9 DATA VARIATIONS

9.1 INTRODUCTION

We have now covered a wide variety of topics ranging from straightforward DEA models and their uses and extending to modifications such as are incorporated in assurance regions or in the treatment of variables with values that have been exogenously fixed to varying degrees. This, however, does not end the possibilities. New uses of DEA with accompanying new developments and extensions continue to appear.

An attempt to cover all of these topics would be beyond the scope of this text — and perhaps be impossible of achievement as still more uses and new developments continue to appear. Hence the course we follow is to provide relatively brief introductions to some of these topics. This will include sensitivity analysis, stochastic-statistical characterizations and probabilistic formulations with their associated approaches. Dynamic extensions of DEA will also be indicated in the form of window analysis.

9.2 SENSITIVITY ANALYSIS

9.2.1 Degrees of Freedom

The topic of sensitivity (= stability or robustness) analysis has taken a variety of forms in the DEA literature. One part of this literature studies the responses when DMUs are deleted or added to the set being considered or when outputs or inputs are added or withdrawn from consideration. See Wilson (1995).[1] See also the discussion of "window analysis" later in this chapter. Another part of this literature deals with the increases or decreases in the number of inputs and outputs to be treated. Analytically oriented treatments of these topics are not lacking[2] but most of this literature has taken the form of simulation studies, as in Banker *et al.* (1996).[3]

Comment : As in statistics or other empirically oriented methodologies, there is a problem involving degrees of freedom, which is compounded in DEA because of its orientation to *relative* efficiency. In the envelopment model, the number of degrees of freedom will increase with the number of DMUs and decrease with the number of inputs and outputs. A rough rule of thumb which can provide guidance is as follows.

$$n \geq \max\{m \times s, 3(m+s)\}$$

where n= number of DMUs, m= number of inputs and s = number of outputs.

9.2.2 Algorithmic Approaches

Attention to this topic of sensitivity analysis in DEA was initiated in Charnes *et al.* (1985)[4] which built on the earlier work in Charnes and Cooper (1968)[5] after noting that variations in the data for the DMU_o under evaluation could alter the inverse matrix used to generate solutions in the usual simplex algorithm computer codes. (See expressions (3.52)-(3.53) in Chapter 3.) Proceeding along the path opened by the latter publication (by Charnes and Cooper) this work is directed to the use of algorithms that avoid the need for additional matrix inversions. Originally confined to treating a single input or output this line of work was extended and improved in a series of papers published by Charnes and Neralic.[6]

We do not pursue these algorithmic approaches here. We turn instead to other approaches where attention is confined to transitions from efficient to inefficient status.

9.2.3 Metric Approaches

Another avenue for sensitivity analysis opened by Charnes *et al.* (1992)[7] bypasses the need for these kinds of algorithmic forays by turning to metric concepts. The basic idea is to use concepts such as "distance" or "length" (= norm of a vector) in order to determine "radii of stability" within which the occurrence of data variations will not alter a DMU's classification from efficient to inefficient status (or *vice versa*).

The resulting classifications can range from "unstable" to "stable" with the latter identified by a radius of some finite value within which no reclassification will occur. Points like E or F in Figure 9.1 provide examples identified as stable. A point like A, however, is unstable because an infinitesimal perturbation to the left of its present position would alter its status from inefficient to efficient.

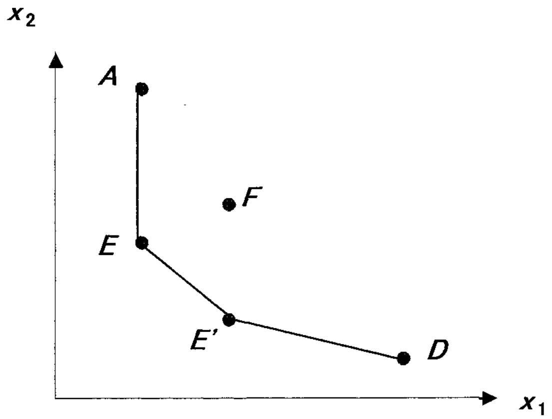

Figure 9.1. Stable and Unstable DMUs

A variety of metrics and models are examined but here attention will be confined to the Chebychev ($= l_\infty$) norm, as in the following model taken from Charnes, Haag, *et al.* (1992, p.795),[7]

$$\max\ \delta \tag{9.1}$$

$$\text{subject to} \quad y_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j - s_r^+ - \delta d_r^+, \quad r = 1, \ldots, s$$

$$x_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j + s_i^- + \delta d_i^-, \quad i = 1, \ldots, m$$

$$1 = \sum_{j=1}^{n} \lambda_j$$

with all variables (including δ) constrained to be nonnegative while the d_r^+ and d_i^- are fixed constants (to serve as weights) which we now set to unity.

With all $d_i^- = d_r^+ = 1$ the solution to (9.1) may be written

$$\sum_{j=1}^{n} y_{rj}\lambda_j^* - s_r^{+*} = y_{ro} + \delta^*, \quad r = 1, \ldots, s \tag{9.2}$$

$$\sum_{j=1}^{n} x_{ij}\lambda_j^* + s_i^{-*} = x_{io} - \delta^*, \quad i = 1, \ldots, m$$

This shows that we are improving all outputs and inputs to the maximum that this metric allows consistent with the solution on the left.

The formulation in (9.2) is for an inefficient DMU which continues to be inefficient for all data alteration from y_{ro} to $y_{ro} + \delta^*$ and from x_{io} to $x_{io} - \delta^*$. This is interpreted to mean that no reclassification to efficient status will occur within the open set defined by the value of $0 \leq \delta^*$ — which is referred to as a "radius of stability." See, for example, the point F in Figure 9.2 which is centered in the square (or box) defined by this C (=Chebyshev) norm which is referred to as a "unit ball."[8]

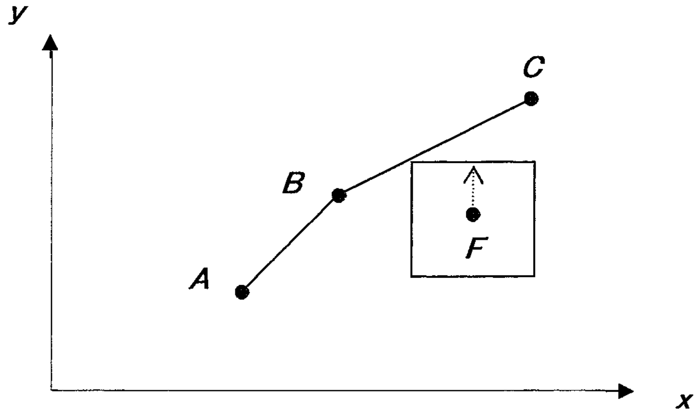

Figure 9.2. A Radius of Stability

The above model dealt with improvements in both inputs and outputs that could occur for an inefficient point before its status would change to efficient — as in the upper left hand corner of the box surrounding F in Figure 9.2. The treatment of efficient points proceeds in the direction of "worsening" outputs and inputs as in the following model.

$$\min \ \delta \tag{9.3}$$

$$\text{subject to} \quad y_{ro} = \sum_{j=1, j\neq o}^{n} y_{rj}\lambda_j - s_r^+ - \delta, \quad r = 1, \ldots, s$$

$$x_{io} = \sum_{j=1, j\neq o}^{n} x_{ij}\lambda_j + s_i^- + \delta, \quad i = 1, \ldots, m$$

$$1 = \sum_{j=1, j\neq o}^{n} \lambda_j$$

where, again, all variables are constrained to be nonnegative.

In this case $j\neq o$ refers to the efficient DMU_o that is being analyzed. Otherwise, as in the following definition, the result will always be unstable.[9]

Definition 9.1 *The coordinates of the point associated with an efficient DMU will always have both efficient and inefficient points within a radius of* $\varepsilon > 0$ *however small the value of* ε.

Definition 9.2 *Any point with the above property is unstable.*

To see that this property is not confined to points associated with efficient DMUs, note that A in Figure 9.1 has this property since a slight variation to the left will change its status from inefficient to efficient. In any case, a solution, δ^*, provides a radius in the Chebychev norm that is to be attained before an efficient DMU is changed to inefficient status.

To see what is happening in the case of an efficient point refer to B in Figure 9.2. The radius of stability for this point would be determined by the "worsenings" allowed in (9.3) until the line connecting A and C is reached. This follows from the fact that worsenings which only move B to a new point which is on the left of this line will not affect its efficiency status.

Comment : The above formulations are recommended only as the *start* for a sensitivity analysis by Charnes *et al.*[10] because, *inter alia*, this norm does not reflect any nonzero slacks which may be present.[11] It might be supposed that omitting the DMU to be evaluated from the right-hand side of (9.3) could lead to non-solution possibilities. This is not the case. Solutions to (9.3) always exist, as is proved in W.W. Cooper, S. Li, L.M. Seiford, K. Tone, R.M. Thrall and J. Zhu (2001) "Sensitivity and Stability Analysis in DEA: Some Recent Developments," *Journal of Productivity Analysis* 15, pp.217-246. See also L.M. Seiford and J. Zhu (1998) "Sensitivity Analysis of DEA Models for Simultaneous Changes in All of the Data," *Journal of the Operational Research Society* 49, pp.1060-1071 as well as Seiford and Zhu (1999) "Infeasibility of Super-Efficiency Data Envelopment Analysis Models," *INFOR* 37, pp.174-187.

9.2.4 Multiplier Model Approaches

The above approaches treat one DMU at a time. However, this needs to be extended for treating cases where the DMUs are numerous and it is not clear which ones require attention. Ideally it should be possible to vary all data simultaneously until the status of at least one DMU is changed from inefficient to efficient or *vice versa.* A third approach initiated by R.G. Thompson and R.M. Thrall[12] and their associates moves in this direction in a manner that we now describe.

