PERGAMON

Socio-Economic Planning Sciences 34 (2000) 1–25

SOCIO-ECONOMIC PLANNING SCIENCES

www.elsevier.com/locate/orms

A unified additive model approach for evaluating inefficiency and congestion with associated measures in DEA

William W. Cooper[a,*], Lawrence M. Seiford[b], Joe Zhu[c]

[a]*Department of Management Science and Information Systems, University of Texas at Austin, Austin, TX 78712-1175, USA*

[b]*Department of Mechanical and Industrial Engineering, University of Massachusetts at Amherst, Amherst, MA 01003, USA*

[c]*Department of Management, Worcester Polytechnic Institute, Worcester, MA 01609, USA*

Abstract

This paper develops a necessary and sufficient condition for the presence of (input) congestion. Relationships between the two congestion methods presently available are discussed. The equivalence between Färe et al. [12,13] and Brockett et al. [2] hold only when the law of variable proportions is applicable. It is shown that the work of Brockett et al. [2] improves upon the work of Färe et al. [12,13] in that it not only (1) detects congestion but also (2) determines the amount of congestion and, simultaneously, (3) identifies factors responsible for congestion and distinguishes congestion amounts from other components of inefficiency. These amounts are all obtainable from non-zero slacks in a slightly altered version of the additive model — which we here further extend and modify to obtain additional details. We also generate a new measure of congestion to provide the basis for a new unified approach to this and other topics in data envelopment analysis (DEA).

Keywords: Congestion; Efficient; Slack; Data envelopment analysis

1. Introduction

Congestion, as used in economics, refers to situations where reductions in one or more

* Corresponding author. Tel.: +1-512-471-3676; fax: +1-512-471-3937.

PII: S0038-0121(99)00010-5

inputs generate an increase in one or more outputs. Färe and Svensson [15] define and develop the topic of congestion in a form that they relate to the law of variable proportions. Färe and Grosskopf [11] subsequently extend and develop this in the context of data envelopment analysis (DEA) which gives the Färe–Svensson operationally implementable form. This is done under assumptions of strong and weak input disposabilities. Finally, Färe et al. [12,13] suggest a procedure for identifying input factors responsible for the congestion.

In Färe and Grosskopf's [11] congestion model, the law of variable proportions requires some inputs be held constant. However, Brockett et al. [2] show that this condition may not be necessary. See the Appendix in [2] which allows congestion to be eliminated by reducing all inputs *without* changing their proportions. For example, in an underground coal mine with too many men and too much equipment, these two inputs may be simultaneously reduced without altering their proportions in order to increase the output of coal.

Nor does this end the problems to be dealt with under the approach by Färe, Grosskopf and their associates. Among other things, their use of the assumption of "free disposal" in either its strong and weak forms, leads to anomalous examples such as situations in which two decision making units (DMUs) may be accorded the same efficiency rating even though one dominates the other and does so strictly in all inputs and outputs that are associated with non-zero slacks. This paper therefore presents results from research directed to a broadened range of possibilities for which we adopt the following definition of congestion from Cooper et al. [9]. See also Brockett et al. [2] and Cooper et al. [7]. The definition we use is:

Definition 1 (congestion)**.** Evidence of congestion is present when reductions in one or more inputs can be associated with increases in one or more outputs — or, proceeding in reverse, when increases in one or more inputs can be associated with decreases in one or more outputs — without worsening any other input or output.

To implement this definition, Brockett et al. [2] developed a new DEA-based approach to capture input congestion and identify its sources and amounts. They also provide an Appendix that examines how their approach relates to that in Färe et al. [12,13]. A supplementary note by Färe and Grosskopf [14] then presents an example to further clarify matters. The current paper studies the relationship between these two DEA congestion approaches. After showing respects in which the two are equivalent, a necessary and sufficient condition is developed for the presence of congestion. A further refinement and extension of the approach in Brockett et al. [2] is then effected to improve upon both of their earlier approaches. Further, it provides a new overall scalar measure of congestion that incorporates all inefficiencies in accordance with the criteria for a satisfactory measure as specified in Cooper et al. [8].

The remainder of this paper is organized as follows. Section 2 provides the congestion approach by Färe and Grosskopf [11], Färe et al. (FGL) [12,13] and related DEA models. It is shown that DEA congestion is related to the non-zero slacks in DEA models. See Definitions 2 and 3. Section 3 studies the relationship between the DEA congestion approaches by FGL [11,12] and Brockett et al. (BCSW) [2]. Section 4 illustrates how to measure congestion by the additive DEA model [5,6]. Section 5 concludes.

2. Congestion and slacks

In this section, we discuss the congestion measure developed by FGL [12,13]. In order to facilitate our development, we first provide the following congestion-related DEA models, which are synthesized as follows.[1]

Suppose we have n DMUs. Each DMU$_j$, $j = 1,2,\ldots, n$, produces s different outputs, y_{rj} ($r = 1,2,\ldots, s$), using m different inputs, x_{ij} ($i = 1,2,\ldots, m$). As given in Charnes et al. [3] the efficiency of a specific DMU$_o$ can be evaluated by either of the following two DEA models:

Input-orientation model

$$
\begin{aligned}
&\theta^* = \min \theta \\
&\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j x_{ij} \leqslant \theta x_{io} \quad i = 1, 2, \ldots, m \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geqslant y_{ro} \quad r = 1, 2, \ldots, s \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \geqslant 0 \quad j = 1, \ldots, n
\end{aligned}
\tag{1}
$$

and

Output-orientation model

$$
\begin{aligned}
&\phi^* = \max \phi \\
&\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j x_{ij} \leqslant x_{io} \quad i = 1, 2, \ldots, m \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geqslant \phi y_{ro} \quad r = 1, 2, \ldots, s
\end{aligned}
\tag{1'}
$$

[1] For more detailed discussion on the properties on these models, refer to Charnes et al. [6].

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geqslant 0 \quad j = 1, \ldots, n.$$

where x_{io} and y_{ro} are, respectively, the ith input and rth output for the DMU_o under evaluation. Associated with the $m+s$ input and output constraints in Eqs. (1) or (1′), some non-zero input and output slacks, s_i^- and s_r^+, may exist in some alternate optimal solutions. In order to handle this possibility, we employ the following model where the efficiency score is fixed via Eqs. (1) (or (1′)) after which the sum of slacks is maximized as follows:

$$\max \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = \theta^* x_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r = 1, 2, \ldots, s \tag{2}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, s_i^-, s_r^+ \geqslant 0$$

or

$$\max \sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi^* y_{ro} \quad r = 1, 2, \ldots, s \tag{2'}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j,\ s_i^-,\ s_r^+ \geqslant 0$$

where θ^* (ϕ^*), the optimal value of Eqs. (1) ((1′)) is fixed as shown for the input (output) constraints. We then have:

Definition 2 (DEA slacks)**.** An optimal value of s_i^- and s_r^+ in Eqs. (2) (or (2′)), which we represent by s_i^{-*} and s_r^{+*}, are respectively called DEA input and output slack values.

