

Technical, Allocative and Overall Efficiency: Estimation and Inference (Appendices B–D)

LÉOPOLD SIMAR

PAUL W. WILSON*

September 2019

*Simar: Institut de Statistique, Biostatistique, et Sciences Actuarielles, Université Catholique de Louvain-la-Neuve, Louvain-la-Neuve, Belgium; email: leopold.simar@uclouvain.be.

Wilson: Department of Economics and School of Computing, Division of Computer Science, Clemson University, Clemson, South Carolina, USA; email: pww@clemson.edu.

Contents

B Revenue Efficiency	1
C Output Allocative Efficiency	4
D Computational Details	8

B Revenue Efficiency

The development here and in Appendix C parallels the discussion in Sections 3.3 and 3.4, respectively. The results in Appendices B–C are stated without formal proofs, which would involve straightforward changes in the notation of the proofs of the results in Sections 3.3–3.4.

In the applied literature, revenue efficiency seems to be estimated with less frequency than cost efficiency. Nonetheless, the usual approach to estimating revenue efficiency is to first estimate the vector of output levels that maximize revenue by employing an empirical analog of (2.9). For a firm operating at $(x_0, y_0) \in \Psi$ and facing the given vector w_y of output prices, replacing Ψ with $\hat{\Psi}_n$ in (2.9) leads to

$$\begin{aligned} \mathcal{R}_{\max}(x_0, y_0 \mid \hat{\Psi}_n, w_y) &= \max_{y, \mathbf{v}} \{w'_y y \mid \mathbf{Y}\mathbf{v} \geq y, \mathbf{X}\mathbf{v} \leq x_0, \mathbf{1}'_n \mathbf{v} = 1, \mathbf{v} \in \mathbb{R}_+^n\} \\ &= w'_y \hat{y}_{\max} \end{aligned} \quad (\text{B.1})$$

where $\hat{y}_{\max}$ is the solution to the optimization problem in the first line of (B.1). Then revenue efficiency is estimated by

$$\mathcal{R}(x_0, y_0 \mid \hat{\Psi}_n, w_y) := \frac{\mathcal{R}_{\max}(x_0, y_0 \mid \hat{\Psi}_n, w_y)}{w'_y y_0} = \frac{w'_y \hat{y}_{\max}}{w'_y y_0}. \quad (\text{B.2})$$

This approach is suggested by Farrell (1957), Färe et al. (1985), Färe and Grosskopf (1995), Coelli et al. (1997), Ray (2004) and Simar and Zelenyuk (2018). However, the statistical properties of the both the estimator of maximum revenue in (B.1) and of revenue efficiency in (B.2) are unknown. Hence researchers typically report only point estimates with no inference in applied studies.

The next result establishes a radial efficiency characterization of the revenue efficiency measure $\mathcal{R}(x_0, y_0 \mid \Psi, w_y)$ defined in (2.10).

Lemma B.1. *Let $r_0 = w'_y y_0$. Then for $(x_0, y_0) \in \Psi$,*

$$\mathcal{R}(x_0, y_0 \mid \Psi, w_y) = \lambda(x_0, r_0 \mid \Psi_{w_y}). \quad (\text{B.3})$$

Now apply the function h_{w_y} to each observation $(X_i, Y_i) \in \mathcal{S}_n$ to transform the set $\mathcal{S}_n$ to a set of identically, independently distributed observations $\mathcal{S}_{w_y, n} = \{(X_i, R_i)\}_{i=1}^n$ on pairs of input quantities and revenue. This leads to the FDH and DEA estimators of Ψ_{w_y} discussed earlier in Section 3.2. As in Sections 3.3–3.4 we omit subscripts “FDH” and “DEA” and let $\hat{\Psi}_{w_y, n}$ denote either estimator of Ψ_{w_y} .

Consider the estimator $\lambda(x_0, r_0 \mid \widehat{\Psi}_{w_y, n})$ of revenue efficiency where r_0 and $\widehat{\Psi}_{w_y, n}$ replace y_0 and $\widehat{\Psi}$ (respectively) in (3.4). The next two results establish the properties of this estimator.

Theorem B.1. *Let $c_0 = w'_x x_0$. Let $\kappa_y = 1/(p+1)$ for the FDH case or $\kappa_y = 2/(p+2)$ for the DEA case. Under Assumptions 2.1–2.7 for the DEA case, and under Assumptions 2.1–2.5 and 2.8 for the FDH case,*

$$n^{2/(p+2)} \left(\lambda(x_0, r_0 \mid \widehat{\Psi}_{w_y, n}) - \lambda(x_0, r_0 \mid \Psi_{w_y}) \right) \xrightarrow{\mathcal{L}} Q_{w_y} \quad (\text{B.4})$$

as $n \rightarrow \infty$, where Q_{w_y} is a non-degenerate distribution with finite variance.

