KEIO ECONOMIC STUDIES **37**(1), 13–32 (2000)

NONPARAMETRIC EFFICIENCY MEASUREMENT UNDER DEMAND AND COST UNCERTAINTY

Jati K. SENGUPTA

Department of Economics, University of California, Santa Barbara, USA

First version received August 1998; final version accepted March 2000

Abstract: Cost inefficiency in firms may arise due to inoptimal input usage in the short run and a decline in market demand leading to excess capacity in the long run. The influence of these two forces: demand uncertainty and slack-ridden costs are explored here in respect of the recent nonparametric techniques which are essentially based on the concept of Pareto efficiency.

1. INTRODUCTION

Recently nonparametric techniques based on Pareto efficiency have been increasingly applied to compare the relative efficiency of public sector organizations. These techniques known as 'data envelopment analysis' (DEA) have seen new developments in recent years in a number of directions, see e.g., Charnes et al. (1994) and Sengupta (1995, 1998a). These new developments however have failed to introduce the DEA tool as a control theoretic device, e.g., how to use the DEA model as a policy tool in determining the optimal level of inputs and outputs. Two types of efficiency measures are usually distinguished at the microlevel. One is technical or production efficiency, which measures the firm's success in producing maximum output from a given set of inputs. The other is the price or allocative efficiency, which measures the firm's success in choosing an optimal set of inputs with a given set of input prices or observed input costs. In our approach we use the allocative efficiency criterion based on observed input prices or costs to determine the optimal level of inputs.

Recently the DEA approach has been applied extensively to measure the relative cost efficiency of competitive firms in the private sector, e.g., the efficiency of international airlines has been analyzed by Schefczyk (1993), Sengupta (1996) and Oum and Yu (1998). These empirical studies have found that gross profit margins are not the sole determinant of firm efficiency, since other factors like market competition and long term viable strategies play an equal and possibly more important role.

Our object here is to develop a class of efficiency models which characterizes both efficiency ranking and the optimal level of inputs for attaining the highest level of relative

Acknowledgement. An anonymous referee's helpful comments are gratefully acknowledged.

efficiency. These models are useful for both the public sector organizations where only input cost data are available but no output prices and also the private sector firms where both input and output prices are available from competitive market data. This second type of application to private sector firms is most important from an economic viewpoint, since it links the DEA approach with the economic theory of market competition and also the concept of organizational slack or X-efficiency discussed by Leibenstein (1966) and more recently by Selten (1986).

2. ALLOCATIVE MODELS OF EFFICIENCY

We consider first a model of relative efficiency, where a firm or decision making unit (DMU) is compared to the cluster of firms or DMUs and only input output data are available with no price data. This is the comparison of production or technical efficiency across firms. Secondly, we consider a model which minimizes overall unit costs subject to input and output constraints. This yields overall efficiency which can be decomposed into technical and allocative efficiency. Finally, we consider a more general class of models where demand considerations are introduced and both demand and cost uncertainty are incorporated. This type of model is generalized to dynamic frameworks, when intertemporal cost functions are minimized. Consider first a DEA model for characterizing the static efficiency of a reference unit k in a cluster of N units, where each unit j or DMU_j has m inputs (x_{ij}) and s outputs (y_{rj}):

$$\begin{aligned}
&\min_{\lambda,\theta} \theta \\
&\text{subject to} \\
&\sum_{j=1}^{N} x_{ij}\lambda_j \leq \theta x_{ik}\,; \quad i = 1, 2, \ldots, m \\
&\sum_{j=1}^{N} y_{rj}\lambda_j \geq y_{rk}\,; \quad r = 1, 2, \ldots, s \\
&\sum_{j=1}^{N} \lambda_j = 1\,, \quad \lambda_j \geq 0\,, \ \theta \geq 0\,.
\end{aligned} \tag{1}$$

In vector matrix form this is:

$$\text{Min}\,\theta\,, \quad \text{s.t.} \quad X\lambda \leq \theta X_k\,; \quad Y\lambda \geq Y_k\,; \quad \lambda' e = 1\,; \ \lambda \geq 0 \tag{2}$$

where e is a column vector with N elements, each of which is unity and the prime denotes a transpose. Here the input (X_j) and output (Y_j) vectors $(j = 1, 2, \ldots, N)$ are all observed and this is called an input oriented model in DEA literature. Here the reference unit k is compared with the other $(N - 1)$ units in the cluster. Let $\lambda^* = (\lambda_j^*)$ and θ^* be the optimal solutions of the above DEA model with all the slack variables zero. Then the reference unit k or DMU_k is technically efficient if $\theta^* = 1$ and the first two sets of inequalities in (1) hold with equality. Thus the optimal value of θ^* provides a measure of technical efficiency (TE). If θ^* is positive but less than unity, then it is not technically efficient at the 100 percent level. Overall efficiency (OE_j) of a DMU or firm j however combines both technical (TE_j) or production efficiency and the allocative (AE_j) or price efficiency as follows:

$$\mathrm{OE}_j = \mathrm{TE}_j \times \mathrm{AE}_j\,; \quad j = 1, 2, \ldots, N\,. \tag{3}$$

Recently a number of research monographs have discussed the current state of research in these areas of efficiency analysis. For example Ganley and Cubbin (1992) has discussed the specific problems of application in the public sector enterprises. Fried et al. (1993) have discussed several economic applications of the concepts of technical and allocative efficiency, whereas Charnes et al. (1994) have presented several types of theoretical generalizations. We present here some of the salient features of this recent research and discuss its relevance to demand and cost uncertainty.

To characterize overall efficiency of the reference unit DMU_k one sets up the linear programming (LP) model as follows:

$$\begin{aligned} &\operatorname*{Min}_{x,\lambda} q'x \\ &\text{s.t.} \\ &\quad \sum_{j=1}^{N} X_j\lambda_j \le x \\ &\quad \sum_{j=1}^{N} Y_j\lambda_j \ge Y_k \\ &\qquad \lambda'e = 1\,; \quad \lambda \ge 0\,; \quad x \ge 0\,. \end{aligned} \tag{4}$$

Here q is an m-element vector of input prices as observed in the competitive market and x is an input vector to be optimally decided by DMU_k along with the weights λ_j. Here X_k and Y_k are the observed input and output vectors for the reference unit k, whereas x is the unknown decision vector to be optimally determined. Let λ^* and x^* be the optimal solution of the LP model (3) with all slacks zero. Then the minimal input cost is given by $c_k^* = q'x^*$, whereas the observed cost of the reference unit is $c_k = q'X_k$. Hence the three efficiency measures are defined as follows:

$$\mathrm{TE}_k = \theta^*\,; \quad \mathrm{OE}_k = c_k^*/c_k \quad \text{and} \quad \mathrm{AE}_k = \mathrm{OE}_k/\mathrm{TE}_k\,. \tag{5}$$

