# DEA in Banking: Analysis and Visualization of Bibliometric Data

Violeta Cvetkoska<sup>1</sup> and Gordana Savic<sup>2</sup>

<sup>1</sup>Faculty of Economics, Ss. Cyril and Methodius University in Skopje, Blvd. Goce Delchev 9V, 1000 Skopje, North Macedonia; vcvetkoska@eccf.ukim.edu.mk
<sup>2</sup>Faculty of Organizational Sciences, University of Belgrade, Jove Ilica 154, 11000 Belgrade, Serbia; gordana.savic@fon.bg.ac.rs

# ABSTRACT

Data Envelopment Analysis (DEA) is a popular nonparametric frontier methodology for measuring the efficiency of decision-making units (DMUs). DEA has been used widely in the banking industry. There is no reference to country co-authorship and keywords co-occurrence network visualization of DEA in the banking literature. In this study, we examine the DEA literature in banking, taking into account articles published in peer-reviewed journals listed in the Scopus database. We have performed an analysis and visualization of bibliometric data related to 791 DEA articles in banking in the period of 34 years (1986–2019). We show the annual trends of published DEA banking articles, the top journals and authors, citation analysis, and country co-authorship. We have also performed an in-depth keyword visualization analysis and discussed the future trends of DEA in banking.

*Keywords:* Data envelopment analysis; banking; bibliometric data; visualization

ISSN 2161-1823; DOI 10.1561/103.00000044 O 2021 V. Cvetkoska and G. Savic

## 1 Introduction

Ever since the seminal article of Charnes *et al.* (1978) was published, where Data Envelopment Analysis (DEA) was introduced as a nonparametric frontier methodology for measuring the efficiency of decisionmaking units (DMUs), there has been a remarkable growth in DEA publications. The banking area is one of the most attractive areas in DEA applications. In this article, we examine the status of DEA articles in banking as well as the development path based on in-depth keywords co-occurrence visualization network analysis. We use the bibliometric data of 791 DEA articles in banking published in peer-reviewed journals listed in the Scopus database in the period between 1986 and 2019. We have found the research hotspots in DEA in banking in a period of 34 years, as well as show and discuss how the keywords change in different time periods, and we suggest future trends in this area.

Efficiency evaluation is crucial for all profit and non-profit organizations. On the one hand, measuring the efficiency of the profit organization can be done by using financial parameters and the basic definition of partial efficiency as the ratio of output and input. Nevertheless, efficiency measurement becomes more complicated when using multiple commensurate inputs, and multiple commensurate outputs are necessary. Having in mind such real-life examples of efficiency evolution of primary schools, Charnes et al. (1978) introduced the mathematical modeling technology DEA, using Farell's definition of efficiency (Farrell, 1957). DEA is a non-parametric frontier methodology that allows relative efficiency evaluation, identification, and assessment of the areas of best performance or best practice within the sample. The main idea is to compare the efficiency of each decision-making unit with others in the observing set. The assumption is that all DMUs consume the same inputs to produce the same outputs. Since 1978, the fields of DEA application have been widely broadened. Emrouznejad and Yang (2018) have found that DEA is most applied in the following five areas: agriculture, banking, supply chain, transportation, and public policy. Their findings refer to published journal articles on DEA in the period 2015 - 2016.

Since the stability of the financial system in one country depends on the stability of its banking system, this makes the banking area quite attractive for research. Also, banks have a central role in less developed

countries where they have the largest share in the total assets of the banking system. According to the study of Berger and Humphrey (1997), where they analyze 130 studies (122 study depository institutions and 8 insurance institutions) that have applied frontier efficiency analysis. it has been found that in the non-parametric and parametric frontier efficiency approaches applied in 69 and 60 studies, respectively, but also in other studies, more than one approach is taken into account. Data envelopment analysis is applied in 62 of the non-parametric frontier analysis studies. Paradi and Zhu (2013) have done a survey on 80 published applications on DEA in bank branches, covering 24 countries. In their analysis, they focus on the models that are used, and they give suggestions for the design of future studies in this area as well as for future areas for research. In the literature, we have not found any reference for DEA in banking regarding the country co-authorship and keywords co-occurrence by using network visualization analysis. In co-authorship analysis, the importance of country co-authorship analysis can be seen in reflection of the degree of communication between countries, as well as the countries that are influential in the current field (Liao et al., 2018). In addition, by keywords co-occurrence analysis, the status of DEA in banking and its research hotspots can be explored. That increases our motivation for this study, and instead of exploring trends of DEA articles in banking by year, top journals, top authors, and citation analysis, we have also done in-depth country and keywords co-occurrence visualization network analysis.

For the purposes of our study, we use bibliometric data downloaded from the Scopus database. Our focus is only on the journal articles, and we have found 791 articles published in the period of 34 years (1986– 2019). In some articles the idea of extended and new DEA models in the banking context is explained, whereas in others, DEA is applied on real banking data. The version of the VOSviewer software — 1.6.15 is used for creating the network, density, and overlay visualization maps. One of the advantages of this version of the VOSviewer is the possibility to make a co-occurrence keywords analysis by extracting keywords from abstracts, and we use this option in our keywords co-occurrence analysis.

We have found that most of the articles on DEA in banking (41.8%) were published in the period 2011–2019, so during the previous year (2019) the highest number of articles (80) in the whole observed period was published. We have identified the top five journals where the DEA

articles in banking have been published according to the following five criteria: (1) the highest number of DEA articles in banking published in the journal in the whole analyzed period; (2) the highest impact factor of the journal (2019 release); (3) the highest SCImago Journal Rank (SJR 2019); (4) the highest number of total citations of the journal received by published articles in the journal in the period 1999–2019; and (5) the highest number of total citations of the journal received by published articles in banking in the journal in the period 1986–2019. For each of the selected criteria respectively, the top journals are identified as the following: the European Journal of Operational Research, Applied Soft Computing, Management Science, Expert Systems with Applications, and Journal of Banking & Finance.

For each of the top five journals, we have identified the top authors of DEA in banking articles, taking into account their total number of citations in the current journal. The most cited authors in DEA articles in banking are: Lovell, A.K., in the European Journal of Operational Research, Tavana, M., in Applied Soft Computing, Banker, R.D., and Morey, R. C. (co-authors) in Management Science, Wu, D., Yang, Z., and Liang, L. (co-authors) in Expert Systems with Applications, and Hall, M.B.J., in Journal of Banking & Finance.

In the identified top five journals there have been published 80 DEA articles in banking, and we present the 10 most cited articles. The most cited DEA article in banking is titled: "The impact of liberalization on the productive efficiency of Indian commercial banks", published in the *European Journal of Operational Research* in 1997 (Bhattacharyya et al., 1997).

