Asia-Pacific Journal of Operational Research
Vol. 35, No. 6 (2018) 1850042 (22 pages)

DOI: 10.1142/S0217595918500422

Evaluating the Efficiencies of Academic Research Groups: A Problem of Shared Outputs

Sonia Valeria Avilés-Sacoto

Universidad San Francisco de Quito USFQ, Campus Cumbayá
Diego de Robles s/n, Quito, Pichincha 170157, Ecuador

Tecnologico de Monterrey, Campus Monterrey
Ave. Eugenio Garza Sada 2501, Monterrey N.L. 64849, México
svaviless@usfq.edu.ec

Wade D. Cook*

Schulich School of Business, York University
Toronto, Ontario, Canada M3J 1P3
wcook@schulich.yorku.ca

David Güemes-Castorena

Tecnologico de Monterrey, Campus Monterrey
Ave. Eugenio Garza Sada 2501, Monterrey N.L. 64849, México
guemes@itesm.mx

Francisco Benita

Singapore University of Technology and Design
8 Somapah Rd, Singapore 487372, Singapore
francisco_benita@sutd.edu.sg

Hector Ceballos

Tecnologico de Monterrey, Campus Monterrey
Ave. Eugenio Garza Sada 2501, Monterrey N.L. 64849, México
ceballos@itesm.mx

Joe Zhu

College of Auditing and Evaluation, Nanjing Audit University
Nanjing, Jiangsu Province 211815, P. R. China

Foisie Business School, Worcester Polytechnic Institute
Worcester, MA 01609, USA
jzhu@wpi.edu

Received 10 December 2017
Revised 11 June 2018
Accepted 19 September 2018
Published 20 November 2018

*Corresponding author.

Data envelopment analysis (DEA) is a methodology for evaluating the relative efficiencies of a set of decision-making units (DMUs), based on their multiple inputs and outputs. The original model is based on the assumption that DMUs operate independently of one another. However, this assumption may not apply in some situations, as in the case we present in this paper, in which DMUs can work together to produce joint outputs. What makes it more interesting is the situation in which this characteristic of sharing outputs among some DMUs differs from one DMU to another; this makes it more challenging to determine independent efficiency scores that cater for this phenomenon. To address this, the current paper presents a methodology for measuring efficiency in situations in which DMUs share outputs with other units. We examine the case of a set of research groups in a Mexican university. For this study, the inputs used are professors belonging to various groups, and outputs are the published journal articles, some of which are produced completely within a group, whereas others arise from collaboration with professors from other research groups. Jointly published articles form a link connecting the groups.

Keywords: DEA; shared outputs; research groups; cooperation; DMU dependence.

1. Introduction

Data envelopment analysis (DEA) was developed by Charnes *et al.* (1978), as a tool for evaluating the relative efficiencies of a set of decision-making units (DMUs) in the presence of multiple inputs and multiple outputs. Over time, extensive research has emerged in terms of both methodology and applications. The traditional setting involving DEA and efficiency measurement is one where each DMU uses its *own* multiple inputs to generate its *own* multiple outputs. There is, in most situations found in the literature, an understood independence among the set of DMUs under evaluation.

The current study involves a set of DMUs with each creating two types of outputs: one type we refer to as "in-house" or internal outputs and the other as "shared" or external outputs. To exemplify, let us consider a given DMU, say DMU_1, and assume that it produces a given output by itself (independent of any collaboration with other DMUs) and, in addition, produces, in a collaborative manner, outputs which it shares with DMU_2 and DMU_5. Correspondingly, DMU_2 produces its own internal outputs, as well as outputs produced collaboratively with DMU_1, DMU_4 and DMU_5, and similarly with the rest of the DMUs. The conventional DEA model as per Charnes *et al.* (1978) is not directly designed to cater for this type of dependence among the DMUs.

In the current paper, we examine the problem of measuring the efficiency of each member of a set of research groups within a university; a setting that gives rise to the collaborative phenomenon discussed earlier. Typical of most academic settings, the "business" of a research group or center is to conduct and publish research, generally in the form of peer-reviewed papers, books, etc. This paper specifically focusses attention on publications in the form of journal articles. In this environment, it is common for a researcher to collaborate with other colleagues, both internal and external to that person's department, faculty and/or institution. In particular, it can happen that not only is there research conducted purely internal to a research group, but as well there can be collaboration among groups. Such collaboration can happen for many reasons: building on the strengths of other researchers with common interests, enhancing the likelihood of publication, enhancing the obtaining of research funding, benefiting from interdisciplinary research, etc. Moreover, it is often the case that the weight or importance attached to papers done jointly between groups can be different from the weight attached to other collaborations.

In the sections mentioned subsequently, we develop and apply a DEA-based methodology that can be used to measure the efficiencies of a set of DMUs while taking into account the above-discussed dependence among those DMUs. Section 2 reviews some of the literature relating to the measurement of efficiency of research groups. Section 3 develops the methodology to be used to evaluate the performance of the DMUs with internal and external outputs, specifically accounting for dependence among DMUs. Section 4 applies the new methodology to a set of research groups. Conclusions and recommendations follow in Sec. 5.

2. Related Work

The problem of modeling the efficiency of a set of research centers, as described earlier, bears some relation to at least two other streams of work, namely, that involving efficiency of academic institutions and departments and that pertaining to the issue of shared inputs and outputs. We briefly discuss these two areas.

2.1. *Research groups and efficiency in academic departments*

As per Hollingshead and Poole (2012), groups can be defined as "... collectives with more than three people whose members share a common goal or purpose, have some degree of interdependence, interact with one another, and generally perceive themselves as a group." In this context, research groups are made up of professors–researchers–scientists whose activities are focussed on developing models, techniques, tools, methodologies and ideas that help to contribute to the advancement and progress of the science. The study of López-Yáñez and Altopiedi (2015) states that research groups are the most important piece of the scientific system in every society and the most delicate one, as they are especially sensitive to the changes and contradictions around them. As a result, research groups establish links with other groups, thereby contributing to professional development. It is here where an evaluation culture should be applied, to help identify the strengths of each group, the strengths of the collaboration, and also boost additional links and networking among the groups.

