

21世纪以来我国经济学研究主题的发展与演变*

——基于经济学期刊摘要的主题挖掘

官国宇 胡汉云

(东北师范大学经济与管理学院, 吉林 长春 130117)

[摘要] 近年来,中国经济快速稳定发展,为经济学研究提供了丰沃的土壤。然而,面对海量文献,读者难以把握经济学的研究热点和发展规律。科学文献的文本挖掘可以为学者们更好地把握学科前沿和研究现状提供参考。经济学类CSSCI来源期刊代表了国内经济学研究的最高水平,其研究主题围绕当代社会经济热点。本文对2000—2018年经济学科CSSCI来源期刊摘要文本数据进行主题挖掘,旨在发现21世纪以来中国经济学研究的主要领域及其发展趋势。首先,利用LDA主题模型计算主题一致性得分,确定最佳主题数;第二,通过计算JS散度得到主题距离矩阵,进而度量各主题间的相关关系;第三,由“文档-主题”的概率分布得出各期刊研究主题的偏好与主题的年度分布规律。研究发现:中国经济学研究大致涵盖20个主题;各研究主题之间呈现出一定的亲疏关系;各期刊对研究主题有所偏好;中国经济学研究热点具有时代特征,反映了中国经济发展不同历史阶段。根据研究结果,经济学科CSSCI来源期刊摘要数据的主题挖掘见证了中国21世纪以来经济学研究的发展历程,可以为青年学者的经济学研究选题和期刊投稿提供参考建议。学者们可以根据“年代-主题”分布审视研究选题的时代意义;根据“期刊-主题”分布选择恰当的投稿期刊,提高研究成果接受发表的概率。

[关键词] 经济学研究 研究热点 文献计量 主题挖掘

[中图分类号] G251 **[文献标识码]** A **[文章编号]** 2096-983X(2021)03-0029-10

一、引言

改革开放四十多年来,中国特色社会主义事业取得了举世瞩目的成就,这为中国特色社会主义经济理论的发展和繁荣提供了丰沃的土壤。2019年3月4日,习近平总书记在第十三届中

国政治协商会议文化艺术界、社会科学委员会的演讲中强调了有意义的学术研究都应该反映现实、关照现实。学术研究应该植根于中国土地、立足于中国现实,展示现代中国的发展和进步,诠释中国精神、中国价值和中國力量。而提高实用性和创造力是当前学术研究面临的主要

收稿日期:2020-10-14

*基金项目:自然科学中央高校基本科研业务费资助项目“汉语文本稀疏性挖掘技术研究”(2412015KJ028)

作者简介:官国宇,副教授,博士,主要从事统计学和计量经济学研究;胡汉云,硕士研究生,主要从事金融计量学研究。

问题。^[1]2018年中国国内生产总值(GDP)增长6.6%,2019年中国GDP增长6.1%,增速比上年下降0.5个百分点,而2020年新冠疫情的全球蔓延无疑对中国经济产生巨大影响,面对国内经济下行压力,经济学研究任重道远。中国的经济学研究工作者正在为之不懈努力,结合中国国情,汲取并借鉴世界经典和前沿的经济学思想和方法,取得了丰硕的研究成果。

CSSCI(中国社会科学引文索引)是南京大学中国社会科学评估中心开发的引文数据库,用于检索中文人文社会科学领域高质量且高影响力的学术文献,很多高等院校和科研院所将其作为衡量学术水平的重要参考标准。^[2]经济学类CSSCI来源期刊代表了国内经济学研究的最高水平,围绕当代热点和前沿经济问题展开,其研究主题与当时社会经济情形紧密相关,是经济发展阶段性特征的重要体现。本文对2000—2018年经济学科的CSSCI来源期刊摘要汉语文本数据进行主题挖掘,借助前沿的文本挖掘与自然语言处理方法,旨在探寻21世纪以来我国经济学研究主题的发展与演变路径,以为为经济学研究学者们选题和期刊投稿提供数据参考支撑和建议。

二、文献综述

主题建模是挖掘文本数据潜在信息的重要工具,而潜在狄利克雷(LDA)模型是主题建模的核心技术之一。LDA模型是Blei等^[3]在2003年提出的用来描述“文档、主题、关键词”三者关系的概率生成模型。该模型假设每个文档是由多个主题组合而成,而每个主题又是由多个词组成,基于“主题-关键词”分布识别大规模文档或者语料库中隐藏的主题信息。词袋(Bag of Words)模型是建立LDA模型的基础,它将每个文档表示为一个词频向量,从而将文档信息转化为易于建模的数值型矩阵。LDA模型可生成两个概率分布矩阵,一个是“文档-主题”概率分布矩阵,另一个是“主题-关键词”的概率分

