Omega 41 (2013) 3–15

Contents lists available at SciVerse ScienceDirect

Omega

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

Data envelopment analysis 1978–2010: A citation-based literature survey

John S. Liu [a,*], Louis Y.Y. Lu [b], Wen-Min Lu [c], Bruce J.Y. Lin [b]

[a] Graduate Institute of Technology Management, National Taiwan University of Science and Technology, 43, Section 4, Keelung Road, Taipei 10607, Taiwan
[b] College of Management, Yuan Ze University, 135 Yuan-Tung Road, Chung-Li, Taoyuan 32003, Taiwan
[c] Department of Financial Management, National Defense University, 70, Section 2, Zhongyang North Road, Beitou, Taipei 112, Taiwan

ARTICLE INFO

Available online 9 February 2012

Keywords:
Data envelopment analysis
Literature survey
Citation analysis
Main path analysis

ABSTRACT

This study surveys the data envelopment analysis (DEA) literature by applying a citation-based approach. The main goals are to find a set of papers playing the central role in DEA development and to discover the latest active DEA subareas. A directional network is constructed based on citation relationships among academic papers. After assigning an importance index to each link in the citation network, main DEA development paths emerge. We examine various types of main paths, including local main path, global main path, and multiple main paths. The analysis result suggests, as expected, that Charnes et al. (1978) [Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European Journal of Operational Research 1978; 2(6): 429–444] is the most influential DEA paper. The five most active DEA subareas in recent years are identified; among them the "two-stage contextual factor evaluation framework" is relatively more active. Aside from the main path analysis, we summarize basic statistics on DEA journals and researchers. A growth curve analysis hints that the DEA literature's size will eventually grow to at least double the size of the existing literature.

1. Introduction

Pioneers of data envelopment analysis (DEA) may not have expected that their ideas have inspired the thinking of a group of researchers and have been developed collectively into a widely accepted academic field. Thirty some years after the publication of the seminal paper by Charnes, Cooper and Rhodes [1], the development continues and has not seen any signs of weakening. In 2009 alone, more than 700 DEA papers were published. Up through the year 2009, the field has accumulated approximately 4500 papers in ISI Web of Science database.

DEA is a non-parametric productive efficiency measurement method for operations with multiple inputs and multiple outputs. According to Seiford [2], DEA in its current form was first described in Charnes et al. [1], who propose a novel method that combines and transforms multiple inputs and outputs into a single efficiency index. This approach first establishes an "efficient frontier" formed by a set of decision making units (DMUs) that exhibit best practices and then assigns the efficiency level to other non-frontier units according to their distances to the efficient frontier. The basic idea has since generated a wide range of variations in measuring efficiency. Today, various DEA efficiency models, such as the constant returns to scale (CRS) model, the variable-returns-to-scale (VRS) model, the additive model, the slacks-based measures and the free disposal hull (FDH) model, etc. are available for different types of measuring requirement. It also has been applied to various industrial and non-industrial contexts, such as banking, education, hospital, etc. [3]. In addition, research targets spread globally to countries such as China [4, 5], Greece [6], Turkey [7], Norway [8] and UAE [9], etc.

Several authors have surveyed the general DEA literature and provide scenarios for DEA methodology development in different time periods from a range of angles. These surveys can be categorized into three types – bibliography listing, qualitative, and quantitative. Seiford [10] and Gattoufi et al. [11] provide bibliography listing with an extensive list of DEA literature. Qualitative survey includes Seiford and Thrall [12], Seiford [2], Cooper et al. [13], and Cook and Seiford [14]. Seiford and Thrall [12] review early-stage DEA development. Seiford [2] traces the evolution of DEA for the period 1978 through 1995, describing the major achievements at each of the four milestones: 1980, 1985, 1990, and 1995. A pictorial evolution map in the article highlighting the timing of major events and the births of new ideas is quite useful. The best part of the map is that it indicates the relationship among ideas, and so it is easy to trace the origin of a new idea. Cooper et al. [13] review, from a theoretical perspective, some DEA models and measures. Cook and Seiford [14] conduct a comprehensive review on the methodological developments since 1978, where most of the important DEA subjects such as generic

* Corresponding author.
E-mail addresses: johnliu@mail.ntust.edu.tw (J.S. Liu), louislu@saturn.yzu.edu.tw (L.Y.Y. Lu), wenmin.lu@gmail.com (W.-M. Lu), s959415@mail.yzu.edu.tw (B.J.Y. Lin).

0305-0483/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.omega.2010.12.006

DEA models, multilevel models, multiplier restrictions, considerations on the status of variables, and data variation, etc. are discussed.

Gattoufi et al. [15] and Emrouznejad et al. [3] both carry out quantitative surveys and present the DEA publication statistics, including summaries of bibliographic data by journals, by authors, etc. Gattoufi et al. [15] gather data from six online professional databases, whereas Emrouznejad et al. [3] maintain a DEA literature database that updates regularly with support from the DEA community. In contrast with these two studies, our data are from ISI Web of Science, which includes the essence of this study – citation.

Citation in academic articles contains rich information on how knowledge disseminates. It has long been used to evaluate the level of contribution a scientist makes to the practice of science. Although there are some arguable problems with such usage of citation data – for example, negative citation and self-citation – it is nevertheless a useful and relatively low cost tool to evaluate scientific performance [16]. Citation count, for example, is commonly used to indicate an academic article's acceptance.

This study adopts two citation-based methods: the main path analysis and the g-index/h-index. The main path analysis [17] is a well-known method that traces the main knowledge flow through citation data. In its original form, it traces the most significant path of a discipline's development. We propose a multiple-path method that includes paths of significance closely next to the most significant path. This new proposed method allows one to examine the development 'branches' to any arbitrary level. The g-index/h-index measures a scientist's academic contribution based on citations of his/her publication, proposed by Egghe [18] and Hirsch [19], respectively. We use these indices to quantify the contribution of DEA authors and journals. Aside from the two citation-based methodologies, we apply the growth curve analysis to hint at the future trend of DEA development.

The main goals of this study are, through quantitative means, to find a set of papers that plays the central role in DEA development and to discover the major DEA activities in recent years. We achieve these goals by applying the main path and multiple main path analysis – that is, tracing the DEA development paths using citation data of academic papers.

This paper is organized as follows. After the introduction we briefly explain the methodology used in this study, in particular the main path analysis, the g-index/h-index, and the growth curve analysis. Section 3 discusses how the data are acquired and presents the basic statistics, followed by a presentation and discussion of the analysis results. The last section concludes.

2. Methodologies

This study applies two citation-based methodologies: the main path analysis and the g-index/h-index. The main path analysis helps comprehend the DEA development to a more detailed level, while the g-index/h-index is used to compare the influence of DEA authors and journals. We also use the growth curve analysis to better grasp the DEA development trend. The following sections briefly introduce these methodologies.

2.1. Main path analysis

In the course of a scientific field's development, new ideas are proposed continuously. Along the way, some ideas stay and some fade away. Those ideas that stay usually raise wide and long-lasting attention. The main research question of this study is: which ideas proposed in the course of DEA development stay and make a significant influence upon the field? Main path analysis is a proper tool to help answer the question.

Hummon and Doreian [17] first introduce main path analysis and use citation information in academic papers or patents to trace the main idea flow in a scientific discipline. When a publication cites a previous work, presumably knowledge flows from the previous work to the citing publication. The method is network based, and the scientific publications are seen as nodes of a network, and citation information is used to establish links among nodes. The citation network thus created is a non-weighted directional network.

Tracing the flow of ideas in a small citation network may be easy, but the difficulty of the task increases as the network grows larger. Hummon and Doreian [17] suggest a way to simplify the task in a large citation network: tracing only the 'main path'. Identifying the importance of each citation link in the network is the first step in finding the main path. The importance of each citation can be measured by counting the times a citation link has been traversed were one exhausts the search from a set of starting nodes to another set of ending nodes. There are several variations of ways to do the count. Node pair projection count (NPPC), search path link count (SPLC), search path nodes pair (SPNP), and search path count (SPC) are mentioned in the literature [17, 20]. These counts are similar, but subtle differences exist among them. It is beyond this paper's scope to discuss the differences. We choose to use SPC as it is recommended by Batagelj [20] as the first choice.

In a citation network, a 'source' is a node that is cited, but cites no other nodes; a 'sink' is a node that cites other nodes, but is not cited. In other words, sources are the origins of knowledge, while sinks are the end points of knowledge dissemination. We use a simple citation network in Fig. 1 to demonstrate how SPCs for each links are calculated. The network has two sources, A and B, and four sinks, C, D, E, and F. There are many alternative paths to go from the sources to the sinks. Assuming that one exhausts searching all paths from all the sources to all the sinks, SPC for each link is defined as the total number of times the link is traversed. For example, link J–C has SPC value of 2, because it is passed through by paths A–H–J–C and B–H–J–C. Link B–I's SPC value is 4 as it is traversed by four paths: B–I–F, B–I–G–D, B–I–G–E and B–I–E. In the example network, B–I and H–J have the largest SPC value. The larger the SPC value is, the more important the link's role is in transmitting the knowledge.

After the SPC value for each citation link is calculated, Verspagen [21], by slightly modifying the method proposed in Hummon and Dorien [17], suggests identifying the main path of a citation network with the following procedures.

