Statistical analysis and econometric modelling of the creditworthiness of non-financial companies

Vladimir I. Malugin*

Faculty of Applied Mathematics and Computer Science,
Department of Mathematical Modelling and Data Analysis,
BelarusianState University,
4, Nezavisimosti avenue,
220030, Minsk, Republic of Belarus
E-mail: malugin@bsu.by
*Corresponding author

Natalia V. Hryn

Department of Mathematical
and Information Support of Economic Systems,
Yanka Kupala State University of Grodno,
22, Ozhesko Street, 230023,
Grodno, Republic of Belarus
E-mail: lebnat@tut.by

Aleksandr Y. Novopoltsev

Faculty of Applied Mathematics and Computer Science,
Department of Mathematical Modelling and Data Analysis,
BelarusianState University,
4, Nezavisimosti avenue,
220030, Minsk, Republic of Belarus
E-mail: fpm.novopolc@bsu.by

Abstract: This paper describes the results of the application of multivariate statistical analysis and econometric modelling to assess the creditworthiness of non-financial companies on the micro and macro levels. On the basis of company's financial reports data we propose a system of credit measures called 'relative statistical credit ratings' (RSCR), which includes: company ratings (CCR), the branch of the economy ratings (BCR) and the integral indicator of creditworthiness of the national economy (ICI). The proposed methodology is applied to evaluate the creditworthiness of Belarusian companies. Using econometric modelling we examine the dependence of the credit measures BCR and ICI on the major macroeconomic factors of the Belarusian economy. We establish also the relations between the integral output indicators of the national economy and the proposed statistical creditworthiness measures. Economic analysis of the obtained statistical and econometric modelling results indicates the informativeness and the economic significance of the proposed indicators.

Keywords: computational economic; creditworthiness of non-financial companies on the micro and macro levels; relative statistical credit ratings; company and branch credit ratings; integral indicator of creditworthiness; econometric modelling; cluster and discriminant analysis; ordered logit model on panel data; Markov dependence of the ratings; creditworthiness of Belarusian companies.

Reference to this paper should be made as follows: Malugin, V.I., Hryn, N.V. and Novopoltsev, A.Y. (2014) 'Statistical analysis and econometric modelling of the creditworthiness of non-financial companies', *Int. J. Computational Economics and Econometrics*, Vol. 4, Nos. 1/2, pp.130–147.

Biographical notes: Vladimir I. Malugin is an Associate Professor. He received Candidate of Science Degree in "Theory of Probability and Mathematical Statistics" from Vilnius University. His current research interests are statistical analysis and classification of a complex multivariate data as well as methods of analysis of multivariate econometric models with heterogeneous structure. He is the author (or co-author) of such books (in Russian) as "The Fundamentals of Simulation and Statistical modelling", "Mathematical and Computer Fundamentals of Statistical Modelling and Data Analysis", "Econometric Modelling", "Securities Market: Quantitative Methods of Analysis", etc. He has publications in *Automation and Remote Control*, *Applied Econometrics*, *Lithuania Statistical Journal*, etc.

Natalia V. Hryn is a Lecturer at the Department of Mathematical and Information Support of Economic Systems at Yanka Kupala State University of Grodno. She received her Bachelor degree in Mathematics and Economy in 2006 and her Master degree in Mathematics in 2013. Her fields of interest include credit scoring, credit risk analysis and applied econometrics.

Aleksandr Y. Novopoltsev is the last year student of the Faculty of Applied Mathematics and Computer Science of the Belarusian State University (specialty – 'Economics Cybernetics'). His research interests are focused on development of the algorithms and software for classification of multidimensional (time series and panel) data under the Markov dependency of class numbers.

This paper is a revised and expanded version of a paper entitled: 'Statistical company's credit ratings and their econometric analysis' presented at the *Proc. of the 10th Intern. Conf. 'Computer Data Analysis and Modelling: Theoretical & Applied Stochastics'*, Belarusian State University, Minsk, Belarus, 10–14 September, 2013.

1 Introduction

The problem of the creditworthiness (solvency) assessment of non-financial companies is one of the most important issues in bank risk management (Jorion, 2009). The Basel Agreements (Basel Committee on Banking Supervision, 2003) indicate the need to use internal bank models based on probabilistic and statistical methods in order to assess the creditworthiness of both non-financial companies and other bank borrowers. Thus, this problem is vital not only for commercial banks but also for the government regulators, who monitor the companies on a regular basis (Milevskiy and Zubovich, 2012). In particular, Central Bank is certainly interested in internal creditworthiness evaluation methodology at the micro-level that takes into account the financial condition of companies and various aspects of their production and business activities. However, such regulators are even more interested in analysis of the state of economy, possible predictive options for its development as well as in working out corrective tools to ensure monetary and economic stability. Therefore, the following problems of creditworthiness analysis on the macro-level are relevant: the classification of companies according to the degree of creditworthiness; dynamic analysis of the average levels of creditworthiness in the branches of economy or in the economy on the whole; analysis of migration between different classes of creditworthiness in different periods of time, etc. The results of such analysis may be used to evaluate the stability of the economy as well as to assess the system risks of the banking sector.

The aim of this research is to develop a methodology for solving the above-mentioned problems based on statistical analysis of the data obtained by means of companies monitoring. The main peculiarities that show the originality and significance of the presented methodology in content of emerging economics are as follows:

- the need for assessment of the creditworthiness for both micro-level (individual companies) and macro-level (branches of the economy and national economy on the whole)
- the majority of national companies do not have the ratings established by financial agencies
- there is no representative statistics on bankruptcies of the companies
- there is quite a small representative database of financial reports data for national companies
- limited capabilities of the official acting procedure for evaluating the creditworthiness of companies.

Due to these features, the statistics on bankruptcies and credit history of companies are not used within the frame of the developed methodology. The proposed system of the credit measures uses only company's financial reports data. As a result it does not take into account the probability of default. At the same time, the obtained credit measures as well as the classes of creditworthiness (ratings of company) have a substantial economic interpretation and may be considered as certain 'relative statistical company's credit ratings' based on the available information.