For this purpose we record the following dual pair from Thompson *et al.* (1996).[13]

Envelopment Model	Multiplier Model
$\text{minimize}_{\theta,\lambda}\ \theta$	$\text{maximize}_{\boldsymbol{u},\boldsymbol{v}}\ z = \boldsymbol{u}\boldsymbol{y}_o$
subject to	subject to
$Y\boldsymbol{\lambda} \geq \boldsymbol{y}_o$	$\boldsymbol{u} \geq \mathbf{0}$
$\theta\boldsymbol{x}_o - X\boldsymbol{\lambda} \geq \mathbf{0}$	$\boldsymbol{v} \geq \mathbf{0}$
$\boldsymbol{\lambda} \geq \mathbf{0}$	$\boldsymbol{u}Y - \boldsymbol{v}X \leq \mathbf{0}$
θ unrestricted	$\boldsymbol{v}\boldsymbol{x}_o = 1,$

(9.4)

where Y, X and $\boldsymbol{y}_o, \boldsymbol{x}_o$ are data matrices and vectors of outputs and inputs, respectively, and $\boldsymbol{\lambda}, \boldsymbol{u}, \boldsymbol{v}$ are vectors of variables ($\boldsymbol{\lambda}$: a column vector, $\boldsymbol{u}$ and $\boldsymbol{v}$: row vectors). θ, a scalar, which can be positive, negative or zero in the envelopment model is the source of the condition $\boldsymbol{v}\boldsymbol{x}_o = 1$ in the multiplier model.

No allowance for nonzero slacks is made in the objective of the above envelopment model. Hence the variables in the multiplier model are constrained only to be nonnegative. That is, the positivity requirement associated with the non-Archimedean element, ε, is absent from both members of this dual pair. Thompson *et al.* refer to Charnes, Cooper and Thrall (1991) [14] to justify this omission of non-Archimedean elements. For present purposes, however, we note only that these sensitivity analyses are centered around the set, E, of efficient extreme points and these points always have a unique optimum with nonzero slacks.

We also note that the analysis is carried forward via the multiplier models[15] by Thompson, *et al.* This makes it possible to exploit the fact that the values $\boldsymbol{u}^*, \boldsymbol{v}^*$ which are optimal for the DMU being evaluated will remain valid over some (generally positive) range of variation in the data.[16]

Following Thompson, *et al.* we try to exploit this property by defining a new vector $\boldsymbol{w} = (\boldsymbol{u}, \boldsymbol{v})$ and a function $h_j(\boldsymbol{w})$ as follows

$$h_j(\boldsymbol{w}) = \frac{f_j(\boldsymbol{w})}{g_j(\boldsymbol{w})} = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}}. \tag{9.5}$$

Next, let

$$h_o(\boldsymbol{w}) = \max_{j=1,\ldots,n} h_j(\boldsymbol{w}) \tag{9.6}$$

so that

$$h_o(\boldsymbol{w}) \geq h_j(\boldsymbol{w}) \quad \forall j. \tag{9.7}$$

It is now to be noted that (9.5) returns matters to the CCR ratio form which was introduced as early as Chapter 2, Section 2.2. Hence we need not be concerned with continued satisfaction of the norm condition, $\boldsymbol{v}\boldsymbol{x}_o = 1$ in (9.4), as we study variations in the data.

When an *optimal* $\boldsymbol{w}^*$ does not satisfy (9.7), the DMU$_o$ being evaluated is said to be "radial inefficient." The term is appropriate because this means that $\theta^* < 1$ will occur in the envelopment model. The full panoply of relations between the CCR ratio, multiplier and envelopment models is thus brought into play without any need for extensive computations or analyses.

Among the frontier points for which $\theta^* = 1$, attention is directed by Thompson *et al.* to "extreme efficient points" — i.e., points in the set E which, for some multiplier $\boldsymbol{w}^*$,

$$h_o(\boldsymbol{w}^*) > h_j(\boldsymbol{w}^*) \quad \forall j \neq o. \tag{9.8}$$

This (strict) inequality will generally remain valid over some range of variation in the data. Hence, in more detail we will have

$$h_o(\boldsymbol{w}^*) = \frac{\sum_{r=1}^{s} u_r^* y_{ro}}{\sum_{i=1}^{m} v_i^* x_{io}} > \frac{\sum_{r=1}^{s} u_r^* y_{rj}}{\sum_{i=1}^{m} v_i^* x_{ij}} = h_j(\boldsymbol{w}^*) \quad \forall j \neq o, \tag{9.9}$$

which means that DMU_o is more efficient than any other DMU_j and hence will be rated as fully efficient by DEA.

Thompson, *et al.* employ a ranking principle which they formulated as: "If DMU_o is more efficient than any other DMU_j *relative* to the vector $\boldsymbol{w}^*$, then DMU_o is said to be top ranked." Holding $\boldsymbol{w}^*$ fixed, the data are then varied and DMU_o is said to be "top ranked" if (9.8) continues to hold.

Thompson, *et al.* allow the data to vary in several different ways which include allowing the data variations to occur at random. Among these possibilities we examine only the following one. For DMU_o, which is extreme efficient, the outputs are all decreased and the inputs are all increased by a stipulated amount (or percentage). This same treatment is accorded to the other DMUs which are efficient. For the other DMU_j, the reverse adjustment is made: All outputs are increased and all inputs are decreased in these same amounts (or percentages). In this way the ratio in (9.8) for DMU_o will be decreased along with the other extreme efficient DMUs while the ratios for the other DMU_j will be increased. Continuing in this manner a reversal can be expected to occur at some point — in which case DMU_o will no longer be "top ranked" and it will then lose the status of being fully (DEA) efficient.

Table 9.1 taken from Thompson *et al.* (1994) will be used to illustrate the procedure in a simple manner by varying only the data for the inputs x_1, x_2 in this table. To start this sensitivity analysis, Table 9.2 records the initial

Table 9.1. Data for a Sensitivity Analysis

	E-Efficient*			Not Efficient		
DMU	1	2	3	4	5	6
Output: y	1	1	1	1	1	1
Input: x_1	4	2	1	2	3	4
Input: x_2	1	2	4	3	2	4

* E-Efficient = Extreme Point Efficient

solutions by applying the multiplier model for (9.8) to these data for each of DMU_1, DMU_2 and DMU_3 which are all extreme point efficient.[17] As can be

Table 9.2. Initial Solutions

DMU	DMU_1 $h_j(w^1)$	DMU_2 $h_j(w^2)$	DMU_3 $h_j(w^3)$
1	1.000	0.800	0.400
2	0.714	1.000	0.714
3	0.400	0.800	1.000
4	0.500	0.800	0.667
5	0.667	0.800	0.550
6	0.357	0.500	0.357

seen these solutions show DMU_1, DMU_2 and DMU_3 to be top ranked in their respective columns.

The gaps between the top and other ranks from these results show that some range of data variation can be undertaken without changing this top-ranked status. To start we therefore introduce 5% increases in each of x_1 and x_2 for DMU_1, DMU_2 and DMU_3 and thereby worsen their performances. Simultaneously we decrease these inputs by 5% for the other DMUs to obtain Table 9.3.

Table 9.3. Results of 5% Data Variations

DMU	DMU_1 $h_j(w^1)$	DMU_2 $h_j(w^2)$	DMU_3 $h_j(w^3)$
1	0.952	0.762	0.381
2	0.680	0.952	0.680
3	0.381	0.762	0.952
4	0.526	0.842	0.702
5	0.702	0.842	0.526
6	0.376	0.526	0.376

The values of the $h_j(w)$ resulting from these data variations are portrayed in Table 9.3. As can be seen, each of DMU_1, DMU_2 and DMU_3 maintain their "top ranked status" and hence continue to be DEA fully efficient (relatively). Nor is this the end of the line. Continuing in this 5% increment-decrement fashion, as Thompson, *et al.* (1994) report, a 15% increment-decrement is needed for a first displacement in which DMU_2 is replaced by DMU_4 and DMU_5. Continuing further, a 20% increment-decrement is needed to replace DMU_1 with DMU_5 and, finally, still further incrementing and decrementing is needed to replace DMU_3 with DMU_4 as top ranked.

Comment: Note that the $h_j(\boldsymbol{w})$ values for all of the efficient DMUs decrease in every column when going from Table 9.2 to Table 9.5 and, simultaneously, the $h_j(\boldsymbol{w})$ values increase for the inefficient DMUs. The same behavior occurs for the other data variations, including the random choices of data changes, used by Thompson, Thrall and their associates in other studies. As noted on page 401 of Thompson *et al.* (1994) this robust behavior is obtained for extreme efficient DMUs which are identified by their satisfaction of the Strong Complementary Slackness Condition (described in Section A.8 of our Appendix A) for which a gap will appear like ones between the top and second rank shown in every column of Table 9.2. In fact, the choice of $\boldsymbol{w}^*$ can affect the degree of robustness as reported in Thompson *et al.* (1996) where use of an interior point algorithm produces a $\boldsymbol{w}^*$ closer to the analytic center and this considerably increases the degree of robustness for the above example. For a more detailed treatment see Cooper, Li, Seiford and Zhu. [18]

9.3 STATISTICAL APPROACHES

Treatment of data variations by statistical methods has taken a variety of forms in DEA and related literatures. More precisely, Banker (1993)[19] and Banker and Natarasan (2004)[20] show that DEA provides a consistent estimator of arbitrary monotone and concave production functions when the (one-sided) deviations from such a production function are regarded as stochastic variations in technical inefficiency.[21] Convergence is slow, however, since, as is shown by Korostolev *et al.* (1995),[22] the DEA likelihood estimator in the single output - m input case converges at the rate $n^{-2/(1+m)}$ and no other estimator can converge at a faster rate.[23]

The above approaches treat only the single output - multiple input case. Simar and Wilson (1998)[24] turn to "bootstrap methods" which enable them to deal with the case of multiple outputs and inputs. In this manner, the sensitivity of θ^*, the efficiency score obtained from the BCC model, can be tested by repeatedly sampling from the original samples. A sampling distribution of θ^* values is then obtained from which confidence intervals may be derived and statistical tests of significance developed.

