# Efficiency: Concepts, Empirical Investigation and Applications in Greek Hospitals

**Kyriaki Sotiriadou,**
Captain, Hellenic Army,
MSc Economics and Management

## Abstract

In this paper the concepts of technical efficiency, distributional efficiency and scale efficiency are presented. Following a brief note on the empirical estimation issues, there is a discussion on the implementation of these concepts in health units and a brief presentation of the main conclusions of the empirical studies on the efficiency of Greek hospitals.

Key - words: technical efficiency, allocative efficiency, scale efficiency, efficiency of Greek hospitals

## Introduction

The concept of efficiency is widely used in economics and refers to the optimum use of resources in a productive process. In other words, a production process (or production unit) is effective when is organized in a way that given the production goals set there is no alternative procedure to produce greater benefit when all costs are taken into account (Shubik, 1978).

According to Farell (1957), efficiency comprises two components: technical efficiency and allocative efficiency. These two elements together make up the overall efficiency. Further, taking into mind the scale of the production process and the form of returns (fixed or variable) that the technology has we come up eith the concept of scale efficiency, which refers to selecting the optimal scale of production. These concepts are linked, and indeed, it is often confusing to use (Rutkauskas and Paulavi ́ien‰, 2005; Shubik, 1978). However, they are widely used to assess the economic units in general and health units in particular. In this paper we analyze these concepts and indicate their relationship diagrammatically. We also make a note on the empirical measurment of efficiency, with emphasis on thre empirical research of evaluating of health units and services in Greece.

## 2. Types of efficiency

### 2.1. Technical efficiency

In general, technical efficiency refers to the possibility of an economic unit to produce the best possible product from a given basket of inputs and given technology. Alternatively, technical efficiency is defined as the ability to produce a given quantity of product with the least possible amount of inputs. Therefore, there is inherent an element of comparison in the definition of technical efficiency. In order to decide whether a production process is technically efficient, it compares the actual with the optimal inputs and outputs. Usually this comparison takes the form of a ratio:

**Y/Y***, where Y is the actual product obtained and Y* is the maximum product obtained with a given input vector X

or

**X*/X**, where X is the actual inputs and X* are the minimum inputs to produce a given output Y

These expressions imply that when an economic unit is technically efficient then it operates on the production frontier, which is defined as the potential production function. This function has the form

$$f(x) = \max\{y: y \in P(x)\} = \max\{y: x \in L(y)\}$$

Where $x = (x_1, x_2, \ldots, x_n) \in R_+^n$ is a vector input, $y = (y_1, y_2, \ldots, y_m) \in R_+^m$ is a vector of outputs, P(x) is total output and L(y) is the set of inputs (Silberberg, 1978; Varian, 1992).

On this basis, the measurement of technical efficiency of an economic unit has to do with measuring the distance of the combination of actual inputs and outputs of the unit from the potential production function. To measure this distance we use the so-called distance functions (Shephard, 1953), which may be defined with respect to either the inputs or the outputs.

The input distance function is defined as (Shephard, 1953; Lovell, 1993; Färe and Primont, 1995; Kumbakhar and Lovell, 2000):

$$D_i(x, y) = \max\left\{\lambda: \frac{x}{\lambda} \in L(y)\right\}$$

The parameter λ (with λ ≥ 1), refers to technical efficiency. If λ = 1 then we have technical efficiency, while if λ > 1 then we have technical inefficiency. Thus, subtracting 1 from λ we have the degree of technical inefficiency and therefore the distance to have a technically efficient production unit.

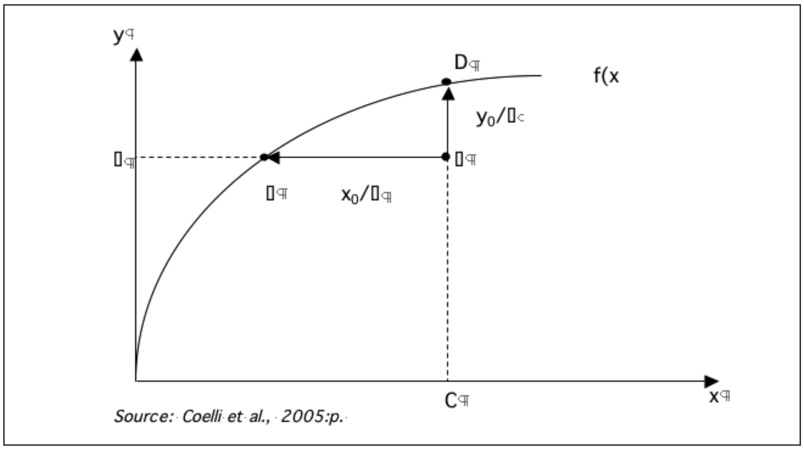
The output distance function is defined as (Färe and Primont, 1995; Kumbakhar and Lovell, 2000):

