University of Connecticut

Department of Economics Working Paper Series

Economic Measures of Capacity Utilization: A Nonparametric Cost Function Analysis

by

Subhash C. Ray
University of Connecticut

John Walden
National Marine Fisheries Service

Lei Chen
Jianghan University

Working Paper 2018-02
March 2018

365 Fairfield Way, Unit 1063
Storrs, CT 06269-1063
Phone: (860) 486-3022
Fax: (860) 486-4463
http://www.econ.uconn.edu/

ECONOMIC MEASURES OF CAPACITY UTILIZATION: A NONPARAMETRIC COST FUNCTION ANALYSIS

Subhash C. Ray
Department of Economics
University of Connecticut
365 Fairfield Way, U-1063
Storrs, CT 06269-1063
Subhash.Ray@Uconn.edu

John Walden
National Marine Fisheries Service
Northeast Fisheries Science Center
166 Water St.
Woods Hole, MA 02543
John.Walden@Noaa.Gov

Lei Chen
Jianghan University, Wuhan, China
Lei.Chen@jhun.edu.cn

Key Words: Data envelopment analysis; Ray Average Cost; Fixed Inputs; Scale economies

Abstract

Cost based measures of capacity utilization and capacity output are important metrics for evaluating firm performance. Understanding where firms are producing on their average cost curve provides information about whether capacity utilization is greater then, less than or equal to one. Most firms are multi-output, multi-input in nature which makes estimation of capacity utilization and capacity output challenging if a cost based measure is desired. For a multi-output firm, the relevant concept is ray average cost (RAC) which can be estimated through non-linear DEA models. This paper offers a simple transformation that linearizes the non-linear DEA program to estimate average, or ray average cost, and to determine capacity utilization and optimal output. The methods are empirically tested on data from a number of U.S. electricity producers for the single output case, and a sample of dental practices for the multi-output case. Results show that for both industries, most firms were operating at less than full capacity, and needed to expand output to minimize their costs. For the dental practices, examination of results from six randomly chosen firms showed the importance of operatories in determining optimal levels of output.

JEL Codes: D24, C61

ECONOMIC MEASURES OF CAPACITY UTILIZATION: A NONPARAMETRIC COST FUNCTION ANALYSIS

Introduction

Capacity is an important economic metric because it yields information about the potential output a firm could produce. Capacity utilization (CU) is also an important measure of performance because it provides information about where a firm is producing relative to its potential (i.e. capacity) output. A competitive firm in long run equilibrium would produce at an output level corresponding to the minimum point of its long run average cost (LRAC) which is tangent to the horizontal demand curve[1] (W. J. Baumol, 1965). This is the only sustainable output level where the firm earns a normal profit (i.e., zero economic profit), unless there is a flat segment at the bottom of the U-shaped LRAC curve. In relation to the LRAC, CU informs whether the plant is producing the correct amount of output, or needs to increase (decrease) output in order to reach the minimum average cost point.

A different (and more popular) definition of capacity output that is independent of any cost criteria is the maximum quantity that the firm can produce given the quantity of its fixed input(s) even with unrestricted availability of variable inputs (Johansen, 1968). A simple example of this measure is the number of passengers that an airline can carry given its fleet of aircrafts. Note that air transportation requires labor, fuel, materials, and planes. However, given a specific number of planes, the airline cannot carry more passengers than the total number of seats in its fleet of aircrafts even though there is no scarcity of labor, fuel, or materials. Capacity output is bounded from above by the physical ability of planes to carry a certain number of passengers, and the measure of CU can never exceed one using the physical definition of capacity. For a group of airlines, capacity and capacity utilization for each airline can be estimated through Data Envelopment Analysis (DEA) as shown by Fare, Grosskopf, and Kokkelenberg (1989). Alternatively, recent work by Cesaroni, Kerstens, and Van de Woestyne (2017) developed an input-oriented physical measure of capacity utilization that can also be estimated with DEA.

Three things should be noted about the output oriented definition of capacity. First, this is a physical upper limit imposed by the available quantity of its fixed input (planes). Second, it is a short run concept because all inputs are not variable. Finally, the capacity output depends on which input is fixed

[1] In a competitive market the demand curve faced by an individual firm is horizontal at the market price even though the market demand curve is downward sloping.

and at what level.[2] We also recognize that the physical measure of capacity is implicitly based on past economic decisions which led to investment in fixed inputs. Note that the physical definition of capacity is widely used by Government statistical agencies. For example, the U.S. Federal Reserve defines capacity as "the greatest level of output a plant can maintain within the framework of a realistic work schedule, after factoring in normal downtime and assuming sufficient availability of inputs to operate the capital in place" (Terry, Walden, & Kirkley, 2008).

Over the past decades, different definitions of a cost based measure of capacity have been proposed in the literature. An early example is one by Cassels (1937), who departed from the LRAC concept of capacity discussed in the first paragraph, and explicitly recognized the role fixed factors play in determining output. He argued that the level of capacity output should correspond to the point where short-run total average costs are minimized. Over two decades later, Klein (1960) questioned the assumption of the U-shaped LRAC curve, but recognized that cost considerations needed to be included in capacity measurement. He also suggested the minimum point on the short-run average total cost curve as the determinant of capacity output, and noted that output prices could factor into the determination of capacity utilization. Ernst R Berndt and Morrison (1981) agreed with Cassels (1937) and Klein (1960) and defined capacity output as the minimum point on the short-run average total cost curve *under constant returns to scale*. Of course, when CRS holds, this would also be the point of tangency between the short run and the long run average cost curves. In the absence of CRS, they prefer the point of tangency between the short and long-run average total cost curves to define capacity output. However, in their empirical application estimating CU for U.S. manufacturing for the period 1958-1977, they assumed CRS reverting thereby to the minimum point of the short run average cost curve. Segerson and Squires (1990) extended the cost based notion of capacity to multi-product industries and developed a CU measure based on ray average costs. Coelli, Grifell-Tatje, and Perelman (2002) proposed a profit based measure of CU for multi-input, multi-output firms where both costs and output prices determined the CU level, and then estimated CU for a group of international airlines using DEA.

With a single output and a parametrically specified cost function, one can analytically derive the capacity output level from the first order condition for a minimum of the average cost. However, in many industries the technology is multi-output, multi-input, which requires a nonparametric approach to estimate capacity output. In the DEA literature, as shown in Fare, Grosskopf, and Lovell (1994) or Subhash C Ray (2004), one can find the value of the minimum average cost under variable returns to

[2] There may be a different capacity output if there are a limited number of pilots. For example, if there are only 7 pilots for 20 planes.

scale from the solution of a cost minimization problem under the assumption of constant returns to scale. While this approach is useful for measuring economic scale efficiency, it does not help to identify the output level where the average cost attains a minimum. Of course, one can usually solve a nonlinear programming problem for minimizing the average cost within the nonparametric framework.

In a recent article published in this journal, Subhash C. Ray (2015) proposed a method that determines long run capacity output from the optimal solution of a simple linear programming problem. In another paper also appearing in this journal and in the same year as Ray (2015), Cesaroni and Giovannola (2015) provided a new measure of scale efficiency which minimizes ray average cost and adjusts for allocative efficiency of the input mix. Their method used a benchmarking approach and presented results primarily for non-convex technologies using the approach of Free Disposal Hull (FDH) analysis.[3]

In this paper we generalize Ray (2015) in several directions to arrive at a measure of short run CU based on short run average cost (SRAC) under alternative returns to scale assumptions and in the presence of multiple outputs and/or multiple fixed inputs. The main contributions of this paper can be summarized as follows:

- First, it provides a *linear programming solution* of the short run average cost minimization problem, which is *an inherently non-linear optimization problem*, for a single output and single fixed input under alternative returns to scale assumptions.
- Then the single output model is extended to allow multiple fixed inputs.
- Finally, we consider the more realistic case of multiple outputs and multiple fixed inputs and identify the scale of a given output mix where the *ray average cost* reaches a minimum.