Research productivity in the context of efficiency of academic institutions has been widely studied since the 1950s, and currently it is a topic of intense debate among politicians, professors and other stakeholders. Table 1 displays features considered in different studies of academic institutions using DEA. Empirical analyses of the performance of educational institutions have been done by means of DEA, Malmquist indexes, bootstrapping, robust frontiers, meta-frontiers and stochastic frontier analysis. We refer the reader to De Witte and López-Torres (2015) and the

Table 1. Features considered in the literature in education.

Study	Institution	Inputs	Outputs	DEA approach
Giménez and Martínez-Parra (2006)	Autonomous University of Barcelona (Spain)	No. of professors and researchers Operational expenditure Teaching and research expenditure	No. of publications Teaching load Quality teaching	Cost efficiency
Tauer *et al.* (2007)	Cornell University (USA)	No. of students Research indicators Leadership and extension measures	No. of credit hours No. of publications No. of extension hours	CCR
Kao and Hung (2008)	National Cheng Kung University (Taiwan)	No. of personal Operating expenses Floor space	No. of credit hours No. of publications External grants	BCC
Kounetas *et al.* (2011)	A representative Greek university	Total expenditures No. of academic staff No. of students	No. of publications	CCR, Tobit
Agasisti *et al.* (2012)	69 academic departments in Lombardy (Italy)	Size of laboratories No. of high qualified staff	No. of publications No. of citations Research funded	BCC
This work	Tecnologico de Monterrey (Mexico)	No. of professors No. of SNI professors	No. of in-house publications No. of external publications	SBM CCR

references therein, for an extensive overview of the literature on efficiency in education. That survey summarizes the use of certain inputs, outputs and contextual variables, including numbers of publications, and as well presents a review of the literature on efficiency in education, covering articles such as the work of Bessent *et al.* (1983), that applied frontier efficiency measurement techniques.

The quantity (and quality) of research publications is a widely used output variable in a number of studies in the education sector, and the number of professors in the institutions being evaluated is a standard input variable. As well, different versions of DEA have been considered, namely, CRS and VRS models or some of their extensions (multi-criteria models, cost efficiency and Tobit regression). In the current paper only, the numbers of professors (at two levels, as explained subsequently) are considered as inputs and the numbers of journal publications (again at two levels) as outputs. The methodological approach of this work takes the form of a slacks-based DEA model as per Tone and Tsutsui (2009) to measure efficiencies.

2.2. *Efficiency models with shared inputs and outputs*

Beasley (1995) introduced one of the first examples of a "shared factor" model in DEA, where only shared inputs are present. It is an interesting fact that in his study, academic departments were the DMUs under consideration. That author focussed on similar departments of different universities which share one of their inputs (total expenditure) and produce independent outputs (number of graduate students.)

Another area where shared inputs and outputs often arise is in *parallel subsystems*. An example of this phenomenon appears in the earlier work of Cook and Hababou (2001) involving bank branch efficiency. Those authors look at the case of two types of branch activities, namely, sales and service components (two subsystems), where on the input side, certain branch staff work on both the sales and service subsystems; such staff would thus constitute *shared* resources. The manner in which such shared resources are generally dealt with in that situation is to *split* them between the two subsystems. Bian *et al.* (2015), Kao (2009) and others present approaches for DMUs that have parallel subsystems. Again, models with inputs and outputs that are shared among such parallel subsystems employ splitting variables to allocate portions of such resources to the parallel units. Recent studies dealing with shared inputs/outputs are shown in Table 2.

While the problem setting in the current research study concerns what might be thought of as a case involving parallel subsystems, splitting of shared factors is not done, nor is it appropriate. To explain, and going back to first principles, in studies involving DEA, it is generally presumed that the DMUs under evaluation use their own inputs to produce their own outputs; moreover, the DMUs are assumed to operate independently of one another. In such a situation, efficiency scores and accompanying projections to the best practice frontier should be a realistic portrayal of the efficiency standing of each DMU; this should mean that targets or projections are realistically achievable.

In is important to recognize that our study herein is very different from previous work, in that it is intended to address processes whereby different departments

Table 2. Related studies dealing with shared input/output models.

Study	Shared inputs	Shared outputs	Application	DEA-type	Formulation techniques
Cook *et al.* (2000)	x		Bank branches	CCR	Aggregate performance measure
Tsai and Molinero (2009)	x	x	UK Health Service	BCC	Weighted objective function
Amirteimoori and Nashtaei (2006)	x	x	Bank branches	CCR	Aggregate performance measure
Chen *et al.* (2010)	x		Banking industry	CCR/BBC two-stage	Weighted average of efficiency scores of stages 1 and 2
Cook and Zhu (2011)	x	x	Bank branches	BBC	Multiple variable proportionality (subgroup-specific behavior)

or research groups *share* outputs, such as the numbers of publications, with other departments. Previous research in which such publications of an academic department j_o were considered as an output *did not* take account of the fact that some of those same publications were done jointly with other institutions (e.g., j_1). As such, the efficiency evaluation of j_o, in the form of recommended enhancements to that institution's publications, is made, assuming that the same publications of j_1 do not change. In such cases, the outputs involved can be considered as being *shared* among the DMUs, and those outputs contribute to the efficiency score of each DMU involved. Our research herein, however, takes this aspect of shared outputs into account. Again, to be clear, research publications considered herein are *not split* across the DMUs that share those publications, but rather are claimed, in their entirety, by all sharing DMUs.

The conventional DEA model fails to provide a proper framework within which to evaluate efficiency in the presence of the *type* of sharing referred to herein. An earlier study by Chen *et al.* (2010) did evaluate a set of DMUs that shared inputs to produce multiple outputs, but in the context of a two-stage process. The characteristic there is that the inputs to the first stage are shared by the two stages. It is important to point out that the sharing phenomenon in that context again involves *splitting* the inputs between the two roles as inputs to both stage 1 and stage 2.

