Bootstrap Estimation for Nonparametric Efficiency Estimates*

Paul W. Wilson**

and

Léopold Simar***

December, 1995

*We have benefited from discussions with Gary Ferrier, Iréne Gijbels, and Max Stinchcombe. Shawna Grosskopf and Richard Grabowski graciously provided data used in the empirical examples. We alone, of course, are responsible for any remaining errors or omissions.

**Department of Economics, University of Texas, Austin, Texas 78712, USA. Research support from the Management Science Group, US Department of Veterans Affairs, is gratefully acknowledged.

***Institut de Statistique and CORE, Université Catholique de Louvain, Voie du Roman Pays 34, Louvain-la-Neuve, Belgium. Research support from the contract "Projet d'Actions de Recherche Concertées" (PARC No. 93/98–164) of the Belgian Government is gratefully acknowledged.

ABSTRACT

This paper develops a consistent bootstrap estimation procedure to obtain confidence intervals for nonparametric measures of productive efficiency. Although the methodology is illustrated in terms of technical efficiency measured by output distance functions, the technique can be easily extended to other consistent nonparametric frontier models. Variation in estimated efficiency scores is assumed to result from variation in empirical approximations to the *true* boundary of the production set. The bootstrap procedure also corrects for bias inherent in nonparametric efficiency estimators.

JEL Classification: C1, C4, C6.

Keywords: Bootstrap, distance function, efficiency, frontier models.

1. INTRODUCTION

A large literature concerning the measurement of efficiency in production has developed since Debreu (1951) and Farrell (1957) provided basic definitions for technical and allocative efficiency in production. Most attempts at estimating efficiency in production fall into two categories: (i) parametric approaches, including models with a noise term along the lines suggested by Aigner *et al.* (1977), Meeusen and van den Broeck (1977), and Jondrow *et al.* (1982), as well as models without a noise term such as those suggested by Winsten (1957), Greene (1980a,b), Stevenson(1980). and Seaver and Triantis (1989). and (ii) nonparametric approaches such as those proposed by Charnes *et al.* (1978, 1979), Färe and Lovell (1978), Burley (1980), Zieschang (1984), Deprins *et al.* (1984), and Färe *et al.* (1985). The nonparametric approaches which rely on convexity assumptions have been termed Data Envelopment Analysis (DEA), and typically rely on linear programming (LP) techniques for solution. The Free Disposal Hull (FDH) model introduced by Deprins *et al.* (1984) does not impose convexity and may be written as an LP, although LP methods are typically not used in computing FDH efficiency scores.

The parametric and nonparametric approaches to efficiency measurement have been labelled *econometric* and *determinisitic* (respectively) by Lovell (1993) and others, as if to suggest that the nonparametric models have no statistical underpinnings. Indeed, although there have been numerous applications of the nonparametric models, they invariably present point estimates of inefficiency, with no measure and only slight or no discussion of uncertainty surrounding these estimates.[1] Yet, in both the parametric and nonparametric approaches, efficiency is measured relative to an *estimate* of an underlying *true* frontier, conditional on observed data resulting from an underlying (and unobserved) data-generating process (DGP).[2] Hence, in both approaches, efficiency estimates are sub-

[1] For example, DEA methods have been used to measure efficiency among banks (Aly *et al.,* 1990; Wheelock and Wilson, 1995), utilities (Byrnes *et al.,* 1986; Färe *et al,* 1986; Färe *et al,* 1989), nursing homes (Nyman and Bricker, 1989; Dusansky and Wilson, 1994, 1995), and hospitals (Burgess and Wilson, 1993, 1995). See Lovell, 1993 for a list of additional DEA applications. Banker (1993) is one of the few to recognize the statistical nature of DEA estimators.

[2] We discuss the nature of the DGP implicit in nonparametric frontier models below and in more detail in Simar and Wilson (1995).

ject to uncertainty due to sampling variation.

This paper presents a consistent bootstrap estimation procedure for obtaining confidence intervals for nonparametric efficiency measures. In developing this methodology, we show that DEA and FDH models involve statistical *estimators* of the unobserved true efficiency. While these estimators are consistent, they are biased in finite samples. Our bootstrap procedure also provides a correction for this bias.

Accounting for sampling variation in nonparametric measures of efficiency is of great importance for typical applications, particularly when measured efficiency scores are to be used by administrators or policy makers to improve efficiency in production among a set of production units. In these settings, measured efficiency scores are typically used to identify units which will then be more closely scrutinized to find the actual source of inefficiency, and to propose ways of eliminating the inefficiency. This exercise is frequently costly, and resources will be wasted if attention is inadvertently directed at the wrong units. Consequently, merely labelling units as either efficient or inefficient, or ranking units by their measured efficiency, is not as useful as being able to say whether particular units are inefficient in a statistically significant sense, whether one unit is significantly more (or less) efficient than another, or being able to provide a confidence intervals for individual efficiency scores.

Simar (1991) appears to be the first to propose a bootstrap methodology for computing confidence intervals for efficiency scores obtained from nonparametric frontier models. Although Simar's bootstrap methodology for nonparametric models is presented in the context of the FDH method, it is easily extended to other nonparametric models. Simar's method views efficiency scores as regression residuals and involves resampling from estimated residual terms that represent stochastic noise. Results from Charnes and Cooper (1980), however, suggest that estimated residuals may not provide a good approximation of the underlying distribution in the case of DEA efficiency scores.

The skeptical reader might well ask, why concern ourselves with nonparametric efficiency models, given their lack of a noise term, when the parametric alternative is available?

We see several answers for such questions, and believe that both paradigms are useful. First, the tradeoffs between the two approaches are more complicated than the question suggests. As observed by Lovell (1993), while parametric models may include a noise term, they also require *a priori* specification of functional forms for both the technology and the noise process, as well as the inefficiency process, and thus confound the effects of misspecification of functional forms with inefficiency. The nonparametric approach is less prone to this type of specification error, but combines any noise that may be present in the data with inefficiency into a single measure. In some applications, data consist of averages over time (or over subunits within a production unit); in this situation, one can appeal to the central limit theorm to argue that the effects of noise have been diminished, reducing the need for an error term.[3]

Second, inclusion of a noise term in parametric models is not a panacea, in that it only allows for measurement error in the dependent variable. In many applications where parametric frontier models have been used (*e.g.,* cost functions for banks and hospitals), substantial measurement error may exist in the independent variables, resulting in biased and inconsistent parameter estimates, with unknown effects on the efficiency terms. Yet, we are aware of no productive efficiency studies which attempt to deal with this problem (probably because it is difficult). In this respect, the presence of a noise term in the parametric models may represent only a slight advantage, if at all. Moreover, careful researchers are aware of the potential for measurement error to affect nonparametric efficiency estimates, and a variety of techniques exist for detecting such errors so that they may be remedied (*e.g.,* Seaver and Triantis, 1989; Dusansky and Wilson, 1994; and Wilson, 1993, 1995).

Third, many researchers continue to use the nonparametric models, particularly in public sector applications where behavioral objectives are ill-defined. Nonparametric models are also useful in situations where units produce multiple outputs from multiple inputs,

[3]Ideally, one might want a nonparamtric model which incorporates noise. However, it does not appear possible to separate the effects of noise and inefficiency in a completely nonparametric model. Kneip and Simar (1995) propose a semi-parametric model, but their approach requires panel data with many time periods, and assumes constant inefficiency over time.

and where price information is either unavailable or unreliable (*e.g.,* as in the case of hospitals or nursing homes), thus preventing estimation of parametric cost functions. Although it is possible in principle to estimate parametric production models with multiple inputs and multiple outputs, data constraints and the resulting complexity of the models often make this approach infeasible.

Finally, in the end parametric models provide only a point estimate of inefficiency in the form of a conditional expectation, with unknown properties, whereas nonparametric models provide consistent estimates of efficiency. As with most nonparametric approaches, the rate of convergence for nonparametric efficiency estimates is rather slow, but as noted above, the bootstrap procedure developed below provides an estimate of the bias in finite samples. Moreover, our experience with empirical applications, also discussed below, indicates that the bias correction may typically be made without increasing mean-square error. Thus, while tradeoffs still exist between the parametric and nonparametric approaches, our bootstrap procedure provides a useful tool to those wishing to use the nonparametric models, and may make the nonparametric approach more palatable to some of those who might otherwise remain skeptical.

We proceed by first discussing in the next section bootstrap estimation in general terms, and the requirements for consistent estimation of bootstrap confidence intervals. The third section discusses consistent nonparametric efficiency estimation, using the Shephard (1970) output distance function for illustration. We demonstrate that the uncertainty in nonparametric efficiency estimation derives from uncertainty regarding the true boundary of the production set. Thus, the problem is related to the estimation of order statistics (Ripley and Rasson, 1977, consider a similar problem). Section four combines the results from sections two and three to develop our bootstrap estimation procedure. Empirical examples are given in section five. Although the discussion throughout is in terms of the output distance function, it is straightforward to adapt the techniques presented below to other nonparametric models, including DEA measures of allocative efficiency and overall efficiency such as those described by Färe *et al.* (1985). Conclusions are discussed in section five.

2. CONSISTENT BOOTSTRAP ESTIMATION

Bickel and Freedman (1981) and Beran and Ducharme (1991) have discussed the requirements for consistent bootstrap estimation. In this section we summarize results from Beran and Ducharme so that we may later demonstrate that our bootstrap estimation procedure is consistent.

Suppose an identically, independently distributed (iid) sample $\mathcal{Z}_N = \{z_1, \ldots, z_N\}$ is drawn from a sample space $\mathcal{X}$; each $z_i \in \mathcal{Z}_N$ has the same finite dimension; *i.e.,* $z_i \in \mathbb{R}^n$. Let $F_{\theta,N}$ denote the unknown exact finite sample distribution of z_i, $i = 1, \ldots, N$; $F_{\theta,N}$ is an unknown member of a parametric family $\{F_\theta | \theta \in \Theta\}$, where Θ is the parameter space in which the unknown parameter θ is restricted. The parameter θ may be either finite- or infinite-dimensional.[4]

Typically, interest lies in some function of θ such as $\tau = T(\theta)$, where $T(\theta) \in \mathcal{J}$, $\mathcal{J} = \{T(\theta)|\theta \in \Theta\}$. For a subset $\mathcal{C} \subset \mathcal{J}$, one may wish to test null hypotheses of the form $H_0 : \tau \in \mathcal{C}$. Alternatively, one might construct a confidence set by obtaining a subset of $\mathcal{J}$ covering the true value τ with high probability.

Assume there exists an estimator of θ, $\hat{\theta}_N = \theta(\mathcal{Z}_N)$, which is consistent on some metric d_Θ on Θ; *i.e.,* $\lim_{N\to\infty} \text{Prob}[d_\Theta(\hat{\theta}_N, \theta) > \varepsilon] = 0 \ \forall \ \varepsilon > 0$. Under certain regularity conditions (*e.g.,* continuity), given $\hat{\theta}_N$, $\hat{\tau}_N = \tau(\hat{\theta}_N)$ is a consistent estimator of τ. In order to construct a confidence set, consider a root $R_N = R_N(\mathcal{Z}_N, \tau)$ for τ with left-continuous cumulative distribution function $H_N(v, \theta) = \text{Prob}[R_N(\mathcal{Z}_N, \tau) < v|\theta, N]$ such that $H_N(v, \theta)$ has nondegenerate weak limit $H_A(v, \theta)$.[5] Thus, the desired confidence set is

$$\mathcal{C}_\tau = \{\tau | R_N(\mathcal{Z}_N, \tau) \leq c_\alpha\}. \tag{2.1}$$

The problem is to determine the critical value c_α such that $\text{Prob}[\tau \in \mathcal{C}_\tau|\theta, N]$ is close to α, the specified level of the coverage probability.

[4]It is understood that in the infinite-dimensional case, F is nonparametric.

[5]The requirement of nondegeneracy in the weak limit is not a real constraint here, as we can always avoid the condition by appropriately scaling the root.