Definition 3 (DEA efficiency)**.** A DMU_o evaluated in the above manner will be found to be DEA efficient if and only if the following two conditions are satisfied: (i) $\theta^*=1$ (or $\phi^*=1$); (ii) $s_i^{-*}=s_r^{+*}=0$ for all i and r.

Model (1) (or (1′)) above satisfies strong disposability and hence do not address the issue of non-zero slacks in some, but not all, optima. If we assume weak input disposability as in FGL [12,13], we have the following DEA models to use in a two-stage evaluation of congestion, viz.

Input-orientation

$$\beta^* = \min \beta$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j x_{ij} = \beta x_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} \geqslant y_{ro} \quad r = 1, 2, \ldots, s \tag{3}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geqslant 0 \quad j = 1, \ldots, n$$

and

Output-orientation

$$\hat{\beta}^* = \max \hat{\beta}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \lambda_j x_{ij} = \tau x_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n}\lambda_j y_{rj} \geqslant \hat{\beta} y_{ro} \quad r=1,\ 2,\ \dots,\ s$$

$$\sum_{j=1}^{n}\lambda_j = 1 \tag{3'}$$

$$0 \leqslant \tau \leqslant 1 \text{ and } \lambda_j \geqslant 0 \quad j=1,\ \dots,\ n$$

Note that, for instance, the difference between Eqs. (1) and (3) is that input inequalities are changed into input equalities so that non-zero slack cannot be associated with any input. Further, every x_{io} is multiplied by the scalar β so the associated value of β^* results from a radial measure that corresponds to the way θ^* is obtained. The input congestion measure is then defined as the following ratio by FGL [12,13]:

$$\text{(Input-orientation)} \quad C(\theta^*,\ \beta^*) = \frac{\theta^*}{\beta^*} \tag{4}$$

$$\text{(Output-orientation)} \quad C(\phi^*,\ \hat{\beta}^*) = \frac{\phi^*}{\hat{\beta}^*}. \tag{4'}$$

Now note that we must have $\theta^* \leqslant \beta^*$ because the latter is associated with equalities in Eq. (3) which are replaced by inequalities in Eq. (1). As shown by FGL, we can use $C(\theta^*,\ \beta^*)$ (or $C(\phi^*,\ \beta^*)$) as a measure of congestion with the following properties. If $C(\theta^*,\ \beta^*)=1$ (or $C(\phi^*,\ \beta^*)=1$), then input is not congested; alternatively, if $C(\theta^*,\ \beta^*) < 1$ (or $C(\phi^*,\ \beta^* > 1)$, then input congestion is present.

Note that the above congestion measure is strongly dependent on the orientation of DEA models employed. We can interpret EQ. (4) by noting that $0 \leqslant (1-C(\theta^*,\ \beta^*)) \leqslant 1$ represents a shortfall in output arising from going from Eqs. (1) to (3) because the latter is required to use all inputs in the production indicated by β. This is not the case for Eq. (1), however, because non-zero slack is allowed when this model is used. See the addendum to BCSW [2] for numerical examples and discussions in the output-orientation case associated with $C(\phi^*,\ \beta^*)$.

Turning to Eqs. (1) and (2), the proportional reduction in all inputs associated with θ^* in Eq. (1) yields $0 \leqslant (1-\theta^*) \leqslant 1$ as a measure of technical inefficiency. As noted in Cooper et al. [8], this inefficiency is to be distinguished from the non-zero slack secured from Eq. (2) because these values will generally change the proportions in which inputs are used or in which outputs are produced. This changes the mix and, as we will see, is later identified as a further source of inefficiency.

Measures that can be used to summarize these mix components of inefficiency are developed and discussed in detail in Cooper et al. [8]. Later in the paper we will provide a measure of congestion inefficiency but, for the moment, we shall simply use condition (ii) in Definition 3 to identify additional sources of inefficiency *after* θ^* has been determined via Eq. (1).

Returning to Eqs. (1) and (3), we may note that matters there proceed in an opposite fashion in that only a measure of congestion has been provided. FGL provide the following model, as taken from [13], to identify sources and amounts of congestion in each input:

$$\begin{aligned} \alpha^* = {} & \min \alpha \\ \text{s.t.} \quad & \sum_{j=1}^{n} \lambda_j x_{ij} = \alpha x_{io} \quad i \in A \\ & \sum_{j=1}^{n} \lambda_j x_{ij} \leqslant \alpha x_{io} \quad i \in \bar{A} \\ & \sum_{j=1}^{n} \lambda_j y_{rj} \geqslant y_{ro} \quad r = 1, 2, \ldots, s \\ & \sum_{j=1}^{n} \lambda_j = 1 \\ & \lambda_j \geqslant 0 \quad j = 1, \ldots, n \end{aligned} \tag{5}$$

where $A \subseteq \{1,2, \ldots, m\}$ and $\bar{A}$ is the complement. Using β^* and α^* for each $A \subseteq \{1,2, \ldots, m\}$, if $C(\theta^*, \beta^*) < 1$, and $\theta^* = \alpha^*$, as obtained from Eqs. (1) and (5), the components of the subvectors associated with $\bar{A}(=\{i | i \notin A\})$ then identify sources and amounts of congestion.

The suggested route requires additional computation which can be onerous because it involves obtaining solutions over all possible partitions of A (see Färe et al. [13], p. 77). Moreover, when done, still more may be required since the use of Eqs. (1) and (5) may not identify the non-zero slack that may be present in some (but not all) alternate optima.

The route followed by FGL emphasizes efficiency measurements with identification of sources and amounts of inefficiencies to be undertaken as an additional job. We proceed in the opposite direction, however, by emphasizing the identification of inefficiencies and their components with measures to be derived thereafter. This will be done after we first prove the following theorem.

Theorem 1 (FGL). *Input congestion as defined by Eq. (4) is not present in the performance of DMU_o* if and only if *an optimal solution is associated with referent frontier DMUs such that non-zero input slack values are* not *detected in DEA model (1) (or (1′)*) or (2) and (2′).

Proof. While we prove this theorem in the case of input-orientation DEA models, the proof for output-orientation DEA models is the same. Recall that the only difference between Eqs. (1) and (3) is that the input inequalities are changed to equalities. The referent frontier DMUs are those in the basis when calculating the strong disposability model (1). If we have some referent DMUs such that no non-zero input slack values are detected for DMU_o, then we have, at optimality,

$$\sum_{j\in B}\lambda_j^* x_{ij} = \theta^* x_{io} \quad \text{for} \quad i = 1, \ldots, m$$

where B represents the set of referent DMUs, $B = \{j | \lambda *_j > 0\}$. Obviously, λ^*_j and θ^* are also optimal for Eq. (3); therefore, $\theta^* = \beta^*$. Thus, no input congestion occurs. This completes the *if* part.