The proposed methodology is examined on the basis of the quarterly financial reports data of Belarusian companies, available from 2008 onwards. Official acting evaluation procedure assumes only two classes of creditworthiness and two financial coefficients with expert threshold values and obviously has limited possibilities for solving the above-mentioned problems.

We submit the following main results:

- The methodology of construction and verification of the relative creditworthiness measures, namely, *relative statistical company's credit ratings* called so because of the above-mentioned reasons. It allows to classify companies into the given number (more than two) of classes of creditworthiness.

- The system of annual and quarterly *relative statistical credit ratings* (RSCR), including company's credit ratings (CCR), average ratings for the branches of the economy (BCR) and integral creditworthiness indicator for the national economy on the whole (ICI).
- The results of verification of the BCR and ICI creditworthiness measures based on econometric models, which show their interdependence on the main economic factors and output indicators of the national economy.

Because of the above-mentioned specific conditions, we use cross-sectional representation of the data with a panel structure. It allows solving the assigned problems using traditional methods of multivariate statistical data analysis, including preliminary statistical analysis and principal component analysis, and cluster and discriminant analysis. In the research, we consider four major classes of creditworthiness that are statistically and economically justified. Since Belarusian companies do not have credit ratings as a rule, we use cluster analysis algorithms for cross-sectional representation of the data to get their preliminary classification.

Thereafter, the classified sample is used as a learning sample to construct the 'plug-in' Bayesian decision rules of discriminant analysis with independent and Markov-dependent class number interpreted as CCR (Kharin, 1996). For this purpose, we also use classification algorithm based on the logit ordered model for panel data with fixed individual effects. Based on the obtained CCR, we calculate the average creditworthiness measures for both the branches of the economy (BCR) and the national economy on the whole (ICI). These aggregative indicators allow evaluating the creditworthiness on the macro-level by means of econometric modelling and forecasting.

2 Mathematical model and used methodology

2.1 Mathematical model of the data

We observe financial characteristics of n companies belonging to one of K branches of the economy $G_1,\dots,G_K$ in the space $\Re^N$ at the moment t $(t=1,\dots,T)$. We use the same set of companies during the entire observation period. For the given moment t, a company (i, k) is characterised by N-vector of observations:

$$x_{i,t}^{(k)} = (x_{i,jt}^{(k)}) \in \Re^N \quad (i \in S_k, k=1,\dots,K,\ t=1,\dots,T), \tag{1}$$

where $x_{i,jt}^{(k)}$ $(j=1,\dots,N)$ is the value of jth financial coefficient characterising the creditworthiness of the company (i, k) at the moment t; S_k is the set of identification codes for companies from the branch $G_k, n_k = |S_k|$ is the number of companies in the branch G_k.

We assume that the degree of creditworthiness of the company (i, k) at the moment t can be described by the discrete random variable $d_{i,t}^{(k)} \in S(L) = \{1,\dots,L\}$, where L is the number of classes of creditworthiness. Thereby all companies in any time can belong to one from L classes of $\Omega_1,\dots,\Omega_L$, where class Ω_1 corresponds to the highest level and class Ω_L to the lowest level of creditworthiness. Creditworthiness measures $\{d_{i,t}^{(k)}\}$ can be interpreted as a set of *relative credit ratings* of the companies.

Thus, the full information about the company (i, k) at the moment t is defined by a staked vector

$$z_{i,t}^{(k)} = \begin{pmatrix} x_{i,t}^{(k)} \\ d_{i,t}^{(k)} \end{pmatrix} \in \Re^N \times S(L) \quad (i \in S_k, t = 1, \ldots, T). \tag{2}$$

We suppose that for the given k and $\forall i \in S_k$ $(k = 1, \ldots, K)$, credit rating $d_{i,t}^{(k)}$ is described by a *hidden homogeneous Markov chain* (HMC) with parameters:

$$\pi^{(k)} = (\pi_l^{(k)}),\ \pi_l^{(k)} = \mathbf{P}\{d_{i,1}^{(k)} = l\} > 0,\ l \in S(L), \tag{3}$$

$$P^{(k)} = (p_{rs}^{(k)}),\ p_{rs}^{(k)} = \mathbf{P}\{d_{i,t+1}^{(k)} = s \mid d_{i,t}^{(k)} = r\} \geq 0, \quad l, r, s \in S(L), t = 1, \ldots, T, \tag{4}$$

where $\pi^{(k)}$ is a vector of initial probabilities, $P^{(k)}$ is a matrix of transition probabilities (*rating migration matrix*) for some given G_k.

The set $\{d_{i,t}^{(k)}\}$ can be represented as a classification matrix of the size $n_k \times T$ for the sample $\{x_{i,t}^{(k)}\}$ $(k = 1, \ldots, K)$:

$$D^{(k)} = (d_{i,t}^{(k)}) = \begin{pmatrix} d_{1,1}^{(k)} & \cdots & d_{1,T}^{(k)} \\ \vdots & \ddots & \vdots \\ d_{n_k,1}^{(k)} & \cdots & d_{n_k,T}^{(k)} \end{pmatrix}. \tag{5}$$

Matrix $D^{(k)}$ allows two types of representations on rows and on columns, which have the substantial economic interpretation. The rows-representation

$$D^{(k)} = \begin{pmatrix} d_1^{(k)\prime} \\ \vdots \\ d_{n_k}^{(k)\prime} \end{pmatrix} \quad (k = 1, \ldots, K) \tag{6}$$

allows to get a vector of relative credit ratings for a company (CCR) $i \in S_k$ during the entire observation period

$$d_i^{(k)} = (d_{i,1}^{(k)}, \ldots, d_{i,T}^{(k)})' \in S^T(L) \quad (i \in S_k).$$

The columns-representation can be written as follows:

$$D^{(k)} = (c_1^{(k)}, \ldots, c_T^{(k)}), \tag{7}$$

where vector

$$c_t^{(k)} = (c_{1,t}^{(k)}, \ldots, c_{n_k,t}^{(k)})' \in S^{n_k}(L)$$

provides the classification of companies for branch k (BCR) at the moment t $(t = 1, \ldots, T)$.