Two important points are to be noted when we compare the LP model (4) with (1). First of all, the input vector x in (4) is a decision vector to be optimally chosen, whereas X_k is the observed data in (1). If $\theta^* X_k = x^*$, then the two models generate identical optimal solutions; otherwise the two optimal solutions are very different. The dual problems corresponding to (4) and (1) appear as follows:

$$\begin{aligned} &\operatorname{Max} \alpha' Y_k + \alpha_0 \\ &\quad \text{s.t.} \quad \beta \le q \;\text{ and }\; \beta' X_j \ge \alpha Y_j + \alpha Y_j + \alpha_0\,; \; j = 1, 2, \ldots, N \\ &\qquad\quad \alpha, \beta \ge 0 \;\alpha_0 \text{ free in sign} \end{aligned} \tag{6}$$

and

$$\begin{aligned} &\text{Max}\, \alpha Y_k + \alpha_0 \\ &\quad \text{s.t.} \quad \beta' X_j \geq \alpha Y_j + \alpha_0\,; \quad j = 1, 2, \ldots, N \\ &\qquad\quad \alpha, \beta \geq 0 \;\; \alpha_0 \text{ free in sign}. \end{aligned} \tag{7}$$

Let asterisks denote optimal values and let DMU_k be efficient. Then it must follow from (6) that the production frontier for the k-th unit is as follows:

$$\alpha^{*\prime} Y_k = \beta^{*\prime} X_k - \alpha_0^*$$

but since β^* is constrained as $\beta^* \leq q$, we must have $\alpha^{*\prime} Y_k \leq q' X_k - \alpha_0$. Thus so long as the actual inputs X_k are not equal to their optimal levels x^*, this efficiency gap measured by $(q' X_k - \alpha_0^{*\prime} - \alpha^{*\prime} Y_k)$ may persist. Thus the constraint $\beta^* \leq q$ reflects the fact that the observed input X_k of the reference unit may or may not be equal to the optimal level x^*, when all firms face the same competitive price q. There is no such constraint for the dual problem (7). Note that if α_0^* is positive (negative or zero), then we have increasing (decreasing or constant) returns to scale.

The second point to note is that the overall efficiency and hence the allocative efficiency in model (4) depend very critically on the observed vector q of market prices, which reflects the pattern of market demand for the whole industry. Farrell (1957), who is the first to develop the nonparametric method of efficiency measurement by an LP model similar to (2) but limited to the case of a single output recommended against the use of price or allocative efficiency, even when the market price data are available. He raised the objection that this efficiency measure would be seriously biased if the observed input prices are widely fluctuating. However, this aspect of random variations in input prices can be easily incorporated in a risk averse efficiency measure, as has been shown by Sengupta (1998b).

3. GENERALIZED EFFICIENCY MODELS

We now consider a more generalized version of the overall efficiency model (4), where both input (q) and output prices (p) are assumed to be available and the optimal vectors of input and output are optimally chosen as follows:

$$\begin{aligned} &\text{Max}\, p' y - q' x \\ &\quad \text{s.t.} \quad \sum_{j=1}^{N} X_j \lambda_j \leq x\,; \quad \sum_{j=1}^{N} Y_j \lambda_j \geq y \\ &\qquad\quad x \leq X_k\,; \;\; y \geq Y_k\,; \;\; \lambda' e = 1 \\ &\qquad\qquad x, y, \lambda \geq 0 \end{aligned} \tag{8}$$

Here (x, y) are the control vectors of inputs and output to be optimally determined and (X_k, Y_k) denote the observed levels for DMU$_k$. The dual of this problem then becomes

$$\begin{aligned} &\text{Max } v'Y_k - u'X_k - \alpha_0 \\ &\quad \text{s.t.} \quad p \le \alpha - v\,, \quad q \ge \beta - u \\ &\qquad\quad \beta' X_j \ge \alpha' Y_j + \alpha_0\,, \quad j = 1, 2, \dots, N \\ &\qquad\quad (u, v, \alpha, \beta) \ge 0\,, \quad \alpha_0 \text{ free in sign} \end{aligned} \tag{9}$$

Two special cases of the generalized model (8) are of great importance. One is the simpler output-oriented model where demand (d_r) for output (y_r) is subject to a probability distribution $F(d_r)$ and the objective function is to maximize the expected value of total revenue minus expected inventory cost. This yields the model

$$\begin{aligned} &\text{Max } E\left[\sum_{r=1}^{s} p_r \min(y_r, d_r) - \sum_{r=1}^{s} h_r (y_r - d_r)\right] \\ &\text{s.t.} \quad X\lambda \le X_k\,; \quad X\lambda \ge y\,, \quad \lambda' e = 1\,, \quad \lambda \ge 0 \end{aligned} \tag{10}$$

where λ and y are the unknown vectors to be optimally solved for and h_r is the observed unit cost of positive inventory for $y_r > d_r$. Denoting optimal values by asterisks, the efficient DMU$_k$ would then satisfy the following marginal condition:

$$\begin{aligned} F(y_r^*) &= (p_r + h_r)^{-1}(p_r - \alpha_r^*) \\ \alpha^{*\prime} Y_k &= \beta^{*\prime} X_k - \alpha_0^* \end{aligned}$$

Clearly higher output price and lower inventory costs would increase the optimal output levels y_r^* which may be compared with the observed outputs y_{rk} in output vector Y_k.

The second case is an input-oriented model, where the input decisions x_i are equal to planned values $\bar{x}_i$ plus an error term ε_I with a zero mean and fixed variance. The errors are disturbances such as mistakes or unexpected difficulties in implementing a planned value $\bar{x}_i$. The planned values $\bar{x}_i$ are the decision variables which have to be optimally chosen by each DMU and the error process ε_i is realized after the planned value of $x_i(t)$ is optimally selected. The input constraints now turn out to be chance constrained

$$\text{Prob}\left[\sum_{j=1}^{N} x_{ij}\lambda_{ij} \le \bar{x}_i + \varepsilon_i\right] = \delta_i\,, \quad 0 < \delta_i < 1$$

where δ_i is the tolerance level of the i-th input constraint. The simpler model then takes the following form:

$$\begin{aligned} &\text{Min} \sum_{i=1}^{m} q_i \bar{x}_i \\ &\text{s.t.} \quad \sum_{j=1}^{N} x_{ij}\lambda_j = \bar{x}_i + w_i\,; \quad w_i = F^{-1}(1-\delta_i) \\ &\qquad \sum_{j=1}^{N} y_{rj}\lambda_j \geq y_{rk}\,; \quad \lambda' e = 1\,, \quad \lambda \geq 0 \\ &\qquad\quad i = 1, 2, \ldots, m\,; \quad r = 1, 2, \ldots, s\,. \end{aligned} \tag{11}$$

Clearly the input uncertainty is here captured by the term w_i which depends on the level δ_I of chance constraint, e.g., the higher the level of w_i, the lower would be the optimal planned inputs $\bar{x}_i^*$.