To establish the *only if* part, we note that if no input congestion is identified when an optimum is associated with a basis B' such that

$$\sum_{j\in B'}\lambda_j^* x_{ij} = \beta^* x_{io} = \theta^* x_{io}$$

then this same optimum provides referent DMUs such that the input constraints are binding in Eq. (1). Therefore, no non-zero input slack values are detected by reference to those DMUs in B'. □

This theorem will help us maintain contact with what has been accomplished by Färe et al. [12,13] and their associates as we shall shortly see when we move toward one version of additive models in which emphasis is centered on the slacks.

3. Relationship

The previous section indicates that there is a strong relationship between input slacks and the FGL measure of input congestion. In fact, we have

Corollary 1. $\tau^* < 1$ in an optimal solution for Eq. (3′) is a necessary and sufficient condition for positive slack to form part of an optimal solution to Eq. (1′) when $x_{io} > 0$, $i = 1, \ldots, m$.

Proof. The condition is sufficient because $\tau^* < 1$ implies

$$x_{io} = \sum_{j=1}^{n}\lambda_j^* x_{ij} + s_i^{-*}$$

with $s_i^{-*} > 0$ for every $i = 1, \ldots, m$. Hence, also, $s_i^{-*} = (1 - \tau^*)\ x_{io}$. To show necessity, we assume $\tau * \geqslant 1$ for Eq. (3′) for every $i = 1, \ldots, m$. We then have

$$\tau^* x_{io} = \sum_{j=1}^{n}\lambda_j^* x_{ij} \geqslant x_{io}.$$

Thus, no positive slack can be added to obtain

$$x_{io} = \sum_{j=1}^{n}\lambda_j^* x_{ij} + s_i^{-*}$$

with $s_i^{-*} > 0$ for every $i = 1, \ldots, m$. □

From the above result, we also have $s_i^{-*}/x_{io} = 1 - \tau^* \leqslant 1$ which we shall shortly relate to a new congestion measure. Here, however, we only note that this measure is dimensionless as is τ^*. As shown in Corollary 1, above, a value of $\tau^* < 1$ is necessary and sufficient for identifying the presence of non-zero slacks in the corresponding solution to Eq. (1′). An input reduction is needed along with an output increase in order for congestion to be identified. The condition $\tau^* < 1$ is therefore necessary but not sufficient for identification of congestion. This can be seen by virtue of the following theorem and an example provided below.

Theorem 2. The FGL ratio measure will not identify congestion, when present, only if the law of variable proportions does not apply.

Proof. Assume that an optimal solution to Eq. (1′) identifies a point that is proportional to the vector of inputs for DMU_o. Let this solution have value ϕ^*. Also, let $j \in S_o$ represent the optimal basis associated with ϕ^*. By virtue of weak disposability as assumed for Eq. (3′) by FGL, we can represent this solution by

$$0 \leqslant \tau^* - \frac{\sum_{j \in S_o} x_{ij}\lambda_j^*}{x_{io}} \leqslant 1$$

since

$$\sum_{j \in S_o} x_{ij}\lambda_j^* \leqslant x_{io}$$

for any i in Eq. (1′). This also satisfies the conditions in Eq. (3′) with $\theta^* = \beta^*$. This follows because the conditions for Eq. (3′) are more restrictive than the conditions for Eq. (1′). Hence, $\theta^* \leqslant \beta^*$ with equality achieved under the assumption of proportionality.

Further, when the law of variable proportions does not apply, we have

$$C(\theta^*, \hat{\beta}^*) = \frac{\theta^*}{\hat{\beta}^*} = 1.$$

The FGL ratio measure will thus fail to identify the congestion that is present even if $\tau^* < 1$. □

We therefore turn to the additive model — as adapted from Cooper et al. [9] by Brockett et al. [2] — in order to develop a DEA formulation to detect the input congestion which we subsequently refine and extend.

The following model is employed by Brockett et al. [2] after solving Eqs. (1) and (2) (or Eqs. (1′) and (2′)), Input-orientation

$$\max \sum_{i=1}^{m} \delta_i^+$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} - \delta_i^+ = \theta^* x_{io} - s_i^{-*} = \hat{x}_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} = y_{ro} + s_r^{+*} = \hat{y}_{ro} \quad r = 1, 2, \ldots, s \tag{6}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geqslant 0, \quad s_i^{-*} \geqslant \delta_i.$$

or

Output-orientation

$$\max \sum_{i=1}^{m} \delta_i^+$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} - \delta_i^+ = x_{io} - s_i^{-*} = \hat{x}_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} = \phi^* y_{ro} + s_r^{+*} = \hat{y}_{ro} \quad r = 1, 2, \ldots, s \tag{6'}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geqslant 0, \quad s_i^{-*} \geqslant \delta_i.$$

Here, θ^* (ϕ^*) is obtained from Eqs. (1) ((1′)) while s_i^{-*} and s_r^{+*} are obtained from Eqs. (2) ((2′)). 'Input-orientation' means that we use an input-oriented DEA model (1) in the first-stage. 'Output-orientation' means that we use an output-oriented DEA model (1′) in the first-stage.

The amount of congestion in each input can then be determined by the difference between each pair of s_i^{-*} and δ_i^{+*}, where δ_i^{+*} are optimal values in Eqs. (6) ((6′)). That is,

$$s_i^c = s_i^{-*} - \delta_i^{+*} \quad i = 1, 2, \ldots, m. \tag{7}$$

Definition 4 (Congestion Slacks). s_i^c defined in Eq. (7) are called input congestion slacks. It should be noted that Eqs. (6) ((6′)) may be modified by altering its objective to

$$\max \sum_{i=1}^{m} \frac{\delta_i^+}{x_{io}} = \sum_{i=1}^{m} \frac{\delta_i^{+*}}{x_{io}}$$

which is independent of the units of measure used for the various inputs. We also have

$$0 \leqslant \delta_i^{+*} \leqslant s_i^{-*} \leqslant x_{io}$$

so

$$0 \leqslant \frac{\sum_{i=1}^{m} \frac{\delta_i^{+*}}{x_{io}}}{m} \leqslant \frac{\sum_{i=1}^{m} \frac{s_i^{-*}}{x_{io}}}{m} \leqslant 1$$

and therefore

$$0 \leqslant \frac{\sum_{i=1}^{m} \frac{s_i^{c*}}{x_{io}}}{m} = \frac{\sum_{i=1}^{m} \frac{s_i^{-*}}{x_{io}}}{m} - \frac{\sum_{i=1}^{m} \frac{\delta_i^{+*}}{x_{io}}}{m} \leqslant 1 \tag{8}$$

This provides a measure that can be interpreted as the average proportion of congestion present in the observed amounts of inputs used by DMU_o.

