Identification of Significant Factors for Improving the Efficiency of Enterprises Based on the Construction of an Efficient Frontier

E.A. Grebenyuk

Doctor of Technical Sciences, Chief Scientist, V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Russia
E-mail: lngrebenk12@yandex.ru

Abstract This study aims at the issues of evaluating the effectiveness of organizations and determining significant factors for its improvement. Effectiveness valuation is measured by the distance of the organization to the efficient production frontier (EPF). We consider two main approaches to the construction of EPF: Data Envelopment Analysis (DEA) and the Stochastic Frontier Analysis (SFA). We analyzed and compared these approaches, applying the developed procedures to evaluate the effectiveness of two groups of organizations from different production sectors (mechanical engineering and oil refining). We conducted 1) building empirical estimates of technological efficiency, 2) identifying factors that contribute to the technological inefficiency of the organization, 3) comparing efficiency factors and inefficiency factors for two groups of enterprises.

Introduction
Estimation of technological efficiency and identification of the main factors for its improvement are the most essential elements of the analysis of the organization's activities. One of the widely used approaches to its estimation is the analysis of the relationship between the resources used and the products produced. Following the concept of technological efficiency, introduced in [1] by Koopmans, we are considering an enterprise’s efficiency as the ability to use the smallest possible combination of resources at a given output level and produce maximum output from a given number of resources [2].

When forming an efficiency estimation, we select the total volume of output, the volume of innovative products, the volume of high-tech products, and the results of scientific activities used to produce the company's products as output objects. So, the set of resources depends on what aspect of the activity we want to analyze: production, scientific, innovative activity. The estimation of technological efficiency is calculated as the distance of an organization to the Efficient Production Frontier (EPF), which is built in the multidimensional space of parameters of the “resources-output” for a group of organizations.

In world practice, there are two main approaches to constructing EPF: the DEA (Data Envelopment Analysis) method [2-3], and the Stochastic Frontier Analysis (SFA) method [4-5]. These approaches are widely used in scientific research, and the objects of assessment can be not only groups of enterprises but various sets of "decision making blocks" [2] with common inputs and outputs. For example, educational institutions, educational programs, etc.

Building adequate estimations of the effectiveness of both the enterprise as a whole and various aspects of its activities: production, innovation, scientific, etc., and identifying factors that affect efficiency are closely related to building models of the production, innovation, and scientific potentials
of the enterprise. We find the factors of inefficiency based on the analysis of the dependences of the levels of development of innovations, technologies, and scientific developments on a certain type of resources. In this study, we examined the following issues: 1) construction of production, innovation, and intellectual potential of high-tech enterprises, 2) identification of inefficiency factors, reducing the potential of an enterprise, and factors contributing to its increase, 3) comparison of two approaches to measuring efficiency estimates.

1. Evaluation of effectiveness by the Data Envelopment Analysis (DEA) method

The definition of technological efficiency was first introduced by Farrell (Farrell-Efficiency) [3] who measured it by the deviation of the observed state of the evaluated object \((x_i, y_i)\) from the boundary (effective) \(\{(x^*, y^*)\}\). The resulting value is defined as the TE efficiency index, \((TE \leq 1)\) Farrell viewed objects with one entrance and one exit. Charnes A. at al. [2] generalized the model to the case of a set of inputs and outputs of a production facility, which they defined as an input-output conversion device or "decision making unit". The method proposed in [2] is called the method of the Data Envelopment Analysis (DEA) and is widely used to measure efficiency.

The DEA method is designed to measure:

1) The economy of the enterprise, which, following [3], we understand as its ability to use the minimum possible number of resources at a given level of outputs.
2) The effectiveness of the enterprise, which we understand as its ability to achieve maximum output for the available set of resources used.

**Problem statement.** Each enterprise \(i\), where \(i = 1, 2, ..., N\), is considered as a transforming element of the resources available to it at the "input": \(x_i = (x_{i1}, ..., x_{ik})\), in the results obtained at the "output" \(y_i = (y_{i1}, ..., y_{ir})\), which characterize the output. An enterprise is considered technologically efficient if the input-output vector \((x_i, y_i)\) lies on the Efficient Production Frontier in the space \(T = \{(x_i, y_i) \mid x_i \text{ realizes } y_i\}\), defined as the set of all realizable sets of vectors \((x_i, y_i)\). By definition, an EPF is a subset of the Pareto-effective points of the set \(T\).