Theorem B.2. *Let $\kappa_y = 1/(p+1)$ for the FDH case or $\kappa_y = 2/(p+2)$ for the DEA case. Under Assumptions 2.1–2.7 and 2.9 for the DEA case, and under Assumptions 2.1–2.5, 2.8 and 2.9 for the FDH case, $\exists$ a constant $D_5 \in (0, \infty)$ such that for all $i, j \in \{1, \dots, n\}$, $i \neq j$,*

$$E \left[\lambda(X_i, R_i \mid \widehat{\Psi}_{W_{y,i}, n}) - \lambda(X_i, R_i \mid \Psi_{W_{y,i}}) \right] = D_5 n^{-\kappa_y} + O \left(n^{-\zeta_4} (\log n)^{\zeta_5} \right), \quad (\text{B.5})$$

$$\text{VAR} \left[\lambda(X_i, R_i \mid \widehat{\Psi}_{W_{y,i}, n}) - \lambda(X_i, R_i \mid \Psi_{W_{y,i}}) \right] = O \left(n^{-\zeta_4} (\log n)^{\zeta_4} \right), \quad (\text{B.6})$$

and

$$\begin{aligned} & \left| \text{COV} \left[\lambda(X_i, R_i \mid \widehat{\Psi}_{W_{y,i}, n}) - \lambda(X_i, R_i \mid \Psi_{W_{y,i}}), \lambda(X_j, R_j \mid \widehat{\Psi}_{W_{y,j}, n}) - \lambda(X_j, R_j \mid \Psi_{W_{y,j}}) \right] \right| \\ &= O \left(n^{-\zeta_6} (\log n)^{\zeta_6} \right) = o \left(n^{-1} \right). \end{aligned} \quad (\text{B.7})$$

where expectations are with respect to (X, Y, W_y) , $(\zeta_4, \zeta_5, \zeta_6) = \left(\frac{3}{p+1}, \frac{p+3}{p+1}, \frac{p+2}{p+1} \right)$ for the FDH case and $(\zeta_4, \zeta_5, \zeta_6) = \left(\frac{3}{p+2}, \frac{p+5}{p+2}, \frac{p+3}{p+2} \right)$ for the DEA case. The value of the constant D_5 depends on the particular estimator (FDH or DEA), the density f_{X,Y,W_y} and the structure of the sets $\mathcal{D} \subset \Psi$ and $\mathcal{D}_W$.

Theorem B.1 ensures that the revenue efficiency estimator converges at rate $n^{2/(p+2)}$ and has a non-degenerate limiting distribution. Existence of a non-degenerate limiting distribution and the knowledge of the convergence rate are sufficient to establish that the sub-sampling methods discussed by Simar and Wilson (2011) provide asymptotically valid inference for the revenue efficiency of an individual firm facing output prices w_y . Simar and Wilson (2011) also discuss a method for choosing the sub-sample size, which is critical to the finite-sample performance of the sub-sampling method. However, the result (B.5) in Theorem B.2 means that standard CLTs (e.g., the Lindeberg-Feller CLT) can be used to make inference about mean revenue efficiency if and only if $p = 1$ due to the

bias term $D_5 n^{-2/(p+2)}$ in the DEA case and not at all in the FDH case (see Kneip et al., 2015 for discussion in a similar setting).

Let $\mu_{W_y} = E[\lambda(X, R | \Psi_{W_y})]$ and $\sigma_{W_y}^2 = \text{VAR}[\lambda(X, R | \Psi_{W_y})] < \infty$ denote the mean and variance of cost efficiency, where expectations are with respect to (X, R, W_y) as noted in Theorem B.2. Let

$$\hat{\mu}_{W_y, n} := n^{-1} \sum_{i=1}^n \lambda(X_i, R_i | \hat{\Psi}_{W_y, i, n}). \quad (\text{B.8})$$

Let $\kappa_y = 1/(p+1)$ for the FDH case and $\kappa_y = 2/(p+2)$ for the DEA case. Define $n_{\kappa_y} := \min(\lfloor n^{2\kappa_y} \rfloor, n) \leq n$. Assume the observations in $\mathcal{S}_{w_y, n}$ are randomly sorted. Define

$$\hat{\mu}_{W_y, n_{\kappa_y}} := n_{\kappa_y}^{-1} \sum_{i=1}^{n_{\kappa_y}} \lambda(X_i, R_i | \hat{\Psi}_{W_y, i, n}). \quad (\text{B.9})$$

Note that the efficiency estimates under the summation sign are computed using the full sample of n observations, but the summation is over only the first n_{κ_y} estimates.

Finally, let $\tilde{B}_{w_y, n, \kappa_y}$ denote the generalized jackknife estimate of the bias term $D_5 n^{-\kappa_y}$ in (B.5) computed from $\mathcal{S}_{w_y, n}$ as described by Kneip et al. (2015, Section 4), and then compute the average

$$\hat{B}_{w_y, n, \kappa_y} = K^{-1} \sum_{k=1}^K \tilde{B}_{w_y, n, \kappa_y, k} \quad (\text{B.10})$$

over $K \ll \binom{n}{n/2}$ splits of the sample $\mathcal{S}_{w_y, n}$ as described above in Section 3.3 and in Kneip et al. (2015, Section 4). Then the next result permits inference about mean cost efficiency for any number q of outputs.