Although the primary focus of our paper is methodological, for illustrative examples, we include two empirical applications of the proposed methodology. In one, we use the well-known Christiansen-Greene data from US electrical utilities to measure capacity utilization in a single output, two variable inputs, and one fixed input case. In the other application we use the data from a sample of dental care practices from Colorado (used earlier by Chen and Ray (2013)) with two outputs, two fixed inputs, and six variable inputs.

The rest of the paper unfolds as follows. Section 2 offers a quick description of the nonparametric characterization of the production technology. Section 3 contains the main methodological contribution

---

[3] Although they mention DEA for convex technologies (in section 4.2 page 126) they never explicitly spell out the underlying model for computation. Nor do they consider fixed inputs and short run analysis.

formulating the various DEA linear programming problem for solving the short run average cost minimization problem under alternative returns to scale assumptions for single and multiple fixed inputs an also for single and multiple outputs. Section 3.1 explains using a simple 1-output 2-input (1 variable and 1 fixed) example why the assumption about returns to scale is relevant even in the short run even when some input remains fixed and there is no scalar variation of the entire input bundle. Sections 3.2 and 3.3 solve the short run average cost minimization problem for the 1-output 1-fixed input case under CRS and VRS, respectively. These are further extended to include multiple fixed inputs (but still considering the single output case) in sections 3.4 and 3.5. Finally, in sections 3.6 and 3.7, the most general case of multiple outputs and multiple fixed inputs under CRS and VRS assumptions are considered. For multiple outputs, of course, there is concept of the short run average cost is replaced by that of the *ray average cost* (for a given mix of outputs). Section 4 includes two illustrative empirical applications of the proposed DEA models. In section 4.1 the Christensen-Greene data are used to measure short run capacity utilization at a sample of US electric utilities in a 1-output 1-fixed input case. In section 4.2 we analyze the data from a number of dental care practices from Colorado each with two outputs (dentist visits and hygienist visits) and two fixed inputs (number of operating rooms and square feet of office space).Section 5 offers a brief summary.

2. Nonparametric Characterization of the Technology

Our starting point is an industry made up of firms producing a single output y using an input bundle $x \in \Re_{+}^{n}$. An input-output bundle (x, y) is considered feasible when y can be produced from x, and the set of all feasible input-output bundles constitutes the production possibility set:

$$T=\{(x,y): y \text{ can be produced from } x\} \qquad (1)$$

Since every observed input-output bundle is feasible, we begin with the observed input-output bundle[4] (x^j, y_j) of firm j $(j = 1, 2,..,, N)$. Under the fairly non-restrictive assumptions that the production possibility set is convex and that both inputs and outputs are freely disposable, one can approximate the set T by the free disposal convex hull $\hat{T}$ of the observed bundles

$$\hat{T}=\left\{(x,y): x \geq \sum_{1}^{N}\lambda_j x^j; y \leq \sum_{1}^{N}\lambda_j y_j; \sum_{1}^{N}\lambda_j = 1; \lambda_j \geq 0\left(j=1,2,...,N\right)\right\}. \qquad (2)$$

[4] In this paper vectors are identified by superscripts and scalars by subscripts.

If, additionally, one assumes CRS, for any $(x, y) \in T$, for any $k \geq 0$, $(kx, ky) \in T$. Correspondingly the empirical construct of the CRS production possibility set would be:

$$\hat{T}^C = \left\{ (x, y) : x \geq \sum_1^N \lambda_j x^j ; y \leq \sum_1^N \lambda_j y_j ; \lambda_j \geq 0 \left(j = 1, 2, ..., N \right) \right\}. \quad (3)$$

Alternatively, the technology can also be characterized in terms of the input requirement sets. For any specific output level (y_0) the input requirement set consists of all input bundles that can produce y_0. This can be expressed as:

$$V(y_0) = \{ x : (x, y_0) \in T \}. \quad (4)$$

Empirically, the input requirement sets under the assumption of variable returns to scale (VRS) will be:

$$\hat{V}(y_0) = \left\{ x : x \geq \sum_1^N \lambda_j x^j ; y_0 \leq \sum_1^N \lambda_j y_j ; \sum_1^N \lambda_j = 1 ; \lambda_j \geq 0 \left(j = 1, 2, ..., N \right) \right\} \quad (5)$$

and

$$\hat{V}^C(y_0) = \left\{ x : x \geq \sum_1^N \lambda_j x^j ; y_0 \leq \sum_1^N \lambda_j y^j ; \lambda_j \geq 0 (j = 1, 2, ..., N) \right\} \quad (6)$$

if constant returns to scale (CRS) is assumed.

When VRS is assumed, the minimum cost of producing any output y at input price (vector) w^0 is:

$$C^{VRS}(w^0, y) = \min w^{0\prime} x : x \in V(y). \quad (7a)$$

Similarly, under CRS, the minimum cost is:

$$C^{CRS}(w^0, y) = \min w^{0\prime} x : x \in V^C(y). \quad (7b)$$

Note that this is a long run model because there is no fixed input. In a standard textbook example (W. J. Baumol, 1965) the long run average cost (LRAC) curve of a firm assuming VRS is U-shaped because the average cost initially declines due to economies of scale, but subsequently rises as diseconomies of scale follow. However, if CRS holds globally, the long run average cost curve is a horizontal straight line (Subhash C. Ray, 2015), and capacity output is not defined. The LRAC assuming VRS (U-shaped curve) and the LRAC assuming CRS (horizontal line) are tangent to one another at the minimum point of the VRS LRAC. The minimum point on the VRS LRAC, which is tangent to the CRS LRAC, would be considered the point of capacity output.

3. Economic Capacity in the Short Run

There is no consensus about what constitutes the economic capacity output of a firm in the short run using the cost based definition of capacity. Some argue that the point of tangency between the long run average cost curve and the short run average cost curve for a given bundle of fixed inputs should define the capacity output (Ernst R. Berndt & Fuss, 1986). Others argue that since we are talking about the short run and a given bundle of fixed inputs is already in place, the long run average cost curve is no longer relevant and capacity is fully utilized at the output level where the short run average cost curve reaches a minimum. That is the point where short run economies of scale have been fully exploited. While it is true that the long run average cost at that output level would be even lower, such lower cost would not be attainable without altering the quantity of fixed inputs. Therefore, it has no direct relevance for measuring capacity output in the short run. Of course, if constant returns to scale prevailed globally, the long run average cost curve (as has already been noted) would be a horizontal line and the tangency between any short run average cost curve and the long run average cost curve would take place at the minimum point of the relevant average cost curve (Subhash C. Ray, 2015). That is the only case when both definitions of the capacity output would lead to the same output level (Ernst R Berndt & Morrison, 1981). Because when the long run average cost curve is horizontal the long run capacity output level is not defined, the minimum point of the short run average cost curve would define the capacity output in the short run.

In order to proceed further, we adopt two lemmas found in Ray (2015), and then progress from the one output-one fixed input scenario to the multiple output, multiple fixed input scenario:

Lemma 1. Locally Constant Returns to Scale holds at the input-output bundle (x^*, y_*) where the average cost reaches a minimum.

Lemma 2. If the technology exhibits Constant Returns to Scale globally, average cost is a constant at all output levels.

3.1 Does the Returns to Scale assumption matter in the short run analysis?

The Returns to Scale (RTS) properties of the technology relate to how the output (bundle) changes in response to an equi-proportionate change in all inputs. Because all inputs change when the scale changes, RTS essentially is a long run concept. However, by definition, at least one input remains fixed in the short run. It might appear therefore, that for short run cost analysis, the RTS assumption is not relevant. We use a simple 1-output (y) and 2-input (K,L) example to show that the short run total cost

(SRTC) functions and the corresponding capacity output levels will be different under alternative RTS assumptions.