1. Find the link with the largest SPC from all possible links emanating from the sources. Assign the beginning node of this link as the start point of the main path. Take the ending node of the link as the start point for the next step. If there are ties, take all the tied links into consideration.
2. Find the link with the largest SPC emanating from the current start point(s). Take the ending node(s) of the link(s) as the start point(s) for the next step. If the ending node is a sink, stop. If there are ties at each start point, take all the tied links into consideration.
3. Continue Step 2 until all the paths hit a sink.

Following the above steps, one finds the main path(s) of the example network as B–I–G–D and B–I–G–E. The main path identified in this fashion is what we call the 'local' main path, as it always selects the current route with the largest SPC value, but the overall (accumulated) SPC value of this local main path may not be the largest among all paths. For example, the overall

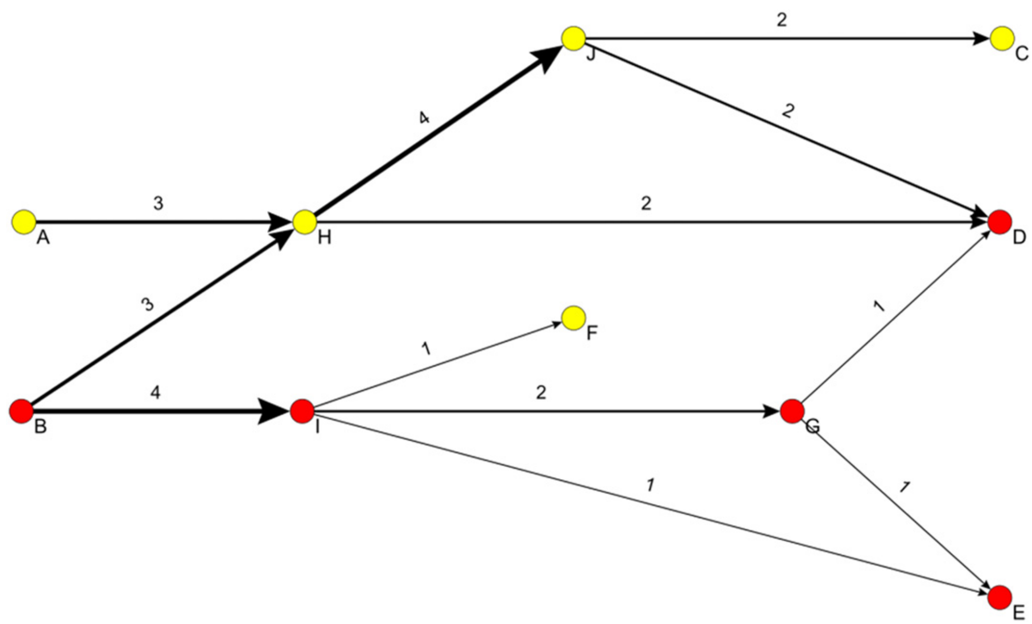

Fig. 1. SPC example.

SPC value of the paths B–H–J–C and B–H–J–D is bigger than that of local main paths B–I–G–D and B–I–G–E. The significance of these two paths should not be ignored. Here, we propose to also examine the 'global' main path, which is the path with the largest overall SPC value. The global main path, in contrast to the local main path, notes the overall importance in knowledge flow. The problem of finding the path with the overall largest SPC value is similar to the longest path problem in graph theory. For a citation network, which is always acyclic, several algorithms to solve the problem are readily available.

The main path not only indicates the development trajectory of a discipline, but also points out influential works. The papers on the main path are significant – in the way that both of their direct and indirect influences are taken into consideration; while the conventional 'citation count' reckons only direct influences. The papers upstream of the main path are more influential than those downstream. For example, node B in the example citation network is more influential than node I as it is the origin of all the later ideas even though it has a citation count less than that of node I.

The global main path methodology can be extended to examine in more detail the progress of an academic discipline. In addition to the most significant path, one can take the next several paths with overall SPC values smaller than the largest one. Including these paths allows one to observe more participants in the development. We call this the 'multiple main paths' method.

2.2. The *g*-index and the *h*-index

Hirsch [19] proposes an index to quantify an individual's scientific research output using citation information. Hirsch-index h is defined as "the number of papers with citation number $\geq h$", i.e., a researcher has index h if h of his or her papers published over a certain years in a certain scientific field have at least h citations each, and his or her other papers in the same period and the same field have $\leq h$ citations each. The index is conceptually simple and has been used successfully to capture scholars' influence in various scientific fields [22–24]. The h-index, nevertheless, does not take citation scores of a researcher's top articles into account. The g-index is an improvement over the h-index on this specific issue. If one lists a researcher's papers in decreasing order of the number of citations, "the g-index is the largest number such that the top g articles received at least g^2 citations" [18]. The g-index, although an improvement over the h-index, is not without flaws. Extremely high citation of an author's paper is still not recognized completely. The g-index and the h-index can also be applied to rate a journal's impact. We take the position that the g-index and the h-index are complementary to each other and therefore apply both of them to compare the influence of DEA authors and journals.

2.3. Growth curve analysis

Growth curve analysis is a method commonly used to project the life cycle of a physical or social system. The method is based on the assumption that the growth of objects in a system – such as human population growth on earth, bacteria expansion in laboratories cannot be unlimited; resource scarcity and environmental factors will inhibit further growth after the size of a system reaches to a certain value. Thus, much of the growth in nature follows an *S*-shaped curve. It is commonly assumed that the growth of scientific literature does the same. It increases a little slow in the beginning. At a certain point, it speeds up exponentially. After passing the 'midpoint' (where the growth rate inflects), the growth slows down and eventually reaches a growth limit. The growth curve analysis fits the given time series data with a logistic function so that one is able to predict the growth limit, the midpoint, and the life cycle of an *S*-shaped curve.

Mathematically, an *S*-shaped curve can be represented with a logistic function [25]:

$$S(t)=\frac{\kappa}{1+\exp[-(\ln(81)/\Delta t)(t-t_m)]}, \tag{1}$$

where κ is the growth limit; t_m is the midpoint of the growth trajectory; and Δt, the life cycle, is defined as the time the development takes to grow from 10% to 90% of the growth limit.

There are situations where the growth is in dual phases, usually caused by technology breakthrough or an unexpected shift of environmental factors. Under these situations the growth trajectory is better represented as the sum of two

S-shaped curves.

$$S(t) = S_1(t) + S_2(t), \tag{2}$$

where

$$S_1(t) = \frac{\kappa_1}{1+\exp[-(\ln(81)/\Delta t_1)(t - t_{m1})]},$$
$$S_2(t) = \frac{\kappa_2}{1+\exp[-(\ln(81)/\Delta t_2)(t - t_{m2})]}, \tag{3}$$

and κ_1, κ_2 are the growth limits; t_{m1}, t_{m2} are the midpoints; Δt_1, Δt_2 are the characteristic durations for the two sub-trajectories, respectively.

3. Data and basic statistics

3.1. Data

We adopt ISI Web of Science (WOS) as the data source of this study. WOS is the world's leading citation database with multi-disciplinary coverage of over 10,000 high impact journals in science, social sciences, as well as international proceedings for over 120,000 conferences. Databases within WOS selected for this study are Science Citation Index Expanded (SCIE), Social Sciences Citation Index (SSCI), Conference Proceedings Index – Science (CPI-S), and Conference Proceedings Index – Social Science and Humanities (CPI-SSH). The data are retrieved in August, 2010 and the data time span is set to range from 1978 to 2010.

DEA papers are searched and retrieved from these sources with great care. The task begins with a query to the databases with properly defined keywords. These keywords are a collection of terms related to data envelopment analysis, including 'DEA', 'data envelopment analysis', 'Malmquist index', 'constant returns to scale', 'variable returns to scale', 'non-parametric efficiency', and 'Farrell efficiency', etc. Papers that contain any of these keywords in the title, abstract, author keyword, or Keywords Plus® fields are retrieved for further examination.

During the search, we found that many variations of a terminology were used by the DEA authors. To achieve a complete search, those variations were thoroughly explored; for example, we also query 'Malmquist indices' and 'Malmquist indexes' for 'Malmquist index'; and 'data envelopment model', 'data envelope analysis', and 'data enveloping analysis' for 'data envelopment analysis'. Another issue is the abbreviation. 'DEA' appears widely in scientific literature, but it has some 30 other meanings than 'data envelopment analysis.' These non-DEA papers were manually examined and excluded from the dataset. In the course of manual checking and screening, the stickiest problem is that there are papers having DEA in the Keywords Plus® field, but discuss nothing about DEA, probably because this field is a result of computer text-mining. Those papers which mention DEA merely for reference purposes are mistakenly marked in the Keywords Plus® field. For these cases, we conducted a partition analysis on the citation network to find out the outliers and then removed them from the dataset.

The last issue is that several important articles in the earlier period are missing from the WOS database. We refer to several review papers [2,3,14] and recover some highly cited articles back into the dataset. In the end, 4936 papers were included in the final dataset for further analysis. Among them, 3503 are articles, 1225 are proceedings papers, 55 are editorial materials, 153 are other document types; and 4848 are English documents, and 88 are in other languages.