Theorem B.3. *Let Assumptions 2.1–2.5, 2.8 and 2.9 hold for the FDH case, or let Assumptions 2.1–2.7 and 2.9 hold for the DEA case. For $p \leq 2$ for the FDH case, or $p \leq 3$ for the DEA case, as $n \rightarrow \infty$,*

$$\sqrt{n} \left(\hat{\mu}_{W_y, n} - \hat{B}_{w_y, n, \kappa_y} - \mu_{W_y} + \xi_{W_y, n, \kappa_y} \right) \xrightarrow{\mathcal{L}} N(0, \sigma_{W_y}^2). \quad (\text{B.11})$$

where $\kappa_y = 2/(p+2)$ and $\xi_{W_y, n, \kappa_y} = O(n^{-\zeta_4} (\log n)^{\zeta_5}) = o(n^{-\kappa_y})$ with ζ_4 and ζ_5 defined as in Theorem B.2 for the FDH and DEA cases. In addition, for $p > 1$ for the FDH case or for $p > 2$ in the DEA case, as $n \rightarrow \infty$

$$n^{\kappa_y} \left(\hat{\mu}_{W_y, n_{\kappa_y}} - \hat{B}_{w_y, n, \kappa_y} - \mu_{W_y} + \xi_{W_y, n, \kappa_y} \right) \xrightarrow{\mathcal{L}} N(0, \sigma_{w_y}^2). \quad (\text{B.12})$$

Moreover, as $n \rightarrow \infty$,

$$\hat{\sigma}_{\mathcal{R}}^2 := \sum_{i=1}^n \left[\lambda(X_i, R_i | \hat{\Psi}_n, W_{y, i}) - \hat{\mu}_{W_y, n} \right]^2 \xrightarrow{p} \sigma_{\mathcal{R}}^2. \quad (\text{B.13})$$

The CLT results in Theorem B.3 can be used to construct confidence intervals for mean cost efficiency or to test hypotheses regarding mean cost efficiency. For example, whenever $p > 2$ (B.12) can be used to construct the asymptotically valid $(1 - \alpha)$ confidence interval

$$\left[\hat{\mu}_{W_y, \kappa_y} - \hat{B}_{w_y, n, \kappa_y} \pm \frac{\hat{\sigma}_{W_y}}{n^{\kappa_y}} z_{(1-\frac{\alpha}{2})} \right] \quad (\text{B.14})$$

where $z_{(1-\frac{\alpha}{2})}$ is the $(1 - \frac{\alpha}{2})$ quantile of the standard normal distribution function. Note that either (B.11) or (B.12) hold for $p = 3$. Intervals based on (B.11) neglect $\sqrt{n}\xi_{w_y, n, \kappa_y} = O(n^{-1/10})$, while those based on (B.12) neglect $n^{\kappa_y}\xi_{w_y, n, \kappa_y} = O(n^{-1/5})$. Hence (B.12) is expected to provide more accurate intervals than (B.11) in finite samples when $p = 3$ and DEA estimators are used, or when $p = 2$ and FDH estimators are used.

C Output Allocative Efficiency

In order to develop properties of an estimator of output allocative efficiency, we first establish properties of the log of the Farrell input-oriented efficiency estimator. Before proceeding, an additional assumption is needed.

Assumption C.1. *There exists a constant $0 < M_y < \infty$ such that $\|y\| \geq M_y$ for all $(x, y) \in \mathcal{D}$.*

Assumption C.1 is necessary to guarantee existence of moments of $\log(\lambda(X, Y \mid \hat{\Psi}_n))$. Although moments necessarily exist for $\lambda(X_i, Y_i) \in [1, \infty)$, $|\log \lambda(X_i, Y_i)|$ is potentially unbounded. While Assumption C.1 could in principle be replaced by a weaker version requiring only existence of all relevant moments, but boundedness of $\|y\|$ greatly simplifies asymptotic arguments used to obtain the results that follow.

Lemma C.1. *Let κ , ζ_1 , ζ_2 and ζ_3 be defined as in Theorem 3.1 for the FDH and DEA cases. Under Assumptions 2.1–2.5, 2.8 and C.1 for the FDH case and under Assumptions 2.1–2.7 and C.1 for the DEA case, for each $(x, y) \in \mathcal{D}$,*

$$n^\kappa \left(\log \left(\lambda(x, y \mid \hat{\Psi}_n) \right) - \log \left(\lambda(x, y \mid \Psi) \right) \right) \xrightarrow{\mathcal{L}} Q_{\lambda, x, y}^{\log} \quad (\text{C.1})$$

where $Q_{\lambda, x, y}^{\log}$ is a non-degenerate distribution with finite variance. In addition, $\exists$ a constant $D_6 \in (0, \infty)$ such that for all $i, j \in \{1, \dots, n\}$, $i \neq j$,

