I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2015/18

Testing the "Separability" Condition in Two-Stage Nonparametric Models of Production

DARAIO, D., SIMAR, L. AND P. WILSON

Testing the "Separability" Condition in Two-Stage Nonparametric Models of Production

Cinzia Daraio Léopold Simar Paul W. Wilson*

August 2015

Abstract

Simar and Wilson (*J. Econometrics,* 2007) provided a statistical model that can rationalize two-stage estimation of technical efficiency in nonparametric settings. Two-stage estimation has been widely used, but requires a strong assumption: the second-stage environmental variables cannot affect the support of the input and output variables in the first stage. In this paper, we provide a fully nonparametric test of this assumption. The test relies on new central limit theorem (CLT) results for unconditional efficiency estimators developed by Kneip et al. (*Econometric Theory,* 2015a) and new CLTs for *conditional* efficiency estimators developed in this paper. The test can be implemented relying on either asymptotic normality of the test statistics or using bootstrap methods to obtain critical values. Our simulation results indicate that our tests perform well both in terms of size and power. We present a real-world empirical example by updating the analysis performed by Aly et al. (*R. E. Stat.,* 1990) on U.S. commercial banks; our tests easily reject the assumption required for two-stage estimation, calling into question results that appear in *hundreds* of papers that have been published in recent years.

Keywords: technical efficiency, conditional efficiency, two-stage estimation, bootstrap, separability, data envelopment analysis (DEA), free-disposal hull (FDH).
JEL classification codes: C12, C14, C44, C46.

*Daraio: Department of Computer, Control and Management Engineering Antonio Ruberti (DIAG), University of Rome "La Sapienza", Rome, Italy; email daraio@dis.uniroma1.it.

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

Wilson: Department of Economics and School of Computing, Clemson University, Clemson, SC 29634; email pww@clemson.edu (corresponding author).

Research support from IAP Research Network P7/06 of the Belgian State (Belgian Science Policy) and from U.S. National Science Foundation grant no. SMA-1243436 is gratefully acknowledged. We have used the Palmetto cluster maintained by Clemson Computing and Information Technology (CCIT) at Clemson University; we are grateful for technical support by the staff of CCIT. Any remaining errors are solely our responsibility.

1 Introduction

Two-stage estimation procedures wherein technical efficiency is estimated by data envelopment analysis (DEA) or free disposal hull (FDH) estimators in the first stage, and the resulting efficiency estimates are regressed on some environmental variables in a second stage, are very popular in the literature. Simar and Wilson (2007) cite 48 published papers that employ this approach and commented that "as far as we have been able to determine, none of the studies that employ this two-stage approach have described the underlying data-generating process." Simar and Wilson go on to (i) define a statistical model where truncated (but not censored, i.e., tobit, nor ordinary least squares) regression yields consistent estimation of model features, (ii) demonstrate that conventional, likelihood-based approaches to inference are invalid, and (iii) develop a bootstrap approach that yields valid inference in the second-stage regression. The model defined by Simar and Wilson rationalizes second-stage regressions of estimated efficiency on environmental variables in the sense that such a regression estimates a feature of the model described by Simar and Wilson. However, as noted by Simar and Wilson, the model contains a crucial feature—and a strong restriction—in the form of a "separability condition" that appears below as Assumption 2.1. Without this condition, second-stage regressions of estimated efficiency do not estimate any meaningful model feature; as Simar and Wilson (2007), this condition should be tested before estimating a second-stage regression, but until now no test has been available. Such a test is provided in this paper.

A number of papers have appeared in recent years using the approach suggested by Simar and Wilson (2007). However, papers that estimate technical efficiency in the first stage and then regress these estimates on some environmental variables in a second-stage tobit model continue to appear. As far as we know, none of these papers present a statistical model in which second-stage tobit estimation would consistently estimate features of the model; the approach is ad hoc in each case. Moreover, regardless of how the second-stage regression is specified, any results from such regressions are meaningless for reasons given below when the separability condition is violated.[1]

[1] A search on Google Scholar on 14 August 2015 using the keywords "dea," "efficiency," "tobit," and "two stage" returned 3,240 papers with dates between 2008 and 2015. As far as we know, none of these papers present a statistical model in which second-stage tobit estimation would consistently estimate features of the model; the approach is ad hoc in each case. Repeating the search after dropping the keyword "tobit"

Recently, Daraio and Simar (2005) develop *conditional* measures of efficiency, which allow nonparametric estimation of technical efficiency conditional on some explanatory variables in a single stage. This raises some important questions for practitioners, such as the question of precisely how environmental variables might affect the production process. In the model presented by Simar and Wilson (2007), environmental variables affect the shape (i.e., mean, variance, etc.) of the distribution of inefficiencies, but not the support of input or output variables. Conceivably, however, environmental variables might have other effects; in particular, they might affect the production possibilities themselves. The statistical model in Simar and Wilson rationalizes second-stage regression of efficiency estimates on some environmental variables, but does not allow for the possibility that environmental variables might affect the production possibilities. If they do, then a different model is needed, and second-stage regression is not appropriate.

In this paper, we present a carefully-developed framework—i.e., a statistical model—in order to make clear how environmental variables might be relevant, and how to test whether two-stage approaches might be meaningful (i.e., whether the separability condition described by Simar and Wilson, 2007 and required by studies that have used the two-stage approach is satisfied). We then extend the CLT results of Kneip et al. (2015a) to *conditional* efficiency estimators; while the new CLTs are useful in their own right for making various hypothesis tests along the lines of Kneip et al. (2015b), they are needed to develop our separability test. We then develop test statistics and prove that they have asymptotic normal limiting distributions from which critical values for implementing the test can be obtained. In addition, we describe a bootstrap method that can be used to assess the significance of test statistics without incurring a large computational burden; results from Monte Carlo simulations suggest that in many cases the bootstrap tests have better power than those relying on asymptotic normality. In cases where two-stage approaches are found to be inappropriate, one can (and should) estimate efficiency conditionally on environmental variables, for reasons given below.

In the next section, we develop the statistical model. Estimators are discussed in Section

returned 19,300 papers over the same years. Even if only half of these hits are relevant, the searches indicate that the practice of regressing nonparametric efficiency estimates on some environmental variables in a second-stage regression is widespread, although perhaps many of these exercises yield meaningless results if the separability condition is frequently violated.

3, and the tests are developed in Section 4. Section 5 describes Monte Carlo experiments used to assess the size and power of our tests as well as results. In Section 6 we provide a real-world example by revisiting the work of Aly et al. (1990) and testing whether the assumptions given by Simar and Wilson (2007) that are required for the two-stage approach used by Aly et al. to be meaningful are satisfied. Conclusions are given in the final section. Appendix A gives technical assumptions used to derive results in Section 4, and Appendix B discusses how one can handle discrete environmental variables.

2 The Production Process in the Presence of Environmental Factors

In this section we formalize a statistical model of the production process along the lines of the probability framework of Cazals et al. (2002). The production process generates random variables (X, Y, Z) in an appropriate probability space, where $X \in \mathbb{R}^p_+$ is the vector of input quantities, $Y \in \mathbb{R}^q_+$ is the vector of output quantities and $Z \in \mathbb{R}^r$ is a vector of variables describing environmental factors. These factors Z are neither inputs nor outputs and are typically not under the control of the manager, but they may influence the production process in different ways as explained below. Let $f_{XYZ}(x, y, z)$ denote the joint density of (X, Y, Z) which has support $\mathcal{P} \subset \mathbb{R}^p_+ \times \mathbb{R}^q_+ \times \mathbb{R}^r$. This joint density can always be decomposed as

$$f_{XYZ}(x, y, z) = f_{XY|Z}(x, y \mid z) f_Z(z). \tag{2.1}$$

Let Ψ^z denote the conditional support of $f_{XY|Z}(x, y \mid z)$, i.e., the support of (x, y) given $Z = z$, and let $\mathcal{Z}$ be the support of $f_Z(z)$. Then Ψ^z is the set of feasible combinations of inputs and outputs for a firm facing the environmental conditions $Z = z$; i.e.,

$$\Psi^z = \{(X, Y) \mid X \text{ can produce } Y \text{ when } Z = z\}. \tag{2.2}$$

The environmental variables in Z can affect the production process either (i) only through Ψ^z, the support of (X, Y), or (ii) only through the density $f_{XY|Z}(x, y \mid z)$, thereby affecting the probability for a firm to be near its optimal boundary, or (iii) through both Ψ^z and $f_{XY|Z}(x, y \mid z)$. Let

$$\Psi = \bigcup_{z \in \mathcal{Z}} \Psi^z. \tag{2.3}$$

Clearly, $\Psi \subset \mathbb{R}_+^{p+q}$, but whether Ψ is useful for benchmarking the performance of a firm producing output levels y from input levels x while facing levels z of the environmental variables depends on whether the "separability" condition described by Simar and Wilson (2007) is satisfied. This condition requires that Z affect production *only* through the conditional density $f_{XY|Z}(x, y \mid z)$ without affecting its support Ψ^z, and is stated explicitly in Assumption 2.1.

Assumption 2.1. *(Separability Condition):* $\Psi^z = \Psi$ *for all* $z \in \mathcal{Z}$.

Clearly, when Assumption 2.1 holds the joint support of (X, Y, Z) can be factorized as

$$\mathcal{P} = \Psi \times \mathcal{Z}, \tag{2.4}$$

and Ψ can be interpreted as the unconditional attainable set

$$\Psi = \{(X, Y) \mid X \text{ can produce } Y\}. \tag{2.5}$$

However, Ψ has the interpretation in (2.5) *if and only if* (iff) Assumption 2.1 holds. The separability condition is very strong and restrictive. Under Assumption 2.1, the environmental factors influence neither the *shape* nor the *level* of the boundary of the attainable set, and the potential effect of Z on the production process is only through the distribution of the inefficiencies. If the separability condition holds, it is meaningful to measure the efficiency of a particular production plan (x, y) by its distance to the boundary of Ψ. For example, under separability, the output-oriented Farrell efficiency score is given by

$$\lambda(x, y) = \sup\{\lambda > 0 \mid (x, \lambda y) \in \Psi\}. \tag{2.6}$$

In this case, it is meaningful to analyze the behavior of $\lambda(x, y)$ as a function of Z by using an appropriate regression model (see Simar and Wilson, 2007, 2011 for details).[2]

Alternatively, if the separability condition does not hold, then we have a more general situation where the factor Z may influence the level and the shape of the boundary of the attainable sets (and may also influence the conditional density $f_{XY|Z}(x, y \mid z)$). The following assumption characterizes this situation explicitly.

[2] We focus the presentation in this paper using output-oriented measures of efficiency such as the one in (2.6), but of course efficiency can be measured in other directions as desired. See the recent surveys by Simar and Wilson (2013, 2015) and the references cited therein for details. All of the results here are easily generalized to input, hyperbolic, and directional distance functions after straight-forward (but perhaps tedious) changes in notation.

Assumption 2.2. *(Non Separability Assumption):* $\Psi^z \neq \Psi$ *for some* $z \in \mathcal{Z}$*, i.e., for some* $z, \tilde{z} \in \mathcal{Z}$*,* $\Psi^z \neq \Psi^{\tilde{z}}$*.*

Note that Assumptions 2.1 and 2.2 are mutually exclusive; one and only one holds in a given situation.

Under Assumption 2.2, the efficiency measure in (2.6) is difficult to interpret; in fact, it is economically meaningless because it does not measure the distance to the appropriate boundary. If Assumption 2.2 holds, the set Ψ can still be defined as in (2.3), but for benchmarking production units, the boundary of Ψ has little interest in this case because it may be unattainable for some firms faced with unfavorable conditions represented described by z. In such cases, the conditional measure

$$\lambda(x, y \mid z) = \sup\{\lambda > 0 \mid (x, \lambda y) \in \Psi^z\} \tag{2.7}$$

introduced by Cazals et al. (2002) and Daraio and Simar (2005) gives a measure of distance to the appropriate, relevant boundary (i.e., the boundary that is attainable by firms operating under conditions described by z).

The distinction between Assumptions 2.1 and 2.2, and their implications for how environmental variables in Z affect the production process, has often been neglected in the literature where researchers analyze the effect of Z on $\lambda(X, Y)$ by estimating some regression of $\lambda(X, Y)$ on Z. Typically, starting with a sample of observations (X_i, Y_i, Z_i), DEA or FDH estimators $\widehat{\lambda}(X_i, Y_i)$ computed in a first stage are regressed on Z_i in a second-stage analysis. Even if Assumption 2.1 holds, additional problems described in Simar and Wilson (2007) remain to be solved in the second stage to obtain sensible inference. Theoretical results on how to make inference in a second stage linear regression, when appropriate, is described in detail by Kneip et al. (2015a). However, if Assumption 2.2 holds, the two-stage approach is almost certain to lead to incorrect results and inferences about the effect of Z on the production process. This explains why it is important, as noted by Simar and Wilson (2007)—indeed, essential—to test Assumption 2.1 against Assumption 2.2. If the test rejects separability in favor of Assumption 2.2, then only a second-stage regression of the conditional measure $\lambda(X, Y \mid Z)$ on Z can be meaningful, as described for example in Bădin et al. (2012).[3]

[3] A search for papers using Google Scholar on July 16, 2015 found approximately 4,500 hits using keywords

In order to derive results below, the efficiency measures in (2.6) and (2.7) must be defined in terms of components of our probability model. Cazals et al. (2002) show that under free disposability (see Assumption 4.2 below) the output-oriented efficiency measure in (2.6) can be written as

$$\lambda(x, y) = \sup\{\lambda > 0 \mid H_{XY}(x, \lambda y) > 0\}, \tag{2.8}$$

where $H_{XY}(x, y) = \Pr(X \leq x,\ Y \geq y)$ is the probability of finding a firm dominating the production unit operating at the level (x, y).[4] This can be factored as $\Pr(X \leq x) \Pr(Y \geq y \mid X \leq x) = F_X(x) S_{Y|X}(y \mid X \leq x)$, where the latter conditional survival function is nonstandard due to the the condition $X \leq x$. For (x, y) such that x is in the interior of its support (i.e., $F_X(x) > 0$), the efficiency score can be written equivalently as

$$\lambda(x, y) = \sup\{\lambda > 0 \mid S_{Y|X}(\lambda y \mid X \leq x) > 0\}. \tag{2.9}$$

Along the same lines, the conditional efficiency score can be expressed as

$$\lambda(x, y \mid z) = \sup\{\lambda > 0 \mid H_{XY|Z}(x, \lambda y \mid z) > 0\}, \tag{2.10}$$

where $H_{XY|Z}(x, y \mid z) = \Pr(X \leq x,\ Y \geq y \mid Z = z)$ is the probability of finding a firm dominating the production unit operating at the level (x, y) and facing environmental conditions z and is the distribution function corresponding to the conditional density $f_{XY|Z}(x, y \mid z)$ introduced earlier. Analogous to (2.9), the conditional efficiency measure can also be written as

$$\lambda(x, y \mid z) = \sup\{\lambda > 0 \mid S_{Y|X,Z}(\lambda y \mid X \leq x, Z = z) > 0\} \tag{2.11}$$

while noting the different roles of X and Z in the conditioning of the conditional survival function $S_{Y|X,Z}(y \mid X \leq x, Z = z) = \Pr(Y \geq y \mid X \leq x, Z = z)$.

3 Non-parametric Efficiency Estimators

The literature on nonparametric statistical inference for efficiency scores is by now well-developed. Here, we summarize the definitions and properties needed to test Assumption 2.1 versus Assumption 2.2. Consider a sample of identically, independently (iid) observations

"dea," "efficiency," and "second-stage regression" while restricting the search to papers dated 2008 through 2015. Apparently, the warnings of Simar and Wilson (2007) have not been heeded.

[4] Note that as usual, inequalities involving vectors are defined on an element-by-element basis.