3.2. Researcher statistics

Many researchers have contributed to the DEA field during its grand development. We apply both the *g*-index and the *h*-index to recognize individual DEA researchers' contribution and influence. Table 1 lists the top 20 DEA authors in order according to their *g*-index. The *h*-index ranking of these authors is also presented. As can be seen from the table, the two rankings are highly correlated. Cooper, Banker, Charnes, Seiford, Grosskopf, and Färe are the top six researchers. The list poses no surprise, except that Charnes is not the top author as everyone would expect, which is most likely caused by three reasons. First and most significant, the *g*-index does not credit the extremely high citation of the paper by Charnes et al. [1]. Second, Charnes has relatively short participation (1978–1997), such that he has published fewer DEA papers than the other pioneers. Third, some of his publications are not included in the WOS database. Nevertheless, we would like to emphasize that in our dataset two papers have extremely high citation in comparison with the other DEA papers. They are the

Table 1
Top 20 DEA researchers according to their *g*-index.

g-Index ranking	*h*-Index ranking	Authors	*g*-Index	*h*-Index	Years active	Total number of papers
1	1	Cooper, WW	82	30	1978–2009	82
2	4	Banker, RD	43	22	1980–2010	43
3	2	Charnes, A	42	25	1978–1997	42
4	5	Seiford, LM	42	22	1982–2009	42
5	3	Grosskopf, S	41	23	1983–2010	69
6	6	Färe, R	40	22	1978–2010	79
7	9	Lovell, CAK	33	17	1978–2007	40
8	10	Thanassoulis, E	30	16	1985–2010	45
9	7	Zhu, J	29	18	1995–2010	70
10	12	Simar, L	29	15	1995–2010	29
11	13	Cook, WD	27	15	1985–2010	63
12	15	Thrall, RM	27	14	1986–2004	27
13	8	Sueyoshi, T	26	18	1986–2010	58
14	11	Golany, B	26	16	1985–2008	26
15	14	Wilson, PW	26	15	1993–2009	26
16	16	Dyson, RG	22	13	1985–2010	22
17	17	Talluri, S	21	13	1997–2007	22
18	18	Athanassopoulos, AD	20	13	1995–2004	23
19	19	Pastor, JT	19	12	1995–2010	25
20	22	Forsund, FR	19	9	1979–2010	22

Note: The authors are listed in the order according to their *g*-index followed by *h*-index and the total number of articles.

Charnes et al. [1] at 2717 citations and Banker et al. [26] at 1468 citations as of August 2010.

3.3. Journal statistics

We apply again the *g*-index and the *h*-index to identify the influential journals that publish DEA papers. Table 2 presents the top 20 journals according to their *g*-index. *European Journal of Operational Research* ranks number one. It is followed by *Management Science*, *Journal of Productivity Analysis*, *Journal of the Operational Research Society*, *Annals of Operations Research*, *Journal of Econometrics,* and *Omega-International Journal of Management*. Some of these journals published several of the highly cited early DEA papers. We would like to make a note that the WOS database kept information on the journal *Socio-Economic Planning Sciences* only up to 1996. The contribution of this journal and the authors who published DEA articles in it are not fully recognized in our study.

There are many journals that are especially supportive to the DEA field. From Table 2 one can find that there are five journals that have published more than 100 DEA papers. They are: *European Journal of Operational Research*, *Journal of Productivity Analysis*, *Journal of the Operational Research Society*, *Applied Economics*, and *OMEGA-International Journal of Management Science*, listed in the order of the number of DEA papers published in them. The rank order of these journals is quite similar to the analysis result of Emrouznejad et al. [3]. A notable difference is that OMEGA is ranked number five rather than number seven as reported earlier. All these five journals have a steady record of publishing DEA papers. One highly influential journal, *Management Science*, however, has a proportionally small amount of DEA papers in the recent 10 years.

3.4. Growth trend

DEA literature was and still is growing at a very fast speed, as is the number of contributing authors. Based on the data collected; there is a total of 216 unique contributors from 1978 to 1990. By 2009, this number increases to 4617. As for the papers, 225 DEA papers are published from 1978 to 1990. Up through 2009, the number becomes 4597. In 2009 alone, 728 new DEA papers are published. How long will this trend continue and what will be the size of the literature in the long run?

In order to answer these questions, we conduct a growth curve analysis on the accumulated number of DEA papers from 1978 to 2009. We use the Loglet Lab software [27], as it has a built-in algorithm to decompose the growth trajectory into two phases when that is proper. It is found that the growth of DEA fits better as a two-phase development rather than one. Fig. 2 presents the result of the analysis. Fig. 2(a) shows the fitted growth curve, while Fig. 2(b) displays it in decomposed form. The figure's upper-left corner lists the estimates of the growth limit, the midpoint, and the life cycle. The two numbers in the parenthesis are the results at the 90% confidence interval for each characteristic.

The first phase of the DEA literature growth began at 1978 and became saturated around 2009. The second phase started in 2001 and is expected to saturate some time in the 2020s. The growth limit of the first phase is 2170 with a 90% confidence interval bracketed in the range of 1925 and 2474. The growth limit of the second phase is estimated at 10,696 and bounded by 7360 and 14,051. Adding them up, the analysis predicts that in the long run the total number of DEA papers could become 12,866 (2170+10,696) and the high and low counts could be in the range of 9286 to 16,525 at the 90% confidence interval. In addition, the midpoint of the second phase is estimated to be in 2012. It should be pointed out that predicting the future through growth curve analysis is an extrapolation of the historical data. We apply the analysis only to provide a point of reference, but not to predict the exact future.

The most interesting result of this analysis is that DEA has been developed in two phases. Shortly after the year 2000 some new driving forces shifted DEA development to a new phase that has grown at a faster speed than the earlier development. One most likely driving force is the availability of DEA software tools. DEA involves solving a set of linear programming equations. Without proper tools it takes some efforts to calculate efficiency for those who are not familiar with linear programming techniques. Around the year 2000 DEA software tools became widely available. Some tools are even bundled with DEA textbooks – examples include DEAP [28], DEA-Solver [29], and DEA Excel Solver (now, DEAFrontier) [30]. They were first available for users in 1996, 2000, and 2003, respectively. These tools have made

Table 2
Top 20 most influential journals in the DEA field.

g-Index ranking	*h*-Index ranking	Journals	*g*-Index	*h*-Index	Years active	Total number of papers	Number of articles since 2000
1	1	European Journal of Operational Research	84	41	1978–2010	500	351
2	2	Management Science	68	35	1981–2008	68	11
3	3	Journal of Productivity Analysis	42	25	1991–2010	226	153
4	4	Journal of the Operational Research Society	38	23	1985–2010	210	145
5	7	Annals of Operations Research	34	16	1985–2010	94	46
6	8	Journal of Econometrics	32	16	1985–2008	32	12
7	5	OMEGA-International Journal of Management Science	29	19	1982–2010	119	76
8	9	Journal of Banking & Finance	24	13	1985–2010	40	32
9	6	Applied Economics	21	17	1985–2010	124	81
10	10	International Journal of Production Economics	21	13	1991–2010	73	51
11	11	Computers & Operations Research	21	13	1987–2010	62	39
12	20	Interfaces	18	8	1984–2008	21	8
13	29	International Journal of Systems Science	17	7	1987–2010	36	4
14	30	Operations Research	17	7	1986–2010	17	13
15	13	Socio-Economic Planning Sciences	16	11	1982–1996	16	0
16	15	Economics Letters	16	9	1988–2006	16	2
17	14	Transportation Research Part A-Policy and Practice	15	10	1992–2009	22	18
18	21	Computers & Industrial Engineering	15	8	1995–2010	24	19
19	12	Medical Care	14	12	1984–2002	14	1
20	31	Transportation Research Part E-Logistics and Transportation Review	14	7	1997–2010	23	22

Note: The journals are listed in the order according to their *g*-index followed by *h*-index and the total number of articles.

