EKOIST Journal of Econometrics and Statistics, 38, 243-264

DOI: 10.26650/ekoist.2023.38.1242702
http://ekoist.istanbul.edu.tr

Received: 26.01.2023
Accepted: 12.04.2023

İSTANBUL ÜNİVERSİTESİ YAYINEVİ

EKOIST Journal of Econometrics and Statistics

RESEARCH ARTICLE / ARAŞTIRMA MAKALESİ

Health Econometrics Research: A Bibliometric Analysis from 1991 to 2020

Ebru Çağlayan Akay*, Merve Ertok Onurlu, Özlem Ergüt*****

Abstract
As a discipline, econometrics provides quantitative insights for many fields of economics, and as a result, many subfields of econometrics have emerged over time. "Health Econometrics" is one of those subfields, which employs econometric theory for the issues in health economics. The number of studies gathering econometrics and health economics, and thereby health econometrics, increased over time, particularly during the 1990s. There is a substantial body of literature in health economics that shares insights on published materials. However, the number of research that use bibliometric analysis to study trends and the present state of health econometrics is limited. This research intends to investigate published materials in health econometrics from a variety of perspectives.

To do this, data from publications with appropriate subject characteristics in the EconLit database were collected between January 1991 and December 2020. The primary methodologies in the study were bibliometric analysis and scientific mapping. The overall findings indicate that the number of publications has grown significantly over the previous 60 years, with the highest contributing writers primarily based in American institutions. In other words, health econometrics is gaining popularity among academics in the United Kingdom and the United States.

Keywords
Health Econometrics, Bibliometric Analysis, Science Mapping, EconLit

* Ebru Çağlayan Akay (Prof. Dr.), Marmara University, Faculty of Economics, Department of Econometrics, İstanbul, Turkiye. E-mail: ecaglayan@marmara.edu.tr ORCID: 0000-0002-9998-5334

** Corresponding author: Merve Ertok Onurlu, (Assist. Prof.), Canakkale Onsekiz Mart University, Biga Faculty of Economics and Administrative Sciences, Çanakkale Turkiye E-mail: m.ertok.onurlu@comu.edu.tr ORCID: 0000-0001-5187-1568

*** Özlem Ergüt, (Assoc. Prof.), Marmara University, Faculty of Economics, Department of Econometrics, Istanbul, Turkiye. E-mail: ozlem.ergut@marmara.edu.tr ORCID: 0000-0002-4076-8343

To cite this article: Caglayan Akay, E., Ertok Onurlu, M., & Ergut, O. (2023). health econometrics research: A bibliometric analysis from 1991 to 2020. *EKOIST Journal of Econometrics and Statistics*, 38, 243-264. https://doi.org/10.26650/ekoist.2023.38.1242702

Introduction

"Health Econometrics" consists of econometric methodology combined with the theory of health economics. The related literature goes back to the 1960s. However, up to the 2000s, no one had actually used the term "Health Econometrics" until Jones (2000) used it for the first time in the literature in his chapter "Health Econometrics", published in the Handbook of Health Economics. In his study, he states that in fact Newhouse (1987) first referred to the term health econometrics. Although Newhouse (1987) did not mention the term explicitly, he reported that there exists a research field importing methods of econometrics into the theory of health economics (Jones, 2000; Newhouse, 1987). From that time onwards, especially since the 1990s, the number of studies gathering econometrics and health economics has proliferated over time (Rubin and Chang, 2003: 403; Wagstaff and Culyer, 2012: 406; Çağlayan et al., 2020: 63). "Health Econometrics" employs econometric methods in evaluating health policies, examining health expenditures and cost-benefit analyses. It further provides econometric evidence in investigating the dynamics of pharmaceutical and health insurance markets. Econometricians have become more interested in using econometric methodology as health data became more accessible and more available to researchers. Health data are usually collected in either the form of large-scale cross-sectional data (i.e. patient level administrative data) or panel data (i.e. data at the hospital level for a specific period). As a result, econometricians would be able to deal with any issues of nonlinearity, individual-level unobserved heterogeneity and cross-sectional dependencies using the econometric methodology and this has increased the interest of applications in health economics of research for applied econometricians in recent years (Basu and Mullahy, 2014).

There is a significant amount of research that provides information on published materials in health economics. However, the number of studies employing bibliometric analysis to investigate the trends and the current situation in the area of health econometrics is limited. In health econometrics literature, there exist two previous exercises explicitly stating the phrase "health econometrics" and examining the published work in the area. The former study is by Wagstaff and Culyer (2012). They investigated the past 40 years of health economics using a bibliographic analysis. The authors used data from the EconLit database broadened by the Google Scholar's citation data along with specific topics created by the authors. "Health statistics and econometrics" was defined as one of the topics authors created. The study indicates an upward trend in the number of published work in the area of health statistics and econometrics. The latter study (Çağlayan et al., 2020) examined health econometrics literature

specifically in the areas health care markets, government health policy and public health for years from 1991 to June 2020 using the EconLit database. As those areas are not completely distinct from each other, the authors aim to cover all publications published in those specific areas suggesting that the majority of publications appeared to be academic articles and the United States employs more than 30% of authors publishing in the field of health econometrics during the given period.

However, there are other studies that specifically come from the health economics content rather than a particular interest in health econometrics (Rubin and Chang, 2003; Greenberg et al., 2010; Hoque et al., 2011; Pitt et al., 2016). An earlier study by Rubin and Chang (403-414) examined the health economics articles published from 1991 to 2000. The metadata in the study were collected from the EconLit database and the authors investigated trends, the network between authors, topics, and concentration in the literature of health economics from 1991 to 2000 and measured by articles published in journals indexed by EconLit. Greenberg et al. (2010) studied the bibliometric properties of cost-effectiveness analysis (CEA) research in health care. To explore trends in publication and co-authorship, authors collected data from the Tufts Medical Center Registry of original CEAs published from 1976 to 2006 (Greenberg et al., 2010: 320). The study reported that Harvard and Tufts Universities and their affiliated hospitals had the most prolific authors and these authors were not academically related. Pitt, Goodmand and Hanson (2016) collected data from 14 literature databases for articles published in the area of health economics between January 2012 and May 2014. They found that high-income countries had the highest publication rate compared to low or middle income countries.

To fill the gap for an extensive bibliometric analysis in health econometrics, this study aims to investigate publication portfolio including academic journals, working papers, collective volume articles, dissertations and books in health econometrics literature and extend the prior work by Wagstaff and Culyer (2012) and Çağlayan et al. (2020). The data were collected from journals indexed by EconLit from 1991 to 2020. To the best of our knowledge, this is the first study examining health econometrics research employing bibliometric analysis and science mapping techniques. The bibliometric analysis in this study would provide an extensive and detailed picture of publications in the field from the 1990s till 2020. Therefore, the results of the study would conclude whether there had been increasing interest in health econometrics which aims to combine health economics with econometric theory. Moreover, detailed intensity pictures were employed to examine co-authorship status, the co-occurrence status of titles and abstracts and regional analysis of abstracts.