From the above discussion, and Corollary 1, we immediately have

Theorem 3. The FGL approach will be equivalent to the BCSW approach in its ability to identify congestion only when the law of variable proportions is present so that no optimum to Eq. (1′) will evaluate DMU_o in a point with coordinates proportional to its inputs.

We now use Fig. 1, which is adapted from BCSW [2], to provide a two-input one-output illustration. The solid lines connecting the points A, B, C, and D represent the unit isoquant for $y = 1$ so the (x_1,x_2) coordinates are capable of producing $y = 1$. Output then rises linearly to 10 units at R along the production surface and then falls back (also linearly) to the level of $y = 1$ at G with coordinates $(x_1,x_2) = (7.5,7.5)$. This is halfway between C and D on the solid line segment connecting them.

To evaluate G's performance, we go back to Eq. (1′) and obtain the following model:

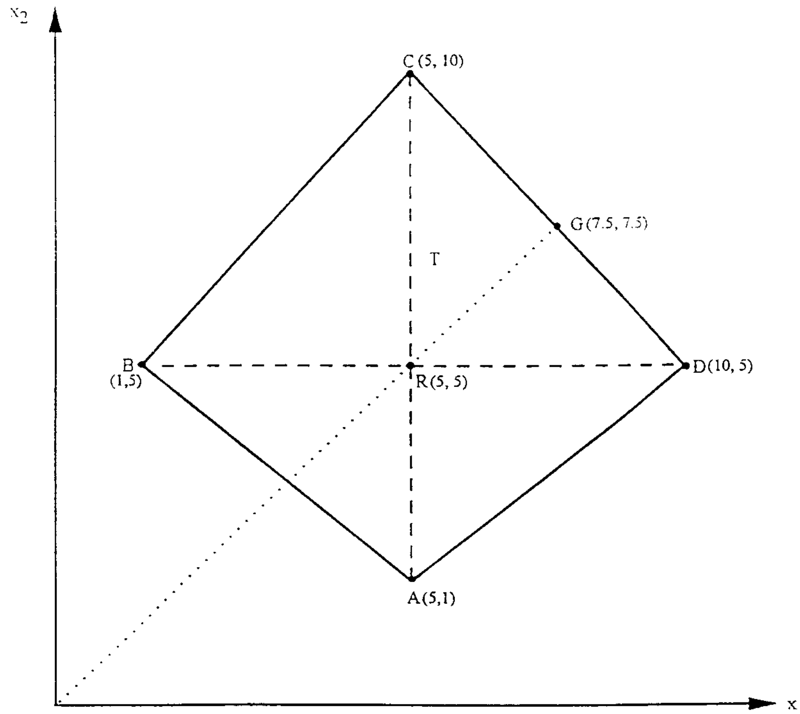

Fig. 1. Isoquant diagram.

$$\max \phi$$

$$\text{s.t.} \quad 1\phi = 1\lambda_A + 1\lambda_B + 1\lambda_C + 1\lambda_D + 1\lambda_G + 10\lambda_R - s^+$$

$$7.5 = 5\lambda_A + 1\lambda_B + 5\lambda_C + 10\lambda_D + 7.5\lambda_G + 5\lambda_R + s_1^-$$

$$7.5 = 1\lambda_A + 5\lambda_B + 10\lambda_C + 5\lambda_D + 7.5\lambda_G + 5\lambda_R + s_2^-$$

$$1 = \lambda_A + \lambda_B + \lambda_C + \lambda_D + \lambda_G + \lambda_R.$$

With all variables constrained to be non-negative, the solution is $\lambda_R^* = 1$, $s_1^{-*} = s_2^{-*} = 2.5$ which, with all other variables zero, gives $\phi^* = 10$. Thus, G is being evaluated by R as indicated by $(7.5 - s_1^{-*}, 7.5 - s_2^{-*}) = (5,5)$ which are the coordinates for R in Fig. 1.

Turning to Eq. (3′) for the second stage of the FGL evaluation, we write

$$\max \hat{\beta}$$

$$\text{s.t.} \quad 1\hat{\beta} = 1\lambda_A + 1\lambda_B + 1\lambda_C + 1\lambda_D + 1\lambda_G + 10\lambda_R - s^+$$

$$7.5\tau = 5\lambda_A + 1\lambda_B + 5\lambda_C + 10\lambda_D + 7.5\lambda_G + 5\lambda_R$$

$$7.5\tau = 1\lambda_A + 5\lambda_B + 10\lambda_C + 5\lambda_D + 7.5\lambda_G + 5\lambda_R$$

$$1 = \lambda_A + \lambda_B + \lambda_C + \lambda_D + \lambda_G + \lambda_R.$$

With all variables constrained to be non-negative, the solution is $\tau^* = 5/7.5 \approx 0.667$, $\lambda_R^* = 1$ and all other variables zero so that $\beta^* = 10$. Hence, the stage 1–stage 2 approach of FGL gives the following ratio,

$$C(\phi^*, \hat{\beta}^*) = \frac{\phi^*}{\hat{\beta}^*} = \frac{10}{10} = 1.$$

This solution fails to identify the congestion that is present in G's record as exhibited in Fig. 1. Thus, in conformance with our theorem, the FGL ratio approach fails to identify congestion when, as in this case, the law of variable proportions is not applicable.

The approach in BCSW [2], however, gives the correct result. This is seen by writing its second-stage problem as follows:

$$\max \delta_1^+ + \delta_2^+$$

$$s.t. \quad 10 = 1\lambda_A + 1\lambda_B + 1\lambda_C + 1\lambda_D + 1\lambda_G + 10\lambda_R$$

$$5 = 5\lambda_A + 1\lambda_B + 5\lambda_C + 10\lambda_D + 7.5\lambda_G + 5\lambda_R - \delta_1^+$$

$$5 = 1\lambda_A + 5\lambda_B + 10\lambda_C + 5\lambda_D + 7.5\lambda_G + 5\lambda_R - \delta_2^+$$

$$1 = \lambda_A + \lambda_B + \lambda_C + \lambda_D + \lambda_G + \lambda_R$$

$$2.5 \geqslant \delta_1^+$$

$$2.5 \geqslant \delta_2^+$$

With all variables constrained to be non-negative, the solution is $\lambda_R^* = 1$ and all other variables zero so that $\delta_1^{+*} = \delta_2^{+*} = 0$. Thus, $s_i^c = s_i^{-*} = 2.5$ as obtained from Eq. (7), for i = 1,2 correctly identifies the congestion amounts in the inputs used by G.