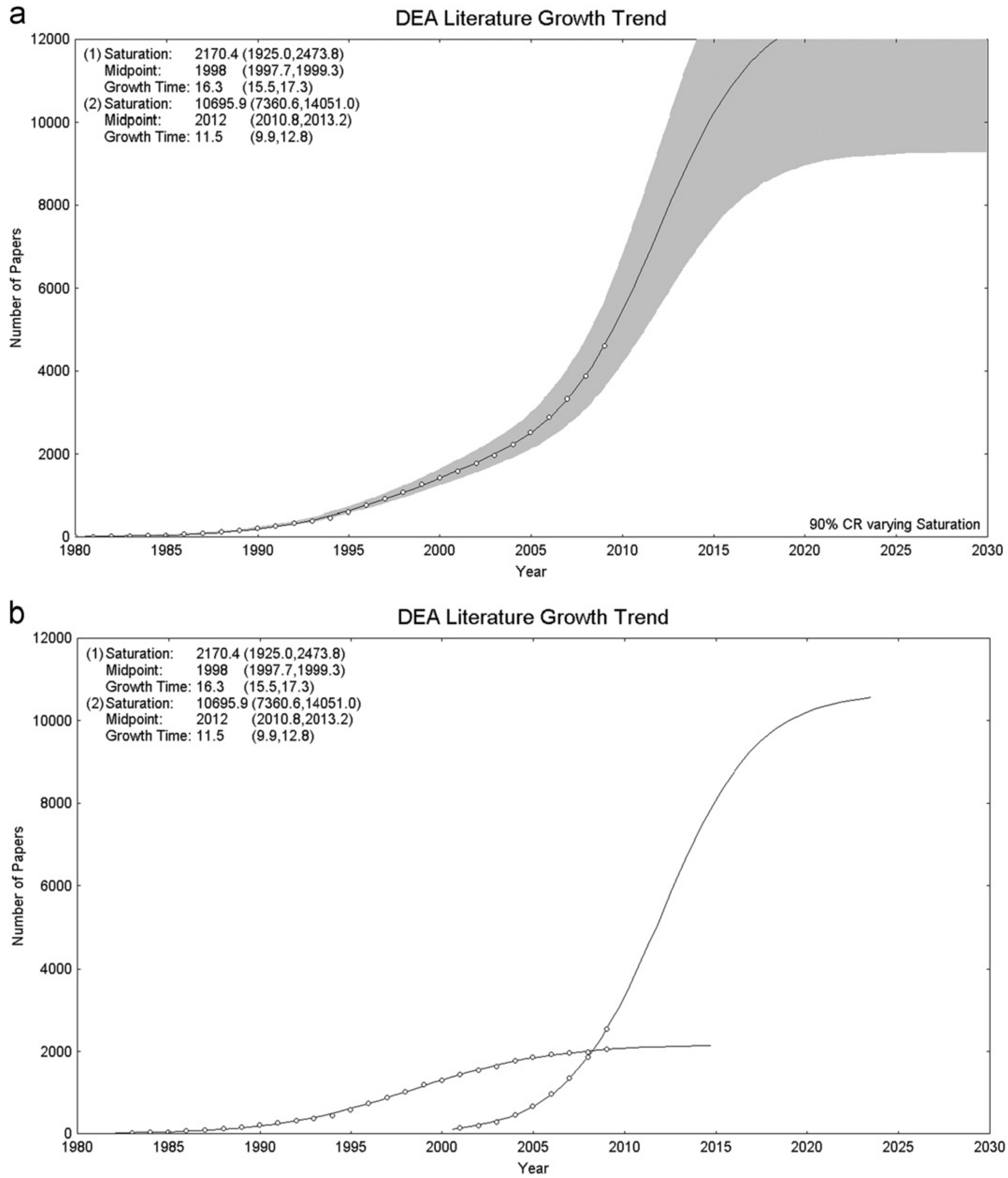

Fig. 2. (a) Growth curve of DEA literature. The solid curve in the middle is the direct estimate from the growth curve analysis. The boundaries of the shaded area enclose the 90% confidence interval; (b) growth curve of DEA literature decomposed into two phases.

efficiency calculation easy, thus removing the hurdle of entering into the DEA field.

4. Main paths

This section discusses several variations of the main paths in DEA development. These variations observe the knowledge diffusion from different angles. They complement each other, thus preventing us from overlooking important DEA papers and subareas.

4.1. Local main path

The local main path indicates the most significant knowledge route at each juncture of knowledge dissemination for a scientific discipline. This is the main path that follows the tradition of Hummon and Doreian [17] and Verspagen [21]. Fig. 3 presents the local main path for DEA development. The figure is drawn with the Pajek software [31]. In the figure, the arrow indicates the direction of knowledge flow, and the line thickness reflects the SPC value. The thicker the line is, the more significant the route is.

The local main path consists of 19 papers. It is quite clear in the figure that the first paper, CharnesCR1978 [1], is the origin of all the following papers. This work lays out the foundation of the DEA methodology. The proposed constant-returns-to-scale model is now commonly referred to as the CCR model. The second paper on the path, CharnesCR1979a [32], is a complement to the CharnesCR1978. It is a one-page note that modifies the constraints of the main formulation in CharnesCR1978. The following paper, CharnesCR1981 [33], applies the then new methodology to evaluate public education programs. It is labeled the "most significant breakthrough" [2] in advancing DEA application at that time. The fourth paper, BankerCC1984 [26], proposes the variable-returns-to-scale model, which is the now well-known

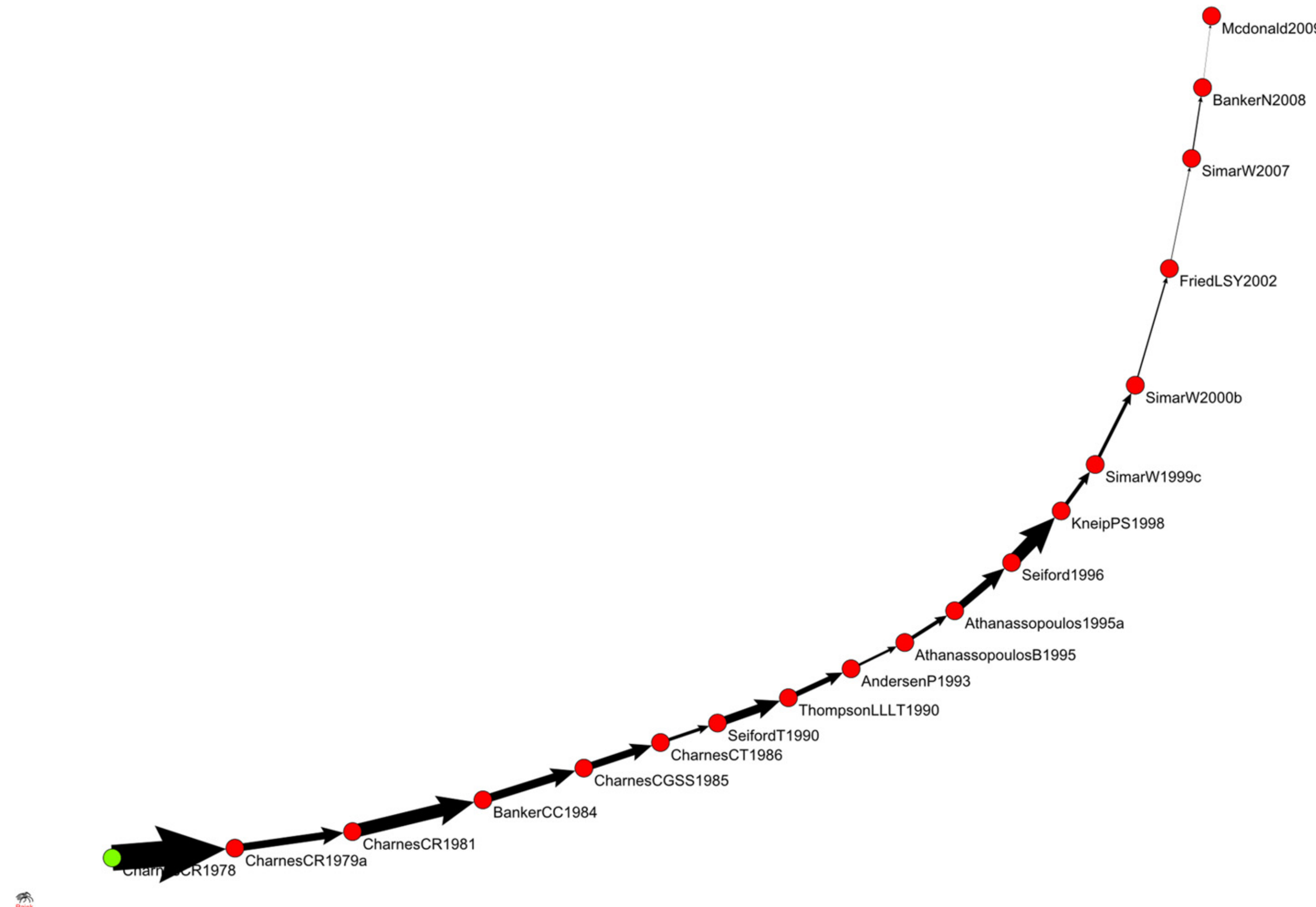

Fig. 3. Local main path of DEA development. Link weights are indicated with different line thickness. Thicker lines indicate heavier weights. The network is drawn with Pajek software.

BCC model. CharnesCGSS1985 [34] introduces an additive model and establishes DEA's link to production theory through analyzing the capabilities of Pareto–Koopmans production function. Seiford [2] suggests that CharnesCR1978, CharnesCR1981, BankerCC1984, and CharnesCGSS1985 are the four most influential DEA papers. The claim is now supported by the main path analysis.

CharnesCT1986 [35] partition DMU's efficiency into six classes. SeifordT1990 [12] review the DEA state of the art as of 1990. ThompsonLLLT1990 [36] define the concept of the assurance region and applies it to evaluate the efficiency of farms. AndersenP1993 [37] propose the super efficiency concept, which is one of the efforts attempting to discriminate efficient units. The subsequent two papers AthanassopoulosB1995 [38] and Athanassopoulos1995a [39] are two application-oriented works that focus on assessing market and sale efficiency in the retail industry. The next paper, Seiford1996 [2], is a review paper with significant impact. KneipPS1998 [40] discusses the consistency and the speed of convergence of DEA estimators under the multiple input and multiple output context.

Two papers afterwards, SimarW1999c [41] and SimarW2000b [42] are a series of works on the idea of bootstrapping DEA data, which makes statistical inference possible when applying DEA. FriedLSY2002 [43] propose a variation of the two-stage analysis that allows incorporating the effects of external factors into DEA-based performance evaluation. SimarW2007 [44] and BankerN2008 [45] both provide statistical foundation for the two-stage analysis approach. Each starts with a substantially different theoretical development and research design and obtains different conclusions. The last paper on the local main path, McDonald2009 [46], is also a work on two-stage analysis.

