Hypercongestion in production correspondences: an empirical exploration

Walter Briec, Kristiaan Kerstens & Ignace Van de Woestyne

Routledge
Taylor & Francis Group

Hypercongestion in production correspondences: an empirical exploration

Walter Briec[a], Kristiaan Kerstens[b] and Ignace Van de Woestyne[c]

[a]University of Perpignan, Perpignan, France; [b]CNRS-LEM (UMR 9221), IESEG School of Management, Lille, France; [c]KU Leuven, Research unit MEES, Brussel, Belgium

ABSTRACT

This empirical contribution reviews the rather limited existing literature measuring congestion in production. It first compares current ways to measure congestion using nonparametric specifications of technologies. In particular, it focuses on the magnitude and incidence of the congestion detected in empirical studies using traditional radial efficiency measures. Thereafter, it shows the limitations of this radial measurement and how alternative measurement schemes may reveal higher amounts of congestion. Then, the new, more general methodology of measuring S-congestion is presented. In particular, we first present a numerical example to illustrate the way the S-disposable technologies allow to capture more extreme forms of congestion by setting empirically determined upper bounds to the wasting of inputs. Then, an empirical illustration is presented based on an existing sample of data. A final section concludes.

I. Introduction

Traditionally, few empirical studies attempt to measure the phenomenon of congestion in production, here intuitively defined as production subject to negative marginal productivity. One of the few streams in the literature where some studies report on congestion makes use of multi-output nonparametric production technologies that impose either ray or free disposability to distinguish between technical inefficiency, i.e. production below the production frontier, and congestion, interpreted as a particular severe form of technical inefficiency. While the empirical analysis of efficiency and productivity has become quite popular (see e.g. Alam and Sickles 2000 or Kumar and Russell 2002), congestion is most often ignored in such studies, despite the fact that some studies find it to be the most important source of poor performance (e.g. Zhengfei and Oude Lansink 2003).

One prominent example of congestion is traffic congestion. In cities worldwide, congestion leads to reduced speeds and traffic flows over a given network. In extreme cases, traffic jams (e.g. due to an accident) can even temporarily destroy the whole throughput along an arc or at a node in a network resulting in a zero traffic flow. The latter extreme case is but one example of what one could label 'hypercongestion', whereby a total loss of output(s) occurs for certain combinations of inputs.

While there is a limited axiomatic literature allowing to reveal and measure some limited forms of congestion in production, it may be surprising to know that currently no economic production model is capable to reveal and measure the above phenomenon of hypercongestion. When there is an upper bound to the wasting of inputs in certain directions, then one can model hypercongestion phenomena leading to the complete destruction of outputs. While the limited forms of congestion in production are known as monotone output-limitational (MOL) congestion, the latter form of hypercongestion is known as output prohibitive (OP) congestion (see Färe and Svensson 1980 for definitions). Briec, Kerstens and Van de Woestyne (2016) are the first to develop a new axiomatic approach allowing for the definition of more general multi-output technologies capable of revealing all currently known (i.e. MOL and OP) congestion concepts.

In this contribution, we offer a complementary empirical perspective to these new theoretical developments. In particular, we set ourselves three targets. First, we want to make a preliminary inventory to document the amounts and incidence of congestion

that are empirically observed in the limited literature available. Second, we want to illustrate how the way one measures congestion affects the amounts that can be revealed. Third, we want to move beyond these forms of MOL congestion and also verify whether OP congestion matters in an empirical context. The latter form of congestion has to the best of our knowledge never been empirically documented.

Considering congestion as an extreme form of technical inefficiency, a key question is how one can think its existence when firms supposedly operate under a high degree of competition. While traditionally technical inefficiency is conceived as incompatible with competitive markets, the framework developed by Allais (1977) and later on reformulated by Luenberger (1995) at least allows to think of the dynamics of market exchanges out of equilibrium and it considers Walrasian equilibria as limiting states where inefficiencies in consumption and production are zero. In this view, inefficiencies and surpluses in the economy determine the dynamics of exchange and the battle to extract surpluses.

The existence of congestion is often related to the law of diminishing returns and it has been presented as both a law and a statistical regularity. In agriculture, crop response models that relate crop yields to essential single nutrients or combinations thereof (e.g. the macronutrients nitrogen, phosphorus, potassium, calcium, magnesium and sulphur, and the micronutrients boron, chlorine, copper, iron, manganese, molybdenum and zinc) almost universally reveal an initial phase with positive marginal product and limited substitution possibilities, the existence of a maximum plateau with zero marginal product and also a declining phase with negative marginal products (in soil science, the latter phase is called the toxic range of nutrients).[1] In hospitals, simulation models have determined a variety of causes contributing to facility congestion (e.g. poor scheduling practices; Johnson and Happ 1977), congested emergency departments due to bottlenecks in long-term care facility (see Patrick 2011). From this scant evidence from the agricultural and hospital sectors, it is clear that while the existence of the congestion phenomenon is beyond doubt, its causes seem to be industry specific.

This article is structured as follows. Section II provides some basic production axioms as well as definitions of the technology and its boundaries. It also discusses the representation of technologies by means of efficiency measures and distance functions. Section III introduces the nonparametric technologies that can be used to model some form of congestion and how congestion has been distinguished from technical efficiency in the literature. We review two types of empirical literatures containing some evidence on the amounts and incidence of congestion. Finally, we illustrate how the traditional radial way of measuring efficiency and congestion actually may underestimate the amounts of congestion.

Having summarized the theoretical and methodological parts, we now turn to the empirical sections. In Section IV, we illustrate the more general notion of S-congestion using a detailed numerical example with two input dimensions generating a single output assuming a convex hull (CH) technology. In particular, we illustrate the notion of S-disposability and measure I-congestion with $I \in S$ by means of a suitable directional distance function. We briefly elaborate on how to define a nonconvex hull (NCH) technology and illustrate its usefulness on the same numerical example. Thereafter, we present a small empirical Section V revisiting an existing data set that further illustrates the new S-congestion concepts. Section VI concludes.

II. Technology: axioms, subsets and representation

Following the literature, this contribution mainly focuses on input efficiency. A production technology describes all available possibilities to transform input vectors $x = (x_1, \ldots, x_m) \in \mathbb{R}^m_+$ into output vectors $y = (y_1, \ldots, y_n) \in \mathbb{R}^n_+$. The production possibility set or technology T summarizes the set of all feasible input and output vectors: $T = \{(x, y) \in \mathbb{R}^{m+n}_+ : x \text{ can produce } y\}$. Given our focus on input-oriented efficiency measurement, technology can be represented by the input correspondence $L : \mathbb{R}^n_+ \rightarrow 2^{\mathbb{R}^m_+}$, where $L(y)$ is the set of all input vectors that yield at least the output vector y:

---

[1]For soil science, see Jones (2001, 216–221) or Munson (1998). For agricultural economics, see Dillon and Anderson (1990, Ch. 2–3) and the survey in Paris (2008).

$$L(y) = \{x : x \text{ can produce } y\}. \qquad (1)$$

The list of axioms imposed on the input correspondence contains first of all the following three regularity properties:

L1: $\forall y \geq 0$ with $y \neq 0$: $0 \notin L(y)$ and $L(0) = \mathbb{R}^m_+$.
L2: $\forall x \in \mathbb{R}^m_+$: $\bigcap_{y \in \mathbb{R}^n_+} L(y) \cap (x - \mathbb{R}^m_+) = \varnothing$.
L3: $L(y)$ is closed $\forall y \in \mathbb{R}^n_+$.

Axiom L1 imposes no free lunch and the possibility of inaction. Assumptions L2 and L3 postulate the boundedness (i.e. infinite outputs can not be obtained from a finite input vector) and closedness of the input set. These regularity axioms are considered as self-evident and not amenable to testing.

In addition, there are other assumptions that may be invoked in various combinations on the input correspondence in the empirical applications:

L4: $L(y)$ is a convex set $\forall y \in \mathbb{R}^n_+$.
L5: If $x \in L(y)$, then $\lambda x \in L(y)$, $\forall \lambda \geq 1$.
L6: Let $u \in \mathbb{R}^m_+$. If there exists a $x \in L(y)$ with $u \geq x$, then $u \in L(y)$.
L7: Let $S \subset 2^{[m]}$ and $u \in \mathbb{R}^m_+$. If for every $I \in S$ there exists a $x_I \in L(y)$ with $u \geq_I x_I$, then $u \in L(y)$. Here,

$$u \geq_I x \iff \begin{cases} u_i \leq x_i & \text{if } i \in I \\ u_i \geq x_i & \text{else} \end{cases}.$$

Axiom L4 postulates that the input correspondence satisfies convexity: that is, linear combinations of activities are feasible. This is a widespread but not innocuous assumption, which is mainly invoked for convenience. Assumption L5 postulates ray or weak disposability of the inputs. Axiom L6 imposes strong or free disposal of inputs. While the former hypothesis allows only for a proportional increase in inputs to produce given amounts of outputs, the latter axiom implies that more inputs can always be used to generate an equal amount of outputs. Finally, axiom L7 imposes S-disposability introduced in Briec, Kerstens and Van de Woestyne (2016). Obviously, when $S = \{\varnothing\}$, S-disposability reduces to the strong disposability axiom L6.

For future reference, we recall some important definitions and results from Briec, Kerstens and Van de Woestyne (2016) concerning S-disposability and the related notions of congestion.

Definition 2.1 Let L be an input correspondence. Let $S \subset 2^{[m]}$. For all $y \in \mathbb{R}^n_+$, $L(y)$ satisfies a minimal S-disposability assumption if:

(a) $L(y)$ satisfies the S-disposal assumption, and
(b) $\nexists\ S' \subset S$ with $S' \neq S$ such that $L(y)$ satisfies the S'-disposal assumption.

Proposition 2.1 Let L be an input correspondence satisfying L1-L3. For all $y \in \mathbb{R}^n_+$, if $L(y)$ is nonempty, then $L(y)$ satisfies the S-disposal assumption if and only if

$$L(y) = \bigcap_{I \in S}(L(y) + K_I),$$

with $K_I = \{x \in \mathbb{R}^m : x \geq_I 0\}$.

Definition 2.2 Let L be an input correspondence and let S be a collection of subsets in $[m]$ that contains $\varnothing$. Let $y \in \mathbb{R}^n_+$. $L(y)$ is said to be S-congested if it is nonempty and fails the S-disposal assumption.