Based on the analysis of DEA articles in banking per country, we have found that Malaysia has the highest number, followed by the United Kingdom and the United States. In addition, the United States has the highest total link strength with other countries, followed by the United Kingdom and Iran. Also, we have found that if the geographical distance between the two countries is shorter, that does not influence greater co-authorship between countries.

According to the in-depth keywords co-occurrence network visualization analysis, we have found that the keywords — crisis, branch, and DEA model — come up the most in the period 1986–2019. By analyzing the links of the DEA model, we have found that one of its greatest link strengths is with multiple outputs and the stage DEA

Multiple outputs are one of the advantages of using DEA model. instead of parametric frontier analysis, where only one output is allowed. Furthermore, we analyze three different periods (pre-2000, 2001–2010, and 2011–2019) in order to see whether the research hotspots change over time. Based on the created three keywords co-occurrence network visualizations, we have found that the number of keywords increases, and also, they differ from one to another period, and this is due to two reasons. One reason is the development of the non-parametric frontier methodology DEA over time as well as its integration with parametric methods, such as multivariate regression, multi-criteria decision-making methods (ELECTRE, AHP, ANP, TOPSIS, VIKOR), and advanced analytics and artificial intelligence (AI) methods, such as data mining and artificial neural networks. The second reason arises from the global situation affecting the banking industry, such as economic and financial crises, development of information technology, internet banking, and e-commerce. We have also analyzed the last three years (2017–2019) to find the research hotspots and suggest future trends on DEA in banking. We have found that the interest of researchers arises regarding: corporate social responsibility, internet banking, non-performing loans, banks' ability, micro-finance institutions, intellectual capital, stochastic DEA, and advanced analytics and AI techniques, as we believe that these keywords will be the new research hotspots on DEA in banking in the future. We hope that this study will help researchers interested in DEA in banking to understand the status of DEA in banking as well as its development path, yet also to do research in the identified and suggested new research areas.

The remainder of this article is organized as follows. Section 2 describes the data and used tools. Sections 3 presents the results on DEA articles in banking regarding the annual trends, journals, authors, citation analysis, and country co-authorship. Section 4 gives the results on the keywords co-occurrence analysis. Section 5 summarizes and concludes the article.

## 2 Data and Tools

The data we use for the purposes of this bibliometric study have been downloaded from the Scopus database. In this database, a search was made for the words: "Data Envelopment Analysis" and "Banking" in Title-Abstract-Keywords, taking into account only articles published in peer-reviewed journals, and the period is from the first article in this database until 2019. A total of 791 articles were found, with the first reference being from 1986. All data (authors, title, year, source title, citations, authors with affiliations, abstract, author keywords) were downloaded on July 8, 2020, as a text file.

VOSviewer version 1.6.15 software is used for the analysis of country co-authorship as well as for keywords co-occurrence. This software was developed by van Eck and Waltman, and it allowed us to create and visualize networks based on bibliometric data; for more details, see Van Eck and Waltman (2010, 2014). Its successful application in bibliometric studies is evidenced by most recent publications in various fields, such as: social networking and academic performance (Doleck and Lajoie, 2018), network-based on co-mention of R-packages (Li and Yan, 2018), medical big data research (Liao *et al.*, 2018), multi-criteria decision-making (Yu *et al.*, 2018), neuroscience (Andersen *et al.*, 2017), corporate social responsibility (Lulewicz-Sas, 2017), and management information systems (Ozkose and Gencer, 2017).

In the version of VOSviewer software from April 2020, which we use in this study, there is an additional option for identifying keywords by searching in the abstract, which was additional motivation for us to make an in-depth analysis of the DEA status in banking, but also to identify trends.

## 3 Results on DEA Articles in Banking

In this section, we present the results on DEA articles in banking regarding the annual trends, top journals and authors, citation, and country co-authorship.

## 3.1 Annual Trends of DEA Articles in Banking

The distribution of DEA articles in banking in the time frame of 34 years (1986–2019) is shown in Figure 1. The total number of published articles is 791. In the period 1986–2000, 5% of the articles were published. In the period 2001–2010, 28% of the articles were published. A double-digit

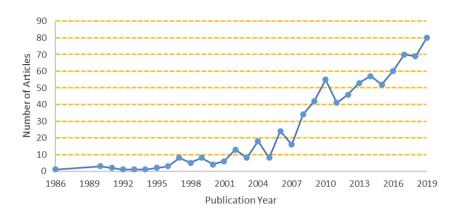


Figure 1: Annual trends of DEA articles in banking.

number of published DEA articles in banking appeared for the first time in 2002 (13 articles). A significant increase is observed from 2008 to 2010, when a total of 131 articles were published in this period, which is one less than the sum of all the previous years in the analysis. In the period 2011–2019, the greatest number of articles was published: 528 (67%). In addition, in the last five years (2015–2019) the average share in the total number of articles is 41.8%, and the most published articles (80) for the entire period of 34 years were published in 2019.

#### 3.2 The Top Journals on Published DEA Articles in Banking

The DEA articles in banking have been published in a total of 322 journals. The identification of the top journals where DEA articles in banking have been published is based on the following five criteria:

- 1. the highest number of DEA articles in banking published in the journal in the period 1986–2019
- 2. the highest impact factor of the journal (2019 release)
- 3. the highest SCImago Journal Rank (SJR 2019)
- 4. the highest number of total citations of the journal received by published articles in the journal in the period 1999–2019

5. the highest number of total citations of the journal received by published DEA articles in banking in the journal in the period 1986–2019

For each of the five criteria, our sample consists of journals that have published two or more DEA articles in banking, and these are 145 journals (45%). For the first and fifth criteria we use the data (title of the articles, source title (the name of the journal), and total citations of each DEA article in banking for the period 1986–2019) downloaded from the Scopus database. For the other three criteria we use data (impact factor for 2019 release, SCImago journal rank (SJR 2019), and total citations of the journal received from published articles in the journal in the period 1999–2019) from the publicly available portal SCImago Journal & Country Rank (SCImago, n.d.).

Based on the performed analysis for each of the five criteria, Table 1 lists the five top journals where DEA articles in banking are published. The European Journal of Operational Research has the highest number of published DEA articles in banking in the period 1986–2019 (39 articles). The journal Applied Soft Computing has the highest impact factor for 2019 release (impact factor: 5.472). The journal Management Science has the highest SCImago Journal Rank, (SJR 2019: 5.44). The journal Expert Systems with Applications has the highest number of total citations received by published articles in the journal in the period 1999–2019 (total citations: 139,646). The highest number of total

Journal	Metric
European Journal of	Number of published DEA articles in banking
<b>Operational</b> Research	(1986–2019): <b>39</b>
Applied Soft Computing	Impact Factor of the journal (2019 release): 5.472
Management Science	SCImago Journal Rank (SJR 2019): 5.44
Expert Systems with Applications	Total citations of the journal received by published articles in the journal in the period 1999–2019:
	139,646
Journal of Banking $\mathfrak{G}$	Total citations of the journal received by published
Finance	DEA articles in banking in the period 1986–2019:
	1,951

Table 1: The top-five journals where DEA articles in banking are published.

citations of the journal received by published DEA articles in banking in the journal in the period 1986–2019 has been identified in the *European Journal of Operational Research*, but this journal is already selected as a top journal based on the number of DEA articles in banking published in the journal in the same period. It is for this reason that we have selected the journal that follows the *European Journal of Operational Research*, according to the fifth criterion, and that is the *Journal of Banking & Finance* (total citations of DEA articles in banking in the period 1986–2019: 1,951).

