INSTITUTE OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

MSAP Working Paper Series

No. 04/2010

Testing over-representation of observations in subsets of a DEA technology.

Mette Asmild

ORMS-group Warwick Business School

University of Warwick

Jens Leth Hougaard

Institute of Food and Resource Economics

University of Copenhagen

Ole B. Olesen

Department of Business and Economics

University of Southern Denmark

Testing over-representation of observations in subsets of a DEA technology

Mette Asmild
Warwick Business School
University of Warwick

Jens Leth Hougaard
Department of Food and Resource Economics
University of Copenhagen

Ole B. Olesen
Department of Business and Economics
University of Southern Denmark

March 17, 2010

Abstract

This paper proposes a test for whether data are over-represented in a given production zone, i.e. a subset of a production possibility set which has been estimated using the non-parametric Data Envelopment Analysis (DEA) approach. A binomial test is used that relates the number of observations inside such a zone to a discrete probability weighted relative volume of that zone. A Monte Carlo simulation illustrates the performance of the proposed test statistic and suggests good estimation of both facet probabilities and the assumed common inefficiency distribution in a three dimensional input space.

Keywords: Data Envelopment Analysis (DEA), Over-representation, Data density, Binomial test, Convex hull.

Correspondence: Mette Asmild, ORMS Group, Warwick Business School, Coventry, CV4 7AL, UK, e-mail: mette.asmild@wbs.ac.uk

1 Introduction

This paper introduces a test for whether observed data points are over-represented in certain production zones, i.e. subsets of (input sets of) the production space. The test is based on considerations of the relative volumes of these zones, weighted by probability estimates of observation frequencies. Specifically we consider the number of observations located in a certain zone, relative to the number that could be expected based on its relative weighted volume.

To motivate the need for such a test, consider attempting to evaluate how relevant different benchmarks are, for example when trying to determine the overall characteristics of the (relevant) benchmarks. Here one might be tempted to ignore benchmarks that do not dominate very many inefficient observations since they are less influential. But the number of observations they dominate really should be viewed in light of how many points they could be expected to dominate. It is here useful to give some thought to the nature of the underlying Data Generating Process (DGP). A simple comparison of the counts of how many observations each benchmark dominates implicitly assumes equal ex ante probabilities of being dominated by every benchmark. However, empirical results are likely to reveal otherwise. In particular there is a tendency of observing fewer points in production zones with either extreme input mixes or with very low efficiency scores. When considering the empirical frequencies we may have some benchmark which dominates a small number of observations but relative to what we expect given its location and the DGP there is actually an over-representation of observations referring to this benchmark. In such cases it might be unwise to ignore this benchmark from further considerations after all. See e.g. Thanassoulis (2001) for the use of the number of observations dominated by a benchmark to determine its robustness.

Another motivational example could be testing the hypothesis of rational inefficiency put forward by Bogetoft and Hougaard (2003). Given a set of common and known input prices, one might expect all observations to be located close to the cost minimizing input combination if the production units are assumed to be rational. Within the rational inefficiency framework, arguments are made, however, that it is still rational to be located inside the cone dominated by the cost minimizing point since production units may derive utility from slack consumption, leading to the notion of rational inefficiency. If we further assume that a Data Envelopment Analysis (DEA)

estimated frontier is a good approximation of the true underlying production possibilities, acceptance of a test for over-representation of data points inside the cone dominated by the cost minimizing input combination provides empirical support for the hypothesis of rational inefficiency.

Since the proposed test considers the number of points in certain zones relative to the weighted volumes of those subsets of the production space, what we call over-representation could also be viewed as higher (weighted) data density. Empirical investigation of data density can be approached in different ways. Statistical cluster analysis aims at identifying groups or clusters of data points. Parametric cluster analysis (Fraley and Raftery 1998, 1999, McLachlan and Peel 2000) is based on the assumption that each group of data points is represented by a density function belonging to some parametric family. The analysis then estimates the number of groups and their parameters from the observed data. In contrast, non-parametric statistical clustering approaches identify the center or mode of various groups and assign each data point to the domain of attraction of a mode. These approaches were originally introduced by Wishart (1969) and have subsequently been expanded on by especially Hartigan (1975, 1981, 1985). Common for the statistical cluster analysis approaches is that they aim at detecting the presence of clusters rather than considering differences in density in pre-specified production zones.

Within the realm of DEA, data density is at least indirectly considered in the recently quite popular bootstrapping approaches (see e.g. Simar and Wilson 1998, 2000). For example a distinction is made between a homogeneous and a heterogeneous bootstrap, reflecting whether or not it is reasonable to assume that the inefficiency distribution is independent of the choice of output levels and of input mix. Bootstrapping in this context analyzes the sensitivity of efficiency estimates to sampling variations of the estimated frontier and is used as a tool for bias correction and statistical inference. As such this literature has a different purpose than the one considered here.

In the present paper we remain within the non-parametric spirit of DEA by relying only on the information contained within the observed data points. We derive a binomial test that relates the number of observations inside certain production zones to a discrete probability weighted relative volume of these zones. This is done by considering the ratio between the volume of the zone and the volume of the total production possibility set, where these volumes are weighted by probability estimates of observations belonging to given facets and efficiency levels. The ratio of volumes provides estimates

of expected frequencies which are then related to the observed frequencies given as the ratio between the number of observations inside the zone and the total number of observations.

2 Methodology

Consider a set of n observed production plans $N = \{(x_i, y_i), i = 1, \ldots, n\}$ originating from a production process where r inputs are used to produce s outputs, i.e. $(x_i, y_i) \in \mathbf{R}_+^{r+s}$. Following the DEA tradition (cf. Banker, Charnes and Cooper 1984) we impose the following set of maintained hypotheses on the true underlying production technology: Convexity, Ray Unboundedness (constant returns to scale), Strong input and output Disposability and Minimal Extrapolation. Furthermore, we need some additional assumptions on the Data Generating Process (DGP). For the purpose of this paper we follow an input oriented version of the DGP suggested in Simar and Wilson (2000). We represent an input vector x in polar coordinates, which means that the angles of an input vector $x \in \mathbf{R}_+^r$ can be expressed as

$$\eta_i = \begin{cases} \arctan x_{i+1}/x_i & \text{if } x_i > 0 \\ \pi/2 & \text{if } x_i = 0 \end{cases} \tag{1}$$

for $i = 1, \ldots r-1$ and the modulus of the input vector is $\omega(x) = \|x\|_2 \equiv \sqrt{x^t x}$. Assume that, given a true technology P, each firm 'draws' an output vector $y \in \mathbf{R}_+^s$ from a distribution with density $f(y)$. Conditioned on this output vector the firm subsequently 'draws' an input mix vector $\eta \in [0, \pi/2]^{r-1}$ from a distribution with density $f(\eta|y)$. Finally, conditioned on the choice of output and input mix vectors the firm 'draws' a modulus $\omega \in \mathbf{R}_+^1$ from a distribution with density $f(\omega|\eta, y)$. Specifically, we maintain that the DGP satisfies the following assumptions:

The observations $(x_i, y_i) \in \mathbf{R}_+^{r+s}, i = 1, \ldots, n$ are realizations of i.i.d. random variables with probability density function $f(x, y)$, which has a support over $P \subset R_+^{r+s}$, where P is a production set defined by

$$P = \{(x, y) \mid x \text{ can produce } y\} \tag{2}$$

and $S(y) = \{x \mid (x, y) \in P\}$ is the input set. We define the radial efficiency measure $\theta(x, y) = \min\{\theta \mid (\theta x, y) \in P\}$. For any given (y, η), the corresponding point on the boundary of P is denoted $x^{\partial}(y, \eta)$ and has a modulus

$$\omega(x^{\partial}(y,\eta)) = \min\left\{\omega \in \mathbf{R}^1_+ : f(\omega|\eta,y) > 0\right\} \tag{3}$$

and the related radial efficiency measure $\theta(x,y)$ can be expressed as

$$0 \leq \theta(x,y) = \frac{\omega(x^{\partial}(y,\eta))}{\omega(x)} \leq 1. \tag{4}$$

Note that the density $f(\omega|\eta,y)$ with support $[\omega(x^{\partial}(y,\eta)),\infty)$ induces a density $f(\theta|\eta,y)$ on $[0,1]$. The advantage of representing the input vector in terms of polar coordinates is that the joint density $f(x,y)$ can now be described as a product of three densities

$$f(\omega,\eta,y) = f(\omega|\eta,y)\, f(\eta|y)\, f(y) \tag{5}$$

where the ordering of the conditioning reflects the assumed sequence of the DGP mentioned above.