If θ were known, an ideal choice for the critical value would be $H_N^{-1}(\alpha, \theta)$, *i.e.,* the largest αth quantile of $H_N(\cdot, \theta)$ which results in $\text{Prob}[\tau \in \mathcal{C}_\tau | \theta, N] = \alpha$.[6] Then an ideal confidence set of level α for τ is given by

$$\begin{aligned} \mathcal{C}_\tau &= \{\tau | R_N(\mathcal{Z}_N, \tau) \leq H_N^{-1}(\alpha, \theta)\} \\ &= \{\tau | H_N(R_N, \theta) \leq \alpha\}. \end{aligned} \tag{2.2}$$

The second line of (2.2) follows from the definition of the quantile function.

Unfortunately, in many applications, θ and hence $H_N(\cdot, \theta)$ are not known. The conventional, asymptotic solution is to approximate $H_N(\cdot, \theta)$ by its weak limit $H_A(\cdot, \theta)$, and estimate $H_A(\cdot, \theta)$ by $\hat{H}_A(\cdot) = H_A(\cdot, \hat{\theta}_N)$. This approach, however, requires parametric knowledge of the limiting distribution $H_A(\cdot, \theta)$ and its inverse, which may also be unknown. Alternatively, the exact distribution $H_N(\cdot, \theta)$ can be estimated by the *bootstrap distribution* of the root R_N,

$$\hat{H}_B(v) = H_N(v, \hat{\theta}_N) = \text{Prob}[R_N^* < v | \mathcal{Z}_N], \tag{2.3}$$

where $R_N^* = R_N(\mathcal{Z}_N^*, \hat{\tau}_N)$ and the distribution of the bootstrap sample $\mathcal{Z}_N^*$, given the original sample $\mathcal{Z}_N$, is $F_{\hat{\theta}_N, N}$, a consistent estimate of $F_{\theta, N}$. Since the bootstrap distribution in (2.3) does not involve any unknown parameters, $\hat{H}_B(\cdot)$ may be used as an estimate of $H_N(\cdot, \theta)$ and $\hat{H}_B^{-1}(\alpha)$ may be computed as an approximation to $H_N^{-1}(\alpha, \theta)$. Thus, the estimated bootstrap confidence interval is

$$\begin{aligned} \mathcal{C}_B &= \{\tau | R_N \leq \hat{H}_B^{-1}(\alpha)\} \\ &= \{\tau | H_N(R_N, \hat{\theta}_N) \leq \alpha\}. \end{aligned} \tag{2.4}$$

Typically, the exact analytical expression for the bootstrap distribution and its quantile function cannot be derived analytically; hence, $\hat{H}_B^{-1}(\alpha)$ must be approximated by Monte-Carlo methods or in some cases by asymptotic expansions.

Bickel and Freedman (1981), Beran (1984), Beran and Ducharme (1991), and others have given general conditions under which $\mathcal{C}_B$ is consistent, asymptotically giving the desired coverage probability. These conditions may be summarized as follows:

[6] Here we use the notation $H_N^{-1}(\alpha, \theta)$ to denote the quantile function; H^{-1} does not denote the inverse of H.

THEOREM 2.1: *Let d_Θ be a metric defined on Θ. Suppose that, for any $\varepsilon > 0$ and $\theta \in \Theta$,*

(i) $\lim_{N\to\infty} \text{Prob}[d_\Theta(\hat{\theta}_N, \theta) > \varepsilon] = 0$;

(ii) For any sequence $\{\hat{\theta}_N\}$ in Θ such that $\lim_{N\to\infty} d_\Theta(\hat{\theta}_N, \theta) = 0$, the distribution $H_N(\cdot, \hat{\theta}_N)$ converges weakly to $H_A(\cdot, \theta)$; and

(iii) The weak limit $H_A(v, \theta)$ is continuous in v for any $\theta \in \Theta$.

Then, for $\hat{H}_B(\cdot)$ computed as in (2.3) and any $\varepsilon > 0$ and $\theta \in \Theta$, the following statements are true:

(a) $\lim_{N\to\infty} \text{Prob}\left[\|\hat{H}_B(\cdot) - H_A(\cdot, \theta)\|_s > \varepsilon | \mathcal{Z}_N\right] = 0$, *where* $\|H(\cdot)\|_s = \sup |H(v)|,\ v \in \mathbb{R}^1$ *denotes the sup-norm, and*

(b) $\lim_{N\to\infty} \text{Prob}[\tau \in \mathcal{C}_B | \mathcal{Z}_N] = \alpha$.

A proof of this theorem appears in Beran and Ducharme (pp. 15–19).

Generation of the bootstrap sample $\mathcal{Z}_N^*$ used to compute $\hat{H}_B(\cdot)$ in (2.3) is a crucial step in bootstrap estimation. In many cases, this may be accomplished by independently sampling from the empirical distribution function of the sample, $\hat{F}_N(z) = N^{-1} \sum_{i=1}^{N} I\{z_i < z\}$, where $I\{\cdot\}$ is the indicator function to form the bootstrap sample $\mathcal{Z}_N^*$, *i.e.*, by independently resampling from $\mathcal{Z}_N$ with replacement such that each $z_i \in \mathcal{Z}_N$ has equal probability of selection.

However, this approach sometimes prevents the weak convergence required in condition (ii) of Theorem 2.1 for consistent bootstrap estimation. The following example illustrates this point, and has particular relevance to sections three and four below.[7]

EXAMPLE 2.1: Let $\mathcal{Z}_N = \{z_i\}_{i=1}^N$, $z_i \in \mathbb{R}^1\ \forall\ i = 1, \ldots, N$ be iid random variables with continuous distribution F supported on the closed interval $[0, \tau]$, where τ is an unknown parameter to be estimated from the data. In this example, θ represents F itself, since F is nonparametric; the parameter of interest is τ, a function of F. The maximum

[7] Example 2.1 has been used by Bickel and Freedman (1981), Beran and Ducharme (1991), and Efron and Tibshirani (1993) to demonstrate bootstrap failure. Other examples are given by Bickel and Freedman (1981), Beran (1982), Bretagnolle (1983), Beran and Srivastava (1985, 1987), and Athreya (1987).

likelihood estimate of τ is $\hat{\tau}_N = \max\{z_i \in \mathcal{Z}_N\}$. Suppose the corresponding bootstrap value is computed as $\hat{\tau}^*_N = \max\{z^*_i \in \mathcal{Z}^*_N\}$, where the elements of $\mathcal{Z}^*_N = \{z^*_1, \ldots, z^*_N\}$ are, conditionally on $\mathcal{Z}_N$, iid from the empirical distribution placing mass $1/N$ on each element of $\mathcal{Z}_N$. Consider the root

$$R_N = \frac{N\left(\tau - \hat{\tau}_N\right)}{\tau} \tag{2.5}$$

and its bootstrap counterpart

$$R^*_N = \frac{N\left(\hat{\tau}_N - \hat{\tau}^*_N\right)}{\hat{\tau}_N}. \tag{2.6}$$

One can easily show that

$$\text{Prob}\left(\hat{\tau}_N = \hat{\tau}^*_N\right) = 1 - \left(1 - N^{-1}\right)^N, \tag{2.7}$$

and hence $\lim_{N\to\infty} \text{Prob}\left(\hat{\tau}_N = \hat{\tau}^*_N\right) = 1 - e^{-1} \approx 0.632$. However, if F is uniform on $[0, \tau]$, then $H_A(v, \theta) = 1 - e^{-v}$, so that $H_A(0, \theta) = 0$. Yet (2.7) shows that $H_N(\cdot, \hat{\theta}_N)$ does not converge to $H_A(\cdot, \theta)$, and therefore does not provide a reasonable approximation to $H_N(\cdot, \theta)$. In terms of Theorem 2.1, condition (ii) is not met. The problem in (2.7) is not unique to the uniform distribution, and in fact will occur with any continuous distribution F on $[0, \tau]$. ■

The apparent failure of the bootstrap in the above example is actually not a failure of the bootstrap method per se, but rather a failure of the empirical density, in placing mass $1/N$ on each element in $\mathcal{Z}_N$, to give a good estimate of the underlying density $f = dF$ in the extreme tail. Fortunately, there are other resampling methods which result in a bootstrap distribution that converges to the appropriate weak limit in the above example. Zelterman (1993) uses a semiparametric technique based on differences between adjacent order statistics to bootstrap the largest k order statistics from a sample $\mathcal{Z}_N$. Swanepoel (1986) proves for the above example that the bootstrap distribution will converge to the appropriate weak limit when bootstrap samples are of size M_N, where for a sequence $\{M_N\}$ and some $\varepsilon \in (0, 1)$, $O(N^\varepsilon) \leq M_N \leq o\left(\left(N^{1+\varepsilon)}/\log(N)\right)^{1/2}\right)$.

Of course, one could also resample from a consistent parametric estimate of F to ensure that the bootstrap sample $\mathcal{Z}^*_N$ has distribution $F_{\hat{\theta}_N, N}$, a consistent estimate of F, as discussed by Bickel and Freedman (1981) and Efron and Tibshirani (1993). In fact, resampling from any smooth, consistent estimate of $f = dF$ (parametric or otherwise) will yield a bootstrap distribution that converges to the appropriate weak limit in the example above (this is the idea behind the smooth bootstrap examined by Silverman, 1986, and Hall *et al.,* 1989). Thus, in cases where the parametric form of f is not known *a priori,* one could use a kernel estimate of f and resample from the kernel estimate, since kernel density estimates are consistent under rather general conditions (*e.g.,* see Parzen, 1962), provided one deals appropriately with the boundary conditions (*e.g.,* see Silverman, 1986). This is the approach taken in section four.

3. CONSISTENT NONPARAMETRIC EFFICIENCY ESTIMATION

To illustrate nonparametric estimation of efficiency in production, define the production set

$$\Psi \equiv \{(\boldsymbol{x}, \boldsymbol{y}) | \boldsymbol{x} \text{ can produce } \boldsymbol{y}\}, \tag{3.1}$$

where $\boldsymbol{x} \in \mathbb{R}_+^n$ represents a vector of inputs and $\boldsymbol{y} \in \mathbb{R}_+^m$ represents a vector of outputs. The boundary of Ψ is sometimes refered to as a technology; DEA and FDH methods attempt to measure technical efficiency relative to an estimate of the boundary of Ψ. The production technology implied by Ψ may be characterized by either the output feasibility set

$$\mathcal{P}(\boldsymbol{x}) \equiv \{\boldsymbol{y} | (\boldsymbol{x}, \boldsymbol{y}) \in \Psi\} \tag{3.2}$$

or the input requirement set

$$\mathcal{L}(\boldsymbol{y}) \equiv \{\boldsymbol{x} | (\boldsymbol{x}, \boldsymbol{y}) \in \Psi\}. \tag{3.3}$$

Knowledge of either $\mathcal{P}(\boldsymbol{x})$ for all $\boldsymbol{x}$ or $\mathcal{L}(\boldsymbol{y})$ for all $\boldsymbol{y}$ is equivalent to knowledge of Ψ; Ψ implies (and is implied by) both $\mathcal{P}(\boldsymbol{x})$ and $\mathcal{L}(\boldsymbol{y})$. Thus, both $\mathcal{P}(\boldsymbol{x})$ and $\mathcal{L}(\boldsymbol{y})$ inherit the properties of Ψ. Various assumptions regarding properties of the production set Ψ; we adopt those of Shephard (1970) and Färe (1988), although our bootstrap methodology does not depend on these assumptions. Specifically, we assume: (i) Ψ is a closed, convex set; (ii) $(\boldsymbol{x}, \boldsymbol{y}) \notin \Psi$ if $\boldsymbol{y} \geq 0, \boldsymbol{y} \neq 0$, *i.e.,* all production requires use of some inputs; and

(iii) for $\tilde{\boldsymbol{x}} \geq \boldsymbol{x}$, $\tilde{\boldsymbol{y}} \leq \boldsymbol{y}$, if $(\boldsymbol{x}, \boldsymbol{y}) \in \Psi$ then $(\tilde{\boldsymbol{x}}, \boldsymbol{y}) \in \Psi$ and $(\boldsymbol{x}, \tilde{\boldsymbol{y}}) \in \Psi$, *i.e.,* both inputs and outputs are strongly disposable.

