

Syracuse University

## SURFACE at Syracuse University

---

Center for Policy Research

Maxwell School of Citizenship and Public  
Affairs

---

2010

### Alternative Technical Efficiency Measures: Skew, Bias, and Scale

William Clinton Horrace

*Syracuse University. Center for Policy Research, whorrace@maxwell.syr.edu*

Qu Feng

*Nanyang Technological University, Division of Economics.*

Follow this and additional works at: <https://surface.syr.edu/cpr>



Part of the [Mathematics Commons](#)

---

#### Recommended Citation

Horrace, William Clinton and Feng, Qu, "Alternative Technical Efficiency Measures: Skew, Bias, and Scale" (2010). *Center for Policy Research*. 45.

<https://surface.syr.edu/cpr/45>

This Working Paper is brought to you for free and open access by the Maxwell School of Citizenship and Public Affairs at SURFACE at Syracuse University. It has been accepted for inclusion in Center for Policy Research by an authorized administrator of SURFACE at Syracuse University. For more information, please contact [surface@syr.edu](mailto:surface@syr.edu).

**Center for Policy Research  
Working Paper No. 121**

**ALTERNATIVE TECHNICAL EFFICIENCY MEASURES:  
SKEW, BIAS, AND SCALE**

**QU FENG AND WILLIAM C. HORRACE**

**Center for Policy Research  
Maxwell School of Citizenship and Public Affairs  
Syracuse University  
426 Eggers Hall  
Syracuse, New York 13244-1020  
(315) 443-3114 | Fax (315) 443-1081  
e-mail: [ctrpol@syr.edu](mailto:ctrpol@syr.edu)**

**March 2010**

**\$5.00**

Up-to-date information about CPR's research projects and other activities is available from our World Wide Web site at [www-cpr.maxwell.syr.edu](http://www-cpr.maxwell.syr.edu). All recent working papers and Policy Briefs can be read and/or printed from there as well.

# CENTER FOR POLICY RESEARCH – Spring 2010

**Christine L. Himes, Director**  
**Maxwell Professor of Sociology**

---

## Associate Directors

Margaret Austin  
Associate Director  
Budget and Administration

Douglas Wolf  
Gerald B. Cramer Professor of Aging Studies  
Associate Director, Aging Studies Program

John Yinger  
Professor of Economics and Public Administration  
Associate Director, Metropolitan Studies Program

## SENIOR RESEARCH ASSOCIATES

Badi Baltagi..... Economics	Andrew London..... Sociology
Robert Bifulco..... Public Administration	Len Lopoo..... Public Administration
Leonard Burman ..Public Administration/Economics	Amy Lutz..... Sociology
Kalena Cortes..... Education	Jerry Miner..... Economics
Thomas Dennison..... Public Administration	Jan Ondrich..... Economics
William Duncombe..... Public Administration	John Palmer..... Public Administration
Gary Engelhardt..... Economics	David Popp..... Public Administration
Deborah Freund.. Public Administration/Economics	Christopher Rohlfs..... Economics
Madonna Harrington Meyer..... Sociology	Stuart Rosenthal..... Economics
William C. Horrace..... Economics	Ross Rubenstein..... Public Administration
Duke Kao..... Economics	Perry Singleton..... Economics
Eric Kingson..... Social Work	Margaret Usdansky..... Sociology
Sharon Kioko.....Public Administration	Michael Wasylenko..... Economics
Thomas Kniesner..... Economics	Jeffrey Weinstein..... Economics
Jeffrey Kubik..... Economics	Janet Wilmoth..... Sociology

## GRADUATE ASSOCIATES

Charles Alamo..... Public Administration	Becky Lafrancois..... Economics
Matthew Baer..... Public Administration	Hee Seung Lee..... Public Administration
Maria Brown..... Social Science	Jing Li..... Economics
Christian Buerger..... Public Administration	Wael Moussa..... Economics
Qianqian Cao..... Economics	Phuong Nguyen..... Public Administration
Il Hwan Chung..... Public Administration	Wendy Parker..... Sociology
Kevin Cook..... Public Administration	Leslie Powell..... Public Administration
Alissa Dubnicki..... Economics	Kerri Raissian..... Public Administration
Katie Francis..... Public Administration	Amanda Ross..... Economics
Andrew Friedson..... Economics	Ryan Sullivan..... Economics
Virgilio Galdo..... Economics	Tre Wentling..... Sociology
Pallab Ghosh..... Economics	Coady Wing..... Public Administration
Clorise Harvey..... Public Administration	Ryan Yeung..... Public Administration
Douglas Honma..... Public Administration	Can Zhao..... Social Science

## STAFF

Kelly Bogart.....Administrative Secretary	Candi Patterson.....Computer Consultant
Martha Bonney.....Publications/Events Coordinator	Roseann Presutti.....Administrative Secretary
Karen Cimilluca.....Office Coordinator	Mary Santy.....Administrative Secretary
Kitty Nasto.....Administrative Secretary	

## Abstract

In the fixed-effects stochastic frontier model an efficiency measure relative to the best firm in the sample is universally employed. This paper considers a new measure relative to the worst firm in the sample. We find that estimates of this measure have smaller bias than those of the traditional measure when the sample consists of many firms near the efficient frontier. Moreover, a two-sided measure relative to both the best and the worst firms is proposed. Simulations suggest that the new measures may be preferred depending on the skewness of the inefficiency distribution and the scale of efficiency differences.

Email: qfeng@ntu.edu.sg, Tel: +65 6592 1543. Division of Economics, School of Humanities and Social Sciences, HSS-04-48, 14 Nanyang Drive, Singapore 637332

†Corresponding Author. Email: whorrace@maxwell.syr.edu, Tel: 315-443-9061. Fax: 315-443-1081. Center for Policy Research, 426 Eggers Hall, Syracuse, NY 13244-1020

**Key Words:** stochastic frontier model, relative efficiency measure, two-sided measure, bias, bootstrap confidence intervals

**JEL Classification:** C15, C23, D24

# 1 Introduction

There are several ways to estimate time-invariant technical efficiency in stochastic frontier models for panel data. Compared to maximum likelihood or generalized least-squares estimation (Battese and Coelli, 1988), fixed-effects estimation (Schmidt and Sickles, 1984) has the advantage of not requiring distributional assumptions on the error components. Without these distributional assumptions, efficiency levels cannot be identified directly. Hence, a measure relative to the *best* firm in the sample is universally employed (Schmidt and Sickles, 1984). In this case only the efficiency distance to the best firms matters. The *worst* firm in the sample is ignored. For example, suppose there are 3 firms with efficiency levels 0.30, 0.90 and 0.99 respectively. It appears that firm 2 is quite efficient. If the efficiency level of the worst firm improves to 0.89 due to technological change, the distance between the firm 2 and the best firm is unchanged. However, firm 2 is now almost as inefficient as the worst firm. This example shows that using the worst firm as a reference point provides a different perspective on technical efficiency. Actually, in competitive settings the worst firm is of particular importance because the marginal cost of this firm may determine price.<sup>1</sup> This paper considers an alternative efficiency measure relative to the *worst* firm in the sample and compares this measure to the traditional relative efficiency measure on a variety of metrics.

More generally, it may be interesting to use both the best firm and the worst firm as reference points. Therefore, a two-sided measure relative to both the best firm and the worst firm in the sample is also proposed. Different from efficiency measures relative to the best or to the worst firms alone, the two-sided measure linearly scales the efficiency level onto the unit interval with efficiency scores of 0 for the worst firm and 1 for the best firm. Consequently, the distance of the efficiency level between *any* firms becomes informative.

This paper discusses fixed-effects estimates of the measure relative to the worst firm and the

---

<sup>1</sup>We would like to thank a referee for pointing this out to us.

two-sided measure (relative to the best and the worst). We focus on estimation bias and inference. The level of the bias of relative efficiency estimates is related to the skewness of the underlying distribution of technical (in)efficiency. Since the “max” operator favors positive noise, the traditional estimate has larger bias when there are more efficient firms in the population. Qian and Sickles (2008) describe this scenario as “mostly stars, few dogs.” When there are "mostly stars" our estimate (relative to the worst firm) is less biased than the traditional estimate. However, not surprisingly, the bias results are reversed when there are "mostly dogs." When the distribution of (in)efficiency is symmetric, the bias results of the two estimators are identical. These results are borne out in our simulations based on three parameterizations of the Beta distribution. The two-sided estimate balances these sources of bias (in a sense) as we shall see in the sequel.

