**Mihai Roman,PH. D., Monica Roman, PH.D,**
**Academy of Economic Studies - Bucharest**

# DATA ENVELOPMENT ANALYSIS METHOD IN LABOUR EFFICIENCY ANALYSIS

*The economic efficiency or the factors productuvity are essential elements of economic analysis. In this paper we propose an economic analysis of labour efficiency in Romania by intermediary of an modern mathematical method – Data Envelopment Analysis (DEA). In the first part of the article we briefly present the basics of the method. In the second part we analyse the Romanian labor efficiency based on DEA method.*
*Key words: efficiency, productivity, factors, envelope, activity*

## 1. INTRODUCTION

The economic efficiency or the factor productivity are essential elements of economic analysis. In the last two decades there was developed some specific methods to analyze efficiency, especially at microeconomic level. But we can use these methods successfully at macroeconomic level, to analyze the national economic efficiency.

Farrel[1] (1957) hypothesis is that firm efficiency consists of two components: technical efficiency and price efficiency. Technical efficiency represents the ability of a firm to obtain maximum level of output from a given set of inputs. Price efficiency represents the ability to use the inputs in optimal proportions for given technologies and prices. The combinations of these two components are a measure of total economic efficiency.

We intend to extend the analysis from microeconomic level to macroeconomic level and we use that to analyze the national economy efficiency.

Literature gives us four major methods to analyze efficiency:

1. Least Squares method
2. Total factor productivity method (TPF)[2]

---

[1] Farrell. M.J. The Measurement of Productive Efficiency, Journal of Royal Statistical Society. A120. 1957. p. 253-290

[2] TPF – Total Factors Productivity – a statistical method based on index method for productivity analyse

3. Data Envelopment Analysis (DEA)
4. Stochastic Frontiers Analysis.

The first two methods are applied to aggregate time series and give us a measurement for technical progress and dynamics of total productivity factors. For both methods we suppose that economy is at full employment level, so it is an efficient economy. The third and fourth methods assume that we haven't efficiency hypothesis, and we can observe the inefficiency causes and its dimension.

We can use total productivity factor method (TPF) to compare the relative efficiency for two economies or for two different periods. DEA and stochastic frontier method are used to observe the technical efficiency variations and the relative efficiency, if a significant panel data is available.

In this paper we shell analyze the labor productivity at macroeconomic level using DEA method.

## 2. DATA ENVELOPMENT ANALYSIS METHOD

Charnes, Cooper şi Rhodes[3] defines data envelopment analysis methodology in 1978-study as a mathematical programming model for data observations wich assure a new way to obtain empirical estimation for externe relationship, such as production function or/and efficient possibilities production surfaces.

*Basic concepts*

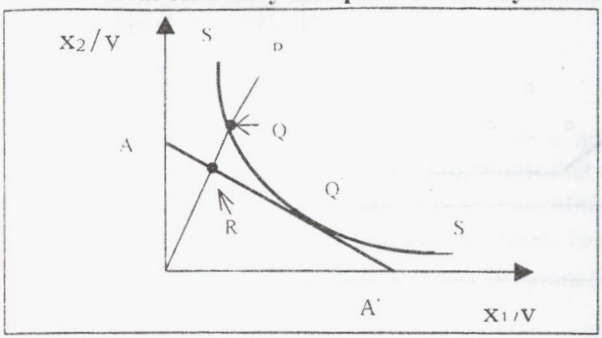
We consider two firms (or branches) who are using two inputs $(x_1,x_2)$ to produce one output, $y$, in constant return to scale hypothesis. We also suppose that we know the isoquants for complete efficient firms (the production frontier is not known, so that we can estimate that by observing firms in considered industry). We use DEA method to estimate this frontier, represented by $SS'$ in Figure 1. If a firm uses quantities of inputs corresponding at point P to produce a unit of output, then the technical inefficiency could be represented by the distance QP. These represents the percentage by which all inputs need to be reduce to obtain technically efficient production. This is represented by the ratio $QP/OP$ .

