# Financial and Economic Analysis of Steel Industry by Multivariate Analysis

*Jana CZILLINGOVÁ* – Igor PETRUŠKA – Michal TKÁČ***

**Abstract**

*Presented paper deals with the implementation of multidimensional statistical methods used to compare the financial health of selected global steel producers. We show the evolution of the steel industry in years 2003 – 2007 using factor analysis, multidimensional scaling and cluster analysis. Results of implemented multidimensional statistical methods are transparently and clearly presented with simple two-dimensional graphical output. The methodology for implementation of various analyses allows us to give satisfactory answers to many questions related to the identification of financial health, but also a business failure or bankruptcy prediction. The above method can be also used for other industries based on available data.*

**Keywords**: *steel industry, principal component analysis, factor analysis, multidimensional scaling, cluster analysis, hierarchical cluster analysis, dendrogram*

**JEL Classification**: B17, C63, G33, L61

## Preliminary

Assessment of the financial situation of the enterprise through financial analysis is a complex expression of levels of all business activities that the company is presented to the market. A very important step in financial analysis is the comparison of business characteristics not only with the average indicators in the relevant field of business, but also with the results of competitors. The fact, what

* Jana CZILLINGOVÁ, University of Economics in Bratislava, Faculty of Business Economics with Seat in Košice, Department of Finance and Accounting, Tajovského 13, 041 30 Košice, Slovakia; e-mail: jana.czillingova@euke.sk

** Igor PETRUŠKA – Michal TKÁČ, University of Economics in Bratislava, Faculty of Business Economics with Seat in Košice, Department of Business Informatics and Mathematics, Tajovského 13, 041 30 Košice, Slovakia; e-mail: igor.petruska@yahoo.com; dekan@euke.sk

*Acknowledgement*: The authors want to thank the anonymous referees for very helpful remarks that have substantially improved the paper.

is the quality and volume of products in the company's innovation activity, the level of commercial activity, as well as other business activities is reflected in the financial situation. Its analysis is the starting point for analysis of the economic performance of the company, its efficiency, capacity utilization, profitability, solvency and liquidity, but also for stocks, receivables management and debt. Through financial analysis we can identify the strengths and weaknesses in the enterprise, diagnose its financial health and to identify the causes that determined the financial situation of the company. Assessment of the financial situation of the companies and industries where they operate is in today's turbulent times matter of great importance.

For the description of the financial situation financial ratio indicators are used, by which we compare results achieved with the results of competitors.

As in any industry many competitors operate and each competitor can be described in a certain period by multivalued absolute or financial ratios, we have to work with large amounts of data. For processing of multidimensional data multidimensional statistical methods should be used. These methods are suitable means for comparing, exploring the links between competing companies with the financial distress prediction of companies. Analysis obtained using these methods are also associated with a vivid and easy to interpret output.

The aim of this paper is to show the possibility of using some of multivariate statistical methods for financial analysis of companies and entire industries.

The contribution also proposes methodology for reduction of 14 financial ratios (FPU) to 4 ratio indicators, so that the reduced group reflects satisfactorily the financial situation of the company and industry.

## Literature Review

Comparing the financial health of companies may be based on either univariate procedures using selected financial ratios (Beaver, 1968) or multivariate procedures (e.g. discriminant analysis) (Altman, 1968) Discriminant analysis, Logit, Probit and neural networks are statistical techniques which are often used to develop predictive models of bankruptcy. In this work (Aziz and Dar, 2006) categorization of prediction models is made: • statistical models; • models based on expert systems and artificial intelligence; • theoretical models.

Discriminant predicate models in the steel sector are dealt with in the work of Trill, Rabidoux and Pesi (2008). 33 financial ratios of 28 steel companies in the 1990 – 2005 time periods were used to develop this model. Multidimensional scaling aimed at examining the links between financial indicators that can be used to describe a financial state of companies in the industry is dealt by the work of Molinero and Ezzame (1991). Multidimensional scaling and cluster analysis to determine the competitive structure of the banking sector (111 banks) were used in the work (DeSarbo at al., 2008). Financial benchmarking in the paper industry analyzed using neural networks (98 companies for the period 1995 to 2002) is devoted by work (Eklund, 2004).

Building on previous results we use some of these statistical techniques to compare the financial health, represented by the selected financial indicators, of some of the world's leading steel companies.

## Data

For the following analyses financial ratios of global steel companies from the period 2003 to 2007 were used. These indicators were derived from financial statements (in particular the balance sheet and profit and loss statement) of the annual reports available on the websites of the companies compared.

Representative sample of 28 global steel companies was used:

Severstal (Se), EVRAZ (EV), Magnitogorsk (MM), Novolipeck (NL), U. S. Steel Corporation (USS), U.S. Steel Košice (USSK), CMC Steel Group(CMC), Nucor Corp (NU), Mittal Steel (Mi), Arcelor (Ar), ArcelorMittal(AM), Svenskt Stal AB (SS), Worthington Group (Wo), Nippon Steel (Ni), JFE (JF), Sumitomo (Su), Kobe Steel (Ko), ThyssenKrupp (Th), Salzgitter (Sa), Corus Group (Co), TATA (TA), Ruukki (Ru), VoestAlpine (Vo), BlueScope (Bl), One Steel (On), CSN (Cs), Gerdau (Ge), POSCO (PO).

Mentioned companies present the largest world steel companies. If we do not consider China, then by the amount of steel produced in 2006, 21 of these companies were located in the first 25 the world's steel producers. Overall, these manufacturers in 2007 produced 456 million. tons of steel which is 53% share of world production (excluding China).

Following financial ratios were used:

Responsive Liquidity (L1), Current Liquidity (L2), Total Liquidity (L3), Stock Turnover (OZAS), Receivables Turnover (OPOH), Turnover Assets (OAKT), Total Indebtedness (CZAD), Cash Flow Level (STSAMFIN), Financial Leverage (FINPAK), Insolvency (PLATNES), Return On Assets (ROA), Return On Equity (ROE), Return On Sales (ROSE), Return On Capital Employed (ROCE).