Specifically, we in the following propose a test for whether data are over-represented in certain subsets of a DEA-estimated technology with input sets

$$\mathcal{S}^{CCR}(y') = \{x \mid \sum_{j\in N}\lambda_j x_j \leq x, \sum_{j\in N}\lambda_j y_j \geq y', \lambda_j \geq 0, j \in N\}, \tag{6}$$

for output level y'. Since the proposed test is based on volumes of production zones, the technology must be bounded. Let θ_j be the radial inefficiency of the j'th production plan and define the projected (possibly weakly) efficient production plans $(\widetilde{x}_j, y_j) \equiv (\theta_j x_j, y_j), j \in N$. Let θ^{α} be defined such that $\Pr(\theta \leq \theta^{\alpha}) = \alpha$. In the following we will focus on the bounded family of input sets given by

$$\begin{aligned}\widetilde{\mathcal{S}}(y',\alpha) =& \{\, x \mid \sum_{i\in N}\lambda_i\widetilde{x}_i + (\theta^{\alpha})^{-1}\sum_{i\in N}\widetilde{\lambda}_i\widetilde{x}_i = x, \\ & \sum_{i\in N}\lambda_i y_i + \sum_{i\in N}\widetilde{\lambda}_i y_i = y', \lambda_i, \widetilde{\lambda}_i \geq 0, i \in N\,\} \\ & \setminus \{\, x \mid (\theta^{\alpha})^{-1}\sum_{i\in N}\widetilde{\lambda}_i\widetilde{x}_i = x, \sum_{i\in N}\widetilde{\lambda}_i y_i = y', \widetilde{\lambda}_i \geq 0, i \in N\,\}\end{aligned} \tag{7}$$

Choosing a small value for α means only ignoring production plans from areas of the production space with low probability when considering only the bounded set $\widetilde{\mathcal{S}}(y,\alpha) \subset \mathcal{S}^{CCR}(y)$ instead of the full set $\mathcal{S}^{CCR}(y)$.

Moreover, for some point $(x', y') \in \mathcal{N}$, denote by

$$\mathcal{K}(x', y') = \{x \in \widetilde{\mathcal{S}}(y', \alpha) | x \geq x'\} \tag{8}$$

the bounded cone of input combinations in $\widetilde{\mathcal{S}}(y', \alpha)$ dominated by x'.

The various concepts are illustrated in Figure 1 below. We have generated 200 data points according to a DGP from the Monte Carlo simulation described in section 7. An envelopment is provided from 10 facets denoted f$i, i = 1, \ldots, 10$. Two of these facets, f9 and f5, are part of unbounded exterior facets. These two facets are spanned by a CCR-efficient observation and one of the two projected inefficient observations indicated by the two arrows in Figure 1. The set $\widetilde{\mathcal{S}}(y', \alpha)$ is the intersection of the 13 halfspaces corresponding to the 13 facets generating the boundary of the convex hull of all 200 observation except for the subset below facet f13 but above the facets f$i, i = 1, \ldots, 10$ expanded by the factor $(\theta^{\alpha})^{-1}$ (approximately equal to 2 in the figure).

2.1 The proposed test

As mentioned in the introduction we aim to identify over-representation of data points in certain subsets of the input set, for instance a dominance cone with vertex in an efficient production plan (x', y'), e.g. a cost minimizing observation. To formally test whether projected data points are over-represented in, for example, the bounded cone $\mathcal{K}(x', y')$, we use a binomial test with the null hypothesis that the probability p of a projected data point being located within the bounded cone is:

$$p = p\,(y') = \int_{(\eta,\omega)\in\mathcal{K}(x',y')} f(\omega, \eta, y') d\,(\eta, \omega) \tag{9}$$

Hence, we maintain that data reflects a DGP as specified in (5). For any set $\mathcal{H}$ let $V(\mathcal{H})$ denote the *volume* of $\mathcal{H}$. To simplify the presentation of the general idea let us initially assume that the projected data are uniformly distributed over the input sets. This is a restrictive assumption but it allows us to directly use the ratios of volumes as a simple estimator of p:

$$\widehat{p} = \frac{V(\mathcal{K}(x', y'))}{V(\widetilde{\mathcal{S}}(y', \alpha))}, \tag{10}$$

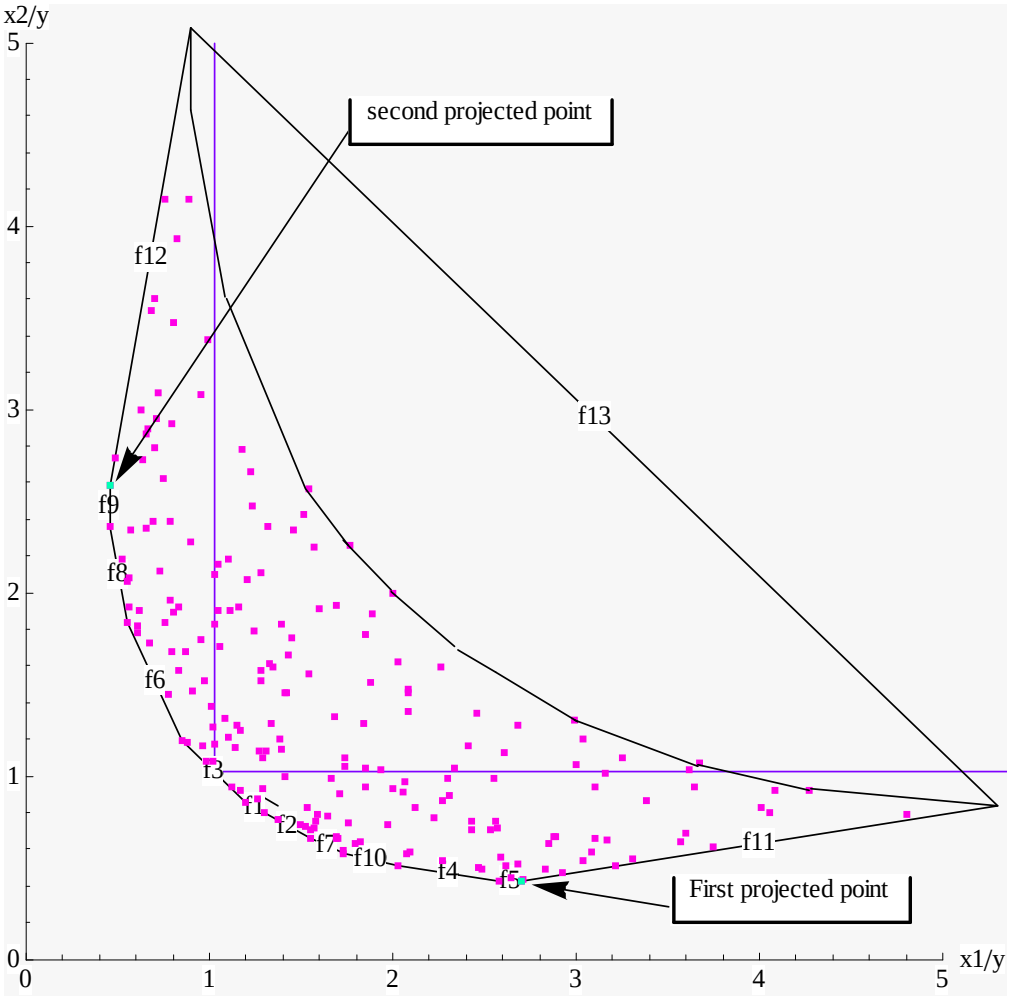

Figure 1: The bounded subset of the input set.

