Management and Business
Administration Doctoral School

THESIS BOOKLET

Tamás Nyitrai

Application of dynamic financial variables in bankruptcy prediction

Ph. D. thesis

Supervisor:

Prof. Dr. Miklós Virág
professor

Budapest, 2015

Department of Enterprise Finances

THESIS BOOKLET

Tamás Nyitrai

Application of dynamic financial variables in bankruptcy prediction

Ph.D. thesis

Supervisor:

Prof. Dr. Miklós Virág

professor

Table of contents

1. About the research

1.1. Research history

I had met the prediction of future solvency of firms as one of the most typical application forms of logistic regression during my studies on statistics at the university for the first time. The interest evolved in those days led me to conduct further studies on this topic after my MSc university degree. I have conducted my doctoral studies under the supervision of the developer of the first Hungarian bankruptcy prediction model, Dr. Miklós Virág, who help me a lot during the early stages of my studies to realize that the field of bankruptcy prediction is much more complex than a simple statistical model fitting task.

On the proposal from my supervisor, I started my scientific research work with reviewing the most important publications appeared in the literature. After the publications which can be considered as milestone studies of bankruptcy prediction, I paid attention to the Hungarian and international literature, especially to reviewing the scientific results achieved in this field during the last 15 years. At this time, I had to face the fact that this topic is really much broader than a simple data fitting task, because the number of publications dealing with bankruptcy prediction is huge and the interest of research community toward this topic has been increasing during the last decade which is indicated by the continuously increasing number of scientific studies devoted to predicting bankruptcy.

With respect to the size of the literature, I had the opportunity to present only the most important publications related to the chosen topic in the thesis due to space limitations. In spite of it, almost 200 scientific journal articles were cited in my work. Having reviewed the literature, my view formed during my university years about bankruptcy prediction has fundamentally changed. I had to face the fact that the outputs of bankruptcy prediction models can not necessarily be interpreted as real predictions. However, it is a common principle in the literature that these models don't provide real predictions. They give only early warning signals regarding the risk of future bankruptcy of firms. That is, a company classified as bankrupt by a model won't necessarily file for bankruptcy in the near future. This result of the model means only that, based on its data, the firm under consideration resembles to the bankrupt firms being in the database used for model building to a greater extent than to the healthy ones (Virág et al. [2013]). So, this fact should be taken into consideration by decision makers when making decisions regarding this firm.

However, in spite of the fact that the outputs of models can't literally be considered as predictions, it is a common practice in the literature that the researchers examine the ratio of bankrupt and healthy firms correctly classified by the model. That is, in some sense, they consider the output of the models as real predictions so that the performance of the models can objectively be measured.

It can be confusing for a reader who is not familiar with the literature of this field that the models published under the headword of "bankruptcy prediction" are often not restricted to predicting only the bankruptcy of firms. The "bankruptcy" to be predicted usually means some kind of realization of insolvency which varies depending on the available data and on the purpose of the research. As a consequence, every study dealing with bankruptcy prediction starts with defining the term "bankruptcy" wished to predict.

In my thesis, I tried to predict the insolvency of firms operating in Hungary. The term "insolvency" was operationalized as the initiation of bankruptcy or liquidation proceeding against the enterprises.

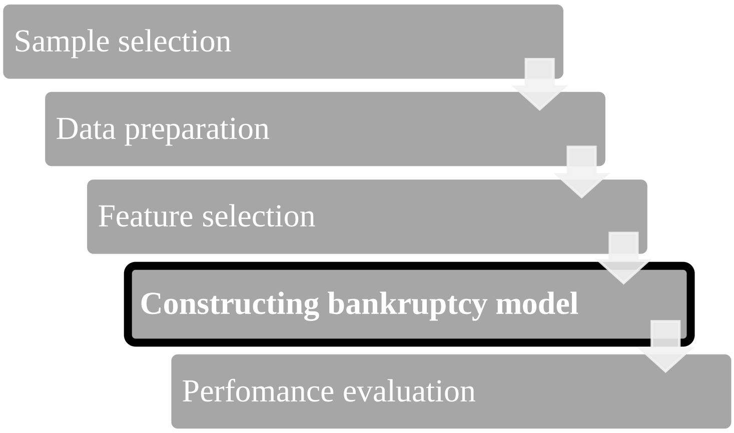
Bankruptcy prediction models are generally developed by means of statistical (data mining) methods. Ratio type financial ratios calculable from financial statements of companies are traditionally used as explanatory variables for models. When developing bankruptcy prediction models, financial ratios for bankrupt and healthy companies from the past are employed in order to explore the possible relationship between the values of ratios and the event of bankruptcy, then the revealed relationship is used to predict the future solvency of firms which were not present in the database used for model building. However, this seemingly easy statistical model fitting task raises much more question than I thought as an undergraduate student.

