Maintaining the Regular Ultra Passum Law
in Data Envelopment Analysis

by

Ole B. Olesen

and

John Ruggiero

Maintaining the Regular Ultra Passum Law in Data Envelopment Analysis

Ole B. Olesen[1] and John Ruggiero[2]

(1) Department of Business and Economics

The University of Southern Denmark, Odense, Denmark, ole@sam.sdu.dk

(2) (correspondig author) School of Business Administration,

University of Dayton, Dayton, Ohio 45469-2251, USA, jruggiero1@udayton.edu

December 20, 2011

Abstract

The variable returns to scale data envelopment analysis (DEA) model is developed with a maintained hypothesis of convexity in input-output space. This hypothesis is not consistent with standard microeconomic production theory that posits an S-shape for the production frontier, i.e. for production technologies that obey the Regular Ultra Passum Law. Consequently, measures of technical efficiency assuming convexity are biased downward. In this paper, we provide a more general DEA model that allows the S-shape.

1 Introduction

The non-parametric data envelopment analysis (DEA) approach envelops observed data with a piecewise linear frontier. The characteristics of a DEA model are derived from a number of maintained assumptions imposed on the technology. A typical estimator used in DEA is the BCC-estimator(Banker, Charnes and Cooper 1984), which assumes the estimated production possibility set is a polyhedral that allows variable returns to scale. As a consequence, the BCC-estimator assumes that marginal product is non-increasing, which violates standard microeconomic theory where marginal product initially increases but diminishing returns eventually set in. In particular, if data reflects the Regular Ultra Passum (RUP) law ((Frisch 1965), Chapter 8), the BCC-estimator will be biased downward.

Definition 1 *The RUP law. Let a single output y be produced from a vector of m inputs x according to a production function $F(x,y) = 0$. This production function obeys the RUP law if $\frac{\partial \varepsilon(x,y)}{\partial x_i} < 0, i = 1, \cdots, m$ where the function $\varepsilon(x,y)$ is the scale elasticity, and for some point (x_1, y_1) we have $\varepsilon(x_1, y_1) > 1$, and for some point (x_2, y_2), where $x_2 > x_1, y_2 > y_1$, we have $\varepsilon(x_2, y_2) < 1$.*

The problem with the BCC-estimator is that the supporting hyperplanes for envelopment can overestimate inefficiency for points that should be projected to the local non-convex segments of the true frontier characterized by increasing returns to scale. In this paper, we are concerned with production technologies satisfying the RUP condition where the BCC-estimator is biased because such technologies are not convex in input-output space. Furthermore, existing measures of scale efficiency will be biased due to the improper projection to production impossibilities. The main contribution of this paper is the development of an approach that is capable of measuring inefficiencies for production possibilities in a non-convex homothetic and S-shaped technology. A non-convex S-shaped technology is characterized as follows: along any expansion path an expanding DMU with low activity will have a high scale elasticity greater than one. As the unit expands its activity the scale elasticity will decrease and will approach optimal scale size with an elasticity equal to one. Further expansion will imply decreasing returns with a scale elasticity less than one and approaching zero.

Several non-convex models exist in the literature (*e.g.*, the FDH-mode of Deprins et al. (1984) and the Petersen-Bogetoft approach (Petersen 1990),(Bogetoft 1996)) but these models are not well-suited to estimate an S-shaped production structure because any non-convex shape can result from these estimation procedures. In other words, we are looking for an estimation procedure that allows ONLY non-convexities that are reflected in an S-shaped production structure. For simplicity, we focus on production technologies that are homothetic. The concept of a homothetic production function was first introduced in ((Shephard 1953), page 30) as a monotonic transformation of a linear homogenous production function. With a homothetic production structure we can smooth the obtained structure of the estimated isoquant because homotheticity implies that the shape of the isoquants are identical. This allows us to maintain convexity in input (and output space) and to allow non-convexities in input-output space.

In order to move between input space and output space, we propose estimating individual isoquants assuming selective input convexity using a simplified order-m estimation procedure (Cazals, Florens and Simar 2002) where we avoid replications. The order-m estimation procedures include a conditional estimation model maintaining selective convexity of the input sets as proposed and formalized in (Daraio and Simar 2005)[1]. Under the assumption of homotheticity, we can aggregate inputs (and outputs) allowing us to move

---

[1] See also (Ruggiero 1996) and (Podinovski 2005).

to aggregate input-output space where we can impose an S-Shape.

The rest of the paper is organized as follows. In section 2 we define the production technology, from an input orientation using an input distance function. The assumption of homotheticity is presented and the implication for input aggregation is discussed. Notably, the assumption of homotheticity allows us to generate any isoquant from a base isoquant and hence, derive a well-defined index of aggregate input. Section 3 is devoted to the estimation of the base isoquant using a conditional estimator. We also discuss criteria for selecting a well-estimated isoquant among all possible base isoquants to aggregate inputs. This isoquant is used for the aggregation of inputs. In section 4, we develop a model to estimate a piecewise linear S-shaped frontier. Using simulated data in section 5, we show that our method overcomes the inherent problems of standard DEA and provides better estimates of inefficiency when the true technology obeys the Regular Ultra Passum Law. The last section concludes with directions for future research.

2   Production Technology

Let us consider a production environment where a vector of s inputs $X = (x_1, \ldots, x_s)$ is used in the production of one output Y. We represent the production technology with the input set $L(Y) = \{X \in \mathbb{R}_+^s : X$ can produce $Y\}$ which has isoquant

$$IsoqL(Y) = \{X : X \in L(Y), \lambda X \notin L(Y), \lambda \in [0, 1)\}. \tag{1}$$

Since we assume that only one output is produced, we can define a production function as

$$\phi(X) = \max\{Y : X \in L(Y)\} \tag{2}$$

The input distance function is then defined as

$$D_I(Y, X) = \max\{\gamma : X/\gamma \in L(Y)\}, \tag{3}$$

which provides an alternative characterization of the technology since $D_I(Y, X) \geq 1 \Leftrightarrow X \in L(Y)$. Finally, the index of technical efficiency proposed by Debreu (1951) and Farrell (1957) that serves as basis for DEA is given as

$$F_I(Y, X) = \min\{\gamma : \gamma X \in L(Y)\}, \tag{4}$$

where $F_I(y, x) = D_I(y, x)^{-1}$.

In this paper, we seek to place additional structure on the production technology. In particular, we assume that production is homothetic.

Definition 2 *A production function $\phi(X)$ is homothetic*

$$Y = \phi(X) = F(g(X))$$

where $F() : R_+ \to R_+$ is monotonic and $g(\lambda X) = \lambda g(X)$ i.e. $g()$ is positive homogeneous of degree one and continuously differentiable (see (Shephard 1970)). $g()$ is denoted the kernel function.

From the definition, we see that a homothetic production function can be represented as a production process whereby the input vector X can be aggregated into a one dimensional input index $g(X)$, i.e. output is determined from the level of aggregate input (see (Färe and Lovell 1988) for a more general result).