## 3.3 The Top Authors of DEA Articles in Banking in Each of the Identified Top Journals

The total number of different authors in all DEA articles in banking (791) is 1,378. To identify the top authors of DEA in banking articles we focus on those authors that have published DEA articles in banking in the identified five top journals (the European Journal of Operational Research, Applied Soft Computing, Management Science, Expert Systems with Applications and Journal of Banking & Finance). In these five top journals there have been published 80 DEA articles in banking, and the total number of different authors of these articles is 206. Most articles (35; 43.8%) are written by two authors, 22 articles (27.5%) are written by three authors, 11 (13.8%) are written by four authors. The greatest number of authors by article is five, and this is identified in four articles. The average number of authors per DEA article in banking is 2.6.

For each of the five top journals, we have identified the top authors of DEA in banking articles, taking into account their total number of citations per article/s in the current journal (Figure 2). In the *European Journal of Operational Research* the most cited author of DEA articles in banking is Lovell, A.K. (total citations: 452). Tavana, M. is the only author who has published two articles in DEA in banking in the journal with the highest impact factor (2019 release), i.e. in the journal *Applied Soft Computing* (total citations: 30). Banker, R.D., and Morey, R.C. (co-authors) are the most cited authors in this area in the journal *Management Science* (total citations: 348). Wu, D., Yang, Z. and Liang, L. (co-authors) are the most cited authors in DEA in banking in the journal *Expert Systems with Applications* (total citations: 184), whereas

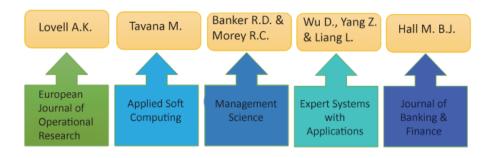


Figure 2: The top authors in DEA in banking in each of the five top journals.

Hall, M.B.J. is the most cited author in this area in the *Journal of* Banking & Finance (total citations: 363).

# 3.4 Citation Analysis

Out of the total number of published DEA articles in banking (791), the number of cited articles is 686 (86.7%). In the citation analysis we have focused only on DEA articles in banking published in the identified five top journals, and in Table 2 we present the 10 most cited articles in these journals. The most cited 10 articles on DEA in banking have been published in three of the five top journals. Half of them (five articles) have been published in the Journal of Banking & Finance, three articles published in the European Journal of Operational Research and two in Management Science. The most cited article is titled: "The impact of liberalization on the productive efficiency of Indian commercial banks", published in the European Journal of Operational Research in 1997 (Bhattacharyya et al., 1997). This article has 367 citations.

# 3.5 The Co-authorship Analysis Based on Countries

By using the software VOSviewer the following has been created: (1) network visualization for country co-authorship, which is presented in Figure 3; and (2) item density visualization, presented in Figure 4.

The type of analysis is co-authorship, and the unit of analysis are countries. A total of 87 countries were identified, but we have set the condition of taking into account only countries with the minimum number

				Total
Title	Source title	Author/s	Year	citations
The impact of liberalization on the productive efficiency of Indian commercial banks	European Journal of Operational Research	Bhattacharyya A., Lovell C.A.K., Sahay P.	1997	367
Use of categorical variables in data envelopment analysis	Management Science	Banker R.D., Morey R.C.	1986	348
The technical efficiency of large bank production	Journal of Banking ど Finance	Miller S.M., Noulas A.G.	1996	252
Banking efficiency in the Nordic countries	Journal of Banking හි Finance	Berg S.A., Førsund F.R., Hjalmarsson L., Suominen M.	1993	206
Evaluating the cost-efficiency of the Italian banking system: What can be learned from the joint application of parametric and non-parametric techniques	Journal of Banking & Finance	Resti A.	1997	200
X-efficiency in Australian banking: An empirical investigation	Journal of Banking ど Finance	Sathye M.	2001	199
Best practice analysis of bank branches: An application of DEA in a large Canadian bank	European Journal of Operational Research	Schaffnit C., Rosen D., Paradi J.C.	1997	196

Table 2: The most cited 10 DEA articles in banking published in the identified top journals.

(Continued)

				Total
Title	Source title	Author/s	Year	citations
An empirical study on measuring operating efficiency and profitability of bank branches	European Journal of Operational Research	Oral M., Yolalan R.	1990	195
Operations, quality, and profitability in the provision of banking services	Management Science	Soteriou A., Zenios S.A.	1999	189
The impact of macroeconomic and regulatory factors on bank efficiency: A non-parametric analysis of Hong Kong's banking system	Journal of Banking හ Finance	Drake L., Hall M.J.B., Simper R.	2006	188

Table 2: (Continued)

of documents (articles) to be five, and as 41 countries have satisfied this condition, they were used to create both types of visualizations.

The network visualization consists of items represented by labels and circles, and lines that present links between items. The larger label and circle refer to the greater weight of that item. The shorter distance between the two items means a stronger relationship. Also, the strength of a link is given by numerical value, which is positive, and if this value is higher, it indicates that the link is stronger, and the total link strength of one item is the strength of the links of one item with other items, calculated as a sum (Van Eck and Waltman, 2010). The same color of items indicates that they are part of the same cluster. The network visualization for country co-authorship in our study is shown in Figure 3. According to the color, it can be seen that items are grouped into nine clusters. The largest label and circle, meaning the greatest weight based on published articles on DEA in banking, belongs to Malaysia. Malaysia has 109 articles in DEA in banking and is the

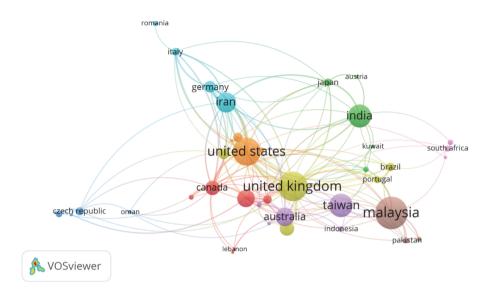


Figure 3: Co-authorship network on DEA articles in banking based on countries.