$\mathcal{S}_n = \{(X_i, Y_i, Z_i) \mid i = 1, \ldots, n\}$. Following Deprins et al. (1984), the FDH of the sample $\mathcal{S}_n$ is the set

$$\widehat{\Psi}_{\text{FDH}}(\mathcal{S}_n) = \bigcup_{(X_i,Y_i)\in\mathcal{S}_n} \left\{(x,y) \in \mathbb{R}_+^{p+q} \mid y \leq Y_i,\ x \geq X_i\right\}. \tag{3.1}$$

The convex hull of $\widehat{\Psi}_{\text{FDH}}(\mathcal{S}_n)$ given by

$$\widehat{\Psi}_{\text{DEA}}(\mathcal{S}_n) = \Big\{(x,y) \in \mathbb{R}_+^{p+q} \,|y \leq \sum_{i=1}^n \omega_i Y_i,\ x \geq \sum_{i=1}^n \omega_i X_i, \\ \sum_{i=1}^n \omega_i = 1,\ \omega_i \geq 0\ \forall\ i = 1, \ldots, n\Big\} \tag{3.2}$$

provides the DEA estimator proposed by Farrell (1957) and popularized by Charnes et al. (1978).[5]

The corresponding efficiency estimators are obtained by plugging these estimators into the definition of $\lambda(x,y)$ in (2.6). Using $\widehat{\Psi}_{\text{FDH}}(\mathcal{S}_n)$ in the FDH case leads to

$$\widehat{\lambda}_{\text{FDH}}(x, y \mid \mathcal{S}_n) = \max_{i=1,\ldots,\, n|X_i\leq x} \left(\min_{j=1,\, \ldots,\, p} \left(\frac{Y_i^j}{y^j}\right)\right), \tag{3.3}$$

where y^j, Y_i^j denote the jth elements of y (i.e., the input vector corresponding to the fixed point of interest) and Y_i (i.e., the output vector corresponding to the ith observation in $\mathcal{S}_n$). This is simply the plug-in version of (2.8), where $H_{XY}(x,y)$ is replaced by its empirical version

$$\widehat{H}_{XY}(x,y) = n^{-1}\sum_{i=1}^n I(X_i \leq x, Y_i \geq y), \tag{3.4}$$

where $I(A)$ is the indicator function equal 1 if A is true and 0 otherwise. In the DEA case, replacing Ψ in (2.6) with $\widehat{\Psi}_{\text{DEA}}(\mathcal{S}_n)$ from (3.2) gives the DEA efficiency estimator

$$\widehat{\lambda}_{\text{DEA}}(x, y \mid \mathcal{S}_n) = \max_{\lambda,\omega_1,\, \ldots,\, \omega_n} \Big\{\lambda > 0 \,|\lambda y \leq \sum_{i=1}^n \omega_i Y_i,\ x \geq \sum_{i=1}^n \omega_i X_i, \\ \sum_{i=1}^n \omega_i = 1,\ \omega_i \geq 0\ \forall\ i = 1, \ldots, n\Big\}. \tag{3.5}$$

For the conditional efficiency scores we need to use a smoothed estimator of $H_{XY|Z}(x, y \mid z)$ to plug in (2.10), which requires a vector of bandwidths for Z. Denoting this r-vector

[5] Note that in (3.1)–(3.2), the data on Z_i are ignored; only the first $(p+q)$ components of the ordered $(p+q+r)$-tuples in $\mathcal{S}_n$ are used.

of bandwidths by h, the conditional distribution function $H_{XY|Z}(x, y \mid z)$ is replaced by the estimator

$$\widehat{H}_{XY|Z}(x, y \mid z) = \frac{\sum_{i=1}^{n} I(X_i \leq x, Y_i \geq y) K_h(Z_i - z)}{\sum_{i=1}^{n} K_h(Z_i - z)}, \tag{3.6}$$

where $K_h(\cdot) = (h_1 \ \ldots \ h_r)^{-1} K\left((Z_i - z)/h\right)$ and the division between vectors is understood to be component-wise. As explained in the literature (e.g., see Daraio and Simar, 2007b), the kernel function $K(\cdot)$ must have bounded support (e.g., the Epanechnikov kernel).[6] This provides the estimator

$$\widehat{\lambda}_{\text{FDH}}(x, y \mid z, \mathcal{S}_n) = \max_{i \in \mathcal{I}(z,h)} \left(\min_{j=1, \ldots, p} \left(\frac{Y_i^j}{y^j} \right) \right), \tag{3.7}$$

where $\mathcal{I}(z, h) = \{i \mid z - h \leq Z_i \leq z + h\}$.

Alternatively, where one is willing to assume that the conditional attainable sets are convex, Daraio and Simar (2007b) suggest a conditional DEA estimator of $\lambda(x, y \mid z)$, namely

$$\widehat{\lambda}_{\text{DEA}}(x, y \mid z, \mathcal{S}_n) = \max_{\lambda, \omega_1, \ldots, \omega_n} \Big\{ \lambda > 0 \mid \lambda y \leq \sum_{i \in \mathcal{I}(z,h)} \omega_i Y_i, \ x \geq \sum_{i \in \mathcal{I}(z,h)} \omega_i X_i, \\ \text{for some } \omega_i \geq 0 \ \text{ such that } \sum_{i \in \mathcal{I}(z,h)} \omega_i = 1, \Big\}. \tag{3.8}$$

Note that the conditional estimators in (3.7) and (3.8) are just localized version of the unconditional FDH and DEA efficiency estimators given in (3.3) and (3.5), where the degree of localization is controlled by the bandwidth in h. Practical aspects for choosing bandwidths are discussed below in Section 4.5.

The properties of nonparametric efficiency estimators have been examined in a number of papers in recent years. Park et al. (2000) and Daouia et al. (2015) derive the rate of convergence and limiting distribution of the FDH efficiency estimator. Kneip et al. (1998) derived the rate of convergence of the DEA estimator in (3.5), while Kneip et al. (2008) derived its limiting distribution. Kneip et al. (2015a) provide results on the moments of both FDH and DEA estimators. See Simar and Wilson (2013, 2015) for comprehensive surveys of the literature. To summarize relevant results for the unconditional efficiency estimators,

[6] An alternative would be, following Bădin et al. (2010), to plug a smoothed estimator of $S_{Y|X,Z}(y \mid X \leq x, Z = z)$ into (2.11), but as shown in Simar et al. (2015), if the two methods are asymptotically equivalent, the latter provides a bandwidth for z that depends on x and the resulting efficiency estimate may not be monotone decreasing in x in finite samples, as the target $\lambda(x, y \mid z)$ is.

under Assumptions 2.1, 4.1, 4.2 and some additional, appropriate regularity conditions (e.g., monotonicity, smoothness of the frontier and smoothness of the density of (X, Y)), for a fixed point (x, y) in the interior of Ψ, as $n \to \infty$,

$$n^{\kappa}\left(\widehat{\lambda}(x, y \mid \mathcal{S}_n) - \lambda(x, y)\right) \xrightarrow{\mathcal{L}} Q_{xy}(\cdot) \tag{3.9}$$

where $Q_{xy}(\cdot)$ is a regular, non-degenerate distribution with parameters depending on the characteristics of the DGP and on (x, y), and κ determines the rate of convergence.[7] For the FDH estimator, $\kappa = 1/(p+q)$ while for the DEA estimator, $\kappa = 2/(p+q+1)$. For the FDH case, the limiting distribution belongs to the Weibull family, but with parameters that are difficult to estimate. For the DEA case, the limiting distribution does not have a closed form. Hence in either case, inference on individual efficiency scores requires bootstrap techniques. In the DEA case, Kneip et al. (2008) provide theoretical results for both a smoothed bootstrap and for subsampling, while Kneip et al. (2011) and Simar and Wilson (2011) provide details and methods for practical implementation. Subsampling can also be used for inference in the FDH case; see Jeong and Simar (2006) and Simar and Wilson (2011).

Jeong et al. (2010) show that the conditional version of the FDH and DEA efficiency estimators share properties similar to their unconditional counterparts whenever the elements of Z are continuous.[8] The sample size n is replaced by the effective sample size used to build the estimates, which is of order $nh_1 \ \ldots \ h_r$, which we write hereafter as nh^r for simplicity (hoping the reader will indulge the abuse of notation, since the individual bandwidths may differ). For a fixed point (x, y) in the interior of Ψ^z, as $n \to \infty$,

$$(nh^r)^{\kappa}\left(\widehat{\lambda}(x, y \mid z, \mathcal{S}_n) - \lambda(x, y \mid z)\right) \xrightarrow{\mathcal{L}} Q_{xy|z}(\cdot) \tag{3.10}$$

where again $Q_{xy|z}(\cdot)$ is a regular, non-degenerate limiting distribution analogous to the corresponding one for the unconditional case. The main argument in Jeong et al. (2010) relies on regularity conditions discussed in the next section, but also on the property that there are enough points in a neighborhood of z, which is obtained with the additional assumption

[7] Here and in the exposition that follows, we omit the subscripts "FDH" and "DEA" from the efficiency estimator in order to describe results in a generic fashion, thereby conserving space.

[8] We discuss below in Appendix B how discrete "environmental" variables can be handled. Otherwise, except in Appendix B, we assume throughout that all elements of Z are continuous.

that $f_Z(z)$ is bounded away from zero at z and that if the bandwidth is going zero, it should not go too fast (see Jeong et al., 2010, Proposition 1; if $h \to 0$, h should be of order $n^{-\alpha}$ with $\alpha < 1/r$). We will return to this point in the discussion following Lemma 4.1 below.

4 Testing Separability

4.1 Basic Ideas

The goal is to test the null hypothesis of separability (Assumption 2.1) against its complement (Assumption 2.2). The idea for building a test statistics is to compare the conditional and unconditional efficiency scores using relevant statistics that are functions of $\widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_n)$ and $\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)$ for $i = 1, \ldots, n$. Note that under Assumption 2.1, $\lambda(X, Y) = \lambda(X, Y \mid Z)$ with probability one, even if Z may influence the distribution of the inefficiencies inside the attainable set, and the two estimators converge to the same object. But under Assumption 2.2, the conditional attainable sets Ψ^z are different and the two estimators converge to different objects. Moreover, under Assumption 2.2, $\lambda(X, Y) \geq \lambda(X, Y \mid Z)$ with strict inequality holding for some $(X, Y, Z) \in \mathcal{P}$.

The approach developed here is similar to those developed in Kneip et al. (2015b) for testing constant versus variable returns to scale or for testing convexity versus non-convexity of the attainable set. Now consider the sample means

$$\widehat{\mu}_n = n^{-1} \sum_{i=1}^{n} \widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_n) \tag{4.1}$$

and

$$\widehat{\mu}_{c,n} = n^{-1} \sum_{i=1}^{n} \widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) \tag{4.2}$$

of unconditional and conditional efficiency estimators. The efficiency estimators in (4.1) and (4.2) could be either FDH or DEA estimators, but for purposes of the following discussion, suppose the same type of estimators (FDH or DEA) are used in both (4.1) and (4.2). By construction $(\widehat{\mu}_n - \widehat{\mu}_{c,n}) \geq 0$, and the null hypothesis of separability should be rejected if this difference is "too big". However, several problems remain to be solved.

In particular, From Kneip et al. (2015a) we know that even under the null hypothesis, standard central limit theorems (e.g., the Lindeberg-Feller theorem) cannot be used with $\widehat{\mu}_n$

to make inferences about population means unless $(p+q) < 3$ in the DEA case or $(p+q) < 2$ in the FDH case. As will be seen below, similar problems exist for $\widehat{\mu}_{c,n}$. Moreover, even if with the applicable CLT from Kneip et al. (2015a) and the CLT proved below for the conditional estimators, the (asymptotic) distribution of the difference $(\widehat{\mu}_n - \widehat{\mu}_{c,n})$ is quite complicated due to the covariance between the two estimators. A viable solution to this problem is to randomly split the sample $\mathcal{S}_n$ into two independent parts consisting of n_1 and n_2 observations (such that $n_1 + n_2 = n$), and compute $\widehat{\mu}_{n_1}$ using the first part $\mathcal{S}_{n_1}$ and $\widehat{\mu}_{c,n_2}$ using the second part $\mathcal{S}_{n_2}$. This provides two independent statistics where it will be possible to apply the results of Kneip et al. (2015a) and additional results proved below to derive the sampling distribution under the null. But this requires some preliminary steps to adapt the existing results to the setup here. We demonstrate below in Section 5 that the procedure works well in practice with finite sample sizes.

4.2 Sampling distribution of averages of the efficiency scores

As noted by Kneip et al. (2015a), availability of the asymptotic results for efficiency estimated at a fixed point (x, y) is useful, but not sufficient for analyzing the behavior of statistics that are function of FDH or DEA estimators evaluated at random points (X_i, Y_i). In the discussion below, we denote the FDH and DEA efficiency estimators by $\widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_n)$ to stress the fact that the estimator is to be evaluated at a random point (X_i, Y_i).

4.2.1 Asymptotic Moments of Efficiency Estimators

Kneip et al. (2015a) prove that for the unconditional FDH and DEA estimators, under some regularity conditions (see Kneip et al., 2015a for details) and as $n \to \infty$,

$$E\left(\widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_n) - \lambda(X_i, Y_i)\right) = Cn^{-\kappa} + R_{n,\kappa} \tag{4.3}$$

$$E\left(\left(\widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_n) - \lambda(X_i, Y_i)\right)^2\right) = o\big(n^{-\kappa}\big), \tag{4.4}$$

and

$$\left|\mathrm{COV}\left(\widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_n) - \lambda(X_i, Y_i),\ \widehat{\lambda}(X_j, Y_j \mid \mathcal{S}_n) - \lambda(X_j, Y_j)\right)\right| = o\big(n^{-1}\big) \tag{4.5}$$

for all $i, j \in \{1, \ \ldots, \ n\}$, $i \neq j$ and where $R_{n,k} = o\left(n^{-\kappa}\right)$. The values of the constant C, the rate κ, and the remainder term $R_{n,\kappa}$ depends on which estimator is used. For the

DEA estimator, $\kappa = 2/(p+q+1)$ and $R_{n,\kappa} = O(n^{-3\kappa/2}(\log n)^{\alpha_1})$; for the FDH estimator, $\kappa = 1/(p+q)$ and $R_{n,\kappa} = O(n^{-2\kappa}(\log n)^{\alpha_2})$. The values of $\alpha_j > 1$, $j = 1, 2$ are given in Kneip et al. (2015a). For purposes of the results needed here, the $\log n$ factor contained in $R_{n,\kappa}$ does not play a role and can be ignored. The results outlined here are valid under a set of corresponding regularity assumptions (see Theorems 3.1 and 3.3 in Kneip et al., 2015a).

Similar results are needed for the asymptotic moments of the conditional efficiency estimators. To achieve this we follow the arguments of Jeong et al. (2010), who note that for a given h, the conditional FDH and DEA estimators in (3.7) and (3.8) do not target $\lambda(x, y \mid z)$, but instead estimate

$$\lambda^h(x, y \mid z) = \sup\{\lambda > 0 \mid (x, y) \in \Psi^{z,h}\}, \tag{4.6}$$

with the *conditional* attainable set given by

$$\begin{aligned}\Psi^{z,h} &= \{(X, Y) \mid X \text{ can produce } Y, \text{ when } |Z - z| \leq h\} \\ &= \left\{(x, y) \in \mathbb{R}^{p+q}_+ \mid H^h_{XY|Z}(x, y \mid z) > 0\right\} \\ &= \left\{(x, y) \in \mathbb{R}^{p+q}_+ \mid f^h_{XY|Z}(\cdot, \cdot \mid z) > 0\right\}\end{aligned} \tag{4.7}$$

where $H^h_{XY|Z}(x, y \mid z) = \Pr(X \leq x, Y \geq y \mid z - h \leq Z \leq z + h)$ gives the probability of finding a firm dominating the production unit operating at the level (x, y) and facing environmental conditions Z in an h-neighborhood of z and $f^h_{XY|Z}(\cdot, \cdot \mid z)$ is the corresponding conditional density of (X, Y) given $|Z - z| \leq h$. Alternatively, (4.6) can be written as

$$\lambda^h(x, y \mid z) = \sup\left\{\lambda > 0 \mid H^h_{XY|Z}(x, \lambda y \mid z) > 0\right\}. \tag{4.8}$$

Moreover, it is clear that $\Psi^{z,h} = \bigcup_{|\widetilde{z} - z| \leq h} \Psi^{\widetilde{z}}$.