1.2. The literature of bankruptcy prediction and the chosen topic

If someone began reviewing the literature of bankruptcy prediction starting with the most recent publications, they would easily think that this research topic is considerably confused because there can be great difference between each studies with respect to the applied methods and to the purposes even if they are published in the same journal. The main reason for this that bankruptcy prediction cannot be considered as a standalone scientific field. It is much more reasonable to locate this topic on the frontier of statistics (data mining) and corporate finance because financial variables are used as explanatory variables in statistical models when building bankruptcy prediction models.

In the absence of consensus related to the applied methods and data used for model building among the researchers of this field, several open research question can be found. This characteristic makes the literature of bankruptcy prediction difficult to understand for a reader who is new to this area. Because of it, I thought reasonable to present the literature in a new way which, in my opinion, could give comprehensive picture for the reader about the current state of research on this field. The essence of this approach is to assign each publications to the main phases of bankruptcy prediction model building, instead of the well-known chronological approach. The process of bankruptcy prediction model building is depicted in Figure 1.

Figure 1. The process of bankruptcy prediction model building



Since statistical and data mining methods are essential for building bankruptcy prediction models, it is not surprising that the development of this research field has been dominated by the methodological modernization. The number of methods applicable in bankruptcy prediction has been increasing in line with the modernization of computer sciences and data mining. This trend raised the question of which classification method can be considered as the best in this field. Since the vast majority of research is devoted to this question (Sánchez-Lasheras et al. [2012]), in my opinion, this topic can be considered as a mainstream research direction in bankruptcy prediction. In spite of the huge number of research attempts, the question has still been unanswered because the researchers often found that each classification methods exhibit different predictive ability on different datasets (Oreski et al. [2012]). Studies comparing the performance of classifiers are assigned to the phase called “Developing bankruptcy model” which are highlighted by black line in Figure 1.

Though, most of the studies has still been dealing with the mainstream research direction, consensus is being formed regarding the statement of Marqués et al. [2012], namely that the so called “best method” probably doesn’t exist because it is frequent that the application of newer methodological approaches doesn’t yield significant increase in the predictive power of models. This is likely the reason for the fact that, in the last 5-10 years, the literature has turned to the alternative ways for enhancing the performance of bankruptcy prediction models. These research possibilities can be assigned to the further (non-methodological) phases of bankruptcy prediction models building. I presented several examples of these research directions in the section dealing with the literature review. The studies cited there point out that the predictive power of models can be enhanced not only from methodological side. This could also be achieved by deeper exploring the alternative research questions of bankruptcy prediction.

Bankruptcy prediction models play important role in the practical rating process of debtors, which fact also increases the importance of alternative ways of enhancing the performance of the models. The main raison for it that, due to the modernization of classification methods, data mining techniques have become dominant in this field. This algorithms are able to achieve outstanding hit rates, but their black box nature is their serious drawback. Black box means that these methods don’t express the relationship between the explanatory variables and the occurrence of bankruptcy in an exact mathematical form. This aspect is judged so serious by Martens et al. [2012] that, in their opinion, it is not likely that, in spite of their high classification ability, these methods will become widespread as decision support systems for lending processes. Based on the tendencies detailed above, during my doctoral studies, I have also oriented to alternative research possibilities of bankruptcy prediction. Having finished reviewing the literature, I identified several open research questions of which I chose the dynamization of bankruptcy prediction models.