where we in the following ignore the trivial cases $\widehat{p} \in \{0,1\}$. By using the constant returns to scale assumption, for sufficiently small α all data points can be projected onto the input set $\widetilde{\mathcal{S}}(y',\alpha)$. Denote by

$$\#\mathcal{K}(x',y') = |\{(x_k,y_k)_{k=1,...N} : x_k \in \widetilde{\mathcal{S}}(y',\alpha) : \exists \kappa \in \mathbf{R}_+,\ \kappa x_k \geq x',\ \kappa y_k \leq y'\}| \tag{11}$$

the number of data points that can be projected into the bounded cone $\mathcal{K}(x',y')$.

The null hypothesis that the ratio of volumes determines the probability of being located in the bounded cone may be tested by,

$$z = \frac{\#\mathcal{K}(x',y') \pm 0.5n\hat{p}}{\sqrt{n\hat{p}(1-\hat{p})}} \tag{12}$$

where $\#\mathcal{K}(x',y') + 0.5$ is used if $\#\mathcal{K}(x',y') < n\hat{p}$ and $\#\mathcal{K}(x',y') - 0.5$ is used if $\#\mathcal{K}(x',y') > n\hat{p}$. The value of z is asymptotically normal distributed with mean 0 and variance 1 (see e.g. Siegel and Castellan 1988). If the null hypothesis is rejected in the corresponding one-tailed test we conclude that there is over-representation of data points inside the bounded cone.

Finally, it should be noted that the two volumes $V(\widetilde{\mathcal{S}})$ and $V(\mathcal{K})$ depend on the units of measurement whereas the number of observations inside the cone $\#\mathcal{K}$, as well as the total number of observations, is independent of the metrics. However, the ratio between the volumes $\hat{p}$, as given by (10), is scale invariant, i.e. the observed data points can be scaled with strictly positive weights without changing $\hat{p}$ (but clearly $\hat{p}$ is not affinely invariant).

3 Discrete approximations of densities

Where the introduction of the proposed test above relied on the assumption of the observations being uniformly distributed we now, perhaps more realistically, relax the distributional assumption on the radial efficiency score as well as on the input mix (η). Let us simplify by considering the one output case and assuming a common distribution of the efficiency scores of some parametric form, i.e. $f(\omega|\eta,1) = f(\omega)$.[1] Consider a discrete approximation of the distribution of efficiency scores given by I intervals (slices) of the

[1] This assumption is often used in the bootstrapping literature and is denoted a homogeneous bootstrap (see Simar and Wilson 2000, p. 64).

bounded support $[\theta^{\alpha}, 1]$: $\{[i_1, i_2], (i_2, i_3], \ldots, (i_I, i_{I+1}]\}$, $i_1 = \theta^{\alpha}, i_{I+1} = 1$ and probability p_i^{θ}, of belonging to the i'th slice for $i = 1, \ldots, I$. To obtain a reasonable precision of the discrete approximation we choose the intervals such that the probabilities are approximately identical. We approximate the probability $\Pr(\theta \leq i_k)$ with $\sum_{i=1}^{k-1} p_i^{\theta}, k = 2, \cdots, I+1$.

As we are now considering the case where multiple inputs are used to produce a single output, let $\widetilde{\Gamma} = \{q_i = \theta_i x_i / y_i\}_{i \in N} \subset \mathbf{R}_+^r$ be the projected data points. Let $conv(\cdot)$ be the convex hull operator, and define the k'th 'slice' of the bounded technology $\widetilde{\mathcal{S}}(1, \alpha)$ (in the following simply denoted by $\widetilde{\mathcal{S}}$) as

$$\widetilde{\mathcal{S}}_k = conv\{(i_{k+1})^{-1}\widetilde{\Gamma} \cup (i_k)^{-1}\widetilde{\Gamma}\} \setminus conv\{(i_k)^{-1}\widetilde{\Gamma}\} \tag{13}$$

Next, let us relax the assumption on the DGP regarding $f(\eta|1)$ using the available (empirical) information from DEA, which has resulted in $j = 1, \ldots, F$ different facets of the estimated frontier. The empirical facets provide a natural discretization of the range of input mixes η. Let p_j^f be the probability of getting an input mix which belongs to the cone spanned by the j'th facet.

Note that the homogeneity assumption $f(\omega|\eta, 1) = f(\omega)$ implies that we approximate the probability $\hat{p}_{ij}$ of getting an observation in the i'th slice intersected by the cone spanned by the j'th facet with $p_i^{\theta} \times p_j^f$.

Figure 2 illustrates the discretization of the efficiency distribution with $I = 10$ and a DEA envelopment frontier with $F = 10$ facets, denoted f$i, i = 1, \ldots, 10$. The dots represent 200 generated observations and the dots marked by the two arrows are projected inefficient observations located on unbounded exterior facets. Adding these two pseudo-observations to the data set results in a bounded subset of the input possibility set with a "lower" envelope consisting of points that dominates all inefficient observations among the generated data.

The figure illustrates the situation where we have generated 200 data points according to the DGP from the Monte Carlo simulation described in section 7. The input mix is uniformly distributed with $\eta \in [0.1, \pi/2 - 0.1]$.[2] The density of the radial efficiency score in the Monte Carlo simulation is assumed to be a uniform distribution on a support $[0.5, 1]$. 10 intervals are

[2]Note that a uniform distribution of $\eta \in [0.1, \pi/2 - 0.1]$ of course implies that projected data points on the boundary are more sparse in the specialized regions closer to the axes compared to the regions in the center with η close to $\pi/4$.