3. Methodology

3.1. *Selecting inputs and outputs*

In a DEA, inputs and outputs should emulate the activity of a DMU under study. They should be complementary and have logical cause–effect relationships (Tong and Liping, 2009). For the current study, we examine a set of research groups within a particular university in Mexico. Professors (the researchers) within these groups constitute the input variables, whereas the research articles, authored or co-authored by those professors, are the output variables of interest. An important characteristic that should be mentioned is that in Mexico there is an association to which some of the Mexican as well as foreign researchers and scientists working in Mexico belong. This association is called "The National Researchers System," from the Spanish "Sistema Nacional de Investigadores — SNI." The SNI was created by a Presidential Agreement published in the Official Journal of the Federation on July 26, 1984 to recognize the work of people dedicated to producing scientific knowledge and technology. The SNI is one of the main tools that the Mexican government has to promote the quality of scientific research. This system endorses the quality of the research performed by the SNI researchers. In order to be accepted to the system, a researcher needs to prove productivity in the forms of peer-reviewed papers, technological developments, patents, books, etc. Each research group under study has both SNI and non-SNI professors. Both types of professors conduct research and develop scientific articles, which will be the outputs in our study; the former (SNI) dedicate more time to research activities than is the case for the latter (non-SNI) and the latter aspire to become SNI researchers (CONACyT, 2017).

On the output side, the "articles" variable is separated into two major classes:

(1) "Internal" or in-house articles, namely, those published by faculty within a group, without involvement of colleagues from other groups;
(2) "External" articles, specifically those articles that were developed by professors from two research groups; hence, professors and researchers from two research centers collaborate. Note that we consider collaborations between say DMU1 and DMU2 as being a different output than collaborations between some other pair, DMU1 and DMU3.

In the particular university under study, there were 41 research groups, each of which normally generates articles of both the "internal" and "external" types. For purposes of the study herein, and due to the privacy of the information provided, we have labeled the research groups "RG1" to "RG41," rather than using actual names. The research groups play the roles of DMUs and are listed randomly in this study. The data used in the study are numerical; for the inputs we used numbers of professors and/or researchers in each research group, whereas the numbers of articles, both internal and external, represent the outputs.

In the study of n research groups, we have used the notation $(x_{ij}, i = 1, 2; j = 1, \ldots, n)$ to denote inputs and $(y_{rj}, r = 1, 2, \ldots, n)$ to denote outputs as follows:

$x_{1j} =$ Professors in research group j who are members of the National Researchers System;
$x_{2j} =$ Professors in research group j who are not members of the National Researchers System;
$y_{rr} =$ Internal articles $(r = 1, 2, \ldots, n)$;
$y_{rj} =$ External or shared articles between research groups r and $j (j \neq r)$ and $(j, r = 1, 2, \ldots, n)$.

Our aim is to assess a set of DMUs, some of which share outputs with others. As well, it is *output expansion* rather than input reduction that is more appropriate for this case; specifically, the efficiency score should focus on required enhancements of outputs (research articles), rather than reduction of inputs (professors), that are needed to bring a DMU to the frontier. This implies that it is the output rather than the input-oriented model that is most appropriate. Regarding the returns to scale technology, the CRS technology was chosen over the VRS technology to reflect that the numbers of publications produced can be viewed as being somewhat proportional to the numbers of professors involved.

3.2. *The conventional DEA model*

The output-oriented CRS model of Charnes *et al.* (1978) in ratio form is given by:

$$\begin{aligned}
&\text{Min } \sum_{i=1}^{I} v_i x_{ij_o} \Big/ \sum_{r=1}^{R} u_r y_{rj_o} \\
&\text{Subject to:} \\
&\sum_{i=1}^{I} v_i x_{ij} - \sum_{r=1}^{R} u_r y_{rj} \geq 0, \quad \forall\, j \\
&v_i, u_r \geq 0, \quad \forall\, i, r
\end{aligned} \tag{1}$$

or in linear programming format

$$\begin{aligned}
&\text{Min } \sum_{i=1}^{I} v_i x_{ij_o} \\
&\text{Subject to:} \\
&\sum_{r=1}^{R} \mu_r y_{rj_o} = 1 \\
&\sum_{i=1}^{I} v_i x_{ij} - \sum_{r=1}^{R} \mu_r y_{rj} \geq 0, \quad \forall\, j \\
&v_i, \mu_r \geq 0, \quad \forall\, i, r.
\end{aligned} \tag{2}$$

The dual of this (multiplier) model (2) is given by the envelopment format shown in model (3):

$$\begin{aligned}
&e = \text{Max}\, \varphi \\
&\text{Subject to:} \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geq \varphi y_{rj_o}, \quad \forall\, r \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{ij_o}, \quad \forall\, i \\
&\lambda_j \geq 0, \quad \varphi \text{ unrestricted in sign.}
\end{aligned} \tag{3}$$

It is important to point out at this stage that an alternative to the radial projection model (3) is a form of *additive* model given by model (4):

$$\begin{aligned}
&e = \text{Max} \sum_{r=1}^{R} \varphi_r / |R| \\
&\text{Subject to:} \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geq \varphi_r y_{rj_o}, \quad \forall\, r
\end{aligned}$$

$$\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{ij_o}, \quad \forall i$$

$$\lambda_j \geq 0, \ \forall j, \varphi_r \geq 1, \ \forall r, \tag{4}$$

where $|R|$ denotes the cardinality of the output set R. Model (4) does not require that the same proportional enhancement factor φ applies (radially) to all outputs, but rather we permit output enhancement factors φ_r specific and unique to the particular output involved. Ultimately, it will be this form of the output-oriented model that we will apply herein. The additive or slacks-based type of model has appeared in various forms over the past several decades, with some of the earliest being that by Färe and Lovell (1978), who discuss the Russel measure model (and gave it this name). One might argue that this early work by Färe and Lovell possibly inspired the slacks-based model of Tone and Tsutsui (2009). Important as well, is the work of Charnes *et al.* (1985). Cooper *et al.* (1999) discuss various forms of the model, as does Thrall (1996). It is immediately clear from the various forms of the additive model that one can adopt pure input or pure output versions as discussed by Zieschang (1984).

To illustrate the difference between the conventional DEA model with independent DMUs and the collaborative or shared output setting to be addressed herein, let us consider the simple situation in which there are two DMUs A and B with outputs as displayed in Table 3.

In this example, DMU A has four internal papers and five papers shared with DMU B. Similarly, DMU B has those same five papers shared with A and 10 internal papers. Note that the matrix of outputs is *symmetric* relative to the main diagonal.