Data base analysis is therefore: 28 x 14 x 5 = 1960 values.

## Methods

Following multidimensional statistical methods were used while processing these data:

- Principal Component Analysis (PCA),
- Factor Analysis (FA),
- Cluster Analysis (CA),
- Multidimensional Scaling (MDS).

Other multi-dimensional statistical methods could be successfully used:

- Regression analysis,
- Correspondence analysis,
- Discriminant analysis.

Suitable are also other procedures:

- Neural Networks,
- Genetic Algorithms.

The aim of principal components analysis is to identify new, hidden variables. We require that the new variables (components, factors) to explain the most of the variability of the original variables. Much statistical software considers this method as part of the PCA. Both methods are used to find hidden variables standing in the background, which are called components or factors. In our case we work with a 14-dimensional datasets and our aim will be to minimize the dimension of space (number of variables) so that the identified factors adequately explain the variability of the original variables and their dependency.

Cluster Analysis is used to split the data (business) in groups (clusters) most closely resembling each other. Data will be divided into clusters so that the companies belonging to the same cluster have been close, and thus similar to companies belonging to different clusters have been away. At the beginning, we will choose the number of categories in which to divide companies. The company is included in the group, which is closely based on those criteria. The calculation uses an iterative method. At each step, the inclusion of companies is reviewed and process ends when there is no transfer. The company, which is assigned at the beginning to one group, may not be there at the end of the process.

Hierarchical Cluster Analysis (HCA) is used to split the business population into groups so that the companies inside the group are most resembled. At the beginning of each enterprise creates one group. Companies are progressively aggregated subject to certain methods, up to the end, when all belong to one group. Hierarchical CA is clearly shown by dendrogram. It is not necessary to know the exact number of clusters at the beginning of the hierarchical procedure. It is used as the initial clustering solution that is supplemented by Non-hierarchical CA; HCA is not used for a large number of objects so the dendrogram is easily readable.

The purpose of MDS is to optimally reduce the size of the data and studied objects relations in the reduced space. In the examined file every brand of objects (business) is described $p$-dimensional vector of values. In our case we have

$n$ = 28 (business) and each business is described by values FPU 14, i.e. $p$ = 14 Multidimensional scaling allows the analyzed objects to be reflected in the plane (space) so as to maintain the distance (dissimilarity) between them. It allows us to draw a map on which businesses are recorded, while businesses with similar FPU values are shown close together and companies with different FPU will be shown further apart.

## Results

In a similar type of analysis from a statistical point of view it is appropriate to exclude those companies whose indicators values are remote from other values. Exclusion such companies from the analysis usually improve strength of the models notice (increase in the percentage of explained variance). However, on the other hand, from the economic perspective, we lose a comprehensive view on group of companies forming a logically closed group in the industry. In our case, it is the choice described in the introductory part (most of steel companies excluding China). Therefore, we further decided to keep businesses with outlying values in the analysis even at the price of reduction of the percentage of variance explained.

In our case, using the FA, we will deal with analyzing the structure of interdependencies FPU with the assumption of some degree of concentration. We therefore assume that these dependencies are sufficiently well explained by the action of a small number of dominant factors. The aim will be the analysis of the impact of the various factors, which are behind the FPU mutually correlated, so that the selected factors significantly clarified the relevant subject but their number was as small as possible. If such factors exist, then they should explain the observed variability in the FPU.

Multivariate statistical methods are common part of the various statistical packages. Outputs from the statistical software SPSS were used to the interpretation of obtained results.

Factor Analysis was applied simultaneously to all businesses for all periods. This means that the number of observations and number of variables was 10 : 1.

The correlation table shows strong correlation among multiple variables. Correlation coefficient ranges from –0.65 to 0.96.

We determine the number of dominant factors using the PCA method.

We consider 14 input variables (FPU) the maximum number of considered factors is 14. Our aim is to select the smallest number of dominant factors to explain most of the original variance. As dominant factors we select those, for which the total value (total eigenvalue) is greater than 1. Other factors explain variance less than the original variable (the fifth factor of only 0.9379).

T a b l e 1

**Correlation Matrix**

| | L1 | L2 | L3 | OZAS | OPOH | OAKT | CZAD | STSAMFIN | FINPAK | PLATNES | ROA | ROE | ROCE | ROSE |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **L1** | 1.00 | 0.93 | 0.84 | 0.31 | 0.06 | –0.04 | –0.34 | 0.56 | –0.12 | –0.14 | 0.47 | 0.29 | 0.22 | 0.52 |
| **L2** | 0.93 | 1.00 | 0.96 | 0.38 | –0.10 | 0.14 | –0.46 | 0.66 | –0.20 | –0.32 | 0.48 | 0.19 | 0.12 | 0.44 |
| **L3** | 0.84 | 0.96 | 1.00 | 0.30 | –0.04 | 0.25 | –0.54 | 0.66 | –0.31 | –0.41 | 0.48 | 0.14 | 0.06 | 0.37 |
| **OZAS** | 0.31 | 0.38 | 0.30 | 1.00 | 0.35 | 0.48 | –0.45 | 0.41 | –0.35 | –0.10 | 0.47 | 0.15 | 0.16 | 0.24 |
| **OPOH** | 0.06 | –0.10 | –0.04 | 0.35 | 1.00 | 0.20 | –0.16 | 0.01 | –0.31 | 0.29 | 0.28 | 0.25 | 0.23 | 0.17 |
| **OAKT** | –0.04 | 0.14 | 0.25 | 0.48 | 0.20 | 1.00 | –0.21 | 0.02 | –0.35 | –0.37 | 0.13 | –0.06 | –0.03 | –0.30 |
| **CZAD** | –0.34 | –0.46 | –0.54 | –0.45 | –0.16 | –0.21 | 1.00 | –0.88 | 0.78 | 0.25 | –0.65 | –0.02 | 0.08 | –0.48 |
| **STSAMFIN** | 0.56 | 0.66 | 0.66 | 0.41 | 0.01 | 0.02 | –0.88 | 1.00 | –0.53 | –0.25 | 0.64 | 0.07 | –0.02 | 0.60 |
| **FINPAK** | –0.12 | –0.20 | –0.31 | –0.35 | –0.31 | –0.35 | 0.78 | –0.53 | 1.00 | 0.14 | –0.44 | 0.08 | 0.12 | –0.21 |
| **PLATNES** | –0.14 | –0.32 | –0.41 | –0.10 | 0.29 | –0.37 | 0.25 | –0.25 | 0.14 | 1.00 | –0.03 | 0.14 | 0.18 | 0.13 |
| **ROA** | 0.47 | 0.48 | 0.48 | 0.47 | 0.28 | 0.13 | –0.65 | 0.64 | –0.44 | –0.03 | 1.00 | 0.68 | 0.54 | 0.85 |
| **ROE** | 0.29 | 0.19 | 0.14 | 0.15 | 0.25 | –0.06 | –0.02 | 0.07 | 0.08 | 0.14 | 0.68 | 1.00 | 0.90 | 0.71 |
| **ROCE** | 0.22 | 0.12 | 0.06 | 0.16 | 0.23 | –0.03 | 0.08 | –0.02 | 0.12 | 0.18 | 0.54 | 0.90 | 1.00 | 0.55 |
| **ROSE** | 0.52 | 0.44 | 0.37 | 0.24 | 0.17 | –0.30 | –0.48 | 0.60 | –0.21 | 0.13 | 0.85 | 0.71 | 0.55 | 1.00 |

