Robustness of Data Envelopment Analysis to Omitted and Irrelevant Production Inputs

Don U. A. Galagedera* and Param Silvapulle

Department of Econometrics and Business Statistics
Monash University
PO Box 197, Caulfield East, Victoria 3145, Australia

Abstract

Data envelopment analysis (DEA) has been widely used for measuring relative performance of universities and schools in the education sector; banks, mutual funds and superannuation funds in the financial sector; and the like. There has been an on going controversial debate about the use of DEA as there is no diagnostic checks, which are in prevalent use in econometrics, for model misspecification generally and for correct input and output variable selection in particular. This paper contributes to this debate by investigating the sensitivity of DEA efficiency estimates to including inappropriate variables and unwittingly omitting some important variables in the DEA model in large samples. Data series is simulated from the Cobb-Douglas production process under different returns-to scale (RS) assumptions. The overall results indicate that omission of relevant inputs and inclusion of irrelevant inputs affect the DEA efficiency estimates with the former affecting them adversely, depending on the RS assumptions. The DEA with irrelevant inputs performs better under the correct RS assumption. Further, under constant RS specification, the DEA with irrelevant inputs tend to over estimate the efficiencies in almost all production units studied in this paper. A similar observation has been made in the presence of irrelevant inputs when the underlying process is decreasing RS but unwittingly assumed to be variable RS.

JEL Classification: C44, C61, C67

Key words: Relative efficiency; non-parametric model; data envelopment analysis; Monte Carlo simulation

*Author for correspondence
Telephone: +61 (03) 9903 1578
E-mail: Tissa.Galagedera@buseco.monash.edu.au

Introduction

Since the seminal paper by Charnes *et al.* (1978), the data envelopment analysis (DEA) methodology – a nonparametric mathematical programming technique for efficiency evaluation, has been given considerable attention during the past two decades. In the survey article by Seiford (1996), popularisation of DEA methodology, its rapid growth and widespread use across many disciplines, health, education, finance and the like are very clear.

Being a non-parametric technique, DEA does not require a structural form for the production frontier, avoiding the danger of miss-specification of the frontier. DEA uses a linear programming technique to determine the level of efficiency of individual production units also referred to as decision-making units (DMUs). An attractive property of the DEA approach is its ability to handle multiple outputs quite easily, while all the parametric approaches have been limited to only a single output.

It has been argued in the literature that, for DEA to be effective, data should be free from statistical noise. If present, applying DEA to estimate the technical efficiency at production unit level may result in compounding inefficiency and statistical noise. See, for example, Charnes et al. 1985; Charnes and Neralie`, 1990; Charnes, Haag, Jaska and Semple, 1992; Charnes, Rousseau and Semple, 1996; Zhu, 1996; Seiford and Zhu, 1998a and Seiford and Zhu, (1998b) for contributing to this view and examining the robustness of DEA models to input data with statistical noise, prompting several studies on sensitivity analysis of DEA efficiency estimates to various issues.

The primary objective of this paper is to investigate the effects of omitting relevant variables from and including irrelevant variables in the model on DEA relative efficiency estimates in a large sample under the three model specification assumptions. Simulating data from a simple production process with increasing returns to scale (IRS), decreasing returns to scale (DRS) or constant returns to scale (CRS) specifications, we examine the effects of multiple input variable omission and multiple input variable inclusion on the DEA efficiencies.

Numerous studies have examined the relative strengths of stochastic frontier models and DEA in estimating the firm-specific technical efficiency. Gong and Sickles (1992) have estimated firm-specific technical efficiencies using panel data and reported that as the misspecification of the functional form becomes more serious and the degree of correlatedness of inefficiency with input variables increases, DEA's appeal becomes more compelling. Banker, Gadh and Gorr (1993) have used a simulated two-input one-output production technology and compared the accuracy of the results of corrected ordinary least squares method and DEA. They have included measurement errors in the model and reported that DEA provides accurate results for small sample sizes and low measurement errors compared with those obtained by corrected OLS method. Read and Thanassoulis (1996), on the other hand, have compared the DEA and stochastic frontier methods in assessing firm-specific technical efficiency when the underlying technology is poorly estimated for certain input-output mixes by the stochastic frontier function. They have reported that even when the classical assumptions are met, there are some situations in which stochastic frontier performs worse than DEA for low measurement errors.

The sensitivity of DEA efficiency estimates to input variable selection and model specification has been another concern for practitioners. Pedraja-Chaparro, Salinas-Jim`enez and Smith (1999) have shown that DEA relative efficiency is influenced by the distribution of true efficiencies, the sample size, the number of inputs and outputs included in the analysis and the degree of correlation between inputs**.** In an empirical study of measuring the relative efficiency of mutual funds using DEA, Galagedera and Silvapulle (2000) have investigated the sensitivity of relative efficiency to various input-output variable combinations, and found the results vary considerably. Although the model misspecification is potentially a serious issue in DEA, it has received very little attention in the literature until recently. The potential for model misspecification in DEA is huge, and it can be attributed to the absence of a credible model-building methodology. Recognising this shortcoming in DEA, some studies have considered alternative specifications; see, for example, Valdmanis (1992).

The principal causes of model misspecification in DEA are known to be omission of relevant variables, inclusion of irrelevant variables and incorrect assumption on returns-to-scale (RS**)** and the sample size, resulting in misleading conclusions. When important variables are omitted the results may be far from reality. On the other hand, increasing the number of variables decreases the ability of the model to differentiate individual production units in terms of efficiency. Smith (1997) has examined the implication for DEA efficiency scores of using a misspecified single-output constant RS model. In particular, Smith (1997) study investigated the robustness of DEA results to omission and inclusion of only a single input using samples of sizes ranging from 10 to 80 and concluded that the dangers of misspecification are most serious when simple models are used and sample sizes are small. Extending this study, Ruggiero (1998a) has investigated the impact of inclusion of inappropriate variables on technical efficiency measurement in a variable RS DEA model with multiple outputs via a simulation study. The Ruggiero (1998a) has used a sample of size 300 and documented that DEA performs well in the presence of additional inputs even though they are production irrelevant.