$$D_o(x,y) = \min\left\{\mu : \frac{y}{\mu} \in P(x)\right\}$$

The parameter μ (with μ ≤ 1), refers to technical efficiency. If μ = 0 then we have technical efficiency, whereas if μ <1 we have technical inefficiency.

In Diagram 1 we have a production function for one input and one output. Output $y_0$ is produced with input $x_0$, but may also be produced with a smaller quantity x1, i.e. a quantity $x_0/\lambda$. Therefore, there is a distance from the technically efficient production. Conversely, $x_0$ can produce output $y_0$, but can also produce greater output y1, i.e. a quantity $y_0/\mu$. Therefore, there is a distance from the technically efficient production.

*Diagram 1. Output and input technical efficiency*




Output technical efficiency is given by the ratio CA/CD and input technical efficiency is given by the ratio EB/EA. As A approaches D the distance AD diminishes and the ratio CA/CD tends to unity and thus production becomes technically efficient in terms of output. Similarly, as A approaches B the distance AB diminishes and the ratio EB/EA tends to unity and thus production becomes technically efficient in terms of inputs. We see therefore that the technical efficiency with respect to either inputs or outputs takes values between zero and one. Fully technically efficient production has the value of one.

The aforementioned refer to the very simple case of a production process which has one input and one output. Diagram 2 shows the measurement of output technical efficiency when one input x is used to produce two outputs y1 and y2. The ratio OA/OB is equal to the degree of technical efficiency, in other words, the distance AB is the technical inefficiency, i.e. whether it could increase output without using a larger quantity of input.

*Diagram 2. Measuring technical efficiency in outputs (one input, two outputs)*

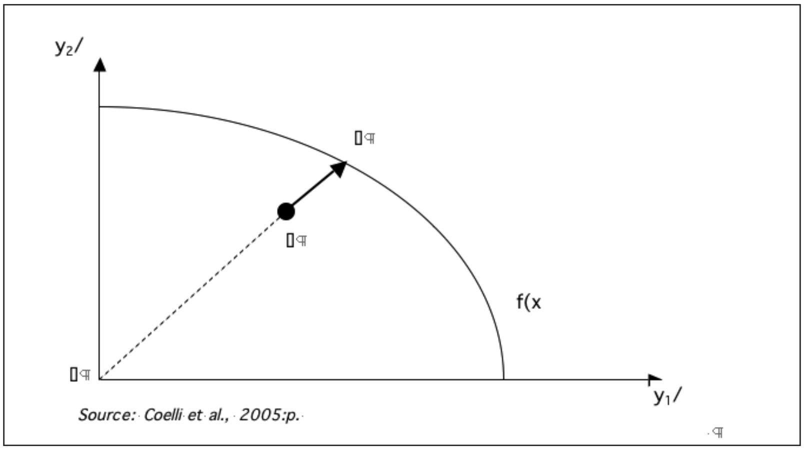
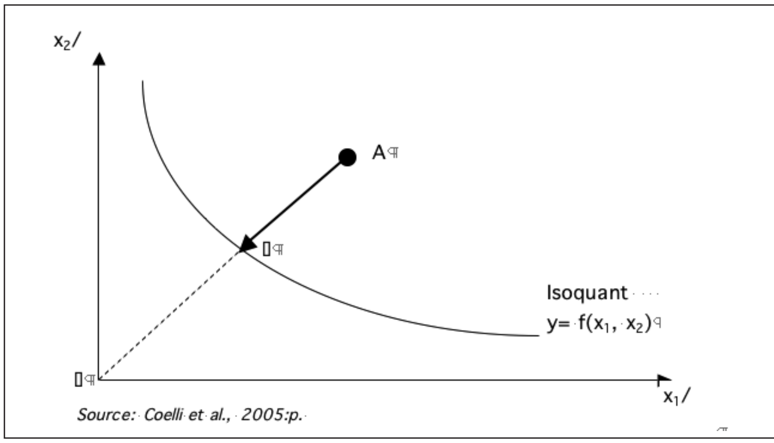



Diagram 3 shows the measurement of input technical efficiency in the case of a production process where two inputs x1 and x2 produce one output y.