布矩阵。其中主题是对文档内容的提炼,该模型可从大规模语料库中提取主题。

LDA模型作为重要的文本主题挖掘技术,在国内早期主要应用于计算机领域。陈江峰、于建军(2008)结合LDA生成概率模型的特点与Web服务的结构化特性,有效地提高了Web服务发现的效率和精确率。^[4]随着Web2.0的发展,在线评论资源日益普及,各大互联网平台积累了数量庞大的UGC用户,一些具有主观情感的网络评论、网络文章蕴含着巨大的潜在商业价值,受到了工业界和商业界的关注。此时LDA主题模型在文本数据的意见挖掘和情感分析中凸显优势。吕韶华(2010)提出了一种依据评论内容对餐馆进行排序的算法,利用LDA模型对评论文本进行服务、环境、价格、口味等方面进行主题抽取和计算得分,最终使此模型可以根据评论对餐馆进行排序。^[5]LDA主题模型在临床医学研究领域同样有着重要的作用。张小平(2011)首次将LDA主题模型应用于中医临床诊疗规律的研究中,为中医临床研究提供一种新颖的理论方法。^[6]杨星等(2012)证实了利用LDA模型研究科学文献的研究热点及趋势的有效性,开启了科学文献主题挖掘热潮。^[7]随后,越来越多的学者运用LDA主题模型来研究文本分类技术。胡吉明、陈果(2014)从LDA模型的主题挖掘原理入手,针对当前网络环境下的文本内容特点,构建适合于动态内容文本主题挖掘的LDA模型,并通过改进的Gibbs抽样估计提高主题挖掘的准确性,进而从主题相似度来研究内容主题随时间的演化问题。^[8]王鹏等(2015)提出了一种基于LDA模型通过Gibbs算法估计文本的主题概率分布,利用Jensen-Shannon(JS)距离作为文本的相似性度量,采用层次聚类进行聚类,说明了LDA-Gibbs模型的文本聚类方法是合理且有效的。^[9]互联网的普及与音乐资源的电子化使人们可以方便地获得音乐资源。但是在音乐如此丰富的时代如何找到自己喜欢的音乐越来越难。李博等(2016)结合LDA主题挖掘模型和音乐试听数据,提出了音乐推荐算法,

可以更加高效地为用户推荐感兴趣的音乐。^[10]在如今信息爆炸的时代,各种真真假假的信息层出不穷,网络谣言的肆虐对人们生活和社会稳定都产生了巨大的负面影响。曾子明、王婧(2019)基于LDA主题模型并采用随机森林算法有效提高了谣言识别的准确率。^[11]

三、数据采集和预处理

(一) 数据采集

本文从中国知网(<https://www.cnki.net/>)上收集了2000—2018年的经济学科CSSCI来源期刊数据。期刊范围为《CSSCI来源期刊(2019—2020)目录》中经济学科的70个CSSCI来源期刊,该范围共包含228415篇文章。全部字段包括发表年份、标题、作者、期刊、摘要和关键词,不包括全文。本文的数据分析仅使用发表年份、期刊、摘要以及关键词四个字段。

(二) 数据预处理

首先是对经济学期刊数据进行去噪。原始数据中含有一些非学术性文章,如征稿启事、会议纪要、新书推介、新年寄语等等。由于这些非学术型文章的存在会影响主题模型的训练,因此对其进行人工去噪。其次,删除了字段不完整的文献数据,如没有关键词、摘要、作者和单位信息的数据,最终保留203748条有效数据。

由于经济类文章所使用的专有名词较多,和各大文本数据词库相差较大,故将期刊的关键词作为此次文本数据挖掘的基础词库。同时,为了提高主题模型的训练效率,剔除摘要数据中的停用词和低频词,得到主题模型的实验语料库。最后将语料库进行向量化处理,形成主题模型的输入文本文档集。

四、经济学期刊摘要的主题分析

(一) 最优主题数的求解

LDA主题模型训练前,需事先确定主题数。

本文将Roder等(2015)^[12]提出的主题一致性得分(Coherence Score)作为确定主题数的标准,该指标越大说明主题的一致性越好。不同主题个数下的主题一致性得分如图1所示。可以看出,主题数为20时,主题一致性得分最高,故将LDA模型的主题数确定为20。

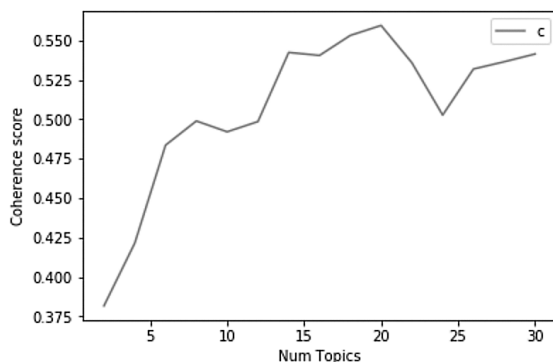


图1 主题一致性得分随主题数变化图

(二) 经济学研究主题及其概率分布

根据LDA模型假设,每个主题是由多个关键词构成的一个概率分布,即“主题-词”,故可以通过归纳每个主题下关键词的共性特征总结出主题。首先,对LDA模型计算出的主题进行可视化(如图2所示),将所有主题以气泡图的形式展示。运用Jensen-Shannon(JS)散度^[13]计算主题之间的距离,进而得到主题间距离矩阵,然后在该矩阵上实施多维缩放(multidimensional scaling)算法,将主题按距离远近排列在二维坐标系上,每个气泡代表一个主题,气泡的大小代表在语料库中该主题所占的比例。