The main implication of the general model (8) is that the input and output gaps measured by $|x_i^* - x_{ik}|$, $|y_r^* - y_{rk}|$ can be quantified as a source of inefficiency. Even if the two constraints $x \leq X_k$, $y \geq Y_k$ are dropped, we would have the dual model

$$\begin{aligned} \text{Min}\,\alpha_0 \quad &\text{s.t.} \quad p \leq \alpha\,, \quad q \geq \beta \\ \text{and} \quad &\beta' X_j \geq \alpha' Y_j + \alpha_0\,, \quad j = 1, 2, \ldots, N \\ &\alpha, \beta \geq 0\,, \quad \alpha_0 \text{ free in sign}\,. \end{aligned}$$

Note that a positive (negative or zero) value of α_0 indicates the size of increasing (decreasing or constant) returns to scale; hence the dual model can quantify the source of inefficiency of a relatively inefficient DMU in terms of returns to scale. Furthermore if $\alpha^* = p$ and $\beta* = q$ then the DEA based profit measure π_j^* for DMU$_j$ becomes

$$\pi_j^* = p'Y_j - q'X_j + \alpha_0^* \leq 0\,, \qquad j = 1, 2, \ldots, N$$

but the efficient DMU$_k$ must have zero profit $\pi_k^* = 0$ always.

On using the optimal solutions λ^*, x^*, y^* of the LP model (8) one could define the composite input $X_c^* = \sum_{j=1}^{N} X_j\lambda_j^*$ and composite output $Y_c^* = \sum_{j=1}^{N} Y_j\lambda_j^*$ and compare with x^* and y^* respectively. This comparison would show if the Pareto efficiency property holds for the optimal input (x^*) and output (y^*) vectors or not.

4. MARKET COMPETITION AND EFFICIENCY

We consider now the role of market competition in the efficiency framework. Hence we assume that each firm or DMU$_j$ produces a single homogenous output denoted by y_j, where the total industry output is denoted by $y_T = \sum_{j=1}^{N} y_j$. If N is large and the firms or DMUs are competitive, then the output price p is a constant, unaffected by the size of each individual firm. In this case the price can be viewed as $p = \bar{p} + \varepsilon$ made up of two components: the expected price $\bar{p}$ and a random part ε with a zero mean and a constant variance σ_ε^2. The total cost of inputs for each firm may now be related to output as

$$c(y_j) = cy_j + F_j$$

assuming a linear form, where F_j is the fixed cost and c is marginal cost that is assumed to be identical for each firm. Maximization of expected profits would then yield the LP model:

$$\begin{aligned} &\operatorname{Max} \bar{\pi} = (\bar{p} - c)y - F \\ &\text{s.t.} \quad \sum_{j=1}^{N} X_j \lambda_j \le X_k \,; \quad \sum_{j=1}^{N} y_j \lambda_j \ge y \,; \quad \lambda' e = 1 \,, \quad \lambda \ge 0 \end{aligned} \tag{12}$$

where y is the unknown decision variable to be optimally selected. In case the market is imperfectly competitive, the price variable then depends on the output supply of different firms. In the homogenous output case the firms are all alike, and the inverted demand function can be written as:

$$\bar{p} = a - bY_T \,, \qquad Y_T = \sum_{j=1}^{N} y_j \,; \quad y_k = y$$

The LP model (12) would then yield the following optimality conditions:

$$\begin{aligned} &(a - c) - bY_T - by^* - \alpha^* \le 0 \\ &\alpha^* y_j - \beta^{*\prime} X_j - \alpha_0^* \le 0 \\ &\alpha^*, \beta^* \ge 0 \,, \quad \alpha_0^* \text{ free in sign} \,. \end{aligned} \tag{13}$$

If firm k is efficient, then one must have

$$\begin{aligned} &y^* = (a_1/b) - Y_T - \frac{\alpha^*}{b} \,; \quad y^* > 0 \\ &a_1 = a - c > 0 \end{aligned}$$

where $y^* = y_k^*$ is the efficient output of the k-th firm. If all firms are efficient, then $Y_T^* = \sum_{j=1}^{N} y_j^*$ and one obtains

$$Y_T^* = (N/b)(1 + N)^{-1}[a_1 - \alpha^*] \tag{14}$$

This analysis of the impact of competitive market pressure on the levels of firm and industry efficiency tends to neglect however a central cost variable in the behavioral theory of the firm. This has been called 'the organizational slack' variable by Cyert and March (1963) and X-efficiency by Leibenstein (1966). In his book *Beyond Economic Man* Leibenstein (1976) has discussed a number of empirical studies which show the importance of the concept of organizational slack. Recently Selten (1986) has used this concept as a part of the cost function of individual firms and its role in imperfect competition. Two hypotheses have been proposed: the strong and the weak slack. The strong slack hypothesis maintains that the slack has a tendency to increase to long as profits are positive; slack can be reduced but only under the threat of losses. This has the consequence that long run profits tend to be zero, regardless of the market structure. The weak slack hypothesis has the implication that long run profits may not tend to zero. Thus in imperfect markets with slack, competition not only reduces profits, it also puts pressure on costs. This aspect will now be discussed in the context of DEA models of efficiency.

Now we introduce organization slack denoted by s_k in the cost function

$$c(y) = (c + s_k)y + F_k\,; \quad s_k \geq 0$$

Following Selten we interpret the slack concept due to Leibenstein's X-efficiency as a part of the cost function and introduce a 'strong-slack' hypothesis which maintains that this type of slack has a tendency to increase so long as profits are positive, i.e., this slack can be reduced only under the threat of loses. Including this slack-ridden cost into the profit maximization model would yield the optimality condition for the efficient output as:

$$y^* = (a_1 - \alpha - s_k)/b - Y_T\,; \quad y^* > 0$$

where $y^* = y_k^*$ is the efficient output of the firm k. If all firms are efficient then

$$\begin{aligned} Y_T^* &= (N/b)(1+N)^{-1}[a_1 - \bar{s} - \alpha^*] \\ \bar{p} &= a + N(N+1)^{-1}(\alpha^* + \bar{s} - a_1) \\ \pi_k^* &= (\bar{p} - cs_k)y^* - F_k \end{aligned} \tag{15}$$

when $\bar{s} = \sum_{k=1}^{N} s_k/N$ is the average rate of slack. Several implications follow from this set (15) of efficiency conditions. First of all, the long run pressure of competition would tend to lead to zero profits $\pi_k^* = 0$ for all $k = 1, 2, \ldots, N$ according to the strong slack hypothesis. In this case the expected price becomes $\bar{p} = c + \bar{s} + (F/y^*)$. This shows that fixed costs have a strong positive role in determining the long run equilibrium price. The higher the average slack rate $\bar{s}$, the higher is the equilibrium expected price. Secondly, as the number N of firms increases, it increases the volume of total industry output Y_T^* and reduces the average price. Finally, as the average slack rate $\bar{s}$ rises (falls), it increases (decreases) the equilibrium price. Note that in case of weak slack hypothesis all profits are not squeezed out and there remains a divergence of individual (s_k) from the average slack rate ($\bar{s}$), when the latter is positive. Thus some inefficiency may persist due to the existence of a positive slack.