$$E \left[\log \left(\lambda(X_i, Y_i \mid \hat{\Psi}_n) \right) - \log \left(\lambda(X_i, Y_i \mid \Psi) \right) \right] = D_6 n^{-\kappa} + O \left(n^{-\zeta_1} (\log n)^{\zeta_2} \right), \quad (\text{C.2})$$

$$\text{VAR} \left[\log \left(\lambda(X_i, Y_i \mid \hat{\Psi}_n) \right) - \log \left(\lambda(X_i, Y_i \mid \Psi) \right) \right] = O \left(n^{-\zeta_1} (\log n)^{\zeta_1} \right), \quad (\text{C.3})$$

and

$$\begin{aligned} \left| \text{COV} \left[\log \left(\lambda(X_i, Y_i \mid \hat{\Psi}_n) \right) - \log \left(\lambda(X_i, Y_i \mid \Psi) \right), \right. \right. \\ \left. \left. \log \left(\lambda(X_j, Y_j \mid \hat{\Psi}_n) \right) - \log \left(\lambda(X_j, Y_j \mid \Psi) \right) \right] \right| = O \left(n^{-\zeta_3} (\log n)^{\zeta_3} \right) \\ = o \left(n^{-1} \right). \end{aligned} \quad (\text{C.4})$$

The value of the constant D_6 depends on the particular estimator, the density f and the structure of the sets $\mathcal{D}$ and $\mathcal{D}_W$.

Now recall the definition of output allocative efficiency in (2.11). Taking logs yields

$$\log (\mathcal{A}_y(x_0, y_0 \mid \Psi, w_y)) = \log (\lambda(x_0, r_0 \mid \Psi_{w_y})) - \log (\lambda(x_0, y_0 \mid \Psi)). \quad (\text{C.5})$$

Similar to the discussion in Section 3.4, a natural estimator of $\log (\mathcal{A}_y(x_0, y_0 \mid \Psi, w_y))$ is obtained by replacing $\theta(x_0, r_0 \mid \Psi_{w_y})$ and $\lambda(x_0, y_0 \mid \Psi)$ on the right-hand side of (C.5) with the corresponding estimators $\theta(x_0, r_0 \mid \hat{\Psi}_{w_y, n})$ and $\lambda(x_0, y_0 \mid \hat{\Psi}_n)$ given by (B.2) and (3.4). The next result establishes the properties of the resulting estimator

$$\log (\mathcal{A}_y(x_0, y_0 \mid \hat{\Psi}_n, w_y)) = \log (\lambda(x_0, r_0 \mid \hat{\Psi}_{w_y, n})) - \log (\lambda(x_0, y_0 \mid \hat{\Psi}_n)). \quad (\text{C.6})$$

Theorem C.1. *Let κ be defined for the FDH and DEA cases as in Lemma A.1. Then under Assumptions 2.1–2.5, 2.8, and C.1 for the FDH case and under Assumptions 2.1–2.7 and C.1 for the DEA case, for each $(x, y) \in \mathcal{D}$,*

$$n^{\frac{2}{p+q+1}} \left(\log (\mathcal{A}_y(x, y \mid \hat{\Psi}_n, w_y)) - \log (\mathcal{A}_y(x, y \mid \Psi, w_y)) \right) \xrightarrow{\mathcal{L}} Q_{\mathcal{A}_y, x, y}^{\log} \quad (\text{C.7})$$

where $Q_{\mathcal{A}_y, x, y}^{\log}$ is a non-degenerate distribution with finite variance.

Similar to Theorem A.1, Theorem C.1 establishes the existence of a non-degenerate limiting distribution and the rate of convergence for FDH and DEA estimators of the log of output allocative efficiency. This permits estimation of confidence intervals with asymptotically correct coverage for the log of output allocative efficiency of individual firms using the sub-sampling methods of Simar and Wilson (2011) while noting that the rate of convergence is $n^{1/(p+q)}$ for the FDH case or $n^{2/(p+q+1)}$ for the DEA case as established by Theorem C.1. Again, since the resulting intervals are transformation-respecting, one can take exponential of the endpoints to obtain an asymptotically valid confidence interval for $\mathcal{A}_y(x_0, y_0 \mid \Psi, w_x)$.

The next results establishes moment properties for FDH and DEA estimators of log output allocative efficiency.