Consequently, for all points (x, y) in the support of $f_{XY|Z}(x, y \mid z)$, the error of estimation can be decomposed as

$$\widehat{\lambda}(x, y \mid z) - \lambda(x, y \mid z) = \underbrace{\widehat{\lambda}(x, y \mid z) - \lambda^h(x, y \mid z)}_{=\Delta_1} + \underbrace{\lambda^h(x, y \mid z) - \lambda(x, y \mid z)}_{=\Delta_2}, \tag{4.9}$$

where the first difference (Δ_1) is due to the estimation error in the localized problem and the second difference (Δ_2) is the non-random bias (≤ 0) introduced by the localization.

Some assumptions are needed to define a statistical model. The next three assumptions are conditional analogs of standard assumptions made by Shephard (1970), Färe (1988), Kneip et al. (2015a) and others.

Assumption 4.1. *For all $z \in \mathcal{Z}$, Ψ^z and $\Psi^{z,h}$ are closed.*

Assumption 4.2. *For all $z \in \mathcal{Z}$, both inputs and outputs are strongly disposable; i.e., for any $z \in \mathcal{Z}$, $\widetilde{x} \geq x$ and $0 \leq \widetilde{y} \leq y$, if $(x, y) \in \Psi^z$ then $(\widetilde{x}, y) \in \Psi^z$ and $(x, \widetilde{y}) \in \Psi^z$. Similarly, if $(x, y) \in \Psi^{z,h}$ then $(\widetilde{x}, y) \in \Psi^{z,h}$ and $(x, \widetilde{y}) \in \Psi^{z,h}$.*

Assumption 4.2 corresponds to Assumption 1F in Jeong et al. (2010), and amounts to a regularity condition on the conditional attainable sets justifying the use of the localized versions of the FDH and DEA estimators. The assumption imposes weak monotonicity on the frontier in the space of inputs and outputs for a given $z \in \mathcal{Z}$, and is standard in micro-economic theory of the firm.

When the DEA estimators are used, the following assumption (corresponding to Assumption 1D in Jeong et al., 2010) is also needed.

Assumption 4.3. *For all $z \in \mathcal{Z}$, Ψ^z and $\Psi^{z,h}$ are convex in $\mathbb{R}_+^{p+q}$.*

The next assumption concerns the regularity of the density of Z and of the conditional density of (X, Y) given $Z = z$, as a function of z in particular near the efficient boundary of Ψ^z (see Assumption 6 in Jeong et al., 2010).

Assumption 4.4. *Z has a continuous density $f_Z(\cdot)$ such that for all $z \in \mathcal{Z}$ $f_Z(z)$ is bounded away from zero. Moreover the conditional density $f_{XY|Z}(\cdot, \cdot \mid z)$ is continuous in z and is strictly positive in a neighborhood of the boundary points.*

A number of additional assumptions are needed to complete the statistical model and to permit statistical analysis of the conditional estimators that have been introduced above as well as the test statistics introduced below. These assumptions are given in Appendix A. Depending on the estimators that are used in a particular situation (i.e., either DEA or FDH), only a subset of the assumptions listed in Appendix A are required.

Our first result establishes smoothness of the potential influence of z on the frontier of Ψ^z. The result is needed in order to control the bias due to the localization, and is expressed in terms of a continuity condition of $\lambda(\cdot, \cdot \mid z)$ as a function of z.

Lemma 4.1. *Under either Assumption A.5 (for FDH case) or under Assumption A.6 (for the DEA case), For all (x, y) in the support of (X, Y),*

$$\lambda^h(x, y \mid z) - \lambda(x, y \mid z) = O(h) \tag{4.10}$$

as $h \to 0$,

Proof. Either assumption A.5 or A.6 is sufficient to establish Lipschitz continuity of $\lambda(x, y \mid z)$ as a function of z. The result follows immediately. ■

Note that if Z is separable and has no effect on the frontier and (4.10) is trivially satisfied for all h. As noted in Bădin et al. (2015), it is easy to show that if $h \propto n^{-\gamma}$ with $1/r > \gamma > 1/(r + \kappa^{-1})$, the difference in (4.10) will be $o\left((nh^r)^{-\kappa}\right)$. We need $\gamma < r^{-1}$ to ensure there are enough observations in the h-neighborhood of z (see Proposition 1 in Jeong et al., 2010). Since we cannot find an explicit expression for the second component Δ_2 in (4.9), and since the Weibull distribution linked to the first component Δ_1 contains unknown parameters, the best that can be done is to determine the order of an optimal bandwidth by balancing the order of the two error terms which leads to $h \propto n^{-1/(r+\kappa^{-1})}$, and then to take, as usual in nonparametric smoothing techniques, a smaller bandwidth to eliminate the bias term due to the localization as suggested in Jeong et al. (2010, Assumption 2). As expected, the order of the optimal bandwidth depends on the dimensions of Z as well as of X and Y. Below, in Section 4.5, we show how to select bandwidths h of appropriate order in applied work (see also the discussions in Bădin et al., 2015).

The following result provides moments for the conditional efficiency estimators.

Theorem 4.1. *Let* $n_h = \min(n, nh^r)$. *Suppose Assumptions 4.1, 4.2, 4.4, A.1, A.2, A.3 and A.4 hold. Then under Assumption A.5 for FDH case, or under Assumptions 4.3 and A.6 for the DEA case, as* $n \to \infty$,

$$E\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) - \lambda^h(X_i, Y_i \mid Z_i)\right) = C_c n_h^{-\kappa} + R_{c,n_h,\kappa}, \tag{4.11}$$

where $R_{c,n_h,\kappa} = o\left(n_h^{-\kappa}\right)$,

$$E\left(\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) - \lambda^h(X_i, Y_i \mid Z_i)\right)^2\right) = o\big(n_h^{-\kappa}\big), \tag{4.12}$$

and

$$\left|COV\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) - \lambda^h(X_i, Y_i \mid Z_i),\ \widehat{\lambda}(X_j, Y_j \mid Z_j, \mathcal{S}_n) - \lambda^h(X_j, Y_j \mid Z_j)\right)\right| = o\big(n_h^{-1}\big) \tag{4.13}$$

for all $i, j \in \{1, \ldots, n\}$, $i \neq j$. *In addition, for the conditional DEA estimator* $R_{c,n_h,\kappa} = O(n_h^{-3\kappa/2}(\log n_h)^{\alpha_1})$ *and for the conditional FDH estimator* $R_{c,n_h,\kappa} = O(n_h^{-2\kappa}(\log n_h)^{\alpha_2})$.

Proof. Under (i) Assumptions 4.1, 4.2, 4.4, A.1, A.2 and two-times differentiability (due to Assumption A.5) of $\lambda(x, y \mid z)$ with respect to x and y for the FDH case, or under (ii) Assumptions 4.1, 4.2, 4.3, 4.4, A.1, A.2 and three-times differentiability (due to Assumption A.6) of $\lambda(x, y \mid z)$ with respect to x and y for the DEA case, Jeong et al. (2010) prove, using the result in Lemma 4.1 and $h = O\left((nh^r)^{-\kappa}\right)$, that the asymptotic behavior of $(nh^r)^{\kappa}\left(\widehat{\lambda}(x, y \mid z, \mathcal{S}_n) - \lambda(x, y \mid z)\right)$ is the same as the asymptotic behavior of $(nh^r)^{\kappa}\left(\widehat{\lambda}(x, y \mid z, \mathcal{S}_n) - \lambda^h(x, y \mid z)\right)$, which leads to the result in (3.10). For any given h, we are in a localized version of the framework of Kneip et al. (2015a) for unconditional efficiencies, except that here $\lambda^h(X_i, Y_i \mid Z_i)$ is the object of interest.

If Z is irrelevant, i.e. if Assumption 2.1 holds, then the optimal $h \to \infty$ and $n_h = n$. Otherwise Assumption 2.2 holds and $h \to 0$ as $n \to \infty$, and the order of the number of observations affecting the estimator is $n_h = nh^r$. Moreover, this is the order of the cardinality of $\mathcal{I}(z, h)$ for all z. Then for the FDH case, the results follow directly from the proof of Theorem 3.3 in Kneip et al. (2015a) after changing notation there to reflect the different number of observations. Similarly for the DEA case, the results follow directly from the proof of Theorem 3.1 in Kneip et al. (2015a). ■

As will be seen, the $\log(n_h)$ factors appearing in the expressions for $R_{c,n_h,\kappa}$ do not play a role in the results that are derived below. The results here should not be surprising since the number of observations used to estimate the moments is reduced by the bandwidths; e.g., the rates n^{κ} for the unconditional estimators are reduced to n_h^{κ} for the conditional estimators.

4.2.2 Central Limit Theorems (CLT)

Here, we use the properties of moments of the conditional efficiency estimators derived in Section 4.2.1 to develop CLTs for means of conditional efficiency estimators.

For the case of means of *un*conditional efficiency estimators, Theorem 4.1 of Kneip et al. (2015a) establishes that

$$\sqrt{n}\left(\widehat{\mu}_n - \mu - Cn^{-\kappa} - R_{n,\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma^2) \tag{4.14}$$

as $n \to \infty$, where $\mu = E\left(\lambda(X, Y)\right)$ and $\sigma^2 = \mathrm{VAR}\left(\lambda(X, Y)\right)$. The theorem also establishes that $\widehat{\sigma}^2 = n^{-1}\sum_{i=1}^{n}\left(\widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_n) - \widehat{\mu}_n\right)^2$ is a consistent estimator of σ^2. Conventional CLTs (e.g., the Lindeberg-Feller CLT) do not account for the bias term $Cn^{-\kappa}$, and hence are

invalid for means of unconditional efficiency estimators unless unless $\kappa > 1/2$. In the case of FDH estimators, $\kappa > 1/2$ iff $(p+q) \leq 1$; in the case of DEA estimators, $\kappa > 1/2$ iff $(p+q) \leq 2$. If $\kappa = 1/2$, the bias is stable as $n \to \infty$, but if $\kappa < 1/2$, the bias explodes asymptotically. Kneip et al. (2015a) solve this problem by incorporating a generalized jackknife estimate of the bias and considering, when needed, test statistics based on averages over a subsample of observations. We use a similar approach below, although with the unconditional efficiency estimators, the problem is rather more complicated than the one in Kneip et al. (2015a) due to the localization in the conditional efficiency estimators.

Define

$$\mu_c^h = E\left(\lambda^h(X, Y \mid Z)\right) = \int_{\mathcal{P}} \lambda^h(x, y \mid z) f_{XYZ}(x, y, z)\, dx\, dy\, dz \tag{4.15}$$

and

$$\sigma_c^{2,h} = \mathrm{VAR}\left(\lambda^h(X, Y \mid Z)\right) = \int_{\mathcal{P}} \left(\lambda^h(x, y \mid z) - \mu_c^h\right)^2 f_{XYZ}(x, y, z)\, dx\, dy\, dz. \tag{4.16}$$

These are the localized analogs of μ and σ^2. Next, let $\overline{\mu}_{c,n} = n^{-1} \sum_{i=1}^n \lambda^h(X_i, Y_i \mid Z_i)$. Although $\overline{\mu}_{c,n}$ is not observed, by the Lindeberg-Feller CLT

$$\sqrt{n}\left(\overline{\mu}_{c,n} - \mu_c^h\right) \xrightarrow{\mathcal{L}} N(0, \sigma_c^{2,h}) \tag{4.17}$$

under mild assumptions.

An obvious solution might be to estimate μ_c^h by $\widehat{\mu}_{c,n}$, but this proves problematic. To see this, define $\zeta_n = \widehat{\mu}_{c,n} - \overline{\mu}_{c,n}$. It is clear that $E(\zeta_n) = C_c n_h^{-\kappa} + R_{c,n_h,\kappa}$ by (4.11), and $\mathrm{VAR}(\zeta_n) = o\left(n_h^{-1}\right)$ due to (4.12) and (4.13). It follows that $\zeta_n - E(\zeta_n) = o_p\left(n_h^{-1/2}\right)$. Now define $\widetilde{\mu}_{c,n} = E\left(\widehat{\mu}_{c,n}\right)$. Then

$$\widetilde{\mu}_{c,n} = \mu_c^h + C_c n_h^{-\kappa} + R_{c,n_h,\kappa}, \tag{4.18}$$

and it follows that

$$\begin{aligned} \widehat{\mu}_{c,n} - \widetilde{\mu}_{c,n} &= \overline{\mu}_{c,n} - \mu_c^h + \zeta_n - E(\zeta_n), \\ &= \overline{\mu}_{c,n} - \mu_c^h + o_p\left(n_h^{-1/2}\right). \end{aligned} \tag{4.19}$$

Clearly $\sqrt{n}(\widehat{\mu}_{c,n} - \widetilde{\mu}_{c,n})$ diverges as $n \to \infty$ since although $\sqrt{n}(\overline{\mu}_{c,n} - \mu_c^h) \xrightarrow{\mathcal{L}} N(0, \sigma_c^{2,h})$, $n^{1/2} o_p\left(n_h^{-1/2}\right)$ diverges if $n_h < n$ since $n_h = nh^r = n^{1-\gamma r}$ with $1/(r + \kappa^{-1}) < \gamma < 1/r$.

Moreover, unless Z is irrelevant, $n_h < n$ for an optimal choice of h. Changing the scaling and considering $n^a(\widehat{\mu}_{c,n} - \widetilde{\mu}_{c,n})$ for some a such that $0 < a < (1 - \gamma r)/2 < 1/2$ does not work because the limiting distribution collapses to a point mass at zero in this case. Consequently, it seems there is no way to develop a CLT for means of conditional efficiency estimators analogous to the one in (4.14) for means of unconditional efficiency estimators.

The following result will be useful for the results developed below.

Lemma 4.2. *Under the assumptions Theorem 4.1, for $\kappa = 1/(p+q)$ in the case of the FDH estimator and for $\kappa = 2/(p+q+2)$ in the case of the DEA estimator,*

$$E\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)\right) = \mu_c^h + C_c n_h^{-\kappa} + R_{c,n_h,\kappa} \tag{4.20}$$

and

$$VAR\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)\right) = \sigma_c^{2,h} + o\left(n_h^{-\kappa/2}\right), \tag{4.21}$$

where $R_{c,n_h,\kappa} = o(n_h^{-\kappa})$.