The paper is organised as follows. The next section provides a brief description of the DEA methodology. The design of the experiment and data generating process are discussed in section 3. The results of the DEA performance are analysed in section 4. Section 5 summarises the findings and gives some directions for future work.

Data envelopment analysis

The DEA methodology is a mathematical programming technique introduced by Charnes, Cooper and Rhodes (1978) to evaluate the relative efficiency of DMUs with no specific functional form of the production frontier and can accommodate multiple inputs and outputs. A DEA run is a series of mathematical programming optimisations applied to an observed data series generated from the past behaviour of each DMU in order to estimate the relative efficiency of DMUs that have common inputs and outputs. Thus, when a 'true' or a theoretical model of efficiency cannot be established, DEA provides a methodology to determine the relative efficiencies based on the known levels of attainment.

Speaking broadly, the DEA technique defines an efficiency measure of a DMU by its position relative to the frontier of the best performance established mathematically by the ratio of weighted sum of outputs to weighted sum of inputs. The estimated frontier of best performance characterises the efficiency of DMUs and identifies inefficiencies. There are several basic DEA models. Essentially each of them establishes subsets of a given set of DMUs that determine parts of an envelopment surface. The specific DEA model used prescribes the geometry of this envelopment surface.

A DEA model can be analysed in two ways, an input orientation and an output orientation. That is, the DEA model may focus on either input reduction or output augmentation to achieve efficiency. An input orientation provides information as to how much proportional reduction of inputs are necessary while maintaining current the levels of outputs for an inefficient DMU to become DEA efficient. An output orientation analysis provides information on how much augmentation to the levels of outputs of an inefficient DMU is necessary while maintaining current input levels for it to become DEA-efficient. Regardless of whether an input or output orientation is used, a DEA efficient DMU will always have 100 percent efficiency.

We use an input orientation version of the DEA model developed by Banker, Charnes and Cooper (1984), denoted BCC hereafter. This model is an extension of the original formulation of the DEA model introduced by Charnes, Cooper and Rhodes (1978). Let us first define the following measures:

y_{rj} = known positive output level of DMU *j*, *r*=1,2,…,*s* where, *s* is the number of outputs

x_{ij} = known positive input level of DMU *j*, *i*=1,2,…,*m* where *m* is the number of inputs

n = total number of DMUs.

An interpretation of the BCC model that estimates the relative efficiency score, θ, of DMU '0' is given by

$$\text{Min } \theta \tag{1}$$

$$\text{subject to} \quad \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{r0}, \qquad r = 1,2,\ldots,s, \tag{2}$$

$$\theta x_{i0} \geq \sum_{j=1}^{n} \lambda_j x_{ij}, \qquad i = 1,2,\ldots,m, \tag{3}$$

$$\sum_{j=1}^{n} \lambda_j = 1, \tag{4}$$

$$\lambda_j \geq 0, \qquad j = 1,2,\ldots,n. \tag{5}$$

The variables in the BCC model are unrestricted θ and λ_j which is non-negative. The variable θ, as evident in constraint (3), is the proportional reduction in all inputs of the designated DMU required to achieve efficiency. The constraints in the model ensure that relative efficiency of the designated DMU never exceeds 1. The sufficient condition for efficiency of the designated DMU is that the optimum value of θ is 1. Otherwise, it is labelled as inefficient compared to the other DMUs in the sample. A DEA run involves solving the above model *n* times; once for each DMU analysed.

The constraint given in (4) is referred to as the convexity constraint and accounts for VRS. When the convexity constraint is removed the resulting model represents the CRS situation. The relative efficiency score obtained for a designated DMU under CRS is a measure of its overall technical efficiency and is always at least as much as the corresponding value obtained under VRS. The relative efficiency score obtained under VRS is a measure of pure technical efficiency. The difference in overall and pure technical efficiencies is attributed to scale efficiency. A measure of scale efficiency is simply the ratio of overall and pure technical efficiencies.

Experimental design and data generation

The production process we have considered is the linearly homogeneous Cobb-Douglas function given by

$$y_{1j} = \beta \prod_{i=1}^{m} x_{ij}^{\alpha_i} \qquad j = 1,2,...,n \tag{6}$$

where, α_i and β are assumed to be known parameters. The assumption of the Cobb-Douglas form for the production function is very common. See, for example, Aigner and Chu (1968), Pedraja-Chaparro, Salinas-Jim`enez and Smith (1999), Ruggiero (1999) and Banker, Gadh and Gorr (1993). With an efficiency term denoted by $\gamma_j \in [0,1]$, which is the technical efficiency of the j th production unit, (6) becomes

$$y_{1j} = \beta \prod_{i=1}^{m} x_{ji}^{\alpha_i} \gamma_j . \qquad j = 1,2,...,n \tag{7}$$

Taking log of (7), an additive representation is given as

$$\ln y_{1j} = \ln \beta + \sum_{i=1}^{m} \alpha_i \ln x_{ij} - u_j \qquad j = 1,2,...,n \tag{8}$$

where $u_j = -\ln \gamma_j$. The variable $u_j > 0$ denotes the efficiency component. The efficiency component is expected to represent the shortfall of output from the production frontier.

Researchers have assumed several distributions for *u*, these being exponential (Banker, Gadh and Gorr, 1993; Meeusen and Van Den Broeck, 1977), half-normal (Jondrow, Lovell, Materov and Schmidt, 1982; Aigner, Lovell and Schmidt, 1977), Gamma (Afriat, 1972; Richmond, 1974; Greene, 1980 and Stevenson, 1980) and truncated normal (Stevenson, 1980). In empirical applications, researchers do not generally have a priori knowledge to justify the shape of the one-sided distribution.