图3列出语料库中与主题10最相关的前30个关键词,每个关键词对应的条形图总长度代表该关键词在语料库中出现的总频次,而深颜色部分代表该关键词在该主题下出现的频次。可以看出,该主题与国际贸易相关,代表该主题的主要关键词有贸易、竞争、出口、产品、中国、市场、竞争力、制造业、比较、工业、产业等。

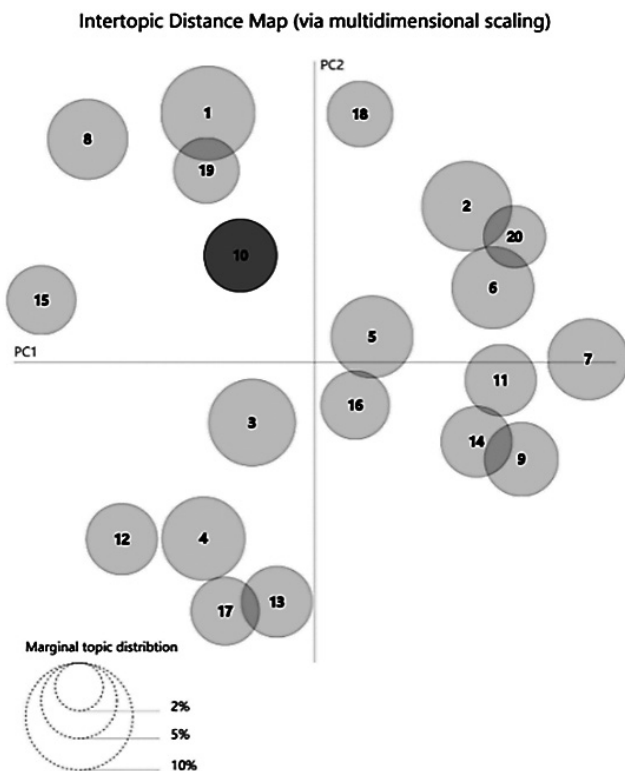


图2 主题分布气泡图

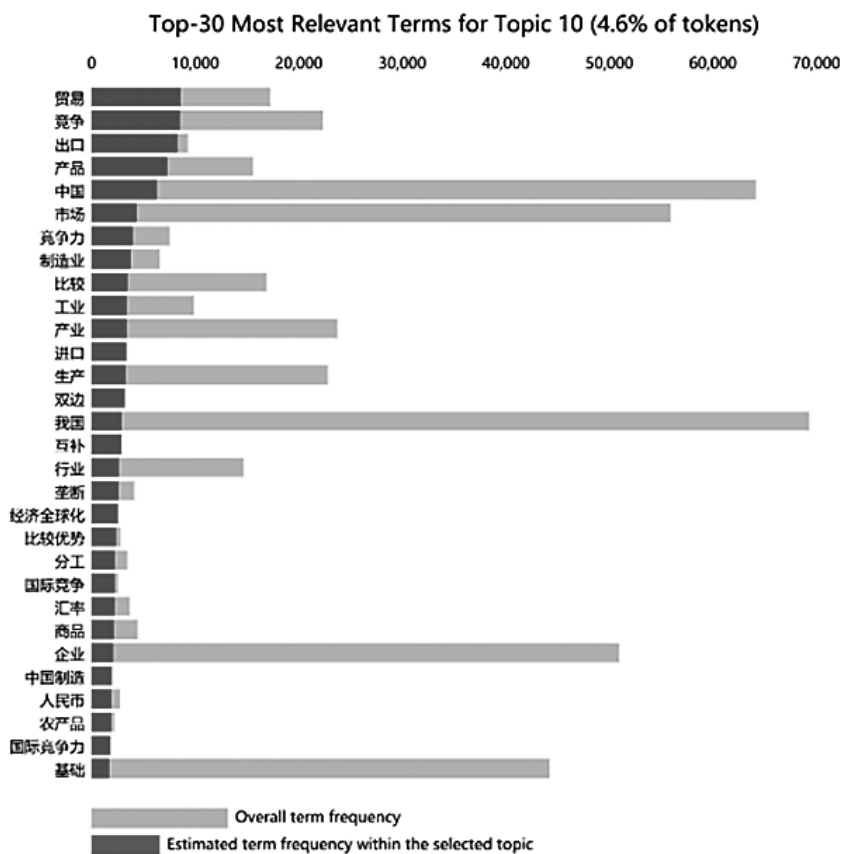


图3 主题10中排名前30的关键词分布

根据“主题-词”概率分布(表1展示了每个主题下概率最大的前10个词)总结出20个主题分别为(按概率大小排序,主题概率见图4):宏观经济计量学、农业经济学、金融市场、企业管理、世界经济、产业经济学、改革开放、微

观计量经济学、马克思主义政治经济学、国际贸易、中国特色社会主义、财政学、金融风险管理、上市公司与资本市场、数理经济学、可持续发展与知识经济、公司金融学、区域经济学、收入分配、环境经济学。