To measure technical efficiency and congestion, it is useful to distinguish between certain subsets of the input set $L(y)$. In particular, one distinguishes between three subsets denoting production units on the boundary. First, one can define the isoquant of an input set as

$$Isoq\ L(y) = \{x \in L(y) : \lambda x \notin L(y), \forall \lambda \in [0, 1[\}. \qquad (2)$$

Second, the weak efficient subset is defined by

$$WEff\ L(y) = \{x \in L(y) : u < x \Rightarrow u \notin L(y)\}. \qquad (3)$$

Finally, the efficient subset of an input set is defined as

$$EffL(y) = \{x \in L(y) : u \leq x \text{ and } u \neq x \Rightarrow u \notin L(y)\}. \qquad (4)$$

It is well known that $Eff\ L(y) \subseteq WEff\ L(y) \subseteq Isoq\ L(y) \subseteq L(y)$.

The subsets introduced in (3) and (4) can be generalized in the context of S-disposability as indicated in Briec, Kerstens and Van de Woestyne (2016). For instance, the generalization of (4) leads to the following definition:

Definition 2.3 Let L be an input correspondence and let $I \subset [m]$. For all $y \in \mathbb{R}^n_+$, the I-congested boundary is the subset $\{x \in L(y): u \leq_I x \text{ and } u \neq x \Rightarrow u \notin L(y)\}$.

Technologies can be characterized using distance functions. These distance functions are related to the efficiency measures defined by Farrell (1957). In the input orientation, this Farrell efficiency measure

$E^i(x, y)$ indicates the minimum contraction of an input vector by a scalar λ while still remaining on the boundary of the input set:

$$E^i(x, y) = \inf_{\lambda}\{\lambda : \lambda x \in L(y), \lambda \geq 0\}. \quad (5)$$

Instead, it is also possible to use the input directional distance function in some arbitrary direction $g \in \mathbb{R}^m$. This input directional distance function $D_L : \mathbb{R}^{m+n}_+ \times \mathbb{R}^m_+ \to \mathbb{R} \cup \{-\infty, +\infty\}$ is defined by

$$D_L(x, y; g) = \sup\{\delta : x - \delta g \in L(y)\}. \quad (6)$$

Obviously, the input directional distance function is more general compared to the Farrell input efficiency measure since it allows for measurement in different directions. When selecting $g = x$, then the input directional distance function is directly linked to the input-oriented Farrell efficiency measure: $D_L(x, y; x) = 1 - E^i(x, y)$.

Bearing in mind the theory dependency of observations (i.e. to observe a certain phenomenon, one needs a theoretical framework that permits to observe it), we need the directional distance function because of the flexibility of its directional vector that allows us to look for congestion in a precise way (see also infra on page 15).

Remark that precise definitions of MOL and OP congestion can be found in Briec, Kerstens and Van de Woestyne (2016, Definition 2.1 (70) and Definition 3.1 (84)).

III. Nonparametric technologies and congestion measurement

Nonparametric technologies: definitions, subsets and congestion measurement

Let us consider a set of J observations $A = \{(x_1, y_1), \ldots, (x_J, y_J)\} \in \mathbb{R}^{m+n}_+$. Nonparametric deterministic specifications of technology can be estimated by enveloping these observations while maintaining some basic production axioms (see Hackman 2008 or Ray 2004).[2]

First, we define both a weak and strong disposable technology under variable returns to scale (VRS).

Under strong input and output disposal (SD), a VRS technology is defined as

$$L(y)^{sd-vrs} = \left\{ x : x \geq \sum_{j=1}^{J} z_j x_j, y \leq \sum_{j=1}^{J} z_j y_j, \right.$$
$$\left. \sum_{j=1}^{J} z_j = 1, z \geq 0 \right\}. \quad (7)$$

From activity analysis, z is the vector of activity variables that indicates the intensity at which a particular activity is employed in constructing the reference technology. Under weak input disposal and strong output disposal (WD), a VRS technology is defined as

$$L(y)^{wd-vrs} = \left\{ x : x = \sum_{j=1}^{J} \mu z_j x_j, y \leq \sum_{j=1}^{J} z_j y_j, \right.$$
$$\left. \sum_{j=1}^{J} z_j = 1, \mu \geq 1, z \geq 0 \right\}. \quad (8)$$

Note that the inequalities on the input dimensions have now been replaced by an equality while the combinations of inputs defining the technology can be scaled up by the scalar μ. Finally, to fully illustrate the notion of S-disposal, we define a simple CH VRS technology that envelops all observations in the inputs as follows:

$$L(y)^{ch-vrs} = \left\{ x : x = \sum_{j=1}^{J} z_j x_j, y \leq \sum_{j=1}^{J} z_j y_j, \right.$$
$$\left. \sum_{j=1}^{J} z_j = 1, z \geq 0 \right\}. \quad (9)$$

The latter technology is based upon the one defined in Charnes et al. (1985), but it is fair to say that it has hardly ever been used in the economic literature (in contrast to the operations research literature).[3] Note that it is not straightforward to transform (9) into a NCH technology if one wishes to dispense with the convexity assumption.[4]

Figure 1 shows typical isoquants for nonparametric input sets with SD and WD under VRS starting from

---

[2]Olesen and Petersen (2016) provide an overview of recent developments to handle measurement errors, sample noise and specification errors when defining these production models.

[3]Note that strong disposability in the outputs is assumed, to have (9) in line with the general notion of an input correspondence stating that it is the set of all input vectors yielding at least the output vector y. However, even if the inequality is replaced by an equality and the notion of input correspondence is adjusted to allow for this case, similar computations as the ones reported here can be made.

[4]We refer to the "Technologies revealing S-congestion: a nonconvex perspective" section for more details on how to obtain a valid description using the notion of S-disposability in combination with nonconvexity.
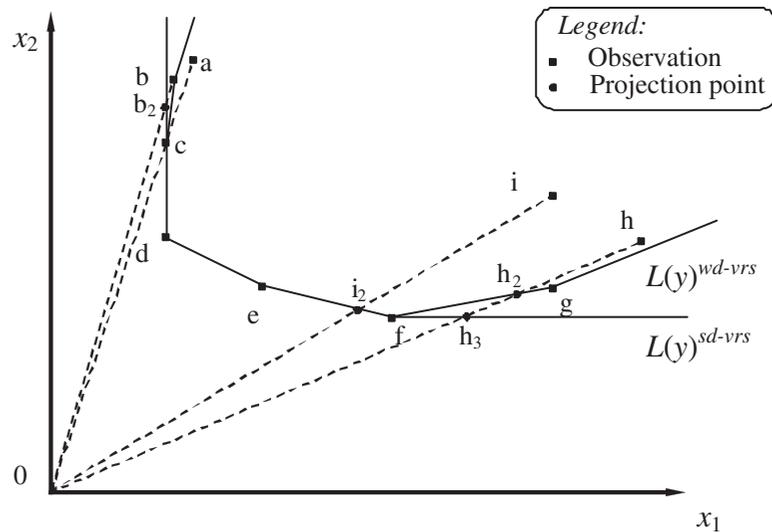
Figure 1. Input set and its subsets.

some basic observations. The SD technology (7) being rather simple, we focus on clarifying the WD technology (8). Note that the WD technology (8) is a subset of the SD technology (7). While the equalities in (8) can explain the line segments bc and fg, the scalar μ larger than or equal to unity is responsible for generating the rays emanating from points b and g, respectively.

To illustrate how the WD technology can model congestion, we start out from observation f. While the SD technology allows to waste additional inputs x_1 at no opportunity cost, the WD technology leaves two options: either the wasting of extra inputs x_1 requires additional costs in terms of extra inputs x_2 to reach, for instance, point g while maintaining current output levels, or the wasting of extra inputs x_1 without any additional inputs x_2 results in reaching another input level set of the WD technology corresponding to a lower level of outputs. In brief, wasting additional inputs x_1 has an opportunity cost in terms of either additional inputs x_2 or less outputs.

Having intuitively explained the notion of congestion, we develop two more issues. First, we clarify the three subsets ((2)–(4)) on these input sets. Then, we explain the radial way of measuring technical efficiency and congestion.

For both SD and WD technologies, the efficient subset $Eff\,L(y)$ consists of the line segments joining points def. For the WD technology, the weak efficient subset $WEff\,L(y)$ contains the connected line segments $cdef$, and its isoquant $Isoq\,L(y)$ is formed by adding the line segments bc and fg to those in

$WEff\,L(y)$. Points on the rays through $0b$ and $0f$ belong to the boundary of the input correspondence, not to any of its three subsets. For the SD technology, the weak efficient subset and the isoquant coincide: both contain the connected line segments $cdef$ and the lines beyond c and f parallel to both axes.

The traditional radial way of measuring technical efficiency and congestion as proposed in Färe, Grosskopf and Lovell (1983) can be illustrated by commenting on observation h situated in the interior of the input set $L(y)^{wd-vrs}$ in Figure 1. Technical efficiency is represented by the ratio of distances $0h_2/0h$ measured relative to the input set $L(y)^{wd-vrs}$. Structural efficiency or congestion is measured by the ratio $0h_3/0h_2$ derived by comparing radial distances between an activity without congestion at point h_3 on the weak efficient subset $WEffL(y)^{sd-vrs}$ and activity with congestion h_2 on the boundary of $L(y)^{wd-vrs}$.

Turning to a comparison of observations a and b, we obtain the following results. Since observation a is projected onto the weak efficient subset of both the SD and WD technologies, it does not suffer from congestion but the ratio of distances $0c/0a$ is just interpreted as technical inefficiency solely. By contrast, applying the same logic, observation b is situated on the $Isoq\,L(y)$ of the WD technology and hence technically efficient. However, the gap between the SD and WD technologies (i.e. $0b_2/0b$) reveals congestion. Noticing that observation a wastes more of both inputs than observation b for identical outputs, one may wonder why the latter is

considered congested but technically efficient, while the former is technically inefficient but uncongested. We return to this issue below.

This distinction between technical efficiency and congestion can be seen against a background of a variety of proposals of static efficiency taxonomies. The seminal article by Farrell (1957) clearly proposed the first basic measurement scheme distinguishing technical and allocative efficiency. Seitz (1970) adds a scale efficiency component based on cost function comparisons. Førsund and Hjalmarsson (1974) and later on Färe, Grosskopf and Lovell (1983) and Banker, Charnes and Cooper (1984) propose a distinction between technology-based technical and scale efficiency, whereby the second team of authors also integrate a congestion component. Färe, Grosskopf and Lovell (1985) were probably the first to offer an extended efficiency decomposition summarizing most of the above developments.