In this study, we set β =1 and assume that the true production process is explained by three inputs only. Our choice of three inputs is to circumvent the difficulties encountered by Smith (1997): a simulation study involving a Cobb-Douglas production function has shown that the DEA performance indicating true efficiencies diminishes with increasing number of inputs in the production process. We consider three sets of parameter values in the Cobb-Douglas production process, obtaining the following production functions and representing three different RS specifications.

$$\text{IRS Model:} \quad y_{1j} = x_{1j} x_{2j}^{1/2} x_{3j}^{1/2} \tag{9}$$

$$\text{DRS Model:} \quad y_{1j} = x_{1j}^{1/8} x_{2j}^{1/4} x_{3j}^{1/4} \tag{10}$$

$$\text{CRS Model:} \quad y_{1j} = x_{1j}^{1/3} x_{2j}^{1/3} x_{3j}^{1/3} \tag{11}$$

Due to the characteristics of the production setting assumed in (9) and (11), a well-defined DEA model will always overestimate the efficiency (Smith, 1997), because DEA implicitly assumes that the production function is concave (Banker, 1993). The extent of the overestimate however appear to be highly dependent on the sample size with large samples being preferred. Therefore, input variables x_1, x_2 and x_3 and corresponding output variable were generated randomly for a sample of 200 DMUs according to the distributions specified in Table 1. To date, the DEA studies have simulated input-output variables from normal and/or uniform distributions. We opt for the uniform distribution and a one-sided distribution, which is half-normal with mean zero and standard deviation 0.3.

Table 1. Input-output variables

Variable	Type	Production relevance	Distribution
x_1	Input	Relevant	Uniform [1,2]
x_2	Input	Relevant	Uniform [30,50]
x_3	Input	Relevant	Uniform [30,50]
x_4	Input	Irrelevant	Uniform [1,2]
x_5	Input	Irrelevant	Uniform [1,2]
x_6	Input	Irrelevant	Uniform [40,60]
y_1	Output	Relevant	Not applicable
U	Efficiency component	Not applicable	$\lvert N(0,0.3)\rvert$

Having generated the efficiency component *u*, the true efficiency, γ, was computed as the exponent of the negative of the absolute value of *u* with restricting the efficiencies of the first twenty-five DMUs equal to one. Thus 29 of the 200 DMUs (14.5%) have an efficiency of at least 0.99. The mean and standard deviation of the generated true efficiency distribution is 0.829 and 0.142 respectively. In addition to the three production relevant input variables, three other production irrelevant input variables were also generated. The Pearson correlation coefficients between the generated input variables (see Table 1) are given in Table 2. None of the generated input variables are correlated with the true efficiency variable, γ. The pair-wise correlations between input variables are within –0.2 and 0.2.

Table 2. Correlations of input variables and true efficiency

Input variable	x_1	x_2	x_3	x_4	x_5	γ
x_1	1					0.011
x_2	0.119	1				-0.042
x_3	-0.078	-0.006	1			0.070
x_4	-0.083	0.120	0.167	1		-0.052
x_5	0.111	0.036	0.109	0.077	1	0.092
x_6	0.051	-0.050	0.089	0.001	-0.132	-0.081

To enhance the inputs by the same proportion we divide the randomly generated inputs (efficient inputs) by the true efficiency**.** Farrell (1957) developed indices of technical efficiency measured as the maximum equiproportional reduction in all inputs consistent with the equivalent production of observed output. The observed output, y_1, for each of IRS, DRS and CRS models were obtained by substituting the efficient inputs in equations (9)-(11) with appropriate set of known parameters. Therefore, considering the enhanced inputs ($x_i / \gamma : i = 1,2,3$) as observed inputs and the computed y_1 as the observed output in the DEA model given by (1)-(5), we obtain the estimates for γ, what is known as the Farrell efficiency.

We investigate the effect of omission of relevant inputs on DEA-efficiency by comparing the true efficiency with efficiencies obtained in the DEA models with at least one relevant input missing. In our experiment, this leads to six different DEA models. The DEA models we considered for comparison of the effects of irrelevant variables on efficiency measurement are: two models with a single irrelevant input; one model with two irrelevant inputs; and another with all three irrelevant inputs; see Table 3 for more details. The model 7 shown in the shaded row in Table 3 includes only the true set of input variables.

Table 3. Input variables considered in DEA models

DEA Model	Input variables	No of omitted variables	No of irrelevant variables
1	x_1	2	0
2	x_2	2	0
3	x_3	2	0
4	x_1, x_2	1	0
5	x_1, x_3	1	0
6	x_2, x_3	1	0
7	x_1, x_2, x_3	0	0
8	x_1, x_2, x_3, x_4	0	1
9	x_1, x_2, x_3, x_5	0	1
10	x_1, x_2, x_3, x_4, x_5	0	2
11	x_1, x_2, x_3, x_4, $x_5 x_6$	0	3

Results and analysis

First, we examine the mean and the standard deviation (SD) of the relative efficiencies of the eleven DEA models. As expected, the best DEA-result (the least absolute difference between the mean DEA-efficiency and the true mean efficiency) is obtained when the DEA model uses the true set of variables together with the correct RS specification; see Table 4. The omission of relevant variables leads to underestimating the mean efficiency, while the inclusion of irrelevant variables leads to over estimation. In both cases however the mean efficiency worsens as the number of variables omitted or included in the model increase with effect of variable omission being detrimental on the mean and SD of the relative efficiency.

