FORECASTING OF MIGRATION MATRICES IN BUSINESS DEMOGRAPHY

Piotr Gurgul[1], Paweł Zając[2]

ABSTRACT

This paper demonstrates that the forecast of migration matrices can be conducted by means of updating procedures, well-known in the I-O theory. The authors use some of the most popular I-O updating procedures (RAS and some non-biproportional approaches) and calculate measures of the ex-post error of predictions. While taking into account the measures of distance between two matrices, a ranking of forecasting methods of migration matrices (forecast horizon one) is established. Finally, the advantages and drawbacks of particular forecasting methods with respect to one-step ex-post forecasts of migration matrices are discussed.

JEL Classification: C02, L25.

Keywords: Births and deaths of enterprises, migration in branches of industry, prediction, updating methods.

1. Introduction

Firm bankruptcies always bring significant economic losses to stockholders, employees, customers, management, and others. This is accompanied by a substantial social and economic cost to the country. The recent collapse of many global corporations, e.g. Enron, WorldCom, Global Crossing, Adelphia Communications, Tyco, Vivendi, Royal Ahold, HealthSouth, and, in Australia, HIH, One.Tel, Pasminco and Ansett, has aroused interest in various aspects of corporate bankruptcy modelling (also corporate disclosure and auditing regulation). A model forecasting corporate bankruptcy would serve to reduce losses by providing an early warning to stockholders. Early signs of probable distress will enable both management and investors to adopt counter measures. The goal is to shorten the length of time of losses.

---

[1] Piotr Gurgul, AGH University of Science and Technology, Department of Computer Science, Cracow, Poland. E-mail: pgurgul@o2.pl.

[2] Paweł Zając, AGH University of Science and Technology, Department of Applications of Mathematics in Economics, Cracow, Poland. E-mail: pzajac@zarz.agh.edu.pl.

The present study is an attempt to forecast the birth and death of enterprises within input-output methodology.

The next section contains a short overview of existing models in the literature and the methods used for forecasting insolvencies. In section 3, the methodology of enterprise migration matrices is presented, in section 4 methods for forecasting migration matrices are shown, and in section 5 the results of computations are presented and discussed in detail. Finally, section 6 presents the conclusions of the paper.

2. Literature on start-ups and insolvencies

According to the Austrian economist Joseph Schumpeter (1982) the failure of company results from coexisting processes of destruction and creation. The same point is shared by Foster and Kaplan (2001). The authors are convinced that processes of creative destruction are often started by the financial markets. This takes place because financial markets support companies by supplying financial capital as well as withdrawing it with decreasing competitiveness of the company. A developing company can get financial resources. However, after signs of a reduction in competitiveness this support is refused. In case of bankruptcy the company not only needs to deal with its current activity, but also with the mutually connected processes of destruction and creation.

In the 1960's quantitative methods were commonly used to predict the risk of bankruptcy. The pioneer of such research was Altman (1968). The author applied discrimination methods. Discrimination methods allowed a classification of firms into two clusters, those vulnerable to bankruptcy and those not. The clustering was conducted on the basis of financial indicators defined for a sample of companies. The main indicators included the working capital to total assets ratio, the retained earnings to total assets ratio (retained earnings - profits which have not been paid out in dividends and which can be re-invested in the business), the EBIT (Earnings Before Interest and Taxes) to total assets ratio, the market value of equity to the book value of total liabilities and sales to total assets. By means of these indicators Altman classified correctly up to 95 percent of companies the year before their bankruptcy and 83 percent two years before.

Beaver (1966) started empirical research into the prediction of business failure. He used a univariate model. In spite of some shortcomings of the univariate approach, especially a lack of integration of the various ratios, Beaver's model achieved a very good level of predictive accuracy. However, later researchers applied a multivariate discriminant analysis (for short MDA) developed by Altman (1968). Deakin (1972) considerably improved the accuracy of the forecasts of the Altman model. He used all the fourteen ratios which Beaver had identified as good predictors of bankruptcy. Moreover, he used a probabilistic classification rule rather than a critical cut-off point. The model by Blum (1974) included variables expressed in terms of change over time. Wilcox (1971, 1973)

developed the variables applied in the discriminant function by deriving a random walk with a drift formulation of the transition to the failure state. Diamond (1976) developed the multivariate model by using realistic prior probabilities. He also conducted an error cost estimation.