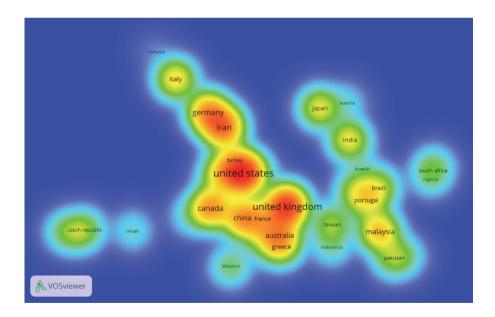


Figure 4: Country density visualization for DEA articles in banking.

No.	Country	Documents	Total link strength
1	Malaysia	109	31
2	United Kingdom	95	75
3	United States	88	81
4	Taiwan	71	14
5	India	68	22
6	Iran	58	53
7	Australia	51	32
8	China	48	35
9	Canada	37	41
9	Greece	37	18
10	Germany	28	41

Table 3: Top countries in DEA articles in banking and their total link strength.

most influential country in this area, followed by the United Kingdom (95 articles) and the United States (88 articles). The 10 countries with the largest number of documents and their total link strength are given in Table 3.

Between Malaysia and the United States, the link strength is 2; between Malaysia and the United Kingdom, and between the United Kingdom and the United States respectively it is 3. Also, the shorter geographical distance is not an indicator for greater co-authorship between countries, e.g. the link strength between Germany and the United Kingdom is 2, while between Germany and the United States it is 13.

In the VOSviewer software for density visualization, we use the variant item density visualization. The country density visualization for DEA articles in banking is presented in Figure 4. Items are shown by their label, i.e. country name. The color of the item is obtained regarding the number of items in its neighborhood, as well as their weights. In this case, the weight is not the number of documents but the total link strength of the country. For Figure 4 we have chosen a rainbow of colors. The countries that are in the red area will have a higher total link strength in comparison to the others. The United States has the highest total link strength, followed by the United Kingdom and Iran.

## 4 Analysis of Keywords in DEA Articles in Banking

In this part, the visualizations created in VOSviewer software for the keyword distribution based on the terms extracted from the 791 abstracts are presented and analyzed. First, the focus is on keywords that occur a minimum of 10 times in the observed period 1986–2019 in order to identify the research hotspots in the whole period. Then, we analyze how the keywords have changed in different time periods (pre-2000, 2001–2010, and 2011–2019), where we have set the minimum number of occurrences of the keywords to be two, because new DEA models have been developed and the interest of researchers might change from one analyzed period to another, and we want to present a detailed picture of the status of DEA in banking. In addition, we present an overlay visualization for the last five years in order to demonstrate the developments of DEA in banking over this period, and therefore suggest future trends.

## 4.1 Keywords Analysis in DEA Articles in Banking in the Period 1986–2019

We have chosen to create a keyword co-occurrence map based on text data from the Scopus database. The terms are extracted from abstracts and the abstract structure labels, as copyright statements are ignored from the abstract fields. 10,226 terms were identified, and we have chosen the minimum number of occurrences to be 10, so 479 terms remained. The default choice by the software is to select the 60% of most relevant terms based on the relevance score, so that 287 terms remain to be verified. Out of these 287 terms, we have excluded the countries (the country analysis is given in part 3.5) as well as the words that do not have meaning for the analysis, such as number, increase, decline, empirical findings, consideration, and whole. The visualization network is presented by items and lines. The items are displayed with labels and circles. The larger the label and the circle, the greater the weight of the item is. For some items, the label might not be shown, and this is due in order to avoid the problem of overlapping labels (Van Eck and Waltman, 2020). A line between two items presents a link. The thicker the line between two items and the shorter the distance, the stronger the relationship between items. The number of different

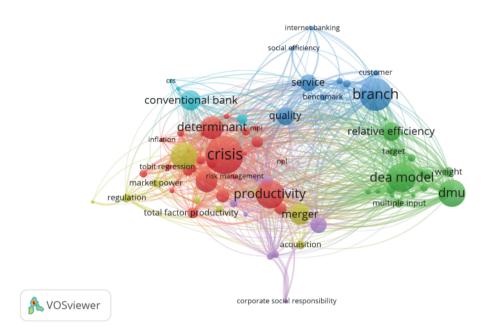


Figure 5: Network visualization for keywords co-occurrence in DEA articles in banking (1986–2019).

colors means the number of different clusters. Items with the same color belong to one cluster, and they are grouped based on how related they are. In Figure 5, the network visualization for keywords co-occurrence in DEA articles in banking for the period 1986–2019 is presented. The number of items for this network is 72, there are 1,016 links, and the total link strength is 5,032. According to the different colors, we can see that the items are classified into six clusters.

The keyword crisis that belongs to the red cluster is presented by the largest label and circle, which means that it has the greatest weight (based on occurrences), followed by branch (blue cluster) and DEA model (green cluster). The number of occurrences of the keyword crisis is 189 with 53 links, and the total link strength is 479. In the same cluster (red color), the highest link strength is between crisis and foreign banks (50), followed by credit risk, determinant, and risk management with link strength of 41, 29, and 23, respectively. The keyword productivity is also part of the red cluster (it has 132 occurrences), and to estimate the productivity change over time the Malmquist index has been used (46 occurrences). The keyword branch has 157 occurrences, 39 links, and 441 is its total link strength. Its highest link strength is with the keyword service (50), followed by relative efficiency (48), behavior (36), and customer (33). In the green cluster, the keyword DEA model has the highest number of occurrences (134), 49 links, and the total link strength is 426. The highest link strength is with DMU (77), relative efficiency (33), multiple output (16), and stage DEA model. In the yellow cluster, the keyword competition has the highest number of occurrences (123), 43 links, and a total link strength of 291. The link strength between competition and market power is highest (38), followed by production (20), service (9), and quality (7). In the turquoise cluster, the conventional bank has the greatest weight based on occurrences (85), 34 links, and a total link strength of 231. Its link strength with revenue is 19, while with quality it is 17. In the purple cluster, the frontier has the highest number of occurrences (66), 47 links, and a total link strength of 259. Banking performance, income, equity, and corporate social responsibility belong in the same cluster, so that the link strength between corporate social responsibility (CSR) and banking performance is 22.

Based on Figure 5 we have found that one of the keywords with high frequency in the cluster, where the DEA model belongs, is multiple outputs, and this term has one of the advantages of using DEA over the parametric frontier approach, where only one output could be part of that analysis. Also, researchers use a stage DEA model, usually with multiple regression to identify determinants of efficiency, especially in a crisis. For assessing the productivity change over time, the Malmquist index has the main role. The attention of the researchers in the banking area is on the crisis, as well as on the global financial crisis that occurred in the period 2007–2009 and its implications on the banking performance. Additionally, the branches are identified as a hotspot area, especially their link with service, customers, and employees, then on mergers and acquisitions in order to make higher financial gains. Taking into account the fact that employees want to be part of firms whose priority is CSR and new customers will be more attracted to give or borrow money from these firms, while current customers will remain loyal, we can see that researchers are attracted to this hotspot in DEA in banking.