The problem and used approach. The aim of the research is to develop the algorithms that allow to solve the following two problems:

- Estimation of the matrices $D^{(k)}$ $(k = 1, \ldots, K)$ based on the sample data $\{x_{i,t}^{(k)}\} \in \Re^N, i \in S_k, t = 1, \ldots, T.$

- Prediction of the credit ratings for the out-of-sample companies $(i \in S'_k, S_k \subseteq S'_k)$ or moments of observations $(t > T)$.

Thus, in general case, we need to define the following mapping:

$$d_{i,t}^{(k)} \equiv d(x_{i,t}^{(k)}) : \Re^N \to S(L) \quad (k = 1, \ldots, K). \tag{8}$$

To solve the first and second problems, we use the algorithms of cluster and discriminant analysis for each branch of the economy G_k $(k = 1, \ldots, K)$, respectively. For the economic interpretation of obtained estimates of the relative statistical ratings $\{d_{i,t}^{(k)}\}$, we apply methods of economic analysis.

We use panel and cross-sectional representations of the data $\{x_{i,t}^{(k)}\}$ that are based on the following assumptions. We assume that the distributions of random vectors $x_{i,t}^{(k)} \in \Re^N$ $(k = 1, \ldots, K)$ described by conditional density $f_l^{(t,k)}(u)$ $(u \in \Re^N)$ provide that company i $(i \in S_k, k = 1, \ldots, K)$ at the moment $t(t = 1, \ldots, T)$ belongs to the class of creditworthiness $\Omega_l (l \in S(L))$. We also suppose that

$$f_l^{(t,k)}(u) \equiv f_l^{(k)}(u) \quad \forall t = 1, \ldots, T. \tag{9}$$

The above described probability model of observations allows to consider the observations as *panel data*. The alternative approach is to consider them as *cross-sectional data*. Transition from *panel data* $X_t^{(k)} = \{x_{i,t}^{(k)}\}$ $(i = 1, \ldots, n, t = 1, \ldots, T)$ to *cross-sectional data* for every kind of objects k is realised by *renumbering of observations*:

$$Y^{(k)} = \{y_j^{(k)}\}(j = 1, \ldots, m_k),\ m_k = n_k T, \tag{10}$$

where $y_j^{(k)} \equiv x_{i,t}^{(k)}$ and

$$\forall t = 1, \ldots, T,\ i = 1, \ldots, n_k : j = (t-1)n_k + i. \tag{11}$$

Initial data representation $X_t^{(k)} = \{x_{i,t}^{(k)}\}$ can be reconstructed in one-valued way from the sample $Y^{(k)} = \{y_j^{(k)}\}$ by using the following inverse transformation

$$x_{i,t}^{(k)} \equiv y_j^{(k)}, \tag{12}$$

where

$$\forall j = 1, \ldots, m_k : (t-1)n_k < i \le t n_k, \tag{13}$$

and for the known values t, j we use the following relation for i

$$i = j - (t-1)n_k. \tag{14}$$

As a result of the classification of the samples $Y^{(k)} (k = 1, \ldots, K)$ by applying statistical algorithms, the classification vector

$$g^{(k)} = (g_1, \ldots, g_{m_k})', g_j^{(k)} \equiv d(y_j^{(k)}) \in S(L) = \{1, \ldots, L\}, \quad j = 1, \ldots, m_k$$

is derived. Further taking into consideration (12)–(14), we carry out the transition from classifications vector $g^{(k)} = (g_1, \ldots, g_{m_k})'$ to a required matrix of classifications $D^{(k)} = (d_{i,t}^{(k)})$ $(k = 1, \ldots, K)$, which considers dynamic character of initial data. In this case, it is considered that

$$d_{i,t}^{(k)} \equiv g_j^{(k)}, \tag{15}$$

where for any given value $j = 1,\dots,m_k$ index i satisfies the condition (13) and for the known values t, j is calculated with the use of formula (14).

2.2 System of RSCR

If classification matrices $\{D^{(k)}\}(k = 1,\dots,K)$ are known, we can calculate different types of creditworthiness indicators for both micro- and macro-levels study:

- Mean rating of company $i \in S_k$ (mean CCR rating) for the entire period of observation

$$\bar{d}_i^{(k)} = \frac{1}{T}\sum_{t=1}^{T} d_{i,t}^{(k)} \quad (1 \le \bar{d}_i^{(k)} \le L) \tag{16}$$

- Mean rating of branch G_k at the moment t $(t = 1,\dots,T)$, which reflects the dynamic of BCR rating for different branches $G_1,\dots,G_K$

$$BCR_t^{(k)} = \frac{1}{n_k}\sum_{i \in S_k} d_{i,t}^{(k)} \quad (1 \le BCR_t^{(k)} \le L) \tag{17}$$

- Statistical integral creditworthiness indicator of the economy as the average value on shares of a value added of branch in GDP.

$$\mathrm{ICI}_t = \sum_{k=1}^{K} \alpha_t^{(k)} \mathrm{BCR}_t^{(k)}. \tag{18}$$

- Weighted coefficients $\alpha_t^{(k)}$ characterises shares of contribution of corresponding branches into integral creditworthiness indicator at the moment t, namely:

$$\alpha_t^{(k)} = \frac{\Delta_t^{(k)}}{\Delta_t},\ \Delta_t = \sum_{k=1}^{K} \Delta_t^{(k)},\ \sum_{k=1}^{K} \alpha_t^{(k)} = 1, \tag{19}$$

  where $\Delta_t^{(k)}$ is share of a value added of kth branch in GDP at the moment t.

For estimating of the parameters (3) for the HMC, we use maximum likelihood estimates (Bhar and Hamori, 2004):

$$\hat{\pi}_l^{(k)} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\hat{d}_{i,1}^{(k)},l},\ \hat{p}_{rs}^{(k)} = \frac{\sum_{i \in S_k}\sum_{t=1}^{T-1} \delta_{\hat{d}_{i,t}^{(k)},r}\delta_{\hat{d}_{i,t+1}^{(k)},s}}{\sum_{i \in S_k}\sum_{t=1}^{T} \delta_{\hat{d}_{i,t}^{(k)},r}}, \tag{20}$$

where estimates $\widehat{D}^{(k)} = (\widehat{d}_{i,t}^{(k)})(k = 1,\dots,K)$ are obtained by means of cluster analysis procedure, $\delta_{s,r}$ is the Kronecker delta, $l, r, s \in S(L)$.

