Literature survey on DEA in the insurance industry with a focus on identification of research hotspots with text mining

Violeta Cvetkoska [a*] , Igor Ivanovski [a] , Marija Tasheva[b]

[a] PhD., Ss. Cyril and Methodius University in Skopje, Faculty of Economics - Skopje, 1) Mathematics and Statistics Department, 2) Management Department, Republic of North Macedonia
[b] Postgraduate student, Ss. Cyril and Methodius University in Skopje, Faculty of Economics - Skopje, Corporate Financial Management Program, Republic of North Macedonia

A B S T R A C T

DEA is a frequently used non-parametric methodology for measuring the relative efficiency of Decision-Making Units (DMUs) that use the same inputs to produce the same outputs. Emrouznejad and Yang (2018) provided a literature survey on DEA with 10,300 peer-reviewed journal articles from 1978 to the end of 2016. Our article focuses on DEA applications in the insurance industry in convergence with the existing relevant literature as Kaffash et al (2020), who have surveyed 132 DEA articles in the insurance industry for the period from 1993 to 2018. We include particular keyword analyses necessary to identify research hotspots in different periods. This article aims to conduct a bibliometric analysis of DEA-published documents (articles in journals and book chapters) in the insurance industry from 1993 to 2021, focusing on identifying research hotspots based on keyword co-occurrence analysis. We have analyzed published documents from relevant databases, such as Scopus, Web of Science, Ebsco and ProQuest. We use descriptive analytics and text mining as the main methods in our analysis. We provide descriptive statistics for articles per year and category of the insurance industry, geographical distribution, top five journals and authors by citations, and citation analysis. An additional qualitative factor of our article is in-depth keyword co-occurrence analysis by using text mining to identify research hotspots in the insurance industry. Our analysis aims to contribute to researchers and insurance practitioners as an empirical and applicative point for initiating and developing research.

1. INTRODUCTION

In the financial landscape of one country, of crucial importance is the stability of its key sectors and stakeholders (banks, insurance companies, pension and health funds and investment funds, among others) to maintain the stability of the financial system. Moreover, the financial system's stability is of key importance for economic growth. Therefore, measuring its efficiency as one of the leading indicators for the successful operation of financial and non-financial institutions attracts the research interest of many academics, researchers and practitioners worldwide.

Data Envelopment Analysis (DEA) is a non-parametric methodology that measures the relative efficiency of Decision Making Units (DMUs) that use the same inputs and produce the same outputs. Despite this, it does not require a production correspondence between the used inputs and the produced outputs. Instead, DEA provides information on the efficiency of each DMU and targets for improvement of the identified inefficiencies. Since DEA was introduced in the operational research (OR) literature by Charnes et al. (1978) and Banker et al. (1984), until the end of 2016, 10,300 DEA journal articles, according to the bibliography by Emrouznejad and Yang (2018), have been published. Cvetkoska and Savić (2021) provided the first bibliometric analysis on DEA in the banking industry by identifying the research hotspots in different periods. Kaffash et al. (2020) provided a literature survey of 132 DEA articles in the insurance industry from 1993 to 2018. Based on our literature survey insights, the insurance industry has not yet conducted a keyword co-occurrence analysis that helps in identifying relevant and essential research hotspots. That conclusion motivated us to conduct such a study and fill in the gap in the academic and expert literature.

We follow Cvetkoska and Savić (2021) in designing our bibliometric analysis research and conducting in-depth keyword co-occurrence analysis in different periods. We have extracted documents (journal articles and book chapters) from four relevant databases: Scopus, Web of Science, Ebsco, ProQuest. We use descriptive analytics to present the articles per year and category of the insurance sector, geographical distribution, relevancy as the top five journals and authors by citations, and the citation analysis. Despite this, our primary focus is on using text mining in the extracted abstracts of DEA insurance-related articles to create keyword co-occurrence maps for the total analyzed period (from 1993 to 2021), then for separate periods as before 2000, 2000-2010 and 2011-2021. Our article contributes to the existing literature by identifying the research hotspots in the analyzed periods and offers substantial future research directions and guidelines in the insurance industry with DEA, which will directly benefit the researchers and management practitioners in the insurance industry in developing conducting contemporary and applicative research.

The rest of the article is organized as follows. Section 2 describes the data and the methodology and tools used. Section 3 presents and analyzes the results obtained on DEA in the insurance industry regarding the annual trends, categories of the insurance industry, geographical distribution, top five journals and authors by citation, and citation analysis. Section 4 provides keyword co-occurrence maps and their analysis. Finally, section 5 provides concluding remarks.

2. DATA, METHODOLOGY AND TOOLS

The data used in our study is based on 55 DEA insurance-related articles from four relevant databases: Scopus, Web of Science, Ebsco, and ProQuest. The distribution of articles is conducted according to their dispersion per year, categories of the insurance sector, geographical distribution of articles, top five journals and authors by citation, and the top five most cited articles, has been made by using Excel. In addition, the keyword co-occurrence analysis is conducted by mining the 55 abstracts with the VOSviewer software version 1.6.17.

Van Eck and Waltman developed this software and more details can be found in Van Eck and Waltman (2010, 2014). Starting from April 2020, it provides the text mining option used in our research. Several VOSviewer applications include multiple criteria decision making (MCDM) by Yu et al. (2018); analyzing the evolution of research using Google Trends by Jun et al. (2018); social networking and academic performance by Doleck and Lajoie (2018); social media in psychology by Zyoud et al. (2018); evolution of entrepreneurship by Chandra (2018); and information sciences by Merigó et al. (2018).

3. RESULTS AND ANALYSIS

3.1 Distribution of DEA articles in the insurance industry by year

In our study, we analyzed 55 DEA insurance-related articles covering 29 (twenty-nine) years (1993-2021). The distribution of articles per year is shown in Figure 1. The highest number of published articles is identified in 2018 to 2019 (5 in each year), while four articles are published in each of the following years: 2012, 2013, and 2020. In addition, one article was published in the first seven years (1993, 1997-1999, 2002-2004), and one article in 2007 and 2017. In the last observed year, i.e. 2021, there were only two articles published, but we believe that this number should increase by the end of the year due to the durability of the issuing procedures.

Figure 1: Distribution of DEA articles in the insurance industry by year

Source: Authors Compilation

3.2 Distribution of DEA articles by insurance industry categories

Based on our analysis, 45 articles (81.8%) consider one of the following seven categories: non-life, life, composite, property-liability, health and family insurance sector, and the ownership structure in the insurance industry. Table 1 presents the distribution of DEA articles in each of the seven insurance industry categories. The highest number, i.e. 12 articles (21.8 %), are applied in the non-life insurance sector, followed by 11 or (20 %) in the non-life insurance sector, 9 (16.4 %) in the composite insurance sector, and so on.