表1 每个研究主题的前10个关键词及其概率

主题名称	关键词1	关键词2	关键词3	关键词4	关键词5	关键词6	关键词7	关键词8	关键词9	关键词10
公司金融学	信息 0.069	投资者 0.035	会计 0.034	财务 0.032	投资 0.025	公司 0.023	监管 0.023	管理 0.022	行业 0.019	我国 0.019
农业经济学	生产 0.032	农业 0.031	农村 0.027	经济 0.025	社会 0.023	劳动 0.021	农民 0.020	城市 0.016	劳动力 0.015	我国 0.014
金融市场	货币 0.029	金融 0.026	市场 0.025	银行 0.025	融资 0.022	我国 0.022	经济 0.021	贷款 0.019	信贷 0.018	流动 0.016
数理经济学	模型 0.045	收益 0.035	价格 0.032	市场 0.031	成本 0.028	利润 0.022	投入 0.021	博弈 0.020	资产 0.019	福利 0.018
宏观经济计量学	增长 0.061	经济 0.049	经济增长 0.041	模型 0.040	中国 0.027	资本 0.022	数据 0.019	投资 0.015	效应 0.014	消费 0.014
马克思主义政治经济学	经济 0.048	商业 0.039	我国 0.031	市场经济 0.031	市场 0.030	社会 0.030	商业银行 0.027	经济学 0.021	保险 0.020	马克思 0.018
收入分配	收入 0.086	调查 0.053	数据 0.045	分配 0.029	居民 0.028	差距 0.023	家庭 0.022	模型 0.021	平等 0.015	中国 0.015
区域经济学	地区 0.086	区域 0.056	西部 0.033	经济 0.028	开发 0.022	资源 0.021	区域经济 0.020	经济发展 0.020	西部地区 0.019	我国 0.016
中国特色社会主义	社会 0.056	经济 0.039	经济学 0.028	基础 0.027	中国 0.027	政治 0.027	文化 0.022	实践 0.020	特色 0.018	历史 0.018
企业管理	企业 0.080	管理 0.043	经营 0.036	资源 0.033	配置 0.026	市场 0.026	组织 0.024	竞争 0.022	国企 0.019	效率 0.019
金融风险 管理	金融 0.093	风险 0.064	机构 0.037	危机 0.036	银行 0.032	信用 0.029	金融危机 0.022	市场 0.021	监管 0.021	我国 0.020
可持续发展 与知识经济	可持续 0.040	消费 0.037	持续发展 0.035	可持续 0.029	产品 0.027	基础 0.024	消费者 0.023	资源 0.023	知识 0.020	企业 0.019
产业经济学	产业 0.062	企业 0.043	小企业 0.031	中小企业 0.031	我国 0.027	技术 0.026	竞争 0.024	服务 0.021	经济 0.020	产业结构 0.018
微观计量 经济学	数据 0.070	检验 0.046	效应 0.044	面板 0.040	面板数据 0.034	模型 0.033	中国 0.030	技术 0.023	实证检验 0.022	效率 0.016
世界经济	国家 0.056	经济 0.049	中国 0.049	世界 0.042	贸易 0.034	自由 0.025	美国 0.024	对外 0.022	发达国家 0.018	开放 0.017
上市公司与 资本市场	公司 0.078	上市 0.062	上市公司 0.051	市场 0.041	证券 0.039	资本 0.036	企业 0.034	我国 0.030	投资 0.022	资本市场 0.021
财政学	政府 0.043	税收 0.034	我国 0.029	完善 0.027	税改 0.026	利益 0.024	主体 0.023	财政 0.021	公平 0.019	税制 0.019
国际贸易	贸易 0.044	竞争 0.043	出口 0.042	产品 0.037	中国 0.032	市场 0.022	竞争力 0.020	制造业 0.019	比较 0.018	产业 0.017
环境经济学	环境 0.056	经济 0.054	政府 0.035	经济发展 0.026	生态 0.025	地方政府 0.022	资源 0.019	变迁 0.019	制度变迁 0.017	城市 0.015
改革开放	改革 0.075	经济 0.041	中国 0.032	开放 0.026	我国 0.024	社会 0.023	改革开放 0.021	市场 0.016	科学 0.015	政府 0.014

