The Royal Veterinary and Agricultural University
Food and Resource Economic Institute

A UNIVERSITY OF LIFE SCIENCES

Unit of Economics Working Papers 2001/10

DEA-Based Incentive Regimes in Health-Care Provision.

Per J. Agrell and
Peter Bogetoft

DEA-BASED INCENTIVE REGIMES IN HEALTH-CARE PROVISION [1]

PER J. AGRELL [2]
Université Catholique de Louvain

PETER BOGETOFT [3]
Royal Agricultural University

ABSTRACT. A major challenge to legislators, insurance providers and municipalities will be how to manage the reimbursement of health-care on partially open markets under increasing fiscal pressure and an aging population. Although efficiency theoretically can be obtained by private solutions using fixed-payment schemes, the informational rents and production distortions may limit their implementation. The healthcare agency problem is characterized by (i) a complex multi-input multi-output technology, (ii) information uncertainty and asymmetry, and (iii) fuzzy social preferences. First, the technology, inherently nonlinear and with externalities between factors, yield parametric estimation difficult. However, the flexible production structure in Data Envelopment Analysis (DEA) offers a solution that allows for the gradual and successive refinement of potentially nonconvex technologies. Second, the information structure of healthcare suggests a context of considerable asymmetric information and considerable uncertainty about the underlying technology, but limited uncertainty or noise in the registration of the outcome. Again, we shall argue that the DEA dynamic yardsticks (Bogetoft, 1994, 1997, Agrell and Bogetoft, 2001) are suitable for such contexts. A third important characteristic of the health sector is the somewhat fuzzy social priorities and the numerous potential conflicts between the stakeholders in the health system. Social preferences are likely dynamic and contingent on the disclosed information. Similarly, there are several potential hidden action (moral hazard) and hidden information (adverse selection) conflicts between the different agents in the health system. The flexible and transparent response to preferential ambiguity is one of the strongest justifications for a DEA-approach. DEA yardstick regimes have been successfully implemented in other sectors (electricity distribution) and we present an operalization of the power-parameter ρ in an pseudo-competitive setting that both limits the informational rents and incites the truthful revelation of information. Recent work (Agrell and Bogetoft, 2002) on strategic implementation of DEA yardsticks is commented in the healthcare context, where social priorities change the tradeoff between the motivation and coordination functions of the yardstick. The paper is closed with policy recommendations and some areas of further work.

KEYWORDS

Data Envelopment Analysis, regulation, health care systems, efficiency.

---

[1] Fourth European Conference on Health Economics, Paris, June 7-10, 2002.

[2] UNIVERSITE CATHOLIQUE DE LOUVAIN Center of Operations Research and Econometrics (CORE), Voie du Roman Pays 34 *Louvain-la-Neuve B-1348, BELGIUM*. *E-mail address:* agrell@core.ucl.ac.be

[3] ROYAL AGRICULTURAL UNIVERSITY, Department of Economics and Natural Resources, Rolighedsvej 25, *DK-1958 Frederiksberg C, DENMARK*. *E-mail address:* pb@kvl.dk

1. INTRODUCTION

Since it was first proposed by Charnes, Cooper and Rhodes(1978,79), Data Envelopment Analyses (DEA) has become a tremendously popular relative performance evaluation tool. The collection of studies in the *Journal of Health Economics*, including early contributions like Grannemann, Brown and Pauly(1986) and Grosskopf and Valdmanis(1987), is only the tip of the iceberg. A recent bibliographic survey (www.deazone.com) identified more than 1000 papers from all sectors of society. Many of these have been published in high quality international journals within economics as well as management science and operations research.

Currently, DEA is also used as the basis for regulation regimes in different areas. In particular, it has been used in incentive regulation of private, semi-private and public utilities. In the regulation of electricity distribution, for example, countries like Norway, Holland, and Finland have introduced DEA based revenue and price cap systems, and DEA has – together with more traditional statistical methods – been used to determine reasonable cost norms in countries like Australia, England, New Zealand and Sweden.

Thus, it is natural to ask whether a DEA based reimbursement scheme would be useful in the design of incentives for the health sector as well.

Recognized Need For Incentives

The need is definitely there. It derives from, among other factors, the ever increasing medical costs, the – partly supply driven – surging demand for health care, and the considerable informational asymmetry between providers and sponsors of health services.

In addition, the need for regulation is *recognized*. Relative performance evaluations, yardsticks and performance based reimbursement schemes have long been employed in the health sector.

In USA, Medicare (covering elderly and disabled) as well as insurance companies before 1982 reimbursed private hospitals retrospectively based on charges for specific services (fee-for-service). This provided little cost reduction incentives and was - during the eighties - substituted by a system where hospitals' reimbursements are determined prospectively based on predetermined prices. Patients in the same Diagnosis Related Group (DRG) are expected to consume similar resources. DRGs can therefore be used to standardize the differences in the case mix of hospitals and thus allow comparisons of hospital efficiency. Also, DRGs can be used as the basis for a prospective payment system. Paying to a DRG unit is effectively a payment for a package of services (medieval treatment, nurse care, laboratory tests etc) and by using federal or national rates rather than local rates, the incentives for a hospital will be to provide the "package" at least possible cost.

In many national health care systems, funding has originally been input based using detailed control of the numbers and types of staff as well as different categories of other expenditures. To improve the incentives, more aggregate control has also been introduced. They may involve defined targets to the hospitals and letting funding be provided in a single category (global budget). More recently, several countries have substituted the input based control approach with an output or activity based one. In the activity based prospective payment system, hospitals recover at least part of their costs using standard rates for different categories of patients.

Often, some version of a DRG system is used to group the activities and hereby account for differences in the case mix.

In short, it is fair to say that there is a recognized need for appropriate performance evaluations and regulatory mechanisms in the health sector. The conception, evaluation and characterization of such systems from an empirical or theoretical economic viewpoint have also attracted a fair amount of interest from the academic community. Hence, we suppress the more general discussion on regulatory design to focus attention a one particular regulatory approach.

DEA BASED INCENTIVE REGULATION ?

The crucial next question is whether the existing systems actually can be improved by using DEA based evaluations and incentives.

Assuming that the hospitals are privately owned and efficiently managed, a DRG based PPS system using federal or national rates will provide incentives to deliver the services at the least possible costs. The question is however if it provides such incentives at the least possible cost? Albeit efficient, the hospitals may end up earning excessive information rents paid by the users or purchasers. Also, the structural development of the sector may be adversely affected by the system such that both health providers and health purchasers end off in a sub optimal situation.

To discuss potential improvements, we needs to be more specific about the context, i.e. about (i) the technology, (ii) the information uncertainty and asymmetry, and (iii) the motivations of the actors.

