

Prediction of Insolvency of Hungarian Micro Enterprises

Daniel Boda

Szent Istvan University, Hungary

Martin Luptak

Bureau van Dijk, Slovakia

Laszlo Pitlik

Szent Istvan University, Hungary

Gabor Szucs

Szent Istvan University, Hungary

Istvan Takacs

Karoly Robert University College, Hungary

Abstract

The aim of the study is to establish insolvency forecast model with the usage of different statistical methods and compare their efficiency. Besides this the relation and direction between indebtedness and financial distress is also part of the examination. With different approaches we nearly reached the same efficiency, the main focus was on the independent testing sample where we did not apply any modification on the dataset supposing realistic circumstances for predicting the probability of default. The research is focusing on small companies, since their number in the economy is considered high, but for this segment such insolvency forecasts are very rare.

Keywords: bankruptcy, market, forecasting

JEL classification: G17, G31, G33

Introduction

The examination of corporate insolvency already has a long history and the development of the technology and the forecasting models always open new perspectives for the researchers. Nowadays this issue mainly affects the banking industry since the loan placements are considered as their main activity however it would be a mistake to shorten the examination only for this segment since all those areas might be interesting where the delayed payment used as the most common payment method. In all cases where the delivery was not followed immediately by the payment, there we have to count with the insolvency risk which reduction is a key factor in preserving the liquidity, in extreme cases in the survival of the company. In most of the cases for the reliable forecasting the data is not available or it would be against the interest of the managers to make it public. But one part of this information is available from public sources thanks for the different legislations and supervisory bodies. These are the financial statements of the companies and this data makes us capable to search for relations between the behaviour of financial ratios and corporate insolvency.

There have been number of studies in this field of research, starting with Beaver (1966) who used univariate model for examining 79 solvent and 79 insolvent companies. The legal entity was considered insolvent when it was in bankruptcy, the

balance of the bank account turned into negative or it could not fill the obligation of bond or dividend payments. The ratios for the forecasting model were calculated from the financial statements of the companies and the cash flow/total assets proved to be the most effective. However this method was not viable due the complexity of the problem but it pointed out the usability of financial ratios which gave new incentives for the further researches. Altman (1968) was the first who used discriminant analysis for identifying the risky companies. In his model he managed to identify with 99 percent accuracy the insolvent company using a sample with 66 elements. The logistic regression was used by Ohlson (1980) first with the same proportion of the insolvent companies in the sample like in the population. The neural networks appeared in the study of Odom and Shrada (1990) and found it at least as effective like the discriminant analysis. In Hungary the first predictive model was developed by Virág and Hajdú (1996) using discriminant analysis and logistic regression on a sample of 156 companies, considers data from the period 1990 to 1991. Arutyunjan (2004) tested both models on agricultural companies and found the latter more effective. Fazekas (2005) in her dissertation tested the main forecasting models extended with the COCO model based on a sample on a five year period from 2000 to 2005.

Methodology and Data Collection

For the current study we used the Orbis database of Bureau van Dijk, where the data of more than 170 million companies is available most of them with updated financial statements. During the selection we made two groups where in the first the main criteria were the active status and availability of the financial statements from 2013. In the second the financials were also set as a necessary criteria but the status had to be bankruptcy or insolvency proceedings where the status date is between the 1st of January and the 31st of December 2014. At the beginning the sample size was 2400 entities than we excluded all of those companies where the turnover did not reach the 5 million HUF or did not have any employees. In such cases we cannot speak about real business activity that's why they should not take part of the analysis. Because of this, the number of active companies was more than the insolvents so we sorted the latter group in alphabetical order than excluded so many companies to have the same group size. Since the company name does not have any effect on the financial ratios than we found this method applicable for the random selection. The next phase was selecting the companies which had a turnover less than 50 Million HUF since the model examines only the micro enterprises. Than the final sample contained 1534 companies which contained the same amount of solvent and insolvent companies. For building the model we used the adjusted values of the variables the value set in many cases can be $-\infty$, or $+\infty$ and it would make the interpretation more difficult. That's why we used the 5th and 95th percentiles as minimum and maximum values. Those ratios were excluded where both the numerator and denominator can have a negative value which would lead to misunderstandings.