One nonparametric efficiency estimator is the Shephard (1970) output distance function, defined for firm i as[8]

$$D_i^{out} \equiv \inf\{\theta > 0 | \boldsymbol{y}_i/\theta \in \mathcal{P}(\boldsymbol{x})\}. \tag{3.4}$$

Clearly, $D_i^{out} \leq 1$, with $D_i^{out} = 1$ indicating that the ith firm lies on the boundary of Ψ and thus is technically efficient. The Shephard (1970) input distance function is another nonparametric efficiency estimator, defined for firm i as

$$D_i^{in} \equiv \sup\{\theta > 0 | \boldsymbol{x}_i/\theta \in \mathcal{L}(\boldsymbol{y})\}. \tag{3.5}$$

Clearly, $D_i^{in} \geq 1$, with $D_i^{in} = 1$ indicating that the ith firm lies on the boundary of Ψ and thus is technically efficient. Unfortunately, neither $\mathcal{P}(\boldsymbol{x})$ nor $\mathcal{L}(\boldsymbol{y})$ are observed. Consequently, the distance functions D_i^{out} and D_i^{in} are also not observed, and hence must be estimated, which involves estimating either $\mathcal{P}(\boldsymbol{x})$ or $\mathcal{L}(\boldsymbol{y})$.[9]

This situation is typical for nonparametric efficiency estimators. To conserve space, we use the output distance function in (3.4) to illustrate how bootstrapping techniques may be used to obtain confidence intervals for nonparametric efficiency estimates. The results are easily extended to the case of input distance functions and other nonparametric efficiency estimators.

For an observed finite sample $\mathcal{Z}_N = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^N$ drawn from Ψ, the *true* output feasibility set in (3.2) subject to the assumptions (i)–(iii) above has been frequently estimated by

$$\widehat{\mathcal{P}}_N(\boldsymbol{x}) = \{\boldsymbol{y} | \boldsymbol{y} \leq \boldsymbol{Y}\boldsymbol{q}, \boldsymbol{x} \geq \boldsymbol{X}\boldsymbol{q}, \ \overrightarrow{\mathbf{1}}\,\boldsymbol{q} \ = 1, \boldsymbol{q} \in \mathbb{R}_+^N\}, \tag{3.6}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1 \ \ldots \ \boldsymbol{y}_N]$, $\boldsymbol{X} = [\boldsymbol{x}_1 \ \ldots \ \boldsymbol{x}_N]$, with each $\boldsymbol{x}_i$, $\boldsymbol{y}_i$ $i = 1, \ldots, N$ denoting the $(n \times 1)$ and $(m \times 1)$ vectors of observed inputs and outputs, respectively, and where $\overrightarrow{\mathbf{1}}$ is a

[8] We shall refer to observations $i = 1, \ldots, N$ as firms, although in practice the observations might represent production units within a single firm, or other production units which do not fit the usual notion of a firm.

[9] The Shephard output distance function may be shown to be the reciprocal of the Debreu-Farrell output oriented measure of technical efficiency, while the Shephard input distance function is the reciprocal of the Debreu-Farrell input oriented measure of technical efficiency.

$(1 \times N)$ vector of ones and $\boldsymbol{q} = [q_1 \ \dots \ q_N]'$ is a $(N \times 1)$ vector of intensity variables which serve to form a piecewise linear estimate of the production set boundary. The constraint $\overrightarrow{\mathbf{1}}\boldsymbol{q} = 1$ imposes variable returns to scale on the reference technology; other returns to scale may be imposed by modifying this constraint (*e.g.,* see Grosskopf, 1986). In other instances, $\mathcal{P}(\boldsymbol{y})$ has been estimated by

$$\widetilde{\mathcal{P}}_N(\boldsymbol{x}) = \{\boldsymbol{y} | \boldsymbol{y} \leq \boldsymbol{Yq}, \boldsymbol{x} \geq \boldsymbol{Xq}, \overrightarrow{\mathbf{1}}\boldsymbol{q} = 1, q_i \in \{0,1\} \ \forall \ i = 1, \dots, N\}, \tag{3.7}$$

where the constraint $q_i \in \{0,1\} \ \forall \ i = 1, \dots, N$ has been added.

Note that $\widehat{\mathcal{P}}_N(\boldsymbol{x})$ implies $\hat{\Psi}_N$, an estimate of Ψ, where $\hat{\Psi}_N$ is the convex hull of the sample points in $\mathcal{Z}_N$. Also note that $\widehat{\mathcal{P}}_N(\boldsymbol{x}) \subseteq \mathcal{P}(\boldsymbol{x})$, which implies $\hat{\Psi}_N \subseteq \Psi$; hence, $\widehat{\mathcal{P}}_N(\boldsymbol{x})$ and $\hat{\Psi}_N$ are said to be biased downward. Similarly, $\widetilde{\mathcal{P}}_N(\boldsymbol{x})$ implies $\tilde{\Psi}_N$, an estimate of Ψ, where the boundary of $\tilde{\Psi}_N$ is the free disposal hull (FDH) of the sample points in $\mathcal{Z}_N$ (first used in efficiency estimation by Deprins *et al.,* 1984, to allow for nonconvexity). Furthermore, $\widetilde{\mathcal{P}}_N(\boldsymbol{x}) \subseteq \mathcal{P}(\boldsymbol{x})$, which implies $\tilde{\Psi}_N \subseteq \Psi$; hence, $\widetilde{\mathcal{P}}_N(\boldsymbol{x})$ and $\tilde{\Psi}_N$ are also biased downward. In addition, $\widetilde{\mathcal{P}}_N(\boldsymbol{x}) \subseteq \widehat{\mathcal{P}}_N(\boldsymbol{x})$. Korostelev *et al.* (1995a, 1995b) examine the downward bias of the DEA and FDH estimators of Ψ and show that both $\widehat{\mathcal{P}}_N(\boldsymbol{x})$ and $\widetilde{\mathcal{P}}_N(\boldsymbol{x})$ are consistent estimators of $\mathcal{P}(\boldsymbol{x})$, with the rates of convergence decreasing as the number of dimensions specified for the input/output space is increased.

Given the data $\mathcal{Z}_N$ and the estimated output feasibility set in (3.6), the *true* output distance function D_i^{out} in (3.4) may be estimated by solving

$$\hat{D}_i^{out} = \max\{\lambda | \boldsymbol{y}_i/\lambda \in \widehat{\mathcal{P}}_N(\boldsymbol{x})\} \tag{3.8}$$

or, given (3.6),

$$\left(\hat{D}_i^{out}\right)^{-1} = \max\{\lambda | \lambda \boldsymbol{y}_i \leq \boldsymbol{Yq}_i, \boldsymbol{x}_i \geq \boldsymbol{Xq}_i, \overrightarrow{\mathbf{1}}\boldsymbol{q}_i = 1, \boldsymbol{q}_i \in \mathbb{R}_+^K\}, \tag{3.9}$$

which is the familiar LP specification for estimating the output distance function.[10] By construction, $\lambda \geq 1$, and hence $0 < \hat{D}_i^{out} \leq 1$, with $\hat{D}_i^{out} = 1$ indicating that firm i is

[10]The DEA literature contains numerous instances where solving an LP similar to (3.9) is refered to as *computing,* as opposed to *estimating,* efficiency. The FDH distance function estimate of efficiency can be obtained by replacing $\widehat{\mathcal{P}}_N(\boldsymbol{x})$ in (3.8) with $\widetilde{\mathcal{P}}_N(\boldsymbol{x})$, although actual computations of FDH efficiency estimates usually do not rely on solution of an LP as in (3.9). In addition, when the FDH estimator is used, the assumption regarding convexity of Ψ is usually dropped; this has no effect on the analysis here, however, provided the boundary of Ψ is monotone.

ostensibly efficient (if $\hat{D}_i = 1$ for firm i, the firm is said to be *ostensibly* efficient since efficiency is being measured relative to the boundary of $\hat{\Psi}_N$ rather than the boundary of the true production set, Ψ). Since $\widehat{\mathcal{P}}_N(\boldsymbol{x})$ and $\hat{\Psi}_N$ are biased downward, $\hat{D}_i^{out}$ must be biased upward; therefore, given a finite sample, the estimated distance function value $\hat{D}_i^{out}$ will overstate the true technical efficiency of firm i. However, $\hat{D}_i^{out}$ is a consistent estimator of D_i^{out} (see Banker, 1993).

To illustrate the output distance function and its estimator $\hat{D}^{out}$, consider a set of firms producing two output quantities $y_1,\ y_2$ from the same inputs as shown in Figure 1. Figure 1 shows two boundaries: the *true* boundary of the set $\mathcal{P}(\boldsymbol{x})$ and the *estimated* boundary correpsonding to $\widehat{\mathcal{P}}(\boldsymbol{x})$. For firm A, $D^{out} = OA/OA''$ and $\hat{D}^{out} = OA/OA'$, with $0 < D^{out} < \hat{D}^{out} < 1$. Given firm A, the stochastic element of the estimator $\hat{D}^{out}$ is due to the distance OA', which is determined by $\widehat{\mathcal{P}}_N(\boldsymbol{x})$. Thus, estimating the distance function in (3.4) is a matter of estimating a boundary much like the example considered in the previous section. In fact, in the trivial case where $n = 0,\ m = 1$, the output distance function can be expressed as $D^{out} \equiv \{\lambda > 0 | y/\lambda \leq \tau, \}$ and its corresponding estimate as $\hat{D}^{out} \equiv \{\lambda > 0 | y/\lambda \leq \hat{\tau}, \}$. In this case, the production set (and the sample space) consist of the interval $[0, \tau]$ in $\mathbb{R}^1$. Furthermore, estimation of the distance function in this case is equivalent to estimation of a function of the parameter τ from a univariate sample distributed on $[0, \tau]$, which is analogous to the case considered in Example 2.1.

4. BOOTSTRAPPING DISTANCE FUNCTION ESTIMATES

To combine the results from the previous two sections, assume once again that a finite sample $\mathcal{Z}_N = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}_{i=1}^N$ is drawn from the unobserved production set Ψ. The distance function estimator $\hat{D}_i^{out}$ is consistent, satisfying condition (i) in Theorem 2.1. In addition, assume that within Ψ, firms deviate from the boundary of Ψ in a radial direction such that the distance measured by $\hat{D}^{out}$ in (3.4) has unknown, continuous distribution $F(v)$; continuity in F ensures that condition (iii) in Theorem 2.1 is satisfied (provided the root R_N is scaled by the appropriate power of N). Finally, as noted near the end of section two, resampling from a smooth, consistent estimate of f (as discussed below) will ensure that the resulting bootstrap distribution converges to the appropriate weak limit, satisfying condition (ii) of Theorem 2.1. With these conditions satisfied, our bootstrap estimation will be consistent; the task in this section is to describe the implimentation of the estimation procedure.[11]

As noted earlier, bootstrapping involves replicating the DGP implied by F to generate a bootstrap sample $\mathcal{Z}_N^* = \{(\boldsymbol{x}_i^*, \boldsymbol{y}_i^*)\}_{i=1}^N$ (as discussed below) and computing the distance function estimator from this pseudo sample to obtain a bootstrap estimate $\hat{D}_i^{out^*}$. This is accomplished by rewriting (3.9) as

$$\left(\hat{D}_i^{out^*}\right)^{-1} = \max\{\lambda | \lambda \boldsymbol{y}_i \leq \boldsymbol{Y}^* \boldsymbol{q}_i, \boldsymbol{x}_i \geq \boldsymbol{X}^* \boldsymbol{q}_i, \overrightarrow{\boldsymbol{1}}\, \boldsymbol{q}_i = 1, \boldsymbol{q}_i \in \mathbb{R}_+^N\}, \tag{4.1}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1^* \ \ldots \ \boldsymbol{y}_N^*]$, $\boldsymbol{X} = [\boldsymbol{x}_1^* \ \ldots \ \boldsymbol{x}_N^*]$, and the remaining variables are defined as before.