Figure 1 shows the positioning of the two points and the result of the application of the radial projection model (3). Let us assume for this simple illustration that the two DMUs have identical inputs, which we can ignore. We observe that DMU B is efficient, and the efficient frontier is defined by CBD. DMU A is located at point A and is clearly inefficient. If we ignore for the moment the fact that the (A, B) and (B, A) slots in the matrix *should be numerically the same*, then in applying the conventional DEA model (3) to DMU A, that point is projected to the weak efficient point A′. The frontier remains fixed in place, the optimal value of the enhancement factor is $\hat{\varphi} = 1.25$ and the coordinates of the new projected point A′ are (5, 6.25).

Table 3. A 2-DMU example.

DMU	A (output 1)	B (output 2)
A	4	5
B	5	10

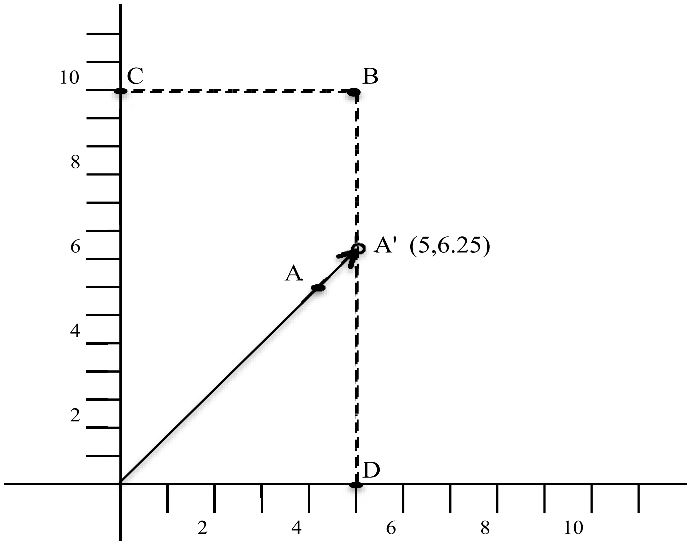

Fig. 1. Efficiency measurement under conventional projection.

3.3. *A collaborative model*

The above output-oriented model (3) is appropriate when DMUs do not collaborate, but rather operate independently. In the case that collaboration does occur, as in the situation of the existence of shared papers as described earlier, then it is necessary to modify Fig. 1; an altered version is given in the form of Fig. 2. Here we note, as mentioned earlier, that since the output (papers) that A shares with B increases from 5 to 6.25 (call this new point A′), then the first coordinate of point B (currently set at 5) should be changed to 6.25 as well, hence B moves to point $\mathrm{B}' = (6.25, 10)$. Thus, the new frontier moves from CBD to CB′D′. Pictorially, this means that if we project point A to A′, that projected point is *not yet* on the frontier (in Fig. 2), unlike the non-collaborating situation pictured in Fig. 1. So, moving the inefficient point A invokes a simultaneous move in the efficient point B. In this particular example, the second component of DMU A is moving up at a rate of 5/4 or 1.25% relative to the first component. The increase φ in the second component of DMU A is restricted to be such that $5\varphi \leq 10$. Thus, the optimal enhancement factor is $\varphi^* = 2$, point B moves to $\mathrm{B}'' = (10, 10)$, the final *altered* frontier is CB″D″ and the final projection of A to that altered frontier is $\mathrm{A}'' = (8, 10)$.

In a collaborative setting such as that described herein, it is necessary to modify the conventional DEA methodology (4) to accommodate the fact that the frontier may not remain fixed, but rather is allowed to *shift* as inefficient points are adjusted "upward." In the interest of creating a transparent description of the analysis of the 41 research groups, we consider a smaller illustrative example involving a set of five DMUs, as displayed in the following table showing inputs and both internal and external outputs.

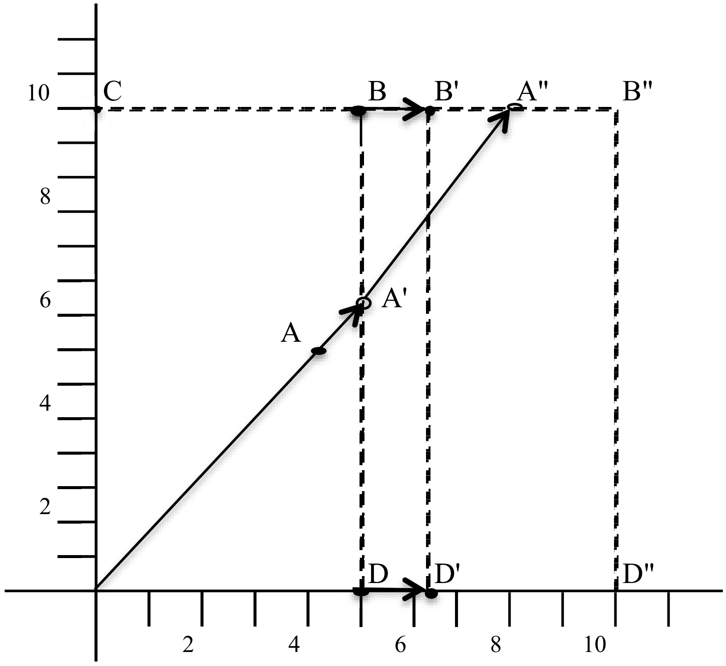

Fig. 2. Efficiency measurement under cooperative projection.

For example, in this hypothetical case, research group #1 produces 20 internal publications, 21 publications jointly with research group 2 and seven publications jointly with research group 5. We can view the research groups 1–5, as the DMUs j; these are listed in the first column. For each DMU, the outputs r are of $n = 5$ types (numbered as 1–5 across the first row of the table). So, in the case of DMU1, for example, the relevant outputs are the 20 papers done internally, 21 papers in collaboration with DMU2, zero papers in collaboration with each of DMUs 3 and 4 and seven papers in collaboration with DMU5. To specify the objective function of any DMU j_o in more correct terms, we ignore any output r for which $y_{rj_o} = 0$.