Table 4. Mean and standard deviation of relative efficiency

Model	IRS Model				DRS Model				CRS Model			
	Mean		SD		Mean		SD		Mean		SD	
	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS
1	0.746	0.696	0.145	0.140	0.670	0.626	0.165	0.157	0.704	0.656	0.156	0.150
2	0.677	0.571	0.142	0.150	0.697	0.662	0.138	0.132	0.684	0.640	0.140	0.130
3	0.697	0.601	0.155	0.170	0.719	0.678	0.144	0.139	0.709	0.687	0.149	0.146
4	0.796	0.722	0.138	0.140	0.773	0.726	0.141	0.138	0.783	0.761	0.139	0.140
5	0.809	0.730	0.143	0.147	0.793	0.754	0.145	0.143	0.801	0.782	0.144	0.144
6	0.763	0.662	0.141	0.168	0.804	0.766	0.137	0.134	0.781	0.754	0.139	0.138
7	0.840	0.747	0.139	0.147	0.842	0.795	0.139	0.136	0.841	0.833	0.139	0.141
8	0.840	0.753	0.137	0.149	0.844	0.804	0.137	0.137	0.844	0.836	0.137	0.140
9	0.843	0.754	0.139	0.150	0.850	0.809	0.138	0.137	0.849	0.839	0.138	0.140
10	0.847	0.758	0.138	0.152	0.852	0.814	0.136	0.137	0.852	0.842	0.137	0.139
11	0.852	0.762	0.139	0.155	0.853	0.831	0.140	0.140	0.856	0.846	0.138	0.139
TRUE	0.829		0.142		0.829		0.142		0.829		0.142	

Second, in order to assess the mismatch between the true efficiency and the DEA efficiency of individual DMUs, several performance measurement criteria have been used. According to Ruggiero (1998b), the correlation coefficient; rank correlation coefficient and mean absolute deviation (MAD) between the true and the estimated efficiency; and the percentage of DMUs for which the estimated efficiency is less than the true efficiency can be used to compare the performance of different efficiency measurement methods. On the other hand, assessing the quality of the DEA models, Pedraja-Chaparro, Salinas-Jim`enez and Smith (1999) have suggested that the proportion of DMUs deemed efficient using DEA; the Spearman rank correlation coefficient between the DEA efficiency and true efficiency; the proportion of DMUs within the sample whose DEA efficiency is within a given percentage of their true efficiencies; and the percentage by which the DEA efficiency exceeds the true efficiency as four useful measures. We have adopted similar measures and two additional ones that measure the accuracy of classification of DMUs as either DEA-efficient or DEA-inefficient; these are the percentage of all DMUs correctly identified in terms of efficiency, and the percentage of efficient DMUs correctly identified as efficient. Our primary interest is in the effect of omission of variables, inclusion of irrelevant variables and incorrect RS specifications on the DEA relative efficiency. In what follows we examine these effects based on several performance measurement criteria.

Mean absolute deviation and mean deviation (MD) between true and DEA efficiency

The first set of measurement criteria we use is MAD and MD between the true and DEA efficiencies; these results are reported in Table 5. Based on the MAD criterion the best DEA-result (with the least MAD) is obtained under the correct model and RS specification. It is noticeable in Table 5 that presence of irrelevant variables in the IRS and DRS production processes tend to increase MAD under the correct RS assumption and decrease it under the incorrect RS assumption; these MAD tend to get higher with increasing number of irrelevant variables. In the CRS production process, however, increasing number of irrelevant variables leads to high MAD whether or not the correct RS assumption is made. Moreover, when irrelevant variables are present, the results of the analysis indicate that MAD between the true and DEA efficiency is lower in all three production processes under the correct assumption on RS compared with those under the incorrect RS assumption. These observations suggest that when using DEA to measure the relative performance, the production process for which the results that are adversely affected (in terms of MAD) in the presence of irrelevant variables and incorrect assumption on RS is the CRS specification. It is worthy of noting that our emphasis is not on the absolute values presented in the tables because to some extent they are dependent on the parameters of the underlying production process. Rather, we seek patterns that might emerge from them.

Table 5. Mean absolute deviation and mean deviation between true and DEA efficiency

Model	IRS Model				DRS Model				CRS Model			
	MAD between true eff and rel eff		Mean deviation between true eff and rel eff		MAD between true eff and rel eff		Mean deviation between true eff and rel eff		MAD between true eff and rel eff		Mean deviation between true eff and rel eff	
	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS
1	0.091	0.133	0.083	0.132	0.169	0.204	0.159	0.203	0.133	0.175	0.125	0.173
2	0.159	0.258	0.152	0.257	0.136	0.167	0.132	0.167	0.149	0.189	0.144	0.189
3	0.143	0.230	0.132	0.228	0.116	0.151	0.110	0.151	0.126	0.142	0.119	0.142
4	0.044	0.107	0.033	0.106	0.068	0.104	0.056	0.102	0.056	0.070	0.046	0.068
5	0.038	0.101	0.019	0.099	0.055	0.079	0.036	0.075	0.045	0.052	0.027	0.047
6	0.080	0.171	0.066	0.167	0.036	0.065	0.025	0.063	0.060	0.077	0.048	0.074
7	0.017	0.087	-0.007	0.082	0.013	0.041	-0.013	0.034	0.012	0.005	-0.012	-0.005
8	0.020	0.084	-0.012	0.076	0.016	0.037	-0.016	0.025	0.016	0.008	-0.016	-0.007
9	0.022	0.086	-0.014	0.075	0.021	0.035	-0.021	0.020	0.020	0.010	-0.020	-0.010
10	0.025	0.085	-0.018	0.071	0.023	0.034	-0.023	0.014	0.023	0.013	-0.023	-0.013
11	0.029	0.085	-0.023	0.067	0.024	0.027	-0.024	-0.003	0.027	0.018	-0.027	-0.018

The variation in MAD under variable omission is similar in all three models. As the number of omitted variables increase the MAD tends to increase whether or not the correct RS specification is used. An exception is that when relevant variables are omitted VRS appears to be a better option even when the true model is CRS.