All of this work represents significant new developments. More remains to be done, however, since neither Banker nor Simar and Wilson make any mention of how to treat nonzero slacks. Thus, it is not even clear that they are estimating efficiency frontiers.

Another line of research proceeds through what are referred to as "stochastic frontier regressions." This line of work has a longer history which can (in a sense) by traced all the way back to Farrell (1957).[25] Subsequently extended by Aigner and Chu (1968)[26] this approach was given its first statistical formulation in Aigner, Lovell and Schmidt (1977) in a form that is now called the "composed error" approach.[27]

To see what is involved in this "composed error" approach we start with the usual formulation of a statistical regression model as in

$$y = f(\boldsymbol{x}) + \varepsilon. \tag{9.10}$$

Here $f(\boldsymbol{x})$ is a prescribed (known) function with parameters to be estimated and ε represents random errors which occur in the dependent (regressand) variable, a scalar, and not in the independent (regressor) variables represented by the vector $\boldsymbol{x}$. The components of $\boldsymbol{x}$, we emphasize, are assumed to be known without error.

The concept of a "composed error" is represented by replacing ε with a 2-component term which we can represent as

$$\varepsilon = \nu - \tau. \tag{9.11}$$

Here ν represents the random error component which may be positive, negative or zero while τ is restricted to nonnegative ranges that represent values of y that fail to achieve the efficient frontier. The term $\tau \geq 0$ is usually assumed to have statistical distributions, such as the exponential or half normal, which are confined to nonnegative ranges that represent inefficiencies.

Following Jondrow, Lovell, Materov and Schmidt (1982),[28] the estimates of technical efficiency are obtained from

$$\widehat{\tau} = -\left[\mu_\tau - \sigma\frac{f^*(\mu_\tau/\sigma)}{F^*(-\mu_\tau/\sigma)}\right] \tag{9.12}$$

where

$$\mu_\tau = \frac{\sigma_\tau^2}{\sigma_\nu^2 + \sigma_\tau^2} \quad \text{and} \quad \sigma^2 = \frac{\sigma_\nu^2\sigma_\tau^2}{\sigma_\nu^2 + \sigma_\tau^2}$$

and where $f^*(\cdot)$ and $F^*(\cdot)$ represent the standard normal density and cumulative normal distribution functions, respectively, with mean μ and variance σ^2. The efficiency corresponding to specified values for the components of $\boldsymbol{x}$ are then estimated from

$$0 \leq e^{-\widehat{\tau}} \leq 1 \tag{9.13}$$

which is equal to unity when $\widehat{\tau} = 0$ and becomes 0 as $\widehat{\tau} \to \infty$.

To see how this measure of efficiency is to be used we employ (9.10) and (9.11) in the following simple (two-input) version of a log-linear (=Cobb-Douglas) production function

$$y = \beta_0 x_1^{\beta_1} x_2^{\beta_2} e^{\varepsilon} = \beta_0 x_1^{\beta_1} x_2^{\beta_2} e^{\nu-\tau} \tag{9.14}$$

so that

$$ye^{\tau} = \beta_0 x_1^{\beta_1} x_2^{\beta_2} e^{\nu}. \tag{9.15}$$

Hence $ye^{\tau} = \widehat{y}$ with $\widehat{y} \geq y$ is estimated stochastically with inefficiency embodied in an output shortfall and not in overuses of either input.

It is possible to view these stochastic frontier regressions as competing with DEA and to study them from this standpoint as is done in Gong and Sickles

(1990),[29] for example, who bring a simulation approach to this task. Carried to an extreme the two approaches, DEA vs. Stochastic Frontier Regressions, can be regarded as mutually exclusive — as in Schmidt (1985).[30] An alternative view is also possible in which the two approaches can be used in complementary fashion. Ferrier and Lovell (1990),[31] for example, use the two approaches to cross-check each other. In this approach, the objective is to avoid what Charnes, Cooper and Sueyoshi (1988)[32] refer to as "methodological bias" when large issues of policy are being addressed. Indeed, it is possible to go a step further and join the two approaches in a common effort as in the example we now discuss.[33]

Arnold *et al.* (1994)[34] describe an experience in which the use of Cobb-Douglas regressions yielded unsatisfactory results in an attempt to use this kind of regression approach in a study conducted under legislative mandate to develop methods for evaluating the performances of public schools in Texas. Using this same functional form, however, and applying it to the same body of data, Arnold *et al.* reported satisfactory results from a two-stage DEA-regression approach which proceeded in the following manner: In stage one all of the 640 schools in this study were submitted to treatment by DEA. The original Cobb-Douglas form was then extended to incorporate these results in the form of "dummy variables." In this approach a school which had been found to be DEA efficient was associated with a value of unity for the dummy variables. A school which had been found to be DEA inefficient was assigned a value of zero. The regression was then recalculated and found to yield very satisfactory results.

The above study was followed by a simulation experiment which we now review for added insight.[35] For this purpose we replaced (9.10) with

$$y = 0.75 x_1^{0.65} x_2^{0.55} e^{\varepsilon}. \tag{9.16}$$

In short, the above expression is used to generate all observations with the Cobb-Douglas form having known parameter values

$$\begin{aligned} \beta_0 &= 0.75 \\ \beta_1 &= 0.65 \\ \beta_2 &= 0.55 \end{aligned} \tag{9.17}$$

and e^{ε} is used to generate random variables which are then used to adjust the thus generated y values to new values which contain these random terms. This procedure conforms to the assumptions of both SF (=Stochastic Frontier) and OLS (=Ordinary Least Squares) regression uses.

The input values for x_1 and x_2 in (9.16) are generated randomly, as a bias avoiding mechanism, and these values are inserted in (9.16) to provide the truly efficient values of y after which the values of $\widehat{y}$ defined in (9.15) are then generated in the previously described manner.

The inputs are then adjusted to new values

$$\widehat{x}_1 = x_1 e^{\tau_1} \quad \text{and} \quad \widehat{x}_2 = x_2 e^{\tau_2} \quad \text{with } \tau_1, \tau_2 \geq 0 \tag{9.18}$$

where τ_1 and τ_2 represent input-specific technical inefficiencies drawn at random to yield the corresponding input inefficiencies embedded in $\widehat{x}_1$ and $\widehat{x}_2$.

This procedure, we note, violates the SF assumption that all inefficiencies are impounded only in the regressand, y. See the discussion immediately following (9.15). It also violates OLS since these $\widehat{x}_1$, $\widehat{x}_2$ are not the input amounts used to generate the y values. Nevertheless it reproduces a situation in which the observed y (or $\widehat{y}$) will tend to be too low for these inputs. Finally, to complete the experimental design, a subset of the observations, chosen at random, used the original x_1, x_2 values rather than the $\widehat{x}_1$, $\widehat{x}_2$ generated as in (9.18). This was intended to conform to the assumption that some DMUs are wholly efficient and it also made it possible (a) to determine whether the first-stage DEA identified the efficient DMUs in an approximately correct manner as well as (b) to examine the effects of such an efficient subset on the derived statistical estimates.

Further details on the experimental design may be found in Bardhan *et al.* (1998).

We therefore now turn to the estimating relations which took a logarithmic form as follows,

$$\ln y = \ln \beta_0 + \beta_1 \ln \widehat{x}_1 + \beta_2 \ln \widehat{x}_2 \tag{9.19}$$

and

$$\ln y = \ln \beta_0 + \beta_1 \ln \widehat{x}_1 + \beta_2 \ln \widehat{x}_2 + \delta D + \delta_1 D \ln \widehat{x}_1 + \delta_2 D \ln \widehat{x}_2 + \varepsilon \tag{9.20}$$

Table 9.4. OLS Regression Estimates without Dummy Variables

Case 1: *EXPONENTIAL* distribution of input inefficiencies

Parameter Estimates	*Case A* $\sigma_\varepsilon^2 = 0.04$ (1)	*Case B* $\sigma_\varepsilon^2 = 0.0225$ (2)	*Case C* $\sigma_\varepsilon^2 = 0.01$ (3)	*Case D* $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.30*	1.58*	1.40*	1.43*
	(0.19)	(0.15)	(0.13)	(0.10)
β_1	0.46*	0.43*	0.45*	0.46*
	(0.024)	(0.02)	(0.016)	(0.013)
β_2	0.48*	0.47*	0.47*	0.46*
	(0.02)	(0.013)	(0.01)	(0.01)

The asterisk "*" denotes statistical significance at 0.05 significance level or better. Standard errors are shown in parentheses.
The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions.

Table 9.5. Stochastic Frontier Regression Estimates without Dummy Variables

Case 1: *EXPONENTIAL* distribution of input inefficiencies

Parameter Estimates	*Case A* $\sigma_\varepsilon^2 = 0.04$ (1)	*Case B* $\sigma_\varepsilon^2 = 0.0225$ (2)	*Case C* $\sigma_\varepsilon^2 = 0.01$ (3)	*Case D* $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.42*	1.62*	1.25*	1.28*
	(0.19)	(0.14)	(0.14)	(0.11)
β_1	0.46*	0.43*	0.48*	0.46*
	(0.024)	(0.02)	(0.017)	(0.01)
β_2	0.48*	0.47*	0.48*	0.47*
	(0.017)	(0.013)	(0.01)	(0.01)
σ_τ	0.15*	0.11*	0.15*	0.15*
	(0.035)	(0.01)	(0.01)	(0.01)
σ_ν	0.15*	0.13*	0.08*	0.04
	(0.01)	(0.02)	(0.01)	(0.025)

The asterisk "*" denotes statistical significance at 0.05 significance level or better. Standard errors are shown in parentheses.
The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions.

where D represents a dummy variable which is assigned the following values

$$D = \begin{cases} 1 & : \text{if a DMU is identified as 100\% efficient in stage 1} \\ 0 & : \text{if a DMU is not identified as 100\% efficient in stage 1.} \end{cases} \tag{9.21}$$

Tables 9.4 and 9.5 exhibit the results secured when (9.19) was used as the estimating relation to obtain parameter values for both OLS and SF regressions. As might be expected, all parameter estimates are wide of the true values represented in (9.17)and significantly so in all cases as recorded in Tables 9.4 and 9.5. Somewhat surprising, however, is the fact that the OLS and SF estimates are very close and, in many cases they are identical.

