REVIEW OF ECONOMIC PERSPECTIVES – NÁRODOHOSPODÁŘSKÝ OBZOR, VOL. 14, ISSUE 1, 2014, pp. 34–54, DOI: 10.2478/revecp-2014-0003

Using Data Envelopment Analysis: a Case of Universities

Tomáš Rosenmayer[1]

Abstract: The aim of this article is to analyse appropriateness and adequacy of use of Data Envelopment Analysis (DEA) in several research papers dealing with effectiveness of economy of universities. The Data Envelopment Analysis is an interesting method used for evaluation of technical efficiency of production units. Comparison is the basic method of this article. At the beginning, basic methodological questions of measurement and evaluation of efficiency are analysed, including definitions of terms efficiency and effectiveness, ways of measurement and formulation of appropriate indicators. Based on the given perquisites for measurement and evaluation of efficiency five articles on evaluation of efficiency of universities using DEA method, published in Canada, Australia, Great Britain, Germany and Spain in 1998 – 2008, will be assessed.

DEA is able to use more parameters of input and output to evaluate which of units under examination is the most effective, and to compare other units with it. For this, it is necessary to have a homogenous group of units. The result of assessment shows that all the examined studies focused rather on way of calculation then the point and reason of measurement. The articles contain a discussion concerning choice of appropriate indicators but do not at all deal with the issue of its construction using interventional logic; the articles do not contain any comparison of objectives of the particular universities.

Evaluation of efficiency of universities is a social construct and it will always be a subjective matter related to objectives of a particular stakeholder. This fact explains how to approach the evaluation of efficiency: it is necessary to set an objective function that means to set the objectives of a given stakeholder and his preferred results and outputs. All the studies lack this basic logic.

Key words: public services, organizational effectiveness, universities, data envelopment analysis

JEL Classification: D240

Introduction

Development of information technologies made possible for scientists in many research fields to use new methods and tools for work with extensive data sets. Similarly to other scientific branches, economy uses mathematical modelling of complex economical phenomena and systems, analyses and verifies such models, and creates predictions and

[1] Masaryk University, Faculty of Economics and Administration, Lipová 41a, 602 00 Brno; rosenmayer@mail.muni.cz

materials for optimal decisions. But the possibility of using sophisticated methods and availability of extensive data sets sometimes prevails the logic of what is examined, and why and if results of complex calculations make any sense.

One of such examples is using of Data Envelopment Analysis (DEA) which seems as one of the most suitable methods for comparison of efficiency of various units providing public services. Its use is, similarly to other methods, limited for several reasons. Generally speaking, public services are always influenced by a public policy (strategy) which is used by a government to influence a specific domain of public services (either directly by public expenditure programmes or through fiscal, legislative and other regulation mechanisms). Public services are provided by various entities where each of them has its own motives for providing the services and which are influenced by a whole range of different stakeholders.

When the author made a bibliographical research of DEA during the work on his dissertation, he discovered that some expert studies using DEA method contain many mistakes, e.g. use of inappropriate data and incorrect or incomplete interpretation of results. Although the research question of the dissertation has a different focus (What is the influence of chosen organisational statute on effectiveness of public service providers), on its concrete level, the dissertation will deal with the same theoretical – methodological problems which arise when using DEA method. The key factor for correct application of complex method is to proceed from correct theoretical basis, and that is why the topic became a theme of a separate article.

The aim of this article is to analyse by means of critical analysis suitability and appropriateness of use of Data Envelopment Analysis (DEA) in several works concerning efficiency of university management, and to suggest how to use and correctly interpret the DEA method. The article has no ambitions to analyse all the studies about university management effectiveness which as well have used DEA method. The article will focus just on the studies best available because the key role of applied sciences is to publish reliable results. Besides that, the article does not use own data, as these can be considered the analysed studies.

At the beginning, it has to be also added that the measurement and evaluation of efficiency in the context of higher education is complicated. Measuring efficiency and productivity in universities provides an indirect evaluation of public funding management informs policy making and improves university productivity and consequently public funding management. Productivity in higher education has an obvious multidimensional character as it relates to both production and dissemination of knowledge through its various activities of teaching, research, and outreach activities (Dundar, Lewis, 1998).

The logical structure of the article is built in the following way: the first part describes the issues of methodology measurement and evaluation of efficiency of public services and universities (higher education). Here, the description of DEA method will be included. The second part will contain analyses of data, methods and results of the particular scientific studies. The last part will focus on the question of what is possible to measure and evaluate with respect to the efficiency in the sector of public services, and whether at all.

Methods

Comparison is the basic method used in the article. First, it is necessary to begin with basic methodological questions of measurement and evaluation of efficiency. Although it seems that many treatises were dedicated to definition of efficiency, methods of measurement or setting of indicators, it is in these very basic building components that research studies usually make most mistakes. The result of this part of the article will thus comprise of determination of conditions for measurement and evaluation of efficiency, and these will be subsequently compared with the texts of the particular studies.

Definitions

Although efficiency is the key concept of economic studies, there are many diverse definitions of this concept which use various explanations. On general level, Mankiw says in his book Principles of economics: "Efficiency means that society is getting the most it can from its scarce resources." (Mankiw, 2007, p.32). Dictionary of terms in the book Economics (Samuelsom, Nordhaus, 1992) explains the same expression: "Use of economic resources that produces the maximum level of satisfaction possible with the given input and technology." Without ambitions of deeper analysis, it is possible to state that both of these definitions coincide in fundamentals.

Macmillan’s dictionary of modern economics mentions two explanations of the term efficiency, the first being "X-efficiency" which is defined negatively (as X-inefficiency): "the situation, where a firm does not minimalist total costs so the real output of given inputs does not reach possible maximum". The second explanation defines the Y-efficiency term which is defined negatively as well, and evaluates market profitability of a firm where upon weakening of the competition the firm fails in its function of product supplier to customers willing to pay profitable price (Pearce, 1992).