Though, the problems coming from the static nature of bankruptcy prediction models have been known for a long time (Abdou-Pointon [2011]), relevant attempts at solving them can only be found in the literature only in the last couple of years. Common practice in the research of bankruptcy prediction is to use only the most recent (static) financial data as explanatory variables in the models. However, this may be problematic because this approach doesn’t take into account the tendencies inherited in the time series of financial ratios. This aspect had been first tried to be incorporated by using survival models belonging to the econometric methods (Hillegeist et al. [2004]). Later, more complex data mining algorithms also appeared in this field (Du Jardin-Séverin [2012]). The lastly cited authors used the time series of financial ratios as “explanatory variables” to explore the typical processes of bankruptcy. Though, this dynamic model exhibited higher predictive power than the traditional static approaches, this promising alternative hasn’t become widespread in the literature, presumably because of its methodological complexity.

Based on the presented tendencies observable in the literature, in my thesis, I proposed a variable which is also capable of incorporating the information content inherited in the time series of financial ratios within the framework of traditional classification methods based on the principles of statistical model building. That is, in my work, I made an attempt at dynamizing bankruptcy prediction models by using the well-known classification methodologies. I named the variable proposed in the thesis as dynamic financial variable. I determined its calculation formula as it can be seen below

$$\frac{X_{i,t-1} - X_{i,min_{[t-2;t-n]}}}{X_{i,max_{[t-2;t-n]}} - X_{i,min_{[t-2;t-n]}}}$$

Where, *X* is the given financial ratio, *I* is the company under consideration, *t* is the year I wish to make prediction for, *n* is the length of the time series.

2. Applied methods

2.1. The hypotheses and the applied classification procedure

I investigated the following hypotheses in my work:

1. In the case of my research sample, there are such dynamic financial variables which have statistically significant discriminating power between the bankrupt and healthy companies.

2. Using decision trees developed with CHAID procedure as classification algorithm for bankruptcy prediction models, the predictive power of models containing static financial ratios from the most recent year in combination with the dynamic variables is significantly higher than those of models in which only static financial ratios from the most recent year were used as input variables.

3. In the case of the available sample, predictive power of models containing dynamic financial variables and developed with CHAID algorithm can be increased if outlier values are replaced by such values which are the closest to them in the same time series but not outliers.

The investigation of the presented hypotheses was conducted by using decision trees based on CHAID algorithm. The reasons for the selection were the followings:

- it is a built-in function in SPSS which is a frequently used data analysis software in empirical investigations;
- this method requires low computational time to build models with relatively high classification performance as it can be seen in the Hungarian (Kristóf-Virág [2012]) as well as in the international literature (Koyuncugil-Ozgulbas [2012]);
- the selection was influenced by the argument cited earlier according to which easily interpretable models are preferred in the practical application of bankruptcy prediction (Martens et al. [2010]);
- the application of decision trees was motivated by the fact that their classification performance is not sensitive to the presence of outlier values (Twala [2010].

The last property is important to the thesis because, in my opinion, outlier values have important explanatory power in identifying bankrupt companies since, during the manual data gathering process, I often found that bankrupt companies exhibit extreme high or low values in some of their financial ratios in the year prior to the beginning of the bankruptcy or liquidation proceeding. Based on this experience, keeping outlier values within the models – in their original form as far as possible – was the core concept of the empirical investigation because there isn't any method in the literature which can be considered as an unambiguously preferred approach for handling outliers, however this is not necessary when employing decision trees.

2.2. Data used for empirical investigations

The hypotheses listed above were investigated by using a sample based on my own manual data gathering process. The number of observations was maximized at 1000 which can be considered as low compared to the total number of firms operating in Hungary. However, financial data for firms included in the sample were collected not only for the most recent year but for all the available years until 2001.[1] This kind of data gathering is a very time-consuming task even in the case of such a small number of observations because the employed public databases don't permit to access data in electronic form. Consequently, financial statement data have to be collected manually.

