Statistical Inference for the Russell Measure of Technical Efficiency

Oleg Badunenko* and Pavlo Mozharovskyi†

December 23, 2018

Abstract

Data Envelopment Analysis (DEA) has become a popular approach to nonparametric efficiency measurement. The statistical inference using bootstrap methods is readily available for the radial DEA estimator, however it is missing for the Russell measure, the nonradial DEA estimator. We propose a bootstrap based procedure for making statistical inference about the individual Russell measures of technical efficiency. We perform simulations to examine finite sample properties of the proposed estimator. Finally, we present an empirical study using proposed bootstrap procedure.

Keywords: Nonradial efficiency; Russell measure; Data envelopment analysis; Frontier Efficiency Models; Statistical inference; Bootstrap

*Corresponding author, Economics and Finance Group, Portsmouth Business School, University of Portsmouth, Portsmouth, United Kingdom. Phone: +44 2392 84 4831, E-mail: oleg.badunenko@port.ac.uk

†LTCI, Telecom ParisTech, University of Paris-Saclay, 46 rue Barrault, 75013 Paris, France.

1 Introduction

Data Envelopment Analysis (DEA) methods are broadly used to evaluate technical efficiency. Radial DEA methods are based on measure of technical efficiency proposed by Debreu (1951) and Farrell (1957). This measure is defined as *proportional* expansion of *all* outputs which is technologically feasible in output-augmenting orientation and as *proportional* reduction of *all* inputs which is technologically feasible in input-conserving orientation. According to the definition of Koopmans (1951) however, technical inefficiency implies that given inputs at least one output can be increased or given outputs at least one input can be reduced. Debreu-Farrell definition is less demanding as efficient observations are required to belong to isoquant, whereas they have to be in the efficiency subset in terms of the Koopman's definition. Often, isoquant and efficient subset differ (Färe et al., 1994). Koopman's idea can be operationalized as a nonradial measure of technical efficiency. One way to accommodate such characteristic is to use directional distance functions introduced by Chambers et al. (1998) and discussed in Färe and Grosskopf (2004) (for the optimal directional vectors). Another option is to rely on the Russell measure, which allows for non-proportional expansions and reductions (Färe et al., 1994).

Statistical inference regarding radial estimators can be based on bootstrap techniques (Simar and Wilson, 2015). Simar and Vanhems (2012) offer probabilistic formulation of the technology described by directional distances. Simar et al. (2012) propose a bootstrap procedure for statistical inference about directional distances. The suggested procedure however assumes that the direction is given and the linear problem is optimized over one parameter. The existing algorithm therefore cannot be directly carried over to the Russell measure. We propose an adjustment to existing bootstrap procedures to enable making inference about the Russell measure of technical efficiency.

The paper is organized as follows. Section 2 introduces Russell measure, the bootstrap procedure for performing statistical inference for Russell mea-

sure is described in Section 3. Section 4 presents a set of simulations to examine the finite sample properties of the proposed bootstrap procedure. Section 5 presents an empirical application and Section 6 concludes.

2 Russell measure of technical efficiency

2.1 Technology

Consider a production process in which multiple inputs produce multiple outputs. If vector $\boldsymbol{x} = (x_1, \ldots, x_N)$ denotes N inputs and vector $\boldsymbol{y} = (y_1, \ldots, y_M)$ denotes M outputs, the production technology $\mathcal{T}$ can then be broadly defined as

$$\mathcal{T} = \{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{y} \text{ are producible by } \boldsymbol{x}\}.$$

The true technology set $\mathcal{T}$ is typically not observed in practice and is usually approximated with the help of activity analysis models and operationalized or estimated via the linear-programming technique (Färe et al., 1994). Given K decision making units and assuming that technology satisfies constant returns to scale (CRS),

$$\begin{aligned} \hat{\mathcal{T}}^{\text{CRS}} = \big\{(\boldsymbol{x}, \boldsymbol{y}) : & \sum_{k=1}^{K} z_k y_{km} \geq y_{km},\ m = 1, \ldots, M, \\ & \sum_{k=1}^{K} z_k x_{kn} \leq x_{kn},\ n = 1, \ldots, N, \\ & z_k \geq 0,\ k = 1, \ldots, K\big\}, \end{aligned} \tag{1}$$

where $\boldsymbol{x} = \{\boldsymbol{x}_k\}_{k=1}^K$ and $\boldsymbol{y} = \{\boldsymbol{y}_k\}_{k=1}^K$, $\boldsymbol{x}_k = (x_{k1}, \ldots, x_{kN})$ and $\boldsymbol{y}_k = (y_{k1}, \ldots, y_{kM})$ denote data vectors of N inputs and M outputs for decision making unit k, $(k = 1, \ldots, K)$ and vector $\boldsymbol{z} = (z_1, \ldots, z_K)$ denotes the intensity variables that help 'envelop' the data with the smallest convex free

disposal cone. Since no parametric assumptions are imposed on the function, the estimator in (1) is referred to as a nonparametric estimator of technology set $\mathcal{T}$, which satisfies CRS, free disposability and convexity. Other returns to scale are modeled by adjusting process intensity levels $\boldsymbol{z}$. If technology satisfies variable returns to scale (VRS) constraint $\sum_{k=1}^{K} z_k = 1$ is added to (1); for non-increasing returns to scale (NIRS), an inequality constraint $\sum_{k=1}^{K} z_k \leq 1$ is added to (1).

2.2 Measure of efficiency

The upper boundary of the technology set $\mathcal{T}$ defines the (technology) frontier. How close a given decision making unit is to the frontier is referred to as its technical efficiency.

Radial measure expands (shrinks) *all* M outputs (*all* N inputs) proportionally until the frontier $\mathcal{T}$ is reached. At the reached frontier point, some but not all outputs (inputs) can be expanded (shrunk) while remaining feasible. If such possibility is available for output m (input n), the reference point is said to have slack in output y_m (input x_n). Nonradial measure of technical efficiency, the Russell measure, accommodates such situations (Färe et al., 1994; Färe and Lovell, 1978). The output-based nonradial measure for decision making unit k is defined by

$$RM_k^o(\boldsymbol{x}_k, \boldsymbol{y}_k|\mathcal{T}) = \max\Big\{M^{-1}\sum_{m=1}^{M}\theta_m : \ (\boldsymbol{x}_k, \theta_1 y_{k1}, \ldots \theta_M y_{kM}) \in \mathcal{T}, \qquad (2)$$
$$\theta_m \geq 1, \ m = 1, \ldots, M\Big\}.$$

The input-based counterpart is given by

$$RM_k^i(\boldsymbol{y}_k, \boldsymbol{x}_k|\mathcal{T}) = \min\Big\{N^{-1}\sum_{n=1}^{N}\lambda_n : \ (\lambda_1 x_{k1}, \ldots \lambda_N x_{kN}, \boldsymbol{y}_k) \in \mathcal{T}, \qquad (3)$$
$$1 \geq \lambda_n \geq 0, \ n = 1, \ldots, N\Big\}.$$

The true $\mathcal{T}$ in (2) and (3) is unobserved and replacing it with its DEA estimate in (1) gives the DEA estimator of output-based Russell measure formulated as

$$\widehat{RM}_k^o(\boldsymbol{x}_k, \boldsymbol{y}_k|\hat{\mathcal{T}}^{\text{CRS}}) = M^{-1} \max_{\theta,z} \sum_{m=1}^{M} \theta_m \tag{4}$$

$$\begin{aligned} \text{s.t. } & \sum_{k=1}^{K} z_k y_{km} \geq y_{km}\theta_m, \ m = 1, \ldots, M, \\ & \sum_{k=1}^{K} z_k x_{kn} \leq x_{kn}, \ n = 1, \ldots, N, \\ & z_k \geq 0, \ k = 1, \ldots, K, \\ & \theta_m \geq 1, \ m = 1, \ldots, M. \end{aligned} \tag{5}$$

and input-based Russell measure, viz.,

$$\widehat{RM}_k^i(\boldsymbol{y}_k, \boldsymbol{x}_k|\hat{\mathcal{T}}^{\text{CRS}}) = N^{-1} \min_{\lambda,z} \sum_{n=1}^{N} \lambda_n \tag{6}$$

$$\begin{aligned} \text{s.t. } & \sum_{k=1}^{K} z_k y_{km} \geq y_{km}, \ m = 1, \ldots, M, \\ & \sum_{k=1}^{K} z_k x_{kn} \leq x_{kn}\lambda_n, \ n = 1, \ldots, N, \\ & z_k \geq 0, \ k = 1, \ldots, K, \\ & 1 \geq \lambda_n \geq 0, \ n = 1, \ldots, N. \end{aligned} \tag{7}$$

If output $y_{km} = 0$, the linear programming problem in (4) is modified and θ_m is set to 1. If input $x_{kn} = 0$, the linear programming problem in (6) is modified and λ_n is set to 1.