The following is the study's research process: (i) Publications were collected from the EconLit database for the period of 1991 to 2020. The identification of documents was based on subject descriptors (JEL codes) I1 and C1, C2 C3, C5, C8 and C41 accordingly. Articles, dissertations, books, book reviews, collective volume articles and working papers were included in the data. 2324 publications in total were selected based on JEL code criteria in the final sample. ii) The studies and scholars of interest in the field of health econometrics were explored by descriptive bibliometric analysis providing a detailed picture for the distribution of publications by year and written language, the concentration of published work by author and the average number of pages in the published research, (iii) VOSviewer was used to provide analyses of science mapping/social network tool, co-authorships, co-occurrence of titles and abstracts and the networks of collaboration between institutions and authors, iii) As a part of science mapping, the geographical regions analysis of abstracts was further depicted in order to examine whether there had been any regional differences in the number of publications. In this way, which subjects were mostly discussed in which regions and which subjects were focused on in health econometrics during 1991-2020 were revealed.

The rest of the paper is organized as follows: the research questions are presented in Section 2. Section 3 introduces research design and bibliometric methodology. Section 4 presents data collection and data analysis. Section 5 provides results and Section 6 reports limitations of the data and constraints in the access to the EconLit database. Future research directions are stated in Section 7 and finally, Section 8 provides concluding remarks.

2. Research Questions

To examine the development of health econometric research, the studies in the area collected from the EconLit database were investigated by bibliometric analysis along with science mapping, and regional and keyword analyses. The study provides extensive research on the number of publications in health econometrics, document type, language, the average number of pages in the document, keywords, authors, and institutions that contributed to the health econometrics literature. The study further investigates if there are any regional variation in health econometrics research between 1991 to 2020.

The following research questions are investigated in the study:

a) Is there an upward trend in the number of publications? How has health econometric research evolved over time by year, document type and by the average number of pages?

b) What is the frequency of the most common language in health econometric research?

(c) Who are the most prominent authors and publications in the field?

(d) Which institutions have the highest number of publications in the field?

(e) Is there a variation in the distribution of authors by their institutional affiliation and therefore by country?

(f) What are the subject descriptors with the highest proportions used in publications for the final sample?

(f) How is the network structure relationship among authors?

(g) What are the most frequently used words in both titles and abstracts of publications?

(f) Is there a regional variation in the distribution of publications in health econometrics?

(g) Is there any regional variation in the distribution of most frequently used words in publications?

3. Research Design And Bibliometric Methodology

This section provides information concerning the research design and the methodology employed in the study.

3.1. Bibliometric Analysis

The increase in conceptual development and new information about a particular subject brings with it a high volume of data. Bibliometric analysis, which is a quantitative analysis technique, helps to characterize large amounts of data. Bibliometrics makes a significant contribution to scientific progress in a variety of ways: allowing for assessing progress made, identifying the most reliable sources of scientific publication, laying the academic foundation for the evaluation of new developments, identifying major scientific actors and developing bibliometric indices to assess academic output (Martínez et al., 2015). Examining the outcomes of bibliometric analysis provides information about previous studies as well as an opportunity to learn about the researched area by shedding light on potential future research areas. Bibliometric analysis has been applied in most studies to determine the main issues in these data and trends over time, and to examine the way the subject is handled between disciplines.

3.2. Science Mapping

Science mapping aims to depict the dynamic and structural aspects of scientific research by defining a research field and identifying, quantifying, and visualizing its thematic subfields (Heradio et al., 2016). Scientific mapping combines quantitative analysis, classification, and visualization to identify structures and interrelationships between bibliographic objects (Andersen and Swami, 2021). With science mapping, relationships and actors that cannot be easily revealed can be discovered. Science mapping analysis focuses on monitoring a scientific field and delimiting research areas to determine its structure and its evolution (Cobo et al., 2013).

Analysis results can be visualized by using various software (SciMAT, VOSviewer, CiteSpace), so that the knowledge structure can be interpreted more easily.

By mapping and analyzing relationships among people, teams, departments or even entire organizations with science mapping analysis, it is possible to examine the cooperation status between the authors and their existing connections, determine the research teams and identify the people who have an important position.

3.3 Co-word Analysis

With co-word analysis, which is used to investigate the structure of many scientific fields, the strength of association between the terms in the documents is determined. One of the primary advantages of co-word analysis introduced by Callon et al. (1983) is that its conceptual structure can be revealed without dealing with the original text and semantic maps can be produced that facilitate the understanding of a discipline.

Co-word analysis can be applied from titles to keywords, or from abstracts to words in full texts in publications. If two or more words co-occur in the same document, it indicates a linkage between them. The greater frequency that two keywords both appear in the same article, the deeper the linkage. By analyzing the co-occurrence relationships between publications, connections between topics, main areas, hot and cold spots can be defined, and the trend of the topics can be determined (Faraji et al., 2022). This analysis, based on the co-occurrence network of words, has also been applied to extract important topics in a specific field (Bai and Li, 2022).

4. Data Collection And Data Analysis

The study aims to comprehensively analyze the current status of publications in health econometrics by using a bibliometric analysis along with science mapping. To achieve the aforementioned purposes, data were collected from the EconLit database. The reason to choose EconLit as the source of bibliographic metadata is that it allows researchers to sort on the JEL classification system. EconLit research results provide six types of records: articles, dissertations, books, book reviews, collective volume

articles and working papers. To classify the publications, subject descriptors are used in EconLit. The data set included publications with subcategories of I1 and C1, C2 C3, C5, C8 and C41 subject descriptors in the EconLit database between January 1991 and December 2020. After dublicates were removed, the final sample consisted of 2324 publications. Figure 1 presents a flowchart for the identification of publications on the field of health econometrics.