So far we have assumed that the expected price $\bar{p}$ is the market clearing price equating market demand and supply. If however this is not the case, then the supply y would differ from demand d, where demand is subject to random fluctuations around the mean level $\bar{d}$. In this framework we have to add to the cost function the costs of inventory and shortage $C(y - d)$. Assuming this cost to be quadratic one may then formalize the decision model

$$\begin{aligned} &\text{Max}\, \bar{\pi} = (\bar{p} - c - s_k)y - (1/2)\gamma E(y - d)^2 - F \\ &\text{s.t.} \quad \text{the same constraints as in (12)}\,. \end{aligned} \tag{16}$$

In this case the optimality conditions for the efficient output becomes

$$y^* = (b + \gamma)^{-1}[(a_1 - \alpha^* - s_k) + \gamma\bar{d} - bY_T] \tag{17}$$

where γ is the unit cost of inventory or shortage and $\bar{d}$ is the expected level of demand. In this case the marginal impact $\partial y^*/\partial\gamma$ of inventory/excess costs may be either negative or positive according as

$$b(Y_T + \bar{d}) > \text{or}, < (a_1 - \alpha^* - s_k)$$

Again this explains the persistence of some inefficiency, when demand is uncertain and the firm chooses its optimal output by the quadratic criterion of adjusted profits. Furthermore the higher the mean demand, the greater the optimal level of efficient output y^*. In case of perfect competition with each firm a price taker, the optimality condition (17) reduces to

$$y^* = \bar{d} + (1/\gamma)(\bar{p} - c - s_k - \alpha^*) \tag{18}$$

which shows unequivocally that higher inventory costs (γ) lead to lower optimal output. Again by comparing the observed output y_k with the optimal output y^*, one could evaluate the impact of inefficiency. Note that we still have the comparative static results: $\partial y^*/\partial\bar{p} > 0$ and $\partial y^*/\partial s_k < 0$. Since $\bar{p} = \gamma(y^* - \bar{d}) + c + s_k + \alpha^*$, we have the results:

$$\bar{p} > \mathrm{MC}_T\,, \quad \text{if } y^* > \bar{d}$$

and (19)

$$\bar{p} < \mathrm{MC}_T\,, \quad \text{if } y^* < \bar{d}$$

where $\mathrm{MC}_T = c + s_k + \alpha^*$ is total marginal cost with three components: production costs (c), cost of slack (s_k) and the cost of discrepancy of observed from optimal output (α^*). Clearly the case of multiple output can be handled in a symmetrical way.

5. COST UNCERTAINTY AND CAPACITY UTILIZATION

Capacity utilization has two basic roles in industrial price and output policies. The first is one of the basic propositions in macroeconomics which says that price inflation accelerates as capacity and resource utilization moves higher. The second is the intertemporal implication of changes in capacity inputs, which affect both the fixed and variable costs in the short run. Since every short run production and cost function is conditional on a fixed supply of capacity inputs, the short run cost minimization model may not ordinarily yield the long run cost frontier. We consider here first, a two-period model of capacity expansion and derive the implication of varying the capacity utilization rates. In the next section the long run implication of optimal capacity expansion and its impact on efficient outputs and prices is investigated in some detail.

The term 'capacity' is often viewed as a ceiling on production or output, that is commonly referred to as the engineering definition of capacity. It has long been recognized that this definition is largely irrelevant to economics. For example, a number of empirical studies have found that the capital stock in the US is idle to a significant degree for most of the time. One of the earliest by Foss (1963) reported an average work week of capital of only 38 hours per week. Presumably most of the idleness is either optimal or useful in the managerial discretionary behavior.

Economists view capacity rather differently. According to Winston (1974), Klein and Long (1973) full capacity describes a firm's planned or intended level of utilization; the level that reflects satisfied expectations and is built into the capital stock and embodied in the normal working schedule. Two empirical measures of capacity are commonly used in applied work in manufacturing industries. One is the US Federal Reserve Board (FRB) series on capacity indexes which attempt to capture the concept of sustainable practical capacity, which is defined as the greatest level of output that a plant can maintain within the framework of a realistic work schedule, taking account of normal downtime and assuming sufficient availability of inputs to operate the machinery and equipment in place. Hence this level of output does not necessarily represent either the maximum that can be extracted from the fixed plant (as indicated by utilization rates that sometimes exceed 100 per cent) or the level associated with the minimum point of the short run average cost curve. More specifically, the first step in estimating capacity indexes is to divide an industrial production index (Q_t) by a utilization rate (CU_t) provided by the Census Department's Survey of Plant Capacity Utilization. This yields an initial estimate of implied capacity: $\mathrm{IC}_t = Q_t/\mathrm{CU}_t$. However the survey is conducted every four years and firms are asked to report utilization in the fourth quarter of that year. This generally leads to cyclical variability in implicit capacity. To eliminate this cyclical volatility the second step is used to regress implied capacity IC_t on capital stock (K_t) and a deterministic function of time as

$$\ln \mathrm{IC}_t = \ln K_t + \alpha + \sum_{i=1}^{\tau} \beta_i f_i(t)\,; \tag{20}$$

where K_t is the year end capital stock and $f_i(t)$ is an i-th order polynomial defined on time t. The fitted values from these regressions provide baselines for the annual FRB estimates of productive capacity (C_t).

A second method in estimating production capacity is to use a filter due to Hodrick and Prescott (i.e., HP filter), which decomposes a time series into a permanent and a transitory component. Hodrick and Prescott (1997) define the permanent component as including those variations which are sufficiently smooth to be consistent with slowly changing demographic and technological factors and the accumulation of capital stocks. The HP permanent component is used as a measure of capacity. Then the capacity utilization is calculated as production (Q_t) divided by the HP permanent component. In the short run, both demand and supply shocks may cause the deviations of actual output from the permanent component. Note that firms may have several options in regard to raising output above its potential level, e.g., by adding shifts, varying the production line speeds, altering the product mix or even bringing mothballed facilities back into use.

Recently Kennedy (1995) used quarterly data (1960I–1992IV) for US manufacturing to regress the rate of producer price index (PPI) on both utilization rates of FRB and HP capacity variables and found the HP variable to be dominant. For example in manufacturing the HP rate coefficient is 29.10 with a t-statistic of 2.8, whereas the FRB coefficient is -1.06 with a t-statistic close to zero. For the disaggregated industries (two digit SIC code industries) the results are similar.