## 4.2 Research Hotspots for DEA in Banking Based on Keywords Analysis in Different Time Periods

In order to gain information on how keywords change over time, in VOSviewer we have created three network visualizations for the following time periods: pre-2000, 2001–2010, and 2011–2019. The network visualization maps for each period are presented in Figures 6–8, respectively. To create these network visualizations, we have chosen the minimum number of occurrences of a keyword to be two.

The network visualization for the first period, i.e. pre-2000, consists of 21 connected items (that are classified into five clusters), 36 links, and the total link strength is 104. For the period 2001–2010, the number of connected items is 98, and they are classified into 13 clusters, so in total there are 300 links, and the total link strength is 1,035. In the last observed period (2010–2019) the network visualization is created with 117 connected items (that are classified into 15 clusters), 263 links, and the total link strength is 924.

The increase in the number and the difference of keywords is due to two reasons: one reason being the development of the DEA methodology

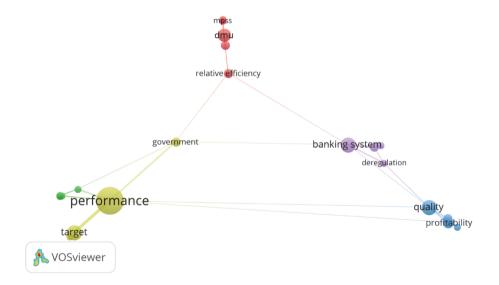


Figure 6: Network visualization for keywords co-occurrence in DEA articles in banking (pre-2000).

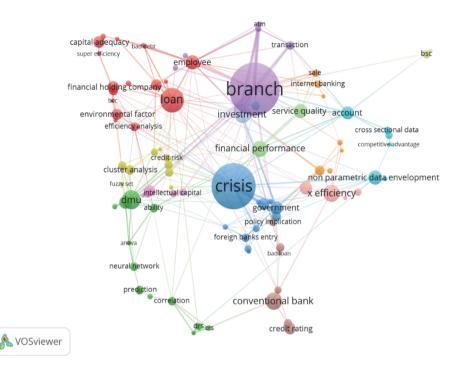


Figure 7: Network visualization for keywords co-occurrence in DEA articles in banking (2001–2010).

over time as well as its integration with parametric methods, such as multivariate regression analysis, multi-criteria decision-making methods (ELECTRE, AHP, ANP, TOPSIS, VIKOR), and advanced analytics and AI methods, such as data mining and artificial neural networks; and the second reason refers to the global situation affecting the banking industry, such as economic and financial crises, the development of information technology, internet banking, and e-commerce.

In the first analyzed period (pre-2000), which is presented in Figure 6, the main keyword is performance with 15 occurrences, 9 links and a total link strength of 35. It belongs in the yellow cluster along with the following three keywords: bank management, government, and target, and its link strength is greatest with the keyword government (the link strength is 5). Based on the number of occurrences, what follows the keyword performance is: banking system, and quality, with

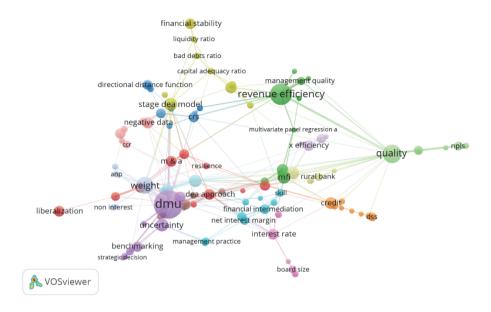


Figure 8: Network visualization for keywords co-occurrence in DEA articles in banking (2011–2019).

six occurrences each. The keyword banking belongs to the purple cluster and it has five links and a total link strength of 15. Its highest link strength is with the keyword in the same purple cluster, more precisely with the technology (the link strength is 8). The keyword quality belongs to the blue cluster, along with multimarket business environment, profitability and services, and it has five links and a total link strength of 13. The decision-making unit is part of the red cluster, it has five occurrences, four links, and a total link strength of 12. Its link with the most productive scale size (mpss) is highest and it is 4. In the green cluster there are four items (financial ratio analysis, liberalization, managerial weakness, and privatization) and each of them occurs two times. The link strength between financial ratio analysis with managerial weakness is 2, and also managerial weakness is linked with performance (2 is the link strength).

In the second period (2001–2010), presented in Figure 7, the keyword with the most occurrences is the branch (60) with 23 links, and it has the highest link strength with customer perception (27). In second place,

based on the number of occurrences is the keyword crisis (50), which has 22 links, and the link strength with the domestic bank counterpart is the highest (12), while with loans intensity it is 4. The loan is the next keyword based on the number of occurrences (21) and 22 links, so that its link strength with the production process is 6, and with Basel II and bad debt it is 2, respectively. In the same cluster with the keyword loan is the keyword managerial efficiency, which has three occurrences and five links, so that its link with capital is 2, and with principal component analysis it is 1. In addition, in the green cluster, DMU is the keyword with the greatest weight, and in this cluster there are included the keywords prediction, neural network, ability. The link strength between neural network and prediction is 2, while between ability and neural network, financial performance, and loan intensity, respectively, is 1.

In the third period (2011-2019) presented in Figure 8, the keyword DMU is with the highest number of occurrences (104), and with 23 links, so that its link strength with the stage DEA model is 16. Revenue efficiency occurs 55 times, the total links are 13, and the link strength with managerial quality is 22. The keyword quality has 39 occurrences and 19 links. The link strength between quality and non-performing loans is 14, and with artificial neural networks is 1. In the pink cluster the link strength between the interest rate and board size is 2. In addition, management practice (which has five occurrences and four links) is shown on this network map. Its link strength with the stochastic DEA is 2, and with net interest margin is 1. The keyword stochastic DEA has two occurrences. Corporate social responsibility occurs 12 times and has 4 links with a total link strength of 57. Its link strength with financial performance is highest (33), while with communication is 1. Keyword skill occurs five times and has seven links. Its link strength with net interest margin and technology innovation, respectively, is 1. In the orange cluster, credit has the greatest weight and its link strength with experience, and logistic regression, respectively, is 2. Also, a data mining technique with two occurrences appears as an orange item (the label is not given in Figure 11).