If we again ignore the collaboration element involved and simply apply envelopment model (4) to say DMU$j_o = 1$, we would obtain the following version labeled as model (4-1).

$$e = \max\,(\varphi_1 + \varphi_2 + \varphi_5)/3$$

Subject to:

Input restrictions:

$$20\lambda_1 + 20\lambda_2 + 4\lambda_3 + 12\lambda_4 + 30\lambda_5 \leq 20$$

$$15\lambda_1 + 12\lambda_2 + 1\lambda_3 + 4\lambda_4 + 10\lambda_5 \leq 15$$

Output restrictions:

$$\begin{aligned}
&20\lambda_1 + 21\lambda_2 + 0\lambda_3 + 0\lambda_4 + 7\lambda_5 \geq 20\varphi_1\\
&21\lambda_1 + 25\lambda_2 + 5\lambda_3 + 5\lambda_4 + 10\lambda_5 \geq 21\varphi_2\\
&0\lambda_1 + 5\lambda_2 + 10\lambda_3 + 0\lambda_4 + 2\lambda_5 \geq 0\varphi_3\\
&0\lambda_1 + 5\lambda_2 + 0\lambda_3 + 20\lambda_4 + 0\lambda_5 \geq 0\varphi_4\\
&7\lambda_1 + 10\lambda_2 + 2\lambda_3 + 0\lambda_4 + 15\lambda_5 \geq 7\varphi_5\\
&\varphi_r \geq 1, \quad \forall\, r, \lambda_j \geq 0, \ \forall j
\end{aligned} \tag{4-1}$$

The resulting output-oriented efficiency score using model (4-1) is $e = 1.223$, and the individual enhancement factors are $(\varphi_1, \varphi_2, \varphi_5) = (1.05, 1.19, 1.43)$. We interpret this to mean that under the assumption that the conventional methodology reflects reality, then the recommended enhancements to DMU 1's *three* classes of publications (needed to render DMU 1 efficient) call for a 5% increase in internal publications, a 19% increase in publications shared with DMU 2 and a 43% increase in publications shared with DMU 5. Note that in forming the constraint set in model (4-1), we replace the values (20, 21, 0, 0, 7) in the *first row* of the output portion of the data table by the altered form $(20\phi_1, 21\phi_2, 0\phi_3, 0\phi_4, 7\phi_5)$; these represent the right-hand sides of the output constraints. Now, since the matrix must remain symmetric, then the elements of the *first column* must be similarly altered. This means that the first output constraint (using the data in the first column) should appear in the altered form as:

$$20\lambda_1\phi_1 + 21\lambda_2\phi_2 + 0\lambda_3\phi_3 + 0\lambda_4\phi_4 + 7\lambda_5\phi_5 \geq 20\phi_1. \tag{5}$$

The other constraints remain unchanged. (Note again that since DMU 1 has only three types of outputs (not five), it is only the enhancement factors φ_r for those 3 that we count.)

It is noted that constraint (5) is *non-linear* by way of the products $\lambda_j\varphi_r$; hence, the GRG code in Solver was applied. The larger efficiency score under this altered form of the model is $e = 1.290$, compared with that arising from the conventional model (i.e., 1.223), reflecting the fact that constraint (1) permits the frontier to shift. In further detail, the individual enhancement factors that yield this large score are given by $(\varphi_1, \varphi_2, \varphi_5) = (1.25, 1.19, 1.43)$. Note that while (φ_2, φ_5) are identical in the two models, it is the size of φ_1 that is different. Specifically, rather than calling for a 5% increase in the number of internal publications under model (4-1), the corrected model requires a 25% increase in internal publications in order for DMU$j_o = 1$ to become efficient.

It is recognized that GRG may yield only a local, and not necessarily a global, optimum in a general non-linear programming problem. We did undertake a number of simulations on small problems, treating the phi variables as parameters, and found that the resulting best values of those parameters came arbitrarily close to the optimal phi values arising from GRG. The phenomenon of a local optimum different from global never arose for the special form of the non-linearity characterizing the problem herein. Note that the form of the non-linearity herein arises only from a single constraint and when expressed in vector format $X^T QP$, where Q is the identity matrix. This may explain why local optima appear to be global as well.

While such simulated results do not constitute proof that local optimal solutions will always be globally optimal, such results add a measure of credibility to the solutions arising from the above model. The simple fact remains, however, that no proof yet exists that local optima are also global, in the shared output problem. Further research is required.

Generalizing this concept for any DMUj_o, we replace model (4-1) by model (4-j_o):

$$
\begin{aligned}
&e = \max \sum_{r \in R^+_{j_o}} \varphi_r / |R^+_{j_o}| \\
&\text{Subject to:} \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geq \varphi_r y_{rj_o}, \quad \forall\, r \in R^+_{j_o},\ r \neq j_o \\
&\sum_{j=1}^{n} \lambda_j \varphi_j y_{j_o j} \geq \varphi_{j_o} y_{j_o j_o} \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{ij_o}, \quad \forall\, i \\
&\lambda_j \geq 0, \quad \forall\, j, \varphi_r \geq 1,\ \forall\, r \in R^+_{j_o}.
\end{aligned}
\tag{4-j_o}
$$

Here $R^+_{j_o}$ denotes those outputs r such that $y_{rj_o} > 0$, and $|R^+_{j_o}|$ is the cardinality of $R^+_{j_o}$.

Again, we point out that the second constraint is imposed purely to insure that the matrix such as that given by Table 4 *remains symmetric*. In other words, it is the DMUs that share publications with the DMUj_o under evaluation that cause the frontier to move as pictured in Figs. 1 and 2. Hence, because A and B share publications, then as A moves so also does that dimension of B representing the publications shared with A. If there had been a DMU C in the figures that did not share publications with DMU A, nothing would happen to that DMU as A

Table 4. A 5-DMU example.

DMU	Output type					Input type	
	1	2	3	4	5	1	2
1	20	21	—	—	7	20	15
2	21	25	5	5	10	20	12
3	—	5	10	—	2	4	1
4	—	5	—	20	—	12	4
5	7	10	2	—	15	30	10

is pushed toward the frontier. Hence, only the left side of the *second* constraint is altered, not those represented by the first constraint.

It is important to emphasize that the set of outputs $J$is the same as the set of all DMUs. The outputs are defined as "shared publications with research group 1," "shared with research group 2," "shared with research group 3,"..., "shared with

Table 5. Data for research groups.