*Technical efficiency* ($TE$) is also mesured by the ratio $TE_i = \frac{OQ}{OP}$ and it provides an indictor of the degree of technical inefficiency:

$$TE_i = \frac{OQ}{OP} = 1 - \frac{QP}{OP} \in [0,1] \qquad (1)$$

[3] Charnes, A., Cooper W.W.M, Rhodes ,E. – *Measuring the Efficiency of Decisions Making Units*, in European Journal of Operations Research, 2, p. 429-444, 1978

**Fig. 1**
**Technical efficiency and price efficiency**




If the input prices ratio (represented by the slope of $AA'$ isoquant in Figure 1.) is known, then we can determine price efficiency.

*The price efficiency* ($AE$) for the firm operating at $P$ is the ratio:

$$AE_i = \frac{OR}{OQ} \qquad (2)$$

and $RQ$ represent cost reduction that would occur if the production occures at the $Q$ point ( this point is pricelly and technically efficient).

Total economic efficiency *(EE)* is defined by ratio:

$$EE_i = \frac{OR}{OP} \qquad (3)$$

where $RP$ indicates cost-reduction possibility.

We observe that :

$$TE_i \times AE_i = \frac{OQ}{OP} \times \frac{OR}{OQ} = \frac{OR}{OP} = EE_i \qquad (4)$$

In order to determine the efficiency coefficients we assume that we know the production function. This hipothesis is not realistic, so that we have to estimate efficient isoquant from a panel data. Farrell suggested the use of a non-parametrical convex isoquant constructed such that observed points does not lie to the left or below it (Figure 2.). Another possibility is to use a Cobb-Douglas parametrical function, such that observed points would not to be at the left or below this curve.

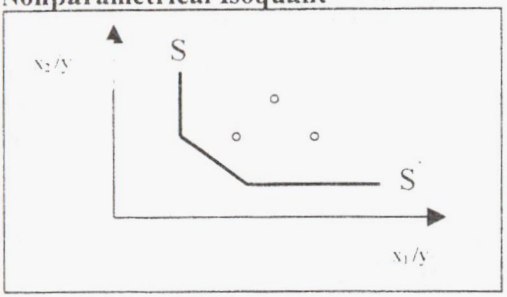
This analysis is ussaly called "input oriented efficiency measure".

One question we can ask here is: *" How much we can diminish inputs without diminishing output quantities produced?"*

Another question is: *" By how much can output quantities be increased without altering the inputs quantities used?"*

Responding on this questions we obtain a measure for output orinted efficiency.

**Fig. 2**
**Nonparametrical Isoquant**




*Enveloped surfaces*

The production functions indicates the frontier of production posibility set. We can evaluate the relative efficiency only if we know this frontier..

Suppose that we analyse $N$ firms (or $N$ activities). Each firm use different inputs ($K$ inputs) to produce $M$ outputs. Colon $K$-dimensional vectors $x_i$ indicates input quantities for $i$-firm and colon $M$-diemnsional vector $y_i$ indicates output quantities for the same $i$ –firm. $X$ and $Y$ are input/output matrix for the whole economy.

By intermediary of DEA methods we search an *enveloped surface* for all production functions. This enveloped surface represents the economy's empirical production function or efficiency frontier. Using DEA we can analyze the relative efficiency for many-inputs ant many-outputs technologies. The efficiency performances of each firm depend on our position relative to this enveloped surface. Each firm wich is on the enveloped surface is an efficient firm, and the rest of them are relative inefficient firms.

Depending on returns to scale, we have two classes of enveloped surfaces: one class for constant returns to scale technology (CRS), and another for variable returns to scale technology (VRS). The kind of these surfaces depend on data panel and economic hypothesis.