In classical textbooks of economy of public sector (Musgrave, Musgrave, 1984; Stiglitz, 2000), the concept of efficiency is identified with the so called Pareto efficiency, where economical decision is efficient when there a change cannot occur, where one subject gains without losing any other thing. Solution is efficient when the benefit of at least one person grows and the benefit of the others remains unchanged. Often we can see an understanding of efficiency as a result of a relation between volume of inputs and volume of outputs. Therefore, solution is considered as efficient when use of all disposable resources brings maximum benefit.

But the crucial question of this article is not "why examine and increase efficiency" but "how to measure and evaluate efficiency". For this purpose it is necessary to make the definitions more specific, as was already done by Dalton and Fitzpatrick (1985, p. 520):

- In their opinion, the term "technical efficiency" means to reach maximum output per given inputs. Efficiency itself is expressed as ratio of outputs and inputs, or it is measured in costs per unit. The term defined in this way refers to the maximization of production or supply of goods, but without any relation to demand, of course.
- The term "allocative efficiency" is explained as a scope of production processes estimating the state of market supply and demand. Relation to the technical

efficiency is here defined in such a way that the technical efficiency should be maximised for production of such goods, allocation of which does not reflect current conditions of demand. To illustrate the term in the public sector, public mass transport is given as an example, providing of which is a form of support for low income population of a town but according to the structure of customers it appears more as a support of middle and high income population. Besides that, it is most often used by non-residents of the town who do not pay local taxes of which the service is subsidized.

- "Social effectiveness" is understood as an integral part (evaluation criterion) of performance evaluation of governmental programs, public sector activities. The object of examination is an adequacy of objectives of public expenditure programs and its fulfilment. In comparison to technical efficiency (of efficiency of tools), the efficiency here refers to indicators related to satisfaction of customer (user), achieving of social results and objectives.

The particular terms defined by Dalton and Fizpatrick more or less comply with general theoretical concepts quoted above. While understanding of the term mostly corresponds with expression "allocative efficiency", definitional expression "technical efficiency" is the most appropriate expression for X efficiency. Although the expression "social effectiveness" does not fully tally with understanding of public sector according the Musgraves couple or Stiglitz, it already uses efficiency evaluation typical for the public sector – evaluation according to objectives. Similarly, efficiency of organisations is examined.

For easier understanding of the difference between the expressions "technical efficiency" and "social (programme, organisational) effectiveness", the following part will explain with the use of interventional logic how, and based on which criteria, it is possible to evaluate efficiency of public expenditure programmes and organisations, eventually.

Indicators of efficiency

In the case of public expenditure programmes, efficiency expenditure programmes is evaluated according to the so-called 3E module, where the three Es (Economy, Efficiency and Effectiveness) are explained in the following way:

1) Economy is the first element of the three Es model, covering the financial aspects of work being done. Economy is measured by looking at the cost of the resources consumed and the value of the output delivered.

2) Efficiency can be measured in terms of the inputs required to generate the outputs. It is about the way in which work is completed. For example, if the same work can be completed using less inputs or resources then efficiency has improved and vice versa.

3) Effectiveness can be explained in terms of what is achieved. It is about whether targets are met or not. Performing effectively means that the right work is being completed. Effectiveness is measured by setting out clear objectives before work starts and then evaluating whether the objectives have been met or not.

Similarly to other quantities, efficiency, too, is measured using appropriate indicators. When evaluating efficiency of public expenditure programmes or organisational efficiency, the following so-called interventional logic is used:

Objectives → Outcomes → Outputs → Activities → Inputs

The relations between the levels of efficiency and indicators are described in figure 1.

Content meaning of these terms is generally accepted and the same, here the definitions will be accompanied by examples from university environment:

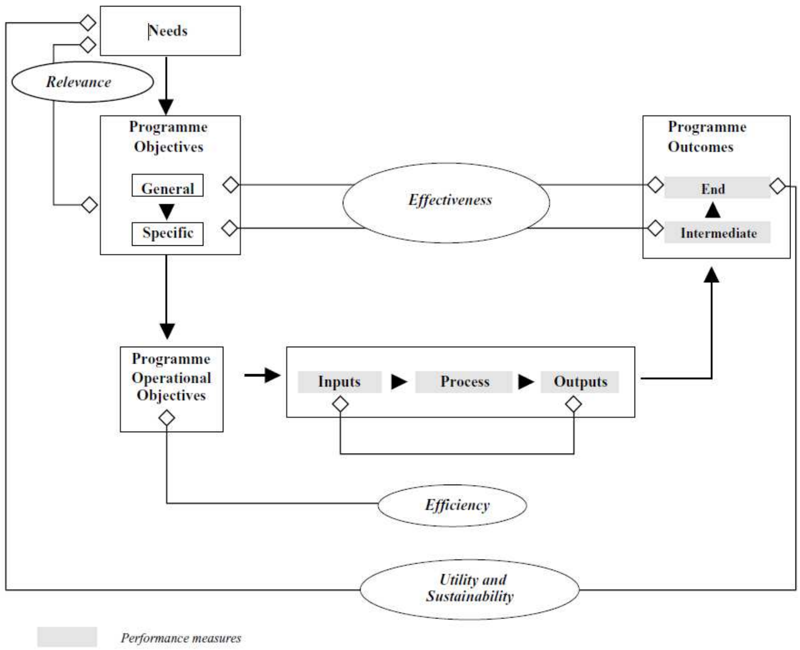
- Outcomes are intended output induced changes in social structure, environment or in characteristic of people (Robinson, 2002. p.3).
- Outputs are goods or services procured by a producer and provided for final consumer (Robinson, 2002, p.3). It is direct output of a programme or an activity (Allen, Tomassi, 463-4).
- Activities are parts of a programme with specific objectives, they make possible measuring and management of efficiency (Allen, Tomassi, 2001).
- Inputs are resources such as people, raw materials, energy, information, or finance that are put into a system to obtain a desired output.

Levels of efficiency, indicators and interventional logic can be easily explained by comparison of a car and a tractor. We use economy level when purchasing both vehicles, lower costs are considered advantageous. When comparing both vehicles as far as efficiency is concerned, we can evaluate performance of the motor at given fuel consumption, or the other way round, by fuel consumption for number of kilometres travelled. Effectiveness relates to the purpose of the vehicle use, which is different for each vehicle. So while a car is used to transport people, and we expect appropriate equipment represented by number of seats, luggage compartment and low consumption, the most important characteristics of a tractor is its use for various field agricultural works and as traction force.