When (9.20) is used — which is the regression with dummy variable values described in (9.21) — the situation is reversed for the efficient, but not for the inefficient DMUs. When the estimates are formed in the manner noted at the bottoms of Tables 9.6 and 9.7, none of the estimate of β_1 and β_2 differ significantly from their true values as given in (9.17). These are the estimates to be employed for $D = 1$. For $D = 0$, the case of inefficient DMUs, the previous result is repeated. All of the estimates differ significantly from their true values in both the empirical and simulation studies we described as can be seen in both of Tables 9.6 and 9.7.

Table 9.6. OLS Regression Estimates without Dummy Variables on DEA-efficient DMUs

Case 1: *EXPONENTIAL* distribution of input inefficiencies

Parameter Estimates	*Case A* $\sigma_\varepsilon^2 = 0.04$ (1)	*Case B* $\sigma_\varepsilon^2 = 0.0225$ (2)	*Case C* $\sigma_\varepsilon^2 = 0.01$ (3)	*Case D* $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.07*	1.47*	1.28*	1.34*
	(0.21)	(0.17)	(0.14)	(0.11)
β_1	0.49*	0.43*	0.46*	0.47*
	(0.03)	(0.02)	(0.02)	(0.01)
β_2	0.48*	0.48*	0.48*	0.46*
	(0.02)	(0.015)	(0.01)	(0.01)
δ	-1.57*	-2.30*	-1.50*	-1.50*
	(0.64)	(0.43)	(0.35)	(0.21)
δ_1	0.155*	0.26*	0.16*	0.16*
	(0.075)	(0.05)	(0.04)	(0.03)
δ_2	0.12*	0.12*	0.10*	0.09*
	(0.05)	(0.04)	(0.03)	(0.02)
Combining Parameters with *Dummy Variables*				
$H_0 : \beta_1 + \delta_1 = 0.65$ $H_a : \beta_1 + \delta_1 \neq 0.65$	$t_1 = 0.07$	$t_1 = 0.87$	$t_1 = -0.72$	$t_1 = 0.82$
$H_0 : \beta_2 + \delta_2 = 0.55$ $H_a : \beta_2 + \delta_2 \neq 0.55$	$t_2 = 1.09$	$t_2 = 1.76$	$t_2 = 1.02$	$t_2 \simeq 0$

The asterisk "*" denotes statistical significance at 0.05 significance level or better. Standard errors are shown in parentheses.
The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions.

The above tables report results for an exponential distribution of the inefficiencies associated with $\widehat{x}_1, \widehat{x}_2$ as defined in (9.18). However, uses of other statistical distributions and other forms of production functions did not alter these kinds of results for either the efficient or the inefficient DMUs. Thus this two-stage approach provided a new way of evaluating efficient and inefficient behaviors in both the empirical and simulation studies where it was used. It also provides an OLS regression as an alternative to the SF regression and this alternative is easier to use (or at least is more familiar) for many uses. See Brockett *et al.* (2004) [36] for an application to advertising strategy and a comparison with other types of statistical regression.

Comment : There are shortcomings and research challenges that remain to be met. One such challenge is to expand these uses to include multiple outputs as well as multiple inputs. Another challenge is to develop ways for identifying and estimating input specific as well as output specific inefficiencies. In order to meet such challenges it will be necessary to develop an analytically based theory in order to extend what can be accomplished by empirical applications and simulation studies.

Table 9.7. Stochastic Frontier Regression Estimates without Dummy Variables on DEA-efficient DMUs

Case 1: *EXPONENTIAL* distribution of input inefficiencies

Parameter Estimates	*Case A* $\sigma_\varepsilon^2 = 0.04$ (1)	*Case B* $\sigma_\varepsilon^2 = 0.0225$ (2)	*Case C* $\sigma_\varepsilon^2 = 0.01$ (3)	*Case D* $\sigma_\varepsilon^2 = 0.005$ (4)
β_0	1.18*	1.50*	0.80*	1.40*
	(0.23)	(0.16)	(0.16)	(0.13)
β_1	0.50*	0.44*	0.53*	0.49*
	(0.03)	(0.02)	(0.02)	(0.01)
β_2	0.48*	0.49*	0.50*	0.47*
	(0.02)	(0.02)	(0.01)	(0.02)
δ	-1.60*	-2.4*	-1.25*	-1.55*
	(0.57)	(0.56)	(0.38)	(0.23)
δ_1	0.16*	0.26*	0.13*	0.15*
	(0.07)	(0.06)	(0.04)	(0.03)
δ_2	0.11*	0.13*	0.086*	0.09*
	(0.05)	(0.04)	(0.04)	(0.03)
σ_ν	0.13(0.01)*	0.09(0.01)*	0.05(0.01)*	0.04(0.01)*
Combining Parameters with *Dummy Variables*				
$H_0 : \beta_1 + \delta_1 = 0.65$ $H_a : \beta_1 + \delta_1 \neq 0.65$	$t_1 = 0.20$	$t_1 = 0.93$	$t_1 = 0.28$	$t_1 = -0.4$
$H_0 : \beta_2 + \delta_2 = 0.55$ $H_a : \beta_2 + \delta_2 \neq 0.55$	$t_2 = 1.03$	$t_2 = 1.90$	$t_2 = 1.16$	$t_2 = 0.45$

The asterisk "*" denotes statistical significance at 0.05 significance level or better. Standard errors are shown in parentheses.
The values for σ_ε^2 shown in the top row of the table represent the true variances for the statistical error distributions.

Fortunately, excellent texts dealing with stochastic frontier and other approaches to efficiency evaluation have become available in the following two books,

1. T. Coelli, D.S.P. Rao and G.E. Battese (1998) *An Introduction to Efficiency and Productivity Analysis* (Boston: Kluwer Academic Publishers).
2. S.C. Kumbhakar and C.A.K. Lovell (2000) *Stochastic Frontier Analysis* (Cambridge: Cambridge University Press).

9.4 CHANCE-CONSTRAINED PROGRAMMING AND SATISFICING IN DEA

9.4.1 Introduction

S. Thore's (1987)[37] paper initiated a series of joint efforts with R. Land, and C.A.K. Lovell[38] directed to joining chance-constrained programming (CCP) with DEA as a third method for treating data uncertainties in DEA. Here we turn to Cooper, Huang and Li (1996)[39] to show how this approach can also be used to make contact with the concept of "satisficing" as developed in the psychology literature by H.A. Simon[40] as an alternative to the assumption of "optimizing" behavior which is extensively used in economics.

9.4.2 Satisficing in DEA

We start with the following CCP formulation that extends the CCR (ratio) model of DEA which was introduced in Section 2.3 of Chapter 2,

$$\begin{aligned} &\max \ P\left(\frac{\sum_{r=1}^{s} u_r \tilde{y}_{ro}}{\sum_{i=1}^{m} v_i \tilde{x}_{io}} \geq \beta_o\right) \qquad (9.22)\\ \text{subject to} \quad & P\left(\frac{\sum_{r=1}^{s} u_r \tilde{y}_{rj}}{\sum_{i=1}^{m} v_i \tilde{x}_{ij}} \leq \beta_j\right) \geq 1-\alpha_j, \quad j=1,\ldots,n \\ & u_r,\ v_i \geq 0 \ \ \forall r,\ i. \end{aligned}$$

Here "P" means "probability" and "$\tilde{}$" identifies these outputs and inputs as random variables with a known probability distribution while $0 \leq \alpha_j \leq 1$ is a scalar, specified in advance, which represents an allowable chance (=risk) of failing to satisfy the constraints with which it is associated.

For "satisficing," the values of β_o is interpreted as an "aspiration level" specified as an efficiency rating which is to be attained. The β_j are also prescribed constants imposed by the individual, or by outside conditions including superior levels of management.

To exhibit another aspect of satisficing behavior we might consider the case of inconsistent constraints. The problem will then have no solution. In such cases, according to Simon, an individual must either quit or else he must revise his aspiration level — or the risk of not achieving this level (or both). Thus, probabilistic (chance-constrained programming) formulations allow for types of behavior which are not present in the customary deterministic models of satisficing.

Now, however, we want to make contact with the deterministic DEA models which we discussed earlier. For this purpose we select the CCR model which was introduced as early as Chapter 2. This model always has a solution and the same is true for (9.22). This can be exhibited by choosing $u_r = 0\ \forall r$ and $v_i > 0$ for some i. Although not minimal for the objective, this choice satisfies all constraints with a probability of unity.

9.4.3 *Deterministic Equivalents*

To align the development more closely with our earlier versions of the CCR ratio model we note that

$$P\left(\frac{\sum_{r=1}^{s} u_r \widetilde{y}_{ro}}{\sum_{i=1}^{m} v_i \widetilde{x}_{io}} \leq \beta_o\right) + P\left(\frac{\sum_{r=1}^{s} u_r \widetilde{y}_{ro}}{\sum_{i=1}^{m} v_i \widetilde{x}_{io}} \geq \beta_o\right) = 1 \tag{9.23}$$

where, for simplicity, we restrict attention to the class of continous distributions. We therefore replace (9.22) with

$$\begin{aligned} & \max\ P\left(\frac{\sum_{r=1}^{s} u_r \widetilde{y}_{ro}}{\sum_{i=1}^{m} v_i x_{io}} \geq \beta_o\right) \qquad (9.24)\\ \text{subject to}\quad & P\left(\frac{\sum_{r=1}^{s} u_r \widetilde{y}_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \leq \beta_j\right) \geq 1-\alpha_j,\quad j=1,\ldots,n\\ & u_r,\ v_i \geq 0\ \ \forall r,\ i. \end{aligned}$$

Here we follow Land, Lovell and Thore (1993) and omit the symbol "$\sim$" from the x_{ij} (and x_{io}) in order to represent the inputs as deterministic. This model corresponds to a situation in which DMU managers choose the inputs without being able to completely control the outputs. Moreover if we also remove the symbol "$\sim$" from the y_{rj} (and y_{ro}), set $\beta_j = 1,\ j = 1,\ldots,n$ and remove β_o from the objective we will reproduce the CCR model that was first encountered in expression (2.3)-(2.6) in Chapter 2.