Inference on estimated technical efficiency is often important, and it proceeds with construction of confidence intervals. When distributional assumptions are made on the two error components (noise and inefficiency), the theory for confidence interval construction is straightforward (Horrace and Schmidt, 1996). The intervals are valid in both finite samples and asymptotically. In the case of fixed-effects estimation, confidence intervals for technical efficiency may be based on asymptotic normality, when the sample size is large (Horrace and Schmidt, 2000). However, when the time dimension of the data is small, the preferred method to construct confidence intervals without distributional assumptions is to perform the bootstrap. Kim, Kim and Schmidt (2007) provide a detailed and intuitive survey on constructing varieties of bootstrap confidence intervals. They argue that the “max” operator of the traditional estimate (relative to the best firm) produces bias, leading to low coverage rates when constructing simple bootstrap confidence intervals. Our proposed estimates suffer from the same source of bias. However, the coverage rates of bootstrap confidence intervals are a function of the magnitude of the bias which is related to the skewness of the underlying distribution of (in)efficiency and to the estimate employed (i.e., relative to best, relative to worst, or two-sided).

Ultimately, the fixed-effects estimates provide information on the skewness of the underlying distribution of (in)efficiency. Given this skewness information there maybe an empirical trade-off between bias and the efficiency measure employed. For example, if the data suggest there are "mostly stars, few dogs", then the traditional estimate has large bias and our estimate (relative to the worst) has small bias. However, a measure relative to the best firm may be of interest. In this case, the empiricist must decide which is more important: bias or the empirical relevance of the measure. (This trade-off also has implications for bootstrap inference.) Of course, if a measure relative to the worst firm is needed, then there is no trade-off in this case. This trade-off underscores the fact that the proposed estimates are alternatives to the traditional estimates and that all three estimates are measuring different quantities (although they are all normalizations to the unit interval, as we shall see). There is no sense in which the estimates are substitutes; they are complements that simply add to the empiricist's toolbox.

The paper is organized as follows. The next section discusses the efficiency measure relative to the worst firm, our proposed estimate of this measure, and performs simulations to compare the bias of the estimate to that of the tradition estimate under different (in)efficiency distributions. Section 3, introduces a two-sided efficiency measure that pegs relative efficiency not only to the most efficient firm in the sample but also to the least efficient firm. A simulation study of its bias is also conducted. In section 4 bootstrap confidence intervals are discussed for the different measures, and a simulation study of coverage rates and interval widths is provided in the spirit of Kim, Kim and Schmidt (2007). Our contribution is to demonstrate how these measures and their estimates perform in finite samples under different skewnesses of the distribution of technical inefficiency. Section 5 applies the estimators to a panel of Indonesian rice farms, and the salient features of all three measures are discussed and compared. The last section summarizes and concludes.

## 2 Relative Efficiency Measures

The stochastic frontier model for panel data is:

$$y_{it} = \alpha + x'_{it}\beta - u_i + v_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T. \quad (1)$$

The error term contains two parts: time-invariant  $u_i \geq 0$ , a measure of technical inefficiency; and  $v_{it} \sim iid(0, \sigma_v^2)$ . A large value of  $u_i$  implies that the firm  $i$  is inefficient. Usually, technical efficiency is defined as  $r_i = \exp(-u_i)$  under a log-linear specification of the Cobb-Douglas production function.<sup>2</sup> Letting  $\alpha_i = \alpha - u_i$ , slope parameter  $\beta$  can be estimated consistently using fixed-effects estimation. Call this estimate  $\hat{\beta}$ . The usual estimate of  $\alpha_i$  is  $\hat{\alpha}_i = \bar{y}_i - \bar{x}'_i \hat{\beta}$ , where  $\bar{y}_i$  and  $\bar{x}_i$  are within-group averages. However,  $u_i$  is unidentified without additional assumptions. The literature suggests an efficiency measure relative to the *best* firm  $u_i^* = \max_j \alpha_j - \alpha_i = u_i - \min_j u_j$  and its estimate  $\hat{u}_i^* = \max_j \hat{\alpha}_j - \hat{\alpha}_i$ , where the  $\hat{\alpha}_i$  are fixed-effects estimates of  $\alpha_i$  (Schmidt and Sickles, 1984). Correspondingly, when output is in logarithms, relative technical efficiency is defined as  $r_i^* \equiv \exp(-u_i^*)$  with its fixed-effects estimate  $\hat{r}_i^* \equiv \exp(-\hat{u}_i^*)$ . We call  $u_i^*$  the **max-measure** or the **traditional measure**.

In the stochastic frontier model (1), the efficient frontier is defined as the best firm using  $\min_j u_j$  or  $\max_j \alpha_j$ . Hence,  $u_i^*$  measures technical inefficiency as the deviation from this frontier. Similarly, we can define the inefficient frontier as the worst firm using  $\max_j u_j$  or  $\min_j \alpha_j$ , and measure technical efficiency as

$$u_i^{**} = \alpha_i - \min_j \alpha_j = \max_j u_j - u_i. \quad (2)$$

This is simply the deviation from the inefficient frontier. We call  $u_i^{**}$  the **min-measure**. Its corresponding technical efficiency score is  $r_i^{**} = 1 - \exp(-u_i^{**})$ . Using  $\exp(-x) \approx 1 - x$  for small  $x$ ,  $r_i^{**}$  is an approximation of  $u_i^{**}$  on the unit interval. To fix ideas, we image that  $u_i$  has some

---

<sup>2</sup>If  $Y = e^{-u} e^v f(x, \beta)$ , then  $e^{-u}$  is technical efficiency for  $y = \ln Y$  and  $f(x, \beta) = \prod_{j=1}^k x_j^{\beta_j}$ , say. Even if this is not the production function in mind, empiricists often use the measure  $r_i$  to normalize  $u_i$  to the unit interval.



upper support bound  $\bar{u}$ , so that realizations of  $u_i$  cannot be too large. Then  $\min_j \alpha_j$  approximates  $\alpha - \bar{u}$ , and  $u_i^{**}$  approximates  $\bar{u} - u_i \geq 0$ , the deviation from the inefficient frontier. The concept of a upper bound for inefficiency was recently considered in Qian and Sickles (2008). Indeed, "using this bound as the inefficient frontier, we may define inverted efficiency scores in the same spirit of Inverted DEA described in Entani, Maeda, and Tanaka (2002)."<sup>3</sup> The corresponding fixed-effects estimates of the min-measure are:

$$\begin{aligned}\hat{u}_i^{**} &= \hat{\alpha}_i - \min_j \hat{\alpha}_j, \\ \hat{r}_i^{**} &= 1 - \exp(-\hat{u}_i^{**}).\end{aligned}\tag{3}$$

Using the same arguments as Schmidt and Sickles (1984),  $\hat{u}_i^{**}$  is consistent for  $u_i^{**} = \bar{u} - u_i$ , as  $T \rightarrow \infty$  and  $N \rightarrow \infty$ . When the production function is Cobb-Douglas and output is in logarithms, the  $r_i^*$  has a natural interpretation: it is the true percentage of the output of firm  $i$  relative to the efficient firm for a fixed set of input, so  $r_i^*$  is the way we would naturally measure efficiency for a Cobb-Douglas production function. In this case the proposed measure,  $r_i^{**}$ , does not have this natural interpretation, however (as already mentioned) performance relative to the inefficient firm may be relevant, because the marginal cost of the inefficient firm may equal price in competitive markets (markets where  $N$  is large and  $\bar{u}$  is small). When output is *not* in logarithms or there is no particular production function in mind,  $r_i^*$ 's interpretation is less clear, and it may be interpreted as a normalization to the unit interval of the measure  $u_i^*$ , as it quantifies inefficiency relative to the most efficient firm. Interpreted this way, the nonlinear exponential normalization of  $r_i^*$  may distort the scale of  $u_i^*$ . The proposed measure  $r_i^{**}$  has a similar interpretation but relative to the least efficient firm, and it too may distort the scale of  $u_i^{**}$ . Either way, the alternative measure,  $r_i^{**}$ , may prove useful to empiricists, particularly if bias and confidence interval coverage are important (as we shall see).