The MDs in Table 5 suggest that when irrelevant variables are present DEA might over estimate the efficiency in a very large number of, if not all, DMUs in the DRS and the CRS Models; this was shown to be stronger through a detailed inspection of the results.

Percentage of efficient and inefficient DMUs identified correctly

We consider two measures only, the percentage of all DMUs correctly identified as efficient or inefficient, and the percentage of efficient DMUs correctly identified as efficient. The results of these measures are presented in Table 6. When variables are omitted the DEA model under VRS assumption seems to out perform those under CRS assumption with respect to the two measures considered in this study; this disparity appears to be profound when identifying the efficient DMUs. Thus, for the three production processes considered in the study VRS assumption appears to be a safer option when classifying DMUs either as efficient or inefficient in the DEA models with omitted variables. In the DEA models with irrelevant variables the same observation was made but only for the two production processes, IRS and DRS specifications. When classifying DMUs either as efficient or inefficient in the DEA with CRS production process, the correct RS assumption appears to yield superior results when irrelevant variables are present.

Table 6. Percentage of DMUs correctly identified

Model	IRS Model				DRS Model				CRS Model			
	% of all DMUs correctly identified		% of eff DMUs correctly identified		% of all DMUs correctly identified		% of eff DMUs correctly identified		% of all DMUs correctly identified		% of eff DMUs correctly identified	
	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS
1	88.5	86.5	27.6	6.9	87.5	85.5	20.7	3.4	88.0	85.5	24.1	3.4
2	87.5	86.0	20.7	3.4	88.5	86.5	24.1	6.9	88.0	86.5	20.7	6.9
3	86.5	85.5	13.8	3.4	87.5	86.0	17.2	3.4	87.0	86.0	17.2	3.4
4	93.0	87.5	62.1	13.8	93.0	87.5	60.1	17.2	93.5	90.0	62.1	34.5
5	90.0	86.5	41.4	10.3	90.0	87.5	37.9	10.3	90.0	86.5	41.4	10.3
6	87.5	86.0	27.6	6.9	89.5	88.0	34.5	17.2	88.5	87.0	31.0	13.8
7	94.0	87.5	75.9	17.2	97.5	89.0	100.0	27.6	98.0	99.0	100.0	100.0
8	94.0	88.5	79.3	27.6	97.5	91.0	100.0	41.4	97.5	99.0	100.0	100.0
9	89.5	87.0	75.9	24.1	94.5	88.5	100.0	34.5	94.5	96.0	100.0	100.0
10	89.5	86.5	79.3	27.6	94.5	90.0	100.0	44.8	94.0	96.0	100.0	100.0
11	88.5	85.5	86.2	34.5	90.5	87.5	100.0	58.6	92.0	95.5	100.0	100.0

Rank correlation between true and DEA efficiency

The higher the Spearman rank correlation between the true and DEA relative efficiency the higher is the DEA performance in measuring DMU efficiency. The results associated with this measure are reported in Table 7, revealing that when DEA model includes irrelevant variables the rank correlation always is higher under the correct RS assumption. However, the rank correlation tends to decrease with increasing number of irrelevant variables included in the DEA model; this is noted under both the VRS and CRS assumptions.

In the DEA models with omitted relevant variables, the two production processes, IRS and CRS specifications, result in higher rank correlations under the correct RS assumption. No clear pattern emerges in the results of the production process with DRS specification. Nevertheless, as the number of omitted variables in the DEA model increases a reduction in rank correlation is observed.

Table 7. Spearman rank correlation coefficients

Model	IRS Model		DRS Model		CRS Model	
	Spearman rank correlation coeff		Spearman rank correlation coeff		Spearman rank correlation coeff	
	VRS	CRS	VRS	CRS	VRS	CRS
1	0.850	0.847	0.637	0.703	0.738	0.769
2	0.725	0.597	0.806	0.845	0.760	0.828
3	0.619	0.525	0.722	0.769	0.670	0.731
4	0.917	0.846	0.860	0.863	0.892	0.917
5	0.919	0.847	0.865	0.869	0.896	0.915
6	0.761	0.582	0.924	0.920	0.839	0.864
7	0.955	0.836	0.967	0.947	0.970	0.996
8	0.950	0.827	0.965	0.942	0.967	0.994
9	0.934	0.814	0.949	0.938	0.951	0.991
10	0.925	0.806	0.948	0.936	0.948	0.988
11	0.919	0.792	0.949	0.931	0.941	0.983

Percentage of DMUs whose DEA efficiency is within 10 percent of true efficiency

The percentage of DMUs for which the DEA efficiency is within a margin of 10% of the true efficiency is considered as an indication of the extent to which the efficiency is either overstated or understated in the DEA. As seen from Table 8 entries, with relevant variables missing the percentage of DMUs whose DEA relative efficiency is within 10% of the true efficiency appear to be higher in the DEA models under the VRS assumption. This provides further support to the earlier observation that when relevant variables are absent it is safer to use the VRS specification. In the case where irrelevant variables are present the CRS and IRS models appear to perform better in terms of this measure under the correct RS assumption.

Our main conclusion based on the performance measures discussed in this section is that when estimating the DEA efficiency of individual DMUs the RS specification becomes important if the DEA model does not include all the variables deemed to be relevant in the analysis. A safer option then appears to be VRS. When the DEA model includes more variables than necessary the true RS specification seems to be crucial. Clearly the adverse impact of misspecification in DEA on individual DMU efficiency is more serious when the relevant variables are omitted from the DEA model compared to the inclusion of irrelevant ones.

