CENTRE FOR HEALTH
PROGRAM EVALUATION

WORKING PAPER 136

# Competing Methods for Efficiency Measurement

## A Systematic Review of Direct DEA vs SFA/DFA Comparisons

**Mr Duncan Mortimer**
Research Associate, School of Medicine,
Health Policy & Practice, University of East Anglia

September, 2002
ISSN 1325-0663
ISBN 1 876662 56 5

## CENTRE PROFILE

The Centre for Health Program Evaluation (CHPE) is a research and teaching organisation established in 1990 to:

- undertake academic and applied research into health programs, health systems and current policy issues;
- develop appropriate evaluation methodologies; and
- promote the teaching of health economics and health program evaluation, in order to increase the supply of trained specialists and to improve the level of understanding of these disciplines within the health community.

The Centre comprises two independent research units, the Health Economics Unit (HEU) which is part of the Faculty of Business and Economics at Monash University, and the Program Evaluation Unit (PEU) which is part of the Department of Public Health at The University of Melbourne. The two units undertake their own individual work programs as well as collaborative research and teaching activities.

## PUBLICATIONS

The views expressed in Centre publications are those of the author(s) and do not necessarily reflect the views of the Centre or its sponsors. Readers of publications are encouraged to contact the author(s) with comments, criticisms and suggestions.

A list of the Centre's papers is provided inside the back cover. Further information and copies of the papers may be obtained by contacting:


The Co-ordinator
Centre for Health Program Evaluation
PO Box 477
West Heidelberg Vic 3081, Australia
**Telephone** + 61 3 9496 4433/4434 **Facsimile** + 61 3 9496 4424
**E-mail** CHPE@BusEco.monash.edu.au
Or by downloading from our website
**Web Address** http://chpe.buseco.monash.edu.au

## ACKNOWLEDGMENTS

The Health Economics Unit of the CHPE is supported by Monash University.

The Program Evaluation Unit of the CHPE is supported by The University of Melbourne.

Both units obtain supplementary funding through national competitive grants and contract research.

The research described in this paper is made possible through the support of these bodies.

## AUTHOR ACKNOWLEDGMENTS

The research reported herein was, in part, completed towards the author's Master of Economics (Honours) degree, awarded by the Department of Economics at Monash University. Professor Jeff Richardson and Dr. Stuart Peacock provided much helpful support and advice during completion of this degree. The Australian Health Ministers Advisory Council (AHMAC), the Monash University Research Fund (MURF), the National Health and Medical Research Council (NH&MRC) and the Centre for Health Program Evaluation (CHPE) provided the necessary financial support to allow completion of this report and related research projects.

# Table of Contents



# List of Tables

# Abstract

Various authors have advised a wait and see approach in evaluating the relative precision of alternative techniques, such as data envelopment analysis (DEA) and stochastic frontier analysis (SFA), in estimating industry-average and firm-specific inefficiency. Chirikos and Sear (2000), for example, contend that "policy-makers may be well advised to wait until additional research clarifies reasons why DEA and stochastic frontier models yield divergent results" (p. 1389).

The main objective of this paper is to highlight the likely trade-offs between competing methods based on direct empirical comparisons using simulated data and to demonstrate the wealth of evidence bearing on a range of real-world applications. Whilst this systematic review indicates that a good deal of evidence is already available, evidence of a different sort may be required to identify a 'correct' approach in addressing specific policy problems. In particular, the now routine practice of cross checking should be taken one step further to include realistic simulation studies along-side real-world DEA *vs* SFA comparisons.

# Competing Methods for Efficiency Measurement

## A systematic Review of Direct DEA vs SFA/DFA Comparisons

### Introduction

Data envelopment analysis (DEA) and stochastic frontier analysis (SFA) employ quite distinct methodologies for frontier estimation and efficiency measurement, each with associated strengths and weaknesses, such that a trade-off exists in selecting the 'correct' approach:

> "…non-statistical approaches such as DEA have the disadvantage of assuming no statistical noise, but have the advantage of being non-parametric and requiring few assumptions about the underlying technology. SFA models on the other hand have the attraction of allowing for statistical noise, but have the disadvantage of requiring strong assumptions as to the form of the frontier" (Jacobs, 2000 p. 3).

Previous studies have attempted to clarify this trade-off, so that the choice of a 'correct' method is rather more clear cut in particular applications. Banker, Gahd and Gorr (1993), for example, report findings from a Monte Carlo experiment to the effect that the relative precision of DEA and SFA is context specific. DEA is favoured where measurement error is unlikely to pose much of a threat and where the assumptions of neoclassical production theory are in question. Conversely, SFA should have the advantage in coping with severe measurement error and where simple functional forms provide a close match to the properties of the underlying production technology. Gong and Sickles (1993) report findings along similar lines so that "...as mis-specification of functional form becomes more serious, DEA's appeal (vis-à-vis SFA) becomes more compelling" (Gong & Sickles, 1993 p. 259).

The purpose of this article is to highlight the likely trade-offs between competing methods for frontier estimation based on a systematic literature review of direct empirical comparisons. Clearly, the set of pair-wise comparisons is steadily growing as new methods for frontier estimation and efficiency measurement arise to address the shortcomings of more traditional methods. In recognition of this fact, the review is restricted to comparisons between data envelopment analysis (DEA) and the two most commonly employed parametric alternatives: deterministic frontier analysis (DFA) and stochastic frontier analysis (SFA).

### Frontier Estimation

To provide some necessary background, a brief review of the methods is provided - focusing on the general approach employed in estimating the minimum cost frontier (rather than technical details). In each case, the choice between methods impacts upon the shape, location and interpretation of the resulting frontier.

Data envelopment analysis (DEA) utilises linear programming to fit a boundary function to observational data for a sample of relatively homogeneous firms. The method is distribution free and allows the data to 'speak for themselves' (Bates, Baines & Whynes, 1996). More specifically, the DEA frontier is floated beneath observed cost-output combinations so that the functional form of the cost frontier is determined by the best extremal fit given convexity constraints and assuming free disposal of both inputs and outputs (Seiford & Thrall, 1990).

Whereas DEA allows the data to 'speak for themselves', parametric methods such as DFA and SFA assume a structure and then fit curve. Several options are available in estimating best-practice frontiers via the parametric approach. The simplest option is to 'correct' the location of the least-squares regression line so as to reflect the behavioural ideal of best-practice rather than an industry-average. An alternative tact, under the parametric approach, is to employ maximum likelihood (ML) methods to estimate the best-practice frontier. ML estimation allows efficient firms to have a greater say in the shape of the frontier, so as to capture any "structural dissimilarity between OLS and frontier technology" (Lovell, 1993 p. 22).

## Efficiency Measurement

The preceding discussion implies that DEA, DFA and SFA frontiers deliver quite different benchmarks describing best-practice. Note, however, that estimating a frontier or benchmark completes only half the job. The next step requires a comparison between actual and frontier production costs to provide a measure of relative efficiency.

Deterministic methods, such as DEA and DFA, characterise deviations from best-practice as entirely due to technical or allocative inefficiency. Inefficient firms are assumed to be capable of producing on the best-practice frontier simply by adopting efficient production methods and all variation in observed cost-output combinations is assumed to be within the control of the firm. More formally, the economic efficiency of a the $j^{th}$ firm ($e_j$) is equal to the ratio of actual ($C_j$) to frontier ($C^*$) production cost: $e_j = C^* / C_j$ where $e_j \in (0, 1)$ and $C_j = C^* + e_j$. In short, measurement error and random variation are simply assumed away and deviations from the frontier are attributed solely to inefficiency.