DMU research groups	Professors who are members of National Researchers System	Professors who are no-members of National Researchers System	Internal papers	External papers
RG 1	11	39	19	0
RG 2	10	13	12	0
RG 3	11	101	55	20
RG 4	12	150	58	15
RG 5	17	4	4	0
RG 6	1	39	5	0
RG 7	6	159	64	3
RG 8	2	11	8	1
RG 9	2	37	13	1
RG 10	5	45	11	1
RG 11	8	32	11	7
RG 12	4	4	3	0
RG 13	18	0	8	1
RG 14	2	12	5	2
RG 15	2	6	4	0
RG 16	2	28	11	6
RG 17	1	1	0	0
RG 18	2	2	0	0
RG 19	3	5	5	1
RG 20	8	543	87	18
RG 21	5	119	33	10
RG 22	14	346	96	7
RG 23	3	349	61	37
RG 24	8	694	133	1
RG 25	5	252	82	32
RG 26	6	410	124	60
RG 27	4	155	38	18
RG 28	8	47	26	4
RG 29	9	428	88	10
RG 30	9	351	88	7
RG 31	7	301	134	45
RG 32	5	136	52	16
RG 33	3	177	78	9
RG 34	10	360	60	15
RG 35	4	108	55	13
RG 36	7	146	81	1
RG 37	3	111	61	46
RG 38	3	91	25	15
RG 39	3	79	23	15
RG 40	4	205	29	7
RG 41	8	370	64	10

Table 6. Distribution of the external papers among the DMUs.

DMU	Internal paper	External papers with								
1	19									
2	12									
3	55	$DMU_4 = 4$	$DMU_{23} = 1$	$DMU_{28} = 2$	$DMU_{33} = 13$					
4	58	$DMU_3 = 4$	$DMU_7 = 1$	$DMU_{10} = 1$	$DMU_{11} = 1$	$DMU_{13} = 1$	$DMU_{23} = 1$	$DMU_{24} = 1$	$DMU_{31} = 3$	$DMU_{35} = 2$
5	4									
6	5									
7	64	$DMU_4 = 1$	$DMU_{14} = 2$							
8	8	$DMU_{21} = 1$								
9	13	$DMU_{20} = 1$								
10	11	$DMU_4 = 1$								
11	11	$DMU_4 = 1$	$DMU_{16} = 6$							
12	3									
13	8	$DMU_4 = 1$								
14	5	$DMU_7 = 2$								
15	4									
16	11	$DMU_{11} = 6$								
17	0									
18	0									
19	5	$DMU_{23} = 1$								
20	87	$DMU_{19} = 1$	$DMU_{21} = 4$	$DMU_{27} = 4$	$DMU_{32} = 7$	$DMU_{37} = 2$				
21	33	$DMU_{18} = 1$	$DMU_{21} = 2$	$DMU_{24} = 2$	$DMU_{27} = 1$	$DMU_{28} = 1$	$DMU_{30} = 2$	$DMU_{32} = 1$		
22	96	$DMU_9 = 1$	$DMU_{22} = 1$	$DMU_{24} = 3$	$DMU_{29} = 1$	$DMU_{31} = 1$				
23	61	$DMU_8 = 1$	$DMU_{18} = 4$	$DMU_{19} = 2$	$DMU_{24} = 3$	$DMU_{27} = 8$	$DMU_{30} = 3$	$DMU_{36} = 11$	$DMU_{37} = 4$	$DMU_{38} = 1$
24	13	$DMU_{20} = 1$								

Table 6. (*Continued*)

DMU	Internal paper	External papers with							
25	82	$DMU_3 = 1$	$DMU_4 = 1$	$DMU_{17} = 1$	$DMU_{24} = 25$	$DMU_{29} = 2$	$DMU_{35} = 2$		
26	124	$DMU_4 = 1$	$DMU_{19} = 2$	$DMU_{20} = 3$	$DMU_{21} = 3$	$DMU_{23} = 25$	$DMU_{25} = 18$	$DMU_{30} = 7$	$DMU_{39} = 1$
27	38	$DMU_{24} = 18$							
28	26	$DMU_{18} = 4$							
29	88	$DMU_{19} = 1$	$DMU_{21} = 8$	$DMU_{32} = 1$					
30	88	$DMU_3 = 2$	$DMU_{19} = 1$	$DMU_{32} = 4$					
31	134	$DMU_{20} = 1$	$DMU_{23} = 2$	$DMU_{30} = 1$	$DMU_{31} = 2$	$DMU_{35} = 37$	$DMU_{36} = 1$	$DMU_{38} = 1$	
32	52	$DMU_{19} = 2$	$DMU_{21} = 3$	$DMU_{24} = 7$	$DMU_{29} = 1$	$DMU_{34} = 1$	$DMU_{35} = 2$		
33	78	$DMU_4 = 3$	$DMU_{20} = 1$	$DMU_{29} = 2$	$DMU_{35} = 3$				
34	60	$DMU_{18} = 7$	$DMU_{19} = 1$	$DMU_{27} = 1$	$DMU_{28} = 4$	$DMU_{39} = 2$			
35	55	$DMU_3 = 13$							
36	81	$DMU_{30} = 1$							
37	61	$DMU_4 = 2$	$DMU_{23} = 2$	$DMU_{29} = 37$	$DMU_{30} = 2$	$DMU_{31} = 3$			
38	25	$DMU_{21} = 11$	$DMU_{29} = 1$	$DMU_{37} = 2$	$DMU_9 = 1$				
39	23	$DMU_{18} = 2$	$DMU_{21} = 4$	$DMU_{36} = 2$	$DMU_{38} = 3$	$DMU_{39} = 4$			
40	29	$DMU_{21} = 1$	$DMU_{29} = 1$	$DMU_{37} = 3$	$DMU_{39} = 2$				
41	64	$DMU_{24} = 1$	$DMU_{32} = 2$	$DMU_{36} = 1$	$DMU_{37} = 4$	$DMU_{38} = 2$			

research group n." So, for example, in Table 4, DMU 1 has 21 shared publications with DMU 2. The other four DMUs 2, 3, 4 and 5 also share publications in amounts 25, 5, 5 and 10, respectively, with DMU 2. Now if DMU 1 were to say have only one set of publications, namely, those shared with DMU 2, then model (3) would allow projection only on that dimension until the frontier is incurred. Hence, on the output dimension "shared papers with DMU 2," there are only those DMUs that also share with DMU 2. Those other shared (with DMU 2) papers do not change as DMU 1 is pushed to the frontier. Hence the frontier will be reached.