Proposition 3 *Assume a homothetic technology with one output. The distance function evaluated at $(1, X)$ is equal to aggregate input defined from the core function in the homothetic production function multiplied by a constant, i.e.*

$$D_I(1, X) = k \times g(X), k \in \mathbb{R}_+$$

Proof. Let $\phi(X) = F(g(X))$ with $F^{-1} = f$. We know that

$$
\begin{aligned}
L(y) &= \{X : F(g(X)) \geq y\} \\
&= \{X : g(X) \geq f(y)\}
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
D_I(1, X) &= \max\{\gamma : X/\gamma \in L(1)\} \\
&= \max\{\gamma : X/\gamma \in \{X : g(X) \geq f(1)\}\} \\
&= \max\{\gamma : g(X/\gamma) \geq f(1)\} \\
&= \max\left\{\gamma : \frac{1}{\gamma}g(X) \geq f(1)\right\} \\
&= \max\{\gamma : g(X) \geq \gamma f(1)\} \\
&= \{\gamma : g(X) = \gamma f(1)\} \\
&= (f(1))^{-1} \times g(X)
\end{aligned}
$$

∎

Proposition 3 establishes that the dimensionality of DEA models can be reduced under the assumption of homotheticity. In addition, homotheticity allows us to span the production technology from $L(1)$ (see (Shephard 1970), page 34).

$$L(Y) = H(Y)L(1), \tag{5}$$

where $L(1)$ is the input set associated with the unit isoquant and $H(Y)$ is the scaling function. Input sets can be theoretically generated from a base input set by the scaling function that depends only on the level of output and not the input mix. From (5), we also have

$$IsoqL(Y) = H(Y)IsoqL(1). \tag{6}$$

This shows that we can generate any isoquant from the unit isoquant. More generally, we could choose any output level and its associated isoquant to serve as the base. Here, we choose the unit isoquant for expositional convenience only. In the next section, we consider the estimation of a base isoquant and provide guidance on selecting a well-estimated base for aggregation purposes.

3   Estimating the Base Isoquant

One useful method for estimating any isoquant is the order-m estimation procedure (Daraio and Simar 2005). The input distance function $D_I(y,x)$, defined in (3) is expressed relative to the input set $L(y)$ and the basic idea in the order-m procedure is to regard this input set $L(y)$ as the support of a conditional density function $L(y) = \{x : F_{X|Y}(x|y) > 0\}$. The corresponding support for the joint input output density $H_{X,Y}(x,y)$ is the production possibility set T, i.e. $T = \{(x,y) : H_{X,Y}(x,y) > 0\}, H_{X,Y}(x,y) = \Pr(X \leq x, Y \geq y) = \Pr(X \leq x|Y \geq y)\Pr(Y \geq y) = F_{X|Y}(x|y)S_Y(y)$, where $S_Y(y) = \Pr(Y \geq y)$. For a fixed level of output y_o let $X_1, \ldots, X_m$ be m i.d.d. random input vectors generated from $F_{X|Y}(.|y_o)$, i.e. all input vectors $X_i, i = 1, \ldots, m$ are random variables that can produce y_o with a strict positive probability. Assuming selective (local) convexity of the input sets, the random input set of order-m[2] for units producing y_o, $L_m^C(y_o)$ is defined as:

$$L_m^C(y_o) = Conv\left[\{x|x \geq X_i, i = 1, \ldots, m\}\right] \tag{7}$$

---

[2] (Daraio and Simar 2005) use the phrase "random producton set of order-m for units producing more than y: $T_m^C(y_o) = Conv\left\{(x,y) \in \mathbb{R}_+^{m+1}|x \geq X_i, y \geq y_0, i = 1, \ldots, m\right\}$

However, this set is not a production set since it is unbounded in the output dimensions. We use the related input set $L_m^C(y_o)$.

The locally convex order-m input efficiency $\theta_m^{LC}(x,y)$ is defined as ((Daraio and Simar 2005), p 17, (3.2.))

$$\theta_m^{LC}(x,y) = E_{X|Y}\left[\widetilde{\theta}_m^{LC}(x,y)\,|Y \geq y\right]$$

where

$$\widetilde{\theta}_m^{LC}(x,y) = \inf\left\{\theta|\theta x \in L_m^C(y)\right\}$$

To obtain the estimator $\widehat{\theta}_m^{LC}(x,y) = \widehat{E}_{X|Y}\left[\widetilde{\theta}_m^{LC}(x,y)\,|Y \geq y\right]$ based on a sample of n observations we plug in the empirical version of $F_{X|Y}(.|y_o)$ as $\widehat{F}_{X|Y,n}(x|y) = \frac{\sum_{i=1}^n 1(X_i \leq x, Y_i \geq y)}{\sum_{i=1}^n 1(Y_i \geq y)}$,where $1()$ is the indicator function. $\widehat{\theta}_m^{LC}(x,y)$ can be approximated by a Monte-Carlo procedure: Sample m observations $X_{1,b},\ldots,X_{m,b}$ conditional on output being greater than $y_o = 1$ with replacement. For each of the n observations find the inverse distance function value $\widetilde{\theta}_m^{LC,b}(X_l,1)$ relative to an input set $Conv\left[\{x|x \geq X_{i,b}, i = 1,\ldots,m\}\right]$. Redo this estimation $b = 1,\ldots,B$ and take the average of the obtained scores as the estimator, i.e. $\widehat{\theta}_m^{LC}(X_l,1) \approx B^{-1}\sum_b \widetilde{\theta}_m^{LC,b}(X_l,1)$. From these scores we obtain an estimated input set $\widehat{L}^{LC}(1)$ as

$$\widehat{L}_m^{LC}(1) = Conv\left[\widehat{\theta}_m^{LC}(X_1,1) \times X_1,\ldots,\widehat{\theta}_m^{LC}(X_n,1) \times X_n\right] + \mathbb{R}_+^m. \tag{8}$$

A simplification of the order-m estimator is the conditional estimator of the base isoquant, which avoids the replications by choosing $m = n$.

In the simulations presented later in the paper, we use this conditional estimator instead of the order-m estimator. The base isoquant can be estimated using this conditional model solving the following linear programs

$$\widehat{\theta}_C^{LC}(X_l, Y_{base}) = \begin{cases} \min & \theta - \varepsilon\,(1,\ldots,1)\,s \\ s.t. & \theta X_l - \sum_{j=1}^{m_1} \lambda_j X_j - s = 0 \\ & \sum_{j=1}^{m_1} \lambda_j = 1 \\ & \lambda_j = 0 \text{ if } Y_j < Y_{base} \\ & \lambda \in \mathbb{R}_+^{m_1}, s \in \mathbb{R}_+^s \end{cases} \tag{9}$$

$l = 1,...,n$, where $Y_{base} = 1$ in this section and where again the estimator $\widehat{L}_C^{LC}(1)$ of the input set is derived as

$$\widehat{L}_C^{LC}(1) = Conv\left[\widehat{\theta}_C^{LC}(X_1,1) \times X_1,\ldots,\widehat{\theta}_C^{LC}(X_n,1) \times X_n\right] + \mathbb{R}_+^m. \tag{10}$$

This model appears in the efficiency literature to control for exogenous inputs (Ruggiero 1996), selective convexity (Podinovski 2005) and as the condition estimator (Daraio and Simar 2005). In this formulation,

units that are not observed producing at least the base amount (in this case, one) are not allowed in the solution space. Hence, we simply envelop all input vectors with observed output at least equal to one. Notably, we replace the standard assumption of convexity with selective input convexity of the input sets:

Axiom: Selective input convexity: If $(X', Y') \in T, (X'', Y'') \in T, Y'' > Y' \Rightarrow \lambda (X', Y') + (1 - \lambda) (X'', Y') \in T, \lambda \in [0, 1]$

Our primary reason for using the conditional model is NOT to estimate efficiencies but to exploit homotheticity to aggregate multiple inputs into a one-dimensional input index. Hence, we estimate each isoquant using the conditional estimator and choose the "best" isoquant that has good coverage in the sense that i) we want as many observations playing an active role of spanning the frontier, ii) we want the cone spanned by these observation to be as large as possible and iii) we want the observations to be spread out across the cone as uniformly as possible. After choosing the isoquant that best meets the desirable criteria, we then estimate the distance of each observation to this isoquant as an index of aggregated input.

