



开放科学
(资源服务)
标识码
(OSID)

金融行业的规模、特征和发展水平 ——基于注册大数据的前向分析

孙广芝¹ 杨秀丹² 张东冉² 王凤琪²

1. 中国标准化研究院 北京 100038;

2. 河北大学管理学院 保定 071002

摘要: 金融业对国家产业经济发展具有重要推动作用,是国民经济发展的重要保障,利用企业注册大数据对金融业进行前向分析有利于评估金融行业发展。本文基于不同城市和地区的金融业企业注册量和注册资金数据,运用时空统计学理论,对金融业近十年的规模、特征和趋势进行分析,研究结果为金融业在空间上呈现条块化,在时间演化上趋向平稳。

关键词: 金融行业;统一社会信用代码;企业注册数据

中图分类号: G35

The Scale, Characteristics and Development of the Financial Industry: A Pro-forecast Analysis Based on the Registration Data of Unified Social Credit Code

SUN Guangzhi¹ YANG Xiudan² ZHANG Dongran² WANG Fengqi²

1. China National Institute of Standardization, Beijing 100038, China;

2. School of Management, Hebei University, Baoding 071002, China

Abstract: The financial industry plays an important role to promote the development of the industrial economy and is an important guarantee for the development of the national economy. Using the big data of unified social credit code to pre-forecast and analyze the development of the financial industry is conducive. Based on the registration data and registered

基金项目: 国家社科基金“新型城镇结构演化中的信息驱动力研究”(14BTQ042)。

作者简介: 孙广芝(1972-), 博士, 副研究员, 研究方向: 数据标准化、数据质量治理与大数据分析, E-mail: sungz@cnis.gov.cn;

杨秀丹(1970-), 博士, 教授, 研究方向: 信息社会、情报学理论与方法, E-mail: poshyang@126.com; 张东冉(1994-), 硕士研究生,

研究方向: 信息分析; 王凤琪(1995-), 硕士研究生, 研究方向: 信息分析。

capital data of financial enterprises in different cities and regions, this paper analyzes the scale, characteristics and trends of the financial industry in the past ten years using the theory of space-time statistics. The results show that financial industry is bar-shaped from the space perspective, and it shows a stable status from the time trend.

Keywords: Financial industry; unified social credit code; enterprise registration data

金融行业是国民经济发展运行的命脉，在国民经济结构中占据重要位置^[1]。2008 年全球金融危机爆发后，我国政府出台了一系列的金融行业调整 and 改革政策以完善金融宏观调控体系。在政府部门提出加大金融行业对经济增长的支持力度的重要举措下，银行业整体实力明显增强，证券业行业空间逐渐打开，保险市场运行也呈现良好的发展态势^[2]。由此，金融危机之后，我国金融行业呈现良好的发展态势，金融体系逐渐步入正轨。据统计，2015 年国内金融行业的增加值为 57500 亿元，增长率达到 15.9%，其增加值占 GDP 的比重也达到 8.5% 的历史峰值，高于美国（7.2%）和英国（7.3%），约为日本、德国、法国、意大利和西班牙的 2 倍^[3]。与此同时，我国在国际金融合作活动中转向主动，积极探索符合自身发展的合作模式，逐步取消对外资金融机构的限制，加强金融业的对外开放，以适应国际金融竞争^[4]。上海、天津和广东等自贸区积极响应国务院号召，制定吸引外资金融机构的新政策来拉动区域经济增长^[5]。金融行业的调整 and 改革举措，不仅为其自身发展注入新的活力，同时为国民经济的稳定发展提供了保障，金融行业在经济发展中的贡献和价值越来越高。但是，近一两年，随着国际竞争以及世界经济的变化，我国金融业