Crucial for our focus on congestion measurement in the remainder are the following remarks. First, we consider congestion as an extreme and unacceptable form of technical efficiency. While technical inefficiency is costly and implies a waste of resources, one can imagine certain reasons justifying its existence (e.g. slack resources and capacity in anticipation of an increasing demand over a product life cycle etc.). However, congestion implies a waste of resources and an additional opportunity cost in terms of additional inputs or wasted outputs. Therefore, it is almost impossible to justify and ideally requires prompt managerial action.

Second, one should clearly distinguish between detecting congestion and summarizing its relative importance as a source of inefficiency within some efficiency decomposition. While the radial efficiency measure (5) is convenient to summarize the relative importance of different efficiency components in a multiplicative decomposition, as illustrated above when comparing points a and b, it need not necessarily be an accurate tool to reveal the incidence of congestion (see also infra).

Radial congestion measurement: amounts and incidence

While congestion is widely cited as a theoretical possibility in most microeconomics textbooks, empirical evidence as to its prevalence is relatively rare. We draw on two different literatures providing some evidence as to its existence and/or incidence.

The merit of the first literature applying this non-parametric efficiency decomposition outlined above is that quite a lot of studies have reported on (parts of) these efficiency decompositions, though relatively few report on congestion. Congestion is the most important source of inefficiency at the sample level in at least four articles we are aware of: Byrnes and Färe (1987) and Byrnes, Färe, Grosskopf and Lovell (1988) both analyse US surface coal mines, Zhengfei and Oude Lansink (2003) assess Dutch agriculture, and Färe, Grosskopf, and Pasurka (1989) analyse US electric utilities. Just to offer some basic idea of the amount of waste involved, Table 1 summarizes for each study the average amount of congestion efficiency as well as its incidence (% of sample). The last column adds the sample size and some remarks whenever needed. Note that the second and fourth study have several entries: in the second article, a basic distinction is made between the interior and western US states, while the fourth study compares two distinct years. Furthermore, for the second study, we also report results for those subsamples for which congestion efficiency is the key component.

Several conclusions can be drawn from Table 1. First, congestion inefficiency can vary from a modest 7.5% (= 1 − 0.925) to a high 29% (= 1 − 0.71) at the sample level. In the second study, for Western non-union mines, one even observes a staggering 57% (= 1 − 0.43) congestion inefficiency. Second, the incidence of congestion inefficiency varies widely: between a low 26.3% and about 75% of the sample.[5] For the second study, two subsamples even record an incidence of 83.3%. Third, congestion inefficiency and incidence need not be correlated. For instance, the lowest incidence coincides with the highest congestion inefficiency (see Byrnes and Färe 1987). By contrast, the second lowest congestion inefficiency goes hand in hand with the highest incidence levels (Byrnes, Färe, Grosskopf and Lovell 1988). Finally, modest congestion inefficiency levels can hide high incidence levels (see Zhengfei and Oude Lansink 2003). In brief, these studies reveal a wide variety of patterns of congestion inefficiency and incidence, even though the sample sizes of most studies are quite modest.

---

[5]Remark that we have not been looking for studies reporting the highest congestion incidence levels relative to the incidence of other sources of poor performance (instead of the highest congestion inefficiency).

Table 1. Congestion efficiency and incidence: literature selection.

Article	Congestion efficiency	Congestion incidence (%)	Remarks
Byrnes and Färe (1987)	0.71	26.3	$N = 186$
Byrnes et al. (1988)	0.74	69.0	$N = 84$, interior states
	0.70	83.3	$N = 54$, interior states; UMWA
	0.77	74.3	$N = 113$, western states
	0.43	83.3	$N = 12$, western states; nonunion
Färe et al. (1989)	0.925	NA	$N = 23$, Year 1969
	0.924	NA	$N = 23$, Year 1975
Zhengfei and Oude Lansink (2003)	0.88	75.0[a]	$N = 1072$

UMWA: Affiliation with United Mine Workers of America;
NA: Not available.
[a]Text states: 'approximately 3/4 of observations' (p. 475).

Furthermore, while in some studies congestion inefficiency does not dominate at the sample level, it may well turn out to be important for specific parts of the sample. For example, Byrnes et al. (1987) document that congestion dominates for Illinois grain farms smaller than 700 ac, representing 72.9% of the sample.

Finally, there is the possibility that congestion plays a negligible role at the sample level or for specific parts of it, but that it is critically important for some particular observations. For instance, evaluating British building societies in 1985 and finding scale inefficiency as the prime source of underperformance, Field (1990) observes that congestion is most important for about 9.9% of observations (with amounts between 0.48 and 0.78).

A second literature worthwhile looking at are the parametric studies using flexible functional forms and testing for the satisfaction of monotonicity and/or curvature conditions. Barnett (2002) stresses that standard second-order conditions for optimizing behaviour fail to hold and duality relations break down if curvature or monotonicity conditions are violated. Sauer (2006) revisits in total eight parametric frontier studies. Three articles fulfil monotonicity in all inputs. Violations of monotonicity occur for a single input in two studies, for two inputs in two articles and even for five out of eight inputs in one study.[6] One obvious interpretation of these monotonicity failures is the existence of congestion.

In conclusion, from these two literatures, it is difficult to deny that congestion may well play a serious role as a source of poor performance in a relatively

wide range of sectors. Furthermore, the sometimes high incidence of congestion seems to indicate that in these samples, a lot of observations are situated close to the isoquant and boundary of the input set. At first sight, this seems to imply that the amounts of congestion measured are not artefacts created by just a few outlying observations enveloped by a particular axiomatic structure imposed on technology.

Radial congestion measurement: limitations

Especially, the sample level results are surprising if one realizes that the use of radial efficiency measures actually underestimates the prevalence of congestion. This can be easily illustrated with the help of Figure 2. Only observations outside the cone spanned by $WEff L(y)^{wd-vrs}$ (i.e. the rays Oc and Of in Figure 2) can be subject to congestion when using radial efficiency measures. Compare, for instance, points g and i in Figure 1. Point g is detected as being congested, since it is efficient relative to $L(y)^{wd-vrs}$ but not relative to $L(y)^{sd-vrs}$. However, point i, identical to g in its use of x_1 but using a higher amount of x_2, is not subjected to congestion, since the radial efficiency measure projects observation i onto the efficient subset at point i_2. Hence, the traditional radial way of measuring congestion may well underestimate its empirical amounts and/or its incidence.

If one is willing to accept the argument that one should distinguish between the detection of congestion and summarizing its relative importance as a source of inefficiency within some static efficiency decomposition, then it is easy to understand that some authors have proposed to measure congestion in a nonradial way.

One obvious possibility is to measure congestion per specific input dimension. This procedure is illustrated in Figure 2 for observation i that remained undetected using the radial measure. By contrast, measuring in the direction of the second input allows detecting its congesting excessive usage of inputs. To be precise, the distances $0g/0i$ and the ratios of distances $0g_2/0g [= (0g_2/0i)/(0g/0i)]$ measure the amount of technical efficiency and congestion in the direction of the second input, respectively. In a similar fashion, observation b

---

[6]We ignore the violations of curvature conditions equally reported by Sauer (2006).
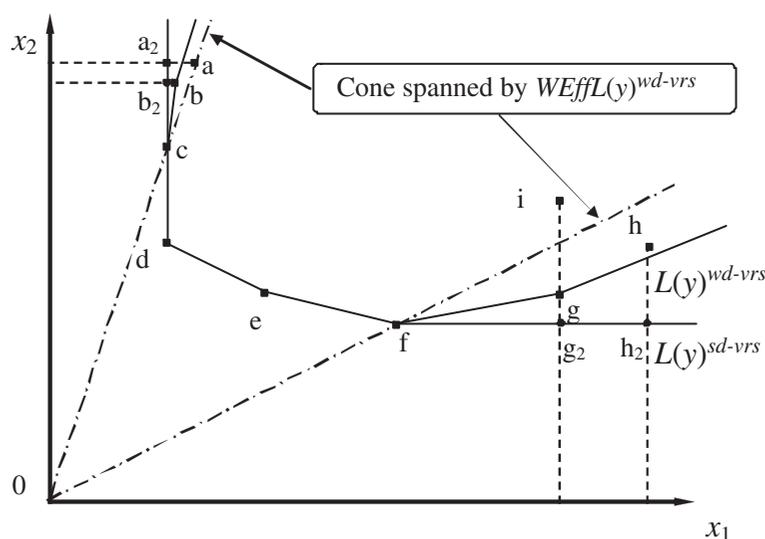
Figure 2. Limits of radial congestion measurement.

which remained uncongested using the radial measure may now be detected as being congested. Two studies are known to us that have implemented such uni-dimensional measurement scheme for congestion: Zhengfei and Oude Lansink (2003) and Färe et al. (1985).

Focusing on the Zhengfei and Oude Lansink (2003) study, while the radial input efficiency measure evaluated over all eight input dimensions leads to on average an amount of 11.7% congestion inefficiency, the use of a subvector measure per input dimension separately leads to average congestion inefficiency levels at the sample level from a minimum of 22.1% for 'other variable inputs' to a maximum of 45.6% for 'nitrogen fertilizer'. Using the radial input efficiency measure, the incidence of congestion is about 75%. The use of subvector measures per input dimension leads to incidence levels varying between a minimum of 35% for 'other pesticides' to 59% for 'nitrogen fertilizer'. One conclusion is clearly that the radial way of measuring congestion may underestimate the amounts of congestion inefficiency relative to an input-specific measurement scheme. The effect on congestion incidence is not clear cut.

Also, more refined measurement schemes have been devised looking for subsets of dimensions responsible for congestion (see Byrnes, Färe, Grosskopf, and Lovell 1988; Färe, Grosskopf, and Lovell 1994). Thus, it remains somewhat an open issue how to best measure congestion: radially, uni-dimensionally or some other way.

Finally, following Färe, Grosskopf and Logan (1987), the presence of congestion can also be interpreted as a violation of the WD assumptions. Observations that are inefficient with respect to a WD technology then simply suggest a lack of fit between the data and the WD assumption. To the extent that the goodness of fit with the WD assumption is low, this may lead to the search for alternative axioms (e.g. S-disposal) and resulting technology specifications yielding an even closer fit with the data. In the next sections, we move beyond MOL congestion to verify whether OP congestion matters from an empirical point of view. We do so by first discussing a numerical example to develop a basic understanding. Thereafter, we discuss some empirical results.

IV. Numerical illustrations of S-congestion

In this section, we first present a numerical example to develop some intuitions for a two input single output technology related to the measurement of S-congestion. To illustrate the notion of S-congestion and the ways of measuring it by means of a suitable directional distance function, we start from an artificial example containing 32 units with two inputs and one output. The data are provided in the first four columns of Table 2. Note that only two output levels are present to simplify the illustrations. Hereafter, all congestion computations are executed for a given output level of 2: therefore, all units can be included in the computations.