## 3. DEA MODELS CLASSIFICATION

*The Constant Returns to Scale DEA Model*

Intuitively we can introduce DEA as a ratio. For each firm we measure the ratio of the value of all outputs over the value of all inputs, that means the ratio $u'y_i / v'x_i$, where $u$ is an $Kx1$-dimensional estimation vector for ouputs and $v$ is an $Mx1$-dimensional estimation vector for outputs.

Optimal estimation results by solving the program:

$$\begin{cases} \max\limits_{u,v} (u^{'} y_i / v^{'} x) \\ u^{'} y_j / v^{'} x \le 1, \qquad j = 1,2,\ldots,N \\ u, v \ge 0 \end{cases} \tag{7}$$

The results are the values output quantities $u_i$ and input quantities $v_i$ that maximize firm $i$ efficiency, subject to the constraints that all efficiency measures are less than or equals to one. This program can have an infinit number of solutions, such as $(\alpha u^{*}, \alpha v^{*})$ where $(u^{*}, v^{*})$ represent system solution. To avoid such a situation, we impose that $v^{'} x_i = 1$, so that we obtain a new program (7):

$$\begin{cases} \max \mu^{'} y_i \\ \mu, v \\ v^{'} x_i = 1 \\ \mu^{'} y_j - v^{'} x \le 0, \qquad j = 1,2,\ldots,N \\ \mu, v \ge 0 \end{cases} \tag{7}$$

This program is called the *multiplier form* of an DEA linnear programing model.

Using duality in linear programing we obtain an equivalent form for this problem:

$$\begin{cases} min\,\theta \\ \theta, \lambda \\ -y_i + Y\lambda \ge 0 \\ \theta x_i - X\lambda \ge 0 \\ \lambda \ge 0 \end{cases} \tag{8}$$

In this model $\theta$ is an real number between $0$ and $1$, and $\lambda$ is an $Nx1$-dimensional vector of constants. If $\theta=1$, then the corresponding firm is technically efficient, and if $\theta < 1$, then the analized firm is not on efficiency frontier.
So, we have to solve the linear programming problem (8) N times, and for each firm we obtain an efficiency parameter or a score of his efficiency.

*The Variable Returns-to-Scale DEA Models (VRS)*

The hipothesis that we have CRS is not very realistic. It is possible like a firm cann't operate at an optimal scale be cause of imperfect competition, imperfect informations, financial restrictions and so on.

Banker, Charnes şi Cooper[4] (1984) suggest an extension of the CRS DEA model to explain the variable returns to scale situations.

We obtain an linear programming problem, based on CRS DEA problem subject on convexity supplementary condition:

$$\begin{cases} \min\limits_{\theta,\lambda} \theta \\ -y_i + Y\lambda \geq 0 \\ \theta x_i - X\lambda \geq 0 \\ N_1\lambda = 1 \\ \lambda \geq 0 \end{cases} \qquad (9)$$

where $N_1$ is an $N\times 1$ – dimensional vector with all components *1*.

An enveloped surface VRS is in this case a convex envelope of points who represent firms efficiency. This convex envelope surface provides a better solution then envelope surface done by CRS DEA model. Using this method we can also observe the type of return –to-scale for each firm, type wich in not necessarly CRS.

We obtain the VRS for the analized firm and if there is a differece between VRS solution *(TE)* and CRS solution, then we have an *inefficient scale*. The inefficiency dimension is done by difference between TE obtained with VRS $(TE_{VRS})$ and *TE* obtained with CRS $(TE_{CRS})$.

In figure 3 we illustrate scale inefficiecy using one input and one output. In CRS hypothesis technical inefficiency input -oriented of point P is $PP_C$ distance, and in VRS hipothesis is $PP_V$ distance.