发展面临着更多的挑战，在相关问题也逐渐呈现出来，因此，如果从全样本数据对金融业发展进行适当的梳理，从经济行为全过程分析和揭示近十年金融业的规模、结构和发展特点，无疑具有重要意义，本文即基于这样的背景，运用时空统计学理论，挖掘未曾充分利用的全样本注册大数据的价值，以经济行为事前阶段数据为基础，尝试前向描述和分析金融业的发展、特征和发展水平。

1 文献综述

目前关于金融行业发展的研究主要集中在两方面：一是发展水平和经济关系研究，如高万东等测量金融深度和金融业贡献率以评价国内金融机构在经济市场的重要程度^[6]，Haykel Zouaoui、Ling Xiong 等以面板数据为基础分析金融发展与经济波动关系^[7-8]；二是区域金融行业研究，如唐雨虹基于金融发展、经济增长、城乡收入差距以及企业绩效等历史数据，阐述西藏金融发展问题^[9]；纪玉俊等人测度金融业集聚水平与区域经济发展之间的关系，认为形成了以北京为中心的环渤海地区金融集聚、以上海为中心的长三角地区金融集聚和珠三角地区的金融集聚，并提出各省份金融业集聚水平

呈橄榄形分布^[10]。

目前关于金融行业发展水平的研究多数是以经济行为的事中和事后情况为研究对象,对经济行为的现状及未来走向作出判断。以企业经营的现实情况为依据所做的经济发展研判,具备一定的真实性和准确性。但是,基于后向数据的宏观经济分析,对政策的制定和研判缺乏必要的预警机制。利用适当的数据分析手段分析事前信息源可以预测社会和经济行为^[11]。企业是经济活动的主体,企业注册信息在一定程度上能够反映经济运行状态的规模和分布,因此对企业注册数据进行全面收集和分析,可以反映一定时间段内经济发展情况。统一社会信用代码数据涵盖各机关、企事业单位等机构的机构代码、机构名称、机构类型、法定代表人、业务范围、经济行业、职工人数、行政区划、单位地址和邮政编码等丰富信息,是当前最权威、最全面的数据资源之一,是国家宏观经济、社会发展等前向观测和分析的重要基础数据来源^[12]。本文提出运用前向分析方法,结合企业统一社会信用代码注册数据对金融业进行分析,试图通过对企业经济行为的事前研究,分析金融业企业的总体规模和空间特征,为相关经济政策的制定提供依据。

2 研究方法 with 样本数据

研究思路,从两个角度进行,一是时空演变的角度,即从时间和空间维度来统计和呈现我国金融业的规模、特征和发展水平。基于时间维度的横向分析是对同一时期数据资料进行研究,有利于概括社会经济现象在特定时期

的变化趋势;基于地理位置和形态特征的空间分析是从地理空间的角度观测具体事物的分布特征,结合空间分析的可视化分析方法可以清晰地呈现金融业的区域分布情况。本文选择克里格方法测度金融业在全国的地区空间分布差异情况。克里格方法是以结构分析以及变异函数理论为基础的一种空间局部插值法,也是在某一区域内对在此范围内的变化量进行的一种无偏最优估计算法,往往根据未知样点有限的邻域中一些已经获知的样本点相关数据,考虑样本点的分布位置、大小情况与形状轮廓,变异函数所体现的结构信息,以及与未知样点之间位置的空间分布关系等因素,综合判断后进行线性无偏最优估计^[13]。

二是经济行为全过程分析角度,企业机构从开始注册到投入生产创造收益需要经历一段时间,因而社会经济生产总值变化必然滞后于企业机构的增长变化情况,从宏观上看,二者之间存在关系,因此本文亦研究企业注册情况与经济发展之间的关联程度,基于企业注册数据和金融业增加值,分析二者的关系。由于金融业的企业注册量、注册资金及其增加值都是连续性变量,为验证其关联程度大小,选择灰色关联法进行测度。灰色关联法是一种通过计算灰色系统中有限数据序列来定量描述和分析系统内部诸因素对系统发展结果的影响程度,以及诸因素之间关联程度的统计方法。同时本文选择对数回归模型进行实证分析,确定金融业注册数量和注册资金的影响程度。