One contextual characteristic of the health sector is its quite *complex technology*. By this we mean that it is difficult to set up technical engineering models and norms of what can be accomplished given a set of resources. In particular, there is no reason to expect simple linear relationships between the activities. The possibilities to substitute one activity for another may depend on the activity level and there may be considerable positive and negative synergies between the different health products. Hence, the cost structure may not be additive. There may be lots of joints costs that one cannot easily allocate on the different products. As we shall argue, the flexibility of the DEA model may prove handy here.

A second characteristic is the *uncertainty and asymmetric information*. The uncertainty is related to the obvious stochastic influence on the provision of health care. On the other hand, the number of cases handled and the general quality of the data available in most developed countries suggest that uncertainty in the usual statistical sense is probably less important. It is then the asymmetry of the information rather than its quality that is key to the problem. The asymmetric information relates to the behavior of the agents and the precise nature of their working conditions. It is therefore important to investigate how to provide incentives in the context of considerable asymmetric information and considerable uncertainty about the underlying technology but limited uncertainty or noise in the registration of the outcome. Again, we shall argue that DEA is suitable for such contexts.

A third important characteristic of the health sector is the somewhat *fuzzy social priorities* and the numerous potential *conflicts* between the stakeholders in the health system. It is difficult and often politically sensitive to make explicit trade offs between different medical conditions. Moreover, the preferences are most likely dynamic and contingent on the possibilities, i.e. the ability to cure one or the other condition will influence the valuation attached to the outcome. Similarly, there are several potential hidden action (moral hazard) and hidden information (adverse selection) conflicts between the different agents in the health system. Professional medical ethics in combination with incomplete contracts with staff may also give raise to complex conflicts of interest between clients, medical staff, management and sponsors. The classical economic conflict is between the rentseeking provider, wanting a high compensation, and the cost minimizing clients/purchasers, wanting a low payment. We shall focus on this basic conflict below, but we note that several other conflicts prevail inside the groups of providers and purchasers. At the provider side, conflicts of interest between the management and the medical staff may significantly affect the costs of implementing cost saving procedures. Similarly, the potential conflicts between the purchasers of health, e.g. the state or insurance company, and the final level users of the health service, the insured citizen, may limit the ability to control the providers. The user will typically demand the best possible treatment irrespectively of cost, while the sponsor will be concerned about the associated costs. Medical staff may obtain rents from this conflict by overproviding healthcare to low-effort clients, while discouraging (balking) clients with above-norm expected cost. It is important to investigate how to handle at least some of these conflicts most efficiently within a context of asymmetric information and a complex technology. To cope with these problems, DEA allows a variety of underlying social value functions. The flexible and transparent response to preferential ambiguity is one of the strongest justifications for a DEA-approach.

Outline

Below we examine the properties and usefulness of DEA based evaluations and incentives in a context with a complex technology, acceptable data quality but non-negible information asymmetry, unclear social priorities and conflicts of interest among the involved agents.

We first give a non-technical introduction to DEA in Section 2. In Section 3, we briefly discuss why DEA has become popular among regulators and some caveats for its regulatory utilization. To explain more convincingly the strengths and weaknesses of DEA in regulation, we offer a survey of the main theoretical results in Section 3. Attention is paid to conditions for DEA giving optimal incentive and mechanism design with DEA under various settings. Some final remarks and suggestions for future work close the paper in Section 5.

2. DATA ENVELOPMENT ANALYSIS

In this Section, we provide a non-technical introduction to the main ideas and constructs in Data Envelopment Analysis (DEA). For a text-book introduction to DEA, see Charnes, Cooper, Lewin and Seiford (1994), Coelli, Rao and Battese (1998) or Cooper Seiford and Tone (2000).

RATIONAL IDEAL EVALUATIONS

Consider the problem of evaluating a given production unit or production plan. In the DEA literature, the evaluated units, hospitals, clinics or even individual physicians, are usually called Decision Making Units (DMUs). One can think of the DMUs as actual organizational units, as production plans or more generally as multiple dimensional performance descriptions at a given time. Note that there is no technical difference between times series and panel data in the DEA framework, past performance of a unit may also be used to evaluate current activity.

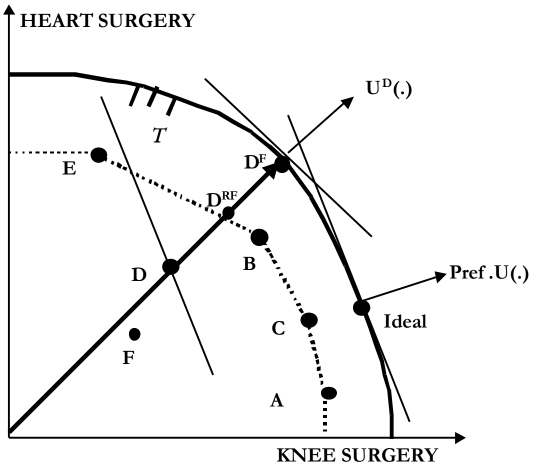
A DMU is characterized by transforming resources into products and services. The transformation is affected by non-controllable variables as well as non-observable skills and efforts in the organization. The inputs in a hospital study could for example be the number of bed-days, physicians, nurses, non-sanitary personnel and the available equipment. The outputs might include the number of patient days in different sections of the hospital, the number of surgical interventions, the number and length of ambulatory visits. In a more detailed study, the output could include also the number of DRG activities delivered. The non-controllable variables will depend on the time horizon etc but could include large parts of the fixed costs, medical status of admitted patients, and the demographic, medical and socio-economic status of the region served by the hospital. We note that establishing a relevant system description of hospital production is not an easy task. Indeed, Newhouse (1994) has argued that the difficulties of measuring output appropriately and the inability to handle the quality dimension satisfactorily make efficiency studies doubtful at best. We acknowledge these concerns but suggest that they may not be more severe than in some other service industries. Also, using a detailed DRG based description combined with a minimum of *a priori* information about the relative prices of the outputs like in Olesen and Petersen (2001) or Chillingerian (1995) can in fact lead to useful insights.

From a standard micro-economic stance, the performance of a given DMU is defined by its ability to choose the best means to pursue its aims. The rational ideal performance evaluation, viz. the absolute effectiveness evaluation, may be summarized by comparing the actually attained goal level to the maximally attainable goal level. Figure 1 below illustrates this idea at a given level of input and exogenous influence.

In real evaluations, it is not entirely easy to apply the micro-economic cookbook recipe. In the typical evaluation we lack clear priorities U as well as clear information about the production possibilities T. DEA provides a way to overcome these fundamental practical problems. The approach is illustrated in Figure 2 below.

FROM EFFECTIVENESS TO EFFICIENCY

Consider first the *lack of clear priorities* concerning how the resource expended and the products created should be evaluated and traded-off relative to each other. In Figures 1 and 2 this corresponds to the U function with the indicated linear indifference curves. The lack of priority information includes the problem of resource allocation for the treatment of patients with different diagnoses and the problem of maximum allocation for a given admitted patient. DEA overcomes this problem by moving from an evaluation of *effectiveness*, i.e. goal attainment, to an evaluation of *efficiency*. Efficiency is here broadly defined as the production of the maxmium output using the minimum input. The efficient plans in Figure 1 are all the plans on the north-east frontier.