Table 8. Percentage of DMUs whose DEA efficiency is within 10% of true efficiency

Model	IRS Model % of DMUs whose abs diff in rel eff with true eff < 10% of true eff		DRS Model % of DMUs whose abs diff in rel eff with true eff < 10% of true eff		CRS Model % of DMUs whose abs diff in rel eff with true eff < 10% of true eff	
	VRS	CRS	VRS	CRS	VRS	CRS
1	48.0	27.5	30.0	21.5	37.0	21.0
2	29.0	7.0	33.5	18.0	32.0	12.5
3	32.0	19.0	34.5	25.0	34.0	25.5
4	82.5	41.5	68.5	46.5	75.5	66.5
5	85.5	45.5	74.5	58.5	82.0	75.5
6	60.0	34.0	87.5	64.0	71.0	57.5
7	97.5	57.5	97.0	89.0	99.0	100.0
8	96.5	69.5	97.0	90.0	98.5	100.0
9	97.0	57.5	92.0	92.0	95.0	98.5
10	95.5	58.5	91.5	92.0	94.5	98.0
11	94.0	57.5	94.5	95.5	100.0	100.0

An empirical example

This section provides empirical estimates of relative efficiencies of 257 Australian mutual funds using the DEA. The aim is to investigate the extent to which adding extra input variables, that were not considered as typical inputs in the previous mutual fund appraisal studies, in the DEA affects the results of the relative efficiency estimates. The input-output variable measures on funds used in this study were collected from ASSIRT Pty Limited, Australia, covering the five-year period 1995 to 1999. ASSIRT provides information on qualitative and quantitative variables associated with a large number of mutual funds. The qualitative variables available in the ASSIRT database (ASSIRT Pty Ltd, 2000) are funds' features, management objectives and strategy, and ASSIRT rating of mutual funds, while the quantitative variables are mainly historical information on such as cash inflow, growth, income, size, asset allocation and fees charged. The reason for considering only 257 funds in this study is the availability of complete information only on these funds. The total value of assets in these 257 funds is approximately AUD 7 billion.

Input and output variable specification for DEA

There is no consensus among researchers and investors as to which input and output variables should be included in a DEA model unambiguously. Murthi, Choi and Desai (1997) analysed 731 mutual funds using DEA with one output: return and four inputs: expense ratio, load, turnover and standard deviation. McMullen and Strong (1998) applied DEA to evaluate the relative performance of 135 US common stock funds using the factors, one, three and five-year annualised returns, standard deviation of returns, sales charge, minimum initial investment and expense ratio. Sedzro and Sardano (1999), on the other hand, analysed 58 US equity funds that exist in Canada using DEA with annual-return, expense ratio, minimum initial investment and a proxy for risk as factors associated with fund performance. Galagedera and Silavapulle (2000) in their DEA analysis of the performance of 257 Australian mutual funds considered the gross performance as the output variable and entry fee, expense ratio, standard deviation of gross performance and minimum initial investment as the input variables.

We consider five-year gross performance defined in terms of growth and income as the output variable. The gross performance, expressed as a percentage per annum, is the sum of annual percentage growth in unit price of a fund and its annual percentage income. Following previous studies in the literature, the input variables considered relevant in a DEA analysis of mutual funds are (i) standard deviation; (ii) entry fee; (iii) operating expenses generally referred to as the management expense ratio; and (iv) minimum initial investment. In addition to these, three more variables: age, size and 12-month net asset flow that reflect a fund's operational characteristics and investor confidence are also considered in the analysis but as irrelevant input variables. Fund characteristics such as the turnover, the size, the managing institution, and the age have been used as determinants of fund performance, measured by relating return to input factors such as proxies for risk, general expenses and management fees, in pervious studies (Annaert, Vander Vennet and Van Den Broeck, 1999; Galagedera and Silvapulle, 2000). The age and size are considered as proxies for fund stability/experience and scale of operation respectively. The 12-month net asset flow is considered as a proxy for the level of investor confidence. Summary statistics of the variables are presented in Table 9. The Pearson correlation coefficients between the variables used are given in Table 10.

Table 9. Input and output variables for DEA runs

Variable	Unit	Mean	SD	Min	Max
<u>Output measure for DEA runs:</u>					
5-yr gross performance (5-yr GP)	%	11.74	5.45	-5	28.9
<u>Input measures for DEA runs:</u>					
Entry fee	%	2.21	1.86	0	6.00
Management expense ratio (MER)	%	1.49	0.57	0.3	3.07
Minimum initial investment (Ini Inv)	AUD	64250	147243	0	500000
5-yr standard deviation (5-yr SD)	%	2.40	1.54	0.07	8.00
Age	Months	109.68	49.11	60	399
Fund size (Size)	AUDm	264.75	584.55	0.42	6272
12-month net funds flow (12m flow)	AUDm	37.78	145.90	-407.6	1275

Results of the DEA analysis

The DEA analysis was carried out using the software DEAP Version 2.1 (Coelli, 1996). A summary of the results of the nineteen DEA runs, each with a different set of input variables, as indicated in Table 11, under VRS and DRS assumptions is given in Table 12. The DEA runs with irrelevant input variables are given in the shaded area in Tables 11 and 12. The number of input variables used in the DEA runs ranges from 1 to 7. It is evident from the results given in Table 12 that the number of funds showing up as DEA-efficient varies considerably across the DEA runs. As expected, the higher the number of input variables used in the DEA model the more funds appears to be efficient. The mean efficiency also increases with the increasing number of input variables used in the DEA model.

When the variables that we consider relevant are omitted from the DEA analysis the mean fund efficiency decreases with the increasing number of omitted variables, which we also observed in the simulation study. It is interesting to note however that the mean efficiency of DEA runs within certain pairs of DEA runs (1 and 3, 5 and 8, 6 and 9, 10 and 12) are almost identical. The only difference between the variables included in any pair of these DEA runs is that in one DEA run the variable entry fee is included as an input and in the other it is not. Further, in the DEA runs 1, 3, 5, 6, 8, 9, 10 and 12 the variable MER is included as an input factor. If two inputs are positively correlated, other things being equal, they contribute less information to the DEA analysis than if they showed zero correlation (Pedraja-Chaparro, Salinas-Jim`enez and Smith, 1999). Thus a plausible reason for our observation is the high correlation (0.752) between MER and entry fee. This observation cannot be unique to our study. The presence of entry fee in a DEA model together with MER is not likely to have a significant impact on DEA analysis of mutual funds in empirical studies and therefore it may not be necessary to include entry fee as an input factor.