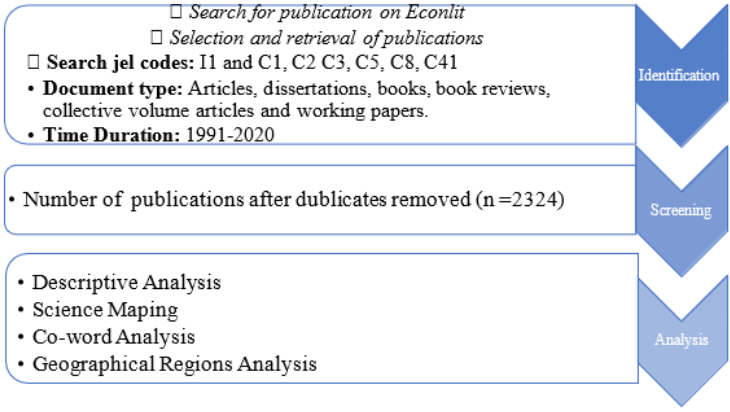


Figure 1. The flowchart of the identification of documents

The data analysis was conducted in two parts including a "bibliometric analysis" and "science mapping" (Figure 2). In the first part, basic descriptive bibliometric statistics such as the number of publications, authors who contributed to the literature the most, top publishing authors, and author affiliations were presented. The second part of the analysis consists of a co-authorship network, a common word analysis of both titles and abstracts. Collaboration networks of authors were examined by co-author analysis. Co-word analysis was applied to find major themes and determine how themes change during the study period. The third part consists of a regional analysis which investigates any regional variations in the distribution of publications and in the most frequently used words in abstracts. VOSviewer was utilized in the analysis process.

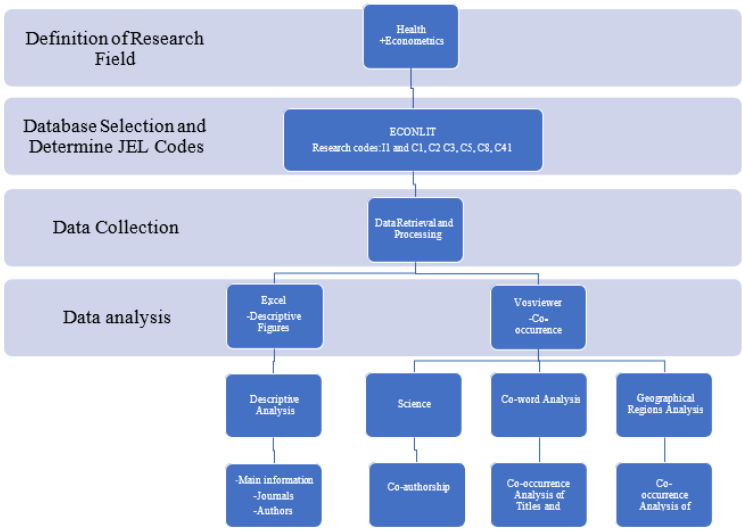


Figure 2. The flowchart of the data analysis process

5. Results of Analyses

The results of the analyses cover three main results. The overview of publications in health econometrics research is initially presented. Our findings are then reported.

5.1. Synopsis of Publications

Via a completed extensive search in the EconLit database, 2324 publications were identified. The distribution of publications by year for the period of 1991 to 2020 is presented in Figure 3.

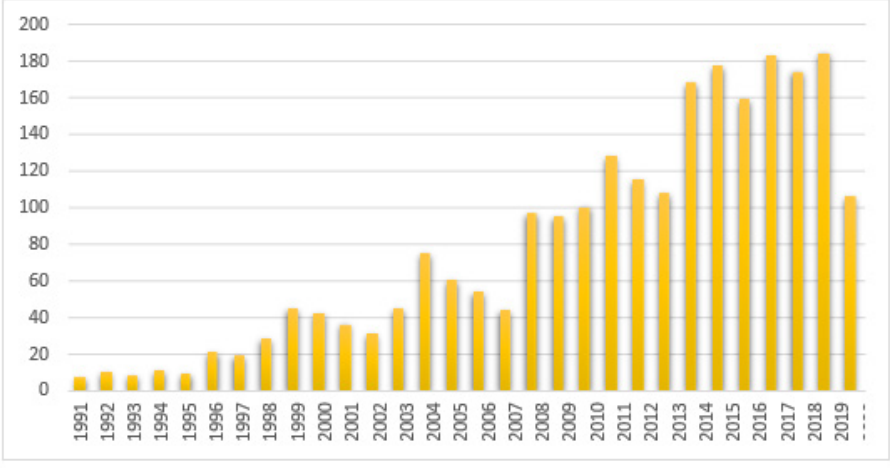


Figure 3. Number of Publications by Year

Figure 3 shows the trends in the number of publications in health econometrics. The figure suggests that the number of publications started to increase in 1999 and this increase continued until 2007. There was a reduction in the number of publications from 2007 to 2008. Starting from 2010 number of publications started to increase. In 2014, a sharp increase in the quantity of publications was observed. The growing number of publications in the area could be considered as a sign that researchers' interest in health econometrics had increased for the analysis period.

Considering the changes in the publications over the years, the analysis period was split into three periods (1991-2000, 2001-2010, and 2010-2020) to be able to observe the changes for every 10-year period more precisely. The distribution of publications according to the document type in a ten-year period is presented in Figure 4.

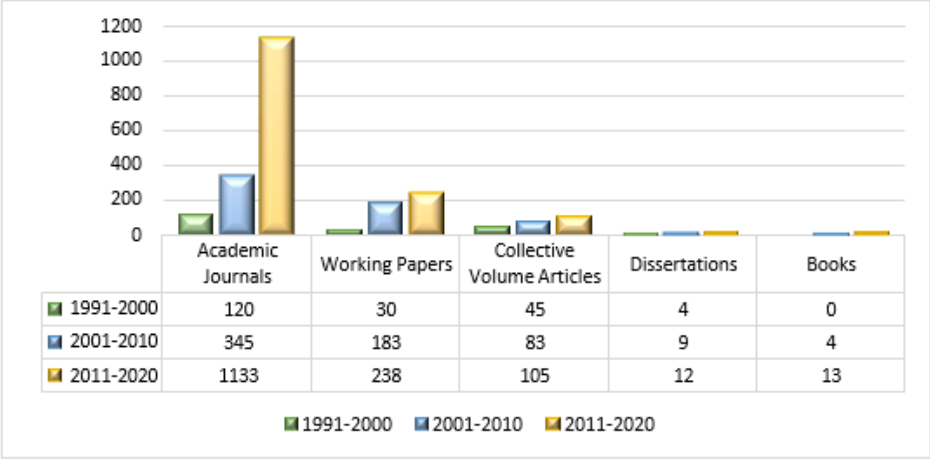


Figure 4. Distribution of Publications by Document Type

Figure 4 shows that there was a steady increase in the number of documents. There were 199 publications in total for the 1991-2000 period, 624 publications for the 2001-2010 period and 1501 publications for the last period. One could clearly see that while the number of articles followed an increase in all three periods and working papers in the second period, there were no major changes in the number of other types of documents.