Effectiveness

$$= \frac{\textbf{Actual Performance}}{\textbf{Ideal Performance}}$$

$$= \frac{\mathbf{U\ (D)}}{\mathbf{U\ (Ideal)}}$$

Figure 1 The rational ideal evaluation model given preferences $U(.)$.

To quantify the extent of inefficiency, the DEA literature uses different measures of the distance between a given DMU, say DMU D, and the frontier of efficient plans. Most studies uses the so-called Farrell (1957) measures that take into account the multiple dimensional character of the inputs and outputs by looking for proportional expansions and contractions. The *Farrell output and input based measures* are verbally defined as:

F = largest proportional expansion of all outputs that is possible without using additional input
E = largest proportional contraction of all inputs that is possible without reducing any output.

The interpretation is straightforward. E.g., $F = 1.2$ means that all output could be increased by 20% while $E = 0.6$ means that all inputs could have been reduce with 40%. In Figure 1, the Farrell base output efficiency is approximately 200% meaning that all outputs could have increased by 100% without introducing additional inputs, namely by moving from D to the Farrell projection plan D^F.

The rationale behind the evaluations may be interpreted in the following manner. The lack of priori information about the social priorities is overcome by *choosing the priorities that puts the evaluated DMU in best possible light*. For DMU D in Figure 1 this would be the priorities corresponding to the stipulated preference structure U^D. Hence, in DEA, each and every DMU, say hospitals, are evaluated according to prices or priorities that makes its effectiveness look as high as possible. The lack of knowledge about priorities is handled by allowing for all possible priorities corresponding to all possible slopes of the indifference curves in Figure 1. Optional access to ordinal or cardinal preference information may be used to refine the evaluations within the model. E.g., assume that consensus can be reached in Figure 1 upon the lesser value of a knee operation compared to a heart surgery. Consequently, the slope of the admissible evaluations, or lines one can stipulate must be smaller (less steep) than –1.

FROM ABSOLUTE EFFICIENCY TO RELATIVE EFFICIENCY

Consider next the other fundamental problem in practice, namely the *lack of sufficient a priori information about the underlying, potentially complex technology*. In Figure 1 this corresponds to the set *T* reflecting the technology in a broad sense, i.e. the socio-technical possibilities to transform combinations of inputs into combinations of outputs. DEA overcomes this problem by estimating the technology *T** from observed historical or cross-sectional actual plans. Performance is then not evaluated relative to an absolute norm but rather relative to the performance of other DMUs.

The idea of substituting an underlying but unknown production possibility set with an estimated one is of course not unique to the DEA approach. It is done also in performance evaluations using traditional statistical methods, accounting approaches etc. What is particular about the DEA approach is the way the approximation of the technology is constructed and the resulting properties of the evaluations.

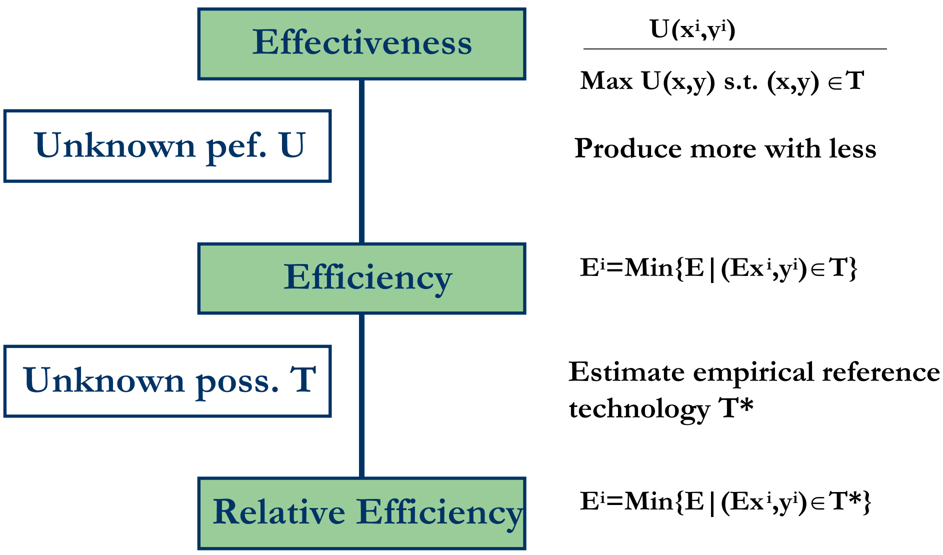


Figure 2 The DEA approach to effectiveness, efficiency and relative efficiency.

The technology is estimated in accordance with the *minimal extrapolation principle*. By this we mean that DEA constructs the smallest possible set of production plans that constraint the observed ones and satisfy a set of (weak) regularity conditions. By constructing the smallest set containing the actual observations, the method extrapolates the least.

The minimal extrapolation principle is illustrated in Figure 1 by the set below the dotted line. Effectiveness, if we know the priorities *U*, or efficiency *F*, in absence of *U*, can now be evaluated relative to *T** rather than *T*. Since we evaluate compared to an empirical norm set by the other DMUs and do not compare to an absolute norm, we say that we evaluate relative effectiveness or efficiency. In the case of DMU D, the relative efficiency is now approximately 1.3 suggesting only a 30% improvement potential since we now compare to the point D^{RF}.

Different DEA models are distinguished by the set of regularities imposed on the set T^* in addition to the requirement that T^* contains the data. The three most often used assumptions are free disposability, convexity and s-returns to scale:

A1: *Free disposability*, i.e. the ability to produce less output using less input

A2: *Convexity*, i.e. the ability to make weighted averages of production plans, and

A3: *s-Return to scale*, i.e. the ability to scale up and down (crs), only down (drs) or not at all (vrs).

The four most commonly used DEA models are the original constant returns to scale (crs) DEA model proposed by Charnes, Cooper and Rhodes (1978, 1979) invoking A1, A2 and A3(crs), the decreasing returns to scale (drs) and (local) variable returns to scale (vrs) models developed by Banker (1984) and Banker, Charnes and Cooper (1984) under A1, A2 and A3(drs) and A1, A2 and A3(vrs), respectively, and the free disposability hull (fdh) model proposed by Deprins, Simar and Tulkens (1984) by invoking only A1. The resulting models in a single input single output case are illustrated in Figure 3 below.

In the multiple output case the vrs, drs and crs models could look like the T^* technology in Figure 1 while the fdh technology would correspond to a step-function between the points A, B, C and E.