When using interventional logic, we will evaluate purchase price (input) during purchase of the tractor (output) according to the number of ploughed hectares of an agricultural land per day (result). These criteria will be absent in the case of a car, of course. For each unique objective there is a unique mix of inputs and outputs. Everybody knows what sort of vehicle is needed for satisfaction of his/her needs.

The relation between result and output is always indirect. Although reaching the output always depends on the participant, the result is influenced by other factors. So in case of ploughed hectares of agricultural land, the role is played not only by the characteristics of the tractor (performance of engine, speed, width of ploughshare) but also by composition of the soil, weather and skills of its driver.

Figure 1: Measuring Efficiency



Source: Allen, R. Tomassi, D. (2001, p. 373)

The interventional logic mentioned above gives clear guidance about how to start when searching for optimal indicators for evaluation of efficiency. First of all, the objectives have to be defined by indicators. In case of public services, objectives are defined by the providers and through public policies, too. In practice, however, the objectives are rarely exactly specified and much less quantified. In the discussion part of the article we will have a look at the reasons why it is so.

Optimal indicators for universities

Unfortunately, the majority of scientific studies focused on measurement and evaluation of efficiency of universities skipped this phase (because of lack of objectives) and started directly to search the optimal mix of inputs and outputs. Therefore, no wonder that as the teaching, research and support actions are considered objectives of universities (e.g. Dundar, Lewis, 1998, McMillan, Datta, 1998). But according the interventional logic, teaching and research are clearly identified as activities. The purpose (objective) of teaching is to increase level of education, and the purpose of research is to create a new knowledge and transfer it into practice. So the indicators of number of successfully graduated students with a certain qualification or number of published papers and patents registered can become quantified results. In both of these cases, these are gross indicators which do not necessarily have to correspond with reality.

For example results of educating should always contain quality parameters. Cohn et al. (1989) maintain that numbers of students that have graduated represent an accumulated output of several years, depending on time length of degree; the efforts of non-graduated students is overlooked and there are no criteria measuring quality. Nonetheless, it must be recognized that students' achievements depend not only on the quality of teaching, but also on ability of the students and their initial qualifications. The research activity should not end just by publication of a result but by its implementation into practice as well. For example cooperation of application sector and research institutions does not necessarily end by patent or any other protection of intellectual property. Finding of optimal indicators is in case of public services always difficult nevertheless mistakes should not occur.

Some authors interchange outputs and results when striving for suitable indicators. For example, Flagg et al. (2003) considers the number of undergraduate degrees awarded a clearly important measure of the output of any university. However, an obvious shortcoming of this measure is that it fails to take any account of the quality of the degrees awarded. According to him, one way of taking quality into account would be to use the graduate unemployment rate, standardized by subject and gender mix, as an index of the quality of degrees awarded.

Indicators for publications such as number of published books, book chapters and refereed journal articles and conference proceedings counts are sometimes used as a measure of research output. Sinuany-Stern et al. (1994) and Tomkins and Green (1988) use both publications counts and grants. Sarafoglou and Haynes (1996) use number of articles and a citation impact factor. In their work on American universities, De Groot et al. (1991) incorporated a measure of both research output (bibliometric) and quality (peer review).

Full-time equivalent student enrolment and student credit hours are considered some of the most appropriate indicators of outcomes of teaching in university environment. The latter has been used (Sinuany-Stern et al., 1994) but it can have the problem that credit hours can differ significantly among programs of full-time students (e.g., science students with labs versus students of humanities), and these differences more likely to reflect input differences than learning differences.

Though various measures of research are commonly taken as measures of university output, they are often measures of an intermediate product. The real output consists only of new knowledge reached, or knowledge transfer according to the focus of the university, respectively. Lacking reliable and easily obtainable output measures, many studies substitute research grants. Tomkins and Green (1988) suggest that research grants reflect the market value of the research conducted and therefore, can be considered a proxy for output. However, the use of research income as a measure of output is problematic since such income may be considered to be an input into the research process rather than an output (Johnes, Johnes, 1993, p. 338). Ahn et al. (1988) blend this approach using state funds allocated to state institutions of higher education as input, and federal and private research funds as output. Research grants may be considered a market price that gives information on the quality as well as on the quantity of research output, e.g. see Johnes (1992).

Inputs pose fewer difficulties. Although there are many kinds of inputs (for example, academic and support staff, student services, libraries, computers, equipment and supplies, maintenance, buildings, etc.), they can usually be defined relatively well in terms of amounts or expenditures. The fact that expenditures can be a relatively complete measure of input, and are well documented, opens the possibility of studying cost efficiency. Variations in input quality, however, may not be easily distinguished (McMillan, Datta, 1998).

Nonetheless, because there is interest in the efficiency of various specific inputs, several critical inputs are usually included. Academic staff is a primary input and the largest item in university costs, and is typically incorporated in full-time equivalent numbers or as salary expenses (Ahn et al., 1988). Other separately designated inputs may include support staff, library expenditures, sometimes certain research costs or student input and plant (Ahn et al., 1988) or space (Bessent et al., 1983).

Data Envelopment Analysis

Generally, productivity depends on production technology, efficiency of production, and production environment. Data Envelopment Analysis (DEA) is focused on measuring the second, that is, production efficiency, for each production unit of a set of decision-making units (DMUs) — universities in this instance. Comparability means that the set of producers has the same objectives and is producing similar outputs using similar inputs with the same technology.

DEA is used to measure efficiency when there are multiple inputs and outputs and there are no generally acceptable weights for aggregating inputs and aggregating outputs. In the case of one input and one output, the output-input ratio reveals efficiency. If prices exist for all inputs and outputs, the value of outputs to the value of inputs (or indexes of these) can be used. A full set of prices may exist in the case of private firms. In the case of public sector production, prices typically do not exist or do not reflect social values; hence the appeal of DEA for the efficiency analysis of public operations.