The difference between these two technical efficiency measures $(P_C P_V)$ is due by scale inefficiency:

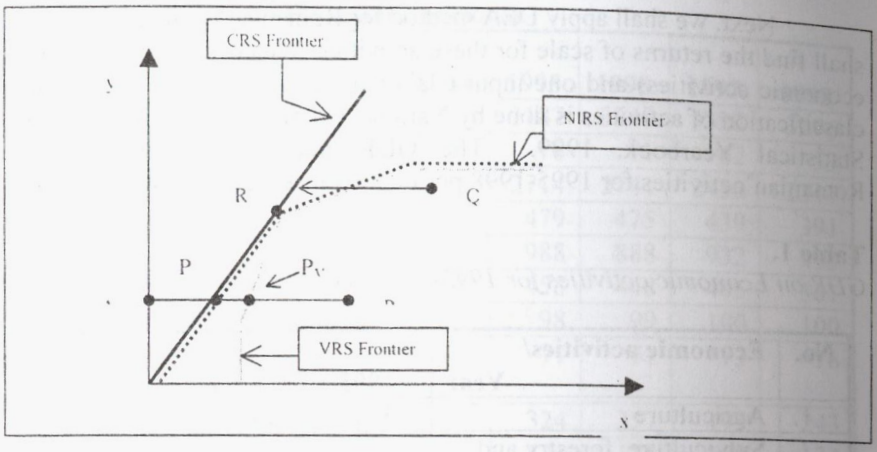
$$TE_{CRS} = \frac{AP_C}{AP}, \quad TE_{VRS} = \frac{AP_V}{AP}, \quad SE = \frac{AP_C}{AP_V} \qquad (10)$$

$$\Rightarrow \quad TE_{CRS} = TE_{VRS} \times SE \qquad (11)$$

So, we can decompose CRS technical efficiency in "pure" technical efficiency and scale efficiency. Scale efficiency is the ratio of the average product at P – point firm and average productivity of efficient firm (R point).

[4] Banker, R.D., Charnes, A. , Cooper, W.W. – *Some Methods for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis*, in Management Science.30:1078-1092. 1984

**Figure 3.**
**Scale economies in DEA models**




The main problem is that we don't know if we have an DRS firm or an ISR firm. To find that, we replace in program (9) condition $N_1\lambda=1$ with $N_1\lambda\leq 1$, and we obtain a condition for Non Increasing Returns to Scale (NIRS).

$$\begin{cases} \min\limits_{\theta,\lambda} \theta \\ -y_i + Y\lambda \geq 0 \\ \theta x_i - X\lambda \geq 0 \\ N_1\lambda \leq 1 \\ \lambda \geq 0 \end{cases} \qquad (12)$$

We can determine the inefficiency scale type for firms comparing techical efficiency obtained in NIRS case with $TE_{VRS}$. If these are not equal, then the P-firm has an Increasing returns to scale (IRS), and if they are equal, then we have an diminishing returns to scale(DRS).

All models presented here are input oriented models. We can identify the inefficiency by DEA VRS model, and thisgive as the proportional input reduction so that the output stay unchanged (input oriented model).

But, it is also possible to measure technical inefficiency as a proportional increase in output for a fixed level of inputs (output oriented model). These two inefficiency measurment are equals under CRS conditions, and differents under VRS conditions.

In real life we use especially input oriented models, because firms have contracts to respect, so that their purpose is to minimize costs (by input reduction) and to obtain the same output.

It also exists the other situation when we try to maximize the output level with a given quantity of inputs. In this case we must use output-oriented models.

## 5. LABOUR EFFICIENCY ANALYSIS IN ROMANIA

Next, we shall apply DEA method for Romanian economic activities and we shall find the returns of scale for these activities. We consider one output (GDP for economic activities) and one input ( labor force involved in these activities). The classification of activities is done by National Comission for Statistics in Romanian Statistical Yearbook, 1999. The GDP and the employed population in Romanian activities for 1995-1998 period are presented in Table 1 and Table 2.