Theorem C.2. *Let κ , ζ_1 , ζ_2 and ζ_3 be defined for the FDH and DEA cases as in Theorem 3.1 Then under the conditions of Lemma C.1 for either the FDH or the DEA case, $\exists$ a constant $D_7 \in (0, \infty)$ such that for all $i, j \in \{1, \dots, n\}$, $i \neq j$,*

$$E \left[\log \left(\mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_{y,i}) \right) - \log \left(\mathcal{A}_y(X_i, Y_i \mid \Psi, W_{y,i}) \right) \right] = D_7 n^{-\kappa} + O \left(n^{-\zeta_1} (\log n)^{\zeta_2} \right), \quad (\text{C.8})$$

$$\text{VAR} \left[\log \left(\mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_{y,i}) \right) - \log \left(\mathcal{A}_y(X_i, Y_i \mid \Psi, W_{y,i}) \right) \right] = O \left(n^{-\zeta_1} (\log n)^{\zeta_1} \right) \quad (\text{C.9})$$

and

$$\begin{aligned} \left| \text{COV} \left[\log \left(\mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_{y,i}) \right) - \log \left(\mathcal{A}_y(X_i, Y_i \mid \Psi, W_{y,i}, W_{y,j}) \right), \right. \right. \\ \left. \left. \log \left(\mathcal{A}_y(X_j, Y_j \mid \widehat{\Psi}_n, W_{y,j}) \right) - \log \left(\mathcal{A}_y(X_j, Y_j \mid \Psi, W_{y,i}, W_{y,j}) \right) \right] \right| = O \left(n^{-\zeta_3} (\log n)^{\zeta_3} \right) \\ = o \left(n^{-1} \right) \end{aligned} \quad (\text{C.10})$$

where expectations are with respect to (X, Y, W_y) . The constant D_7 depends on the particular estimator (FDH or DEA), the density f , the structure of the set $\mathcal{D} \subset \Psi$ and the function h_{w_x} .

Among other things, Theorem C.1 establishes that $\log \left(\mathcal{A}_y(x, y \mid \widehat{\Psi}_n, w_y) \right)$ is a consistent estimator of $\log \left(\mathcal{A}_y(x, y \mid \Psi, w_y) \right)$. Confidence intervals for the log of input allocative efficiency of individual firms can be estimated using the sub-sampling methods described by Simar and Wilson (2011) while noting that the rate of convergence is $n^{2/(p+q+1)}$ as established by Theorem C.1. Since the resulting intervals are transformation-respecting, one can take exponentials of the endpoints to obtain an asymptotically valid confidence interval for $\mathcal{A}_y(x_0, y_0 \mid \Psi, w_y)$. Alternatively, the logarithmic transformation is avoided in the following pair of results.

Theorem C.3. *Let κ be defined for the FDH and DEA cases as in Lemma A.1. Then under Assumptions 2.1–2.5, 2.8, 2.9 and C.1 for the FDH case and under Assumptions 2.1–2.7, 2.9 and C.1 for the DEA case, for each $(x, y) \in \mathcal{D}$,*

$$n^{\frac{2}{p+q+1}} \left(\mathcal{A}_y(x, y \mid \widehat{\Psi}_n, w_y) - \mathcal{A}_y(x, y \mid \Psi, w_y) \right) \xrightarrow{\mathcal{L}} Q_{\mathcal{A}_y, x, y} \quad (\text{C.11})$$

where $Q_{\mathcal{A}_y, x, y}$ is a non-degenerate distribution with finite variance.

Theorem C.4. Let κ , ζ_1 , ζ_2 and ζ_3 be defined for the FDH and DEA cases as in Theorem 3.1. Then under the conditions of Lemma C.1 for either the FDH or the DEA case, $\exists$ a constant $D_8 \in (0, \infty)$ such that for all $i, j \in \{1, \dots, n\}$, $i \neq j$,

$$E \left[\mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_y) - \mathcal{A}_y(X_i, Y_i \mid \Psi, W_y) \right] = D_8 n^{-\frac{2}{p+q+1}} + O \left(n^{-\frac{3}{p+q+1}} (\log n)^{\frac{p+q+4}{p+q+1}} \right), \quad (\text{C.12})$$

$$\text{VAR} \left[\mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_y) - \mathcal{A}_y(X_i, Y_i \mid \Psi, W_y) \right] = O \left(n^{-\frac{3}{p+q+1}} (\log n)^{\frac{3}{p+q+1}} \right), \quad (\text{C.13})$$

and

$$\begin{aligned} & \left| \text{COV} \left[\mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_y) - \mathcal{A}_y(X_i, Y_i \mid \Psi, W_y), \right. \right. \\ & \quad \left. \left. \mathcal{A}_y(X_j, Y_j \mid \widehat{\Psi}_n, W_y) - \mathcal{A}_y(X_j, Y_j \mid \Psi, W_y) \right] \right| = O \left(n^{-\frac{p+q+2}{p+q+1}} (\log n)^{\frac{p+q+2}{p+q+1}} \right) \\ & \quad = o(n^{-1}). \end{aligned} \quad (\text{C.14})$$

where expectations are with respect to (X, Y, W_y) . The value of the constant D_8 depends on the particular estimator (FDH or DEA), the density f_{X,Y,W_y} and the structure of the sets $\mathcal{D}$ and $\mathcal{D}_W$.