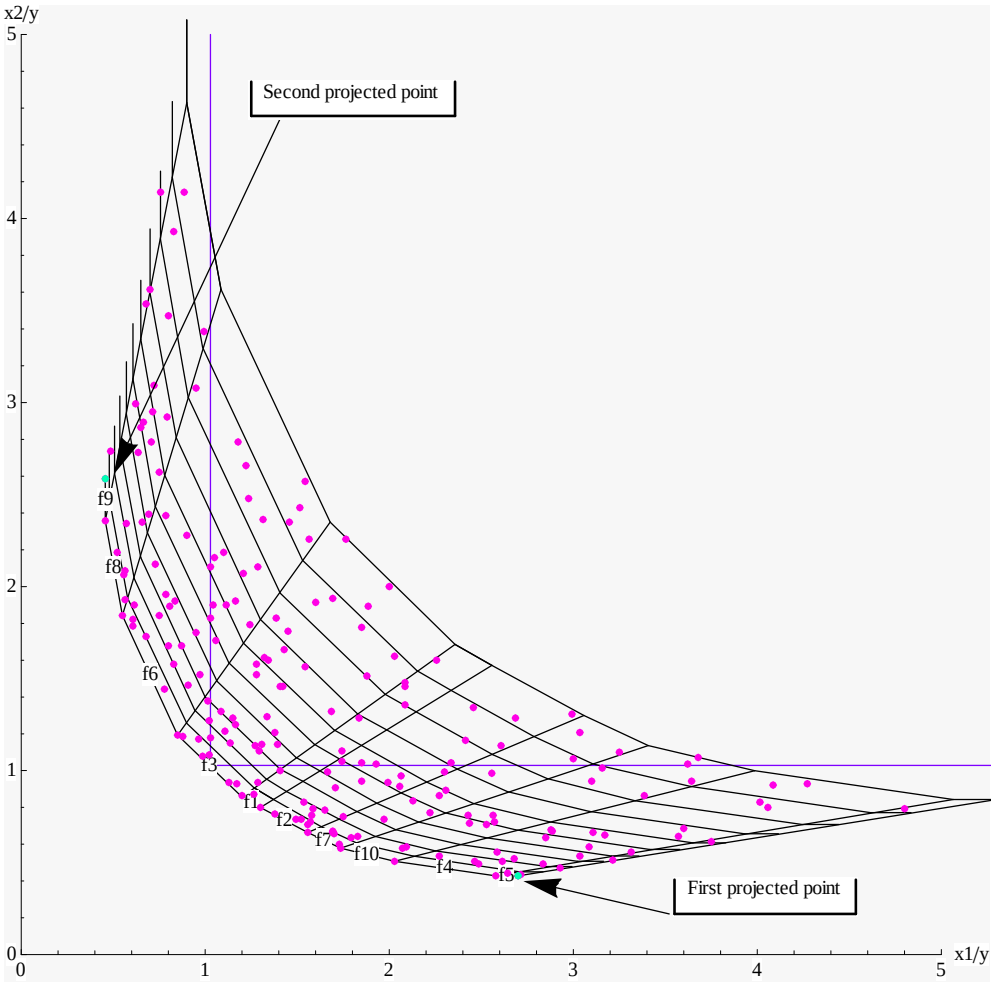

Figure 2: The discretization of the efficiency distribution

used in the discrete approximation. Note that a uniform distribution of the efficiency scores combined with of a uniform η does not imply that data points are distributed uniformly on some bounded subset of $\mathcal{S}^{CCR}$.

Now, consider a bounded cone $\mathcal{K}(q') \subset \widetilde{\mathcal{S}}$ with vertex $q' \in \widetilde{\Gamma}$. Let a $I \times F$ matrix be given as $m_{ij} = \left\{ \frac{V(\mathcal{K}_{ij}(q'))}{V(\widetilde{\mathcal{S}}_{ij})} \right\}$ where $\widetilde{\mathcal{S}}_{ij}$ is the i'th slice of the cone spanned by the j'th facet and $\mathcal{K}_{ij}(q') = \mathcal{K}(q') \cap \widetilde{\mathcal{S}}_{ij}$. Hence, to calculate a more general estimator $\widehat{\widehat{p}}$ of the probability of getting an observation within the bounded cone $\mathcal{K}$ we use

$$\widehat{\widehat{p}} = \begin{bmatrix} \widehat{p}_1^{\theta} \\ \vdots \\ \widehat{p}_I^{\theta} \end{bmatrix}^T \begin{bmatrix} m_{11} & \dots & m_{1F} \\ \vdots & \dots & \vdots \\ m_{I1} & \dots & m_{IF} \end{bmatrix} \begin{bmatrix} \widehat{p}_1^{f} \\ \vdots \\ \widehat{p}_F^{f} \end{bmatrix} \tag{14}$$

To formally test, in this more general setup, whether data points $\widetilde{\Gamma}$ are over-represented in the bounded cone $\mathcal{K}(q')$ we simply use this more general estimator $\widehat{\widehat{p}}$ as given in (14) instead of the $\widehat{p}$ given in (12). Note that since $\mathcal{K}_{ij}(q')$ in some cases will contain very few data points one should keep in mind that this test is only meaningful for situations where $n\hat{p}_{ij}(1-\hat{p}_{ij})$ is not too small (Siegel and Castellan 1988 advocate that $n\hat{p}_{ij}(1-\hat{p}_{ij}) > 9$).

4 Practical solution procedure

Consider a DEA where multiple inputs are used to produce a single output and the "lower" envelope of the estimated input set has F different facets. Further choose a disretization involving I different intervals of efficiency scores. To get the estimator $\widehat{\widehat{p}}$ in (14) we need i) an estimator $\widehat{p}_j^f$ of p_j^f, ii) an estimator $\widehat{p}_i^{\theta}$ of p_i^{θ} and iii) the $I \times F$ matrix given as $\{m_{ij}\}$ for all i, j.

For a given empirical data set $N = \{(x_i, y_i), i = 1, \dots, n\} \subset \mathbf{R}_+^{r+1}$ the tests described previously can be performed using the following procedure:

Step I: First identify the input set $\widetilde{\mathcal{S}}$. Assuming constant returns to scale enables a projection of all observations onto the level set of $y = 1$ by transforming the observations $\{(x_i, y_i)\}_{i \in N}$ into the data points $\{x_i/y_i\}_{i \in N} \subset \mathbf{R}_+^r$ and calculating the corresponding efficiency scores $\theta_i, i \in N$.

Step II: Let $\theta^{\alpha} = \min_{i \in N} \theta_i$ and let $\widetilde{\Gamma} = \{\theta_i x_i / y_i\}_{i \in N}$. Define in accordance

with (7) the estimated input set as the convex hull

$$\widehat{\widetilde{\mathcal{S}}} = conv\{\widetilde{\Gamma} \cup (\theta^{\alpha})^{-1}\widetilde{\Gamma}\} \setminus conv\{(\theta^{\alpha})^{-1}\widetilde{\Gamma}\} \tag{15}$$

Decompose $\widehat{\widetilde{\mathcal{S}}}$ into $\widehat{\widetilde{\mathcal{S}}}_{ij}, \forall i, j$ using the following procedure. For the j'th facet let $\widetilde{\Gamma}_j$ be the subset of data points in $\widetilde{\Gamma}$ belonging to facet j, i.e. $\cup_{j=1}^{F}\widetilde{\Gamma}_j = \widetilde{\Gamma}$. Hence, we can decompose $\widehat{\widetilde{\mathcal{S}}}$ into the following F convex subsets:

$$\begin{aligned} \widehat{\widetilde{\mathcal{S}}} &= \cup_{j=1}^{F}\widehat{\widetilde{\mathcal{S}}}_j, \text{ where} \\ \widehat{\widetilde{\mathcal{S}}}_j &= conv\{\widetilde{\Gamma}_j \cup (\theta^{\alpha})^{-1}\widetilde{\Gamma}_j\} \end{aligned}$$

Finally we can decompose each of these $\widehat{\widetilde{\mathcal{S}}}_j$ into the following I convex subsets:

$$\begin{aligned} \widehat{\widetilde{\mathcal{S}}}_j &= \cup_{i=1}^{I}\widehat{\widetilde{\mathcal{S}}}_{ij}, \text{ where} \\ \widehat{\widetilde{\mathcal{S}}}_{ij} &= conv\{\kappa_i(\theta^{\alpha})^{-1}\widetilde{\Gamma}_j \cup \kappa_{i-1}(\theta^{\alpha})^{-1}\widetilde{\Gamma}_j\}, i = 1, \ldots, I \end{aligned}$$

where $\kappa_i = 1 + \frac{I-i}{I}\left((\theta^{\alpha})^{-1} - 1\right), i = 0, 1, \ldots, I$

Step III: Determine the volume $V(\widehat{\widetilde{\mathcal{S}}}_{ij})$ using, for instance, the Qhull software (www.qhull.org), which employs the Quickhull algorithm for convex hulls, as suggested by Barber, Dobkin and Huhdanpaa (1996).

