European Journal of Operational Research 247 (2015) 996–1002

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Cost constrained industry inefficiency

Antonio Peyrache*

Centre for Efficiency and Productivity Analysis (CEPA), School of Economics, The University of Queensland, Australia

A R T I C L E I N F O

Article history:
Received 8 May 2014
Accepted 29 June 2015
Available online 9 July 2015

Keywords:
Data envelopment analysis (DEA)
Industry inefficiency
Indirect directional distance function
Productivity
optimal configuration

A B S T R A C T

In this paper a definition of industry inefficiency in cost constrained production environments is introduced. This definition uses the indirect directional distance function and quantifies the inefficiency of the industry in terms of the overall output loss, given the industry cost budget. The industry inefficiency indicator is then decomposed into sources components: reallocation inefficiency arising from sub-optimal configuration of the industry; firm inefficiency arising from a failure to select optimal input quantities (given the prevalent inputs prices); firm inefficiency due to lack of best practices. The method is illustrated using data on Ontario electricity distributors. These data show that lack of best practices is only a minor component of the overall inefficiency of the industry (less than 10 percent), with reallocation inefficiency accounting for more than 75 percent of the overall inefficiency of the system. An analysis based on counter-factual input prices is conducted in order to illustrate how the model can be used to estimate the effects of a change in the regulation regime.

Introduction

In this paper a definition of industry inefficiency in cost constrained production environments is introduced. The idea behind cost constrained production is that firms or decision making units (DMU) are allocated a certain cost and are supposed to produce as much output as possible with this given cost (this implies optimizing when selecting input quantities). It is assumed that data on the inputs, the prices of the inputs and the outputs produced by each firm in a given industry are available. With such data, it is possible to determine the overall cost budget of the industry by looking at the inputs used and the input prices faced by each firm. The problem that the central planner (or a market) now faces is how to allocate this overall budget across the different production units (given a certain number of constraints) in order to maximize the overall output. This problem corresponds to the implicit determination of the optimal structure of the industry via the determination of the optimal number of firms that should populate the industry and the optimal allocation of resources across these firms. Once this optimization problem is solved, one is able to quantify inefficiency in terms of the directional distance function (DDF), where inefficiency is measured in terms of the overall output loss due to different sources: *(i)* the inefficiency of the firms actually operating in the industry (lack of best practice) and *(ii)* the inefficiency arising from a sub-optimal configuration of the industry.

To the best of our knowledge the first paper to address this issue explicitly in a linear programming framework is due to Ray and Hu (1997). This contribution introduced the basic model in a primal context, where only input and output quantities are observed. Later on Lozano and Villa (2004) introduced the centralized resource allocation model which uses the same idea (and it is in fact a special case of Ray and Hu (1997) where the number of firms is fixed to the observed one). Following these attempts a literature developed to accommodate alternative empirical settings (see Aparicio & Pastor, 2012; Aparicio, Pastor, & Ray, 2013; Asmild, Paradi, & Pastor, 2009; 2012; Fang, 2013; Fang & Zhang, 2008; Giménez-García, Martínez-Parra, & Buffa, 2007; Lotfi, Noora, Johanshahloo, Gerami, & Mozaffari, 2010; Lozano & Villa, 2004; 2005; Lozano, Villa, & Adenso-Díaz, 2004; Lozano, Villa, & Braennlund, 2009; Lozano, Villa, & Canca, 2011; Mar-Molinero, Prior, Segovia, & Portillo, 2012; Ray, 2007; Ray & Mukherjee, 1998). Ray, Chen, and Mukherjee (2008) extended the industry efficiency model using a cost function approach with input prices varying across locations.

The method introduced in this paper is illustrated using data on Ontario electricity distributors. The data show that lack of best practices at the firm level is only a minor component of the overall inefficiency of the industry (less than 10 percent). The bulk of the industry inefficiency is accounted for by severe deviations from the optimal configuration. In this empirical part a counter-factual analysis is also provided.

The rest of the paper is organized as follows. Section 1 presents the definition of firm technology and inefficiency. Section 2 extends these

* Tel.: +61 (0) 420214263.
E-mail address: a.peyrache@uq.edu.au

http://dx.doi.org/10.1016/j.ejor.2015.06.067
0377-2217/

notions to the industry level. Section 3 is dedicated to the empirical illustration. Finally Section 4 concludes.

1. Firm technology and inefficiency

Consider an industry where $\mathbf{x} \in R_+^N$ inputs produce $\mathbf{y} \in R_+^M$ outputs. It is assumed that data on inputs and outputs are available for a number K of firms ($k = 1, \ldots, K$). The data can be collected into two matrices: an input matrix $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_K]'$ and an output matrix $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_K]'$, where the observations relative to each firm are collected on the rows of these matrices. The dataset is the collection of these matrices:

$$(\mathbf{X}, \mathbf{Y}) \tag{1}$$

The data generated firm technology set (production set, or production possibilities set) is given by the picece-wise linear envelop of the observations available (Banker, Chang, & Natarajan, 2005; Banker, Charnes, & Cooper, 1984):

$$\Psi = \left\{ (\mathbf{x}, \mathbf{y}) : \lambda \mathbf{X} \le \mathbf{x}', \ \lambda \mathbf{Y} \ge \mathbf{y}', \ \sum_k \lambda_k = 1, \ \lambda \ge \mathbf{0} \right\} \tag{2}$$