Repeating this process B times yields B bootstrap estimates $\hat{D}_i^{out^*}(b)$, $b = 1, \ldots, B$. As noted earlier, $\hat{D}_i^{out}$ is necessarily biased upward relative to D_i^{out} since $\hat{\Psi} \subseteq \Psi$. Similarly, $\hat{D}_i^{out^*}$ is biased upward relative to $\hat{D}_i^{out}$ since $\hat{\Psi}^* \subseteq \hat{\Psi}$; furthermore, $\hat{D}_i^{out} \leq \hat{D}_i^{out^*}(b) \leq 1 \ \forall \ b = 1, \ldots, B$. Thus, the bootstrap values $\{\hat{D}_i^{out^*}(b)\}_{b=1}^B$ provide a biased approximation of the sampling distribution of $\hat{D}_i$. The *bias* of $\hat{D}_i^{out}$ is defined as the difference between the expectation of $\hat{D}_i^{out}$ and the true value D_i^{out}. The bootstrap bias estimate (Efron, 1979; Efron and Tibshirani, 1993) is given by

$$\widehat{bias}_i = B^{-1} \sum_{b=1}^{B} \hat{D}_i^{out^*}(b) - \hat{D}_i^{out}, \tag{4.2}$$

and so a bias-corrected estimate of D_i^{out} may be computed as

$$\hat{\hat{D}}_i^{out} = \hat{D}_i^{out} - \widehat{bias}_i. \tag{4.3}$$

Of course, the bias-corrected estimate $\hat{\hat{D}}_i^{out}$ may have larger mean square error than the uncorrected estimate $\hat{D}_i^{out}$; hence, the correction in (4.3) should be used with caution.

[11] Alternatively, one might assume the data are distributed in $(n+m)$-space by some unknown distribution and attempt to compute a kernel density estimate of this $(m+n)$-dimensional distribution. However, the number of observations required to obtain a reasonable estimate is likely to make the entire exercise infeasible.

However, if the estimated variance of $\hat{D}_i^{out}$ is much less than the square of the estimated bias, the bias correction in (4.3) is unlikely to increase mean square error. The variance of $\hat{D}_i^{out}$ may be estimated by the sample variance of the $\{\hat{D}_i^{out^*}(b)\}_{b=1}^{B}$.

The bootstrap values $\{\hat{D}_i^{out^*}(b)\}_{b=1}^{B}$, together with the bias estimate in (4.2), may be used to obtain a bias-corrected approximation to the sampling distribution of the bias-corrected distance function estimate $\hat{\hat{D}}_i^{out}$. To remove the bias, compute

$$\hat{\hat{D}}_i^{out^*}(b) = \hat{D}_i^{out^*}(b) - \left(2 \times \widehat{bias}_i\right) \quad \forall\, b = 1 \ldots, B. \tag{4.4}$$

Note that merely subtracting $\widehat{bias}_i$ in (4.4) would center the distribution of the $\hat{D}_i^{out^*}$ on $\hat{D}_i^{out}$, but $\hat{D}_i^{out}$ must also be adjusted to remove bias; hence $2 \times \widehat{bias}_i$ is the correct adjustment factor. Thus the values $\{\hat{\hat{D}}_i^{out^*}(b)\}_{b=1}^{B}$ provide a bias-corrected approximation of the sampling distribution of $\hat{\hat{D}}_i^{out}$.[12]

Efron's (1982) percentile method involves sorting the values $\{\hat{\hat{D}}_i^{out^*}(b)\}_{b=1}^{B}$ by algebraic value and deleting $(1-\alpha/100)/2 \times B$ elements from either end of the sorted array to obtain two-sided α-percent confidence intervals for the distance function value D_i^{out}.[13] However, some of the values in $\{\hat{\hat{D}}_i^{out^*}(b)\}_{b=1}^{B}$ may be greater than the original value of the estimate $\hat{D}_i^{out}$. Since all estimation is conditional on the sample data, the true distance function value D_i^{out} cannot exceed the observed value $\hat{D}_i^{out}$. Thus, the appropriate confidence intervals, conditional on the estimated value $\hat{D}_i^{out}$ (which is merely a function of the observed data) may be obtained by first truncating the sorted values $\{\hat{\hat{D}}_i^{out^*}(b)\}_{b=1}^{B}$ by deleting all values greater than $\hat{D}_i^{out}$. Then, two-sided α-percent confidence intervals for the distance function value D_i^{out} may be obtained by deleting $(1-\alpha/100)/2 \times B'$ elements from either end of the truncated, sorted array, where B' denotes the number of elements in $\{\hat{\hat{D}}_i^{out^*}(b) | \hat{\hat{D}}_i^{out^*}(b) \leq \hat{D}_i^{out},\ b = 1, \ldots, B\}$.[14]

[12] The reader can indeed verify that $B^{-1} \sum_{b=1}^{B} \hat{\hat{D}}_i^{out^*}(b) = \hat{\hat{D}}_i^{out}$ to demonstrate that the correct adjustment factor is used in (4.4).

[13] Alternatively, one could employ the bias-corrected and accelerated (BC_a) described by Efron and Tibshirani (1993). However, since additional jackknife estimates are required to estimate the acceleration parameter in this method, the computational burden is likely to be too great for datasets with more than 100–200 observations.

[14] In effect, we are merely conditioning on the original data. The values in $\{\hat{\hat{D}}_i^{out^*}(b)\}_{b=1}^{B}$ approximate

As discussed in section two, the replication of the DGP is a crucial step in the bootstrap methodology. As also discussed in the previous sections, resampling from a kernel estimate of the denstity $f = dF$ will ensure consistency of the bootstrap estimates for the output distance function. The naive kernel density estimator is given by

$$\hat{f}(v|w) = \frac{1}{Nw} \sum_{i=1}^{N} K \left[w^{-1} \left(v - \hat{D}_i^{out} \right) \right], \tag{4.5}$$

where $K(\cdot)$ is the kernel function and w is a fixed constant (termed the "bandwidth"). The kernel function K is typically specified as some parametric density function. Unfortunately, the naive kernel density estimator ignores the boundary conditions $0 < D_i^{out} \leq 1$, and will underestimate the density f at the right boundary.[15]

The reflection method proposed by Silverman (1986) avoids the boundary problem. First, for a set of output distance function estimates $\{\hat{D}_i^{out}\}_{i=1}^N$, reflect each element of the set about unity to create a set $\{\hat{D}_i^{out},\ 2 - \hat{D}_i^{out}\}_{i=1}^N$. Then a consistent estimator of $f = dF$ is given by

$$\tilde{f}(u|w) = \begin{cases} 2\hat{h}(v|w) & \text{for } 0 < v \leq 1 \\ 0 & \text{otherwise,} \end{cases} \tag{4.6}$$

where

$$\hat{h}(v|w) = \frac{1}{2Nw} \sum_{i=1}^{N} \left\{ K \left[w^{-1} \left(v - \hat{D}_i^{out} \right) \right] + K \left[w^{-1} \left(v - 2 + \hat{D}_i^{out} \right) \right] \right\}. \tag{4.7}$$

To generate pseudo random deviates from $\tilde{f}$, construct $\{\delta_i^*\}_{i=1}^N$ by drawing independently from $\{\hat{D}_i^{out}\}_{i=1}^N$ N times with repacement, such that each element of $\{\hat{D}_i^{out}\}_{i=1}^N$ has equal probability of selection on each draw. Let $d_i = \delta_i^* + w\varepsilon_i$, where ε_i is drawn from $K(\cdot)$. Similarly, let $d_i' = 2 - \delta_i^* - w\varepsilon_i$, the reflection of d_i about unity. Using the convolution theorem, it is easy to show that

$$d_i \sim \hat{h}_1(v|w) = \frac{1}{Nw} \sum_{i=1}^{N} K \left[w^{-1} \left(v - \hat{D}_i^{out} \right) \right] \tag{4.8}$$

the distribution of the unbiased *estimator* (a random variable), while the value computed via (3.9) for $\hat{D}_i^{out}$ is an *estimate* (*i.e.,* a realization of the random variable).

[15] Typically there will be no observations at the left boundary. Thus, the density f will be zero at the left boundary, which does not create a problem for the naive kernel estimator.

and

$$d_i' \sim \hat{h}_2(v|w) = \frac{1}{Nw} \sum_{i=1}^{N} K\left[w^{-1}\left(v - 2 + \hat{D}_i^{out}\right)\right], \tag{4.9}$$

with

$$\hat{h}(v|w) = \hat{h}_1(v|w) + \hat{h}_2(v|w). \tag{4.10}$$

Therefore, setting

$$\gamma_i = \begin{cases} \delta_i^* + w\varepsilon_i & \text{if } \delta_i^* + w\varepsilon_i \leq 1 \\ 2 - \delta_i^* - w\varepsilon_i & \text{otherwise} \end{cases} \tag{4.11}$$

yields values $\gamma_i \sim \tilde{f}(v|w)$.

The γ_i generated from (4.11) have variance $V(\gamma_i|\hat{D}_1^{out}, \ldots, \hat{D}_N^{out}) = \hat{\sigma}^2 + w^2$, where $\hat{\sigma}^2$ is the sample variance of the original distance function estimates $\{\hat{D}_i^{out}\}_{i=1}^N$. As is typical with kernel estimators, the γ_i must be rescaled by computing

$$\gamma_i^* = \bar{\delta}^* + \left(1 + w^2\hat{\sigma}^{-2}\right)^{-1/2}\left(\gamma_i - \bar{\delta}^*\right), \tag{4.12}$$

where $\bar{\delta}^* = N^{-1}\sum_{i=1}^N \delta_i^*$. The variance of the transformed deviates γ_i^* is

$$V(\gamma_i^*|\hat{D}_1^{out}, \ldots, \hat{D}_N^{out}) = \hat{\sigma}^2\left[1 + w^2N^{-1}(\hat{\sigma}^2 + w^2)^{-1}\right], \tag{4.13}$$

which will approach $\hat{\sigma}^2$ as the sample size is increased.

The remaining issue for the kernel estimation is the choice of the bandwidth w. Tapia and Thompson (1978), Silverman (1978, 1986), and Härdle (1990) discuss considerations relevant to the choice of w; in general, for a given sample size, larger values of w produce more diffuse (*i.e.,* less efficient) estimates of the density, while very small values produce estimated densities with multiple modes. Thus a natural approach suggested by Efron and Tibshirani (1993) is to sequentially test for the number of modes $m = 1, 2, \ldots$ until the null hypothesis H_0: number of modes $= m$ cannot be rejected, then setting w equal to the smallest bandwidth that yields the number of modes given by the last value of m. This can be accomplished in several steps:

[1] Set $m = 1$.

[2] Find the smallest bandwidth $\hat{w}_m$ which yields m modes in order to test the null hypothesis H_0: number of modes $= m$.

[3] Draw B bootstrap samples of size N from $\tilde{f}(v|\hat{w}_m)$ using (4.12).

[4] For each bootstrap sample drawn in step [3], compute $\hat{w}_m^*$, the smallest bandwidth yielding a density estimate with m modes.

[5] Approximate the P-value (achieved significance level) by $\#\{\hat{w}_m^* \geq \hat{w}_m\}/B$. If the null hypothesis in [1] cannot be rejected, then set $w = \hat{w}_m^*$ and stop; otherwise, increment m by 1 and go to step [2].[16]

An alternative approach to bandwidth selection is to plot estimated densities for various bandwidths and choose one that looks somehow sensible. However, Silverman (1986) discusses the desirability of having an "automatic" bandwidth selection criteria which would allow two researchers independently examining the same data to arrive at the same value for w. The bootstrap test for multimodality described above is one such automatic criterion. Silverman (1986) suggests setting $w = 0.9AN^{-1/5}$, where

$$A = \min(\hat{\sigma}, R/1.34), \tag{4.14}$$

$\hat{\sigma}$ is the standard deviation of the distance function estimates and R is the interquartile range of the estimates. An alternative approach would be to use cross-validation methods, but these are more complex and computationally intensive.

