the first step, we applied “kml” clustering to values of Equity and obtained 7 clusters. The means of trajectories of created clusters and the percentage of values in each cluster are shown in Figure 1. The letter V in Figure 1 stands for value of Equity.

It can be seen, that clusters labeled as E, F, G contain the decreasing values of Equity criterion. In comparison with a static expert’s point of view, there are another 25 companies that would be not labeled as financially distressed, because their Equity is not negative, but we found them as at the risk of financial distress, because their Equity values decrease over the time.

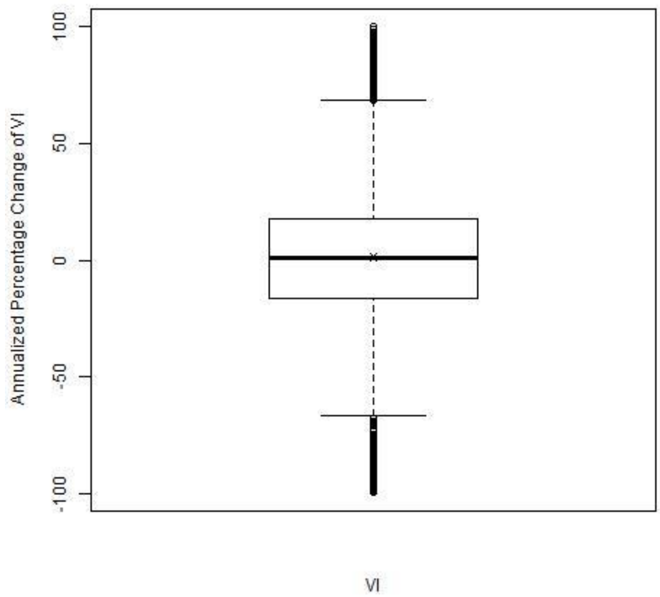
The Figure 2 displays the distribution of annualized percentage change of Equity calculated from the values measured in years 2010 and 2013. Almost half of companies (median of yearly change equals to 0.75, see Table 1) has decreasing trend. According to this fact the “kml” algorithm should create more balanced clusters with the higher number of members. Unfortunately, the result of partitioning is contradictory. We get only one large cluster which contains 90.1% of values and 6 small clusters.

Fig. 1: The K-means clustering of Equity values



Source: the authors.