Table 1
Overview of Publications (N = 2324)

Panel A: Distribution of Publications According to Written Language				
Language	*Number of Articles*			
English	2292			
Chinese	13			
Spanish	10			
Turkish	5			
French	4			
Panel B: Distribution of Publications by Document Type				
Document Type	*Number of Documents*			
Academic Journals	1598			
Working Papers	451			
Collective Volume Articles	233			
Dissertations	25			
Books	17			
Panel C: Average Number of Pages for Articles				
Periods	*Journal Count*	*Number of Articles*	*Total Number of Pages*	*Average Length*
1991-2000	32	120	1548	12,9
2001-2010	86	345	5137	14,88
2011-2020	198	1133	19395	17.11
1991-2020	316	1598	26080	16.32

Table 1 provides a detailed picture of the overview of publications in health econometrics from 1991 to 2020. Panel A of Table 1 reports findings regarding the language of publications. 98% of publications were written in English and this was followed by Chinese and Spanish and

Turkish, respectively. Among those publications, academic journals had the highest frequency whereas books had the lowest frequency for the analysis period (Panel B, Table 1).

Journal counts, the average page length of articles for both the 1991-2020 time period and three periods of ten years are reported in Panel C (Table 1). The number of journals, which was 32 in the first period, increased over time and reached its highest level in the last period. Furthermore, the average page length increased from 12.9 to 17.11, accordingly.

5.2. Synopsis of Authors

This section provides information about the detailed overview of the authors. The distribution of the publications according to the number of authors is shown in Table 2.

Table 2
Distribution of publications by number of authors

1 Author	*2 Authors*	*3 Authors*	*4 Authors*	*5 Authors*	*6 or More Authors*
%23	%33	%24	%12	%5	%3

The table suggests that 33% of them were studies with 2 authors, followed by studies with 3 authors and one author. There was only one study with 7 and 8 authors each.

Table 3
The Authors Who Have the Highest Contribution to the Health Econometrics Literature

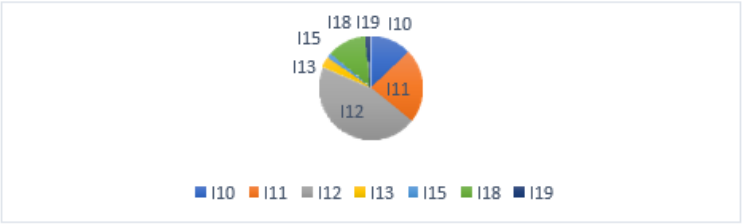
Authors	***Number of Articles***
Jones, Andrew M.	20
van Soest, Arthur	16
Basu, Anirban	16
Rice, Nigel	15
Kapteyn, Arie	14
Mullahy, John	13
O'Hare, Colin	13
Li, Johnny Siu-Hang	12
Ryan, Mandy	12
Trivedi, Pravin K.	11
Manski, Charles F.	11
Rosenbaum, Paul R.	10
Gaiha, Raghav	10
Small, Dylan S.	10
Rosenman, Robert	10
Lindeboom, Maarten	10
Windmeijer, Frank	10
van Ours, Jan C.	9
Kreider, Brent	8
Imai, Katsushi S.	8

Table 3 presents the top 20 most active authors with at least 8 publications, and suggests that the author, Andrew M. Jones, who contributed most to the health econometrics literature, had 20 publications. 82% of authors had only one publication whereas 11% had two publications. Only 4% of authors had three publications for the analysis period.

5.3. Synopsis of Subject Descriptors

The data set consists of 2324 publications with subcategories of I1 and C1, C2 C3, C5, C8 and C41 subject descriptors in the EconLit database between January 1991 and December 2020. Figure 4 provides information regarding the distribution of publications by subject descriptors in the EconLit database.[1] In the identification of publications, Panel A of Figure 5 suggests that I12 (Health Behavior) had the largest share whereas I15 (Health and Economic Development) appears to have the smallest share among all other I1 codes. Moreover, Panel B (Figure 5) shows the distribution of publications with the first 14 C codes. The largest share in C-coded subject identifiers was C51 (Model Construction and Estimation) and the second-largest share was C83 (Survey Methods; Sampling Methods). These results suggest that C51 and C83 were the most commonly preferred subject descriptors among 2324 publications in the field of health econometrics.

Panel A: Distribution of Publications by Codes I1*



Panel B: Distribution of Publications According to C Codes

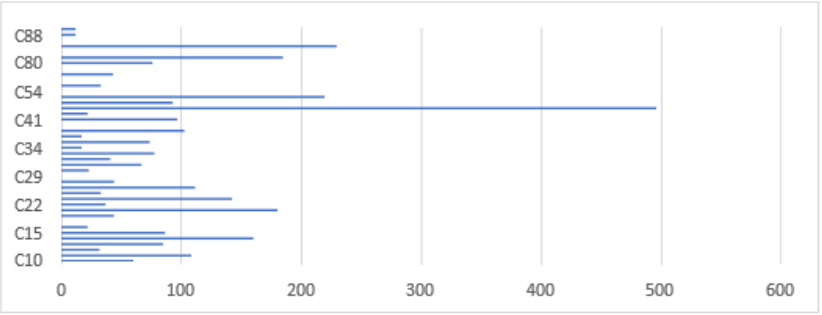


Figure 5. Distribution of Publications by Subject Descriptors (N = 2324)

[1] For more information about JEL Classification System / EconLit Subject Descriptors, please visit the following website: https://www.aeaweb.org/econlit/jelCodes.php.

5.4. Synopsis of Institutions and Countries

The frequency of authors by their institutions along with the country where the institution is established was also examined. Table 4 shows the distribution of the authors by institutional affiliation. The table reports the top 20 universities with the highest frequency of authors published in the area of health econometrics. Among the top 20 universities, Harvard University, based in the US, had the highest frequency of authors (63 authors), while Bristol University, based in the UK, had 23 the lowest frequency of authors published in the field of health econometrics during 1991-2020. One could clearly suggest that institutions or universities located in the United States had the highest share of authors, followed by the UK, Australia and the Netherlands.

Table 4
Distribution of Authors by Their Institutional Affiliation

Name of the Institution	*Number of Authors*	*Country*
Harvard University	63	USA
University of Pennsylvania	61	USA
University of York	57	UK
Johns Hopkins University	48	USA
University of Sheffield	41	UK
University of California Berkeley	36	USA
University of Chicago	35	USA
Monash University	31	Australia
London School of Hygiene and Tropical Medicine	31	UK
University of Southampton	31	UK
University of North Caroline	30	USA
University of Aberdeen	29	Australia
Tilburg University//Center for Economic Research	29	Netherlands
Duke University	29	USA
University of Washington	27	USA
City University London	26	UK
Stanford University	26	USA
University of Wisconsin	24	USA
University of Minnesota	24	USA
University of Bristol	23	UK

5.5. Synopsis of Co-authorship Status

The co-authorship status of documents was analyzed to see the academic cooperation among the authors and networks between their participants by using science mapping analysis. The analysis results show that 783 authors had published at least 2 papers each. The network created to see the co-authorship status is given in Figure 6. In the network map, the nodes, each representing an author, grow with the increase in the number of publication of that author. Once the collaboration between the authors increases, the lines (links) become thicker. As can be detected from the figure, lines demonstrating the cooperation among authors are not thick. Micro-level collaborations existed between authors suggesting that certain author groups worked together in a loose network structure.