In both (4) and (6), θ_m's and λ_n's should be restricted by 1 from below and above, respectively. If they are not, the Russell measure can still be mathematically calculated, but it will represent something that conceptually is not the Russell measure and will not have the desired properties (Färe et al., 1985). Figure 1 shows the direction toward the frontier in which inefficient observations move when λ_n's are restricted and when they are not restricted.

Consider one-output-two-inputs technology. In Figure 1, the technology is defined by points $ABCDE$. All units produce the same amount of output. We also observe inefficient points F, G, H, and I. Some points will not be affected whether λ_n's are restricted or not. For example, F to be efficient would need to progress to A, G to C and I to D. Points that are unaffected by restrictions are in green area. Point H and others in non-green areas differ in their behavior depending on whether λ_n's are restricted by 1 or not. If λ_n's are not restricted by 1 from above, H will tend to C, while if they are restricted, H will move either in horizontal or vertical direction depending on the shape of the frontier. Not restricting is undesirable as passage from H to C implies that using more of one of the inputs leads to being more efficient.

The Russell measure is conceptually a DEA estimator. The maximizing values $\hat{\theta}_m$'s and minimizing $\hat{\lambda}_n$'s can be roughly thought of as output- and input-specific measures of efficiency. By setting the restriction $\hat{\theta}_m = \hat{\theta}, \forall m = 1, \ldots, M$ or $\hat{\lambda}_n = \hat{\lambda}, \forall n = 1, \ldots, N$, the Russell measure collapses to the Debreu-Farrell measure. Furthermore, in case of one input (output), the input (output)-based Russell measure is equal to the Debreu-Farrell radial measure. Technologies under non-increasing and variable returns to scale can be modeled by imposing respective restrictions on the intensity vector, $\boldsymbol{z}$, in the piecewise linear technology, that is in (4) and (6). Then the Russell measure can be calculated relative to these technologies.

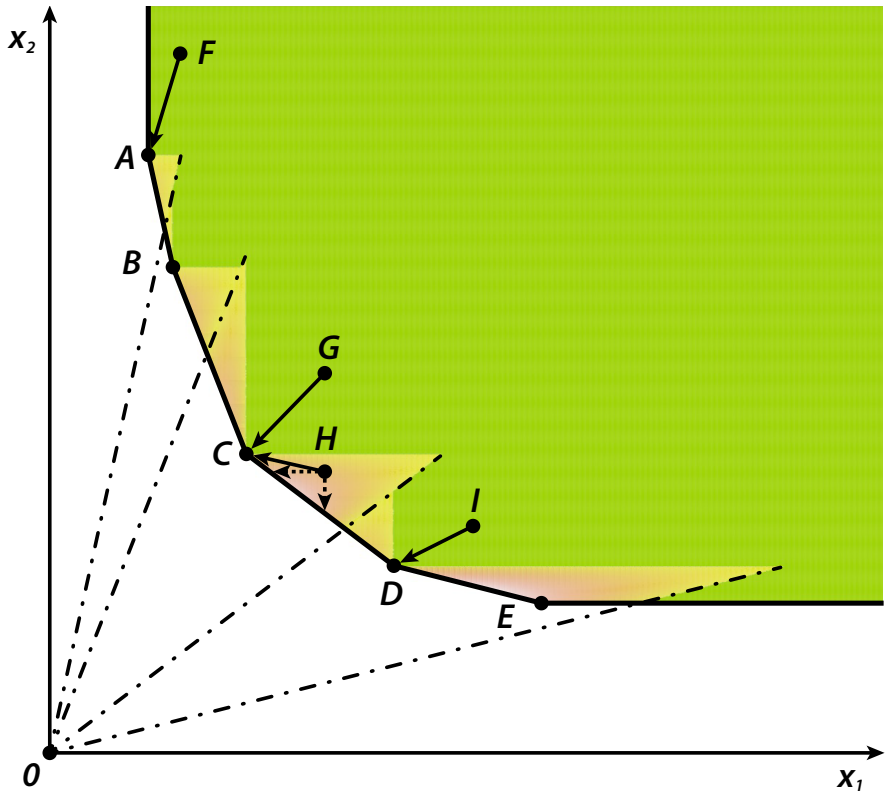

Figure 1 The Russell measure

Notes: Figure shows one-output-two-inputs technology, which is defined by points $ABCDE$. All units produce the same amount of output. Points F, G, H, and I are inefficiency points. F to be efficient would need to proceed to A, G to C and I to D. Points that are unaffected by restrictions in λ_n's are in green area. Point H and others in non-green areas are different. If λ_n's are not restricted by 1 from above, H will tend to C, while if they are restricted, H will move either in horizontal or vertical direction.

3 Statistical inference

The estimates in (4) and (6) are point estimates that are calculated relative to the finite sample DEA *estimate* of the true frontier and they ignore the issue of statistical significance. The efficiency scores estimated using a finite sample are subject to sampling variation of the estimated frontier. In what follows,

we employ a *smoothed* bootstrap procedure to provide statistical inference on point estimates of the nonradial efficiency measure. The following exposition describes input-based nonradial efficiency measurement. The procedure for providing statistical inference about output-based measures can be easily extended.

We first apply the estimator introduced in (6) to the original observed sample, $\{(\boldsymbol{y}_k, \boldsymbol{x}_k)\}_{k=1}^K$ under the technology defined in (1) to obtain estimates of the efficiency scores $\widehat{\boldsymbol{RM}} = \left\{\widehat{RM}_k\right\}_{k=1}^K$. In contrast to radial efficiency measurement, where $\lambda_{kn} = \lambda_k, \forall n = 1, \ldots, N$, the Russell measure does not restrict $\hat{\lambda}_{kn}$'s; nevertheless the non-proportional reductions of inputs are likely to be correlated to varying degree. If they are perfectly correlated, we are back to the radial measure. But if they are not, this needs to be taken into account. We make use of the bootstrap methods based on homogeneity of the efficiency distribution assumption and adjust the existing algorithms to preserve the degree of correlation between λ_{kn}'s (Badunenko et al., 2014; Simar and Wilson, 1999). Since $\widehat{RM}_k$ depends on $\left(\hat{\lambda}_{k1}, \ldots, \hat{\lambda}_{kN}\right)$, the data generating process of $\widehat{\boldsymbol{RM}}$ necessarily depends on $\hat{\boldsymbol{\Lambda}} = \left[\hat{\boldsymbol{\lambda}}_1 \ldots \hat{\boldsymbol{\lambda}}_N\right]$, a $(K \times N)$ matrix, where $\hat{\boldsymbol{\lambda}}_j = (\hat{\lambda}_{j1}, \ldots, \hat{\lambda}_{jK})'$. Providing statistical inference on $\widehat{\boldsymbol{RM}}$ involves generating bootstrap samples $\{(\boldsymbol{y}_k^*, \boldsymbol{x}_k^*)\}_{k=1}^K$, which takes correlation between $\hat{\lambda}_{kn}$'s into account. Ignoring this correlation may result in violation of the data generating process of $\widehat{\boldsymbol{RM}}$. The bootstrap procedure we propose assumes that the density of $\boldsymbol{\Lambda}|(\boldsymbol{y}, \boldsymbol{x})$ is homogeneous (similar to Badunenko et al., 2014; Simar and Wilson, 1999). Simar and Wilson (1998) have shown that the naïve bootstrap, i.e., sampling from the empirical distribution of the data $\{(\boldsymbol{y}_k, \boldsymbol{x}_k)\}_{k=1}^K$, or equivalently from the efficiency scores, $\hat{\boldsymbol{\Lambda}}$, will yield inconsistent results because the efficiency score is truncated at one. Recall that λ_n's are all bounded at 1 from above in (6). Using the reflection method and the smoothed bootstrap results in consistent estimation of confidence intervals (see Silverman, 1986, p. 29–32). We discuss both the reflection technique and the smoothing approach in the multivariate case in greater detail below.