This identification having been made, we restore the symbol "$\sim$" to the y_{rj} (and y_{ro}) — thereby characterizing them as random variables — but we continue to treat the x_{ij} (and x_{io}) as deterministic. Then using vector-matrix notation we restate the constraints in (9.24) via the following development,

$$P\left(\frac{\boldsymbol{u}^T\widetilde{\boldsymbol{y}}_j}{\boldsymbol{v}^T\boldsymbol{x}_j} \leq \beta_j\right) = P\left(\boldsymbol{u}^T\widetilde{\boldsymbol{y}}_j \leq \beta_j \boldsymbol{v}^T\boldsymbol{x}_j\right). \tag{9.25}$$

Now let $\bar{\boldsymbol{y}}_j$ be the vector of output means and let Σ_j represent the variance-covariance matrix. We assume that this matrix is positive definite so we can represent the variance by $\boldsymbol{u}^T\Sigma_j\boldsymbol{u}$, a scalar, which is also positive for all choices of $\boldsymbol{u} \neq \boldsymbol{0}$. We then subtract $\boldsymbol{u}^T\bar{\boldsymbol{y}}_j$ from both sides of the right-hand inequality in (9.25) and divide through by $\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}$ to obtain

$$P\left(\frac{\boldsymbol{u}^T\widetilde{\boldsymbol{y}}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j}{\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}} \leq \frac{\beta_j\boldsymbol{v}^T\boldsymbol{x}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j}{\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}}\right) \geq 1 - \alpha_j \tag{9.26}$$

for each $j = 1, \ldots, n$. Now we note that the expression on the right in the parenthesis does not contain any random elements.

To simplify our notation we introduce a new random variable defined by

$$\widetilde{z}_j = \frac{\boldsymbol{u}^T\widetilde{\boldsymbol{y}}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j}{\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}}. \tag{9.27}$$

We then replace (9.26) with

$$P\left(\widetilde{z}_j \leq k_j(\boldsymbol{u}^T, \boldsymbol{v}^T)\right) \geq 1 - \alpha_j, \ j = 1, \ldots, n \tag{9.28}$$

where

$$k_j(\boldsymbol{u}^T, \boldsymbol{v}^T) = \frac{\beta_j\boldsymbol{v}^T\boldsymbol{x}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j}{\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}}$$

so we can write

$$\int_{-\infty}^{k_j(\boldsymbol{u}^T,\boldsymbol{v}^T)} f(z_j)dz_j = \Phi\left(\frac{\beta_j\boldsymbol{v}^T\boldsymbol{x}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j}{\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}}\right) \geq 1 - \alpha_j \tag{9.29}$$

in place of (9.26).

We now assume that Φ is the normal distribution which has been standardized via

$$\widetilde{z}_j = \frac{\boldsymbol{u}^T(\widetilde{\boldsymbol{y}}_j - \bar{\boldsymbol{y}}_j)}{\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}}. \tag{9.30}$$

Assuming $\alpha_j \leq 0.5$ we can utilize the property of invertibility associated with this distribution and apply it to (9.29) to obtain

$$\frac{\beta_j\boldsymbol{v}^T\boldsymbol{x}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j}{\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}} \geq \Phi^{-1}(1 - \alpha_j) \tag{9.31}$$

where Φ^{-1} is the fractile function associated with the standard normal distribution. Hence also

$$\beta_j\boldsymbol{v}^T\boldsymbol{x}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j \geq \Phi^{-1}(1 - \alpha_j)\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}. \tag{9.32}$$

We now employ what Charnes and Cooper (1963)[41] refer to as "splitting variables" which we symbolize by η_j in

$$\beta_j\boldsymbol{v}^T\boldsymbol{x}_j - \boldsymbol{u}^T\bar{\boldsymbol{y}}_j \geq \eta_j \geq \Phi^{-1}(1 - \alpha_j)\sqrt{\boldsymbol{u}^T\Sigma_j\boldsymbol{u}}. \tag{9.33}$$

For every $j = 1, \ldots, n$ this variable is nonnegative by virtue of the expression on the right. Provided this nonnegativity is preserved we can therefore use this variable to split the expression in (9.33) into the following pair

$$\begin{aligned} &\beta_j \boldsymbol{v}^T \boldsymbol{x}_j - \boldsymbol{u}^T \bar{\boldsymbol{y}}_j \geq \eta_j \geq 0 \\ &K^2_{(1-\alpha_j)} \boldsymbol{u}^T \Sigma_j \boldsymbol{u} \leq \eta_j^2 \end{aligned} \tag{9.34}$$

where

$$\begin{aligned} &K_{(1-\alpha_j)} = \Phi^{-1}(1-\alpha_j) \\ &j = 1, \ldots, n. \end{aligned}$$

We have thus separated the conditions in (9.33) into a pair for which the first relation refers to a valuation effected by multipliers assigned to the inputs and outputs while the second relation treats the "risks" as in a portfolio analysis of the Markowitz-Sharpe type used in finance.[42]

In place (9.24) we now have

$$\begin{aligned} &\max\ P\left(\frac{\boldsymbol{u}^T \widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^T \boldsymbol{x}_o} \geq \beta_o\right) \\ \text{subject to}\quad &\beta_j \boldsymbol{v}^T \boldsymbol{x}_j - \boldsymbol{u}^T \bar{\boldsymbol{y}}_j - \eta_j \geq 0 \\ &K^2_{(1-\alpha_j)} \boldsymbol{u}^T \Sigma_j \boldsymbol{u} - \eta_j^2 \leq 0 \\ &\boldsymbol{v},\ \boldsymbol{u} \geq \boldsymbol{0},\ \eta_j \geq 0,\ j = 1, \ldots, n. \end{aligned} \tag{9.35}$$

The constraints, but not the objective, are now deterministic. To bring our preceding development to bear we therefore replace (9.35) with

$$\begin{aligned} &\max\ \gamma_o \\ \text{subject to}\quad &P\left(\frac{\boldsymbol{u}^T \widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^T \boldsymbol{x}_o} \geq \beta_o\right) \geq \gamma_o \\ &\beta_j \boldsymbol{v}^T \boldsymbol{x}_j - \boldsymbol{u}^T \bar{\boldsymbol{y}}_j - \eta_j \geq 0 \\ &K^2_{(1-\alpha_j)} \boldsymbol{u}^T \Sigma_j \boldsymbol{u} - \eta_j^2 \leq 0 \\ &\boldsymbol{v},\ \boldsymbol{u} \geq \boldsymbol{0},\ \eta_j \geq 0,\ j = 1, \ldots, n. \end{aligned} \tag{9.36}$$

Proceeding as before we then have

$$\begin{aligned} &\max\ \gamma_o \\ \text{subject to}\quad &\boldsymbol{u}^T \bar{\boldsymbol{y}}_o - \beta_o \boldsymbol{v}^T \boldsymbol{x}_o \geq \Phi^{-1}(\gamma_o)\sqrt{\boldsymbol{u}^T \Sigma_j \boldsymbol{u}} \\ &\eta_j + \boldsymbol{u}^T \boldsymbol{y}_j - \beta_j \boldsymbol{v}^T \boldsymbol{x}_j \leq 0 \\ &\eta_j^2 - K^2_{(1-\alpha_j)} \boldsymbol{u}^T \Sigma_j \boldsymbol{u} \geq 0 \\ &\boldsymbol{u},\ \boldsymbol{v} \geq \boldsymbol{0},\ \eta_j \geq 0,\ j = 1, \ldots, n \\ &0 \leq \gamma_o \leq 1. \end{aligned} \tag{9.37}$$

This is a "deterministic equivalent" for (9.24) in that the optimal values of $\boldsymbol{u}^*$, $\boldsymbol{v}^*$ in (9.37) will also be optimal for (9.24).

9.4.4 Stochastic Efficiency

Although entirely deterministic, this problem is difficult to solve because, by virtue of the first constraint, it is nonconvex as well as nonlinear. As shown in Cooper, Huang and Li (1996), it is possible to replace (9.37) with a convex programming problem but we do not undertake the further development needed to show this. Instead, we assume that we have a solution with

$$\gamma_o^* = P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \geq \beta_o\right), \tag{9.38}$$

where γ_o^* is obtained from (9.36).