Proof. The result in (4.20) follows directly from Theorem 4.1. In addition,

$$\begin{aligned}
\mathrm{VAR}(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)) =& E\left(\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) - \lambda^h(X_i, Y_i \mid Z_i)\right)^2\right) \\
&+ E\left(\left(\lambda^h(X_i, Y_i \mid Z_i) - E\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)\right)\right)^2\right) \\
&+ 2E\Big(\left(\lambda^h(X_i, Y_i \mid Z_i) - E\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)\right)\right) \\
&\quad \left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) - \lambda^h(X_i, Y_i \mid Z_i)\right)\Big).
\end{aligned} \tag{4.22}$$

Using the result in (4.11) from Theorem 4.1,

$$\begin{aligned}
E\left(\left[\lambda^h(X_i, Y_i \mid Z_i) - E\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)\right)\right]^2\right) &= \sigma_c^{2,h} + \left[E\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) - \lambda^h(X_i, Y_i \mid Z_i)\right)\right]^2 \\
&= \sigma_c^{2,h} + C_c^2 n_h^{-2\kappa} + o\left(n_h^{-2\kappa}\right).
\end{aligned} \tag{4.23}$$

Applying the Cauchy-Schwartz inequality, the result in (4.21) in Theorem 4.1 and (4.23), the last term in (4.22) is bounded by $o\left(n_h^{\kappa/2}\right)$, establishing the result in (4.21). ■

Next, suppose $n_h < n$ (i.e., Z is relevant), and consider a random subsample $\mathcal{S}_{n_h}^*$ from $\mathcal{S}_n$ of size n_h where for simplicity we use the optimal rates for the bandwidths so that $n_h = \lfloor n^{1/(\kappa r+1)} \rfloor$ where $\lfloor a \rfloor$ denotes floor(a), i.e., the integer part of a. Define

$$\widehat{\mu}_{c,n_h} = \frac{1}{n_h} \sum_{\{(X_i,Y_i,Z_i) \in \mathcal{S}_{n_h}^*\}} \widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n), \tag{4.24}$$

and let $\widetilde{\mu}_{c,n_h} = E(\widehat{\mu}_{c,n_h})$. Note that the estimators on the right-hand side of (4.24) are computed relative to the full sample $\mathcal{S}_n$, but the summation is over elements of the subsample $\mathcal{S}^*_{n_h}$.

The next result provides our first CLT for means of conditional efficiency estimators.

Theorem 4.2. *Under the assumptions of Theorem 4.1, the following conditions hold as* $n \to \infty$ *with* $\kappa = 1/(p+q)$ *for the FDH case and* $\kappa = 2/(p+q+1)$ *for the DEA case: (i)* $\widetilde{\mu}_{c,n_h} = \mu_c^h + C_c n_h^{-\kappa} + R_{c,n_h,\kappa}$*; (ii)* $\widehat{\mu}_{c,n_h} - \widetilde{\mu}_{c,n_h} = \overline{\mu}_{c,n_h} - \mu_c^h + o\left(n_h^{-1/2}\right)$*; (iii)* $\sqrt{n_h}\left(\widehat{\mu}_{c,n_h} - \mu_c^h - C_c n_h^{-\kappa} - R_{c,n_h,\kappa}\right) \xrightarrow{\mathcal{L}} N(0, \sigma_c^{2,h})$*; and (iv)* $\widehat{\sigma}^{2,h}_{c,n} = n^{-1}\sum_{i=1}^n \left[\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) - \widehat{\mu}_{c,n}\right]^2 \xrightarrow{p} \sigma_c^{2,h}$.

Proof: Let

$$\overline{\mu}_{c,n_h} = \frac{1}{n_h} \sum_{(X_i,Y_i,Z_i)\in\mathcal{S}^*_{n_h}} \lambda^h(X_i, Y_i \mid Z_i). \tag{4.25}$$

By the Lindeberg-Feller CLT, $\sqrt{n_h}(\overline{\mu}_{c,n_h} - \mu_c^h) \xrightarrow{\mathcal{L}} N(0, \sigma_c^{2,h})$. Define $\zeta_{n_h} = \widehat{\mu}_{c,n_h} - \overline{\mu}_{c,n_h}$. Using Lemma 4.2, we have $E(\zeta_{n_h}) = C_c n_h^{-\kappa} + R_{c,n_h,\kappa}$, $\mathrm{VAR}(\zeta_{n_h}) = o\left(n_h^{-1}\right)$ and $\zeta_{n_h} - E(\zeta_{n_h}) = o_p\left(n_h^{-1/2}\right)$.

It can be shown that $\widetilde{\mu}_{c,n_h} = \mu_c^h + E(\zeta_{n_h})$, and part (i) of the results is obtained by substitution for $E(\zeta_{n_h})$. Next, note that $\widehat{\mu}_{c,n_h} - \widetilde{\mu}_{c,n_h} = \left(\zeta_{n_h} + \overline{\mu}_{c,n_h}\right) - \left(\mu_c^h - E(\zeta_{n_h})\right) = \overline{\mu}_{c,n_h} - \left(\mu_c^h + (\zeta_{n_h} + E(\zeta_{n_h})\right)$. The last term in parentheses is $o_p\left(n_h^{-1/2}\right)$, establishing the result in (ii). Part (iii) follows directly from part (ii). Finally,

$$\begin{aligned}
\widehat{\sigma}^{2,h}_{c,n} &= n^{-1}\sum_{i=1}^n (\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n))^2 - \widehat{\mu}^2_{c,n_h} \\
&\xrightarrow{p} E[(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n))^2] - (\mu_c^h)^2 \\
&= \mathrm{VAR}(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)) + \left[E\left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n)\right)\right]^2 - (\mu_c^h)^2.
\end{aligned}$$

The result obtains after applying the results of Lemma 4.2. ■

There are no cases where standard CLTs with rate $\sqrt{n}$ may be used with means of conditional efficiency estimators, unless Z is irrelevant (i.e., unless Assumption 2.1 holds). Theorem 4.2 provides a CLT for means of conditional efficiency estimators, but the convergence rate is $\sqrt{n_h}$ as opposed to $\sqrt{n}$, and the result is of practical use only if $\kappa > 1/2$. If $\kappa = 1/2$, the bias term $C_c n_h^{-\kappa}$ does not vanish, and if $\kappa < 1/2$, the bias term explodes as $n \to \infty$. These cases are addressed below.

4.2.3 Bias corrections and subsample averaging

For the unconditional case, all necessary details can be found in Kneip et al. (2015a, Theorems 4.3 and 4.4). Here, we derive corresponding results for conditional efficiency estimators. Assume the observations in $\mathcal{S}_n$ are randomly ordered, and to simplify notation, assume n is even. Let $\mathcal{S}_{n/2}^{(1)}$ denote the set of the first $n/2$ observations from $\mathcal{S}_n$, and let $\mathcal{S}_{n/2}^{(2)}$ denote the set of remaining $n/2$ observations from $\mathcal{S}_n$.[9] Next, for $j \in \{1,\ 2\}$ define

$$\widehat{\mu}_{c,n/2}^{j} = (n/2)^{-1} \sum_{(X_i,Y_i,Z_i)\in\mathcal{S}_{n/2}^{(j)}} \widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_{n/2}^{(j)}). \tag{4.26}$$

Let $\widetilde{\mu}_{c,n/2} = E\left(\widehat{\mu}_{c,n/2}^{1}\right) = E\left(\widehat{\mu}_{c,n/2}^{2}\right)$ and define

$$\overline{\mu}_{c,n/2}^{j} = \frac{2}{n} \sum_{(X_i,Y_i,Z_i)\in\mathcal{S}_{n/2}^{(j)}} \lambda^{h}(X_i, Y_i \mid Z_i). \tag{4.27}$$

By (4.19),

$$\widehat{\mu}_{c,n/2}^{j} - \widetilde{\mu}_{c,n/2} = \overline{\mu}_{c,n/2}^{j} - \mu_c^h + o_p(n_h^{-1/2}) \tag{4.28}$$

for $j \in \{1,\ 2\}$. Now define $\widehat{\mu}_{c,n/2}^{*} = \left(\widehat{\mu}_{c,n/2}^{1} + \widehat{\mu}_{c,n/2}^{2}\right)/2$. Clearly,

$$\widehat{\mu}_{c,n/2}^{*} - \widetilde{\mu}_{c,n/2} = \overline{\mu}_{c,n} - \mu_c^h + o_p(n_h^{-1/2}). \tag{4.29}$$

Subtracting (4.19) from (4.29) and re-arranging terms yields

$$\widehat{\mu}_{c,n/2}^{*} - \widehat{\mu}_{c,n} = \widetilde{\mu}_{c,n/2} - \widetilde{\mu}_{c,n} + o_p(n_h^{-1/2}). \tag{4.30}$$

Since $\widetilde{\mu}_{c,n/2} - \widetilde{\mu}_{c,n} = C_c(2^{\kappa}-1)n_h^{-\kappa} + R_{c,n_h,\kappa}$ we obtain an estimator

$$\widetilde{B}_{\kappa,n_h}^{c} = (2^{\kappa}-1)^{-1}\left(\widehat{\mu}_{c,n/2}^{*} - \widehat{\mu}_{c,n}\right) = C_c n_h^{-\kappa} + R_{c,n_h,\kappa} + o_p(n_h^{-1/2}), \tag{4.31}$$

of the leading bias term $C_c n_h^{-\kappa}$ in Theorem 4.2, part (iii), noting that the remainder term $R_{c,n_h,\kappa} = o(n_h^{-\kappa})$ can be neglected.

Of course, for n even there are $\binom{n}{n/2}$ possible splits of the sample $\mathcal{S}_n$. As noted by Kneip et al. (2015b), the variation in $\widetilde{B}_{\kappa,n_h}^{c}$ can be reduced by repeating the above steps

[9] If n is odd, $\mathcal{S}_{n/2}^{(1)}$ can contain the first $\lfloor n/2 \rfloor$ observations and $\mathcal{S}_{n/2}^{(2)}$ can contain remaining $n - \lfloor n/2 \rfloor$ observations from $\mathcal{S}_n$. The fact that $\mathcal{S}_{n/2}^{(2)}$ contains one more observation than $\mathcal{S}_{n/2}^{(1)}$ makes no difference asymptotically.

$K \ll \binom{n}{n/2}$ times, shuffling the observations before each split of $\mathcal{S}_n$, and then averaging the bias estimates. This yields a generalized jackknife estimate

$$\widehat{B}^c_{\kappa,n_h} = K^{-1} \sum_{k=1}^{k} \widetilde{B}^c_{\kappa,n_h,k}, \tag{4.32}$$

where $\widetilde{B}^c_{\kappa,n_h,k}$ represents the value computed from (4.31) using the kth sample split.

Combining results yields the following:

Theorem 4.3. *Under the Assumptions of Theorem 4.1, with $\kappa = 1/(p+q) \geq 1/3$ in the FDH case or $\kappa = 2/(p+q+1) \geq /2/5$ in the DEA case,*

$$\sqrt{n_h} \left(\widehat{\mu}_{c,n_h} - \mu^h_c - \widehat{B}^c_{\kappa,n_h} - R_{c,n_h,\kappa} \right) \xrightarrow{\mathcal{L}} N(0, \sigma^{2,h}_c) \tag{4.33}$$

as $n \to \infty$.

Proof. The result follows by substituting (4.32) in Theorem 4.2, part (iii), and noting that for the indicated ranges of values for κ, $\sqrt{n_h} R_{c,n_h,\kappa} = o(1)$. ■

If κ is smaller than $1/3$ in the FDH case, or $2/5$ in the DEA case, then the remainder term does not vanish fast enough and $\sqrt{n_h} R_{c,n_h,\kappa} \to \infty$ as $n \to \infty$. In such cases, the approach of averaging efficiency scores over a subsample of smaller size as in Kneip et al. (2015a) must be employed.

Define $n_{h,\kappa} = \lfloor n_h^{2\kappa} \rfloor$ so that $\sqrt{n_{h,\kappa}} < n_h^{1/2}$ when $\kappa < 1/2$. Then define

$$\widehat{\mu}_{c,n_{h,\kappa}} = \frac{1}{n_{h,\kappa}} \sum_{(X_i,Y_i,Z_i) \in \mathcal{S}^{**}_{n_{h,\kappa}}} \widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_n) \tag{4.34}$$

where $\mathcal{S}^{**}_{n_{h,\kappa}}$ is a random subsample of size $n_{h,\kappa}$ from $\mathcal{S}_n$.

Theorem 4.4. *Under the Assumptions of Theorem 4.1, with $\kappa = 1/(p+q)$ in the FDH case or $\kappa = 2/(p+q+1)$ in the DEA case,*

$$\sqrt{n_{h,\kappa}} \left(\widehat{\mu}_{c,n_{h,\kappa}} - \mu^h_c - \widehat{B}^c_{\kappa,n_h} - R_{c,n_h,\kappa} \right) \xrightarrow{\mathcal{L}} N(0, \sigma^{2,h}_c), \tag{4.35}$$

as $n \to \infty$ whenever $\kappa < 1/2$.

Proof. let

$$\overline{\mu}_{c,n_{h,\kappa}} = \frac{1}{n_{h,\kappa}} \sum_{(X_i,Y_i,Z_i) \in \mathcal{S}^{**}_{n_{h,\kappa}}} \lambda^h(X_i, Y_i \mid Z_i). \tag{4.36}$$

Clearly,

$$\widehat{\mu}_{c,n_{h,\kappa}} - \mu_c^h = \overline{\mu}_{c,n_{h,\kappa}} - \mu_c^h + \frac{1}{n_{h,\kappa}} \sum_{(X_i,Y_i,Z_i)\in\mathcal{S}^{**}_{n_{h,\kappa}}} \left(\widehat{\lambda}(X_i,Y_i \mid Z_i,\mathcal{S}_n) - \lambda^h(X_i,Y_i \mid Z_i)\right). \quad (4.37)$$

Since $n_{h,\kappa} \to \infty$ as $n \to \infty$, $\sqrt{n_{h,\kappa}}\left(\overline{\mu}_{c,n_{h,\kappa}} - \mu_c^h\right) \xrightarrow{\mathcal{L}} N(0,\sigma_c^{2,h})$. By Lemma 4.2, the third term on the right-hand side of (4.37) has expectation $\mu_c^h + C_c n_h^{-\kappa} + R_{c,n_h,\kappa}$ and variance $\sigma_c^{2,h} + o(n_h^{-\kappa/2})$. Replacing $C_c n_h^{-\kappa}$ with $\widehat{B}^c_{\kappa,n_h}$ and then multiplying both sides by $\sqrt{n_{h,\kappa}}$ yields the result. ■

Remark 4.1. *Kneip et al. (2015a) note that for selected values of $p+q$, two different CLTs are available for means of unconditional efficiency estimators. The same is true for the conditional cases. With the DEA estimator when $p+q=4$ (so that $\kappa = 2/5$), using Theorem 4.3 neglects a term $\sqrt{n_h}R_{c,n_h,\kappa} = O\big(n_h^{-1/10}\big)$, whereas using Theorem 4.4, and an average over a subsample we neglect a term $\sqrt{n_{h,\kappa}}R_{c,n_h,\kappa} = O\left(n_h^{-1/5}\right)$ and we might expect a better approximation. For the conditional FDH estimator when $p+q=3$ (and hence $\kappa = 1/3$), using Theorem 4.3 implies an error of order $O\left(n_h^{-1/6}\right)$, and using an average over a subsample implies, by Theorem 4.4, an error of the smaller order $O\left(n_h^{-1/3}\right)$.*

4.3 Test Statistics

As noted above, in order to test the hypothesis that Z is separable, i.e., to test H_0: Assumption 2.1 holds versus H_1: Assumption 2.2 holds, one might consider the difference between estimators of $\mu = E(\lambda(X,Y))$ and $\mu_c^h = E(\lambda^h(X,Y \mid Z))$, which under the null estimate the same quantity. When the null is true, $\lambda(X,Y) \equiv \lambda^h(X,Y|Z)$ with probability one, for all values of h. Under the null, The two estimators $\widehat{\mu}_n$ and $\widehat{\mu}_{c,n_h}$ have (when appropriately rescaled, depending on the value of κ), an asymptotic normal distribution with mean $\mu = \mu_c^h$ and variance $\sigma^2 = \sigma_c^{2,h}$ for all h, and so both are consistent estimators of the common μ. As explained in the preceding section, we can also, in both cases, correct for the inherent bias of the estimators.