*Source:* Used websites (in References); own computation.

When we consider four appropriate factors, they will explain up to 82.31% of the original variance. Considering only the dominant factors we achieve a substantial reduction in size of the original area of 14 to 4 dimensions in maintaining a good degree of clarification of the original variance.

T a b l e 2

**Extraction Method: Principal Component Analysis**

| Total Variance Explained | | | |
|---|---|---|---|
| **Component** | **Eigenvalues** | | |
| | total | % of variance | cumulative % |
| 1 | 5.5206 | 39.43 | 39.43 |
| 2 | 2.7277 | 19.48 | 58.92 |
| 3 | 1.9009 | 13.58 | 72.49 |
| **4** | 1.3746 | 9.82 | **82.31** |
| 5 | 0.9379 | 6.70 | 89.01 |
| 6 | 0.5628 | 4.02 | 93.03 |
| 7 | 0.4102 | 2.93 | 95.96 |
| 8 | 0.2207 | 1.58 | 97.54 |
| 9 | 0.1560 | 1.11 | 98.65 |
| 10 | 0.0660 | 0.47 | 99.12 |
| 11 | 0.0508 | 0.36 | 99.49 |
| 12 | 0.0400 | 0.29 | 99.77 |
| 13 | 0.0280 | 0.20 | 99.97 |
| 14 | 0.0039 | 0.03 | 100.00 |

*Source:* Own computation.

According to Meloun, Militký and Hill (2005), we choose the number of factors so that the explained variance percentage ranged from 70% to 90%. In the social sciences 60% is sufficient.

T a b l e 3

**Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization**

| **Rotated Component Matrix** | **Component** | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| L1 | 0.82 | 0.31 | 0.18 | –0.05 |
| L2 | 0.93 | 0.16 | 0.23 | 0.07 |
| L3 | 0.89 | 0.08 | 0.29 | 0.16 |
| OZAS | 0.18 | 0.25 | 0.40 | 0.60 |
| OPOH | –0.37 | 0.41 | 0.33 | 0.39 |
| OAKT | 0.10 | –0.10 | 0.06 | 0.93 |
| CZAD | –0.30 | –0.01 | –0.91 | –0.11 |
| STSAMFIN | 0.54 | 0.07 | 0.77 | –0.09 |
| FINPAK | 0.00 | 0.08 | –0.84 | –0.30 |
| PLATNES | –0.54 | 0.34 | 0.03 | –0.37 |
| ROA | 0.30 | 0.68 | 0.57 | 0.10 |
| ROE | 0.10 | 0.95 | –0.06 | 0.01 |
| ROCE | 0.04 | 0.91 | –0.16 | 0.07 |
| ROSE | 0.29 | 0.72 | 0.47 | –0.32 |

*Source:* Own computation.

Drastical reduction of the original dimension of the space allows us to easily search for links between the business in the investigation industry and simplify the use of other methods. We found that the original 14 variables (FPU) face four dominant factors. To identify these factors FA will be used.

We select the box FPU in each column, with the highest absolute values. Best if these values are just under 1. Based on the above choice, we get the following dominant factors: • liquidity factor; • profitability factor; • factor of indebtedness; • activity factor.

We can say that most variance in the steel industry is explained by liquidity factor, the least variance by a factor of performance. For each company we calculate factors (scores) in each year. Average values of "scores" for the industry present us following table.

Table 4

**Average Values of "Scores"**

| Year | Liquidity | Return | Liability | Activity |
|---|---|---|---|---|
| 2003 | –0.015 | –0.480 | –0.169 | –0.078 |
| 2004 | –0.030 | 0.162 | 0.027 | 0.014 |
| 2005 | 0.057 | 0.322 | –0.033 | 0.213 |
| 2006 | 0.023 | 0.058 | 0.083 | 0.060 |
| 2007 | –0.038 | –0.067 | 0.106 | –0.231 |

*Source:* Own computation.

The table shows that liquidity, profitability and efficiency in the sector peaked in 2005. Indebtedness was even lowest at the end of the period. Overall, from the view of all 4 factors, the financial situation was better at the end of the period than at the beginning. Factor analysis can also be used to determine the position of a company in the industry.