To develop the concepts of stochastic efficiency we assume that $\beta_o = \beta_{j_o}$ so the level prescribed for DMU_o in its constraints is the same as the β_o level prescribed in the objective. Then we note that $\gamma_o^* > \alpha_{j_o}$ is not possible because this would fail to satisfy this constraint. To see that this is so we note, as in (9.23), that

$$P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \geq \beta_o\right) + P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \leq \beta_o\right) = 1 \tag{9.39}$$

because $P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} = \beta_o\right) = 0$ for a continuous distribution. Hence

$$\begin{aligned} P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \leq \beta_o\right) &= 1 - P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \geq \beta_o\right) \\ &= 1 - \gamma_o^* \ < \ 1 - \alpha_{j_o} \end{aligned} \tag{9.40}$$

which fails to satisfy the constraint

$$P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \leq \beta_o\right) \geq 1 - \alpha_{j_o}.$$

Now if $\beta_{j_o} = \beta_o = 1$, which will usually be the case of interest, the above development leads to the following,

Theorem 9.1 (Cooper, Huang and Li (1996)) *If $\beta_{j_o} = \beta_o = 1$ then DMU_o will have performed in a stochastically efficient manner if and only if $\gamma_o^* = \alpha_{j_o}$.*

This leaves the case of $\gamma_o^* < \alpha_{j_o}$ to be attended to. In this case the risk of failing to satisfy the constraints for DMU_{j_o} falls below the level which was specified as satisfactory. To restate this in a more positive manner, we return to (9.40) and reorient it to

$$\begin{aligned} P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \leq \beta_o\right) &= 1 - P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} \geq \beta_o\right) \\ &= 1 - \gamma_o^* \ > \ 1 - \alpha_{j_o}. \end{aligned} \tag{9.41}$$

This leads to the following corollary to the above theorem,

Corollary 9.1 *If $\beta_{j_o} = \beta_o = 1$ then DMU$_o$'s performance is stochastically inefficient with probability $1-\gamma_o^*$ if and only if $\gamma_o^* < \alpha_{j_o}$.*

We justify this by noting that (9.41) means that the probability of falling below β_o exceeds the probability that was specified as being satisfactory.

To return to the satisficing model and concepts discussed in Section 9.4.2, we assume that $0 \leq \beta_o = \beta_{j_o} < 1$. The above theorem and corollary then translate into: "satisficing was attained or not according to whether $\gamma_o^* = \alpha_{j_o}$ or $\gamma_o^* < \alpha_{j_o}$." To see what this means we note that (9.39) consists of opposed probability statements except for the case $P\left(\frac{\boldsymbol{u}^{*T}\widetilde{\boldsymbol{y}}_o}{\boldsymbol{v}^{*T}\boldsymbol{x}_o} = \beta_o\right) = 0$. Hence failure to attain satisficing occurs when the goal specified in the objective can be attained only at a level below the risk that is specified for being wrong in making this inference.

Returning to our earlier choice of $u_r = 0\ \forall r$ and some $v_i > 0$ for (9.22) we note that this assigns a positive value to some inputs and a zero value for all outputs. The objective and the associated constraints in (9.22) will then be achieved with probability one because of refusal to play and hence there will be a zero probability of achieving the objective in (9.24). See (9.39). This is an important special case of Simon's "refusal to play" behavior that was noted in the discussion following (9.22).

Comment : This is as far as we carry our analyses of these chance constrained programming approaches to DEA. We need to note, however, that this analysis has been restricted to what is referred to as the "P-model" in the chance-constrained programming literature. Most of the other DEA literature on this topic has utilized the "E-model," so named because its objective is stated in terms of optimizing "expected values." None has yet essayed a choice of "V-models" for which the objective is to minimize "variance" or "mean-square error." See A. Charnes and W.W. Cooper (1963) "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints," *Operations Research* 11, pp.18-39.

The formulations here (and elsewhere in the DEA literature) are confined to a use of "conditional chance constraints." A new chapter was opened for this research, however, by Olesen and Petersen (1995)[43] who used "joint chance constraints" in their CCP models. In a subsequent paper, Cooper, Huang, Lelas, Li and Olesen (1998)[44] utilize such joint constraints to extend the concept of "stochastic efficiency" to a measure called "α-stochastic efficiency" for use in problems where "efficiency dominance" is of interest.

Like the rest of the DEA literature dealing with CCP, we have restricted attention to the class of "zero order decision rules." This corresponds to a "here and now" approach to decision making in contrast to the "wait and see" approach that is more appropriate to dynamic settings in which it may be better to delay some parts of a decision until more information is available. To go further in this direction leads to the difficult problem of choosing a best decision rule from an admissible class of decision rules. To date, this problem has only been addressed in any detail for the class of linear (first order) decision rules even in the general literature on CCP and even this treatment was conducted under restrictive assumptions on the nature of the random variables and their statistical distributions.[45] See also Charnes, Cooper and Symods (1958), [46] an article which originated the (as-yet-to-be named) "chance constraint programming." This topic is important and invites further treatments as an example of the "information processing" that plays an important role in studying use of the term "satisficing behavior." See Gigerenzer (2004). [47] Finally we come to the treatments which, by and large, have assumed that these probability distributions are known. Hence there is a real need and interest in relaxing this assumption. Only bare beginnings on this topic have been made as in R. Jagannathan (1985) "Use of Sample Information in Stochastic Recourse and Chance Constrained Programming Models," *Management Science* 31, pp.96-108.

9.5 WINDOW ANALYSIS

9.5.1 An Example

Although now used in many other contexts we here revert to the applications in army recruiting efforts that gave rise to "window analysis." This example is built around 3 outputs consisting of recruits such as male and female high school graduates and degree candidates (such as high school seniors not yet graduated). The 10 inputs consist of number of recruiters (such as recruiting sergeants) and their aids, amount spent on local advertising, as well as qualities such as recruiter experience which are all discretionary variables as well as non-discretionary variables such as local unemployment, number of high school seniors in the local market and their propensity to enlist.[48]

These variables are used (among others) by the U.S. Army Recruiting Command (USAREC) to evaluate performance in its various organization units which Klopp (1985, p.115)[48] describes as follows. USAREC recruits for the entire United States. To facilitate control of the recruiting process USAREC divides the U.S. into 5 regions managed by entities referred to as "Recruiting Brigades." Each Recruiting Brigade is responsible for some of the 56 "Recruiting Battalions" that operate in the U.S. and the latter are responsible, in turn, for the "Recruiting Stations" where recruiters are assigned specific missions. For the example we employ, it was decided to use the Recruiting Battalions each of which was designated as a DMU.

The DEA analyses took a variety of forms. One consisted of standard (static) DEA reports in the manner of Table 3.7 expanded to include detailed information on the inefficiencies present in each DMU. Both discretionary and non-discretionary variables were included along with the members of the peer group used to effect the evaluation.

Something more was wanted in the form of trend analyses of the quarterly reports that USAREC received. A use of statistical regressions and time series analyses of the efficiency scores proved unsatisfactory. Experimentation was therefore undertaken which led to the window analysis we now describe.

9.5.2 Application

Table 9.8, as adapted from Klopp (1985) will be used for guidance in this discussion. The basic idea is to regard each DMU as if it were a different DMU in each of the reporting dates represented by Q1, Q2, etc., at the top of the table. For instance, the results in row 1 for Battalion 1A represent four values obtained by using its results in Q1, Q2, Q3 and Q4 by bringing this DMU into the objective for each of these quarters. Because these 56 DMUs are regarded as different DMUs in each quarter, these evaluations are conducted by reference to the entire set of $4\times 56 = 224$ DMUs that are used to form the data matrix. Thus the values of $\theta^* = 0.83$, 1.00, 0.95 and 1.00 in row 1 for Battalion 1A represent its quarterly performance ratings as obtained from this matrix with $13 \times 224 = 2{,}912$ entries.

Similar first-row results for other DMUs are obtained for each of the 10 DMUs in Table 9.8 which we have extracted from a larger tabulation of 56 DMUs to obtain compactness. After first row values have been similarly obtained, a new 4-period window is obtained by dropping the data for Q1 and adding the data for Q5. Implementing the same procedure as before produces the efficiency ratings in row 2 for each Battalion. The process is then continued until no further quarters are added — as occurs (here) in "row" 5 with its 4 entries.

To help interpret the results we note that the "column views" enable us to examine the stability of results across the different data sets that occur with these removal and replacement procedures. "Row views" make it possible to determine trends and/or observed behavior with the same data set. Thus, the column view for Battalion 1A shows stability and the row view shows steady behavior after the improvement over Q1. At the bottom of Table 9.8, however, Battalion 1K exhibits deteriorating behavior and this same deterioration continues to be manifested with different data sets. Moreover, to reenforce this finding the values in each column for Battalion 1K do not change very much and this stability reenforces this finding.

Finally we augment this information by the statistical values noted on the right side of Table 9.8. Here the average is used as a representative measure obtained from the θ^* values for each DMU and matched against its variance. Medians might also be used and matched with ranges and so on. Other summary statistics and further decompositions may also be used, of course, but a good deal of information is supplied in any case.

Weaknesses are also apparent, of course, such as the absence of attention to nonzero slacks. However, the principle and formulas to be supplied in the next section of this chapter may be applied to slack portrayals, too, if desired and other similar measures may be used such as the SBM (Slacks Based Measure) given in Section 4.4 of Chapter 4.

Another deficiency is apparent in that the beginning and ending period DMUs are not tested as frequently as the others. Thus the DMUs in Q1 are examined in only one window and the same is true for Q8. In an effort to address this problem Sueyoshi (1992)[49] introduced a "round robin" proce-

Table 9.8. Window Analysis: 56 DMUs in U.S. Army Recruitment Battalions 3 Outputs - 10 Inputs