2001年11月中国加入世界贸易组织,这是顺应经济全球化具有里程碑意义的重大举

措,标志着中国对外开放进入了新阶段。在加入WTO过程中,学者们认为这会对国家宏观

经济产生冲击,涉及金融业、商业、农业、信息业等诸多产业。因此,经济学研究主题也聚焦在宏观经济、全要素生产率、金融市场与企业

管理等方面(见图4)。此外,中国经济在经历了三十多年的高速增长后,正在发生阶段性转变。党的十九大报告指出,中国经济已由高速

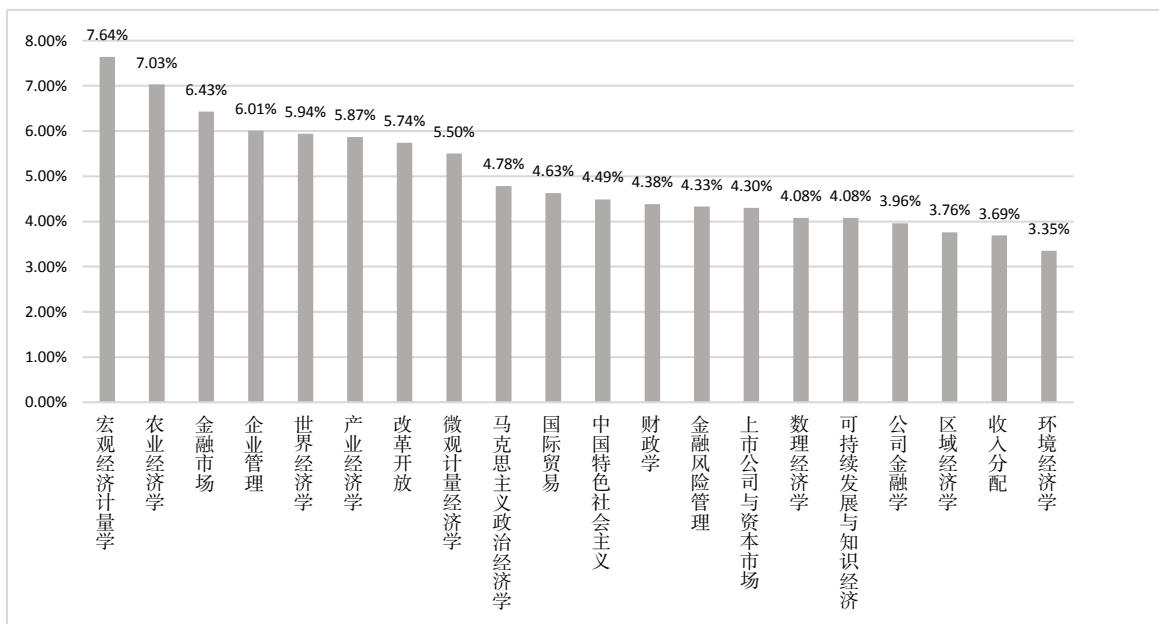


图4 20个主题的概率分布

增长阶段转向高质量发展阶段。高质量发展在经济学意义上,也可从不同的视角去探讨和解读。在宏观层面,主要指国民经济的整体质量和效率,通常用全要素生产率来衡量;在中观层面,主要指产业和区域发展的质量;在微观层面,主要指企业产品和服务的质量。要推动中国高质量发展,关键在于深化改革,不断探索找出符合国情的经济发展模式,这更彰显出宏观经济研究的重要性。

(三) 经济学研究主题的层次聚类

在确定了每个主题下关键词的概率分布后,运用JS散度计算出主题间的距离矩阵。JS散度越小则表示概率分布越相似,说明主题越相关。采取凝聚式的层次聚类方法(一种自下而上逐渐汇集的聚类方法),基于主题间的距离矩阵,通过不断将距离最小的对象合并为同一簇,最后构建出聚类二叉树,其中树叶为20个经济学研究主题。该二叉树刻画了20个经济

学研究主题之间的相关关系(见图5)。

图5展示了21世纪以来经济学科CSSCI期刊中热点关注的研究主题及其相关关系。可见,20个经济学研究主题可以概括为4大类:中国特色社会主义政治经济学、世界经济与产业发展、计量经济学、金融市场与企业管理。中国特色社会主义政治经济学主要探讨中国特色社会主义、马克思主义政治经济学、改革开放等主题。以习近平同志为核心的党中央全面推进中国特色社会主义政治经济学建设,习近平总书记首次提出了中国特色社会主义政治经济学的范畴,强调要立足我国国情和我国发展实践,学习、研究、运用马克思主义政治经济学,发展当代中国马克思主义政治经济学,并概括说明了党的十一届三中全会以来形成的当代马克思主义政治经济学的诸多重要理论成果。他还多次从不同的角度提出了发展中国特色社会主义政治经济学的必要性和路径,将

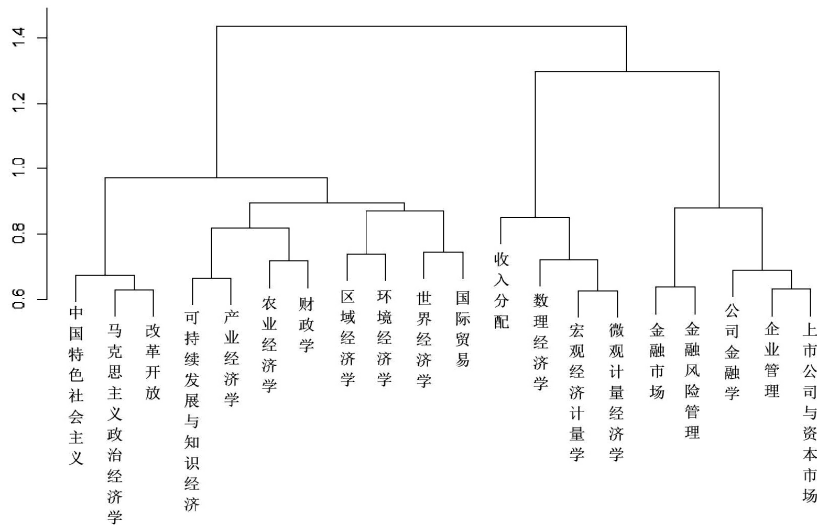


图5 20个经济学研究主题的层次聚类图

党中央对建设中国特色社会主义政治经济学的重视与指导提升到前所未有的高度。从主题的层次聚类图也可以看出经济学期刊正在积极地承担起推动中国特色社会主义政治经济学发展的职责。