Table 1: Distribution of DEA articles in one of the identified insurance industry categories

	Category	Number of articles
1	non-life	12
2	life	11
3	composite insurance	9
4	property-liability	8
5	health	2
6	family	2
7	ownership structure	1
	Total	**45**

Source: Authors Compilation

In addition, in the rest of the analyzed articles, i.e. 10 in total, DEA is applied in two or more insurance industry categories. The distribution of those articles per insurance category is presented in Table 2. According to Table 2, we can see that in the different categories, the articles dominate life insurance (in 5 articles), followed by non-life insurance (in 4 articles). Only once the family insurance article has been used, along with the non-insurance.

Table 2: Distribution of DEA articles in two or more of the insurance industry categories

Categories	Number of articles
property-liability insurance and ownership structure	2
life, non-life, and composite insurance	2
health and life insurance sector	1
life insurance and ownership structure	1
life, non-life and health insurance	1
non-life and family insurance	1
non-life and composite insurance	1
life, health, and property-liability insurance	1
Total	**10**

Source: Authors Compilation

3.3 Geographical distribution of DEA insurance-related articles

Based on our analysis, in 49 articles, the research considers only one country (Table 3), while in 6 articles, DEA is applied in two or more countries (Table 4). Table 3 shows that 23 countries are identified so that the data set from the U.S. is applied in 12 articles, followed by India in 6 articles, Malaysia in 4 articles, China in 3 articles, Austria, Japan, Taiwan, UAE and Pakistan in two countries each and the rest in one country. Table 4 shows that in 2 articles are from countries within the European Union, while the others are as follows: Global insurance industry from the GCC countries (UAE, Saudi Arabia, Qatar, Oman, Kuwait and Bahrain), Malaysia and the Gulf states, and the BRIC countries (Brazil, Russia, India, China and South Africa).

Table 3: Geographical distribution on DEA research within a single country

	Research in one country	Number of articles
1	U.S.	12
2	India	6
3	Malaysia	4
4	China	3
5	Austria	2
6	Japan	2
7	Taiwan	2
8	UAE	2
9	Pakistan	2
10	France	1
11	Greece	1
12	South Africa	1
13	Ghana	1
14	Ukraine	1
15	Spain	1
16	Ethiopia	1
17	North Macedonia	1
18	Ecuador	1

19	Iran	1
20	Australia	1
21	Portugal	1
22	Korea	1
23	Canada	1
	Total	**49**

Source: Authors Compilation

Table 4: Geographical distribution of DEA research in at least two countries

	Research in two or more countries	**Number of articles**
1	European Union	2
2	Global	1
3	GCC countries	1
4	Malaysia and the Gulf States	1
5	BRIC countries	1

Source: Authors Compilation

3.4 The top five journals in the insurance industry based on DEA articles

According to the total number of citations of each of the 55 articles published in 42 journals, we have identified the top five journals. Those journals are presented in Table 5. The first rank belongs to the *Journal of Productivity Analysis,* with a total of 294 citations, followed by *The Geneva Papers on Risk and Insurance-Issues and Practice* (269 citations), the Journal of *Empirical Economics* (176 citations), the *Journal of Banking & Finance* (175 citations), and *Omega* (105 citations).

Table 5: The top five journals in the insurance industry based on DEA articles

	Journal	**Citations**
1	*Journal of Productivity Analysis*	294
2	*The Geneva Papers on Risk and Insurance - Issues and Practice*	269
3	*Empirical Economics*	176
4	*Journal of Banking & Finance*	175
5	*Omega*	105

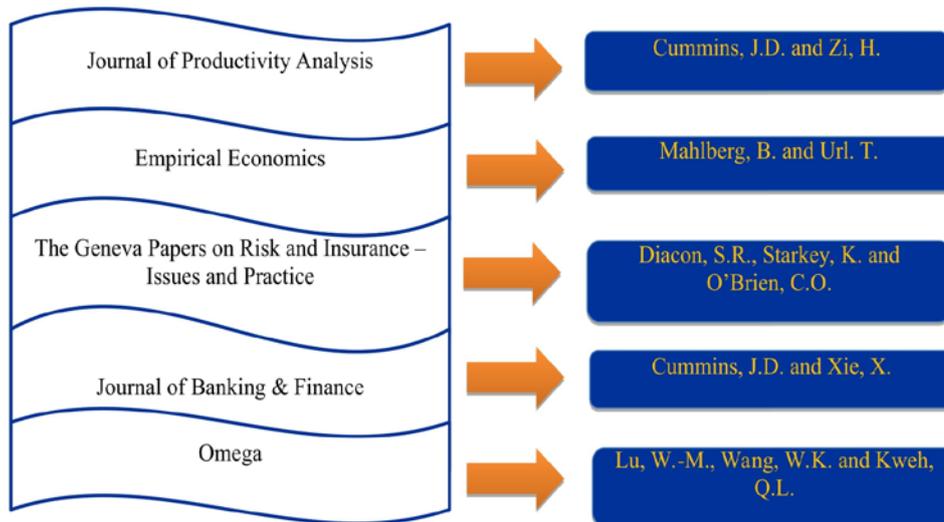## 3.5 Top DEA insurance related-article authors based on citations

Table 6 presents the top authors based on the total number of citations of DEA insurance-related articles published in the identified top five journals in Section 3.4. According to Table 6, we can see that Cummins, J.D. has the highest number of citations (322) and has a rank of 1, followed by Zi, H. (196 citations), Mahlberg, B. and Url, T. (176 citations), Diacon, S.R, Starket, K. And O'Brien C.O. (150 citations) and Xie, X (126 citations).

Table 6: Top DEA insurance related-article authors based on citations

	Author(s)	Citations
1	Cummins, J.D.	322
2	Zi, H.	196
3	Mahlberg, B. and Url, T.	176
4	Diacon, S.R, Starkey, K. and O'Brien C.O.	150
5	Xie, X.	126

Source: Authors Compilation

Also, in Figure 2 below, we present the most cited authors in DEA related articles in the insurance industry in each of the top five journals. Cummins, J.D., and Zi, H. are most cited in the *Journal of Productivity Analysis*; Mahlberg, B., and Url, T. in the *Empirical Economics*; Cummins, J.D., and Xie, X. in the *Journal of Banking and Finance*; Diacon, S.R., Starkey, K., and O'Brien, C.O. in *The Geneva Papers on Risk and Insurance – Issues and Practice*; and Lu, W.-M., Wang, W.K., and Kweh, Q.L. in *Omega*. Based on Table 6, we have identified that Cummins, J. D. is the most-cited author, and from Figure 2, we can see that he is the most cited in two of the identified top five journals.

Figure 2: Top DEA insurance-related article authors based on citations

Source: Authors Compilation

3.6 The top five most cited DEA articles in the insurance industry

Table 7 presents the top five most-cited DEA-related articles in the insurance industry in the top identified journals.