Results

One of the used methods was the binary logistic regression since the observed events can only have two status, 0=solvent, 1=insolvent. The independent variables can contain constant and nominal data, to the probability of the occurrence of the event the logistic regression gives estimation. The big advantage of the approach against the classification methods that the here the normal distribution of the

variables is not an obligation, additionally the regression coefficients can be explained as probabilities. The odds namely probability is the ratio of the solvency and insolvency which natural logarithm is the linear function of the independent variables.

$$\ln(\text{odds}) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$$

The probability of default ($P=1$) is the following:

$$P = \text{odds} / (1 + \text{odds}), \text{ where } \text{odds} = \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k),$$

$$P = \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k) / (1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k))$$

Where the probability of default can be in this format:

$$P = 1 / (1 + \exp(-(\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k)))$$

In order to presuppose non-monotonic ceteris paribus relationships not using any kind of premises, it worth pursuing an explorative modelling. In this context the explorative modelling is the capability that the function fitting can take rational formations between X_i and Y so the ceteris paribus relations could freely build polynomial figures....without external intervention of human experts.

One of the applied non-linear method in the credit management are the artificial neural networks. They provide an alternative to discriminant analysis and linear regression mostly in those cases where the variables show non-linear behaviour. The main critics against NN are the long training process and the risk of overlearning. The input layers are the financial ratios where the inputs are being selected and distribute the result to the next level neurons (Pacelli, Azzollini [2011]). Then they forward the information to the middle layer's neurons. The output in this case is binary, because the companies have only two statuses, solvent or insolvent.

One of the similarity analyses (COCO¹-MCM) has such kind of modelling characteristics which point is: We rank all X_i sets in a separate way based on the more (is the raw value of an attribute) the less (is the ranking value) principle, where the objects are the companies. The ranking values will be used as a kind of input for a stair case function, where the stair levels will be derived in frame of a (non-) linear programming as substitution values of the ranks. The objective function for the (n)LP is the minimizing of the difference between estimations and insolvency codes. The estimation can be produced in an additive way based on the adequate substitution values of attributes pro object.

If we do not apply additional restrictions in the model related to the evolution of the steps, then the result will be such staircase function which contains exchange/substitution value at some particular steps but at others not necessarily. The exchange/substitution value can be considered as a basic rule which states if the input size which determines the given step of the given X_i is valid than for the development of the conclusion a given quantity of part impact was detected, where the resultant of the part impact is the final conclusion².

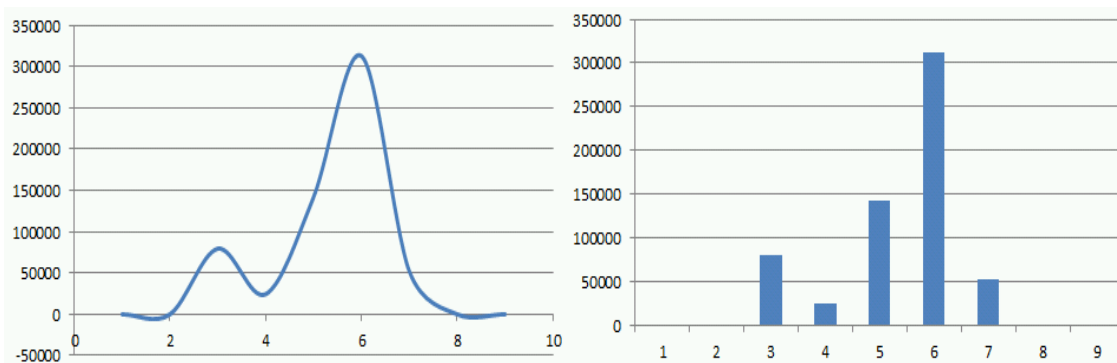
¹ coco = component-based object comparison for objectivity

² <http://miau.gau.hu/myx-free/>

The Xi specific result of the explorative modelling is a conjecture or suspicion which express how could the effect of the presumptively average ceteris paribus of the Xi and Y relation, which should be also known that can be any kind depending on the general conditions (f.e. in case of good market growth potentials the effect of the capital subventions can boost the expected turnover of the company however in another case, where the market is not interested new products to absorb, a capital subvention can generate irrational internal processes in a company (e.g. developing and launching new productstoo fast - with the negative result of falling turnover.