3.1 Multivariate kernel density estimation of efficiency measures

Let $\hat{\mathbf{\Lambda}}_i$ be the ith row of the $(K \times N)$ matrix $\hat{\mathbf{\Lambda}}$, the columns of which being $\hat{\boldsymbol{\lambda}}_1, \ldots, \hat{\boldsymbol{\lambda}}_N$. Given sample realizations, $\hat{\mathbf{\Lambda}}$, from a population with unknown multivariate density f, a multivariate nonparametric estimate of the joint density that accounts for the possibility of the correlation across input-specific efficiencies is given by

$$\hat{f}\left(\hat{\boldsymbol{\lambda}}, h\right) = \frac{1}{Kh^N} \sum_{k=1}^{K} \mathcal{K}_h\left(\hat{\mathbf{\Lambda}}_i, \hat{\boldsymbol{\lambda}}\right), \tag{8}$$

where $\hat{\boldsymbol{\lambda}} = (\hat{\lambda}_1, \ldots, \hat{\lambda}_N)$ is of dimension $(1 \times N)$, $\mathcal{K}_h(\cdot, \cdot)$ is the multivariate kernel function and bandwidth h is the smoothing parameter (e.g., see Henderson and Parmeter, 2014). As a kernel function $\mathcal{K}_h(\cdot, \cdot)$, one might use different choices for the multivariate (i) non-negative, (ii) radially symmetric, (iii) unimodal probability density function that integrates to one. We use the standard multivariate normal density function that is scaled to have the same covariance matrix as the data, viz.,

$$\mathcal{K}_h\left(\boldsymbol{U}_i, \boldsymbol{u}\right) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{\Sigma}|}} \exp\Big(-\frac{1}{2h^N}\left(\left(\boldsymbol{U}_i - \boldsymbol{u}\right)' \mathbf{\Sigma}^{-1} \left(\boldsymbol{U}_i - \boldsymbol{u}\right)\right)\Big),$$

where $\boldsymbol{u}$ is of dimension $(1 \times N)$, $\boldsymbol{U}$ is a $(K \times N)$ matrix, $\mathbf{\Sigma}$ is symmetric estimated covariance matrix of $\boldsymbol{U}$, and $|\mathbf{\Sigma}|$ is the determinant of $\mathbf{\Sigma}$. We use a single smoothing parameter $h = K^{-1/(N+4)}$ as a bandwidth for the multivariate data since we scale the kernel function to have the same covariance matrix as the estimated covariance matrix of the original data (see Silverman, 1986, p. 86–87).

Since the DEA efficiency scores are truncated at one, the density estimate using (8) will be inconsistent and asymptotically biased. In the univariate case (e.g. where we have $\hat{\boldsymbol{\lambda}}_1$ only), this truncation can be overcome by

reflecting the truncated data observations $\left\{\hat{\lambda}_{1k}\right\}_{k=1}^{K}$ in the boundary (i.e. about one), $\left\{2-\hat{\lambda}_{1k}\right\}_{k=1}^{K}$, and estimating the density using the resulting set of $2K$ observations, $\left\{\hat{\lambda}_{1k}\ 2-\hat{\lambda}_{1k}\right\}_{k=1}^{K}$ (see Silverman, 1986, p. 29–32). However, we have a multivariate case and hence there are N truncations in each of $\hat{\boldsymbol{\lambda}}_1,\ldots,\hat{\boldsymbol{\lambda}}_N$. This yields $Q=2^N$ different combinations of the reflections. Each block q of dimension $(K\times N)$ of $(QK\times N)$ matrix $\hat{\boldsymbol{\Lambda}}^R$ contains unique combination of original and reflected $\hat{\boldsymbol{\lambda}}$'s. For example, in the case $N=2$, the matrix with all possible combinations will consist of $2^2=4$ blocks, $q=1$: $[\hat{\boldsymbol{\lambda}}_1\ \hat{\boldsymbol{\lambda}}_2]$, $q=2$: $[2-\hat{\boldsymbol{\lambda}}_1\ \hat{\boldsymbol{\lambda}}_2]$, $q=3$: $[\hat{\boldsymbol{\lambda}}_1\ 2-\hat{\boldsymbol{\lambda}}_2]$, and $q=4$: $[2-\hat{\boldsymbol{\lambda}}_1\ 2-\hat{\boldsymbol{\lambda}}_2]$,

$$\hat{\boldsymbol{\Lambda}}^R=\begin{bmatrix}\hat{\boldsymbol{\lambda}}_1 & \hat{\boldsymbol{\lambda}}_2 \\ 2-\hat{\boldsymbol{\lambda}}_1 & \hat{\boldsymbol{\lambda}}_2 \\ \hat{\boldsymbol{\lambda}}_1 & 2-\hat{\boldsymbol{\lambda}}_2 \\ 2-\hat{\boldsymbol{\lambda}}_1 & 2-\hat{\boldsymbol{\lambda}}_2\end{bmatrix}\begin{matrix}(q=1)\\(q=2)\\(q=3)\\(q=4)\end{matrix}.$$

For $N=3$ $(Q=2^3=8)$, the matrix with reflected combinations is

$$\hat{\boldsymbol{\Lambda}}^R=\begin{bmatrix}\hat{\boldsymbol{\lambda}}_1 & \hat{\boldsymbol{\lambda}}_2 & \hat{\boldsymbol{\lambda}}_3 \\ 2-\hat{\boldsymbol{\lambda}}_1 & \hat{\boldsymbol{\lambda}}_2 & \hat{\boldsymbol{\lambda}}_3 \\ \hat{\boldsymbol{\lambda}}_1 & 2-\hat{\boldsymbol{\lambda}}_2 & \hat{\boldsymbol{\lambda}}_3 \\ 2-\hat{\boldsymbol{\lambda}}_1 & 2-\hat{\boldsymbol{\lambda}}_2 & \hat{\boldsymbol{\lambda}}_3 \\ \hat{\boldsymbol{\lambda}}_1 & \hat{\boldsymbol{\lambda}}_2 & 2-\hat{\boldsymbol{\lambda}}_3 \\ 2-\hat{\boldsymbol{\lambda}}_1 & \hat{\boldsymbol{\lambda}}_2 & 2-\hat{\boldsymbol{\lambda}}_3 \\ \hat{\boldsymbol{\lambda}}_1 & 2-\hat{\boldsymbol{\lambda}}_2 & 2-\hat{\boldsymbol{\lambda}}_3 \\ 2-\hat{\boldsymbol{\lambda}}_1 & 2-\hat{\boldsymbol{\lambda}}_2 & 2-\hat{\boldsymbol{\lambda}}_3\end{bmatrix}\begin{matrix}(q=1)\\(q=2)\\(q=3)\\(q=4)\\(q=5)\\(q=6)\\(q=7)\\(q=8)\end{matrix}.$$

The number of blocks $Q=2^N$ in $\hat{\boldsymbol{\Lambda}}^R$ will thus depend on the number of inputs N. For each block we let $\widehat{\boldsymbol{\Sigma}}_q$ denote the estimated covariance matrix of the columns of block q in $\hat{\boldsymbol{\Lambda}}^R$.

Let $\hat{\mathbf{\Lambda}}_j^R$ be the j^{th} row of the $(QK \times N)$ matrix $\hat{\mathbf{\Lambda}}^R$. The multivariate nonparametric estimate of the joint density of the QK reflected observations of $\hat{\mathbf{\Lambda}}^R$ is given by

$$\hat{g}(\hat{\boldsymbol{\lambda}}, h) = \frac{1}{QKh^N} \left(\sum_{q=1}^{Q} \sum_{j=(q-1)K+1}^{qK} \mathcal{K}_h \left(\hat{\mathbf{\Lambda}}_j^R, \hat{\boldsymbol{\lambda}} \right) \right). \tag{9}$$

Note that the estimate in (9) is the additive kernel estimate, which consists of Q terms. The q^{th} term is the multivariate normal density function with the estimated covariance matrix $\widehat{\mathbf{\Sigma}}_q$, $q = 1, \ldots, Q$. The consistent estimate of the density of the original data $\hat{\mathbf{\Lambda}}$ is given by (see Silverman, 1986, p. 30)

$$\hat{g}^*(\hat{\boldsymbol{\lambda}}, h) = \begin{cases} Q\widehat{g}(\hat{\boldsymbol{\lambda}}, h), & \text{for } \lambda_i \leq 1,\ i = 1, ...N, \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