The value of the stochastic methods such as SFA lies in the assumption that the actual performance of each evaluated firm reflects a range of factors that relate to good fortune as well as to good practice. "Observed hospital costs may deviate from an efficient cost frontier due to events that are both within and outside of the hospital's control" (Zuckerman, Hadley and Iezzoni, 1994 p. 274). More formally, the stochastic frontier approach treats deviations ($e_j$) from best-practice as 'composed residuals' comprising two components: a one-sided inefficiency term ($u_j$) reflecting managerial competence; and a symmetric random error ($v_j$) reflecting omitted variables, measurement error and stochastic elements beyond managerial control. Comparisons between the stochastic frontier ($C^* + v_j$), actual production cost ($C_j$), and the deterministic frontier ($C^*$), reflect the share of excess production cost attributable to random error and inefficiency. More precisely, the economic efficiency of the $j^{th}$ firm is given as: $u_j = (C^* + v_j) / C_j$ where $u_j \in (0, 1)$, $C_j = C^* + e_j$, and $e_j = v_j + u_j$.

## Competing Paradigms?

It should be obvious from the above discussion that DFA provides an unhappy compromise between DEA and SFA (at least at a conceptual level). Whereas DFA and SFA adopt a similar approach to frontier estimation; DEA and DFA have much in common when it comes to efficiency measurement. In short, DFA "combines the bad features of the econometric and programming approaches to frontier construction: it is deterministic *and* parametric" (Lovell, 1993 p. 21). Note, however, that pairwise comparisons involving DFA should isolate the differences along one of two relevant dimensions: **deterministic *vs* stochastic efficiency measurement** *or* **non-parametric *vs* parametric frontier estimation**. In contrast, DEA and SFA efficiency scores are expected to differ due to both dimensions.

Recall that previous studies have attempted to clarify the trade-off between competing paradigms so that the choice of a 'correct' method is rather more clear-cut in certain applications. DEA is favoured where measurement error is unlikely to pose much of a threat and where the assumptions of neoclassical production theory are in question. Conversely, SFA should have the advantage in coping with measurement error and where simple functional forms provide a close match to the properties of the underlying production technology. Unfortunately, neither DEA nor SFA is ideally suited across all applications. In estimating hospital efficiency, for example, both measurement error and functional form are likely to cause problems. In short, the analysis of hospital efficiency really requires a hybrid approach that copes well with random variation and measurement error; but that is also flexible in modelling both the underlying production technology and the objectives and conduct of individual hospitals. Some progress has been made towards making DEA stochastic (see Sengupta, 1987; 1998) and SFA more flexible (see Lovell, 1993). In the meantime, the relative precision and policy value of alternative measures of hospital efficiency remains an empirical question and it is not possible to install a 'gold standard' based solely on *a priori* deliberation.

## Search Strategy

The review of empirical comparison that follows is based on literature identified from an initial search of citation databases, together with supplementary searches of the authors own citation databases, review article bibliographies, and web-based resources. The initial search included:

***ISI Web of Science***

- Science Citation Index (SCI-X) Expanded, 1981- April 2002: 25 records
- Social Sciences Citation Index (SSCI), 1981- April 2002: 51 records

***OVID***

- Pre-MEDLINE & MEDLINE, 1966- April 2002: 2 records
- Journals@Ovid Full Text, April 2002 edition: 1 record

Searches were conducted using the following search terms: *("STOCHASTIC FRONTIER" OR "FRONTIER ESTIMATION" OR "DETERMINISTIC FRONTIER") AND ("DATA ENVELOPMENT ANALYSIS")*, yielding just over 100 articles after removal of duplicates. Articles published in a language other than English were excluded from the review. Abstracts (and, if necessary, articles) were then scanned to identify 51 articles and unpublished papers reporting at least one pair-wise empirical comparison between DFA, DEA and SFA.

## Empirical Evidence

Both DEA and SFA have the potential to deliver biased estimates of inefficiency due to specification errors of one sort or another. Biases in opposite directions raise the possibility of fairly substantial divergence between DEA-, DFA- and SFA-based estimates. In such circumstances, correspondence between real-world DEA, DFA and SFA efficiency scores provides some reassurance that competing methods are accessing similar latent variables. Note, however, that correspondence alone provides no guarantee that competing methods are accessing the 'target' construct and might simply reflect biases in the same direction. Fortunately, the use of simulated data delivers a criterion against which to quantify potential errors and biases under a wide range of different conditions. The simulation studies reviewed below include all possible pair-wise comparisons[1] across the two relevant dimensions: non-parametric *vs* parametric frontier estimation and deterministic *vs* stochastic efficiency measurement.

### Accuracy: Simulation Studies

For the most part, well-established variants of alternative frontier estimation techniques provide a fairly accurate picture of the relative efficiency of individual production units. Banker, Chang & Cooper (1996), for example, reported mean absolute deviations between DEA-/DFA-based estimates of output-oriented technical efficiency and 'actual' efficiency scores that ranged from 0.006 to 0.054 efficiency points (depending on sample size and returns to scale). Banker, Charnes, Cooper & Maindiratta (1987) found a similar correspondence (mean absolute deviations of between 0.003 and 0.049) between DEA-/DFA-based estimates and actual efficiency scores. Note, however, that mean absolute deviations between DEA- and DFA-based estimates (rather than between estimates and actual scores) might be somewhat larger if competing methods diverge in opposite directions. Moreover, the above findings are based on simulated data in which deviations between the efficient frontier and observed production points really are *entirely* due to inefficiency.

In real-world applications, measurement error is the rule rather than the exception. Inclusion of measurement error to simulate 'noisy' data would therefore seem an obvious step in making simulated production environments more realistic. Findings from Banker, Gadh & Gorr (1993) suggest that choice of the 'correct' or 'best' estimation method is likely to be much more important in the presence of measurement error. At low levels of measurement error, mean absolute deviations between DEA-/SFA-based estimates and 'actual' efficiency scores varied between 0.03 and 0.11 efficiency points (depending on sample size, technology, and the distributions of inefficiency and measurement error). At higher levels of measurement error, the gap between estimates and 'actual' efficiency scores widened, with mean absolute deviations of between 0.08 and 0.40. In short, "neither method performed satisfactorily for high measurement errors" (Banker, Gahd & Gorr, 1993 p. 332).

In contrast, Yu (1998) estimated DEA-/SFA-based efficiency scores against a background of fairly high levels of measurement error and, for lower values of exogenous variables, found mean absolute deviations between estimated and actual scores no higher than 0.161 and correlations ranging from 0.62 to 0.89. Similarly, Resti (2000) estimated DEA-/SFA-based scores for overall economic efficiency in the presence of 'low', 'medium' and 'high' levels of noise. Resti reported mean absolute deviations from actual scores ranging from 0.004 to 0.063, and correlations between estimated and actual scores of between 0.63 and 1.00. In short, "all the 'classic' techniques performed rather satisfactorily in measuring the amount of inefficiency, although their performance can worsen in some specific situations" (Resti, 2000 p. 568).

Under relatively benign conditions, frontier methods also manage to get reasonably close to the mark in characterising the properties of the underlying production technology. Banker, Chang & Cooper (1996) correctly identified returns to scale properties in up to 87.9% of observations for some DEA-based estimates (but the proportion misclassified reached 38.33% for some DFA-based estimates). Banker, Charnes, Cooper & Maindiratta (1987) reported mean absolute

[1] In particular: SFA *vs* SDEA (Resti, 2000), DFA *vs* DEA (Banker, Chang & Cooper, 1996; Banker, Charnes, Cooper & Maindiratta, 1987), SFA *vs* DEA (Banker, Gadh & Gorr, 1993; Bojanic, Caudill & Ford, 1998; Gong & Sickles, 1992; Resti; 2000; Yu, 1998), and SFA *vs* DFA (Ruggerio, 1999).

deviations between DEA-based estimates and actual rates of technical substitution as low as 0.085 (but as high as 0.24 for some DFA-based estimates). Similarly, the proportion of observations misclassified with respect to scale effects was as low as 6.4% for some DEA-based estimates (but as high as 40.4% for some DFA-based estimates).