The research started by Altman and continued by others led to the foundation of the American Bankruptcy Institute (ABI) in the USA in 1982. This institute supplies the Congress and public opinion with reports describing cases of company bankruptcy. In these reports its causes and results are analyzed. This Institute employs attorneys, accountants, auctioneers, assignees, bankers, lenders and academic professors and conducts not only research activities concerning bankruptcy in the USA, but is also involved in many educational projects on bankruptcy. Moreover, the institute owns an extensive empirical dataset with insolvency data. It is also responsible for permanent cooperation with the media. The institute tries to forecast the future of companies. This is very important because of changes in the corporate environment and very complex intercorporate connections that may pose difficulties in formulating a long-term growth strategy. The Institute discusses past events, but the most important part of the ABI's activity is the investigation and forecast of the future performance of American companies.

Greiner and Schein (1988) stressed that the flexibility of a company is a reflection of the creativity of the owner. The improvements resulted from the problems which a company encounters in existence should lead to improvements. The management should recognize problems early on. Otherwise, the company will tend to an end. General uncertainty, a rapidly changing situation, unexpected problems and new goals have an impact on the immunity to change business life cycles, which as a result become shorter. Uncompetitive companies are eliminated from the market.

There are similar ideas in Handy (1995), who formulated what is known as the 'S' shape. According to this entrepreneurs should already be preparing the company for the crisis period even in the time of its prosperity. They should provoke artificial symptoms of crisis and undertake proper countermeasures. These actions can weaken the organization in the short term. However, the changes developed by the fake crisis can strengthen and immunize the company.

Frederick et al. (1988) stressed that company activities are always connected with a company's social function as its resources come from society. If the organization terminates the proper exercise of these functions by not using those resources in a way determined as adequate by society, it starts to tend towards its end. This situation is implied by the decline of the financial and human resources which are necessary for the proper functioning and existence of the company. The insolvency of a company is the result of the social necessity to extend the efficiency of economic processes.

The removal of unprofitable firms, the protection of debtees, employees and the whole country against dishonest debtors as well as the elimination from the market of companies which have lost their cost-effectiveness are main

advantages. As a result of this process a bad company is either reorganized or disposed of which is supposed to protect it and make possible a better usage of social resources.

In practice bankruptcy is mainly a tool of control for the protection of the market. Davydenko and Franks (2008) compared British, French, and German insolvency codes. They checked whether outside creditors can adjust low protection of creditors through changing their rules for granting loans. Kahl (2002) created a model explaining different options for creditors in a dynamic context. This problem became particularly important during the last banking and economic crisis which resulted in a credit crunch. It led to the bankruptcy of many small and middle sized companies.

That is why, in order to answer the question whether and when a company will default, a structured approach was needed. Over the years a wide range of credit risk and distress prediction models have been developed. This created a new economic subfield of corporate bankruptcy prediction. The most recent studies give a good overview of developments in the field of insolvency research. An introduction to the methods of corporate failure prediction are presented by Thiele and Lohmann (1995) or Baetge and Ströher (2005).

In Matschke (1979) previous research contributions on the problems of company classification by MDA in international studies are reviewed. In the seventies and eighties the number of insolvencies and bankruptcies rose. The use of early warning signals of insolvency and distress in the forecasts became very important. Numerous studies, e.g. Eisenbeis (1977), Moyer (1977), Altman et al. (1981), Zmijewski (1984), Platt and Platt (1991) concerned the methodological and statistical problems arising from the application of multiple discriminant models. Thus, new methods, e.g. logistic analysis were developed and became popular.

The last method uses a firm's financial ratios, weighs each of the ratios by its respective weight and aggregate products to a probability measure of insolvency for a company. Corporate failure predictions were constructed with the use of logistic analysis, e.g. Ohlson (1980), Zavgren (1985) and Peel and Peel (1988).

Ohlson (1980) used a logit of the maximum likelihood method to build and analyze a model which sampled 105 failed companies and 2058 non-failed companies from 1970 to 1976. He set up 3 models from 9 explanatory variables to predict corporate failure.

Keasey and Watson (1987) employed a logit to build a prediction model. They sampled 73 failed companies and 73 non-failed companies from 1970 to 1983, using 28 financial variables and 18 non-financial variables in their study.

Logit techniques are free of the restrictive assumptions typical of the MDA. Moreover, they allow an interpretation of individual coefficients. Other methods without restrictive statistical assumptions are non-parametric techniques. The best known of these is the neural networks method. This tries to improve computerized pattern recognition. It develops models based on the functioning of human brain.