To ease the presentation of the proposed methodology, we chose the unit isoquant as the base in our discussion above. We now provide guidelines for how to choose the output level with the most useful information. Using the conditional estimator relative to a given output level y we only include input vectors from observations with an output level at least equal to this y. We would like to have as many observations as possible available for spanning the isoquant, which tends to suggest a low output level. However, observations producing output much larger than y may not provide any additional information. If we knew the positions and the shapes of the true isoquants we would look for a specific isoquants (a y level) where i) we have many observed points on or just above the isoquant and ii) where the points are spread out evenly along the full isoquants. Unfortunately, we do not know the locations and the shape of the true isoquants. Hence, we have to rely on an estimator, and in this case we will use the conditional estimator defined above. For each observed output level $Y_j, j = 1, \ldots, n$, we use the conditional estimators $\widehat{\theta}_C^{LC} (X_l, Y_j), l = 1, \ldots, n$ which provides us with the estimators $\widehat{L}_C^{LC}(Y_j)$ of all n input sets corresponding to all n outputs. As base isoquant we now choose the specific output level which performs reasonably well according to the following two criterias[3]:

1. A distribution of the angle coordinates of the observed data points on the conditional piecewise linear estimator of the isoquant, which mimics the uniform distribution on the empirical support of the angle coordinates for the whole data set. As a measure of the amount of deviation of the empirical distribution from the uniform distribution we suggest the area between the two distribution functions.

2. A large number of observed data points is located on the conditional piecewise linear estimator of the

---

[3]Of course, other criteria may be relevant. We leave for future research the evaluation of isoquant coverage.

isoquant[4].

The selection process for choosing the "best" base isoquant is implemented as follows. Estimate all $n \times n$ conditional scores providing $\widehat{\theta}_C^{LC}(X_l, Y_j)$, $j, l = 1, \ldots, n$, where this score is missing, if $Y_l < Y_j$. For a given observed output level Y_{j_o} keep the input vectors X_l if $\widehat{\theta}_C^{LC}(X_l, Y_{j_o})$ is non-missing and if $\widehat{\theta}_C^{LC}(X_l, Y_{j_o}) = 1$. Let the number of input vectors that satisfy these two conditions be $n_{Y_{j_o}}$ and let the different angles-vectors in the polar representation of these input vectors be η_l, $l = 1, \ldots, n_{Y_{j_o}}$. The resulting set of angles corresponds to observations on the estimated isoquant at output level Y_{j_o}. For the case of two inputs (this is the case covered in the included simulations) we only have one angle in the polar representation of the input vectors. Hence, we sort the angles and plot the points $\left(\frac{\eta_j - S_L}{S_R - S_L}, \frac{j}{n_{Y_{j_o}}} \right)$, $j = 1, \ldots, n_{Y_{j_o}}$, where S_L, S_R are the left and right endpoint of the support of the angle distribution (empirical estimates), see Figure 1. The deviation of this empirical distribution from the uniform distribution is measured as the area between the 45 degree line and the piecewise linear curve going through these $n_{Y_{j_o}}$ points, starting at the origin and ending at $(1,1)$[5].

Figure 1 illustrates this deviation for isoquant 750 used as the best base isoquant in the simulation study presented below in Section 5.

---

[4] An order m estimation of the base isoquant would allow us to expand the relevant angles η_l, of the input vectors X_l, where $\widehat{\theta}_{m_1}^{LC}(X_l, Y_{j_o})$ is non-missing and where $\widehat{\theta}_{m_1}^{LC}(X_l, Y_{j_o})^{-1} < 1 + \delta$, for some small $\delta \in \mathbb{R}_+$, and where $\widehat{\theta}_{m_1}^{LC}(X_l, Y_{j_o})$ is the corresponding order m estimator

[5] The area under the piecewise linear curve determined from the absolute deviation between the two curves can be determined as a sum of areas of a combined rectangle and a triangle. Consider three point in this deviation space $(\eta_j, z_j), j = 1, 2, 3$ where $\eta_{j+1} > \eta_j$ and let $z_2 > z_1$ and $z_3 < z_2$. The area spanned by $(\eta_j, z_j), j = 1, 2$ is $\left(\frac{1}{2}(z_2 - z_1) + z_1 \right)(\eta_2 - \eta_1) = \frac{1}{2}(z_2 + z_1)(\eta_2 - \eta_1)$. The area spanned by $(\eta_j, z_j), j = 2, 3$ is $\left(\frac{1}{2}(z_2 - z_3) + z_3 \right)(\eta_3 - \eta_2) = \frac{1}{2}(z_3 + z_2)(\eta_3 - \eta_2)$. Hence, the areas, except for the first and the last traingle are determined as $\sum_{j=2}^{n_{Y_o}} \frac{1}{2}(z_j + z_{j-1})(\eta_j - \eta_{j-1})$. Similarities to the Gini coefficient are apparent.
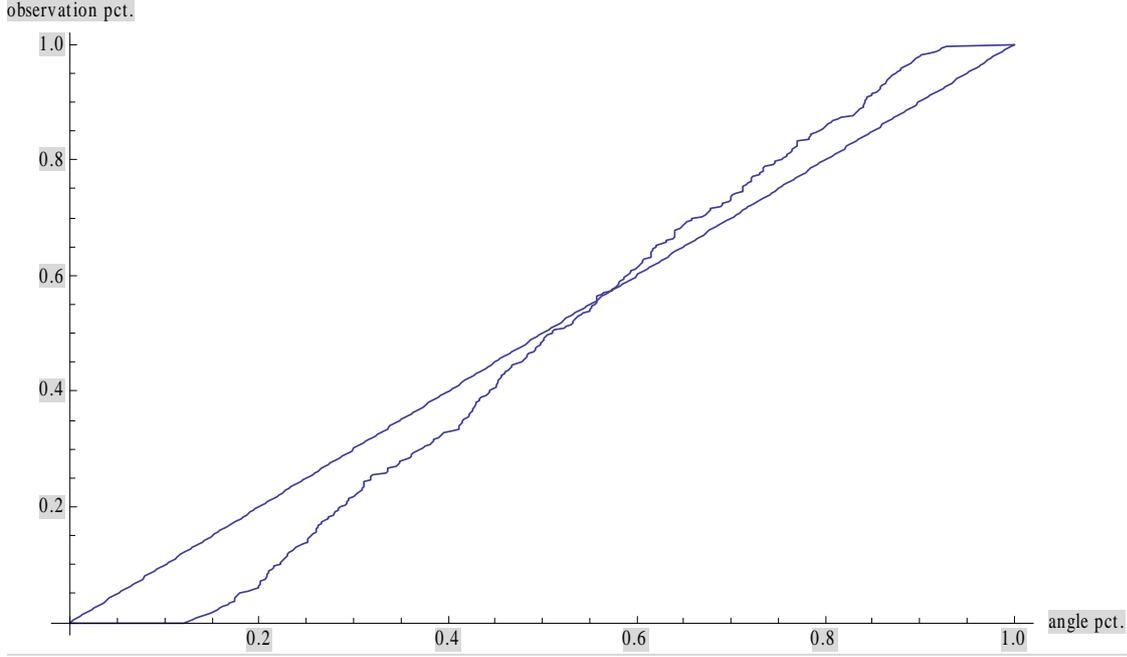
Figure 1. An empirical angle distribution and a uniform distribution.

Suppose we have identified $Y^* = Y_k$ as the output level associated with our chosen base isoquant. Then, an index measure of aggregate[6] input is $\widehat{x}(X_l, Y_l) = \left(\theta_c^{LC}(X_l, Y^*)\right)^{-1}$ for $l = 1, \ldots, n$ where $\theta_c^{LC}(X_l, Y^*)$ is determined as the solution to the following linear program:

$$\theta_c^{LC}(X_l, Y^*) = \begin{cases} \min & \theta - \varepsilon(1, \ldots, 1)s \\ s.t. & \theta X_l - \sum_{j=1}^{m_1} \lambda_j X_j - s & = & 0 \\ & \sum_{j=1}^{m_1} \lambda_j & = & 1 \\ & \lambda_j = 0 \text{ if } Y_j < Y^* \\ & \lambda \in \mathbb{R}_+^{m_1}, s \in \mathbb{R}_+^s, \theta \in \mathbb{R} \end{cases} \tag{11}$$

$, l = 1, ..., n.$

4 An estimator of a piecewise linear S-shaped frontier and the inflection point.

Our estimate of aggregate input (11) allows us to analyze an estimator of the S-shaped technology in the single (aggregate) input single output case.[7] With homothetic production, of course, we can also aggregate

---

[6] We distinguish between an index of aggregated inputs from (11) with or without non-zero slacks present in the optimal solution. As we will see in Section 5 we can nicely recover the true efficiency for data points without non-zero slacks present.