(四) 期刊的主题分布

根据LDA主题模型得出的“文档-主题”概率分布,可以计算出“期刊-主题”的概率分布,进一步归一化处理后得到70个期刊在各个主题下的概率分布。如图6所示,用颜色深浅代表期刊中各研究主题的概率大小,即期刊对研究主题的偏好程度。

横向来看,期刊的发文偏好呈现以下特点:一是有些期刊的主题分布较集中,如《农村经济》的主要研究主题是农业经济学,《政治经济学评论》的主要研究主题是马克思主义政治经济学,《国际贸易》的主要研究主题是世界经济,《税务研究》的主要研究主题是财政学。这类期刊属于某一具体研究领域的经济类期刊;二是有些期刊的主题分布较分散,如《经济研究》涵盖了几乎所有研究主题,属于综合性经济类期刊。

从纵向来看,各研究主题在各期刊中的分布不均衡:一是有的研究主题主要集中发表于一种期刊,如马克思主义政治经济学主题主要发表于《政治经济学评论》中,收入分配主题

主要发表于《劳动经济研究》中;二是有的研究主题分散发表于多个期刊,如世界经济主题,在《国际贸易》、《世界经济与政治论坛》、《国际经济评论》、《亚太经济》等期刊中均有较大占比。

由中国人民大学主办的《政治经济学评论》的发刊宗旨和方针是以马克思主义为指导,高举中国特色社会主义的伟大旗帜,推进马克思主义政治经济学的中国化和时代化。在政治经济学研究中,注重理论和现实问题的结合,鼓励创新与学术自由,积极推动国内外经济学界的交流合作。图6中《政治经济学评论》的研究主题主要围绕中国特色社会主义以及马克思主义政治经济学,而这与其发刊宗旨是一致的。同样由上海财经大学主办的《财经研究》期刊的研究主题主要是宏观经济计量学、微观计量经济学以及数理经济学,该期刊主要致力于用量化工具探索经济市场发展运行规律的文献。综上所述,各期刊对研究主题的偏好也正反映了期刊的创办宗旨和研究重点。

(五) 经济学研究主题的时间演变

根据LDA主题模型输出的“文档-主题”概率分布,再结合各文章(即文档)的发表时间,计算出“年代-主题”分布。图7展示了各经济学研究主题占比的逐年变化规律,其中每一种颜色代表一个研究主题。