To understand the flexibility of the DEA approach it is worthwhile to note that the fdh model merely assumes that production and cost functions are increasing. The vrs model modestly assumes that production functions are increasing and concave and that cost functions are increasing and convex. The remaining models add the scaling possibilities but even the crs model is much more flexible than a linear model (given multiple inputs or outputs)

To summarize, DEA convincingly addresses two fundamental problems in practical evaluations. The lack of clear preference or priority information is handled by moving from effectiveness to efficiency and the lack of *a priori* technological information is handled by making weak *a priori* assumptions, by estimating using the minimal extrapolation principle, and by evaluating efficiency relative to be best practice.

3. WHY IS DEA SO POPULAR?

We have argued that relative performance evaluations in regulation are advantageous in several sectors, including the health sector. We have also described how DEA solves some of the fundamental information asymmetry problems in real evaluations. We shall now look a little closer at some of the implied pros and cons of DEA.

CAUTIOUS IMPROVEMENT ESTIMATES?

The way we estimate the production possibilities in DEA has several implications. The use of the minimal set containing the observations suggests that DEA provides an inner approximation of the underlying production possibility set. The efficiency estimates are therefore conservative in the sense that the potential output expansions or input savings are underestimated. We have already seen this for DMU *D* in Figure 1 where the expansion possibility was estimated to 30% with *T** and to 100% with *T*. It is said that DEA by means of the minimal approximation and the construction of an inner approximation generates *conservative* or *cautious estimates*.

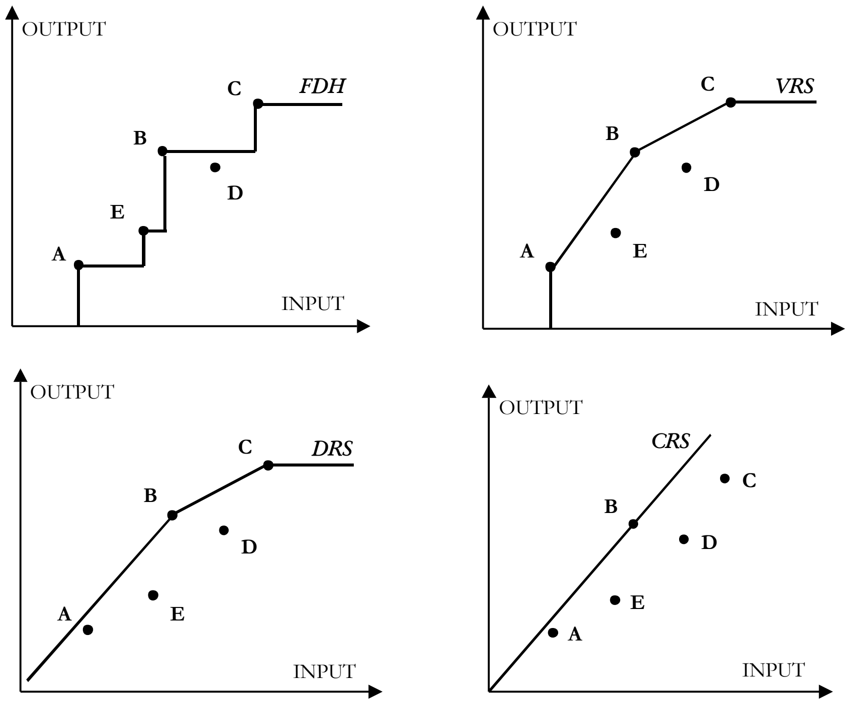


Figure 3. The free disposable hull (fdh), variable returns -to-scale (vrs), decreasing returns -to-scale (drs) and constant returns-to-scale (crs) technologies for DMU A, B, C, D, E .

Best Practice

The use of the minimal extrapolation principle and hereby the construction of the largest inner or smallest outer approximation also imply that the technology identifies the so -called *best practice*. The DEA approach captures what the units having performed the best - using the least inputs to produce the largest amounts of outputs – have been able to achieve. This is attractive in many settings since the methods and procedures of the best units are more likely targets to other units. Thus, for example if *D* in Figure 1 is to learn, it would probably find little to learn from looking at *F*. It would be more interesting to inquire what units like *B* and perhaps *E* have done differently. A further consequence of using the DEA approach is that *real peers* are identified. In Figure 1, the unit D has two peers, *B* and *E* since D^{RF} is located on the line between these two units. *B* is the primary peer since D^{FR} is located close to *B*. Of course, the construction of best practice norms should also influence the way we design a possible incentive scheme.

Flexible Production Structures

The third and in many cases most important implication of the DEA estimation approach is its ability to work with *weak a priori assumptions and flexible production opportunity models*. DEA models generally allows for the underlying best practice production structure to take many different forms. If we estimate a cost function using DEA for example, we may assume that it is simply any increasing function – or any increasing convex function. We do not need to assume that the substitution possibilities between the outputs are fixed for example. No parametric statistical model or accounting cost function allow for a similar flexibility in the technology. There are many versions of the DEA approach corresponding to the introduction of different combinations of *a priori* assumptions. In all cases, however, the imposed *a priori* structure is mild compared to competing approaches.

Noise or Efficiency

The single most problematic feature of DEA is the risk of mistaking noise for efficiency or inefficiency. And similarly, to mistake best practice for most lucky practice.

If a DMU by chance faces particularly favourable circumstances, not accounted for in the model, or if the registration of the outputs by luck (or intent) is biased upwards and the inputs downwards, the unit will appear to have performed particularly well and have little or no inefficiency. Similarly, there is a risk of non-favorable circumstances or registrations leading to groundless accusations of inefficiency in a DEA analysis. The first case is particularly problematic since it might influence the evaluation of others that may now face tougher standards by being compared to a unit with a wind fall gain.

These observations have lead some theorists as well as practitioners to question DEA and to forward the use of statistical methods, including so-called stochastically frontier analyses SFA. SFA is like a traditional statistical model except that the noise is composed of two terms, a one sided efficiency terms and a two sided traditional noise term.

The consultants' answers to these problems – unless they have worked for firms challenging the DEA regulation – have usually been optimistic. They might claim that "DEA puts everyone in their best light" or even that "DEA bends itself backwards to make everyone look as good as possible." Some theoretical basis for these answers exists as we have seen, namely in the weak a priory regulatory imposed, the minimal extrapolation principle applied and in the interpretation of efficiency as allowing each DMU to choose its own priorities.

On the other hand, this is of course only part of the story. The inclusion of all relevant factors and the exclusion of any noise in the registration of performance are seldom guaranteed in any formal model, DEA not being an exception.

As any other operations research technique, DEA is just a tool that can be used with success if put in the right hands and used sensibly. To be well executed, a DEA analysis must involve careful data collection, serious sensitivity analysis (using Monte Carlo techniques, peeling techniques, testing alternative technologies etc), perhaps stochastic programming and if possible hypothesis testing. Unfortunately, the statistical theory underlying DEA is not fully developed to eliminate the impact of noisy registrations. There are by now numerous contributions (involving sensitivity analysis, stochastic programming, resembling, bootstrapping, and asymptotic test theory, cf. for example Simar and Wilson (2000) for a recent survey) but the state of the art is still lacking compared to what can be done in a parametric models.