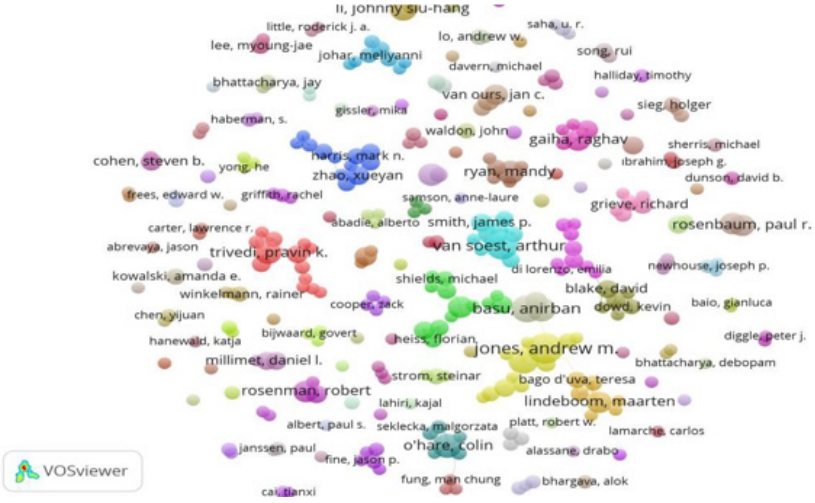


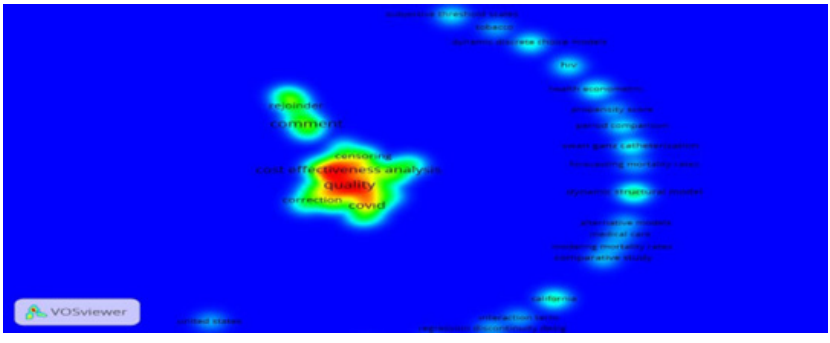
Figure 6. Network Structure among Authors

Andrew Jones, Anirban Basu, and Arthur van Soet had the highest contribution to the literature on health econometrics. Arthur von Soest and Nigel Rice were the authors with the highest total link strength, which indicates the total strength of the co-authorship links of a given researcher with other researchers. The author Arthur von Soest frequently worked with Arie Kapteyn, while Nigel Rice's best collaborator was Andrew M. Jones.

5.6. Synopsis of Co-Word Analysis of Abstracts and Titles

A co-word analysis was conducted on the titles and abstracts of the papers to investigate the hot research areas, explore the intensity of the words employed in the titles, and evaluate the changes over time. Panel A in Figure 7 shows that the mainly focused words were life quality, covid, China, Germany, France, big data, obesity, cannabis use, cost-effectiveness analysis, casual effect, discrete choice experiment, dynamic analysis, mortality modelling, additive nonparametric regression, meta-analysis, selection bias, comment, survival data, moral hazard, and case study.

Panel A: The Most Frequently Used Words in the Titles (1991-2020)



Panel B: Words Used in the Titles by Years

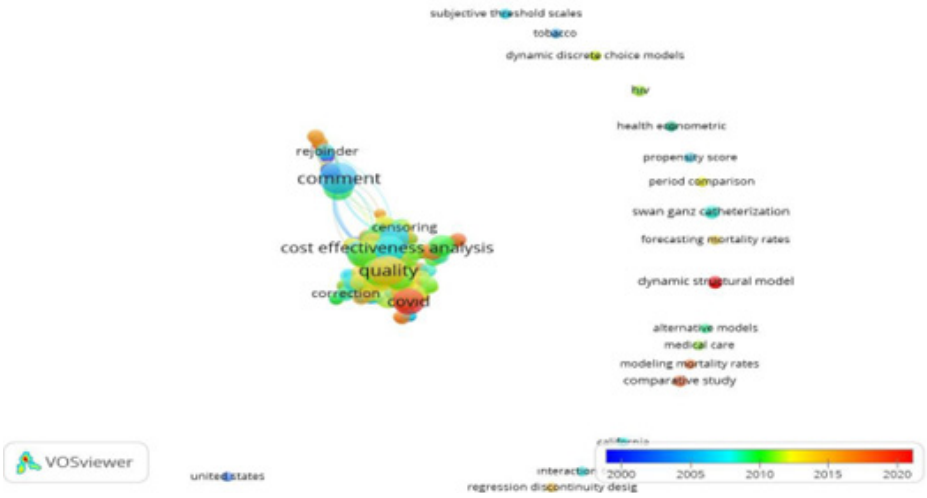


Figure 7. Co-Word Analysis of Abstracts and Titles

Panel B in Figure 7 provides information on the change of titles by year. The figure indicates that these words were frequently used especially in the titles of the studies conducted after 2014. The red color in the figure shows the topics that had been concentrated on in recent years. Work disability, health status, causal effect, fertility, and count data models were in the headings in the early 2000s. These were replaced by "uncertainty, cannabis use, adverse selection, education, mortality projection, survival model, moral hazard, health care reform, England, United States, Europe" in 2010. Towards 2015 "mental health, Monte Carlo simulation, natural experiments, inequality, health shocks, poverty, household, econometric analysis, challenge, longevity risk, infant mortality, survival analysis" words were frequently encountered. Between 2015 and 2020, the words, "nutrition, pitfalls, financial incentives, quantiles, big data, economic evaluations, fixed effect, nonparametric conditional approach, observational data, Covid, outbreak, spread, China, treatment comparison, dynamic structural model, generalized linear models, meta-analysis, systematic review, comparative study, modelling mortality rates," seemed to be associated with an increasing density.