Depending on a calculated estimation of effectiveness: economy or performance, the EPF is constructed in different ways:

- if we consider the economy as the ability to minimize resources for a given release, the EPF includes the points \((x_i, y_i)\) such that the value of any resource \((x_{i1}, ..., x_{ik})\) cannot be reduced without reducing one or more release values \((y_{i1}, ..., y_{ir})\);
- if we consider effectiveness as the ability to maximize output for a given resource, then the EPF includes such points \((x_i, y_i)\) at which the value of none of the outputs \((y_{i1}, ..., y_{ir})\) can be increased without increasing one or more resource values \((x_{i1}, ..., x_{ik})\).

To evaluate the economy of the i-th enterprise, we are looking for an estimate of \(\theta_i\) as a solution to a linear programming problem of the form:

\[
\min \theta_i \\
\text{under conditions} \\
- y_i + \lambda Y \geq 0, \\
\theta x_i - \lambda X \geq 0, \\
\lambda \geq 0, \ 0 \leq \theta \leq 1, i = 1, 2, ..., N, \tag{1}
\]

where \(\theta_i\) is a scalar estimate of the economy of the i-th enterprise, and \(\lambda\) is an \(N \times 1\) vector, \(i = 1, 2, ..., N\). To evaluate the performance of the i-th enterprise, we are looking for an estimate of \(\phi_i\) as a solution to a linear programming problem of the form:

\[
\max \phi_i \\
\text{under conditions}
\]
where \(1/\varphi_i\) is a scalar estimate of the effectiveness of the \(i\)-th enterprise, and \(\lambda\) is an \(N \times 1\) vector, \(i = 1, 2, ..., N\).

If \(\theta_i = 1\), \((\varphi_i = 1)\), then the \(i\)-th object lies on the efficient production frontier. For objects that have \(\theta_i < 1\), \((\varphi_i > 1)\), as a result of solving problems (1) and (2), target values of resources (outputs) can be determined. When these target values are reached, the objects will be on the efficient production frontier, if they are able to provide all output volumes under the given decrease in resources.

As a result of solving linear programming problems (1) and (2), we calculate the estimates of the efficiency of enterprises with constant returns to scale. If the compared enterprises have variable returns to scale, then we need to introduce an added constraint into problems (1) and (2): \(\sum \lambda_i = 1\).

The results of solving problems (1) and (2) include:

1. The set of Pareto - efficient enterprises belonging to the EPF, i.e., those where no resource can be reduced without reducing output (when evaluating economy) or no output can be increased without increasing at least one of the resources (when evaluating performance). The efficiency of these enterprises is equal to 1.

2. The set of inefficient enterprises and the evaluation of technological efficiency for each of them.

3. The magnitude of the excess resource for every enterprise: resource values from a set of \(x_i, i = 1, 2, ..., N\), on which they can be reduced without changing the values of other resources and outputs that will enhance technology efficiency (if we estimate the economy). Similarly, the values of all outputs for each enterprise - the values of outputs from the set \(y_i, i = 1, 2, ..., N\), by which they can be increased without changing the values of other outputs and resources, which will increase the technological efficiency of the enterprise (if we estimate the effectiveness).

4. A set of reference enterprises for each of the inefficient ones (closest found to it on the effective hypersurface and, so, close in structure to the analyzed enterprise).

5. A method for increasing efficiency (reaching an effective hypersurface) by finding the amount of reduction in the number of resources spent when evaluating the efficiency of an inefficient enterprise or the amount of increase in output (when evaluating its performance).