We first consider a CRS production function:

$$y = \sqrt{KL} \qquad (8)$$

and assume that K is fixed in the short run at $K = K_0$. Then, the short run production function becomes

$$y = \left(\sqrt{K_0}\right)\sqrt{L}. \qquad (9)$$

Hence, the minimum quantity of *L* required to produce *y* in the short run is

$$L = \frac{y^2}{K_0}. \qquad (10)$$

At input prices *w* for *L* and *r* for *K*, the short run total cost is

$$SRTC_{CRS} = w\frac{y^2}{K_0} + rK_0 \qquad (11)$$

The corresponding short run average cost is

$$SAC_{CRS} = w\frac{y}{K_0} + r\frac{K_0}{y}. \qquad (12)$$

From the first order condition for a minimum for *SAC*, one gets

$$y^*_{CRS} = \sqrt{\frac{r}{w}}K_0. \qquad (13)$$

Next we discard the CRS assumption and consider the production function

$$y = K^{\frac{1}{3}}L^{\frac{1}{2}}. \qquad (14)$$

This time, with *K* fixed at $K = K_0$. the minimum required quantity of *L* is

$$L = \frac{y^2}{K_0^{\frac{2}{3}}} \qquad (15)$$

Now the short run total cost is

$$SRTC_{VRS} = w\frac{y^2}{K_0^{\frac{2}{3}}} + rK_0 \qquad (16)$$

The corresponding average cost is

$$SAC_{VRS} = w\frac{y}{K_0^{\frac{2}{3}}} + r\frac{K_0}{y}. \qquad (17)$$

The output level where the average cost reaches a minimum is

$$y^*_{VRS} = \sqrt{\frac{r}{w}}K_0^{\frac{5}{6}}. \qquad (18)$$

A comparison of (12) with (17) shows that the short run average cost under CRS is lower than the short run average cost under VRS at every level of the output *y*. Also, the short run capacity output (where the average cost reaches a minimum) is smaller under VRS than under CRS.

For a numerical example, assume $(r = 36, w = 4; K_0 = 64)$. Then,

$$SAC_{CRS} = \frac{y}{16} + \frac{2304}{y}; y^*_{CRS} = 192;$$
$$SAC_{VRS} = \frac{y}{8} + \frac{2304}{y}; y^*_{VRS} = 96.$$

Thus one can see that RTS assumption *does make a difference* for capacity utilization. Even though we are considering the short run and some input remains fixed, one needs to build appropriate DEA model under alternative RTS assumptions.

3.2 One Output, One Fixed input, CRS

First consider a single fixed input *K* along with a vector of variable inputs *v* used to produce a single output *y*, and a CRS technology. Suppose the price vector of the variable inputs is *w*, the quantity of the fixed input is K_0 and its price is r_K. Then the short run total cost is $SRTC = w'v + r_K K_0$ and the minimum of the short run average cost (SAC) is attained at the value of *y* from the solution of the following minimization problem:

$$
\begin{aligned}
&\min_{y,v,\lambda} \frac{w_v{}'v + r_k K_0}{y} \\
&s.t. \sum_{j=1}^{N} \lambda_j y_j \geq y; \\
&\sum_{j=1}^{N} \lambda_j v^j \leq v; \\
&\sum_{j=1}^{N} \lambda_j K_j \leq K_0; \\
&v \geq 0; y \geq 0; \lambda_j \geq 0; (j = 1, 2, \ldots N)
\end{aligned}
\tag{19}
$$

In this problem, apart from the λs, the output level *y* and the vector of variable inputs *v* are decision variables. The objective function is clearly non-linear. But it can be converted into an LP problem through appropriate scaling of variables. Note that the problem in (19) can be written as:

$$
\begin{aligned}
&\min_{y,v,\lambda} w_v{}' \left(\tfrac{v}{y}\right) + r_k \left(\tfrac{K_0}{y}\right) \\
&s.t. \\
&\sum_{j=1}^{N} \left(\tfrac{\lambda_j}{y}\right) y_j \geq 1; \\
&\sum_{j=1}^{N} \left(\tfrac{\lambda_j}{y}\right) v^j \leq \left(\tfrac{v}{y}\right); \\
&\sum_{j=1}^{N} \left(\tfrac{\lambda_j}{y}\right) K_j \leq \left(\tfrac{K_0}{y}\right); \\
&v \geq 0; y \geq 0; \lambda_j \geq 0; (j = 1, 2, \ldots, N)
\end{aligned}
\tag{20}
$$

Next, define $q = \frac{1}{y} v, z = \frac{1}{y} K_0,$ and $\mu_j = \frac{1}{y} \lambda_j$. The problem in (19) or (20) then becomes

$$
\begin{aligned}
&\min_{\mathrm{q,z},\mu,\sigma} w_v{}' q + r_k z \\
&s.t. \sum_{j=1}^{N} \mu_j y_j \geq 1; \\
&\sum_{j=1}^{N} \mu_j v^j \leq q; \\
&\sum_{j=1}^{N} \mu_j K_j \leq z; \\
&q \geq 0; z \geq 0; \mu_j \geq 0; (j = 1, 2, \ldots N)
\end{aligned}
\tag{21}
$$

It may be emphasized that even though K_0 is a parameter, because *y* is a choice variable, $z = \frac{K_0}{y}$ is also a choice variable. From the optimal solution of (10) we get $z^* = \frac{K_0}{y^*}$. Thus, the short run capacity output is

$$y^* = \frac{K_0}{z^*}. \qquad (22)$$

One can compare the actual output of a firm (y_0) with the capacity output y^* to measure the rate of capacity utilization as

$$CU = \frac{y_0}{y^*}. \qquad (23)$$

The intuition behind this model can be explained using a simple diagram. The short run cost (SAC) curves corresponding to different levels of the fixed input along with the long run average cost curve are shown in Figure 1. The tangency between any SAC curve and the LAC curve takes place at the minimum point of the relevant SAC curve. In Figure 2A we focus on the SAC curve for the fixed input level K_0. Suppose that the actual output level of the firm is y_0 and its actual bundle of variable inputs is v^0. At variable input prices w_v and fixed input price r_k, its short run total cost (SRTC) is $C_0 = w^{v\prime} v^0 + r_k K_0$ and its actual average cost is shown by the point A_0. Minimization of the SRTC for the given fixed input K_0, would yield the variable input bundle v_0^* and its SAC would be A_0^* on the SAC curve corresponding to K_0. If, however, both *K* as well as the variable inputs *v* were optimally chosen, it would select bundle (v_0^{**}, K_0^*). Because *all inputs are freely chosen* it would imply long run cost minimization, and the average cost for the output y_0 would correspond to the point B_0 on the horizontal LAC curve. Now suppose that while the fixed input remained K_0, the firm produced the output y_1. Then its short run cost minimizing bundle would be (v_1^*, K_0) and the corresponding point on the SAC curve would be A_1^*. If, however, the firm could select the level of *K* optimally for output y_1, it would select the bundle (v_1^{**}, K_1^*) and its LAC for output y_1 would be shown by the point B_1. Similarly, for output y_3, its short run input bundle would be (v_3^*, K_0) while the long run bundle would be (v_3^{**}, K_3^*). The corresponding points would be A_3^* on the SAC curve and B_3 on the LAC curve. For the output level y_2, the long run optimal level of *K* is also the short run level K_0, Thus, at this output level the short run and long run bundles are identical and the SAC and LAC are both shown by the point A_2^*. Hence, $y_2 = y^*$ is short run capacity output.

