Publication Efficiency at DSI FEM CULS – An Application of the Data Envelopment Analysis

Martin Flégl, Helena Brožová [1]

Abstract. The education and research efficiency at universities has always been very important factor. Educational institutions as university departments receive financial resources according their successful performances. In our contribution we compiled the model of the Data Envelopment Analysis for evaluating publication and research efficiency at Department of Systems Engineering, Faculty of Economics and Management, Czech University of Live Sciences Prague in the period 2008-2011. Measured units are department academic staff divided into four categories as Ph.D. students, Lecturer staff + Technical workers, Associate professors and Professors. The attention is also paid to changes in the staff position during the evaluated period. The model outputs are points from publication and research activity. According to the input and output form various versions of the basic model were calculated and the results were analysed.

Keywords: Data Envelopment Analysis, Research and Development results, publication activity, Malmquist index.

JEL Classification: C61, C65
AMS Classification: 90B50, 90C90

1 Introduction

The methods for evaluation of the Research and Development results (R&D) are intensively discussed within the field of scientific policy. The main goal of evaluation is to provide information on research results that were created due to financial support from public resources, and also to gain an insight into the efficiency of such financing. The problem of R&D performance is also discussed at the Czech University of Life Sciences Prague (CULS) and its faculties and departments. The result of this discussion at the Faculty of Economics and Management (FEM) is the Motivation Programme [9] which was introduced in 2010. The program aims are to stimulate publication and research activity of the all academics.

The quantitative evaluation of the organisation has direct implications for financing universities, research organisations and others. From this point of view, the achieved scores indicate the scientific productivity of the organisation. Despite the fact that the official evaluation has many weaknesses, a different tool is not available to enable R&D results to be quantitatively evaluated on the same level of exactness and complexity as the current system.

The official evaluation process is based on formalised procedure which distinguishes between two categories of results [1]:

- Results of basic research – books, papers in scientific journals, conference proceedings;
- Results of applied research – patents, prototypes, industrial designs, maps, certified methods, software.

Each of these results is ascribed a score, such as 20 points for a book, a paper in a journal of the impact factor (IF) receives a score within the interval 10 – 305 (according to the journal ranking), and certified methods approved by a State administration body are valued at 40 points, etc. The evaluation is carried out for each organisation (such as a university), whereby the organisation gains the relative share equal to the share of the authors who created the outcome and are affiliated to the given organisation.

In our contribution we propose to use the model of the Data Envelopment Analysis (DEA) for evaluating publication and research efficiency and we show its application at Department of Systems Engineering, Faculty of Economics and Management, Czech University of Live Sciences Prague (DSI FEM CULS) in the period

[1] Czech University of Live Sciences in Prague, Faculty of Economics and Management, Department of Systems Engineering, Kamycka 129, 165 21 Praha 6 – Suchdol, flegl@pef.czu.cz, brozova@pef.czu.cz.

2008-2011. Academic staff of department is divided into four categories as Ph.D. students, lecturer staff + technical workers, associate professors and professors are Decision Making Units (DMUs). The attention is also paid to changes in the staff position during the evaluated period. The model outputs are points from publication and research activity. DEA models for three periods and the Malmquist index are calculated. The results of the models are also interpreted in a graphical form.

2 Materials and Methods

2.1 Efficiency measuring in educational institutions

The reliability of the results depends on the accurate selection of the data best adapted to the objective of the study. Bessent et al [2] pointed out the major problems in educational efficiency measurement which are also important in process of R&D evaluation:

- Obtaining data to specify adequate input measures;
- Obtaining data to specify outputs that were not limited to cognitive test results;
- Difficulties in communicating the results to those affected by them.

It is possible to find many scientific studies based on measuring efficiency in educational environmental. Worthington [11] summarises the approaches which have been used for measuring efficiency in educational institutions (high schools, universities, study programme efficiency, etc.) between 1981 and 1998. The DEA method was the dominant method of the educational efficiency measurement. These DEA models mainly contain the number of teaching, administrative and support staff as the inputs.

