Multivariate statistical analysis in determining the enterprises' behavior with resources

Sergey Kravchenko¹,²,∗, and Mariia Pankova²

¹Institute for International Cooperation Development, 24-26 Kazimierza Wielkiego Str, 61-863 Poznan, Poland
²Enterprise Economics Department, SHEE “Donetsk National Technical University”, 2 Shybankova Square, 85300 Pokrovsk, Donetsk region, Ukraine

Abstract. The threefold concept of sustainable development is largely based on the optimal use of limited resources. In this case, the development of an effective toolkit for the conscious influence of business entities to the effectiveness of using its resource base deserves special attention. The study attempts to substantiate the existence of several basic (characteristic) enterprises types in terms of resources usage results and to identify the key parameters that form a certain style of behavior. This approach includes the methods of multivariate statistical analysis (cluster and discriminant). The results obtained make it possible to assert the existence of four basic enterprises types with different styles of resource management, which affects the level of their resource security. Special attention was paid to identifying indicators that have the greatest separation power. In addition, the classification functions were built, and their quality was confirmed, which makes it possible to model the situation of targeted impact on individual indicators (forecasting the future situation for correction). The research results should contribute to an increase in the degree of the scientific validity recommendations for business entities on the rational use of their resource base (effective regulating the level of resource security).

1 Introduction

In modern aggravated competition conditions, the state of each enterprise largely depends on the nature of the resources' usage, that is, the ability to effectively use and develop its own resource base. At the same time, under the influence of numerous factors, the situation in all areas is constantly changing, objectively causing the need for various regulatory measures [1,2].

One of the main regulation aspects of any process is the object's current state analysis, which allows identifying weaknesses, the reasons for their occurrence and the direction of the necessary impact (correction) [3-5]. Definitely, each enterprise, having a set of certain operating features, in the course of such an analysis will receive its own individualized

∗ Corresponding author: sergey.iv.kravchenko@gmail.com
results. However, from a scientific point of view, general features and trends are of greater interest [6-73].

Given the above, to develop an effective toolkit for conscious influence on the effectiveness of using one's own resource base, the study made an attempt to substantiate the existence of several basic (generalized) enterprises' types based on their resources' management results, as well as to identify the key parameters that largely shape the style of behavior and thus, they are a priority for purposeful correction.

2 Methods

In the current study, the methods of multivariate statistical analysis (cluster and discriminant) were used: (1) cluster analysis – for dividing a certain set of enterprises into basic types; (2) discriminant analysis – for determining the variables with the greatest separation power and the formation of classification functions that can determine the enterprise's type by the nature of its behavior with resources.

The combination of these methods was carried out by using in the discriminant analysis as a dependent variable the enterprise's type, which was previously determined by cluster analysis, and as independent variables – 32 input indicators of personnel, financial, and production enterprise's subsystems.

3 Results

For the enterprises typologization (clustering) by the resource security level, the activities of 81 enterprises were analyzed, which belong to various sectors of the Ukrainian economy, are characterized by different sizes and functioning efficiency, as well as territorially located in different regions. For this purpose, three groups of indicators have been formed that characterize the financial condition, the level of personnel, material and technical security of their activities – that is, reflect the state and efficiency of the resources' main types use. In total, for the analysis 32 indicators, which represent the average value over the five years (before the COVID-19 coronavirus pandemic), were used.

The clustering of the study sample was performed by Ward's method of combining, which minimizes the intragroup variance for any two (hypothetical) clusters that can be formed at each step. The squared Euclidean distances were chosen as a measure of distance. Figure 1 shows the results of the hierarchical classification.

According to the data obtained, it is possible to distinguish from 4 to 6 clusters, which are not equal in objects number. After determining the number of possible clusters, the clustering algorithm, the k-means method, was implemented. The obtained results confirmed the feasibility of dividing the studied population into 4 clusters.

The first cluster included 8 enterprises, the second and the third – 20 for each one's, the fourth cluster included 33 enterprises. Figure 2 shows the cluster centers in a standardized dimension for each type of enterprise.

Thus, it can be argued that there are four basic enterprises types with different resource management styles, which affect the level of their resource security.

Based on the standardized values of the five-year average of those indicators that have the greatest scope beyond the clusters, and their absolute values, the enterprises' distinctive features of each basic type are revealed. These features are the consequences of their different typical behavior in resource management:

Type1 – predominantly high-profitable enterprises, with a large prevalence of current assets in the property structure and a low level of labor productivity;
**Fig. 1.** The findings of the cluster analysis (hierarchical tree diagram for 81 cases).