As should now be seen, the equivalences that have been discovered between FGL and BCSW hold only when the law of variable proportions is applicable. Thus, before FGL is used to evaluate any DMU_o, it is desirable to see whether this law is applicable. An easy way to do this is as follows.

Note, first, that congestion will be present in DMU_o's performance only if there is a DMU_R which strictly dominates DMU_o in both its inputs and outputs. For each such DMU_R, one can then examine whether the following relations hold for any i and l

$$\frac{x_{io}}{x_{i\mathrm{R}}} = \frac{x_{lo}}{x_{l\mathrm{R}}}.$$

If there is one, then it is prudent to use BCSW. Unfortunately, the converse is not true. The fact that these relations do not hold for any $\mathrm{DMU_R}$ does not foreclose the possibility that such a point will be generated from a non-negative combination of other DMUs.

To continue our study of relations with the approach of FGL, we proceed as follows. Let $x(s_i^c)$ be an input subvector in which its ith component corresponds to $s_i^c \neq 0$, i.e. $x(s_i^c)$ is a congesting subvector. Next, let $\mathbf{X^C}$ be the set of all congesting subvectors obtained via Eq. (5). We then have:

Theorem 4. $x(s_i^c) \in \mathbf{X^C}$. *Furthermore, if Eq. (6) (or (6′)) yields a unique optimal solution, then* $\mathbf{X^C} = \{x(s_i^c)\}$.

Proof. We prove this theorem on the basis of Eq. (6). Let $A = \{i | s_i^c = 0\}$ *and* $\bar{A} = \{i | s_i^c \neq 0\}$. Then the constraints of Eq. (6) become

$$\sum_{j=1}^{n} \lambda_j x_{ij} = \theta^* x_{io} \quad i \in A$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} \leqslant \theta^* x_{io} \quad i \in \bar{A}$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} = y_{ro} + s_r^{+*} \quad r = 1, 2, \ldots, s \tag{9}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j \geq 0, \quad s_i^{-*} \geq \delta_i.$$

where θ^* is the optimal value to Eq. (1). This implies that $\theta *$ is a feasible solution to Eq. (5). Thus, $\alpha^* \leqslant \theta^*$, where α^* is the optimal value to (5) associated with A and $\bar{A}$. On the other hand, any optimal solution to (5) is a feasible solution (1); therefore, $\alpha^* \geqslant \theta^*$. Thus, $\theta^* = \alpha^*$ indicating that the input subvector associated with $\bar{A}$, $x(s_i^c)$, is a source of congestion. Therefore, $x(s_i^c) \in \mathbf{X^C}$.

Moreover, if Eq. (6) yields a unique optimal solution, then the solution in Eq. (9) is also unique. This means that $\theta^* = \alpha^*$ does not hold for other input subvectors. Thus, $\mathbf{X^C} = \{x(s_i^c)\}$. □

Theorem 4 indicates that under the condition of uniqueness, congestion will occur in the

BCSW approach if and only if it appears in the FGL formulation as well. However, the BCSW approach identifies technical or mix inefficiencies and distinguishes these from congestion components via Eq. (7).

We may observe that the use of Eq. (5) may result in different congestion factors because of possible multiple optimal solutions. The above theorem indicates that the results from Eq. (6) then yield one of the congesting subvectors obtained from Eq. (5). As a result, the procedure by FGL for detecting the factors responsible for congestion may be replaced by model Eq. (6). One can thus more easily find and identify congestion and its sources without having to conduct a series of solutions as required for Eq. (5).

Next, we recall that all DMU evaluations can be classified into four groups, E, E′, F, N, (Charnes et al. [4]): (i) set E contains all extreme efficient DMUs, (ii) set E′ contains efficient DMUs that can be expressed as convex combinations of DMUs in set E; (iii) set F contains weakly efficient DMUs that are frontier DMUs but have non-zero DEA slacks, and (iv) set N contains DMUs not on the frontier.

Cooper et al. [9] refer to F as the 'extended frontier' in order to distinguish it from the 'efficiency frontier'. The presence of F can create problems because of the existence of weakly efficient DMUs consisting of frontier points with non-zero slacks that must be reduced to zero to attain full DEA efficiency. However, if the DEA frontier is composed solely of extreme efficient DMUs, as described in (i) above, we then have the following result.

Theorem 5. *If the observed values on the efficient frontier are composed only of extreme efficient DMUs, then congestion can occur if and only if non-zero DEA slack values are detected in Eq. (2) (or (2′)*). Furthermore, the sources of congestion can then only be found in these non-zero DEA slack values.

Proof. The proof is obvious from the results stated in theorem 1. □

This theorem can be important in real world applications, since the frontiers (defined by the first three groups of DMUs) in most real world data sets contain only the extreme efficient DMUs. Consequently, the input congestion and its amount can simply be represented by the DEA slacks defined in Eqs. (2) (or (2′)) (see Ray et al. [16]).

To illustrate the use of theorem 5, we here revisit the Chinese data set analysed by BCSW in a study of congestion in Chinese production before and after the 1978 economic reforms.[2] In order to see whether the frontier of this data set consists of only extreme efficient DMUs, we employ the following super-efficiency DEA model[3] in which the DMU under evaluation is excluded from the reference set to classify the data set into different efficiency groups:

$$\tilde{\theta}^* = \min \tilde{\theta}$$

[2] See Cooper et al. [10] for a complete discussion of this data set.

[3] This is due to Andersen and Petersen [1] who use it to rank DMUs. It has also been used for sensitivity analysis [17,20]. For a complete discussion of this type od DEA model and its uses, see [18].

$$
\begin{aligned}
s.t. \quad & \sum_{j \neq o,\ j=1}^{n} \lambda_j x_{rj} \leqslant \tilde{\theta} x_{io} \quad i = 1, 2, \ldots, m \\
& \sum_{j \neq o,\ j=1}^{n} \lambda_j y_{rj} \geq y_{ro} \quad r = 1, 2, \ldots, s \\
& \sum_{j \neq o,\ j=1}^{n} \lambda_j = 1 \\
& \lambda_j \geq 0 \ j \neq o, \quad j = 1, \ldots, n.
\end{aligned} \tag{10}
$$

Thrall [19] shows that (i) if $\tilde{\theta}^* > 1$ or Eq. (10) is infeasible, then DMU_o belongs to E; (ii) if $\tilde{\theta}^* = 1$, then DMU_o belongs to the union of E′ and F; and (iii) if $\tilde{\theta}^* < 1$, then DMU_o belongs to N. We can illustrate by applying Thrall's theorem to the data on Chinese production in Cooper et al. [10] for textiles, chemicals and metallurgies during the period 1966–1988. Table 1

