Methods frontier analysis as a tool for comparative studies

Natalya Poluektova
Academy of Bioresources and Nature Management, V.I. Vernadsky Crimean Federal University
Department of System Analysis and Informatization
Simferopol

n-poluektova@yandex.ru

Abstract **— The article provides definitions and approaches to the classification of methods and models of comparative analysis. The work represents disadvantages of the most common methods are given. Methods of frontier analysis we proposes to use as tools for comparative studies. These techniques allow evaluate the deviation of the effectiveness of the test object or phenomenon with respect to a boundary that generates based on the analysis of similar objects. This gives grounds to judge the possibility of their use in comparative analysis.**

This paper describes the essence of the nonparametric method of frontier analysis. DEA (Data Environment Analyzes). The focus of the method is the evaluation of the comparative technical efficiency of the objects.

We have considered an example of the use of the DEA method for a comparative study of the problem of the effectiveness of investments in various branches of production in the region. The results of the application of the approach make it possible to judge the prospects of its application as one of the elements of the comparative analysis of complex phenomena. The use of DEA in this task allows us to overcome the problems of lack of accurate quantitative data and accounting for the qualitative results of the processes that arise when using other methods of comparative analysis.

Keywords — comparative analysis, frontier analysis methods, DEA, innovation effectiveness

I. INTRODUCTION

According to C. Pickvance, comparative analysis is a set of theoretical concepts and practical methods the purpose of which is to explain and better understand the causes of the processes or phenomena involved in creating an event, function, or relationship, usually by combining variations of an explanatory variable or variables [1].

Comparative research has a long history of development. In modern practice of economic research, they moved from sociology. In the beginning of the 20th century, M. Weber singled out the application of quantitative comparative methods from traditional methods of historical analysis in scientific practice. [2]. At present, such research methods are becoming more common in political science and economics due to globalization, technological progress and the development of information and communication technologies.

C. Tilli [3] and R. Azarian [4] in their works propose the following classification of approaches in comparative analysis.

a. Individualization of contrasts comparison. This approach involves the study "in depth" that is a detailed study of a small number of cases with the most complete study of their characteristics. This approach is not typically comparative, although it is extremely useful for studying specific aspects of problems.

б. Universal comparison. This approach involves the establishment of the essence of the investigated processes, phenomena or objects, by correlating them with other similar elements. This approach is applicable to the study of fundamental phenomena, paradigms or concepts.

в. Comparison with the search for variations establishes the principle of a change in the nature or intensity of a phenomenon in different conditions (in different regions, in different time periods). That is, by comparing the numerous forms of a single phenomenon, it is possible to detect important differences between instances and establish a standard for changing the nature or intensity of these changes.

Д. Comprehensive comparison of phenomena in different conditions within a single system. This method allows us to identify the influence of the environment on the change in the characteristics of the phenomenon, to explain the features of the behavior of objects by the characteristics of their relationship to the system as a whole.

Depending on the method and source of information about the object of study, all methods are conventionally divides into the following groups. [5]:

1) Theoretical methods are methods based on rational knowledge and logical inference procedures. These methods are associated with the comprehension and processing of the material obtained empirically.

2) Logical-intuitive methods are the means of cognition based on the use of knowledge, experience, skills, logical methods, intuition by a person.

3) Empirical methods - methods of knowledge of reality at the level of experience.

4) Complex-combined methods are research methods that combine several methods belonging to different classification groups.

In addition, comparative research methods include those that focus on qualitative and quantitative tools.

In the qualitative methodologies of comparative analysis, various expert approaches, methods of multi-criteria optimization, and approaches based on the use of fuzzy sets are actively used. The use of these approaches determines possible problems with the use of qualitative methods of comparative analysis - the omission of important details and nuances with the wrong choice or application of rating scales, the biasedness of expert opinions, etc.

For quantitative comparative analysis, multifactorial regression and cluster analysis are most often used. They are also not without certain flaws, some of which, when applied to comparative analysis, are analyzed in [6].

In our opinion, methods of frontier analysis could have been used in the study of complex problems in the comparative analysis along with traditional tools and methods.