From Figures 6 to 8 we have seen that different keywords appear in different time periods, so that at the beginning (pre-2000) the articles in DEA in banking focus on measuring the performance, the efficiency of the banking system by using only DEA models. In the next period (2001–2010), the main focus for researchers is on the crisis, especially the

global financial crisis that is part of this period. They analyze the loans, capital adequacy, and bad debts that are linked with the crisis. The second research hotspot is branches, where they focus on transactions, customers, employees, ATMs, and quality. This period shows the term stage DEA model, where DEA is integrated with regression analysis. In addition, in this period the keywords fuzzy DEA, bootstrap, cluster analysis, balanced scorecard, and neural networks are found. In addition, attention is paid on ability and its link with financial performance, loan intensity, and crisis. In the last period (2011–2019) we found keywords, such as fuzzy DEA, network DEA, inverse DEA, and stochastic DEA. As we live in a world of big data, the data analytics and AI techniques as data mining techniques and artificial neural networks are found as keywords, but with a small number of occurrences. In this period most of the clusters are linked with the keywords, such as management quality, management practice, a board of directors, also with a small number of occurrences.

#### 4.3 Future Trends for DEA in banking

Figure 9 presents the overlay visualization of keywords, showing the development of DEA in banking over the period 2015–2019 (the colors of the rainbow are used for its creation), and our focus will be for the last three years.

We have found that 2017 uses the keywords: corporate social responsibility, internet banking, non-performing loans, and banks' ability. In 2018 the following keywords are used: micro-finance institution, social efficiency, intellectual capital, and stochastic DEA, while in 2019, the keywords are monitoring activity, data mining technique, and logistic regression. The keywords: banks ability, stochastic DEA, data mining techniques occur only two times, but instead of the other abovementioned, we believe that they will be the new research hotspots in DEA in banking.

Demerjian *et al.* (2012) have developed a measure for managerial ability by using DEA, but they have not taken into account the banking sector. In this sector there is no generic production function, and based on Berger and Humphrey (1997), one way that banks could be seen is as producers of transactions for customers, while the other is as depositors. Banker and Tripathi (2021) have developed and validated a measure of

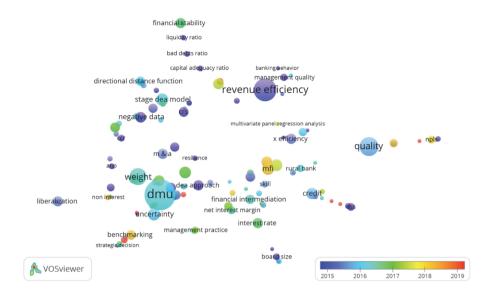


Figure 9: Overlay visualization for keywords in DEA in banking in the period 2015–2019.

organizational ability in the context of the banking sector in the U.S. They have found that this measure is persistent and directly associated with performance.

The stochastic DEA (SDEA) was introduced by Banker (1988) as an extension of a traditional DEA, and it allows two-sided deviations from the estimated production frontier. For a Stochastic DEA review, see Olesen and Petersen (2016).

Despite the integration of advanced analytical and AI techniques with DEA, we suggest that it will be interesting to examine how banks that apply data analytics and AI transform their operations and how efficient they are.

#### 4.4 Managerial Implications

Managers might use the efficiency and productivity study of financial institutions as a supporting tool in the decision-making process. Generally, the primary purpose of data envelopment analysis is to classify units under observation into relatively efficient and inefficient groups. Furthermore, individual weakness and improvement goals can be found for each inefficient unit. Malmquist DEA based on total factor productivity indices reveal the dynamics of relative efficiency and productivity, and their potential trend may lead to a development direction (Camanho and Dyson, 2006). Those main properties of DEA methodology might be beneficial in the banking industry.

The green cluster (Figure 5) sheds light on DEA methodology and shows that the keywords, such as DEA models, relative efficiency, and DMUs, are the most frequent in the surveyed studies. The DMUs selection depends on the level and context of efficiency evaluations. They can be countries, regions, mergers, micro-finance institutions, conventional banks, or bank branches. Regardless of the level, DEA models help in classifying DMUs into efficient and inefficient groups, hence DEA results help in identifying the 'best' practices, as the most efficient DMUs, operating on or near the efficient frontier. Their common management practices and policies might be used as benchmarks to identify directions of the performance improvement by managers of inefficient units. In contrast, the policies and operating models of the units assessed as the "worst practices" should be avoided. Furthermore, the main role of the institution might be determined using DEA. Banks offer different services, and products for the customers relate to collecting deposits, storing, transferring and extending credit, and managing other transactions. Therefore, the most common approaches consider that units in banking industry play the intermediate or productive role in order to gain profit. In that context, the effect of deposits on bank efficiency depends on the efficiency in both stages of the bank production process (Holod and Lewis, 2011). Resolving the role of deposits is vital to setting the goal of improving profitability. In addition, by evaluating the cost and profit efficiency of banks (Ray and Das, 2010; Resti, 1997; Shiraz et al., 2016), managers could take proper actions to continuously improve banks' operation, which is of crucial importance for the stability of the financial system.

A blue cluster (Figure 5) indicates that a financial institution may successfully use DEA analysis for internal ranking and comparison of bank branches. Usually the managers pay special attention to cost and revenue efficiency (Cook *et al.*, 2000). The cost efficiency evaluation is proven to be robust and might indicate strategy improvement actions, even in situations of price uncertainty (Camanho and Dyson, 2005). Profitability in banking is closely related to satisfied customers. The internal and external satisfaction of customers might be improved by service efficiency and quality raising. The service efficiency and quality are measured by tangible and intangible determinants (Chang *et al.*, 2017; Mukherjee *et al.*, 2003; Soteriou and Zenios, 1999). The impact of those determinants varies for one to another DMU under evaluation, depending on its own achievements, which determines the strategy of efficiency improvement for each of them. Furthermore, the nonparametric methodology DEA could help the management of banks to reallocate resources among the network of branches with the aim of increasing their efficiency (Oral and Yolalan, 1990).

During and after the world economic crisis, managing the financial stability, risks, and crisis together became of great importance in the banking industry (red, yellow, and purple clusters; Figure 5). The determinants of profitability and marketability and overall technical efficiency have proven to help predict and manage the risk of failure (Luo, 2003), while the risk factors impact efficiency (Chang and Chiu, 2006). For example, increasing loan loss reserve and participation of banks in the Private Sector Involvement, as an intangible factor, could imply changing of loan approval and disbursement policy (Tsolas and Charles, 2015). Managerial decisions need to be influenced by external changes and factors, such as regulations, deregulations, crisis, market restructuring, and financial freedom (Chortareas et al., 2013). Their impacts are usually investigated by two-stage, and multiple-stage DEA combined with ordinary least squares (OLS), and Tobit regression (Banker et al., 2010; Drake et al., 2006; Sturm and Williams, 2004; Thoraneenitiyan and Avkiran, 2009).