The neural network attempts to implement learning behaviour into computing systems.

In spite of the application of new methods to bankruptcy forecasts, MDA is still the most frequently used and developed method in different insolvency problems. In a more recent study Hesselmann (1995) reported that a discriminant approach based on "z-score" model of Altman (1968) achieved a predictive accuracy of 70 percent. Thus, the MDA method became an integral part of the monitoring system in the banking sector. The drawback of the MDA model is, according to the author, the lack of all qualitative factors. Baetge and Ströher (2005), stressed that MDA is not a proper tool for forecasting an acute crisis of a company. They suggested the usage of neural networks combined with neuro-fuzzy systems for an evaluation of a firm's survival chances. However, MDA can help to detect the symptoms of problems which can raise the awareness of a coming crisis. MDA can also help to find the relevant measures against coming distress. The future development of a company can be frequently derived from historical data because certain patterns of the symptoms which can lead to insolvency are visible several years before the death of a company.

Some contributors criticize MDA because of a lack of an economic theory to support this method. Muche (2007) tried to relate his stochastic model of insolvency prediction with the extraction and loading of common financial ratios from financial statements. Agarwal and Taffler (2011) were concerned with accounting-ratio-based models. In their opinion the traditional models were no worse than market-based models for credit risk assessment. Moreover, they bettered them in terms of potential bank profitability, especially in the case of error misclassification costs. Diacogiannis (1996) expressed his doubts concerning taking into account security prices for corporate insolvency forecasts. He argued that macro-economic variables such as inflation could improve the predictive power of the known models.

Jacobson and Lindé (2000) checked whether macro-economic variables can improve the accuracy of prediction models. They worked on credit ratings produced by the two biggest Swedish credit rating agencies monitoring Swedish financial markets. The focus of the research was on the assessment of the stability of a financial system. The authors advised taking into account macroeconomic developments in order to determine risk in the banking system. In their opinion macroeconomic variables have large explanatory power in the explanation of the actual proportion of company bankruptcies. This is important because companies are responsible for a large part of a bank's credit losses. In their recent studies Jones and Hensher (2008) reviewed different modern methods used in the field of credit risk and corporate bankruptcy prediction. They included probit models, advanced logistic regression models, survival analysis models, non-parametric techniques, structural models and a model in reduced form. The contributors advised using several new methods, like the mathematical and theoretical systems known as "belief functions" and insolvency modelling for public sector entities.

The financial statements of companies are often the basis for statistical forecast models like univariate or MDA. The assessment of the likelihood of a firm's insolvency is frequently conducted by market-based econometric models which are based on security prices. This spectrum of models includes logistic regression models and non-parametric techniques, e.g. neural networks and recursive partitioning models. The results of insolvency research can significantly differ from one another. There are many approaches to predicting bankruptcy though the results depend to large extent on the method used.

Recently Gurgul and Zając (2011) introduced the bankruptcy prediction model which was used for empirical purposes. It did not refer to financial indicators but described the processes of destruction and creation of companies in a national economy.

In the next section we will apply different techniques aimed at forecasting the births and deaths of companies in different sectors of national economy.

3. Enterprise migration matrices

Define matrix **X** as follows:

$$\mathbf{X}^{(t)} = \left[x_{ij}^{(t)} \right] = \begin{bmatrix} \mathbf{M}^{(t)} & \mathbf{d}^{(t)} \\ \mathbf{b}^{(t)} & 0 \end{bmatrix}$$

Matrix **X** is formed using the following matrices:

$\mathbf{M}^{(t)} = \left[m_{ij}^{(t)} \right]$ - migration matrix representing the number of enterprises crossing between sectors i and j during a period from t-1 to t.

$\mathbf{d}^{(t)} = \left[d_{i}^{(t)} \right]$ - death vector presenting the number of dead enterprises in each sector; by a dead enterprise we understand an enterprise that existed in year t-1 and simultaneously did not exist in t.

$\mathbf{b}^{(t)} = \left[b_{j}^{(t)} \right]$ - a birth vector containing the number of new enterprises in each sector; by a new enterprise we understand an enterprise that did not exist in year t-1 and existed in t.

A zero in matrix **X** indicates that moving from 'birth' state to 'death' in the same year is not possible by this model.