However, the properties of $(\widehat{\mu}_n - \widehat{\mu}_{c,n_h})$ (and their bias-corrected versions) are complicated due to the covariance between the two estimators, and this covariance is hard to estimate. Even in the limiting case where h is big enough so that $n_h = n$, it is clear that under the

null, the asymptotic distribution of $(\widehat{\mu}_n - \widehat{\mu}_{c,n_h})$ will be degenerate with mass one at zero.[10]

The solution used here is analogous to the method used in the test for convexity of Ψ described by Kneip et al. (2015b). In particular, the sample $\mathcal{S}_n$ can be split into two independent, parts $\mathcal{S}_{1,n_1}$, $\mathcal{S}_{2,n_2}$ such that $n_1 + n_2 = n$, $\mathcal{S}_{1,n_1} \bigcup \mathcal{S}_{2,n_2} = \mathcal{S}_n$, and $\mathcal{S}_{1,n_1} \bigcap \mathcal{S}_{2,n_2} = \emptyset$. The n_1 observations in $\mathcal{S}_{1,n_1}$ are used for the unconditional estimates, while the n_2 observations in $\mathcal{S}_{2,n_2}$ are used for the conditional estimates. Recall that the unconditional efficiency estimators converge at rate n^{κ}, while the conditional efficiency estimators converge at rate $(nh^r)^{\kappa}$. The optimal bandwidths are of order $n^{-\kappa/(r\kappa+1)}$, giving a rate of $n^{\kappa/(r\kappa+1)}$ for the conditional efficiency estimators. The full sample $\mathcal{S}_n$ can be split so that the estimators in the two subsamples achieve the same rate of convergence by setting $n_1^{\kappa} = n_2^{\kappa/(r\kappa+1)}$. This gives $n_1 = n_2^{1/(r\kappa+1)}$. Value of $n_1,\ n_2$ are obtained by finding the root η_0 in $n - \eta - \eta^{1/(r\kappa+1)}$ and setting $n_2 = [\eta_0]$ and $n_1 = n - n_2$, where $[a]$ denotes the integer nearest a.

After splitting the sample, compute the estimators

$$\widehat{\mu}_{n_1} = n_1^{-1} \sum_{(X_i,Y_i)\in\mathcal{S}_{1,n_1}} \widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_{1,n_1}) \tag{4.38}$$

and

$$\widehat{\mu}_{c,n_{2,h}} = n_{2,h}^{-1} \sum_{(X_i,Y_i,Z_i)\in\mathcal{S}^*_{2,n_{2,h}}} \widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_{2,n_2}), \tag{4.39}$$

where as above in Section 4.2.2, $\mathcal{S}^*_{2,n_{2,h}}$ in (4.39), is a random subsample from $\mathcal{S}_{2,n_2}$ of size $n_{2,h} = \min(n_2, n_2 h^r)$. Consistent estimators of the variances are given in the two independent samples by

$$\widehat{\sigma}^2_{n_1} = n_1^{-1} \sum_{(X_i,Y_i)\in\mathcal{S}_{1,n_1}} \left(\widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_{1,n_1}) - \widehat{\mu}_{n_1}\right)^2 \tag{4.40}$$

and

$$\widehat{\sigma}^{2,h}_{c,n_2} = n_2^{-1} \sum_{(X_i,Y_i,Z_i)\in\mathcal{S}_{2,n_2}} \left(\widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_{2,n_2}) - \widehat{\mu}_{c,n_2}\right)^2 \tag{4.41}$$

(respectively), where the full (sub)sample $\mathcal{S}_{2,n_2}$ to estimate the variance $\sigma_c^{2,h}$ of the conditional efficiency measures.

[10] As observed by Hall et al. (2004), if Z is irrelevant in the production process (independent of (X, Y)), the optimal value of the bandwidth is infinity. This limiting case is more restrictive that the hypothesis to be tested here, but may arise in practice.

The estimators of bias corresponding to (4.31) for a single split of each subsample for the unconditional and conditional cases are given by

$$\widetilde{B}_{\kappa,n_1} = (2^\kappa - 1)^{-1} \left(\widehat{\mu}^*_{n_1/2} - \widehat{\mu}_{n_1} \right) \tag{4.42}$$

and

$$\widetilde{B}^c_{\kappa,n_{2,h}} = (2^\kappa - 1)^{-1} \left(\widehat{\mu}^*_{c,n_2/2} - \widehat{\mu}_{c,n_2} \right). \tag{4.43}$$

For the unconditional case in (4.42), $\widehat{\mu}^*_{n_1/2} = \left(\widehat{\mu}^1_{n_1/2} + \widehat{\mu}^2_{n_1/2} \right)/2$, and for $j \in \{1,\ 2\}$, $\widehat{\mu}^j_{n_1/2} = (n_1/2)^{-1} \sum_{(X_i,Y_i,Z_i) \in \mathcal{S}^{(j)}_{n_1/2}} \widehat{\lambda}(X_i, Y_i \mid \mathcal{S}^{(j)}_{n_1/2})$, where $\mathcal{S}^{(j)}_{n_1/2}$ is the jth part of a random split of the full (sub)sample $\mathcal{S}_{n_1}$. Details are given in Kneip et al. (2015a). For the conditional case in (4.43), $\widehat{\mu}^*_{c,n_2/2} = \left(\widehat{\mu}^1_{c,n_2/2} + \widehat{\mu}^2_{c,n_2/2} \right)/2$, and for $j \in \{1,\ 2\}$, $\widehat{\mu}^j_{c,n_2/2} = (n_2/2)^{-1} \sum_{(X_i,Y_i,Z_i) \in \mathcal{S}^{(j)}_{n_2/2}} \widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}^{(j)}_{n_2/2})$, where $\mathcal{S}^{(j)}_{n_2/2}$ is the jth part of a random split of the full (sub)sample $\mathcal{S}_{n_2}$. The bias estimates in (4.42)–(4.43) can then be averaged over K random splits of the two subsamples $\mathcal{S}_{n_1}$ and $\mathcal{S}_{n_2}$ to obtain bias estimates $\widehat{B}_{\kappa,n_1}$ for the unconditional case and $\widehat{B}^c_{\kappa,n_{2,h}}$ for the conditional case.

For small values of $(p+q)$ such that $\kappa \geq 1/3$ in the FDH case or $\kappa \geq 2/5$ when DEA estimators are used, Theorem 4.3 and Kneip et al. (2015a, Theorem 4.3) can be used to construct an asymptotically normal test statistic for testing the null hypothesis of separability. In particular, since our bias-corrected sample means are independent due to splitting the original sample into independent parts, and since two sequences of independent variables each with normal limiting distributions have a joint bivariate normal limiting distribution with independent marginals, if follows that for the values of $(p+q)$ given above

$$T_{1,n} = \frac{\left(\widehat{\mu}_{n_1} - \widehat{\mu}_{c,n_{2,h}} \right) - \left(\widehat{B}_{\kappa,n_1} - \widehat{B}^c_{\kappa,n_{2,h}} \right)}{\sqrt{\frac{\widehat{\sigma}^2_{n_1}}{n_1} + \frac{\widehat{\sigma}^{2,h}_{c,n_2}}{n_{2,h}}}} \xrightarrow{\mathcal{L}} N(0,1) \tag{4.44}$$

under the null. Alternatively, for $\kappa < 1/2$, similar reasoning with Theorem 4.4 and Kneip et al. (2015a, Theorem 4.4) leads to

$$T_{2,n} = \frac{\left(\widehat{\mu}_{n_{1,\kappa}} - \widehat{\mu}_{c,n_{2,h,\kappa}} \right) - \left(\widehat{B}_{\kappa,n_1} - \widehat{B}^c_{\kappa,n_{2,h}} \right)}{\sqrt{\frac{\widehat{\sigma}^2_{n_1}}{n_{1,\kappa}} + \frac{\widehat{\sigma}^{2,h}_{c,n_2}}{n_{2,h,\kappa}}}} \xrightarrow{\mathcal{L}} N(0,1) \tag{4.45}$$

under the null, where $n_{1,\kappa} = \lfloor n_1^{2\kappa} \rfloor$ with $\widehat{\mu}_{n_{1,\kappa}} = n^{-1}_{1,\kappa} \sum_{(X_i,Y_i) \in \mathcal{S}^*_{n_{1,\kappa}}} \widehat{\lambda}(X_i, Y_i \mid \mathcal{S}_{n_1})$, and $\mathcal{S}^*_{n_{1,\kappa}}$ is a random subsample of size $n_{1,\kappa}$ taken from $\mathcal{S}_{n_1}$ (see Kneip et al., 2015a for details). For the

conditional part, we have similarly and as described in the preceding section, $n_{2,h,\kappa} = [n_{2,h}^{2\kappa}]$, with $\widehat{\mu}_{c,n_{2,h,\kappa}} = n_{2,h,\kappa}^{-1} \sum_{(X_i,Y_i,Z_i)\in\mathcal{S}^*_{n_{2,h,\kappa}}} \widehat{\lambda}(X_i, Y_i \mid Z_i, \mathcal{S}_{n_2})$ where $\mathcal{S}^*_{n_{2,h,\kappa}}$ is a random subsample of size $n_{2,h,\kappa}$ from $\mathcal{S}_{n_2}$.

Given a random sample $\mathcal{S}_n$, one can compute values $\widehat{T}_{1,n}$ or $\widehat{T}_{2,n}$ depending on the value of $(p+q)$. From the discussion in Section 4.1, it is clear that a one-sided test is appropriate; hence the null should be rejected whenever null whenever $1-\Phi(\widehat{T}_{1,n})$ or $1-\Phi(\widehat{T}_{2,n})$ is less than the desired test size, e.g., .1, .05, or .01, where $\Phi(\cdot)$ denotes the standard normal distribution function.

4.4 Bootstrap Approximation

Under the null hypothesis of separability, the test statistics $T_{1,n}$ and $T_{2,n}$ in (4.44) and (4.45) are asymptotically pivotal as well as asymptotically normally distributed. It is well-known that bootstrap methods sometimes provide better performance than tests based on asymptotic normality, particularly when asymptotically pivotal statistics are available.

Given a sample $\mathcal{S}_n$, a bootstrap test of separability can be implemented by estimating a one-sided bootstrap confidence interval for one of the statistics in (4.44) or (4.45) and rejecting the null hypothesis if this estimated interval does not cover zero. The test is very fast from a computational viewpoint, although implementation requires re-ordering the computations leading to the bias corrections as discussed in Kneip et al. (2015b). Since the test statistics in (4.44) and (4.45) involve differences in sample means, a "naive" bootstrap can be used; i.e., once the original firm-specific efficiencies have been estimate, no new efficiency estimates have to be computed. See Kneip et al. (2015b) for details.

4.5 Bandwidth Optimization

As noted above, explicit expressions for the two components Δ_1 and Δ_2 of the estimation error in (4.9) are not available. Consequently, the best that can be done is to determine the order of optimal bandwidths by balancing the order of the two error terms yielding $h \propto n^{-1/(r+\kappa^{-1})}$ as explained earlier. Although the order by itself is of little help in applications, following the suggestion of Jeong et al. (2010) one can select optimal bandwidths for estimating the conditional distribution $H_{XY|Z}(x, y \mid z)$ by $\widehat{H}_{XY|Z}(x, y \mid z)$ given in (3.6). This can be accomplished using the least-squares cross-validation (LSCV) procedure de-

scribed by Li et al. (2013), smoothing only on the r conditioning variables in Z, and not the dependent variables (X, Y). Note that, as proved by Hall et al. (2004), if one component of Z is irrelevant, then the corresponding bandwidth obtained by LSCV will converge to infinity as $n \to \infty$; but for relevant components of Z, LSCV gives a bandwidth with optimal rate $h \propto n^{-1/(r+4)}$ for estimating $H_{XY|Z}(x, y \mid z)$.

Recall that if Z is relevant, the optimal bandwidths for estimating $\lambda(x, y \mid z)$ have a different order ($h \propto n^{-1/(r+\kappa^{-1})}$, as opposed to $h \propto n^{-1/(r+4)}$) due to the presence of the localizing bias. In practice, one can optimize bandwidths using LSCV, and then correct the resulting bandwidths by multiplying by the scaling factor $n^{1/(r+4)} n^{-1/(r+\kappa^{-1})} = n^{(\kappa^{-1}-4)/((r+4)(r+\kappa^{-1})}$ to obtain optimal bandwidths h for estimating $\lambda(x, y \mid z)$. To avoid numerical difficulties, for the jth element Z_i^j of Z_i, $j = 1, \ldots, r$, $i = 1, \ldots, n$, one should in practice bound the LSCV search between a small factor, say 0.01, times the normal reference rule bandwidth (i.e., $0.01 \times 1.06 \widehat{\sigma}_j n^{1/5}$, where $\widehat{\sigma}_j$ is the sample standard deviation of the observations Z_i^j, $j = 1, \ldots, n$) and 2 times the difference $(\max_i(Z_i^j) - \min_i(Z_i^j))$. If Z_i^j is irrelevant, LSCV will drive the jth element h_j of h to its upper bound; using a bounded kernel (e.g., the Epanechnikov kernel), no smoothing will be done in the jth dimension of Z when this happens. In such cases, there is no need to apply the scaling factor above to h_j.

5 Monte Carlo Evidence

We perform Monte Carlo experiments to gauge the performance of the separability test described in Section 4. In each experiment, we simulate $n \in \{100,\ 200,\ 1000\}$ observations with $r = 1$ and $(p, q) \in \{(1,1),\ (2,1),\ (2,2),\ (3,2),\ (3,3)\}$ so that $(p+q) \in \{2,\ 3,\ 4,\ 5,\ 6\}$. To generate an observation $(X_i, Y_, Z_i)$, we first simulate a draw $Z_i \sim N(0,1)$. Next, we generate a $(p+q)$-tuple $u = \left[u_p', u_q'\right]'$ uniformly distributed on a unit sphere centered at the origin in $\mathbb{R}^{p+q}$, where u_p and u_q are column vectors of length p and q, respectively. We then set $X = (1 - \text{abs}(u_p))$ and $Y = \text{abs}(u_q)\left(|Z'\beta|^{\delta}\lambda^{-1}\right)$ where Z is $(r \times 1)$, β is an $(r \times 1)$ vector of ones, $\lambda \geq 1$ is a scalar-valued pseudo random variable such that $(\lambda - 1) \sim N^{+}(0,1)$, $\text{abs}(a)$ denotes the vector containing the absolute values of elements of a vector a, and $\delta \in \{0,\ 0.1,\ \ldots,\ 0.9,\ 1.0,\ 1.5,\ 2.0\}$. When $\delta = 0$, Z plays no role and Assumption 2.1 (separability of Z) holds. Otherwise, when $Z > 0$, separability does not hold and instead

Assumption 2.2 holds.[11]

The results of our experiments are shown in Tables 1–4. In Tables 1-2, we test for separability using DEA estimators. In Table 1 we rely on the asymptotic normality of the test statistics in (4.44)–(4.45), while in Table 2 we use bootstrap methods as described in Section 4.4. In Tables 3–4 give results for the corresponding experiments using FDH estimators. Each table contains 3 groups of results corresponding to 100, 200, or 1,000 observations. Within each of these groups, we show, for various values of δ, rejection rates for the separability tests for nominal test sizes of .10, .05, and .01. The first row in each group corresponds to $\delta = 0$, where the null hypothesis is true; the remaining rows give rejection rates with increasing departures from the null, corresponding to increasing values of δ.

Overall, the results in Tables 1–4 confirm that the tests tend to reject the null hypothesis of separability at increasing rates both with increasing departure from the null and as sample size increases. Comparing the results in Tables 1–2 where DEA estimators are used with the corresponding results in Tables 3–4 where FDH estimators are used reveals that the tests have greater power when DEA estimators are used than when FDH estimators are used. Given the slower convergence rate of the FDH estimator, this is as expected.