The scope of these methods is the evaluation of the effectiveness of complex objects functioning. Usually such researches study technical efficiency. Technical efficiency is formally determined as follows: it is the ability to use the minimum resource vector for the production of a given release vector or, conversely, to obtain the maximum release vector for a given resource. All known methods of frontier analysis in one way or another offer to determine the deviation of the indicator of the effectiveness of the object under study from some maximum possible efficiency, which is achievable with a given amount of resources or other production factors. Such a frontier builds based on studying the activities of a multitude of similar objects in similar conditions.

All such methods divides into parametric and non-parametric. V ariants of parametric methods are Stochastic Frontier Analysis (SFA), the method Distribution-Free Approach (DFA) and Thick Frontier Approach (TFA). [7]

For example, the SFA Model uses the Cobb-Douglas production function, which represents as follows.

$$\ln(y^i) = \beta X^i + v^i - u^i, \quad i \in [1, N] \qquad (1)$$

where:

y^i - is the resultant performance indicator of an object i,

X^i is a vector whose elements are natural logarithms of the values of efficiency factors for an object i:$X^i=\{1,\ln(x_1^i), \ln(x_2^i),\ldots \ln(x_k^i)\}$,

β is a vector of model parameters that you need to evaluate,

v^i – is a random symmetrically distributed error,

u^i – is a non-symmetrically distributed error, which is a measure of inefficiency.

Calculating, with the help of special statistical methods, the indicator u^i for a set of homogeneous objects makes it possible to identify the limit of efficiency, and the deviation from this limit for the object under study.

It is worth noting that in these models, inefficiency is a random variable, which means that it is possible to find factors that influence its decline. This is an important practical task. Such models applied to the cost function too.

DEA (Data Environment Analyzes) method is most known among non-parametric methods. This paper proposes to consider in more detail the nonparametric method of boundary analysis of DEA.

II. Materials and Methods (Model)

Data Environment Analyzes (DEA) is a method that uses to analyze the effectiveness of various objects and systems. based on the determination of the effectiveness boundary. The ideology of the method use the definition of technical efficiency, when the result of the object's activity compares with the maximum possible result for a given amount of resources.

M. Farrell [8] described the basics of the method. He proposed an approach to assessing the technical efficiency of single-input and single-output systems. However, the method became widespread only after the creation in 1978 in the work of A. Charnes et al. of the base model CCR and further research [9, 10], which led to the appearance of multiplicative and additive models of this type. To determine the effectiveness taken to consider DEA technical efficiency as a special ratio of the weighted sum of outputs (results) to the weighted sum of inputs (costs).

$$Z_j = \frac{\sum_{r=1}^{t}(w_r y_{rj})}{\sum_{i=1}^{m}(v_i x_{ij})}, \; j \in [1, n] \qquad (2)$$

where:

y_{rj} – the quantitative assessment of the r-th in t results,

x_{ij} – the quantitative assessment of the i-th type in m cost elements,

w_r – the desired values of weight coefficients by types of results,

v_i – the desired coefficients by type of cost,

n –the number of the objects (j.

Determination of efficiency according to Pareto-Kumpans-Farrell means that the functioning of each of the objects is fully effective if the functioning of other objects does not provide evidence that some inputs or outputs of the evaluated

object improve without deteriorating other inputs or outputs. This expresses by the following relationship

$$Z = \frac{\sum_{r=1}^{t}(w_r y_{ro})}{\sum_{i=1}^{m}(v_i x_{i0})} = \max\{ Z_j = \frac{\sum_{r=1}^{t}(w_r y_{rj})}{\sum_{i=1}^{m}(v_i x_{ij})}, \quad (3)$$
$$j \in [1,n]\}$$

where:

y_{r0} – the quantitative assessment of the r-th in t results for object under study,

x_{i0} – the quantitative assessment of the i-th type in m cost elements for object under study.

An object or process will became effective when the value of technical efficiency reaches a unit value in relation (3). Otherwise, the value of the relation (3), which is less than one, indicates that the functioning of the rest of the objects proves the relative inefficiency of the object or process under study. This leads to the formulation of the objective function:

$$\max Z_0 = \frac{\sum_{r=1}^{t}(w_r y_{r0})}{\sum_{i=1}^{m}(v_i x_{i0})} \quad (4),$$

As well as a system of restrictions that show that no single object can have an efficiency, that is greater than one:

$$\frac{\sum_{r=1}^{t}(w_r y_{rj})}{\sum_{i=1}^{m}(v_i x_{ij})}, j \in [1,n] \quad (5)$$

$v_i, w_r > 0, i \in [1,m], r \in [1,t]$,

which maximize the value of the effectiveness of the object, in relation to the activities of other objects.