2.3 Development and economic verification of the RSCR

We use quarterly database of financial reports of companies belonging to four major branches of the Belarusian economy for the period 2008–2011: industry, construction, transport and trade. The dimension of the initial data is described by the following characteristics: $N = 43, n \approx 1000, T = 24, K = 4.$ The choice of statistical methods and algorithms is determined by the features of the used data, including:

- the lack of classified training sample
- the presence of outliers and missing observations
- a given set of available financial coefficients
- time-varying data.

The proposed methodology consists of the following main stages.

- *Preliminary stage*. It includes statistical analysis of the sample and formation of the set of classification features. At this stage, we perform the following main steps: processing of outliers and missing observations; normalisation and censoring of financial coefficients to the interval [0, 1]; choice of classification features (financial coefficients) on the base of correlation and principal component analysis as well as economic analysis and interpretations. As a result, the sample size is reduced to $n \approx 750$ companies and the number of features is reduced to $N = 14$ financial coefficients (see Table 1) according to national system of financial reporting (Malugin et al., 2009).
- *Cluster analysis stage*. At this stage, the first problem from Section 2.1 is being solved. To classify the sample obtained at the first stage, we use L-means algorithm for cross-sectional presentation of the sample in reduced normalised features space (Kharin, 1996). The main result of this stage is the estimates of classification matrices $\widehat{D}^{(k)} = (\widehat{d}_{i,t}^{(k)})(k = 1,\dots,K).$

  Thereafter, using the estimates $\{\widehat{D}^{(k)}\},$ we calculate relative ratings CCR, BCR, ICI as well as the estimate $\widehat{P}^{(k)} = (\widehat{p}_{rs}^{(k)})$ of the rating migration matrix by formulas (16)–(18) and (20).
- *Discriminant analysis stage* is connected with the solving of the second problem from Section 2.1. Discriminant analysis is used to classify out-of-sample companies, e.g. companies excluded at the first stage; new companies; initial sample companies at the new moments.

The following discriminant analysis algorithms are available:

- bayesian quadratic ‘plug-in-rule’ that takes into account Markov-dependent ratings (DA1)
- bayesian quadratic ‘plug-in-rule’ that does not take into account Markov-dependent ratings (DA2)
- classification algorithm based on the logit ordered model on panel data with fixed individual effects (DA3) (Mátyás and Sevestre, 2008).

We can apply a special modification of the above-mentioned algorithms that allow to extend the number of classes to 7: major classes 1, 2, 3, 4 and additional intermediate classes 1.2, 2.3, 3.4. For the economic analysis and interpretation of the obtained classifications, we apply a formalised financial statements analysing procedure.

Table 1 Used financial coefficients

Name	*Economic interpretation*
K1	Current ratio
K2	Cash ratio
K4	Equity in working assets ratio
K5	Liabilities to total assets ratio
K6	Overdue accounts receivable to total accounts receivable ratio
K7	Overdue accounts payable to total accounts payable ratio
K9	Debt-to-equity ratio
K10	Current assets to noncurrent assets ratio
K11	Earnings growth ratio
K13	Assets turnover ratio
K14	Payable turnover ratio
K15	Receivables turnover ratio
K18	Return on sales
K19	Return on assets

3 The empirical results for the statistical evaluation of the creditworthiness of companies

3.1 Cluster analysis results and its economic interpretation

For visualisation of statistical and economic inference, it is convenient to use a various graphic representations of the obtained results. Figure 1 illustrates the results of the cluster analysis of the quarterly data for the industry companies in the period from 2008 to 2011 in the space of canonical discriminant functions, including classes (clusters) and their centres (Huberty and Olejnik, 2006).

Based on Figure 1, we can say about a quite good separability of the classes Ω_1 and Ω_4 that correspond to the companies with the highest and lowest rates of creditworthiness. Classes Ω_2 and Ω_3 occupy the intermediate position. Table 2 contains the coordinates of the centres for the correspondent clusters in the space of 14 initial (non-normalised) financial coefficients. According to the used economic procedure of the analysis, economic interpretations of these centres are given in Table 3. For this purpose, all considered financial coefficients are grouped according to five main directions of companies' economic activity:

- liquidity (coefficients K1 and K2)
- financial soundness (K4, K5, K6, K7, K9 and K10)
- dynamics of development (K11)
- business activity (K13, K14 and K15)
- financial efficiency and profitability (K18 and K19).

The development dynamics ratio (K11) is not included in Table 3 since it is not informative in the considered period, which is characterised by essential two-stage devaluation of Belarusian ruble (in 2009 and 2011).

Figure 1 The classified observations in the space of canonical discriminant functions

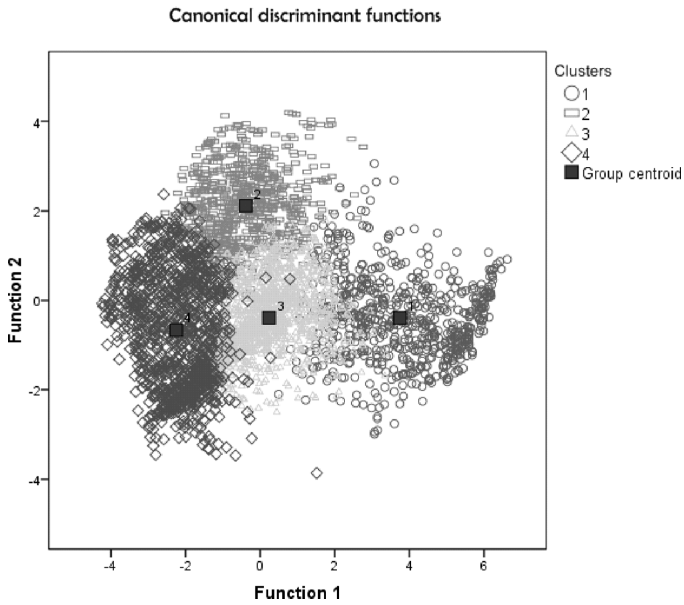


Table 2 Coordinates of the obtained classes' centres

	Class			
Coefficients	*1*	*2*	*3*	*4*
K1	4.932	1.496	1.323	1.307
K2	0.942	0.050	0.084	0.041
K4	0.659	0.112	–0.039	–0.233
K5	0.147	0.299	0.553	0.356
K6	0.189	0.142	0.236	0.384
K7	0.043	0.055	0.172	0.389
K9	12.770	7.685	1.730	5.335
K10	0.971	0.626	1.564	0.521
K11	3.804	1.384	1.390	1.212
K13	1.498	1.424	1.158	0.084
K14	17.897	13.121	3.259	0.978
K15	15.214	15.601	9.322	8.193
K18	11.042	6.702	5.430	–4.979
K19	16.456	8.530	3.580	–8.024

Figures 2 and 3 illustrate the dynamics of quarterly time series $\mathrm{BCR}_t^{(k)} (k = 1, 2, 3, 4)$ of the branches credit ratings estimated by formula (17).