The article with the highest number of citations (196) is "Comparison of Frontier Efficiency Methods: An Application to the U.S. Life Insurance Industry" by Cummins and Zi (1998) that is published in the *Journal of Productivity Analysis*. This article uses a sample of 445 life insurance companies in the U.S. between 1988 to 1992. With 20 citations less, it is the second most cited DEA article in the insurance industry, entitled: "Effects of the single market on the Austrian insurance industry" by Mahlberg and Url (2003), published in *Empirical Economics*. It considers 52 companies from the following three categories of the insurance sector: health, life, and property-liability insurance in Austria, and the observed period is from 1992 to 1999. The article "Size and efficiency in

European long-term insurance companies: An international comparison" by Diacon et al. (2002), published in *The Geneva Articles on Risk and Insurance-Issues and Practice*, has 150 citations and is ranked third. It is applied in 15 European countries, and the sample consists of 454 insurance companies observed between 1996 and 1999. The article "Mergers and acquisitions in the US property-liability insurance industry: Productivity and efficiency effects" by Cummins and Xie (2008), with 126 citations, ranks fourth. This article was published in the *Journal of Banking & Finance*. It considers U.S. companies in the property-liability insurance sector observed in the period from 1994 to 2003 so that the total insurers-year observations are 14,592 and on average 768 insurers per year. The article by Fecher et al. (1993), entitled "Productive performance of the French Insurance Industry," published in the *Journal of Productivity Analysis*, is cited 98 times and ranked in fifth place. It considers a sample of 327 insurance companies (84 from the life and 243 from the non-life insurance sector) in France.

Table 7: The top five most cited DEA articles in the insurance industry

	Authors	Year	Title	Journal	Citations
1	Cummins, J.D., Zi, H.	1998	Comparison of Frontier Efficiency Methods: An Application to the U.S. Life Insurance Industry	*Journal of Productivity Analysis*	196
2	Mahlberg, B., Url, T.	2003	Effects of the single market on the Austrian insurance industry	*Empirical Economics*	176
3	Diacon, S.R., Starkey, K., O'Brien, C.O.	2002	Size and efficiency in European longterm insurance companies: An international comparison	*The Geneva Papers on Risk and Insurance - Issues and Practice*	150
4	Cummins, J.D., Xie, X.	2008	Mergers and acquisitions in the US property-liability insurance industry: Productivity and efficiency effects	*Journal of Banking & Finance*	126
5	Fecher, F., Kessler, D., Perelman, S., Pestieau, P.	1993	Productive performance of the French insurance industry	*Journal of Productivity Analysis*	98

Source: Authors Compilation

4. Identification of research hotspots for DEA in the insurance industry based on mining keywords in the abstracts

The first bibliometric analysis on DEA in banking to identify the research hotspots based on mining keywords in the 791 abstracts in 34 years (1986–2019) was carried out by Cvetkoska and Savić (2021). In this article, we follow Cvetkoska and Savić (2021) work and provide in-depth keyword co-occurrence analysis on DEA articles in the insurance industry in the period under review (1993-2021) by setting the minimum number of keyword co-occurrence to be 2 (two). In addition, we created separate keyword co-occurrence maps for the periods before-2000, 2000-2010, and 2011-2021 to identify whether and how researchers changed, transformed or integrated their evolving research interests. We have also created an overlay visualization map for the last six years (2016-2021) to identify the research hotspots in each year. In this case, we do not set a minimum number of keyword occurrences. Instead, we use the VOSviewer software 1.6.17 to conduct abstract mining and create and visualize the maps.

4.1 DEA research hotspots in the insurance industry (1993-2021)

The creation of the keywords co-occurrence map based on text data, more specifically abstracts from the 55 journal articles and book chapters from Scopus, Web of Science, Ebsco and ProQuest, is made by using the VosViewer Software. We have chosen to create a co-occurrence map based on text data, and in the 55 abstracts, we ignore structured abstract labels and copyright statements (if any). The counting method is binary counting. The total number of extracted keywords is 1318, and we have set the minimum number of keyword occurrences to be two so that 243 keywords satisfy the set threshold. Then, a relevance score is calculated for each keyword, and 60% of the most relevant keywords are selected by default, in our case, 146 keywords. We have thoroughly checked the extracted 146 keywords and have excluded the country names (because we provided geographical distribution in Section 3.3) and keywords that are not adequate for our analysis (presented in Appendix 1a). As a result, a total of 42 keywords were eliminated from our side, and our keyword co-occurrence map consists of a final 104 keywords. The network visualization for the keywords co-occurrence map for the observed period is given in Figure 3.

Based on Figure 3, we can see that the keywords are grouped into seven clusters so that those that have the same colour belong to the same cluster. The network visualization consists of items (labels and circles) and lines. The larger the item and the label, the higher the weight of the keyword. The founders of the VosViewer software, van Eck and Waltman (2020) point out that if some labels are not presented, that does not appear to be an overlapping problem. If between two items, the line is thicker and the distance is shorter, that indicates that their relationship is stronger (Cvetkoska and Savic, 2021). Also, the strength of a link is given by a binary number, which is a positive value. Furthermore, if this value is higher, it indicates that the link is stronger, and the total link strength (TLS) of one item is the strength of the links of one item related to the other items, calculated as a sum (van Eck and Waltman, 2020).
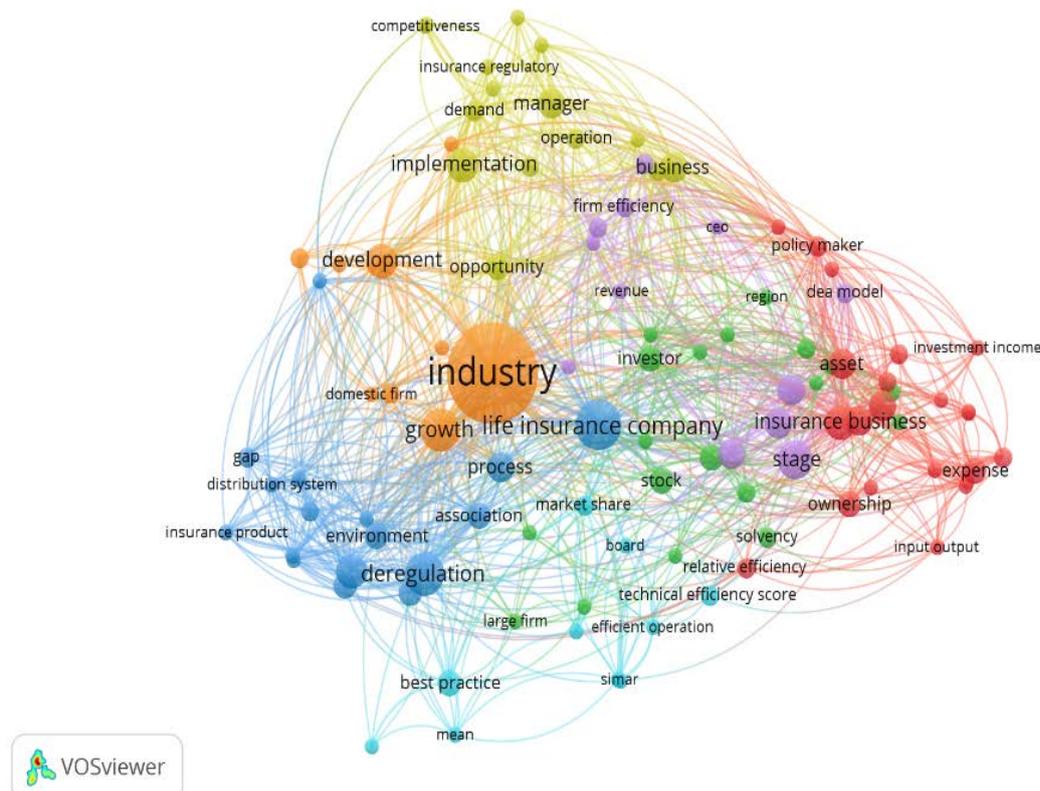
Source: Authors Compilation