Figure 1

Impact of shareholders funds/total assets ratio described in 9 stairs (X) to the solvency (Y) measured with fitting scores



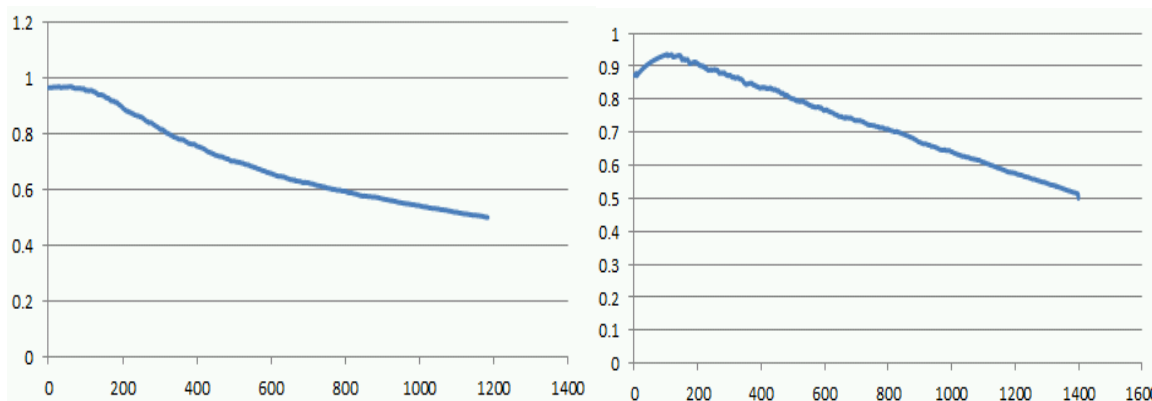
Source: Orbis

Chart x: example for the suspicion generated by the explorative modelling in case of the shareholders funds/total assets ratio, where instead of the monotone interaction an optimum can be supposed which was formulated between the generated categories (9 pieces with homogenous ratios) on the X axis and the probability of the insolvency evasion (Y). With other words: the low shareholders funds ratio is just as risky as the too high shareholders funds in relation to the total assets. The too low ratio was confirmed by 5/9 of many observations, while the risk of the too high ratio was only confirmed by 2/9 of the observations.

The next question related to the logistic regression is the possibility of the modelling of the insolvency on a non-numeric way so/it means: using rulesets. I.e. within the non-categorical variables can such limits detected where the behaviour of the object is different (suitable for forecasting). If these rules exist then they can be detected with pivot wizards where in the rows stays the Xi ascending set of values while in the columns the existence of insolvency. If we count by rows the proportion of the insolvency considering the values of the previous rows we will get the following figures:

Figure 2

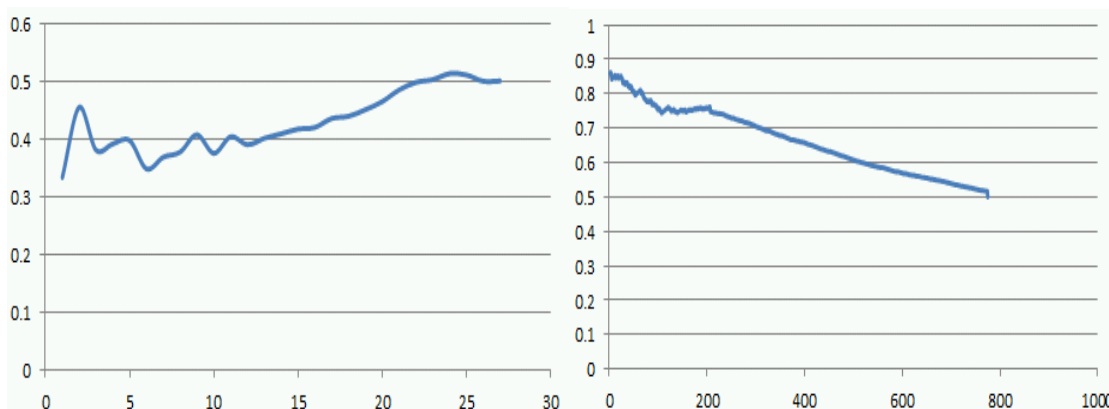
LN Turnover, Y: cumulated insolvency risk and X: liquidity quick ratio, Y: cumulated insolvency risk