---

<sup>3</sup>Qian and Sickles (2008). Indeed, Qian and Sickles consider cross-sectional ( $T = 1$ ) and random-effects estimation of  $\bar{u}$ . Hence, the current paper and the Qian and Sickles papers are complements.

Ultimately we subject these estimates to finite sample simulations and compare their performance under a variety of assumptions on the distribution of inefficiency. (In all cases the distribution has bounded support from above and below, so the min-measure has a population interpretation.) However, it is useful to consider the theoretical biases. In particular, the biases of  $\hat{u}_i^*$  and  $\hat{u}_i^{**}$  are directly comparable, even though they estimate different measures.<sup>4</sup> The bias of the max-measure estimate is:

$$\begin{aligned} b_{\max} &= E(\hat{u}_i^*) - u_i^* = E \left[ \max_j \hat{\alpha}_j - \hat{\alpha}_i - (\max_j \alpha_j - \alpha_i) \right] \\ &= E(\max_j \hat{\alpha}_j - \max_j \alpha_j) - E(\hat{\alpha}_i - \alpha_i) \\ &= E(\max_j \hat{\alpha}_j) - \max_j \alpha_j, \end{aligned}$$

since  $E(\hat{\alpha}_j) = \alpha_j$  for each  $j$ . Notice that the bias is not firm-specific. The bias will be largest when there is much uncertainty over the identity of the best firm ( $\max_j \alpha_j$ ) in the population. Per Horrace and Schmidt (1996), this occurs when  $T$  is small or when the variability of the  $u_i$  is large. Uncertainty over the best firm is also worse when there are many firms in the population ( $\alpha_i$ ) close to being best ( $\max_j \alpha_j$ ). This is likely to occur when the distribution of  $u_i$  is skewed to the right: "mostly stars, few dogs". In this paper, we use the Beta distribution  $B(a, b)$  to model three cases of the distribution of  $u_i$ :  $B(2, 8)$  "mostly stars, few dogs",  $B(8, 2)$  "mostly dogs, few stars", and the symmetric distribution  $B(2, 2)$  "few stars, few dogs". See Figure 1. The discussion above suggest that *ceteris paribus*, the bias,  $b_{\max}$ , is small in the case of "mostly dogs, few stars". (It is interesting to note that most Monte Carlos studies of the stochastic frontier model involve the truncated normal distribution which can only be skewed in the opposite direction: "mostly stars, few dogs. See, for example, KKS, 2007.)

---

<sup>4</sup>It is not entirely clear how to compare the theoretical bias of the  $\hat{r}_i^*$  and  $\hat{r}_i^{**}$ . Also, per Kim, Kim and Schmidt (2007) coverage rates for bootstrap confidence intervals on  $\hat{u}_i^*$  (and  $\hat{u}_i^{**}$ ) converted to intervals on  $\hat{r}_i^*$  (and  $\hat{r}_i^{**}$ ) are better than coverage rates on  $\hat{r}_i^*$  (and  $\hat{r}_i^{**}$ ) directly, so understanding the bias of the  $\hat{u}_i^*$  (and  $\hat{u}_i^{**}$ ) helps us better understand the coverage rates of the preferred intervals.

Similarly, the bias of the min-measure estimate is  $b_{\min} = E(\min_j \hat{\alpha}_j) - \min_j \alpha_j$ . Here, bias is large in magnitude when there is uncertainty over the worst firm in the population, which will be worse when there are many firms in the population ( $\alpha_i$ ) close to being worst ( $\min_j \alpha_j$ ). This corresponds to the case where the distribution of  $u_i$  is "mostly dogs, few stars". While we cannot know which of the biases,  $b_{\max}$  or  $b_{\min}$ , will be larger in magnitude in any empirical analysis, it would be easy to speculate based on one's knowledge of the relative frequencies of dogs and stars that occur in the sample. Obviously when the relative frequency of dogs and stars is equal, one would speculate that the biases be equal in magnitude. These types of results are borne out in simulations that follow.

Table 1 reports the simulation results on the biases  $b_{\max}$  and  $b_{\min}$ . Ignoring regressors in equation 1, simulations are performed with  $v_{it} \sim iidN(0, \sigma_v^2)$  and  $u_i$  distributed  $B(8, 2)$  or  $B(2, 2)$  or  $B(2, 8)$ . Each Beta distribution represents different efficiency scenarios as described above. As is standard in SF model simulations, we define  $\gamma = Var(u)/(\sigma_v^2 + Var(u))$  and vary  $\sigma_v^2$  so that  $\gamma = 0.1, 0.5$  and  $0.9$ . The  $\gamma$  is a "signal to noise ratio" measure, so small  $\gamma$  indicates a particularly noisy experiment. We focus on the cases where  $T$  is small and bias of the estimates of the min- and max-measures will be largest, so we fix  $T = 10$ . We consider four values of  $N = 10, 20, 50$  and  $100$ .

The bias analysis in Table 1 contains no surprises. Bias for both estimates is increasing in  $N$  (for fixed  $T$ ) and decreasing in  $\gamma$  as uncertainty over the best and worst firms in the population increases. Varying the skew of the distribution of inefficiency also produces predictable results. The max-measure estimate outperforms our min-measure estimate when the distribution of inefficiency is  $B(8, 2)$ , while the min-measure estimate outperforms the max-measure estimate when the distribution of inefficiency is  $B(2, 8)$ . Compare the 0.077 of  $B(8, 2)$  and 0.122 of  $B(2, 8)$  for measure  $u_i^*$  (first row of results) to the 0.125 of  $B(8, 2)$  and 0.077 of  $B(2, 8)$  for measure  $u_i^{**}$  (second row). This near-perfect symmetry of the results across the two inefficiency distributions occurs every-

where in the table for obvious reasons. When the inefficiency distribution is symmetric,  $B(2, 2)$ , the estimates perform equally with any differences in bias being caused by sampling variability of the simulations. Compare 0.097 for  $u_i^*$  to 0.099 for  $u_i^{**}$  in the first and second rows of results for  $B(2, 2)$ . The implications are clear: in an industry marked with mostly stars and few dogs, the "min" operator in the min-measure estimate induces a smaller bias than the max operator of the max-measure estimate. Put more generally, the min-estimator is less biased than the max-estimator when the industry under study has many efficient firms. (In competitive markets this may be the relevant case.) In any empirical exercise, if bias concerns outweigh the choice of the inefficiency measure employed ( $u_i^*$  vs.  $u_i^{**}$ ), then the choice of estimator should be based on prevailing efficiency market conditions in the industry under study. Knowledge of the distribution of inefficiency (up to location) is contained in the distribution of the estimated  $\alpha_i$  and can be used to inform these empirical choices.

### 3 Two-Sided Measure

We now consider a two-sided measure that incorporates both the max operator and the min operator. The motivation of the two-sided measure is the issue of scale. By scale we mean the way in which estimators are normalized (transformed) to the unit interval. For the max-measure we have the normalization  $r_i^*$ , which rescales (distorts) efficiency differences with the exponential function.<sup>5</sup> Due to the non-linearity of the exponential function, technical efficiency differences between firms in the low range of  $\widehat{r}_i^*$  are smaller than those in the high range, for a given difference in  $\widehat{\alpha}_i$ , and are, therefore, not comparable. Hence, efficiency differences in the low range of  $\widehat{r}_i^*$  are less informative. This creates a distortion in the efficiency differences for  $\widehat{r}_i^*$  when it serves as a normalization for  $-\widehat{u}_i^*$ . The normalization of the min-measure,  $r_i^{**}$ , also nonlinearly rescales (distorts) efficiency

---

<sup>5</sup>This idea of distortion is based on the idea that  $r_i^* \approx u_i^*$ , for small  $u_i^*$ . Obviously if output is in logarithms, then  $r_i^*$  is not simply a normalization; it is the true percentage of the output of firm  $i$  relative to the efficient firm for a fixed set of inputs. However, the normalization could magnify the bias associated with estimating  $u_i^*$ . If output is *not* in logarithms, then  $r_i^*$  is simply a normalization.

differences.