2. Estimation of efficiency using the SFA

2.1 The Stochastic Frontier Analysis method

**Problem Statement 1.** Let \(\{\Pi_1, \Pi_2, ..., \Pi_N\}\) be a set of \(N\) objects with \(k\) inputs \(x_i = (x_{i1}, ..., x_{ik})\), and an input \(y_i\). Depending on the properties of the analyzed objects, \(i = 1, 2, ..., N\), for the specification of its production function we choose either the Cobb – Douglas function:

\[
ln y_i = \beta_0 + \sum_{j=1}^{k} \beta_j ln x_{ij} + v_i - u_i,
\]

or the translogarithmic function:

\[
ln y_i = \beta_0 + \sum_{j=1}^{k} \beta_j ln x_{ij} + \frac{1}{2} \sum_{j=1}^{k} \sum_{\ell=j+1}^{k} \beta_{j\ell} ln x_{ij} ln x_{i\ell} + v_i - u_i,
\]

where \(\beta_0, \beta_1, ..., \beta_k, \beta_{11}, ..., \beta_{kk}\) are model coefficients; \(v_e\) is a stochastic error, depending on uncontrolled external parameters, which is described by a normal distribution \(v_i \sim N(0, \sigma_v^2)\); \(u_i\) is a technological inefficiency, independent of the random component \(v_i\):

\[u_i \approx N^+(\mu, \sigma_u^2)\]
where $N^+(\mu, \sigma^2_\mu)$ is the normal distribution truncated at zero with variance $\sigma^2_\mu$, $\mu > 0$ being its mathematical expectation, the value of which figures out the average effect of the combination of all factors that reduce the output efficiency. To solve the problem, you need:

a) to evaluate model parameters: the coefficients $\beta_0, \beta_1, ..., \beta_k$, the variance of the random part $\sigma^2_\mu$, the variance of the inefficiency $\sigma^2_\delta$, the value of the mathematical expectation of the cumulative effect of inefficiency factor $\mu$,

b) to build EPF that defines the enterprise's potential: a production function in the absence of inefficiency.

**Problem Statement 2.** Under the conditions of Problem 1, model (3b) is replaced by the model:

$$u_i \approx N^+(g_i, \sigma^2_\mu), \hspace{1cm} (3c)$$

where $g_i = Z_i \delta, Z_i = (1, Z_{i1}, ..., Z_{im})$ is a vector of values of $m$ factors of the inefficiency of the $i$-th enterprise, $\delta = (\delta_0, \delta_1, ..., \delta_m)$ is the vector of coefficients of the inefficiency function, $\sigma^2_\mu$ is the variance of inefficiency. To solve the problem, you need:

a) to estimate the model parameters: coefficients $\beta_0, \beta_1, ..., \beta_k$, the variance of the random part $\sigma^2_\mu$, the variance of inefficiency $\sigma^2_\delta$, the coefficients of the inefficiency function $\delta = (\delta_0, \delta_1, ..., \delta_m)$;

b) and c) are the same as in Problem 1.

### 2.2 Estimation of the potential of the enterprise

**Definition [6].** Frontier production potential is the volume of production that is possible with fixed volumes of the main production factors under conditions of the random influence of uncertainty factors corresponding to the observed production results and the absence of inefficiency factors.

We will examine the following tasks:

- construction of estimates of the innovative, production, and scientific potential of an enterprise, considering random impacts and the cumulative effect of accompanying factors that reduce the real potential,

- construction of estimates of the innovative, production, and scientific potential of an enterprise, considering random factors and finding factors of inefficiency, the elimination of which will increase the estimated potential.

The definitions of various potentials of an enterprise differ only in the choice of output parameters and production factors. When building production potential, we estimate the value of output by the volume of products produced, resources by the volume of labor costs, the cost of raw materials and equipment. The choice of inefficiency factors depends on the properties of the group of evaluated objects and the conditions in which they work. For example, we can choose such parameters as the age of the company, its size, the share of foreign capital, the volume of exports, the volume of research and development, the volume of government subsidies, etc.

The innovative potential determines the ability of an enterprise to produce innovative products. For its construction, we can use the following parameters:

- the output is income from innovative products,

- resources are the volume of labor costs, the cost of innovative equipment, external and internal costs of research and development,

- factors that change the efficiency are the volume of export supplies, wear and tear of equipment, the level of personnel qualifications.