The lack of prices means that DEA analysis measures technical efficiency, not economic efficiency. That is, the DEA reveals how efficiently are inputs used to produce outputs, but not whether even the efficient units could reduce costs or enhance the value of outputs by choosing different combinations of inputs or outputs. Nevertheless, information on technical efficiency is valuable for assessing and improving the performance of DMUs when price information is absent or limited.

As a technical analysis, DEA is relative. From the set of DMUs analyzed, it determines an efficient group. It still might be possible, however, to improve the technical efficiency of even those efficient units which were the best production possibilities known. However, the actual production function is not known and none is assumed. The efficient units in DEA are the most efficient of those observed, not in comparison to some ideal. Thus, the DEA efficient group is that subset demonstrating the "best practices" among a group of operating units. Inefficient DMUs are compared to those units demonstrating superior performance (McMillan, Datta, 1998).

Mathematically, DEA is a linear programming procedure for a frontier analysis of inputs and outputs. DEA assigns a score of 1 to a unit only when comparisons with other relevant units do not provide evidence of inefficiency in the use of any input or output. DEA assigns an efficiency score less than one to (relatively) inefficient units. A score lesser than one means that a linear combination of other units from the sample could produce the same vector of outputs, using a smaller vector of inputs. The score reflects the radial distance from the estimated production frontier to the DMU under consideration.

Conditions for measurement and evaluation of efficiency

The overview mentioned above of definitions of efficiency and short outline of way of measurement of efficiency, introduction of possible indicators and DEA method leads to setting conditions for making evaluation and measurement of efficiency possible:

1) Efficiency is always a relative quantity. Efficiency as a concrete value has its information value only when compared with efficiency of other alternatives; without this comparison, there is no appropriate information value. The conclusion concerning the level of relative efficiency is possible to find only when comparing the obtained value with indicator of the same construction, that elaborated:
   a. for different (past) time period for the same entity,
   b. for different entity (in the same time period),
   c. for different entity and different time period, but provided that there is similarity, comparability of conditions.

2) To express the efficiency in numbers we have to have quantifiable (numeral) values of inputs and outputs at our disposal. Although it is easier to find appropriate indicators in profitable environment where quantity of profit (in monetary units) is the indicator of efficiency, it does not mean that it is impossible to measure efficiency in public sector. Given that it is not always necessary to relate the efficiency in public services to the objectives, the indicators should be set on the basis of interventional logic.
3) The Data Envelopment Analysis is a method intended for evaluation of production efficiency using technical efficiency. The DEA method selects from a file of units the most efficient and compares with them indicators of inputs and outputs of the other units. The basic condition for performing of the DEA method is a file of homogenous units with the same production (objective) function.

Data and Results

Articles on evaluation of efficiency of universities using DEA method, published in 1998 – 2008, will be assessed herein. The studies were searched for through Google search using keywords “data envelopment analysis” and “universities”. Although this way of searching is not a sophisticated method, it corresponds with the lay approach. That means, this approach also corresponds with the article assignment which was “to analyse the best available studies”.

The following studies (full bibliographical references are available at the end of the article) were chosen:

- McMillan, M., Datta, D. (1998)
- Abott, M., Doucouliagos, C. (2003)
- Flegg, A. T; Allen D.O; Field K., Thurlow, T.W. (2003)
- Kempkes, gG, Pohl, C. (2006)
- Garcia-Aaracil, A., Palomares-Montero, D. (2008)

Data

Regarding selective file of examined universities, the evaluated articles can be divided into two groups. In the first group, there are articles examining efficiency of universities just in one year (Canada, Australia), the second group deals with a longer time horizon. The standard DEA approach has a disadvantage that it cannot distinguish between changes in relative efficiency brought about by movements towards or away from the efficiency frontier in a given year and shifts in this frontier over time. To capture these two sources of change in efficiency, all the three articles from the second group computed Malmquist indices.

The use of Malmquist indices cannot solve the methodology mistake of articles of the second group either. All of them are actually in breach with the condition that it is possible to compare efficiency between various units in different time periods only provided that there are comparable conditions. The British article even confesses that "the period 1980/81 to 1992/93 was chosen because it was characterized by major changes in public funding and in student: staff ratios."… "real funding was, in fact, cut by 8.7% between 1980/81 and 1984/85, what is more, the cuts were applied highly selectively" (Flegg et al., 2003). The British article regrettably does not explain the background of the cuts. What if the governmental cuts were caused by bigger support of research activities at the expense of education? Did universities adapt to the new way of financing? Did they change their strategy in order to fulfil objectives of public policy more efficiently? If such changes really happened it would be necessary to change the monitored indicators as well. The correct methodological approach should be based on separation of selective indicators according the change of governmental policy.

The Spanish article informs about a change in number of students: at first, the number rose in the monitored period, and descended as a result of demographic trends from 1998 (Garcia-Aracil, Palomares-Montero, 2008). It is not possible to draw conclusions regarding the influence on examined file and efficiency evaluation without additional information about impact of demographic changes on universities (number of students admitted) and the way of financing (student ratio).

The German study does not mention any changes in the environment in the years monitored, with exception of different GDP values in Western and Eastern Germany, but at least provides some information about distribution of the file of monitored units. "Due to differing faculty compositions, the universities in our sample are quite heterogeneous… Specifically, universities with engineering and medical faculties seem to have a differential and personnel structure than universities without these two faculties… On average, a single university spends about 66,700 euro per year and per graduate in the considered time period 1998-2003. When restricting the sample to universities that have engineering and a medical department, we find expenditures of 92,200 euro per graduate. Accordingly, higher education institutions without such cost-intensive faculties clearly spend the least money per graduate (17,800 euro)" (Kempkes, Pohl, 2006).