In addition, by **r** and **c** we understand vectors holding the sums of rows and columns of matrix **X**:

$$\mathbf{r}^{(t)} = \left[r_{i}^{(t)} \right] = \left[\sum_{j} x_{ij}^{(t)} \right]$$

$$\mathbf{c}^{(t)} = \left[c_{j}^{(t)} \right] = \left[\sum_{i} x_{ij}^{(t)} \right]$$

The main properties of matrix $\mathbf{X}^{(t)}$ are:

- the sum of each row ($r_i^{(t)}$) describes the number of enterprises in corresponding sectors in year t-1,

- the sum of each column ($c_j^{(t)}$) describes the number of enterprises in sectors in year t,

- the diagonal values of matrix $\mathbf{X}^{(t)}$ represent the number of enterprises that survived year t and remained in the same sector,

- for each sector the sum of a column in year t becomes the sum of a row in year t+1 ($c_i^{(t-1)} = r_i^{(t)}$).

An example of the matrices defined above is presented in table 1. Table 1 holds data that describes migration, births and deaths in Belgian enterprises between the years 2000 and 2001. Belgian companies are divided into sectors such as agriculture, industry, construction trade and transport, financial activities, other services and activities. Those enterprises that do not belong to these sectors are gathered in the group named unknown.

Table 1. Matrix $\mathbf{X}^{(2001)}$ for Belgian companies in the year 2001

	Agriculture	Industry	Construction	Trade and Transport	Financial activities	Other services and activities	Unknown	Death	Total
Agriculture	26165	6	13	42	13	12	104	1527	27882
Industry	2	42731	48	148	50	29	183	2713	45904
Construction	18	60	65718	55	53	7	310	5035	71256
Trade and Transport	171	151	77	229481	280	104	1434	20839	252537
Financial activities	23	63	55	321	137713	100	536	10479	149290
Other services and activities	4	6	4	59	96	70519	226	4107	75021
Unknown	46	70	166	497	2310	314	12628	5299	21330
Birth	1689	2317	5288	16936	15298	6062	2469	0	50059
Total	28118	45404	71369	247539	155813	77147	17890	49999	

Source: Crossroads Bank for Enterprises.

   In the first row one can see that 27882 enterprises were active in "Agriculture" branch in 2000 (Total - sum of the first row). Between the year 2000 and 2001: 26165 of them remained in the "Agriculture" branch in 2001, 6 of them migrated to the "Industry" branch, 2 enterprises migrated to "Agriculture" from "Industry", 1527 enterprises from "Agriculture" died (Death – first row), which means they were no longer economically active in 2001, 1689 new enterprises in "Agriculture" were created (Birth – the first column). Finally, in 2001 the "Agriculture" branch had 28118 active enterprises.

   In our next step we will define a matrix $\mathbf{P}^{(t)}$ that holds the probability of company migration between sectors, "birth" and "death" states in year t. For each year a matrix will be defined as a fraction of enterprises migrating between possible states. Every element in matrix $\mathbf{X}$ would be divided by the sum of the corresponding row. In this way we obtain a probability matrix which is analogous to input matrix in the context of IO analysis:

$$\mathbf{P}^{(t)} = \left[p_{ij}^{(t)} \right] = \left[\frac{x_{ij}^{(t)}}{\sum_i x_{ij}^{(t)}} \right]$$

   Matrix $\mathbf{P}^{(2001)}$ (in %) calculated from matrix $\mathbf{X}^{(2001)}$ is presented in table 2.

Table 2. Matrix $\mathbf{P}^{(2001)}$ derived from matrix $\mathbf{X}^{(2001)}$

	Agriculture	Industry	Construction	Trade and Transport	Financial activities	Other services and activities	Unknown	Death	Total
Agriculture	93.8419	0.0215	0.0466	0.1506	0.0466	0.0430	0.3730	5.4767	27882
Industry	0.0044	93.0877	0.1046	0.3224	0.1089	0.0632	0.3987	5.9102	45904
Construction	0.0253	0.0842	92.2280	0.0772	0.0744	0.0098	0.4351	7.0661	71256
Trade and Transport	0.0677	0.0598	0.0305	90.8702	0.1109	0.0412	0.5678	8.2519	252537
Financial activities	0.0154	0.0422	0.0368	0.2150	92.2453	0.0670	0.3590	7.0192	149290
Other services and activities	0.0053	0.0080	0.0053	0.0786	0.1280	93.9990	0.3012	5.4745	75021
Unknown	0.2157	0.3282	0.7782	2.3301	10.8298	1.4721	59.2030	24.8429	21330
Birth	3.3740	4.6285	10.5635	33.8321	30.5599	12.1097	4.9322	0	50059
Total	28118	45404	71369	247539	155813	77147	17890	49999	

Source: Crossroads Bank for Enterprises, own calculations.