Table 1
Super-efficiency scores in the textile, chemical, and metallurgy industries (1966–1988)

Year	Textile	Chemical	Metallurgy
1966	1.03374	1.03732	1.03681
1967	1.06422	1.03020	0.97543
1968	1.00017	1.03328	0.91046
1969	1.02344	1.05570	0.87747
1970	1.03198	1.01878	0.88253
1971	1.01987	0.99468	0.97405
1972	0.98439	0.98099	1.13660
1973	0.95807	0.97602	1.01185
1974	0.94559	0.98392	1.00539
1975	0.93955	0.98788	1.03259
1976	0.91511	0.98359	0.92188
1977	0.89158	0.95898	0.90363
1978	1.01286	1.06580	1.03261
1979	1.00458	0.99094	1.02684
1980	1.03478	0.94167	0.99712
1981	1.28058	0.95448	0.97904
1982	0.83021	1.00130	0.99712
1983	0.84044	1.00088	1.02660
1984	0.97463	1.00984	0.99723
1985	1.03691	1.00237	0.94756
1986	0.95112	1.02410	1.00147
1987	1.00174	0.98796	1.01637
1988	Infeasible	Infeasible	Infeasible

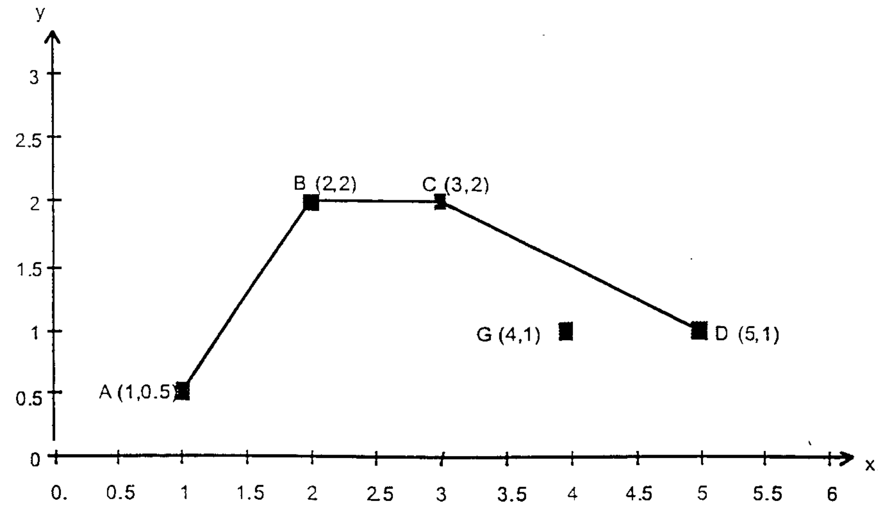

Fig. 2. Illustration of congestion.

gives the super-efficiency scores in these industries, where, as can be seen, no DMUs are in $E' \cup F$, with $\tilde{\theta}^* = 1$. The only points that occur are in N, with $\tilde{\theta}^* < 1$, or in E, with $\tilde{\theta}^* > 1$ or infeasible. Therefore, the congestion amounts can be directly obtained as slacks from Eq. (1). This is confirmed by the results in BCSW [2] where the labor congestion amount is equal to the corresponding labor slack value in each industry.

Finally, one may notice that results from the congestion measures used may depend on the orientation of the DEA model used. We illustrate this with Fig. 2, which is adapted from BCSW [2] with y representing the single output amount produced by each DMU and x representing the single input used. Only the line segment connecting A and B is efficient. C is not efficient because it has one unit of slack in its input. D, another point on the frontier, is congested because its input amount $x = 5$ is associated with a reduction of output from $y = 2$, at C or B, to a value of $y = 1$ at D.

The performance of DMU associated with G also exhibits congestion by the same output reduction as D, but with a smaller value for its input congestion. The latter needs to be reduced from $x = 4$ to $x = 3$ in order to achieve coincidence with the corresponding input amount in C. However, if an input-oriented DEA model is used, as in Eq. (1), for instance, the following solution is obtained for the evaluation of DMU_G: $\lambda^*_A = 2/3$, $\lambda^*_B = 1/3$, $\theta^* = 1/3$ and all zero slacks. Yet, as observed, production is congested at this point even though $C(\theta^*, \beta^*) = 1$ as indicated by the FGL measure.[4] Also, if we use an input-oriented BCSW approach (6), we will have $s^c = s^{-*} - \delta^{+*} = 0 - 0 = 0$, indicating no congestion. However, if we turn to the output-

[4] The output-oriented FGL ratio measure (Eq. (4′)) also fails to identify congestion at G (see Appendex).

oriented BCSW approach as given in Eqs. (6′) and (7), we obtain $s^c = s^{-*} - \delta^{+*} = 2 - 1 = 1$, as the congesting amount of input. This is a correct characterization.

4. A unified additive model approach

We now provide an approach in which additive models are used for both congestion and inefficiency analyses. First, we recall that Eq. (2) may be interpreted as a version of the additive model as first provided in Charnes et al. [5]. In fact, setting $\theta^* = 1$ in Eq. (2) produces the ordinary form of the additive model. Following BCSW [2], we would employ Eqs. (1) and (2) followed by Eq. (7) to obtain an extension of both slack (=mix or technical inefficiency) and congestion via Eq. (8). This, however, is not the end of the trail. We can, in fact, replace this use of a mixture of models with a unified approach based entirely on the additive model.

To see how this can be done, we start with the following model

$$\max \frac{\sum_{r=1}^{s} \frac{s_r^+}{y_{ro}}}{s} + \varepsilon \frac{\sum_{i=1}^{m} \frac{s_i^-}{x_{io}}}{m}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = y_{ro} \quad r = 1, 2, \ldots, s \tag{11}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, s_i^-, \quad s_r^+ \geqslant 0$$

where $\varepsilon > 0$ is a non-Archimedean element that is smaller than any positive real number and remains so no matter how large the value of the real number $\sum_{i=1}^{m} \frac{s_i^-}{x_{io}}$.

The formulation in Eq. (11) accords preemptive priority to maximizing $\sum_{r=1}^{s} \frac{s_r^+}{y_{ro}}$. This modification of the usual additive model is employed because the latter seeks to maximize the distance to the efficient frontier in a manner that maximizes inputs and outputs simultaneously in the sense of a vector optimization. Here, however, we mean to accord priority to output maximization after which we try to identify all input congestion that may be present.