One could suggest that the subject titles did not vary much in the 2000s. In the period between 2005 and 2010, the information for which country the study was applied was included in the titles of the studies, and various words were also emphasized in the titles regarding subjects such as cannabis use, health care reform and mortality. A wide variety of terms were reported to be included in the titles of the publications in the 2010-2015 period. After 2015, as well as the type of data (census data, observational data) used in the titles and information about which analysis was employed. In studies conducted in 2020, the words Covid, China, influenza, virus diffusion, determining factors, and treatment comparison are included in the headings

and were an indication that the studies conducted after the epidemic have shifted in this direction.

Co-word analysis in the abstracts provides information on the counts of occurrences of words, thereby a higher number of the same words used by different authors indicates certain significance hotspot points in the subject area. The mapping based on common words used in the abstracts helps researchers find main topics, future trends and new fields of study by visualizing different sub-study areas and the relationship between them.

5.7. Synopsis of Keywords

A keyword analysis was undertaken in order to investigate which research topics were employed in the publications between 1991 and 2020.

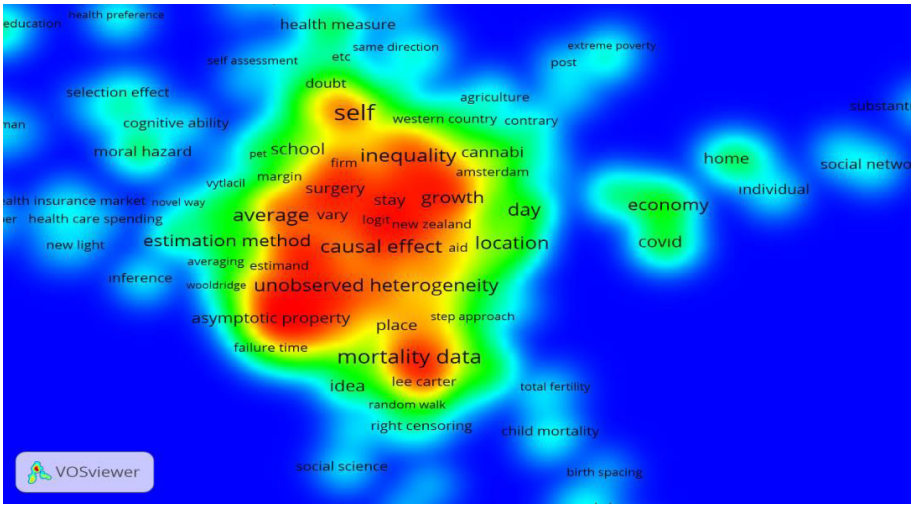


Figure 8. Keyword Analysis of the Abstracts

Based on the keyword analysis, frequently mentioned and especially emphasized words were reported as mortality data, unobserved heterogenetiy, causal effect, treatment effect, Lee Carter model, Covid, virus, fertility, cost-effective analysis, surgery, birth weight, Moral Hazard, poverty, Oregon health insurance, retirement, cohort effect, health care utilization, private sector, copula model, longevity risk, Netherlands, inequality, economic growth, health data, BMI, cannabis use, concentration index, life satisfaction, health care expenditure, California, UK, Canada, Europe, Amsterdam. Analysing the words in abstracts also provided information on the type of data used and the countries studied. In terms of data source, survey data were the most widely used data source (household survey, British household panel survey).

The methods/elements used in health econometric research included probit, dynamic panel data models, conditional quantile, MCMC (Markov Chain Monte Carlo), discrete choice experiment, structural equation model, latent variable, Bayesian analysis, stochastic mortality model, dynamic model, functional data analysis, adverse selection effect, principal component analysis, nonparametric estimator.

5.8. Synopsis of Regional Analysis

One of the main contributions of this study is to provide evidence for regional variation for co- word analysis along with the distribution of the minimum number of documents of an author by geographical regions. Figure 9 presents the most frequently used words in abstracts divided by different geographical regions.

Figure 9 demonstrates the most frequently used words in the studies by regions as well as by countries discussed in the studies conducted with the particular region. If the studies based in Africa are considered, one can clearly see that HIV prevalence and food security issues are frequently discussed. Mortality rate and life expectancy are prominent in studies on Asia, and cancer and quality of life in Europe. The table and the figure further show which econometric techniques are used in the studies based in North America and Europe.

Table 5 shows the minimum number of documents of an author by geographic region.

Table 5
Minimum Number of Documents per Author

	Minimum Number of Documents from an Author					
	2	***3***	***4***	***5***	***6+***	***TOTAL***
Africa	5					158
Asia	16	7				335
Europe	101	22	10	5	10	968
Latin America and the Caribbean						75
North America	113	35	11	5		1156
Oceania	15	4				97