In order to make inference about mean output allocative efficiency, more work is needed due to the bias term $D_8 n^{-\kappa}$ in (C.12). Let $\mu_{\mathcal{A}_y} = E[\mathcal{A}_y(X, Y \mid \Psi, W_y)]$ and $\sigma_{\mathcal{A}_y}^2 = \text{VAR}[\mathcal{A}_y(X, Y \mid \Psi, W_y)] < \infty$ denote the mean and variance of input allocative efficiency, where again expectations are with respect to (X, Y, W_y) . As discussed earlier, output prices are typically taken as fixed for each firm, but are often allowed to vary across firms. In other words, firms are implicitly assumed to face perhaps different, but fixed output prices. Consequently, let $W_{y,i}$ denote the vector of stochastic output prices faced by firm i . Let

$$\widehat{\mu}_{\mathcal{A}_y, n} := n^{-1} \sum_{i=1}^n \mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_{y,i}). \quad (\text{C.15})$$

Let $\kappa = 1/(p+q)$ for the FDH case and $\kappa = 2/(p+q+1)$ for the DEA case, and define $n_\kappa := \min(\lfloor n^{2\kappa} \rfloor, n) \leq n$. Assume the observations in $\mathcal{S}_n$ are randomly sorted. Define

$$\widehat{\mu}_{\mathcal{A}_y, n_\kappa} := n_\kappa^{-1} \sum_{i=1}^{n_\kappa} \mathcal{A}_y(X_i, Y_i \mid \widehat{\Psi}_n, W_{y,i}). \quad (\text{C.16})$$

Analogous to (3.18), the estimates of output allocative efficiency under the summation sign in (C.16) are computed using the full sample of n observations, but the summation is over only the first n_κ estimates.

Finally, let $\tilde{B}_{\mathcal{A}_y, n, \kappa}$ denote the generalized jackknife estimate of the bias term $\mathcal{A}_y(x, y \mid \Psi, w_y) D_8 n^{-\kappa}$ in (C.12) computed as described by Kneip et al. (2015, Section 4). Analogous to (B.10), compute the average

$$\hat{B}_{\mathcal{A}_y, n, \kappa} = K^{-1} \sum_{k=1}^K \tilde{B}_{\mathcal{A}_y, n, \kappa, k} \quad (\text{C.17})$$

over $K \ll \binom{n}{n/2}$ random splits of the sample to reduce the variance of the bias estimate. The next result gives a CLT for mean input allocative efficiency.

Theorem C.5. *Assume the conditions of Lemma C.1 hold for either the FDH or DEA cases. For $(p+q) \leq 3$ in the FDH case or $(p+q) \leq 4$ in the DEA case, as $n \rightarrow \infty$,*

$$\sqrt{n} \left(\hat{\mu}_{\mathcal{A}_y, n} - \hat{B}_{\mathcal{A}_y, n, \kappa} - \mu_{\mathcal{A}_y} + \xi_{\mathcal{A}_y, n, \kappa} \right) \xrightarrow{\mathcal{L}} N \left(0, \sigma_{\mathcal{A}_y}^2 \right) \quad (\text{C.18})$$

where $\xi_{\mathcal{A}_y, n, \kappa} = O \left(n^{-\frac{3}{p+q+1}} (\log n)^{\frac{p+q+4}{p+q+1}} \right) = o(n^{-\kappa})$. Alternatively, for $(p+q) > 2$ in the FDH case or $(p+q) > 3$ in the DEA case, as $n \rightarrow \infty$

$$n^{\kappa} \left(\hat{\mu}_{\mathcal{A}_y, n, \kappa} - \hat{B}_{\mathcal{A}_y, n, \kappa} - \mu_{\mathcal{A}_y} + \xi_{\mathcal{A}_y, n, \kappa} \right) \xrightarrow{\mathcal{L}} N \left(0, \sigma_{\mathcal{A}_y}^2 \right). \quad (\text{C.19})$$

In addition, as $n \rightarrow \infty$,

$$\hat{\sigma}_{\mathcal{A}_y}^2 := \sum_{i=1}^n \left[\mathcal{A}_y(X_i, Y_i \mid \hat{\Psi}_n, w_{y,i}) - \hat{\mu}_{\mathcal{A}_y, n} \right]^2 \xrightarrow{p} \sigma_{\mathcal{A}_y}^2. \quad (\text{C.20})$$

The CLT results in Theorem C.5 can be used to construct confidence intervals for mean input allocative efficiency or to test hypotheses about mean input allocative efficiency. Similar to Theorem 3.2, either (C.18) or (C.19) can be used when $(p+q) = 4$. Intervals based on (C.18) neglect $\sqrt{n} \xi_{\mathcal{A}_y, n, \kappa} = O(n^{-1/10})$, while those based on (C.19) neglect $n^{\kappa} \xi_{\mathcal{A}_y, n, \kappa} = O(n^{-1/5})$. Hence (C.19) is expected to provide more accurate intervals than (C.18) when $(p+q) = 4$ and DEA estimators are used. Similar reasoning applies when $(p+q) = 3$ and FDH estimators are used.

D Computational Details

As noted in Section 4, the estimates in Table 1 are obtained using *R* with the FEAR library due of Wilson (2008). The following shows the *R* statements and the resulting output used to obtain the estimates in the first panel of Table 1 of the full set of 322 banks. The file “randomize.R” contains the *R* code given by Daraio et al. (2018) for randomly shuffling observations.