For clarity, the following notations are used: INDUSTR – for industry branch $(\mathrm{BCR}_t^{(1)})$, CONSTR – for construction branch $(\mathrm{BCR}_t^{(2)})$, TRANSPR – for transport branch $(\mathrm{BCR}_t^{(3)})$, TRADER – for trade branch $(\mathrm{BCR}_t^{(4)})$.

The graph of quarterly integral creditworthiness indicator of the national economy ICI_t, which is calculated according to equations (18) and (19), is represented in Figure 4.

Table 3 Economic interpretation of the classes' centres

Direction of economic activity	*Class 1*	*Class 2*	*Class 3*	*Class 4*
Liquidity position, liquid reserve	High, large	Moderate, sufficient	Low, insufficient	Low, insufficient
Financial soundness	High	Average	Below an average	Below an average
Business activity	Very high	Very high	High	Moderate
Efficiency, financial result	High, stable	Average, stable	Below an average, unstable	Low, close to unprofitability

Figure 2 Quarterly time series of the BCR's for industry and construction

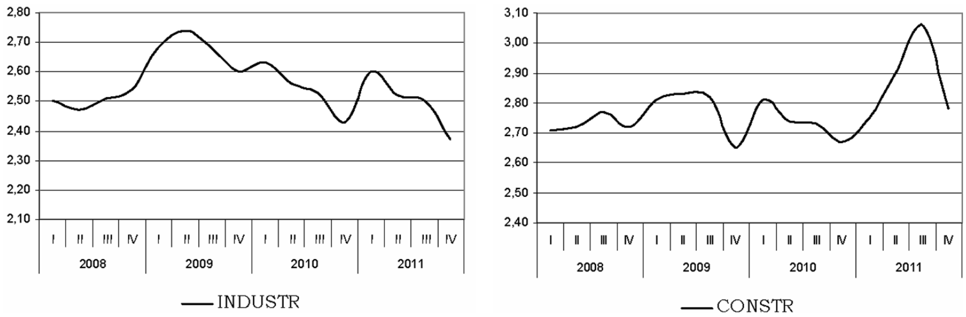
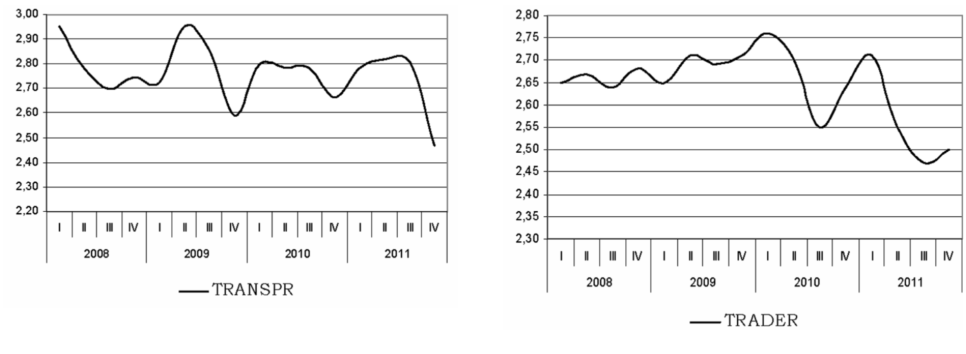
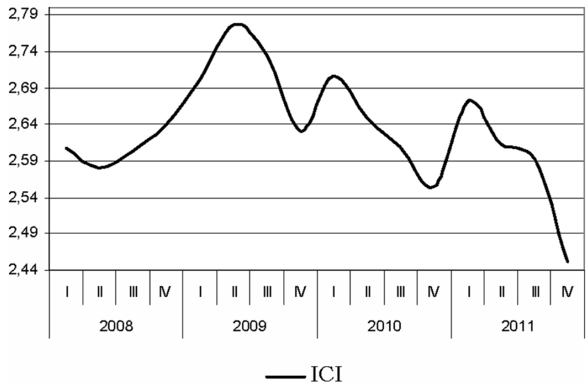


Figure 3 Quarterly time series of the BCR's transport and trade



Economic interpretations of the obtained results. The obtained statistical results have substantial economic interpretation. Since the second quarter of 2009, the dynamics of branch credit ratings got a positive trend, indicating a general deterioration in the creditworthiness of companies. This fact can be interpreted as a result of negative impact of World financial crisis on Belarusian economy. The Belarusian ruble was *devalued* against the US dollar by *20*% in January of 2009, the GDP expanded only 0.2% this year.

Figure 4 Quarterly time series of integral creditworthiness indicator of the economy ICI_t



This period was also characterised by the decrease in the volume of foreign trade and deterioration in the balance of trade. The peak of crisis was in the third, fourth quarters of 2009. Since 2010, the macro-economic situation has begun to improve and the ratings of the companies have increased.

Thus, the second quarter of 2009 can be considered as the moment of structural change of analysed indicators of dynamics model that is confirmed by the results of econometric modelling (see Section 4).

The decline of creditworthiness in economy on the whole in the first quarter of 2011 can be explained by the essential change of supply price conditions of Russian gas to the Republic of Belarus that influenced the financial results of enterprises activity. However, since the second quarter of 2011, there is a trend of creditworthiness improvement of the considered companies can be noticed. It can be considered as the result of accounting peculiarities in Belarus, fixing significant increase of profit volume.