Table 2. Numerical example: data and S-congestion results from CH and NCH technologies and proportional projections with output set to 2.

				CH			NCH		
Unit	Input 1	Input 2	Output	{1}	{2}	{1, 2}	{1}	{2}	{1, 2}
1	3.5	10.0	2.0	3.5714	0.0000	0.2000	3.5714	0.0000	0.2000
2	3.0	7.0	2.0	4.2222	0.0000	0.7143	4.0000	0.0000	0.7143
3	3.0	5.0	2.0	4.0000	0.4000	1.4000	4.0000	0.4000	1.4000
4	4.0	3.0	2.0	2.3500	2.4000	2.4000	2.2500	2.3333	2.3333
5	5.0	2.0	2.0	0.6000	4.3000	2.2000	0.6000	4.0000	2.2000
6	8.0	2.0	2.0	0.0000	4.7500	1.0000	0.0000	4.5000	1.0000
7	13.0	2.5	2.0	0.0000	3.8000	0.2308	0.0000	3.8000	0.2308
8	13.0	3.0	2.0	0.0308	3.0000	0.2308	0.0000	3.0000	0.2308
9	13.0	4.0	2.0	0.0923	2.0000	0.2308	0.0769	2.0000	0.2308
10	4.0	4.0	2.0	2.5500	1.5500	1.8333	2.5000	1.5000	1.5000
11	7.0	6.0	2.0	1.1905	0.8750	0.7895	1.1429	0.8333	0.6667
12	8.0	6.0	2.0	0.9167	0.9167	0.7000	0.8750	0.8333	0.6667
13	4.0	10.0	2.0	3.0000	0.0200	0.2000	3.0000	0.0000	0.2000
14	5.0	10.0	2.0	2.2000	0.0600	0.2000	2.2000	0.0000	0.2000
15	6.0	11.0	2.0	1.6667	0.0000	0.0909	1.6667	0.0000	0.0909
16	10.0	12.0	2.0	0.6000	0.0000	0.0000	0.6000	0.0000	0.0000
17	14.0	10.0	2.0	0.1429	0.2000	0.0000	0.1429	0.2000	0.0000
18	16.0	8.0	2.0	0.0000	0.5000	0.0000	0.0000	0.5000	0.0000
19	15.0	5.0	2.0	0.0000	1.4000	0.0667	0.0000	1.4000	0.0667
20	14.5	9.0	2.0	0.1034	0.3333	0.0213	0.1034	0.3333	0.0000
21	14.0	4.0	2.0	0.0143	2.0000	0.1429	0.0000	2.0000	0.1429
22	11.0	10.0	2.0	0.4545	0.2000	0.0968	0.4545	0.2000	0.0000
23	5.0	10.0	4.0	2.2000	0.0600	0.2000	2.2000	0.0000	0.2000
24	4.0	7.0	4.0	2.9167	0.4571	0.7143	2.7500	0.4286	0.7143
25	4.0	4.0	4.0	2.5500	1.5500	1.8333	2.5000	1.5000	1.5000
26	8.0	3.0	4.0	0.6750	2.8333	1.0000	0.6250	2.6667	1.0000
27	10.0	3.0	4.0	0.3400	3.0000	0.6000	0.3000	3.0000	0.6000
28	14.0	5.0	4.0	0.0714	1.4000	0.1429	0.0714	1.4000	0.1429
29	15.0	7.0	4.0	0.0444	0.7143	0.0667	0.0000	0.7143	0.0667
30	14.0	9.0	4.0	0.1429	0.3333	0.0435	0.1429	0.3333	0.0000
31	10.0	11.0	4.0	0.6000	0.0909	0.0625	0.6000	0.0909	0.0000
32	9.0	8.0	4.0	0.7778	0.4688	0.3600	0.7778	0.3750	0.2500

Technologies revealing S-congestion

To measure congestion, one needs a technology that allows observing some form of congestion. Obviously, assuming SD (7) does not allow revealing congestion by definition. WD (8) allows for certain types of S-congestion and could be used if interest is limited to MOL congestion. But since there is no upper bound to wasting inputs in certain directions, it cannot detect OP congestion. Therefore, to allow for a full range of S-congestion measurements, we start from a CH VRS technology in the inputs for which the input correspondence is defined by (9).

Figure 3 displays the input correspondence $L(2)$ for output level 2. The boundary of this input correspondence is clearly visible as the region bounded by the union of lines between two consecutive points of the list of points labelled 1, 2, 3, 4, 5, 6, 7, 19, 18, 17, 16, 15 and 1. Obviously, $L(2)$ satisfies minimal S-disposability with $S = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}$. Consequently, three meaningful boundaries and corresponding directions can be identified for measuring congestion leading to $\{1\}$-, $\{2\}$- and $\{1, 2\}$-congestion measures. We first describe these three boundaries in great detail.

First, in Figure 3, the $\{1\}$-congested boundary is represented by the dashed line starting at the left-side horizontally towards point 5, then connecting points 5, 6, 7, 19 and 18 and from there continuing vertically. The set $L_{\{1\}}(2)$ contains all points located left and above of this $\{1\}$-congested boundary. Second, the $\{2\}$-congested boundary is shown by the dashed-dotted line starting at the left-side vertically towards point 3, then connecting points 3, 2, 1, 15 and 16 and from there continuing horizontally. The set $L_{\{2\}}(2)$ contains all points located right and below of this $\{2\}$-congested boundary. Finally, the $\{1, 2\}$-congested boundary is represented by the dotted line starting at the left-side horizontally towards point 16, then connecting points 16, 17 and 18 and from there continuing vertically. The set $L_{\{1,2\}}(2)$ contains all points located left and below of this $\{1, 2\}$-congested boundary.

Two remarks can be made at this point. First, note that also the $\varnothing$-congested boundary and its corresponding set $L_{\varnothing}(2)$ could be considered, which corresponds with the traditional SD boundary. However, measuring with respect to this boundary implies a reduction of at
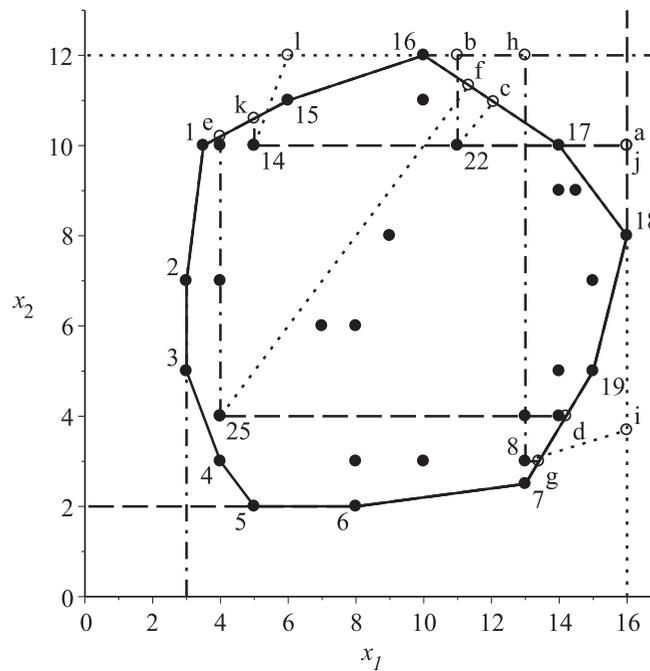
Figure 3. *S*-congested boundaries from a CH technology and proportional projections.

least one of the inputs. This makes it hard to interpret such a measurement as revealing congestion. Second, observe that the intersection of the sets $L_{\varnothing}(2)$, $L_{\{1\}}(2)$, $L_{\{2\}}(2)$ and $L_{\{1,2\}}(2)$ exactly corresponds with $L(y)$ demonstrating that $L(2)$ satisfies the *S*-disposal assumption according to Proposition 2.1.

Measuring S-congestion

Recall the input directional distance function D_L defined by (6). For efficiency measuring, $L(y)$ is the classical input correspondence at output level y and $g \in \mathbb{R}_+^m$. The directional distance function measures the maximal possible factor of g by which the input x can be decreased and still remains capable of producing y.

However, in the context of congestion measuring, this definition needs to be adjusted to allow projections in opposite (potentially congesting) directions. The adapted input directional distance function D_{L_I} with $I \in S \backslash \{\varnothing\}$ is now defined by

$$D_{L_I}(x, y; g_I) = \sup\{\delta : x - \delta g_I \in L_I(y)\}, \qquad (10)$$

with $g_I \in \mathbb{R}^m$ such that $g_{I_i} \leq 0$ for all $i \in I$ and $g_{I_i} = 0$ otherwise. This choice of direction vector g_I guarantees an increment of the input in the maximization process of the directional distance function and making this optimization process bounded as

well. The latter means that the corresponding *I*-congested boundary is hit by the directional distance function.

Thus, the definition of the direction vector g_I is complementary to the definition of the *I*-congested boundary in that a feasible finite solution of the adapted input directional distance function D_{L_I} is always guaranteed. This simply reflects the theory dependency of observations: the adapted input directional distance function D_{L_I} and *I*-congested boundary are theoretical constructs allowing to observe potential forms of OP congestion for any given empirical configuration of the production data, because the adapted input directional distance function D_{L_I} measures into a direction where it always will meet the *I*-congested boundary.

This adapted directional distance function D_{L_I} measures the maximal possible factor of $-g_I$ by which the input x can be increased until the boundary of $L_I(y)$ is hit. Or put differently, until the current level of production y can no longer be maintained. Since at this stage, OP congestion in the direction opposite of g_I is observed, the value of D_{L_I} can be used as a measurement of the presence or absence of *I*-congestion. The larger this value, the more further increases in inputs are needed before congestion occurs in the direction opposite of g_I. Values closer

to zero indicate that congestion occurs with only a slight increase of inputs in the direction opposite of g_I. A value of zero means that congestion in the direction opposite of g_I is already a fact.