This is done as follows. First, we use the results from Eq. (11) to form the following model in which we normalize the input slacks:

$$\max \frac{\sum_{i=1}^{m} \frac{s_i^+}{x_{io}}}{m}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} = \hat{y}_{ro} \quad r = 1, 2, \ldots, s \tag{12}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, \quad s_i^-, \geqslant 0$$

where $\hat{y}_{ro} = y_{ro} + s_r^{+*}$, and s_r^{+*} represent the output slacks obtained from Eq. (11).

This yields a new set of maximal input slacks consistent with the thus adjusted outputs. We then attempt to 'back out' the maximal inputs which are interpreted as technical or mix inefficiencies of the ordinary variety. This backing out is accomplished by means of the following modification of Eq. (6):

$$\max \frac{\sum_{i=1}^{m} \frac{\delta_i^+}{x_{io}}}{m}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} - \delta_i^+ = \hat{x}_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} = \hat{y}_{ro} \quad r = 1, 2, \ldots, s \tag{13}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, \quad \delta_i^- \geqslant 0$$

where $\hat{x}_{io} = x_{io} - s_i^{-*}$ and s_i^{-*} are the optimal slacks obtained in the second stage optimization implied by Eq. (11).

This accounts for the congestion (as well as the technical and mix inefficiency) components of each input. The measures to be used are in the objectives. Note that we have separated the measures of output and input inefficiencies into two averages, in part because we cannot guarantee a bound of unity for the output proportions.

Recourse may be had to the measures described in Cooper et al. [8] when this is desired. Here, however, we use these outputs only in the form of a 'driving function' to obtain the input congestion measures of interest.

This brings us back to Eq. (7) for determining s_i^c, the amount of congestion in input i, as determined by subtracting δ_i^{+*}, the technical or mix inefficiencies that do not reduce any output.

To provide a rationale, we need only observe that the conditions

$$\begin{cases} \hat{x}_{io} = \sum_{j=1}^{n} x_{ij}\lambda_j & i = 1, \ldots, m \\ \hat{y}_{ro} = \sum_{j=1}^{n} y_{rj}\lambda_j & r = 1, \ldots, s \end{cases} \tag{14}$$

are associated with the maximization of output and maximization of input slacks, respectively, in Eq. (11). Hence, forcing a positive value of δ_i^{+*} into Eq. (14) to alter $\hat{x}_{io}$ to $\hat{x}_{io} + \delta_i^{+*}$ would require replacing the equalities with inequalities to produce a new solution with

$$\hat{x}_{io} < \sum_{j=1}^{n} x_{ij}\lambda_j,$$

for some i and

$$\hat{y}_{ro} > \sum_{j=1}^{n} y_{rj}\lambda_j$$

for some r.

In many cases, the output reduction resulting from congestion will be apparent from inefficiency. For a formal development that will handle all cases, however, we now replace Eq. (12) with

$$\max \frac{\sum_{r=1}^{s} \frac{\delta_r^-}{y_{ro}}}{s}$$

$$s.t. \quad \sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io} \quad i = 1, 2, \ldots, m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} + \delta_r^- = \hat{y}_{ro} \quad r = 1, 2, \ \ldots, s \tag{15}$$

$$\sum_{j=1}^{n} \lambda_j = 1$$

$$\lambda_j, \quad \delta_r^- \geqslant 0$$

where $\hat{y}_{ro}$ is defined as in Eq. (12) and the x_{io} are original data as in the inputs for Eq. (11). When the optimal solution for Eq. (11) is unique, the solution to Eq. (15) will simply reproduce the original data via $\hat{y}_{ro} - \delta_r^{-*} = y_{ro}$. When not unique, however, other possibilities may be present.

Finally, we compare the FGL approach with our own. Both input-orientation and output-orientation approaches are employed. Tables 2, 3 and 4 report the results on the three Chinese industries studied by BCSW [2]. In the FGL approach, the congestion is represented by

Table 2
Congestion in textile industry

Year	FGL		$\sum_{i=1}^{m} \frac{s_i^c}{x_{io}} / m$	
	Input-orientation	Output-orientation	Input-orientation	Output-orientation
1966	0	0	0	0
1967	0	0	0	0
1968	0	0	0	0
1969	0	0	0	0
1970	0	0	0	0
1971	0	0	0	0
1972	ε	0.01791	0.02101	0.02101
1973	ε	0.02549	0.03739	0.03701
1974	0.01295	0.02230	0.01409	0.03357
1975	0.00048	0.01307	0.00106	0.02183
1976	0.00747	0.03500	0.00809	0.03793
1977	0.02239	0.09763	0.02186	0.06104
1978	0	0	0	0
1979	0	0	0	0
1980	0	0	0	0
1981	0	0	0	0
1982	ε	0.03533	0.00715	0.06157
1983	ε	0.03793	0.00611	0.06562
1984	0	0.00955	0	0.02789
1985	0	0	0	0
1986	0.02649	0.01635	0.00930	0.02153
1987	0	0	0	0
1988	0	0	0	0

ε is a very small number.

Table 3
Congestion in chemical industry

Year	FGL		$\sum_{i=1}^{m} \frac{s_i^c}{x_{io}} / m$	
	Input-orientation	Output-orientation	Input-orientation	Output-orientation
1966	0	0	0	0
1967	0	0	0	0
1968	0	0	0	0
1969	0	0	0	0
1970	0	0	0	0
1971	0	0	0	0
1972	ε	0.00422	0.02120	0.01738
1973	ε	0.00142	0.00146	0.00370
1974	0	0	0	0
1975	0	0	0	0
1976	0.00578	0.01661	0.07463	0.05690
1977	0.03170	0.04258	0.11247	0.07469
1978	0	0	0	0
1979	0	0	0	0
1980	0.05833	0.03198	0.01376	0.02647
1981	0.02004	ε	0.00576	0.06506
1982	0	0	0	0
1983	0	0	0	0
1984	0	0	0	0
1985	0	0	0	0
1986	0	0	0	0
1987	0.01204	0.00583	0.00114	0.00642
1988	0	0	0	0

ε is a very small number.

$1-C(\theta^*, \beta^*)$ (input-orientation) and $C(\phi^*, \beta^*)-1$ (output-orientation). In our approach, 'input-orientation' means that we use an input-oriented DEA model first and then back out the congestion; 'output-orientation' means that we use the unified additive approach.

It can be seen that in nine cases (six input-orientation and three output-orientation), the input congestion is a very small number (close to zero) in FGL measure, whereas our approach shows that, on average, there is substantial input congestion amounts. Also, note that different orientations lead to different results on input congestion. This indicates that the focus on an empirical study, e.g. measuring output losses, should be carefully set. Finally, note that the congestion results under output-orientation are consistent with those in BCSW.