The collected sample was divided evenly between bankrupt and healthy observations. From statistical point of view, it would be desirable to use a representative sample but, for such a small sample size, it would have resulted in a very low number of bankrupt companies which could have raised the question of whether such a database contains enough information for a model to effectively identify bankrupt observations which is the primary aim of bankruptcy prediction models (Du Jardin [2010]). So, in line with the common practice in the literature, I used bankrupt and healthy firms in equal proportions in the sample.

During the data gathering process, firms under liquidation or bankruptcy proceedings according to the Hungarian company register were considered as bankrupt observations, all the other firms as healthy ones. Firms publishing announcement in some randomly chosen number of the Hungarian Company Gazette had chance to get incorporated into the sample. Financial data for the selected firms were manually obtained from the electronic financial reporting portal[2] provided by the Ministry of Justice.

---

[1] In order to ensure the comparability of financial data, sample period ends in 2001 because there was a significant change in the Hungarian accounting rules in 2000.

[2] http://e-beszamolo.im.gov.hu/kereses-Default.aspx

Sampling considerations applied during the gathering process:

- financial data for firms under consideration must be available at least three consecutive years beginning from the most recent year (in the case of bankrupt firms beginning from the year prior to bankruptcy);
- firms having financial ratios with the same value[3] over the entire sample period were excluded;
- companies with more than one year without sales were also excluded because some kind of firms are probably not doing any significant business activity.

Since in the case of the applied public databases it is not possible to restrict the search on the basis of the activity, size and age of the firms, so the resulting sample is considerably heterogeneous with respect to these aspects. Taking into account the fact that the aim of the thesis is only to demonstrate the concept and the use of dynamic financial variables in bankruptcy prediction and not to develop ready-to-use models for rating debtors, such a heterogeneous sample could be advantageous in terms of the purpose of the thesis because if the application of the proposed dynamic variables can enhance the predictive power of models in the case of such a heterogeneous sample, it is likely that this will be the case for a more homogenous sample as well.

Since, ratio type financial variables are commonly used explanatory variables in bankruptcy prediction models (Chen [2012]), furthermore the employment of financial ratios successfully applied in past researches is also a frequent practice in the literature (Du Jardin [2010]), I also followed these principles during my research. The independent variables of the first Hungarian bankruptcy prediction model (Virág et al. [2013]) were applied in the thesis along with some other variables based on my own consideration. The name and the calculation formula of the employed explanatory variables are given in the following table.

[3] with zero standard deviation over time

Table 1. The name and formula of financial ratios used in empirical research

Variable name	Formula
Liquidity ratio I	Current assets/Current liabilities
Liquidity ratio II	(Current assets – Inventories)/Current liabilities
Ratio of cash	Cash/Current assets
Cash flow/Liabilities	(After-tax profit + Depreciation)/Liabilities
Cash flow/Current liabilities	(After-tax profit + Depreciation)/Current liabilities
Equity financing ratio	(Fixed assets + Inventories)/Equity
Assets turnover	Net sales/Assets
Inventory turnover	Net sales/Inventories
Receivable turnover (time)	Receivables/Net sales
Indebtedness	Liabilities/Assets
Equity ratio	Equity/Assets
Creditworthiness	Liabilities/Equity
Return on sales	After-tax profit/Net sales
Return on assets	After-tax profit/Assets
Receivables/Current liabilities	Receivables/Current liabilities
Ratio of working capital	(Current assets – Current liabilities)/Assets
Size	Natural logarithm of Assets
Years	Number of observed years

As a result of the presented data gathering process, there is database consisting of 1000 Hungarian enterprises for which 17 explanatory variables were calculated. This means altogether 7592 firm-year observations over the 2001-2012 period. Since the exact age of the firms in the sample was not available in the applied public databases, I used a proxy variable for it which is the number of years for which financial data for the firms contained in the sample were available in the public database mentioned earlier. Financial ratios for the most recent year were considered as static financial ratios in the thesis. The proposed dynamic variable was calculated for each of the financial ratios presented in Table 1. The dynamic variables express the value of the financial ratios from the most recent year in the mirror of those from the previous period. So, there were 35 potential independent variables for bankruptcy prediction models. The descriptive statistics for these variables are given in Table F.1. which can be found in the Appendix of the thesis.