The special situation was observed in 2011 in the construction. Since the second quarter of 2010, the average credit rating of construction companies CONSTR has had a permanent positive trend. Such situation is explained by the sufficient influence of energy price increase on construction branch (that was typical for the period of 2010 and the first half of 2011). In addition, rather sharp deterioration of CONSTR in the third quarter of 2011 was caused by almost threefold devaluation of Belarusian ruble which significantly reduced purchasing power of the population and the companies. In addition, cuts in housing construction as well as high interest rates on loans for the purchase of real estate had a negative impact on the creditworthiness of construction companies.

Table 4 shows the estimates of the rating migration matrix for industrial companies $P^{(1)} = (\widehat{p}_{rs}^{(1)})\ (r, s \in S(L) = \{1, \ldots, L\},\ L = 4)$ over the entire observation period 2008–2011 and for the period 2010–2011. To estimate the matrices, the formulas (20) were used.

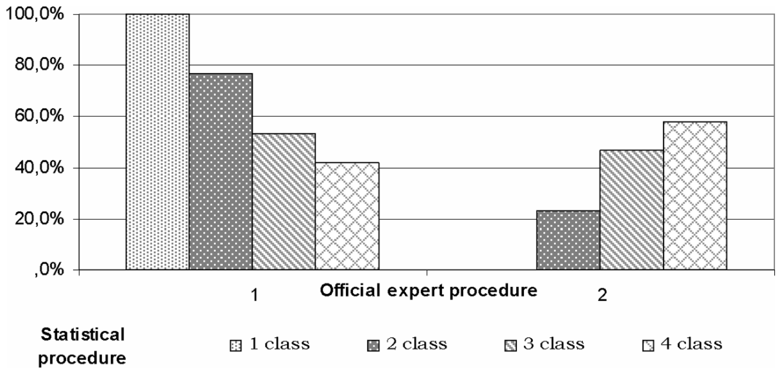
As would be expected in 2011, there is an increase of migration of industrial enterprises from the first class with the highest credit rating due to the currency crisis.

Table 4 The matrices of ratings migration for industrial companies

2008–2011				*2010–2011*			
0.719	0.171	0.067	0.043	0.641	0.192	0.090	0.077
0.139	0.637	0.116	0.108	0.150	0.630	0.126	0.094
0.066	0.160	0.667	0.107	0.038	0.188	0.675	0.100
0.043	0.188	0.087	0.682	0.081	0.202	0.111	0.606

Conformity with the official expert procedure. Official expert evaluation procedure classifies all companies only into two classes ('creditworthy' and 'not creditworthy') on the basis of two coefficients K1 and K4 with expert thresholds. Obviously, it gives a very rough estimate of creditworthiness. For this reason, such a simplified procedure cannot be considered as some benchmark. The proposed statistical procedure uses 14 coefficients and allows classifying companies into 4 classes. Our extensive experimental study of the effectiveness of statistical procedure proves its greater flexibility. At the same time, it shows a certain coherence with the official procedure. So, analysis of Spearman's and Kendall's rank correlation coefficients (Johnson and Wichern, 2007) demonstrates a coherence in classification results received by using statistical algorithms and official acting procedure for evaluating the creditworthiness of companies. The values of these coefficients are 0.484 and 0.542, respectively, and p-values in both cases are much smaller than 0.05. The analysis of relation between classification variables on the basis of cross tables also indicates consistency of decisions: quantitative measures of contingency φ-coefficient and Cramer's V-statistics are 0.663 and 0.469, respectively, and they are statistically significant on 0.05 level of significance. More profound statistical analysis shows that the expert procedure is too 'rigid' that causes more pessimistic forecasts of creditworthiness of companies (see illustration of comparative analysis of the distributions of companies on creditworthiness classes by means of the expert and proposed statistical evaluation procedures in Figure 5).

Figure 5 Distributions of companies on creditworthiness classes by means of the expert and statistical procedures

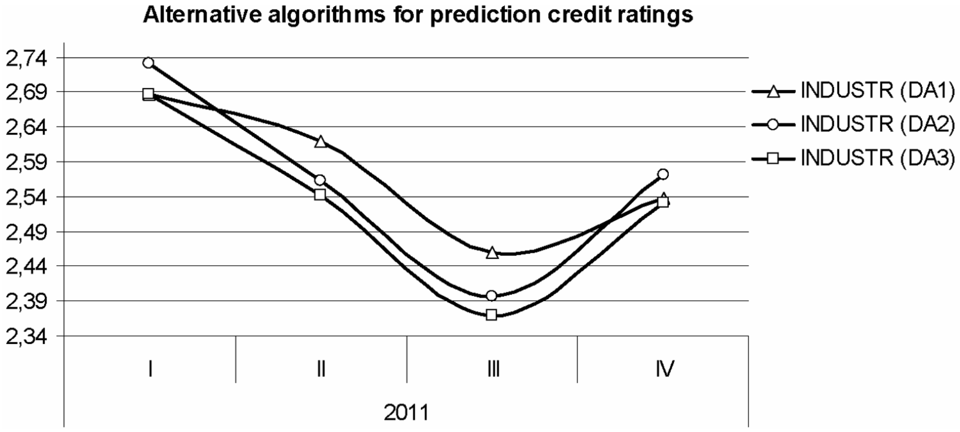


3.2 Discriminant analysis of out-of-sample data

For classification of new observations (out-of-sample data), the above-mentioned discriminant analysis algorithms DA1, DA2 and DA3 are used. The parameters of algorithms were estimated on the classified training sample for 2008–2010, which was obtained by *L*-means algorithm on the first stage. Figure 6 represents the dynamics of quarterly time series for average ratings of the industry companies INDUSTR obtained on the basis of alternative algorithms of discriminant analysis.

The usage of three algorithms gives similar results. However, the credit ratings received on the basis of DA1 seems to be more pessimistic.

Figure 6 Quarterly industrial rating dynamics calculated using alternative algorithms



4 Econometric modelling and forecasting of statistical credit ratings

One more stage of the offered statistical methodology verification is building and analysis of various types' econometric models:

- econometric models of BCR which use corresponding economic indicators as explanation variables (factors) and which take into account both structural and seasonal changes
- econometric models of dependences of output branch indicators on branch credit ratings (BCR).