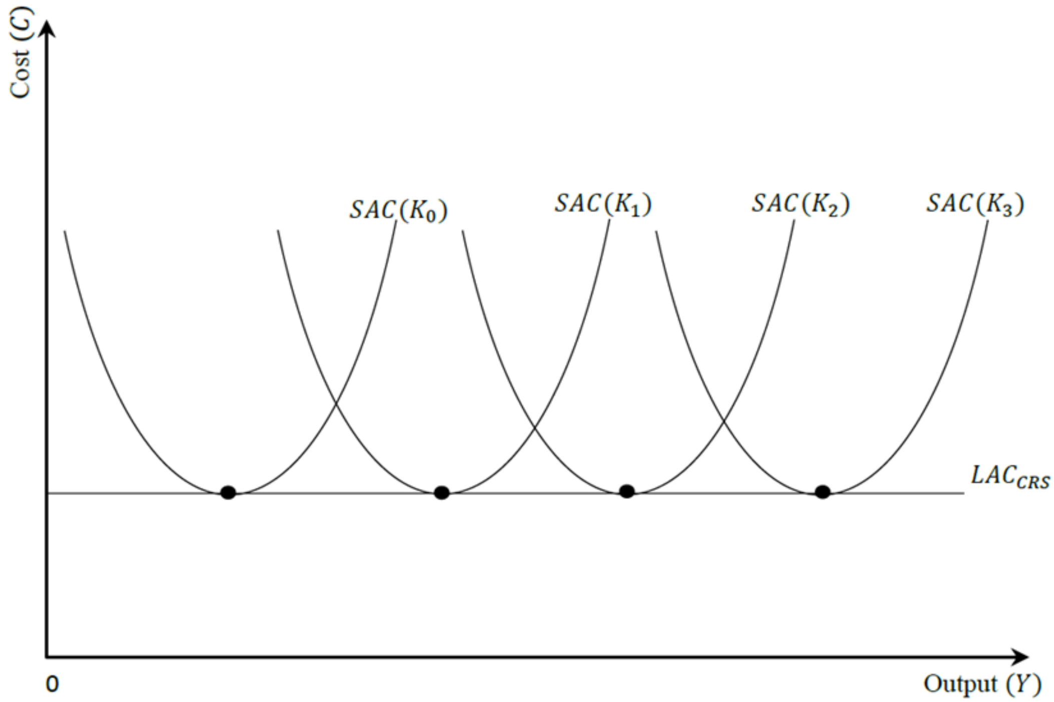

Figure 1. Short and Long-Run Average Cost

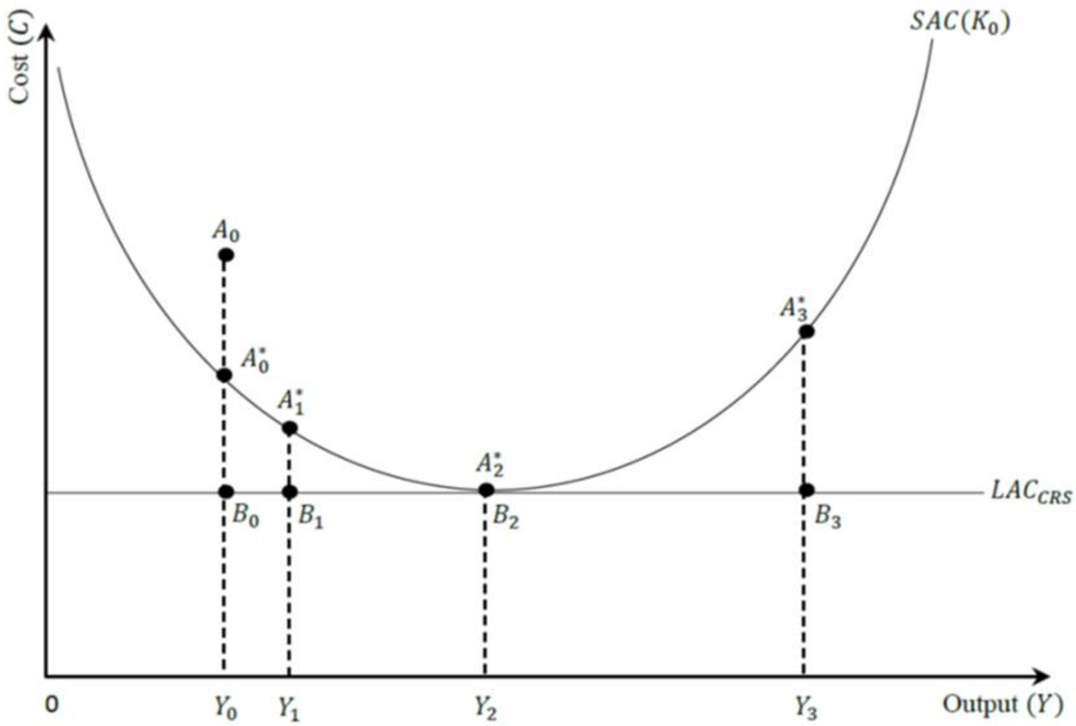

Figure 2A. Short-run average cost

Figure 2B provides an alternative geometric illustration. The radial expansion path OR shows the cost minimizing combinations of (v, K) for the different output levels on the corresponding isoquants. The tangency points B_0 through B_3 are the long run cost minimizing bundles for the outputs y_0 through y_3 under CRS. On the other hand, the points A_0 through A_3 are the short run cost minimizing bundles. Here the fixed input K is held constant at K_0 and only the variable input v changes from v_0^{**} through v_3^* as the output level changes. The point of intersection between the expansion path and horizontal line through K_0 determines the capacity output level y_2 where the (total and average) short run and the long run costs coincide.

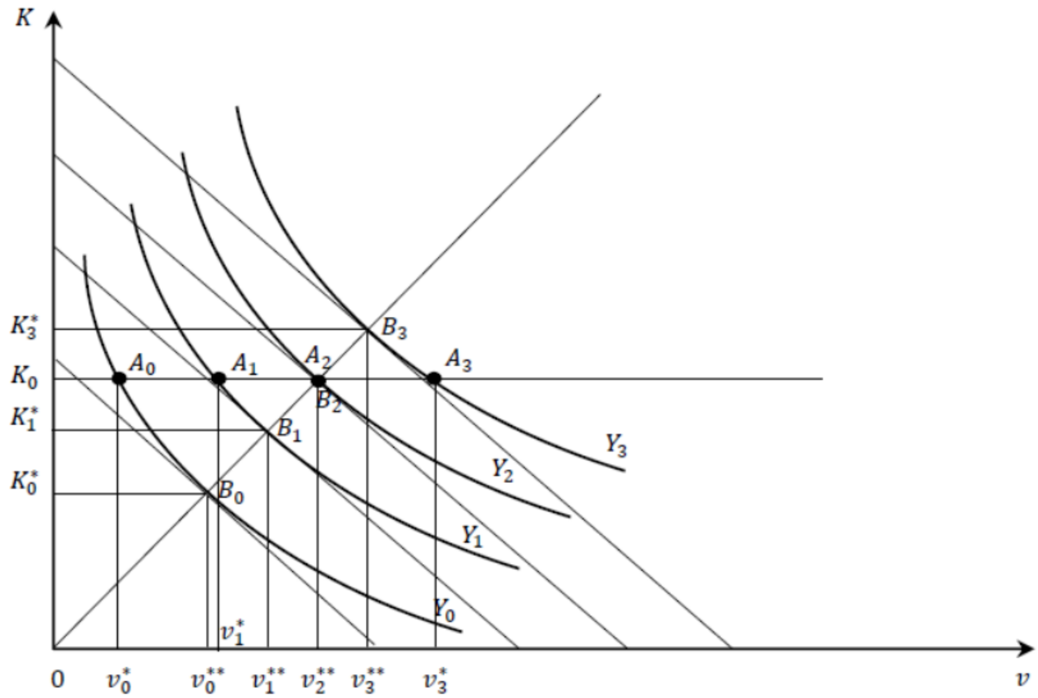

Figure 2B: Expansion Path under CRS.

Compare the input-output bundles $(v_0^{**}, K_0^*; y_0)$ at B_0 and $(v_2^{**}, K_2^*; y_2)$ at B_2.

Due to CRS:

$$\frac{v_0^{**}}{v_2^{**}} = \frac{K_0^*}{K_2^*} = \frac{y_0}{y_2} \qquad (24)$$

Since $K_2^* = K_0$, $y_2 = y^*(K_0)$ is the capacity output for K_0.

Hence the capacity output is

$$y^*(K_0) = y_2 = \left(\frac{K_0}{K_0^*}\right) y_0. \qquad (25)$$

The corresponding capacity utilization rate is

$$CU(K_0) = \frac{y_0}{y^*(K_0)} = \frac{K_0^*}{K_0} \qquad (26)$$

Thus, utilizing the two lemmas stated above, we have found both capacity output and CU for the short-run problem assuming CRS with one output and a single fixed input.