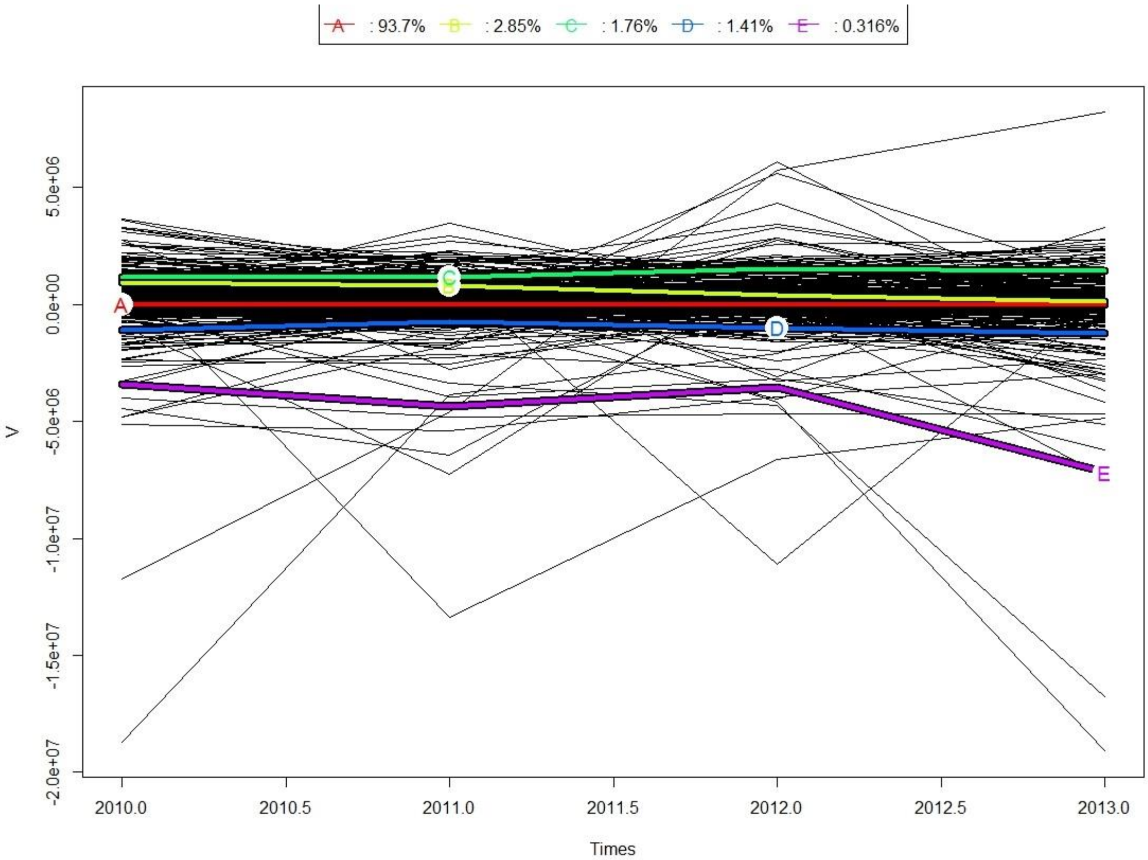
Fig. 2: The box plot of Equity values



Source: the authors.

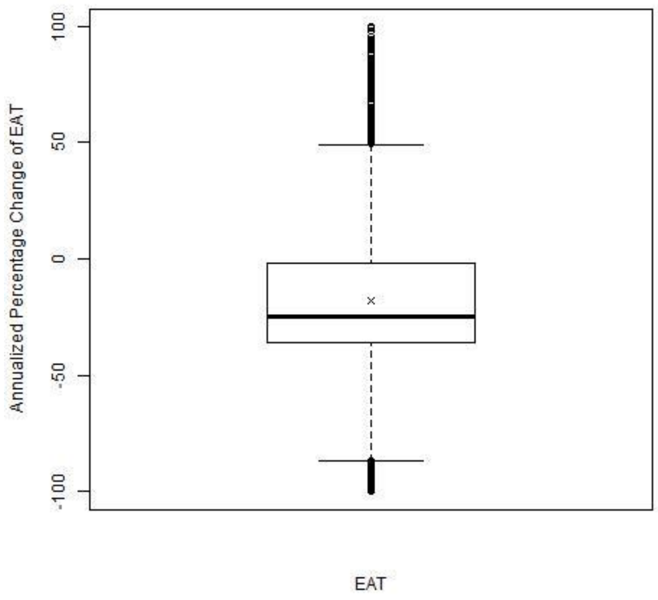
Secondly, we clustered the EAT values and we obtained 5 clusters. Figure 3 shows, that clusters named with letters B, D, E contain the companies with decreasing values of EAT criterion. Another 110 companies were labeled as financially distressed according to this clustering. A box plot in Figure 4 and measurements of central tendencies (mean = -17.86, median = -24.5) again shows, that the “kml” should classify more companies into risky clusters with negative trend.