An alternative (or complementary) efficiency measure that does not distort efficiency differences yet normalizes efficiency scores on the unit interval is the two-sided:

$$e_i \equiv \frac{\alpha_i - \min_j \alpha_j}{\max_j \alpha_j - \min_j \alpha_j},$$

with estimate

$$\hat{e}_i \equiv \frac{\hat{\alpha}_i - \min_j \hat{\alpha}_j}{\max_j \hat{\alpha}_j - \min_j \hat{\alpha}_j}.$$

Compared to  $\hat{r}_i^*$  or  $\hat{r}_i^{**}$ , technical efficiency differences  $(\hat{e}_i - \hat{e}_j)$  are not distorted:

$$\hat{e}_i - \hat{e}_j = \frac{\hat{\alpha}_i - \hat{\alpha}_j}{\max_j \hat{\alpha}_j - \min_j \hat{\alpha}_j}.$$

Since  $\hat{\alpha}_i$  is a consistent estimator in  $T$  for  $\alpha - u_i$ , then the difference  $\hat{e}_i - \hat{e}_j$  is consistent for  $-u_i - (-u_j)$ . Since the denominator is constant for each pair of firms, efficiency differences have the same scale across the sample.

To demonstrate the distortion induced by  $\hat{r}_i^*$  or  $\hat{r}_i^{**}$  relative to  $\hat{e}_i$  we again consider a beta distribution for technical inefficiency. By considering different levels of skew, we are considering different levels of efficiency differences between ranked sample realizations at the high and low ends of the rank statistic. For example any ranked sample from the  $B(8, 2)$  or  $B(2, 8)$  distributions will have larger differences at one end of the rank statistic and smaller differences at the other (on average). The  $B(2, 2)$  distribution will have symmetrical differences at either end of the ranked sample, because the probability mass is symmetric about the mean. There are many ways that we could illustrate these differences and the distortions created by normalization to the unit interval. One way would be to use the distribution of  $u_i$  to calculate the theoretical distributions of the transformations  $r_i^*$ ,  $r_i^{**}$  and  $e_i$ . Then the distortions could be compared simply by comparing plots

of the distributions. However, rather than calculate these distributions (a not trivial task), we simulate and estimate them using kernel techniques.

We simulate each beta distribution with 100,000 draws of  $u_i$ . With this many draws the maximal draw is arbitrarily close to 1 and the minimal draw is arbitrarily close to 0, so "uncertainty" over the population maximum and minimum is essentially zero, and any "bias" caused by this uncertainty is mitigated. Our purpose is to get a fairly accurate picture of the distribution and not to understand the effects of sampling variability on efficiency estimation, which we investigated in the last section. We estimate the distributions of  $u_i$ ,  $r_i^*$ ,  $r_i^{**}$  and  $e_i$  using the Gaussian kernel and an arbitrarily selected bandwidth of 0.1. The estimated distributions are in Figures 2*a-c* for  $u_i \sim B(8, 2)$ ,  $u_i \sim B(2, 8)$ , and  $u_i \sim B(2, 2)$ , respectively. Obviously, the density estimates are only approximate at the boundaries (there is no boundary bias correction). However, this is fine for the purposes of scale comparisons. Beginning with panel *a* in Figure 2, we see that when the distribution of  $u_i$  (thick dashed line) is "mostly dogs", the estimated distribution of the max-measure estimate,  $r_i^*$ , is fairly close to that of  $u_i$ , while that of the min-measure,  $r_i^{**}$ , is not. In this case the scale distortion of the max-measure is small relative to that of the min-measure. Also, the two-sided measure,  $e_i$ , is comparable to the max-measure in terms of scale distortion. The two-sided measure over-scales in the center of the distribution while the max-measure over-scales in the right tail of the distribution. This makes sense as the two-sided measure is (in some sense) a "middle ground" between the max-measure and the min-measure. The min-measure,  $r_i^{**}$ , clearly over-scales in the left tail of the distribution. Of course things are reversed in the "mostly stars" case of  $u_i \sim B(2, 8)$ , contained in panel *b* of Figure 2. Here the min-measure outperforms the max-measure in terms of scale preservation. Again, the two-sided measure also preserves scale fairly well and is comparable to the min-measure. In panel *c* we see that the distribution of the two-sided measure,  $e_i$ , is nearly identical to the distribution of  $u_i$  (thick dashed line), while the max- and min-measures exhibit large scale distortions. (Again, the reader is reminded that these are merely kernel density

estimates with no end-point correction.) This is not surprising, given the way the two-sided measure is constructed, but the point should be clear on its usefulness when scale preservation of efficiency scores is important. Regardless of the skewness of the inefficiency distribution, the two-sided measure reliably preserves scale (or differences in the rank statistic), while the performance of the max- and min-measures is a function of the distributional skewness.

For completeness we now examine the bias of the two-sided measure with a brief simulation study. The simulated bias results for the estimator  $\hat{e}_i$  in Table 2 use the same parameterizations as the bias results of Table 1. Unlike the  $\hat{u}_i^*$  and  $\hat{u}_i^{**}$  estimates, the bias results for  $\hat{e}_i$  are firm specific, so average biased across firms are reported. A few results are noteworthy. First, the direction of the bias is a function of the skewness of the efficiency distribution. For  $u \sim B(8, 2)$  (mostly dogs) the bias is positive, for  $u \sim B(2, 8)$  (mostly stars) the bias is negative, and for  $u \sim B(2, 2)$  (few stars or dogs) the bias is close to zero. This may suggest that for efficiency distributions with centralized mass or symmetric efficiency distributions, the two sided estimator is the appropriate choice.<sup>6</sup> Indeed, in the symmetric case, the average bias for the two-sided measure is always smaller in absolute value than the biases in the one-sided measures in Table 1. (For these different measures bias comparisons of estimates are not entirely meaningless, because all three measures are essentially unitless percentages.) Second, bias is (not surprisingly) increasing in  $N$  and decreasing in  $\gamma$  for all levels of skewness.

## 4 Confidence Intervals

Per Schmidt and Sickles (1984),  $\hat{\beta}$  converges to  $\beta$  for large  $N$  or  $T$ , while  $\hat{\alpha}_i$  converges to  $\alpha_i$  for large  $T$  only. Therefore, when  $T$  is small (the usual panel case) asymptotic approximations for confidence intervals on functions of  $\alpha_i$  are inappropriate, and a bootstrap method should be employed. See Kim, Kim and Schmidt (2007) for a detailed survey of methods for bootstrap confidence intervals

---

<sup>6</sup>Again this may not be a "choice" *per se*, but the two-sided measure may simply be a convenient way to report efficiency scores,  $u_i$ .

on technical efficiency and a comprehensive simulation of the coverage rates and confidence interval widths of a variety of bootstrap techniques. Our purpose here is two-fold. First, we would like to replicate the salient features of the Kim, Kim and Schmidt (KKS) confidence interval simulations, while experimenting with the skewness of the technical inefficiency distributions using our three parameterizations of the beta distribution. Second (and simultaneously), we extend the simulations to include our min-measure and the two-sided measure. Again, all the estimates considered are for different measures and cannot be consider direct substitutes, but it is useful to empiricists to know which measures are better in a statistical sense when information on the skew of the efficiency distribution is known or can be approximated from the fixed-effects estimates. For simplicity the underlying data generation mechanism for our confidence intervals is identical to that of our bias analysis of section 2. Our overall finding is that the bias associated with max and min operators erodes the coverage rates of the bootstrap confidence intervals, so the relationship between coverage rates and distributional skewness is similar to that between bias and skewness.