Coppens and Verduyn (2009) assumed that matrix $\mathbf{P}^{(t)}$ is constant over time. Based on this assumption they used a matrix $\overline{\mathbf{P}} = \dfrac{1}{6} \displaystyle\sum_{i=2001}^{2006} \mathbf{P}^{(i)}$ for all their calculations. But in fact this assumption is a simplification. We will show methods for the prediction of further values of matrix $\mathbf{X}$ based on input-output matrix updating procedures. In our studies we use the same data as Coppens and Verduyn (2009). Our research is based on data from the Crossroads Bank for Enterprises published by Belgostat about Belgian enterprise migration in industry sectors in the period between 2000-2006 (see: http://www.belgostat.be - select option: "Business demography").

4. Forecasts of migration matrices

The problem of forecasting migration matrices can be conducted by means of updating procedures well-known in I-O theory. In this paper, we use some of the most popular I-O updating algorithms (RAS and some non-biproportional approaches) and measure the ex-post error of prediction. Non-biproportional algorithms generally define a measure of distance between the elements of two matrices and minimize it. Potentially a measure of the distance between two matrices can be identified in many ways, but those based on absolute differences and/or squared differences are most popular among researchers (Jackson and Murray (2004)).

For simplicity of notation, let $p_{ij} \in \mathbf{P}$ be the last known probability coefficient matrix, and let $q_{ij} \in \mathbf{Q}$ be the prediction of the probability matrix for the next year. Similarly as in the previous section, $\mathbf{c}$ and $\mathbf{r}$ represent known vectors containing the sum of columns and the sum of rows in matrix $\mathbf{Q}$.

1. *Constant Value Matrix*

The method is based on research done by Coppens and Verduyn (2009).

$$\mathbf{P} = \overline{\mathbf{P}} = \frac{1}{6} \sum_{i=2001}^{2006} \mathbf{P}^{(i)}$$

2. *Absolute Differences*

The main goal of this method is to minimize the sum of absolute differences between the elements of the last known matrix and the prediction of the matrix for the following year. Unfortunately, some zeros may appear in the algorithmic solution.

$$\sum_{i} \sum_{j} \left| p_{ij} - q_{ij} \right| \to \min$$

$$s.t. \ \sum_{j} q_{ij} = 1 \ \text{for all } j$$

$$\sum_{i} q_{ij} r_i = c_j \ \text{for all } i$$

$$q_{ij} > 0 \ \text{for all } i, j$$

3. *Weighted Absolute Differences*

This method assigns weight to absolute differences by the size of coefficients. A large (small) coefficient implies a large (small) assigned weight. In the case of our data, the diagonal values of matrix **P** are more "important" than the rest of the matrix. The percentage of enterprises remaining in the same sector is more stable than migrating portion.

$$\sum_{i} \sum_{j} p_{ij} \left| p_{ij} - q_{ij} \right| \rightarrow \min$$

$$s.t. \ \sum_{j} q_{ij} = 1 \ \text{for all } j$$

$$\sum_{i} q_{ij} r_i = c_j \ \text{for all } i$$

$$q_{ij} > 0 \ \text{for all } i, j$$

4. *Normalized Absolute Differences (Matuszewski 1964)*

The fourth method also assigns weights to absolute differences, but in contrast to method 3 small (large) coefficients are assigned to large (small) weights. In this approach the diagonal elements of matrix **P** which are relatively high in level are significantly less "important" than the other coefficients of this matrix. The percentage of migrating enterprises is more stable than the ratio of enterprises remaining in the same sector.

$$\sum_{i} \sum_{j} \frac{\left| p_{ij} - q_{ij} \right|}{p_{ij}} \rightarrow \min$$

$$s.t. \ \sum_{j} q_{ij} = 1 \ \text{for all } j$$

$$\sum_{i} q_{ij} r_i = c_j \ \text{for all } i$$

$$q_{ij} > 0 \ \text{for all } i, j$$

5. *Squared Differences (Almon 1968)*

The fifth method was formulated by Almon (1968). It is worth noticing that in this method we minimize the Euclidean distance between matrices **P** and **Q**. This fifth method in fact minimizes the sum of squares, and because of that has the properties of ordinary least squares (Durieux and Payen, 1976).