Using the pseudo random deviates γ_i^* $i = 1, \ldots, N$ generated from (4.12), pseudo data $(\boldsymbol{x}_i^*, \boldsymbol{y}_i^*)$, where $\boldsymbol{x}_i^* = \boldsymbol{x}_i$ and $\boldsymbol{y}_i^* = \gamma_i^* \hat{D}_i^{-1} \boldsymbol{y}_i$ can be computed.[17] The pseudo data $\{(\boldsymbol{x}_i^*, \boldsymbol{y}_i^*)\}_{i=1}^N$ comprise a bootstrap sample $\mathcal{Z}_N^*$. The complete bootstrap procedure based on the kernel density estimation may now be described by the following steps:

[16] Note that since the procedure involves a sequence of nested tests, the critieria for rejecting the null hypothesis should be adjusted for successive values of m to avoid increasing probability of type I error. In each of our empirical examples which appear below, we were unable to reject H_0 for $m = 1$ in each case, and so this was not a critical issue in our examples.

[17] Note that $\boldsymbol{y}_i/\hat{D}_i^{out}$ gives the ostensibly efficient output level for firm i; *i.e.,* the point $(\boldsymbol{x}_i, \boldsymbol{y}_i/\hat{D}_i^{out})$ falls on the estimated boundary of $\tilde{\Psi}$. Multiplying by γ_i^* produces a random deviation away from the frontier, reflecting the underlying notion that the data were generated by the process characterized by $\hat{f}(v)$ in (4.5). The output distance function measures the feasible proportionate expansion in outputs for a given input vector, and hence the bootstrap values $\boldsymbol{x}_i^*$ are the same as the original data. If the input orientation were being used as with the Shephard (1970) input distance function, one would generate a pseudo datum $(\gamma_i^* \boldsymbol{x}_i/\hat{D}_i^{in}, \boldsymbol{y}_i)$. Note that psuedo random deviates must be generated for each observation in the sample.

[1] For each firm $i = 1, \ldots, N$ in $\mathcal{Z}_N$ drawn from Ψ, compute $\hat{D}_i^{out}$ from (3.9).

[2] Generate random deviates $\gamma_i^* \; \forall \; i = 1, \ldots, N$ using (4.12).

[3] Compute $\mathcal{Z}_N^* = \{(\boldsymbol{x}_i, \gamma_i^* \boldsymbol{y}_i / \hat{D}_i^{out})\}_{i=1}^N$.

[4] Compute bootstrap estimates $\hat{D}_i^{out^*} \; \forall \; i = 1, \ldots, N$, using (4.1).

[5] Repeat [2]–[4] B times to produce N sets $\{\hat{D}_i^{out^*}(b)\}_{b=1}^B$, $i = 1, \ldots, N$.

Once the sets $\{\hat{D}_i^{out^*}(b)\}_{b=1}^B$ have been obtained in step [5], $\widehat{bias}_i$ can be computed using (4.2). Then, the $\hat{\hat{D}}_i^{out^*}(b)$ computed from (4.4) and truncated at $\hat{D}_i^{out}$ allow construction of bias-corrected α-percent confidence intervals for D_i^{out}. Empirical examples are given in the next section.

Before proceeding to empirical examples from the DEA literature, note that consistency is a desirable, but perhaps weak property; one might reasonably ask what is the convergence rate of the bootstrap estimates proposed above. Bootstrap percentile confidence limits are root-N consistent (Efron and Tibshirani, 1993), which is well above the convergence rate of DEA and FDH frontier estimators (see Korostelev *et al.,* 1995a, b). Noise from resampling in the bootstrap can be reduced to arbitrarily low order by appropriately choosing the kernel bandwidth and using a large number of bootstrap replications B. Thus, noise from our bootstrap procedure will be of lower order than the noise from the original efficiency estimation.

5. EMPIRICAL EXAMPLES

For each of the examples presented in this section, we specified the kernel function as standard normal. In both examles, the number of bootstrap replications B was set equal to 1000. Detailed descriptions of the data used in each example may be found in the original articles cited below.

EXAMPLE 5.1: Charnes *et al.* (1981) report data for five inputs and three outputs in 70 educational organizations. Applying (3.9) yields the distance function estimates $\hat{D}_i^{out}$ shown in the second column of Table 1. These results indicate that 27 firms are ostensibly efficient as indicated by $\hat{D}_i^{out} = 1$, with the distance function estimates ranging as low as

0.7883 for firm #36. The distance function estimates have a standard deviation of 0.05335, and an interquartile range of 0.0801.

Using the bandwidth selection criteria in (4.14), we set the bandwidth $w = 0.05335$ and Bootstrapped the distance function estimates as described in section 4 to obtain the bias-corrected distance function estimates shown in the third column of Table 1. The fourth and fifth columns of Table 1 give the the bootstrap bias estimates and the standard deviation of the bootstrap estimates, respectively. Comparing the bias estimates with the standard deviations indicates that the bias correction is quite unlikely to increase the mean-square error of the distance function estimates; the bias estimates are typically 1–2 orders of magnitude larger than the standard deviations. The bias estimates range from 0.0254 to 0.0721, with a mean of 0.0463. Thus, on average, the conventional distance function estimates in column two overstate the level of efficiency in these data by more than 4.6 percent.

The confidence intervals in columns 6 and 7 of Table 1 indicate considerable uncertainty surrounding the distance function estimates. For example, firm #5 was initially indicated as ostensibly efficient by the conventional estimator $\hat{D}^{out}$. After applying the bias correction in (4.3), this firm is found to produce only about 93 percent of its potential output. Yet, the confidence interval in column 6 of Table 1 suggests that with 90 percent probability, firm #5 is producing between about 88 percent and 98 percent of its potential output—a range of 10 percent. At the 95 percent level in column 7, the range is slightly wider. For some firms, however, the confidence intervals are much narrower; for example, the bias-corrected distance function estimate for firm #36 indicates that this firm is producing only about 76 percent of its potential output, with the 95 percent confidence interval ranging from about 74 percent to about 78 percent.

Next, we used the bootstrap test for multimodality in the density of the distance function estimates. With number of modes $m = 1$, the P-value from step [4] was 0.3420; hence we could not reject the null hypothesis on one mode. The minimum bandwidth with one mode was 0.01611, smaller than the bandwidth used to produce the results in Table 1 by a factor of greater than 3. Using this new bandwidth, we again bootstrapped the distance function estimates from the Charnes *el al.* data to produce the results shown in Table 2. The estimated biases are generally smaller, with an average of 0.0318 in Table 2, and ranging from 0.0109 to 0.0619. Hence the bias corrected efficiency estimates are slightly higher than before, but the confidence intervals are roughly similar, with a great deal of overlap between those in Table 1 and the corresponding intervals in Table 2. Also as before, the individual bias estimates are generally 1–2 orders of magnitude larger than the corresponding standard errors, suggesting that the bias correction is unlikely to increase mean-square error. Thus the bootstrap procedure appears to be somewhat robust with respect to the choice of bandwidth. ■

EXAMPLE 5.2: This example uses data from Färe *et al.* (1986), who report data on 5 inputs and 5 outputs for 100 electric power utilities. The conventional distance function estimates from (3.9) shown in the second column of Table 3 indicate that only 18 firms in the sample are inefficient, while 82 are ostensibly efficient. The distance function estimates have a standard deviation of 0.053544, but the interquartile range is 0.0 due to the large number of values at unity. The smallest bandwidth yielding one mode in the kernel density estimate for the distance function estimates was 0.02429, with a P-value of 0.5140; hence the null hypothesis of one mode could not be rejected. Using this bandwidth to bootstrap as described in section 4 yielded the bias-corrected distance function estimates in the third column of Table 2, and the bias estimates in the fourth column. The bias estimates in column 4 range from 0.0044 to 0.0387, with a mean of 0.0244. Thus, the conventional estimates in column 2 overstate technical efficiency by only about 2.4 percent on average. The bias estimates exceed the standard deviation of the bootstrap estimates shown in column 5 by an order of magnitude or more, indicating that applying the bias correction in (4.3) is unlikely to increase mean square error. The standard deviations of the bootstrap estimates shown in Table 3 are in many instances smaller than those encountered in the previous example, and consequently the confidence intervals in columns 6 and 7 of Table 3 are tighter in these cases. ■

EXAMPLE 5.3: As a third example, the output distance function was estimated for the data used by Aly *et al.* (1990), who specify 3 inputs and 5 outputs for 322 commercial banks. Computing $\hat{D}_i^{out}$ for each bank in the sample yielded estimates ranging from 0.2765 to unity, with an average value of 0.80135 (the results for each bank are omitted to conserve space, but are available from the authors on request). Of the 322 observations, 89 (26.7 percent) are ostensibly efficient as indicated by $\hat{D}_i^{out} = 1$. The minimum bandwidth producting one mode in the kernel density estimate for the $\{\hat{D}_i^{out}\}_{i=1}^N$ was 0.03857, with a P-value of 0.6190; hence the null hypothesis of one mode could not be rejected. The distance function estimates have a standard deviation of 0.1810, and an interquartile range of 0.3410. Using the bandwidth selection criteria in (4.14) would lead to oversmoothing in this example, reducing statistical efficiency.

Setting $w = 0.03857$ and bootstrapping the Aly *et al.* data as described in section four yielded an average bias estimate of 0.1044, with individual estimates ranging from 0.0227 to 0.3574. As in the other examples, the bias estimates for individual observations typically exceeded the standard deviation of the bootstrap estimates of the distance function by 1-2 orders of magnitude, suggesting that the bias correction in (4.3) is unlikely to increase mean square error. The bias-corrected distance function estimates obtained from (4.3) range from 0.2429 to 0.9284, with a mean of 0.6970. The confidence intervals indicate a large amount of uncertainty surrounding the distance function values, with the 95 percent confidence intervals ranging from 0.0397 to 0.6533 in width, with an average width of 0.1909. Confidence intervals tend to be tighter for observations with low distance function estimates as opposed to those with higher distance function estimates. ■

All computations for examples 5.1 and 5.2 were performed using optimized fortran code running on a SUN Sparcstation 10 desktop workstation. Computations for example 5.3 were performed using the same code, but running on a SUN Sparcstation 20 desktop workstation.[18] LP problems for the distance function estimates were solved using a standard

[18] Code for example 5.3 was compiled under SunOS Release 4.1.3; thus, the resulting object code could not take advantage of the parallel processing capabilities of the Sparcstation 20.

revised simplex algorithm, without attempting to reduce the reference set (*e.g.,* by only including observations belonging to the FDH efficient subset). The Charnes *et al.* data with 70 observations used in example 5.1 required roughly 14 minutes elapsed time to bootstrap the distance function estimates; the Färe *et al.* data used in example 5.2 required 55 minutes, while the Aly *et al.* data in example 5.3 required approximately $3'46''$.

Using the bootstrap test of multimodality to set the bandwidth for the kernel estimation required somewhat longer computation times. For each of the examples 5.1–5.3, the elapsed times were $8'15''$, $11'7''$, and $21'25''$, respectively. Given that the cost of CPU cycles continues to decline, these times do not seem formidable. However, for researchers facing severe computing constraints, graphical methods suggested by Silverman (1978) can be used to avoid this part of the computational burden.

6. CONCLUSIONS

This paper provides a bootstrap methodology which yields statistical confidence intervals for nonparametric efficiency measures. Although the methodology has been illustrated using the Shepard output distance function, the methodology can be adapted for use in input-oriented models as well. Indeed, it is straightforward to adapt the methodology to any consistent nonparametric frontier estimator. The methodology also allows correction of an inherent bias which causes the output distance function and other DEA and FDH efficiency scores to understate the amount of inefficiency in individual firms.

The empirical examples in the previous section indicate that the magnitude of the bias in conventional distance function estimation can be substantial, causing conventional estimates to overstate efficiency by more than 35 percent for individual production units in some instances. In much of the DEA literature, researchers have treated DEA efficiency scores as devoid of statistical content. Yet, bootstrapped confidence intervals for the examples considered above suggest considerable uncertainty surrounding distance function estimates due to sampling variation.

REFERENCES

Aigner, D., C. A. K. Lovell, and P. Schmidt (1977), Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics* 6, 21–37.

Aly, H. Y., R. G. Grabowski, C. Pasurka, and N. Rangan (1990), Technical, scale, and allocative efficiencies in U. S. banking: an empirical investigation, *Review of Economics and Statistics* 72, 211–218.

Athreya, K. B. (1987), Bootstrap of the mean in the infinite variance case, *Annals of Statistics* 15, 724–731.