The KKS simulation study considers both *direct* bootstrap confidence intervals from the distribution of  $\hat{r}_i^*$  and *indirect* bootstrap confidence intervals from the distribution of  $\hat{u}_i^*$ , which are transformed to confidence intervals on  $r_i^*$ . When the indirect and direct confidence intervals are the same, the interval is said to be *transformation-respecting* (Efron and Tibshirani 1993, p.175). The empirical advantage of transformation-respecting confidence intervals are obvious: the choice of estimator to report (transformed or not transformed) does not affect the coverage probabilities of the intervals. Of all the bootstrap intervals considered by KKS, only the percentile bootstrap (*percentile*) is transformation respecting. However, the bias-corrected with acceleration ( $BC_a$ ) intervals are approximately transformation-respecting (KKS, p.169). They find that the bias-corrected percentile ( $BC$  percentile) intervals are generally not transformation-respecting but conclude that they have better coverage rates than the other bootstrap confidence intervals that they consider. They also find that, when the intervals are not transformation-respecting, the indirect method for inter-



val construction on  $\widehat{r}_i^*$  has better coverage rates than direct methods. Therefore, in what follows we only consider these three confidence interval construction techniques and only for the indirect method. Our bootstrap confidence interval construction procedures are EXACTLY those of KKS, so we do not detail their procedures here. The reader is referred to the KKS study for details on indirectly constructing percentile, bias-corrected percentile and bias-corrected with acceleration intervals. While our coverage rate results are slightly different than those of KKS, our overall findings are the same: for the indirect method, the bootstrap *BC* percentile intervals (Simar and Wilson, 1998) are generally better in terms of coverage rates than either the percentile or the *BC<sub>a</sub>* intervals.<sup>7</sup>

Coverage results for the percentile and the two bias-corrected bootstraps for  $r_i^*$  and for  $r_i^{**}$  using the indirect method are reported in Table 3 . Here the nominal coverage rate is 0.90. Generally speaking, the BC percentile coverage rates appear to be best for all scenarios considered. (This is the general finding of the KKS study.) When the distribution of inefficiency is symmetric,  $B(2, 2)$ , the coverage rates and interval widths for the  $r_i^*$  and for  $r_i^{**}$  measures are identical up to sampling variability for all the bootstrap techniques. This corresponds to the case where the biases of the two measures are the same (few stars, few dogs). This is particularly clear when the signal to noise ratio (sampling variability) is large (small). Not surprisingly the coverage rates are always decreasing in  $N$ , as uncertainty over the best or worst firms in the sample is increasing (as is bias). Coverage rates are uniformly better for  $r_i^*$  when inefficiency is distributed  $B(8, 2)$  and better for  $r_i^{**}$  when inefficiency is distributed  $B(2, 8)$ . For example, when  $N = 100$ ,  $\gamma = 0.1$ , and  $B(8, 2)$  the *BC* percentile coverage rate for  $r_i^*$  and  $r_i^{**}$  are 0.826 and 0.760 in Table 3. However, when inefficiency is distributed  $B(2, 8)$  the respective coverage rate are 0.750 and 0.821. Again these results are driven by the relative level of bias of the two estimates under the different inefficiency

---

<sup>7</sup>A referee also pointed out that the distributions of interest for construction of the bootstrap confidence intervals are invariant to the true values of the model parameters  $\alpha$  and  $\beta$ .

regimes (mostly stars or mostly dogs). In terms of the different interval construction techniques, it appears that both bias-corrected techniques have coverage rates of about 0.7-0.8 regardless of the skew of the distribution, the size of  $N$ , or the signal to noise ratio (except in the noisiest cases). The uncorrected percentile bootstrap intervals do surprisingly well relative to the bias-corrected intervals except in the noisiest cases (large  $N$  and small  $\gamma$ ). For example, for  $N = 100$ ,  $\gamma = 0.1$ , and  $B(2, 2)$  the percentile coverages for  $r_i^*$  and  $r_i^{**}$  are 0.359 and 0.349, respectively. However, even then the bias-corrected intervals are not very impressive. For example, the  $BC_a$  intervals are 0.624 and 0.611, respectively in this case. The worst coverage probability is the uncorrected percentile bootstrap in the noisiest case ( $N = 100$ ,  $\gamma = 0.1$ ) for  $r_i^{**}$  and  $B(8, 2)$ . In this case, the coverage rate is only 0.254.

For completeness the bootstrap confidence intervals for the two-sided measure,  $e_i$ , are provided in Table 4. Notice that the coverage rates for this measure are fairly stable across the different technical inefficiency distributions. For example, in the least noisy setting ( $N = 10$ ,  $\gamma = 0.9$ ) the coverage rates for the  $BC_a$  intervals are 0.847, 0.854 and 0.851 for  $B(8, 2)$ ,  $B(2, 2)$  and  $B(2, 8)$ , respectively. The interval widths are also about the same. Again, this is due to the fact that scale distortion is minimal for the two-sided measure across the different levels of skew. Also, the uncorrected percentile bootstrap has generally higher coverage rates than the bias-corrected intervals, and in the least noisy cases it comes very close to achieving the nominal coverage rate of 0.90. While the measures in this study are all different, it is interesting to note that, when inefficiency is distributed  $B(2, 2)$ , the coverage probabilities on the two-sided measure (Table 4) are uniformly higher than those of the one-sided measures (Table 3). This illustrates the effects of scale distortions of the exponential transformations of the  $u_i^*$  and  $u_i^{**}$  when the distribution is symmetric. This is depicted in panel *c* of Figure 2.

## 5 Rice Farm Application

The quintessential example of large  $N$  and small  $T$  in the stochastic frontier literature is the Indonesian rice farm data set with  $N = 171$  and  $T = 6$ . This particular data set has been analyzed a number of times, starting with Erwidodo (1990) and, most recently, with Kim, Kim and Schmidt (2007). See Horrace and Schmidt (1996, 2000) for a detailed description. In our example we ignore the issue of bias caused by unidentifiable time-invariant inputs in fixed-effects estimation (Feng and Horrace, 2007). The form of the production function and the parameter estimates are precisely those contained in Horrace and Schmidt (2000), but what is important to know is that output is in logarithms of kilograms of rice. Our purpose is to highlight the different efficiency measures considered. The distribution of the  $\hat{\alpha}_i$  in Figure 3 is produced using the `ksdensity(x)` command in MATLAB and a Gaussian kernel. It has a normalized positive skewness of 0.4740. Therefore, the distribution of the  $\hat{u}_i^*$  (up to location) is its mirror image and has a skewness of -0.4740. However, the distribution in Figure 3 is actually quite symmetric (except for a small wiggle in the right tail). Table 5 presents efficiency estimates for 7 of the rice farms. These are ranked by  $\hat{\alpha}_i$  and correspond to the 2 best farms, the 75%ile farm, the median farm, the 25%ile farm and the two worst farms. Each entry in the last three columns contains the estimate and the indirect 90% bias-corrected percentile confidence interval based on 999 bootstrap replications.

In any empirical exercise a discussion of potential bias is difficult. However, given our simulation results and the fact that the distribution of the  $\hat{\alpha}_i$  is nearly symmetric (or  $u_i^*$  have slight negative skewness), perhaps the two-sided measure (or traditional measure) will have a less biased estimate than that of the the measure relative to the least efficient farm  $r_i^{**}$ . Even though the measures are different, they are all unitless, so their biases will be unitless and, perhaps, comparisons are not entirely unreasonable. However, for this particular data set the efficiency measures are quite imprecise, judging from the confidence intervals. This is particularly telling, when one considers the

poor coverage rates of the bootstrap confidence intervals that arose in our (and KKS's) simulation study. The implication being that the confidence intervals in Table 5 may only achieve 70-80% coverage rates, even after bias correction.