For practical computations, it is often convenient to opt for a so-called input position dependent direction g_I related to input x for which $g_{I_i} = -x_i$ for all $i \in I$ and $g_{I_i} = 0$ otherwise. As a consequence, $1+$

To compute this adapted input directional distance function D_{L_I} with $I \in S \setminus \{\varnothing\}$ with input position dependent direction g_I relative to the CH and WD technologies, it suffices to solve a linear programme. First, applied to the case of the CH technology, this input directional distance function (10) can be computed using the following linear programming model (11):

$$D_{L_I}^{CH-VRS}(x^o, y^o, g_I) = \max_{\delta, z_j} \delta$$

$$\text{subject to} \quad \sum_{j=1}^{J} z_j x_{j,i} \geq x_i^o - \delta g_{I_i} \qquad i \in I$$

$$\sum_{j=1}^{J} z_j x_{j,i} \leq x_i^o \qquad i \in \{1, \ldots, m\} \setminus I$$

$$\sum_{j=1}^{J} z_j y_{j,k} \geq y_k^o \qquad k \in \{1, \ldots, n\} \tag{11}$$

$$\sum_{j=1}^{J} z_j = 1$$

$$z_j \geq 0 \qquad j \in \{1, \ldots, J\}.$$

D_{L_I} represents the factor by which inputs x_i with $i \in I$ can be scaled until I-congestion occurs, giving it a convenient proportional interpretation.[7]

When we consider the technology WD, the input directional distance function (10) requires solving the following mathematical programming model (12):

$$D_{L_I}^{WD-VRS}(x^o, y^o, g_I) = \max_{\mu, \delta, z_j} \delta$$

$$\text{subject to} \quad \sum_{j=1}^{J} \mu z_j x_{j,i} \geq x_i^o - \delta g_{I_i} \qquad i \in I$$

$$\sum_{j=1}^{J} \mu z_j x_{j,i} \leq x_i^o \qquad i \in \{1, \ldots, m\} \setminus I$$

$$\sum_{j=1}^{J} z_j y_{j,k} \geq y_k^o \qquad k \in \{1, \ldots, n\} \tag{12}$$

$$\sum_{j=1}^{J} z_j = 1$$

$$\mu \geq 1$$

$$z_j \geq 0 \qquad j \in \{1, \ldots, J\}.$$

---

[7]Obviously, much more can be said on this choice of direction vector. We just refer to the recent literature discussing various choices of direction vectors and their consequences: see e.g. Atkinson and Tsionas (2016), Daraio and Simar (2016), and Peyrache and Daraio (2012), among others.

Note that the WD technology defined by (8) can be linearized using the substitution $z_{j'} = \mu z_j$, $(j \in \{1, \ldots, J\})$, which is applied to obtain (13):

$$
D_{L_I}^{WD-VRS}(x^o, y^o, g_I) = \max_{\mu, \delta, z_{j'}} \delta
$$

subject to

$$
\sum_{j=1}^{J} z_{j'} x_{j,i} \geq x_i^o - \delta g_{I_i} \qquad i \in I
$$

$$
\sum_{j=1}^{J} z_{j'} x_{j,i} \leq x_i^o \qquad i \in \{1, \ldots, m\} \backslash I
$$

$$
\sum_{j=1}^{J} z_{j'} y_{j,k} \geq \mu y_k^o \qquad k \in \{1, \ldots, n\} \tag{13}
$$

$$
\sum_{j=1}^{J} z_{j'} = \mu
$$

$$
\mu \geq 1
$$

$$
z_{j'} \geq 0 \qquad j \in \{1, \ldots, J\}.
$$

Returning to the numerical example, for the fixed output level of 2, the adapted directional distance function values can be computed for all units with respect to the $\{1\}$-, $\{2\}$- and $\{1, 2\}$-congested boundaries. The directional distance function values based on the input position-dependent direction are recorded in the last three columns of Table 2. We explain the interpretation of the directional distance function values using the two points 22 and 25. Both these observations are also depicted in Figure 3.

To measure $\{1\}$-congestion for point 22, we compute $D_{L_{\{1\}}}(x, 2; g_{\{1\}}) = \delta^* = 0.4545$ with $x = (11, 10)$ and $g_{\{1\}} = (-11, 0)$. It follows from the definition that $x_1 - \delta^* g_{\{1\}_1} = x_1 + \delta^* x_1 \quad = x_1(1 + \delta^*) = 11(1 + 0.4545) = 16$. Obviously since $g_{\{1\}_2} = 0$, $x_2 - \delta^* g_{\{1\}_2} = x_2 = 10$. Therefore, the optimal projection of point 22 has coordinates $(16, 10)$. Looking at Figure 3, observe that point 22 is projected following the horizontal dashed line onto the point labelled a located at the $\{1\}$-congested boundary. Note that the coordinates of this point exactly correspond with the above computations. The value $1 + D_{L_{\{1\}}} = 1 + \delta^* = 1 + 0.4545 = 1.4545$ indicates the factor by which the first input of point 22 can be multiplied before $\{1\}$-congestion occurs. Analogously, the $\{1\}$-congestion measurement

for point 25 results in a value of 2.55 meaning that the first input 4 can be multiplied with a factor $1 + 2.55 = 3.55$ before hitting the $\{1\}$-congested boundary. This yields a first input of $3.55 \times 4 = 14.2$ which is the first coordinate of the point labelled d in Figure 3.

The measurement of $\{2\}$-congestion for point 22 is computed by $D_{L_{\{2\}}}(x, 2; g_{\{2\}}) = \delta^* = 0.2$ with $x = (11, 10)$ and $g^{\{2\}} = (0, -10)$. Again by definition, $x_2 - \delta^* g_{\{2\}_1} = x_2 + \delta^* x_2 = x_2(1 + \delta^*) = 10(1 + 0.2) = 12$. Since $g_{\{2\}_1} = 0$, $x_1 - \delta^* g_{\{2\}_1} = x_1 = 11$. Thus, the optimal projection of point 22 now has coordinates $(11, 12)$ which corresponds in Figure 3 with vertical projection of point 22 following the vertical dash-dotted line onto the point labelled b located at the $\{2\}$-congested boundary. The value $1 + D_{L_{\{2\}}} = 1 + \delta^* = 1 + 0.2 = 1.2$ provides the factor by which the second input of point 22 can be multiplied before $\{2\}$-congestion occurs. Analogously, the $\{2\}$-congestion measurement for point 25 results in a value of 1.55 meaning that the second input 4 can be multiplied with a factor $1 + 1.55 = 2.55$ before reaching the $\{2\}$-congested boundary. The resulting second input yields a value of $2.55 \times 4 = 10.2$ which is the second coordinate of the point labelled e in Figure 3.

For the measurement of $\{1, 2\}$-congestion for point 22, we compute $D_{L_{\{1,2\}}}(x, 2; g_{\{1,2\}}) = \delta^* = 0.0968$ with $x = (11, 10)$ and $g_{\{1,2\}} = (-11, -10)$. It follows from the definition that $x_1 - \delta^* g_{\{1,2\}_1} = x_1 + \delta^* x_1 =$

$x_1 \ (1 + \delta^*) = 11(1 + 0.0968) = 12.0465$. Again by definition, $x_2 - \delta^* g_2^{\{1,2\}} = x_2 + \delta^* x_2 \ (1 + \delta^*) = 10 \ (1 + 0.0968) = 10.968$. Therefore, the optimal projection of point 22 has coordinates $(12.0465, 10.968)$. Looking again at Figure 3, observe that point 22 is now projected following the diagonal dotted line onto the point labelled c located at the $\{1, 2\}$-congested boundary. The value $1 + D_{L_{\{1,2\}}} = 1 + \delta^* = 1 + 0.0968 = 1.0968$ indicates the factor by which both inputs of point 22 can be multiplied before $\{1, 2\}$-congestion occurs. In the same way, the $\{1, 2\}$-congestion measurement for point 25 results in a value of 1.8333 meaning that both inputs 4 can be multiplied with a factor $1 + 1.8333 = 2.8333$ before $\{1, 2\}$-congestion is observed. This yields for both inputs a value of $2.8333 \times 4 = 11.3332$ which is the value for both coordinates of the point labelled f in Figure 3.

Note that the measures of $\{1\}$- and $\{2\}$-congestion on the one hand, and $\{1, 2\}$-congestion on the other hand, need not have some monotonic relation to one another. It suffices to compare the results for points 22 and 25: while for point 22 the measure of $\{1, 2\}$-congestion is situated below the measures of $\{1\}$- and $\{2\}$-congestion, for point 25 the measure of $\{1, 2\}$-congestion is lower than the measure of $\{1\}$-congestion, but higher than the measure of $\{2\}$-congestion.

Figure 3 contains two more points, 8 and 14, with their projections on the relevant I-congested boundaries. For these points, the projection points onto the $\{1, 2\}$-congested boundary labelled i and l, respectively, are not located on the CH, but rather at the cone.

Thus, to reiterate once more the complementarity between adapted input directional distance function D_{L_I} and I-congested boundary, this numerical example has shown how $\{1\}$-congestion ($\{2\}$-congestion) is detected by looking into the direction opposite of $g_{\{1\}}$ ($g_{\{2\}}$), while $\{1, 2\}$-congestion is revealed by simultaneously looking into the direction opposite of $g_{\{1,2\}}$.

Technologies revealing S-congestion: a nonconvex perspective

The convexity assumption maintained so far can be replaced by a nonconvex alternative. This cannot be done in a direct approach by merely requiring the activity vector z in the CH technology (9) to be binary.[8] However, note that Proposition 2.1 offers a characterization of any S-disposal input correspondence $L(y)$ as the intersection of all possible subsets $L(y) + K_I$ without any assumption regarding convexity. It is easy to verify that in general the CH input correspondence satisfies the minimal S-disposal assumption with $S = 2^{[m]}$. In particular, this is demonstrated on the numerical example in the "Technologies revealing S-congestion" section. Thus, this CH input correspondence can be obtained via an indirect approach as the intersection of all possible subsets $L^{CH}(y) + K_I$. This indirect approach for generating the CH input correspondence $L^{CH}(y)$ can also be used for introducing the NCH input correspondence $L^{NCH}(y)$. Indeed, one can define the NCH input correspondence as the intersection of all subsets $L^{NCH}(y) + K_I$ with $I \in S = 2^{[m]}$.[9] Applied to

$$
\begin{aligned}
D_{L_I}^{NCH-VRS}(x^o, y^o, g_I) = \quad & \max_{\delta, z_j} \ \delta \\
\text{subject to} \quad & \sum_{j=1}^{J} z_j x_{j,i} \geq x_i^o - \delta g_{I_i} \qquad i \in I \\
& \sum_{j=1}^{J} z_j x_{j,i} \leq x_i^o \qquad\qquad i \in \{1, \ldots, m\} \backslash I \\
& \sum_{j=1}^{J} z_j y_{j,k} \geq y_k^o \qquad\qquad k \in \{1, \ldots, n\} \\
& \sum_{j=1}^{J} z_j = 1 \\
& z_j \in \{0, 1\} \qquad\qquad\quad j \in \{1, \ldots, J\}.
\end{aligned}
\tag{14}
$$

---

[8]This leads to infeasibilities in the corresponding optimization programme. Consequently, the NCH technology cannot be defined in way similar to the convex setting.