Focusing on Tables 1–2, where DEA estimators are used, our experiments suggest that both the tests based on asymptotic normality as well as those based on bootstrap methods are conservative in the sense that they tend to reject the null at rates less than the nominal size when the null is true. for the cases where $(p+q) \leq 4$ and the statistic $T_{1,n}$ can be used, the tests based on asymptotic normality and on bootstrap methods provide very similar rejection rates for given values of δ, $(p+q)$, and nominal test size. However, for $(p+q) > 4$ where the statistic $T_{2,n}$ based on subsamples must me used, the bootstrap tests are seen to provide greater power than the tests based on asymptotic normality in many cases, particularly for

[11] Given a $(p+q)$-vector v of draws from the uniform distribution on $[0,1]$, $u = v(v'v)^{-1/2}$ is a vector of coordinates from a uniform distribution on the unit sphere in $\mathbb{R}^{p+q}$. Setting $Y = |u_q|$ amounts to reflecting any point that lies below one or more of the u_p axes around those axes. Similarly, $-|u_p|$ reflects around the u_q axes, but in negative directions; adding 1 shifts the resulting points to the positive orthant in $\mathbb{R}^{p+q}$. This amounts to generating uniform points on a unit sphere centered at $\left[\mathbf{1}'_p, \mathbf{0}'_q\right]'$, reflecting the points so that all lie on the part of the sphere in the unit hypercube with in the positive orthant with a corner at the origin. We then projecting points away from this "frontier" in the output directions. We use the massively parallel Palmetto Cluster at Clemson University for our experiments, generating pseudo-random uniform deviates using independent Mersenne Twister generators on each processor; see Matsumoto and Nishimura (2000) for details. Standard normal deviates are generated from uniform $(0,1)$ deviates using the transformation method.

larger values of δ. In a number of cases where $(p+q>4$, the power of the bootstrap tests is almost twice that of the corresponding tests based on asymptotic normality.

Similar remarks hold for the results in Tables 3–4 where FDH estimators are used. Using FDH estimators, the results in Table 3 suggest that the tests provide reasonable power when $(p+q) \leq 3$ and the statistic $T_{1,n}$ can be used. However, for larger dimensionality where $(p+q) > 3$, even with $n = 1,000$ and $\delta = 2$, the tests that rely on asymptotic normality have almost no power. This is not true for the bootstrap tests. The results in Table 4 reveal that for $(p+q) = 5$ or 6, the tests using FDH and bootstrap methods result in greater power than the corresponding tests using DEA and relying on asymptotic normality.[12]

The simulation results in Tables 1–4 show rejection rates when $r = 1$. One should expect the power of the tests to decrease with increasing values of r for given values p, q, and n. However, as the empirical example in the next section illustrates, one can perform marginal tests for each element of Z, ignoring the other elements when $r > 1$, before performing a joint test with all the elements of Z. If one is testing the separability condition in order to justify a second-stage regression, any rejection of Assumption 2.1 should rule out use of a second-stage regression. In other words, if one of the marginal tests rejects Assumption 2.1, there is no need to incur the computational expense of further marginal or joint tests, and any plans for a second-stage regression should be abandoned. Moreover, when confronted with evidence that an element of Z affects the shape of the frontier (as is the case whenever Assumption 2.1 is rejected), one should use *conditional* efficiency estimators instead of unconditional efficiency estimators.

6 Empirical Illustration using Bank Data

As a final exercise, we revisit the empirical examples provided by Simar and Wilson (2007), where estimated efficiency of U.S. Banks is regressed on some explanatory variables in a second-stage analysis. We start with the same data used by Simar and Wilson (2007), and consider both the subsample of 322 banks as well as the full sample of 6,955 banks examined by Simar and Wilson. The data include observations on 3 inputs (purchased funds, core deposits, and labor) and 4 outputs (consumer loans, business loans, real estate loans, and

[12] The favorable comparison does not carry over to $(p+q) = 4$. But in this case, $T_{1,n}$ can be used with DEA estimators, but $T_{2,n}$ must be used with FDH estimators.

securities held). The data also include observations for two continuous explanatory variables used by Simar and Wilson (2007), namely *SIZE* (i.e., the log of total assets, reflecting banks' sizes) and *DIVERSE* (i.e., a measure of diversity of banks' loan portfolios). Specific definitions of variables and other data details are given in Simar and Wilson (2007).

Our empirical examples here and in Simar and Wilson (2007) are motivated by Aly et al. (1990), who similarly estimate efficiency for a sample of 322 U.S. banks operating during the fourth quarter of 1986, and then attempt to explain variation in the first-stage efficiency estimates in a second-stage regression by regressing estimated efficiency on continuous variables reflecting bank size and loan-type diversity, as well as binary dummy variables reflecting membership in a multi-bank holding company and presence in a metropolitan statistical area. Whereas Aly et al. used the second-stage regression in an attempt to better understand the performance of U.S. banks' operations, Simar and Wilson carefully note that their second-stage regressions are only for purposes of illustrating the bootstrap methods for inference developed in their paper. As discussed above, and as noted by Simar and Wilson, such second-stage regressions can only be meaningful if the separability condition in Assumption 2.1 holds. Simar and Wilson also noted that this condition should be tested before employing a second-stage regression, but until now no such test has been available.

It is well-known that the distribution of U.S. bank sizes is heavily skewed to the right; in fact, the distribution of total assets of U.S. banks is roughly log-log-normal (e.g., see Wheelock and Wilson, 2001 for discussion). In order to use global bandwidths, as opposed to adaptive bandwidths (which would increase computational burden), we first eliminate very large banks and other outliers from the sub-sample of 322 observations as described by Florens et al. (2014) (who used the same data in an empirical illustration), leaving 303 observations for analysis. Similarly, we omit the largest 5-percent of banks from the full sample of 6,955 observations, leaving 6,607 observations. To further reduce computational burden, we exploit multicollinearity among the input and output variables by aggregating inputs into a single measure and also aggregating outputs into a single measure using eigensystem techniques employed by Florens et al. (2014) in their analysis of the subsample of our data and as described by Daraio and Simar (2007a, pp. 148–150). Due to the high degrees of correlation among the original input and output variables, little information is lost by this aggregation, while dimensionality is reduced from $(p+q) = 7$ to 2.

We test the separability condition (Assumption 2.1) using both the subsample of 303 observations and the "full" sample of 6,607 observations using DEA estimators in both input and output directions, with bandwidths optimized by least-squares cross-validation and then adjusted to obtain the optimal order as discussed above. We first test separability marginally by considering only *SIZE*, and then by considering only *DIVERSE* so that $r = 1$. We also perform joint tests ($r = 2$) considering both *SIZE* and *DIVERSE*.

Results for the tests for both samples are shown in Table 5. With the subsample, in the input orientation, both the asymptotic normal and the bootstrap tests yield p values smaller than 0.05 for *SIZE*; in the output orientation, the asymptotic normal test also yields a p-value less than 0.05, while the bootstrap test gives a p-value just larger than 0.10. There is no evidence against separability in the marginal tests with *DIVERSE*. The joint test yields one p-value less than 0.10, while two of the other three p-values are just larger than 0.10. With the sample of 6,607 observations, it becomes even more clear that *SIZE* violates separability, while there is no evidence that *DIVERSE* violates the condition.

Again, the second-stage regression in Simar and Wilson (2007) was used only to illustrate how one might apply the bootstrap methods proposed there. But, results from the second-stage regression in Aly et al. (1990), and those from similar exercises in other papers that have regressed estimates of bank efficiency on total assets, are rendered dubious and likely meaningless by the results obtained here.

7 Conclusions

We have provided a test of the separability condition described by Simar and Wilson (2007) on which many papers that regress estimated efficiency scores on some environmental variables depend. The condition is a restrictive, but can now be tested empirically. In our empirical example in Section 6, patterned after the application by Aly et al. (1990), we easily reject separability. This suggests that results of the second-stage regression in Aly et al. (1990) are meaningless, or at best very difficult to interpret. Furthermore, it raises the question of whether separability would similarly be rejected in the hundreds or thousands of papers that have regressed estimated efficiencies on environmental variables in a second stage regression. It is perhaps too much to expect that all of these studies be re-examined, but now that an easily-implemented test of separability has been made available, researchers

should employ the test before proceeding to a second-stage regression. Moreover, whenever the test rejects separability, the researcher should use conditional efficiency estimators instead of unconditional estimators in order to estimate distance to the relevant frontier (i.e., to the frontier of Ψ^z instead of the frontier of Ψ which has no particular economic meaning when separability does not hold).

Of course, failure to reject the null hypothesis of separability does not by itself imply that separability holds. As is always the case, our test can do only one of two things: it can either reject, or fail to reject the null hypothesis. Failure to reject might be due to other factors, such as insufficient data, or too many dimensions. In the later case, we have shown in our empirical example how dimensionality can be reduced before testing separability.

It should be remembered, as noted in Section 3, that the conditional efficiency estimators provide consistent estimates regardless of whether separability holds, but the unconditional efficiency estimators provide meaningfully consistent estimates if and only if separability holds. Of course, if separability holds, the unconditional estimators converge faster than their unconditional counterparts. But when testing separability, these points argue in favor of a *conservative* test. Whereas one might ordinarily test a null hypothesis at the 10, 5, or 1-percent level, here one might want to test at a 20, 30, 40, or even larger percentage level. The cost of a type-I error is slower convergence due to subsequent use of the conditional efficiency estimators, whereas the cost of a type-II error is statistical inefficiency due to subsequent inappropriate use of unconditional efficiency estimators. The cost of a type-II error here is arguably greater than the cost of a type-I error, which is the reverse of the usual situation in hypothesis testing. Here, however, reversing things by testing a null hypothesis of non-separability versus an alternative hypothesis of separability would result in a test with poor size and power properties, as separability is a much more restrictive condition than non-separability.

Appendix A Technical Details

The assumptions listed here impose regularity conditions on the data-generating process. The first assumption appears as Assumption 4 in Jeong et al. (2010).

Assumption A.1. *The joint density $f_{XYZ}(\cdot,\cdot,\cdot)$ of (X,Y,Z) is continuous on its support.*

The next assumptions are needed to establish results for the moments of the conditional FDH and DEA estimators in Section 4.2.1. The assumptions here are conditional analogs of Assumptions 3.1–3.4 and 3.6 (respectively) in Kneip et al. (2015a). Assumption A.2, part (iii) and Assumption A.3, part(iii) appear as Assumption 5 in Jeong et al. (2010).

Assumption A.2. *For all $z \in \mathcal{Z}$, (i) the conditional density $f_{XY|Z}(\cdot,\cdot \mid z)$ of $(X,Y) \mid Z = z$ exists and has support $\mathcal{D}^z \subset \Psi^z$; (ii) $f_{XY|Z}(\cdot,\cdot \mid z)$ is continuously differentiable on $\mathcal{D}^z$; and (iii) $f^h_{XY|Z}(\cdot,\cdot \mid z)$ converges to $f_{XY|Z}(\cdot,\cdot \mid z)$ as $h \to 0$.*

Assumption A.3. *(i) $\mathcal{D}^{z*} := \{(x, \lambda(x,y \mid z)y) \mid (x,y) \in \mathcal{D}^z\} \subset \mathcal{D}^z$; (ii) $\mathcal{D}^{z*}$ is compact; and (iii) $f_{XY|Z}(x, \lambda(x,y \mid z)y \mid z) > 0$ for all $(x,y) \in \mathcal{D}^z$.*

Assumption A.4. *For any $z \in \mathcal{Z}$, $\mathcal{D}^z$ is almost strictly convex; i.e., for any (x,y), $(\widetilde{x},\widetilde{y}) \in \mathcal{D}^z$ with $\left(\frac{x}{||x||}, y\right) \neq \left(\frac{\widetilde{x}}{||\widetilde{x}||}, \widetilde{y}\right)$, the set $\{(x^*,y^*) \mid (x^*,y^*) = (x,y) + \alpha((\widetilde{x},\widetilde{y})) \text{ for some } \alpha \in (0,1)\}$ is a subset of the interior of $\mathcal{D}^z$.*

Assumption A.5. *For all $z \in \mathcal{Z}$, (i) $\lambda(x,y \mid z)$ is twice continuously differentiable on $\mathcal{D}^z$; and (ii) all the first-order partial derivatives of $\lambda(x,y \mid z)$ with respect to x and y are nonzero at any point $(x,y) \in \mathcal{D}^z$.*

Assumption A.6. *For any $z \in \mathcal{Z}$, $\lambda(x,y \mid z)$ is three times continuously differentiable with respect to x and y on $\mathcal{D}^z$.*

When the conditional FDH estimator is used, Assumption A.5 is needed; when the conditional DEA estimator is used, this is replaced by the stronger Assumption A.6.

Note that under the separability condition in Assumption 2.1, the assumptions here reduce to the corresponding assumptions in Kneip et al. (2015a) due to the discussion in Section 2.

Appendix B Discrete Environmental Variables

In applied work, it is often the case that researchers include binary or categorical variables in second-stage regressions of estimated efficiency on environmental variables. All of the results obtained in the main part of this paper assume Z is continuous. However, in order for second-stage regressions to estimate any useful, meaningful feature, the separability condition in Assumption 2.1 must also hold with respect to discrete environmental variables.

Testing the separability condition in the case of discrete variables can be done using results and ideas from Kneip et al. (2015b), where a test of equivalent mean efficiency across two groups of producers is developed. To illustrate, suppose $r = 1$ and Z is a binary dummy variable. To test separability, first shuffle the observations, and then divide into two groups of size $n_1 = \lfloor n/2 \rfloor$ and $n_2 = n - n_1$. Apply the unconditional efficiency estimator to group 1. For group 2, a conditional efficiency estimator is needed, but since Z is discrete, there is no smoothing to be done.[13] Since Z is binary, there are only two sets Ψ^z. Hence, in the second group, divide observations into two sub-groups according to whether $Z = 0$ or $Z = 1$; observations in each sub-group, estimate efficiency using the same *unconditional* efficiency estimator used with group 1, ignoring observations in the other group. This will yield a set of n_2 *conditional* efficiency estimates since the n_2 observations have been divided into sub-groups.

Note that the conditional estimates from group 2 have the usual convergence rate of the unconditional efficiency estimator since no bandwidth is involved since Z is discrete. One can now apply the difference-in-means test as described in Kneip et al. (2015b), taking care to compute the bias-correction terms for group 2 separately and independently for observations in the subgroup (of group 2) where $Z = 0$ and the subgroup where $Z = 1$. This will necessitate splitting each sub-group (of group 2) to compute the generalized jackknife estimates of bias for observations in each sub-group. See Kneip et al. (2015b) for details.

[13] The problem here is rather different from the problem of nonparametric estimation of regressions or densities, where one can smooth across discrete categories of data using the methods discussed by Li and Racine (2007). Here, we are interested in boundaries of support, as opposed to densities or conditional mean functions.

References

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

Bădin, L., C. Daraio, and L. Simar (2010), Optimal bandwidth selection for conditional efficiency measures: A data-driven approach, *European Journal of Operational Research* 201, 633–664.

— (2012), How to measure the impact of environmental factors in a nonparametric production model, *European Journal of Operational Research* 223, 818–833.

— (2015), Bandwidth selection issues for nonparametric conditional efficiency. Working Paper, Institut de Statistique Biostatistique et Sciences Actuarielles, Université Catholique de Louvain, Louvain-la-Neuve, Belgium.

Cazals, C., J. P. Florens, and L. Simar (2002), Nonparametric frontier estimation: A robust approach, *Journal of Econometrics* 106, 1–25.

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

Daouia, A., L. Simar, and P. W. Wilson (2015), Measuring firm performance using nonparametric quantile-type distances, *Econometric Reviews* Forthcoming.

Daraio, C. and L. Simar (2005), Introducing environmental variables in nonparametric frontier models: A probabilistic approach, *Journal of Productivity Analysis* 24, 93–121.