We now discuss scale considerations. The difference in the  $\hat{\alpha}_i$  of the two best farms is 0.070 log-points. For the measure  $r_i^*$  we see that the second-best farm has technical efficiency 6.8 percentage points below the best farm (1.000-0.932), a good approximation for the log-point difference in the  $\hat{\alpha}_i$ . Ceteris paribus, efficiency differences suggest that the second-best farm will produce 6.8% fewer kilograms of rice (0.070 fewer log-points of rice) than the best farm. For  $r_i^{**}$  we see that the second-best farm is 27 percentage points below the best, so its approximation for the the log-point difference in the  $\hat{\alpha}_i$  is poor at this end of the order statistic (where  $\bar{u} - u_i$  is large). Things are reversed at the other end of the order statistic. The log-point difference in the  $\hat{\alpha}_i$  for the two worst firms is approximated well by  $r_i^{**}$  and poorly by  $r_i^*$ . By definition the two-sided measure,  $e_i$ , will always approximate these differences well, particularly at the ends of the order statistic. All three measures approximate the log-point differences less precisely in the middle of the order statistic, but it is clear what the two-sided measure is doing: it is a middle ground between the two exponential measures. For the median farm  $r_i^* = 0.544$ ,  $r_i^{**} = 0.340$  and  $e_i = 0.413$ . Hence, for reporting purposes, the two-sided measure is a simple log-point normalization that facilitates discussion of relative efficiency over the entire range of the order statistic and that approximates well the percentage change of output ( $r_i^*$  and  $r_i^{**}$ ) at both ends of the order statistic.

## 6 Conclusions

The goal of this research is to consider the performance of various technical efficiency measures under different skewness of the distribution of technical inefficiency. We find that the traditional one-sided estimate relative to the sample maximum,  $\hat{r}_i^*$ , performs best in terms of bias and confidence interval coverage rates when the distribution of inefficiency consists of "few stars, mostly dogs." In

highly competitive markets where inefficient firms are rare, estimators of the traditional measure may not be reliable (large bias and poor interval coverage). On the other hand, estimators of the traditional measure may be reliable in markets where competitive forces are weak, and this may be the empirically relevant case for efficiency measurement in general. That is, estimating technical efficiency may only be meaningful in markets or industries where it might already exist to a great extent (like utility industries, where capital barriers to entry limit competitive forces). The proposed min-measure,  $\widehat{r}_i^{**}$ , has small bias and better confidence interval coverage when the inefficiency distribution has "mostly stars, few dogs, which may correspond to highly competitive industries. The majority of economic theory would suggest that this corresponds to the more frequently encountered case. Of course in competitive markets, technical efficiency estimation may be difficult from the start, but that does not diminish the potential importance of the min-measure. For example, the marginal cost of the least efficient firm in the sample may equal the market price, so a measure relative to the least efficient firm in the industry may be useful. Estimates of the two-sided measure,  $e_i$ , are particularly appealing when the distribution of inefficiency is symmetric and when issues of scale are important. That is, when the magnitude of the differences of the  $u_i^*$  must be preserved, the two-sided measure normalizes scores to the unit intervals without the nonlinear scale distortion induced by the exponent operator.

We reiterate that all the measures are different, so comparisons between the measures are sometimes difficult to interpret. However, this study adds a few new measures to the empiricist's toolbox that may prove useful in the future. Our example suggests that the inefficiency distribution of Indonesian rice farms is fairly symmetric; this suggests that the two-sided estimator may be preferred in terms of bias and confidence interval coverage of the measure. Symmetry aside, the two-sided measure necessarily preserves the scale of the  $\widehat{u}_i^*$ , better than the traditional measure in this (or any) particular example.

## References

- [1] Battese GE, Coelli TJ. 1988. Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics* 38: 387-399.
- [2] Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. Chapman and Hall: New York, NY.
- [3] Entani T, Maeda Y, Tanaka H. 2002. Dual models of interval DEA and its extension to interval data, *European Journal of Operational Research* 136: 32-45.
- [4] Erwidodo. 1990. *Panel data analysis on farm-level efficiency, input demand and output supply of rice farming in West Java, Indonesia*. PhD dissertation. Department of Agricultural Economics. Michigan State University.
- [5] Feng Q, Horrace WC. 2007. Fixed-effect Estimation of Technical Efficiency with Time-invariant Dummies. *Economics Letters* 95: 247-252.
- [6] Hall P, Hardle W, Simar L. 1993. On the inconsistency of bootstrap distribution estimators. *Computational Statistics and Data Analysis* 16: 11-18.
- [7] Horrace WC, Schmidt P. 1996. Confidence statements for efficiency estimates from stochastic frontier models, *Journal of Productivity Analysis*, 7, 257-282.
- [8] Horrace WC, Schmidt P. 2000. Multiple comparisons with the best, with economic applications. *Journal of Applied Econometrics* 15: 1-26.
- [9] Kim M, Kim Y, Schmidt P. 2007. On the accuracy of bootstrap confidence intervals for efficiency levels in stochastic frontier models with panel data. *Journal of Productivity Analysis* 28: 165–181.

- [10] Koop G, Osiewalski J, Steel MFJ. 1997. Bayesian efficiency analysis through individual effects: hospital cost frontiers. *Journal of Econometrics* 76: 77-105.
- [11] Qian J, Sickles RC. 2008. Stochastic Frontiers with Bounded Inefficiency. Unpublished manuscript. Department of Economics. Rice University.
- [12] Satchachai P, Schmidt P., 2008. Estimates of Technical Inefficiency in the Stochastic Frontier Model with Panel Data: Generalized Panel Jackknife Estimation. Unpublished manuscript. Michigan State University.
- [13] Schmidt P, Sickles RC. 1984. Production frontiers and panel data. *Journal of Business and Economic Statistics* 2: 367-374.
- [14] Simar L. 1992. 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: 171-203.
- [15] Simar L, Wilson P. 1998. Sensitivity of efficiency scores: How to bootstrap in nonparametric frontier models, *Management Science* 44: 49-61.

Table 1. Biases of Estimates of the Min- and Max-Measures.

Measure	$T$	$\gamma$	$N$	Bias		
				$B(8, 2)$	$B(2, 2)$	$B(2, 8)$
$u_i^*$	10	0.1	10	0.077	0.097	0.122
$u_i^{**}$				0.125	0.099	0.077
$u_i^*$	10	0.1	20	0.092	0.129	0.160
$u_i^{**}$				0.162	0.130	0.088
$u_i^*$	10	0.1	50	0.106	0.174	0.212
$u_i^{**}$				0.211	0.174	0.106
$u_i^*$	10	0.1	100	0.117	0.208	0.249
$u_i^{**}$				0.248	0.207	0.112
$u_i^*$	10	0.5	10	0.021	0.029	0.046
$u_i^{**}$				0.043	0.029	0.022
$u_i^*$	10	0.5	20	0.023	0.046	0.064
$u_i^{**}$				0.062	0.042	0.022
$u_i^*$	10	0.5	50	0.029	0.068	0.094
$u_i^{**}$				0.094	0.065	0.028
$u_i^*$	10	0.5	100	0.029	0.087	0.120
$u_i^{**}$				0.119	0.085	0.031
$u_i^*$	10	0.9	10	0.002	0.004	0.009
$u_i^{**}$				0.007	0.004	0.002
$u_i^*$	10	0.9	20	0.004	0.007	0.012
$u_i^{**}$				0.011	0.006	0.004
$u_i^*$	10	0.9	50	0.002	0.011	0.019
$u_i^{**}$				0.019	0.013	0.003
$u_i^*$	10	0.9	100	0.005	0.017	0.026
$u_i^{**}$				0.026	0.016	0.004



Table 2. Average Bias of the Two-sided Estimate,  $\hat{e}_i$

Measure	$T$	$\gamma$	$N$	Average Bias		
				$B(8, 2)$	$B(2, 2)$	$B(2, 8)$
$e_i$	10	0.1	10	0.0555	-0.0028	-0.0554
	10	0.1	20	0.0725	0.0002	-0.0743
	10	0.1	50	0.1012	0.0012	-0.0983
	10	0.1	100	0.1124	0.0013	-0.1148
$e_i$	10	0.5	10	0.0616	0.002	-0.0153
	10	0.5	20	0.0220	-0.006	-0.0241
	10	0.5	50	0.0314	0.0014	-0.0300
	10	0.5	100	0.0381	0.0013	-0.0374
$e_i$	10	0.9	10	0.0027	-0.0009	-0.0025
	10	0.9	20	0.0025	0.0008	-0.0029
	10	0.9	50	0.0055	-0.0006	-0.0051
	10	0.9	100	0.0068	0.0004	-0.0069