Table 5

**Extraction Method: Principal Component Analysis**

| Communalities | Extraction |
|---|---|
| L1 | 0.81 |
| L2 | 0.94 |
| L3 | 0.91 |
| OZAS | 0.62 |
| OPOH | 0.56 |
| OAKT | 0.88 |
| CZAD | 0.94 |
| STSAMFIN | 0.91 |
| FINPAK | 0.80 |
| PLATNES | 0.54 |
| ROA | 0.89 |
| ROE | 0.92 |
| ROCE | 0.87 |
| ROSE | 0.93 |

*Source:* Own computation.

The requirement to practice – understandable interpretation of analytical results requires to deal with 2 – 3 factors. Location in 2 – 3 dimensional graph is relatively easy to interpret. Although, if any two dominant factors could be explained enough variance, we can depict individual firms in two-dimensional graph. The table Communalities determine how much variance of individual variables (FPU) is explained by PCA method.

We skip FPU OZAS, OPOH, FINPAK, PLATNES, as the proportion of explained variance by the Principal Component Analysis is lowest. We repeat the PCA with 10 remaining FPU.

T a b l e 6

**Extraction Method: Principal Component Analysis**

| Total Variance Explained | | | |
|---|---|---|---|
| **Component** | **Eigenvalues** | | |
| | total | % of variance | cumulative % |
| 1 | 4.97 | 49.74 | 49.74 |
| **2** | **2.29** | **22.93** | **72.67** |
| 3 | 1.17 | 11.69 | 84.35 |
| 4 | 1.12 | 11.25 | 95.60 |
| 5 | 0.17 | 1.69 | 97.29 |
| 6 | 0.09 | 0.92 | 98.21 |
| 7 | 0.08 | 0.78 | 98.99 |
| 8 | 0.05 | 0.48 | 99.47 |
| 9 | 0.04 | 0.38 | 99.85 |
| 10 | 0.01 | 0.15 | 100.00 |

*Source:* Own computation.

If we consider the remaining 10 FPU and reduce the dimension of considered space from 10 to 2, and we would be able to explain nearly 73% of the original variance, which is according to Meloun, Militký and Hill (2005) sufficient amount.

T a b l e 7

**Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization**

| **Rotated Component Matrix(a)** | **Component** | |
|---|---|---|
| | 1 | 2 |
| L1 | 0.79 | 0.23 |
| L2 | 0.90 | 0.09 |
| L3 | 0.91 | 0.02 |
| OAKT | 0.21 | –0.23 |
| CZAD | –0.79 | 0.00 |
| STSAMFIN | 0.88 | 0.07 |
| ROA | 0.61 | 0.67 |
| ROE | 0.07 | 0.96 |
| ROCE | –0.03 | 0.92 |
| ROSE | 0.50 | 0.76 |

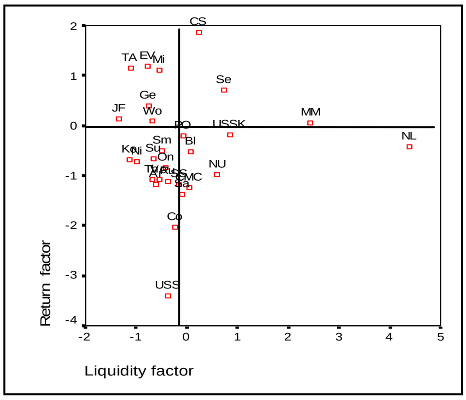
*Source:* Own computation.

The first factor is therefore liquidity again and the second factor is profitability. We show the state of the industry in a plane for each year by these two factors:

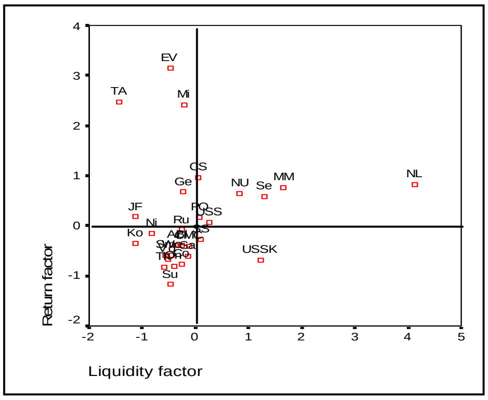
Figure 1

**State of Industry in Plane**

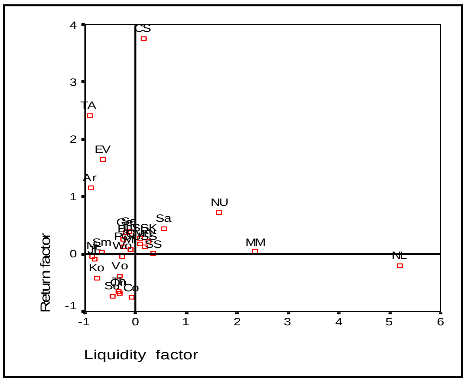
**2003**



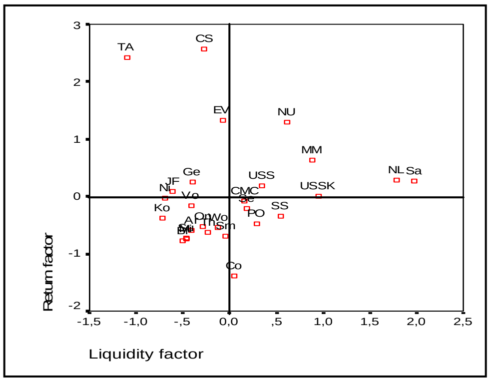

**2004**



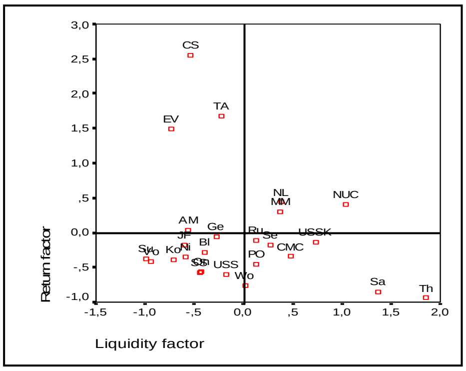

**2005**




**2006**




**2007**

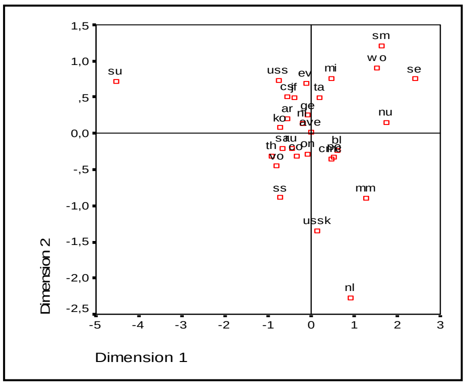



*Source:* Own construction.