Battalion	Efficiency Scores								Summary Measures			
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Mean	Var	Column Range	Total Range
1A	0.83	1.00	0.95	1.00								
		1.00	0.95	1.00	1.00				0.99	0.03	.05	.17
			1.00	1.00	1.00	1.00						
				1.00	1.00	1.00	1.00					
					1.00	1.00	1.00	1.00				
1B	0.71	1.00	0.91	0.89								
		1.00	0.91	0.89	0.77				0.92	0.21	.11	.29
			1.00	1.00	0.76	1.00						
				1.00	0.75	1.00	1.00					
					0.80	1.00	1.00	1.00				
1C	0.72	1.00	0.86	1.00								
		1.00	0.87	1.00	0.74				0.93	0.21	.26	.28
			0.87	1.00	0.74	1.00						
				1.00	1.00	1.00	1.00					
					0.80	1.00	1.00	1.00				
1D	0.73	1.00	1.00	1.00								
		1.00	1.00	1.00	1.00				0.99	0.07	.00	.27
			1.00	1.00	1.00	1.00						
				1.00	1.00	1.00	1.00					
					1.00	1.00	1.00	1.00				
1E	0.69	1.00	0.83	1.00								
		0.83	0.83	1.00	0.77				0.88	0.17	.17	.31
			0.83	1.00	0.79	0.93						
				1.00	0.77	0.95	0.83					
					0.82	0.88	0.88	1.00				
1F	1.00	1.00	1.00	1.00								
		1.00	1.00	1.00	1.00				0.99	0.04	.20	.20
			1.00	1.00	1.00	1.00						
				1.00	1.00	1.00	1.00					
					0.80	1.00	1.00	1.00				
1G	0.65	0.73	0.68	0.77								
		0.73	0.68	0.77	0.68				0.79	0.32	.32	.35
			0.68	0.77	0.68	0.72						
				0.76	0.74	0.75	1.00					
					1.00	1.00	1.00	1.00				
1H	0.82	1.00	1.00	1.00								
		1.00	1.00	1.00	0.71				0.90	0.32	.29	.30
			1.00	1.00	0.71	0.70						
				1.00	1.00	0.72	0.72					
					0.79	0.81	1.00	1.00				
1I	1.00	1.00	1.00	1.00								
		1.00	1.00	1.00	1.00				0.99	0.02	.14	.14
			0.86	1.00	1.00	1.00						
				1.00	1.00	1.00	1.00					
					1.00	1.00	1.00	1.00				
1K	0.81	1.00	1.00	1.00								
		1.00	1.00	1.00	0.73				0.86	0.37	.11	.33
			1.00	1.00	0.73	0.67						
				1.00	0.85	0.67	0.67					
					0.67	0.67	0.67	0.67				

dure which proceeds as follows: First each period is examined independently. This is followed by a 2-period analysis after which a three-period analysis is used. And so on. However, this analysis becomes unwieldy since the number of combinations grows to $2^p - 1$ so that some care is needed with this approach.

9.5.3 *Analysis*

The following formulas adapted from D.B. Sun (1988)[50] can be used to study the properties of these window analyses. For this purpose we introduce the following symbols

$$\begin{aligned} n &= \text{number of DMUs} \\ k &= \text{number of periods} \\ p &= \text{length of window } (p \le k) \\ w &= \text{number of windows.} \end{aligned} \tag{9.42}$$

We then reduce the number of DMUs to the 10 in Table 9.8 so we can use it for numerical illustrations.

	Formula	Application (to Table 9.8)
no. of windows:	$w = k - p + 1$	$8 - 4 + 1 = 5$
no. of DMUs in each window:	$np/2$	$10 \times 4/2 = 20$
no. of "different" DMUs:	npw	$10 \times 4 \times 5 = 200$
Δ no. of DMUs:	$n(p-1)(k-p)$	$10 \times 3 \times 4 = 120$

Here "Δ" represents an increase compared to the $8 \times 10 = 80$ DMUs that would have been available if the evaluation had been separately effected for each of the 10 DMUs in each of the 8 quarters.

An alternate formula for deriving the total number of DMUs is given in Charnes and Cooper (1990)[51] as follows.

$$\text{Total no. of "different" DMUs}: \; n(k-p+1)p = 10 \times (8-4+1) \times 4 = 200 \tag{9.43}$$

Differentiating this last function and equating to zero gives

$$p = \frac{k+1}{2} \tag{9.44}$$

as the condition for a maximum number of DMUs. This result need not be an integer, however, so we utilize the symmetry of (9.43) and (9.44) and modify the latter to the following,

$$p = \begin{cases} \frac{k+1}{2} & \text{when } k \text{ is odd} \\ \frac{k+1}{2} \pm \frac{1}{2} & \text{when } k \text{ is even.} \end{cases} \tag{9.45}$$

To see how to apply this formula when k is even we first note that

$$n(k-p+1)p = n\left[(k+1)p - p^2\right].$$

Hence by direct substitution we obtain

$$\begin{aligned} & n\left[(k+1)\left(\frac{k+1}{2}-\frac{1}{2}\right)-\left(\frac{k+1}{2}-\frac{1}{2}\right)^2\right] \\ = \; & n\left[(k+1)\left(\frac{k+1}{2}+\frac{1}{2}\right)-\left(\frac{k+1}{2}+\frac{1}{2}\right)^2\right] = \frac{n}{4}\left[(k+1)^2-1\right] \end{aligned}$$

Then, for $k = 8$, as in our example, we find from (9.44) that

$$p = \frac{8+1}{2} = 4.5$$

which is not an integer. Hence using $[p]$ to mean "the integer closest to p" we apply the bottom expression in (9.45) to obtain

$$[p] = \begin{cases} 4 = 4.5 - 0.5 \\ 5 = 4.5 + 0.5 \end{cases}$$

and note that substitution in (9.43) produces 200 "different" DMUs as the maximum number in either case.

9.6 SUMMARY OF CHAPTER 9

In this chapter we have treated the topic of data variability in the following manner. Starting with the topic of sensitivity and stability analysis we moved to statistical regression approaches which we aligned with DEA in various ways. We then went on to probabilistic formulations using the P-model of chance-constrained programming which we could relate to our CCR and BCC models which we had previously treated in deterministic manners. Finally we turned to window analysis which allowed us to study trends as well as stability of results when DMUs are systematically dropped and added to the collection to be examined.

The topics treated in this chapter are all under continuing development. Hence we do not provide problems and suggested answers like those we presented in preceding chapters. Instead we have provided comments and references that could help to point up issues for further research. We hope that readers will respond positively and regard this as an opportunity to join in the very active research that is going on in these (and other) areas.

We here note that research on the topics treated in this chapter were prompted by problems encountered in attempts to bring a successful conclusion to one or many attempts to use DEA in actual applications. Indeed a good deal of the very considerable progress in DEA has emanated from actual attempts to apply it to different problems. This, too, has been an important source of progress in that new applications as well as new developments in DEA are being simultaneously reported.

To help persons who want to pursue additional topics and uses we have supplied an extensive bibliography in the disk that accompanies this book as well as in this chapter. We hope this will be helpful and we hope readers of our text will experience some of the fun and exhilaration that we have experienced as we watch the rapid pace of developments in DEA.

9.7 RELATED DEA-SOLVER MODELS FOR CHAPTER 9

Window-I(O)-C(V) These codes execute Window Analysis in Input (Output) orientation under constant (CRS) or variable (VRS) returns-to-scale assumptions. See the sample data format in Section B.5.10 and explanation on results in Section B.7 of Appendix B.

Notes

1. P.W. Wilson (1995), "Detecting Influential Observations in Data Envelopment Analysis," *Journal of Productivity Analysis* 6, pp.27-46.

2. See, for instance, R.M. Thrall (1989), "Classification of Transitions under Expansion of Inputs and Outputs," *Managerial and Decision Economics* 10, pp.159-162.

3. R.D. Banker, H. Chang and W.W. Cooper (1996), "Simulation Studies of Efficiency, Returns to Scale and Misspecification with Nonlinear Functions in DEA," *Annals of Operations Research* 66, pp.233-253.

4. A. Charnes, W.W. Cooper, A.Y. Lewin, R.C. Morey and J.J. Rousseau (1985), "Sensitivity and Stability Analysis in DEA," *Annals of Operations Research* 2, pp.139-156.

5. A. Charnes and W.W. Cooper (1968), "Structural Sensitivity Analysis in Linear Programming and an Exact Product Form Left Inverse," *Naval Research Logistics Quarterly* 15, pp.517-522.

6. For a summary discussion see A. Charnes and L. Neralic (1992), "Sensitivity Analysis in Data Envelopment Analysis 3," *Glasnik Matematicki* 27, pp.191-201. A subsequent extension is L. Neralic (1997), "Sensitivity in Data Envelopment Analysis for Arbitrary Perturbations of Data," *Glasnik Matematicki* 32, pp.315-335. See also L. Neralic (2004), "Preservation of Efficiency and Inefficiency Classification in Data Envelopment Analysis," *Mathematical Communications* 9, pp.51-62.

7. A. Charnes, S. Haag, P. Jaska and J. Semple (1992), "Sensitivity of Efficiency Calculations in the Additive Model of Data Envelopment Analysis," *International Journal of System Sciences* 23, pp.789-798. Extensions to other classes of models may be found in A. Charnes, J.J. Rousseau and J.H.Semple (1996) "Sensitivity and Stability of Efficiency Classifications in DEA," *Journal of Productivity Analysis* 7, pp.5-18.

8. The shape of this "ball" will depend on the norm that is used. For a discussion of these and other metric concepts and their associated geometric portrayals see Appendix A in A. Charnes and W.W. Cooper (1961), *Management Models and Industrial Applications of Linear Programming* (New York: John Wiley & Sons).

9. This omission of DMU_o is also used in developing a measure of "super efficiency" as it is called in P. Andersen and N.C. Petersen (1993), "A Procedure for Ranking Efficient Units in DEA," *Management Science* 39, pp.1261-1264. Their use is more closely associated with stability when a DMU is omitted, however, so we do not cover it here. See the next chapter, Chapter 10, in this text. See also R.M. Thrall (1996) "Duality Classification and Slacks in DEA," *Annals of Operations Research* 66, pp.104-138.

10. A. Charnes, J.J. Rousseau and J.H. Semple (1996), "Sensitivity and Stability of Efficiency Classification in Data Envelopment Analysis," *Journal of Productivity Analysis* 7, pp.5-18.

11. In fact, Seiford and Zhu propose an iterative approach to assemble an exact stability region in L. Seiford and J. Zhu, "Stability Regions for Maintaining Efficiency in Data Envelopment Analysis," *European Journal of Operational Research* 108, 1998, pp.127-139.

12. R.G. Thompson, P.S. Dharmapala and R.M. Thrall (1994), "Sensitivity Analysis of Efficiency Measures with Applications to Kansas Farming and Illinois Coal Mining," in A. Charnes, W.W. Cooper, A.Y. Lewin and L.M. Seiford, eds., *Data Envelopment Analysis: Theory, Methodology and Applications* (Norwell, Mass., Kluwer Academic Publishers) pp.393-422.