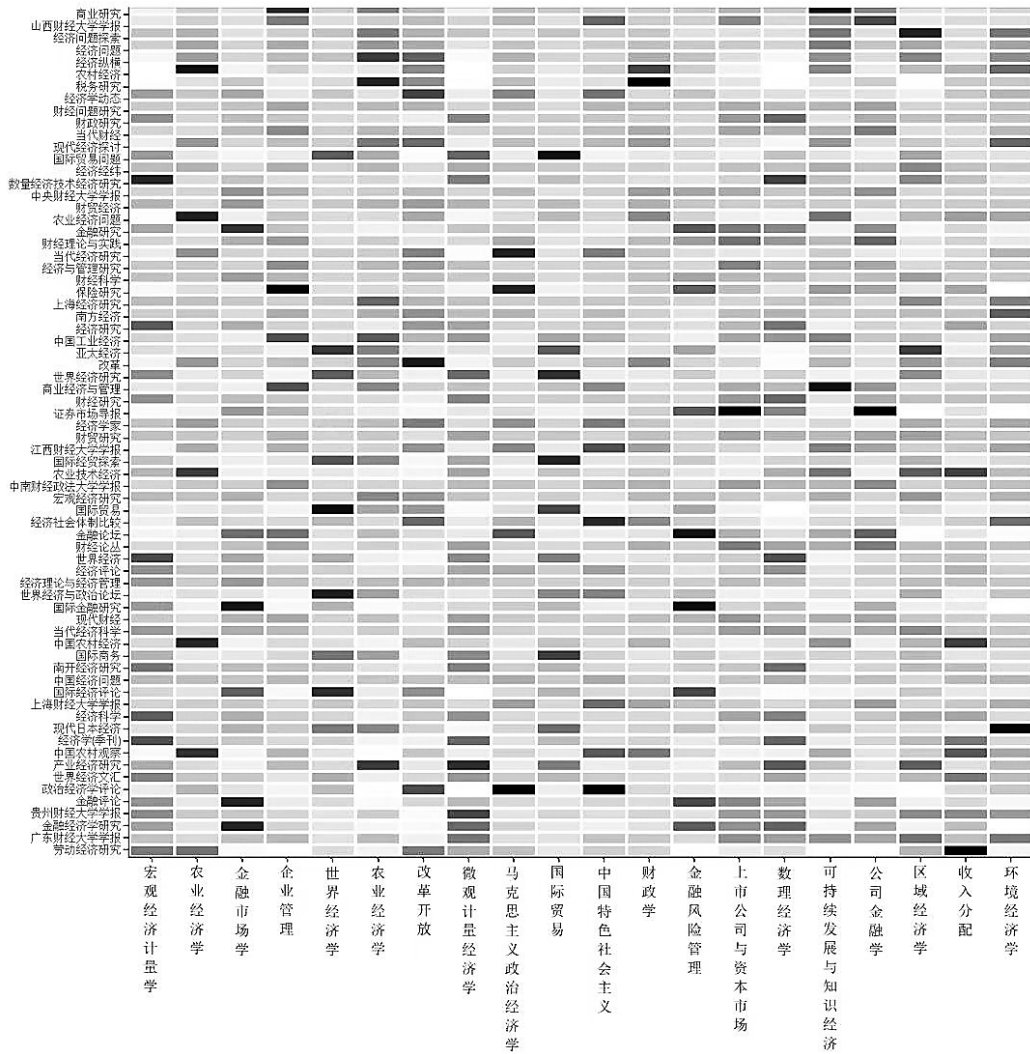


图6 各经济学期刊的研究主题分布

横向来看,从2000年至2018年,各研究主题的发展趋势差别迥异:一是有些研究主题逐年占比相对稳定,如农业经济学、中国特色社会主义在各时间段占比变化不大,而且一直是比较重要的研究主题;二是有些主题呈现出快速增长的势头,如微观计量经济学和宏观经济计量学

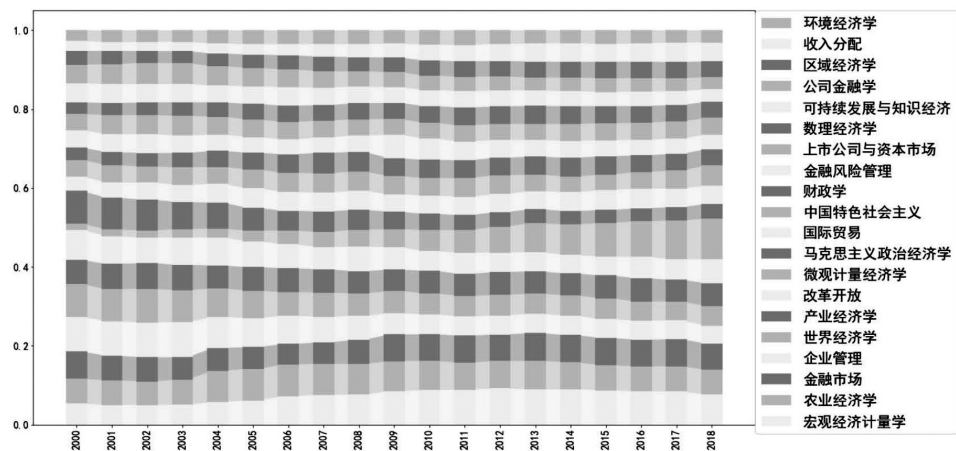


图7 各经济学研究主题占比的逐年变化图(图例与纵坐标顺序一致)

主题占比逐年增加；三是有些主题则呈现衰退趋势，如马克思主义政治经济学研究主题，然而这并不意味着该主题的研究成果减少了，而是该主题的部分研究成果在哲学类期刊中发表。

纵向来看，不同时段对各经济学研究主题有所侧重。经济学研究与国家的重大社会经济问题密切相关，与中央政治决策高度统一。具体而言，将经济学研究主题的变化趋势大致划分为三个阶段。

第一个阶段是2000年至2005年，即第九个五年计划的最后一年和第十个五年计划期间。随着中国加入世界贸易组织，经济全球化为中国进一步融入世界经济体系，深化改革开放提供了难得的机遇。WTO所推行的市场经济运行体制有利于我国加快市场化改革进程，全球性的资源配置有利于国内经济结构的调整和优化，促进产业升级。随着中国经济与国际逐步接轨，中国经济越来越面临多方面的严峻挑战。例如，如何控制扩大开放所可能带来的经济风险，尤其是1997年亚洲金融危机所引发的对金融安全的关注。这一时期对外经济迅猛发展，对外开放水平不断提高。从经济学研究主题的年度分布也不难看出，国内学者在此时期也将研究重心转向世界经济学和改革开放。在完善社会主义市场经济体制的同时，重视对马克思主义政治经济学的研究，探索中国道路。

第二个阶段是2006年至2010年，即第十一个五年规划期间。“十一五”规划时期复杂多变的国内外经济环境和日趋激烈的市场竞争，对我国经济社会发展和安全也提出了诸多新的挑战。我国正处于并长期处于社会主义初级阶段，生产力还不发达，城乡区域发展不平衡，经济结构不够合理，使经济社会发展与资源、环境之间的矛盾凸显；技术储备不足、自主创新能力不强，直接影响了我国经济国际竞争能力的提高和可持续发展；解决“三农”问题的任务相当艰巨，就业压力依然较大；公平与效率、经济与社会发展之间不协调的矛盾日益显现。在这一时期农业经济学和产业经济学逐渐成为经济

学的研究热点，符合当时的经济现状。

第三个阶段是2011年至2018年，即第十二个五年规划期间和第十三个五年规划的前三年。

“十二五”规划期间，我国经济发展中不平衡、不协调、不可持续问题依然突出，主要是经济增长的资源环境约束强化，投资和消费关系失衡，收入分配差距较大，产业结构不合理，城乡区域发展不平衡，就业总量压力大和结构化矛盾并存，物价上涨压力大。在这一时期宏观经济计量学蓬勃发展，对宏观经济数据进行计量分析，探索影响宏观经济发展的关键因素，挖掘其中蕴含的潜在规律，解决宏观经济发展所面临的问题与矛盾。“十三五”规划期间，国际竞争日益激烈，贸易保护主义抬头，中美贸易摩擦持续升温，对国内企业的发展环境产生巨大影响。在这种严峻的国际形势下，学者们则是更多地从微观计量经济学角度寻找解决问题的出路。