$$\sum_i \sum_j \left(p_{ij} - q_{ij}\right)^2 \rightarrow \min$$

$$s.t. \ \sum_j q_{ij} = 1 \text{ for all } j$$

$$\sum_i q_{ij} r_i = c_j \text{ for all } i$$

$$q_{ij} > 0 \text{ for all } i, j$$

6. *Weighted Squared Differences*

The sixth method assigns weights to squared differences. The weights are analogous to those applied in the third method (small coefficient implies small weight).

$$\sum_i \sum_j p_{ij} \left(p_{ij} - q_{ij}\right)^2 \rightarrow \min$$

$$s.t. \ \sum_j q_{ij} = 1 \text{ for all } j$$

$$\sum_i q_{ij} r_i = c_j \text{ for all } i$$

$$q_{ij} > 0 \text{ for all } i, j$$

7. *Normalized Squared Differences (Friedlander 1961)*

The seventh method was formulated by Friedlander. This method is a normalized version of Almon's method and has weights similar to those of the fourth method.

$$\sum_i \sum_j \frac{\left(p_{ij} - q_{ij}\right)^2}{p_{ij}} \rightarrow \min$$

$$s.t. \ \sum_j q_{ij} = 1 \ \text{ for all } j$$

$$\sum_i q_{ij} r_i = c_j \ \text{ for all } i$$

$$q_{ij} > 0 \ \text{ for all } i, j$$

The main difficulty connected to last three methods is resolving the non-linear optimization problem. Solutions can often be local rather than global. In our simulations we used the Simplex method and a Generalized Reduced Gradient (GRG2) algorithm. As a starting point actual numbers from target matrices were taken.

8. *RAS method*

Last but not least is the biproportional method RAS (Bacharach,1970). The iterative proportional fitting procedure IPFP, known as biproportional fitting in statistics, the RAS algorithm in economics and matrix scaling in computer science, is an iterative algorithm for estimating cell values of a contingency table such that the marginal totals remain fixed and the estimated table decomposes into an outer product.

The RAS method was developed by Stone (1961). It is based on a biproportional technique which is an iterative procedure that allows the computation (on condition that the marginal totals remain fixed) of non-negative matrix elements. At the beginning the estimated matrix is equal to the prior matrix. In further stages the rows and columns of the estimated matrix are scaled alternately in order to achieve the desired properties. The scaling is conducted by the multiplication of each element in a row (column) by the proportion of the desired and actual sum of that row (column). In the case of our data about 200 iterations were performed for single migration matrix.

5. Empirical results

In order to decide which method gives better predictions we used four matrix comparison methods. We used Theil's U statistic (Theil 1971), the weighted absolute difference (WAD, Lahr, 2004), index C (Roy et al., 1982) and finally standardized total percentage error (STPE) for multiplier assessment (Miller and Blair, 1985). In fact these comparison methods are measures of the ex-post error of prediction. In order to present the formula of the ex-post statistics, we used elements $a_{ij} \in \mathbf{A}$, where $\mathbf{A}$ denotes the actual matrix from a certain year.

Numbers $q_{ij} \in \mathbf{Q}$ are the forecasts of the elements of an actual matrix $\mathbf{A}$. We applied the following measures of ex-post errors:

$$U = \sqrt{\frac{\sum\sum (a_{ij} - q_{ij})^2}{\sum\sum a_{ij}^2}}$$

$$WAD = \frac{\sum\sum a_{ij}|a_{ij} - q_{ij}|}{\sum\sum (a_{ij} + q_{ij})}$$

$$C = \frac{\left|\sum\sum q_{ij}\log q_{ij} - \sum\sum a_{ij}\log a_{ij}\right|}{\sum\sum a_{ij}\log a_{ij}}$$

$$STPE = 100\frac{\sum\sum |a_{ij} - q_{ij}|}{\sum\sum a_{ij}}$$

We calculated these statistics for all the methods used. Migration matrices from years 2000-2006 were used. In table 3 we present the best methods of predicting the following year's migration according to ex-post error statistics for the matrices.

Table 3. Best methods for predicting the following year's migration according to ex-post error statistics

	2001/2002	2002/2003	2003/2004	2004/2005	2005/2006
STPE	M8	M2	M8	M1	M8
Theil's U	M4	M2	M8	M1	M8
WAD	M8	M2	M7	M1	M8
C	M8	M2	M4	M1	M1

Source: Own calculations.