In Table 8, we present the largest item in each cluster, its links, occurrences and the total link strength. For example, the largest item in the orange cluster is industry, and it is also the largest on the whole map. It has 84 links, 26 occurrences and a TLS of 160.

**Table 8. The largest item in each cluster for the period 1993-2021**

	Clusters	Keywords	Links	Occurrences	Total link strength (TLS)
1	orange	industry	84	26	160
2	violet	stage	40	7	58
3	blue	life insurance company	59	82	10
4	yellow	implementation	30	38	6
5	red	insurance business	36	46	6
6	green	investor	26	27	4
7	light blue	best practice	15	18	4

Source: Authors Compilation

4.2 DEA research hotspots in the insurance industry (before 2000, 2000-2010 and 2011-2021)

In order to show whether and how the keywords in the analyzed DEA articles in the insurance industry changed in the different periods (before 2000, 2000-2010, and 2011-2021), we created a keyword co-occurrence map for each period as described in Section4.1, but the minimum number of keywords co-occurrence is set to 1 (one).

From the abstracts of 5 articles published before 2000, 180 keywords are extracted, out of which 108 (60%) are selected by default. By eliminating 32 that consist of country names and irrelevant keywords for our analysis, the sample of keywords is decreased to 76. The eliminated keywords are presented in Appendix 1b. The network visualization for keyword co-occurrence map on DEA insurance related articles for the period before 2000 is shown in Figure 4. From Figure 4, we can see that the keywords are grouped into five clusters. In this period, we have investigated the efficiency of life and non-life insurers and the ownership structure (stock and mutual). Despite DEA, some studies also used parametric methods, and Cummins et al. (1999) set agency theoretic hypothesis about the organizational form. The data sets considered included the insurance companies' data from the developed markets of the U.S., France, Italy, and Japan.
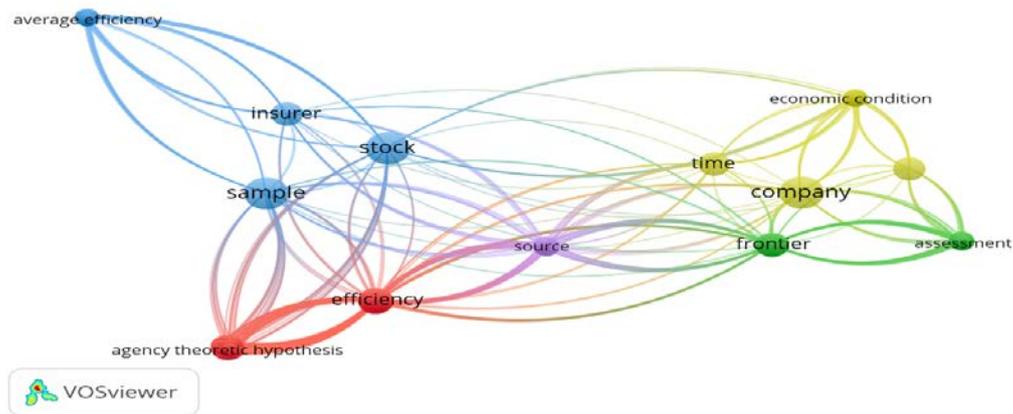
Figure 4: A network visualization of keyword co-occurrence in DEA insurance-related articles prior to 2000
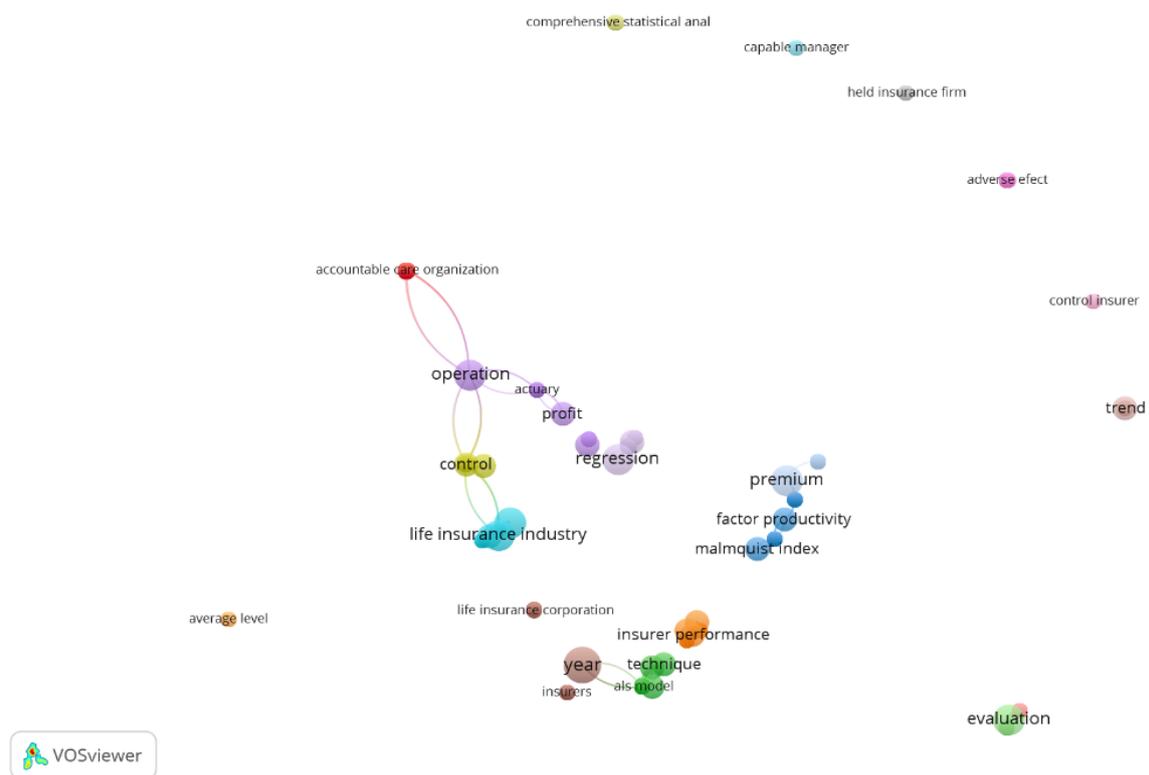
Source: Authors Compilation