In our approach we combine the two methods above to define a series of capacity levels CAP_{jt} for j-th unit at time t. This is based on two steps. In the first step we assume an additive decomposition of implied capacity into a permanent component ($\text{IC}^{\text{P}}_{jt}$) and a transitory component (ζ_{jt}):

$$\text{IC}_{jt} = \text{IC}^{\text{P}}_{jt} + \zeta_{jt} .$$

A filtering method (e.g., Kalman filter) is applied here to estimate the permanent component, until the random component ζ_{jt} turns out to be a white noise process. In the second step we use the data on capital stock (K_{jt}) and the time variable to regress $\text{IC}^{\text{P}}_{jt}$ on K_{jt} and $f_i(t)$ as defined in (20):

$$\ln \widehat{\text{IC}}_{jt} = \ln K_t + \alpha + \sum_{i=1}^{\tau} \beta_i f_{ij}(t)$$

Taking antilogs of the dependent variable we obtain the estimate $\widehat{\text{CAP}}_{jt}$ of capacity. On using this capacity series $z_{jt} = \widehat{\text{CAP}}_{jt}$ we set up two overall cost minimization models in the DEA framework: one involving the optimal utilization rate ψ^* and the other the optimal capacity z^* and optimal variable inputs x^*.

$$\begin{aligned} &\text{Min}\, \theta + \psi \\ &\text{s.t.} \quad \sum_{j=1}^{N} \lambda_j x_j \leq \theta X_k ; \quad \sum_{j=1}^{N} \mu_j z_{jt} \leq \psi z_{kt} \\ &\qquad \sum_{j=1}^{N} \lambda_j Y_j \geq Y_k ; \quad \sum \lambda_j = 1 = \sum \mu_j \\ &\qquad\qquad \lambda_j, \mu_j \geq 0 \end{aligned} \tag{21}$$

and

$$\begin{aligned} &\text{Min}\, q'x + z \\ &\text{s.t.} \quad \sum_j \lambda_j X_j \leq x ; \quad \sum_j \mu_j z_{jt} \leq z ; \quad \sum_j \lambda_j Y_j \geq Y_k \\ &\qquad \sum \lambda_j = 1 = \sum \mu_j ; \quad (\lambda, \mu) \geq 0 \end{aligned} \tag{22}$$

Here capacity z is a scalar variable, (X_j, Y_j) are input output vectors for unit j and q' is a row vector denoting unit costs (prices) for the variable inputs x. Denote by asterisks the optimal values of the decision variables. Then unit k is relatively inefficient in the use of capacity inputs if $\psi^* < 1.9$, whereas it is inefficient in the use of current inputs if $\theta^* < 1.0$. The optimal values x^*, z^* of current and capacity input may also be compared with the actual levels X_k, z_k used by unit k in order to locate efficiency gap if any. In case market price data (p) are available for the output vector y and a two-period framework is assumed, then the optimal inputs and outputs can be determined from the

LP model as follows:

$$\underset{y,x,z}{\text{Max}}\ \pi = p_t'y - q_t'x - w_t'z + (1+r)^{-1}w_{i+1}'\hat{z}$$

$$\text{s.t.}\quad X\lambda \leq x\,; \quad Y\lambda \geq y\,; \quad \mu'z_t \leq z\,; \quad \mu'\hat{z}_t \geq \hat{z} \tag{23}$$

$$\lambda'e = 1 = \mu'e\,; \quad \lambda \geq 0\,, \quad \mu \geq 0$$

Here z_t is a vector of durable inputs purchased at the beginning g of period t at prices w_t, $\hat{z}_t$ is a vector of depreciated durable inputs that will be available to the firm at the beginning of the subsequent period, w_{t+1} is the vector of durable input prices that the firm anticipates will prevail during period $t+1$, and r is an appropriate discount rate exogenously given. Here the capacity-related inputs are the durable inputs and their unit costs are the input prices. With observed values $(X, Y, z_t, \hat{z}_t)$ of inputs and outputs the firm could now determine the optimal inputs and outputs $(x^*, y^*, z^*, \hat{z}^*)$. We note however some basic differences of this formulation from the traditional DEA models. First of all, the vector of spot prices w_{t+1} is not observed at time t and hence the producer's anticipation of future price is needed. In this sense this model yields anticipated or expected efficiency. Since the anticipated prices are uncertain, the firm's attitude towards uncertainty must be modeled. This is the framework where the rational expectations (RE) hypothesis may be introduced. Secondly, the durable inputs are used here to approximate the stock of capacity inputs, but for certain stocks like natural resources and goods inventories there may be no natural market prices. Finally, the relevant discount rate r must be common to all the firms and also known. In the static DEA models these basic questions are not addressed at all.

6. THE CASE OF STOCHASTIC DEMAND

Uncertainty of demand affects the pattern of capacity utilization both in the short and the long run. Inventories in the form of unsold outputs may often trigger this process.

In case of stochastic demand d_t the DEA model (23) can be transformed as:

$$\text{Max}\ E\pi = E[p_t' \min(y, d_t) - q_t'x + (1+r)^{-1}\hat{z} - w_t'z]$$

$$\text{s.t.}\quad \text{the same constraints as in (23)}$$

where E denotes expectation. On using the Lagrangian expression:

$$\begin{aligned} L =& E\pi + \beta'(x - X\lambda) + \alpha'(Y\lambda - y) + \gamma'(z - \mu'z_t) + \delta'(\mu'\hat{z}_t - \hat{z}) \\ &+ \beta_0(1 - \lambda'e) + \gamma_0(1 - \mu'e) \end{aligned}$$

We must have for the efficient unit:

$$\begin{aligned} &p_t(1 - F(y^*)) - \alpha^* = 0 \\ &\gamma^* = w_t\,; \quad \beta^* = q_t\,; \quad \delta^* = (1+r)^{-1} \\ &Y'\alpha^* = X'\beta^* + \beta_0^* e\,; \quad \gamma^{*\prime} z_t = \delta^{*\prime}\hat{z}_t - \gamma_0^* e\,. \end{aligned}$$

This shows that the unit exhibits output inefficiency if $\sum_{j=1}^{N} Y_j\lambda_j^* > y^*$, input efficiency if $\sum X_j\lambda_j^* < x^*$, capacity inefficiency if $\mu'^*z_t < z$ or $\mu^{*\prime}z_t > \hat{z}^*$. Clearly there are five sources of inefficiency in this framework: the input, output, capacity and inefficiency due to market demand uncertainty. The theory of organizational slack deals specifically with the demand and capacity oriented sources of inefficiency which may apparently inflate the marginal costs.

In case we have a time horizon it is simpler to introduce investment variables denoted by a vector I_t and rewrite the long run profit function as

$$\begin{aligned} \text{Max } E & \left\{ \sum_{t=0}^{\infty} (1+r)^{-1} [p_t' \min(y_t, d_t) - q_t' x_t - \rho_t' I_t - w_t' z_t] \right\} \\ \text{s.t. } & X_t \lambda_t \le x_t\,; \quad Y_t \lambda_t \ge y_t\,; \quad I_t \le (1+\delta_0) z_{t+1} - z_t \\ & \lambda_t' e = 1\,; \quad \lambda_t \ge 0 \end{aligned}$$

where investment is constrained by changes in capacity inputs with δ_0 denoting fixed rates of depreciation. The theory of adjustment costs which relates current production to capital stock and investment in new capital along with the variable inputs is implicit in this formulation and its implications have been discussed by Artus and Muet (1990) in an empirical framework and by Sengupta (1995b) in the DEA framework.