3.3 One Output, One Fixed input, VRS

Now we drop the CRS assumption and assume a U-shaped (rather than a flat) long run average cost curve. This time tangency between the SAC and LAC for any given quantity of the fixed input does not occur at the minimum point of the relevant SAC curve. This is shown in Figure 3 where the curves labeled LAC_{VRS} is the long run average cost curve for a VRS technology and $SAC_{VRS}(K_0)$ is a short run average cost for the fixed input level K_0. The minimum of $SAC_{VRS}(K_0)$ occurs at output level $y_V^*(K_0)$, which is different from the output where these two curves are tangent to one another. By contrast, the horizontal LAC_{CRS} is the long run average cost curve for a CRS technology and the corresponding short run average cost curve $SAC_{CRS}(K_0)$ reaches a minimum at the output level $y_C^*(K_0)$ where it is also tangent to LAC_{CRS}. In the case of VRS, one needs to need to determine the short run capacity output $y_V^*(K_0)$ without referring to the long run average cost curve.

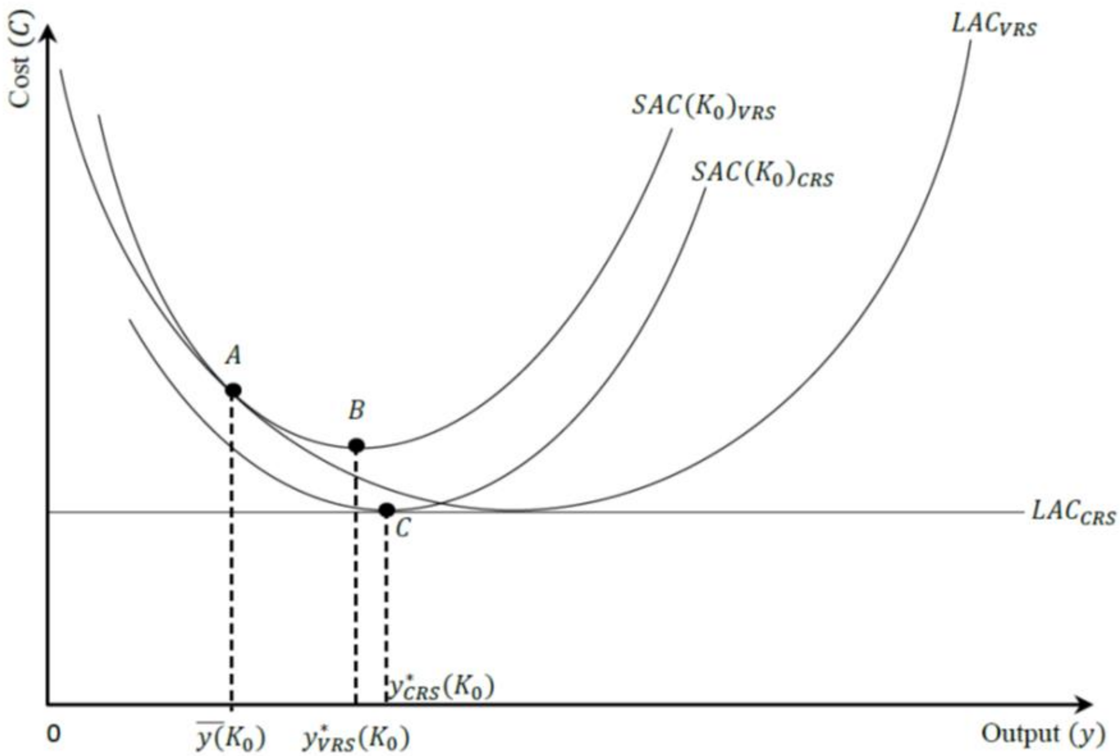

Figure 3. Short-run Capacity Output under VRS and CRS

In the case of VRS, due to inclusion of the constraint $\sum_{j=1}^{N} \lambda_j = 1$ the problem in (19) becomes:

$$
\begin{aligned}
&\min_{y,v,\lambda} w_v{}' \left(\tfrac{v}{y}\right) + r_k \left(\tfrac{K_0}{y}\right) \\
&s.t. \\
&\sum_{j=1}^{N} \left(\tfrac{\lambda_j}{y}\right) y_j \geq 1; \\
&\sum_{j=1}^{N} \left(\tfrac{\lambda_j}{y}\right) v^j \leq \left(\tfrac{v}{y}\right); \\
&\sum_{j=1}^{N} \left(\tfrac{\lambda_j}{y}\right) K_j \leq \left(\tfrac{K_0}{y}\right); \\
&\sum_{j=1}^{N} \left(\tfrac{\lambda_j}{y}\right) = \tfrac{1}{y}; \\
&v \geq 0; y \geq 0; \lambda_j \geq 0; (j = 1, 2, \ldots, N)
\end{aligned}
\tag{27}
$$

As before, define $q = \frac{1}{y} v, z = \frac{1}{y} K_0,$ and $\mu_j = \frac{1}{y} \lambda_j$. Further, define $\sigma = \frac{1}{y}$. This implies the additional restriction $z = \sigma K_0$ and $\sum_{j=1}^{N} \mu_j = \sigma.$ Hence, (16) can be written as

$$
\begin{aligned}
&\min_{q,z,\mu,\sigma} w_v{}' q + r_k z \\
&s.t. \sum_{j=1}^{N} \mu_j y_j \geq 1; \\
&\sum_{j=1}^{N} \mu_j v^j \leq q; \\
&\sum_{j=1}^{N} \mu_j K_j \leq z; \\
&z = \sigma K_0; \\
&\sum_{j=1}^{N} \mu_j = \sigma; \\
&q \geq 0; z \geq 0; \sigma \geq 0; \mu_j \geq 0; (j = 1, 2, \ldots N)
\end{aligned}
\tag{28}
$$

The optimal or capacity output is obtained as:

$$
y^*_{VRS}(K_0) = \tfrac{1}{\sigma^*}. \tag{29}
$$

Note further that by definition $z^* = \frac{K_0}{y^*}$ and an alternative way to measure y^* is to consider $\frac{K_0}{z^*}$.

However, the constraint $z = \sigma K_0$ ensures that the two measures are identical.

3.4 One Output, Multiple Fixed inputs, CRS

We now consider a single output, and multiple fixed inputs. For simplicity assume that there are 2 fixed inputs: $K^0 = (K_1^0, K_2^0)$ and the corresponding rental prices are $r = (r_1, r_2)$. As before the vector of variable inputs is v with the input price vector w_v. Define $\delta_0 = \frac{K_1^0}{K_2^0}$.

The short run average cost with the given bundle of fixed inputs is

$$SAC = \frac{w_v' v + r_1 K_1^0 + r_2 K_2^0}{y}$$

And the corresponding short run CRS cost minimization problem is:

$$\begin{aligned}
&\min_{y,v,\lambda} \frac{w_v' v + r_1 K_1^0 + r_2 K_2^0}{y} \\
&s.t. \sum_j \lambda_j v^j \le v; \\
&\sum_j \lambda_j K_1^j \le K_1^0; \\
&\sum_j \lambda_j K_2^j \le K_2^0; \\
&\sum_j \lambda_j y_j \ge y; \\
&y, v \ge 0; \lambda_j \ge 0; (j = 1, 2, ..., N).
\end{aligned} \tag{30}$$

As before, define $q = \frac{1}{y} v$ and $\mu_j = \frac{1}{y} \lambda_j$. Also $z_1 = \frac{1}{y} K_1^0$ and $z_2 = \frac{1}{y} K_2^0$. Further define $\delta_0 = \frac{K_1^0}{K_2^0}$. Then, clearly. $z_1 = \delta_0 z_2$. The minimization problem in (30) would then become

$$
\begin{aligned}
&\min{}_{z,q,\mu} w_v{}'q + r_1 z_1 + r_2 z_2 \\
&s.t. \sum_j \mu_j v^j \le q; \\
&\sum_j \mu_j K_1^j \le z_1; \\
&\sum_j \mu_j K_2^j \le z_2; \\
&\sum_j \mu_j y_j \ge 1; \\
&z_1 = \delta_0 z_2; \\
&\mathrm{z}_1, z_2 \ge 0; \mu_j \ge 0; (j = 1, 2, \ldots, N).
\end{aligned}
\tag{31}
$$