## 5 Conclusion

In our study, we have done an analysis and visualization of bibliometric data on DEA articles in banking published in peer-reviewed journals listed in the Scopus database. The analysis covers 791 DEA articles in banking, published in a period of 34 years (from 1986 to 2019).

We have summarized the findings from this study in six points: (1) Based on the distribution of DEA articles in banking in the observed period, the greatest number of articles (528) has been published in the period 2011–2019, and the last year is characterized as containing the highest number of published DEA articles in banking (80 articles). (2) The five top journals where DEA articles on banking have been published are: the European Journal of Operational Research (based on the highest number of published DEA articles in banking in the period 1986–2019), Applied Soft Computing (based on the highest impact factor for 2019 release). Management Science (according to the highest SCImago Journal Rank (SJR 2019)), Expert Systems with Applications (according to the highest number of total citations of the journal received by published articles in the journal in the period 1999-2019), and the Journal of Banking & Finance (based on the highest number of total citations of the journal received by published DEA articles in banking in the journal in the period 1986–2019). (3) In the identified five top journals there have been published 80 DEA articles in banking, and the total number of different authors in these articles is 206. Most of the articles (35) are written by two authors. For each of the five top journals the top authors on DEA articles in banking have been identified based on their total number of citations per article/s in the current journal. Top authors in DEA in banking are: Lovell, A. K. (in the European Journal of Operational Research), Tavana, M. (in *Applied Soft Computing*), Banker, R.D., and Morey, R.C. (co-authors) (in Management Science), Wu. D., Yang, Z. and Liang, L. (co-authors) (in *Expert Systems with Applications*), and Hall, M.B.J. (in Journal of Banking & Finance). (4) Taking into consideration only DEA articles in banking published in the five top journals we have listed the 10 most cited articles. The most cited DEA article in banking is: "The impact of liberalization on the productive efficiency of Indian commercial banks" by the authors Bhattacharyya et al. (1997), published in the European Journal of Operational Research. (5) Based on the country co-authorship analysis, we have found 87 countries in total, and 41 countries have a minimum of five DEA articles in banking. The country with the highest number of DEA articles in banking is Malaysia, followed by the United Kingdom and the United States. In addition, our findings show that the shorter geographical distance between countries does not influence greater co-authorship between them. The United States has the highest total link strengths (sum of the total strength of links with other countries), followed by the United Kingdom and Iran. (6) By using the software VOSviewer, we have made an in-depth visualization analysis for keywords extracted

from 791 abstracts, hence finding that the keywords: crisis, branch, and DEA model occur the most in the whole observed period. One of the greatest link strengths of the DEA model is with multiple outputs and stage DEA models. By analyzing three different periods (pre-2000, 2001–2010, and 2011–2019) we have found that the number of keywords increases, and also they differ from one to another period, as this is due to two reasons. One reason is the development of the non-parametric frontier methodology DEA over time as well as its integration with parametric methods, such as multivariate regression, multi-criteria decision-making methods (ELECTRE, AHP, ANP, TOPSIS, VIKOR), and advanced analytics and AI methods, such as data mining and artificial neural networks. The second reason arises from the global situation affecting the banking industry, such as economic and financial crises, the development of information technology, internet banking, and e-commerce. Additionally, based on the visualization analysis of keywords in the last three years (2017-2019), we have found that the interest of researchers arises in: corporate social responsibility, internet banking, non-performing loans, banks ability, micro-finance institutions, intellectual capital, stochastic DEA, and advanced analytics and AI techniques, and we believe that these keywords will be the new research hotspots in DEA in banking in the future. In addition, we emphasize how managers might use the efficiency study of banking institutions as a supporting tool in the decision-making process. Despite the interesting findings from our study as a limitation, we point out that we have focused on DEA articles in banking in one database, therefore it may not cover all articles related to this area.

## References

- Andersen, J. P., K. Krogsgaard, A. M. Engel, and J. W. Schneider (2017), "Mapping International Impact of Danish Neuroscience from 2004 to 2015 Using Tailored Scientometric Methodology", *European Journal* of Neuroscience, 47(3), 193–200, https://doi.org/10.1111/ejn.13814.
- Banker, R. D. (1988), "Stochastic Data Envelopment Analysis", Carnegie Mellon University.

- Banker, R. D., H. Chang, and S. Y. Lee (2010), "Differential Impact of Korean Banking System Reforms on Bank Productivity", *Journal of Banking & Finance*, 34, 1450–60, https://doi.org/10.1016/j.jbankfin. 2010.02.023.
- Banker, R. D. and M. K. Tripathi (2021), "Measurement of Organizational Ability in US Commercial Banking", Accepted for publication in *Data Envelopment Analysis Journal*, 5(1), 183–242.
- Berger, A. N. and D. B. Humphrey (1997), "Efficiency of Financial Institutions: International Survey and Directions for Future Research", *European Journal of Operational Research*, 98(2), 175–212, https: //doi.org/10.1016/S0377-2217(96)00342-6.
- Bhattacharyya, A., C. A. K. Lovell, and P. Sahay (1997), "The Impact of Liberalization on the Productive Efficiency of Indian Commercial Banks", *European Journal of Operational Research*, 98(2), 332–45, https://doi.org/10.1016/S0377-2217(96)00351-7.
- Camanho, A. S. and R. G. Dyson (2005), "Cost Efficiency Measurement with Price Uncertainty: A DEA Application to Bank Branch Assessments", *European Journal of Operational Research*, 161(2), 432–46, https://doi.org/10.1016/j.ejor.2003.07.018.
- Camanho, A. S. and R. G. Dyson (2006), "Data Envelopment Analysis and Malmquist Indices for Measuring Group Performance", *Journal* of Productivity Analysis, 26(1), 35–49, https://doi.org/10.1007/ s11123-006-0004-8.
- Chang, M., H. B. Jang, Y. M. Li, and D. Kim (2017), "The Relationship between the Efficiency, Service Quality and Customer Satisfaction for State-Owned Commercial Banks in China", *Sustainability*, 9(12), 2163, https://doi.org/10.3390/su9122163.
- Chang, T. C. and Y. H. Chiu (2006), "Affecting Factors on Risk-Adjusted Efficiency in Taiwan's Banking Industry", *Contemporary Economic Policy*, 24(4), 634–48, https://doi.org/10.1093/cep/byl008.
- Charnes, A., W. W. Cooper, and E. Rhodes (1978), "Measuring the Efficiency of Decision Making Units", European Journal of Operational Research, 2(6), 429–44, https://doi.org/10.1016/0377-2217(78)90138-8.
- Chortareas, G. E., C. Girardone, and A. Ventouri (2013), "Financial Freedom and Bank Efficiency: Evidence from the European Union", *Journal of Banking & Finance*, 37(4), 1223–31, https://doi.org/10. 1016/j.jbankfin.2012.11.015.