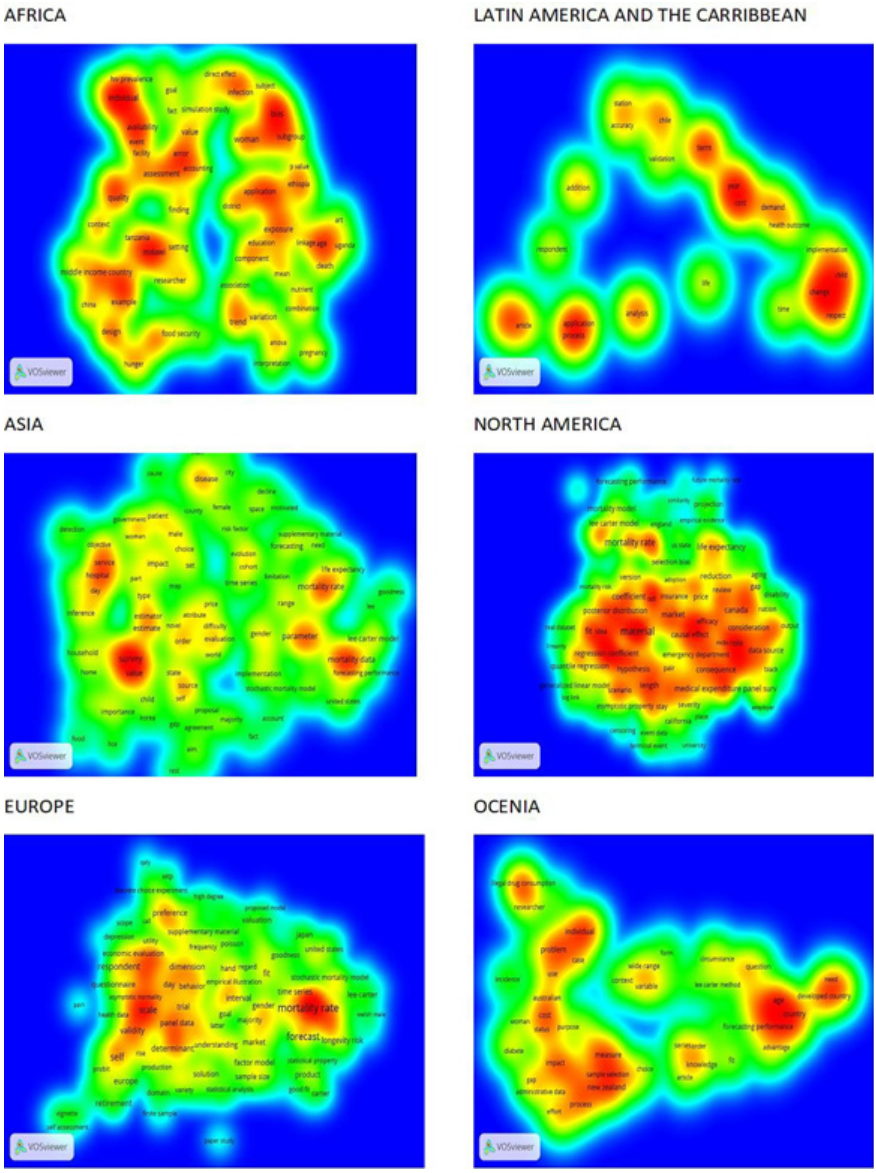


Figure 9. Density Map of the Most Frequently Used Words in Abstracts by Geographical Regions

One striking feature of Table 5 is that North America had the highest number of authors (1156 authors in total) and Europe had the highest number of authors having a minimum of 6 documents. This was followed by North America, 5 authors had a minimum of 5 documents. In addition, Africa had only 5 authors with a minimum of 2 documents.

Figure 10 and Figure 11 show the distribution of publications by I and C codes in different geographic regions. Figure 9 reports that the C21 subject descriptor was

the least common C code among all other regions whereas C51 and C53 were the most common descriptors across all regions. Regarding descriptors starting with the I code, I11 and I12 were the most frequently used codes among 6 different regions. However, I15 was the least common I code and only a small proportion of I15 was employed by Asian countries.

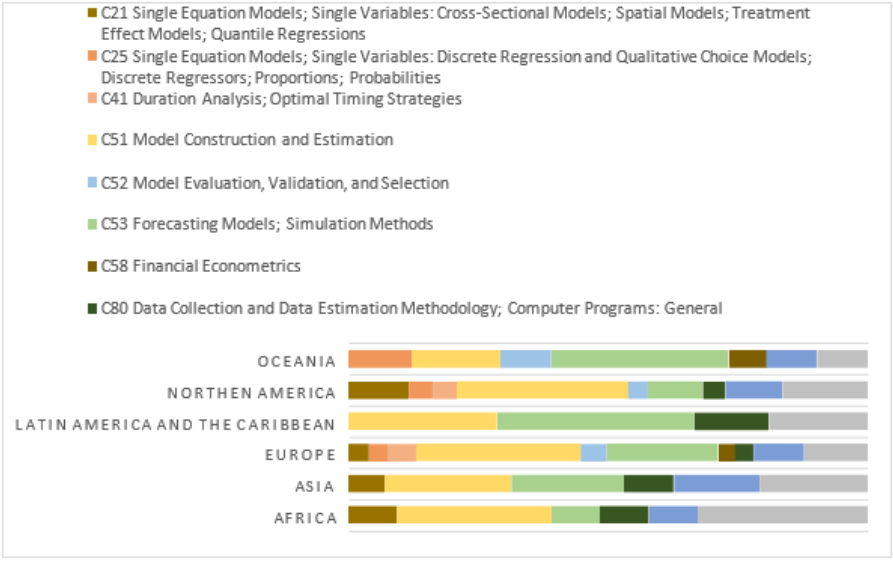


Figure 10. Distribution of C Subject Descriptors by Geographical Region

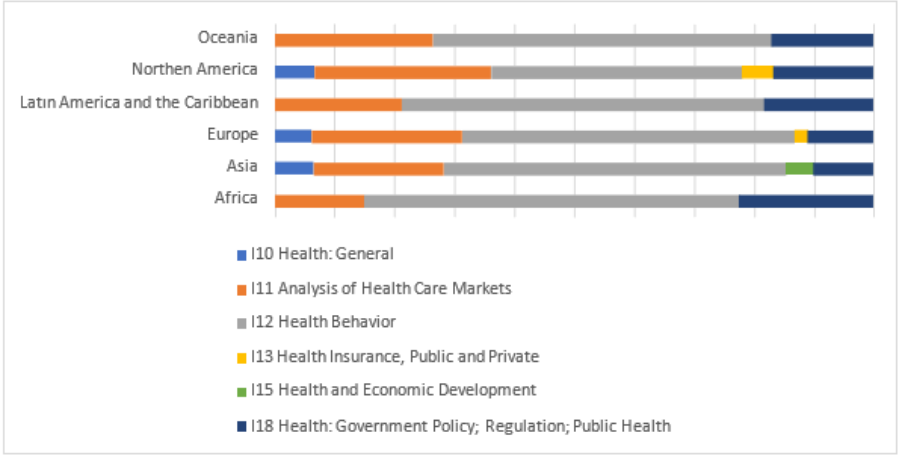


Figure 11. Distribution of I Subject Descriptors by Geographical Region

6. Limitations of The Study

In the study, before bibliometric analysis, subject descriptors on both health economics and econometrics were filtered out from the EconLit database. One of

the main drawbacks of deriving data on publications from the EconLit database is the limitation to search not for a keyword, but only JEL codes. Another weakness of the Econlit database is that the data format is limited when extracting data from Econlit, and accordingly, the number of suitable programs that could analyze this data is limited.

Furthermore, the EconLit database does not provide any information concerning the citation of articles and authors. Therefore, our study was unable to cover any citation analysis for publications in health econometrics. As a result, we lacked the information for which authors or publications had the highest number of citations in health econometrics. Moreover, it could be the case that our study might be missing publications with no JEL subject descriptors or codes when deriving data from the EconLit database for both health economics and econometrics publications. However, we were able to capture all publications with relevant JEL codes available in the EconLit database.

7. Future Research Directions

For future research directions, alternative databases such as Google Scholar or Web of Science could be employed in order to do a bibliometric analysis of publications in the area of health econometrics. With the aim to minimize the limitations of the EconLit database, to capture all publications (including those with no JEL subject descriptors), a search on publications could be performed using keywords for health econometrics as these aforementioned databases allow researchers to sort by keywords not only by JEL subject descriptors.