From the optimal solution we can extract the capacity output level

$$
y^*_{CRS}(K_1^0, K_2^0) = \tfrac{K_1^0}{z_1^*} = \tfrac{K_2^0}{z_2^*}. \tag{32}
$$

3.5 One Output, Multiple Fixed inputs, VRS

When the constraint for VRS is included in (19), $\sum_{j=1}^{N} \lambda_j = 1$ implies $\sum_{j=1}^{N} \mu_j = \frac{1}{y}$ and defining $\sigma \equiv \frac{1}{y}$ the optimization problem in (20) may be revised as:

$$
\begin{aligned}
&\min_{y,v,\lambda} w_v' q + r_1 z_1 + r_2 z_2 \\
&s.t. \sum_j \mu_j v^j \le q; \\
&\sum_j \mu_j K_1^j \le z_1; \\
&\sum_j \mu_j K_2^j \le z_2; \\
&\sum_j \mu_j y_j \ge 1; \\
&\sum_j \mu_j = \sigma; \\
&z_1 = \sigma K_1^0; \\
&z_2 = \sigma K_2^0; \\
&v, \lambda \ge 0; \sigma, z_1, z_2 \ge 0; \lambda_j \ge 0; (j = 1, 2, ..., N).
\end{aligned} \tag{22}
$$

The capacity output can be obtained as

$$y_{VRS}^*(K_1^0, K_2^0) = \tfrac{1}{\sigma^*} = \tfrac{K_1^0}{z_1^*} = \tfrac{K_2^0}{z_2^*}. \tag{34}$$

3.6 Multiple Outputs, Multiple Fixed inputs, CRS

One dilemma with defining the capacity output in terms of the minimum of the average cost curve (whether in the short run or in the long run) for a multiple output firm is that ‘a multiproduct cost function possesses no natural scalar quantity over which costs may be "averaged" (W. Baumol, Panzar, & Willig, 1982). One may, however, treat a specific output bundle y^0 as one unit of a composite good so that a different output bundle *with the same output mix* $y^1 = ty^0$ can be regarded as t units of the composite good. One can, then, define the *Ray average cost*

$$RAC(t; y^0, w^0) = \tfrac{C(ty^0, w^0)}{t}. \tag{35}$$

as the average cost of that composite good (Coelli et al., 2002; Segerson & Squires, 1990). For example, if a basket of fruits contains 2 apples and 3 oranges, instead of adding apples and oranges, we simply count the number of baskets each of which has 2 apples and 3 oranges and divide the total cost by the number of baskets to get the *cost per basket* holding the mix of apples and oranges in each basket constant. We may define the optimal or efficient *scale* of production as t^* which minimizes $RAC(t; y^0, w^0)$. If a firm is actually producing the output bundle y^0 its efficient output bundle would be $y^* = t^* y^0$ and its capacity utilization rate would be $\frac{1}{t^*}$.

Consider now the minimum point of the short run ray average cost curve for the output mix corresponding to the vector y^0 when the firm produces multiple outputs and has a vector of fixed inputs K^0 in the short run. Assume the technology exhibits CRS, and for simplicity assume 2 outputs $y^0 = (y_1^0, y_2^0)$ and 2 fixed inputs $K^0 = (K_1^0, K_2^0)$. The short run ray average cost minimization problem is:

$$
\begin{aligned}
&\min \frac{w_v{}'v + r_1K_1^0 + r_2K_2^0}{t} \\
&s.t. \sum_j \lambda_j y_1^j \geq t y_1^0; \\
&\quad \sum_j \lambda_j y_2^j \geq t y_2^0; \\
&\quad \sum_j \lambda_j v^j \leq v; \\
&\sum_j \lambda_j K_1^j \leq K_1^0; \\
&\sum_j \lambda_j K_2^j \leq K_2^0; \\
&t \geq 0; v \geq 0; \lambda_j \geq 0 \, (j = 1, 2, \ldots, \mathrm{N})
\end{aligned}
\tag{36}
$$

As before, we seek to convert this nonlinear model to a linear programming problem. Define $\mu_j = \frac{\lambda_j}{t} \, (j = 1, 2, \ldots, N), q = \frac{1}{t} v, z_1 = \frac{K_1^0}{t}$, $z_2 = \frac{K_2^0}{t}$ and $\delta_0 = \frac{K_1^0}{K_2^0}$. . Problem (36) can then be written as the following LP problem:

$$
\begin{aligned}
&\min_{q,z,\mu}\, w_v{}'q + r_1 z_1 + r_2 z_2 \\
&s.t. \sum_j \mu_j y_1^j \geq y_1^0; \\
&\sum_j \mu_j y_2^j \geq y_2^0; \\
&\sum_j \mu_j v^j \leq q; \\
&\sum_j \mu_j K_1^j \leq z_1; \\
&\sum_j \mu_j K_2^j \leq z_2; \\
&z_1 = \delta_0 z_2; \\
&q, z_1, z_2 \geq 0; \mu_j \geq 0\,(j = 1, 2, ..., \mathrm{N})
\end{aligned} \tag{37}
$$

As in (20), the efficient production scale is

$$t^*_{CRS}(y^0; K_1^0, K_2^0; w^0) = \frac{K_1^0}{z_1^*} = \frac{K_2^0}{z_2^*}. \tag{38}$$

The corresponding capacity utilization rate is

$$CU_{CRS}(y^0; K_1^0, K_2^0; w^0) = \frac{1}{t^*_{CRS}(y^0; K_1^0, K_2^0; w^0)}. \tag{39}$$

3.7 Multiple Outputs, Multiple Fixed inputs, VRS

Next we relax the CRS assumption and impose the constraint $\sum_j \lambda_j = 1$ in (36) above.

Correspondingly the problem (37) is reformulated as:

$$
\begin{aligned}
&\min w_v{}' q + r_1 z_1 + r_2 z_2 \\
&s.t. \sum_j \mu_j y_1^j \geq y_1^0; \\
&\quad \sum_j \mu_j y_2^j \geq y_2^0; \\
&\quad \sum_j \mu_j v^j \leq q; \\
&\quad \sum_j \mu_j K_1^j \leq z_1; \\
&\quad \sum_j \mu_j K_2^j \leq z_2; \\
&\quad z_1 = \delta_0 z_2; \\
&\sum_j \mu_j = \sigma; \\
&\quad q, \sigma, z_1, z_2 \geq 0; \mu_j \geq 0 \, (j = 1, 2, ..., \mathrm{N})
\end{aligned} \tag{40}
$$

The optimal production scale is

$$t_{VRS}^*(y^0; K_1^0, K_2^0; w^0) = \tfrac{1}{\sigma^*} = \tfrac{K_1^0}{z_1^*} = \tfrac{K_2^0}{z_2^*} \tag{41}$$

and the capacity utilization rate

$$CU_{VRS}(y^0; K_1^0, K_2^0; w^0) = \tfrac{1}{t_{VRS}^*(y^0; K_1^0, K_2^0; w^0)} = \sigma^*. \tag{42}$$

4. Empirical Applications

In this paper we include two empirical applications of the proposed methodology for measuring capacity output and capacity utilization rate in the short run. For both examples, we assume the VRS technology. The first example utilizes data for a number of US electrical utilities used by Christensen and Greene (1976) in their translog cost function paper. Here, we measure short run capacity utilization for the 1-output 1-fixed input case considered in section 3b above. The other example uses a 117 observation data set for general practice dental care groups in the state of Colorado during 2005 constructed from responses to a survey conducted by the American Dental Association (Chen & Ray, 2013). This example considers two outputs and two fixed inputs and solves the DEA LP problem from section 3f.

4.1 Example 1: Short Run Capacity Utilization in Electric Utilities

The Christensen and Greene (1976) data is from 99 utilities producing electricity in 1970. Each utility produces a single output, kilowatt hours of power generated (KWH), using three inputs – labor (L), fuel (F) and capital (K). We treat labor and fuel as variable inputs and capital as a fixed input in the short run. Quantities of labor, fuel, and capital are indices in our data. Output is measured in 1000s of kilowatt hours and ranged between eight and 53,918, with a mean value of 9,000 (Table 1). Short-run average cost ($srac_0$) across all utilities was between $23.88 and $1,340.42, with an average value of $94.46.