Step IIIa: Convex hull generation and calculation of volumes is typically difficult if facets are over-determined, i.e. if facets are generated by more data points than the dimension of the space. Hence generating the sets $\widehat{\widetilde{\mathcal{S}}}_{ij}$ from $\widetilde{\Gamma} = \{\theta_i x_i / y_i\}_{i \in N}$ where all inefficient data points are projected to facets complicates the subsequent use of Qhull to determine facets and volumes. Hence for practical purpose we only project a subset of inefficient points to the facets. To determine which points are needed the following procedure is used:

Project all inefficient observations dominated by points on exterior facets to the frontier. Inefficient observation dominated by points on interior facets are ignored since they provide no additional information. Let $I_{CCR} \subset N$ be an index set of CCR-efficient DMUs. Solve a modified CCR-DEA model

evaluating each of the projected data points where the objective function minimizes the sum of input slacks and the evaluated DMU is excluded from the set of potential peers. We only project DMUs having a strictly positive sum of slacks, i.e. points that can not be expressed as a convex combination of other projected points or the points in I_{CCR}. Such projected DMUs are needed to delimit bounded representations of the relevant (unbounded) exterior facets.

Step IV: Select a point q' to be the vertex of the bounded cone $\mathcal{K}(q')$, such that the intersection with the bounded input set $\widetilde{\mathcal{S}}$ is non-empty (an obvious choice could be a cost minimizing observation). One way to calculate $V(\mathcal{K}(q'))$ would involve both identification of extreme points dominated by q' and identification of all the extreme points of the intersection between the cone $\{q|q \geq q'\}$ and the input set $\widetilde{\mathcal{S}}$. (see Muller and Preparata 1978). However, an easier way to calculate $V(\mathcal{K}(q'))$ using QHULL is based on the so-called Minkowski-Weyl's Theorem (see Appendix), which states that every polyhedron has both a (halfspace) H-representation and a (vertex) V-representation.

In the specific case where we use the estimator $\widehat{\widetilde{\mathcal{S}}}$, to find the volume of the intersection of the bounded cone and $\widehat{\widetilde{\mathcal{S}}}_{ij}$ we suggest the following subprocedure:

- Use QHULL to generate a H-representation of $\widehat{\widetilde{\mathcal{S}}}_{ij}$, as defined by the extreme points (cf. the set-up in Olesen and Petersen 2003).
- Augment the H-representation of $\widehat{\widetilde{\mathcal{S}}}_{ij}$ with r halfspaces defining the bounded cone. Each of these halfspaces is characterized by having one normal vector component equal to zero and all halfspaces contain the vertex of the cone.
- Use Qhull to calculate the volume of this H-representation of $\widehat{\widetilde{\mathcal{S}}}_{ij} \cap \mathcal{K}(q')$.

Step V: The number of data points that can be projected onto the bounded cone $\#\mathcal{K}(q')$ is determined by a simple count of data points in $\widetilde{\mathcal{S}}$ dominated by q'.

Step VI: It is now possible to establish the null hypothesis of the binomial test. The ratios between $V(\widehat{\widetilde{\mathcal{S}}}_{ij} \cap \mathcal{K}(q'))$ and $V(\widehat{\widetilde{\mathcal{S}}}_{ij})$ is an estimator of the probability of a projected data point being located within $\widehat{\widetilde{\mathcal{S}}}_{ij} \cap \mathcal{K}(q')$ if the location of the data points is determined by the relative probability weighted volumes alone. With n observations this means that we should expect $n\widehat{\widehat{p}}$ observations inside the bounded cone, where $\widehat{\widehat{p}}$ is given by (14) resulting in the test statistic z given by (12) and corresponding test probability.

It should be noted that the proposed algorithm is well defined and can be performed using a standard LP-solver combined with QHULL, which is sufficient for most data sets. The Monte Carlo studies reported below is performed using a combination of CPLEX (step I, and saving results on files) and a Mathematica code (step II-VI) reading the CPLEX results on scores and dominating vertices and calling a QHULL-code for getting the H-representation of each intersection with the bounded cone and each volume calculation.

5 Monte Carlo simulations, 3 inputs and one output

A minor change in the representation of the input vector x expressed in polar coordinates is introduced for the Monte Carlo simulation. We express an input vector $x \in R^r_+$ as $\eta_i = \arctan x_{i+1}/x_1$ for $x_i > 0$ and $\pi/2$ if $x_i = 0$ for $i = 1, \ldots r-1$ and the modulus of the input vector is $\omega(x) = \|x\|_2 \equiv \sqrt{x^t x}$. Hence, for $r = 3$, $x_2 = (\tan \eta_1)\, x_1$ and $x_3 = (\tan \eta_2)\, x_1$. We assume that the true isoquant has the following form $x_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} = 10$, where $\alpha_1 + \alpha_2 + \alpha_3 = 1$. Hence, $x_1 = 10\,(\tan \eta_1)^{-\alpha_2}\,(\tan \eta_2)^{-\alpha_3}$. $\omega^2 = x_1^2 + x_2^2 + x_3^2$ from which it follows that $\omega^2 = x_1^2 + x_1^2\,(\tan^2 \eta_1) + x_1^2\,(\tan^2 \eta_1) = x_1^2\,[1 + (\tan^2 \eta_1) + (\tan^2 \eta_2)]$ or $\omega = 10\,(\tan \eta_1)^{-\alpha_2}\,(\tan \eta_2)^{-\alpha_3}\left[\sqrt{1 + (\tan^2 \eta_1) + (\tan^2 \eta_2)}\right]$. Hence, we use a DGP where[3]

- $\tan \eta_i \equiv \frac{x_{i+1}}{x_1} \sim U\left[\tan(0.1), \tan\left(\frac{\pi}{2} - 0.1\right)\right], i = 1, 2$

[3]The conversion from polar coordinates to Euclidian coordinates follows as $x_1 = \omega\sqrt{\left[1 + \left(\tan^2 \eta_1\right) + \left(\tan^2 \eta_2\right)\right]^{-1}}$, $x_2 = (\tan \eta_1)\, x_1$, $x_3 = (\tan \eta_2)\, x_1$.