Table 3: 90% Indirect Bootstrap Confidence Intervals for Technical Efficiency Measures

Measure	T	gamma	N	Percentile						BCa						BC percentile					
				B(8,2)		B(2,2)		B(2,8)		B(8,2)		B(2,2)		B(2,8)		B(8,2)		B(2,2)		B(2,8)	
				cover	width	cover	width	cover	width	cover	width	cover	width	cover	width	cover	width	cover	width	cover	width
r*	10	0.1	10	0.780	0.342	0.754	0.347	0.702	0.348	0.823	0.332	0.819	0.329	0.810	0.327	0.828	0.387	0.838	0.399	0.838	0.402
r**				0.712	0.348	0.757	0.348	0.785	0.342	0.757	0.298	0.773	0.304	0.787	0.304	0.847	0.404	0.841	0.399	0.832	0.389
r*	10	0.1	20	0.734	0.331	0.640	0.335	0.563	0.339	0.794	0.319	0.773	0.315	0.750	0.309	0.823	0.383	0.824	0.396	0.823	0.406
r**				0.565	0.340	0.651	0.335	0.735	0.331	0.715	0.299	0.750	0.301	0.775	0.305	0.820	0.408	0.831	0.397	0.829	0.384
r*	10	0.1	50	0.673	0.306	0.482	0.313	0.372	0.320	0.758	0.294	0.697	0.289	0.637	0.286	0.816	0.364	0.804	0.384	0.783	0.401
r**				0.371	0.320	0.486	0.313	0.672	0.304	0.626	0.283	0.689	0.284	0.756	0.290	0.786	0.401	0.812	0.385	0.823	0.360
r*	10	0.1	100	0.630	0.284	0.359	0.295	0.254	0.303	0.748	0.275	0.624	0.270	0.545	0.265	0.826	0.342	0.780	0.371	0.750	0.388
r**				0.254	0.303	0.349	0.295	0.627	0.284	0.542	0.262	0.611	0.268	0.740	0.272	0.760	0.390	0.775	0.371	0.821	0.343
r*	10	0.5	10	0.858	0.214	0.848	0.228	0.830	0.241	0.839	0.223	0.833	0.236	0.835	0.245	0.824	0.223	0.824	0.241	0.832	0.258
r**				0.825	0.241	0.844	0.228	0.860	0.211	0.801	0.224	0.822	0.217	0.840	0.203	0.827	0.259	0.824	0.241	0.834	0.221
r*	10	0.5	20	0.841	0.198	0.809	0.220	0.766	0.237	0.831	0.204	0.818	0.223	0.804	0.232	0.831	0.208	0.827	0.237	0.825	0.260
r**				0.759	0.238	0.816	0.219	0.833	0.198	0.785	0.224	0.797	0.211	0.818	0.195	0.828	0.260	0.818	0.235	0.820	0.208
r*	10	0.5	50	0.837	0.171	0.746	0.203	0.649	0.222	0.820	0.175	0.797	0.200	0.764	0.212	0.825	0.181	0.828	0.224	0.821	0.251
r**				0.645	0.223	0.750	0.203	0.824	0.168	0.758	0.208	0.790	0.197	0.818	0.169	0.827	0.253	0.831	0.224	0.827	0.178
r*	10	0.5	100	0.817	0.151	0.663	0.190	0.525	0.210	0.809	0.154	0.767	0.187	0.711	0.197	0.821	0.161	0.820	0.214	0.803	0.243
r**				0.523	0.211	0.677	0.191	0.826	0.150	0.704	0.196	0.764	0.184	0.814	0.151	0.807	0.244	0.824	0.216	0.827	0.160
r*	10	0.9	10	0.856	0.083	0.862	0.094	0.856	0.106	0.844	0.086	0.846	0.098	0.839	0.110	0.832	0.084	0.831	0.095	0.821	0.107
r**				0.860	0.106	0.862	0.094	0.860	0.083	0.838	0.103	0.850	0.093	0.849	0.082	0.822	0.107	0.832	0.095	0.837	0.083
r*	10	0.9	20	0.855	0.075	0.844	0.091	0.851	0.105	0.845	0.077	0.831	0.094	0.840	0.108	0.838	0.076	0.824	0.092	0.835	0.108
r**				0.842	0.105	0.852	0.092	0.854	0.076	0.822	0.102	0.836	0.091	0.842	0.075	0.824	0.107	0.833	0.093	0.836	0.076
r*	10	0.9	50	0.845	0.062	0.838	0.087	0.806	0.101	0.833	0.063	0.822	0.087	0.807	0.100	0.834	0.062	0.827	0.088	0.821	0.104
r**				0.815	0.101	0.834	0.086	0.843	0.062	0.806	0.098	0.819	0.086	0.831	0.062	0.824	0.104	0.823	0.088	0.830	0.063
r*	10	0.9	100	0.845	0.054	0.819	0.082	0.773	0.098	0.837	0.055	0.813	0.082	0.806	0.095	0.839	0.054	0.826	0.085	0.834	0.102
r**				0.770	0.098	0.817	0.082	0.846	0.053	0.795	0.094	0.812	0.081	0.836	0.054	0.827	0.101	0.826	0.084	0.837	0.054

Table 4: 90% Indirect Bootstrap Confidence Intervals for the Two-sided Measure

T	gamma	N	Percentile						BCa						BC percentile					
			B(8,2)		B(2,2)		B(2,8)		B(8,2)		B(2,2)		B(2,8)		B(8,2)		B(2,2)		B(2,8)	
			cover	width	cover	width	cover	width	cover	width	cover	width	cover	width	cover	width	cover	width	cover	width
10	0.1	10	0.809	0.567	0.823	0.570	0.806	0.568	0.714	0.487	0.729	0.489	0.733	0.487	0.714	0.567	0.712	0.570	0.703	0.568
10	0.1	20	0.750	0.497	0.786	0.501	0.753	0.496	0.681	0.437	0.699	0.440	0.696	0.435	0.686	0.497	0.692	0.501	0.687	0.496
10	0.1	50	0.667	0.424	0.748	0.428	0.669	0.423	0.640	0.383	0.674	0.388	0.651	0.382	0.653	0.424	0.673	0.428	0.656	0.423
10	0.1	100	0.601	0.380	0.714	0.385	0.599	0.380	0.612	0.348	0.657	0.356	0.611	0.349	0.634	0.380	0.659	0.385	0.627	0.380
10	0.5	10	0.861	0.305	0.867	0.299	0.865	0.300	0.805	0.289	0.819	0.289	0.829	0.288	0.795	0.305	0.788	0.299	0.798	0.300
10	0.5	20	0.823	0.270	0.846	0.274	0.827	0.270	0.787	0.257	0.798	0.263	0.806	0.257	0.790	0.270	0.789	0.274	0.793	0.270
10	0.5	50	0.769	0.230	0.831	0.242	0.772	0.228	0.772	0.217	0.787	0.232	0.782	0.215	0.788	0.230	0.788	0.242	0.789	0.228
10	0.5	100	0.709	0.207	0.809	0.224	0.705	0.205	0.748	0.194	0.768	0.215	0.749	0.193	0.775	0.207	0.772	0.224	0.749	0.205
10	0.9	10	0.876	0.114	0.874	0.111	0.873	0.113	0.847	0.113	0.854	0.112	0.851	0.116	0.814	0.114	0.819	0.111	0.817	0.113
10	0.9	20	0.856	0.102	0.861	0.104	0.864	0.103	0.829	0.102	0.838	0.105	0.847	0.104	0.816	0.102	0.819	0.104	0.829	0.103
10	0.9	50	0.839	0.088	0.850	0.096	0.832	0.089	0.820	0.087	0.824	0.096	0.825	0.088	0.826	0.088	0.818	0.096	0.823	0.089
10	0.9	100	0.815	0.081	0.846	0.091	0.817	0.080	0.813	0.079	0.818	0.091	0.820	0.079	0.824	0.081	0.819	0.091	0.828	0.080

Table 5. Rice Farm Efficiency Estimates.