The theorists' answer would therefore be more cautious. The appropriateness of DEA depends to a large extent on how it is applied and for which purposes. Careful decision analysis can determine whether the underlying assumptions regarding the controllability of performance are undermining the viability of the incentive system. Despite the likely entangling of efficiency and noise in any analysis, its consequences may be moderated by choice of variables, aggregation unit and sensitivity analysis.

In a regulatory context, it will usually be the regulators discretion how much inefficiency to eliminate in the coming periods, cf. also below. In such case, by acting generous, he may effectively create a safeguard against noise.

Further, given the flexibility in the production structure, individual noise or outlier problem may only have local impacts.

Finally, the impact of noisy registrations and mis-specified models should be also be viewed against the uncertainty associated with the underlying (mean) production structure. With sparse *a priori* information about a potentially complex technology, the DEA approach offers clear advantages over parametric statistical methods and mechanic accounting models.

Figure 4 below illustrates the differences with respect to best practice identification, flexibility of the model, and the handling of noise in a technical ("true") benchmarking approach, a DEA approach and an econometric approach.

Intuitively, it seems natural to conclude that if one faces a simple technology and very noisy data, the use of parametric statistical models are preferable from an inference perspective. If on the other hand, we have access to relatively high quality data but a complex technology with considerable uncertainty about the structure of the input-output correspondences (the rates of substitution etc), DEA is preferable.

More formal models of the comparative advantages of DEA in regulation in presence of considerable structural uncertainty will be surveyed below.

EASY TO USE AND DEFEND?

A couple of more pragmatic observations are relevant here as well.

DEA is easy to use given the existing software, the limited *a priori* assumptions needed and – somewhat counterintuitive – the lack of good standard indicators of model mis-specification!

Moreover, DEA may also be considered easy to justify. Again, this rests on the mild regularity assumptions, the ability to handle multiple inputs and outputs and its clear connection to production theory. Albeit informative in a collaborative setting, the identification of explicit peers actually opens the model for criticism. If for example a regulated firm appeals the regulator's decision in court, the existence of explicit peers limits the burden of proof on the firm. Rather than having to invalidate the entire sample, as in an econometric model, the DEA -regulated firm concentrates on a few selec ted comparators. The statistical "black -box" approach gives the regulator a strategic advantage. Hence, although transparency and managerial informativ e-ness are virtues of any formal method, it does prompt for strategic behavior in regulation. Also, the ri sk possibly entangling slack and noise, inefficiency and noise, may make the DEA approach harder to d efend.

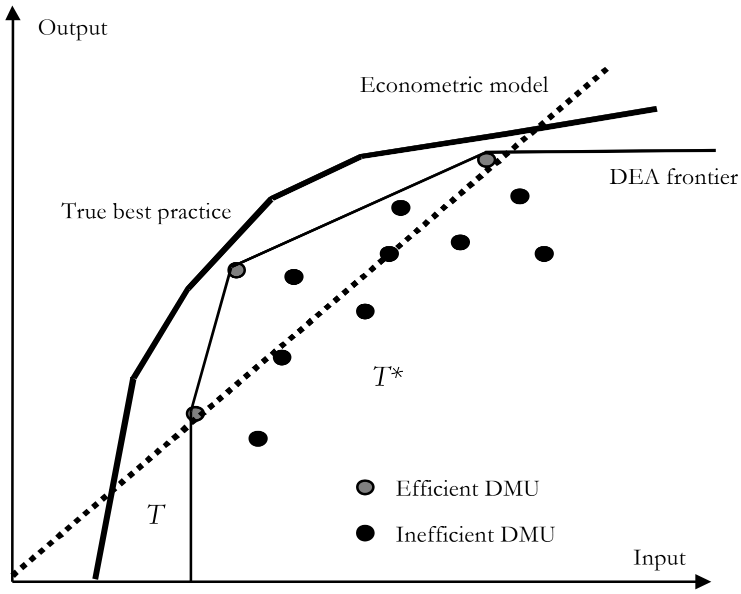


Figure 4. Alternative Benchmarking Approaches

Preference Information Again

A second drawback of the DEA approach, as of efficiency studies in general is the *lack of focus on organizational objectives*. The impressing progress that can be made in the evaluation without much preference information should not lead one to forget that importance of "doing the right things" and n ot just "doing things right". It may be preferable to move slowly in the right direction than to run fast in the wrong direction The importance and potential gains from gi v-ing more attention to preference modeling and less to the evaluation with possibly n aïve priorities have been emphasized in the multiple criteria literature, cf. Bogetoft and Pr uzan (1991) and have gradually been included into the DEA literature as well, cf. e.g. Ali, Cook and Seiford (1991), Golany (1988a,b), Halme, Joro, Korhonen and Walle nius (1999), and Joro, Korhonen and Wa llenius (1998).

Pros and Cons of DEA

To summarize our discussion, we have identified a series of strengths and weaknesses of DEA.

The *strengths* include

- Requires no or little preference, price or priority information
- Requires no or little technological information

- Makes weak *a priori* assumptions
- Handles multiple inputs and multiple outputs
- Provides reel peers
- Identifies best practice
- Cautious or conservative evaluations (minimal extrapolation)
- Supports learning and - as we shall argue in the next Section - planning and motivation

The *weaknesses* include

- Relatively weak theory of significance testing (sensitivity, resampling, bootstrapping, asymptotic theory)
- Lack of focus of goals

4. LESSONS FROM THEORY

We now give a survey of the main results about incentive provision using DEA analysis.

Formal analysis of the likely usefulness of DEA in incentive provision was initiated by Prof. Peter Bogetoft in 1990. The present survey gives a non-technical introduction to a series of papers – often somewhat technical –- written primarily by Bogetoft and his associates. The main results concern the use of so-called super-efficiency, the design of static incentives and the design of dynamic incentives in context with considerable technological uncertainty and asymmetric information about a regulated agent's actions (moral hazard) and working conditions (adverse selection). We commence with a few words about the research approach.

THE BASIC PROBLEM AND CONTEXT

The basic problem addressed in this literature is the following: Given a cross section, a time series or panel information on the multiple inputs and outputs used by n DMUs

$$(\text{inputs}^i, \text{outputs}^i) \text{ for DMUs } i=1,\ldots,n$$

what should we ask the DMUs to do in the future, how should we motivate and compensate them for their private costs ?

The answer to these questions depends intimately on the organizational *context* and in particular on the technological, informational and preferential assumptions of the parties, i.e. the regulator (principal) and DMUs (agents).

In general, we consider the case where the principal (regulator) faces considerable *uncertainty about the technology*. In a single input multiple output cost setting he may for example know that the cost function is increasing and convex, but otherwise have no *a priori* information about the cost structure. In the pure moral hazard models, we also assume that the agents face a similar uncertainty.