3.2 Smoothed bootstrap

We do not actually have to estimate the density in (10). The following *smoothed* bootstrap procedure can be applied to simulate from the density estimates (see Silverman, 1986, p. 142–144). Draw $\mathbf{\Lambda}^A = (\lambda_{k1}^A, \ldots, \lambda_{kN}^A)$, $k = 1, \ldots, K$, randomly with replacement from $\hat{\mathbf{\Lambda}}^R$ where each row is drawn with equal probability, $\mathbf{\Lambda}^A$ is a $(K \times N)$ matrix and $\hat{\mathbf{\Lambda}}^R$ is a $(QK \times N)$ matrix. Then compute the $(K \times N)$ matrix

$$\mathbf{\Lambda}^B = \overline{\mathbf{\Lambda}}^A + \left(\mathbf{\Lambda}^A - \overline{\mathbf{\Lambda}}^A + h\boldsymbol{\epsilon}^* \right) / \left(1 + h^2 \right)^{1/2},$$

where $\overline{\mathbf{\Lambda}}^A$ is a $(K \times N)$ matrix in which the elements of the n^{th} column are the mean of the n^{th} column of $\mathbf{\Lambda}^A$ and $\boldsymbol{\epsilon}^*$ is the $(K \times N)$ matrix of deviates from the multivariate normal distribution with the estimated covariance matrix $\widehat{\mathbf{\Sigma}}_q$ for the rows of $\mathbf{\Lambda}^B$ that were drawn from q^{th} block of $\hat{\mathbf{\Lambda}}^R$ matrix. Matrix $\mathbf{\Lambda}^* = [\lambda_{k1}^* \ \ldots \ \lambda_{kN}^*]$ $(k = 1, \ldots, K)$ contains simulated efficiency scores that

are set as follows:

$$\lambda_{kj}^* = \begin{cases} \lambda_{kj}^B, & \text{if } \lambda_{kj}^B \leq 1 \\ 2 - \lambda_{kj}^B, & \text{otherwise} \end{cases},$$

where λ_{kj}^B are the elements of $\boldsymbol{\Lambda}^B$ for $k = 1, \ldots, K,\ j = 1, \ldots, N$. The bootstrap sample $(\boldsymbol{y}^*, \boldsymbol{x}^*)$, where $\boldsymbol{x}^* = \{\boldsymbol{x}_k^*\}_{k=1}^K$, $\boldsymbol{y}^* = \{\boldsymbol{y}_k\}_{k=1}^K$, and $\boldsymbol{x}_k^* = (x_{k1}^*, \ldots, x_{kN}^*)$, can then be constructed as

$$x_{kn}^* = x_{kn} \frac{\hat{\lambda}_{kn}}{\lambda_{kn}^*},\ n = 1, \ldots, N,\ k = 1, \ldots, K$$

For the given bootstrap sample construct the technologies:

$$\begin{aligned} \hat{\mathcal{T}}^{\text{CRS},*} = \big\{(\boldsymbol{y}^*, \boldsymbol{x}^*) :\ & \sum_{k=1}^K z_k^* y_{km}^* \geq y_{km}^*,\ m = 1, \ldots, M, \\ & \sum_{k=1}^K z_k^* x_{kn}^* \leq x_{kn}^*,\ n = 1, \ldots, N, \\ & z_k^* \geq 0,\ k = 1, \ldots, K\big\}, \end{aligned} \tag{11}$$

Compute the Russell measure that appear in (6) using the bootstrap technology defined in (11). Adjust the returns to scale as necessary. Apply the procedure above B times (B should be large in practice) to obtain B bootstrap Russell measures, $\widehat{\boldsymbol{RM}}^*$.

3.3 Bias correction and confidence intervals

By definition bias is given by,

$$\text{bias}\, \widehat{RM}(\boldsymbol{y}, \boldsymbol{x}) \equiv E\left[\widehat{RM}(\boldsymbol{y}, \boldsymbol{x})\right] - RM(\boldsymbol{y}, \boldsymbol{x}).$$

Empirically, we approximate it by a bootstrap bias estimate, viz.,

$$\widehat{\text{bias}}\,\widehat{RM}(\boldsymbol{y},\boldsymbol{x}) = \frac{1}{B}\sum_{b=1}^{B}\widehat{RM}_b^* - \widehat{RM}(\boldsymbol{y},\boldsymbol{x}), \tag{12}$$

where $\widehat{RM}_b^*$ is the bootstrap Russell measure calculated relative to $\mathcal{T}_b^*$, $b = 1,\ldots,B$. We can now compute bias-corrected estimator of Russell measure $\widetilde{RM}$ by subtracting bootstrap bias estimate from the original Russell measure in (6) (see Efron and Tibshirani, 1993). Efron and Tibshirani (1993) however caution that bias correction introduces additional noise thus increasing mean square error of the bias corrected estimator, and therefore suggest performing bias correction only if $\widehat{\text{bias}}\,\widehat{RM}(\boldsymbol{y},\boldsymbol{x})/\hat{\sigma}_{\widehat{\text{bias}}} > 0.25$, where $\hat{\sigma}_{\widehat{\text{bias}}}$ is the standard error of the bootstrap Russell measures.

We can also use bootstrap Russell measures to construct confidence intervals. For an unknown distribution of $\left(\widehat{RM}_k - RM_k\right)$, construction of the $(1-\alpha)$–percent confidence interval boils down to finding values $l_{k\alpha}$ and $u_{k\alpha}$ in

$$\text{Prob}\left(-u_{k\alpha} \leq \widehat{RM}_k - RM_k \leq -l_{k\alpha}\right) = 1-\alpha. \tag{13}$$

Using B bootstrap values $\widehat{RM}_i^*$ we can find $l_{k\alpha}^*$ and $u_{k\alpha}^*$ such that the probability that the equation

$$\text{Prob}\left(-u_{k\alpha}^* \leq \widehat{RM}_k^* - \widehat{RM}_k \leq -l_{k\alpha}^*|S\right) = 1-\alpha$$

holds approaches one as $B \to \infty$. Since the bootstrap involves approximating the unknown distribution of $\left(\widehat{RM}_k - RM_k\right)$ by that of $\left(\widehat{RM}_k^* - \widehat{RM}_k\right)$ conditional on original sample S, we can substitute $l_{k\alpha}$ and $u_{k\alpha}$ in (13) with $l_{k\alpha}^*$ and $u_{k\alpha}^*$ to obtain the bootstrap approximation based on the original sample as follows

$$\text{Prob}\left(-u_{k\alpha}^* \leq \widehat{RM}_k - RM_k \leq -l_{k\alpha}^*|S\right) \approx 1-\alpha.$$

This relationship allows us to construct the approximated $(1-\alpha)$–percent confidence interval as

$$\widehat{RM}_k + l^*_{k\alpha} \leq RM_k \leq \widehat{RM}_k + u^*_{k\alpha}. \tag{14}$$

4 Monte Carlo evidence

We examine the finite sample properties of the proposed bootstrap based estimator of the input-based Russell measure. Specifically, we first discuss the design of the experiment. We define the data generating processes accounting for potential correlation, and the distributional assumptions on the data used in the experiment. Further, we discuss the methods to compare the performance of the proposed efficiency estimator. We take several standard measures as well as one that we propose just for this experiment. Third, we run our simulations and discuss the performance of the proposed estimator under various scenarios.

4.1 Design of the experiment

We conduct Monte Carlo simulations for a production process which employs two inputs (x_1 and x_2) to produce a single output (y). We consider the following CES production frontier, viz.,

$$y = \left(\alpha \tilde{x}_1^{\rho} + \beta \tilde{x}_2^{\rho}\right)^{1/\rho}, \tag{15}$$

where $\tilde{x}_1$ and $\tilde{x}_2$ are individually efficient levels of inputs, more specifically, $\tilde{x}_n = x_n \lambda_n$, $n = 1, 2$. Then, x_1 and x_2 are the 'observed' inefficient levels of outputs. For each Monte Carlo trial, the efficiencies are generated in the following way. First, we draw the efficient levels of inputs from a doubly