For each method of comparison the mean value of statistics based on the matrices were calculated. In order to complete the ranking of the methods applied, we assessed them according to all four ex-post error statistics. The results are presented in table 4.

Table 4. Results of matrix updating. Average index values for years 2000-2006

	STPE	rank	Theil's U	rank	WAD	rank	C	rank	Average Rank	Combined Rank
M1	1.7876	4	0.0427	2	0.0052	3	0.0227	1	2.5	2
M2	1.7822	3	0.041	1	0.0049	1	0.0271	2	1.75	1
M3	2.1837	7	0.0431	3	0.0052	2	0.0289	3	3.75	3
M4	1.8037	5	0.0544	8	0.0063	8	0.0353	7	7	8
M5	2.0286	6	0.0542	7	0.006	6	0.0358	8	6.75	7
M6	2.2417	8	0.054	6	0.006	7	0.0344	5	6.5	6
M7	1.7684	2	0.0531	5	0.0058	5	0.0335	4	4	5
M8	1.6440	1	0.0518	4	0.0056	4	0.0347	6	3.75	3

Source: Own calculations.

Based on these results, we see that all predictions have comparable statistical values. That is why on the basis of these rankings we recommend the use of minimizing absolute differences and the RAS method.

6. Conclusions

As a first conclusion we must draw attention to the size of prediction errors made by using these methods. The standard prediction error (STPE) for all the years and methods is between 0.39% and 3.8%, which means that effectiveness of this methodology is very high.

According to the ranking based on the mean error statistics the second method is the best one for predicting migration matrices. The second method is a non-biproportional algorithm minimizing absolute differences. The unexpectedly good ex-post forecasts were achieved by the average matrix (i.e. mean value of matrices from the period under study). One of the reasons for that may be that the target matrix is also included in the average matrix calculation (as one of six). It is worth noting that although the RAS method was ranked as third it was most often the best method, and exhibited the lowest values of STPE statistics. The latter observation means that the probability of wrong classification of an enterprise by this method is lower in comparison to the others.

The comparison of methods that favour large diagonal values of a matrix and methods which attach large weights to relatively small matrix elements is also interesting. Our investigation shows that generally the second group of algorithms gives better results. This means that the migration process which applies to a small percentage of enterprises is stable in time. Irregularity connected to changes in the numbers of enterprises involves mostly companies that stick to a certain branch of industry.

REFERENCES

AGARWAL V., TAFFLER R. (2011). Comparing the performance of market-based and accounting based bankruptcy prediction models, Journal of Banking and Finance, 32, pp. 1541-1551.

ALMON C. (1968). Recent methodological advances in input-output in the United States and Canada, Fourth International Conference on Input-Output Techniques, Geneva.

ALTMAN E. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, Journal of Finance, 23, pp. 589-609.

ALTMAN E., AVERY R., EISENBEIS R., SINKEY J. (1981). Application of classification techniques in business, banking, and finance, Greenwich, Conn, JAI Press.

BAETGE J., STROEHER T. (2005). Empirische Insolvenzforschung zur Beurteilung der Bestandsfestigkeit von Unternehmen, in: Burmann, Chr. (Eds.), Management von Ad-hoc-Krisen: Grundlagen, Strategien und Erfolgsfaktoren, Gabler, Wiesbaden, pp. 151-167.

BEAVER W.H. (1966). Financial ratios as predictors of failure, in: Empirical research in accounting: Selected studies, Supplement for Journal of Accounting Research, 4, pp. 71-111.

BACHARACH M. (1970). Biproportional Matrices and Input-Output Change, Cambridge, U.K., Cambridge University Press.

BLUM M. (1974). Failing Company Discriminant Analysis, Journal of Accounting Research, Spring, pp. 1 -25.

COPPENS F., VERDUYN F. (2009). Analysis of business demography using Markov chains: an application to Belgian data, Working Paper Research 170, National Bank of Belgium.

DAVYDENKO S., FRANKS J. (2008). Do bankruptcy codes matter? A study of defaults in France, Germany, and the U.K., Journal of Finance, 63, pp. 565-608.

DEAKIN E.B. (1972). A Discriminant Analysis of Predictors of Business Failure, Journal of Accounting Research, Spring, pp. 167- 179.

DIACOGIANNIS G.P. (1996). The usefulness of share prices and inflation for corporate failure prediction, University of Piraeus Journal of Economics Business, Statistics and Operations Research, 46, pp. 135-156.