The general case also empowers agents to take *private actions*, which the principal cannot observe. The action could for example be to reduce costs or increase the quality of the work done. This leads to a usual *moral hazard problem* since the principal and the agents may conflict as to which actions the agents should take. The traditional setting depicts the agents as work averse, tempted to rely on their good luck and to explain possibly bad performances with unfavorable circumstances. In general, however, it is simply one way to model the underlying conflicts giving rise to a motivation problem. However , the zealous physician who pledges his Hippocratic oath by providing top -class treatment to any and all patients illustrates another moral hazard. The conflict might also be that the medical staff have diverging preferences that induce them to work (too) hard, to treat (groups of) patients below cost and to accomodate the patients' desire to undergo multiple treatments.

In some models, we also consider the possibility that the agents have *superior information* about the working conditions before contracting with the principal. A hospital manager may for example have good information about the primary cost drivers at his hospital while the Ministry of Health may have little information about what causes the total bill to increase. This leads to the classical *adverse selection problem* where an agent will try to extract information rents by claiming to be under less favorable conditions.

As regards the specific *preference* of the parties, we generally assume that the principal is risk neutral and that the agent is either risk averse or risk neutral. The principal's aim is to minimize the costs of inducing the agent to take the desired (hidden) actions in the relevant (hidden) circumstances. An agent's aim is usually to maximize the utility from payment minus the disutility form private effort. In the combined moral hazard and adverse selection models, we usually make a simplifying assumption about the structure of the agent's trade offs between effort and payment. We assume that his aim is to maximize a weighted sum of profit and slack:

$$\begin{aligned}\text{DMU's Utility} &= \text{Utility from Payment} - \text{Cost of Effort}\\ &= \text{Profit} + \rho \cdot \text{Slack}\end{aligned}$$

where slack is a measure of the extent to which input utilization exceeds the minimal possible and where $1 \geq \rho \geq 0$ is the relative value of slack

The general set-up and timeline is illustrated in Figure 5 below.

SUPER EFFICIENCY

The first simple lesson is that it is important to modify the simple Farrell measure slightly for use in incentive regulation.

The Farrell measures E and F give all units on the relative efficient frontier a score of 1. This severely limits the ability to give high-powered incentives based on Farrell measures. As demonstrated in particular in Bogetoft (1990,1994a and 1995) in a simple moral hazard context, the Farrell measures can give incentives to match others, but not to surpass the norm and push out the frontier. Combining this with the multiple dimensional character of the typical DEA model and hereby with the ability to be special in different ways, the Nash Equilibria (NE) that can implemented using the Farrell measure will often involve minimal effort and maximal slack.

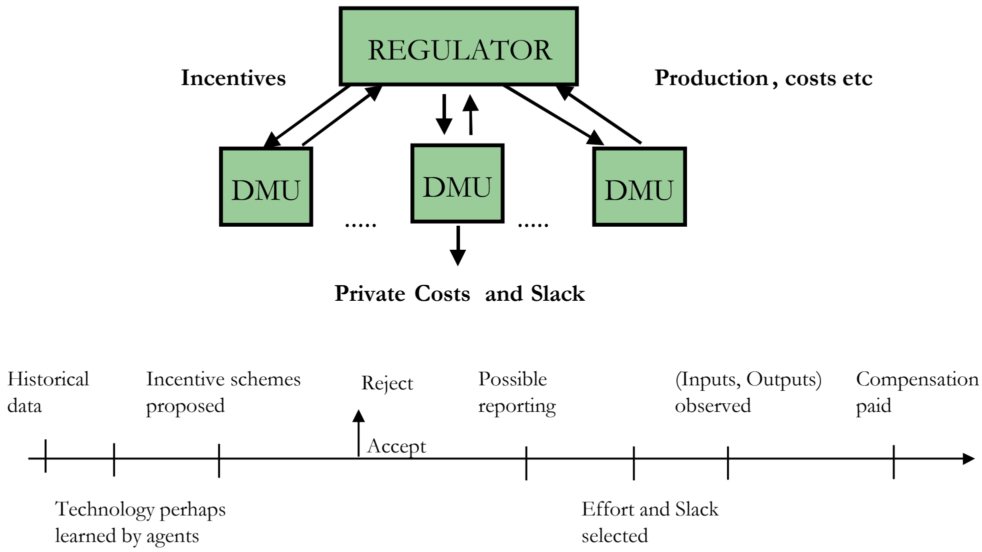


Figure 5. Agency Structure with Timeline of Events.

Figure 6 below illustrates this. Assume that the cost to the agents are proportional to the length of the production vectors and that payment is decreasing in the F score such that maximal payment is received when a DMU is efficient with a score of F =1. Now if DMU 1, planned to produce at A, moves from A to C, he would get the same payment but supply less effort. Doing A is therefore not a best response. Next, DMU 2 could move from the planned B to a more easy life in D, again reducing his private costs of effort without affecting his pa y-ment. This procedure can continue until they both supply the min imal effort and receive the maximal payment.

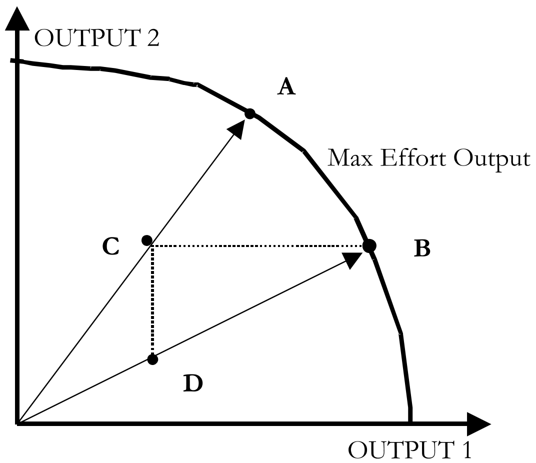


Figure 6. Incentive Provision with Efficiency and Super Efficiency

This somewhat discouraging outcome can easily be remedied. By eliminating the evaluated unit from the set of comparators, one can measure th e so-called super-efficiency of frontier units.

This was suggested by Andersen and Petersen (1993) as a means to differentiate among frontier units. Presently, it is motivated by a desire to provide incentives. The super-efficiency measure on the inputs side, $E^{SUPER,}$ can now be larger than one. A score of for example $E^{SUPER} = 1.2$ suggests that the DMU could have increased all of its inputs with 20% and still have been efficient. Similarly, a super-efficiency score on the output side of $F^{SUPER} = 0.9$ suggests that the DMU could have reduced all of its outputs with 10% and still have been on the frontier. In Figure 6, the output based super efficiency for DMU 1 in A is approximately 0.6 but if the payment is sufficiently decreasing in F^{SUPER}, it would not pay to reduce the effort. Indeed, its does not pay to reduce the effort if the marginal reduction in payment exceeds the marginal decease in the cost of effort.

More generally, using super-efficiency, it turns out that one can support the implementation of most plans – even in so-called un-dominated Nash-equilibria.

All the papers discussed here use super-efficiency as the basis for designing contracts.