R version 3.6.0 (2019-04-26) -- "Planting of a Tree"
Copyright (C) 2019 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

```
> #
> # empirical example for sw-cost using data from
> # Aly et al. (1990), "Technical, Scale, and Allocative Efficiencies
> #   in U.S. Banking: An Empirical Investigation," Review of Economics
> #   and Statistics 72, 211--218.
> # Since sub-sample means are reported in the second column of Table 1,
> # the data are randomly shuffled using the randomization function
> # 'randomize.R' given by
> # Daraio et al. (2018), "Central Limit Theorems for Conditional
> #   Efficiency Measures and Tests of the 'Separability' Condition in
> #   Nonparametric, Two-Stage Models of Production," Econometrics
> #   Journal 21, 170--191.
> #
> #
> # In the analysis below,
> #   t1 = VRS-DEA estimates;
> #   t2 = cost efficiency;
> #   t3 = allocative efficiency.
> # Here, the analysis is for the full sample, i.e., for both banks with
> # and without branches.
> #
> #
> setwd(Sys.getenv("PWD"))
> require(FEAR)
Loading required package: FEAR
FEAR (Frontier Efficiency Analysis with R) version 2.9 installed
Copyright Paul W. Wilson 2019
  Type "fear.license()" to view the software license for FEAR
  Type "fear.cite()" to view the proper citation for FEAR
> require(digest)
Loading required package: digest
> source("randomize.R")
> infile="../Data/grab.data"
> tdata=t(matrix(scan(file=infile),nrow=14,ncol=322))
```

```

Read 4508 items
> dim(tdata)
[1] 322 14
> n=nrow(tdata)
> n1=floor(n/2)
> n2=n-n1
> tmp=randomize(tdata)
> tmp=t(tmp)
> x=tmp[3:5,]
> wx=tmp[6:8,]
> y=tmp[9:13,]
> dim(x)
[1] 3 322
> dim(wx)
[1] 3 322
> dim(y)
[1] 5 322
> np=nrow(x)
> nq=nrow(y)
> ones=matrix(1,nrow=1,ncol=n)
> #
> # estimate efficiencies for each bank:
> t1=dea(x,y,ORIENTATION=1,RTS=1,METRIC=2)
> t2=vector(length=n)
> for (i in 1:n) {
+   cost=matrix(apply(x * (wx[,i] %*% ones),2,sum),nrow=1,ncol=n)
+   t2[i]=dea(XOBS=matrix(cost[1,i],nrow=1,ncol=1),
+             YOBS=matrix(y[,i],nrow=nq,ncol=1),
+             XREF=cost,YREF=y,ORIENTATION=1,RTS=1,METRIC=2)
+ }
> t3=t2/t1
> #
> # compute means and variances:
> tm1=mean(t1)
> tm2=mean(t2)
> tm3=mean(t3)
> sig2.1=var(t1)
> sig2.2=var(t2)
> sig2.3=var(t3)
> #
> # generalized jackknife bias estimates:
> kappa=2/(np+nq+1)
> kappa.c=2/(nq+2)
> bc.fac=1/(2**kappa - 1)
> bc.fac.c=1/(2**kappa.c - 1)
> tbar1=rep(0,n)
> tbar2=rep(0,n)
> tbar3=rep(0,n)
> d2a=vector(length=n1)
> d2b=vector(length=n2)
> for (j in 1:100) {
+   if (j==1) {