The full set of scores under the conventional model is (1.223, 1, 1, 1, 8.938). Under the collaborative model, these become (1.290, 1, 1, 1, 8.938). In the following section, we examine the application of model (4-j_o) to the full set of research units in a Mexican university.

4. Evaluating Research Groups at a Mexican University: An Application

Tables 5 and 6 display data on 41 research groups at a university in Mexico. Specifically, Table 5 provides the input data (the two types of professors) and the outputs in terms of the total numbers of internal and external papers. Table 6 provides specific details as to the numbers of academic papers written collaboratively between pairs of research groups. As mentioned previously, we have codified each of the research groups with the letters "RG" and with a number.

Table 7. Conventional CRS results.

DMUs: Research groups	Efficiency score: PHI	DMUs: Research groups	Efficiency score: PHI
1	1	22	1
2	1	23	1
3	1	24	1
4	1	25	1
5	1	26	1
6	1	27	1
7	1.953	28	1.367
8	1	29	1
9	1	30	1.042
10	1	31	1
11	1	32	1
12	1	33	1.128
13	1	34	1
14	1	35	1.427
15	1	36	1
16	1	37	1.079
17	0	38	1
18	0	39	1
19	1	40	1
20	1.036	41	1
21	1		

It is important to point out that we have assumed herein that all collaborations are between *pairs* of research centers as opposed to say collaborations among three or more centers. Consider the case in which research center 1 shares papers with two other centers say 2 and 3. What we have done in this paper is to put the same weight on three-center papers as on two-center papers. So, shared papers between DMUs 1 and 2 are taken as being the sum of those papers shared *only* between the

Table 8. Results from applying the new methodology.

DMUs: Research groups	Number of DMUs sharing outputs with this DMU	Efficiency score
RG 1	0	1.371
RG 2	0	1
RG 3	4	1.064
RG 4	9	1.42
RG 5	0	2.438
RG 6	0	4.17
RG 7	2	1.304
RG 8	1	1
RG 9	1	1.652
RG 10	1	2.457
RG 11	2	1.934
RG 12	0	1.57
RG 13	1	1
RG 14	1	1.704
RG 15	0	1.239
RG 16	1	1.529
RG 17	0	0
RG 18	0	0
RG 19	1	1
RG 20	5	1.54
RG 21	7	2
RG 22	5	1.396
RG 23	9	1.279
RG 24	1	1.008
RG 25	6	1.297
RG 26	8	1
RG 27	1	2.054
RG 28	1	1.123
RG 29	3	1.523
RG 30	3	1.523
RG 31	7	1
RG 32	6	1.415
RG 33	4	1
RG 34	5	2.233
RG 35	1	1.088
RG 36	1	1
RG 37	5	1
RG 38	4	2.019
RG 39	5	1.92
RG 40	4	3.146
RG 41	5	2.094

two centers, 1 and 2, and say those among centers 1, 2 and 3. Otherwise, if we wish to distinguish these two types of papers, we would need to create three new output measures, "shared with 2 and 3," "shared with 1 and 2" and "shared with 1 and 3." If we were to attempt to consider all possible combinations, the number of output dimensions explodes. Therefore, we have taken the alternative route of keeping the output set at the n values described.

When running the conventional model, most of the DMUs are rated as efficient (have a score of unity); only seven (7) of the research groups show an efficiency score less than one. This may be due primarily to the fact that the number of internal publications for a DMU tends to dominate any collaborative publications that the DMU may possess. In addition, if a DMU "a" collaborates with another unit "b," but no other center collaborates with "b," then group "a" can put a heavy weight on that collaboration and hence be scored as efficient. Table 7 displays the results obtained from the conventional model. We point out that two of the research groups, RG17 and RG18, have no publications and hence received an efficiency score of 0 (zero).

When we apply the new model, the results obtained appear rather more realistic. Table 8 displays the results. We note that the shifting frontier phenomenon results in only nine efficient research groups, with 30 of the groups being designated as inefficient.

We believe that this new methodology helps one to better evaluate each research group by taking into account its own "in house" productivity, as well as its collaborative productivity. The latter element is difficult to capture by way of classical DEA.

5. Conclusions

This paper has addressed the important issue in DEA of dependence among the set of DMUs under investigation. We look at the specific problem of measuring the efficiencies of a set of research groups where dependence arises from research articles done jointly by pairs of those groups. In simple terms, under dependence, as inefficient DMUs are projected toward the frontier, portions of that frontier can shift at the same time. As a consequence, output enhancement factors φ_r (in the adjusted model) tend to be larger (less efficient) than under the conventional model. This means that the conventional model, which assumes that the frontier is fixed, tends to overstate efficiencies.

One shortcoming of the moving frontier model is that it is non-linear, meaning that there is potential for local optimal solutions that may not be globally optimal. Future research will examine this aspect of the problem.

A related problem is that dependence among DMUs is instigated by the input side of the DEA model. Such might be the case when there is say a common input servicing a subgroup of the DMUs. The problem there is that in reducing that factor, while projecting a particular inefficient DMU toward the frontier (in an input-oriented sense), other efficient DMUs that are also serviced by that same factor will cause the frontier to move.

Acknowledgments

Wade Cook acknowledges financial support from NSERC Grant No. A8966. This research is also supported by National Natural Science Funds of China (No. 71828101) and Natural Sciences and Engineering Research Council of Canada (NSERC) (No. A8966), and Poligrant #5487 from the USFQ. The authors also wish to express their appreciation to the referee and Area Editor for their constructive comments.

References

Agasisti, T, G Catalano, P Landoni and R Verganti (2012). Evaluating the performance of academic departments: An analysis of research-related output efficiency. *Research Evaluation*, 21(1), 2–14.

Amirteimoori, A and RA Nashtaei (2006). The role of time in multi-component efficiency analysis: An application. *Applied Mathematics and Computation*, 177(1), 11–17.

Beasley, J (1995). Determining teaching and research efficiencies. *Journal of the Operational Research Society*, 46, 441–452.

Bessent, A, W Bessent, A Charnes, WW Cooper and N Thorogood (1983). Evaluation of educational program proposals by means of data envelopment analysis. *Educational Administration Quarterly*, 19(2), 82–107.

Bian, Y, M Hu and H Xu (2015). Measuring efficiencies of parallel systems with shared inputs/outputs using data envelopment analysis. *Kybernetes*, 44(3), 336–352.