Banker, R. D. (1993), Maximum likelihood, consistency and data envelopment analysis: A statistical foundation, *Management Science* 39, 1265–1273.

Beran, R. (1982), Estimated sampling distributions: The bootstrap and competitors, *Annals of Statistics* 10, 212–225.

Beran, R. (1984), Bootstrap methods in statistics, *Jahresbericht Deutsche Mathematiker-Vereinigung* 86, 14–30.

Beran, R. (1988), Prepivoting test statistics: A bootstrap view of asymptotic refinements, *Journal of the American Statistical Association* 83, 687–697.

Beran, R., and G. Ducharme (1991), *Asymptotic Theory for Bootstrap Methods in Statistics,* Montreal: Centre de Reserches Mathematiques, University of Montreal.

Beran, R., and M. S. Srivastava (1985), Bootstrap tests and confidence regions for functions of a covariance matrix, *Annals of Statistics* 13, 95–115.

Beran, R., and M. S. Srivastava (1985), Correction to "Bootstrap tests and confidence regions for functions of a covariance matrix," *Annals of Statistics* 15, 470–471.

Bickel, P. J., and D. A. Freedman (1981), Some asymptotic theory for the bootstrap, *Annals of Statistics* 9, 1196–1217.

Bretagnolle, J. (1983), Lois limites du bootstrap des certaines fonctionnelles, *Annales de l'Institut Henri Poincare* (Section B) 19, 223–234.

Burgess, J. F. Jr., and P. W. Wilson (1993), "Technical Efficiency in VA Hospitals," in *The Measurement of Productive Efficiency: Techniques and Applications,* ed. by Harold O. Fried, C. A. Knox Lovell, and Shelton S. Schmidt, Oxford University Press, Inc., Oxford, pp. 335–351.

Burgess, J. F. Jr., and P. W. Wilson (1995), Determinants of technical efficiency in the US hospital industry," *Management Science,* forthcoming.

Burley, H. T. (1980), Productive efficiency in U. S. manufacturing: a linear programming approach, *Review of Economics and Statistics* 62, 619–622.

Byrnes, P., S. Grosskopf, and K. Hayes (1986), Efficiency and ownership: Further evidence, *Review of Economics and Statistics* 68, 337–341.

Charnes, A. and W. W. Cooper (1980), Management science relations for evaluation and management accountability, *Journal of Enterprise Management* 2, 143–162.

Charnes, A., W. W. Cooper, and E. Rhodes (1978), Measuring the efficiency of decision making units, *European Journal of Operational Research* 2, 429–444.

Charnes, A., W. W. Cooper, and E. Rhodes (1979), Measuring the efficiency of decision making units, *European Journal of Operational Research* 3, 339.

Charnes, A., W. W. Cooper, and E. Rhodes (1981), Evaluating program and managerial efficiency: an application of data envelopment analysis to program follow through, *Management Science* 27, 668–697.

Debreu, G. (1951), The coefficient of resource utilization, *Econometrica* 19, 273–292.

Deprins, D., L. Simar, and H. Tulkens (1984), "Measuring labor inefficiency in post offices," in *The Performance of Public Enterprises: Concepts and Measurements,* ed. by M. Marchand, P. Pestieau, and H. Tulkens, North-Holland, Amsterdam, 243–267.

Dusansky, R., and P. W. Wilson (1994), Technical efficiency in the decentralized care of the developmentally disabled, *Review of Economics and Statistics* 76, 340–345.

Dusansky, R. and P. W. Wilson, (1994b), On the relative efficiency of alternative modes of producing public sector output: The case of the developmentaly disabled, *European Journal of Operational Research* 80, 608–618.

Efron, B. (1979), Bootstrap methods: another look at the jackknife, *Annals of Statistics* 7, 1–16.

Efron, B. (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans,* DBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, Society for Industrial and Applied Mathematics, Philadelphia.

Efron, B., and R. J. Tibshirani (1993), *An Introduction to the Bootstrap,* New York: Chapman and Hall.

Färe, R. (1988), *Fundamentals of Production Theory,* Berlin: Springer-Verlag.

Färe, R., S. Grosskopf, and C. A. K. Lovell (1985), *The Measurement of Efficiency of Production.* Boston: Kluwer-Nijhoff Publishing.

Färe, R., S. Grosskopf, and E. C. Kokkelenberg (1989), Measuring plant capacity, utilization and technical change: A nonparametric approach, *International Economic Review* 30, 655–666.

Färe, R., S. Grosskopf, and C. Pasurka (1986), Effects on relative efficiency in electric power generation due to environmental controls, *Resources and Energy* 8, 167–184.

Färe, R., and C. A. K. Lovell (1978), Measuring the technical efficiency of production, *Journal of Economic Theory* 19, 150–162.

Farrell, M. J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society A* 120, 253–281.

Greene, W. H. (1980a), Maximum likelihood estimation of econometric frontier functions, *Journal of Econometrics* 13, 27–56.

Greene, W. H. (1980b), On the estimation of a felxible frontier production model, *Journal of Econometrics* 13, 101–115.

Grosskopf, S. (1986), The role of the reference technology in measuring productive efficiency, *The Economic Journal* 96, 499–513.

Hall, P. (1986), On the number of bootstrap simulations required to construct a confidence interval, *The Annals of Statistics* 14, 1453–1462.

Hall, P., T. J. DiCiccio, and J. P. Romano (1989), On smoothing and the bootstrap, *Annals of Statistics* 17, 692–704.

Härdle, W. (1990), *Applied Nonparametric Regression,* Cambridge: Cambridge University Press.

Jondrow, J., C. A. K. Lovell, I. S. Materov, and P. Schmidt (1982), On the estimation of technical inefficiency in the stochastic frontier production function model, *Journal of Econometrics* 19, 233–238.

Kneip, A. and L. Simar (1995), "A General Framework for Frontier Estimation with Panel Data," unpublished working paper, L'Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Korostelev, A. P., L. Simar, and A. B. Tsybakov (1995a), On estimation of monotone and convex boundaries, *Publications de l'Institut de Statistique de l'Université de Paris* 39, 3–18.

Korostelev, A. P., L. Simar, and A. B. Tsybakov (1995b), Efficient estimation of monotone boundaries, *Annals of Statistics,* forthcoming.

Lewis, P. A., A. S. Goodman, and J. M. Miller (1969), A pseudo-random number generator for the System/360, *IBM Systems Journal* 8, 136–146.

Lovell, C. A. K. (1993), "Production Frontiers and Productive Efficiency," in *The Measurement of Productive Efficiency: Techniques and Applications,* ed. by Hal Fried, C. A. Knox Lovell, and Shelton S. Schmidt, Oxford University Press, Inc., Oxford, pp. 3–67.

Meeusen, W. and J. van den Broeck (1977), Efficiency estimation from Cobb-Douglas production functions with composed error, *International Economic Review* 18, 435–444.

Nyman, J. A., and D. L. Bricker (1989), Profit incentives and technical efficiency in the production of nursing home care, *Review of Economics and Statistics* 71, 586–594.

Parzen, E. (1962), On estimation of a probability density function and mode, *Annals of Mathematical Statstics* 33, 1065.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling (1986), *Numerical Recipes,* Cambridge University Press, Cambridge.

Ripley, B. D., and J. P. Rasson (1977), Finding the edge of a Poisson forest, *Journal of Applied Probability* 14, 483–491.

Seaver, B. L., and Triantis, K. P. (1989), The implications of using messy data to estimate production-frontier-based technical efficiency measures, *Journal of Business and Economic Statistics,* 7, 49–59.

Simar, L. (1991), Estimating efficiencies from frontier models with panel data: a comparison of parametric, non-parametric and semi-parametric methods with bootstrapping, *Journal of Productivity Analysis* 3, 171–191.

Simar, L. and P. W. Wilson (1995), "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models," Discussion Paper #9503, L'Institut de Statistique, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Shephard, R. W. (1970), *Theory of Cost and Production Functions,* Princeton University Press, Princeton.

Silverman, B. W. (1978), Choosing the window width when estimating a density, *Biometrika* 65, 1–11.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis,* Chapman and Hall Ltd., London.

Stevenson, R. (1980), Likelihood functions ofr generalized stochastic frontier estimation, *Journal of Econometrics* 13, 58–66.

Swanepoel, J. W. H. (1986), A note on proving that the (modified) bootstrap works, *Communications in Statistics: Theory and Methods* 15, 3193–3203.

Tapia, R. A., and J. R. Thompson (1978), *Nonparametric Probability Density Estimation,* Baltimore: Johns Hopkins University Press.

Wheelock, D. C. and P. W. Wilson (1995), Explaining bank failures: Deposit insurance, regulation, and efficiency, *Review of Economics and Statistics,* forthcoming.

Wilson, P. W. (1993), Detecting outliers in deterministic nonparametric frontier models with multiple outputs, *Journal of Business and Economic Statistics* 11, 319–323.

Wilson, P. W. (1995), Detecting influential observations in data envelopment analysis, *Journal of Productivity Analysis* 6, 27–45.

Winsten, C. B. (1957), Discussion on Mr. Farrell's paper, *Journal of the Royal Statistical Society* Series A, 120, 282–284.

Zelterman, D. (1993), A semiparametric bootstrap technique for simulating exterme order statistics, *Journal of the American Statistical Association* 88, 477–485.

Zieschang, K. O. (1984), An extended Farrell technical efficiency measure, *Journal of Economic Theory* 33, 387–396.

TABLE 1

Bootstrap Results for Charnes *et al.* (1981) Data
(Bandwidth $w = 0.05335$ chosen by equation (4.14))

(1) i	(2) $\hat{D}_i^{out}$	(3) $\hat{\hat{D}}_i^{out}$	(4) $\widehat{bias}_i$	(5) std. dev.	(6) 90% conf. int.	(7) 95% conf. int.
1	0.9687	0.9238	0.0449	0.000511	0.8953—0.9580	0.8938—0.9624
2	0.9015	0.8642	0.0373	0.000167	0.8423—0.8850	0.8407—0.8887
3	0.9360	0.8969	0.0390	0.000277	0.8742—0.9240	0.8725—0.9279
4	0.9030	0.8751	0.0279	0.000074	0.8622—0.8902	0.8606—0.8924
5	1.0000	0.9315	0.0685	0.002171	0.8809—0.9843	0.8783—0.9908
6	0.9049	0.8675	0.0374	0.000186	0.8460—0.8914	0.8440—0.8955
7	0.8936	0.8646	0.0290	0.000063	0.8513—0.8768	0.8495—0.8790
8	0.9056	0.8652	0.0404	0.000403	0.8408—0.8932	0.8395—0.9008
9	0.8615	0.8307	0.0308	0.000104	0.8152—0.8479	0.8135—0.8505
10	0.9483	0.9064	0.0419	0.000425	0.8797—0.9370	0.8784—0.9413
11	1.0000	0.9405	0.0595	0.000980	0.8983—0.9844	0.8966—0.9922
12	1.0000	0.9397	0.0603	0.000960	0.8971—0.9876	0.8948—0.9920
13	0.8650	0.8396	0.0254	0.000044	0.8292—0.8516	0.8271—0.8535
14	0.9847	0.9430	0.0417	0.000283	0.9190—0.9702	0.9168—0.9759
15	1.0000	0.9288	0.0712	0.002594	0.8744—0.9816	0.8724—0.9895
16	0.9506	0.9116	0.0390	0.000362	0.8891—0.9387	0.8877—0.9440
17	1.0000	0.9298	0.0702	0.002341	0.8764—0.9791	0.8748—0.9904
18	1.0000	0.9444	0.0556	0.000700	0.9060—0.9864	0.9046—0.9907
19	0.9532	0.9155	0.0377	0.000235	0.8937—0.9415	0.8921—0.9460
20	1.0000	0.9342	0.0658	0.001714	0.8857—0.9807	0.8826—0.9892
21	1.0000	0.9354	0.0646	0.001451	0.8882—0.9880	0.8861—0.9934
22	1.0000	0.9543	0.0457	0.000297	0.9254—0.9805	0.9239—0.9859
23	0.9754	0.9423	0.0331	0.000155	0.9253—0.9636	0.9238—0.9669
24	1.0000	0.9319	0.0681	0.001817	0.8811—0.9830	0.8790—0.9907
25	0.9791	0.9494	0.0296	0.000052	0.9368—0.9612	0.9347—0.9630
26	0.9432	0.9121	0.0311	0.000082	0.8978—0.9275	0.8966—0.9310
27	1.0000	0.9380	0.0620	0.001042	0.8939—0.9877	0.8921—0.9928
28	0.9877	0.9483	0.0394	0.000229	0.9254—0.9720	0.9236—0.9760
29	0.8477	0.8127	0.0351	0.000207	0.7923—0.8351	0.7911—0.8394
30	0.8950	0.8631	0.0319	0.000107	0.8465—0.8802	0.8449—0.8840
31	0.8383	0.8107	0.0276	0.000063	0.7978—0.8238	0.7956—0.8256
32	1.0000	0.9300	0.0700	0.002434	0.8770—0.9813	0.8749—0.9911
33	0.9531	0.9186	0.0345	0.000160	0.9002—0.9401	0.8987—0.9457
34	0.8615	0.8316	0.0299	0.000110	0.8163—0.8499	0.8147—0.8533
35	1.0000	0.9457	0.0543	0.000733	0.9083—0.9868	0.9066—0.9922