The purpose of the method is to find a subset of those points that create a border of efficiency. Objects that are not on the border are inefficient, and the method allows determining the causes of this inefficiency.

The method uses linear programming as a way to achieve an optimal result.

The authors of the study [10] showed that the problem of the optimization above is equivalent to the following form.

$$G_0 = \sum_{r=1}^{t}(w_r y_{ro}) \to \max$$

$$\sum_{i=1}^{m}(v_i x_{i0}) = 1$$

$$\sum_{r=1}^{t}(w_r y_{rj}) - \sum_{i=1}^{m}(v_i x_{ij}) \le 0$$

$$j \in [1,n]$$

$$w_r, v_i > \varepsilon, i \in [1,m], r \in [1,t] \quad (6)$$

This problem solves by known methods for solving linear programming problems.

We propose to consider the application of the method based on the study of the problem of determining the comparative effectiveness of innovations.

III. Results and Discussion

The basis of this study is the classification of innovations, which identifies the following categories.

- Institutional innovations that involve changing policies, standards, regulations, processes, agreements, models, ways of organizing, institutional practices or relationships with other organizations. As a result, a more dynamic environment creates a more efficient production or sales processes.

- Technological innovations consist in the application of new ideas, scientific know-how, new technological methods for the development, production or improvement of products, the reorganization or improvement of production processes.

- Social innovations cause the development or significant improvement of concepts, ideas or organizations and lead to a more complete satisfaction of the social needs of certain groups of individuals or the society as a whole (increased employment, consumption, improved quality of life).

- Marketing innovations are associated with changes in the methods or conditions for marketing or distribution of products.

- Organizational innovations cause changes in the organization's management structure, production structure, business processes of operations or processes and methods of relationships with other stakeholders.

We managed to get data from official Russian Federation statistical reports [11], which indirectly characterize only technological, marketing and organizational innovations. In particular, we were analyzing the four sectors: crop and animal production (with the provision of services in this area) (1), mining (2), processing industry (3), operating in the telecommunications sector (4), with the 2017 data include:

- indicator 1 (I1): innovative goods and services shipped (mln. rub.);

- indicator 2 (I2): costs of technological innovations (mln. rub.);

- indicator 3 (I3): the proportion of organizations implementing organizational innovations (%);

- indicator 4 (I4): the proportion of organizations that implemented marketing innovations (%).

The main data for analysis we summarized in table 1.

TABLE I. STATISTICAL DATA FOR DEA ANALYSIS

	industry (1)	**industry (2)**	**industry (3)**	**industry (4)**
I1	28446	489447.06	2832804.41	74036.18
I2	15806.0	184811.2	610218.09	35378.64
I3	3.2	7.4	13.3	8.8
I4	0.5	0,.	3.0	4.7

To assess how effectively different types of innovations are used, it is proposed to consider the DEA model as a result - innovative goods shipped and services shipped, and the remaining three indicators as costs.

The linear programming task corresponding to model 6 for determining the effectiveness of innovations in the agricultural industry consists of the following expressions.

$$2844w \rightarrow max$$
$$15806v_1+3,2v_2+0,5v_3=1$$
$$28446w-(15806v_1+3,2v_2+0,5v_3)\leq 0$$
$$48446,06w-(184811v_1+7,4v_2+0,3v_3)\leq 0$$
$$2832804,4w-(610218v_1+13,3v_2+3v_3)\leq 0$$
$$74036w-(35378,64v_1+8,8v_2+4,7v_3)\leq 0$$
$$V_i, w_r>0,0001$$

In addition, three similar problems were solving for the three other industries to determine those that form the boundary of efficiency.

The results of solving problems represent the values of the objective function G. (table 2).

TABLE II. RESULTS OF THE DEA METHOD APPLICATION FOR THE RESEARCH OF INNOVATIVE ACTIVITY EFFICIENCY IN SOME BRANCHES OF PRODUCTION OF RUSSIAN FEDERATION

	industry (1)	**industry (2)**	**industry (3)**	**industry (4)**
G	0,38	1	1	0,03

The interpretation of the results is that at the border of efficiency (compared to other sectors) in terms of the ratio of the costs of innovation and sales of innovative products, there are two industries: mining and processing industries (the objective function is equal to one).