This production set is built under the assumption that the technology satisfies convexity and free disposability of inputs and outputs. The constraint on the intensity vector $\sum \lambda_k = 1$ means that the technology allows for variable returns to scale (VRS). An equivalent representation of the technology is given by the output sets, which are the collection of all producible outputs given a certain input vector $P(\mathbf{x}) = \{\mathbf{y} : (\mathbf{x}, \mathbf{y}) \in \Psi\}$ and can be represented in functional form via the directional output distance function (DDF) (see Chambers, Chung, & Färe, 1996; 1998):

$$TE = D(\mathbf{x}, \mathbf{y}; \mathbf{g}_y) = \sup_\beta \{\beta : (\mathbf{y} + \beta \mathbf{g}_y) \in P(\mathbf{x})\} \tag{3}$$

The DDF provides a measure of technical efficiency (TE) or, more precisely, of technical inefficiency because it represents the total loss in output due to inefficient use of the inputs available to the firm (with respect to the benchmark represented by technology (2)). This measure of inefficiency is expressed in terms of the *numeraire* $\mathbf{g}_y$ which is therefore assumed to be fixed across firms and time periods (comparisons of technical inefficiency using different *numeraires* would be equivalent to comparing apples with oranges). Function (3) can also be interpreted as a shortage function, inasmuch it is a measure of total output loss with respect to a potential output benchmark.

It is now interesting to consider an alternative representation of the technology (2) which assumes availability of information on input prices in the form of a row vector $\mathbf{w} \in \mathbb{R}_+^N$. In this case, one may think of the production possibilities as the collection of all the possible output vectors which are feasible when the cost budget is set at level C. This gives rise to the indirect output sets $IP(\mathbf{w}/C) = \{\mathbf{y} : (\mathbf{x}, \mathbf{y}) \in \Psi \ , \ \mathbf{wx} \le C\}$ and their functional representation via the indirect directional output distance function (IDDF):

$$CE = ID(\mathbf{w}/C, \mathbf{y}; \mathbf{g}_y) = \sup_\beta \{\beta : (\mathbf{y} + \beta \mathbf{g}_y) \in IP(\mathbf{w}/C)\} \tag{4}$$

This alternative measure of inefficiency involves the use of input prices and it measures how much output production could be expanded, given that the overall cost the firm is facing is given. Since the constraint here is the overall cost rather than a specific input vector, this type of inefficiency has also been called cost constrained inefficiency (see Färe & Grosskopf, 1994; Grosskopf, Hayes, Taylor, & Weber, 1997). It should be emphasized that, though the cost is constraining production, the inefficiency measure is defined on the output side for a given *numeraire* $\mathbf{g}_y$. This means that the quantities in Eqs. (3) and (4) share the same underlying *numeraire* and they can be compared. Invoking the duality theorem proved in Färe and Primont (2006), it holds that $ID(\mathbf{w}/C, \mathbf{y}; \mathbf{g}_y) \ge D(\mathbf{x}, \mathbf{y}; \mathbf{g}_y)$, where the difference between the two indicators is an allocative efficiency component, interpreted as the loss in output due to the choice of a non-optimal input mix:

$$AE = ID(\mathbf{w}/C, \mathbf{y}; \mathbf{g}_y) - D(\mathbf{x}, \mathbf{y}; \mathbf{g}_y) \tag{5}$$

The allocative inefficiency definition embedded in Eq. (5) determines the quantity of output which is lost because the firm fails to choose the optimal input mix given the prevailing input prices. Given the duality between the direct and the indirect DDF, it is thus possible to decompose the firm level cost constrained inefficiency (CE) into the two components defined in Eqs (5) and (3):

$$CE = TE + AE \tag{6}$$

The left hand side of this equation is a measure of the overall loss in output for a firm operating at the specified cost level. The first component on the right hand side attributes part of this inefficiency to a less than optimal use of the given input vector; while the second component is a measure of loss in output attributable to the firm failing to choose an optimal input vector.

2. Industry technology and inefficiency

The previous section introduced the main representation of the firm technology set and the firm inefficiency. The purpose of this section is to extend these notions to the industry level. For the purposes of this study the industry is defined in terms of inputs and outputs homogeneity, so that all the firms operating in the industry uses the same set of inputs to produce the same set of outputs. It is assumed that any number of firms can operate in the industry (entry and exit of firms is allowed) and all the firms in the industry (already operating in it or potentially entrant) face the same technology set Ψ defined in Eq. (2); in other words all the firms in the industry face the same production trade-offs. Under these assumptions the industry technology set is defined as (see Peyrache, 2013):

$$\Psi_I = \cup_{S=1}^{+\infty} \left(\sum_{s=1}^{S} \Psi \right) \tag{7}$$

The summation in parentheses is a special case of the aggregation discussed in Li and Ng (1995) and Zelenyuk (2006)[1]. It should be noted that, though the firm technology set is convex, the industry technology set may show some non-convexity. This non-convexity at the industry level arises because of the indivisibility of the firm: only an integer number of firms can operate in the industry (of course this non-convexity becomes less important as the number of firms grows large). The set defined in (7) can be also written as follow:

$$\Psi_I = \left\{ (\mathbf{x}, \mathbf{y}) : \lambda \mathbf{X} \le \mathbf{x}', \ \lambda \mathbf{Y} \ge \mathbf{y}', \ \sum_k \lambda_k = S, \ \lambda \ge \mathbf{0}, \ S \in \mathbb{N} \right\} \tag{8}$$

Contrary to Definition (2), the intensity vector is now constrained to sum up to the the integer number S which represents the number of firms operating (actually or potentially) in the industry. Interestingly enough, this set collapses to the firm technology set (2) when $S = 1$. The industry technology returns all the possible input and output combinations which are feasible at the industry level and it is an enlargement of the firm production set. Since the industry as a whole is operating within a given cost budget, it is interesting to describe the industry technology set in terms of the associated indirect output sets which represent all the output combinations the

[1] The reader should note that if the production set is non-convex, then $\sum_{s=1}^{S} \Psi \ne S\Psi$. The proof of this is in Li and Ng (1995). In the case of a convex production set the equality holds.