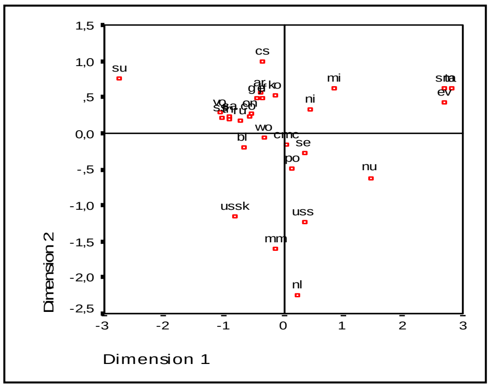
Figure 2

**Multidimensional Scaling**

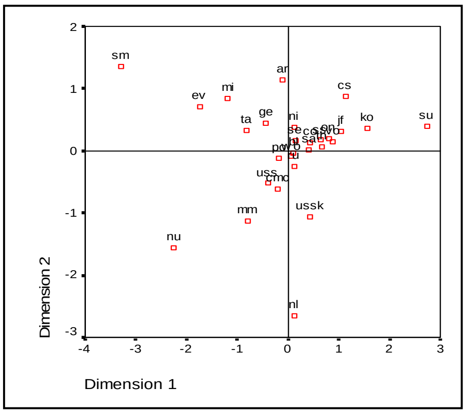
**2003**



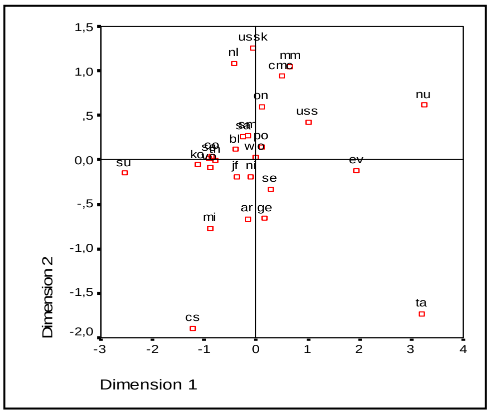

**2004**



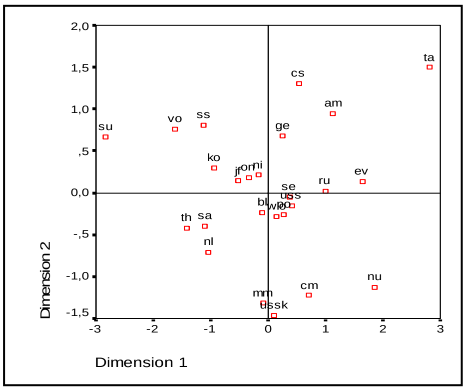

**2005**




**2006**




**2007**




*Source:* Own construction.

In interpreting the results of the analysis we will neither address the comparison of mutual position of steel companies in the sector during the period, nor the identification of companies with extreme values of factors. We can say that the industry as a whole made a significant move "in a northeasterly direction" to the higher value of liquidity and profitability in first three years. In 2006 – 2007 there was a shift in the opposite direction, but the final position is better than the baseline.

Factor analysis thus allows us to reduce the number of FPU and thus facilitate the evaluation of the whole sector, but also the situation of individual companies.

Multidimensional scaling can be also used to the analysis of the status of companies in the sector, in which the similarity of FPU companies translates into a distance between enterprises in two-dimensional or three dimensional space. Companies which have close FPU appear as points close together. If companies have different FPU, the geometric distance of points assigned to them increases. In determining the position of companies in the industry, we can use the fact, that the company (Ave) is located at the intersection of coordinate axes, whose FPU was obtained as a FPU arithmetic average of all surveyed firms.

For the calculation SPSS – Multidimensional scaling (MDS) procedure Alscal – nonmetric scalling was used.

If we want to deal with the situation of only Slovak representative – U.S. Steel Košice (USSK) in the industry, it is easy to find that USSK was "close" to the companies Magnitogorsk (MM) Novolipeck (NL), SSAB (SS), POSCO (PO), NUCOR (NU), CMC during the whole period. This means that the U.S. Steel Košice has similar FPU as mentioned creditworthy companies in selected period. Therefore, it is gratifying, that U.S. Steel Košice belongs to this important group of pairs.

Cluster analysis is rather used to separation of firms into groups (clusters). For reasons of brevity, we address only the outcomes of cluster analysis for 2003 and hierarchical cluster analysis at the end of 2007 period. Entry into the CA were the original data.

We apply non-hierarchical factor analysis using the data from 2003 (K-Means Cluster Analysis). If we choose the number of clusters 2 then we get allocation (27, 1), 3 clusters (6, 1, 21), 4 (4, 20, 3, 1), 5 (4, 1, 1, 5, 17). So we divide companies into four clusters (groups) so that differences within the clusters were small and the differences among businesses in different clusters were substantial. Each group then has one representative (Cluster centre).

Number of companies in each cluster is described in following table. In the case of Sumitomo (Su), the distance from the center of a cluster is zero. Center of the cluster thus merges with its sole member.