Fig. 3: The K-means clustering of EAT values



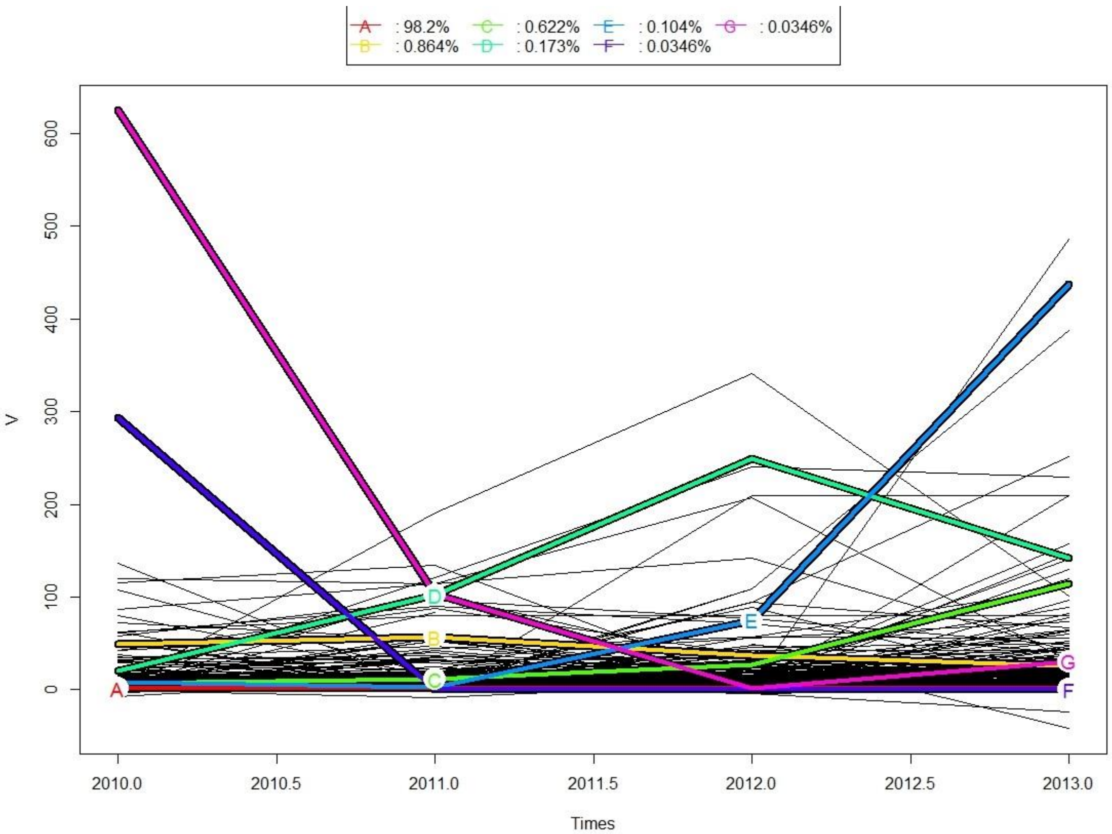
Source: the authors.

Fig. 4: The box plot of EAT values



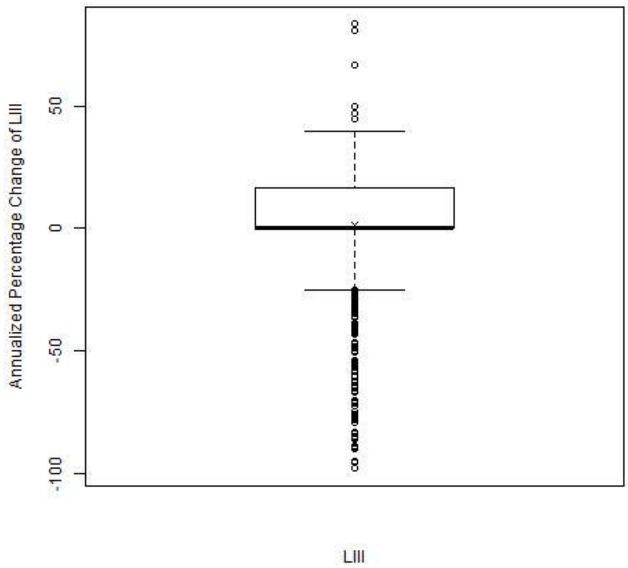
Source: the authors.

Fig. 5: The K-means clustering of Current ratio value



Source: the authors.

Fig. 6: The box plot of Current ratio values



Source: the authors.