The keyword co-occurrence map on DEA related articles in the insurance industry presented in Figure 5 was created by mining the abstracts in this period, so there are 470 keywords in total, out of which 60% (282 keywords) are selected as relevant by default, and 104 are eliminated. The eliminated keywords are given in Appendix 1c. The map consists of 178 keywords grouped into 12 clusters. Between 2000 and 2010, we considered more specialized lines of business life, non-life, health, and property-liability insurers. We measured their efficiency, the effects of mergers and acquisitions, liberalization, and deregulation on efficiency, and we also compared the efficiency between domestic and foreign insurers, namely the influence of the ownership structure. In the analysis, we have used DEA, DEA with Malmquist, stochastic frontier and DEA, a two-stage DEA procedure proposed by Simar and Wilson (2007) and DEA with econometric methods to investigate the determinants of efficiency. The DEA study is applied to data sets from Austria, China, the United Arab Emirates (UAE), Portugal, India, Greece, Ukraine, and Taiwan and in one study, a set of various 15 European insurance industries are analyzed.
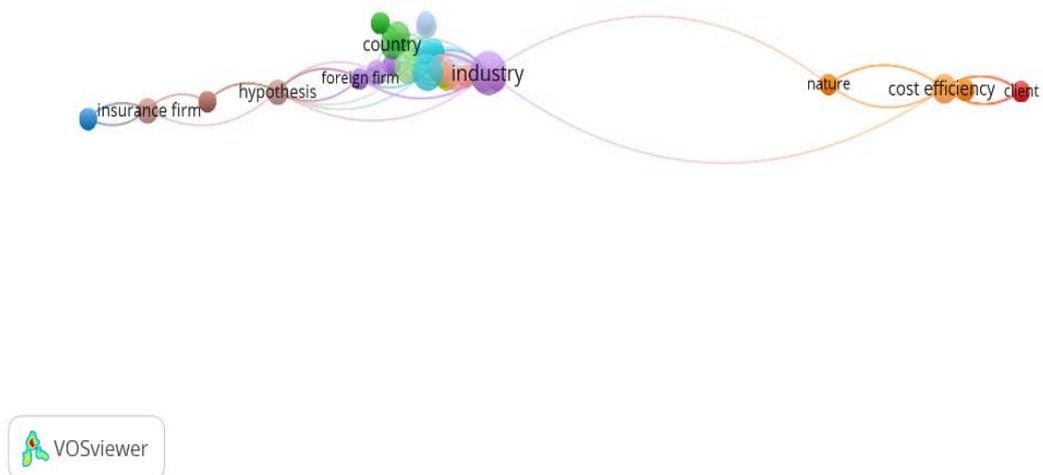
Figure 5: A network visualization of keyword co-occurrence in DEA insurance-related articles in the period 2000-2010

Source: Authors Compilation

From the 34 abstracts in the last observed period 2011–2021, 879 keywords were extracted, out of which 527 (60%) were selected by default. By eliminating 125 (presented in Appendix 1d), 402 keywords remained and were used to create the network visualization of the co-occurrence map on DEA insurance-related articles presented in Figure 6. From Figure 4, we can see that the keywords are grouped into 19 (nineteen) clusters. This period covers a variety of life, non-life, property-liability, private health insurers, general insurance, life-micro insurance portfolio of insurance companies, and ownership (local and foreign). Furthermore, we investigate efficiency, change in the productivity of insurers, risk management efficiency, the impact of the characteristics of corporate governance as well as the financial sector reforms on insurance efficiency, whether the insurers that have applied ERP systems to improve their efficiency have increased their efficiency, the link between CEO compensation and the insurers' performance, and managerial ability. In the analyzed studies, two-stage DEA (DEA and truncated bootstrapped regression, DEA and Panzar-Rosse H statistic method, DEA and Simar-Wilson (2007) procedure, two-stage network DEA, DEA and Tobit regression, DEA-translog model), but also DEA and Malmquist, DEA and the Analytic Hierarchy Process (AHP), DEA-Malmquist-Tobit were used. In addition, the variety of countries is the highest in this period, so DEA is applied in the U.S., China, India, Malaysia, Pakistan, Korea, Ghana, Ecuador, Taiwan, Japan, Republic of Macedonia, Australia, South Africa, the BRIC countries, Gulf Cooperation Council Countries, and on a global data set.
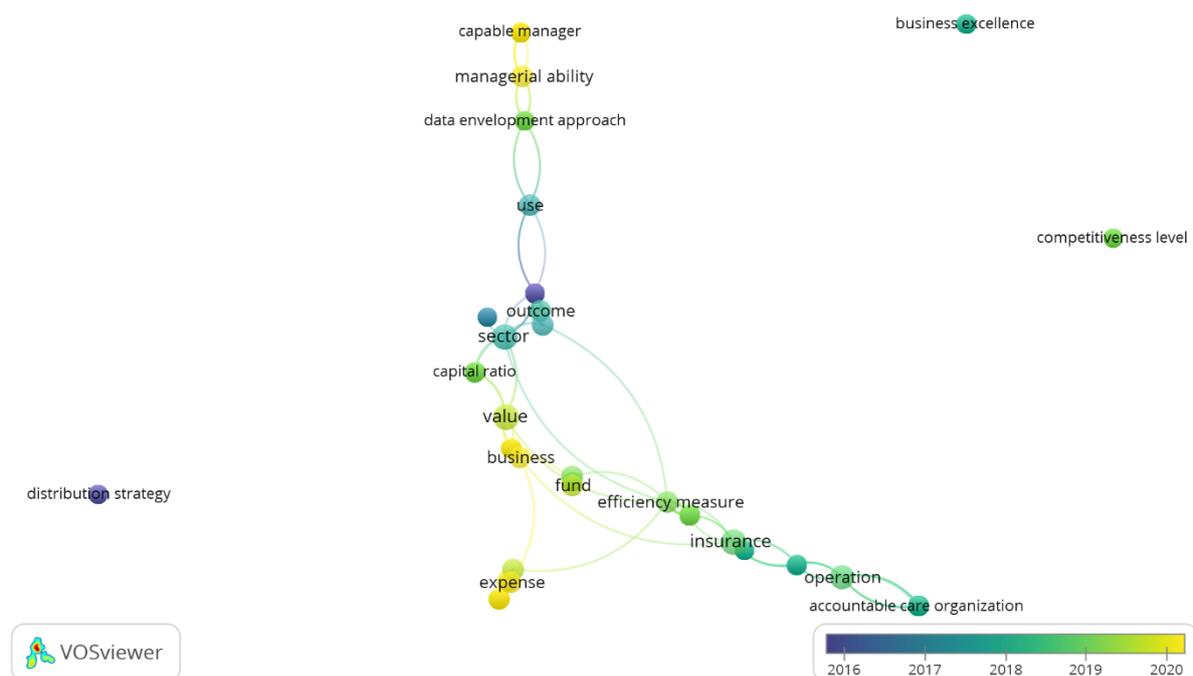