In other industries, the same level of efficiency will achieve even with a decrease in input investment by a factor equal to the value of the objective function for this industry.

The results of the analysis also show that innovations in the telecommunications industry are the least effective, although this may be due to the innovativeness of the industry itself and the incorrect assessment of the results of introducing innovation in it.

IV CONCLUSION

The use of traditional comparative analysis tools is complicate by many factors, including:

- difficulty quantifying parameters;
- the complexity of the correlation of parameters available to different nature, etc;

The example given in the paper showed that the frontier analysis toolkit we can be used as a component of the comparative analysis of complex objects or phenomena, in the tasks of evaluating the effectiveness of their activities.

The DEA method does not impose any restrictions on the functional form of the relationship between inputs and outputs and does not require a priori specification of weights for variables corresponding to input and output parameters.

The results of this analysis will also be more effective when used in combination with other methods and organizational practices of comparative research.

In further studies, the method will allow to take into account variables external to the system under consideration - environmental factors. To do this, we need to formulate a linear regression problem in which the dependent variable is the performance indicator obtained when solving the DEA problem and the parameters are environmental variables. Signs with regression coefficients indicate the “direction” of influence: the “plus” sign indicates the positive influence of the environment on efficiency; the “minus” sign indicates the negative influence of the environment. At the same time, we need to conduct mandatory traditional tests to confirm statistical hypotheses.

On the other hand, the results of the effectiveness of a variety of objects we can divide in groups, according to the degree of deviation from the efficiency boundary, and carry out further studies of these clusters.

Along with the considered method of nonparametric analysis in the framework of comparative studies, we should consider in more detail the methods of boundary analysis based on the application of the frontier production function.

References

[1] Pickvance C. “The four varieties of comparative analysis: the case of environmental regulation,” Journal of Housing and the Built Environment, 2005, vol. 16, pp.7-28.

[2] Ragin C., Zaret D. “Theory and Method in Comparative Research: Two Strategies,” Social Forces, Vol. 61, No. 3 (Mar., 1983), pp. 731-754.

[3] Tilly C. “Big structures, large processes, huge comparisons,”. Contemporary Sociology, Vol. 14, No. 6 (Nov., 1985), pp. 693-695

[4] Azarian R. “Potentials and limitations of comparative method in social science,”. International Journal of Humanities and Social Science, 2005, vol.1(4), pp.113-125.

[5] Subal C. Kumbhakar. Stochastic Frontier Analysis [Електронний ресурс]/ Subal C. Kumbhakar, Lovell C.A.K. // Cambridge University Press, 2003. – Available at:: :http://books.google.ru/books?id=wrKDztxLWZ8C&dq=stochastic+frontier+analysis&printsec=frontcover&source=bl&ots=L0KvEYKJ_3&sig=jTH07kcNfF4IT43lwLsNxTbSW_Q&hl=ru&ei=ptIkSq7uGsqQsAbcy

NSABg&sa=X&oi=book_result&ct=result&resnum=3. Reference date: 17.02.2019.

[6] Mishin V.M. Issledovanie sistem upravleniya: Uchebnik dlya vuzov / V.M. Mishin -2- e izd. - M.: YuNITI-DANA, 2012. - 527 s.

[7] Rubinson C., Ragin C.C. "New methods for comparative research?", Capitalisms Compared Comparative Social Research, Volume 24, pp.373–389

[8] Farrell M. J. "The measurement of productive efficiency," Journal of the Royal Statistical Society,1957, Series A, 120. pp. 253–281.

[9] Charnes A. "Measuring the efficiency of decision making units,". European Journal of Operational Research, 1978, № 2., pp. 429–444.

[10] Charnes A. A, Cooper W. W, . Seiford L. M and oth. "Multiplicative model for efficiency analysis", Socio-Economic Planning Sciences, 1982., № 16(5). pp. 223–224

[11] Innovatsii. Federal'naya sluzhba gosudarstvennoi statistiki Available at:: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/science_and_innovations/science/#8. Reference date: 17.02.2019.