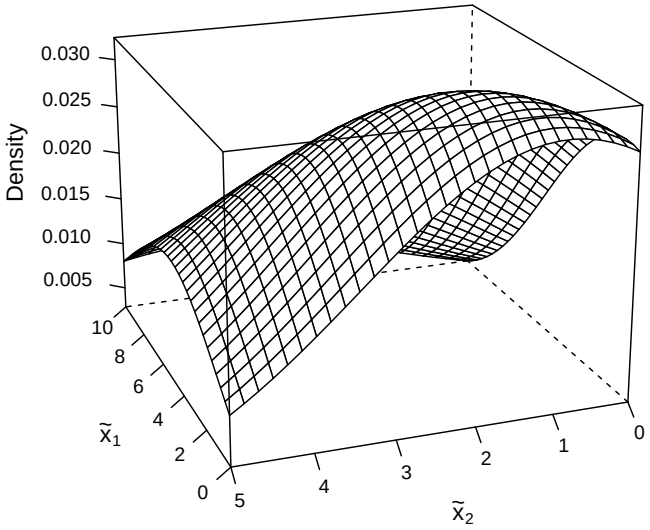

Figure 2 Density of $\tilde{x}_1$ and $\tilde{x}_2$

truncated normal distribution, viz.,

$$\begin{bmatrix}\tilde{x}_1\\ \tilde{x}_2\end{bmatrix} \sim \mathcal{N}_{\left((0,0)',(10,5)'\right)}\left(\begin{bmatrix}2\\ 1.5\end{bmatrix}, \begin{bmatrix}\sigma^2_{\tilde{x}_1} & 0.5\sqrt{\sigma^2_{\tilde{x}_1}\sigma^2_{\tilde{x}_1}/3}\\ 0.5\sqrt{\sigma^2_{\tilde{x}_1}\sigma^2_{\tilde{x}_1}/3} & \sigma^2_{\tilde{x}_1}/3\end{bmatrix}\right), \quad (16)$$

where $\mathcal{N}_{\left((0,0)',(10,5)'\right)}$ denotes bivariate normal distribution truncated at $(0,0)'$ from below and $(10,5)'$ from above. Drawing as in (16) allows inputs to be moderately correlated. We set $\sigma^2_{\tilde{x}_1} = 25$. The density of $\tilde{x}_1$ and $\tilde{x}_2$ is shown in Figure 2. Such density reflects real world situation where distribution of inputs is skewed with most decision making units having small inputs and only some – very large.

$\boldsymbol{\lambda}_0 = (\lambda_{10}, \lambda_{20})'$ are drawn from a bivariate log-normal distribution truncated to the unit square $(\lambda_{10}, \lambda_{20})' \in [0,1]^2$

$$\begin{bmatrix}\log\lambda_{10}\\ \log\lambda_{20}\end{bmatrix} = \begin{bmatrix}u\\ v\end{bmatrix} \sim \mathcal{N}_{(-\infty,0]^2}\left(\begin{bmatrix}0\\ 0\end{bmatrix}, \begin{bmatrix}\sigma^2_u & \sigma_{uv}\\ \sigma_{uv} & \sigma^2_v\end{bmatrix}\right),$$

where $\mathcal{N}_{(-\infty,0]^2}$ stands for the normal distribution truncated in its positive tail at $(0,0)'$. Variances σ_u^2 and σ_v^2 are set in advance while σ_{uv} is chosen to give correlation coefficient ρ_{12} of $(\lambda_{10},\lambda_{20})'$. Moments of order (s_u,s_v) of $(\lambda_{10},\lambda_{20})'$ can be obtained (by simple manipulation of formulas from Lien (1985)) as

$$\begin{aligned}\mathbb{E}\left[\lambda_{10}^{s_u},\lambda_{20}^{s_v}\,|\,(\lambda_{10},\lambda_{20})'\in[0,1]^2\right] &= \mathbb{E}\left[e^{s_u u+s_v v}\,|\,(u,v)'\in(-\infty\ 0]^2\right] \qquad (17)\\ &= e^{-\frac{D}{2Q}}\times\frac{\Phi\left(\begin{bmatrix}0\\0\end{bmatrix}\,\middle|\,\begin{bmatrix}h\\k\end{bmatrix},\begin{bmatrix}\sigma_u^2 & \sigma_{uv}\\ \sigma_{uv} & \sigma_v^2\end{bmatrix}\right)}{\Phi\left(\begin{bmatrix}0\\0\end{bmatrix}\,\middle|\,\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\sigma_u^2 & \sigma_{uv}\\ \sigma_{uv} & \sigma_v^2\end{bmatrix}\right)},\end{aligned}$$

where

$$\begin{aligned}h &= s_u\sigma_u^2+s_2\sigma_v^2\,,\\ k &= s_u\sigma_u^2+s_2\sigma_v^2\,,\\ Q &= \sigma_u^2\sigma_v^2-\sigma_{uv}^2\,,\\ D &= -Q(s_u^2\sigma_u^2+2s_us_v\sigma_{uv}+s_v^2\sigma_v^2)\,.\end{aligned}$$

Equation (17) is used to find σ_{uv}, which sets the correlation between λ_{10} and λ_{20} to a desired level ρ_{12}. Then, the 'observed' inefficient inputs x_1 and x_2 are computed by dividing $\tilde{x}_n$ by λ_{n0}, $n=1,2$. $\tilde{x}_1$ and $\tilde{x}_2$ are in effect pushed inside the frontier defined in (15).

To analytically calculate the true efficiencies, we minimize the average of λ_1 and λ_2 subject to constraint that $\tilde{x}_n = x_n\lambda_n$, $n=1,2$ are on the frontier. Thus, the true efficiencies are computed as optimal values in the following optimization problem:

$$\min_{\lambda_1,\lambda_2}\frac{\lambda_1+\lambda_2}{2}\quad\text{s.t.}\quad y=\left[\alpha\,(x_1\lambda_1)^\rho+\beta\,(x_2\lambda_2)^\rho\right]^{1/\rho}. \qquad (18)$$

Denote

$$
\begin{aligned}
a_1 &= 2\alpha\rho x_1^{\rho}\,,\\
b_1 &= 2\beta\rho x_2^{\rho}\,,\\
s_1 &= \alpha x_1^{\rho} a_1^{(-\rho/(\rho-1))} + \beta x_2^{\rho} b_1^{(-\rho/(\rho-1))}\,,\\
\lambda_3 &= (y^{\rho}/s_1)^{(-(\rho-1)/\rho)}\,,\\
\lambda_{1_1} &= (1/\lambda_3/a_1)^{(1/(\rho-1))}\,,\\
\lambda_{2_1} &= (1/\lambda_3/b_1)^{(1/(\rho-1))}\,.
\end{aligned}
$$

Note that λ_{1_1} and λ_{2_1} are not restricted separately. To ensure that the optimal λ_1 and λ_2 are smaller than 1 in the spirit of the Russell measure, the components of the true efficiency are computed as follows:

$$
\lambda_1 = \begin{cases} 1, & \lambda_{1_1} > 1 \\ \lambda_{1_2}, & \text{otherwise} \end{cases}\,, \tag{19}
$$

where

$$
\lambda_{1_2} = \begin{cases} ((y^{\rho} - \beta x_2^{\rho})/\alpha/x_1^{\rho})^{(1/\rho)}, & \lambda_{2_1} > 1 \\ \lambda_{1_1}, & \text{otherwise} \end{cases}
$$

and

$$
\lambda_2 = \begin{cases} 1, & \lambda_{2_1} > 1 \\ \lambda_{2_2}, & \text{otherwise} \end{cases}\,, \tag{20}
$$

where

$$
\lambda_{2_2} = \begin{cases} ((y^{\rho} - \alpha x_1^{\rho})/\beta/x_2^{\rho})^{(1/\rho)}, & \lambda_{1_1} > 1 \\ \lambda_{2_1}, & \text{otherwise} \end{cases}\,.
$$

The true efficiency is then computed as an average of λ_1 and λ_2.

All experiments consist of 999 Monte Carlo trials. In each trial we set $B = 1999$. Within each set of experiments, we analyze seven sample sizes, $K = 30, 50, 100, 200, 500,$ 1000 and 2000. Since we are interested in a nonradial measure, we would like to analyze the performance of the estimator when the true frontier has a different degree of symmetry. We set parameters $\rho = 0.5$, $\alpha = 0.5$ and $\beta = 0.5^c$ in (15), where $c = 1, 2, 3$. Figure 3 shows various degree of symmetry of the true frontiers with these parameters.

We consider $\sigma_v^2 = 0.02$ and $\sigma_u^2 = 10^{\varphi}$, where $\varphi = -2, -1, 0$, and select σ_{uv} using (17) to ensure correlation between λ_{10} and λ_{20} is equal to $\rho_{12} = 0.5$.[1] Due to the shape of the frontier, the correlation between λ_1 and λ_2 is different, but we account for this dependence by drawing from the multivariate normal density in the smoothed bootstrap procedure. To give a better feel about λ_1 and λ_2 in (19) and (20), Figure 4 shows scatter plots of λ_1 and λ_2 for chosen values of c and φ. Each of 9 panels in Figure 4 shows a single draw when $K = 3000$. The larger is the φ, the more inefficient observations are allowed to be in the sample. In the case of $\varphi = -2$, especially in asymmetric case, there is very little inefficiency on average. One observation that stands out is that as the true frontier becomes more asymmetric, the individual efficiencies become more correlated, especially when variation in inputs becomes larger. For a given value of c, ρ and φ we compute the 'observed' output observations as in (15).