[7] Maintaining the RUP law requires that the scale elasticity is monotonically decreasing for increasing production. However, it is well known that along any facet below mpss we will have constant marginal product and decreasing average product

multiple outputs into a single output aggregate. Let the true production possibility set (PPS) be denoted T^S, and we will assume that the boundary of T^S is S-shaped in the sense that we can divide the input axis into two parts $[0, x^*]$ and $[x^*, \infty)$ where the production function is convex (concave) on the first (second) interval. Hence, the marginal product is monotonically non-decreasing in $[0, x^*]$ and monotonically non-increasing in $[x^*, \infty)$. (see (Ginsberg 1974) for an example of such a convex-concave production function). We know of course that the convex hull estimator $\widehat{T}^{BCC}$ (Banker et al. 1984) of the PPS is too large below x^*, but we also know that for input and output above the inflection point x^* this estimator works well, because of the true concave shape of the production function. Hence, in the following we will remove or "dig out" the part of the estimator $\widehat{T}^{BCC}$, that violates the S-shape. To be more precise, we will dig out a certain convex hull of observed data point that satisfies the following:

- the convex hull is spanned by points below (and on) the inflection point, i.e. points that are supposed to reflect the convex IRS part of the technology

- the convex hull is constructed such that no point is located above the frontier (or equivalently, no points are located in the interior of this hull)

Figure 2 illustrates this idea using 6 input output observations generated from an "S-shaped" data generating process (DGP). Observations, A, E and F are BCC-efficient and observation E is most productive scale size (mpss). In this small illustrative example we use the mpss as an estimator of the inflection point[8]. In other words we assume that the production function is convex up to data point E and concave to the right of this point. The basic idea behind the digging approach is to determine a subset of all FDH-efficient DMUs "below" mpss which determines a convex hull $\widehat{T}^{Dig}$, where none of these DMUs belongs to the interior of this hull. An estimator $\widehat{T}^S$ of the PPS with an S-shape with an efficient boundary being piecewise linear is now available as $\widehat{T}^{BCC} \backslash \widehat{T}^{Dig} \equiv \widehat{T}^S$, i.e. the convex hull BCC estimator of the PPS minus the convex hull $\widehat{T}^{Dig}$. In Figure 2 the estimated $\widehat{T}^{BCC}$ is the convex hull of observations A, E, F set added to $\mathbb{R}_+ \times \mathbb{R}_-$ (strong input and output disposability). $\widehat{T}^S$ is estimated as $\widehat{T}^{BCC} \backslash \widehat{T}^{Dig}$, where $\widehat{T}^{Dig}$ is the convex hull of the observations A, B, C, D, E.

(Førsund and Hjalmarsson 2004), which seems to imply that the RUP-law is violated. However, this violation disappears asymptotically, see (Olesen and Petersen 2011).

[8] In section 5 below we will propose a more general estimator of the inflection point.
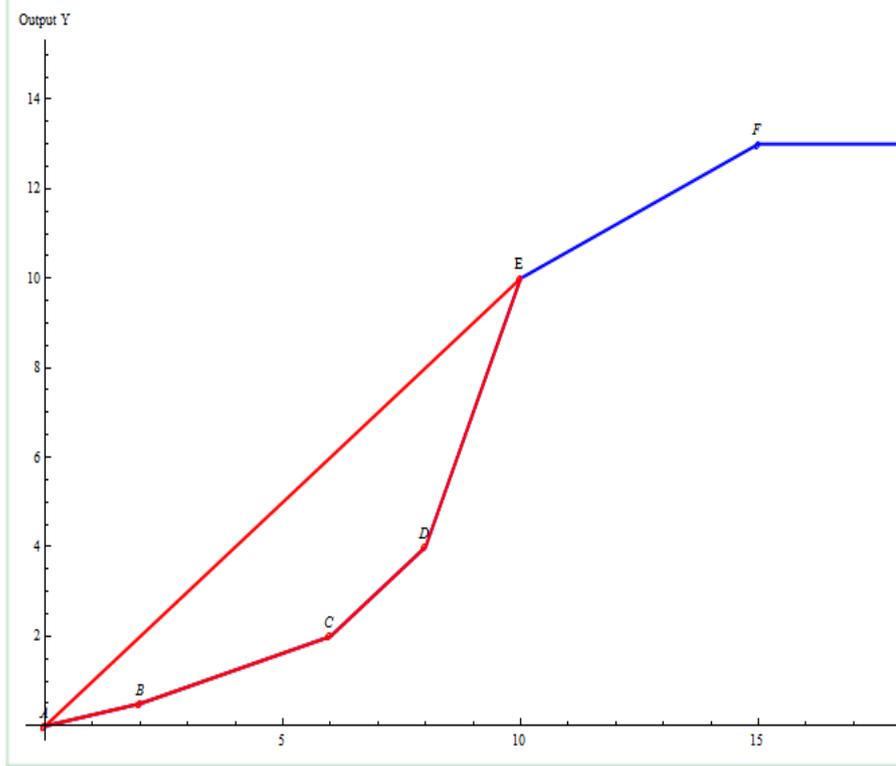
Figure 2. Illustration of the BCC- and the S-shaped frontier

The piecewise linear strongly efficient frontier in Figure 2 of $\widehat{T}^S$ is $ABCDEF$. In Figure 3 the FDH step function is included. Since only FDH efficient points are allowed to influences the estimation we may have data points present only within the four triangles bounded by the FDH-steps and the frontier $ABCDE$. Observations $H, I, J,$ and K are indicated below facet CD. Observation I is consistent with the frontier $ABCDE$ in the sense that if I is present then we simply "dig" a deeper hole into $\widehat{T}^{BCC}$ providing the efficient frontier as $ABCIDE$. This is possible without changing the rest of the hull because observation I is above the extension of both facet BC and facet DE. Hence, including I on the frontier still gives us a monotonic non-decreasing marginal product moving from C to I to D. Notice however, that neither H nor J, K are consistent with the frontier $BCDE$. These three additional observations share the characteristic of being either below the extension of facet BC or below the extension of facet DE, or both.

The determination of this inverted convex hull is unfortunately not unique. This is illustrated in Figure 3 by the point L being below the extension of the facet CD. Hence, we cannot include L on the S-shaped frontier and at the same time maintain that A, B, C, D and E all are on the frontier. However, we may include L as being on the frontier if we remove C from the frontier, as indicated by the dashed convex hull. Hence, we have a choice. Either C or L is efficient, but not both. Below we will partly resolve this non-uniqueness of the solutions by searching for the solution that maximizes the number of FDH-efficient