**Table 1.**
*GDP on Economic activities for 1995-1998 period (billion lei, curent prices)*

| No. | Economic activities/ Year | 1995 | 1996 | 1997 | 1998 |
|---|---|---|---|---|---|
| 1. | Agriculture | 13941,3 | 20459,7 | 44582,7 | 52419,0 |
| 2. | Sylviculture, forestry and hunting | 328,0 | 489,5 | 950,1 | 1262,9 |
| 3. | Industry | 23711,3 | 36181,5 | 78093,8 | 101165,8 |
| 4. | Construction | 4755,1 | 7067,4 | 13230,0 | 19611,8 |
| 5. | Trade | 7570,2 | 12722,3 | 28767,5 | 50610,0 |
| 6. | Transport and storage | 4567,8 | 7658,9 | 17270,7 | 22429,3 |
| 7. | Post and telecommunications | 1008,8 | 2145,8 | 5394,6 | 9697,1 |
| 8. | Financial, banking and insurance activities | 3588,1 | 3243,1 | 4298,7 | 5656,9 |
| 9. | Real estate and other services | 3452,6 | 5871,1 | 25860,1 | 43406,4 |
| 10. | Public administration | 2691,8 | 3355,2 | 6763,3 | 12190,8 |
| 11. | Education | 1853,9 | 2776,5 | 5152,7 | 8182,2 |
| 12. | Health and social assistance | 1304,6 | 2105,9 | 3709,4 | 6067,7 |

**Source**: *Romanian Statistical Yearbook, 1999*, National Commission for Statistics

Using DEA metod we obtain a classificaton for economic activities, depending on efficiency of labor.

Using data from Tables 1 and 2 we determined efficiency parametres $\theta$ for each activity and for each year. We calculate efficiency parametres $\theta$ *by* intermediary of QM program and the results are in Table 3.

Table 2.

**Employed population in economic activities (1995-1998)**
**(Thousands persons)**

| No. | Economic activities/ Year | 1995 | 1996 | 1997 | 1998 |
|---|---|---|---|---|---|
| 1. | Agriculture | 3187 | 3249 | 3322 | 3296 |
| 2. | Sylviculture, forestry and hunting | 78 | 71 | 62 | 53 |
| 3. | Industry | 2714 | 2741 | 2450 | 2317 |
| 4. | Construction | 479 | 475 | 439 | 391 |
| 5. | Trade | 988 | 888 | 932 | 933 |
| 6. | Transport and storage | 458 | 448 | 405 | 361 |
| 7. | Post and telecommunications | 98 | 99 | 100 | 100 |
| 8. | Financial, banking and insurance activities | 71 | 71 | 73 | 76 |
| 9. | Real estate and other services | 324 | 257 | 199 | 243 |
| 10. | Public administration | 131 | 125 | 130 | 134 |
| 11. | Education | 437 | 441 | 426 | 426 |
| 12. | Health and social assistance | 333 | 337 | 315 | 317 |
| **13** | **Total** | **9298** | **9202** | **8853** | **8647** |

Source: *Romanian Statistical Yearbook, 1999*, National Commission for Statistics

DEA model is:

$$\begin{cases} \min\limits_{\theta,\lambda} \theta \\ -y_i + Y\lambda \geq 0 \\ \theta x_i - X\lambda \geq 0 \\ N_1\lambda \leq 1 \\ \lambda \geq 0 \end{cases} \qquad (13)$$

where:

$\theta$ - is efficiency parameter;
$n$ - is the number of activities (12);
$Y$ - is an n-dimensional array of GDP activities, for each year;
$X$ - is an n-dimensional array of input (labor force / activity);
$N_1$ - is an n-dimensional array with 1;
$\lambda$ - is an array of working variable for our problem.

Table 3.