- Cook, W. D., M. Hababou, and H. J. H. 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(3), 209–24, https://doi.org/10.1023/A:1026598803764.
- Demerjian, P., B. Lev, and S. McVay (2012), "Quantifying Managerial Ability: A New Measure and Validity Tests", *Management Science*, 58(7), 1229–48, https://doi.org/10.1287/mnsc.1110.1487.
- Doleck, T. and S. Lajoie (2018), "Social Networking and Academic Performance: A Review", *Education and Information Technologies*, 23(1), 435–65, https://doi.org/10.1007/s10639-017-9612-3.
- Drake, L., M. J. B. Hall, and R. Simper (2006), "The Impact of Macroeconomic and Regulatory Factors on Bank Efficiency: A Non-parametric Analysis of Hong Kong's Banking System", *Journal of Banking & Finance*, 30(5), 1443–66, https://doi.org/10.1016/j.jbankfin.2005.03. 022.
- Emrouznejad, A. and G. I. Yang (2018), "A Survey and Analysis of the First 40 Years of Scholarly Literature in DEA: 1978–2016", Socio-Economic Planning Sciences, 61, 4–8, https://doi.org/10.1016/j. seps.2017.01.008.
- Farrell, M. J. (1957), "The Measurement of Productive Efficiency", Journal of the Royal Statistical Society, 123(3), 253–81.
- Holod, D. and H. F. Lewis (2011), "Resolving the Deposit Dilemma: A New DEA Bank Efficiency Model", Journal of Banking & Finance, 35(11), 2801–10, https://doi.org/10.1016/j.jbankfin.2011.03.007.
- Li, K. and E. Yan (2018), "Co-mention Network of R Packages: Scientific Impact and Clustering Structure", Journal of Informetrics, 12(1), 87–100, https://doi.org/10.1016/j.joi.2017.12.001.
- Liao, H., M. Tang, L. Luo, C. Li, F. Chiclana, and X. J. Zeng (2018), "A Bibliometric Analysis and Visualization of Medical Big Data Research", *Sustainability*, 10(1), 166, https://doi.org/10.3390/ su10010166.
- Lulewicz-Sas, A. (2017), "Corporate Social Responsibility in the Light of Management Science — Bibliometric Analysis", Procedia Engineering, 182, 412–7, https://doi.org/10.1016/j.proeng.2017.03.124.

- Luo, X. (2003), "Evaluating the Profitability and Marketability Efficiency of Large Banks: An Application of Data Envelopment Analysis", Journal of Business Research, 56(8), 627–35, https://doi.org/ 10.1016/S0148-2963(01)00293-4.
- Mukherjee, A., P. Nath, and M. Pal (2003), "Resource, Service Quality and Performance Triad: A Framework for Measuring Efficiency of Banking Services", *Journal of the Operational Research Society*, 54(7), 723–35, https://doi.org/10.1057/palgrave.jors.2601573.
- Olesen, O. B. and N. C. Petersen (2016), "Stochastic Data Envelopment Analysis — A Review", European Journal of Operational Research, 251(1), 2–21, https://doi.org/10.1016/j.ejor.2015.07.058.
- Oral, M. and R. Yolalan (1990), "An Empirical Study on Measuring Operating Efficiency and Profitability of Bank Branches", *European Journal of Operational Research*, 46(3), 282–94, https://doi.org/10. 1016/0377-2217(90)90002-S.
- Ozkose, H. and C. Gencer (2017), "Bibliometric Analysis and Mapping of Management Information Systems Field", *Gazi University Journal* of Science, 30(4), 356–71.
- Paradi, J. C. and H. Zhu (2013), "A Survey on Bank Branch Efficiency and Performance Research with Data Envelopment Analysis", Omega, 41(1), 61–79, https://doi.org/10.1016/j.omega.2011.08.010.
- Ray, S. C. and A. Das (2010), "Distribution of Cost and Profit Efficiency: Evidence from Indian Banking", *European Journal of Operational Research*, 201(1), 297–307, https://doi.org/10.1016/j.ejor.2009.02. 030.
- Resti, A. (1997), "Evaluating the Cost-Efficiency of the Italian Banking System: What Can Be Learned from the Joint Application of Parametric and Non-parametric Techniques", Journal of Banking & Finance, 21(2), 221–50, https://doi.org/10.1016/S0378-4266(96)00036-2.
- SCImago (n.d.), "SJR-SCImago Journal and Country Rank [Portal]", Retrieved August 1–10, 2020: http://www.scimagojr.com.
- Shiraz, R. K., H. Fukuyama, M. Tavana, and D. D. Caprio (2016), "An Integrated Data Envelopment Analysis and Free Disposal Hull Framework for Cost-Efficiency Measurement Using Rough Sets", *Applied Soft Computing*, 46, 204–19, https://doi.org/10.1016/j.asoc. 2016.04.043.

- Soteriou, A. and S. A. Zenios (1999), "Operations, Quality, and Profitability in the Provision of Banking Services", *Management Science*, 44(9), 1221–38, https://doi.org/10.1287/mnsc.45.9.1221.
- Sturm, J. E. and B. Williams (2004), "Foreign Bank Entry, Deregulation and Bank Efficiency: Lessons from the Australian Experience", *Journal of Banking & Finance*, 28(7), 1775–99, https://doi.org/10. 1016/j.jbankfin.2003.06.005.
- Thoraneenitiyan, N. and N. K. Avkiran (2009), "Measuring the Impact of Restructuring and Country-Specific Factors on the Efficiency of Post-crisis East Asian Banking Systems: Integrating DEA with SFA", *Socio-Economic Planning Sciences*, 43(4), 240–52, https://doi.org/ 10.1016/j.seps.2008.12.002.
- Tsolas, I. E. and V. Charles (2015), "Incorporating Risk into Bank Efficiency: A Satisficing DEA Approach to Assess the Greek Banking Crisis", *Expert Systems with Applications*, 42(7), 3491–500, https: //doi.org/10.1016/j.eswa.2014.12.033.
- Van Eck, N. J. and L. Waltman (2010), "Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping", *Scientometrics*, 84(2), 523–38, https://doi.org/10.1007/s11192-009-0146-3.
- Van Eck, N. J. and L. Waltman (2014), "Visualizing Bibliometric Networks", in *Measuring Scholarly Impact: Methods and Practice*, ed. Y. Ding, R. Rousseau, and D. Wolfram, https://doi.org/10.1007/978-3-319-10377-8 13, Springer, Cham., 285–320.
- Van Eck, N. J. and L. Waltman (2020), "Manual for VOSviewer Version 1.6.15".
- Yu, D., W. Wang, W. Zhang, and S. Zhang (2018), "A Bibliometric Analysis of Research on Multiple Criteria Decision Making", *Current Science*, 114(4), 747–58.