Source: Orbis

Figure 3

X: year of foundation, Y: cumulated insolvency risk and X: shareholders funds/total assets ratio, Y: cumulated insolvency risk



Source: Orbis

As shown, the basic graphs of the rule setup are given and these graphs can have a completely different runoff. The direction of the graphs express of the ceteris paribus logic of the formations between the X_i and Y : the higher is the turnover the lower is the probability of default and the elder is a company the lower is the probability of default etc. The relation between the liquidity quick ratio and the shareholders funds/total assets in this context seems to be more complex as well.

If somebody gives a limit related to the discovered relations (e.g. rules cannot be built under the threshold of 80% for the contingency coefficient), then only the following layers have to be examined: the rule is valid for all learning objects or if for one learning object more rules are valid then are they have the same direction regarding the consequences if not what kind of aggregation procedures can provide rational consequences. The WIZWHY³ software is able to perform on an automated way the above mentioned procedures. The staircase functions explored in the frame of the optimization are also describing a kind of rule system which

³ <http://miau.gau.hu/miau2009/index.php3?x=e0&string=wizwhy>

transform into hybrid expert systems since they lead to the conclusions not only on the way of logic but through the way of numeric aggregation of the partial results.

Considering the goodness of the logistic regression model it worth comparing it with other modelling logics in order to be able to control the overlearning phenomenon between the learning and testing and make the further steps on the way of model type optimization⁴.

The examined modelling procedures were the following:

Logistic regression

Neural Networks

Wizwhy

COCO mcm

COCOstd

Table 1

Results of the examined models

input scale	unlimited	unlimited	1--4	1--4	1--4	1--4	1--4	1--4	1--9	1--3
input variables	49	35	31+1	31+1	31+1	31+1	31+1	31+1	31+1	8+1
%	LR	NN	WW	WW	WW	WW	WW	WW-Hybrid	COCO_MCM	COCO_YO
Learning samples (pc)	1534	861	1000	1000	200	200	200	200	200	200
Insolvent companies (50%)	90	94	95	95				n.a.	81	88
Solvent companies (50%)	91	91	92	92				n.a.	100	89
Total	91	92	94	94	97	97	88	93	90	88,5
Test samples (pc)	3181	3181	534	3181	1334	3181	3181	3181	1334	1334
Insolvent companies	81	68	84	85		56	80	80	65	64,6
Solvent companies	69	78	91	58		74	50	74	88	91,5
Total	74	77	88	68	75	67	61	76	76	78,0
All records	80	80	92	74	78	69	63	77	78	79,4

Source: Orbis

Conclusion

The motivation of the study was to examine the relation between the insolvency and the financial ratios and then establish a predictive model which can give significant help for credit decisions in case of micro enterprises. During the history of insolvency prediction there was no consequence within the researchers regarding the financial ratios which can provide to most reliable result. The question remained open and we did not try to define a methodological unified system of criteria. That why the study stays at comparative level, using a significant dataset, number of different approaches were compared in order to see the validity of predictive models for those companies where the company size and the long term operation as presumably determinant factors are missing.

The hybridization has to be considered as a model making possibility and it is able to forecast the insolvency with high probability. The demand for information can be

⁴ <http://miau.gau.hu/miau2009/index.php3?x=e0&string=occam>

decreased in the learning phase without deteriorating the testing results. So the problem of prediction of insolvency has to be handled on a segmented way, the numeric approach with the interaction of all variables seems to be less applicable.