Scientific and technical potential determines the ability of an enterprise to use new technical and scientific ideas and technologies in its activities, to conduct their technological and design study and implementation.

In the model of scientific and technical potential, we estimate the output by the volume of science-intensive products, and in the composition of production factors, in addition to labor and capital, we considered intellectual capital. In [7], the authors proposed to choose one of the following indicators as the values of intellectual capital: the volume of intangible assets of an enterprise, the value of the
difference between capitalization and the book value of tangible assets, and Tobin's Q-indicator. As a result of the analysis of empirical data for the analysis of intellectual potential based on experimental data, we have chosen the number of intangible assets of the enterprise.

The scientific potential is determined by the level of fundamental and applied scientific research conducted to create new knowledge about the basic laws of the world and/or high-tech products and technologies. The estimation of scientific potential depends on the profile of the group of organizations that we want to evaluate. For educational institutions, the evaluation parameters can be the volume of research, the number of publications of results in indexed journals and conference materials, the level of qualification of employees and graduates. In the case of corporations, publications are usually not available in the open press, and the parameters that characterize the scientific level of the enterprise team can be the volume of developments of high-tech products, new scientific developments and patents, the number of implementations using them, and the economic effect of implementation. As factors of inefficiency, researchers usually use data on the research team: the number and level of education of researchers, the presence of an academic degree, age, etc.

3. Estimation efficiency and identification of inefficiency factors based on empirical data using DEA and SFA methods

3.1 Goals of the experiment and tasks to be solved

The main goals of the experiment were to analyze the possibilities of identifying the factors of inefficiency in various areas of the organization's activities, the elimination of which will increase their effectiveness, and to compare the effectiveness estimates obtained by the DEA and SFA methods. When conducting an experiment on real data, we solved the following problems:

1) Measurement of production, innovation, scientific and technical potentials with unidentified efficiency factors.
2) Identification of removable factors of inefficiency that increase measurable production, innovation, scientific and technical potentials.
3) Comparison of the values of technological efficiency for groups of enterprises obtained by the DEA and SFA methods.

Using the stochastic frontier method for two groups of enterprises in different industries, we evaluated the production potential with unidentifiable efficiency factors (Problem Statement 1, (3) - (3b)) and the production potential with identifiable efficiency factors (Problem Statement 2, (3) - (3c)). For one group of companies, we measured and analyzed the inefficiency factors of innovation and scientific and technical potentials.

3.2 Empirical analysis of production potential and identification of factors of inefficiency.

To conduct an empirical analysis, we used data for two groups of enterprises: sixty-two enterprises of the mechanical engineering industry (aviation industry) and fifty-two enterprises of the oil refining industry. Individual enterprises within each group are homogeneous by type of activity and are part of vertically integrated companies.

To build the production function, we selected the following indicators: the volume of work performed, and products produced (output), material costs, and the number of production personnel. We selected the following indicators as the main factors of inefficiency: the share of public funds raised, the volume of export deliveries, the volume of research and development, and the level of staff qualification.

The values of production potential, output logarithm, and measured technological efficiency for two groups of enterprises are the results of solving problem 1 and problem 2. These values are shown in Fig.1 and Fig.2, respectively. We show graphical characteristics of the efficiency of a group of mechanical engineering enterprises in the left part of each of the figures, and graphs of characteristics of a group of oil refining enterprises on the right.

All graphs have the potential and output logarithm values on the left axis and the efficiency values on the right axis.
When calculating using the model (3) - (3c), which specifies the impact of inefficiency factors, the values of the potential and production function converge and the efficiency increases, which says that these dependencies describe the model more accurately. The efficiency of oil refining enterprises is on average higher than that of mechanical engineering enterprises, however, with partial elimination of inefficiency factors, it increases slightly. The efficiency of machine-building enterprises in the case of solving the problem (3) - (3c) with identified inefficiency factors is comparable to the efficiency of oil refining enterprises.