$\alpha_i$	%ile	$r_i^*$	$r_i^{**}$	$e_i$
5.556	best	1.000 [ 0.912, 1.000 ]	0.635 [ 0.493, 0.667 ]	1.000 [ 0.929, 1.000 ]
5.486	99 <sup>th</sup>	0.932 [ 0.837, 1.000 ]	0.608 [ 0.463, 0.658 ]	0.930 [ 0.836, 1.000 ]
5.072	75 <sup>th</sup>	0.616 [ 0.423, 0.890 ]	0.407 [ 0.000, 0.566 ]	0.519 [ 0.156, 0.780 ]
4.970	50 <sup>th</sup>	0.554 [ 0.448, 0.737 ]	0.340 [ 0.078, 0.450 ]	0.413 [ 0.197, 0.566 ]
4.859	25 <sup>th</sup>	0.498 [ 0.439, 0.608 ]	0.266 [ 0.000, 0.387 ]	0.308 [ 0.070, 0.436 ]
4.586	1 <sup>st</sup>	0.379 [ 0.306, 0.477 ]	0.036 [ 0.000, 0.215 ]	0.036 [ 0.000, 0.790 ]
4.550	worst	0.366 [ 0.275, 0.531 ]	0.000 [ 0.000, 0.189 ]	0.000 [ 0.000, 0.157 ]

Bracketed values are 90% bias-corrected percentile intervals.

Figure 1. Skewness of Various Beta Distributions

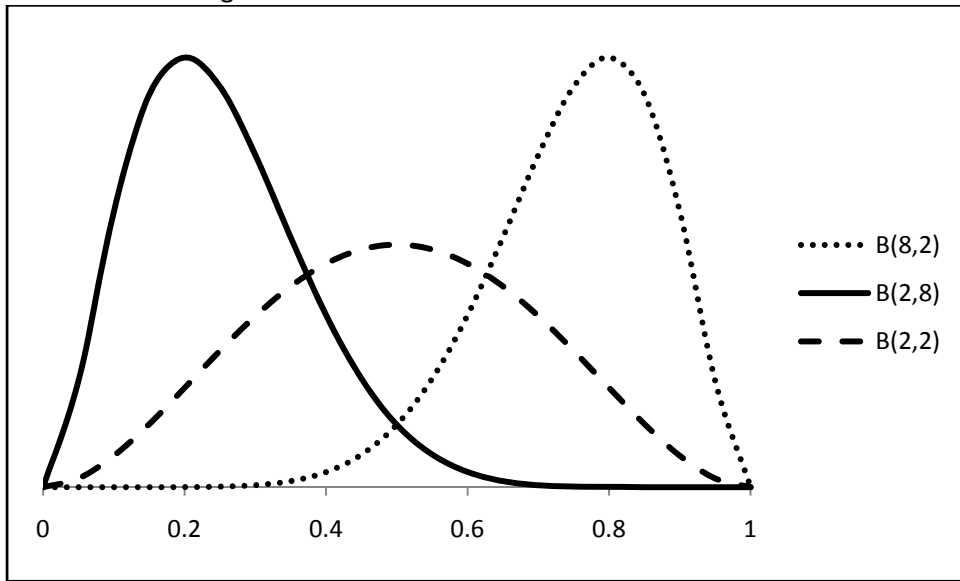


Figure 2a. Efficiency Distribution Estimates – Mostly Dogs B(8,2)

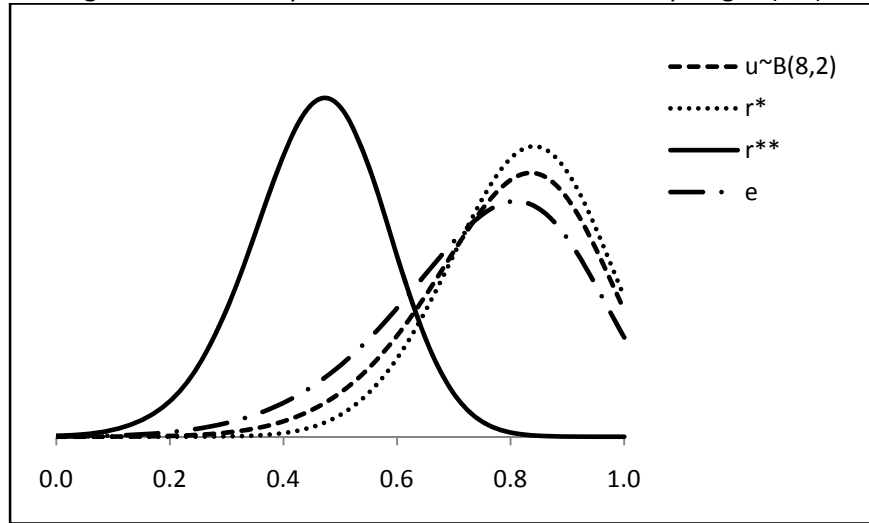


Figure 2b. Efficiency Distribution Estimates – Mostly Stars B(2,8)

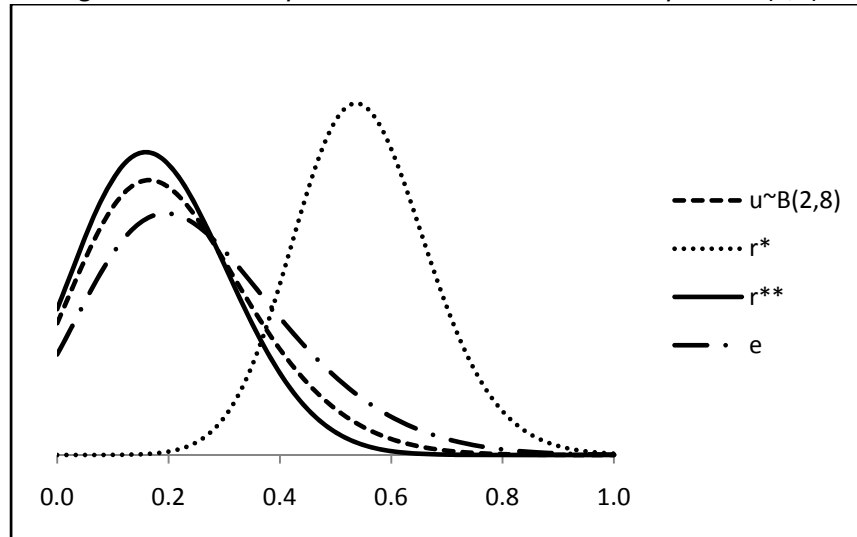
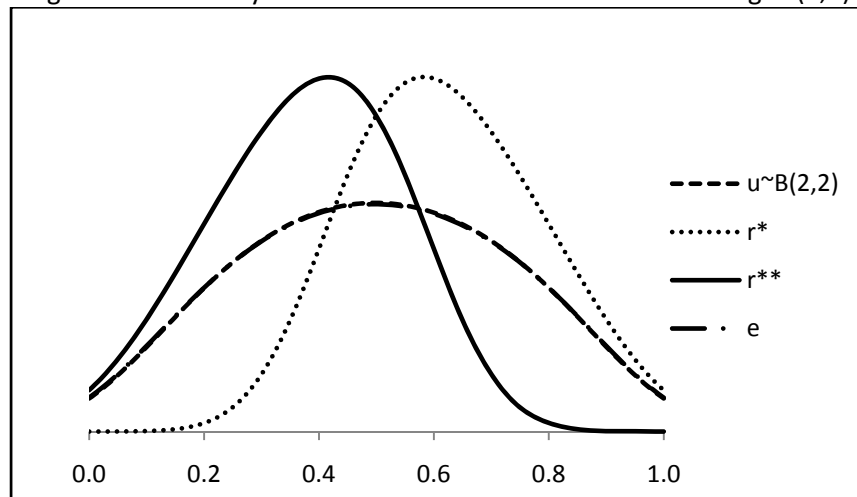


Figure 2c. Efficiency Distribution Estimates – Few Stars or Dogs B(2,2)



Note: the curves  $u$  and  $e$  are indistinguishable.

Figure 3. Distribution of  $\alpha_i$  for 171 Rice Farms. Skewness = 0.4740