[1]Though we do not directly set the correlation between λ_1 and λ_2, our experiments with different constellations gave us a sense that setting the correlation between λ_{10} and λ_{20} directly influenced the correlation between λ_1 and λ_2.

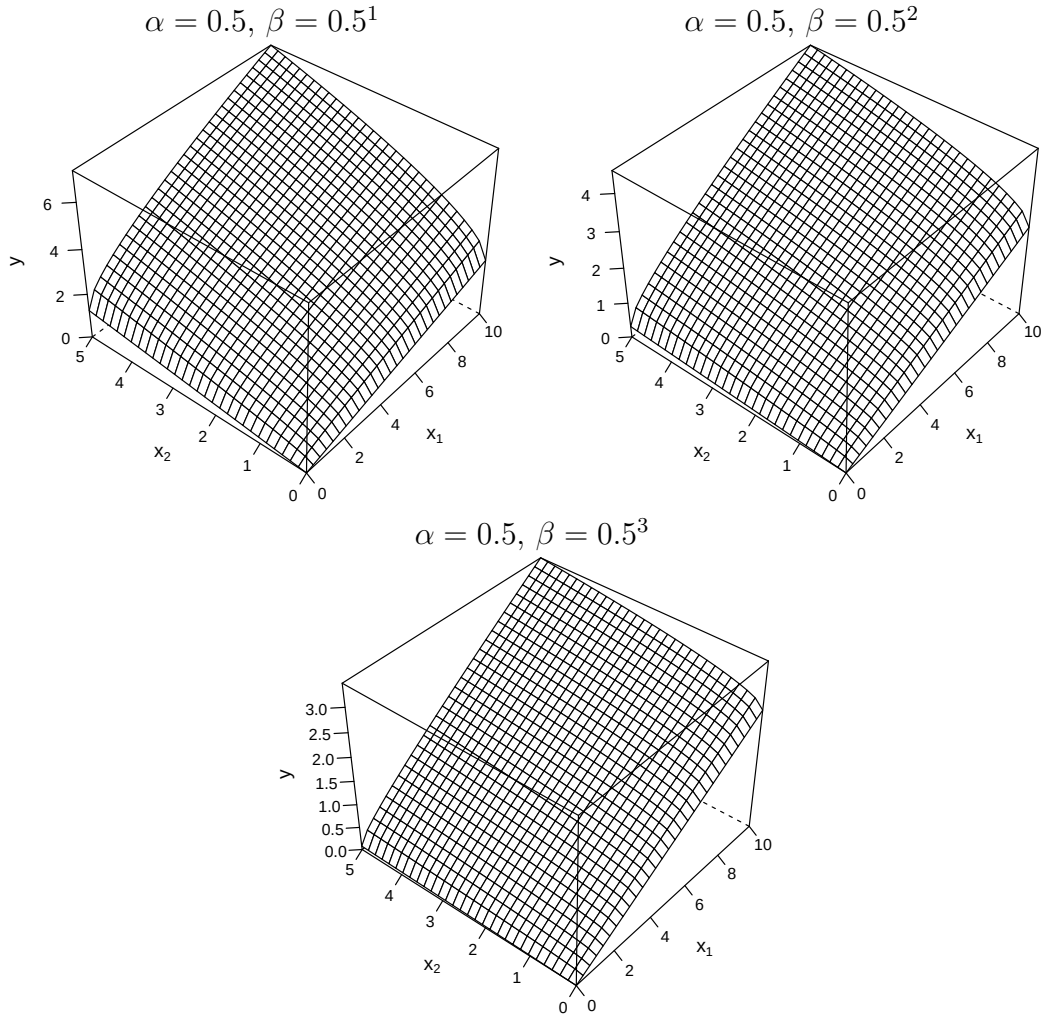

Figure 3 Shape of the true frontier in (15) with $\rho = 0.5$, $\alpha = 0.5$, and $\beta = 0.5^c$, where $c = 1, 2, 3$

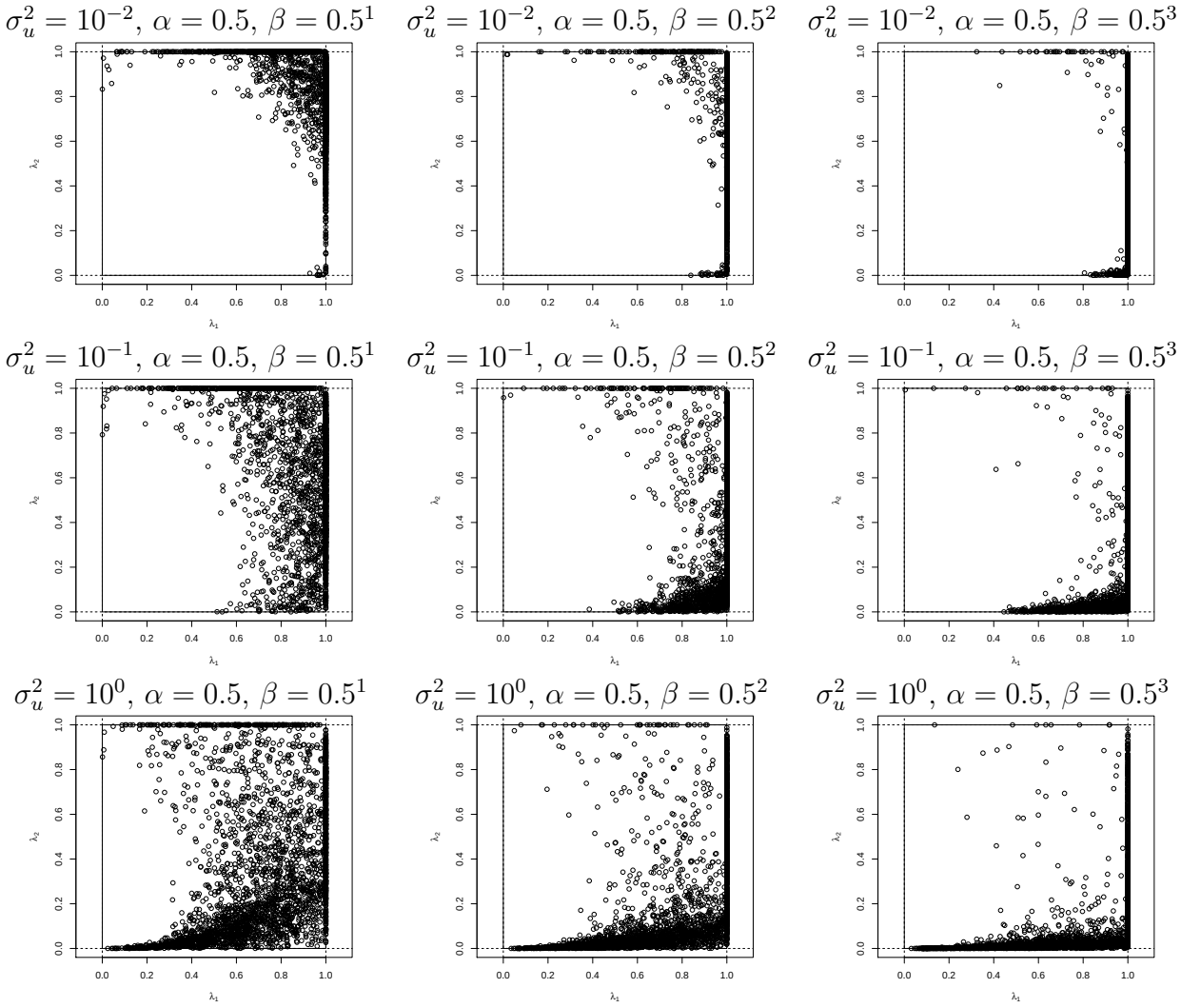

Figure 4 Scatter plot of λ_1 and λ_2 with $c = 1, 2, 3$ and $\varphi = -2, -1, 0$

To compare the finite sample performance of our estimators we consider the following mean (over 999 Monte Carlo trials) measures

- $\text{Bias}(RM) = \frac{1}{K}\sum_{k=1}^{K}\left(\widetilde{RM}_k - RM_k\right)$,
- $\text{MSE}(RM) = \frac{1}{K}\sum_{k=1}^{K}\left(\widetilde{RM}_k - RM_k\right)^2$,
- $\text{Upward Bias}\,(RM) = \frac{1}{K}\sum_{k=1}^{n}\mathbf{1}\left(\widetilde{RM}_k > RM_k\right)$
- $\text{Kendall's}\ \tau\,(RM) = \frac{K_c - K_d}{0.5K\,(K-1)}$,

where K_c is the number of concordant pairs, and K_d is the number of discordant pairs in the data set (efficiency ranks of $\widetilde{RM}$) and $\mathbf{1}\,(A)$ denotes the indicator function that the event A is true. Specific for this study is the upward bias, i.e., the share of predicted Russell measures strictly larger than the true Russell measures. The desired value of upward bias is 0.5. The values less and greater than 0.5 indicate systematic underestimation and overestimation respectively of Russell measures.