The German study is the only one which informs about the way how the selective file was determined: "private universities in Germany are highly specialized, i.e. oriented towards business management and/or medical studies, so that their inclusion represents a possible source of bias. For the same reason we also drop other specialized universities such as universities of fine arts and music. Universities of applied science are also excluded from our investigation since these are more oriented towards teaching instead of research. In particular, universities of applied science are in general not enabled to train doctoral students so that their consideration would have created a more heterogeneous sample" (Kempkes, Pohl, 2006). The Spanish study mentions elimination of 5 institutions from the file, "because of lack of data for some of the years in the period under study and due to their different structures" (García-Aracil, Palomares-Montero, 2008).

The Canadian study admits different characteristics of particular universities. "Differences among universities arise from the presence or absence of medical schools, large versus small graduate programs relative to undergraduate programs, and more versus less emphasis on research" (McMillan, Datta, 1998). The question whether or not it would be better to evaluate the efficiency for each individual group the study is answered negatively, and though the scores are comparable among categories, it is often convenient to discuss the results for the separate categories.

Regrettably, no single study contains any information concerning objectives of the particular universities. Therefore, we do not know whether the condition of homogeneity of examined institutions has been met.

Table 1: Basic information about research carried out

Country (year) / Authors	Sampling set	Years	Data processing
Canada (1998) McMillan, Datta	45	1992/3	-
Australia (2003) Abbott, Doucauliagos	36	1995	-
UK (2003) Flegg, Allen, Field, Thurlow	45	1980/81 – 1992/93	Malmquist indeces
Germany (2006) Kempkes, Pohl	72	1998 - 2003	Malmquist indeces
Spain (2008) García-Aracil, Palomares-Montero	43	1994/95 – 2004/5	Malmquist indeces

Indicators

Although there is no research study analysed herein that would contain any information about objectives of the particular universities, all of them at least contained a discussion regarding setting of the particular indicators. All studies work on the presumption that teaching and research is the main objective of universities. The Australian study furthermore adds community services, the Canadian speaks just about services, the Spanish adds, using quotation, outreach activities, and the British consultancy and other educational services. Yet all the studies further work just with indicators related to teaching and research, which can be considered as a correct approach because the so-called other activities are in fact just an intermediate.

Table 2: Comparison of inputs and outputs

Country (year) / Authors	Inputs	Outputs
Canada (1998) McMillan, Datta	· total number of full-time faculty in the three professorial ranks · total expenditure less faculty salaries and benefits · total operating expenditure and sponsored research expenditure	· total undergraduate student enrolment · graduate enrolment in master's level programs · graduate enrolment in doctoral stream programs · total sponsored research expenditures · number of active grants as a percentage of eligible faculty
Australia (2003) Abbott, Doucauliagos	· total number of academic staff · number of non-academic staff · expenditure on all other inputs than labour · value of non-current assets	· number of equivalent full-time students · number of post-graduate and under-graduate degrees enrolled · number of post-graduate degrees conferred · number of under-graduate degrees conferred · Research Quantum Allocation
UK (2003) Flegg, Allen, Field, Thurlow	· number of staff · number of undergraduate students · number of postgraduate students · aggregate departmental expenditure	· income from research and consultancy · number of undergraduate degrees awarded, adjusted for quality · number of postgraduate degrees awarded
Germany (2006) Kempkes, Pohl	· number of technical personnel · number of research personnel · financial means · total costs - research grants	· number of graduates · amount of research grants
Spain (2008) García-Aracil, Palomares-Montero	· expenses · number of academic staff · number of non-acad. staff	· number of graduates · number of publications · applied research

As for the indicators, with exception of the Canadian study, all other studies included in the 'outputs' the numbers of undergraduate or postgraduate students. Author of the article considers the numbers of undergraduate or postgraduate students as the 'results' (as was explained in the methodological part). Besides that, the Canadian study probably dealt in the best way with the outputs for research when using indicators "total sponsored research expenditures" and "number of active grants as a percentage of eligible faculty." The other studies use less specific indicators for research.

For outputs, most of the studies use standard indicators in the form of costs, and these either in more detailed classification, or in total, and numbers of employees, here again at least classified as academic and non-academic employees. With the exception of the Canadian and the Australian study, all other studies always use the above-mentioned inputs altogether, which is an incorrect procedure because inputs will surely be included into the total costs as well salary of academic and non-academic employees, thereby they will be de facto counted twice.

Interesting, yet methodologically incorrect approach is mentioned in the British study which uses numbers of undergraduate or postgraduate students as inputs. However, these are target groups, not inputs. The British study regrettably does not explain why this unusual approach was chosen.

Data modelling

The differences between the studies are in the way of processing. The German study, for example, uses two models - one contains all universities, the second excludes universities with engineering and medical faculties. This classification influences results, too. “Several of the interaction terms of faculty dummies with output/wage variables are significant, indicating that universities with medical and/or engineering faculties not only have different cost levels but also different marginal cost structures. For instance, research grants only substitute for state money in universities with engineering faculties whereas this effect is not significant in universities without medical and engineering faculty. In universities with medical faculties, research grants even crowd in additional state funds. Not controlling for the faculty composition of universities significantly biases the estimation results and thus the predictions for university level efficiency scores... For instance, in the model without controls for faculty composition, research grants have a highly significant and highly positive effect on total costs less research grants. However, after controlling for faculty structure, this effect decreases significantly in size and turns insignificant. Moreover, a negative effect of the number of graduates on total costs disappears after controlling for faculty structure” (Kempkes, Pohl, 2006).

Australian study which monitors separately the results of the use of medical and non-medical research income (as two separate research output measures) used a very similar methodology to that taken advantage of in the German study. Interestingly enough, all other combinations of outputs and inputs led to results that were similar to the basic data set (Abbott, Doucauliagos, 2003).

The Spanish study, too, worked with models but a different approach was chosen. To evaluate Spanish public universities, first, we analyze a “general model” taking total expenses, number of academic and non-academic staff as inputs, and graduates, publications and applied research as outputs. Then, in order to understand the sources of productivity changes, three additional specifications of university productivity are examined. The first focuses on “teaching-only” productivity, the second on “research-only" productivity, and the third on "industry-only" productivity. Variable definitions in both instances are identical to the "general model", but the "teaching-only" specification does not include the output publications and applied research, the "research-only" specification excludes the output graduates and applied research, and the "industry-only" specification does not include graduates and publications" (García-Aracil, Palomares-Montero, 2008). The results for the particular models differ. So while the general model showed an annual mean increase in total factor productivity of 4.6 percent for the period 1994 to 2004 across the university sector, the teaching model caught 3.8, research model 9.5 and industry model only 1.8 percent.