5. Conclusions

This paper establishes an equivalence between the congestion measures used by Brockett et al. [2] and Färe et al. (FGL) [12,13]. FGL provides a measure of output loss estimated in the manner described in the addendum to Brockett et al. [2] whereas the approach used in our

Table 4
Congestion in metallurgy industry

Year	FGL		$\sum_{i=1}^{m} \frac{s_i^c}{x_{io}} / m$	
	Input-orientation	Output-orientation	Input-orientation	Output-orientation
1966	0	0	0	0
1967	0.00161	0	0.00407	0
1968	0.00496	0	0.01261	0
1969	0.01849	0	0.03587	0
1970	0.11747	0.42124	0.08125	0.08864
1971	0	0	0	0
1972	0	0	0	0
1973	0	ε	0	0.04284
1974	0	0	0	0
1975	0	0	0	0
1976	0	ε	0	0.01873
1977	0	0.01947	0	0.04424
1978	0	0	0	0
1979	0	0	0	0
1980	0	0	0	0
1981	0	0	0	0
1982	0	0	0	0
1983	0	0	0	0
1984	0	0	0	0
1985	0	0	0	0
1986	0	0	0	0
1987	0	0	0	0
1988	0	0	0	0

ε is a very small number.

paper identifies all the congesting (and other) amounts and provides measures of congestion, technical (mix) and total (= technical + congestion) inefficiency as well.

Also, as noted in Brockett et al. [2], the FGL measure only gives a greatest lower bound for the case of multiple outputs, whereas our measure, covers any finite number of inputs.

Here, we have focused on the aspects of production as represented in technical, mix and congestion inefficiencies. Other aspects of inefficiencies can be handled by associating prices and costs with output and input slacks, as is done in the appendix to Cooper et al. [8]. Hence, nothing appears to be lost and much appears to be gained by using varieties and extensions of the additive model, as is done here, to obtain a unified approach to all these aspects of inefficiencies.

Acknowledgements

Support for this research from the IC^2 Institute of the University of Texas at Austin is gratefully acknowledged

Appendix A. FGL output-oriented congestion ratio

Consider the FGL output-oriented congestion ratio Eq. (4′) for point G in Fig. 2. We use Eq. (1′) to obtain the following alternate optima in the first stage evaluation for the performance of G.

$$\phi^* = 2, \quad \lambda_{\mathrm{B}}^* = 1, \quad s^{-*} = 2$$

or, alternatively

$$\phi^* = 2, \quad \lambda_{\mathrm{C}}^* = 1, \quad s^{-*} = 1.$$

A similar situation applies at the second stage, i.e., model Eq. (3′), since

$$\hat{\beta}^* = 2, \quad \lambda_{\mathrm{B}}^* = 1, \quad \tau^* = 1/2$$

or, alternatively

$$\hat{\beta}^* = 2, \quad \lambda_{\mathrm{C}}^* = 1, \quad \tau^* = 3/4.$$

However, the combination represented by

$$C(\phi^*, \hat{\beta}^*) = \frac{\phi^*}{\hat{\beta}^*} = \frac{2}{2} = 1$$

fails to identify G's performance with congestion. Thus, although the values of $\phi^* = \beta^* = 2$ correctly identify the output that should have been achieved, the ratio test suggested by FGL fails in this case. One reason for the failure is the inapplicability of the 'law of variable proportions', which cannot be applied to the single input-single output case. Another shortcoming is the inability to distinguish between points like B and C, as shown by the above alternate optima, because of recourse to free disposal assumptions. To see that this is so, we turn to the fact that our approach identifies the first of the stage-one solutions as uniquely optimal because its slack value is maximal. Stage-two then decomposes these two units of total slack into a congesting component of $s^c = 1$ so that the residual (also = 1) represents the technical (or mix)[5] inefficiency of one unit involved in moving from point C to point B.

References

[1] Andersen P, Petersen NC. A procedure for ranking efficient units in data envelopment analysis. Management Science 1993;39:1261–4.
[2] Brockett PL, Cooper WW, Shin HC, Wang Y. Inefficiency and congestion in Chinese production before and after the 1978 economic reforms. Socio-Econ Plann Sci 1998;32:1–20.
[3] Charnes A, Cooper WW, Rhodes E. Measuring efficiency of decision making units. Eur J Opl Res 1978;2:429–44.

[5] In the one input-one output case, it is not possible to distinguish between technical and mix inefficiencies.

[4] Charnes A, Cooper WW, Thrall RM. A structure for classifying and characterizing efficiencies and inefficiencies in DEA. J of Productivity Analysis 1991;2:197–237.
[5] Charnes A, Cooper WW, Golany B, Seiford LM, Stutz J. Foundations of data envelopment analysis for Pareto-Koopman's efficient empirical production functions. J of Econometrics 1985;30:1–17.
[6] Charnes A, Cooper WW, Lewin AY, Seiford LM. Data envelopment analysis: theory, methodology and applications. Kluwer Academic Publishers, 1994.
[7] Cooper WW, Gu B, Li S. Comparisons and evaluations of alternative approaches to the treatment of congestion in DEA. European Journal of Operational Research 1999; (submitted).
[8] Cooper WW, Park KS, Pastor JT. RAM: A range adjusted measure of inefficiency for use with additive models, and relations to other models and measures in DEA. Journal of Productivity Analysis 1999;11:5–42.
[9] Cooper WW, Thompson RG, Thrall RM. Introduction: extensions and new developments in DEA. Annals of Operations Research 1996;66:3–45.
[10] Cooper WW, Kumbhakar S, Thrall RM, Yu X. DEA and stochastic frontier analysis of the 1978 Chinese economic reforms. Socio-Econ Plann Sci 1995;29:85–112.
[11] Färe R, Grosskopf S. Measuring congestion in production. J of Economics 1983;43:257–71.
[12] Färe R, Grosskopf S, Lovell CAK. The measurement of efficiency of production. Kluwer-Nijhoff Publishing, 1985.
[13] Färe R, Grosskopf S, Lovell CAK. Production Frontiers. Cambridge University Press, 1994.
[14] Färe R, Grosskopf S. Congestion: A note. Socio-Econ Plann Sci 1998;32:21–4.
[15] Färe R, Svensson L. Congestion of production factors. Econometrica 1980;48:1745–53.
[16] Ray R, Seiford LM, Zhu J. Chinese state-owned enterprises behave as market entities. Omega 1998;16:263–78.
[17] Seiford LM, Zhu J. Stability regions for maintaining efficiency in data envelopment analysis. European Journal of Operational Research 1998;108(1):127–39.
[18] Seiford LM, Zhu J. Infeasibility of super-efficiency DEA models. INFOR 1999;37:174–187.
[19] Thrall RM. Duality, classification and slacks in DEA. Annals of Operations Research 1996;66:109–38.
[20] Zhu J. Robustness of the efficient DMUs in data envelopment analysis. European Journal of Operational Research 1996;90:451–60.