TABLE 1 (continued)

(1) i	(2) $\hat{D}_i^{out}$	(3) $\hat{\hat{D}}_i^{out}$	(4) $\widehat{bias}_i$	(5) std. dev.	(6) 90% conf. int.	(7) 95% conf. int.
36	0.7883	0.7588	0.0295	0.000094	0.7431—0.7754	0.7418—0.7781
37	0.8391	0.8054	0.0337	0.000145	0.7865—0.8255	0.7847—0.8289
38	1.0000	0.9334	0.0666	0.001998	0.8840—0.9805	0.8820—0.9887
39	0.9371	0.8944	0.0427	0.000385	0.8675—0.9253	0.8656—0.9293
40	0.9498	0.9138	0.0360	0.000210	0.8942—0.9390	0.8924—0.9426
41	0.9534	0.9279	0.0255	0.000040	0.9179—0.9386	0.9163—0.9415
42	0.9476	0.9064	0.0412	0.000315	0.8811—0.9334	0.8793—0.9382
43	0.8648	0.8286	0.0362	0.000226	0.8086—0.8546	0.8061—0.8586
44	1.0000	0.9299	0.0701	0.002320	0.8776—0.9820	0.8755—0.9911
45	1.0000	0.9460	0.0540	0.000721	0.9104—0.9858	0.9081—0.9918
46	0.9143	0.8831	0.0313	0.000106	0.8674—0.9009	0.8660—0.9049
47	1.0000	0.9322	0.0678	0.001778	0.8818—0.9850	0.8800—0.9925
48	1.0000	0.9279	0.0721	0.002467	0.8737—0.9813	0.8714—0.9939
49	1.0000	0.9368	0.0632	0.001185	0.8914—0.9890	0.8895—0.9949
50	0.9583	0.9191	0.0391	0.000410	0.8954—0.9486	0.8939—0.9532
51	0.9199	0.8882	0.0317	0.000110	0.8720—0.9075	0.8701—0.9107
52	1.0000	0.9325	0.0675	0.002017	0.8819—0.9835	0.8804—0.9890
53	0.8722	0.8412	0.0310	0.000096	0.8245—0.8574	0.8230—0.8603
54	1.0000	0.9308	0.0692	0.002212	0.8783—0.9807	0.8769—0.9895
55	0.9994	0.9555	0.0439	0.000359	0.9292—0.9862	0.9277—0.9905
56	1.0000	0.9342	0.0658	0.001705	0.8857—0.9898	0.8841—0.9946
57	0.9281	0.8886	0.0395	0.000391	0.8650—0.9191	0.8633—0.9234
58	1.0000	0.9291	0.0709	0.002704	0.8748—0.9785	0.8724—0.9872
59	1.0000	0.9307	0.0693	0.002274	0.8782—0.9821	0.8762—0.9913
60	0.9809	0.9483	0.0325	0.000080	0.9342—0.9628	0.9318—0.9657
61	0.8815	0.8454	0.0361	0.000302	0.8238—0.8685	0.8223—0.8744
62	1.0000	0.9333	0.0667	0.002182	0.8837—0.9829	0.8817—0.9921
63	0.9632	0.9279	0.0353	0.000144	0.9087—0.9491	0.9072—0.9529
64	0.9314	0.9015	0.0299	0.000072	0.8880—0.9160	0.8856—0.9190
65	0.9737	0.9372	0.0365	0.000156	0.9173—0.9577	0.9154—0.9608
66	0.9368	0.9105	0.0263	0.000044	0.9005—0.9220	0.8987—0.9239
67	0.9476	0.9167	0.0309	0.000084	0.9016—0.9322	0.9002—0.9352
68	1.0000	0.9327	0.0673	0.001854	0.8820—0.9820	0.8802—0.9897
69	1.0000	0.9316	0.0684	0.002275	0.8804—0.9811	0.8779—0.9886
70	0.9647	0.9336	0.0311	0.000077	0.9199—0.9476	0.9182—0.9498

TABLE 2

Bootstrap Results for Charnes *et al.* (1981) Data
(Bandwidth $w = 0.01611$ chosen by bootstrapping)

(1) i	(2) $\hat{D}_i^{out}$	(3) $\hat{\hat{D}}_i^{out}$	(4) $\widehat{bias}_i$	(5) std. dev.	(6) 90% conf. int.	(7) 95% conf. int.
1	0.9687	0.9394	0.0293	0.000578	0.9137—0.9619	0.9130—0.9653
2	0.9015	0.8805	0.0209	0.000149	0.8632—0.8975	0.8622—0.8995
3	0.9360	0.9126	0.0234	0.000282	0.8928—0.9301	0.8921—0.9323
4	0.9030	0.8899	0.0132	0.000051	0.8799—0.8996	0.8791—0.9009
5	1.0000	0.9426	0.0574	0.003167	0.8893—0.9858	0.8879—0.9940
6	0.9049	0.8829	0.0220	0.000176	0.8644—0.8999	0.8637—0.9023
7	0.8936	0.8793	0.0143	0.000051	0.8688—0.8901	0.8680—0.8915
8	0.9056	0.8797	0.0259	0.000487	0.8572—0.8980	0.8566—0.9005
9	0.8615	0.8457	0.0158	0.000083	0.8328—0.8577	0.8323—0.8594
10	0.9483	0.9224	0.0259	0.000431	0.8999—0.9413	0.8990—0.9443
11	1.0000	0.9574	0.0426	0.001101	0.9184—0.9918	0.9177—0.9946
12	1.0000	0.9570	0.0430	0.001034	0.9180—0.9935	0.9170—0.9960
13	0.8650	0.8531	0.0120	0.000028	0.8449—0.8611	0.8437—0.8628
14	0.9847	0.9596	0.0251	0.000300	0.9388—0.9787	0.9379—0.9817
15	1.0000	0.9381	0.0619	0.003765	0.8802—0.9848	0.8789—0.9925
16	0.9506	0.9272	0.0234	0.000378	0.9078—0.9450	0.9068—0.9475
17	1.0000	0.9407	0.0593	0.003419	0.8850—0.9864	0.8844—0.9925
18	1.0000	0.9613	0.0387	0.000723	0.9267—0.9950	0.9258—0.9977
19	0.9532	0.9309	0.0223	0.000228	0.9123—0.9481	0.9112—0.9501
20	1.0000	0.9473	0.0527	0.002447	0.8985—0.9811	0.8973—0.9914
21	1.0000	0.9509	0.0491	0.001855	0.9052—0.9862	0.9046—0.9931
22	1.0000	0.9721	0.0279	0.000266	0.9483—0.9937	0.9472—0.9964
23	0.9754	0.9582	0.0172	0.000145	0.9443—0.9719	0.9434—0.9735
24	1.0000	0.9453	0.0547	0.002427	0.8945—0.9865	0.8935—0.9906
25	0.9791	0.9652	0.0139	0.000034	0.9549—0.9741	0.9540—0.9761
26	0.9432	0.9279	0.0153	0.000064	0.9164—0.9394	0.9154—0.9412
27	1.0000	0.9534	0.0466	0.001190	0.9113—0.9942	0.9101—0.9965
28	0.9877	0.9657	0.0220	0.000198	0.9475—0.9827	0.9466—0.9851
29	0.8477	0.8272	0.0205	0.000216	0.8101—0.8423	0.8094—0.8449
30	0.8950	0.8778	0.0172	0.000086	0.8644—0.8914	0.8634—0.8935
31	0.8383	0.8244	0.0139	0.000048	0.8137—0.8349	0.8131—0.8361
32	1.0000	0.9412	0.0588	0.003432	0.8860—0.9829	0.8852—0.9893
33	0.9531	0.9343	0.0188	0.000149	0.9189—0.9483	0.9181—0.9508
34	0.8615	0.8459	0.0156	0.000095	0.8336—0.8581	0.8329—0.8599
35	1.0000	0.9615	0.0385	0.000887	0.9266—0.9934	0.9256—0.9968

TABLE 2 (continued)

(1) i	(2) $\hat{D}_i^{out}$	(3) $\hat{\hat{D}}_i^{out}$	(4) $\widehat{bias}_i$	(5) std. dev.	(6) 90% conf. int.	(7) 95% conf. int.
36	0.7883	0.7720	0.0164	0.000076	0.7590—0.7840	0.7580—0.7855
37	0.8391	0.8195	0.0196	0.000127	0.8032—0.8351	0.8025—0.8367
38	1.0000	0.9451	0.0549	0.003179	0.8939—0.9808	0.8930—0.9894
39	0.9371	0.9107	0.0265	0.000398	0.8876—0.9315	0.8868—0.9340
40	0.9498	0.9299	0.0198	0.000179	0.9140—0.9456	0.9132—0.9480
41	0.9534	0.9425	0.0109	0.000021	0.9351—0.9496	0.9341—0.9512
42	0.9476	0.9227	0.0249	0.000311	0.9014—0.9428	0.9004—0.9445
43	0.8648	0.8436	0.0212	0.000233	0.8260—0.8596	0.8252—0.8620
44	1.0000	0.9397	0.0603	0.003495	0.8836—0.9858	0.8825—0.9927
45	1.0000	0.9626	0.0374	0.000900	0.9296—0.9938	0.9285—0.9967
46	0.9143	0.8983	0.0161	0.000090	0.8857—0.9110	0.8847—0.9126
47	1.0000	0.9458	0.0542	0.002292	0.8954—0.9839	0.8948—0.9919
48	1.0000	0.9389	0.0611	0.003578	0.8818—0.9851	0.8808—0.9919
49	1.0000	0.9528	0.0472	0.001408	0.9095—0.9957	0.9085—0.9973
50	0.9583	0.9343	0.0239	0.000435	0.9136—0.9509	0.9130—0.9549
51	0.9199	0.9028	0.0170	0.000103	0.8892—0.9149	0.8885—0.9169
52	1.0000	0.9450	0.0550	0.002693	0.8937—0.9815	0.8929—0.9896
53	0.8722	0.8560	0.0162	0.000074	0.8427—0.8682	0.8420—0.8694
54	1.0000	0.9413	0.0587	0.003015	0.8860—0.9811	0.8854—0.9904
55	0.9994	0.9724	0.0270	0.000413	0.9493—0.9935	0.9485—0.9961
56	1.0000	0.9466	0.0534	0.002227	0.8971—0.9893	0.8965—0.9954
57	0.9281	0.9047	0.0234	0.000379	0.8846—0.9206	0.8842—0.9246
58	1.0000	0.9394	0.0606	0.003668	0.8823—0.9816	0.8817—0.9884
59	1.0000	0.9413	0.0587	0.003265	0.8860—0.9782	0.8851—0.9895
60	0.9809	0.9643	0.0165	0.000062	0.9523—0.9761	0.9512—0.9776
61	0.8815	0.8598	0.0217	0.000343	0.8411—0.8750	0.8406—0.8774
62	1.0000	0.9452	0.0548	0.003135	0.8939—0.9859	0.8930—0.9920
63	0.9632	0.9443	0.0189	0.000117	0.9290—0.9586	0.9282—0.9608
64	0.9314	0.9172	0.0142	0.000047	0.9071—0.9273	0.9059—0.9292
65	0.9737	0.9533	0.0204	0.000141	0.9368—0.9692	0.9357—0.9713
66	0.9368	0.9252	0.0116	0.000023	0.9174—0.9325	0.9164—0.9341
67	0.9476	0.9327	0.0149	0.000058	0.9211—0.9433	0.9205—0.9450
68	1.0000	0.9452	0.0548	0.002487	0.8939—0.9858	0.8930—0.9929
69	1.0000	0.9422	0.0578	0.003380	0.8878—0.9859	0.8873—0.9925
70	0.9647	0.9496	0.0151	0.000058	0.9385—0.9600	0.9378—0.9621