The Canadian study used in total six models:

- Model 1 is the basic model. It has five variables: (total undergraduate students, total student enrolment in graduate programs, total sponsored research expenditures for outputs and number of academic staff and total expenditure less faulty salaries as inputs.
- In Model 2, the graduate student variable is divided into master's level and PhD level students.
- In Model 3, faculty are divided into two groups.
- Model 4 provides the results of including a parallel subdivision of students taking science programs and those taking other programs.
- Model 5 represents efforts to introduce quality of research and of presumably (primarily) graduate education into the analysis. This is done by adding variables reflecting the faculty's success at obtaining council research grants
- Model 6 demonstrates the effect of deleting indicator "total sponsored research expenditures". Otherwise the specification is the same as for Model 5.

The authors mostly tend to the variants 3, 4, and 5, and the results bear this out. Distinguishing between master's and doctoral graduate students, distinguishing between the arts and sciences, and possibly recognizing grant success as an indicator of research quality is important in determining relative efficiency (McMillan, Datta, 1998). The Canadian study is probably the most interesting in the way how the authors examine the informative ability of the particular models.

In the British study sensibility analysis has been carried out but no interesting conclusions were brought.

Conclusions and statements

The British study focused not only on evaluation of productivity of universities in the particular years, but on search for structure of efficiency as well. "As a first step, a technical efficiency score was computed for each university for each academic year. These scores were then aggregated by calculating the weighted geometric mean. The Malmquist analysis revealed a rise of 51.5% in total factor productivity over the study period. What is interesting about this growth in total factor productivity is that it was brought about predominantly by a marked outward shift in the efficiency frontier rather than by enhanced technical efficiency whereas frontier technology improved by 39.1%, TE rose by only 8.8%" (Flegg et al., 2003).

The least concrete results were published in the Australian study. “Taken as a group, the Australian universities are performing very well against each other. The results are not substantially different between the four, three and two input models. Overall, the level of technical efficiency in the Australian university system appears to be high. However, it cannot be concluded that there is no scope for improvement in efficiency. The results indicate homogeneity in performance across the university system. They do not rule out the possibility that the entire system may be under performing. Nor can it be concluded that the Australian university system is efficient when compared to institutions overseas” (Abbott, Doucauliagos, 2003).

Results are similarly assessed by the Spanish study. “The results indicate that annual productivity growth was largely attributable to technological progress rather than efficiency improvements. Gains in scale efficiency appear to have played only a minor role in productivity gains. The fact that technical efficiency contributes little suggests that most universities are operating near the best practice frontier.

The separate analyses of teaching-only, research-only and industry-only productivity suggest that most productivity growth was associated with improvements in research rather than teaching and knowledge transfer. In turn, the increase in teaching productivity is mainly sourced from technological gains and little efficiency improvement, whereas the research gains are mostly associated with the removal of inefficiency rather than technological improvements” (García-Aracil, Palomares-Montero, 2008).

Although the Canadian study admitted that there are different categories of universities, its results for particular categories do not differ, and similar deviations were measured in all the categories. “A subset of universities - including universities from each of the three categories (comprehensive with medical school, comprehensive without medical school, and primarily undergraduate) - are regularly found efficient and a subset quite inefficient but, overall, the efficiency scores are relatively high. The average university is about 94 percent efficient but there is a possibility that, due to the modest number of observations, efficiency scores are upwardly biased” (McMillan, Datta, 1998).

And finally, the German study assesses a different state of universities in western and eastern federal republics. “With respect to German universities, our estimation results indicate that total factor productivity has been increasing more rapidly in East German universities compared to their West German counterparts. Due to the upcoming demographic changes –the number of high-school graduates in East Germany is expected to decrease by about 40% within the next 15 years the universities in East Germany must continue their good dynamic efficiency performance. On a university level, West German universities appear at the top end of our efficiency rankings” (Kempkes, Pohl, 2006).

Discussion

The Data Envelopment Analysis is an interesting method used for evaluation of technical efficiency of production units. DEA is capable of using more parameters of input and output to evaluate which of units examined is the most effective (or rather, which unit is able to transform given inputs into given outputs in the most efficient way) and to compare other units with it. To do so, a homogenous group of units is needed, and it is necessary to follow the rules for evaluation of efficiency – either to compare the measured values between the units for the same unit in different points in time, or it is possible to apply both time and unit perspective respecting condition of ceteris paribus.

Surprisingly, the given conditions were met most of all in the oldest studies - the Canadian and Australian; all other studies evaluate the data at the same time for more years and under changing conditions. In addition to that, the Canadian study admits heterogeneity of the file, yet no special measures for elimination are taken. Although the German study monitors data of universities with medical and technical faculty separately, no attempt to find any other indicators for these subjects has regrettably been done.

The Australian and Canadian studies dealt well with the input indicators; moreover, the Canadian study (as the only one) did not mistake the outputs for results (number of graduates). The worst solution was chosen in the British study which considers students as an input. The Canadian study is interesting for its approach to input and output modelling, too. On the other hand, the British study besides other significant mistakes interprets results in the worst way, too. Using the analogy from the first part of the article, what is worthwhile of knowing information that the performance of the engine of all types of vehicles increased by 10 % in the last ten years? This information can be for sure learnt by searching established innovation in the car industry, and there is no need to do any extensive research and make exacting analyses.

In cases of all studies, a treatise on roles of universities is absent; there is no comparison of their objectives. It is hard to believe that all universities have the same or very similar objectives. Some of them target more on preparing their students for fulfilment on labour market, others prefer basic research and adjust curricula accordingly, whereas technical universities cooperate with industry in the field of innovations. The way of processing by models distinguishing orientation of universities and the results of German and Spanish study confirm the assumption. As the previous part has shown, rather than focusing on the point and reason of measurements, all studies under examination herein focused on the way of calculation. The articles contain a discussion concerning choice of appropriate indicators but do not deal with the basic logic of its construction at all.