The values of statistical efficiency characteristics obtained in the calculations according to model 1 ((3) - (3b)) and according to model 2 ((3) - (3c)) are given in Table 1.
Table 1. The values of statistical efficiency characteristic

|                          | Enterprises of the mechanical engineering industry | Enterprises of the oil refining industry |
|--------------------------|-----------------------------------------------------|-----------------------------------------|
|                          | Model one   | Model two   | Model one   | Model two   |
| Mean efficiency          | 0.80        | 0.92        | 0.88        | 0.93        |
| Variance efficiency      | 0.11        | 0.04        | 0.06        | 0.04        |

The factors (having significantly different from zero positive coefficients in (3c)) that increase the efficiency for enterprises in the engineering industry include the level of qualifications of personnel, the share of export deliveries, and the volume of research and development. For enterprises of the oil refining industry, a significant factor that increases efficiency is the volume of export deliveries. The share of attracted public funds does not affect the increase in efficiency in both groups considered.

3.3 Analysis of technological efficiency of enterprises using the DEA method

For enterprises of the mechanical engineering industry, we estimated the production potential by the DEA method and compared with the results obtained during the analysis by the SFA methods according to models (3) - (3b) and (3) - (3c).

Variant one. When comparing the efficiency by the DEA method with the results of solving problem 1 according to model (3b), we used the following indicators: output $y_i$, $i = 1, 2, ..., N$, estimated by the volume of work performed and products produced, inputs $x_i$, $i = 1, 2$, representing the logarithms of material costs and the number of production personnel.

Variant two. When comparing DEA with the results of problem 2 according to model (3) - (3c), we added three resources to the parameters used in variant 1: logarithms of the volumes of attracted public funds, volumes of export supplies, and the volume of research and development costs.

3.4 Analysis of technological efficiency of enterprises using the DEA method

The correlation coefficient between the efficiency calculated by the model (3) - (3b) and the efficiency calculated by the DEA in variant one is 0.72, the correlation coefficient between the SFA model (3) – 3(c) and the DEA in variant two is 0.8. Fig. 3 shows the scatter diagrams of the efficiency correlation coefficient illustrating the calculations performed.

Both approaches, generally, showed comparable results, which reveals the possibility of their joint application, which is advisable in many cases, since each of these methods has its own limitations and expands the scope of application.
Fig. 3. The efficiency correlation coefficients: the SFA model (3) - (3b) and DEA (left), the SFA model (3) – 3(c), and DEA (right).

4 Conclusion
Experiments conducted on real data to evaluate technological efficiency and identify significant factors for its improvement using the methods discussed above have demonstrated their effectiveness and the existence of a relationship between the results obtained by various methods. As a direction for further research, we consider it proper to expand the field of choice of inefficiency factors to include environmental factors such as crises, catastrophes, climate disasters, accidents, etc.

References
[1] Koopmans T.C. An Analysis of Production as an Efficient Combination of Activities, in Koopmans, T. C. (Ed.) Activity Analysis of Production and Allocation, Proceeding of a Conference, John Wiley and Sons Inc., London. 1951. P. 33-97.
[2] Charnes A., Cooper W.W., Rhodes E. Measuring of efficiency of decision making units, European Journal of Operational Research., 1978. Vol. 2. №6. P. 429-444
[3] Farrell, M.J. (). The measurement of productive efficiency, Journal of the Royal Statistical Society Series A (General), 1957. 120(3), P. 253-290. https://doi.org/10.2307/2343100
[4] Coelli, T.J. A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program, CEPA Working Paper 96/08, Department of Econometrics, 1996. University of New England, Armidale. Another reference
[5] Meeusen W., van den Broeck J., Efficiency Estimation from Cobb-Douglas Production Functions WithComposed Error, International Economic Review. 1976. p. 435–444.
[6] S. Aivazyan, M. Afanasiev, Evaluation of measures aimed at managing factors of production inefficiency//Applied Econometrics, 2007. №4, P. 28-41.
[7] S. Aivasian. M. Afanasiev. V. Rudenko. Problems of a company’s production potential modeling with consideration of intellectual capital. European Researcher 2013, vol. 48, N 5, P. 1220-1225