**Efficiency parameters for economic activities in 1995-1998 period**

| No. | Economic activities/ Year | 1995 | 1996 | 1997 | 1998 |
|---|---|---|---|---|---|
| 1. | Agriculture | 0.449 | 0.447 | 0.335 | 0.172 |
| 2. | Sylviculture, forestry and hunting | 0.083 | 0.151 | 0.141 | 0.133 |
| 3. | Industry | 1.000 | 1.000 | 1.000 | 1.000 |
| 4. | Construction | 0.468 | 0.747 | 0.277 | 0.281 |
| 5. | Trade | 0.601 | 0.922 | 0.52 | 0.538 |
| 6. | Transport and storage | 0.436 | 0.901 | 0.393 | 0.384 |
| 7. | Post and telecommunications | 0.204 | 0.475 | 0.497 | 0.543 |
| 8. | Financial, banking and insurance activities | 1.000 | 1.000 | 0.542 | 0.417 |
| 9. | Real estate and other services | 0.211 | 1.000 | 1.000 | 1.000 |
| 10. | Public administration | 0.407 | 0.631 | 0.497 | 0.509 |
| 11. | Education | 0.084 | 0.138 | 0.111 | 0.108 |
| 12. | Health and social assistance | 0.078 | 0.137 | 0.108 | 0.107 |

In Table 4 we present the ordered activities depending on labor efficiency for 1998 year.

In year 1998 the best efficiency parameters was registered in Industry and in Real estate and other services, coefficient 1.

The others activities registered lower efficiency parameters reporting on in Industry and in Real estate and other services.

**Table 4**

**Ordered activities depending on labor efficiency**

| No. | Economic activities | Efficiency parameter |
|---|---|---|
| 1. | Agriculture | 1.000 |
| 2. | Sylviculture, forestry and hunting | 1.000 |
| 3. | Industry | 0.543 |
| 4. | Construction | 0.538 |
| 5. | Trade | 0.509 |
| 6. | Transport and storage | 0.417 |
| 7. | Post and telecommunications | 0.384 |
| 8. | Financial, banking and insurance activities | 0.281 |
| 9. | Real estate and other services | 0.172 |
| 10. | Public administration | 0.133 |
| 11. | Education | 0.108 |
| 12. | Health and social assistance | 0.107 |

Trade, Post and telecommunication, and Public administration are efficiency parametrers between 0.5 and 0.6, and follow the best efficiency of industry. Surprising, Financial, banking and insurance activities are only in sixth position , with an 0.417 efficiency coefficient. This parameter goes down together with GDP in Romania.

The lowest efficiency parameters are in Education and Health and social assistance, with one tenth parts of industry efficiency. In the same way we find that Agriculture and Sylviculture, forestry and hunting registered very low levels of efficiency in Romania, that means that a large part of Romanian population are involved in these activities and there are a poor productivity.

Dynamic analysis of efficiency parameters suggests that:

- Industry and Real estate and other services was during the observed period the best activities, with a 1-coefficient for efficiency for each yearbetween 1995-1998;
- Financial, banking and insurance activities was in 1995 and 1996 on efficiency frontier, and then goes down to 0.417 in 1998;
- Another activity who decrease his efficiency is Agriculture (from 0.446 in 1995 to 0.172 in 1998);
- Only Post and telecommunications registered a permanent growth, from 0.204 in 1995 to 0.543 in 1998;
- The others activities registered different oscillations in analyzed period.

## REFERENCES

**[1] Banker, R.D.,Charnes, A. ,Cooper, W.W. - *Some Methods for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis*, in Management Science,30:1078-1092, 1984**

**[2] Charnes, A., Cooper W.W.M, Rhodes ,E. - *Measuring the Efficiency of Decisions Making Units*, in European Journal of Operations Research, 2, p. 429-444, 1978**

**[3] Farrell, M.J. - *The Measurement of Productive Efficiency*, Journal of Royal Statistical Society, A120, 1957, p. 253-290**

**[4] Greene, W. H. - *The Econometric Analysis to Efficiency Analysis*, in Friend, H.O., Lowell,C. and Schmidt, S.S., (Eds), Oxford University Press, New York, 1993**

**[5] Roman, M.,Roman, M. - *Analiza echilibrului pe piața forței de muncă*, în volumul "Rolul statisticii în economia de piață ", Ed. Centrul Tipografic ASE, București, 1999, pag. 245-249**

**[6] *** - *Anuarul Statistic al României 1999*, Comisia Națională pentru Statistică,2000**