五、结论与启示

本文运用LDA主题模型对2000—2018年的经济学科CSSCI来源期刊摘要汉语文本数据进行主题挖掘，通过概率估计和可视化分析，总结出我国经济学研究的热点以及发展演变趋势。主要结论有：一是根据经济学期刊摘要文本数据挖掘出20个研究主题，分别为：宏观经济计量学、农业经济学、金融市场、企业管理、世界经济、产业经济学、改革开放、微观计量经济学、马克思主义政治经济学、国际贸易、中国特色社会主义、财政学、金融风险管理、上市公司与资本市场、数理经济学、可持续发展与知识经济、公司金融学、区域经济学、收入分配、环境经济学；二是根据主题的相似性进行层次聚类，主题间呈现出一定的亲疏关系。20个经济学研究主题可概括为4大类，同一类的各主题之间相关性较强，而不同类的主题之间相关性较弱；三是各经济学期刊对研究主题有所偏好，有些期刊侧重于某个研究主题，而有些期刊涵盖多个研究主题；四是根据研究主题的年代分布发现，经济

学研究主题紧紧围绕当时的社会经济现实,随着经济社会的发展逐渐变化。“十五”计划期间中国刚刚加入世界贸易组织,世界经济、改革开放成为了当时经济学研究的热点。“十一五”规划期间生产力发展不平衡促使经济学研究重点转向农业经济学和产业经济学。“十二五”和“十三五”规划期间,国际形势严峻,宏观和微观经济环境都在发生巨变,以计量方法研究经济发展规律成为经济学研究的又一主流趋势。

经济学期刊摘要数据的主题挖掘还能够为青年学者选择研究题目和投稿期刊提供参考建议。“年代-主题”分布展示了不同的历史阶段各经济学研究主题的占比情况及演化规律,学者可以据此审视研究选题的时代意义;“期刊-主题”分布展示了各期刊对研究主题的偏好,学者可以据此选择恰当的投稿期刊,使文章投稿能够有的放矢,进而提高接受发表的可能性。

参考文献:

- [1]习近平. 一个国家、一个民族不能没有灵魂[J]. 内蒙古宣传思想文化工作, 2019(4): 4-6.
 [2]李姣. CSSCI及其评价功用解析[J]. 科技信息, 2011(14): 355.
 [3]Blei D M, Ng A Y, Jordan M I, Lafferty J. Latent dirichlet allocation[J]. Journal of Machine

Learning Research, 2003(3): 993-1022.

- [4]陈江锋, 于建军. 基于主题模型的结构化Web服务发现机制[J]. 北京航空航天大学学报, 2008(6): 734-738.
 [5]吕韶华. 面向中文评论文本的情感倾向性研究[D]. 大连: 大连理工大学, 2010.
 [6]张小平. 主题模型及其在中医临床诊疗中的应用研究[D]. 北京: 北京交通大学, 2011.
 [7]杨星, 李保利, 金明举. 基于LDA模型的研究领域热点及趋势分析[J]. 计算机技术与发展, 2012, 22(10): 66-69+74.
 [8]胡吉明, 陈果. 基于动态LDA主题模型的内容主题挖掘与演化[J]. 图书情报工作, 2014, 58(2): 138-142.
 [9]王鹏, 高铨, 陈晓美. 基于LDA模型的文本聚类研究[J]. 情报科学, 2015, 33(1): 63-68.
 [10]李博, 陈志刚, 黄瑞, 郑祥云. 基于LDA模型的音乐推荐算法[J]. 计算机工程, 2016, 42(6): 175-179+184.
 [11]曾子明, 王婧. 基于LDA和随机森林的微博谣言识别研究——以2016年雾霾谣言为例[J]. 情报学报, 2019, 38(1): 89-96.
 [12]Roder M, Both A, Hinneburg A. Exploring the space of topic coherence measures[C]. Web Search and Data Mining, 2015: 399-408.
 [13]Lin J. Divergence measures based on the Shannon entropy[J]. IEEE Transactions on Information Theory, 1991, 37(1): 145-151.

【责任编辑 许鲁光】

The Development and Evolution of Economic Research Topics in China since the 21st Century: Topic Mining Based on Abstracts of Economic Journals

GUAN Guoyu & HU Hanyun

Abstract: This article aims to discover the main fields and development trends of China's economic research since the 21st century, based on the abstracts of CSSCI source journals of economics from 2000 to 2018. It shows that, (1) there are about 20 topics of China's economic research, (2) there are certain relationships among the topics, and (3) each journal has a preference for the topics. The highlights of China's economic research reflect the different historical stages of China's economic development. The topic mining of the abstracts of journals have witnessed the development of economic research in China since the 21st century, and provided suggestions for young scholars to choose research topics and journals to submit to.

Keywords: economics research; research hotspots; bibliometrics; topic mining