Let us present a brief summary of some econometric models on the base of quarterly data since *Q*1 2008 till *Q*4 2011. Unfortunately, the available time series is quite short. It does not allow to build large multi-factor econometric models. For this reason, we restrict ourselves to the construction of feasible rather simple models. However, these models are statistically significant and give consistent economic conclusions.

We shall use the following notation: RATEDISC − average interest rate on bank loans in national currency (not taking into account Interbank market) including loans of the resources of the National Bank and the Government; USDI − growth ratio of an exchange rate of Belarusian ruble to US dollar; OILI − growth rate of oil price; GASI − growth rate of gas price; *S*1, *S*2, *S*3, *S*4 − seasonal dummy variables for quarters 1, 2, 3, 4; DUM_2009_2 and DUM_2011_1 − dummy variables for structural breaks at the moments *Q*2 2009 and *Q*1 2011.

4.1 The relationships between output indicators and BCR

We use the following quarterly *output indicators for branches of national economy* (Statistical Yearbook of the Republic of Belarus, 2011):

- *INDUSTI* – index of industrial production, which is a ratio characterising changes in the volume of products produced over periods compared (growth rate of volume of industrial production related to the previous period)
- *CONSTRI* – index of volume of contract works performed in construction economic activity, which includes work under construction contracts (agreements) classified in section F 'Construction' (growth rate of volume of contract works related to the previous period)
- *TRANSPI* – index of freight turnover, which is the volume of work of transport to carry freights (growth rate of freight turnover related to the previous period)
- *TRADERI* – volume index of retail trade turnover (growth rate of retail trade turnover related to the previous period).

Models for the index of industrial production. For this indicator, two models are built: one-dimensional model of dependence of the index of industrial production INDUSTI from a branch credit rating INDUSTR and multidimensional vector autoregressive model VAR (Lütkepohl, 2005), which allows to consider linear interdependencies among these indicators.

Univariate model

$$\begin{aligned}\mathrm{INDUSTI} = &\underset{(-4.2)}{-66.389}\cdot \mathrm{INDUSTR}(-1) + \underset{(5.0)}{0.299}\cdot \mathrm{OILI}\\ &+\underset{(3.9)}{11.586}\cdot \mathrm{DUM\_2009\_2} + \underset{(6.0)}{233.647},\end{aligned}$$

$$R^2 = 0.81,\ P_{LM} = 0.47,\ P_{JB} = 0.51.$$

Vector autoregressive model with exogenous variables VARX(1)

$$\begin{aligned}\mathrm{INDUSTI} = &\underset{(0.77)}{0.088}\cdot \mathrm{INDUSTI}(-1) - \underset{(-2.17)}{35.550}\,\mathrm{INDUSTR}(-1) + \underset{(3.21)}{141.599}\\ &+\underset{(5.13)}{0.322}\cdot \mathrm{OILI} + \underset{(3.36)}{0.998}\cdot T,\\ \mathrm{INDUSTR} = &\underset{(-2.09)}{-0.003}\cdot \mathrm{INDUSTI}(-1) + \underset{(2.08)}{0.499}\cdot \mathrm{INDUSTR}(-1)\\ &+\underset{(0.67)}{0.0006}\cdot \mathrm{OILI} - \underset{(-1.68)}{0.007}\cdot T + \underset{(2.53)}{1.635},\end{aligned}$$

$$R^2 = 0.62(\text{for the first equation}),\quad R^2 = 0.80(\text{for the second equation}),$$

$$P_{\text{VAR LM}} = 0.2,\ P_{\text{Joint JB}} = 0.51.$$

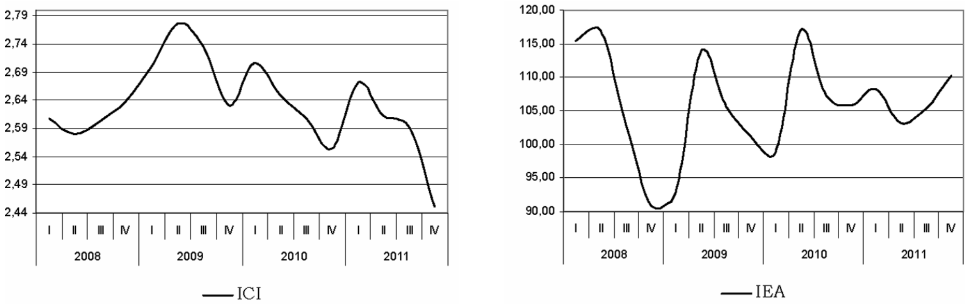
In the above formulas, we use the following notations: below the parameter estimates the values of *t*-statistics are given in parentheses; R^2 – determination coefficient; $P_{\text{VAR LM}}$ – *p*-value of Lagrange multiplier test for VAR when testing autocorrelation in the *residuals*, $P_{\text{Joint JB}}$ – *p*-value of Jarque–Bera test on the joint normal distribution of the *residuals* of two equations. Both models demonstrate statistically significant relationship between index of industrial production and corresponding credit rating, which has substantial economic interpretation: *if creditworthiness of branch on the whole deteriorates the* volume of industrial production *decreases*. Thus, the branch rating plays a role of *leading indicator* for the index of industrial production. Similar models were also built for variables CONSTRI, RANSPI and TRADERI.

4.2 Econometric relation between integral indicators for output and creditworthiness of the national economy

The following integral indicators are used: IEA – index of economic activity which was worked out by the National Bank of the Republic of Belarus; ICI – integral creditworthiness indicator of the economy developed within this research. Index IEA is calculated as the weighed sum of output indicators for branches of national economy which were described before. It represents the aggregated indicator of economic activity in the country.

The visual analysis of graphs for quarterly time series ICI and IEA (Figure 7) allows to assume the existence of the inverse statistical relationship between these indicators as well as seasonal and structural changes.

Figure 7 Dynamic of quarterly time series ICI and IEA



Let us present a short description of the models for ICI and IEA in the class of trend stationary models.