The ratio OB/OA is equal to the degree of technical efficiency, in other words, the distance AB is the technical inefficiency, i.e. by how much inputs could be proportionately reduced without reducing the quantity of output procuced.

*Diagram 3. Measuring technical efficiency in inputs (two inputs, one output)*

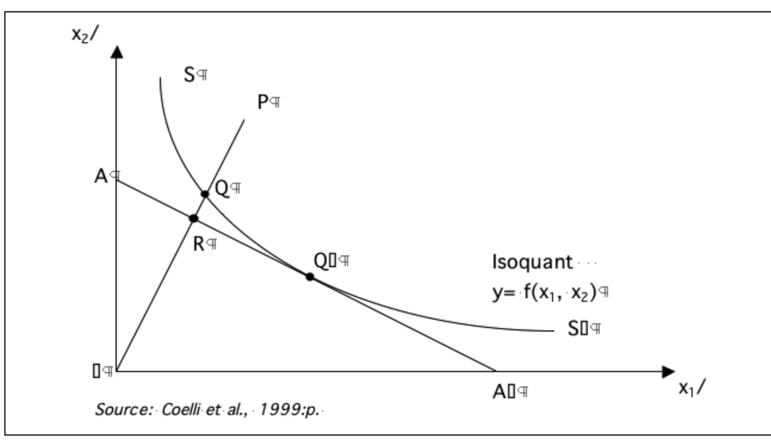


*Source: Coelli et al., 2005:p.*

## 2.2. Allocative efficiency

The allocative efficiency, or price efficiency according to Farell (1957), refers to the ability of an economic unit to use the optimal amounts and ratios of inputs given their cost. In other words, we have allocative efficiency when given the input prices an economic unit produces outputs that maximize revenue, or when the mix of inputs minimizes cost (Farell, 1957). Diagram 4 shows a production process with two inputs x1 and x2 which produce an output y. Line AAÏ is the ratio of input prices, while isoquant y = f(x1, x2) dirives form the production function and represents all possible combinations of inputs x1 and x2 in order to produce a given output level.

*Diagram 4. Distributional efficiency*



*Source: Coelli et al., 1999:p.*

All points along the isoquant are technically efficient, as, by default, it represents production of the highest possible product from the inputs. Points to the right of the isoquant are technically inefficient. For example, at point P we have technical inefficiency which is reflected in the distance PQ. At point Q, however, although we have technical efficiency, we have allocative inefficiency because we can produce the same quantity of product with less cost at point QÏ. The distance RQ is the reduction in costs we would have if production was at the allocatively efficient point QÏ (Lovell, 1993; Colli et al., 2005).

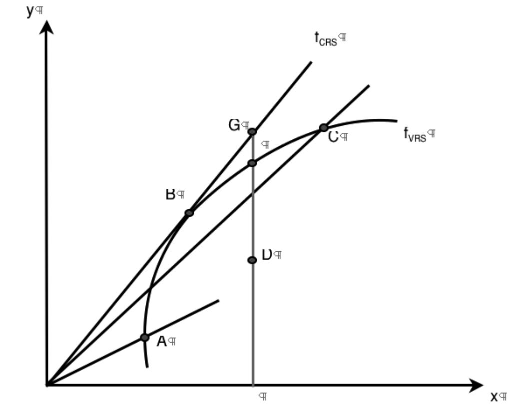
## 2.3. Total efficiency

The combination of technical and allocative efficiency gives us the overall efficiency. In terms of Diagram 4, the overall efficiency is the ratio OR/OP, whereas the cost that can be saved is given by the distance RP. As the point of production moves from P to QÏ, more cost is saved.

## 2.4. Scale efficiency

A production unit could operate at a small scale, where the production function exhibits increasing returns or at a large scale, where the production function exhibits decreasing returns. Efficiency can be improved in both cases if the scale of productions changes (Hannesson, 2005). Diagram 5 shows a production technology (production function) with constant returns to scale fCRS and a production technology with variable returns to scale fVRS.

Diagram 5. Scale efficiency




Source: Coelli et al., 1999:p 137; 2005:p.