Table 1. Summary Statistics of the Christensen-Greene Data

Variable	Obs.	Mean	Std. dev.	Min	Max
Kwh	99	8999.73	10315.14	8	53918
labor	99	69.91	86.68	1.02	440.53
fuel	99	960.77	1087.21	2.97	5541.24
capital	99	132.28	151.71	0.67	851.13
SRAC Observed	99	94.46	160.07	23.88	1340.42

Results from the model (3b) yielded a mean SRAC of $47.38, with output increasing from a mean of 8,999 kwh to 10,647 kwh (Table 2). This suggests that on the whole the utilities were producing below their short run capacity output levels. Increasing output would have lowered average cost from 47.38 to 36.17. The mean capacity utilization rate was 71.34%. There was a wide variation in capacity utilization rates across utilities ranging from a low of 4% to a high of 176%. Removing two units with extremely small levels of output raised the minimum CU rate to 21.8%. Out of the 99 units in the sample, 84 were operating below their respective short run economic capacity levels and could lower their average costs by producing more (Table 3). In all, 36 utilities are producing within ± 25% of their respective levels of short run capacity output. Only three utilities are producing at their economic capacity level. An interesting point to note is that while units producing well above their capacity output levels are, in general, among the bigger ones measured by kilowatt hours of output produced, the one with the highest capacity utilization rate (178%) is one of the smallest units producing only 374 (compared to an overall mean of 9000) kilowatt hours for the entire sample. This apparent paradox arises because its capital input is also very small (3.01 against a sample average of 132.38). Similarly, among the 3 units with capacity utilization rate of 100% are the 3rd smallest and the 3rd largest in terms of output levels with the smallest producing 30 kilowatt hours and the largest 38,343 kilowatt hours. Their corresponding quantities of the capital input are 0.67 and 562.13 units.

Table 2. Results from model 3b using the Christensen-Greene Data

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Actual Output	99	8999.73	10315.14	8	53918
Capacity Output	99	10647.79	9508.52	50	38343
Capacity Utilization	99	0.71	0.30	.05	1.77
SRAC	99	47.38	104.11	23.85	894.09
Minimum SRAC	99	36.18	36.72	22.24	373.66

Table 3. Frequency Distribution of Capacity Utilization

Capacity Utilization Rate	Frequency
Below 50%	20
50% - 75%	41
75% - 95%	18
95% - 100%	5
100%	3
100% - 125%	7
125% - 150%	3
Greater than 150%	2

Most utilities in our sample were operating below 100% CU. For those utilities operating above 100% CU, the short run economic capacity output level depended on the quantity of the fixed input (capital). We found that even a small level of output may be located in the region of short run diseconomies of scale (i.e., the upward rising segment of the short run average cost curve) if the quantity of the fixed input is very small. Operating where there are short run diseconomies of scale means that the utilities should cut back output in order to lower average cost.

4.2 Example 2: Short Run Capacity Utilization in Dental Care

In this empirical application we use the Chen and Ray (2013) dental care provider data. Typically, patients are treated by dental hygienists for periodic cleaning and by dentists for palliative, restorative, prosthetic, or orthodontic care (like tooth extraction, fillings, root canals, implants, etc.). We measure the two kinds of outputs of a practice by the numbers of dentist visits and hygienist visits. The inputs included are (i) dentist hours, (ii) hygienist hours, (iii) chairside assistant hours, (iv) other staff hours, (v) lab expenses, (vi) supply expenses, (vii) square feet of office space, and (viii) number of equipped examination and operating rooms (operatories). The last two inputs, office space and operatories, are treated as fixed in the short run. Table 4 reports the summary statistics of the output and input quantities and the input prices[5]. Out of the 117 practices covered in the study, 2 had no dental hygienists. The number of operatories varied from 1 to 20 with a mean of 4.65. Office floor space varied from 800 to 13,000 square feet, with an average floor space of 2,482 square feet. Apparently, large practices with 20 operatories or 13,000 square feet of office space were multi-location dental care groups. The output mix (measured by the ratio of dentist visits and hygienist visits) averaged 1.453 with a minimum of 0.395 and a maximum of 6.349. Similarly, the mix of the fixed inputs varied from a minimum of 225 square feet per operatory to a maximum of 1250 square feet. The average was 465.197. As will be shown below, in the multiple output, multiple fixed input case (3f), capacity utilization depends on both output mix and the ratio of fixed inputs.

Results from model 3f showed a capacity utilization rate of 73.2%, with a minimum of 25% and a maximum of 150%. Table 5 shows the frequency distribution of capacity utilization rates for the dental care practices. A majority of the observations (89 out of 117) showed capacity utilization below 100%. All of these practices were operating in the downward sloping segment of their ray average cost curves in regions of positive short run scale economies. Sixteen practices were producing at their short run optimal scale (i.e., at the minimum of the short run ray average cost curve). The remaining 12 (with capacity utilization rates in excess of 100%) were experiencing short run diseconomies of scale.

[5] For details of how the input prices have been constructed, see Chen and Ray (2012).

Table 4. Summary Statistics of Dental Care Outputs and Inputs

Variable	Mean	Std. Dev.	Min	Max
Outputs				
Dentist visit	3216.752	2432.793	884	18356
Hygienist visit	2504.581	1772.043	0	10400
Inputs				
Dentist (hrs)	2120.436	1142.134	800	9214
Hygienist (hrs)	2352.709	1472.062	400	8820
Chairside assistant (hrs)	3085.137	1838.019	360	11210
Other staff (hrs)	3187.085	2403.095	400	16400
Lab expenses ($)	57389.85	36890.59	5278	200000
Supply expenses ($)	49524.52	34613.57	3676	236816
Square feet office space	2155.504	1534.343	800	13000
Operatories (number)	4.650	2.482	1	20
Sq ft per Operatory	465.197	154.887	225	1250
Input Prices				
Dentist hourly wage	131.51	79.74	25.39	717.06
Hygienist hourly wage	37.28	8.0	18.79	65.10
Chair-side hourly wage	19.31	6.70	10.00	62.12
Other staff hourly wage	17.01	1.31	13.50	19.50
Per square footage price	24.37	15.57	5.16	116.67
Operatory Price	10000	10000	10000	0
Lab expense price	1	1	1	0
Supply expense price	1	1	1	0
Dentist Visit per Hygienist Visit	1.453	0.896	0.395833	6.349206

Table 5. Frequency Distribution of Capacity Utilization from Dental Care Practices

Capacity Utilization Rate	Frequency
Below 50%	20
50% - 75%	56
75% - 95%	11
95% - 100%	2
100%	16
100% - 125%	5
125% - 150%	7

Table 6 illustrates how the capacity utilization rate and the capacity output bundle depend on the output mix and/or the capital mix of any unit. Begin by comparing units #487 and #431. Both are relatively small units and their fixed inputs are more or less comparable. Unit #431 has 1 more operatory and 76 square feet more of office space than unit #487. Their output mix is the same but unit #487 has 512 square feet of office space per operatory compared to 367 square feet for unit #431. With substantially more of both dentist visits and office visits compared to #487, unit #431 naturally has a higher capacity utilization rate. However, when we look at the output bundles at the optimal scale, they are 3354 dentist visits and 2964 hygienist visits for unit #431 compared to 1768 dentist visits and 1560 hygienist visits for unit #487. A comparison of these two practices suggests that devoting more space to operatories leads to both higher outputs and capacity utilization.