Figure 6: A network visualization of keyword co-occurrence in DEA insurance-related articles in the period 2011-2021

Source: Authors Compilation

4.3 Research hotspots in the insurance industry in the period 2016-2021 and future trends

In Figure 7, we present an overlay visualization regarding the eighteen DEA mined abstracts in the period 2016-2021, based on which we can see the research hotspots in the most recent period. In creating this map, we have considered 254 keywords (562 were the total, 60% of the most relevant were selected by default, which gives a total of 337, and we have eliminated 83). The eliminated keywords by us are given in Appendix 1e. In this period we have investigated the following issues: how the financial sector reforms affect the efficiency and productivity (Noreen and Ahmad, 2016), efficiency and productivity of property-liability insurers (Cummins and Xie, 2016), efficiency-solvency linkage (Sinha, 2017) the performance evaluation of life, general insurers and its' ownership (Nourani et al. 2018), the corporate performance (Tone et al. 2019) and the corporate governance effect (board characteristics) on efficiency of life insurance companies (Alhassan and Boakye, 2020), the efficiency and competitiveness (Lee et al. 2018), efficiency of the Medicare plans (Brockett et al. 2018), efficiency of the publicly held insurance industry (Copeland and Cabanda, 2018), the determinants of the insurance profitability (Camino-Morgo and Bermudez-Barrezueta, 2019), the effect of risk-based capital on efficiency, productivity and competitiveness (Lee et al. 2019), the life micro-insurance portfolio of insurance companies (Savitha et al. 2019), the relationship between efficiency and profitability of insurance companies (Eling and Jia, 2019), the performance of non-life companies and its determinants (Ilyas and Rajasekaran, 2019), the measure and rank efficiency of general insurance companies (Meher et al. 2020), the risk-management efficiency and determinants of non-life insurance companies (Noreen, 2020), private health insurance (Nguyen and Worthington, 2021) and lastly managerial ability (Banker et al. 2021).

In most studies, the two-stage DEA (bootstrap DEA, network DEA, DEA+Tobit regression, DEA+Panzar-Rosse H statistic method, and bootstrap Malmquist index) and the three methods: DEA-Malmquist-Panzar-Rosse, DEA-Malmquist-Tobit regression were used.

Figure 7: Overlay visualization for keywords in DEA insurance-related articles in the period 2016-2021

Source: Authors Compilation

As for future research directions in the insurance industry, researchers could investigate the multi-factor performance of the insurance companies before, during, and post-Covid-19 to explore whether the use of modern technologies impacts the performance of insurers by using stage-DEA (DEA with parametric methods, DEA+multi criteria decision making (MCDM) methods, DEA + machine learning models as well as measure their organizational (managerial) ability and identify the determinants of that ability by using the Banker et al. (2021) measure.

CONCLUSION

In our study, we provide analysis and visualization of the bibliometric data on DEA articles in the insurance industry published in journals and as book chapters listed in the following relevant databases: Scopus, Web of Science, Ebsco and ProQuest, with a focus on the identification of research hotspots concerning the analysis of keywords. We use descriptive statistics and text mining as the main methods in our analysis. The analysis covers 55 DEA insurance articles published over 29 years (from 1993 to 2021).

We have summarized the main findings from this study analysis in 7 points: **(1)** Based on the distribution of DEA articles in insurance in the observed period, the highest number of articles (5) has been published in the period 2018-2019, and four articles each were published in the following years: 2012, 2013 and 2020; **(2)** In 45 articles, DEA is applied in one of the seven identified insurance categories, which leads the non-life with 12 articles, life with 11 articles, and composite insurance with nine articles. In addition, in the other ten articles, DEA is applied in at least two insurance categories; **(3)** Based on the geographical distribution of the data sets used, the data set is used in a single country in 23 articles, so that the highest number is applied in the U.S. (12 articles), followed by India (6), and Malaysia (4). Also, we have found five studies where DEA in the insurance industry is applied to a data set from at least two countries, so that in two studies, the EU countries are considered, followed by the BRIC and GCC countries, Malaysia, and the Gulf States, and on a global data set (each considered in one separate article); **(4)** The five top journals where DEA insurance-related articles are published, based on citations, are: the *Journal of Productivity Analysis* in first place with 294 citations, followed by *The Geneva Papers on Risk and Insurance - Issues and Practice with 269 citations, Empirical Economics with 176 citations, and Journal of Banking & Finance with 175 citations, and Omega with 105 citations;* **(5)** The top authors on DEA articles in insurance based on their citations in the top identified journals are: Cummins, J.D. with the highest number of citations (322) and has a rank 1, followed by Zi, H. (196 citations), Mahlberg, B. and Url, T. (176 citations), Diacon, S.R, Starket, K. And O'Brien C.O. (150 citations), and Xie, X (126 citations); **(6)** We have identified the top five most-cited DEA insurance-related articles so that the article with the highest number of citations (196) is "Comparison of Frontier Efficiency Methods: An Application to the U.S. Life Insurance Industry" by Cummins and Zi (1998) published in *Journal of Productivity Analysis*. This article uses a sample of 445 life insurance companies in the U.S. and observes the period between 1988 and 1992; **(7)** Based on in-depth keyword co-occurrence analysis by mining the 55 abstracts, we provide network visualization for the whole period (1993-2021), but also each of the following periods: before 2000, 2000-2010, and 2011-2021. The number of keywords is large and diverse, which implies that the interest of researchers has increased and changed over the years. Our overlay visualization for the last six years (2016-2021) highlights the most recent research hotspots.

Our detailed analysis shows that efficiency and its determinants, corporate governance, risk management efficiency, and managerial ability, have been investigated during this time. In most studies, two-stage DEA is used, i.e., bootstrap DEA, network DEA, DEA+Tobit regression, DEA+Panzar-Rosse H statistic method, and bootstrap Malmquist index), but also three methods are used: DEA-Malmquist-Panzar-Rosse, DEA-Malmquist-Tobit regression.

We do not claim that all DEA insurance-related articles (published in journals and as book chapters) listed in the four databases are included so that some data may be omitted. Instead, the findings of our analysis aim to contribute to researchers and insurance practitioners as an empirical and applicative point for initiating and developing ongoing and future research.

REFERENCES

Alhassan, A.L. and Boakye, M.-A.A. (2020), "Board characteristics and life insurance efficiency in South Africa", *Pacific Accounting Review*, Vol. 32 No. 2, pp. 217–237.

Banker, R.D., Charnes, A. and Cooper, W.W. (1984), "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*, Vol. 30 No. 9, pp. 1078–1092.