The outputs are thus the papers and letters in academic journals, authored/edited books, published works [6]. Kao and Hung [7] compiled model for measuring department efficiency. Model contains personnel, operating expenses and floor space as the inputs. The outputs were total credit hours, publications and external grants. Jablonský [4] presented the DEA model for measuring resources' allocation among university departments. The model contains the number of hours of direct and indirect teaching and the quality of research as the outputs. DEA models for measuring departmental efficiency can also be focused on improving teaching performance. Montoneri et al [8] used the richness of course contents and the diversity of accessed multiple teaching channels as the inputs. The outputs were thus the positive degree of teaching attitude and students' learning performance.

2.2 Data Envelopment Analysis Method

DEA evaluates DMUs against the best DMUs with the idea: if some DMU can produce a certain level of output utilizing certain level of input, another DMU of equal scale should be capable of doing relatively the same. DEA is a nonlinear programming model for the estimation of productive efficiency of DMUs based on relationship between multiple outputs and multiple inputs. These outputs and inputs are usually of various characters and of variety of forms which are difficult to measure. The DEA measure of the efficiency of any DMU is obtained as the maximum of a ratio of weighted outputs to weighted inputs subject to the condition that the similar ratio for every DMU is less than or equal to 1.

The simplest DEA model assumes constant returns to scale, this model is called CCR model according to its authors Charnes, Cooper, and Rhodes [3]. Let y_{jk} be the amount of the j^{th} output from unit k, and x_{ik} be the amount of the i^{th} input to the k^{th} unit. Using the CCR model the DMU efficiency of a particular unit H is calculated using the following linearization of original DEA model. Primal and dual CCR output oriented models are formulated as:

Primal model
$$\Phi_H = \sum_{i=1}^{m} v_{iH} x_{iH} \to MIN \tag{1}$$

subject to
$$\sum_{j=1}^{n} u_{jH} y_{jH} = 1$$
$$\sum_{i=1}^{m} v_{iH} x_{ik} - \sum_{j=1}^{n} u_{jH} y_{jk} \geq 0, \quad k = 1, 2, \ldots, p \tag{2}$$
$$u_{jH} \geq 0, \; j = 1, 2, \ldots, n \; a \; v_{iH} \geq 0, \; i = 1, 2, \ldots, m.$$

Dual model
$$z_H \rightarrow MAX \tag{3}$$

subject to
$$\sum_{k=1}^{p} \lambda_{kH} x_{ik} \le x_{iH},\ i = 1,\ 2,\ldots,\ m$$
$$z_H y_{jH} - \sum_{k=1}^{p} \lambda_{kH} y_{jk} \le 0,\ j = 1,\ 2,\ldots,\ n \tag{4}$$
$$\lambda_{kH} \ge 0,\ k = 1,\ 2,\ldots,\ p$$

The decision variables $u = (u_1,\ldots,u_m)$ and $v = (v_1,\ldots,v_n)$ are the weights given to the m outputs and to the n inputs respectively. To obtain the relative efficiencies of the all units, the model is solved for one unit at a time. The decision variables $\lambda = (\lambda_1,\ldots,\lambda_p)$ are the weights given to the efficient DMUs for creating virtual (efficient) DMU corresponding to non efficient DMU. The inputs and outputs of virtual DMU are calculated used the formulas:

$$x'_{iH} = \Phi_H x_{iH} - s_{iH}^{-},\ i = 1,\ldots,m$$
$$y'_{jH} = y_{jH} + s_{jH}^{+},\ j = 1,\ldots,n \tag{5}$$

or

$$x'_{iH} = \sum_{k=1}^{p} \lambda_{kH} x_{ik},\ i = 1,\ldots,m$$
$$y'_{jH} = \sum_{k=1}^{p} \lambda_{kH} y_{jk},\ j = 1,\ldots,n \tag{6}$$

where s_{iH}^{-} and s_{jH}^{+} are slacks in the dual constraints.

The constant returns to scale describes the individual constant ability of publication and research work. Output orientation of the model means that results explicitly show the necessary augmentation of outputs with the same amount of inputs. This model orientation reflects the possibility of DMUs to improve his/her research activity.