We have two remarkable observations for the papers on the main path. First, most of the papers on the main path are theoretical works, with the minority exceptions of CharnesCR1981, AthanassopoulosB1995, and Athanassopoulos1995a. Theoretical works are essential for any scientific discipline to have long lasting development, and DEA is no exception. These theoretical DEA works deserve recognition. Second, the SPC values and hence the significance of the routes towards the end of the main path are much less than the beginning routes. This is because the papers close to the tail of the main path usually have few citations. They are on the main path, because they have relatively high citation, as compared with the other papers that cite papers on the main path. In other words, they currently receive the most attention among the followers of mainstream works. The importance of these tail papers has yet to be verified in time.

There are undoubtedly many important works missing from this path, as the local main path selects only the route with the highest SPC value at every branching point. In the following section we adopt a different perspective in tracing the path in order to recognize more papers of significance in DEA development.

4.2. Global main path

The global main path is a development trajectory that holds the overall largest SPC values. Fig. 4 presents the global main path of DEA evolution. Table 3 summarizes the papers on both the local and global

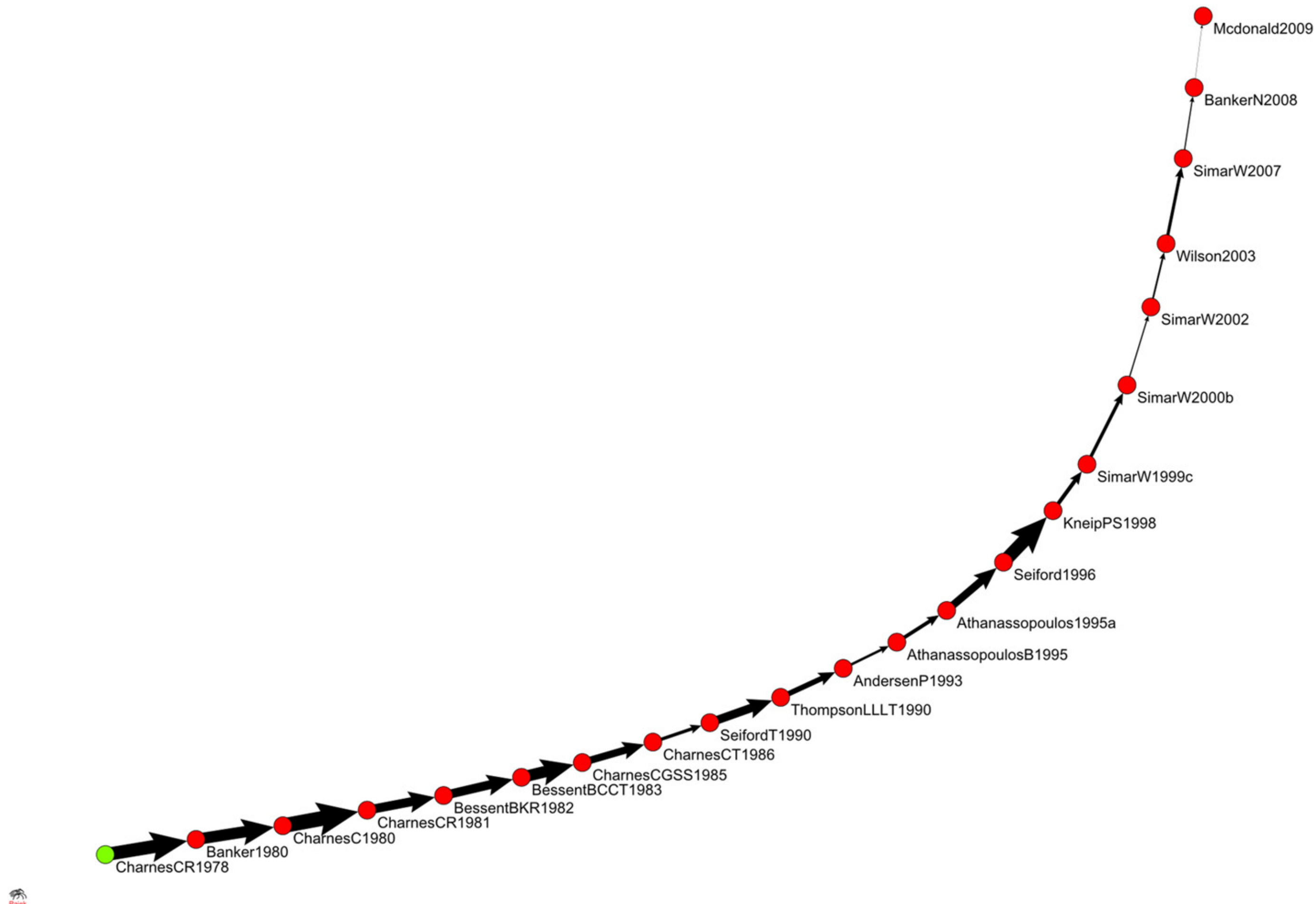

Fig. 4. Global main path of DEA development. Link weights are indicated with different line thickness. Thicker lines indicate heavier weights. The network is drawn with Pajek software.

main paths. As expected, many of the works on the global main path, exactly 16, are the same as those on the local main path. These works are labeled Local/Global in Table 3. The resemblance indicates that these works no doubt have great significance to DEA development.

Six additional works are uncovered from the global viewing angle. Banker1980 [47] reestablishes the CCR model through a game theoretical approach. CharnesC1980 [48] present an early application work that precedes the well-known application paper, CharnesCR1981. BessentBKR1982 [49] and BessentBCCT1983 [50] are application papers focusing on educational applications. SimarW2002 [51] discuss various statistics for testing the hypothesis regarding returns to scale. Wilson2003 [52] provides a method to test if homogeneous bootstrap or heterogeneous bootstrap should be used for inference in non-parametric models.

The global main path indicates that efficiency evaluation of educational programs and organizations are at the center of early DEA development, as CharnesC1980, BessentBKR1982, and BessentCCT1983 can be categorized under such applications. Another observation is that the last seven papers on the global main path are all statistical works. This suggests that studying DEA methodology from the statistical aspect is the focus of attention in the decades since the year 2000. Daraio and Simar [53] provide the details in this subarea.

4.3. Latest DEA development activities

The global main path method provides the ground work for us to look at the multiple main paths and to examine DEA developments in more detail. By selecting paths with smaller overall SPC values, one is able to see paths other than the most significant one. As the number of paths selected increases, the details of the citation network surface little by little. The effect is like zooming in with a variable-focal-length camera lens. The number of paths selected can be arbitrary, depending on how detailed the network that one is interested in inspecting. The topics of papers surfacing this way provide a good indication on the recent major DEA activities. We thus are able to identify recent DEA active subareas through a proper selection of the number of multiple paths. In this study we gradually increase the number of paths in order to discover proper streams of DEA papers. For the discussion in this section, the number of paths is set at 200, as increasing the number beyond the number uncovers no more significant branches.

Fig. 5 presents the result of multiple global main path analysis. The ‘sink’ papers are shown in a dark gray level. In Fig. 5, along the main path, there are five major branches of literature, or in other words, five active DEA subareas. After examining the title, abstract, and keywords of these papers, these five subareas are identified to be “two-stage contextual factor evaluation framework”, “extending models”, “handling special data”, “examining the internal structure”, and “measuring environmental performance.” The branches that have their details surface earlier before the others can be regarded as more active. Among the five branches of literature, “two-stage contextual factor evaluation framework” is the subarea that surfaces first and thus can be considered as relatively more active. The papers in the first four subareas are mostly studies of theoretical oriention and the

Table 3
Papers on the main paths.