T a b l e 8

**Cluster Analysis – 2003**

| Final Cluster Centers | | | | |
|---|---|---|---|---|
| | 1 | 2 | **3** | 4 |
| L1 | 0.376 | 0.170 | **1.557** | 0.155 |
| L2 | 0.927 | 0.681 | **2.699** | 0.871 |
| L3 | 1.718 | 1.300 | **3.502** | 1.068 |
| OZAS | 8.796 | 6.017 | **9.242** | 3.840 |
| OPOH | 13.158 | 7.244 | **6.596** | 1.061 |
| OAKT | 1.214 | 0.986 | **0.908** | 0.326 |
| CZAD | 0.465 | 0.647 | **0.251** | 0.931 |
| STSAMFIN | 1.260 | 0.638 | **3.448** | 0.074 |
| FINPAK | 1.924 | 3.342 | **1.349** | 14.506 |
| PLATNES | 0.715 | 0.836 | **0.111** | 0.899 |
| ROA | 0.067 | 0.040 | **0.191** | 0.016 |
| ROE | 0.116 | 0.121 | **0.256** | 0.239 |
| ROCE | 0.178 | 0.198 | **0.332** | 0.229 |
| ROSE | 0.068 | 0.049 | **0.214** | 0.050 |

| | Cluster | Number of cases |
|---|---|---|
| | 1 | 4 |
| | 2 | 20 |
| | 3 | 3 |
| | 4 | 1 |
| Valid | | 28 |

*Source:* Own computation.

T a b l e 9

**Cluster Membership – 2003**

| Comp | Cluster Membership | Distance from cluster centre | Comp | Cluster Membership | Distance from cluster centre |
|---|---|---|---|---|---|
| Se | 1 | 2.03 | Su | 4 | 0.00 |
| EV | 2 | 2.65 | Ko | 2 | 1.84 |
| NL | 3 | 3.36 | Th | 2 | 2.63 |
| MM | 3 | 2.21 | Sa | 2 | 2.00 |
| USS | 2 | 4.28 | Co | 2 | 1.87 |
| USSK | 3 | 3.25 | TA | 2 | 2.79 |
| CMC | 2 | 3.69 | On | 2 | 1.71 |
| NU | 1 | 3.38 | Ru | 2 | 1.41 |
| Mi | 2 | 3.61 | Vo | 2 | 2.61 |
| Ar | 2 | 1.84 | Bl | 2 | 3.41 |
| SS | 2 | 3.51 | Sm | 1 | 2.72 |
| Wo | 1 | 1.33 | CS | 2 | 4.48 |
| Ni | 2 | 1.00 | Ge | 2 | 1.47 |
| JF | 2 | 2.21 | PO | 2 | 3.38 |

*Source*: Own computation.

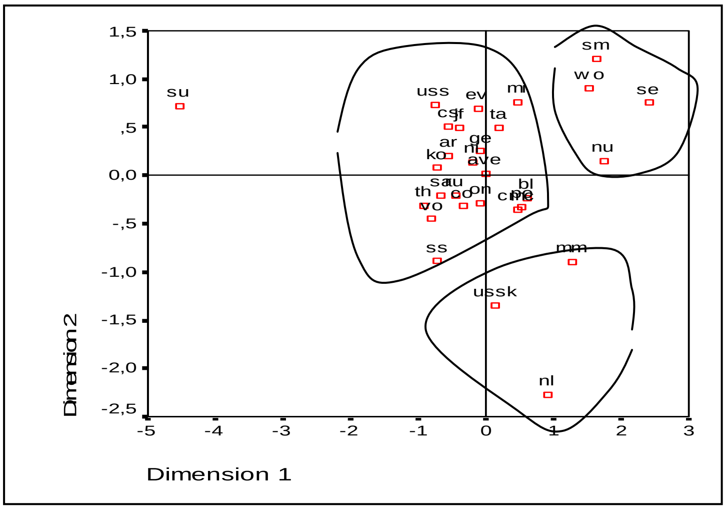
So we have 4 clusters. The first cluster consists of four companies, the second cluster of 20 companies, third cluster of 3, and the fourth cluster of sole company (Sumitomo). Each cluster is represented by the centre and the FPU values of cluster centre are presented in table. If we want to deal again with the situation of USSK by CA, we find that USSK belong to the third, excellent, cluster, along with Novolipeck and Magnitogorsk in 2003.

We compare the results obtained by methods of CA and MDS. Membership to clusters, which was acquired by CA, is shown on the graph obtained through the MDS.

F i g u r e 3

**Cluster Analysis versus MDS**

**2003**




*Source*: Own construction.

We can therefore say that the CA (if we consider all the variables – 14) gave the identical results with MDS

Number of companies should be greater than $2^{\#FPU}$ (where #FPU means the number of used FPU´s). With 28 analyzed companies, we should use four indicators. Thus we reduce the number of indicators. We choose 4 characteristics, one for each factor, so that they correlate with the respective factor more than other indicators. We get the following indicators:

- Current Liquidity (L2),
- Turnover of Assets (OA),
- Total Indebtedness (CZAD),
- Return On Equity (ROE).

If again illustrate the correlation matrix for a reduced group of variables we get

T a b l e 10

**Correlation Matrix for Reduced Group**

| Correlation Matrix | | | | |
|---|---|---|---|---|
| | L2 | OAKT | CZAD | ROE |
| L2 | 1.00 | 0.14 | –0.46 | 0.19 |
| OAKT | 0.14 | 1.00 | –0.21 | –0.06 |
| CZAD | –0.46 | –0.21 | 1.00 | –0.02 |
| ROE | 0.19 | –0.06 | –0.02 | 1.00 |

*Source:* Own computation.

The table shows that the correlation coefficients range from –0.46 to + 0.19, which means that we have selected a group of indicators that are "more orthogonal" than the previous "big" group.

If we choose the number of clusters 3, we got allocation (3, 23, 2), then 4 (23, 3, 1, 1), 5 (10, 14, 1, 1, 2) and 6 (4, 2, 1, 1, 19). We chose 5 clusters.