Analysis of changes of DMU's efficiency over time is based on the Malmquist index [4], which can be used for the investigation of the causes of efficiency change. Malmquist index is defined with constant returns to scale, which allows supposing the same technology in both periods. This convention enables Malmquist index with output orientation quantifies change of efficiency of DMU between period *t* and period *t*+1 and can be formulated as follows

$$M(x^t, y^t, x^{t+1}, y^{t+1}) = \left(\frac{D^t(x^t, y^t) D^t(x^{t+1}, y^{t+1})}{D^{t+1}(x^t, y^t) D^{t+1}(x^{t+1}, y^{t+1})} \right) \tag{7}$$

where $D^t(x^t, y^t)$ is efficiency in the period *t* and $D^{t+1}(x^{t+1}, y^{t+1})$ efficiency in period *t*+1,

$D^t(x^{t+1}, y^{t+1})$ is efficiency in the period *t*+1 considering efficiency frontier in period *t* and

$D^{t+1}(x^t, y^t)$ is efficiency in period *t* considering efficiency frontier in period *t*+1.

Malmquist index greater than 1 indicates productivity gain; Malmquist index less tahn 1 indicates productivity loss; and Malmquist index equal to 1 means no change in productivity from time *t* to *t* + 1.

Authors used Efficiency Measurement System (EMS) software for calculation DEA model [10].

2.3 Model Data

This study fructifies the secondary data from Rejstřík informací o výsledcích/Information Register of R&D results (RIV), which is the key database for the evaluation of scientific work in the Czech Republic. The evaluation is carried out by the Rada pro výzkum, vývoj a inovace/Research, Development and Innovation Council (RVVI). All the results are evaluated by the Metodika hodnocení výsledků výzkumných organizací/Methods for evaluating R&D results [1] which are focused on results that were produced by each research organisation in the last five years. The study is based on the most up-to-date files that refer to R&D results published between 2006 and 2010. These results were officially published by the RVVI in January 2012.

Background data for the DEA model contains DMUs, inputs and outputs (**Table 1**).

DMU	Input	Output (2008)	Output (2009)	Output (2010)	Output (2011)
P1	8	3.153	0	0	30.253
P2	12	8.536	8.278	8.518	28.333
P3	12	18.920	0	5.845	31.666
P4	16	17.738	0	0	5
P5	8	21.286	0	15.305	24.334
P6	8	21.286	0	15.305	15
P7	12	8.536	7.096	0	19.252

Table 1 Input data for DEA model

Evaluated DMUs are expressed for 7 employees of Department of Systems Engineering who published at least one paper in 2008-2009 and 2010-2011 periods. This constraint is important for possibility of DEA model application and for Malmquist index calculation.

Each DMU has only one input expressing the position at the department, i.e. Ph.D. student express one point per year, Lecturer staff + Technical workers two points, Associated professors three points and finally Professors four points for each year. The data were obtained from the university's databases during the period of 2008 - 2011. Data had to be cleaned from imprecise data to guarantee accurate results. We noted changes of each DMU, e.g. if a Ph.D. student graduated in 2009 we must calculate its points as follows: first two years 2 points (one point per year) and from years 2010 and 2011 4 points (two pints per year). From this point of view, the position of authors is precisely classified to a year in which the person published his/her paper. Inputs are calculated by position-year measure with regard to four categories.

The outputs of our DEA model represent the evaluation of publication activities divided into the four categories: publications in 2008, 2009, 2010 and 2011. The points evaluation represents received points from publication/research activities during this period. This publication/research activity was evaluated by RVVI [1]. Points were then summarized for each author and its position with regard to the published year. In the case of more than two authors of a paper, we divided the points in proportion.

3 Results and Discussion

In our contribution, the output-oriented DEA model is used. The reason for output orientation is because the authors want to evaluate publication/research activities. The results will give us information as to who is efficient and who is not. The recommendation for an inefficient employee is going to be an improvement of the publication/research activity. It is also necessary to mention that this model is calculated for the 2008-2011 period.

Table 2 summarizes the efficiency results of all seven employees. In the first column Output Score 2008-2011 shows the efficiency of the DSI's employees. Efficient employees are P1, P2, P5 and P6 which have the efficiency result of 100 %. The column Output score I shows the efficiency result in 2008-2009 period. The efficient DMUs are P2, P5 and P6. The other DMUs are inefficient although the DMU P7 is closed to the efficient line. The results for the same DMUs but for the 2010-2011 period are shown in column Output score II. In this period the group of the efficient DMUs partly changed. P5 and P6 are still efficient. P1 became efficient instead of the P2 which is now inefficient. P2 and P3 are relatively closed to the efficient line.