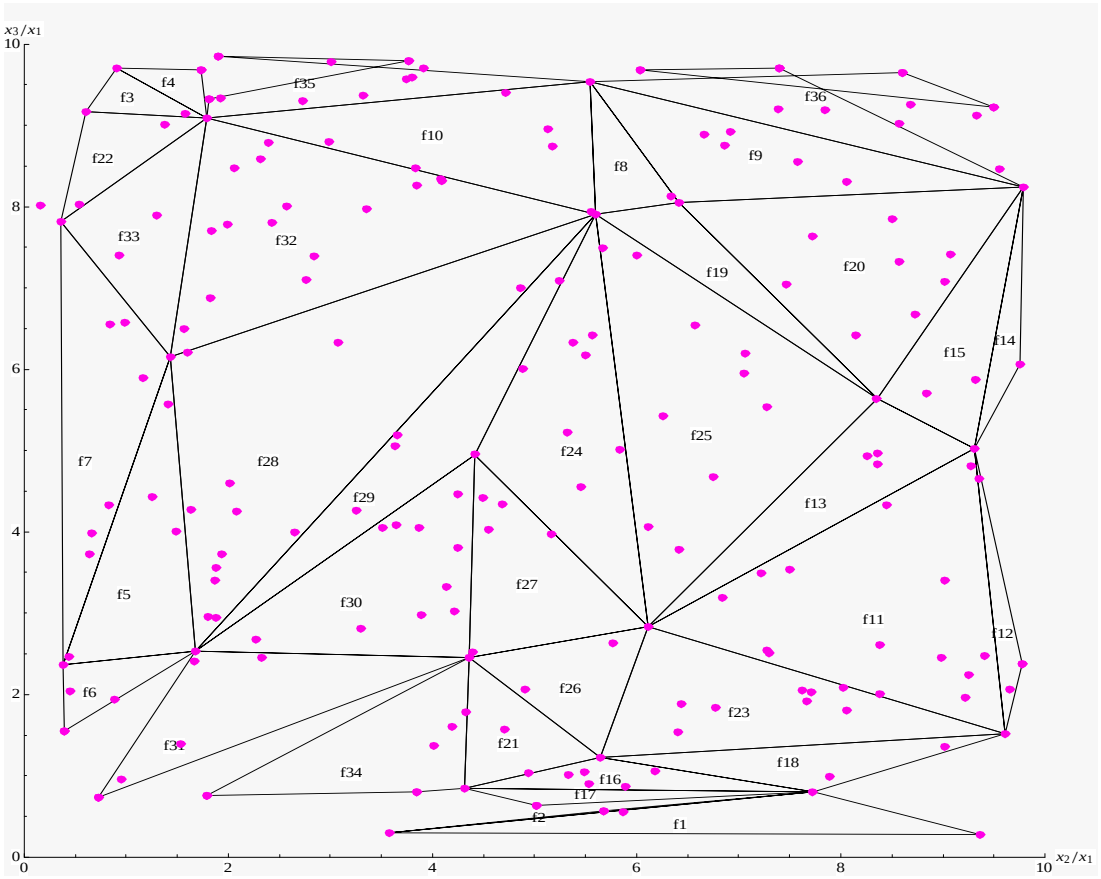

Figure 3: 200 observations generated in $(x_2/x_1) - (x_3/x_1)$ space

- $\theta^{-1} \sim U[1,2]$
- $\omega = \theta^{-1} 10 \left(\tan \eta_1\right)^{-\alpha_2} \left(\tan \eta_2\right)^{-\alpha_3} \left[\sqrt{1 + \left(\tan^2 \eta_1\right) + \left(\tan^2 \eta_2\right)}\right]$

Figure 3 illustrates this DGP for one of the replications of the simulation. In $\left(\tan \eta_1 = \frac{x_2}{x_1}, \tan \eta_2 = \frac{x_3}{x_1}\right)$ space 200 observations are generated uniformly distributed on $\left[\tan\left(0.1\right), \tan\left(\frac{\pi}{2} - 0.1\right)\right]^2$. 36 facets are spanning the frontier, where the last two facets (f35 and f36) are spanned by more than 3 data points (exterior facets).

Results from the Monte Carlo simulation are reported in Table 1-4. We have analyzed the sensitivity of the results from the simulation with regard to two and five different estimators of p_j^f and p_i^θ, respectively. $\widehat{p}_j^f(1)$ is estimated as the relative number of DMUs projected to the j'th facet[4] and $\widehat{p}_j^f(2)$ is estimated as the relative volume of the j'th facet in $\frac{x_2}{x_1} - \frac{x_3}{x_1}$ space. $\widehat{p}_i^\theta(l), l = 1, 2$ are estimated as the relative number of DMUs with scores corresponding to each of the ten slices, where we for $l = 1$ disregard all scores of one and use the original scores and for $l = 2$ use a set of bias corrected scores (for bias correction, see Wilson (2008) and Simar and Wilson (1998)). $\widehat{p}_i^\theta(l), l = 3, 4$ are estimated like $\widehat{p}_i^\theta(l), l = 1, 2$ but using a kernel estimator of the density with a Gaussian kernel and bandwidth in $\{0.1, 0.15, 0.2\}$ using the reflection method to avoid bias at the boundaries of the bounded support for θ (Silverman 1986). Optimal bandwidth (approximately 0.15 for $l = 3$) has been estimated using crossvalidation (see Daraio and Simar (2007), Ch. 4, and Silverman (1986), Ch. 3). The score function is rather flat on the interval $[0.1, 0.2]$ for $l = 3$. Hence in the tables we have included results from more than one bandwidth in this interval. The optimal bandwidth for $l = 4$ is approximately 0.1. Finally, $\widehat{p}_i^\theta(5) = 0.1, \forall i$ reflecting the "true" generation of θ in the DGP. We use $B \in \{50, 100, 150, 164\}$ replications in this Monte Carlo simulation. For each replication and each combination of $\widehat{p}_j^f$ and $\widehat{p}_i^\theta$ we estimate the probability $\widehat{\widehat{p}}(l, k)$ in (14):

$$\widehat{\widehat{p}}(l,k) = \left[\widehat{p}_i^\theta(l)\right]^t \left[\frac{V(\mathcal{K}_{ij}(q'))}{V(\widetilde{\mathcal{S}}_{ij})}\right] \left[\widehat{p}_j^f(k)\right], k = 1, 2, l = 1, \ldots, 5$$

and the corresponding z_{lk} in (12). Since data in the simulation is generated from a DGP that does not reflect any tendency to having an overrepresentation of data points in the restricted cone we expect to see an empirical distribution of the values of z_{lk} closely resembling a standard N(0,1) distribution. The results from the simulations show in general that $\widehat{p}_j^f(1)$ performs rather poorly compared to $\widehat{p}_j^f(2)$ when combined with $\widehat{p}_i^\theta(5)$. Using $\widehat{p}_j^f(2)$ in combination with the true information on the score distribution of the DGP ($\widehat{p}_i^\theta(5)$) provides a test statistic z_{25} that very nicely recovers the characteristics of the DGP.

[4]Efficient DMUs spanning the facets are added with a fraction to each facet reflecting how many facets such a DMU is spanning. The estimator reflects only the generated 200 DMUs.

Summary Statistics, Monte Carlo results					
z_{ij}	$Mean$	$Variance$	Min	Max	$Skewness$
z_{11}	-1.58144	0.623905	-4.23251	0.372523	-0.0990964
z_{12}	-1.75651	0.757589	-4.16705	0.993659	0.209076
z_{21}	0.304115	0.600193	-1.88618	2.84391	0.457486
z_{22}	0.154575	0.721779	-1.92377	3.16947	0.442015
$bandwidth = 0.1$					
z_{31}	0.183046	1.48998	-3.65963	3.19543	-0.235727
z_{32}	0.0255714	1.80728	-3.59005	2.9542	-0.17745
$bandwidth = 0.15$					
z_{31}	0.194465	1.30411	-3.45035	2.93507	-0.250046
z_{32}	0.0365467	1.5956	-3.31926	2.84367	-0.181991
$bandwidth = 0.2$					
z_{31}	0.197774	1.13229	-3.21822	2.66221	-0.251012
z_{32}	0.0427142	1.40117	-3.03304	2.72701	-0.174128
–					
z_{51}	0.213105	0.796296	-2.61312	2.31405	-0.228829
z_{52}	0.0566723	1.01885	-2.46096	2.43443	-0.0902907

Table 1: Summary statistics of the testors

$z_{l,1}$ facet probabilities estimated from relative number of projected data points

$z_{l,2}$ facet probabilities estimated from relative volume of facets

$z_{l,k}, l = 1, 2$ score interval probabilities estimated from empirical score distributions

$z_{l,k}, l = 3, 4$ score interval probabilities estimated from kernel density score distributions

$z_{5,k}$ score interval probabilities equal to 0.1

Simulation Monte Carlo results

Nominal levels 0.8,0.9,0.95,0.975,0.99

z_{ij}	0.8	0.9	0.95	0.975	0.99
z_{11}	0.349693	0.509202	0.699387	0.797546	0.91411
z_{12}	0.294479	0.429448	0.552147	0.699387	0.822086
z_{21}	0.889571	0.944785	0.969325	0.981595	0.981595
z_{22}	0.895706	0.95092	0.969325	0.97546	0.993865
$bandwidth = 0.1$					
z_{31}	0.717791	0.815951	0.90184	0.92638	0.96319
z_{32}	0.680982	0.773006	0.834356	0.907975	0.95092
$bandwidth = 0.15$					
z_{31}	0.723926	0.846626	0.91411	0.93865	0.98773
z_{32}	0.693252	0.797546	0.877301	0.92638	0.95092
$bandwidth = 0.2$					
z_{31}	0.760736	0.883436	0.92638	0.969325	0.98773
z_{32}	0.723926	0.815951	0.889571	0.932515	0.96319
–					
z_{51}	0.846626	0.920245	0.96319	0.98773	0.993865
z_{52}	0.809816	0.889571	0.932515	0.969325	1.