3. The results of the thesis

3.1. The examination of the first hypothesis

The first hypothesis of my work was formulated as shown below:

In the case of my research sample, there are such dynamic financial variables which have statistically significant discriminating power between the bankrupt and healthy companies.

The examination of the hypothesis was carried out by using the T-test for comparing means obtained from samples of two independent populations. Variables[4] showing significant differences between the two examined groups at 5 % significance level are highlighted by * in Table F.1. of the appendix in the thesis. Based on the results of the T-test, I drew the following conclusions:

1. Out of the examined 35 variables 13 were proven to be significant. Out of the significant variables 7 were static and 5 were dynamic, that is by applying dynamic variables, the set of variables able to discriminate between bankrupt and healthy companies in the sample under consideration has definitely increased.[5]
2. Based on the fact that dynamic variables were calculated for each financial variable, three groups can be distinguished within the set of significant variables. There were financial ratios where
    a. only the static value from the most recent year was significant (Ratio of cash, Ratio of working capital);
    b. both the static and dynamic variables were also significant (Receivables turnover time, Indebtedness, Equity ratio, Return of assets);
    c. only the dynamic variable were significant in terms of discriminating between the two groups (Cash flow/Debt, Inventories turnover).

Based on the obtained results, I accepted the first hypothesis of thesis.

---

[4] Dynamic variables were indicated by the prefix „D_” in Table F.1. of the thesis.
[5] The 13th significant variable was the Years variable which shows the number of years for which financial data were available for the firms in the sample.

3.2. The examination of the second hypothesis

The second of my hypotheses was formulated as shown below:

Using decision trees developed with CHAID procedure as classification algorithm for bankruptcy prediction models, the predictive power of models containing static financial ratios from the most recent year in combination with the dynamic variables is significantly higher than those of models in which only static financial ratios from the most recent year were used as input variables.

During the empirical investigation, the procedure of tenfold cross-validation was applied. The first step of this procedure is the partition of the sample into ten equal parts of which nine parts are joined to form a training sample and the remaining one part is used for testing the model developed on the training sample. The second step is to repeat the previous modelling so that each of the ten parts will be used once as testing sample. In other words, ten models are generated during the ten-fold cross-validation but every model is based on different training samples and tested on different testing sets. The aim of the procedure is to avoid the bias coming from an arbitrarily chosen training and testing sets. The performance of each model is recorded and averaged. The conclusions coming from the experiment are based on the average of the ten models, so the bias from the partition of the sample into training and testing sets can be mitigated. Having carried out the presented cross-validation procedure, the following results were obtained.

Table 2. The average hit rates of CHAID decision trees obtained by using ten-fold cross-validation

Sample	Group	Input variables		
		Static financial ratios	Dynamic variables	Static and dynamic variables
Training	Bankrupt	85.7%	83.2%	84.6%
	Non-bankrupt	80.1%	77.8%	**83.8%**
	Total	82.9%	80.5%	**84.2%**
Testing	Bankrupt	76.6%	75.0%	**76.8%**
	Non-bankrupt	71.0%	71.6%	**74.6%**
	Total	73.8%	73.3%	**75.7%**

Based on the results in the table, the following conclusions can be drawn:

1. The predictive power didn't increase by applying only dynamic variables as independent variables in the models relative to the performance of models based only on static financial ratios.
2. When dynamic and static variables were simultaneously employed, the predictive power of models increased in the case of bankrupt and healthy companies as well. In comparison with the performance of models based only on static financial ratios, the average increment is 1.9 percentage point in terms of the testing samples.

The results suggest that there may be complementary relationship between the proposed dynamic variables and static financial ratios, that is, the dynamic variables bring such information into the models which is not present in the values of static financial ratios. On the basis of the presented results, I also accepted the second research hypothesis of my work.