## Pair-wise Comparisons: Simulation Studies

Recall that comparisons between DEA and DFA isolate differences due to parametric *vs* non-parametric frontier estimation. Banker, Charnes & Cooper (1996) conclude that 'DEA generally shows superior performance' (p. 233) with variable returns to scale (VRTS) DEA models dominating DFA-based alternatives under all conditions. Note, however, that DFA was often 'closer to the mark' than constant returns to scale (CRTS) DEA models and it is not possible to specify a 'best' model for all conditions. Under a more restricted set of conditions Banker, Charnes, Cooper & Maindiratta (1988) found DEA-based estimates of output-oriented technical efficiency to be more accurate and more stable than their DFA-based counterparts. In short, comparisons between DEA and deterministic variants of the parametric approach simply confirm what was already suspected: the move from DEA to DFA comes at the cost of reduced flexibility in modelling the production environment and offers nothing by way of compensation.

In assuming an environment characterised by measurement errors and stochastic variation, SFA appears (at least at the conceptual level) less of a straw man. Even so, results drawn from Banker, Gahd & Gorr's (1993) DEA *vs* SFA comparison "show DEA to produce more accurate efficiency estimates …even with remarkably high measurement errors present" (p. 341). SFA only gains the upper hand when measurement errors reach a threshold of between ±17% to 45% of observed output values (depending on sample size, technology, and the distribution of inefficiency). Results also suggest that SFA is more accurate whenever sample size reaches a threshold of 50 units and distributional *assumptions* mirror 'actual' distributions of noise and inefficiency[2]. In other words, the expected trade-off applies with relatively few caveats: SFA has the advantage in coping with *severe* measurement error or when the distributional assumptions required in separating measurement error from inefficiency accurately reflect the properties of the underlying production environment.

More recent studies have sought to generalise decision-rules to more realistic production environments. Yu (1998), for example, simulated the impact of exogenous factors (beyond managerial control) with varying degrees of influence over output. One-step, two-step and first-stage SFA-based estimates of technical efficiency were generally more accurate than any of the DEA-based alternatives. Within the SFA camp, the one-step procedure strictly dominates two-step and first-stage options in accounting for the impact of exogenous variables. Within the DEA camp, first-stage and two-step options are preferred whenever the impact of exogenous factors is relatively minor. As the influence of exogenous factors on observed performance increases, the Banker and Morey (1986) one-step procedure edges in front of the DEA-based competition. In every case, the one-step SFA procedure yields lower mean absolute deviations and higher rank correlations between estimated and actual scores than the next best option.

[2] Results also suggest that, for small sample sizes and small measurement errors, DEA and SFA are biased in *opposite* directions. As such, "…a combination frontier may be more accurate than either estimate by themselves" (Banker, Gahd & Gorr, p. 341).

In a similar vein, Resti (2000) generated simulated production units that mimicked the scale and scope of production, concentration of market power, and various other features of a real-world banking sector. Results suggest that SFA-based estimates of overall economic efficiency are always more accurate for the large sample size of 500 units and, as expected, the relative advantage of the stochastic frontier approach increases with the amount of noise in the data. Even for Resti's small sample size of 50 units, SFA occasionally edges out the DEA-based competition (depending on the form of the frontier, the magnitude of between-unit differences in actual efficiency scores and the amount of noise in the composed residuals). For the small sample size, DEA is most accurate and most likely to outstrip SFA when RTS assumptions mirror RTS properties of the simulated production technology (Resti, 2000).

## Impact of Specification Error: Simulation Studies

Specification error is expected to have two effects: firstly, it should weaken the correspondence between estimates and true values and secondly, certain sorts of errors should favour one or other of the competing approaches. More specifically:

- DEA is usually thought to be less accurate and more erratic at 'corner points' where few, if any, observations are available to provide a reliable standard of comparison in estimating the best-practice frontier. This is largely because measurement error is more likely to influence the location and shape of the frontier around corner points, but also because VRTS models tend to confuse technical efficiency with scale effects when regions on the feasible set remain 'hidden'. Competing parametric methods encounter a similar problem because it is difficult to fit a well-specified curve over sparsely populated regions of the feasible set (Read and Thanassoulis, 1996).

  In line with expectations, Resti (2000) and Yu (1998) found that VRTS DEA models tended to overestimate the efficiency of atypical and outlying production units. Similarly, Read and Thanassoulis (1996) reported higher mean absolute deviations around atypical and outlying observations for SFA-/DEA-based estimates of technical efficiency. Banker, Chang and Cooper (1996) reported an increase in mean absolute deviations for both DEA- and DFA-based estimates in the vicinity of corner points. Moreover, efficiency scores drawn from Banker, Chang and Cooper's (1996) VRTS DEA models were most erratic when faced with sparse comparison sets around corner points.

- Banker, Charnes, Cooper & Maindiratta (1988) encountered problems due to the non-parametric equivalent of incorrect functional form. More specifically, "DEA performs poorly for observations that fall in the region where the 'true' production possibility set is non-convex, violating the convexity axiom in DEA" (p. 50). In a similar vein, Gong and Sickles (1992) reported that failure to correctly specify the form of *parametric* frontiers, relying instead on overly flexible functional forms such as the translog, leads to imprecise estimates of technical efficiency. Banker, Charnes and Cooper (1996), Ruggiero (1999) and Resti (2000) encountered much the same problem when relying on flexible functional forms in smaller samples. Resti (2000) attributed this loss of precision to multicollinearity between cross products and higher-order terms in the unrestricted translog.

- A number of studies (eg. Banker, Gahd & Gorr, 1993; Read & Thanassoulis, 1996; Ruggiero, 1999; Resti, 2000; Ondrich & Ruggiero, 2001) have reported a loss of precision for DEA, DFA and/or SFA in the presence of higher levels of measurement error. Recall that deterministic methods, such as DEA and DFA, are usually regarded as particularly susceptible to measurement error because deviations from the frontier are attributed *solely* to inefficiency. Results from Resti (2000) confirm the superior performance of SFA in the presence of higher levels of measurement error. Read and Thanassoulis (1996) encountered problems (mean absolute deviations as high as 0.284) in obtaining accurate DEA-based estimates of technical efficiency from noisy data. In contrast, SFA-based estimates from the same study remained reasonably close to the mark even at 'high' levels of noise (Read & Thanassoulis, 1996).

Note, however, that a number of findings call into question SFA's comparative advantage in handling noisy data. Banker, Gahd and Gorr (1993), for example, "show DEA to produce more accurate efficiency estimates… even with remarkably high measurement errors present" (p. 341). In particular, COLS-based SFA models frequently characterised deviations from the frontier as entirely due to inefficiency, "leading to overall poor performance relative to DEA" (Banker, Gahd & Gorr, 1993 p. 337). Ruggiero (1999) and Ondrich and Ruggiero (2001) report similar results in comparing DFA- and SFA-based models. Even in the presence of 'high' levels of measurement error, the deterministic approach sometimes out-performs the SFA-based competition (Ruggiero, 1999; Ondrich and Ruggiero, 2001). In each case, SFA failed to accurately separate noise and inefficiency: calling into question one of the chief selling points of the stochastic frontier approach - the ability to cope with noisy data. In other words, these results suggest that the 'poor-cousin' status of DFA might not stand up to scrutiny.

- Bojanic, Caudill & Ford (1998) allowed the extent of measurement error to increase as output increased (ie. heteroskedastic measurement error). Both DEA and SFA systematically overestimated the inefficiency parameter in the presence of severe and heteroskedastic measurement errors[3]. Whilst neither method delivered satisfactory estimates for technology and efficiency parameters, SFA-based estimates consistently outperformed their DEA-based counterparts[4].