Points A, B, C and E, which are on the function fVRS are technically efficient, in contrast to point D, which displays technical inefficiency equal to the distance DE. Points A, B and C have different average product which is equal to yA/xA, yB/xB and yC/xC, respectively. The average product can also be measured by the slope of the ray passing through each point. The difference in the average cost at these points is due to the existence of economies of scale. In point A we have increasing returns to scale and average product may increase as we move to point B. The opposite is true in point C, where we have decreasing returns to scale and average product can be increased if we decrease the scale of production to point B. Therefore, Point B corresponds to the maximum average product and is the point of optimal production scale. The production function with constant returns to scale is identical to the radius through point B and is essentially the tangent of the production function with variable returns to scale.

## 3. Methods of empirically estimating efficiency

In the empirical research there are two main approaches to estimate efficiency, as shown in Diagram 6 (Fried et al. 2008; Greene, 2008; Thanassoulis et al., 2008; Färe et al., 2008).

The main feature of the parametric and econometric methods is the use of stochastic equations, which separate the effect of random error and inefficiency (Fried et al., 2008). On the other hand, non-parametric methods help to avoid confusion between specification errors and inefficiency (Fried et al., 2008). The non-parametric methods include the Data Envelopment Analysis (DEA), developed by Charnes, Cooper and Rhodes (1978) and the Free Disposal Hull method (FHD), developed by Deprins, Simar and Tulkens (1984).

Understandably, a question arises as to which method is best and whether the results would be different if the same set of data is used with the parametric and non-parametric method. Several empirical studies have raised this question; the answer is that the better and more complete the datasets the closer are the results of both methods (Fried et al., 2008). For example, in a study of U.S. banks (Bauer et al. 1998), a very high degree of correlation was found between the results of two alternative parametric methods and between two alternative non-parametric methods, while a relatively high degree of correlation was found between the results of parametric and non-parametric methods.

## 4. Efficiency in health units

The literature of empirical studies on efficiency is very extensive and covers all sectors of the economy. Specifically in health, these studies started in the decade of 1960 (Lave and Lave, 1970), while in the last two decades are multiplied at an increasing rate. Hollingsworth (2003) presents a review of 188 relevant studies up to 2002, most of which refer to the estimation of technical efficiency using non-parametric methods. In a previous review, about five years ago (Hollingsworh et al., 1999), the number of studies reviewed was less than half. Both these reviews highlight the difficulties concerning the nature of the data and in particular the difficulty of measuring actual outcome, as well as an increased likelihood of severe bias due to omission of variables. Most studies are dealing with the efficiency of hospital units and use as inputs the number of physicians, the number of nurses, the number of beds and the cost of medicine and sanitary materials, whereas as outputs they use the length of stay in the hospital and the number of hospitalized patients. Quite common is the use of dummy variables to characterize the legal status of the hospital (private or public), the location (urban or regional), size, etc.

### 4.1. Empirical studies on the efficiency of hospitals in Greece

Most studies on the efficiency of hospitals in Greece are using the method of DEA, while some use the econometric - parametric approach (Table 1). There are studies that generally referred to hospitals and others that focus on some particular departments, e.g. nephrology, cardiology, etc. (Economou et al., 2007).

It is perhaps worth noting at this point some limitations of the analysis. First, the results of the studies are hardly comparable across countries, as not only the methodologies are different from country to country, but also the type of hospitals and the broader institutional framework. It should also be noted that because of the difficulty of measuring the actual effect of treatment, which is to improve the health status of citizens, it is common practice to use proxies such as the number of incidents, diagnostic checks, etc. The same happens with the inputs, with commonly used proxies the number of doctors, nurses, etc. This of course does not include factors of quality of service, which is critical in achieving health outcomes for the patient and in general for the efficiency of the hospital.

## 5. Conclusion

The measurement of efficiency as a component of economic evaluation is increasingly widespread application in health economics worldwide, especially in light of the increasing health costs that create fiscal problems in all countries and bring forth the requirement to all stakeholders and the society at large to know whether the resources allocated ensure the achievement of optimal results.

The analysis of efficiency is a necessity for Greece, where there are gaps and significant cost inefficiencies, as shown by studies of individual researchers at academic level. Public health authorities at the central government level but also management authorities of health units should incorporate the relevant measurements on a regular basis in their daily operation, and to adopt practices of economic evaluation for health programs and the various alternative therapies. This will provide valuable tools in both the management of health units to make rational decisions, and central government to effectively exercise their supervising role, leading ultimately to the efficient functioning of the health system and enhance the services provided to the citizens.