11

points on the frontier[9].
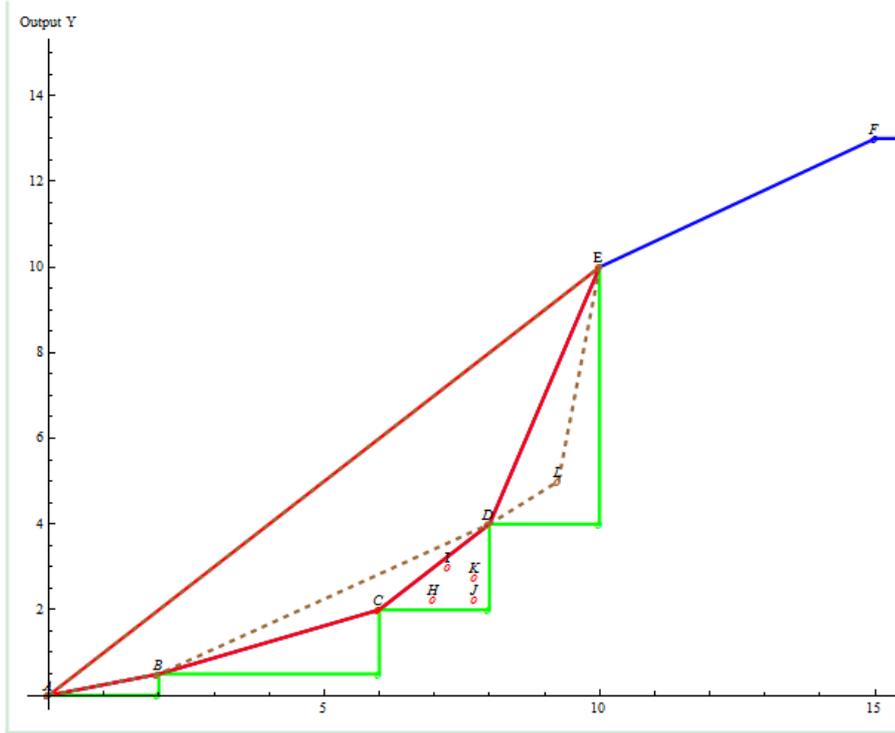
Figure 3. Illustration of the BCC- and the S-shaped frontier

Assuming that we know the position of the inflection point a simple procedure to determine an arbitrary inverted convex hull is as follows:

Step 1. Generate the FDH efficiency scores for all units below the inflection point. Remove all FDH-inefficient point.

Step 2. W.l.o.g. let $(X_j, Y_j), j = 1, \ldots, n'$ be the FDH-efficient point and project each of these point towards higher input levels within the convex hull of these DMUs, using the following program

$$
\begin{aligned}
\max \quad & \sum_{k=1}^{n'} \theta_k \\
s.t. \quad & \sum_{j=1}^{n'} \lambda_{j,k} X_j - \theta_k X_k & = \quad & 0 & k = 1, \ldots, n' \\
& \sum_{j=1}^{n'} \lambda_{j,k} Y_j \theta_k - Y_k & = \quad & 0 & k = 1, \ldots, n' \\
& \sum_{j=1}^{n'} \lambda_{j,k} & = \quad & 1 & k = 1, \ldots, n' \\
& \lambda \in \mathbb{R}_+^{m \times n}, \theta_k \in \mathbb{R}, \forall k
\end{aligned}
\tag{12}
$$

---

[9]In the case of a tie between several alternative solutions we suggest that each alternative solution is used to provide an efficiency evaluation of the observed points.

Step 3. For each optimal $\theta_k^* > 1$, we know that under mild regularity conditions at most two components among $\lambda_{1,k}, \ldots, \lambda_{n',k}$ are strictly positive. Remove one of these two FDH-efficient DMUs from the sample, but never remove the two DMUs with the largest and the smallest input value

Step 4. Do step 2-3 until $\theta_k^* = 1$, for all index k in the remaining set of DMUs.

We denote this approach *the filling approach*, because after "digging" an inverted convex hull including *all* FDH-efficient DMU, we modify the hull by making the hull smaller by removing FDH-efficient DMUs that force some other FDH efficient DMUs to be located in the interior of the hull.

Assuming that we know the position of the inflection point this approach will provide us with an inverted hull $\widehat{T}^{Dig}$ and will thereby provide a piecewise linear estimator of the S-shaped technology given as $\widehat{T}^S \equiv \widehat{T}^{BCC} \backslash \widehat{T}^{Dig}$. However, since the estimator is not unique the determination of which FDH-efficient points to include on the increasing returns to scale part will depend on the which point we choose to delete in step 3 above and in what sequence such points are removed.

We now consider an alternative procedure that provides a piecewise linear estimator of the S-shaped technology using the inverted convex hull that maximizes the number of FDH efficient points on the S-shaped frontier, i.e that maximizes the number of "S-shaped efficient" points. An integral part of this procedure is an estimation of the inflection point; we seek as an inflection point below mpss one that allows for an estimated S-shaped frontier with a maximum number of FDH efficient points on the frontier. Testing a given point as a candidate for the inflection point involves several conditions: i) the marginal products along the facets from the origin to the inflection point must be increasing, ii) the marginal products must be non-increasing on facets above the inflection point and iii) all the points have to be on or below the frontier.

Let us consider FDH-efficient points below an inflection point candidate (X_n, Y_n). $(X_j, Y_j) \in R_+^2, j = 2, \ldots, n+1$, $(X_1, Y_1) = (0,0)$ and $\frac{Y_{j+1}}{X_{j+1}} > \frac{Y_j}{X_j}, j = 1, \ldots, n-1, \frac{Y_{n+1}}{X_{n+1}} < \frac{Y_n}{X_n}$. We are looking for a convex shape as a graph through a subset of the points $(X_j, Y_j), j = 1, 2, \ldots, n+1$, starting at $(X_1, Y_1) = (0,0)$ and ending at the estimator of the inflection point (X_n, Y_n). This problem resembles the so-called traveling salesman problem (TSP)(Dantzig, Fulkerson and Johnson 1954), which consists of finding the shortest path through a set of points, never visiting a point more than once and returning to the starting point. In our problem the length of the path does not matter and it is not required that we return to the starting point. But it is required that we start at point 1 and end at point n. Secondly, we do not require that the path covers all points. In fact, we expect only a subset of points to be covered, but we maximizes the number of points visited upto and including the inflection point. Thirdly, the $(i+1)$*th* edge is required to have a larger marginal product compared to the i*th*. Finally, all points have to be on or below the frontier.

As in the TSP we use binary variables b_{ij}, where $b_{ij} = 1$ if the edge from point i to j is used, otherwise $b_{ij} = 0, i, j \in \{1, \ldots, n\}$. We only consider a $b_{ij} = 1$ as feasible if $j > i$, since data are sorted and a feasible

convex path never will go from i to j, where $j < i$. A requirement of at most one path into the $k'th$ point and at most one path out of the l'th point is modelled with slightly modified assignment constraints:

$$\sum_{k=1}^{n} b_{kl} + s_{l}^{IntoNode} = 1, l = 2, \ldots n \tag{13}$$

$$\sum_{l=1}^{n} b_{kl} + s_{k}^{OutOfNode} = 1, k = 1, \ldots n-1 \tag{14}$$

where $s_{l}^{IntoNode} \geq 0, \forall l$, $s_{k}^{OutOfNode} \geq 0, \forall k$. If $s_{l}^{IntoNode} = 1$ or $s_{l}^{OutOfNode} = 1$ then the l'th point is not on the "path" that constitutes the convex part of the S-shaped estimator of the production function[10].