- Banker, Charnes and Cooper (1996) simulated the impact of common specification errors via the inclusion of an irrelevant variable and the omission of a relevant variable. Inclusion of an irrelevant variable frequently increased mean absolute deviations but left the relative standing of DEA- and DFA-based models largely intact. In contrast, omitting a relevant variable *always* increased mean absolute deviations and dramatically eroded the relative standing of estimates drawn from CRTS DEA models (relative to the DEA- and DFA-based competition) with respect to both accuracy and consistency (Banker, Charnes & Cooper, 1996). Ruggiero (1999) reported similar problems due to omitted variables when

[3] DEA-based estimates were, however, much more inflated than their SFA-based counterparts. For the case of moderate inefficiency and moderate heteroskedasticity, SFA overestimated the inefficiency parameter by about five times and DEA by about ten times. For the case of low inefficiency and high heteroskedasticity, SFA overestimated the inefficiency parameter by about 25 times and DEA by about 50 times (Bojanic, Caudill & Ford, 1998).

[4] Note that the relative standing of DEA and SFA was (at least partly) a result of 'rigging' the simulation in favour of SFA. Specifically, the relatively poor performance of DEA is less than surprising given the presence of fairly severe levels of measurement error. In addition, the functional form of the true frontier and the distributions of both error and inefficiency were known in advance; eliminating two sources of specification error that are specific to the stochastic frontier approach.

estimating SFA production frontiers. Because any loss of precision can be attributed to an increase in noise (that would otherwise be explained by the omitted variable), and because the performance of all methods declines as noise increases; Ruggiero (1999) argues that the adverse impact of omitting a relevant variable is common to all competing methods.

## Real-world Comparisons

Simulated studies have the advantage of providing a *criterion* against which to compare the performance of competing methods for frontier estimation and efficiency measurement. In particular, the true values of technology and inefficiency parameters are known and available for comparison against estimated values. Note, however, that simulation studies tend to assume away the sort of complications that arise for real-world applications. In attempting to provide a neutral setting and a fair comparison, simulated production units, technologies, and production environments might not be all that realistic in capturing the multi-plant, multi-product organisation of many real-world firms[5] and might not provide a true test of the techniques (Resti, 2000). The stability of, and correspondence between, competing methods *in real-world applications* therefore make an important contribution to the weight of empirical evidence.

The majority of direct DEA *vs* SFA/DFA comparisons reported in the literature have found moderate to strong correspondence when ranking financial institutions (eg. Drake & Weyman, 1996; Resti, 1997), railways (Coelli & Perelman, 1999), social security offices (Bjurek, Hjalmarsson & Forsund, 1990), cement plants (Hjalmarsson, Kumbhakar & Heshmati, 1996), pig farms (Sharma, Leung & Zaleski, 1997), school districts (Ruggiero & Vitaliano, 1999), local governments (de Borger & Kerstens, 1996), and acute care hospitals (eg. Linna, 1998; Webster, Kennedy & Johnson, 1998). A smaller number of studies found only mediocre to poor correspondence in ranking acute care hospitals (eg. Chirikos & Sear, 2000), financial institutions (eg. Ferrier & Lovell, 1990), and electrical utilities (eg. Ray & Mukherjee, 1995). A summary review of the 41 real-world DEA *vs* SFA/DFA comparisons identified in the literature is provided in table 1 below.

[5] For example, hospitals are typically multi-product firms that might operate on several campuses and employ a broad range of heterogeneous inputs in the production of teaching, research, hotel services, and patient care. Moreover, episodes of care are likely to differ with respect to c ase-mix factors such as diagnosis, treatment, severity, age, complications and co-morbidity.

## Table 1 - Direct DEA *vs* SFA/DFA Comparisons

| Author / Date | Country | Units / Period | DEA Model: Orientation | DEA Model: RTS | SFA/DFA Model: Form | SFA/DFA Model: Frontier | Findings: DEA | Findings: SFA/DFA | Findings: Comparison |
|---|---|---|---|---|---|---|---|---|---|
| Banker, Chang & Cooper (1996) | N / A | simulation study: 50 - 200 units | output-increasing | CRTS | Cobb-Douglas | deterministic prod. frontier | "Both DEA and COLS generally give good results at all sample sizes. In evaluating efficiency, DEA generally shows superior performance" (p. 233). | | |
| | | | | VRTS | translog | | | | |
| Banker, Charnes, Cooper & Maindiratta (1987) | N / A | simulation study: 100 & 500 units | output-increasing | VRTS | translog | deterministic prod. frontier | "DEA outperforms the parametric approach ...the piecewise linear production function is more flexible in approximating the true frontier… simulation results reported in this study bear this out" (pp. 42 & 50). | | |
| Banker, Conrad & Strauss (1986) | US | 114 acute care hospitals 1978 / 79 | input-reducing | CRTS | CRTS translog | deterministic TC frontier | 45 units $TE_x = 1$<br>37 units $0.9 \leq TE_x \leq 1$<br>32 units $TE_x < 0.9$ | 45 units $EE_x = 1$<br>37 units $0.9 \leq EE_x \leq 1$<br>32 units $EE_x < 0.9$ | ($\chi^2 = 11.79$, df = 4, $p < 0.05$) ie. 'broad agreement' |
| Banker, Gahd & Gorr (1993) | N / A | simulation study: 25 - 200 units | output-increasing | VRTS | translog | stochastic prod. frontier | In most cases, "DEA produces more accurate efficiency estimates than COLS, even with remarkably high measurement errors" (p. 341). | | |
| Bates (1997) | England | 96 local education authorities, 1980s | output-increasing | CRTS | Cobb-Douglas | stochastic prod. frontier | 23 units $TE_y > 0.95$<br>55 units $0.95 \leq TE_y \leq 0.9$<br>18 units $TE_y < 0.9$ | 16 units $TE_y > 0.95$<br>71 units $0.95 \leq TE_y \leq 0.9$<br>9 units $TE_y < 0.9$ | "…much agreement between DEA and SF results" (p. 92). |
| Bauer, Berger, Ferrier & Humphrey (1998) | US | 683 banks, 1977 - 1988 | cost-reducing | VRTS | translog | stochastic TC frontier | $EE_x$ means: 0.21-0.385 | $EE_x$ means: 0.875-0.88 | Spearman's rho: 0.10-0.17 |
| Bjurek, Hjalmarsson & Forsund (1990) | Sweden | 400 social security offices 1974 - 1984 | input-reducing | CRTS | Cobb-Douglas | deterministic input-req. frontier | potential input-saving: 0.17-0.31 | potential input-saving: 0.18-0.26 | Spearman's rho > 0.83 |
| | | | | DRTS NITRS | quadratic | | potential input-saving: 0.16-0.25 | potential input-saving: 0.12-0.20 | Spearman's rho > 0.87 |
| Bojanic, Caudill & Ford (1998) | N / A | simulation study: 25 - 200 units | cost-reducing | CRTS | Cobb-Douglas | stochastic TC frontier | "…heteroskedasticity in the two-sided error term introduces substantial biases into ML, COLS and DEA estimators. Although none perform well, both ML and COLS are found to be superior to DEA" (p. 140). | | |
| Brümmer (2001) | Slovenia | 147 private farms, 1995 & 1996 | revenue-increasing | VRTS | translog | stochastic prod. frontier | $TE_y$ means: 0.43-0.45 | $TE_y$ means: 0.74-0.75 | Spearman's rho: 0.69 |