For public sector enterprises however the market prices of output are generally unavailable and the profit maximization objective does not apply, since these are not for profit organizations. Hence in this case we may restrict ourselves to the cost frontier alone and use the theory of adjustment costs to develop a model of capacity utilization. Consider the production function

$$y = f(v, x)$$

of a firm, which produces a single output y by means of the vector v of variable inputs and the vector x of service flows from the quasi-fixed inputs (i.e., these inputs are fixed in the short run but variable in the long run). Since the production function may exhibit increasing returns to scale, the usual profit maximization principle may not yield determinate results. Hence we adopt the cost minimization model, where in the short run the firm minimizes variable costs $q'v$ in the short run subject to the production constraint $y \le f(v, x)$, where x is fixed. This yields the short run cost function $C_v = g(y, q, x)$. Denoting by w the vector of rental prices for the quasi-fixed inputs, the total cost $C = C_v + C_x$ may be defined with $w'x = C_x$ as the fixed cost. Capacity output $\hat{y}$ is now defined by that level of output for which total cost C above is minimized, i.e.,

$$\hat{y} = h(q, x, w) \tag{24}$$

with the associated cost function for capacity output

$$\hat{C} = G(q, w, \hat{y}) \tag{25}$$

Two implications of this concept of optimal capacity output must be noted. One is that the capacity output $\hat{y}$ may be viewed as a point of tangency between the short and the long run average total cost curves. Secondly, one can now define the rate of capacity utilization as $u = y/\hat{y}$ where $0 \le u \le 1$. Morrison and Berndt (1981) used this type of a dynamic cost function model with a single quasi-fixed input called capital to estimate the patterns of capacity utilization of US manufacturing over the period 1958–77 by using a regression model. One can also use a DEA model to specify a cost frontier as follows:

$$\begin{aligned} &\text{Min}\, s_k + \hat{s}_k \\ &\text{s.t.} \quad a_0 + \sum_i a_i q_{ij} + \sum_i b_i x_{ij} + d y_j + s_j = C_j \\ &\qquad \alpha_0 + \sum_i \alpha_i q_{ij} + \sum_i h_i w_{ij} + \delta \hat{y}_j + \hat{s}_j = \hat{C}_j \\ &\qquad j = 1, 2, \ldots, N \end{aligned}$$

where C_j and $\hat{C}_j$ are observed short run and long run costs for unit j and the observed data consist of input and output prices and the two outputs. If unit k is efficient then we must have s_k^* and $\hat{s}_k^*$ to be zero implying full capacity utilization.

If the short and long run cost components can be separately obtained as C_{ij} and $\hat{C}_{ij}$, then these could be used more directly to characterize DEA efficiency as follows:

$$\begin{aligned} &\text{Min}\, \varepsilon + \zeta \\ &\text{s.t.} \quad \sum_{j=1}^{N} C_{ij}\lambda_j \le \varepsilon C_{ik}\,; \quad \sum \lambda_j = 1\,, \quad \lambda_j \ge 0 \\ &\qquad \sum_{j=1}^{N} \hat{C}_{ij}\mu_j \le \zeta \hat{C}_{ik}\,; \quad \sum \mu_j = 1\,, \quad \mu_j \ge 0 \\ &\qquad j = 1, 2, \ldots, N\,. \end{aligned}$$

In this framework unit k is efficient in the short run if $\varepsilon^* = 1.0$, but not efficient in the long run if ζ^* is less than one. The fact that some inputs are fixed in the short run makes it clear that the rate of capacity utilization may influence short and long run costs differently.

From an economic viewpoint the most important source of excess capacity is due to a fall in market demand, i.e., demand uncertainty and the existence of excess capacity tends to inflate the short and the long run costs of output. For public sector enterprises the competitive market pressure is very weak, hence the probability of incurring dead weight losses and hence inefficiency due to organizational slack is much higher.

7. EMPIRICAL APPLICATIONS

Two types of empirical applications to the world airlines industry are considered here by way of illustration. One is the application of a cost competitiveness model in a static DEA framework involving input output data set of 14 airlines averaged over the period 1988–1990 taken from Schefczyk (1993). The second is an application of a dynamic production frontier model formulated in (9), where the annual data set for 11 Latin America based airlines and 6 US based airlines over the 8 year period 1981–88 is

Table 1. Cost efficiency based on the DEA model.

Airline	θ^* (efficiency score)	Input x_1 Actual	Input x_1 Optimal	Input x_2 Actual	Input x_2 Optimal	Input x_3 Actual	Input x_3 Optimal
Air Canada	0.893	5,723	5,111	3,239	2,892	2,003	1,788
AU Nippon	0.844	5,895	4,975	4,225	3,566	4,557	3,846
American	0.948	24,099	22,846	9,560	9,063	6,267	5,941
British Air	0.959	13,565	13,008	7,499	7,191	3,213	3,081
Cathay Pacific	1.000	5,183	5,183	1,880	1,880	783	783
Delta	0.977	19,080	18,641	8,032	7,847	3,272	3,197
Iberia	0.999	4,603	4,598	3,457	3,453	2,360	2,358
Japan	0.859	12,097	10,391	6,779	5,823	6,474	5,561
KLM	0.973	6,587	6,409	3,341	3,251	3,581	3,484
Korean Air	1.000	5,654	5,654	1,878	1,878	1,916	1,916
Lufthansa	1.000	12,559	12,559	8,098	8,098	3,310	3,310
Quantas	1.000	5,728	5,728	2,481	2,481	2,254	2,254
Singapore	1.000	4,715	4,715	1,792	1,792	2,485	2,485
UAL Corporation	1.000	22,793	22,793	9,874	9,874	4,145	4,154

Note: x_1 = available ton kilometer, x_2 = operating cost, x_3 = nonflight assets.

considered. This application was previously considered for measuring technical change by Sengupta (1998c).

To illustrate the first application we have utilized the time series data set over the period 1988–90 for 14 airlines previously utilized by Schefczyk (1993). Each airline has three inputs (x_1, x_2, x_3) and two outputs (y_1, y_2) all measured in logarithmic units. The input output data set exhibits widespread fluctuations for the airline industry due to various regulatory controls and cost uncertainties. One main reason for cost uncertainty is the relative fixity of the capacity-related cost elements, e.g., acquisition of aircraft, development of route systems etc. which have a multiperiod impact on costs and revenues. This is why capital cost is considered an important long run factor in airline operations.

In this application three input costs (x_i) and two output revenues (y_r) are considered as follows: x_1 = available aircraft capacity in ton kilometers, x_2 = operating cost defined as total operating expenses minus rent, depreciation and amortization, x_3 = cost of total nonflight assets, y_1 = passenger kilometer revenue and y_2 = nonpassenger ton-kilometer revenue. Besides these input costs and revenues, the other instrument variables which directly affect airlines efficiency are the following, that are reported by Schefczyk: z_1 = gross profit margin and z_2 = international passenger load.