## References


Aletras, Vassilios, Nick Kontodimopoulos, Athanasios Zagouldoudis and Dimitris Niakas (2007), "The short-term effect on technical and scale efficiency of establishing regional health systems and general management in Greek NHS hospitals," Health Policy, 83, pp. 236-245.

Athanassopoulos, A.D., C. Gounaris and A. Sissouras (1999), "A descriptive assessment of the production and cost efficiency of general hospitals in Greece," Health Care Management Science, 2, pp. 97-106.

Athanassopoulos, Antreas and Chrysostomos Gounaris (2001), "Assessing the technical and allocative effciency of hospital operations in Greece and its resource allocation implications," European Journal of Operational Research, 133, pp. 416-431.

Bauer, P.W., A. N. Berger, G. D. Ferrier and D. B. Humphrey (1998), "Consistency Conditions for Regulatory Analysis of Financial Institutions: A Comparison of Frontier Methods," Journal of Economics and Business 50(2), pp. 85-114.

Charnes, A., W.W. Cooper and E. Rhodes (1978), "Measuring the efficiency of decision making units," European Journal of Operational Research, 2(6), pp. 429-444.

Coelli T., D.S. P. Rao and G.E. Battese (1999), An Introduction to Efficiency and Productivity Analysis, Kluwer Academic Publishers, Boston.

Coelli T., D.S. P. Rao, G.E. Battese and C. O'Donnell (2005), An Introduction to Efficiency and Productivity Analysis, 2nd edition, Springer, New York.

Deprins, D., L. Simar and H. Tulkens (1984), Measuring Labor-Efficiency in Post Offices, in M. Marchand, P. Pestieau and H. Tulkens (eds.), The Performance of Public Enterprises: Concepts and Measurement, North-Holland, Amsterdam.

Economou, Charalampos and Claude Giorno (2009), Improving the performance of the public health care system in Greece, OECD Economics Department Working Paper No. 722, Paris.

Economou N.A. J. Tountas and D. Niakas (2007), "Greek Studies of economic evaluations and efficiency in health care," Archives of Hellenic Medicine, 24(1), pp. 48-57, (in Greek).

Färe, Rolf and Daniel Primont (1995), Multi-output Production and Duality: Theory and Applications, Kluwer Academic Publishers, Boston.

Färe, Rolf, Ahawna Grosskopf and Dimitri Margaritis (2008), Eficiency and Productivity: Malmquist and More, in Harold Fried, C.A. Knox Lovell and Shelton Schmidt (eds.), The Measurement of Productive Efficiency and Productivity Growth, Oxford University Press, NY.

Farrell, M.J. (1957), "The Measurement of Productive Efficiency," Journal of the Royal Statistical Society, Series A (General), 120(3), pp. 253-290.

Fried, Harold, C.A. Knox Lovell and Shelton Schmidt (2008), Efficiency and Productivity, in Harold Fried, C.A. Knox Lovell and Shelton Schmidt (eds.) The Measurement of Productive Efficiency and Productivity Growth, Oxford University Press, NY.

Giokas, Dimitris I. (2001), "Greek hospitals: how well their resources are used," Omega, 29, pp. 73-83.

Greene, William (2008), The Ecnometric Approach to Efficiency Analysis, in Harold Fried, C.A. Knox Lovell and Shelton Schmidt (eds.), The Measurement of Productive Efficiency and Productivity Growth, Oxford University Press, NY.

Hannesson, Rögnvldur (2005), Does Scale Efficiency Tell Us Anything About Optimal Scale?, SNF Working Paper No. 23/05, Institute for Research in Economics and Business Administration, Bergen, Norway.

Hollingsworth, Bruce (2003), "Non-Parametric and Parametric Applications Measuring Efficiency in Health Care," Health Care Management Science, 6, pp. 203-218.

Hollingsworth, Bruce, P.J. Dawson and N. Maniadakis (1999), "Efficiency measurement of health care: a review of non-parametric methods and applications," Health Care Management Science, 2, pp. 161-172.

Kontodimopoulos, Nick, Giorgos Moschovakis, Vassilis Aletras and Dimitris Niakas (2007), "The effect of environmental factors on technical and scale efficiency of primary health care providers in Greece," Cost Effectiveness and Resource Allocation, 5(14).

Kumbhakar, Subal and C.A. Knox Lovell (2000), Stochastic Frontier Analysis, Cambridge University Press, UK.

Lave, Judith R. and Lester B. Lave (1970), "Economic Analysis for Health Service Efficiency: A Review Article," Applied Economics, 1(4), pp. 293-305.