3.3. The examination of the third hypothesis

When investigating longer time series of financial ratios for each firms, there can be found such years which don't fit into the tendency formed by the values from the other years. Such outlier values can seriously distort[6] the values of the proposed dynamic variable, so I think, it is reasonable to handle the outlier values occurring in the time series of financial ratios for each firms. I carried out this by replacing the outlier values. When a value could be considered as outlier, then this value was replaced by the value which is the closest to it in the same time series but not outlier.

To do this, it has to be defined when to consider a value as outlier. In the absence of objective definition, I employed statistical rules of thumb. The values in the time series of each financial ratios were standardized[7] by the mean and the standard deviation of the given time series. The financial position of the firms was described by 17 financial ratios, so there were 17 time series for each firms in the database. The standardization procedure was carried out for each of the 17 financial ratio individually in the case of every company. As the database contained 1000

[6] This distorting effect was illustrated on the basis of a real life example in section 5.3. of the thesis.

[7] In line with the concept outlined in section 2.1 of this thesis booklet, the values from the most recent year were used neither in the standardization nor in the replacement in order to preserve the information content of these values in their original form for the models.

observations, this means 17000 standardizations. Then, the values outside the range of two standard deviations were considered as outlier in every time series.

However, it is open to doubt whether the predictive power of models will be enhanced by handling outlier values in the way presented above compared to the models based on only static financial ratios. If the answer is yes, the extent of the improvement is a further question. The hypothesis aiming at this was formed in the thesis as shown below:

In the case of the available sample, predictive power of models containing dynamic financial variables and developed with CHAID algorithm can be increased if outlier values are replaced by such values which are the closest to them in the same time series but not outliers.

After conducting all the necessary replacements, with the resulting new dynamic variables, the same tenfold cross-validation procedure as it was used in the examination of the second hypothesis was repeated in order to see whether the dynamic variables after handling outliers can enhance the predictive power of the models to a greater extent than as it was seen in the case of the second hypothesis where the outlier values weren't handled. The results are given in the following table. The values in the rows called "Original" show the results without replacing outliers and the rows called "After replacement" show the results of models where such dynamic variables were employed for which, during the calculation, the outliers were replaced in the way presented earlier.

Table 3. The hit rates of models developed in the empirical analysis

Dataset	Sample	Group	Input variables		
			Static financial ratios	Dynamic variables	Static and dynamic variables
Original	Training	Bankrupt	85.7%	83.2%	84.6%
		Non-bankrupt	80.1%	77.8%	83.8%
		Total	82.9%	80.5%	84.2%
	Testing	Bankrupt	76.6%	75.0%	76.8%
		Non-bankrupt	71.0%	71.6%	74.6%
		Total	73.8%	73.3%	75.7%
After replacement	Training	Bankrupt	85.7%	80.0%	86.9%
		Non-bankrupt	80.1%	81.8%	83.1%
		Total	82.9%	80.9%	85.0%
	Testing	Bankrupt	76.6%	71.6%	80.2%
		Non-bankrupt	71.0%	72.4%	75.8%
		Total	73.8%	72.0%	78.0%

Based on the results, the following conclusions can be drawn:

1. The predictive power of models containing only dynamic variables obtained after replacing outliers weren't enhanced when compared to the models based on only static financial ratios. So, the dynamic variables are not worth employing alone as independent variables for bankruptcy prediction models.
2. However, considerable improvement can be observed when static financial ratios and the proposed dynamic variables after handling outliers were simultaneously used as independent variables. Compared to the predictive power of models containing only static financial ratios, the average extent of the improvement is 4.2 percentage points in the case of the testing samples. This improvement is higher than it was in the case of models in which static financial ratios and dynamic variables without handling outliers were used as explanatory variables. In the latter case, only 1.9 percentage point was the average improvement relative to the models based only on static financial ratios as it could be seen in the case of the second research hypothesis.