It is worthwhile to note that if we extend the set-up of the papers considered here by introducing more uncertainty, including individual noise terms and a lack of common knowledge among the agents being motivated, even traditional efficiency will be able to provide sufficient incentives to implement many plans, among those the optimal (cost minimal). The logic is that if an DMU is uncertain about the motives of the other DMU or there is considerable noise in the outcome, an agent cannot rely on the fine strategic behavior in Figure 6 because he cannot with certainty foresee the results.

Static Incentives with Individual Noise

In Bogetoft (1994a, 94b) we considered pure moral hazard contexts with

- Considerable technological uncertainty,
- Risk averse DMUs
- Individual uncertainty (noise) in the DMUs performances.

The technological uncertainty was represented by a large class of *a priori* possible technologies, e.g. the set of productions functions that are increasing and concave or the set of functions that are simply increasing. We addressed the question of when the associated DEA based frontiers, e.g. the vrs frontier or the fdh frontier from Figure 3, contained all the relevant information for contracting:

Optimal Compensation to DMU ^{i} = B^i (input^i, output^i, DEA frontier based on other DMUs)

This is the case where optimal relative performance evaluations can be made by comparing the performance of a given DMU against the DEA best practice frontier estimated from the performance of the other DMUs.

It turned out that

- DEA frontiers support optimal contracts when the distributions of the individual noise terms are exponential or truncated,

- DEA frontiers based on large samples supports optimal contracts when noise is monotonic in the sense that small noise terms are more likely than large.

Hence, even when we have individual noise elements and not just the structural uncertainty which intuitively seems to favor DEA, will DEA based contacts be optimal for special distributional assumptions – and for rather general assumptions if the sample is sufficiently large.

STATIC INCENTIVES WITH ADVERSE SELECTION

In Bogetoft (1997, 2000) we considered combined adverse selection and moral hazard contexts with

- Considerable asymmetric information about the technology
- Risk neutral DMUs,
- DMUs seeking to maximizes {Profit + ρ•Slack}.

The DMUs are supposed to have superior technological information. In the extreme case, they know the underlying true cost function with certainty while the regulator only knows the general nature of the cost function. Thus for example, the regulator may know that there are fixed unit costs of the different outputs but not the exact unit cost, being the DMU's private information. Alternative assumptions may be made about the information available to the regulator. We may assume for example that he know simply that the cost function is increasing and convex.

The optimal solution in this case depends on whether the actual costs, i.e. the minimal possible cost plus the slack introduced by DMU, can or cannot be verified and hence contracted upon.

If the actual costs cannot be contracted upon, the optimal solution is to use the following revenue cap with non-verifiable cost information

Optimal Reimbursement B^i to DMU ^{i}

$$= k + C^{DEA}(y^i)$$

= Lump Sum Payment + DEA-Estimated Cost Norm Based on the Other DMUs

The size of the lump sum payment depends on the DMU alternatives, including his profit potentials in other markets or the surplus from contracting with other regulators, say private insurance companies.

If instead we assume that the actual costs of the DMU can be contracted upon, the optimal reimbursement scheme becomes

Optimal Reimbursement B^i to DMU ^{i}

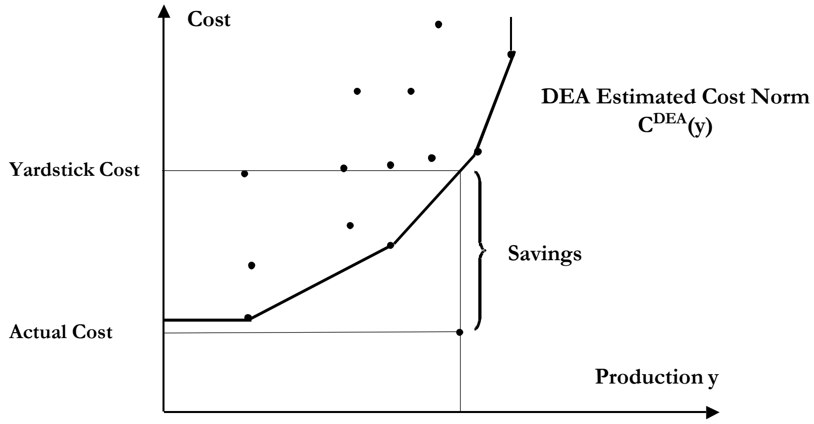
$$= k + c^i + \rho\bullet(C^{DEA}(y^i) - c^i)$$

= Lump Sum Payment + Actual Costs + Fraction ρ of DEA-Estimated Cost Savings

The structure of this payment schemes can be interpreted as a *DEA based yardstick model*: Using the performance of the other DMUs, the regulator creates a cost yardstick and the regulated DMU is allowed to keep a fraction ρ of his saving compared to the yardstick costs as his effective compensation. Figure 7 illustrates this reimbursement scheme.

These results provides an incentive rationale for using DEA based revenue cap systems in contexts where the regulator face considerable uncertainty about the underlying cost structure.

Several extensions and generalizations of these results are provided in Bogetoft (1997,2000). In particular, it is shown how the structure of the schemes are essentially unaffected by introducing decentralized decision making (where the DMUs, e.g. the hospitals, decide on the output mix) as well as participatory budgeting arrangements. Also, the impact of introducing genuine social benefit functions, alternative costs of slack reduction models, rationing etc is investigated.



Payment = Lump Sum + Actual Cost + ρ•Savings

Figure 7. The DEA Yardstick Model in the Cost – Production Space.

DYNAMIC INCENTIVES WITH RATCHET AND LIMITED CATCH UP

In Agrell, Bogetoft and Tind (2000a,b) we extended the above adverse selection and moral hazard context by introducing a time dimension. The dynamic perspective gives rise to new issues, including

- The possibility to accumulate and use new information
- The need to avoid the ratchet effect, i.e. deliberate subperformance in early periods to avoid facing too tough standards in the future
- The possibility of technical progress (or regress)

Nevertheless, the structure of the optimal dynamic scheme is similar to the ones developed above. Thus the optimal revenue cap for a DMU is found by a DEA based yardstick norm. Assuming verifiable actual costs, the optimal scheme taking into account the generation of new information, the ratchet effect and the possible technical progress becomes:

Optimal Reimbursement B_t^i to DMU^i in period t

$$= k + c_t^i + \rho\bullet(C^{DEA}{}_{1\text{-}t}\,(y_t^i) - c_t^i)$$

= Lump Sum Payment + Actual Costs in Period t

+ Fraction ρ of DEA-Estimated Cost Savings in Period t using all the information from the other DMUs generated in periods 1 through t.