### Table 1 (cont.) - Direct DEA *vs* SFA/DFA Comparisons

| Author / Date | Country | Units / Period | DEA Model | | SFA/DFA Model | | Findings | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Orientation | RTS | Form | Frontier | DEA | SFA/DFA | Compare |
| Chirikos & Sear (2000) | US | 186 acute care hospitals, 1982 - 1993 | cost-reducing | CRTS | translog | stochastic TC frontier | $EE_x$ mean: 0.80 | $EE_x$ mean: 0.85 | Pearson's r: 0.33 |
| | | | | | CRTS translog | | | $EE_x$ mean: 0.75 | Pearson's r: 0.13 |
| | | | | | hybrid | | | $EE_x$ mean: 0.82 | Pearson's r: 0.26 |
| Coelli & Perelman (1999) | Europe | 17 railway firms, 1988 -1993 | input-reducing | CRTS | CRTS translog | deterministic input dist $f$ | $TE_x = TE_y$ means: 0.81-0.88 | $TE_x = TE_y$ mean: 0.78 | Pearson's r: 0.56-0.71 |
| | | | output-increasing | | | deterministic output dist $f$ | | | |
| | | | input-reducing | VRTS | translog | deterministic input dist $f$ | $TE_x$ means: 0.86-0.93 | $TE_x$ mean: 0.90 | Pearson's r: 0.29-0.43 |
| | | | output-increasing | | | deterministic output dist $f$ | $TE_y$ means: 0.88-0.93 | $TE_y$ mean: 0.89 | Pearson's r: 0.26-0.43 |
| Cooper, Kumbhakar, Thrall & Yu (1995) | China | 3 industries (textiles, chemicals, metals), 1966 - 1988 | output-increasing | CRTS | Cobb-Douglas | stochastic prod. frontier | "Although there are discrepancies, results from DEA and SFA generally tend toward confirmation. …this consistency increases confidence that behaviour is 'in the observations' and not simply a reflection of our models" (p. 101). | | |
| Cummins & Zi (1998) | US | 445 life insurers, 1988 -1992 | cost-reducing | VRTS | translog | stochastic TC frontier | $EE_x$ means: 0.42-0.50 | $EE_x$ means: 0.44-0.86 | Spearman's rho: 0.56-0.60 |
| De Borger & Kerstens (1996) | Belgium | 589 local govts, 1985 | cost-reducing | VRTS | translog | deterministic TC frontier | $EE_x$ mean: 0.727 | $EE_x$ mean: 0.570 | Pearson's r: 0.81 Spearman's rho: 0.81 |
| | | | | | | stochastic TC frontier | | $EE_x$ means: 0.78-0.81 | Pearson's r: 0.82-0.83 Spearman's rho: 0.82 |
| Dismuke & Sena (1999) | Portugal | 2 DRGs in 52 central & general hospitals, 1992 -1994 | input-reducing | CRTS | linear | stochastic input-req. frontier | "…technical efficiency change and technological change computed using DEA… results are generally consistent with those obtained from parametric methods" (pp. 112-113). | | |
| Drake & Weyman-Jones (1996) | UK | 46 building societies, 1988 | cost-reducing | VRTS | translog | stochastic TC frontier | $EE_x$ mean: 0.876 | $EE_x$ mean: not reported | Spearman's rho: 0.97 |
| Eisenbeis, Ferrier & Kwan (1999) | US | 254 banks, 1986 - 1991 | cost-reducing | VRTS | restricted translog | stochastic TC frontier | $EE_x$ means: 0.60-0.72 | $EE_x$ means: 0.81-0.92 | Spearman's rho: 0.444-0.589 |

### Table 1 (cont.) - Direct DEA *vs* SFA/DFA Comparisons

| Author / Date | Country | Units / Period | DEA Model | | SFA/DFA Model | | Findings | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Orientation | RTS | Form | Frontier | DEA | SFA/DFA | Compare |
| Ferrier & Lovell (1990) | US | 575 financial institutions, 1984 | cost-reducing | VRTS | translog | stochastic TC frontier | $TE_x$ mean: 0.84<br>$AE_x$ mean: 0.95<br>$EE_x$ mean: 0.79 | $TE_x$ mean: 0.91<br>$AE_x$ mean: 0.83<br>$EE_x$ mean: 0.74 | Spearman's rho: 0.014-0.017 |
| Gong & Sickles (1992) | N / A | simulation study: 50 units over 10 to 50 periods | input-reducing | VRTS | translog, CES translog & generalised Leontief stochastic prod. frontiers | | "…if the employed form is close to the underlying technology, stochastic frontier models outperform DEA. As specification error becomes more serious, DEA's appeal becomes more compelling" (p. 259). | | |
| Guiffrida & Gravelle (1998) | England | 90 FHSAs, 1993/94 - 1994/5 | cost-reducing | CRTS | hybrid | deterministic TC frontier | $EE_x$ means: 0.89-0.97 | $EE_x$ means: 0.86-0.91 | Spearman's rho: 0.26-0.70 |
| | | | | | | stochastic TC frontier | | $EE_x$ means: 0.89-0.99 | Spearman's rho: 0.22-0.71 |
| Hjalmarsson, Kumbhakar & Heshmati (1996) | Colombia | 15 cement plants, 1968 - 1988 | output-increasing | CRTS | Zellner-Revankar | deterministic prod. frontier | $TE_y$ means: 0.72-0.89 | $TE_y$ means: 0.68-0.78 | Spearman's rho: 0.52-0.80 |
| | | | | VRTS | | | $TE_y$ means: 0.75-0.97 | | Spearman's rho: 0.35-0.77 |
| | | | | CRTS | translog | stochastic prod. frontier | $TE_y$ means: 0.72-0.89 | $TE_y$ means: 0.71-0.97 | Spearman's rho: -0.08-0.75 |
| | | | | VRTS | | | $TE_y$ means: 0.75-0.97 | | Spearman's rho: -0.37-0.92 |
| Jacobs (2001) | UK | 232 NHS hospitals, 1995/96 | cost-reducing | VRTS | linear | stochastic TC frontier | $EE_x$ means: 0.65-0.94 | $EE_x$ means: 0.83-0.88 | Pearson's r: 0.43-0.63 |
| Johnes (1998) | UK | 50 universities, 1989/90 | cost-reducing | CRTS<br>VRTS | quadratic | stochastic TC frontier | "…comparison of efficiency measures obtained by SFA and DEA confirms that these are broadly in accord with one another. …the magnitude of the rank correlation coefficient, at 0.133, is rather low, however" (p. 206). | | |
| Linna (1998) | Finland | 43 acute care hospitals, 1988 - 1994 | cost-reducing | CRTS | Box-Cox transform | stochastic TC frontier | $EE_x$ means: 0.70-0.93 | $EE_x$ means: 0.84-0.93 | Spearman's rho: 0.57-0.75 |
| | | | | VRTS | | | $EE_x$ means: 0.79-0.96 | | Spearman's rho: 0.54-0.72 |
| Linna & Häkkinen (1999) | Finland | 95 acute care hospital units, 1994 | cost-reducing | CRTS | Box-Cox transform | stochastic TC frontier | $EE_x$ means: 0.90-0.92 | $EE_x$ mean: 0.86 | Spearman's rho: 0.41-0.58 |
| | | | | VRTS | | | $EE_x$ means: 0.85-0.87 | | Spearman's rho: 0.59-0.61 |