References

1. Altman, E. I. (1968) "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *The Journal of Finance*, Vol. 23, No. 4, pp. 589-609.
2. Arutyunjan, A. (2002) "A mezőgazdaságivállalatokfizetési képességének előrejelzése", Gödöllő: Szent István Egyetem, Közgazdaságtudományi Doktori Iskola (Ph.D. thesis).
3. Beaver, W. (1966) "Financial Ratios as Predictors of Failure, Empirical Research in Accounting: Selected Studies", *Journal of Accounting Research*, Supplement to Vol. 5, pp. 1-111.
4. Brealey, R. A., Myers, S. C. (1999) "Modern vállalatipénzügyek", Panem Kft., Budapest
5. Fazekas, B. (2007) "Vállalati változtatásérváltság menedzsmentés kalmazott pénzügyi módszerek", Thesis, Gödöllő.
6. My-X.hu, available at: <http://miau.gau.hu/myx-free/> (01/05/2016)
7. Odom, M. D., Sharda, R. (1990) "A Neural Network Model for Bankruptcy Prediction", In: *Proceeding of the International Joint Conference on Neural Networks*, San Diego, 17-21 June 1990, Volume II. IEEE Neural Networks Council, Ann Arbor, pp. 163-171.
8. Ohlson, J. (1980) "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, Vol. 18, No. 1, pp. 109-131.
9. Pacelli, V., Azzollini, M. (2011) "An Artificial Neural Network Approach for Credit Risk Management", *Journal of Intelligent Learning Systems and Applications*, pp. 103-112.
10. Ross, Stephen A., (1977) "The Determination of Financial Structure: The Incentive Signaling Approach", *Bell Journal of Economics* 8, 23-40.
11. Sajtos L., Mitev A. (2007), "SPSS kutatási adatalemzési kézikönyv", Alenia Kiadó, Budapest.
12. Székelyi M., Barna I. (2002) "Túlélőkészlet az SPSS-hez", Typotex Kiadó, Budapest.
13. Virág M., Kristóf T., Fiáth A., Varsányi J. (2013) "Pénzügyelemzés, csődelőrejelzés", Kossuth Kiadó, Budapest.

About the authors

Daniel Boda graduated at the Faculty of Finance at Szent István University-Gödöllő. Nowadays he works at Erste Bank in the Risk-Management Department, dealing with corporate risk and in the same time he is developing a probability of default model for Eastern European companies. His MSc thesis won a special award on the so called Scientific Students Associations (TDK) Competition. He starts his Ph.D. studies from this September at Gödöllő. Author can be contacted at bodadaniel1990@gmail.com.

Martin Luptak works for Bureau Van Dijk since 2007. Currently he is working as Central and Eastern Europe Key Account Manager. He is responsible for developing Compliance, Transfer Pricing, Credit risk markets in the region and managing the South sales team. Martin started his career in 2003 as Fundraising manager at Greenpeace right After finishing the Political science and International Relations faculty in Banská Bystrica. Author can be contacted at martin.luptak@bvinfo.com.

László Pitlik, Ph.D., head of the research team for interdisciplinarity and innovation University Gödöllő, Faculty of Economics and Social Sciences Institute of Social Sciences and Teacher Training Research topics: artificial intelligence, socio-cybernetics, big data, similarity analysis, data-driven policy making, automation of decision support, machine learning.

Gabor Szucs is a managing director of Credit reform Hungary; previously he worked in the market research industry. He graduated at the Corvinus University Budapest at the faculty of sociology and economics than on the Budapest Business School at the faculty of Risk Management. He started his Ph.D. studies last year on the field of capital structure and corporate finance. Being involved in several international projects within the Credit reform group, from 2015 January he also holds the position Sales and Marketing Director at AM Holding, owner of 6 Credit reform offices worldwide. Author can be contacted at szucs.gabor@creditreform.hu.

Dr. habil Istvan Takacs Ph.D., Professor at Robert Karoly University and head of Institute of Business Sciences. Since 2007 he is member of the International Assotiation of Agricultural Economists (IAAE) and the National Association of Budget Inspectors where he became a board member in 2010. Member of the Committee of Agricultural Economics at the Hungarian Academy of Sciences since 2014. His fields of research focus on the capital structure and corporate finance.