In Agrell, Bogetoft and Tind(2000a,b) we have described how to modify the schemes above to take into account also the

- Possibly limited catch-up capacity, i.e. the fact that it may take time for a DMU to learn the best practice
- Possible cost of innovation (frontier movements) and loss from dissemination (sharing) of information

STRUCTURAL DEVELOPMENTS

Before closing, it is important to emphasize the complexity involved in the design of a regulation mechanism. We have focused above on how DEA supports to setting of appropriate cost norms in a reimbursement scheme. However, this is but one of many relevant concerns. The stress is here on the reduction of short run cost via relevant improvements and adaption to best practice. In a long run perspective, the restructuring of the industry to optimal scale and scope economies may be far more important than such short run concerns. An important problem is therefore how alternative short run schemes affect the DMUs incentives to make such adjustment, e.g. through mergers. Another issue of interest is how the regulated DMUs through restructuring of the business and accounting procedures may strategically try to adjust their structure to the regulatory scheme. Some results on these issues are available in Agrell and Bogetoft (2001) and Bogetoft and Wang(1999).

5. CONCLUSIONS

In this paper, we have discussed the strengths and weaknesses of DEA in the design of activity based reimbursement schemes. In particular we have argued that DEA yardstick schemes may be useful in contexts with a complex technology, reasonable precise data, fuzzy social priorities and numerous potential conflicts between the regulators and the regulated, as well as considerable amounts of asymmetric information. The potential presence of ambigous or even conflicting preference functions, that characterizes health care provision, invalidates some of the underlying microeconomic assumptions in contracting. A system based on the flexible production structure of DEA may facilitate reimbursement, as well as organizational learning, since it allows for alternative preference functions. In the light of recent developments to incorporate dynamic frontier shifts, ratchet effects and innovation, in addition to the classical single-period adverse selection and moral hazard problems, DEA has become a mature component in modern incentive menus. Not a panacea, but certainly benefical if correctly administered, as the doctor would say.

LITERATURE

1. Agrell, P. and P. Bogetoft, I ncentive Regulation of Swedish Electricity Distributors, Working Papers, 2001
2. Agrell, P., P. Bogetoft and J.Tind, Dynamic DEA Yardsticks in Scandinavian Electricity Distribution, under revision for *Journal of Productivity Analysis*, 2002a.
3. Agrell, P., P. Bo getoft and J.Tind, Incentive Plans for Productive Efficiency, Innovation and Learning, *Int.Journal of Production Economics*, 78, 1 -11, 2002b.
4. Ali, A.I., W.D.Cook and L.M. Seiford, Strict vs. Weak Ordinal Relations for Multipliers in Data Envelopment Analysis, *Management Science*, 37, pp. 733-738, 1991
5. Andersen, P. and N.C. Petersen, A Procedure for Ranking Efficient Units in Data Envelop ment Analysis, Managements Science, 39, pp.1261 -1264, 1993.
6. Bogetoft, P., Strategic Responses to DEA Control, Working Pape r, 1990.
7. Bogetoft, P., *Non-Cooperative Planning Theory*, Springer -Verlag, 1994a.
8. Bogetoft, P , Incentive Efficient Production Frontiers: An Agency Perspective on DEA, *Management Science*, 40, pp.959 -968, 1994b.
9. Bogetoft, P, Incentives and Productivity Measu rements, *International Journal of Production Economics*, 39, pp. 67 - 81, 1995.
10. Bogetoft, P, DEA -Based Yardstick Competition: The Optimality of Best Practice Regulation, *Annals of Opera-tions Research*, 73, pp. 277 -298, 1997.
11. Bogetoft, P., DEA and Activity Plan ning under Asymmetric Information, *Journal of Productivity Analysis*, 13, pp. 7 -48, 2000.
12. Bogetoft, P. and P. Pruzan, *Planning with Multiple Criteria*, North -Holland, 1991.
13. Bogetoft, P. and D. Wang, Estimating the Potential Gains from Mergers, Working Pape r, 1999.
14. Charnes, A., W. Cooper, A.Y.Lewin and L.M.Seiford, *Data Envelopment Analysis: Theory, Methodology and Appl i-cation*, Kluwer Academic Publishers, 1994.
15. Charnes, A., W.W. Cooper and E. Rhodes, Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, 2, pp. 429 -444, 1978.
16. Charnes, A., W.W. Cooper and E. Rhodes, Short Communication: Measuring the Efficiency of Decision Making Units, *European Journal of Opera tional Research*, 3, pp. 339, 1979.

1. Chilingerian, J. A. (1995) Evaluating Physician Efficiency in Hospitals: A Multivariate Analysis of Best Pra c-tice. *European Journal of Operational Research,* **80**(3), 548 -574.

17. Coelli, T., D.S.Prasada Rao, and G. Battese, *An Introduction to Efficiency and Productivity Analysis*, Kluwer Aca-demic Publishers, 1998.
18. Cooper, W.W., L.M.Seiford, and K.Tone, *Data Envelopment Analysis*, Kluwer Academic Pu blishers, 2000.
19. Deprins, D., D. Simar, and H. Tulkens, Measuring Labor Efficiency in Post Offices, pp. 243 -267 in M. Marchand, P. Pestieau, and H. Tulkens, "*The Performance of Public Enterprises: Concepts and Measurements*", North Ho lland, 1984.
20. Farrell, M.J. , The measurement of productive efficiency, *Journal of the Royal Statistical Society*, Series A, III, pp. 253-290, 1957.
21. Golany, B., A Note on Including Ordinal Relations Among Multipliers in Data Envelopment Analysis, *Manage-ment Science*, 34, pp. 1029 -1033, 1988a.
22. Golany, B., An Interactive MOLP Procedure for Extension of DEA to Effectiveness Analysis, *Journal of Opera-tional Research Society*, 39, pp. 725 -734, 1988b.
23. Granneman, T., R.S. Brown and M.V.Pauly, Estimating Hospital Costs: A Multiple -Output Analysis, *Journal of Health Economics*, 5, pp. 107 -128, 1986.
24. Grosskopf, S. and V. Valdmanis, Measuring Hospital Performance: A Non -Parametric Appro ach, *Journal of Health Economics*, 6, pp. 89 -107, 1987.
25. Halme, M., T.Joro, P.Korhonen, S.Salo and J.Wallenius, A Value Efficiency Approach to Incorporating Pre f-erence Information in Data Envelopment Analysis, *Management Science*, 45, pp. 103 -115, 1999.
26. Joro, T., P.Korhonen, and J.Wallenius, Structural Comparison of Data Envelopment Analysis and Multiple Objective Linear Programming, *Management Science*, 44, pp. 962 -970, 1998.
27. Newhouse, J.P. , Frontier Estimation: How Useful a Tool for Health Economics, *Journal of Health Economics*, 13, pp. 317 -322, 1994.
28. Olesen, O. and N.C.Petersen, The Use of Data Envelopment Analysis with Probabilistic Assurance Regions for Measuring Hospital Efficiency, *Journal of Productivity Analysis*, 2001, to appear.
29. Simar, L. and P.Wilso n, Statistical Inference in Nonparametric Frontier Models: The State of the Art *, Journal of Productivity Analysis*, 13, pp.- 49 -78, 2000.