**Table 1 (cont.) - Direct DEA *vs* SFA/DFA Comparisons**

| Author / Date | Country | Units / Period | DEA Model | | SFA/DFA Model | | Findings | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Orientation | RTS | Form | Frontier | DEA | SFA/DFA | Compare |
| Linna & Häkkinen (1998) | Finland | 95 acute care hospital units, 1994 | cost-reducing | CRTS | Cobb-Douglas | stochastic TC frontier | $EE_x$ mean: 0.84-0.89 | $EE_x$ mean: 0.89-0.93 | Spearman's rho: 0.52-0.63 |
| | | | | | translog | | | $EE_x$ mean: not reported | Spearman's rho: 0.28 |
| | | | | | Box-Cox transform | | | $EE_x$ mean: 0.86-0.89 | Spearman's rho: 0.55-0.68 |
| Meibodi (1998) | developing countries | 26 electricity industries, 1987 - 1988 | input-reducing | CRTS | Cobb-Douglas | stochastic prod. frontier | $TE_x$ mean: 0.72 | $TE_y$ mean: 0.77 | Pearson's r : 0.48<br>Spearman's rho: 0.36 |
| | | | | VRTS | | | $TE_x$ mean: 0.79 | | Pearson's r : 0.71<br>Spearman's rho: 0.66 |
| Mortimer (2001) | Australia | 38 public hospitals, 1993 | cost-reducing | VRTS | Cobb-Douglas | stochastic TC frontier | $EE_x$ means: 0.83-0.86 | $EE_x$ means: 0.81-0.86 | Pearson's r : 0.61-0.65<br>Spearman's rho: 0.48-0.55 |
| Odeck (2001) | Norway | 170 rock-blasting units, 1993 | input-reducing | CRTS | Cobb-Douglas | deterministic prod. frontier | $TE_x = TE_y$ mean: 0.36 | $TE_x = TE_y$ mean: 0.12 | Spearman's rho: 0.85 |
| | | | output-increasing | | | | | | |
| | | | input-reducing | VRTS | | | $TE_x$ mean: 0.47<br>$SE_x$ mean: 0.77 | $TE_x$ mean: 0.23<br>$SE_x$ mean: 0.60 | Spearman's rho: 0.78<br>Spearman's rho: 0.71 |
| | | | output-increasing | | | | $TE_y$ mean: 0.44<br>$SE_y$ mean: 0.85 | $TE_y$ mean: 0.19<br>$SE_y$ mean: 0.76 | Spearman's rho: 0.81<br>Spearman's rho: 0.66 |
| Ondrich & Ruggiero (2001) | N / A | simulation study: 200 & 1000 units | N / A | N / A | Cobb-Douglas | stochastic *vs* deterministic prod. frontier | "using the rank correlation criterion, the COLS deterministic model performs as well as the stochastic frontier model regardless of actual measurement error variance and sample size" (p. 441) | | |
| Park & Lesourd (2000) | South Korea | 64 power plants, 1990 | input-reducing | CRTS | translog | stochastic prod. frontier | $TE_x$ mean: 0.904 | $TE_y$ mean: 0.761 | Pearson's r : 0.636 |
| | | | | VRTS | | | $TE_x$ mean: 0.934 | | Pearson's r : 0.612 |
| Ray & Mukherjee (1995) | US | 123 electricity utilities, 1970 | cost-reducing | VRTS | hybrid | stochastic TC frontier | $EE_x$ means: 0.82-0.92 | $EE_x$ means: 0.88-0.97 | Spearman's rho: 0.21-0.55 |
| Read & Thanassoulis (1996) | N / A | simulation study: 500 units | output-increasing | CRTS | Cobb-Douglas | stochastic prod. frontier | "…when the SF function does specify the underlying technology closely, SF estimates are… much better than DEA" (p. 28). | | |

### Table 1 (cont.) - Direct DEA *vs* SFA/DFA Comparisons

| Author / Date | Country | Units / Period | DEA Model | | SFA/DFA Model | | Findings | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Orientation | RTS | Form | Frontier | DEA | SFA/DFA | Compare |
| Reinhard, Lovell & Thijssen (2000) | Holland | 613 dairy farms, 1991- 1994 | input-reducing | VRTS | restricted translog | stochastic prod. frontier | $TE_x$ means: 0.81-0.82<br>$EnvE_x$ means: 0.51-0.53 | $TE_x$ means: 0.90<br>$EnvE_x$ means: 0.79-0.8 | Spearman's rho: 0.70<br>Spearman's rho: 0.49 |
| | | | output-increasing | | | | $TE_y$ means: 0.78-0.79 | $TE_y$ means: 0.89 | Spearman's rho: 0.76 |
| Resti (2000) | N / A | simulation study: 50 & 500 'banks' | cost-reducing | CRTS<br>VRTS | restricted translog | stochastic TC frontier | "…different techniques do not lead to dramatically different results when used in the same methodological framework (as far as selection of variables and of the relevant concept of efficiency are concerned)" (p. 574). | | |
| Resti (1997) | Italy | 270 banks, 1988 - 1992 | cost-reducing | CRTS | restricted translog | stochastic TC frontier | $EE_x$ means: 0.66-0.69 | $EE_x$ means: 0.69-0.70 | Pearson's r: 0.87<br>Spearman's rho: 0.89 |
| | | | | VRTS | | | $EE_x$ means: 0.73-0.76 | | Pearson's r: 0.71<br>Spearman's rho: 0.73 |
| Ruggiero (1999) | N / A | simulation study: 25 - 200 units | N / A | N / A | translog & Cobb-Douglas | stochastic *vs* deterministic prod. frontier | "…the parametric deterministic model outperformed the stochastic frontier model in nearly all of the model situations considered. …In addition, the deterministic frontier approach was more consistent" (p. 562). | | |
| Ruggiero & Vitaliano (1999) | US | 520 school districts, 1990 / 91 | cost-reducing | VRTS | Cobb-Douglas | stochastic TC frontier | $EE_x$ mean: 0.875 | $EE_x$ mean: 0.860 | Spearman's rho: 0.86 |
| Scarsi (1999) | Italy | 76 electricity firms, 1994 - 1996 | output-increasing | VRTS | translog | stochastic prod. frontier | $TE_y$ means: 0.60-0.71 | $TE_y$ mean: 0.62 | Paired t: -0.65-6.41<br>$p \leq 0.518$ |
| Sharma, Leung & Zaleski (1997) | US | 53 pig farms, 1994 | output-increasing | CRTS | Cobb-Douglas | stochastic prod. frontier | $TE_y$ mean: 0.644 | $TE_y$ mean: 0.749 | Spearman's rho: 0.883 |
| | | | | VRTS | | | $TE_y$ mean: 0.726 | | Spearman's rho: 0.745 |
| Sickles & Streitwieser (1992) | US | 14 natural gas firms, 1977 - 1985 | input-reducing | CRTS | translog | stochastic prod. frontier | $TE_x = TE_y$ means: 0.78-0.86 | $TE_y$ means: 0.70-0.78 | Pearson's r: 0.53-0.54<br>Spearman's rho: 0.46-0.61 |
| Singh, Coelli & Fleming (2000) | India | 23 dairy plants, 1992/93 - 1996/97 | cost-reducing | VRTS | Cobb-Douglas | stochastic prod. frontier | $TE_x$ mean: 0.91<br>$AE_x$ mean: 0.68<br>$EE_x$ mean: 0.62 | $TE_x$ mean: 0.89<br>$AE_x$ mean: 0.91<br>$EE_x$ mean: 0.80 | "…choice of method can have a significant impact on results" (p.25). |
| Souza, Alves & Avila (1999) | Brazil & Argentina | 34 research units, 1996 | output-increasing | CRTS | Cobb-Douglas | stochastic prod. frontier | $TE_x = TE_y$ mean: 0.65 | $TE_y$ means: 0.67-0.68 | Pearson's r > 0.91<br>Spearman's rho > 0.90 |