		Authors	Title	Journal	Total citations	Year published
Local/Global	CharnesCR1978	Charnes, A; Cooper, WW; Rhodes, E	Measuring the efficiency of decision-making units	European Journal of Operational Research	2717	1978
Local	CharnesCR1979b	Charnes, A; Cooper, WW; Rhodes, E	Measuring the efficiency of decision-making units (short communication)	European Journal of Operational Research	171	1979
Global	Banker1980	Banker, RD	A game theoretic approach to measuring efficiency	European Journal of Operational Research	19	1980
Global	CharnesC1980	Charnes, A; Cooper, WW	Auditing and accounting for program efficiency and management efficiency in not-for-profit entities	Accounting, Organizations and Society	25	1980
Local/Global	CharnesCR1981	Charnes, A; Cooper, WW; Rhodes, E	Evaluating program and managerial efficiency – an application of data envelopment analysis to Program Follow Through	Management Science	319	1981
Global	BessentBKR1982	Bessent, A; Bessent, W; Kennington, J; Reagan, B	An application of mathematical-programming to assess productivity in the Houston independent school-district	Management Science	77	1982
Global	BessentBCCT1983	Bessent, A; Bessent, W; Charnes, A; Cooper, WW; Thorogood, NC	Evaluation of educational-program proposals by means of DEA	Educational Administration Quarterly	38	1983
Local	BankerCC1984	Banker, Rd; Charnes, A; Cooper, WW	Some models for estimating technical and scale inefficiencies in data envelopment analysis	Management Science	1468	1984
Local/Global	CharnesCGSS1985	Charnes, A; Cooper, WW; Golany, B; Seiford, L; Stutz, J	Foundations of data envelopment analysis for Pareto–Koopmans efficient empirical production-functions	Journal of Econometrics	316	1985
Local/Global	CharnesCT1986	Charnes, A; Cooper, WW; Thrall, RM	Classifying and characterizing efficiencies and inefficiencies in data envelopment analysis	Operations Research Letters	62	1986
Local/Global	SeiforT1990	Seiford, LM; Thrall, RM	Recent developments in DEA – The mathematical-programming approach to frontier analysis	Journal of Econometrics	355	1990
Local/Global	ThompsonLLLT1990	Thompson, RG; Langemeier, LN; Lee, CT; Lee, E; Thrall, RM	The role of multiplier bounds in efficiency analysis with application to Kansas farming	Journal of Econometrics	203	1990
Local/Global	AndersenP1993	Andersen, P; Petersen, NC	A procedure for ranking efficient units in data envelopment analysis	Management Science	393	1993
Local/Global	AthanassopoulosB1995	Athanassopoulos, AD; Ballantine, JA	Ratio frontier analysis for assessing corporate performance – evidence from the grocery industry in the UK	Journal of The Operational Research Society	21	1995
Local/Global	Athanassopoulos1995a	Athanassopoulos, AD	Performance improvement decision aid systems (PIDAS) in retailing organizations using data envelopment analysis	Journal of Productivity Analysis	10	1995
Local/Global	Seiford1996	Seiford, LM	Data envelopment analysis: The evolution of the state of the art (1978–1995)	Journal of Productivity Analysis	198	1996
Local/Global	KneipPS1998	Kneip, A; Park, BU; Simar, L	A note on the convergence of nonparametric DEA estimators for production efficiency scores	Econometric Theory	71	1998
Local/Global	SimarW1999c	Simar, L; Wilson, PW	Some problems with the Ferrier/Hirschberg bootstrap idea	Journal of Productivity Analysis	15	1999
Local/Global	SimarW2000b	Simar, L; Wilson, PW	Statistical inference in nonparametric frontier models: The state of the art	Journal of Productivity Analysis	145	2000
Local	FriedLSY2002	Fried, HO; Lovell, CAK; Schmidt, SS; Yaisawarng, S	Accounting for environmental effects and statistical noise in data envelopment analysis	Journal of Productivity Analysis	52	2002
Global	SimarW2002	Simar, L; Wilson, PW	Non-parametric tests of returns to scale	European Journal of Operational Research	25	2002

Table 3 (*continued*)

		Authors	Title	Journal	Total citations	Year published
Global	Wilson2003	Wilson, PW	Testing independence in models of productive efficiency	Journal Of Productivity Analysis	3	2004
Local/Global	SimarW2007	Simar, L; Wilson, PW	Estimation and inference in two-stage, semi-parametric models of production processes	Journal of Econometrics	102	2007
Local/Global	BankerN2008	Banker, RD; Natarajan, R	Evaluating contextual variables affecting productivity using data envelopment analysis	Operations Research	17	2008
Local/Global	McDonald2009	McDonald, J	Using least squares and tobit in second stage DEA efficiency analyses	European Journal of Operational Research	4	2009

papers in the last subarea are basically application works. In the following discussion of important and potential works in each subarea, we focus on the theoretical aspects of the DEA development and elaborate only upon the first four active subareas.

4.3.1. *Two-stage contextual factor evaluation framework*

There are many DEA studies that evaluate the effect of contextual variables on production efficiency through a two-stage procedure. A typical two-stage study first obtains efficiency scores through DEA and then correlates these scores with various contextual factors either by ordinary least squares (OLS), Tobit regression analysis, or maximum likelihood estimation (MLE). There is, however, no theoretical justification for the statistical validity for such method. SimarW2007 and BankerN2008, the two papers mentioned earlier in the main path discussion, independently provide a statistical foundation for the approach.

SimarW2007, in particular, spawn many new works as seen from the explosive pattern surrounding the paper in Fig. 5. These works can be further categorized into three groups. The first group consists of empirical works that apply the methodology. For example, Latruffe et al. [54] and Barros and Dieke [55] take the two-stage procedure of SimarW2007 to industrial settings such as farms and airports. The second group either extends or modifies the method proposed in SimarW2007 and earlier bootstrap works. Examples include Daraio and Simar [56], Johnson and McGinnis [57], and Balcombe et al. [58]. The third group contains works partially inspired by the concept mentioned in SimarW2007, but the focuses are not on the two-stage procedure.

4.3.2. *Extending models*

This branch of literature includes a group of works extending the existing models that deal with assurance regions on multipliers and with flexible variables. The concept of assurance region has widespread usage in DEA. It restricts the upper and lower bounds of multipliers to a relatively proper size such that unacceptable efficiency scores can be avoided. The original work of Thompson et al. [36] imposes uniform restrictions across all DMUs. Cook and Zhu [59] extend the model so that multiple sets of restrictions can be applied to reflect the context for each subset of DMUs.

Cook et al. [60] and Cook and Zhu [61] improve existing DEA models to handle the case where factors simultaneously play both input and output roles. Thus, the ambiguous role of factors such as ‘research funding’ in evaluating university performance can be clarified. Two review articles, Cooper et al. [13] and Cook and Seiford [14], mentioned in an earlier section review the recent development of DEA models in great detail.

4.3.3. *Handling special types of data*

The classical DEA models assume that all data have specific and positive numerical values. This may not be the case in some real life applications. Data can be bounded, ordinal, qualitative, negative, fuzzy, etc. Various models and methods are developed to deal with such types of data. Cook et al. [62, 63] first incorporate rank order data within the DEA structure. Cooper et al. [64] develop the imprecise DEA (IDEA) model to handle applications with interval or ordinal data. Zhu and Cook [65] present detailed descriptions of all these types of models and methods.

The latest development in this subarea includes some new approaches, mostly by Zhu and his colleagues [66, 67, 68, 69]. Zhu [66] discusses an approach that converts imprecise data into exact data, thus allowing the use of the standard linear DEA model. This is in comparison to the approach of scale transformations and variable alternations that convert the non-linear IDEA

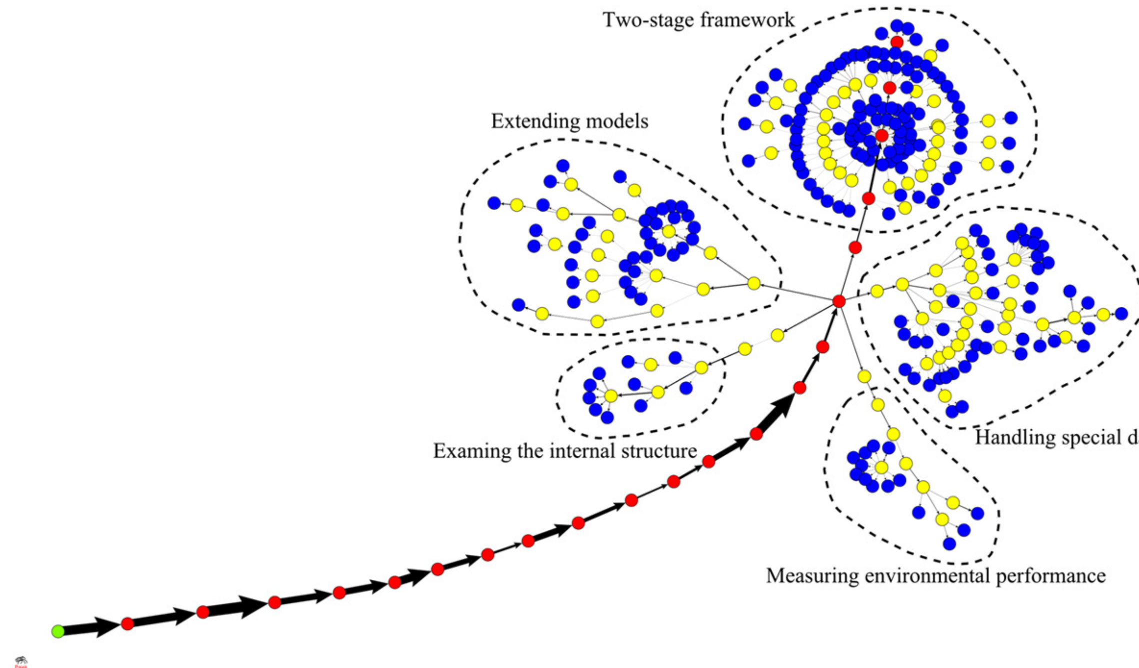

Fig. 5. Multiple global main paths of DEA development. Darker dots indicate end nodes. Link weights are indicated with different line thickness. Thicker lines indicate heavier weights. The network is drawn with Pajek software.

model into a linear program. Cook and Zhu [69] develop a unified structure for embedding rank order or Likert scale data into the DEA framework. Wang et al. [70] propose a general model to deal with interval, ordinal, and fuzzy data. Portela et al. [71] propose a range directional model (RDM) to handle situations with negative data.

Data can also be random in nature. Land et al. [72] adapt chance constrained programming to DEA to deal with such data. New research studies for this topic include Cooper et al. [73, 74], where they introduce chance constrained models to handle technical inefficiencies and congestions in stochastic situations. Cooper et al. [75] provide an overview for this topic.

4.3.4. *Examining the internal structure of DMUs*

In the early development of DEA, the internal structure of the DMU was not an issue. It was viewed as a black box. Färe and Grosskopf [76] propose a network DEA model to allow the examination of the inner workings of the ‘black box’. The model treats the process under study as several interconnected sub-processes and looks for efficiencies in each process by solving all the efficiency equations as a whole. Many variations of the concept are suggested thereafter, mostly under the label of a network DEA, multilevel models, and two-stage DEA.