— (2007a), *Advanced Robust and Nonparametric Methods in Efficiency Analysis*, New York: Springer Science+Business Media, LLC.

— (2007b), Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach, *Journal of Productivity Analysis* 28, 13–32.

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

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

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

Florens, J. P., L. Simar, and I. Van Keilegom (2014), Frontier estimation in nonparametric location-scale models, *Journal of Econometrics* 178, 456–470.

Hall, P., J. S. Racine, and Q. Li (2004), Cross-validation and the estimation of conditional probability densities, *Journal of the American Statistical Association* 99, 1015–1026.

Jeong, S. O., B. U. Park, and L. Simar (2010), Nonparametric conditional efficiency measures: asymptotic properties, *Annals of Operations Research* 173, 105–122.

Jeong, S. O. and L. Simar (2006), Linearly interpolated FDH efficiency score for nonconvex frontiers, *Journal of Multivariate Analysis* 97, 2141–2161.

Kneip, A., B. Park, and L. Simar (1998), A note on the convergence of nonparametric DEA efficiency measures, *Econometric Theory* 14, 783–793.

Kneip, A., L. Simar, and P. W. Wilson (2008), Asymptotics and consistent bootstraps for DEA estimators in non-parametric frontier models, *Econometric Theory* 24, 1663–1697.

— (2011), A computationally efficient, consistent bootstrap for inference with non-parametric DEA estimators, *Computational Economics* 38, 483–515.

— (2015a), When bias kills the variance: Central limit theorems for DEA and FDH efficiency scores, *Econometric Theory* 31, 394–422.

— (2015b), Testing hypotheses in nonparametric models of production, *Journal of Business and Economic Statistics* Forthcoming.

Li, Q., J. Lin, and J. S. Racine (2013), Optimal bandwidth selection for nonparametric conditional distribution and quantile functions, *Journal of Business and Economic Statistics* 31, 57–65.

Li, Q. and J. Racine (2007), *Nonparametric Econometrics*, Princeton, NJ: Princeton University Press.

Matsumoto, M. and T. Nishimura (2000), Dynamic creation of pseudorandom number generators, in H. Niederreiter and J. Spanier, eds., *Monte Carlo and Quasi Monte Carlo Methods 1998*, Berlin: Springer-Verlag, pp. 56–69.

Park, B. U., L. Simar, and C. Weiner (2000), FDH efficiency scores from a stochastic point of view, *Econometric Theory* 16, 855–877.

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

Simar, L., A. Vanhems, and I. Van Keilegom (2015), Unobserved heterogeneity and endogeneity in nonparametric frontier estimation, *Journal of Econometrics* Forthcoming.

Simar, L. and P. W. Wilson (2007), Estimation and inference in two-stage, semi-parametric models of productive efficiency, *Journal of Econometrics* 136, 31–64.

— (2011), Two-Stage DEA: Caveat emptor, *Journal of Productivity Analysis* 36, 205–218.

— (2013), Estimation and inference in nonparametric frontier models: Recent developments and perspectives, *Foundations and Trends in Econometrics* 5, 183–337.

— (2015), Statistical approaches for non-parametric frontier models: A guided tour, *International Statistical Review* 83, 77–110.

Wheelock, D. C. and P. W. Wilson (2001), New evidence on returns to scale and product mix among U.S. commercial banks, *Journal of Monetary Economics* 47, 653–674.

Table 1: Rejection Rates for Separability Test using DEA, Asymptotic Normality ($r = 1$)

		$T_{1,n}$									$T_{2,n}$					
		$p=1,\ q=1$			$p=2,\ q=1$			$p=2,\ q=2$			$p=3,\ q=2$			$p=3,\ q=3$		
n	λ_2	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
100	0.00	0.055	0.028	0.012	0.047	0.030	0.009	0.054	0.037	0.011	0.044	0.029	0.008	0.046	0.026	0.008
	0.10	0.074	0.038	0.008	0.052	0.033	0.011	0.048	0.033	0.016	0.042	0.025	0.010	0.043	0.028	0.009
	0.20	0.124	0.083	0.034	0.094	0.069	0.027	0.084	0.053	0.019	0.043	0.027	0.012	0.053	0.041	0.016
	0.30	0.193	0.135	0.040	0.123	0.080	0.033	0.098	0.066	0.027	0.070	0.044	0.020	0.077	0.046	0.015
	0.40	0.311	0.241	0.108	0.149	0.105	0.054	0.132	0.096	0.049	0.071	0.049	0.017	0.083	0.056	0.024
	0.50	0.383	0.291	0.160	0.213	0.149	0.059	0.204	0.132	0.063	0.083	0.054	0.030	0.080	0.063	0.029
	0.60	0.475	0.364	0.179	0.225	0.166	0.072	0.203	0.163	0.078	0.104	0.074	0.031	0.098	0.069	0.038
	0.70	0.509	0.387	0.197	0.269	0.193	0.096	0.261	0.204	0.109	0.097	0.071	0.031	0.087	0.066	0.031
	0.80	0.539	0.416	0.218	0.317	0.242	0.114	0.296	0.225	0.130	0.115	0.078	0.025	0.117	0.083	0.044
	0.90	0.561	0.421	0.202	0.323	0.238	0.113	0.321	0.266	0.158	0.128	0.084	0.035	0.101	0.077	0.040
	1.00	0.583	0.455	0.237	0.329	0.253	0.126	0.336	0.278	0.160	0.142	0.108	0.054	0.122	0.090	0.052
	1.50	0.639	0.495	0.225	0.395	0.321	0.177	0.393	0.340	0.198	0.145	0.109	0.064	0.137	0.098	0.046
	2.00	0.738	0.588	0.266	0.446	0.396	0.197	0.476	0.417	0.235	0.173	0.116	0.065	0.154	0.111	0.066
200	0.00	0.073	0.039	0.017	0.035	0.018	0.003	0.045	0.032	0.011	0.030	0.013	0.003	0.040	0.019	0.005
	0.10	0.075	0.044	0.019	0.036	0.025	0.009	0.074	0.045	0.019	0.028	0.016	0.003	0.034	0.016	0.002
	0.20	0.187	0.133	0.063	0.099	0.073	0.027	0.093	0.066	0.027	0.039	0.022	0.010	0.037	0.018	0.005
	0.30	0.368	0.290	0.146	0.178	0.129	0.070	0.164	0.121	0.059	0.051	0.030	0.012	0.069	0.040	0.011
	0.40	0.541	0.429	0.246	0.296	0.226	0.126	0.273	0.194	0.099	0.085	0.051	0.016	0.053	0.035	0.012
	0.50	0.630	0.521	0.314	0.338	0.261	0.141	0.373	0.307	0.179	0.127	0.078	0.031	0.089	0.059	0.025
	0.60	0.694	0.581	0.352	0.444	0.336	0.204	0.381	0.318	0.182	0.161	0.107	0.047	0.106	0.069	0.026
	0.70	0.719	0.602	0.369	0.472	0.366	0.220	0.450	0.361	0.214	0.149	0.108	0.041	0.106	0.073	0.037
	0.80	0.706	0.603	0.364	0.541	0.429	0.250	0.504	0.425	0.255	0.176	0.123	0.053	0.138	0.100	0.057
	0.90	0.733	0.608	0.377	0.533	0.435	0.245	0.488	0.410	0.260	0.174	0.124	0.066	0.125	0.098	0.051
	1.00	0.750	0.615	0.351	0.530	0.422	0.243	0.532	0.439	0.260	0.169	0.125	0.064	0.136	0.099	0.062
	1.50	0.765	0.593	0.296	0.609	0.510	0.267	0.567	0.490	0.289	0.206	0.152	0.097	0.136	0.102	0.055
	2.00	0.913	0.732	0.358	0.727	0.615	0.339	0.713	0.652	0.402	0.282	0.220	0.137	0.185	0.140	0.101

Table 1: Rejection Rates for Separability Test using DEA, Asymptotic Normality ($r = 1$, continued)

n	λ_2	$T_{1,n}$ — $p=1,\ q=1$ — .10	.05	.01	$T_{1,n}$ — $p=2,\ q=1$ — .10	.05	.01	$T_{1,n}$ — $p=2,\ q=2$ — .10	.05	.01	$T_{2,n}$ — $p=3,\ q=2$ — .10	.05	.01	$T_{2,n}$ — $p=3,\ q=3$ — .10	.05	.01
1000	0.00	0.065	0.041	0.004	0.021	0.010	0.001	0.040	0.024	0.006	0.025	0.012	0.007	0.030	0.017	0.010
	0.10	0.156	0.120	0.048	0.057	0.037	0.016	0.074	0.039	0.014	0.033	0.019	0.003	0.047	0.029	0.008
	0.20	0.515	0.449	0.317	0.323	0.254	0.175	0.330	0.268	0.164	0.091	0.057	0.017	0.080	0.050	0.021
	0.30	0.737	0.707	0.584	0.632	0.567	0.441	0.618	0.564	0.441	0.249	0.173	0.083	0.189	0.124	0.041
	0.40	0.789	0.778	0.711	0.755	0.719	0.621	0.761	0.712	0.621	0.373	0.276	0.134	0.268	0.189	0.092
	0.50	0.822	0.812	0.757	0.791	0.759	0.688	0.809	0.790	0.706	0.418	0.326	0.176	0.287	0.199	0.100
	0.60	0.836	0.813	0.744	0.829	0.807	0.701	0.825	0.798	0.714	0.479	0.381	0.218	0.313	0.229	0.123
	0.70	0.856	0.827	0.728	0.802	0.777	0.657	0.803	0.764	0.667	0.455	0.354	0.217	0.287	0.204	0.101
	0.80	0.833	0.785	0.679	0.780	0.733	0.617	0.790	0.744	0.614	0.419	0.320	0.186	0.253	0.182	0.097
	0.90	0.815	0.769	0.635	0.777	0.723	0.583	0.786	0.725	0.582	0.352	0.271	0.156	0.242	0.163	0.093
	1.00	0.806	0.741	0.576	0.784	0.699	0.538	0.760	0.697	0.524	0.344	0.253	0.157	0.208	0.146	0.093
	1.50	0.819	0.689	0.481	0.842	0.749	0.507	0.854	0.800	0.556	0.323	0.247	0.160	0.193	0.145	0.100
	2.00	0.796	0.685	0.385	0.851	0.810	0.558	0.869	0.857	0.710	0.432	0.319	0.208	0.297	0.220	0.142

Table 2: Rejection Rates for Separability Test using DEA, Bootstrap ($r = 1$)

		$T_{1,n}$									$T_{2,n}$					
		$p=1,\ q=1$			$p=2,\ q=1$			$p=2,\ q=2$			$p=3,\ q=2$			$p=3,\ q=3$		
n	λ_2	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
100	0.00	0.056	0.030	0.006	0.050	0.029	0.008	0.048	0.033	0.007	0.021	0.010	0.003	0.027	0.014	0.005
	0.10	0.070	0.027	0.002	0.050	0.021	0.007	0.045	0.027	0.011	0.020	0.006	0.002	0.036	0.017	0.004
	0.20	0.122	0.064	0.016	0.100	0.059	0.018	0.084	0.051	0.013	0.029	0.017	0.005	0.041	0.023	0.005
	0.30	0.176	0.099	0.023	0.119	0.077	0.022	0.095	0.063	0.024	0.064	0.035	0.009	0.047	0.027	0.012
	0.40	0.281	0.192	0.063	0.142	0.102	0.045	0.130	0.096	0.048	0.076	0.050	0.026	0.081	0.050	0.023
	0.50	0.358	0.250	0.098	0.200	0.140	0.052	0.192	0.128	0.063	0.100	0.069	0.036	0.098	0.072	0.037
	0.60	0.441	0.310	0.119	0.214	0.156	0.072	0.206	0.165	0.088	0.143	0.108	0.059	0.139	0.097	0.057
	0.70	0.477	0.340	0.138	0.260	0.181	0.091	0.250	0.204	0.125	0.162	0.116	0.056	0.152	0.113	0.068
	0.80	0.520	0.365	0.161	0.310	0.223	0.110	0.288	0.224	0.146	0.157	0.116	0.075	0.164	0.115	0.072
	0.90	0.544	0.364	0.151	0.319	0.231	0.119	0.320	0.272	0.183	0.197	0.151	0.104	0.188	0.148	0.100
	1.00	0.567	0.414	0.184	0.323	0.243	0.134	0.333	0.282	0.187	0.201	0.152	0.097	0.206	0.166	0.122
	1.50	0.633	0.433	0.180	0.392	0.319	0.187	0.392	0.340	0.242	0.246	0.220	0.165	0.282	0.242	0.201
	2.00	0.736	0.533	0.221	0.451	0.394	0.233	0.469	0.434	0.304	0.280	0.250	0.191	0.293	0.260	0.218
200	0.00	0.079	0.041	0.010	0.036	0.019	0.004	0.050	0.027	0.008	0.003	0.000	0.000	0.006	0.002	0.000
	0.10	0.069	0.043	0.008	0.038	0.021	0.006	0.068	0.041	0.016	0.012	0.002	0.001	0.007	0.005	0.000
	0.20	0.162	0.103	0.044	0.088	0.066	0.023	0.084	0.060	0.026	0.022	0.008	0.002	0.019	0.008	0.002
	0.30	0.323	0.240	0.115	0.168	0.121	0.057	0.157	0.110	0.049	0.035	0.014	0.002	0.033	0.015	0.005
	0.40	0.486	0.363	0.178	0.269	0.219	0.114	0.256	0.189	0.106	0.067	0.039	0.017	0.064	0.039	0.013
	0.50	0.589	0.466	0.242	0.325	0.244	0.138	0.366	0.304	0.189	0.133	0.087	0.033	0.116	0.078	0.038
	0.60	0.646	0.523	0.295	0.428	0.327	0.201	0.374	0.312	0.204	0.178	0.133	0.061	0.153	0.095	0.045
	0.70	0.685	0.566	0.329	0.460	0.367	0.223	0.442	0.362	0.240	0.214	0.170	0.087	0.182	0.143	0.075
	0.80	0.698	0.566	0.331	0.527	0.432	0.260	0.495	0.427	0.286	0.255	0.199	0.117	0.229	0.176	0.106
	0.90	0.728	0.574	0.327	0.522	0.432	0.271	0.481	0.410	0.299	0.274	0.218	0.141	0.249	0.205	0.142
	1.00	0.740	0.589	0.328	0.529	0.427	0.276	0.520	0.460	0.309	0.282	0.231	0.161	0.275	0.226	0.143
	1.50	0.764	0.570	0.293	0.611	0.526	0.325	0.562	0.511	0.365	0.389	0.347	0.249	0.362	0.314	0.254
	2.00	0.919	0.710	0.352	0.725	0.648	0.412	0.710	0.665	0.512	0.481	0.436	0.322	0.466	0.429	0.356