When the variables that we consider irrelevant are included in the DEA analysis there is a sharp increase in mean efficiency irrespective of the RS assumption. So does the number of efficient funds. Our observation here is somewhat different from those that of the simulation study. In the simulation study we observed that inclusion of irrelevant variables makes little difference to mean efficiencies; see Table 4. Such implications for mean efficiency have been observed regardless of the complexity of the DEA model (Smith, 1997). Our input variables can be categorised under transaction costs (minimum initial investment, expense ratio, entry fee), risk (standard deviation) and other characteristics (age, size and 12-month net asset flow). The relevant variables capture the transaction costs and the risk while the irrelevant variables capture other fund characteristics. Thus a reason for the sharp increase in mean efficiency we observe here may be due to the differences in the characteristics of the variables labelled as relevant and irrelevant in the DEA analysis; a feature that can not be incorporated in a simulated experiment.

Table 10. Correlation of variables used in empirical study

	5-yr GP	Entry Fee	M.E.R.	Ini Inv	5-yr SD	Age	12m flow	Size
5-yr GP	1.000							
Entry Fee	0.050	1.000						
M.E.R.	0.046	0.752	1.000					
Ini Inv	0.139	-0.473	-0.528	1.000				
5-yr SD	0.543	0.110	0.181	0.138	1.000			
Age	0.113	0.127	0.137	-0.104	0.050	1.000		
12m flow	0.284	-0.155	-0.201	0.265	0.071	-0.039	1.000	
Size	0.237	-0.167	-0.148	0.127	0.012	0.320	0.580	1.000

Table 11. Variables considered in different DEA runs

DEA Run	5-yr GP	Entry Fee	MER	Ini Inv	5-yr SD	Age	Size	12m flow
1	x		x					
2	x				x			
3	x	x	x					
4	x	x			x			
5	x		x	x				
6	x		x		x			
7	x			x	x			
8	x	x	x	x				
9	x	x	x		x			
10	x		x	x	x			
11	x	x		x	x			
12	x	x	x	x	x			
13	x	x	x	x	x	x		
14	x	x	x	x	x		x	
15	x	x	x	x	x			x
16	x	x	x	x	x	x	x	
17	x	x	x	x	x	x		x
18	x	x	x	x	x		x	x
19	x	x	x	x	x	x	x	x

Note: Relevant input variables are in columns 3-6, and irrelevant input variables are in columns 7-9.

The conflicting results of the empirical example and the simulation study may be attributed to reasons other that the choice of input variables. The simulated experiment results are based on a specified production frontier and the mutual fund feature might not be conforming to it. In such situations, a comparison of the results of the simulated experiment and the empirical study is inappropriate. Further, we assume that by taking a large sample some compensation is provided for not replicating. This however, is not uncommon. Many simulation studies that compare the performance of estimating DMU specific efficiency using econometric methods and DEA report the results based on a single sample without replication. Nevertheless, our observations should be regarded as suggestive rather than definitive.

Table 12. Summary results of DEA runs

DEA Run	No of input variables	Mean Efficiency		Efficient funds (%)	
		VRS	CRS	VRS	CRS
1	1	0.342	0.313	1.56	0.39
2	1	0.441	0.127	2.33	0.39
3	2	0.342	0.313	1.56	0.39
4	2	0.477	0.127	2.72	0.39
5	2	0.480	0.410	5.84	2.33
6	2	0.536	0.385	3.50	1.17
7	2	0.501	0.154	5.84	1.17
8	3	0.481	0.410	5.84	2.33
9	3	0.537	0.385	3.50	1.17
10	3	0.633	0.523	11.67	5.84
11	3	0.559	0.240	8.56	2.33
12	4	0.635	0.532	17.12	5.84
13	5	0.826	0.737	24.12	14.01
14	5	0.787	0.723	24.51	14.79
15	5	0.813	0.773	19.07	13.23
16	6	0.862	0.798	31.52	21.79
17	6	0.895	0.810	31.13	20.23
18	6	0.911	0.840	33.07	24.12
19	7	0.932	0.854	39.69	29.18

Conclusion

In this paper, using a simulation study, we have investigated the effects of omission of production relevant input variables and inclusion of production irrelevant input variables in an input orientation DEA model on the results of technical efficiency estimates of individual production units. Moreover, an illustrative example is given in that we compare the efficiencies of 257 Australian mutual funds with what we observed in the simulation study.

A simulation experiment was conducted assuming the Cobb-Douglas form for the production frontier. The results based on a large sample reveal that omission of relevant variables adversely affects the DEA performance. When relevant variables are omitted in the DEA analysis VRS assumption appears to be a safer option. This is true even when the underlying specification is CRS. Inclusion of irrelevant variables in the analysis also affects the DEA performance to a lesser extent; correct RS assumption appears to be very crucial in this case.

We have provided an illustrative example in that we have compared the variation in the DEA mean efficiencies observed in the simulation study when the relevant (irrelevant) variables were omitted (included) with those observed in the real data analysis. Towards this, the efficiencies of 257 Australian mutual funds were estimated using DEA. Prompted by the previous studies that used the same set of four input variables, namely, standard deviation of funds return, management expense ratio, minimum initial investment and entry fee in the DEA analysis to evaluate mutual fund efficiencies, we considered these four input variables as the relevant ones for the mutual fund appraisal, and another three input variables, namely, age, size and 12-month net funds flow as irrelevant input variables. The results are not consistent with those observed in the simulated experiment when the irrelevant input variables are included in the DEA analysis. A sharp increase in mean fund efficiency is observed in this case. This is noticeable whether or not the returns-to-scale specification is variable or constant. However, when the relevant variables are omitted a high variation in mean DEA efficiency similar to that of simulated results is observed. Therefore, DEA's performance in the absence of relevant variables is evidently unacceptable. We believe that the results of the simulation study conducted in this paper would be useful to DEA model builders. Much more could be learned about the sensitivity of the DEA model by expanding the simulation study to incorporate multiple outputs, which is a topic for the future research.