Two-stage DEA addresses the simpler case where there are only two sub-processes, and outputs of the first stage are the only inputs to the second stage. This two-stage process should not be confused with that mentioned in Section 4.3.1, where production efficiency is evaluated through a ‘two-stage’ procedure. Chen and Zhu [77], Kao and Hwang [78], and Chen et al. [79, 80] propose a variety of models under different returns-to-scale assumptions. Chen et al. [81] discuss the correspondence of two of the models. Liang et al. [82] propose a game-theoretic approach.

The network DEA model involves two or more sub-processes and more complicated interconnections among sub-processes. Lewis and Sexton [83], Yu and Lin [84], Avkiran [9], Chen et al. [85], Liang et al. [86], and Cook et al. [87] are examples of theoretical and empirical works on the subject. Kao [88] develops a model that treats the process as a series of sub-processes, yet each sub-process can be divided into a parallel structure. Tone [89] extends the network DEA model to the slacks-based measure framework. Dynamic DEA is an idea similar to network DEA in which the processes are interconnected in time [90]. The latest development in dynamic DEA includes the works of Chen [91] and Tone [92].

This subarea is relatively active in recent years. A more detailed literature survey of this subarea can be found in Cook and Seiford [14], Cook et al. [93], and Casstelli et al. [94].

5. Conclusion

The strong growth of DEA research in recent years has increased the DEA literature to a scale in which it is not easy to conduct a general review without quantitative methodologies. We survey the DEA literature with the assistance of the main path method. The method is quantitative and citation based. It helps identifies significant paths, important papers, and recent active subareas in DEA development. The method first assigns a search path count to each citation and then traces the paths with the largest search path counts. Search path count is the exhausted count of the routes for knowledge in all the sources to disseminate to all the sinks. The local, global, and multiple main paths are examined. Each of them provides us with different views on the DEA evolution.

From the local main path, we find support for the claim that Charnes et al. [1], Charnes et al. [33], Banker et al. [26], and Charnes et al. [34] are the four most influential papers in DEA development. The global main path indicates that measuring the efficiency of educational institutions was the focus of attention on practical applications in early DEA development, and that the statistical aspect plays an important role in recent decades. The multiple main paths suggest five recent active DEA sub-areas. Among them, “two-stage contextual factor evaluation framework” attracts the most attention.

There are several limitations to this study. First, the dataset is taken from the WOS database. Although it is the largest citation-based academic database available, there are, however, some DEA papers published in journals not included in the WOS. Presentation and interpretation of the results should be accompanied by a warning on the limitation of the data source. Second, albeit much effort has been made to select correct DEA papers from the database, two situations may still exist: missing DEA papers and an incorrect inclusion of non-DEA papers. We believe that these papers are a very small percentage of the total papers and do not change the major analysis results. Third, a situation we call 'remote' citation occurs occasionally when a paper cites others, not because of a close connection with the main subject, but merely because of a connection in a broad sense such as the same application area, the same general method, or even just because of applying DEA methodology. Citations of this type are noises and may cause the true main paths to be surpassed by the noise paths. The tail portions of the multiple paths are especially sensitive to these noises as the number of citation count is few there. Thus, one needs to be more careful in interpreting the results close to the tail. Another way to overcome this issue is taking more multiple paths to let more of the true paths appear.

The main contribution of this study is two-fold. First, we present the DEA development scenario from a perspective different from previous studies. The main DEA development path is presented the first time in the DEA literature. Second, we demonstrate a novel way of analyzing an academic discipline through citation data. The proposed multiple path method complements and increases the value of the traditional main path methodology.

Acknowledgement

The authors would like to thank two anonymous reviewers and the editors of this special issue for their constructive comments that have much improved the accuracy and readability of this article. This work is partially supported by Taiwan's National Science Council grant NSC 98-2410-H-155-037-MY2 and NSC 98-2410-H-011-010.

References

[1] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European Journal of Operational Research 1978;2(6):429–44.
[2] Seiford LM. Data envelopment analysis: the evolution of the state of the art (1978–1995). Journal of Productivity Analysis 1996;7(2–3):99–137.
[3] Emrouznejad A, Parker BR, Tavares G. Evaluation of research in efficiency and productivity: a survey and analysis of the first 30 years of scholarly literature in DEA. Socio-Economic Planning Sciences 2008;42(3):151–7.
[4] Avkiran NK. Association of DEA super-efficiency estimates with financial ratios: investigating the case for Chinese banks. OMEGA-International Journal of Management Science 2011;39(3):323–34.
[5] Zhong W, Yuan W, Li SX, Huang Z. The performance evaluation of regional R&D investments in China: an application of DEA-based on the first official China economic census data. OMEGA-International Journal of Management Science 2011;39(4):447–55.
[6] Siriopoulos C, Tziogkidis P. How do Greek banking institutions react after significant events? – a DEA approach OMEGA-International Journal of Management Science 2010;38(5):294–308.
[7] Demirbag M, Tatoglu E, Glaister KW, Zaim S. Measuring strategic decision making efficiency in different country contexts: a comparison of British and Turkish firms. OMEGA – International Journal of Management Science 2010;38(1–2):95–104.
[8] Odeck J. Statistical precision of DEA and Malmquist indices: a bootstrap application to Norwegian grain producers. OMEGA – International Journal of Management Science 2009;37(5):1007–17.
[9] Avkiran NK. Opening the black box of efficiency analysis: an illustration with UAE banks. OMEGA – International Journal of Management Science 2009;37(4):930–41.
[10] Seiford LM. A bibliography for data envelopment analysis (1978–1996). Annals of Operations Research 1997;73:393–438.
[11] Gattoufi S, Oral M, Reisman A. Data envelopment analysis literature: a bibliography update (1951–2001). Socio-Economic Planning Sciences 2004;38:159–229.
[12] Seiford LM, Thrall RM. Recent developments in DEA – the mathematical-programming approach to frontier analysis. Journal of Econometrics 1990;46(1–2):7–38.
[13] Cooper WW, Seiford LM, Tone K, Zhu J. Some models and measures for evaluating performances with DEA: past accomplishments and future prospects. Journal of Productivity Analysis 2007;28(3):151–63.
[14] Cook WD, Seiford LM. Data envelopment analysis (DEA) – thirty years on. European Journal of Operational Research 2009;192(1):1–17.
[15] Gattoufi S, Oral M, Kumar A, Reisman A. Epistemology of data envelopment analysis and comparison with other fields of OR/MS for relevance to applications. Socio-Economic Planning Sciences 2004;38(2–3):123–40.
[16] Garfield E. Is citation analysis a legitimate evaluation tool? Scientometrics 1979;1(4):359–75.
[17] Hummon NP, Doreain P. Connectivity in a citation network: the development of DNA theory. Social Networks 1989;11(1):39–63.
[18] Egghe L. Theory and practice of the *g*-index. Scientometrics 2006;69(1):131–52.
[19] Hirsch JE. An index to quantify an individual's scientific research output. PNAS 2005;102(46):16569–72.
[20] Batagelj V. Efficient Algorithms for Citation Network Analysis. University of Ljubljana, Institute of Mathematics, Physics and Mechanics, Preprint series; 2003. 41, 897.
[21] Verspagen B. Mapping technological trajectories as patent citation networks: a study on the history of fuel cell research. Advances in Complex Systems 2007;10(1):93–115.
[22] Bornmann L, Daniel H-D. Does the *h*-index for ranking of scientists really work? Scientometrics 2005;65(3):391–2.
[23] Mingers J. Measuring the research contribution of management academics using the Hirsch-index. Journal of the Operational Research Society 2009;60(9):1143–53.
[24] Saad G. Applying the *h*-index in exploring bibliometric properties of elite marketing scholars. Scientometrics 2010;83(2):423–33.
[25] Meyer PS, Yung JW, Ausubel JH. A primer on logistics growth and substitution: the mathematics of the Loglet Lab software. Technological Forecasting and Social Change 1999;61(3):247–71.
[26] Banker RD, Charnes A, Cooper WW. Models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 1984;30(9):1078–92.
[27] Yung JW, Meyer PS, Ausubel JH. The Loglet Lab software: a tutorial. Technological Forecasting and Social Change 1999;61(3):273–95.
[28] Coelli TJ, Rao DSP, O'Donnell CJ, Battese GE. An Introduction to Efficiency and Productivity Analysis. 2nd edn. New York: Springer; 2005. Science+Business Media.
[29] Cooper WW, Seiford LM, Tone K. Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software. 2nd edn. New York: Springer; 2007. Science+Business Media.
[30] Zhu J. Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets. 2nd edn. New York: Springer; 2009. Science+Business Media.
[31] Batagelj V, Mrvar A. Pajek – program for large network analysis. Connections 1998;21(2):47–57.
[32] Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. European Journal of Operational Research 1979;3(4):339.
[33] Charnes A, Cooper WW, Rhodes E. Evaluating program and managerial efficiency – an application of data envelopment analysis to Program Follow Through. Management Science 1981;27(6):668–97.
[34] Charnes A, Cooper WW, Golany B, Seiford LM, Stutz J. Foundations of data envelopment analysis for Pareto–Koopmans efficient empirical production-functions. Journal of Econometrics 1985;30(1–2):91–107.
[35] Charnes A, Cooper WW, Thrall RM. Classifying and characterizing efficiencies and inefficiencies in data envelopment analysis. Operations Research Letters 1986;5(3):105–10.
[36] Thompson RG, Langemeier LN, Lee CT, Lee E, Thrall RM. The role of multiplier bounds in efficiency analysis with application to Kansas farming. Journal of Economitrics 1990;46(1–2):93–108.
[37] Andersen P, Petersen NC. A procedure for ranking efficient units in data envelopment analysis. Management Science 1993;39(10):1261–5.
[38] Athanassopoulos AD, Ballantine JA. Ratio frontier analysis for assessing corporate performance – evidence from the grocery industry in the UK. Journal of The Operational Research Society 1995;46(4):427–40.
[39] Athanassopoulos AD. Performance improvement decision aid systems (PIDAS) in retailing organizations using data envelopment analysis. Journal of Productivity Analysis 1995;6(2):153–70.
[40] Kneip A, Park BU, Simar L. A note on the convergence of nonparametric DEA estimators for production efficiency scores. Econometric Theory 1998;14(6):783–93.
[41] Simar L, Wilson PW. Some problems with the Ferrier/Hirschberg bootstrap idea. Journal of Productivity Analysis 1999;11(1):67–80.
[42] Simar L, Wilson PW. Statistical inference in nonparametric frontier models: the state of the art. Journal of Productivity Analysis 2000;13(1):49–78.
[43] Fried HO, Lovell CAK, Schmidt SS, Yaisawarng S. Accounting for environmental effects and statistical noise in data envelopment analysis. Journal of Productivity Analysis 2002;17(1–2):157–74.
[44] Simar L, Wilson PW. Estimation and inference in two-stage, semi-parametric models of production processes. Journal of Econometrics 2007;136(1):31–64.
[45] Banker RD, Natarajan R. Evaluating contextual variables affecting productivity using data envelopment analysis. Operations Research 2008;56(1):48–58.