Table 1 presents the nonparametric estimates of cost efficiency θ_k^* for the 14 airlines with 6 efficient and 8 inefficient. However, since one airline (e.g., Iberia) attains the level 0.999 which can be rounded to 1.00, this one may be included in the efficient set S_1, in which case half of the total is efficient, the other half being inefficient. The least efficient airline is AU Nippon with a value of efficiency score $\theta^* = 0.844$. This implies that this airline would have to reduce its input costs to 84.4 percent of the current level to become efficient.

To compare different airlines belonging to the efficient set S_1, determined by model (2) where each has efficiency score $\theta^* = 1.0$, we estimated the optimal value α_0^* in (8), as given by the dual LP model (7) and the results are as follows:

Table 2. Estimates of returns to scale.

Airline	α_0^*	Returns to scale
Cathay Pacific	0.426	IRS
Lufthansa	0.015	IRS
Singapore	1.264	IRS
Korean Air	0.0	CRS
Quantas	0.327	IRS
UAL	0.0	CRS
Iberia	1.0	IRS

Clearly Singapore Airlines tops the list in terms of the size of IRS and Lufthansa the least, with Korean Air and UAL displaying CRS.

Table 3 presents a comparative view of production function estimates of the two sets S_1, S_2, the efficient and inefficient respectively. Here only the most important output y_1

Table 3. Regression estimates of the linear production function (dependent variable: y_1).

Sample	Intercept	x_1	x_2	x_3	x_4	$\bar{R}^2$
1. Total	−49105	6.91	−3.62	0.39	63.44	0.961
($N = 14$)	($t = -0.86$)	(7.86)	(−1.71)	(0.22)	(0.82)	
2. Efficient set	−78519	8.32	−7.12	5.48	87.80	0.902
($N_1 = 6$)	(−0.39)	(2.94)	(−0.99)	(0.57)	(0.33)	
3. Inefficient set	61609	4.24	1.17	0.55	−92.53	0.977
($N_2 = 8$)	(0.62)	(1.81)	(0.24)	(0.25)	(−0.63)	
4. Total	−4301	5.20	—	—	—	0.956
($N = 14$)	(−1.12)	(16.92)				
5. Efficient set	−4987	5.23	—	—	—	0.902
($N_1 = 6$)	(−0.57)	(6.87)				
6. Inefficient set	−3601	5.16	—	—	—	0.986
($N_2 = 8$)	(−1.18)	(22.39)				
7. Total	−64292	5.59	—	—	80.63	0.957
($N = 14$)	(−1.13)	(11.76)			(1.06)	
8. Total	−70873	5.76	—	−0.86	91.56	0.954
($N = 14$)	(−1.17)	(9.36)		(−0.47)	(1.11)	

Note: $\bar{R}^2$ denotes adjusted R^2, adjusted for degrees of freedom.

Table 4. Efficiency regression on inputs and other instrument variables.

Dependent variable	Intercept	Inputs				Instrument variable		R^2
		x_1	x_2	x_3	x_4	z_1	z_2	
θ^*	0.737	1.00 E-05	−1.26 E-05	01.26 E-05	0.0003	—	—	0.232
	(t = 1.62)	(1.42)	(−0.74)	(−0.84)	(0.52)			
θ^*	0.944	1.64 E-06	—	—	—	—	—	0.042
	(33.46)	(0.73)						
			Inputs and instrument variables in logs					
$\log(100\theta^*)$	4.361	—	—	—	—	0.063	0.008	0.320
	(27.99)					(1.99)	(0.19)	
$\log(100\theta^*)$	4.391	—	—	—	—	0.064	—	0.317
	(89.92)					(2.16)		
$\log y_1$	−1.08	1.611	−1.371	0.894	—	—	—	0.496
	(−0.32)	(2.013)	(−1.551)	(1.679)				

Note: z_1 = gross profit margin, z_2 = volume of international passenger demand, y_1 = passenger revenue; E-05 = 10^{-5}.

is considered as the dependent variable; also x_4 is added as an extra explanatory variable representing passenger load factors. Three interesting points come out very clearly. One is that the capacity variable x_1 emerges as the major explanatory variable; other explanatory variables x_2, x_3, x_4 have either insignificant coefficients or wrong signs. Secondly, the intercept term for the efficient set S_1 is always negative, thus implying IRS in a consistent fashion. Finally, we tested by Chow test the difference in the coefficients between sets S_1 and S_2 for two cases: four inputs (x_1 through x_4) and one input (x_1) and the results are: 0.318 ($F_{6,2} = 5.14$ at 5% level) and 0.021 ($F_{2,10} = 7.56$ at 5% level) respectively. This implies that the null hypothesis that the two coefficient structures are equal is not rejected at 5% level. Note that the sample sizes ($N_1 = 6$, $N_2 = 8$) are quite small here and this may bias the tests. The difference in the average DEA efficiency scores is however $1 - 0.931 = 0.049$ if Iberia is included in the efficient set and $1 - 0.922 = 0.078$ if it is not. Thus the DEA estimate is more discriminating than the least squares estimate.

Finally, we have in Table 4 the estimated results on the possible sources of efficiency, where the efficiency measure (θ^*) or its log equivalent is regressed on the four inputs and two instrument variables z_1 and z_2 representing gross profit margin and international passenger demand. Only gross profit margin (z_1) and the capacity input (x_1) turned out to be positively correlated with efficiency score, but only the gross profit margin has a significant coefficient at 5% level of t test, when z_1 alone is used as the explanatory variable. This suggests that the profit margin alone does not indicate a measure of higher efficiency, i.e., there is a tradeoff of short run profits to other goals like retaining market share and the competitive edge in international air travel market.

To illustrate the second application we have used the airline panel data set from Cooper and Gallegos (1992). Here output is measured by ton-kilometers performed and the thee inputs are: labor, fuel and capacity. Unlike the previous data set from Schefczyk (1993) this data set did not have detailed unit costs such as operating expenses net of rents and depreciation and the cost of nonflight assets. Sengupta (1998c) used this data set previously in terms of DEA models for analyzing changes in efficiency over time and also for filtering of systematic efficiency measures.

The capacity variable here (x_3) is measured by available capacity in terms of ton-kilometers. Labor (x_1) is measured by the volume of employment and fuel (x_2) by expenditure in US dollars. The inputs and output are all measured in logarithmic units. The following production function in log units is estimated by ordinary least squares over the period 1981–88 as a whole:

$$y = \alpha_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

The results are as follows:

Table 5. Least squares estimates of production function.