Lovell, Knox (1993), Production frontiers and productive efficiency, in Harold O. Fried, C.A. Knox Lovell and Shelton S. Schmidt (eds.), The Measurement of Productive Efficiency, Oxford University Press, UK.

Polyzos, Nicholas M. (2002), Striving towards efficiency in the Greek hospitals by reviewing case mix classifications," Health Policy, 61, pp. 305-328.

Rutkauskas, Jonas and Eimen‰ Paulavi´ien‰ (2005), "Concept of Productivity in Service Sector," Engineering Economics, 3(43), pp. 29-34.

Shephard, R.W. (1953), Cost and Production Functions, Princeton University Press, Princeton.

Shubik, Martin (1978), "On Concepts of Efficiency," Policy Sciences, 9, pp. 121-126.

Silberberg, Eugene (1978), The Structure of Economics: A mathematical analysis, McGraw-Hill, New York.

Thanassoulis, Emmanuel, Maria C.S. Portela and Ozren Despiç (2008), Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis, in H. Fried, C.A. K. Lovell and S. Schmidt (eds.), The Measurement of Productive Efficiency and Productivity Growth, Oxford University Press, NY.

Varian, Hal R. (1992), Microeconomic Analysis, 3rd edition, W.W. Norton & Company, New York.

Zavras, Athanasios I., Georgios Tsakos, Charalabos Economou and John Kyriopoulos (2002), "Using DEA to Evaluate Efficiency and Formulate Policy Within a Greek National Primary Health Care Network," Journal of Medical Systems, 26(4), pp. 285-292.

**Table 1. Empirical studies of efficiency of Greek hospitals**

| Study | Methodology | Number of hospitals | Time period | Main results |
|---|---|---|---|---|
| Aletras et al., 2007 | DEA<br>Inputs: number of physicians, number of other staff, number of beds<br>Outputs: number of inpatients, number of outpatients, number of operations | 51 NHS hospitals | Years 2000 and 2003 | Technical efficiency and scale efficiency are reduced after the NHS reforms introduced by Law 2889/2001 |
| Athanassopoulos et al., 1999 | DEA<br>Inputs: number of pathologists, surgeons, microbiologists, nursing and administrative staff number of beds<br>Outputs: number of patients, number of diagnostis tests, number of clinical tests | 98 NHS hospitals | 1992 | It is possible to improve both technical and cost efficiency |
| Athanassopoulos and Gounaris, 2001 | DEA<br>Inputs: number of physicians, number of nursing and administrative staff, administration costs, cost of medicines<br>Outputs: number of patients, number of diagnostis tests, number of clinical tests | 98 NHS hospitals | 1992 | Significant inefficiencies in the system that cost about 100 million Ecu in 1992 prices |
| Giokas, 2001 | DEA and econometric estimation of cost functions | 91 public hospitals | 1992 | Around 20% of the cost can be saved, while the inefficient hospitals can provide the same services at a reduced daily per patient cost by 26%. Also, 4,1% of health expenditure as a percentage of GDP is due to inefficient funcioning of the hospitals |
| Kontodimopoulos et al., 2007 | DEA<br>Inputs: number of physicians, number of nursing and administrative staff Outputs: number of patients, number of diagnostis tests | 133 NHS health centers and 118 IKA clinics | 2004 | The IKA clinics have higher technical and scale efficiency than the health centers. About 75% of the units operate with increasing returns to scale |
| Economou and Giorno, 2009 | DEA | Health system | 1990 – 2006 | The Greek health system form the 3rd potition in 1990 among the OECD countries fell in the 12th place in 2006 in terms of efficiency |
| Polyzos, 2002 | Econometric estimation of two equations with dependent variable the average lenth of stay and cost of treatment per patient. Number of observations 104.688 people | 22 hospitals (average length of stay equation)<br>127 NHS hospitals (per patient cost of treatment equation) | 1995 and 1993 for the two equations respectively | Hospitals with 250 – 400 beds are the most efficient while the main factors affecting the average length of stay are hospital size and type, number of specialists, access to outpatient clinics, age, sex and family status of the patient etc. |
| Zavras et al., 2002 | DEA<br>Inputs: number of staff by caztegory, population covered Output: annual number of visits | 133 IKA clinics | 1998 – 1999 | Clinics which have diagnostic equipment are more efficient, as well as those that cover 10.000 to 50.000 people |