Source: own computations.

**Fig. 2.** Cluster centers by standardized indicators.

Source: own computations.

Type 2 – predominantly medium-profitable enterprises with equal shares of current and non-current assets and a high level of labor productivity;

Type 3 – predominantly medium-unprofitable enterprises with a large prevalence of current assets in the property structure and a high level of labor productivity;

Type 4 – predominantly high-unprofitable enterprises with a slight prevalence of current assets in the property structure and a low level of labor productivity.

According to study objectives, the next stage founds the variables that had the greatest impact on the distribution of enterprises by type. The tools of multiple discriminant analysis used for this purpose allow studying the differences between two or more groups of objects for several variables simultaneously.

The discriminant analysis is performed by the Forward stepwise method, according to which all variables are analyzed at each step, and the one that makes the greatest contribution to the difference between the populations is found. This variable is included in the model at this stage and the process proceeds to the next step. The procedure ends when all variables that have an $F$-value greater than indicated will include in the model [74,75].

The nature of discrimination is investigated by canonical analysis, that is, the contribution of each discriminant function to the distribution between types of enterprises is determined. Table 1 provides information on the number of constructed discriminant functions (roots) and descriptive statistics indicating their information content.

**Table 1.** The findings of the discriminant analysis.

<table>
<thead>
<tr>
<th>Roots Removed</th>
<th>Eigen-value</th>
<th>Canonical correlation</th>
<th>% of Variance</th>
<th>Wilks’ Lambda</th>
<th>Chi-square</th>
<th>Degrees of freedom</th>
<th>p-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.02</td>
<td>0.90</td>
<td>45.00</td>
<td>0.02</td>
<td>271.52</td>
<td>60</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>3.18</td>
<td>0.87</td>
<td>35.50</td>
<td>0.09</td>
<td>163.36</td>
<td>38</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>1.74</td>
<td>0.80</td>
<td>19.50</td>
<td>0.37</td>
<td>67.56</td>
<td>18</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Source: own computations.

Table 1 shows that the first discriminant function explains 45.0% of the initial variability, the second – 35.5%, the third – 19.5%. Another characteristic that allows us to evaluate the utility of the discriminant function is the canonical correlation coefficient, the maximum value of which is one. The larger the value of the canonical $R$, the better the separation power of the discriminant function, that is, all constructed functions have high discriminant power. Tests of discriminant functions on the criteria of Wilks’ Lambda, $\chi^2$-Pearson, and $p$-level confirmed their significance, that is, dividing the studied set of enterprises into four types is expedient.

The determination of the predictors that have the greatest value in the intergroup differences is based on a comparison of the absolute values of the standardized canonical discriminant function coefficients (Table 2).

Table 2 shows the standardized coefficients that indicate the relative importance of all 20 selected predictors in determining the enterprise's type for its resource management. The coefficients with the highest absolute values (highlighted in bold italics) are the variables with the greatest discriminatory power. In general, labor productivity has the greatest power (in terms of the absolute values amount – 4.17), while the share of fixed assets in assets is the lowest one (0.48).

The obtained results confirm that the classification is correct, and the selected types of enterprises differ significantly from each other. Thus, using the selected most significant indicators, classification functions that have the following form were built:
\[ DF = b_1 x_1 + \cdots + b_i x_i + \cdots + b_p x_p + C \]  

where \( DF \) is discriminant function value; \( x_i \) is the numerical value of the \( i \)-th variable (predictor); \( b_i \) is contribution of the \( i \)-th variable to the function's value; \( p \) is number of variables; \( C \) is constant.