Table 11

**Cluster Analysis for Reduced Group**

| Cluster | | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| L2 | 1.06 | 0.53 | 3.94 | 2.73 | 0.79 |
| OAKT | 1.05 | 0.84 | 0.80 | 0.93 | 1.87 |
| CZAD | 0.52 | 0.69 | 0.15 | 0.29 | 0.59 |
| ROE | 0.13 | 0.13 | 0.25 | 0.26 | 0.08 |

| Cluster | Number of cases |
|---|---|
| 1 | 10 |
| 2 | 14 |
| 3 | 1 |
| 4 | 1 |
| 5 | 2 |

*Source:* Own computation.

Table 12

**Cluster Membership for Reduced Group**

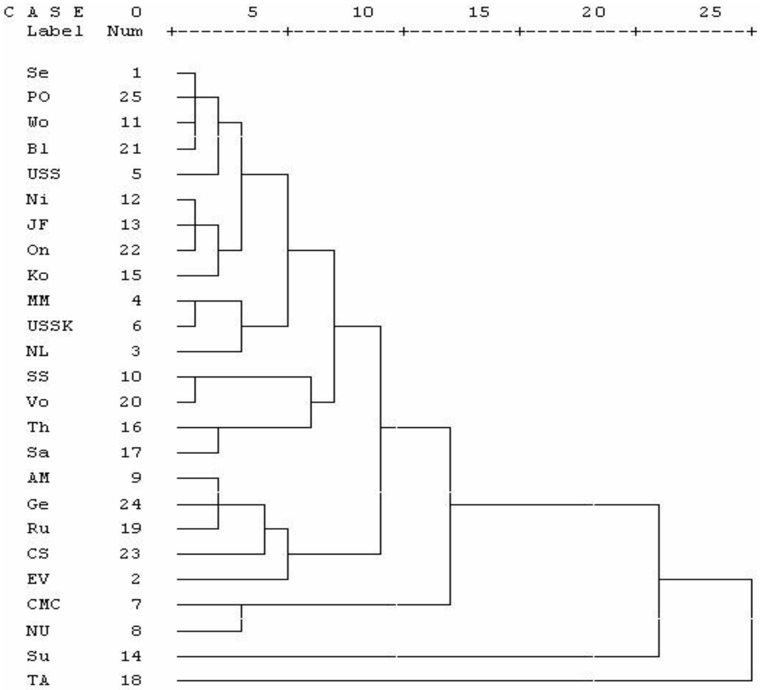
| Comp | Cluster membership | Distance from cluster centre | Comp | Cluster membership | Distance from cluster centre |
|---|---|---|---|---|---|
| Se | 1 | 0.38 | Su | 2 | 0.66 |
| EV | 2 | 0.35 | Ko | 2 | 0.35 |
| NL | 3 | 0.00 | Th | 2 | 0.23 |
| MM | 4 | 0.00 | Sa | 1 | 0.31 |
| USS | 2 | 0.68 | Co | 1 | 0.32 |
| USSK | 1 | 0.44 | TA | 2 | 0.35 |
| CMC | 5 | 0.45 | On | 1 | 0.36 |
| NU | 1 | 0.55 | Ru | 1 | 0.28 |
| Mi | 2 | 0.36 | Vo | 2 | 0.21 |
| Ar | 2 | 0.17 | Bl | 1 | 0.36 |
| SS | 1 | 0.25 | Sm | 2 | 0.37 |
| Wo | 5 | 0.45 | CS | 1 | 0.79 |
| Ni | 2 | 0.26 | Ge | 2 | 0.11 |
| JF | 2 | 0.46 | PO | 2 | 0.37 |

*Source:* Own computation.

If we compare the obtained result with the MDS for the year 2003 can not find such a clear match, as in "full" number of FPU. MDS shows three companies of three clusters in one place Bl (1), Po (2), CMC (5). There is also "collision" of CS (1) with companies from the second cluster. On 2007 figures (14 FPU) we will use different type of cluster analysis, Hierarchical Cluster Analysis (Cluster Method – Between Groups linkage, Measure – Squared Euclidean Distance). The HCA is a distinct process of formation of clusters. First, each object forms a separate cluster. Finally, all the objects merged into one cluster. According the shape of dendrogram we can decide on the best number of clusters.

S c h e m e 1

**Dendrogram**

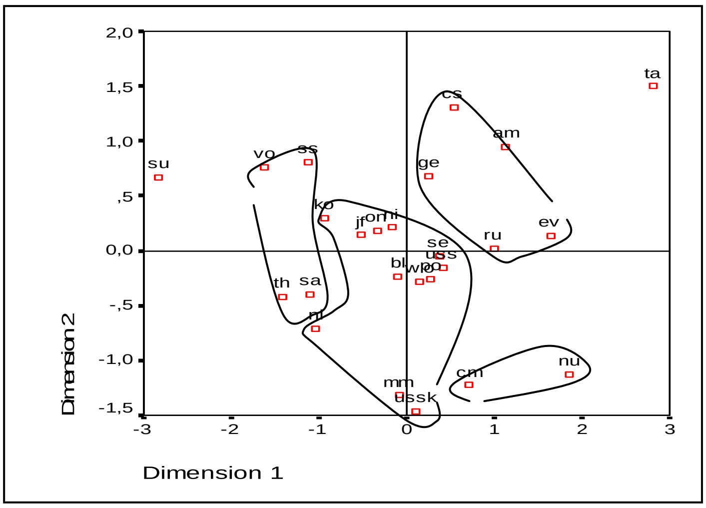



*Source:* Own construction.

F i g u r e 4

**Hierarchical Cluster Analysis versus MDS**

**2007**




*Source:* Own construction.