13. R.G. Thompson, P.S. Dharmapala, J. Diaz, M.D. Gonzales-Lina and R.M. Thrall (1996), "DEA Multiplier Analytic Center Sensitivity Analysis with an Illustrative Application to Independent Oil Cos.," *Annals of Operations Research* 66, pp.163-180.

14. A. Charnes, W.W. Cooper and R.M.Thrall (1991), "A Structure for Classifying and Characterizing Efficiency in Data Envelopment Analysis," *Journal of Productivity Analysis* 2, pp.197-237.

15. An alternate approach to simultaneous variations in all data effected by the envelopment model is available in L.M. Seiford and J. Zhu (1998), "Sensitivity Analysis of DEA Models for Simultaneous Changes in All Data," *Journal of the Operational Research Society* 49, pp.1060-1071. See also the treatments of simultaneous changes in all data for additive models in L. Neralic (2004), "Preservation of Efficiency and Inefficiency Classification in Data Envelopment Analysis," *Mathematical Communications* 9, pp.51-62.

16. The ranges within which these dual variable values do not change generally form part of the printouts in standard linear programming computer codes.

17. This status is easily recognized because $\theta_1^* = \theta_2^* = \theta_3^* = 1$ are all associated with uniquely obtained solutions with zero slacks. See A. Charnes, W.W. Cooper and R.M. Thrall (1991) "A Structure for Classifying and Characterizing Efficiency in Data Envelopment Analysis," *Journal of Productivity Analysis* 2, pp.197-237.

18. W.W. Cooper, S. Li, L.M. Seiford and J. Zhu (2004), Chapter 3 in W.W. Cooper, L.M. Seiford and J. Zhu, eds., *Handbook on Data Envelopment Analysis* (Norwell, Mass., Kluwer Academic Publishers).

19. R.D. Banker (1993), "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation," *Management Science* 39, pp.1265-1273.

20. R. Banker and R. Natarasan (2004), "Statistical Tests Based on DEA Efficiency Scores," Chapter 11 in in W.W. Cooper, L.M. Seiford and J. Zhu, eds., *Handbook on Data Envelopment Analysis* (Norwell, Mass., Kluwer Academic Publishers).

21. Quoted from p.139 in R.D Banker (1996), "Hypothesis Tests Using Data Envelopment Analysis," *Journal of Productivity Analysis* pp.139-159.

22. A.P. Korostolev, L. Simar and A.B. Tsybakov (1995), "On Estimation of Monotone and Convex Boundaries," *Public Institute of Statistics of the University of Paris,* pp.3-15. See also Korostolev, Simar and Tsybakov (1995), "Efficient Estimation of Monotone Boundaries," *Annals of Statistics* 23, pp.476-489.

23. See the discussion in L. Simar (1996), "Aspects of Statistical Analysis in DEA-Type Frontier Models," *Journal of Productivity Analysis* 7, pp.177-186.

24. L. Simar and P.W. Wilson (1998), "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models," *Management Science* 44, pp.49-61. See also Simar and Wilson (2004) "Performance of the Bootstrap for DEA Estimators and Iterating the Principle," Chapter 10 in W.W. Cooper, L.M. Seiford and J. Zhu, eds. *Handbook on Data Envelopment Analysis* (Norwell, Mass: Kluwer Academic Publishers).

25. M.J. Farrell (1951), "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society* Series A, 120, pp.253-290.

26. D.J. Aigner and S.F. Chu (1968), "On Estimating the Industry Production Frontiers," *American Economic Review* 56, pp.826-839.

27. D.J. Aigner, C.A.K. Lovell and P. Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Models," *Journal of Econometrics* 6, pp.21-37. See also W. Meeusen and J. van den Broeck (1977) "Efficiency Estimation from Cobb-Douglas Functions with Composed Error," *International Economic Review* 18, pp.435-444.

28. J. Jondrow, C.A.K. Lovell, I.S. Materov and P. Schmidt (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Model," *Journal of Econometrics* 51, pp.259-284.

29. B.H. Gong and R.C. Sickles (1990), "Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data," *Journal of Econometrics* 51, pp.259-284.

30. P. Schmidt (1985-1986), "Frontier Production Functions," *Econometric Reviews* 4, pp.289-328. See also P.W. Bauer (1990), "Recent Development in Econometric Estimation of Frontiers," *Journal of Econometrics* 46, pp.39-56.

31. G.D. Ferrier and C.A.K. Lovell (1990), "Measuring Cost Efficiency in Banking — Econometric and Linear Programming Evidence," *Journal of Econometrics* 6, pp.229-245.

32. A. Charnes, W.W. Cooper and T. Sueyoshi (1988), "A Goal Programming/Constrained Regression Review of the Bell System Breakup," *Management Science* 34, pp.1-26.

33. See R.S. Barr, L.M. Seiford and T.F. Siems (1994), "Forcasting Bank Failure: A Non-Parametric Frontier Estimation Approach," *Recherches Economiques de Louvain* 60, pp. 417-429. for an example of a different two-stage DEA regression approach in which the DEA scores from the first stage served as an independent variable in the second stage regression model.

34. V. Arnold, I.R. Bardhan, W.W. Cooper and S.C. Kumbhakar (1994), "New Uses of DEA and Statistical Regressions for Efficiency Evaluation and Estimation — With an Illustrative Application to Public Secondary Schools in Texas," *Annals of Operations Research* 66, pp.255-278.

35. I.R. Bardhan, W.W. Cooper and S.C. Kumbhakar (1998), "A Simulation Study of Joint Uses of Data Envelopment Analysis and Stochastic Regressions for Production Function Estimation and Efficiency Evaluation," *Journal of Productivity Analysis* 9, pp.249-278.

36. P.L. Brockett, W.W. Cooper, S.C. Kumbhakar, M.J. Kwinn Jr. and D. McCarthy (2004), "Alternative Statistical Regression Studies of the Effects of Joint and Service-Specific Advertising on Military Recruitment," *Journal of the Operational Research Society* 55, pp.1039-1048.

37. S. Thore (1987), "Chance-Constrained Activity Analysis," *European Journal of Operational Research* 30, pp.267-269.

38. See the following three papers by K.C. Land, C.A.K. Lovell and S. Thore: (1) "Productive Efficiency under Capitalism and State Socialism: the Chance Constrained Programming Approach" in Pierre Pestieau, ed. in *Public Finance in a World of Transition* (1992) supplement to *Public Finance* 47, pp.109-121; (2) "Chance-Constrained Data Envelopment Analysis," *Managerial and Decision Economics* 14, 1993, pp.541-554; (3) "Productive Efficiency under Capitalism and State Socialism: An Empirical Inquiry Using Chance-Constrained Data Envelopment Analysis," *Technological Forecasting and Social Change* 46, 1994, pp.139-152. In "Four Papers on Capitalism and State Socialism" (Austin Texas: The University of Texas, IC^2 Institute) S. Thore notes that publication of (2) was delayed because it was to be presented at a 1991 conference in Leningrad which was cancelled because of the Soviet Crisis.

39. W.W. Cooper, Z. Huang and S. Li (1996), "Satisficing DEA Models under Chance Constraints," *Annals of Operations Research* 66, pp.279-295. For a survey of chance constraint programming uses in DEA, see W.W. Cooper, Z. Huang and S. Li (2004), "Chance Constraint DEA," in W.W. Cooper, R.M. Seiford and J. Zhu, eds., *Handbook on Data Envelopment Analysis* (Norwell, Mass., Kluwer Academic Publishers).

40. See Chapter 15 in H.A. Simon (1957), *Models of Man* (New York: John Wiley & Sons, Inc.)

41. A. Charnes and W.W. Cooper (1963), "Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints," *Operations Research* 11, pp.18-39.

42. See W.F. Sharpe (1970), *Portfolio Theory and Capital Markets* (New York: McGraw Hill, Inc.)

43. O.B. Olesen and N.C. Petersen (1995), "Chance Constrained Efficiency Evaluation," *Management Science* 41, pp.442-457.

44. W.W. Cooper, Z. Huang, S.X. Li and O.B. Olesen (1998), "Chance Constrained Programming Formulations for Stochastic Characterizations of Efficiency and Dominance in DEA," *Journal of Productivity Analysis* 9, pp.53-79.

45. See A. Charnes and W.W. Cooper (1963) in footnote 41, above.

46. A. Charnes, W.W. Cooper and G.H. Symods (1958), "Cost Horizons and Certainty Equivalents," *Management Science* 4, pp.235-263.

47. G. Gigerenzer (2004), "Striking a Blow for Sanity in Theories of Rationality," in M. Augier and J.G. March, eds., *Models of a Man: Essays in Memory of H.A. Simon* (Cambridge: MIT Press).

48. As determined from YATS (Youth Attitude Tracking Survey) which obtains this information from periodic surveys conducted for the Army. See also G.A. Klopp (1985), "The Analysis of the Efficiency of Productive Systems with Multiple Inputs and Outputs," Ph.D. Dissertation (Chicago: University of Illinois at Chicago). Also available from University Microfilms, Inc., in Ann Arbor, Michigan.

49. T. Sueyoshi (1992), "Comparisons and Analyses of Managerial Efficiency and Returns to Scale of Telecommunication Enterprises by using DEA/WINDOW," (in Japanese) *Communications of the Operations Research Society of Japan* 37, pp.210-219.

50. D.B. Sun (1988), "Evaluation of Managerial Performance in Large Commercial Banks by Data Envelopment Analysis," Ph.D. Thesis (Austin, Texas: The University of Texas, Graduate School of Business). Also available from University Microfilms, Inc.

51. A. Charnes and W.W. Cooper (1991), "DEA Usages and Interpretations" reproduced in *Proceedings of International Federation of Operational Research Societies 12th Triennial Conference* in Athens, Greece, 1990.