TABLE 3

Bootstrap Results for Färe *et al.* (1986) Data
(Bandwidth $w = 0.02429$ chosen by bootstrapping)

(1) i	(2) $\hat{D}_i^{out}$	(3) $\hat{\hat{D}}_i^{out}$	(4) $\widehat{bias}_i$	(5) std. dev.	(6) 90% conf. int.	(7) 95% conf. int.
1	1.0000	0.9700	0.0300	0.002799	0.9424—0.9856	0.9418—0.9908
2	1.0000	0.9651	0.0349	0.004611	0.9328—0.9773	0.9319—0.9839
3	1.0000	0.9846	0.0154	0.000073	0.9720—0.9961	0.9713—0.9977
4	1.0000	0.9748	0.0252	0.000963	0.9520—0.9908	0.9514—0.9927
5	1.0000	0.9674	0.0326	0.004089	0.9380—0.9798	0.9369—0.9857
6	1.0000	0.9905	0.0095	0.000032	0.9831—0.9975	0.9826—0.9986
7	1.0000	0.9810	0.0190	0.000197	0.9646—0.9957	0.9636—0.9973
8	1.0000	0.9790	0.0210	0.000251	0.9602—0.9955	0.9595—0.9971
9	1.0000	0.9689	0.0311	0.003377	0.9401—0.9838	0.9394—0.9901
10	1.0000	0.9648	0.0352	0.004944	0.9320—0.9756	0.9313—0.9827
11	1.0000	0.9770	0.0230	0.000543	0.9561—0.9929	0.9555—0.9961
12	1.0000	0.9848	0.0152	0.000083	0.9720—0.9966	0.9715—0.9984
13	0.9579	0.9432	0.0147	0.000080	0.9310—0.9542	0.9303—0.9559
14	1.0000	0.9685	0.0315	0.003673	0.9396—0.9833	0.9389—0.9888
15	1.0000	0.9764	0.0236	0.000578	0.9550—0.9928	0.9545—0.9960
16	1.0000	0.9708	0.0292	0.001775	0.9440—0.9867	0.9431—0.9904
17	1.0000	0.9672	0.0328	0.003615	0.9374—0.9803	0.9364—0.9872
18	1.0000	0.9724	0.0276	0.001833	0.9472—0.9849	0.9466—0.9900
19	1.0000	0.9770	0.0230	0.000437	0.9563—0.9930	0.9558—0.9964
20	1.0000	0.9691	0.0309	0.003526	0.9410—0.9842	0.9403—0.9894
21	1.0000	0.9691	0.0309	0.003364	0.9410—0.9838	0.9400—0.9884
22	1.0000	0.9800	0.0200	0.000177	0.9625—0.9969	0.9618—0.9985
23	0.9994	0.9868	0.0127	0.000058	0.9768—0.9959	0.9759—0.9971
24	1.0000	0.9781	0.0219	0.000537	0.9589—0.9948	0.9580—0.9971
25	0.9931	0.9836	0.0095	0.000019	0.9767—0.9897	0.9761—0.9905
26	1.0000	0.9863	0.0137	0.000116	0.9749—0.9964	0.9745—0.9975
27	1.0000	0.9825	0.0175	0.000138	0.9675—0.9955	0.9668—0.9982
28	1.0000	0.9653	0.0347	0.004466	0.9334—0.9777	0.9326—0.9838
29	1.0000	0.9687	0.0313	0.003719	0.9399—0.9836	0.9393—0.9894
30	1.0000	0.9636	0.0364	0.004854	0.9297—0.9736	0.9289—0.9821
31	1.0000	0.9652	0.0348	0.004824	0.9332—0.9781	0.9325—0.9840
32	0.9797	0.9653	0.0144	0.000079	0.9534—0.9754	0.9526—0.9775
33	1.0000	0.9656	0.0344	0.004616	0.9336—0.9778	0.9328—0.9817

TABLE 3 (continued)

(1) i	(2) $\hat{D}_i^{out}$	(3) $\hat{\hat{D}}_i^{out}$	(4) $\widehat{bias}_i$	(5) std. dev.	(6) 90% conf. int.	(7) 95% conf. int.
34	1.0000	0.9723	0.0277	0.001284	0.9471—0.9874	0.9464—0.9911
35	1.0000	0.9668	0.0332	0.003778	0.9363—0.9809	0.9355—0.9856
36	1.0000	0.9683	0.0317	0.003693	0.9390—0.9803	0.9384—0.9869
37	1.0000	0.9656	0.0344	0.004337	0.9341—0.9789	0.9334—0.9843
38	1.0000	0.9761	0.0239	0.000586	0.9546—0.9912	0.9539—0.9948
39	1.0000	0.9669	0.0331	0.003740	0.9365—0.9815	0.9357—0.9883
40	1.0000	0.9613	0.0387	0.005611	0.9251—0.9722	0.9245—0.9769
41	1.0000	0.9807	0.0193	0.000160	0.9639—0.9961	0.9632—0.9974
42	1.0000	0.9673	0.0327	0.003908	0.9372—0.9816	0.9363—0.9872
43	1.0000	0.9693	0.0307	0.003306	0.9408—0.9834	0.9402—0.9886
44	0.7304	0.7195	0.0109	0.000037	0.7103—0.7280	0.7097—0.7290
45	1.0000	0.9755	0.0245	0.000629	0.9537—0.9925	0.9532—0.9955
46	0.9985	0.9910	0.0075	0.000015	0.9860—0.9963	0.9853—0.9970
47	1.0000	0.9680	0.0320	0.003068	0.9385—0.9836	0.9378—0.9885
48	0.9987	0.9862	0.0124	0.000122	0.9764—0.9950	0.9757—0.9963
49	1.0000	0.9660	0.0340	0.004546	0.9344—0.9782	0.9338—0.9837
50	0.9615	0.9474	0.0142	0.000073	0.9357—0.9574	0.9349—0.9586
51	1.0000	0.9685	0.0315	0.003461	0.9396—0.9806	0.9390—0.9881
52	1.0000	0.9888	0.0112	0.000038	0.9802—0.9972	0.9794—0.9984
53	1.0000	0.9670	0.0330	0.004005	0.9369—0.9789	0.9361—0.9862
54	0.9987	0.9893	0.0094	0.000026	0.9821—0.9963	0.9816—0.9973
55	1.0000	0.9825	0.0175	0.000123	0.9676—0.9962	0.9668—0.9980
56	0.7988	0.7891	0.0097	0.000026	0.7814—0.7969	0.7809—0.7976
57	1.0000	0.9786	0.0214	0.000247	0.9596—0.9960	0.9590—0.9978
58	1.0000	0.9770	0.0230	0.000492	0.9566—0.9936	0.9559—0.9961
59	1.0000	0.9691	0.0309	0.003639	0.9407—0.9841	0.9399—0.9915
60	1.0000	0.9675	0.0325	0.004187	0.9379—0.9796	0.9373—0.9854
61	0.9985	0.9941	0.0044	0.000003	0.9918—0.9968	0.9913—0.9972
62	0.8899	0.8783	0.0115	0.000034	0.8692—0.8867	0.8685—0.8885
63	1.0000	0.9771	0.0229	0.000413	0.9566—0.9923	0.9560—0.9956
64	0.9996	0.9947	0.0049	0.000006	0.9916—0.9979	0.9913—0.9986
65	0.9973	0.9824	0.0149	0.000264	0.9698—0.9934	0.9691—0.9950
66	1.0000	0.9654	0.0346	0.004548	0.9332—0.9777	0.9326—0.9854

TABLE 3 (continued)

(1) i	(2) $\hat{D}_i^{out}$	(3) $\hat{\hat{D}}_i^{out}$	(4) $\widehat{bias}_i$	(5) std. dev.	(6) 90% conf. int.	(7) 95% conf. int.
67	1.0000	0.9657	0.0343	0.003947	0.9337—0.9771	0.9330—0.9832
68	1.0000	0.9676	0.0324	0.003841	0.9377—0.9805	0.9371—0.9849
69	1.0000	0.9877	0.0123	0.000145	0.9777—0.9965	0.9771—0.9980
70	1.0000	0.9739	0.0261	0.000991	0.9502—0.9891	0.9497—0.9941
71	1.0000	0.9818	0.0182	0.000121	0.9664—0.9955	0.9654—0.9973
72	1.0000	0.9739	0.0261	0.001006	0.9505—0.9890	0.9496—0.9933
73	1.0000	0.9666	0.0334	0.004044	0.9356—0.9795	0.9349—0.9869
74	0.9977	0.9875	0.0102	0.000038	0.9794—0.9950	0.9790—0.9959
75	1.0000	0.9663	0.0337	0.004716	0.9347—0.9775	0.9341—0.9836
76	1.0000	0.9833	0.0167	0.000099	0.9690—0.9955	0.9682—0.9969
77	1.0000	0.9675	0.0325	0.004063	0.9382—0.9817	0.9372—0.9880
78	1.0000	0.9676	0.0324	0.003391	0.9380—0.9811	0.9374—0.9856
79	1.0000	0.9678	0.0322	0.003487	0.9382—0.9821	0.9374—0.9878
80	1.0000	0.9823	0.0177	0.000334	0.9671—0.9956	0.9667—0.9974
81	1.0000	0.9688	0.0312	0.003453	0.9403—0.9835	0.9396—0.9906
82	1.0000	0.9693	0.0307	0.003513	0.9412—0.9842	0.9407—0.9907
83	1.0000	0.9754	0.0246	0.000608	0.9533—0.9905	0.9527—0.9954
84	1.0000	0.9670	0.0330	0.004040	0.9365—0.9803	0.9359—0.9864
85	0.7430	0.7328	0.0103	0.000051	0.7244—0.7404	0.7239—0.7413
86	1.0000	0.9699	0.0301	0.002065	0.9423—0.9852	0.9416—0.9904
87	1.0000	0.9684	0.0316	0.003101	0.9393—0.9821	0.9387—0.9875
88	1.0000	0.9681	0.0319	0.003606	0.9390—0.9799	0.9381—0.9853
89	1.0000	0.9751	0.0249	0.000906	0.9525—0.9929	0.9517—0.9959
90	1.0000	0.9671	0.0329	0.003635	0.9367—0.9823	0.9361—0.9874
91	1.0000	0.9819	0.0181	0.000148	0.9661—0.9959	0.9653—0.9978
92	0.9902	0.9829	0.0073	0.000023	0.9775—0.9883	0.9771—0.9891
93	1.0000	0.9679	0.0321	0.003931	0.9380—0.9814	0.9375—0.9860
94	1.0000	0.9672	0.0328	0.003946	0.9367—0.9820	0.9360—0.9867
95	1.0000	0.9678	0.0322	0.003626	0.9381—0.9810	0.9375—0.9874
96	1.0000	0.9785	0.0215	0.000213	0.9595—0.9965	0.9586—0.9981
97	1.0000	0.9768	0.0232	0.000593	0.9562—0.9923	0.9555—0.9954
98	1.0000	0.9662	0.0338	0.004375	0.9349—0.9782	0.9342—0.9849
99	0.6754	0.6710	0.0044	0.000003	0.6680—0.6739	0.6677—0.6744
100	1.0000	0.9797	0.0203	0.000186	0.9619—0.9958	0.9612—0.9975

FIGURE 1

The Output Distance Function