3.4. Utilization possibilities of the results obtained in the thesis

In my work, I made an attempt at dynamizing bankruptcy prediction models within the framework of decision trees commonly used in the practice and in the literature of the topic. The results of the thesis can be utilized in the scientific research of bankruptcy prediction as well as during the development of credit scoring models supporting decisions on granting credit in the real world.

The empirical research presented in the thesis was carried out by taking into account the need emerging from the side of practical application of bankruptcy prediction models. Hence, I tried to dynamize bankruptcy prediction models by using such a classification algorithm which generates easily interpretable "if-then" rules. The presented results suggest that the proposed dynamic variables have their own raison d'être in bankruptcy prediction since the predictive power, which is the primary measure of the practical usefulness of the models, increased by incorporating dynamic variables into the models.

The results of the empirical investigations supported the main hypothesis of my work, namely that, besides the static financial ratios, relevant information can be found in the evolution of financial ratios over time. I proposed an easy and, at the same time, efficient possibility for taking this information into account. The concept of dynamic variables are easy to implement within the practical process of rating debtors because the calculation is simple and it can be automated, furthermore, it doesn't need using complex methodological solutions.

Since the advantageous impact of dynamic variables was observed only in the case when those were applied along with the static financial ratios, complementary relationship can be assumed between the two sets of variables. Though, several future research are needed for examining the robustness of the results, furthermore, the limitations coming from the way of the empirical research and from the available data should be taken into account, it can be concluded that the potential set of explanatory variables applicable in bankruptcy prediction has considerably increased by the concept of dynamic variables. This fact opens several completely new future research direction for scientists interested in this topic.

The importance of the dynamic aspect is highlighted by the fact that five dynamic variables exhibited statistically significant difference between the bankrupt and healthy companies contained in the available sample. The occurrence of bankruptcy is generally the final outcome of a process, however, this process is often left out during the modeling. The proposed dynamic variables offer possibility for taking this process into account. By analyzing this process, we can get deeper understanding in this research field, so the proposed dynamic variables could play role in developing the theoretical framework of bankruptcy prediction which has been missing for decades.

4. References

Abdou, H. – Pointon, J. [2011]: Credit scoring, statistical techniques and evaluation criteria: a review of the literature. *Intelligent Systems in Accounting, Finance and Management.* Vol. 18., p. 59-88., DOI: http://dx.doi.org/10.1002/isaf.325.

Altman, E. I. [1968]: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance.* Vol. 23., No. 4., p. 589-609, DOI: http://dx.doi.org/0.1111/j.1540-6261.1968.tb00843.x.

Beaver, W. H. [1966]: Financial ratios as predictors of failure. *Journal of Accounting Research.* Vol. 4., Supplement, p. 71-111, DOI: http://dx.doi.org/10.2307/2490171.

Chen, J. H. [2012]: Developing SFNN models to predict financial distress of construction companies. *Expert Systems with Applications.* Vol. 39., p. 823-827, DOI: http://dx.doi.org/10.1016/j.eswa.2011.07.080.

Du Jardin, P. – Séverin, E. [2012]: Forecasting financial failure using a Kohonen map: A comparative study to improve model stability over time. *European Journal of Operational Research.* Vol. 221., p. 378-396, DOI: http://dx.doi.org/10.1016/j.ejor.2012.04.006.

Du Jardin, P. [2010]: Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing.* Vol. 73., p. 2047-2060, DOI: http://dx.doi.org/10.1016/j.neucom.2009.11.034.

Hillegeist, S. A. – Keating, E. K. – Cram, D. P. – Lundstedt, K. G. [2004]: Assessing the probability of bankruptcy. *Review of Accounting Studies.* Vol. 9., p. 5-34., DOI: http://dx.doi.org/10.2139/ssrn.307479.

Koyuncugil, A. S. – Ozgulbas, N. [2012]: Financial early warning system model and data mining application for risk detection. *Expert Systems with Applications.* Vol. 39., p. 6238-6253, DOI: http://dx.doi.org/10.1016/j.eswa.2011.12.021.