What is actually the use of discovered findings? This is a separate question. When returning back to our analogy, technical efficiency used by Data Envelopment Analysis measures performance of engine in relation to a fuel unit consumed. It can be an interesting benchmark for two vehicles of the same type but using performance of engine, it is hardly possible to compare a car and tractor. As has been mentioned already, setting the indicators for a car or tractor is easy because we know for what purpose we want use the vehicle. Nevertheless, the situation is more complicated in the case of public services. We have already mentioned that the purpose of non-profit organisations or public expenditure programmes is usually unknown because the objectives are not formulated at all, or are intentionally formulated vaguely. Why is this so? Each organisation is result of a social construct created by various parties and interactions between particular stakeholders. Each stakeholder can then perceive the objectives of a public programme or a specific organisation, and then efficiency, too, in a different way.

Balser and McClusky (2005, p.296) correctly point out the fact that particular stakeholders evaluate efficiency of behaviours of a specific organisation according to fulfilment of their expectations – objectives. Generally, the stakeholder theory emerged as a response to principal-agent problem of profit-making firms. Although the stakeholder theory perceives stakeholders as lords, they become only one group among various lords that is company stakeholders (Abzug and Webb, 1999, p.419). Simple relation of shareholders and managers was thus substituted by multilayered relations of an organisation and its stakeholders. In the case of universities, it is possible to divide the stakeholders into internal – academic and non-academic employees, who have the most direct claims of organisational resources, and external ones – students, parents, graduates, government, unions, local communities and the public in general. This group does neither own the organisation, nor work for it and yet they have an interest in it.

It is more than logical that each group of stakeholders will force management of the university to accomplish just their objectives. As Steinberg (2006, p. 133) refers, objectives in one organisation are often multiple and opposed without considering what weight one objective has against the other. The objectives are often formulated intentionally vaguely because too specific objectives can lead to alienation of a specific group of stakeholders, or to conflicts between various stakeholders. These thoughts add complexity into the models of nonprofits organisations, which is generally unavailable in models of for profit companies, where the decisions are considered in order to maximise profit (Brown and Sliwinski, 2006, p.141).

On one hand, the stakeholder perspective makes the complicated situation for measurement of public services efficiency even more complex. On the other hand, it makes it possible to understand what is possible to measure, and evaluate and what is not. Therefore, it makes sense to evaluate efficiency from a subjective viewpoint as well.

For example, from the viewpoint of an organisation it is possible to investigate to what extent particular subjects are able to achieve their objectives. But such a comparison has a meaning just in discovering how organisational behaviours in case of particular subjects work, i.e. how organisational objectives are set, how precisely are they formulated, how are they evaluated and what consequences are drawn from the evaluation.

It will be surely interesting for donors, clients, governmental agencies but for general public as well to compare how particular economic subjects contribute to achieving of objectives chosen by particular stakeholders. Comparison of efficiency will make it possible to establish objective function of a third party (arbiter judge) which would monitor efficiency in the strict sense, i.e. what is the ratio of chosen outputs against inputs in a particular organisation. This should have been the approach of the studies examined: first to set objective function and then, using interventional logic, to search for appropriate indicators. Certainly, some studies would not have made the mistake of confusing the outputs and inputs.

In the similar way, the efficiency of universities and other public service providers should be evaluated by the government. If a government intends to evaluate efficiency of invested means, then at the first there have to be specified the objectives to be achieved. It is certainly not appropriate to use Data Envelopment Analysis to evaluate efficiency of the very public expenditure programmes. For this purpose, it is much more appropriate to use for example cost-benefit analysis. It is possible to evaluate technical efficiency in case of public expenditure programmes, too, but it always will be just evaluation of procedure, or how efficiently were inputs transformed into outputs.

But Data Envelopment Analysis can tell the government which organisations contributed to achievement of objectives of public expenditure programme in best and worst way, the only correct procedure then being evaluation of efficiency based on granted subsidies as the input, and evaluating of outputs of a public expenditure programme. Further investigation should be then focused on searching for causes of the state. Do universities accommodate their strategy to the objectives of governmental programmes? Do they use the same technologies? Do they evaluate their organisational behaviours on regular basis? Do they motivate their employees towards achieving the objectives? This information is useful for further improvement of public expenditure programmes setting and therefore, it is in the interest of government to examine organisational efficiency of subjects, especially when providing public services financed from governmental resources.

Conclusion

Social (Programme, Organizational) Effectiveness is always related to some objectives and differs from the technical efficiency by focusing on purpose. In the case of Social Effectiveness, it is always a subjective quantity. If two different economic units have their unique objectives, then in the case of two distinct subjects, two identical approaches can have totally different impact on objectives. That clearly follows that from an objective viewpoint, it is impossible to compare if a one economic subject is more or less efficient than the other. For each objective function of the economic subject there is a unique mix of outputs. The term "subjective", besides that, means that each stakeholder can perceive organisational objectives in a different way, and therefore perceive the efficiency in a different way. This fact explains how to approach evaluation of efficiency: although the Social Effectiveness is a subjective quantity, it is possible to measure it and evaluate it from the viewpoint of a third party. But it is necessary to create an objective function that intends to set objectives of a given stakeholder and his preferred results and outputs.

From the conclusion it is possible to deduce three general recommendations. For universities as non-profit organisations, it means to work with all stakeholder groups on daily basis. This is in the concrete extreme difficult because specific persons can play many roles (academic employee, parent, and taxpayer) and these can change in time. But for evaluation of success rate of universities it is necessary that these organisations specified their objectives (if possible as SMART: specific, measurable, accurate, realistic and time-bounded). It is the necessary to explain the objectives to the particular stakeholders to let them adopt the objectives as theirs.

There is a similar task for public administration, too. If public administration should be transparent towards citizens - taxpayers, it has to publish objectives of all public expenditure programmes and evaluate these on regular basis. Otherwise it is not possible to find out if public means were used efficiently. Using DEA method for processing of efficiency makes it possible to arrange the supported units in order according to the rate of its contribution to achieving of the objectives.