The main objective of our contribution was to calculate Malmquist index and measure the efficient development during the 2008-2011 period. Malmquist index values are in the last column in the **Table 2**. The DMUs with the Malmquist index greater than 1 have increased their efficiency from the first period to the second period. The highest improvement from measured DMUs reached the DMU P1. Malmquist index value 5,724 means efficiency increased almost 6 times, this employee was really inefficient in the first period and strongly increased his efficiency in the second period. P2, P3 and P7 were almost inefficient in the both periods; nevertheless their efficiency was increasing in the group of analysed DMUs. On the other side those DMUs which have the Malmquist index lower than 1, have decreased their efficiency in monitored periods. This is only a situation of P4. The DMUs which have the Malmquist index equal to 1 are efficient in both periods.

DMU	Output Score (2008-2011)	Output Score I (2008-2009)	Output Score II (2010-2011)	Malmquist index
P1	100.00%	675.02%	100.00%	5.724
P2	100.00%	100.00%	143.48%	1.369
P3	124.00%	168.75%	133.76%	1.237
P4	240.01%	240.01%	1210.12%	0.394
P5	100.00%	100.00%	100.00%	1
P6	100.00%	100.00%	100.00%	1
P7	111.69%	111.69%	235.71%	1.129

Table 2 DMUs' efficiency and Malmquist index

The **Figure 1** describes the DMUs' efficiency changes during the measured periods. Figure contains two efficiency frontiers, the efficiency frontier I for the period 2008-2009 and then the efficiency frontier II for period 2010-2011. Each employee is described by the dotted line; its initial point shows the position (efficiency) in the period 2008-2009, its ending point with arrow represents the position in the second period 2010-2011.