Table 2: The accuracy of the coverage of the testors

$z_{l,1}$ facet probabilities estimated from relative number of projected data points

$z_{l,2}$ facet probabilities estimated from relative volume of facets

$z_{l,k}, l = 1, 2$ score interval probabilities estimated from empirical score distributions

$z_{l,k}, l = 3, 4$ score interval probabilities estimated from kernel density score distributions

$z_{5,k}$ score interval probabilities equal to 0.1

Summary Statistics, Monte Carlo results					
z_{ij}	$Mean$	$Variance$	Min	Max	$Skewness$
$bandwidth = 0.1, B = 50, 100, 150, 164$					
z_{32}	-0.0423542	1.76272	-3.22304	2.8622	0.122281
z_{32}	-0.112964	1.80697	-3.37765	2.9542	0.0683355
z_{32}	-0.004464	1.81732	-3.59005	2.9542	-0.193337
z_{32}	0.0255714	1.80728	-3.59005	2.9542	-0.17745
$bandwidth = 0.2, B = 50, 100, 150, 164$					
z_{32}	-0.00535393	1.32712	-2.68905	2.32881	0.0971093
z_{32}	-0.0768455	1.39895	-2.94285	2.61226	0.0459719
z_{32}	0.0211947	1.41388	-3.03304	2.61226	-0.203502
z_{32}	0.0427142	1.40117	-3.03304	2.72701	-0.174128
$B = 50, 100, 150, 164$					
z_{52}	0.0242642	0.925966	-2.01797	1.87989	0.114405
z_{52}	-0.048325	1.0211	-2.36023	2.23369	0.126139
z_{52}	0.0445987	1.03427	-2.46096	2.26337	-0.136333
z_{52}	0.0566723	1.01885	-2.46096	2.43443	-0.0902907

Table 3: Summary statistics of selected testors for varying sample size

Simulation Monte Carlo results

Nominal levels 0.8,0.9,0.95,0.975,0.99

z_{ij}	0.8	0.9	0.95	0.975	0.99
$bandwidth = 0.1, B = 50, 100, 150, 164$					
z_{32}	0.68	0.76	0.84	0.92	0.96
z_{32}	0.7	0.78	0.84	0.9	0.94
z_{32}	0.68	0.773333	0.833333	0.906667	0.953333
z_{32}	0.680982	0.773006	0.834356	0.907975	0.95092
$bandwidth = 0.2, B = 50, 100, 150, 164$					
z_{32}	0.74	0.82	0.9	0.96	0.98
z_{32}	0.75	0.82	0.88	0.93	0.96
z_{32}	0.726667	0.813333	0.886667	0.933333	0.966667
z_{32}	0.723926	0.815951	0.889571	0.932515	0.96319
$B = 50, 100, 150, 164$					
z_{52}	0.8	0.9	0.98	1.	1.
z_{52}	0.8	0.88	0.93	0.99	1.
z_{52}	0.806667	0.886667	0.933333	0.973333	1.
z_{52}	0.809816	0.889571	0.932515	0.969325	1.

Table 4: The coverage of the testors for varying sample size

Results are presented for the accuracy of the proposed test statistic. The summary statistics and the empirical coverage accuracy of both z_{31} and z_{32} with bandwidth equal to 0.2 are behaving reasonably, partly recovering the characteristics of the DGP. Hence, these are the test statistics with the best characteristics. The results in Table 1 and 2 show that the test statistic z_{32} (based on the kernel based score distribution from the non bias-corrected scores $\widehat{p}_i^{\theta}(3)$ in combination with $\widehat{p}_j^{I}(2)$) is almost unbiased but apparently the smoothening from the kernel induces a variance above one even for a sample size above 150. Combining $\widehat{p}_i^{\theta}(3)$ with with $\widehat{p}_j^{I}(1)$ for the test statistic z_{31} decreases the bias of the variance but at the expense of a somewhat positive biased mean value. Comparing z_{32} in Table 2 with z_{52} for increasing bandwidth one can observe that increasing the smoothening implies a decrease of the variance towards the expected value of one.

Table 2 presents empirical coverage accuracy of simple percentile intervals for the 10 different estimators z_{lk} from nominal standard $N(0,1)$ confidence intervals. Hence coverages at the nominal level of $\alpha \in \{0.8, 0.9, 0.95, 0.97, 0.99\}$ show the relative numbers of the estimator z_{lk} that fall within the intervals $\left[-\Phi^{-1}\left(\frac{\alpha}{2}\right), \Phi^{-1}\left(\frac{\alpha}{2}\right)\right]$. The empirical coverage accuracy of z_{52} shows a nice recovering of the DGP, but as mentioned above this test statistic relies on the use of the true information of the distribution of the scores. Replacing this true information with the kernels based information in z_{32} with a bandwidth of 0.2 provides a coverage somewhat below the nominal value, which is to be expected because the variance is biased upwards even for a sample size above 150. Figure 4 shows the variance for z_{32} for increasing sample size and suggests that the variance is indeed biased. Hence the Monte Carlo study seems to suggest (accepting the premises in the form of the used assumption behind the DGP) that testing H_o using z_{32} we should refrain from rejecting H_o at e.g. 5 percent confidence with $|z_{32}| > 1.96$. We should allow $|z_{32}|$ to be as extreme as $\Phi^{-1}(0.9875) = 2.24$.

Table 3 and 4 present the summary statistics and coverage accuracy for z_{32} and z_{52} for varying sample size and bandwidth. We again see that increasing the bandwidth tends to decrease the variance of z_{32} towards one. We have also experimented with a kernel estimation of the bias corrected empirical score distribution. These results are not encouraging. The bias correction (Wilson (2008), Simar and Wilson (1998)) apparently imply a structural over-representation of scores in the lower part $[0.5, 0.75]$ of the support compared

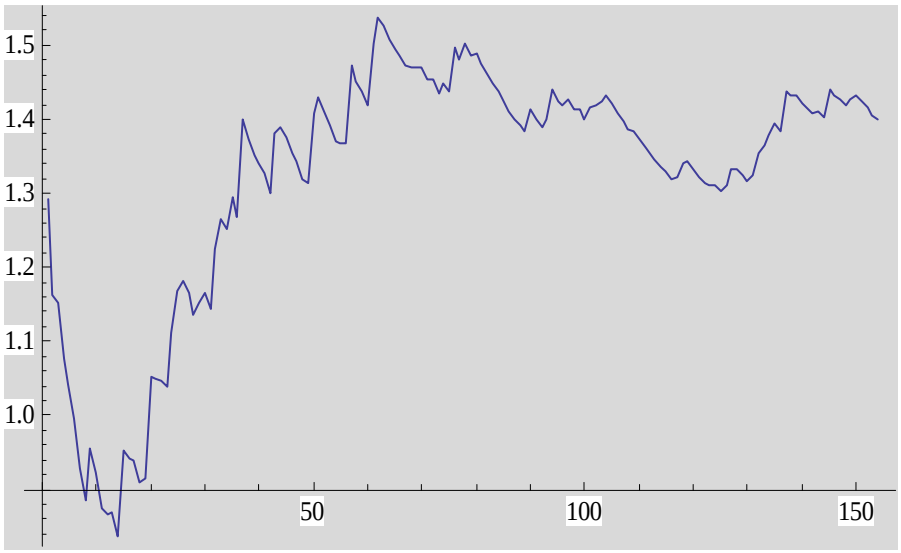

Figure 4: Variance for z_{32} for increasing sample size.

to the upper part $[0.75, 1]$. As illustrated in table A1 and A2 in Appendix this structural error implies that both the summary statistics and the coverage accuracy are far from being satisfactory. It is beyond the scope of this paper to analyze why the bias correction has this peculiar impact on the kernel based density estimation, and this is left for future research.