Banker, R.D., Luo, J. and Oh, H. (2021), "Measuring Managerial Ability in the Insurance Industry", *Data Envelopment Analysis Journal*, Vol. 5 No. 1, pp. 115–143.

Brockett, P.L., Golden, L.L. and Yang, C.C. (2018), "Potential 'Savings' of Medicare: The Analysis of Medicare Advantage and Accountable Care Organizations", *North American Actuarial Journal*, Vol. 22 No. 3, pp. 458–472.

Camino-Mogro, S. and Bermúdez-Barrezueta, N. (2019), "Determinants of profitability of life and non-life insurance companies: evidence from Ecuador", *International Journal of Emerging Markets*, Vol. 14 No. 5, pp. 831–872.

Chandra, Y. (2018), "Mapping the evolution of entrepreneurship as a field of research (1990–2013): A scientometric analysis", *PLOS ONE*, Vol. 13 No. 1, e0190228.

Charnes, A., Cooper, W.W. and Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research*, Vol. 2 No. 6, pp. 429–444.

Copeland, M.K. and Cabanda, E. (2018), "Efficiency Analysis of the U.S. Publicly Held Insurance Industry: A Two-Stage Efficiency Model", *International Journal of Information Systems in the Service Sector (IJISSS)*, Vol. 10 No. 1, pp. 1–15.

Cummins, J.D. and Xie, X. (2008), "Mergers and acquisitions in the US property-liability insurance industry: Productivity and efficiency effects", *Journal of Banking & Finance*, Vol. 32 No. 1, pp. 30–55.

Cummins, J.D. and Xie, X. (2016), "Efficiency and Productivity in the US Property-Liability Insurance Industry: Ownership Structure, Product and Distribution Strategies", in Zhu, J. (Ed.), *Data Envelopment Analysis: A Handbook of Empirical Studies and Applications*, Springer US, Boston, MA, pp. 113–163.

Cummins, J.D. and Zi, H. (1998), "Comparison of Frontier Efficiency Methods: An Application to the U.S. Life Insurance Industry", *Journal of Productivity Analysis*, Vol. 10 No. 2, pp. 131–152.

Cummins, J.D., Weiss, M.A. and Zi, H. (1999), "Organizational Form and Efficiency: The Coexistence of Stock and Mutual Property-Liability Insurers", *Management Science*, Vol. 45 No. 9, pp. 1254–1269.

Cvetkoska, V. and Savić, G. (2021), "DEA in Banking: Analysis and Visualization of Bibliometric Data", *Data Envelopment Analysis Journal*, Vol. 5 No. 2, pp. 455–485.

Diacon, S.R., Starkey, K. and O'Brien, C. (2002), "Size and Efficiency in European Long–term

Insurance Companies: An International Comparison", *The Geneva Papers on Risk and Insurance - Issues and Practice*, Vol. 27 No. 3, pp. 444–466.

Doleck, T. and Lajoie, S. (2018), "Social networking and academic performance: A review", *Education and Information Technologies*, Vol. 23, No. 1, pp. 435–465.

Eling, M. and Jia, R. (2019), "Efficiency and profitability in the global insurance industry", *Pacific-Basin Finance Journal*, Vol. 57, p. 101190.

Emrouznejad, A. and Yang, G. (2018), "A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016", *Socio-Economic Planning Sciences*, Vol. 61, pp. 4–8.

Fecher, F., Kessler, D., Perelman, S. and Pestieau, P. (1993), "Productive performance of the French insurance industry", *Journal of Productivity Analysis*, Vol. 4 No. 1–2, pp. 77–93.

Ilyas, A.M. and Rajasekaran, S. (2019), "An empirical investigation of efficiency and productivity in the Indian non-life insurance market", *Benchmarking: An International Journal*, Vol. 26 No. 7, pp. 2343–2371.

Jun, S.P., Yoo, H.S. and Choi, S. (2018), "Ten years of research change using Google Trends: From the perspective of big data utilization and applications", *Technological Forecasting and Social Change*, Vol. 130, pp. 69-87.

Kaffash, S., Azizi, R., Huang, Y. and Zhu, J. (2020), "A survey of data envelopment analysis applications in the insurance industry 1993–2018", *European Journal of Operational Research*, Vol. 284 No. 3, pp. 801–813.

Lee, H.S., Chang, F.F., Nassir, A.M. and Razak, N.H.A. (2018), "Does efficienct promote competitiveness of the insurance industry?", *Journal of Business Economics and Management*, Vol. 19 No. 4, pp. 566–591.

Lee, H.S., Chang, F.F., Nassir, A.M. and Razak, N.H.A. (2019), "Impacts of Risk-Based Capital Regulation in Malaysian Islamic Insurers (Takaful)", *Asian Academy of Management Journal of Accounting and Finance (AAMJAF)*, Vol. 15 No. 1, pp. 27–59.

Mahlberg, B. and Url, T. (2003), "Effects of the single market on the Austrian insurance industry", *Empirical Economics*, Vol. 28 No. 4, pp. 813–838.

Meher, K., Asfawu, A., Muthuraman, M. and Satapathy, S.K. (2020), "Assessment of measurement and ranking of technical efficiencies of Ethiopian general insurers", *Problems and Perspectives in Management*, Vol. 18 No. 4, pp. 334–350.

Merigó, J. M., Pedrycz, W., Weber, R.. and de la Sotta, C. (2018), "Fifty years of Information Sciences: A bibliometric overview", *Information Sciences*, Vol. 432, pp. 245-268.

Nguyen, L. and Worthington, A.C. (2021), "Industry regulation, fund characteristics, and the efficiency of Australian private health insurers", *Accounting & Finance*, Vol. 61 No. 1, pp. 781–801.

Noreen, U. (2020), "The Estimation of Risk Management Efficiency and its Determinants", *Jurnal Ekonomi Malaysia*, Vol. 54 No. 1, pp. 69–79.

Noreen, U. and Ahmad, S. (2016), "Cost Efficiency and Total Factor Productivity: An Empirical Analysis of Pakistan's Insurance Sector", *The Lahore Journal of Economics,* Vol. 21 No. 1, pp. 123–150.

Nourani, M., Devadason, E.S. and Chandran, V. (2018), "Measuring technical efficiency of insurance companies using dynamic network DEA: An intermediation approach", *Technological and Economic Development of Economy*, Vol. 24 No. 5, pp. 1909–1940.

Savitha, B., Banerjee, S. and Shetty, A. (2019), "Product diversification versus technical efficiency of conglomerate life microinsurance companies: evidence from India", *The Geneva Papers on Risk and Insurance - Issues and Practice*, Vol. 44 No. 3, pp. 527–547.

Simar, L. and Wilson, P.W. (2007), "Estimation and inference in two-stage, semi-parametric models of production processes", *Journal of Econometrics*, Vol. 136 No. 1, pp. 31–64.

Sinha, R.P. (2017), "Efficiency-solvency linkage of Indian general insurance companies: a robust non-parametric approach", *Eurasian Economic Review*, Vol. 7 No. 3, pp. 353–370.