Charnes, A, WW Cooper and E Rhodes (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.

Charnes, A, WW Cooper, B Golany, LM Seiford and J Stutz (1985). Foundations of data envelopment analysis and Pareto–Koopmans empirical production functions. *Journal of Econometrics*, 30, 91–107.

Chen, Y, J Du, D Sherman and J Zhu (2010). DEA model with shared resources and efficiency decomposition. *European Journal of Operational Research*, 207(1), 339–349.

CONACyT (2017). Sistema Nacional de Investigadores. Retrieved from https://www.conacyt.gob.mx/index.php/el-conacyt/sistema-nacional-de-investigadores.

Cook, WD and M Hababou (2001). Sales performance measurement in bank branches. *OMEGA*, 29(4), 299–307.

Cook, WD, M Hababou and HJH Tuenter (2000). Multicomponent efficiency measurement and shared inputs in data envelopment analysis: An application to sales and service performance in bank branches. *Journal of Productivity Analysis*, 14, 209–224.

Cook, WD and J Zhu (2011). Multiple variable proportionality in data envelopment analysis. *Operations Research*, 59(4), 1024–1032.

De Witte, K and L López-Torres (2015). Efficiency in education: A review of literature and a way forward. *Journal of the Operational Research Society*, 1–15.

Färe, RS and CAK Lovell (1978). Measuring the technical efficiency of production. *Journal of Economic Theory*, 19, 150–162.

Giménez, V and JL Martínez-Parra (2006). Cost efficiency in the university: A departmental evaluation model. *Economics of Education Review*, 25(5), 543–553. doi:10.1016/j.econedurev.2005.05.006.

Hollingshead, AB and MS Poole (2012). Groups research methods: An introduction. In *Research Methods for Studying Groups and Teams: A Guide to Approaches, Tools,* *and Technologies*, AB Hollingshead and MS Poole (Eds.), pp. 1–5. New York, USA: Routledge.

Kao, C and HT Hung (2008). Efficiency analysis of university departments: An empirical study. *Omega*, 36(4), 653–664.

Kao, C (2009). Efficiency measurement for parallel production systems. *European Journal of Operational Research*, 196(3), 1107–1112.

Kounetas, K, A Anastasiou, P Mitropoulos and I Mitropoulos (2011). Departmental efficiency differences within a Greek University: An application of a DEA and Tobit analysis. *International Transactions in Operational Research*, 18(5), 545–559.

López-Yáñez, J and M Altopiedi (2015). Evolution and social dynamics of acknowledged research groups. *Higher Education*, 70(4), 629–647.

Tauer, LW, HO Fried and WE Fry (2007). Measuring of academic departments within a college. *Education Economics*, 15, 473–489.

Thrall, RM (1996). The lack of invariance of optimal dual solutions under translation. *Annals of Operations Research*, 66, 103–108.

Tone, K and M Tsutsui (2009). Network DEA: A slacks-based measure approach. *European Journal of Operational Research*, 197(1), 243–252.

Tong, L and C Liping (2009). *Research on the Evaluation of Innovation Efficiency for China's Regional Innovation System by Utilizing DEA*. Paper presented at the 2009 International Conference on Information Management, Innovation Management and Industrial Engineering. Retrieved from http://0-ieeexplore.ieee.org.millenium.itesm.mx/stamp/stamp.jsp?tp=&arnumber=5369923.

Tsai, PF and CM Molinero (2002). A variable returns to scale data envelopment analysis model for the joint determination of efficiencies with an example of the UK health service. *European Journal of Operational Research*, 141(1), 21–38.

Zieschang, K (1984). An extended Farrell efficiency measure. *Journal of Economic Theory*, 33(2), 387–396.

Biography

Sonia Valeria Avilés-Sacoto is an Associate Professor in the School of Industrial Engineering at Universidad San Francisco de Quito in Ecuador. She holds a Ph.D. degree in Engineering Sciences with a major in Industrial Engineering at Tecnológico de Monterrey — Mexico, an M.Sc. in Quality and Productivity Systems from Tecnológico de Monterrey and a B.Sc. in Industrial Engineering from the Universidad de Cuenca — Ecuador. Her research interests include operations research, modeling, production/operations management, tourism modeling, inventory control and methodology development of data envelopment analysis. Her research articles appear in a range of journals.

Wade D. Cook is Professor Emeritus at the Schulich School of Business, York University, Toronto, Canada, having received his doctorate from Dalhousie University. He is an internationally recognized expert in Data Envelopment Analysis, performance evaluation and benchmarking and multi-criterion decision modeling. He has published several books and authored or co-authored more than 170 articles in a wide range of top-quality journals. He holds a number of editorial appointments and has held numerous visiting scholar positions.

David Güemes-Castorena is a Full Professor of Strategic Management of Technology and Innovation at Tecnologico de Monterrey, Campus Monterrey. He received his D.Sc. from The George Washington University. His research interest involves technological strategy and applied technological foresight. He is currently a visiting scholar at the MIT Sloan School of Management.

Francisco Benita is a Postdoctoral Research Fellow at Singapore University of Technology and Design. He received his Ph.D. in Engineering Sciences in 2016 from Tec de Monterrey. His work lies on the intersection of optimization, data science and micro-econometrics.

Hector Ceballos is head of the Scientometrics Office at the Research Vice-Rectory of Tecnologico de Monterrey, and he is ascribed to the Intelligent Systems group within the National School of Engineering and Science. He has a Masters and a Ph.D. in Intelligent Systems from Tecnologico de Monterrey, and his main research interests include social network analysis, process mining and agent theory, applied to scientometrics and cultural heritage. He is author of more than 30 papers in specialized journals and conferences.

Joe Zhu is a Professor of Operations Analytics in the Foisie School of Business at Worcester Polytechnic Institute, USA. He is the founding Dean of the College of Auditing and Evaluation at Nanjing Audit University. He is an internationally recognized expert in methods of performance evaluation and benchmarking, and his research interests include operations and business analytics, productivity modeling and performance evaluation and benchmarking. He has more than 26,000 google scholar citations, has published more than 15 books related to data envelopment analysis and 130 peer-reviewed articles, many in top-tier journals. He holds a number of editorial appointments and has had numerous prestigious visiting scholar appointments.