Finally, we clustered the Current ratio values. We identified 7 clusters as it is plotted in Figure 5. The decreasing trajectories of this criterion are placed in the clusters B and F. We found another 26 companies at the risk of financial distress that would not be recognized in static expert's point of view. However, the box plot in Figure 6 shows that the percentage of created clusters is not what we would have expected as median of yearly change is 0.

When we combine the results obtained in each step of our analysis, we found overall 115 companies (app. 6 %) that would not be labeled as financially distressed according to the three criteria mentioned in Introduction of this paper, but they are recognized as at the risk of financial distress according to their decreasing values of Equity, or EAT, or Current ratio over the time.

Conclusion

The main aim of this contribution was to investigate if it is possible to identify homogeneous clusters regarding the companies' financial distress by using the financial longitudinal data. The *k*-means partitioning of companies based on time trajectories of the three criteria was applied. The results show, that there is an evidence of companies that should be recorded as being at the risk of financial distress according to their decreasing values of selected criteria.

These companies would not be found using the static expert method that takes into account only positivity/negativity of one-year indicator values. However, we also found, that the "kml" method of clustering of trajectories is very rough and it is no able to identify the majority of companies with negative trend. This fact leads us to investigate in our future work more appropriate clustering algorithm that would recognize better existing patterns of the development of repeated measures in time.

Acknowledgment

Mária Stachová has been supported by the project VEGA 1/0647/14. Lukas Sobisek received financial support for research activity from long term institutional support of research activities by Faculty of Informatics and Statistics, University of Economics, Prague.

References

Altman, E. I., (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. The Journal of Finance, 23(4), 583-609.

Altman, E. I. (1983). *Corporate Financial Distress: A Complete Guide to Predicting, Avoiding, and Dealing with Bankruptcy.* New York: WIley.

Balcaen, S., & Ooghe, H. (2006). 35 Years of Studies on Business Failure: an Overview of the Classic Statistical Methodologies and Their Related Problems.The British Accounting Review, 38(1), 63–93.

Boďa, M., & Úradníček, V. (2016). The portability of Altman's Z-score model to predicting corporate financial distress of Slovak companies. Technological and Economical Development in Economy. Submitted.

Baixauli, J. S., & Modica-Milo, A. (2010). The bias of unhealthy SMEs in bankruptcy prediction models. Journal of Small Business and Enterprise Development, 17(1), 60-77.

Brezigar - Masten, A., & Masten, I. (2012). CART-based selection of bankruptcy predictors for the logit model. Expert Systems with Applications, 39(11), 10153–10159.

Genolini, Ch., Alacoque, X., Sentenac, M, & Arnaud, C. (2015). kml and kml3d: R Packages to Cluster Longitudinal Data. Journal of Statistical Software, 65(4), 1-34. URL http://www.jstatsoft.org/v65/i04/.

Kráľ, P., Stachová, M. & Sobíšek, L. (2014). Utilization of repeatedly measured financial ratios in corporate financial distress prediction in Slovakia: In the 17th AMSE, international scientific conference, conference proceedings, Poland, 156-163.

Li, D., & Liu, J. (2009). Determinants of Financial Distress of ST and PT Companies: A Panel Analysis of Chinese Listed Companies. Retrieved February 12, 2009, from http://ssrn.com/abstract=1341795

MacQueen, J., B. (1967). Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, University of California Press, 281-297.

R Core Team 2013: R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2012, http://www.r-project.org/.

Stachová, M., Kráľ, P., Sobíšek, L., & Kakaščík, M. (2015). Analysis of financial distress of Slovak companies using repeated measurement. In 18th AMSE, International Scientific Conference Proceedings, Czech Republic.

Contact

Maria Stachova

Faculty of Economics, Matej Bel University, Banska Bystrica, Slovakia

Tajovskeho 10, 975 90 Banska Bystrica, Slovakia

maria.stachova@umb.sk

Lukas Sobisek

Faculty of Informatics and Statistics, University of Economics, Prague, Czech Republic

W. Churchill Sq. 4, 130 67 Prague 3, Czech Republic

lukas.sobisek@vse.cz