4.2 Simulation results

Table 1 shows the performance measures of the technical efficiency estimates for various parameters of the experiment. The table shows the 'average' performance of the estimator for observations that have different levels of true efficiency. We split our discussion of the results into two cases, determined by amount of true inefficiency (row-wise movement in the Figure 4) and shape of the true technology (column-wise movement in the Figure 4).

Recall that the lower is the σ_u, the smaller is the inefficiency on average. The first three panels of Table 1 show the results for $\sigma_u^2 = 10^{-2}$, i.e., the cases that are shown in the first row of Figure 4. Irrespective of the shape of the true frontier, mean squared error goes down in number of observations albeit

Table 1 Finite sample performance of the estimates under different technologies and true efficiencies

K	Bias[a]	MSE[b]	Upward bias[c]	Correlation[d]
$\sigma_u^2 = 10^{-2}$, $\alpha = 0.5$, $\beta = 0.5^1$				
30	−0.0438	0.0068	0.1600	0.4585
50	−0.0271	0.0049	0.1859	0.5417
100	−0.0134	0.0035	0.2103	0.6460
200	−0.0050	0.0026	0.2372	0.7289
500	0.0009	0.0018	0.2883	0.8086
1000	0.0023	0.0039	0.3222	0.8510
2000	0.0027	0.0010	0.3408	0.8844
$\sigma_u^2 = 10^{-2}$, $\alpha = 0.5$, $\beta = 0.5^2$				
30	−0.0489	0.0483	0.2240	0.4913
50	−0.0408	0.0141	0.2167	0.5600
100	−0.0369	0.0311	0.1869	0.6420
200	−0.0306	0.0077	0.1645	0.7105
500	−0.0188	0.0033	0.1484	0.7898
1000	−0.0094	0.0021	0.1480	0.8394
2000	−0.0030	0.0014	0.1696	0.8771
$\sigma_u^2 = 10^{-2}$, $\alpha = 0.5$, $\beta = 0.5^3$				
30	−0.0468	0.9935	0.3550	0.4930
50	−0.0420	0.0488	0.3162	0.5520
100	−0.0381	0.0151	0.2690	0.6228
200	−0.0385	0.0103	0.2166	0.6916
500	−0.0366	0.0068	0.1600	0.7702
1000	−0.0319	0.0049	0.1322	0.8156
2000	−0.0261	0.0034	0.1128	0.8522
$\sigma_u^2 = 10^{-1}$, $\alpha = 0.5$, $\beta = 0.5^1$				
30	0.0003	0.0403	0.4277	0.4927
50	0.0098	0.0098	0.4530	0.5710
100	0.0168	0.0894	0.4966	0.6607
200	0.0196	0.0193	0.5247	0.7314
500	0.0187	0.0041	0.5385	0.8001
1000	0.0169	0.0033	0.5444	0.8366
2000	0.0144	0.0025	0.5345	0.8667

(*continued on next page*)

Table 1 (*Continued*)

K	Bias[a]	MSE[b]	Upward bias[c]	Correlation[d]
$\sigma_u^2 = 10^{-1}$, $\alpha = 0.5$, $\beta = 0.5^2$				
30	0.0184	0.0583	0.5400	0.4594
50	0.0296	0.0335	0.5489	0.5338
100	0.0334	0.0341	0.5526	0.6164
200	0.0309	0.0097	0.5397	0.6897
500	0.0248	0.0067	0.5212	0.7658
1000	0.0210	0.0051	0.5126	0.8107
2000	0.0178	0.0039	0.5050	0.8435
$\sigma_u^2 = 10^{-1}$, $\alpha = 0.5$, $\beta = 0.5^3$				
30	0.0679	0.0687	0.6878	0.4112
50	0.0644	0.0964	0.6830	0.4670
100	0.0511	0.0174	0.6613	0.5561
200	0.0391	0.0129	0.6392	0.6417
500	0.0271	0.0085	0.6087	0.7357
1000	0.0214	0.0063	0.5902	0.7858
2000	0.0174	0.0048	0.5731	0.8239
$\sigma_u^2 = 10^{0}$, $\alpha = 0.5$, $\beta = 0.5^1$				
30	0.0777	0.1127	0.7010	0.4955
50	0.0833	0.0524	0.7226	0.5767
100	0.0781	0.0246	0.7262	0.6529
200	0.0682	0.0187	0.7230	0.7108
500	0.0567	0.0132	0.7139	0.7675
1000	0.0480	0.0102	0.6954	0.8015
2000	0.0421	0.0082	0.6827	0.8262
$\sigma_u^2 = 10^{0}$, $\alpha = 0.5$, $\beta = 0.5^2$				
30	0.1276	0.0568	0.7946	0.4921
50	0.1194	0.0422	0.8009	0.5510
100	0.0946	0.0301	0.7848	0.6247
200	0.0787	0.0230	0.7746	0.6801
500	0.0581	0.0153	0.7455	0.7452
1000	0.0483	0.0116	0.7363	0.7816
2000	0.0409	0.0091	0.7241	0.8092
$\sigma_u^2 = 10^{0}$, $\alpha = 0.5$, $\beta = 0.5^3$				

(*continued on next page*)

Table 1 (*Continued*)

K	Bias[a]	MSE[b]	Upward bias[c]	Correlation[d]
30	0.1528	1.8291	0.8667	0.4971
50	0.1399	0.0475	0.8657	0.5410
100	0.1107	0.0341	0.8572	0.6044
200	0.0823	0.0233	0.8460	0.6741
500	0.0577	0.0150	0.8326	0.7453
1000	0.0449	0.0110	0.8203	0.7875
2000	0.0350	0.0081	0.8108	0.8205

[a]The mean deviation between the estimated and true technical efficiency. Reported in this table is the mean of such deviations across all Monte Carlo simulations.
[b]The the squared mean deviation between the estimated and true technical efficiency. Reported in this table is the mean of such deviations across all Monte Carlo simulations.
[c]Upward bias is the share of predicted technical efficiencies strictly larger than the true efficiencies. The desired value of upward bias is 0.5. The values less and greater than 0.5 indicate systematic underestimation and overestimation respectively of technical efficiencies. Reported in this table is the mean of such shares across all Monte Carlo simulations.
[d]Kendall correlation coefficient between predicted and true technical efficiencies. Reported in this table is the mean of such coefficients across all Monte Carlo simulations.

at a relatively slow pace. The bias measure is not very big even for $K = 30$, but as number of observations gets larger it gets still smaller quite quickly and already for $K = 200$ is at a very low level. What bias measure means is that if true efficiency is for example 0.95, we estimate it to be 0.92. When the amount of inefficiency is small, the estimator tends to underestimate. In all three cases, the upward bias is below the desired level of 0.5. As the technology becomes more asymmetric, the estimator underestimates in more cases on average when the number of observations increases. As number of observations increases the correlation between true and estimated efficiency is increasing reaching 0.8 for $K = 500$. This is important as the estimator gets better in estimating the rank of decision making units, which, depending on the purpose of the efficiency analysis maybe the only thing researcher or policy maker is interested in.

With moderate amount of inefficiency ($\sigma_u^2 = 10^{-1}$), the estimator tends to overestimate on average. Bias does go down as the sample size grows, but it does so slowly. It would be difficult to put much faith in estimator for a sample size smaller than 200. The tendency of MSE is very similar to the case where average inefficiency is rather small. The upward bias in this case slowly but surely approaches the desired value of 0.5. Since the bias is positive, on average, the size of overestimation is bigger then the size of underestimation.