Table 2: Rejection Rates for Separability Test using DEA, Bootstrap ($r = 1$, continued)

n	λ_2	$T_{1,n}$ — $p=1,\ q=1$ — .10	.05	.01	$T_{1,n}$ — $p=2,\ q=1$ — .10	.05	.01	$T_{1,n}$ — $p=2,\ q=2$ — .10	.05	.01	$T_{2,n}$ — $p=3,\ q=2$ — .10	.05	.01	$T_{2,n}$ — $p=3,\ q=3$ — .10	.05	.01
1000	0.00	0.074	0.039	0.007	0.021	0.012	0.001	0.044	0.022	0.004	0.000	0.000	0.000	0.001	0.000	0.000
	0.10	0.140	0.099	0.034	0.057	0.033	0.014	0.069	0.034	0.011	0.003	0.000	0.000	0.000	0.000	0.000
	0.20	0.444	0.371	0.241	0.288	0.221	0.131	0.297	0.232	0.137	0.041	0.018	0.003	0.019	0.007	0.003
	0.30	0.698	0.648	0.514	0.587	0.522	0.390	0.588	0.517	0.410	0.201	0.116	0.037	0.101	0.048	0.018
	0.40	0.773	0.748	0.677	0.736	0.692	0.582	0.731	0.678	0.605	0.400	0.286	0.120	0.242	0.141	0.046
	0.50	0.815	0.796	0.749	0.778	0.745	0.682	0.802	0.772	0.695	0.533	0.398	0.204	0.369	0.251	0.110
	0.60	0.829	0.807	0.731	0.818	0.795	0.704	0.812	0.791	0.728	0.615	0.511	0.298	0.461	0.329	0.173
	0.70	0.852	0.824	0.730	0.789	0.766	0.677	0.784	0.752	0.681	0.621	0.519	0.337	0.481	0.374	0.203
	0.80	0.834	0.779	0.690	0.770	0.724	0.644	0.776	0.737	0.650	0.614	0.523	0.356	0.488	0.390	0.244
	0.90	0.819	0.767	0.653	0.776	0.713	0.621	0.775	0.729	0.622	0.585	0.516	0.367	0.507	0.404	0.284
	1.00	0.808	0.734	0.605	0.769	0.718	0.594	0.744	0.710	0.589	0.606	0.529	0.385	0.523	0.432	0.323
	1.50	0.837	0.700	0.520	0.841	0.782	0.609	0.853	0.827	0.681	0.731	0.655	0.522	0.676	0.588	0.474
	2.00	0.814	0.694	0.441	0.855	0.837	0.678	0.873	0.864	0.822	0.854	0.797	0.648	0.833	0.750	0.619

Table 3: Rejection Rates for Separability Test using FDH, Asymptotic Normality ($r = 1$)

n	λ_2	$T_{1,n}$						$T_{2,n}$								
		$p=1,\ q=1$			$p=2,\ q=1$			$p=2,\ q=2$			$p=3,\ q=2$			$p=3,\ q=3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
100	0.00	0.062	0.034	0.009	0.075	0.050	0.024	0.077	0.045	0.011	0.085	0.064	0.041	0.086	0.064	0.035
	0.10	0.077	0.057	0.021	0.056	0.040	0.019	0.068	0.047	0.022	0.080	0.058	0.031	0.104	0.075	0.034
	0.20	0.158	0.118	0.048	0.090	0.066	0.034	0.080	0.054	0.019	0.083	0.061	0.030	0.121	0.083	0.047
	0.30	0.246	0.185	0.092	0.114	0.085	0.033	0.090	0.061	0.028	0.102	0.068	0.037	0.109	0.075	0.046
	0.40	0.318	0.253	0.132	0.133	0.105	0.058	0.102	0.066	0.030	0.093	0.060	0.031	0.116	0.085	0.047
	0.50	0.435	0.346	0.197	0.200	0.165	0.090	0.097	0.073	0.037	0.079	0.060	0.024	0.097	0.064	0.040
	0.60	0.485	0.389	0.213	0.231	0.174	0.098	0.120	0.082	0.042	0.091	0.071	0.052	0.089	0.065	0.039
	0.70	0.538	0.427	0.255	0.255	0.204	0.126	0.125	0.082	0.045	0.094	0.075	0.039	0.086	0.052	0.028
	0.80	0.576	0.462	0.268	0.259	0.211	0.131	0.140	0.092	0.050	0.080	0.056	0.035	0.085	0.061	0.036
	0.90	0.579	0.470	0.274	0.279	0.230	0.141	0.137	0.100	0.061	0.096	0.068	0.036	0.093	0.058	0.035
	1.00	0.596	0.494	0.280	0.291	0.245	0.147	0.138	0.098	0.066	0.089	0.054	0.032	0.105	0.072	0.039
	1.50	0.620	0.510	0.275	0.361	0.317	0.221	0.142	0.113	0.078	0.117	0.081	0.061	0.078	0.043	0.026
	2.00	0.611	0.517	0.243	0.396	0.363	0.240	0.140	0.098	0.077	0.093	0.056	0.038	0.081	0.048	0.040
200	0.00	0.047	0.029	0.011	0.031	0.024	0.013	0.032	0.019	0.008	0.052	0.032	0.011	0.078	0.053	0.029
	0.10	0.074	0.037	0.019	0.043	0.030	0.013	0.057	0.027	0.007	0.051	0.029	0.008	0.064	0.043	0.015
	0.20	0.209	0.153	0.082	0.070	0.052	0.023	0.064	0.041	0.016	0.058	0.032	0.014	0.072	0.042	0.026
	0.30	0.412	0.327	0.166	0.149	0.106	0.046	0.094	0.059	0.024	0.062	0.040	0.024	0.071	0.050	0.021
	0.40	0.568	0.464	0.281	0.188	0.149	0.077	0.111	0.072	0.033	0.061	0.045	0.022	0.045	0.029	0.017
	0.50	0.681	0.587	0.389	0.264	0.206	0.110	0.127	0.081	0.043	0.059	0.037	0.017	0.063	0.044	0.021
	0.60	0.712	0.610	0.432	0.359	0.294	0.175	0.153	0.110	0.066	0.072	0.058	0.023	0.061	0.044	0.023
	0.70	0.728	0.634	0.409	0.387	0.317	0.202	0.163	0.118	0.067	0.074	0.058	0.032	0.053	0.039	0.021
	0.80	0.747	0.645	0.428	0.393	0.341	0.207	0.170	0.111	0.064	0.069	0.046	0.030	0.060	0.045	0.028
	0.90	0.724	0.641	0.408	0.430	0.355	0.215	0.156	0.119	0.064	0.074	0.050	0.033	0.056	0.037	0.028
	1.00	0.720	0.631	0.406	0.456	0.382	0.243	0.139	0.104	0.061	0.071	0.056	0.037	0.058	0.038	0.021
	1.50	0.683	0.569	0.328	0.423	0.377	0.250	0.130	0.104	0.070	0.062	0.044	0.032	0.058	0.039	0.032
	2.00	0.701	0.562	0.270	0.493	0.456	0.290	0.124	0.102	0.080	0.057	0.044	0.038	0.066	0.048	0.036

Table 3: Rejection Rates for Separability Test using FDH, Asymptotic Normality ($r = 1$, continued)

n	λ_2	$T_{1,n}$						$T_{2,n}$								
		$p=1,\ q=1$			$p=2,\ q=1$			$p=2,\ q=2$			$p=3,\ q=2$			$p=3,\ q=3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
1000	0.00	0.035	0.014	0.003	0.006	0.003	0.001	0.017	0.008	0.000	0.024	0.013	0.001	0.033	0.012	0.003
	0.10	0.133	0.090	0.029	0.022	0.019	0.006	0.020	0.010	0.002	0.022	0.012	0.003	0.037	0.023	0.004
	0.20	0.544	0.488	0.350	0.215	0.162	0.105	0.085	0.058	0.020	0.041	0.026	0.008	0.046	0.031	0.006
	0.30	0.756	0.730	0.634	0.488	0.429	0.318	0.174	0.122	0.050	0.059	0.028	0.012	0.055	0.032	0.019
	0.40	0.802	0.792	0.745	0.679	0.623	0.496	0.226	0.153	0.074	0.076	0.054	0.022	0.066	0.046	0.026
	0.50	0.838	0.827	0.789	0.761	0.710	0.608	0.271	0.172	0.077	0.083	0.055	0.028	0.072	0.044	0.012
	0.60	0.829	0.817	0.767	0.792	0.751	0.660	0.297	0.201	0.098	0.098	0.066	0.031	0.074	0.054	0.033
	0.70	0.863	0.834	0.748	0.766	0.730	0.627	0.227	0.158	0.083	0.099	0.064	0.032	0.037	0.028	0.017
	0.80	0.830	0.794	0.709	0.753	0.724	0.597	0.223	0.151	0.087	0.084	0.061	0.034	0.053	0.035	0.018
	0.90	0.806	0.763	0.646	0.732	0.686	0.576	0.213	0.140	0.070	0.069	0.051	0.031	0.046	0.036	0.023
	1.00	0.807	0.757	0.609	0.754	0.708	0.568	0.179	0.127	0.070	0.065	0.051	0.023	0.051	0.036	0.017
	1.50	0.717	0.608	0.393	0.632	0.568	0.384	0.108	0.080	0.056	0.035	0.026	0.016	0.035	0.030	0.025
	2.00	0.649	0.514	0.243	0.575	0.500	0.304	0.068	0.058	0.046	0.041	0.031	0.024	0.019	0.016	0.013

Table 4: Rejection Rates for Separability Test using FDH, Bootstrap ($r = 1$)

n	λ_2	$T_{1,n}$						$T_{2,n}$								
		$p=1,\ q=1$			$p=2,\ q=1$			$p=2,\ q=2$			$p=3,\ q=2$			$p=3,\ q=3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
100	0.00	0.056	0.024	0.006	0.069	0.048	0.020	0.072	0.035	0.018	0.182	0.142	0.103	0.210	0.195	0.139
	0.10	0.073	0.040	0.011	0.056	0.039	0.020	0.063	0.041	0.024	0.173	0.141	0.102	0.204	0.184	0.135
	0.20	0.143	0.089	0.034	0.087	0.065	0.033	0.094	0.063	0.032	0.176	0.143	0.105	0.207	0.192	0.133
	0.30	0.232	0.162	0.070	0.117	0.084	0.033	0.110	0.075	0.046	0.182	0.136	0.106	0.208	0.190	0.124
	0.40	0.296	0.226	0.107	0.132	0.110	0.071	0.150	0.102	0.061	0.172	0.135	0.102	0.201	0.178	0.123
	0.50	0.406	0.319	0.181	0.201	0.169	0.106	0.158	0.108	0.081	0.184	0.146	0.099	0.194	0.171	0.117
	0.60	0.464	0.362	0.195	0.229	0.185	0.123	0.178	0.113	0.076	0.201	0.156	0.109	0.204	0.181	0.131
	0.70	0.522	0.407	0.244	0.258	0.218	0.149	0.210	0.165	0.107	0.199	0.159	0.113	0.226	0.197	0.147
	0.80	0.562	0.460	0.270	0.264	0.226	0.152	0.253	0.190	0.125	0.202	0.152	0.111	0.230	0.201	0.146
	0.90	0.566	0.458	0.276	0.284	0.247	0.165	0.277	0.214	0.156	0.228	0.192	0.140	0.236	0.218	0.161
	1.00	0.587	0.475	0.285	0.295	0.258	0.181	0.292	0.223	0.160	0.197	0.169	0.127	0.238	0.221	0.155
	1.50	0.618	0.512	0.299	0.361	0.328	0.263	0.371	0.310	0.268	0.293	0.252	0.212	0.248	0.239	0.196
	2.00	0.612	0.524	0.291	0.398	0.376	0.305	0.408	0.367	0.335	0.295	0.268	0.241	0.318	0.305	0.239
200	0.00	0.049	0.030	0.010	0.033	0.022	0.012	0.023	0.010	0.003	0.076	0.054	0.034	0.096	0.066	0.039
	0.10	0.067	0.034	0.010	0.042	0.030	0.011	0.022	0.007	0.002	0.056	0.038	0.024	0.100	0.082	0.052
	0.20	0.189	0.133	0.060	0.064	0.047	0.025	0.039	0.017	0.007	0.061	0.045	0.029	0.096	0.065	0.049
	0.30	0.385	0.287	0.139	0.140	0.106	0.052	0.077	0.044	0.017	0.088	0.063	0.035	0.097	0.070	0.044
	0.40	0.536	0.425	0.266	0.189	0.149	0.087	0.141	0.081	0.039	0.108	0.075	0.047	0.109	0.083	0.059
	0.50	0.650	0.549	0.370	0.263	0.202	0.136	0.165	0.110	0.053	0.121	0.098	0.063	0.132	0.108	0.069
	0.60	0.688	0.600	0.405	0.365	0.300	0.205	0.237	0.169	0.091	0.142	0.105	0.078	0.165	0.135	0.093
	0.70	0.714	0.621	0.412	0.384	0.328	0.238	0.293	0.222	0.126	0.177	0.135	0.095	0.165	0.127	0.099
	0.80	0.728	0.635	0.450	0.398	0.350	0.251	0.301	0.236	0.157	0.203	0.157	0.118	0.190	0.161	0.114
	0.90	0.708	0.622	0.431	0.432	0.374	0.263	0.343	0.285	0.190	0.216	0.166	0.129	0.198	0.165	0.135
	1.00	0.707	0.618	0.444	0.448	0.407	0.297	0.341	0.286	0.217	0.252	0.210	0.158	0.232	0.191	0.139
	1.50	0.691	0.591	0.394	0.432	0.390	0.323	0.446	0.403	0.336	0.291	0.257	0.227	0.308	0.282	0.249
	2.00	0.708	0.594	0.355	0.498	0.471	0.382	0.499	0.458	0.404	0.336	0.311	0.285	0.339	0.321	0.290

Table 4: Rejection Rates for Separability Test using FDH, Bootstrap ($r = 1$, continued)

n	λ_2	$T_{1,n}$						$T_{2,n}$								
		$p=1,\ q=1$			$p=2,\ q=1$			$p=2,\ q=2$			$p=3,\ q=2$			$p=3,\ q=3$		
		.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01	.10	.05	.01
1000	0.00	0.035	0.014	0.002	0.006	0.004	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.001	0.000
	0.10	0.115	0.071	0.022	0.022	0.017	0.006	0.001	0.000	0.000	0.002	0.000	0.000	0.011	0.004	0.002
	0.20	0.495	0.410	0.286	0.193	0.147	0.096	0.034	0.016	0.004	0.011	0.005	0.001	0.018	0.007	0.003
	0.30	0.735	0.685	0.589	0.462	0.417	0.304	0.118	0.075	0.030	0.036	0.022	0.010	0.045	0.025	0.013
	0.40	0.791	0.774	0.726	0.650	0.600	0.493	0.266	0.166	0.068	0.102	0.062	0.022	0.081	0.055	0.031
	0.50	0.830	0.820	0.786	0.746	0.696	0.613	0.392	0.242	0.109	0.162	0.105	0.046	0.107	0.070	0.037
	0.60	0.825	0.811	0.774	0.774	0.739	0.663	0.431	0.315	0.149	0.214	0.157	0.076	0.166	0.114	0.070
	0.70	0.852	0.827	0.773	0.751	0.717	0.653	0.449	0.343	0.190	0.281	0.189	0.129	0.215	0.160	0.096
	0.80	0.820	0.789	0.727	0.740	0.713	0.640	0.498	0.411	0.263	0.293	0.217	0.137	0.244	0.179	0.118
	0.90	0.805	0.760	0.683	0.725	0.690	0.616	0.516	0.427	0.273	0.301	0.243	0.163	0.288	0.228	0.151
	1.00	0.802	0.761	0.656	0.745	0.715	0.618	0.543	0.455	0.345	0.336	0.270	0.203	0.277	0.229	0.159
	1.50	0.732	0.636	0.471	0.631	0.599	0.485	0.532	0.476	0.413	0.391	0.359	0.312	0.373	0.337	0.276
	2.00	0.681	0.555	0.339	0.578	0.543	0.423	0.568	0.529	0.481	0.442	0.411	0.364	0.423	0.393	0.348

Table 5: p-values for Tests of Separability on Banking Data

	Input		**Output**	
	Asy. Normal	Bootstrap	Asy. Normal	Bootstrap
$n = 303$				
SIZE	0.0013	0.0135	0.0495	0.1085
DIVERSE	0.8869	0.7910	0.8294	0.7295
joint test	0.1043	0.1760	0.0874	0.1180
$n = 6,607$				
SIZE	0.0000	0.0000	0.0000	0.0000
DIVERSE	0.9999	0.9980	0.9998	0.9925
joint test	0.0000	0.0000	0.0000	0.0000