If $b_{kl} = 1, \forall k, l, k < l, k \neq 1, l \neq n$ then an edge out of k and into l is used and we require that an edge into k and an edge out of l must be used, i.e. $\sum_{i=1}^{n} b_{ik} = 1$, $\sum_{j=1}^{n} b_{lj} = 1 \Longrightarrow s_{k}^{IntoNode} = s_{l}^{OutOfNode} = 0$, or

$$s_{k}^{IntoNode} + s_{l}^{OutOfNode} \leq (1 - b_{kl}) M_2, \forall k, l, k < l, k \neq 1, l \neq n \tag{15}$$

where M_2 is a large number (here M_2 must be greater than 2). If $b_{1l} = 1$ then an edge into point l is used and we require that an edge out of l must be used, i.e. $\sum_{j=1}^{n} b_{lj} = 1$ implying that $s_{l}^{OutOfNode} = 0$ or

$$s_{l}^{OutOfNode} \leq (1 - b_{1l}) M_2, l \in \{2, \ldots, n-1\} \tag{16}$$

If $b_{kn} = 1$ then an edge out of k is used, and we require that an edge into k must be used, i.e. $\sum_{i=1}^{n} b_{ik} = 1 \Longrightarrow s_{k}^{IntoNode} = 0$, or

$$s_{k}^{IntoNode} \leq (1 - b_{kn}) M_2, l \in \{2, \ldots, n-1\} \tag{17}$$

We prefer an estimator of the convex part of the frontier with as many points on the frontier as possible. Hence, we maximize the sum of the binary variables b_{ij}. Model (18) presents the full optimization problem:

---

[10]Notice that if $s_{l}^{IntoNode} = s_{l}^{OutOfNode} = 0, \forall l$ these constraints are the well known assignment constraints from the LP-formulation of the TSP-problem making sure that one arc into and out of every node are used in a feasible solution.

$$\max \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}$$

s.t.

$$
\begin{array}{llll}
\frac{Y_k - Y_j}{X_k - X_j} - \frac{Y_j - Y_i}{X_j - X_i} - s_{ijk} & = & 0 & i,j,k \in \{1,\ldots,n\}, i < j, j < k \quad (18.1) \\[2mm]
s_{ijk} + (2 - b_{ij} - b_{jk}) M_1 & \geq & 0 & i,j,k \in \{1,\ldots,n\}, i < j, j < k \quad (18.2) \\[2mm]
& & & \hspace{5.5cm} (18.3) \\[2mm]
\sum_{k=1}^{n} b_{kl} + s_l^{IntoNode} & = & 1 & l \in \{2,\ldots,n\} \hspace{2.2cm} (18.4) \\[1mm]
& & & \text{into } l \text{ if } s_l^{Into} = 0 \\[2mm]
\sum_{l=1}^{n} b_{kl} + s_k^{OutOfNode} & = & 1 & k \in \{1,\ldots,n-1\} \hspace{1.3cm} (18.5) \\[1mm]
& & & \text{out of } k \text{ if } s_k^{OutOf} = 0 \\[2mm]
s_k^{IntoNode} + s_l^{OutOfNode} & & & \hspace{5.5cm} (18) \\[1mm]
\quad - (1 - b_{kl}) M_2 & \leq & 0 & \forall k,l, k < l, k \neq 1, l \neq n \hspace{0.8cm} (18.6) \\[2mm]
s_l^{OutOfNode} - (1 - b_{1l}) M_2 & \leq & 0 & l \in \{2,\ldots,n\} \ \ (\text{only } b_{1l}) \hspace{0.6cm} (18.7) \\[2mm]
s_k^{IntoNode} - (1 - b_{kn}) M_2 & \leq & 0 & k \in \{1,\ldots,n-1\} \ (\text{only } b_{kn}) \quad (18.8) \\[2mm]
\frac{Y_t - Y_n}{X_t - X_n} - \frac{Y_n - Y_k}{X_n - X_k} + d_k^{term} & = & 0 & k \in \{1,\ldots,n-1\} \hspace{1.3cm} (18.9) \\[2mm]
d_k^{term} + (1 - b_{kn}) M_1 & \geq & 0 & k \in \{1,\ldots,n-1\} \hspace{1.2cm} (18.10) \\[2mm]
b_{ij} & = & 0 & i,j \in \{1,\ldots,n\}, i \geq j \\[2mm]
b_{ij} \in \{0,1\} \, \forall, i, j \\[2mm]
s_l^{IntoNode} \geq 0, s_k^{OutOfNode} \geq 0 \\[2mm]
d_k^{term} \in \mathbb{R}^{n-1}
\end{array}
$$

The constraints (18.1-2) are included to only allow a sequence of edges with increasing marginal product as feasible, where M_1 is a large number. The constraints (18.9-10) are included to require that the marginal product starts to decrease when passing through the inflection point, i.e. when moving from the convex part to the concave part of the S-shaped frontier. The structure in (18.8) allows only one edge into point n. Let us denote the starting point of this edge into point n as k_o, i.e. $b_{k_o n} = 1$. For all $k \neq k_o$ (18.9) implies redundant constraints, $d_k^{term} \geq -M_1$. Given the inflection point, the "termination" point of the convex part denoted (X_t, Y_t) is determined as the point with the maximal rate of transformation relative to the estimator of the inflection point, i.e.

$$(X_t, Y_t) = \arg \max_{j > n} \left(\frac{Y_j - Y_n}{X_j - X_n} \right).$$

Unfortunately, (18) does not provide a path from the origin to the inflection point with no uncovered point above the path. Hence we have to supplement (18) with the following cutting procedure.

1. Solve (18).

2. For each $b_{ij} = 1$ loop through all points with the input component larger than X_i and smaller than X_j and check if any such point is located "above" the facet spanned by $[(X_i, Y_i), (X_j, Y_j)]$.

3. If any such observation is above this facet then add the constraint $b_{ij} = 0$ to (18) and goto step 1.

This procedure will either terminate with a feasible solution providing a convex part from origin to the inflection point with all points on or below the path, or with a status being integer infeasible, in which case no path exists with the required characteristics.

This approach was applied to simulated data in the next section. In anticipation of our results see Figure 4, which shows estimate of the S-shape technology with the endogenous inflection point identified as point 66. In the next section, we analyze our approach using simulated data.

5   Simulation

We will show that our approach is capable of recovering the true S-shaped technology while simultaneously providing better estimates of technical inefficiency. Assuming one output, two inputs and homotheticity we generate data according to the following data generating processes (DGP). We specify a generalized production function (Zellner and Revankar 1969)

$$Y = \phi(X) = F(g(X))$$

where the scaling law is $F(z) = \frac{15}{1+e^{-5\ln(z)}}$ and the linear homogenous core function is CES: $g(x_1, x_2) = \left(\beta x_1^{\frac{\sigma-1}{\sigma}} + (1-\beta) x_2^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$, with $\beta = 0.45$ and $\sigma = 1.51$. Data are generated for 1000 DMUs as follows; inputs are generated in polar coordinates as angles η and modulus ω uniformly distributed on $\left[0.05, \frac{\pi}{2} - 0.05\right]$ and $[0, 2.5]$, respectively. Output is generated from the generalized production function $F(g(\omega\cos\eta, \omega\sin\eta))$. Inefficiency is added to the input vectors with $X = e^{\theta} \times (\omega\cos\eta, \omega\sin\eta)$, where θ is a random variable from a truncated normal distribution with standard deviation 0.2.

We sort the data on output and estimate component and estimate (11) for each of the 1000 output levels; the solution space for each esimated isoquant is conditioned such that only DMUs with outputs greater than or equal to the i'th DMUs output, $i = 1, \dots, 1000$ is included. We thus obtain 1000 input oriented scores for each isoquant. If the input oriented score has additional slack in either input dimensions the score is assigned the value "missing". Based on the these results, we identify for each potential base isoquant only those points that span the conditional isoquants (i.e. only observations with input oriented score equal to one with no additional slack are included). For this simulation, we consider two criteria for choosing our base isoquant. Firstly, we are looking for the particular isoquant with as many observation on the

frontier as possible. Secondly, we search for an isoquant with an empirical distribution of the angles of these points spanning the frontier as close as possible to a uniform distribution. Taken together, we are looking for isoquants that have a lot of points that uniformly span the isoquant. For our simulation, we identified numerous isosquants that performed well on both criteria. This result, while not surprising, is encouraging: the selection of a base isoquant for our input aggregation is robust. For our analysis, we chose isoquant 750 as our base isoquant.