Correlation column shows exactly the same pattern as before, which will also be true for the case where the amount of inefficiency is large. More observations translate, irrespective of the true average inefficiency or the shape of the true technology, into better prediction of ranks.

Finally, if the amount of true inefficiency is large ($\sigma_u^2 = 1$), the estimator overestimates on average even more than in previous case. As in previous cases, MSE is slowly decreasing. Upward bias is now larger than before suggesting that not only estimator overestimates, but that it overestimates by more than it underestimates.

If we look at the results in Table 1 from the 'shape of the true frontier' point of view (column-wise movement in the Figure 4), the big picture is that the finite sample properties of the proposed estimator depend only on the amount of the true inefficiency. Irrespective of the amount of true inefficiency or shape of the frontier, Table 1 suggests that both bias and mean squared error of our estimator decrease in sample size, thus providing evidence for consistency of the proposed estimator

The upshot of our experiment is twofold. First, as sample size gets bigger, the bias and mean squared error decrease. Second, our estimator gets better at predicting ranking of decision making units as sample size grows. This leaves little hope for practitioners, who use relatively small sample sizes in their efficiency analysis.

5 Empirical application: efficiency of US electric utilities

In this section, we illustrate the described procedure using data on fossil-fuel-fired steam electric power-generating plants in the USA over the period 1986–1999 (Kumbhakar and Tsionas, 2011). We consider the year 1998 (the year with the most observations), where the sample size n is 81 firms. Production technology consists of four variables: output (net steam electric power generation in megawatt-hours), labor, fuel and capital. Quantities of labour are calculated by dividing the aggregate costs of labor by a cost-share-weighted price for labor. The fuel quantities are calculated by dividing the fuel expenses by the Tornqvist price of fuel aggregate. The values of capital stocks are calculated by the valuation of base and peak load capacity at replacement cost to estimate capital stocks in a base year and then updating it in the subsequent years on the basis of the value of additions and retirements to steam power plants.

Table 2 presents the input-based Russell measures calculated assuming that technology satisfies variable returns to scale and bootstrap results for the first 20 observations in the sample ($B = 1999$). Positive bias indicates that all Russell measures are as expected upward biased. For several observations, Russell measures are overestimated by as much as 0.21. Bias over standard error of the bias is well above quarter for all except for the second observation, where it cannot be calculated. This means that bias corrected measures for all except for the second observation are preferred to the original Russell measures.

Table 2 Russell measure, bootstrap diagnostic, bias corrected measure, and confidence interval

Lower bound of the 99% confidence interval[a]	95%	90%	$\widehat{RM}$[b]	B[c]	$\widehat{Bias}$[d]	$\widehat{Bias}/\hat{\sigma}$[e]	$\widetilde{RM}$[f]	Upper bound of the 90% confidence interval	95%	99%
0.538	0.549	0.600	0.762	1784	0.106	0.838	0.656	0.762	0.762	0.762
1.000	1.000	1.000	1.000	436	0.000	–	1.000	1.000	1.000	1.000
0.686	0.693	0.700	0.875	1314	0.108	1.042	0.767	0.875	0.875	0.875
0.528	0.570	0.581	0.763	1980	0.112	0.761	0.651	0.763	0.763	0.763
0.549	0.564	0.569	0.806	1325	0.165	1.012	0.641	0.799	0.805	0.806
0.337	0.358	0.368	0.548	1935	0.068	0.604	0.479	0.548	0.548	0.548
0.423	0.492	0.524	0.599	1998	0.041	0.619	0.558	0.599	0.599	0.599
0.870	0.885	0.893	1.000	1999	0.039	0.633	0.961	1.000	1.000	1.000
0.462	0.474	0.480	0.536	1999	0.033	1.019	0.503	0.535	0.536	0.536
0.406	0.414	0.419	0.469	1999	0.028	1.049	0.441	0.466	0.469	0.469
0.987	0.987	0.987	1.000	803	0.005	0.487	0.995	1.000	1.000	1.000
0.578	0.585	0.589	0.728	1999	0.071	0.682	0.657	0.720	0.725	0.728
0.211	0.216	0.258	0.405	1787	0.055	0.420	0.350	0.405	0.405	0.405
0.426	0.486	0.504	0.600	1999	0.060	0.895	0.540	0.600	0.600	0.600
0.427	0.476	0.495	0.736	1988	0.217	1.258	0.519	0.736	0.736	0.736
0.627	0.673	0.689	0.813	1994	0.054	0.640	0.759	0.813	0.813	0.813
0.614	0.624	0.629	0.755	1999	0.054	0.770	0.701	0.755	0.755	0.755
0.441	0.450	0.462	0.633	1998	0.098	0.802	0.535	0.633	0.633	0.633
0.459	0.469	0.475	0.645	1998	0.075	0.482	0.570	0.644	0.645	0.645
0.378	0.380	0.466	0.558	1886	0.057	0.663	0.501	0.558	0.558	0.558

[a]Confidence intervals are computed using (14).
[b]Russell measure is caclulated using (6).
[c]Number of bootstrap replications out of $B = 1999$, where observation was feasible.
[d]Bias is caclulated using (12).
[e]The numerator of this statistic is caclulated using (12). $\hat{\sigma}$ the standard error of the bootstrap Russell measures.
[f]$\widetilde{RM}$ is obtained by subtracting bootstrap bias estimate from the Russell measure.

6 Concluding remarks

The voluminous DEA literature uses radial Debreu-Farrell measures of technical efficiency. This is partly because statistical inference regarding individual efficiency is readily available. The nonradial measure of technical

efficiency, the Russell measure, has received less attention and statistical inference for it is non-existent. In this paper we make use of bootstrap methods under homogeneity of the efficiency distribution assumption and adjust existing algorithms to provide statistical inference about individual Russell measures. Development of more general bootstrap procedures is left for future research.

The Monte Carlo evidence suggests that the proposed estimator possesses some attractive finite sample properties. As sample size increases, the bias and mean squared error drop, while prediction of ranking of decision making units improves. Finally, when sample size is smaller than 100, the results tend to be rather unreliable.

Acknowledgments

The authors wish to thank two anonymous Referees and the Editor for their valuable comments. The usual disclaimers apply. The authors acknowledge the Lebesgue Center of Mathematics who supported the work of Second Author according to the program PIA-ANR-11-LABX-0020-01 during his postdoc at Agrocampus Ouest, when a substantial part of this project was carried out.

References

Badunenko, O., Henderson, D. J., and Houssa, R. (2014). Significant drivers of growth in Africa. *Journal of Productivity Analysis*, 42(3):339–354.

Chambers, R. G., Chung, Y., and Färe, R. (1998). Profit, directional distance functions and nerlovian efficiency. *Journal of Optimization Theory and Applications*, 98(2):351–364.

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

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

Färe, R. and Grosskopf, S. (2004). *New Directions: Efficiency and Productivity*. Kluwer Academic Publishers, Boston.

Färe, R., Grosskopf, S., and Lovell, C. A. K. (1985). *The Measurement of Efficiency of Production*. Springer Science Business Media, New York.

Färe, R., Grosskopf, S., and Lovell, C. A. K. (1994). *Production Frontiers*. Cambridge University Press, Cambridge, U.K.

Färe, R. and Lovell, C. A. K. (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. Series A (General)*, 120(3):253–290.

Henderson, D. J. and Parmeter, C. F. (2014). *Applied Nonparametric Econometrics*. Cambridge University Press.

Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities. In Koopmans, T. C., editor, *Activity Analysis of Production and Allocation*. Wiley, New York.

Kumbhakar, S. C. and Tsionas, E. G. (2011). Stochastic error specification in primal and dual production systems. *Journal of Applied Econometrics*, 26(2):270–297.

Lien, D.-H. D. (1985). Moments of truncated bivariate log-normal distributions. *Economics Letters*, 19(3):243–247.

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

Simar, L. and Vanhems, A. (2012). Probabilistic characterization of directional distances and their robust versions. *Journal of Econometrics*, 162(2):342–354.

Simar, L., Vanhems, A., and Wilson, P. W. (2012). Statistical inference with dea estimators of directional distances. *European Journal of Operational Research*, 220(3):853–864.

Simar, L. and Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44:49–61.

Simar, L. and Wilson, P. W. (1999). Estimating and bootstrapping Malmquist indices. *European Journal of Operational Research*, 115(3):459–471.

Simar, L. and Wilson, P. W. (2015). Statistical approaches for non-parametric frontier models: A guided tour. *International Statistical Review*, 83(1):77–110.