[9]To the best of our knowledge, the definition of a NCH technology is not available in the economic literature.
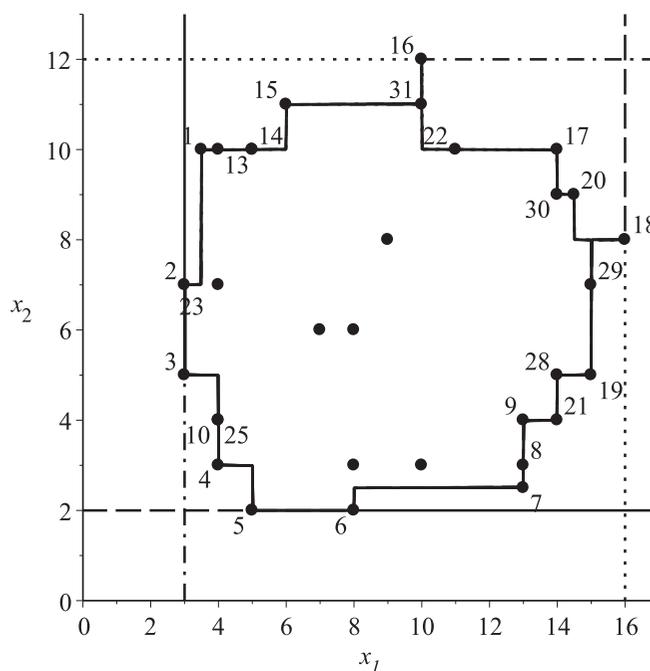
Figure 4. *S*-congested boundaries from a NCH technology.

the case of the NCH technology, the adapted input directional distance function (10) is obtained by the following linear programming model (14):

In particular in the case of the example, this means that $I \in S$ with $S = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}$. The corresponding optimization programmes do not generate any infeasibility, whereby the extension with the cones K_I proves crucial. The boundary of these intersecting subsets yields the NCH level set that can be observed in Figure 4.

First, there is the subset $L^{NCH}(y) + K_\emptyset$ with its boundary represented by the solid line starting vertical towards point 3, then continuing with the nonconvex connection to points 4 and 5 and finally ending in the horizontal line from point 5 onwards. Second, the subset $L^{NCH}(y) + K_{\{1\}}$ has its boundary represented by the dashed line starting horizontal towards point 6, then connecting points 6, 7, 9, 21, 28, 19 and 18 in a nonconvex manner and ending vertically from point 18 upwards. Third, the boundary of subset $L^{NCH}(y) + K_{\{2\}}$ is depicted by the dash-dotted line starting horizontally towards point 16, then connecting points 16, 31, 15, 1 and 2 and then continuing vertically from point 2 downwards. Finally, the boundary of subset $L^{NCH}(y) + K_{\{1,2\}}$ is shown by the dotted line starting horizontally towards point 16, then connecting points

16, 22, 17, 30, 20 and 18 in a nonconvex way and ending vertically from point 18 downwards.

The intersection of these four subsets provides exactly the NCH of which the boundary is shown in solid. Note the very particular shape of this boundary, especially the 'antenna shaped' extensions leading towards points 16 and 18. These one-dimensional extensions of the NCH input correspondence may partially explain the computational difficulties encountered when trying a direct reconstruction approach.

Obviously, the dimension-wise measurements relative to the *I*-congested boundaries presented in the previous subsection could be duplicated. Comparing Figures 3 and 4, we limit ourselves to pointing out that, for instance, points 8, 13 and 22 that are situated in the interior in the convex case are situated on the boundary in the nonconvex case. In general, the NCH thus offers a better fit to the data relative to the CH resulting in lower *I*-congestion values. Focusing on observation 22, one can notice that the measures of $\{1\}$- and $\{2\}$-congestion are identical with respect to CH and NCH on the one hand and that $\{1,2\}$-congestion is smaller with respect to NCH (i.e. 0) compared to CH (i.e. 0.0968) on the other hand.

This succinct display of a NCH input set suffices to illustrate the main point that this new approach is compatible with both the convexity axiom and its

absence. Obviously, more remains to be done to fully explore the impact of convexity and its absence on the relative amounts of MOL- and OP-congestion one can reveal.

V. Empirical illustration

We now perform the congestion computations on the data set provided in Färe, Grosskopf, Logan, et al. (1985) containing the input–output combinations of 32 electric power generating plants.[10] The single output is electricity generated (expressed in 10^6 kWh). The three inputs are capital (in megawatt capacity), fuel (in 10^{10} BTU) and labour (in average annual employees). Färe, Grosskopf, Logan, et al. (1985) report an average congestion inefficiency of 3% using traditional radial efficiency measures. Thus, we are at least certain that some limited amount of congestion is present in this data set.

While the causes of congestion in electricity generation may be a priori unclear, we can add two types of evidence on the issue of congestion in an energy setting. First, several other studies have reported congestion when analysing samples of electricity generating firms. Examples include Färe, Grosskopf and Logan (1987); Färe, Grosskopf and Pasurka (1989); and Zeitsch and Lawrence (1996). Recall that the Färe, Grosskopf and Pasurka (1989) study even reports congestion as the main source of poor performance. Second, there is a literature focusing on energy efficiency (see Boyd 2014; Lin and Wang 2014; Pardo Martnez 2011, among others). In at one of these studies, congestion in the input factor energy has been detected (see Wu, Zhou, and Zhou 2016).

As indicated in the numerical example above, one has to start from a technology allowing for some form of congestion so as to be able to detect it. Examples of such technologies are the CH (9) and the WD (8) technologies. Assuming a CH VRS technology (9), any input correspondence $L(y)$ satisfies minimal S-disposability with

$$S = \{\varnothing, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}.$$

When assuming only WD, $L(y)$ satisfies minimal S^{wd} disposability with $S^{wd} = S \setminus \{\{1,2,3\}\}$ because of the unboundedness. Therefore, meaningful $\{1\}$-, $\{2\}$-, $\{3\}$-, $\{1,2\}$-, $\{1,3\}$-, $\{2,3\}$- and $\{1,2,3\}$-congestion measures can be computed, but the latter only in the case of the CH technology.

When using again a position dependent projection scheme as described before, we obtain the congestion results for all power plants reported in Table 3. The columns represented with a normal font provide the congestion values based on the CH technology, while the columns in italics give the amount one has to add to the CH congestion values to obtain the congestion measures based on a WD technology. The bottom lines contain some basic descriptive statistics. Note that contrary to the numerical example, the output is not set to a single value for all observations: instead, the actual output level of each individual observation is used.

Starting with some basic descriptive statistics at the sample level, the following conclusions can be drawn. First, in the single input dimensions $\{1\}$, $\{2\}$ and $\{3\}$, there are 14, 8 and 10 observations that are situated on the upper bound of technology. In the twin input dimensions $\{1,2\}$, $\{1,3\}$ and $\{2,3\}$, there are 10, 11 and 7 observations on the upper bound. Finally, for the triplet dimensions $\{1,2,3\}$, only one observation is situated on the upper bound. Comparing the CH technology versus the WD technology results, the majority of observations add nothing to the CH results. Second, from the differences reported in italics in Table 3, one can observe that the WD results are all greater than or equal to the CH results, except for the $\{3\}$-congestion measure of the unique unit 2. Thus, the CH seems to provide an almost uniformly better fit to the data compared to the WD results. Third, for all observations not situated on the upper bound of technology, in the single input dimensions $\{1\}$, $\{2\}$ and $\{3\}$, the inputs can on average be proportionally increased between 24% and 31% before reaching the upper bound. In the twin input dimensions $\{1,2\}$, $\{1,3\}$ and $\{2,3\}$, these percentages amount on average between 21% and 54%. For the triplet dimensions $\{1,2,3\}$, this percentage increase even becomes a staggering 292%.

To further illustrate the interpretation of the congestion results, we focus on some specific observations. Starting with unit 14 and the CH-based computations, a $\{1\}$-congestion of 0.5097 is observed meaning that

---

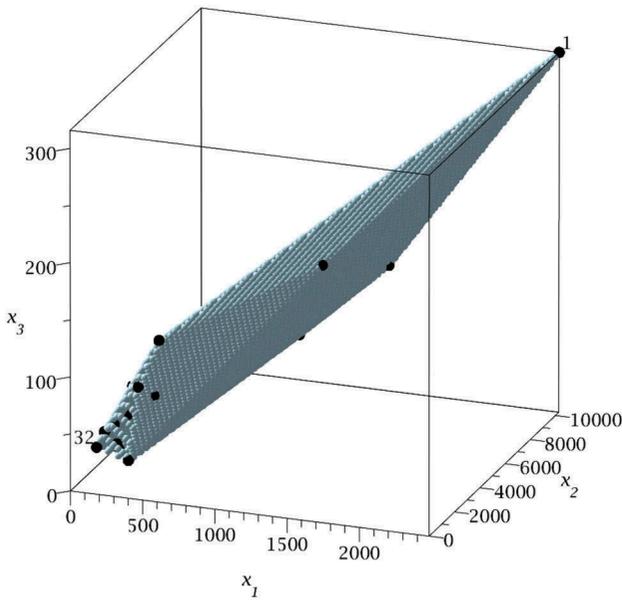[10]See their Table 8–1 on pp. 201–202.

Table 3. Empirical example: *S*-congestion results from a CH technology and proportional projections (normal) and differences with *S*-congestion results from a WD technology (italics).