If we want to have an adequate number of clusters (e.g. 6), then we get the following composition of groups:

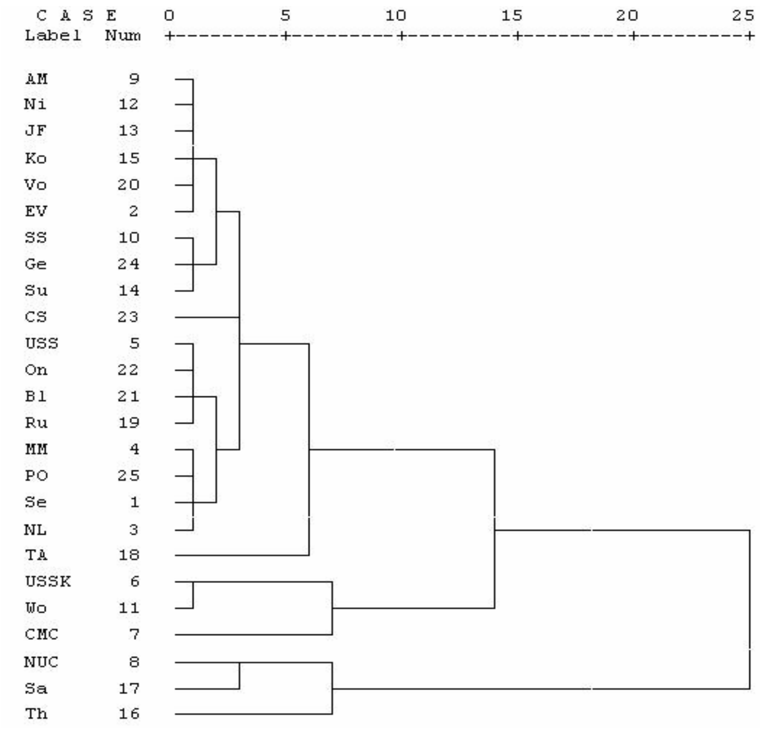
- Severstal to Novolipeck (12 companies),
- SSAB to Salzgitter (4 companies),
- ArcelorMittal to EVRAZ (5 companies),
- CMC and Nucor (2 companies),
- Sumitomo,
- TATA.

If we show this distribution in the output of MDS, we can say that both methods give us very similar results.

We repeat Hierarchical Cluster Analysis with 4 selected variables (L2, OAKT, CZAD, ROE).

S c h e m e 2

**Dendrogram for Reduced Group**




*Source:* Own construction.

Unlike the dendrogram for 14 FPU again we have less consistent results with MDS and a larger number of clusters consisting of one company. One possibility is for example (9, 1, 8, 1, 2, 1, 1, 1, 1).

## Conclusions

Using multivariate statistical methods provides a considerable help in assessing the status of industry and business in it. It facilitates and accelerates the processing of large amounts of data, allows reduction in the number of dimensions of data and thus easier using of other analytical procedures. Moreover, outcomes achieved by mentioned methods are illustrative and easy to interpret. The selected group of FPU (L2, OAKT, CZAD, ROE) sufficiently reflects the financial situation in the industry at selected time.

## References

ALTMAN, E. I. (1968): Financial Ratios, Discriminant Analysis and the Predictions of Corporate Bankruptcy. The Journal of Finance, *XXIII*, No. 4, pp. 589 – 609.

AZIZ, M. A. – DAR, H. A. (2006): Predicting Corporate Bankruptcy: Where We Stand? Corporate Governance, *6*, No. 1, pp, 18 – 33.

BEAVER, W. H. (1968): Alternative Accounting Measures as Predictor of Failure. The Accounting Review, January, pp. 113 – 122.

DeSARBO, W. S. – GREWAL, R. – HWANG, H. – WANG, Q. (2008): The Simultaneous Identification of Strategic/Performance Groups and Underlying Dimensions for Assessing an Industry's Competitive Structure. Journal of Modelling in Management, *3*, No. 3, pp. 220 – 248.

DILLON, W. R. – GOLDSTEIN, M. (1984): Multivariate Analysis – Methods and Applications. New York: John Wiley & Sons. ISBN 0-471-08317-8.

EKLUND, T. (2004): The Self-Organizing Map in Financial Benchmarking. [TUCS Dissertations No. 56.] Ǎbo: Turku Centre for Computer Science.

HEBÁK, P. a kol. (2005): Vícerozměrné statistické metody (3). Praha: Informatorium. ISBN 80-7333-039-3.

MARCUS, P. F. – KIRSIS, K. M. (2006): World Steel Dynamics – Financial Dynamics of International Steelmakers. New Jersey: World Steel Dynamics Inc.

MELOUN, M. – MILITKÝ, J. – HILL, M. (2005): Počítačova analýza vícerozměrných dat v příkladech. Praha: Academia.

Metal Bulletin Directories (2007): Iron& Steel Works of the World 2007 on CD-ROM. 7 th Edition. London: World Steel Dynamics Inc. ISBN 1-904333-29-X.

MOLINERO, M. C. – EZZAMEL, M. (1991): Multidimensional Scaling Applied to Corporate Failure. Manchester: University of Manchester.

STANKOVIČOVÁ, I. – VOJTKOVÁ, M. (2007): Viacrozmerné štatistické metódy s aplikáciami. Bratislava: Iura Edition. ISBN 978-80-8078-152-1.

TRILL, B. – RABIDOUX, R. – PESI, J. A. (2008): Predicting Bankruptcy in Iron and Steel Mills Industry, Advances in Accounting. Finance and Economics, *1*, No. 2, pp. 117 – 122.

***Used websites:***

<www.arcelormittal.com>, <www.bluescopesteel.com>, <www.cmcsg.com>, <www.corusgroup.com>, <www.csc.com.tw>, <www.csn.com.br>, <www.evraz.com>, <www.gerdau.com>, <www.mmk.ru>, <www.nlmksteel.com>, <www.nucor.com>, <www.onesteel.com>, <www.posco.com>, <www.ruukki.com>, <www.salzgitter-ag.de>, <www.severstal.com>, <www.smorgonsteel.com.au>, <www.ssab.se>, <www.tatasteel.com>, <www.thyssenkrupp.com>, <www.usske.sk>, <www.ussteel.com>, <www.voestalpine.com>, <www.worthingtonindustries.com>.