## Table 1 (cont.) - Direct DEA *vs* SFA/DFA Comparisons

| Author / Date | Country | Units / Period | DEA Model | | SFA/DFA Model | | Findings | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | **Orientation** | **RTS** | **Form** | **Frontier** | **DEA** | **SFA/DFA** | **Compare** |
| Stone (2000) | Spain | 21 High Courts, 1991 | input-reducing<br>cost-reducing | CRTS | Cobb-Douglas | stochastic TC frontier | $TE_x$ mean: 0.77<br>$EE_x$ mean: 0.66 | $EE_x$ means: 0.66-0.84 | Pearson's r: 0.85-0.90<br>Pearson's r: 0.82-0.89 |
| Uri (2001) | US | 19 telecom LECs 1988 - 1998 | input-reducing | CRTS<br>VRTS | Cobb-Douglas | stochastic prod. frontier | "…based on DEA results there was no identifiable improvement in technical efficiency over the 1988-1998 period … results from a stochastic frontier approach confirm no change in technical efficiency" (p. 844). | | |
| van den Broek, Førsund, Hjalmarsson & Meeusen (1980) | Sweden | 28 dairy plants, 1964 -1973 | N / A | N / A | Zellner-Revankar | deterministic prod. frontier<br>stochastic prod. frontier | N / A | $TE_y$ means: 0.46-0.80<br>$TE_y$ means: 0.79-0.90 | SFA $TE_y$ > DFA $TE_y$ |
| Wadud & White (2000) | Bangladesh | 150 rice farms, 1997 | output-increasing | CRTS<br>VRTS | translog | stochastic prod. frontier | $TE_y$ means: 0.789<br>$TE_y$ means: 0.858 | $TE_y$ means: 0.791 | Spearman's rho: 0.777<br>Spearman's rho: 0.747 |
| Webster, Kennedy & Johnson (1998) | Australia | 301 private hospitals, 1991/92 - 1994/5 | input-reducing | VRTS | translog & Cobb-Douglas | stochastic prod. frontier | $TE_x$ means: 0.39-0.90 | $TE_y$ means: 0.71-0.79 | Pearson's r: 0.29-0.79<br>Spearman's rho: 0.32-0.80 |
| Whiteman (2000) | worldwide | 41 electricity, 51 natural gas & 31 telecom firms, 1996 & 1996/97 | input-reducing | CRTS<br>VRTS<br>CRTS<br>VRTS | Cobb-Douglas | deterministic dist*ƒ*s & ray prod. frontier<br>stochastic dist*ƒ*s & ray prod. frontier | "…the existence of stochastic error appears to have little impact on estimates of TE. Also there appear to be major problems in estimating multiple output stochastic frontier models which make the parametric deterministic models or DEA our preferred methodologies" (pp. 10 & 11). | | |
| Yin (2000) | worldwide | 102 softwood pulp mills, 1996 | cost-reducing | VRTS | Cobb-Douglas<br>translog | stochastic TC frontier | $EE_x$ means: 0.857<br>$EE_x$ range: 0.681-1.00 | $EE_x$ means: 0.922<br>$EE_x$ range: 0.751-0.976<br>$EE_x$ means: 0.951<br>$EE_x$ range: 0.801-0.992 | Spearman's rho: 0.7<br>Spearman's rho: 0.5 |
| Yu (1998) | N / A | simulation study: 250 units | output-increasing | CRTS<br>VRTS | translog | stochastic prod. frontier | "…the one-step stochastic frontier method has a dominant advantage over other methods in dealing with exogenous variables…as long as exogen-ous variables are correctly identified and accounted for" (pp. 569 & 579). | | |

## Conclusions & Directions for Future Research

This paper reviews empirical results drawn from published simulation studies with the aim of highlighting the pros and cons of competing methods for frontier estimation and efficiency measurement. In summary, the conclusions drawn by Resti (2000) seem appropriate:

> "…none of these articles demonstrates that either DEA or econometric models have an absolute advantage over their competitors. Nevertheless, …they succeed in indicating a range of specific situations (depending for example on the number of units in the sample, or on the amount of inefficiency and noise in the data) where some estimation technique proves superior" (Resti, 2000 p. 560).

In other words, results drawn from the simulation studies reviewed above confirm, clarify and/or contradict expected trade-offs, so that the choice of a 'correct' method is rather more clear cut in certain specific situations. The problem is that these specific situations (defined over key features of the simulated technology and production environment such as dimensionality of input/output space, economies of scale and scope, heterogeneity of production units, the extent of measurement error and random shocks, et cetera) are usually much simplified and occur relatively infrequently in the real world.

Our review of some 41 real-world DEA *vs* SFA/DFA comparisons suggests that calls for the parallel application of competing methods (to cross check results) have already been heeded. Unfortunately, real-world comparisons aren't much help either unless actual efficiency scores (like those underlying simulated data) are available for comparison purposes. It seems that the best of both worlds (realism *and* a criterion measure) is required in determining the relative precision (and policy value) of DEA, DFA and SFA. Both Yu (1998) and Resti (2000) have already made some progress in this regard by making their simulated data "more realistic and closer to the characteristics of existing industries" (Resti, 2000 p. 560). To provide further clarification as to the relative precision of competing methods, the now routine practice of cross checking should be taken one step further to include realistic simulation studies along-side real-world DEA *vs* SFA/DFA comparisons.

## References


Banker, R.D., Chang, H. and Cooper, W. (1996) Simulation studies of efficiency, returns to scale and mis-specification with non-linear functions in DEA. *Annals of Operations Research* 66, 233-254.

Banker, R.D., Charnes, A., Cooper, W. and Maindiratta, A. (1988) A comparison of DEA and translog estimates of production frontiers using simulated observations from a known technology. In: Dogramaci, A. and Färe, R., (Eds.) *Applications of Modern Production Theory: Efficiency and Productivity,* Massachusetts: Norwell, 33-55.

Banker, R.J. and Morey, R.C. (1986) Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research* 34, 513-521.

Banker, R.D., Conrad, R.F. and Strauss, R.P. (1986) A Comparative Application of Data Envelopment Analysis and Translog Methods: An Illustrative Study of Hospital Production. *Management Science* 32, 30-44.

Banker, R.D., Gadh, V.M. and Gorr, W.L. (1993) A Monte Carlo comparison of two production frontier estimation methods: Corrected ordinary least squares and data envelopment analysis. *European Journal of Operational Research* 67, 332-343.

Bates, J.M. (1997) Measuring Predetermined Socioeconomic 'inputs' When Assessing the Efficiency of Educational Outputs. *Applied Economics* 29, 85-93.

Bates, J.M., Baines, D. and Whynes, D.K. (1996) Measuring the efficiency of prescribing by general practitioners. *Journal of the Operational Research Society* 47, 1443-1451.

Bauer, P.W., Berger, A., Ferrier, G. and Humphrey, D. (1998) Consistency conditions for regulatory analysis of financial institutions: A comparison of frontier efficiency methods. *Journal of Economics and Business* 50, 85-114.

Bjurek, H., Hjalmarsson, L. and Forsund, F.R. (1990) Deterministic Parametric and Nonparametric Estimation of Efficiency in Service Production: A Comparison. *Journal of Econometrics* 46, 213-227.

Bojanic, A.N., Caudill, S.B. and Ford, J.M. (1998) Small-Sample Properties of ML, COLS, and DEA Estimators of Frontier Models in the Presence of Heteroscedasticity. *European Journal of Operational Research* 108, 140-148.

Brummer, B. (2001) Estimating Confidence Intervals for Technical Efficiency: the Case of Private Farms in Slovenia. *European Review of Agricultural Economics* 28, 285-306.

Chirikos, T.N. and Sear, A.M. (2000) Measuring hospital efficiency: A comparison of two approaches. *Health Services Research* 34, 1389-1408.

Coelli, T. and Perelman, S. (1999) A comparison of parametric and non-parametric distance functions: With application to European railways. *European Journal of Operational Research* 117, 326-339.

Cooper, W., Kumbhakar, S.C., Thrall, R.M. and Yu, X. (1995) DEA and stochastic frontier analyses of the 1978 Chinese economic reforms. *Socio-Economic Planning Sciences* 29, 85-112.

Cummins, J.D. and Zi, H. (1998) Comparison of Frontier Efficiency Methods: An Application to the U.S. Life Insurance Industry. *Journal of Productivity Analysis* 10, 131-52.

De Borger, B. and Kerstens, K. (1996) Cost Efficiency of Belgian Local Governments: A Comparative Analysis of FDH, DEA, and Econometric Approaches. *Regional Science and Urban Economics* 26, 145-70.

Dismukea, C. and Senab, V. (1999) Has DRG payment influenced the technical efficiency and productivity of diagnostic technologies in Portugese public hospitals? An empirical analysis using parametric and non-parametric methods. *Health Care Management Science* 2, 107-116.