The last recommendation is to carry out the Data Envelopment Analysis cautiously. It is always necessary to pay attention to fulfilment of all its conditions when defining the task. Only then it is reasonable to carry out the analysis and its use can have an informative value.

References

ABOTT, M., DOUCOULIAGOS, C. (2003). The efficiency of Australian universities: a data envelopment analysis. *Economics of Education Review*. 22, pp. 89-97. DOI: 10.1016/S0272-7757(01)00068-1

ABZUG, R., WEBB, N. J. (1999). Relationships Between Nonprofit and For-Profit Organizations: A Stakeholder Perspective. *Nonprofit and Voluntary Sector Quarterly*, 28 (4), pp. 416-431.

AHN, T., CHARNES, A., COOPER, W.W. (1988). Some Statistical and DEA Evaluations of Relative Efficiencies of Public and Private Institutions of Higher Education, *Socio-Economic Planning Sciences*. 6, pp. 259-269.

ALLEN, R., TOMASSI, D. (2001). *Managing public expenditure*. Paris : OECD.

BALSER, D., MCCLUSKY, J. (2005). Managing Stakeholder Relationships and Nonprofit Organization Effectiveness. *Nonprofit Management & Leadership*, 15(3), pp. 295-315. DOI: 10.1002/nml.70

BESSENT, A.M., BESSENT E.W., CHARNES A., COOPER W.W., THOROGOOD, N.C. (1983). Evaluation of Educational Program Proposals by Means of DEA, *Educational Administration Quarterly*. 19, pp. 82-107.

BROWN, E., SLIWINSKI, A. (2006). Nonprofit Organizations and the Market. In POWELL, W.W., STEINBERG, R. (eds.) The Nonprofit sector. A Research Handbook. 2nd. ed. New Haven and London : Yale Univerzity Press.

COHN, E., RHINE, S. L.W., SANTOS, M.C (1989). Institutions of Higher Education as Multi-Product Firms: Economies of Scale and Scope," *Review of Economics and Statistics*. 71, pp. 275-290.

DALTON, T. C., FITZPATRICK, L. C. (1985). Productivity measurement and the public organization. *Review of Policy Research*. 4, pp. 519–531. DOI: 10.1111/j.1541-1338.1985.tb00252.x

DE GROOT, H., MCMAHON, W.W. , VOLKWEIN, J.F. (1991). The Cost Structure of American Research Universities. *Review of Economics and Statistics*. 73, pp. 424-431.

DUNDAR, H., LEWIS, D.R. (1998). Determinants of research productivity in higher education. *Research in Higher Education*. 39(6), pp. 607-631.

DENISTON, O. L., ROSENSTOCK, I. M., WELCH, W., GETTING, V. A. (1968). Evaluation of Program Effectiveness. *Public Health Reports*. 83(4), pp. 323-335. DOI: 10.2307/4593279

FLEGG, A. T; ALLEN D.O; FIELD K., THURLOW, T.W. (2003). Measuring the Efficiency and Productivity of British Universities: An Application of DEA and the Malmquist Approach. University of the West of England, Department of Economics, series Discussion Papers n. 304.

GARCIA-ARACIL, A., PALOMARES-MONTERO, D. (2008). Evaluation of Spanish Universities: Efficiency, Technology and Productivity Change. Paper presented in the Prime-Latin America Conference at Mexico City, September 24-26th 2008.

JOHNES, G. (1992). Performance Indicators in Higher Education: A Survey of Recent Work, *Oxford Review of Economic Policy*. Vol. 8, pp. 19-34. DOI: 10.1093/oxrep/8.2.19

JOHNES, J., JOHNES, G. (1995) Research Funding and Performance in UK University Departments of Economics: A Frontier Analysis. *The Economics of Education Review*. 14, pp. 301-14. DOI: 10.1016/0272-7757(95)00008-8

KEMPKES, G., POHL, C. (2006). The Efficiency of German Universities – Some Evidence form Non-parametric and Parametric Methods. Ifo Working Paper No. 36.

MANKIW N. G. (2007). *Principles of economics*. Mason : Thomson Higher Education.

MCMILLAN, M., DATTA, D. (1998). *The Relative Efficiencies of Canadian Universities: A DEA Perspective. Canadian Public Policy*. University of Toronto Press, 24(4), pp. 485-511.

MUSGRAVE, R. MUSGRAVE, B. (1984). *Public finance in theory and practice*. 4. ed. New York : McGraw-Hill.

PEARCE, D. (ed.) (1992). *Macmillan dictionary of modern economics*. 4th ed. London ; Basingstoke : Macmillan.

ROBINSON, M. (2002). Output-Purchase Funding and Budgeting Systems in the Public Sector. *Public Budgeting & Systems*. 22, pp. 17-33

SAMUELSON P., NORDHAUS, W. (1992). *Economics*. 14th ed. New York : McGraw-Hill.

SARAFOGLOU, N., HAYNES, K.E. (1996). University Productivity in Sweden: A Demonstration and Explanatory Analysis for Economics and Business Programs, Annals of Regional Science. 30, pp. 285-304. DOI: 10.1007/BF01580523

SINUANY-STERN, Z., MEHREZ, A., BARBOY, A. (1994). Academic Departments' Efficiency via DEA, *Computers and Operations Research*. 21, pp. 543-56. DOI: 10.1016/0305-0548(94)90103-1

STEINBERG, R. (2006). *Economic Theories of Nonprofit Organizations*. In POWELL, W.W., STEINBERG, R. (eds.) *The Nonprofit Sector: A Research Handbook*. New Haven & London : Yale University Press.

STIGLITZ, J. (2000). *Economics of the public sector*. 3rd ed. New York : W. W. Norton & Company.

TOMKINS, C., GREEN, R. (1988). An Experiment in the Use of Data Envelopment Analysis for Evaluating the Efficiency of UK University Departments of Accounting. *Financial Accountability and Management*. 4, pp. 147-164. DOI: 10.1111/j.1468-0408.1988.tb00066.x