Tone, K., Kweh, Q.L., Lu, W.-M. and Ting, I.W.K. (2019), "Modeling investments in the dynamic network performance of insurance companies", *Omega*, Vol. 88, pp. 237–247.

van Eck, N. J. and Waltman, L. (2020), "Manual for VOSviewer Version 1.6.15".

van Eck, N.J. and Waltman, L. (2010), "Software survey: VOSviewer, a computer program for bibliometric mapping", *Scientometrics*, Vol. 84 No. 2, pp. 523–538.

van Eck, N.J. and Waltman, L. (2014), "Visualizing Bibliometric Networks", in Ding, Y., Rousseau, R. and Wolfram, D. (Eds.), *Measuring Scholarly Impact*, Springer International Publishing, Cham, pp. 285–320.

Yu, D., Wang, W., Zhang, W. and Zhang, S. (2018). "A bibliometric analysis of research on multiple criteria decision making", *Current Science*, Vol. 114 No. 4, pp. 747-758.

Zyoud, S.H., Sweileh, W.M., Awang, R. and Al-Jabi, S.W. (2018), "Global trends in research related to social media in psychology: Mapping and bibliometric analysis", *International Journal of Mental Health Systems*, Vol. 12 No. 4.

Appendix 1a. Eliminated keywords for the co-occurrence map in the period 1993-2021

Eliminated keywords
adoption, assessment, average, beginning, cause, challenge, China, choice, claim, consistent, DEA result, decline, empirical result, end, entry, European country, European Union, further development, future development, Germany, important issue, India, influence, interest, key concern, kind, long run, main finding, Malaysia, MPI, nature, new information, number, percent, point, reduction, research limitation implications , significant increase, Takaful insurer, Takaful operator, UAE, view

Source: Authors Compilation

Appendix 1b. Eliminated keywords for the co-occurrence map in the before-2000 period

Period	Eliminated keywords
Pre-2000	article, choice, consistent, data basis, difference, dispersion, display, distribution, efficiency study, estimation method, French company, incentive, Japanese life insurance company, main finding, mutual, new information, non-parametric result, paper, previous, reference, research, researcher, return, richer interpretation, sample, scale, significant effect, Spanish savings bank, technique, US life insurer, use, wide range

Source: Authors Compilation

Appendix 1c. Eliminated keywords for the co-occurrence map in the 2000-2010 period

Period	Eliminated keywords
2000-2010	adoption, Anglo minimal regulated UK industry, approach firm, approach use, article, Austrian experience,, Austrian insurance market, Austrian insurance regulatory, Austrian insurer, authority, benefit, broad agreement, Canadian i&h insurance industry, Canadian life, central, challenge, Chinese life insurance industry, Chinese market, claim, combination, concern, conflict, consequence, continental maximal regulated German, course, Denmark, direct implication, direction, dual impact, Eastern European country, efficient value, empirical result, entry, European counterpart, European country, European insurance industry, European insurer, European specialist, examination, firm specific data, firm specific estimate, first attempt, flow, further development, future competition, future development, German industry, German insurance company, German regulation, Germany, good stead, greater competition, hand UK insurer, i&h, importance, important issue,

	important light, India, interest, interpretation, intra, investigation, issue, key concern, key determinant, Korea, low level, main finding, major structural change, mutual form, mutuality, new variant, non state, number, optimism, Philippine life insurance industry, Philippines, positive effect, potential, potential improvement, primary concern, recent year, relationship, significant growth, significant heterogeneity, significant increase, similar benefit, similar pattern, Spanish insurance industry, Spanish insurance market, strong evidence, Sweden, Taiwan, Taiwanese, Thai life insurance business, Thailand, theoretical concern, UK cost efficiency, UK Spain, Ukrainian insurance firm, Ukrainian insurance industry, Unit price, United Arab Emirates, Wide variation

Source: Authors Compilation

Appendix 1d. Eliminated keywords for the co-occurrence map in the 2011-2021 period

Period	Eliminated keywords
2011-2021	adjacent year, adoption, African insurance market, alternative, annual basis, attempt, attention, Australian private health insurance industry, Australian private health insurer, benefit, Brazil, BRIC country, Caribbean insurer, cent, central objective, chapter, China, claim, comparing, consideration, consistent, context, conversion, coordination, creation, distinguishing, document, Ecuadorian insurance sector, efficiency pre, empirical evidence, empirical finding, entry, European country, European life insurance industry, extant evidence, first empirical assessment, first research, first result, first time, following variable, France, Germany, ground, handbook, high level, higher level, idea, inappropriate use, incorporation, Indian general insurance sector, Indian insurance market, Indian insurance statistic, Indian life, interaction, interest, Iranian insurance market, Islamic country, Italy, Japan, Japanese non-life insurance market, key reason, lack, large extent, Latin American, light, link, long run, Macedonia, Macedonian insurance company, main determinant, main objective, main source, Malaysian insurance company, marginal impact, means, monitor, new finding, number, opportunity, order, overall result, Pakistan insurance sector, percent, Philippine, point, positive correlation, post, post conversion, potential negative association, pre, present study, previous study, reason, regard, relative influence, republic, respect, response, Russia, sample public, separation, shift, side, significant effect, significant finding, significant impact, significant relationship, source, South African life insurance industry, Spain, study period, Taiwan, Takaful industry, Takaful insurance, Takaful insurance market, Takaful insurer, Takaful operations, understanding, underwriting discipline, unfavorable impact, US insurance industry, US property liability insurer, whole, whole observed period, world the BRIC

Source: Authors Compilation

Appendix 1d. Eliminated keywords for the co-occurrence map in the 2016-2021 period

Period	Eliminated keywords
2016-2021	adjacent year, adverse effect, attention, Australian private health insurance industry, Australian private health, insurer, better efficiency, Caribbean insurer, central objective, chapter, claim, consistent, context, contributing factor, DMU3, document, Ecuadorian insurance sector, efficient implementation, empirical evidence, empirical, finding, extant evidence, family Takaful operators, first research, first time, following variable, general segment, greater benefit, handbook, high usage, higher managerial ability, idea, inappropriate use, increase, India, Indian general insurance sector, Indian insurance market, Indian insurance statistic, key reason, large extent, Latin American, light, luck, Malaysian insurance company, marginal impact, means, Medicaid, Medicare, Medicare advantage, Medicare beneficiary, Medicare market penetration, Medicare obtainable, Medicare plan, Medicare plan, Medicare privatization, Medicare reform, Medicare reform effort, Motivation, new finding, nonspecialized segment, outstanding claim, Pakistan insurance, percent, Philippine, present study, previous study, reasonable level, response, sample public, sector, side, significant effect, significant finding, significant impact, significant relationship, structural change, substantial scope, Takaful industry, Takaful insurance, Takaful insurer, Takaful market, Takaful operation, underwriting, US insurance industry, year

Source: Authors Compilation