The inverse of the input oriented efficiency scores relative to isoquant 750, $\left(\theta_c^{LC}(X_l, Y_{750})\right)^{-1}, \forall l$ are now used as indexes of aggregated input. Initiating the estimation of the inflection point and the S-shaped piecewise linear production function (one aggregated input, one output) we first remove the observations with positive slacks present in the estimation of the aggregate input index (181 observations) and observations that are FDH inefficient (732). That leaves us with a data set of 88 observations of which 16 (71) are above (below) mpss. We estimate the BCC efficiency scores based on the sample of observations from 72 to 88 since this concave part of the production function is unaffected of the exact choice of inflection point. Next, we endogenously determine which point below mpss is the inflection point. For fixed $i \in \{1, 71\}$ we go through the following steps:

1. The BCC model is solved including only points above point $72 - i$ and on or below mpss. We count the number of points on the frontier and estimate the termination marginal product of the facet from point $72 - i$ to point $72 - i + 1$.

2. Focusing on $72 - i$ as the candidate for the inflection point we solve (18) above (including the sequence of cuts) to determine a sequence of binary variables indicating a path through a number of points below point $72 - i$ starting at the origin and ending at point $72 - i$ and with a monotone non decreasing marginal product along the path and no points above the path. The optimal solution (assuming that one exists) will provide the count of points on this convex part of the frontier.

3. In addition we require that the marginal product on the facet from point $72 - i - 1$ to point $72 - i$ is greater than or equal to the termination marginal product estimated in step 1.

4. Finally, we add the counts of points on the frontier from step 1 and step 2.

Table 1 summarized the results for 10 candidate points.

Inflection Point	71	70	69	68	67	66	65	64	63	62
The concave part up to mpss	2	3	2	3	4	5	6	7	6	7
The convex part	13*	–	17	21	25	27	17	–	25	–
The frontier up til mpss	–	–	19	24	29	32	23	–	31	–
Notes: – indicate the (18) is integer infeasible										
* indicate that the cutting procedure has stopped after 50 cuts										

Table 1: The number of points one the different parts of the frontier with different choices of the inflection point.

The results of the analysis indicate that either point 63 or point 66 is a good choice for the inflection[11].Next, we solve (18) to estimate the S-shaped production frontier using point 66 as the inflection point. As shown, we are able to obtain a good approximation of the true underlying S-shaped technology.
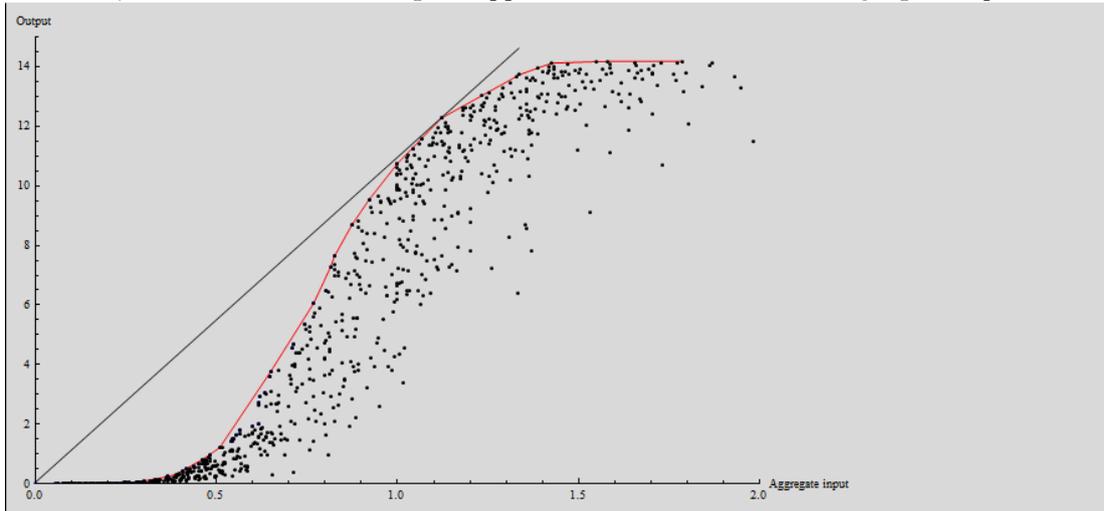
Figure 4. The estimated S-shape and FDH efficient points below and above mpss.

The next step is an analysis of the performance of the aggregation procedure to recover the true inefficiency. For this analysis we use point 66 as the inflection point. The relationship between the distance function and the core of the production function is expressed in proposition (3), i.e. assuming a homothetic structure, $D_I(y_o, X) = k \times g(X), k \in \mathbb{R}_+$, where $k = (f(y_o))^{-1}, f() = F^{-1}()$, and $F()$ is the scaling function. To compare the estimated radial input oriented efficiency scores based on the aggregated input with the true efficiency scores based on the original two dimensional input vector we need to estimate the

---

[11]Analyzing point 71 as candidate for the inflection point was terminated after 50 cuts. At termination, only 15 points were on this infeasible frontier, implying that point 71 is not a promising candidate for the inflection point.

conversion factor $k = (f(output_{750}))^{-1}$, $f(x) = F^{-1}(x)$. $F(x) = \frac{15}{1+e^{-5\log x}} \Rightarrow f(y) = \left(e^{-\frac{1}{5}\log\frac{15-y}{y}}\right)$, and $k = \left(e^{-\frac{1}{5}\log\frac{15-9.84283}{9.84283}}\right)^{-1} = 0.87874$, where 9.84283 is the level of output of the 750'th data point.
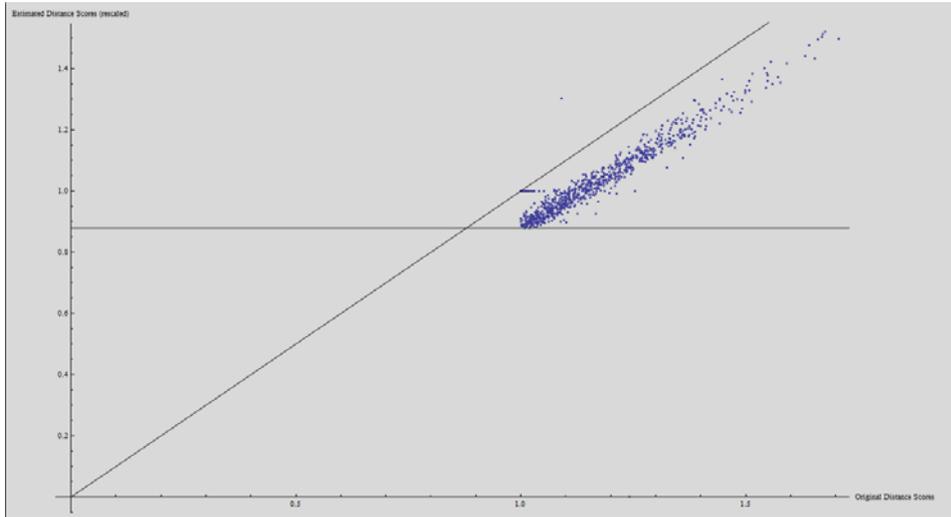
Figure 5a. Plotting the true distance function scores against the estimated scores (rescaled), only "good" observations
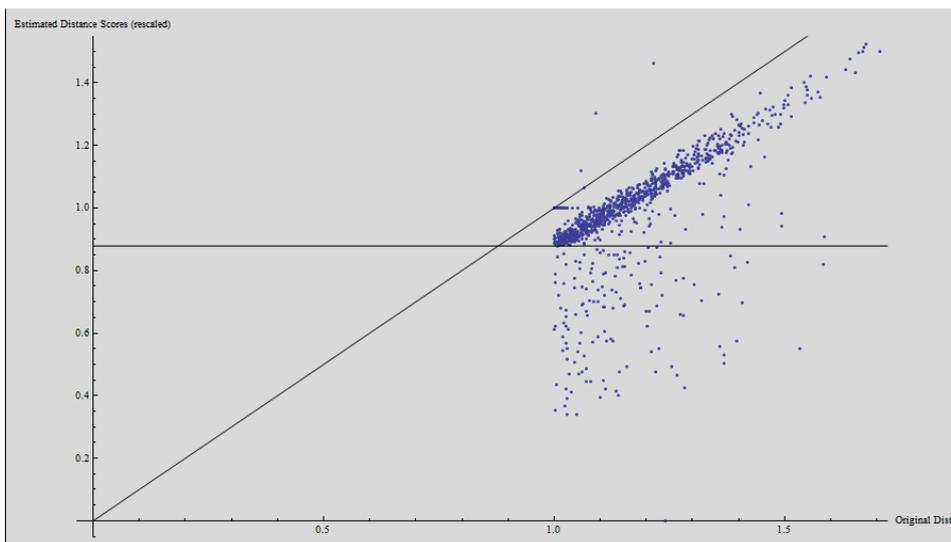
Figure 5b. Plotting the true distance function scores against the estimated scores (rescaled)