Drake, L. and Weyman Jones, T.G. (1996) Productive and Allocative Inefficiencies in U.K. Building Societies: A Comparison of Non-parametric and Stochastic Frontier Techniques. *Manchester School of Economic and Social Studies* 64, 22-37.

Eisenbeis, R., Ferrier, G. and Kwan, S. (1999) The informativeness of stochastic frontier and programming frontier efficiency scores: Cost efficiency and other measures of bank holding company performance. *Federal Reserve Bank of Atlanta Working Paper #99-23.*

Ferrier, G.D. and Lovell, C.A.K. (1990) Measuring cost efficiency in banking: Econometric and linear programming evidence. *Journal of Econometrics* 46, 229-245.

Giuffrida, A. and Gravelle, H. (1999) Measuring performance in primary care: Econometric analysis and DEA. *Discussion Paper #1999/36, Department of Economics, University of York.*

Gong, B.H. and Sickles, R.C. (1992) Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data. *Journal of Econometrics* 51, 259-84.

Hjalmarsson, L., Kumbhakar, S.C. and Heshmati, A. (1996) DEA, DFA and SFA: A Comparison. *Journal of Productivity Analysis* 7, 303-27.

Jacobs, R. (2000) Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. *University of York: Centre for Health Economics (CHE) Discussion Paper #177.*

Jacobs, R. (2001) Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. *Health Care Management Science* 4, 103-15.

Johnes, G. (1998) The costs of multi-product organisations and the heuristic evaluation of industrial structure. *Socio-Economic Planning Sciences* 32, 199-209.

Linna, M. (1998) Measuring hospital cost efficiency with panel data models. *Health Economics* 7, 415-27.

Linna, M. and Häkkinen, U. (1998) A Comparative Application of Econometric Frontier and DEA Methods for Assessing Cost Efficiency of Finnish Hospitals. In: Zweifel, P. (Ed.) *Health, the Medical Profession and Regulation,* Boston; Dordrecht and London: Kluwer Academic.

Linna, M. and Häkkinen, U. (1999) Determinants of cost efficiency of Finnish hospitals: A comparison of DEA and SFA. *Helsinki University of Technology: Systems Analysis Laboratory Research Report #A78.*

Lovell, C.A.K. (1993) Production frontiers and productive efficiency. In: Fried, H.O., Lovell, C.A.K. and Schmidt, S.S., (Eds) *The Measurement of Productive Efficiency: Techniques and Applications*, New York & Oxford: Oxford University Press, 3-67.

Meibodi Ali, E. (1998) Efficiency considerations in the electricity supply industry: the case of Iran. *Surrey Energy Economics Centre: SEEDS Discussion Paper #95.*

Mortimer, D. (2001) Methods for the Measurement of Hospital Efficiency: A Comparison of Frontier Estimation Techniques in a Sample of Victorian Public Hospitals. *Department of Economics, Monash University: Unpublished Master of Economics (Honours) thesis.*

Odeck, J. (2001) Comparison of Data Envelopment Analysis and Deterministic Parametric Frontier Approaches: an Application in the Norwegian Road Construction Sector. *Transportation Planning and Technology* 24, 111-134.

Ondrich, J. and Ruggiero, J. (2001) Efficiency Measurement in the Stochastic Frontier Model. *European Journal of Operational Research* 129, 434-442.

Park, S.-U. and Lesourd, J.-B. (2000) The Efficiency of Conventional Fuel Power Plants in South Korea: A Comparison of Parametric and Non-parametric Approaches. *International Journal of Production Economics* 63, 59-67.

Ray, S.C. and Mukherjee, K. (1995) Comparing Parametric and Non-parametric Measures of Efficiency: A Reexamination of the Christensen-Greene Data. *Journal of Quantitative Economics* 11, 155-68.

Read, L.E. and Thanassoulis, E. (1996) A comparison of data envelopment analysis (DEA) and stochastic frontiers (SF) under variation of fit. *Warwick Business School: Research Bureau Research Paper #234*.

Reinhard, S., Lovell, K. and Thijssen, G. (2000) Environmental efficiency with multiple environmentally detrimental variables: estimated with SFA and DEA. *European Journal of Operational Research* 121, 287-303.

Resti, A. (1997) Evaluating the cost-efficiency of the Italian banking system: What can be learned from the joint application of parametric and non-parametric techniques. *Journal of Banking & Finance* 21, 221-250.

Resti, A. (2000) Efficiency measurement for multi-product industries: A comparison of recent techniques based on simulated data. *European Journal of Operational Research* 121, 559-578.

Ruggerio, J. (1999) Efficiency estimation and error decomposition in the stochastic frontier model: A Monte Carlo analysis. *European Journal of Operational Research* 115, 555-563.

Ruggiero, J. and Vitaliano, D.F. (1999) Assessing the efficiency of public schools using data envelopment analysis and frontier regression. *Contemporary Economic Policy* 17, 321-331.

Scarsi, G.C. (1999) Local electricity distribution in Italy: Comparative efficiency analysis and methological cross-checking. *Fondazione Eni Enrico Mattei (FEEM) Nota Di Lavoro #16/99.*

Seiford, L.M. and Thrall, R.M. (1990) Recent Developments in DEA: The Mathematical Programming Approach to Frontier Analysis. *Journal of Econometrics* 46, 7-38.

Sengupta, J.K. (1987) Data Envelopment Analysis for Efficiency Measurement in the Stochastic Case. *Computers & Operations Research* 14, 117-129.

Sengupta, J.K. (1998) Stochastic Data Envelopment Analysis: A New Approach. *Applied Economics Letters* 5, 287-90.

Sharma, K.R., Leung, P. and Zaleski, H.M. (1997) Productive Efficiency of the Swine Industry in Hawaii: Stochastic Frontier vs. Data Envelopment Analysis. *Journal of Productivity Analysis* 8, 447-59.

Sickles, R.C. and Streitwieser, M.L. (1992) Technical Inefficiency and Productive Decline in the U.S. Interstate Natural Gas Pipeline Industry under the Natural Gas Policy Act. *Journal of Productivity Analysis* 3, 119-33.

Singh, S., Coelli, T. and Fleming, E. (2000) Performance of dairy plants in the cooperative and private sectors in India. *University of New England: Centre for Efficiency and Productivity Analysis (CEPA) Working Paper #2/2000.*

Stone, M. (2000) Questions of efficiency estimation in public services: Is the state of the art really 'state of the art'? *University College London: Department of Statistical Science Research Report #208.*

Uri, N.D. (2001) The Effect of Incentive Regulation on Productive Efficiency in Telecommunications. *Journal of Policy Modeling* 23, 825-846.

van den Broek, J., Førsund, F.R., Hjalmasson, L. and Meeusen, W. (1980) On the Estimation of Deterministic and Stochastic Frontier Production Functions: A Comparison. *Journal of Econometrics* 13, 117-38.

Wahud, A. and White, B. (2000) Farm household efficiency in Bangladesh: A comparison of stochastic frontier and DEA methods. *Applied Economics* 32, 1665-1673.

Webster, R., Kennedy, S. and Johnson, L. (1998) Comparing techniques for measuring the efficiency and productivity of Australian Private Hospitals. *Australian Bureau of Statistics Working Paper #98/3.*

Whiteman, J. (1999) The measurement of efficiency where there are multiple outputs. *Monash University: Centre for Policy Studies, General Paper #134*.

Yin, R. (2000) Alternative measurements of productive efficiency in the global bleached software sector. *Forest Science* 46, 558-569.

Yu, C. (1998) The effects of exogenous variables in efficiency measurement: A Monte Carlo study. *European Journal of Operational Research* 105, 569-580.

Zuckerman, S., Hadley, J. and Iezzoni, L. (1994) Measuring Hospital Efficiency with Frontier Cost Functions. *Journal of Health Economics* 13, 255-80.