Unit	{1}	*{1}*	{2}	*{2}*	{3}	*{3}*	{1,2}	*{1,2}*	{1,3}	*{1,3}*	{2,3}	*{2,3}*	{1,2,3}
1	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000
2	0.0000	*0.0000*	0.0000	*0.0000*	0.1967	*−0.0084*	0.0000	*0.0000*	0.1961	*0.0000*	0.0000	*0.0221*	0.2964
3	0.0000	*0.0000*	0.0788	*0.0484*	0.3741	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.2100	*0.2289*	0.5203
4	0.0000	*0.0000*	0.2081	*0.2620*	0.0000	*0.0000*	0.1959	*0.1662*	0.0000	*0.0000*	0.0783	*0.0211*	0.5266
5	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.9652
6	0.0000	*0.0000*	0.0504	*0.0291*	0.7555	*0.0000*	0.0000	*0.0000*	0.1065	*0.0000*	0.2877	*0.3825*	1.0386
7	0.0932	*0.0000*	0.3640	*0.3938*	0.1082	*0.0001*	0.4790	*0.1788*	0.0695	*0.0000*	0.0699	*0.0427*	0.9506
8	0.0000	*0.0000*	0.4481	*0.4435*	0.1307	*0.0295*	0.2894	*0.1469*	0.0000	*0.0000*	0.1807	*0.1279*	0.8810
9	0.3733	*0.0000*	0.0000	*0.0000*	0.4844	*0.0143*	0.0000	*0.0000*	0.8875	*0.0022*	0.0883	*0.0000*	1.3672
10	0.0000	*0.0000*	0.0838	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	0.5461	*0.0382*	2.0013
11	0.0000	*0.0000*	0.0000	*0.0000*	0.3638	*0.0000*	0.0000	*0.0000*	0.4130	*0.0000*	0.0611	*0.0000*	2.0521
12	0.0113	*0.0000*	0.1281	*0.0000*	0.9238	*0.0000*	0.0116	*0.0001*	0.0497	*0.0000*	0.5650	*0.1013*	1.9858
13	0.1117	*0.0000*	0.6478	*0.1598*	0.0476	*0.0052*	0.8913	*0.2048*	0.0362	*0.0019*	0.0293	*0.1049*	1.1351
14	0.5097	*0.0000*	0.3252	*0.0000*	0.1384	*0.0205*	0.5814	*0.1007*	0.4271	*0.0000*	0.0630	*0.0414*	1.9811
15	0.4560	*0.0000*	0.4024	*0.0000*	0.2760	*0.0269*	0.5199	*0.0548*	0.4123	*0.0000*	0.1426	*0.0723*	2.1026
16	0.2923	*0.0000*	0.5069	*0.0000*	0.5097	*0.0219*	0.5171	*0.0114*	0.2887	*0.0000*	0.2844	*0.0696*	2.7094
17	0.5826	*0.0000*	0.1456	*0.0000*	0.5405	*0.0000*	0.2058	*0.0000*	0.5783	*0.0000*	0.1899	*0.0017*	3.0477
18	0.8088	*0.0092*	0.1934	*0.0000*	0.2632	*0.0000*	0.6351	*0.0000*	0.6500	*0.0146*	0.0734	*0.0391*	3.3542
19	0.9604	*0.0080*	0.0719	*0.0000*	0.3011	*0.0000*	0.3640	*0.0000*	0.8681	*0.0058*	0.0681	*0.0039*	3.5833
20	0.1446	*0.0000*	0.8671	*0.0000*	0.6031	*0.0599*	0.7012	*0.0622*	0.1130	*0.0000*	0.4097	*0.1223*	3.0000
21	0.3607	*0.0000*	0.5872	*0.0000*	0.6672	*0.0000*	0.6914	*0.0008*	0.3111	*0.0000*	0.3329	*0.0896*	3.6471
22	0.0000	*0.0000*	0.8531	*0.0000*	0.0000	*0.0000*	1.8752	*0.2399*	0.0000	*0.0000*	0.0000	*0.0000*	1.6333
23	0.7037	*0.0000*	0.2017	*0.0000*	0.1736	*0.0000*	0.8565	*0.0000*	0.6529	*0.0019*	0.0879	*0.0263*	4.7455
24	0.5354	*0.0000*	0.4474	*0.0000*	0.0000	*0.0000*	1.7743	*0.0182*	0.2046	*0.0000*	0.0000	*0.0000*	2.9012
25	0.1127	*0.0071*	1.1479	*0.0000*	0.3258	*0.0000*	0.9355	*0.1955*	0.0022	*0.0172*	0.4249	*0.1684*	3.2133
26	0.1272	*0.0000*	0.9954	*0.0000*	0.0000	*0.0000*	1.6574	*0.2087*	0.0000	*0.0000*	0.1479	*0.0328*	2.8072
27	0.0000	*0.0000*	0.0000	*0.0000*	1.1938	*0.0000*	0.0000	*0.0000*	0.9559	*0.0107*	0.1663	*0.0000*	6.6660
28	0.0000	*0.0000*	0.3405	*0.0000*	0.0000	*0.0000*	1.3376	*0.0000*	0.0000	*0.0000*	0.1223	*0.0324*	5.3200
29	0.9239	*0.0171*	0.1367	*0.0000*	0.8668	*0.0158*	0.3219	*0.0000*	0.7750	*0.0200*	0.2608	*0.0000*	7.4850
30	0.6643	*0.0000*	0.0670	*0.0000*	0.0000	*0.0000*	1.3879	*0.0000*	0.1757	*0.0000*	0.0000	*0.0000*	6.1818
31	0.0000	*0.0000*	0.0000	*0.0000*	0.7319	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	1.7216	*0.0350*	6.6042
32	0.0000	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	1.1192	*0.0000*	0.0000	*0.0000*	0.0000	*0.0000*	8.8750
#0 Obs.	14	28	8	26	10	22	10	18	11	24	7	10	1
Avg.	0.2429	0.0013	0.2906	0.0418	0.3117	0.0058	0.5421	0.0497	0.2554	0.0023	0.2066	0.0564	2.9243
St. Dev.	0.3061	0.0037	0.3176	0.1109	0.3238	0.0132	0.5661	0.0793	0.3061	0.0053	0.3137	0.0805	2.2258
Max.	0.9604	0.0171	1.1479	0.4435	1.1938	0.0599	1.8752	0.2399	0.9559	0.0200	1.7216	0.3825	8.8750

the first input can be increased with 0.5097 times its actual input before a negative effect on the output can be expected while keeping the other inputs at their initial level. Put differently, the first input can be multiplied with a factor $1 + 0.5097 = 1.5097$ before congestion occurs. Analogously, the {1, 2}-congestion of 0.5814 indicates that 0.5814 times the current values of the first two inputs can be added before congestion occurs. Expressed in terms of proportionality, this means that the current values of the first two inputs can be multiplied with a factor 1.5814 before effects on the output can be expected. Note that the {1, 2, 3}-congestion measure of unit 14 equals 1.9811 meaning that all inputs can be multiplied simultaneously with a factor 2.9811 before congestion is expected.

From Table 3, it becomes clear that congestion depends on the direction in which it is measured. Again observing the CH results for unit 14, one observes that congestion is much less of an issue when all inputs are simultaneously increased proportionally rather than when only the second and the third input is increased proportionally (since $1.9811 > 0.0630$). An even more drastic result is observed for unit 32. For this unit, congestion is detected for all three inputs separately (value of 0). However, the {1, 2, 3}-congestion measure equals 8.8750 meaning that all three inputs can be increased simultaneously with 8.8750 times their initial value before congestion is observed. To understand this somewhat strange behaviour, observe the location of unit 32 compared to the input correspondence $L(778.5)$ when assuming a CH technology in Figure 5. Note that the output level of unit 32 equals 778.5. This unit is located at the lower left side near the origin located at boundary of $L(778.5)$. Therefore, it is impossible to stay within $L(778.5)$ when only increasing one of the inputs. However, when all inputs are increased simultaneously, the input combination can move up towards unit 1 while remaining in the input correspondence $L(778.5)$.

produces the highest possible output and is therefore efficient. However, since no information is available regarding possibly even higher output levels (assuming all information is provided in the original data), also congestion is registered for this unit.

Duplicating the same results with the nonconvex NCH technology, again using a position dependent projection scheme one obtains the congestion results reported in Table 4. Two main conclusions emerge. First, the number of observations situated at the upper bound of technology is always higher under nonconvexity compared to convexity. For instance, while for the triplet dimensions $\{1, 2, 3\}$ the difference is 0, for the single input dimension $\{3\}$ and the twin dimensions $\{1, 3\}$ and $\{2, 3\}$, this difference amounts to 20. Second, for all observations not situated on the upper bound of technology, the proportional increases in inputs are always lower or equal under nonconvexity compared to convexity. This simply shows the better fit of the NCH technology compared to the CH technology.

Figure 5. Input correspondence of a CH technology in the inputs of empirical data with minimal output set to 700.

From the congestion computations of unit 1, one observes congestion for all inputs and all possible combinations of inputs. This unit actually

Table 4. Empirical example: *S*-congestion results from a NCH technology and proportional projections.

Unit	{1}	{2}	{3}	{1, 2}	{1, 3}	{2, 3}	{1, 2, 3}
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2964
3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5203
4	0.0000	0.0000	0.0000	0.0295	0.0000	0.0000	0.5266
5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9652
6	0.0000	0.0374	0.0000	0.0000	0.0000	0.0000	1.0386
7	0.0000	0.1190	0.0000	0.3962	0.0000	0.0000	0.9506
8	0.0000	0.0812	0.0000	0.2374	0.0000	0.0000	0.8810
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	1.3672
10	0.0000	0.0000	0.0000	0.0000	0.0000	0.4925	2.0013
11	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	2.0521
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.4286	1.9858
13	0.0000	0.5484	0.0000	0.3956	0.0000	0.0000	1.1351
14	0.0000	0.0011	0.0000	0.4988	0.0000	0.0000	1.9811
15	0.0000	0.0166	0.0000	0.0391	0.0000	0.0154	2.1026
16	0.0000	0.2154	0.0000	0.2424	0.0000	0.0000	2.7094
17	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	3.0477
18	0.0000	0.0000	0.0000	0.4583	0.0000	0.0000	3.3542
19	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	3.5833
20	0.0000	0.6008	0.0000	0.4532	0.0000	0.0000	3.0000
21	0.0000	0.2466	0.0000	0.6518	0.0000	0.0441	3.6471
22	0.0000	0.3977	0.0000	1.5077	0.0000	0.0000	1.6333
23	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.7455
24	0.0000	0.0472	0.0000	1.2814	0.0000	0.0000	2.9012
25	0.0000	0.6548	0.0800	0.7324	0.0000	0.0000	3.2133
26	0.0000	0.7133	0.0000	1.1710	0.0000	0.0000	2.8072
27	0.0000	0.0000	0.0000	0.0000	0.6337	0.0000	6.6660
28	0.0000	0.0000	0.0000	0.0532	0.0000	0.0000	5.3200
29	0.0000	0.0000	0.5625	0.0226	0.0000	0.0000	7.4850
30	0.0000	0.0000	0.0000	0.3665	0.0000	0.0000	6.1818
31	0.0000	0.0000	0.0000	0.0000	0.0000	0.8469	6.6042
32	0.0000	0.0000	0.0000	0.7338	0.0000	0.0000	8.8750
#0 Obs.	32	19	30	14	31	27	1
Avg.	0.0000	0.1150	0.0201	0.2897	0.0198	0.0571	2.9243
St. Dev.	0.0000	0.2141	0.0984	0.4082	0.1103	0.1804	2.2258
Max.	0.0000	0.7133	0.5625	1.5077	0.6337	0.8469	8.8750

VI. Conclusions

This article has started by describing the axiomatic production literature revealing and measuring a variety of congestion concepts. While traditionally only MOL forms of congestion could be revealed, the article of Briec, Kerstens and Van de Woestyne (2016) has laid the foundations to measure also OP congestion in a multiple output setting, eventually even dispensing with the axiom of convexity. This contribution has offered a complementary empirical perspective to these theoretical developments.