Our study includes health-related JEL subject descriptors starting with only I1 (Health). Our findings suggest that there has been an increase in interest in health econometrics since the 1990s. However, there are other health-related JEL subject descriptors such as H51 (Government Expenditures and Health) or H75 (State and Local Government: Health, Education, Welfare, Public Pensions) which were excluded from this study. Our future research directions could be to examine the current situation in those categories of health and econometric-related JEL codes together and provide a bibliometric analysis.

8. Concluding Remarks

This study has investigated almost 60 years of published research in health econometrics literature starting from 1991 until the end of 2020. We aimed to identify different properties of published materials and therefore to draw a detailed picture of recent publications. The data were collected from the EconLit database and consisted of publications with subcategories of I1 and C1, C2 C3, C5, C8 and C41 subject descriptors. First of all, all publications were examined by year, type of document

and written language. An upward trend was observed in the number of published documents since 1991. This clearly shows that health econometrics has increasingly been in the interest of researchers although as a term it was first used by Jones (2000) in the 2000s. Authors were further examined in terms of their contribution to the recent literature. Andrew M. Jones (University of York, UK) was the author with the highest contribution to the Health Econometrics Literature. However, in general the majority of authors who have contributed to the literature are based in the USA. In a bibliometric approach, co-authorship status is an important part of the analysis. Surprisingly, micro-level collaborations were found between authors suggesting that only certain author groups work together for a loose network structure. In addition, survey data were the most common type of data used in publications.

From the perspective of common word analysis by geographical regions, it was interesting to see that most of the common words in documents were region specific. For instance, African studies were more likely to use words such as HIV prevalence, infection, woman, Middle-income country, health status, food security, simulation study, policy maker, and education whereas this was slightly different for other regions such as Europe or North America. The regional examination of the most common subject descriptors used in the documents also showed a similar pattern, especially for subject descriptors referring to the econometric methods. For example, Northern American studies were more likely to employ C51. However, for African studies, C52 was the most common subject descriptor. We assume that this could be a result of the data type available in each region and the health-related issue examined in each published document.

Certainly and not least, since December 2019, Covid-19 has become a major pandemic countries are dealing with worldwide. Therefore, one can clearly see how important measuring and examining health outcomes and health-related policies/ implications are. This could only be more effective by integrating health economics into econometric methods. As a result, we hope to see health econometrics becomes more accepted and more commonly known by more researchers in the future and hope that this specific area of research would continue developing.

Peer-review: Externally peer-reviewed.
Conflict of Interest: The author has no conflict of interest to declare.
Grant Support: The author declared that this study has received no financial support.

References

Andersen, N., & Swami, V. (2021). Science mapping research on body image: a bibliometric review of publications in body image, 2004–2020. *Body Image*, *38*, 106-119. https://doi.org/10.1016/j.bodyim.2021.03.015.

Bai, Y., & Li, H. (2022). Mapping the evolution of e-commerce research through co-word analysis: 2001–2020. *Electronic Commerce Research and Applications*, 101190. https://doi.org/10.1016/j.elerap.2022.101190.

Basu, A., & Mullahy, J. (2014). Health econometrics: An overview. In A. J. Culyer (Ed.), *Encyclopedia of health economics* (pp.355-356). Elsevier.

Callon, M., Courtial, J.-P., Turner, W. A., & Bauin, S. (1983). From translations to problematic networks: An introduction to co-word analysis. Social Science Information, 22(2), 191–235. https://doi.org/10.1177/053901883022002003.

Cobo, M.J., Chiclana, F., Collop, A., de Ona, J., & Herrera-Viedma, E. (2013). A bibliometric analysis of the intelligent transportation systems research based on science mapping. *IEEE transactions on intelligent transportation systems*, *15*(2), 2013, 901-908. 10.1109/TITS.2013.2284756.

Çağlayan Akay, E., Ergüt, Ö., & Ertok Onurlu, M. (2020). Health econometrics publications in the area of health care policy and health care markets: a bibliometric analysis. In Özge Uysal Şahin & Merve Ertok Onurlu (Eds.) *Pandemi ve Ekonomi Kıskacında Sağlık* (59-78). Çanakkale: Rating Academy Press.

Faraji, O., Ezadpour, M., Dastjerdi, A. R., & Dolatzarei, E. (2022). Conceptual structure of balanced scorecard research: A co-word analysis. *Evaluation and Program Planning*, *94*, 102128. https://doi.org/10.1016/j.evalprogplan.2022.102128.

Greenberg, D., Rosen, A.B., Wacht, O., Palmer, J., & Neumann, P.J. (2010). A bibliometric review of cost-effectiveness analyses in the economic and medical literature: 1976–2006. *Medical Decision Making*, 30, 320–327. https://doi.org/10.1177/0272989X09360066.

Heradio, R., Perez-Morago, H., Fernandez-Amoros, D., Cabrerizo, F. J., & Herrera-Viedma, E. (2016). A bibliometric analysis of 20 years of research on software product lines. *InformationandSoftwareTechnology*,1-15. https://doi.org/10.1016/j.infsof.2015.11.004.

Hoque, M. E., Khan, J. A., Hossain, S. S., Gazi, R., Rashid, H. A., Koehlmoos, T. P., & Walker, D. G. (2011). A systematic review of economic evaluations of health and health- related interventions in Bangladesh. *Cost Effectiveness and Resource Allocation*, *9*(1), 1-8. https://doi.org/10.1186/1478-7547-9-12.

Jones, A.M. (2000). Chapter 6 - Health Econometrics. In Anthony J. Culyer, Joseph P. Newhouse(Eds),*HandbookofHealthEconomics*(pp.265-344).Elsevier. https://doi.org/10.1016/S1574-0064(00)80165 (https://www.sciencedirect.com/science/article/pii/S1574006400801651).

Martínez, M. A., Cobo, M. J., Herrera, M., & Herrera-Viedma, E. (2015). Analyzing the scientific evolution of social work using science mapping. *ResearchonSocialWorkPractice*, *25*(2),257-277. https://doi.org/10.1177/1049731514522101.

Newhouse, J. P. (May, 1987). *Health Economics and Econometrics.* Papers and Proceedings of the Ninety-Ninth Annual Meeting of the American Economic Association.

Pitt, C., Goodman, C., & Hanson, K. (2016). Economic evaluation in global perspective: a bibliometric analysis of the recent literature. *Health Economics.* 25, 9-28. https://doi.org/10.1002/hec.3305.

Rubin, R. M. and Chang, C. F. (2003). A bibliometric analysis of health economics articles in theeconomics literature: 1991-2000. *HealthEconomics*,12(5),403-414. https://doi.org/10.1002/hec.802.

Wagstaff, A. & Culyer, A. J. (2012) Four decades of health economics through a bibliometric lens. *JournalofHealthEconomics*,31(2),406-439. https://doi.org/10.1016/j.jhealeco.2012.03.002.