DIAMOND U.S. (1976). Pattern Recognition and the Detection of Corporate Failure, Ph.D. dissertation, New York University.

EISENBEIS R.A. (1977). Pitfalls in the application of discriminant analysis in business, finance, and economics, Journal of Finance, 32, pp. 875-900.

FREDERICK W.C., DAVIS K., POST J.E. (1988). Corporate Social Responsibility and Business Ethics, McGraw-Hill Publishing Company, New York, pp. 28-29.

FOSTER R., KAPLAN S. (2001). Creative Destruction, Financial Times.

GURGUL H., ZAJAC P. (2011). The dynamic model of birth and death of enterprises, Statistics in Transition, 12/ 2, pp. 381–400.

GREINER L.E., SCHEIN V.E. (1988). Power and Organization Development, Addison-Wesley.

HANDY CH. (1995). The Age of Paradox, Harvard Business School Press.

HESSELMANN S. (1995). Insolvenzprognose mit Hilfequalitativer Faktoren, Shaker Verlag, Aachen.

JACKSON R.W., MURRAY A.T. (2004). Alternative input–output matrix updating formulations, Economic Systems Research, 16, pp. 135–14.

JACOBSON T., LINDE J. (2000). Credit rating and the business cycle: can bankruptcies be forecast?, Sveriges Riksbank economic review, 4, pp. 11-33.

JONES S., HENSHER D.A. (2008). Advances in credit risk modeling and corporate bankruptcy Prediction, University Press, Cambridge, UK.

KAHL M. (2002). Economic distress, financial distress and dynamic liquidation, Journal of Finance, 57, pp. 135-168.

KEASEY K., WATSON R. (1987). Non-financial symptoms and the prediction of small company failure: a test or Argenti's hypotheses, Journal of Business Finance and Accounting, 14 (3), 335-354.

LAHR M., de MESNARD L. (2004). Biproportional Techniques in Input-Output Analysis: Table Updating and Structural Analysis, Economic Systems Research, 16 (2), 115-34.

MATSCHKE M.J. (1979). Insolvenzprognose aus vergangenheitsorientierten Jahresabschlüssen als Basis von Kreditentscheidungen, Betriebswirtschaftlich eForschung und Praxis, 31, pp. 485-504.

MILLER R.E., BLAIR P.D. (1985). Input-output analysis: Foundations and extensions, Prentice-Hall, Englewood Cliffs, N.J.

MOYER R.C. (1977). Forecasting financial failure: a re-examination, Financial Management, 6, pp.11-17.

MUCHE T. (2007). Ein stochastisches Modell zur Insolvenzprognose auf der Basis von Jahresabschlußdaten, Betriebswirtschaftliche Forschung und Praxis, 59, pp. 376-399.

OHLSON J.A. (1980). Financial Ratios and Probabilistic Prediction of Bankruptcy, Journal of Accounting Research, Spring, pp. 109-131.

PEEL M.J., PEEL D.A. (1988). A Multilogit Approach to Predicting Corporate Failure: Some Evidence for the UK Corporate Sector, Omega, pp. 309-318.

PLATT H., PLATT M. (1991). A note on the use of industry-relative ratios in bankruptcy prediction, Journal of Banking and Finance, 15, pp. 1183-1194.

ROY J., BATTEN D., LESSE P. (1982). Minimizing information loss in simple aggregation, Environment and Planning, 14, 973-980.

SHUMPETER J. (1982). The Theory of Economic Development: An inquiry into profits, capital, credit, interest and the business cycle, Transaction Publisher.

STONE R.A. (1961). Input-Output Accounts and National Accounts. Paris, Organization for European Economic Cooperation.

THEIL H. (1971). Applied Economic Forecasting. Amsterdam, North-Holland.

THIELE O., LOHMANN K. (1995). Möglichkeiten und Grenzen der Insolvenzprognose auf der Grundlage von Jahresabschluss informationen, Freiberger Arbeitspapiere, 95/17.

WILCOX J.W. (1971), A Simple Theory of Financial Ratios as Predictors of Failure, Journal of Accounting Research, Autumn, pp. 389-395.

ZAVGREN C.V. (1985). Assessing the Vulnerability to the Failure of American Industrial Firms: A Logistic Analysis, Journal of Business Finance and Accounting, pp. 19-45.

ZMIJEWSKI M. (1984). Methodological issues related to the estimation of financial distress prediction models, Journal of Accounting Research (Supplement), pp. 59-82.