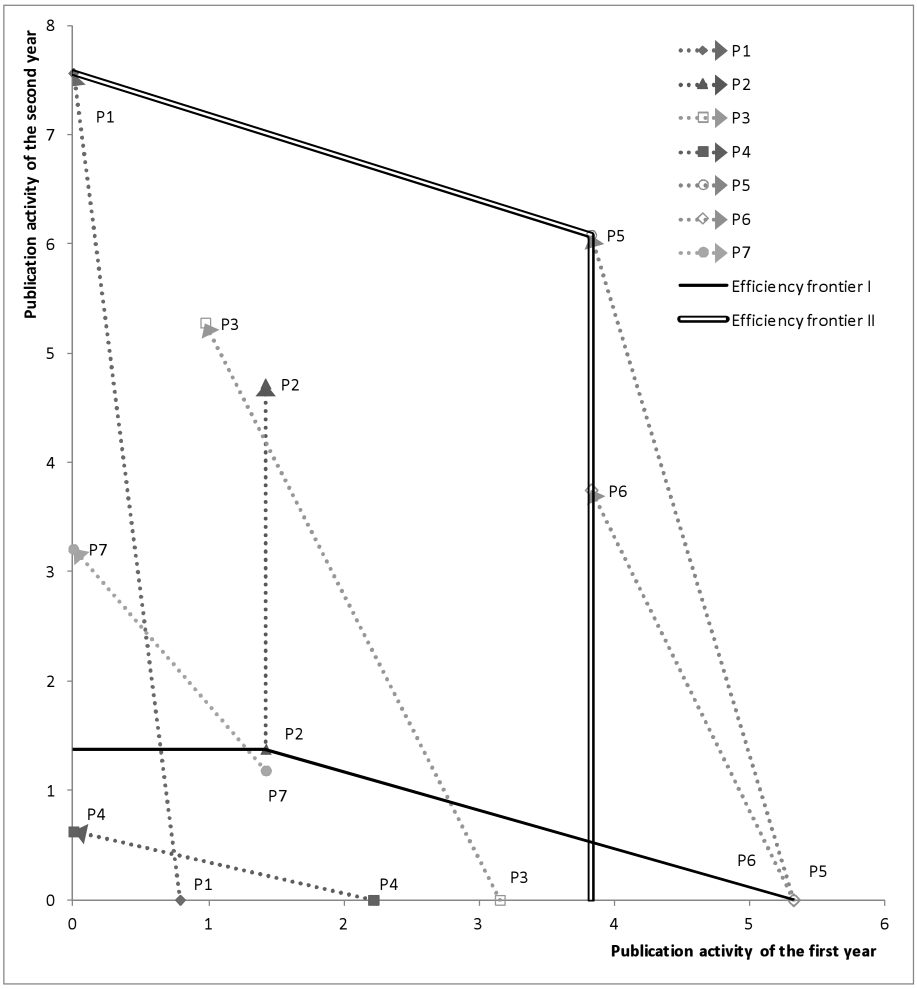

Figure 1 Changes of DMUs' efficiency during the measured periods

According to the DEA methodology all those DMUs which are lying on the frontier are efficient. The DMUs P5 and P6 are efficient in both periods, both DMUs lie on the efficiency frontiers. The unit P6 however is dominated by P5, because P5 is better in the second period, P5 has more publications then P6. The representation of units P3, P4 and P7 explains inefficiency of these units, but also the differences among them. Unit P4 shows significant deterioration of its efficiency, because the ending point of its line is very far from the efficiency frontier of the second period. On the other hand unit P1 is shifted from inefficient position to the efficient point on the efficiency frontier of the second period. And unit P2 is shifted from the efficient point on the efficiency frontier of the first period to inefficient position in the second period.

4 Conclusion

Presented DEA analysis of the efficiency of R&D publication activity of employees shows reasonability of this approach even though the number of analysed DMUs in this case is small.

The important requirement of this approach is simple structure of inputs and outputs system allowing two-dimensional description of the results.

- Inputs data describes category of employee at a department during two-year period;
- Output data specification is based on RIV evidence and analysed period can be only two-year long;
- Graphical representation is the best way how to explain the results of the DEA model.

The efficiency score indicates how the employees have to improve their research activity. These results depend on the selected radial measure, so it is possible for some employees improve their publication activity in other proportion. Malmquist indices calculated for analysed employees show individual position at a department and show the changes of their efficiency. Generally it is not possible to expect that employees will be efficient in following periods. This is caused by the character of research work and long time of research finalisation.

Acknowledgements

This work was supported by the grant project of the Grant Agency of CULS Prague, No. 201221029, "Využití metody DEA pro alokaci veřejných zdrojů mezi produkční jednotky".

References

[1] RVVI: Metodika hodnocení výsledků výzkumných organizací/Methods for evaluating R&D results. [Online], Available: http://www.vyzkum.cz [12 April 2012], 2010.

[2] Bessent, A., Bessent, W., Kennington, J., Reagan, B.: An application of mathematical programming to assess productivity in the Houston independent school district. *Management Science*, 28, pp. 1355-1367, 1982

[3] Charnes, A., Cooper, W., Rhodes, E.: Measuring the efficiency of decision-making units. *European Journal of Operational Research,* vol. 2, pp. 429–444, 1987.

[4] Cooper, W., Seiford, L. M., Tone, K.: *Data Envelopment Analysis.* 1st ed. New York: Springer Publisher, 2006.

[5] Jablonský, J.: Models for efficiency evaluation in education. In: *Proceedings of the 8th International Conference on Efficiency and Responsibility in Education (ERIE 2011),* Prague, pp. 88-97, 2011.

[6] Johnes, J., Johnes, G.: Research funding and performance in U.K. university departments of economics: A frontier analysis. *Economics of Education Review*, 14, pp. 301-314, 1995.

[7] Kao, Ch., Hung, H.T.: Efficiency analysis of university departments: An empirical study. *Omega*, 36, pp. 653-664, 2006.

[8] Montoneri, B., Lin, T. T., Lee, Ch. Ch., Huang, S.-L.: Application of data envelopment analysis on the indicators contributing to learning and teaching performance. *Teaching and Teacher Education*, 28, pp. 382-395, 2012.

[9] Motivační program pro pracovníky a doktorandy PEF ČZU v Praze v roce 2012/Incentive program for employees and postgraduates of FEM CULS in Prague in 2012. Pef ČZU, Praha, 2012.

[10] Scheel, H.: *EMS: Efficiency Measurement System.* [Online], Available: http://www.holger-scheel.de/ems/ [12 April 2012], 2000.

[11] Worthington, A.: An Empirical Survey of Frontier Efficiency Measurement Techniques in Education. *Education Economics*, 9 (3), pp. 245-268, 2001.