Univariate models with linear trend, structural and seasonal changes

$$\text{ICI} = \underset{(3.5)}{0.0295}\cdot T - \underset{(-6.2)}{0.06}\cdot \text{DUM}\_2009\_2\cdot T + \underset{(7.9)}{0.422}\cdot \text{DUM}\_2009\_2$$
$$+ \underset{(2.3)}{0.005}\cdot \text{DUM}\_2011\_1\cdot T + \underset{(2.1)}{0.036}\cdot S1 - \underset{(-3.2)}{0.056}\cdot S4 + \underset{(88.7)}{2.534}$$

$R^2 = 0.93$, $P_{LM} = 0.11$, $P_{JB} = 0.98$;

$$\text{IEA} = -\underset{(-4.3)}{6.443}\cdot T + \underset{(-4.2)}{6.582}\cdot \text{DUM}\_2009\_2\cdot T - \underset{(-2.6)}{18.158}\cdot \text{DUM}\_2009\_2$$
$$+ \underset{(2.4)}{6.845}\cdot S2 + \underset{(23.8)}{121.715},$$

$R^2 = 0.74$, $P_{LM} = 0.12$, $P_{JB} = 0.35$.

Univariate models with explanation for economic factors

$$\text{ICI} = -\underset{(2.0)}{0.012} \cdot \text{RATEDISC}(-7) - \underset{(-2.6)}{0.006} \cdot \text{OILI} - \underset{(-3.2)}{0.003} \cdot \text{USDI} + \underset{(12.5)}{3.697},$$

$$R^2 = 0.82,\ P_{LM} = 0.11,\ P_{JB} = 0.65.$$

$$\text{IEA} = \underset{(2.3)}{0.212} \cdot \text{OILI} - \underset{(2.1)}{0.145} \cdot \text{GASI} + \underset{(-3.2)}{100.0},$$

$$R^2 = 0.5,\ P_{LM} = 0.56,\ P_{JB} = 0.64.$$

Vector autoregressive model with exogenous variables VARX(1)

$$\text{IEA} = -\underset{(-1.43)}{38.841} \cdot \text{ICI}(-1) + \underset{(0.28)}{0.118} \cdot \text{IEA}(-1) + \underset{(3.27)}{0.227} \cdot \text{OILI} - \underset{(-2.29)}{0.163} \cdot \text{GASI}$$
$$+ \underset{(0.94)}{0.387} \cdot \text{RATEDISC}(-1) + \underset{(2.28)}{186.416},$$

$$\text{ICI} = \underset{(2.2)}{0.644} \cdot \text{ICI}(-1) - \underset{(-1.59)}{0.005} \cdot \text{IEA}(-1) + \underset{(0.08)}{0.00034} \cdot \text{OILI} - \underset{(-0.16)}{0.00039} \cdot \text{GASI}$$
$$- \underset{(-2.43)}{0.0101} \cdot \text{RATEDISC}(-1) + \underset{(1.49)}{1.656},$$

$$R^2 = 0.59\left(\text{for the first equation}\right),\ R^2 = 0.67\left(\text{for the second equation}\right),$$

$$P_{\text{VAR LM}} = 0.19, P_{\text{Joint JB}} = 0.29.$$

All models are statistically significant and have substantial economic interpretation. According to expectations, indicators ICI and IEA are connected by the inverse statistical relationship. Oil and gas prices have essential influence on the index of economic activity. In its turn, the change of credit interest rates has expected the influence on ICI.

The obtained results of econometric modelling of BCR and ICI show that these measures adequately take into account the impact of the major economic factors of national economy. Moreover, they have interpreted dependences with branches output indicators. This makes it possible to talk about the positive results of verification of the constructed system of credit ratings on the macro-level.

5 Conclusion

This paper presents the system of RSCR to assess the creditworthiness of non-financial companies. It is based on the methods of multivariate statistical analysis as well as on a set of multivariate econometric models used to assess the creditworthiness on the macro-level. The proposed methodology is applied to evaluate the creditworthiness of Belarusian companies on the basis of their financial reports data obtained within the National Bank's monitoring system. The software developed on the proposed statistical methodology has passed a preliminary approbation on an expanded database in the National Bank of the Republic of Belarus. The approbation results confirm the acceptable effectiveness of this approach in solving some problems which the Central Bank faces. They include the evaluation of the stability of the national economy as well as the assessment of the system risks of the banking sector. A promising direction for the development of this methodology is the remote analysis of companies similar to that performed for banks (Peresetsky, 2012).

Acknowledgement

The work is conducted with the financial support of the National Bank of the Republic of Belarus.

References

Basel Committee on Banking Supervision (2003) *Consultative Document: Overview of The New Basel Capital Accord*, Bank for International Settlements, Basel.

Bhar, R. and Hamori. S. (2004) *Hidden Markov Models. Application to Financial Economics*, Kluwer Academic Publishers, Dordrecht.

Huberty, C.J. and Olejnik, S. (2006) *Applied MANOVA and Discriminant Analysis*, 2nd ed., John Wiley and Sons, Hoboken, New Jersey.

Johnson, R.A. and Wichern, D.W. (2007) *Applied Multivariate Statistical Analysis*, 6th ed., Pearson Education, Texas A&M University.

Jorion, P. (2009) *Financial Risk Manager Handbook*, 5th ed., John Wiley & Sons, New York.

Kharin, Yu. (1996) *Robustness in Statistical Pattern Recognition*, Kluwer Academic Publishers, Dordrecht, Boston, London.

Lütkepohl, H. (2005) *New Introduction to Multiple Time Series Analysis*, Springer, Berlin.

Malugin, V., Korchagin, O. and Hryn, N. (2009) 'Examining efficiency of algorithms of classifying bank borrowers on the basis of balance ratios', *Bank Bulletin, Minsk*, Vol. 7, pp.27–32 (in Russian).

Mátyás, L. and Sevestre, P. (Eds.) (2008) *The Econometrics of Panel Data*, 3rd ed., Springer, Berlin, Heidelberg.

Milevskiy, P. and Zubovich, A. (2012) 'Monitoring of non-financial enterprises – as the national bank's function', *Bank Bulletin, Minsk*, Vol. 25, pp.52–60 (in Russian).

Peresetsky, A.A. (2012) *Econometric Methods in the Remote Analysis of Russian Banks*, Publishing House HSE (in Russian).

Statistical Yearbook of the Republic of Belarus (2011) National Statistical Committee of the Republic of Belarus, Minsk (in Russian).