### Table 2. Standardized canonical discriminant function coefficients.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Function</th>
<th>The amount of the absolute values of standardized coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Labor productivity</td>
<td>1.23</td>
<td>-2.08</td>
</tr>
<tr>
<td>Current and non-current assets ratio</td>
<td>-1.13</td>
<td>2.07</td>
</tr>
<tr>
<td>Product profitability</td>
<td>-0.86</td>
<td>1.57</td>
</tr>
<tr>
<td>Fixed assets turnover ratio</td>
<td>1.18</td>
<td>-1.81</td>
</tr>
<tr>
<td>Average monthly wages</td>
<td>-1.22</td>
<td>1.17</td>
</tr>
<tr>
<td>Wages to turnover ratio</td>
<td>-0.62</td>
<td>1.86</td>
</tr>
<tr>
<td>Return on sales</td>
<td>0.81</td>
<td>-0.91</td>
</tr>
<tr>
<td>Assets turnover ratio</td>
<td>-0.50</td>
<td>1.13</td>
</tr>
<tr>
<td>The share of revolving production capital in current assets</td>
<td>0.43</td>
<td>-0.57</td>
</tr>
<tr>
<td>Material intensity of products</td>
<td>-0.18</td>
<td>1.22</td>
</tr>
<tr>
<td>Return on assets</td>
<td>0.08</td>
<td>-0.79</td>
</tr>
<tr>
<td>Asset coverage ratio</td>
<td>-1.09</td>
<td>-0.29</td>
</tr>
<tr>
<td>Accounts payable turnover ratio</td>
<td>-0.70</td>
<td>0.24</td>
</tr>
<tr>
<td>Cash ratio</td>
<td>0.87</td>
<td>0.17</td>
</tr>
<tr>
<td>The share of wages arrears in the debt capital</td>
<td>-0.57</td>
<td>-0.27</td>
</tr>
<tr>
<td>Equity ratio</td>
<td>0.09</td>
<td>0.38</td>
</tr>
<tr>
<td>The share of wages in the production cost</td>
<td>-0.01</td>
<td>0.55</td>
</tr>
<tr>
<td>Long-term debt to capitalization ratio</td>
<td>-0.03</td>
<td>0.45</td>
</tr>
<tr>
<td>Depreciation of fixed assets</td>
<td>0.05</td>
<td>0.43</td>
</tr>
<tr>
<td>The share of fixed assets in assets</td>
<td>-0.16</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

Source: own computations.

In this case, the dependent variable is the type of enterprise, which was previously determined using cluster analysis, and the independent variables are 32 indicators of personnel, financial and production subsystems. Thus, for the four characteristic types of enterprises, the classification functions (based on the previously selected 20 predictors) will be as follows:

\[ Type_i = -137.007 - 0.047x_1 + 9.345x_2 + 86.853x_3 + 80.984x_4 + 7.444x_5 + 7.407x_6 + 1.267x_7 + 3.191x_8 - 8.212x_9 + 29.022x_{10} + 1.553x_{11} - 0.661x_{12} + 0.639x_{13} - 0.63x_{14} + 34.337x_{15} - 7.292x_{16} + 80.716x_{17} + 30.093x_{18} - 59.631x_{19} + 7.452x_{20} \]  

(2)
Type_2 = -78,486 - 0.038x_1 + 7,566x_2 + 71,346x_4 - 122,335x_6 + 6,219x_7 + 10,071x_8 + 1,14x_10 + + 1,214x_1 - 2,335x_13 + 25,346x_14 + 0,543x_15 - 0,556x_22 + 0,538x_23 - 0,616x_25 + 28,021x_26 - (3)

Type_3 = -117,461 - 0.045x_1 + 8,444x_2 + 91,072x_4 - 207,079x_6 + 7,642x_7 + 12,172x_8 + 1,611x_10 + + 0,472x_1 - 0,672x_13 + 30,441x_14 + 0,723x_15 - 0,625x_22 + 0,653x_23 - 0,825x_25 + 39,284x_26 - (4)

Type_4 = -81,817 - 0.037x_1 + 6,919x_2 + 82,2x_4 - 129,63x_6 + 6,269x_7 + 7,698x_8 + 1,044x_10 + + 0,753x_11 - 1,212x_13 + 23,168x_14 + 0,598x_15 - 0,479x_22 + 0,465x_23 - 0,522x_25 + 33,675x_26 - (5)

Company will belong to the type for which the calculated value will be the maximum.

It should be noted that the use of the obtained discriminant functions made it possible to correctly classify 96.3% of the analyzed enterprises (Table 3).

<table>
<thead>
<tr>
<th>Observed types</th>
<th>Percent Correct</th>
<th>Predicted types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1</td>
<td>100,0</td>
<td>8</td>
</tr>
<tr>
<td>Type 2</td>
<td>95,0</td>
<td>0</td>
</tr>
<tr>
<td>Type 3</td>
<td>95,0</td>
<td>0</td>
</tr>
<tr>
<td>Type 4</td>
<td>97,0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>96,3</td>
<td>8</td>
</tr>
</tbody>
</table>

A case is considered correctly classified if, according to the discriminant function, it is included in the group to which it actually belongs. In the study, the classification error is only 3.7%, which indicates the high efficiency of the constructed tools.