6 Final remarks

This paper has introduced the idea of testing for over-representation of data points in specific production zones. The approach was then operationalized based on discrete approximations of the densities of both the efficiency scores and the input mixes (angles). The test is non-parametric and being ratio-based it is scale (but not affinely) invariant. It relates to estimated technologies using only standard assumptions of convexity, ray unboundedness and minimal extrapolation. For practical applications the assumption of ray unboundedness (constant returns to scale in production) probably seem limiting, but in fact the DEA literature is full of empirical studies where the

constant returns to scale assumption seems justified, at least within a reasonable range of input-output values. Furthermore, well-established theoretical approaches, like the DEA based Malmquist index of productivity change, rely on constant returns to scale (see e.g. Wheelock and Wilson 1999).

That the practical test procedure in this paper is presented in a multiple input-single output setting alone, is simply for notational and computational convenience. The idea can easily be generalized to the multiple output scenario, but note that the increased dimensionality increases the probabilities of getting thinly populated combinations of facets and slices (for a given sample size), which reduces the strength of the test.

Several takeaways are available from the Monte Carlo simulation. Based on different choices of both facet probabilities and of probabilities of the discrete approximation to the assumed common inefficiency distribution we have illustrated the performance of the proposed test statistic. The best estimator of the common inefficiency distribution is apparently the estimator based on a kernel density estimation from the empirical score distribution (denoted $\widehat{p}_i^{\theta}(3)$) which has not been bias corrected. The estimator of probabilities of input mix or rather of the discrete approximation to the (assumed) common mix distribution based on the relative volumes of the facets in $\frac{x_2}{x_1} - \frac{x_3}{x_1}$ space (that is $\widehat{p}_j^{f}(2)$) apparently performs slightly better than the estimator based on the relative number of DMUs projected on to each facet (that is $\widehat{p}_j^{f}(1)$). The combination of $\widehat{p}_j^{f}(1)$ and $\widehat{p}_i^{\theta}(3)$ provides however a test statistic with the best coverage, but this test statistic has a somewhat positively biased mean value. Finally, it seems that the variance of the test statistic from the combination $\widehat{p}_i^{\theta}(3)$ and $\widehat{p}_j^{f}(2)$ is somewhat biased even for sample sizes above 150 in a three dimensional input space. Further research is needed to determine if that is in fact a general tendency, but the result seems to suggest caution when testing H_o using this test statistic.

References

[1] Banker, R.D., A Charnes and W.W. Cooper (1984), Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis, *Management Science*, 30, 1078-1092.

[2] Barber, C.B., D.P. Dobkin and H.T. Huhdanpaa (1996), The Quickhull Algorithm for Convex Hulls, *ACM Transactions on Mathematical*

Software, 22, 469-483.

[3] Bogetoft, P. and J.L. Hougaard (2003), Rational inefficiencies, *Journal of Productivity Analysis*, 20, 243-271.

[4] Daraio, C. and Simar, L. (Eds) (2007), Advancex Robust and Nonparametric Methods in Efficiency Analysis - Methodology and Applications, Springer New York.

[5] Fraley, C. and A. Raftery (1998), How Many Clusters? Which Clustering Method? – Answers Via Model-Based Cluster Analysis, *The Computer Journal*, 41, 578-588.

[6] Fraley, C. and A. Raftery (1999), Mclust: Software for Model-Based Clustering, *Journal of Classification*, 16, 297-306.

[7] Hartigan, J.A. (1975), Clustering Algorithms, John Wiley & Sons, New York.

[8] Hartigan, J.A. (1981), Consistency of Single-Linkage for High-Density Clusters, *Journal of the American Statistical Association*, 76, 388-294.

[9] Hartigan, J.A. (1986), Statistical Theory in Clustering, *Journal of Classification*, 2, 63-76.

[10] McLachlan, G.J. and D. Peel (2000), Finite Mixture Models, Wiley Series in Probability and Statistics.

[11] Muller, D.E. and Preparata, F.P. (1878), Finding the Intersection of two Convex Polyhedra, *Theoretical Computer Science*, 7, 17-236.

[12] Olesen, O.B and N.C. Petersen (2003), Identification and use of efficient faces and facets in DEA, *Journal of Productivity Analysis*, 20, 323-360.

[13] Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis, London: Chapman and Hall

[14] Siegel S. and N.J. Castellan (1988), Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill.

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

[16] Simar, L., and P.W. Wilson (2000), Statistical inference in nonparametric frontier models: The state of the art, *Journal of Productivity Analysis*, 13, 49-78

[17] Thanassoulis (2001), Introduction to the Theory and Application of Data Envelopment Analysis: A foundation text with integrated software, Kluwer Academic Publishers.

[18] Wheelock, D.C. and P.W. Wilson (1999), Technical Progress, Inefficiency, and Productivity Change in U.S. Banking, 1984-1993, *Journal of Money, Credit and Banking*, 31, 212-234.

[19] Wilson, P. W. (2008), "FEAR 1.0: A Package for Frontier Efficiency Analysis with R," *Socio-Economic Planning Sciences* 42, 247–254

[20] Wishart, D. (1969), Mode Analysis: A Generalization of Nearest Neighbor which Reduces Chaining Effects, in Cole, A.J. (Ed), Numerical Taxonomy, Academic Press, 282-311.

Appendix

The Minkowski-Weyl Theorem: *For a subset P of $\mathbf{R}^n$, the following statements are equivalent:*
(a) P is a polyhedron which means that there exist some fixed real matrix A and a real vector b such that $\{P = x : Ax \leq b\}$
(b) There are finite real vectors $v_1, v_2, \ldots . v_s$ and $r_1, r_2, \ldots, r_t$ in $\mathbf{R}^n$ such that

$$P = conv\{v_1, v_2, \ldots . v_s\} + cone\{r_1, r_2, \ldots, r_t\}$$

Results based on bis corrected scores:

Summary Statistics, Monte Carlo results					
z_{ij}	$Mean$	$Variance$	Min	Max	$Skewness$
$bandwidth = 0.1$					
z_{41}	1.66901	1.85947	-2.31623	4.72554	-0.128694
z_{42}	1.52079	2.17503	-2.23857	4.73	-0.0813843
$bandwidth = 0.2$					
z_{41}	0.987224	1.26812	-2.46287	3.46981	-0.170587
z_{42}	0.837708	1.53803	-2.27749	3.66136	-0.0971653

Table A1 Summary statistics of the testors (bias corrected scores)

Simulation Monte Carlo results					
	Nominal levels 0.8,0.9,0.95,0.975,0.99				
z_{ij}	0.8	0.9	0.95	0.975	0.99
$bandwidth = 0.1$					
z_{41}	0.386503	0.496933	0.539877	0.619632	0.730061
z_{42}	0.429448	0.478528	0.570552	0.680982	0.754601
$bandwidth = 0.2$					
z_{41}	0.558282	0.699387	0.785276	0.865031	0.920245
z_{42}	0.601227	0.699387	0.785276	0.846626	0.907975

Table A2 The accuracy of the coverage of the testors (bias corrected scores)