Figure 5a-b illustrates the performance of the recovery of the true scores after aggregating the inputs. In Figure 5b all 1000 observations are included, while only the 819 observations with no slacks in the score estimation are included in Figure 5a. It is clear from the plots that the estimated scores are biased downwards as expected for most of the observations. Especially in Figure 5a we observe that all estimated scores are below the true scores except for one "outlier", which turns out to be the smallest observation in the sample.

19

Given our data generating process, we know that the BCC model will not perform well given the assumption of convexity. Nonetheless, we compared our estimates and the BCC estimates of efficiency with the true efficiency for contextual purposes. Four measures were used for the comparisons: mean squared and mean absolute deviations between estimated and true efficiency and the Pearson and Spearman rank correlation coefficients. The performance of the BCC estimator was poor; the mean squared (absolute) deviation was 0.195 (0.345) and the correlation (rank correlation) between estimated and true efficiency was only 0.178 (0.197). The results for our estimator were much better. Using all observations, the mean squared (absolute) error was 0.059 (0.093) while the correlation (rank correlation) was 0.428 (0.757). However, when we include only those observations when the aggregate input was defined without additional slack, the results are much better. In this case, the mean squared (absolute) error was only 0.001 (0.017) while the correlation (rank correlation) was 0.958 (0.964). While this is not surprising given the data generating process, the results suggest that the degree of bias assuming convexity can be very high.

6 Conclusion and further research

A maintained hypothesis of convexity in input-output space is often used in DEA estimations of efficiency scores. However, convexity is not consistent with standard microeconomic production theory that posits an S-shape for the production frontier. In this paper we have outlined an approach that allows for an estimation of efficiency from an S-shaped technology for the multiple inputs and one output case. To simplify, we have assumed that the technology is input homothetic. This assumption has allowed us to split the estimation procedure into two parts, i) an aggregation procedure based on the structure of input homotheticity, and ii) a joint estimation of the inflection point and a piecewise linear S-shaped structure for one aggregated input and one output.

As an estimation procedure for individual isoquants we propose, assuming selective input convexity, the use of a simplified order-m estimation procedure. In theory, any input isoquant can be used as the base isoquant used to aggregate inputs utilizing the input homotheticity. Relative to this base isoquant, an index of aggregated input can be estimated as the inverse distance function value of any observed input vector. We have argued that in practice it is important to choose an isoquant which performs reasonably well according to the following two criteria: i) the empirical distribution of the angle coordinates of the observed data points should mimic a uniform distribution, and ii) a large number of observed data points should be located on the conditional piecewise linear estimator of the isoquant. To facilitate the choice of a base isoquant with good coverage we have proposed an estimation of all possible isoquants using a simplified version of the order-m estimation procedure. The simplification used is a conditional estimator of the base isoquant, which avoids

the replications.

Taking advantage of the reduced dimensionality (one aggregated input and one output) we have developed a model to estimate a piecewise linear S-shaped frontier where the aggregate input axis is divided to allow a production frontier that is concave and convex. In other words, we have assumed that the boundary of the true PPS is S-shaped in the sense that we can divide the input axis into two parts, where the frontier is convex (concave) on the first (second) part. Consequently, the convex hull estimator is too large and we have proposed a "digging approach" where we remove the part of the PPS that violates this S-shape. This digging approach is formulated as a joint estimation of the inflection point and the convex part of the frontier from the origin to the inflection point.

Using simulated data in section 5, we have shown that our method overcomes the inherent problems of standard DEA and provides better estimates of inefficiency when the true technology obeys the Regular Ultra Passum Law. future research.

The approach proposed in this paper has two apparent shortcomings. First and foremost we have assumed input homotheticity which may or may not be a reasonable assumption. Hence an important extension of the approach is to allow at least for some kind of deviation from pure input homotheticity. Secondly, to simplify the presentation we have assumed only one output. Generalizing the approach to the case of multiple input multiple outputs is another important area for future research. Unfortunately, a straight forward approach based on the joint assumption of both input and output homotheticity requires some rather restrictive additional assumptions. As noted in (Färe and Primont 1995) the notion of inverse homotheticity was introduced by ((Shephard 1970), page 255-57), where it is shown that this structure is sufficient for both input and output homotheticity. This result is generalized in (Färe and Primont 1995), where it is shown that we have inverse homotheticity if and only if the technology exhibits simultaneous input and output homotheticity[12]. Hence, extending our approach to multiple inputs and multiple outputs is straight-forward if the technology simultaneously exhibits input and output homotheticity, i.e., inverse homotheticity and if these "mild" additional conditions are maintained. This seems to be a natural starting point for a generalization of the approach in this paper to the case of multiple inputs and multiple outputs.

References

Banker, R. D., Charnes, A. and Cooper, W. W.: 1984, Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science* **30**(9), 1078–1092.

Bogetoft, P.: 1996, DEA on relaxed convexity assumptions., *Management Science* **42**, 457–465.

---

[12] Färe and Primont 1995 present additional "mild" conditions under which this is true.

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

Dantzig, G., Fulkerson, R. and Johnson, S.: 1954, Solution of a large scale travelling-salesman problem, *Journal of the Operations Research Society* **2**(4), 393–410.

Daraio, C. and Simar, L.: 2005, Conditional nonparametric frontier models for convex and nonconvex technologies: A unified approach., *Journal of Productivity Analysis* **28**, 13–32.

Färe, R. and Lovell, C. A. K.: 1988, Aggregation and efficiency, *in Measurement in Economics ed. by W. Eichhorn. Physica-Verlag Heidelberg 1988* pp. 639–647.

Frisch, R.: 1965, *Theory of Production.*, D. Reidel Publ. Comp., Dodrecht-Holland.

Färe, R. and Primont, D.: 1995, On inverse homotheticity, *Bulletin of Economic Research* **47**(2), 161–166.

Førsund, F. R. and Hjalmarsson, L.: 2004, Are all scales optimal in DEA? theory and empirical evidence., *Journal of Productivity Analysis* **21**(1), 25–48.

Ginsberg, W.: 1974, The multiplant firm with increasing returns to scale, *Journal of Economic Theory* **9**, 283–292.

Olesen, O. B. and Petersen, N. C.: 2011, Scale properties in data envelopment analysis, *Workingpaper, forthcomming in Omega* pp. 1–36.

Petersen, N. C.: 1990, Data envelopment analysis on a relaxed set of assumptions, *Management Science* **36**(3), 305–314.

Podinovski, V. V.: 2005, Selective convexity in dea models, *European Journal of Operational Research* **161**, 552–563.

Ruggiero, J.: 1996, On the measurement of technical efficiency in the public sector., *European Journal of Operational Research* pp. 553–565.

Shephard, R. W.: 1953, *Cost and Production Functions*, Princeton University Press, 1953.

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

Zellner, A. and Revankar, N. S.: 1969, Generalized production functions, *Review of Economics and Statistics* **36**(2), 241–250.