The obtained classification functions can be used to identify the type of enterprises that were not involved in the study, and/or to make a forecast for future periods. Accordingly, it is also possible to implement a targeted impact on individual indicators to ensure the resource security of the enterprise.

4 Conclusions

In order to develop effective toolkit for the conscious influence of the enterprise to the effectiveness of using its own resource base in the study:

(1) four basic types of enterprises in terms of resource management, which can be used to analyze the individual characteristics of an individual business entity's behavior in the context of general and inter-cluster trends have been identified and described;

(2) the predictors that have the greatest separation power in terms of the enterprises' typology have been identified that makes it possible to prioritize adjustment measures the resource management style (the transition of an enterprise from one cluster to another);

(3) classification functions have been built and confirmed the its quality, based on which it becomes possible to carry out modeling of a targeted impact on individual indicators (forecasting a future situation for adjustment).

The multivariate statistical analysis considers individual variables as well as the interrelationships between them that allows obtaining more accurate information about the
key regulatory parameters. The findings make it possible to improve the quality of management decisions to ensure resource security of business entities.

References
19. R. Miskiewicz, Energies, 13(22), 6106 (2020)
21. R. Miśkiewicz, R. Wolniak, Sustainability, 12(14), 5776 (2020)
23. O. Prokopenko, R. Miskiewicz, Entrepreneurship and Sustainability Issues, 8(2), 269-284 (2020)
26. H. Dźwigoł, Virtual Economics, **4(1)**, 98-117 (2021)
33. V. Boiko, A. Kwilinski, M. Misiuk, L. Boiko, Economic Annals-XXI, **175(1-2)**, 68-72 (2019)
34. S. Bogachov, A. Kwilinski, B. Miethlich, V. Bartosova, A. Gurnak, Entrepreneurship and Sustainability Issues, **8(2)**, 487-499 (2020)
35. O. Chygryn, Y. Bilan, A. Kwilinski, Marketing and Management of Innovations, **3**, 356-368 (2020)
37. V.V. Dementyev, A. Kwilinski, Journal of Institutional Studies, **12(1)**, 100-116 (2020)
38. V. Koibichuk, N. Ostrovskva, F. Kashiyeva, A. Kwilinski, Marketing and Management of Innovations, **1**, 253-265 (2021)
40. A. Kuzior, A. Kwilinski, V. Tkachenko, V. Entrepreneurship and Sustainability, **7(2)**, 1353-1376 (2019)
41. A. Kuzior, A. Kwilinski, I. Hroznyi, Energies, **14(9)**, 2572 (2021)
42. A. Kuzior, O. Lyulyov, T. Pimonenko, A. Kwilinski, D. Krawczyk, Sustainability, **13(15)**, 8145 (2021)
43. A. Kwilinski, Virtual Economics, **1(1)**, 7-25 (2018)
44. A. Kwilinski, Marketing and Management of Innovations, **4**, 116-128 (2018)
45. A. Kwilinski, Academy of Accounting and Financial Studies Journal, **23(S12)**, 1-6 (2019)
52. A. Kwilinski, Y. Zaloznova, N. Trushkina, N. Rynkevych, E3S Web of Conferences, 168, 00031 (2020)
59. O. Lyulyov, T. Pimonenko, A. Kwilinski, Y. Us, E3S Web of Conferences, 250, 03006 (2021)
64. V. Tkachenko, A. Kwilinski, I. Tkachenko, P. Puzyrova, Marketing and Management of Innovations, 2, 228-238 (2019)
70. J. Polcyn, Management, 22(2), 171-186 (2018)
71. J. Polcyn, Sustainability, 13(15), 6846 (2021)
72. A. Zielińska-Chmielewska, A. Olszańska, J. Kaźmierczyk, E.V. Andrianova, 
Agronomy, 11(2), 299 (2021)
73. R. Abazov, Herald of Journalism, 58(4), 34-43 (2021)
74. N. Bureeva, Multivariate statistical analysis using “STATISTICA” (Nizhny Novgorod, 2007)
75. K. Kocisova, M. Misankova, Discriminant analysis as a tool for forecasting 
company’s financial health, in Proceedings of the 2-dn International Scientific 
conference “Contemporary Issues in Business, Management and Education 2013”, 110, 
1148-1157 (2014)