[46] McDonald J. Using least squares and tobit in second stage DEA efficiency analyses. European Journal of Operational Research 2009;197(2):792–8.
[47] Banker RD. A game theoretic approach to measuring efficiency. European Journal of Operational Research 1980;5(4):262–6.
[48] Charnes A, Cooper WW. Auditing and accounting for program efficiency and management efficiency in not-for-profit entities. Accounting, Organizations and Society 1980;5(1):87–107.
[49] Bessent A, Bessent W, Kennington J, Reagan B. An application of mathematical-programming to assess productivity in the Houston independent school-district. Management Science 1982;28(12):1355–67.
[50] Bessent A, Bessent W, Charnes A, Cooper WW, Thorogood NC. Evaluation of educational-program proposals by means of DEA. Educational Administration Quarterly 1983;19(2):82–107.
[51] Simar L, Wilson PW. Non-parametric tests of returns to scale. European Journal of Operational Research 2002;139(1):115–32.
[52] Wilson PW. Testing independence in models of productive efficiency. Journal of Productivity Analysis 2003;20(3):361–90.
[53] Daraio C, Simar L. Advanced Robust and Nonparametric Methods in Efficiency Analysis. New York: Springer; 2007. Science+Business Media.
[54] Latruffe L, Davidova S, Balcombe K. Application of a double bootstrap to investigation of determinants of technical efficiency of farms in Central Europe. Journal of Productivity Analysis 2008;29(2):183–91.
[55] Barros CP. Dieke PUC. Measuring the economic efficiency of airports: a Simar–Wilson methodology analysis. Transportation Research Part E-Logistics and Transportation Review 2008;44(6):1039–51.
[56] Daraio C, Simar L. Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. Journal of Productivity Analysis 2007;28(1–2):13–32.
[57] Johnson AL, McGinnis LF. Outlier detection in two-stage semiparametric DEA models. European Journal of Operational Research 2008;187(2):629–35.
[58] Balcombe K, Davidova S, Latruffe L. The use of bootstrapped Malmquist indices to reassess productivity change findings: an application to a sample of Polish farms. Applied Economics 2008;40(16):2055–61.
[59] Cook WD, Zhu J. CAR-DEA: context-dependent assurance regions in DEA. Operations Research 2008;56(1):69–78.
[60] Cook WD, Green RH, Zhu J. Dual-role factors in DEA. IIE Transactions 2006;38(2):105–15.
[61] Cook WD, Zhu J. Classifying inputs and outputs in data envelopment analysis. European Journal of Operational Research 2007;180:692–9.
[62] Cook WD, Kress M, Seiford L. On the use of ordinal data in data envelopment analysis. Journal of the Operational Research Society 1993;44(2):133–40.
[63] Cook WD, Kress M, Seiford L. Data envelopment analysis in the presence of both quantitative and qualitative factors. Journal of the Operational Research Society 1996;47:945–53.
[64] Cooper WW, Park KS, Yu G. IDEA and AR-IDEA: models for dealing with imprecise data in DEA. Management Science 1999;45(4):597–607.
[65] Zhu J, Cook WD. Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. New York: Springer; 2007. Science+Business Media.
[66] Zhu J. Imprecise data envelopment analysis (IDEA): a review and improvement with an application. European Journal of Operational Research 2003;144(3):513–29.
[67] Zhu J. Efficiency evaluation with strong ordinal input and output measures. European Journal of Operational Research 2003;146(3):477–85.
[68] Zhu J. Imprecise DEA via standard linear DEA models with a revisit to a Korean mobile telecommunication company. Operations Research 2004;52(2): 323–9.
[69] Cook WD, Zhu J. Rank order data in DEA: a general framework. European Journal of Operational Research 2006;174(2):1021–38.
[70] Wang YM, Greatbanks R, Yang JB. Interval efficiency assessment using data envelopment analysis. Fuzzy Sets and Systems 2005;153(3):347–70.
[71] Portela MCAS Thanassoulis E, Simpson G. Negative data in DEA: a directional distance approach applied to bank branches. Journal of the Operational Research Society 2004;55(10):1111–21.
[72] Land KC, Lovell CAK, Thore S. Chance-constrained data envelopment analysis. Managerial and Decision Economics 1993;14(6):541–54.
[73] Cooper WW, Deng H, Huang Z, Li SX. Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. Journal of the Operational Research Society 2002;53(12): 1347–56.
[74] Cooper WW, Deng H, Huang Z, Li SX. A chance constrained programming approach to congestion in stochastic data envelopment analysis. European Journal of Operational Research 2002;53:1–10.
[75] Cooper WW, Huang Z, Li S. Chance constraint DEA. In: Cooper WW, Seiford LM, Zhu J, editors. Handbook on Data Envelopment Analysis. Norwell MA: Kluwer Academic Publishers; 2004.
[76] Färe R, Grosskopf S. Network DEA. Socio-Economic Planning Sciences 2000;34(1):35–49.
[77] Chen Y, Zhu J. Measuring information technology's indirect impact on firm performance. Information Technology & Management Journal 2004;5(1–2): 9–22.
[78] Kao C, Hwang SN. Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. European Journal of Operational Research 2008;185(1):418–42.
[79] Chen Y, Cook WD, Li N, Zhu J. Additive efficiency decomposition in two-stage DEA. European Journal of Operational Research 2009;196(3):1170–6.
[80] Chen Y, Cook WD, Zhu J. Deriving the DEA frontier for two-stage processes. European Journal of Operational Research 2010;202(1):138–42.
[81] Chen Y, Liang L, Zhu J. Equivalence in two-stage DEA approaches. European Journal of Operational Research 2009;193(2):600–4.
[82] Liang L, Cook WD, Zhu J. DEA models for two-stage processes: game approach and efficiency decomposition. Naval Research Logistics 2008;55:643–53.
[83] Lewis H, Sexton T. Network DEA: efficiency analysis of organizations with complex internal structure. Computers and Operations Research 2004;31(9): 1365–410.
[84] Yu MM, ETJ Lin. Efficiency and effectiveness in railway performance using a multi-activity network DEA model. OMEGA-International Journal of Management Science 2008;36(6):1005–17.
[85] Chen Y, Du J, Sherman HD, Zhu J. DEA model with shared resources and efficiency decomposition. European Journal of Operational Research 2010;207(1):339–49.
[86] Liang L, Li ZQ, Cook WD, Zhu J. Data envelopment analysis efficiency in two-stage networks with feedback. IIE Transactions 2011;43(5):309–22.
[87] Cook WD, Zhu J, Bi G, Yang F. Network DEA: additive efficiency decomposition. European Journal of Operational Research 2010;207(2):1122–9.
[88] Kao C. Efficiency measurement for parallel production systems. European Journal of Operational Research 2009;196(3):1107–12.
[89] Tone K, Tsutsui M. Network DEA: a slacks-based measure approach. European Journal of Operational Research 2009;197(1):243–52.
[90] Färe R, Grosskopf S. Intertemporal Production Frontiers: With Dynamic DEA. Boston: Kluwer Academic Publishers; 1996.
[91] Chen CM. A network-DEA model with new efficiency measure to incorporate the dynamic effect in production networks. European Journal of Operational Research 2009;194(3):687–99.
[92] Tone K, Tsutsui M. Dynamic DEA: a slacks-based measure approach. OMEGA – International Journal of Management Science 2010;38(3–4):145–56.
[93] Cook WD, Liang L, Zhu J. Measuring performance of two-stage network structures by DEA: a review and future perspective. Omega – International Journal of Management Science 2010;38(6):423–30.
[94] Castelli L, Pesenti R, Ukovich W. A classification of DEA models when the internal structure of the decision making units is considered. Annals of Operations Research 2010;173(1):207–35.