Kristóf, T. – Virág, M. [2012]: Data reduction and univariate splitting – Do they together provide better corporate bankruptcy prediction? *Acta Oeconomica.* Vol. 62., No. 2., p. 205-227, DOI: http://dx.doi.org/10.1556/AOecon.62.2012.2.4.

Marqués, A. I. – García, V. – Sánchez, J. S. [2012]: Two-level classifier ensembles for credit risk assessment. *Expert Systems with Applications.* Vol. 39., p. 10916-10922, DOI: http://dx.doi.org/10.1016/j.eswa.2012.03.033.

Martens, D. – Van Gestel, T. – De Backer, M. – Haesen, R. – Vanthienen, J. – Baesens, B. [2010]: Credit rating prediction using ant colony optimization. *Journal of Operational Research Society.* Vol. 61., No. 4., p. 561-573, DOI: http://dx.doi.org/10.1057/jors.2008.164.

Oreski, S. – Oreski, D. – Oreski, G. [2012]: Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment. *Expert Systems with Applications.* Vol. 39., p. 12605-12617, DOI: http://dx.doi.org/10.1016/j.eswa.2012.05.023.

Sánchez-Lasheras, F. – De Andrés, J. – Lorca, P. – De Cos Juez, F. J. [2012]: A hybrid device for the solution of sampling bias problems in the forecasting of firms' bankruptcy. *Expert Systems with Applications.* Vol. 39., p. 7512-7523, DOI: http://dx.doi.org/10.1016/j.eswa.2012.01.135.

Twala, B. [2010]: Multiple classifier application to credit risk assessment. *Expert Systems with Applications.* Vol. 37., p. 3326-3336, DOI: http://dx.doi.org/10.1016/j.eswa.2009.10.018.

Virág, M. – Kristóf, T. – Fiáth, A. – Varsányi, J. [2013]: *Pénzügyi elemzés, csődelőrejelzés, válságkezelés.* Kossuth Kiadó, Budapest.

5. The author's own publications related to this topic

<u>*Publications in Hungarian:*</u>

Journal articles:

1. Nyitrai, T. [2014]: Növelhető-e a csőd-előrejelző modellek előrejelző képessége az új klasszifikációs módszerek nélkül? *Közgazdasági Szemle,* Vol. 61., No. 5., p. 566-585.

2. Nyitrai, T. [2014]: Validációs eljárások a csődelőrejelző modellek teljesítményének megítélésében. *Statisztikai Szemle*, Vol. 92., No., 4., p. 357-377.

3. Nyitrai, T. [2015]: Hazai vállalkozások csődjének előrejelzése a csődeseményt megelőző egy, két, illetve három évvel korábbi pénzügyi beszámolók adatai alapján. *Vezetéstudomány,* Vol. 46., No. 5., p. 55-65.

4. Virág, M. – Nyitrai, T. [2014]: Metamódszerek alkalmazása a csődelőrejelzésben. *Hitelintézeti Szemle,* Vol. 13., No. 4., p. 180-195.

Study published in book:

1. Virág, M. – Nyitrai, T. [2015]: Csődelőrejelző modellek dinamizálása – a szakértői tudás megjelenítése a csődelőrejelzésben. *Vezetés és szervezet társadalmi kontextusban. Tanulmányok Dobák Miklós 60. születésnapja tiszteletére.* Editors: Bakacsi, Gy. – Balaton, K. Akadémiai Kiadó, Budapest, ISBN 978 963 05 9634 3, p. 284-304.

<u>*Publications in English*</u>

Journal articles:

1. Virág M. – Nyitrai T. [2014]: Is there a trade-off between the predictive power and the interpretability of bankruptcy models? The case of the first Hungarian bankruptcy prediction model, *Acta Oeconomica*, Vol. 64, No. 4., p. 419-440, DOI: http://dx.doi.org/10.1556/AOecon.64.2014.4.2.

2. Virág, M. – Nyitrai, T. [2013]: Application of support vector machines on the basis of the first Hungarian bankruptcy model. *Society and Economy.* Vol. 35., No. 2., p. 227-248, DOI: http://dx.doi.org/10.1556/SocEc.35.2013.2.6.