Airline	α_0	β_1	β_2	β_3	$\bar{R}^2$
US	−0.691	−0.006	−0.110	1.165	0.996
	(t = −6.14)	(−0.22)	(−2.98)	(28.22)	
Latin American	1.632	0.723	−0.005	0.224	0.975
	(9.41)	(12.20)	(0.095)	(3.54)	

where $\bar{R}^2$ is squared multiple correlation coefficient adjusted for degrees of freedom and the t-values are in parentheses below each regression coefficient. It is clear that the capacity input (x_3) is the most significant of all the inputs over the whole period. Since the sum ($\beta_1 + \beta_2 + \beta_3$) denotes the scale of returns, it is clear the US airlines exhibit IRS, while the Latin airlines exhibit DRS. The output elasticity of the capacity input for the US airlines is more than five times that of the Latin airlines. The changes in efficiency ranking measured by the optimal score θ^* defined in model (2) may be shown in two ways. One is the trend in values of θ_t^* for selected airlines for three subperiods 1981–88, 1983–88 and 1985–88. The other is in terms of changes in the parameters values α_0^*, β_1^*, β_2^*, β_3^* of a typical airlines in the efficient and inefficient category, e.g., American and Varig. Two points clearly emerge. One is that the airlines do not usually retain their 100% efficiency in every year, although some airlines maintain their efficiency score on or above 96%. This is clearly exhibited by the results in Table 7. If one estimates the proportion of efficient units for the three subperiods above the results are as follows:

Table 6. Proportion of efficient airlines.

Time average	All inputs	Capacity omitted
1981–88	70	80
1983–88	40	60
1985–88	20	50

Table 7. Trend of efficiency score (θ^*).

Airlines	1981–88	1983–88	1985–88
Amex	1.000	0.986	0.844
Amer	1.000	0.992	0.963
Varig	0.980	0.969	0.981
Arge	0.975	0.817	0.909
Mexi	0.966	0.944	0.830
Eastern	0.960	0.864	0.955
Delta	0.838	0.678	0.633
Peru	0.802	0.531	0.556
Conti	0.775	0.643	0.638

Table 8. Change in production frontier parameters for Varig and Amer.

	α_0^*	β_1^*	β_2^*	β_3^*
1981 (Varig)	2.226	0.385	0.485	0.135
(Amer)	(1.912)	(0.012)	(0.002)	(0.994)
1983	0.943	0.266	0.337	0.431
	(0.065)	(0.024)	(0.045)	(0.987)
1985	−1.548	0.001	0.192	0.917
	(−0.951)	(0.000)	(0.051)	(1.090)
1987	−0.554	0.272	0.001	0.749
	(−0.102)	(0.058)	(0.010)	(1.012)
1988	0.064	0.246	0.005	0.800
	(0.012)	(0.079)	(0.041)	(1.004)

Note: The values for Amer are in parentheses.

Clearly the omission of the capacity input increases the proportion of efficient airlines, which implies that the capacity is not fully utilized in many cases.

Table 8 shows very clearly the dominance of the capacity utilization factor in generating increasing returns to scale for the American airlines. The output elasticity of the capital input has increased over the years in case of Varig Airlines, although American airlines have always outperformed Varig very consistently.

8. CONCLUDING REMARKS

The impact of demand and cost uncertainty on the efficiency evaluation by the Pareto criterion is investigated here in respect of the recent technique of data envelopment analysis. The role of organizational slack in the cost frontier and the influence of excess capacity of the production frontier are discussed here both theoretically and empirically. Some empirical applications to international airlines industry serve to illustrate the dominant role of the capacity variable, which changes in the long run and thereby affect the returns to scale in the production frontier.

REFERENCES

Artus, P. and Muet, P., *Investment and Factor Demand*, Amsterdam, North Holland Publishing (1990).

Charnes, A., Cooper. W. W., Lewin, A., and Seiford, L. M., *Data Envelopment Analysis: Theory, Methodology and Applications*, Dordrecht, Kluwer Academic Publishers (1994).

Cooper, W. W., and Gallegos, A., "A combined DEA-stochastic frontier approach to Latin American airline efficiency evaluations", Working Paper, Univ. of Texas at Austin, 1992.

Cyert, R. M., and March, J. G., *A Behavioral Theory of the Firm*, Englewood Cliffs, Prentice Hall (1963).

Farrell, M. J., "The measurement of productive efficiency", *Journal of Royal Statistical Society*, Series A, Vol. 120 (1957), 253–290.

Foss, M. F., "The utilization of capital equipment: postwar compared with prewar", *Survey of Current Business*, Vol. 43 (1963), 8–16.

Fried, H., Lovell, C., and Schmidt, S., *The Measurement of Productive Efficiency*, New York, Oxford Univ. Press (1993).

Ganley, J., and Cubbin, J., *Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis*, Amsterdam, North Holland (1992).

Hodrick, R., and Prescott, "Post-war US business cycles: an empirical investigation", *Journal of Money, Credit and Banking*, Vol. 29 (1997), 1–16.

Kennedy, J. E., "An analysis of time series estimates of capacity utilization", Discussion Paper No. 95-37, Federal Research Board, Washington, D.C. (1995).

Klein, L. R., and Long, V., "Capacity utilization: concept, measurement and recent estimates", *Brookings Papers on Economic Activity*, Vol. 3 (1973), 743–761.

Leibenstein, H., "Allocative efficiency versus X-efficiency", *American Economic Review*, Vol. 66 (1966), 392–415.

Leibenstein, H., *Beyond Economic Man*, Cambridge, Harvard Univ. Press (1976).

Morrison, C. J., and Berndt, E. R., "Short-run labor productivity in a dynamic model", *Journal of Econometrics*, Vol. 16 (1981), 339–365.

Oum, T. H. and Yu, C., *Winning Airlines: Productivity and Cost Competitiveness of the World's Major Airlines*, Dordrecht, Kluwer Academic Publishers (1998).

Schefczyk, M., "Operational performance of airlines: an extension of traditional measurement paradigms", *Strategic Management Journal*, Vol. 14 (1993), 301–317.

Selten, R., "Elementary theory of slack-ridden imperfect competition". In *New Developments in the Analysis of Market Structure*, Cambridge, Massachusetts: MIT Press (1986).

Sengupta, J. K., "A dynamic efficiency model using data envelopment analysis", *International Journal of Production Economics*, Vol. 62 (1998), 209–218.

Sengupta, J. K., "A model of dynamic efficiency measurement", *Applied Economics Letters*, Vol. 1 (1994), 119–121.

Sengupta, J. K., *Dynamics of Data Envelopment Analysis: Theory of Systems Efficiency*, Dordrecht, Kluwer (1995a).

Sengupta, J. K., "Models of dynamic efficiency in the public sector", *Keio Economic Studies*, Vol. 32 (1995b), 37–46.

Sengupta, J. K., "Recent models in data envelopment analysis: theory and applications", *Applied Stochastic Models and Data Analysis*, Vol. 12 (1996), 1–26.

Sengupta, J. K., "New efficiency theory: extensions and new applications of data envelopment analysis", *International Journal of Systems Science*, Vol. 29 (1998a), 255–265.

Sengupta, J. K., "Stochastic data envelopment analysis: a new approach", *Applied Economics Letters*, Vol. 5 (1998b), 287–290.

Sengupta, J. K., "Testing for technical change: a nonparametric application", *Applied Economics Letters*, Vol. 5 (1998c), 43–46.

Sengupta, J. K., "A dynamic efficiency model using data envelopment analysis", *International Journal of Production Economics*, Vol. 62 (1999), 209–218.

Winston, G. C., "The theory of capital utilization and idleness", *Journal of Economic Literature*, Vol. 12 (1974), 28–40.