```

```

+     ind=c(1:n)
+   } else {
+     ind=sample(ind,size=n)
+     x[,1:n]=x[,ind]
+     y[,1:n]=y[,ind]
+     wx[,1:n]=wx[,ind]
+   }
+   d1a=dea(matrix(x[,1:n1],nrow=np),
+           matrix(y[,1:n1],nrow=nq),
+           RTS=1,ORIENTATION=1,METRIC=2)
+   d1b=dea(matrix(x[, (n1+1):n],nrow=np),
+           matrix(y[, (n1+1):n],nrow=nq),
+           RTS=1,ORIENTATION=1,METRIC=2)
+   for (i in 1:n1) {
+     cost=matrix(apply(x * (wx[,i] %*% ones),2,sum),nrow=1,ncol=n)
+     d2a[i]=dea(XOBS=matrix(cost[1,i],nrow=1,ncol=1),
+               YOBS=matrix(y[,i],nrow=nq,ncol=1),
+               XREF=matrix(cost[,1:n1],nrow=1),
+               YREF=y[,1:n1],ORIENTATION=1,RTS=1,METRIC=2)
+   }
+   for (i in 1:n2) {
+     cost=matrix(apply(x * (wx[,i] %*% ones),2,sum),nrow=1,ncol=n)
+     ii=n1+i
+     d2b[i]=dea(XOBS=matrix(cost[1,ii],nrow=1,ncol=1),
+               YOBS=matrix(y[,ii],nrow=nq,ncol=1),
+               XREF=matrix(cost[, (n1+1):n],nrow=1),
+               YREF=y[, (n1+1):n],ORIENTATION=1,RTS=1,METRIC=2)
+   }
+   d3a=d2a/d1a
+   d3b=d2b/d1b
+   tbar1[ind[1:n1]]=tbar1[ind[1:n1]] + d1a - t1[ind[1:n1]]
+   tbar1[ind[(n1+1):n]]=tbar1[ind[(n1+1):n]] + d1b - t1[ind[(n1+1):n]]
+   tbar2[ind[1:n1]]=tbar2[ind[1:n1]] + d2a - t2[ind[1:n1]]
+   tbar2[ind[(n1+1):n]]=tbar2[ind[(n1+1):n]] + d2b - t2[ind[(n1+1):n]]
+   tbar3[ind[1:n1]]=tbar3[ind[1:n1]] + d3a - t3[ind[1:n1]]
+   tbar3[ind[(n1+1):n]]=tbar3[ind[(n1+1):n]] + d3b - t3[ind[(n1+1):n]]
+ }
> tbar1=0.01*bc.fac*tbar1
> tbar2=0.01*bc.fac.c*tbar2
> tbar3=0.01*bc.fac*tbar3
> bc1=mean(tbar1)
> bc2=mean(tbar2)
> bc3=mean(tbar3)
> #
> # compute the re-centered bounds:
> nk=floor(n**(2*kappa))
> nk.c=floor(n**(2*kappa.c))
> tmk1=mean(t1[1:nk])
> tmk2=mean(t2[1:nk.c])
> tmk3=mean(t3[1:nk])
> ts1=sqrt(sig2.1/nk)
> ts2=sqrt(sig2.2/nk.c)

```

```

> ts3=sqrt(sig2.3/nk)
> crit=qnrm(p=c(0.95,0.975,0.99,0.05,0.025,0.005))
> bounds1=matrix((tm1-bc1-ts1*crit),nrow=3,ncol=2)
> bounds2=matrix((tm2-bc2-ts2*crit),nrow=3,ncol=2)
> bounds3=matrix((tm3-bc3-ts3*crit),nrow=3,ncol=2)
> #
> # make first panel of Table 1:
> tm=matrix(c(tm1,tm2,tm3,tmk1,tmk2,tmk3,sqrt(sig2.1),sqrt(sig2.2),
+           sqrt(sig2.3),bc1,bc2,bc3),nrow=3,ncol=4)
> res=cbind(tm[,1:2],
+           rbind(bounds1[2,],bounds2[2,],bounds3[2,]),
+           tm[,3:4])
> tab=matrix(nrow=3,ncol=6)
> for (k in 1:6) {
+   for (j in 1:3) {
+     tab[j,k]=formatC(res[j,k],format="f",width=6,digits=4)
+   }
+ }
> print(noquote(tab))
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,] 0.8021 0.7760 0.4565 0.6506 0.1785 0.2485
[2,] 0.7078 0.6790 0.4530 0.5968 0.1906 0.1829
[3,] 0.8819 0.8979 0.7985 0.9284 0.1195 0.0184
>
> proc.time()
      user  system elapsed
 32.696   0.124  32.818

```

The last line in the output indicates that all of the computations were done in less than 33 seconds. The code was run on a single core of a 3.5GHz 6-Core Intel Xeon E5 processor using MacOS version 10.13.6.

References

- Coelli, T., D. S. P. Rao, and G. E. Battese (1997), *An Introduction to Efficiency and Productivity Analysis*, Boston: Kluwer Academic Publishers.
- Daraio, C., L. Simar, and P. W. Wilson (2018), Central limit theorems for conditional efficiency measures and tests of the ‘separability condition’ in non-parametric, two-stage models of production, *The Econometrics Journal* 21, 170–191.
- Färe, R. and S. Grosskopf (1995), Nonparametric tests of regularity, Farrell efficiency, and goodness-of-fit, *Journal of Econometrics* 69, 415–425.
- Färe, R., S. Grosskopf, and C. A. K. Lovell (1985), *The Measurement of Efficiency of Production*, Boston: Kluwer-Nijhoff Publishing.
- Farrell, M. J. (1957), The measurement of productive efficiency, *Journal of the Royal Statistical Society A* 120, 253–281.
- Kneip, A., L. Simar, and P. W. Wilson (2015), When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores, *Econometric Theory* 31, 394–422.
- Ray, S. C. (2004), *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Reserach*, Cambridge: Cambridge University Press.
- Simar, L. and P. W. Wilson (2011), Inference by the m out of n bootstrap in nonparametric frontier models, *Journal of Productivity Analysis* 36, 33–53.
- Simar, L. and V. Zelenyuk (2018), Central limit theorems for aggregate efficiency, *Operations Research* 66, 137–149.
- Wilson, P. W. (2008), FEAR: A software package for frontier efficiency analysis with R, *Socio-Economic Planning Sciences* 42, 247–254.