References

Afriat SN (1972). Efficiency estimation of a production function. *Intl Econ Review* **13**: 568-598.

Aigner D and Chu SF (1968). On estimating the industry production function. *American Econ Review* **58**: 826-839.

Aigner D, Lovell CAK and Schmidt P (1977). Formulation and estimation of stochastic frontier production models. *J Econometrics* **6**: 21-37.

Annaert J, Vander Vennet R and Van Den Broeck J (1999). Determinants of mutual fund performance: A Bayasian stochastic frontier approach. Presented at SFA, Sixth European Workshop on Efficiency and Productivity Analysis, Denmark. 31/10-31/10.

ASSIRT Pty Ltd., *ASSIRT Investment Library CD-ROM*, Sydney, Australia, 2000.

Banker RD (1993). Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. *Mgt Sci* **39**: 1265-1273.

Banker R, Charnes A and Cooper WW (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Mgt Sci* **30**: 1078-1092.

Banker RD, Gadh VM and Gorr WL (1993). A Monte Carlo simulation of two production frontier estimation methods: corrected ordinary least squares and data envelopment analysis. *Eur J Opl Res* **67**: 332-343.

Charnes A, Cooper WW, Levin AY, Morey RC and Rousseau JJ (1985). Sensitivity and stability analysis in DEA. *Annals of Opns Res* **2**: 139-156.

Charnes A, Cooper WW and Rhodes E (1978). Measuring the efficiency of decision making units. *Eur J Opl Res* **2**: 429-444.

Charnes A, Haag S, Jaska P and Semple J (1992). Sensitivity of efficiency classifications in the additive model of data envelopment analysis. *Int J Systems Sci* **23**: 789-798.

Charnes A and Neralie` L (1990). Sensitivity analysis of the additive model in data envelopment analysis. *Eur J Opl Res* **48**: 332-341.

Charnes A, Rousseau J and Semple J (1996). Sensitivity and stability of efficiency classifications in data envelopment analysis. *J Prod Anal* **7**: 5-18.

Coelli TJ (1996). A guide to DEAP Version 2.2: A data envelopment analysis (computer) program. *CEPA Working Paper 96/8*, Department of Econometrics, University of New England.

Farrell MJ (1957). The measurement of productive efficiency. *J Royal Statistical Society* **120**: 253-290.

Galagedera DUA and Silvapulle P (2000). Australian mutual fund performance appraisal using data envelopment analysis. *13th Australasian Finance & Banking Conference*, Sydney, Australia.

Gong B and Sickles RC (1992). Finite sample evidence on the performance of stochastic frontiers and data envelopment analysis using panel data. *J Econometrics* **51**: 259-284.

Greene WH (1980). Maximum likelihood estimation of econometric frontier functions. *J Econometrics* **13**: 27-56.

Jondrow J, Lovell CAK, Materov IS and Schmidt P (1982). On the estimation of technical efficiency in the stochastic frontier production function model. *J Econometrics* **19**: 233-238.

McMullen PR and Strong RA (1998). Selection of mutual funds using data envelopment analysis. *J Bus and Econ Studies* **4**: 1-12.

Meeusen W and Van Den Broeck J (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *Intl Econ Review* **18**: 435-444.

Murthi BPS, Choi YK and Desai P (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. *Eur J of Operational Res* **98**: 408-418.

Pedraja-Chaparro F, Salinas-Jim`enez J and Smith P (1999). On the quality of the data envelopment analysis model. *J Opl Res Sc* **50**: 636-644.

Read L and Thanassoulis E (1996). A comparison of data envelopment analysis and stochastic frontiers under variation of fit. Research Paper No 234 Warwick Business School, Warwick University.

Richmond J (1974). Estimating the efficiency of production. *Intl Eco Review* **15**: 515-521.

Ruggiero J (1998a). A new approach for technical efficiency estimation in multiple output production. *Eur J Opl Res* **111**: 369-380.

Ruggiero J (1998b). Non-discretionary inputs in data envelopment analysis. *Eur J Operational Res* **111**: 461-469.

Ruggiero J (1999). Efficiency estimation and error decomposition in the stochastic frontier model: a Monte Carlo analysis. *Eur J Opl Res* **115**: 555-563.

Sedzro K and Sardano D (1999). Mutual fund performance evaluation using data envelopment analysis. Working Paper, School of Business, University of Quebec at Montreal, Canada.

Seiford LM (1996). Data envelopment analysis: the evolution of the state of the art (1978-1995). *J Product Anl* **7**: 99-137.

Seiford LM and Zhu J (1998a). Stability regions for maintaining efficiency in data envelopment analysis. *Eur J Opns Res* **108**: 127-139.

Seiford LM and Zhu J (1998b). Sensitivity analysis of DEA models for simultaneous changes in all the data. *J Opl Res Society* **49**: 1060-1071.

Smith P (1997). Model misspecification in data envelopment analysis. *Annals of Operations Research* **73**: 233-252.

Stevenson RE (1980). Likelihood functions for generalized stochastic frontier estimation. *J Econometrics* **13**: 57-66.

Valdmanis V (1992). Sensitivity analysis for DEA models: An empirical example using public vs NFP hospitals. *J of Public Economics* **48**: 185-205.

Zhu J (1996). Robustness of the efficient DMUs in data envelopment analysis. *Eur J Opl Res* **90**: 451-460.