Three empirical goals have been achieved. First, a rather detailed overview of the empirical literature has revealed that the amounts and incidence of congestion reported can be surprisingly high at the level of the sample, some subsample or for specific individual observations. Second, it has been illustrated that the traditional radial way of measuring congestion yields lower amounts of congestion compared to a measurement scheme per input dimension. Third, moving beyond MOL congestion, we have offered the first empirical evidence that forms of OP congestion may well matter in an empirical setting.

The main theoretical limitation of the analysis is that generalizations to congestion phenomena in the outputs space or to the input and output space are still needed. Note that the use of the directional distance function allows for a relatively straightforward extension of our proposals to define congestion in the output space or in the input–output space. Other theoretical limitations are listed in Briec, Kerstens and Van de Woestyne (2016).

From an empirical point of view, it is worthwhile underscoring two main limitations. First, this empirical article has been limited to a primal approach focusing on technologies capable to reveal several forms of congestion. For reasons of space, the dual approach focussing on the cost function in a congested setting as developed in Briec, Kerstens and Van de Woestyne (2016) has been completely ignored. This also implies that the whole literature on generating appropriate shadow prices for bad outputs using technologies modelling the joint production of good and bad outputs (see the recent survey in Dakpo, Jeanneaux, and Latruffe 2016) has equally been ignored.

Second, we have not managed to establish a link with a related literature in operational research focusing on alternatives to the WD approach (see Cooper, Thompson, and Thrall 1996 for the seminal alternative proposal and Kao 2010 for a recent overview). As summarized by Kao (2010), these different approaches are distinct in terms of their focus on input space, output space or input–output space. There is an urgent need to extend the S-disposal approach into the output space and into the input–output space in order to be able to measure S-congestion in the outputs and in the input–output space. Only thereafter, meaningful comparisons between the WD, S-disposal and operational research approaches in terms of the incidence and amounts of congestion can eventually be established. This necessitates substantial future work.

Acknowledgements

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Alam, I. M. S., and R. C. Sickles. 2000. "Time Series Analysis of Deregulatory Dynamics and Technical Efficiency: The Case of the US Airline Industry." *International Economic Review* 41 (1): 203–218.

Allais, M. 1977. "Theories of General Economic Equilibrium and Maximum Efficiency." In *Equilibrium and Disequilibrium in Economic Theory*, edited by G. Schwödiauer, 129–201. Dordrecht: Reidel.

Atkinson, S. E., and M. G. Tsionas. 2016. "Directional Distance Functions: Optimal Endogenous Directions." *Journal of Econometrics* 190 (2): 301–314.

Banker, R. D., A. Charnes, and W. W. Cooper. 1984. "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science* 30 (9): 1078–1092.

Barnett, W. A. 2002. "Tastes and Technology: Curvature Is Not Sufficient for Regularity." *Journal of Econometrics* 108 (1): 199–202.

Boyd, G. A. 2014. "Estimating the Changes in the Distribution of Energy Efficiency in the U.S. Automobile Assembly Industry." *Energy Economics* 42: 81–87.

Briec, W., K. Kerstens, and I. Van de Woestyne. 2016. "Congestion in Production Correspondences." *Journal of Economics* 119 (1): 65–90.

Byrnes, P., and R. Färe. 1987. "Surface Mining of Coal: Efficiency of US Interior Mines." *Applied Economics* 19 (12): 1665–1673.

Byrnes, P., R. Färe, S. Grosskopf, and C. A. K. Lovell. 1988. "The Effect of Unions on Productivity: U.S. Surface Mining of Coal." *Management Science* 34 (9): 1037–1053.

Byrnes, P., R. Färe, S. Grosskopf, and S. Kraft. 1987. "Technical Efficiency and Size: The Case of Illinois Grain Farms." *European Review of Agricultural Economics* 14 (4): 367–381.

Charnes, A., W. W. Cooper, B. Golany, L. M. Seiford, and J. Stutz. 1985. "Foundations of Data Envelopment Analysis for Pareto-Koopmans Efficient Empirical Production Functions." *Journal of Econometrics* 30 (1–2): 91–107.

Cooper, W. W., R. Thompson, and R. M. Thrall. 1996. "Chapter 1 Introduction: Extensions and New Developments in DEA." *Annals of Operations Research* 66 (1): 1–45.

Dakpo, K. H., P. Jeanneaux, and L. Latruffe. 2016. "Modelling Pollution-Generating Technologies in Performance Benchmarking: Recent Developments, Limits and Future Prospects in the Nonparametric Framework." *European Journal of Operational Research* 250 (2): 347–359.

Daraio, C., and L. Simar. 2016. "Efficiency and Benchmarking with Directional Distances: A Data-Driven Approach." *Journal of the Operational Research Society* 67 (7): 928–944.

Dillon, J. L., and J. R. Anderson. 1990. *The Analysis of Response in Crop and Livestock Production*. 3rd ed. Oxford: Pergamon.

Färe, R., S. Grosskopf, and C. A. K. Lovell. 1983. "The Structure of Technical Efficiency." *The Scandinavian Journal of Economics* 85 (2): 181–190.

Färe, R., S. Grosskopf, and C. A. K. Lovell. 1985. *The Measurement of Efficiency of Production*. Boston, MA: Kluwer.

Färe, R., S. Grosskopf, and C. A. K. Lovell. 1994. *Production Frontiers*. Cambridge: Cambridge University Press.

Färe, R., S. Grosskopf, and J. Logan. 1987. "The Comparative Efficiency of Western Coal-Fired Steam-Electric Generating Plants: 1977-1979." *Engineering Costs and Production Economics* 11 (1): 21–30.

Färe, R., S. Grosskopf, J. Logan, and C. A. K. Lovell. 1985. "Measuring Efficiency in Production: With an Application to Electric Utilities." In *Managerial Issues in Productivity Analysis*, edited by A. Dogramaci and N. Adam, 185–214. Boston, MA: Kluwer.

Färe, R., S. Grosskopf, and C. A. Pasurka. 1989. "The Effect of Environmental Regulations on the Efficiency of Electric Utilities: 1969 versus 1975." *Applied Economics* 21 (2): 225–235.

Färe, R., and L. Svensson. 1980. "Congestion of Production Factors." *Econometrica* 48 (7): 1745–1753.

Farrell, M. 1957. "The Measurement of Productive Efficiency." *Journal of the Royal Statistical Society. Series A (General)* 120 (3): 253–281.

Field, K. 1990. "Production Efficiency of British Building Societies." *Applied Economics* 22 (3): 415–426.

Førsund, F. R., and L. Hjalmarsson. 1974. "On the Measurement of Productive Efficiency." *The Swedish Journal of Economics* 76 (2): 141–154.

Hackman, S. T. 2008. *Production Economics: Integrating the Microeconomic and Engineering Perspectives*. Berlin: Springer.

Johnson, G. M., and W. W. Happ. 1977. "Digital Simulation for Detecting Congestion in Hospital Facilities." In *Winter Simulation Conference Proceedings*, 849–853. Institute of Electrical and Electronics Engineers (IEEE).

Jones, J. B. 2001. *Laboratory Guide for Conducting Soil Tests and Plant Analysis*. Boca Raton, FL: CRC.

Kao, C. 2010. "Congestion Measurement and Elimination under the Framework of Data Envelopment Analysis." *International Journal of Production Economics* 123 (2): 257–265.

Kumar, S., and R. R. Russell. 2002. "Technological Change, Technological Catch-Up, and Capital Deepening: Relative Contributions to Growth and Convergence." *American Economic Review* 92 (3): 527–548.

Lin, B., and X. Wang. 2014. "Exploring Energy Efficiency in China's Iron and Steel Industry: A Stochastic Frontier Approach." *Energy Policy* 72: 87–96.

Luenberger, D. G. 1995. *Microeconomic Theory*. Boston, MA: McGraw-Hill.

Munson, R. D. 1998. "Principles of Plant Analysis." In *Handbook of Reference Methods for Plant Analysis*, edited by Y. P. Kalra, 1–24. Boca Raton, FL: CRC.

Olesen, O. B., and N. C. Petersen. 2016. "Stochastic Data Envelopment Analysis—A Review." *European Journal of Operational Research* 251 (1): 2–21.

Pardo Martínez, C. I. 2011. "Energy Efficiency Development in German and Colombian Non-Energy-Intensive Sectors: A Non-Parametric Analysis." *Energy Efficiency* 4 (1): 115–131.

Paris, Q. 2008. "Law of the Minimum." In *Encyclopedia of Soil Science*, edited by W. Chesworth, 431–437. New York: Springer.

Patrick, J. 2011. "Access to Long-Term Care: The True Cause of Hospital Congestion?" *Production and Operations Management* 20 (3): 347–358.

Peyrache, A., and C. Daraio. 2012. "Empirical Tools to Assess the Sensitivity of Directional Distance Functions to Direction Selection." *Applied Economics* 44 (8): 933–943.

Ray, S. C. 2004. *Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research*. Cambridge: Cambridge University Press.

Sauer, J. 2006. "Economic Theory and Econometric Practice: Parametric Efficiency Analysis." *Empirical Economics* 31 (4): 1061–1087.

Seitz, W. D. 1970. "The Measurement of Efficiency Relative to a Frontier Production Function." *American Journal of Agricultural Economics* 52 (4): 505–511.

Wu, F., P. Zhou, and D. Q. Zhou. 2016. "Does There Exist Energy Congestion? Empirical Evidence from Chinese Industrial Sectors." *Energy Efficiency* 9 (2): 371–384.

Zeitsch, J., and D. Lawrence. 1996. "Decomposing Economic Inefficiency in Base Load Power Plants." *Journal of Productivity Analysis* 7 (4): 359–378.

Zhengfei, G., and A. Oude Lansink. 2003. "Input Disposability and Efficiency in Dutch Arable Farming." *Journal of Agricultural Economics* 54 (3): 467–478.