Table 6. Data and Results for Selected Dental Care Practices

survey_id	*CU*	*K1*	*K2*	*y1*	*y2*	*y1**	*y2**	K-mix	y-mix
487	0.50	1024	2	884	780	1768	1560	512	1.13
431	0.67	1100	3	2236	1976	3354	2964	367	1.13
290	0.60	2200	5	2652	2340	4420	3900	440	1.13
295	0.66	4800	8	3796	3328	5788	5075	600	1.14
247	1.30	6000	10	13520	6760	10428	5214	600	2.00
229	1.25	8000	10	5990	9360	4792	7488	800	0.64

Notes: *CU* = Capacity Utilization Rate; *K1* = square feet of office space; *K2* = Number of operatories; *y1* = Dentist Visits; *y2* = Hygienist Visits; K-mix = square feet per operatory; y-mix = Dentist visit per Hygienist visit

We now shift focus to unit #290, which has twice as much office space and 67% more operatories than unit #431. It has the same output-mix as unit #431 and unit #487, and an intermediate level of office space per operatory in relation to the other two (440 compared to 512 and 367). The CU for unit #290 is lower than that of unit #431, while the optimal level of hygienist and dentist visits are higher. Again, this suggests that configuring office space to have more operatories leads to higher optimal output levels and lower RAC.

Next we compare #295 and #247 which have the same mix of the fixed inputs, although #247 is 25% bigger in terms of both of the fixed inputs. However, their output mixes are quite different as can be seen in the y-mix ratio. Unit #247 has twice the number of dentist visits as hygienist visits compared to 14% more for unit #295. The two units deviate from their respective optimal scale of output approximately by the same proportion - #295 is 34% smaller and #247 is 30% bigger than the optimal scale. Interestingly, when we compare their optimal output bundles, the number of hygienist visits does not differ much (5214 for #247 against 5075 for #295). However, the numbers of dentist visits differ widely (10428 for #247 against 5788 for #295). Since dentist visits are more likely to be associated with using an operatory this again suggests that the number of operatories is the critical input determining optimal output levels.

Lastly, compare unit #229 with #247, which have equal number of operatories. Unit #229 has both a higher square footage of office space, and a lower ratio of dentist visits to hygienist visits (y-mix). Compared to all other units, #229 is the only practice with more hygienist visits than dentist visits, as seen in a y-mix of 0.64. Both units are operating above their optimal scale, although the CU for unit #229 is slightly lower than unit #247 (1.25 vs. 1.30). Since the number of operatories were the same for both, this again suggests that operatories were more important than the total square feet of office space in determining optimal output levels.

This example underscores the point that in the multiple output multiple fixed input case, the short run capacity output bundle depends crucially on both the output mix and the proportion of the fixed inputs. Another critical factor is the relative prices of inputs which we have not discussed. Unless the production technology is homothetic, the curvature (and the optimal scale) depends on the relative prices of inputs. In the present example, there was not enough variation of input prices across units to significantly alter the shape of the ray average cost.

5 Summary

Firms producing either a single output from multiple inputs, or multiple outputs from multiple inputs are ubiquitous in a modern economy. Measuring capacity utilization and optimal output for these firms is important because it yields information about the firms' ability to lower costs by expanding production, and whether capital is being fully utilized. Many studies have estimated capacity and capacity utilization based on maximal output levels without consideration of costs in a nonparametric DEA framework because the models are easy to construct and apply, particularly when data are limited. While this may provide useful information for managers and policy makers, a measure which ties output and cost together can provide richer insights.

This research has focused on estimating capacity output and capacity utilization for single and multi-output technologies which use multiple inputs based on minimum average or ray average cost. These estimates are based on non-linear DEA models which were made linear through a simple transformation. Progressing from a single output, one fixed input, one variable input model to multiple output, multiple fixed input models we demonstrated how to transform the non-linear programming cost minimizing models to their linear programming counterparts. We then show how to obtain the optimal cost minimizing output levels, along with observed capacity utilization levels based on the model chosen. In order to demonstrate our approach, two different models were constructed and estimated (models 3b and 3f).

The first application was a single output, three input model (3b) used to estimate capacity utilization and optimal output levels based on data from the Christensen and Greene (1976) study of U.S. electric utilities. The single output was electricity produced and the three inputs were capital, labor and energy, where capital was considered a fixed input. Findings showed that most utilities needed to increase output in order to produce at minimum average cost. Interestingly, results also indicated that even small firms with low output and low levels of capital can be producing in areas where there are diseconomies of scale.

The second application was a two output, eight input model of dental practices located in the U.S. state of Colorado. This application differed from the first in that there was more than one output (dentist hours and hygienist hours) and there were two fixed inputs (square feet of office space and number of operatories). Findings showed that the majority of practices were operating below capacity and could have lowered costs by expanding output. Examination of results from six of the practices showed that the number of operatories, a fixed factor, was the key input in determining where the practice was operating in terms of CU.

Converting the non-linear RAC minimization problem to a linear programming problem to obtain optimal output levels and capacity utilization was a key contribution of this research. While a non-linear programming problem can be readily solved with good solvers available through commercial packages such as GAMS, and good starting points, converting the problem to a L.P. model means that there are more software packages available to solve the problem. This can be particularly advantageous in less developed countries where budgets may limit the ability to purchase advanced software packages.

References

Baumol, W., Panzar, J., & Willig, R. (1982). *Contestable markets and the theory of market structure*. New York: Harcourt Brace Javanovich, Inc.

Baumol, W. J. (1965). *Economic theory and operations analysis*: Prentice-Hall, NJ. 606p.

Berndt, E. R., & Fuss, M. A. (1986). Productivity measurement with adjustments for variations in capacity utilization and other forms of temporary equilibrium. *Journal of Econometrics*, *33*(1–2), 7-29. doi: http://dx.doi.org/10.1016/0304-4076(86)90025-4

Berndt, E. R., & Morrison, C. J. (1981). Capacity utilization measures: underlying economic theory and an alternative approach. *The American Economic Review*, *71*(2), 48-52.

Cassels, J. M. (1937). Excess capacity and monopolistic competition. *The Quarterly Journal of Economics*, 426-443.

Cesaroni, G., & Giovannola, D. (2015). Average-cost efficiency and optimal scale sizes in non-parametric analysis. *European Journal of Operational Research*, *242*(1), 121-133.

Cesaroni, G., Kerstens, K., & Van de Woestyne, I. (2017). A New Input-Oriented Plant Capacity Notion: Definition and Empirical Comparison. *Pacific Economic Review*.

Chen, L., & Ray, S. (2013). Cost efficiency and scale economies in general dental practices in the US: a non-parametric and parametric analysis of Colorado data. *Journal of the Operational Research Society, 64*(5), 762-774.

Christensen, L. R., & Greene, W. H. (1976). Economies of scale in US electric power generation. *Journal of political Economy*, *84*(4, Part 1), 655-676.

Coelli, T., Grifell-Tatje, E., & Perelman, S. (2002). Capacity utilisation and profitability: A decomposition of short-run profit efficiency. *International Journal of Production Economics, 79*(3), 261-278.

Fare, R., Grosskopf, S., & Kokkelenberg, E. C. (1989). Measuring plant capacity, utilization and technical change: a nonparametric approach. *International Economic Review*, 655-666.

Fare, R., Grosskopf, S., & Lovell, C. K. (1994). *Production frontiers*: Cambridge University Press.

Johansen, L. (1968). Production functions and the concept of capacity. *Recherches récentes sur la fonction de production, Collection, Economie mathématique et économétrie*, *2*, 52.

Klein, L. R. (1960). Some theoretical issues in the measurement of capacity. *Econometrica: Journal of the Econometric Society*, 272-286.

Ray, S. C. (2004). *Data envelopment analysis: theory and techniques for economics and operations research*: Cambridge University Press.

Ray, S. C. (2015). Nonparametric measures of scale economies and capacity utilization: An application to U.S. manufacturing. *European Journal of Operational Research, 245*(2), 602-611. doi: http://dx.doi.org/10.1016/j.ejor.2015.03.024

Segerson, K., & Squires, D. (1990). On the measurement of economic capacity utilization for multi-product industries. *Journal of Econometrics, 44*(3), 347-361.

Terry, J. M., Walden, J., & Kirkley, J. E. (2008). *National assessment of excess harvesting capacity in federally managed commercial fisheries* (Vol. NMFS-F/SPO-93). Silver Spring, MD: US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service.