本文数据来源为全国组织机构代码统一社会信用代码数据服务中心的企业注册大数据,选择2008—2014年期间的金融业企业注册数量、注

册资金以及相应的外资情况进行收集整理作为研究的样本数据。同时,以直辖市、省会城市、计划单列市和重点城市等规模以上城市为研究对象,同步考虑其地域代表性,选取144个城市作为区域分析样本。其中,华北地区选取北京、天津、石家庄等城市;华中地区选取郑州、南昌、武汉、长沙、株洲等城市;华东地区选取上海、广州、深圳、杭州、宁波等城市;西南地区选取重庆、北海、南宁、海口、三亚等城市;东北地区选取包括大连、沈阳、长春、哈尔滨等城市。

3 金融业规模、特征和趋势分析

3.1 金融业规模

金融是现代经济的核心,分析金融行业的规模,认识金融业在整体经济中的优势,可以促进资源的合理配置,助力实体经济的发展^[14]。本研究对2008—2014年的金融业的企业整体注册情况和外资企业注册情况进行统计,并与其他行业企业注册数据进行比较分析,得到截

止2014年金融业的总体情况,结果如图1所示,其中,内环表示金融业企业注册数量占比,外环表示金融业企业注册资金占比情况。截止2014年底,金融业的企业整体注册量累积达到172317家,注册资金量为1290万亿元,外资企业的注册量为3755家,注册资金达到3070亿元。在十一个行业^①中,金融业的企业注册量占比为2.48%,但企业注册资金占比例将近三分之一。从外资企业注册情况来看,注册数量占比为1.10%,注册资金数量占比1.76%。在所有行业中,金融业外资企业的注册规模偏小,这和早期我国金融领域对外资的准入政策和限制条件有关。种种限制导致外资企业基数小,发展也受到一定程度的制约。

综合来看,金融业企业注册数量占比小而注册资金占比较大,说明金融行业小微企业较少,对稳定金融秩序较为有利,但对金融市场的活跃程度有一定影响。金融业的体量在全国行业里较小,但其掌握的资本在整个国民经济体系中占据重要地位。

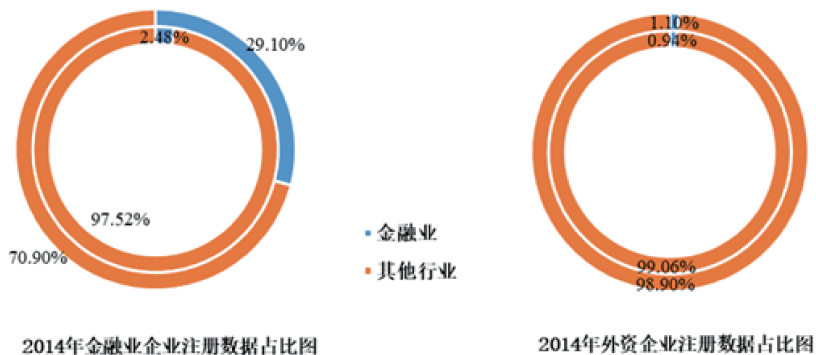


图1 金融业企业注册占比图

数据来源:作者根据全国组织机构代码数据服务中心相关数据整理

①本研究参考《国民经济行业分类》(GB/T 4754),将企业所属经济行业划分为仓储业、交通运输业、金融业、零售业、批发业、信息传输业、软件和信息技术服务业、居民服务、修理和其他服务业、文化、体育和娱乐业、邮政业和制造业。

金融业的发展与实体经济发展密切相关,尤其是制造业。制造业是国民经济发展的支柱型产业,也是国家现代化的主导力量^[15]。二者的经济方式、产业构成等存在较大的差异,在社会中的作用也各有优势,本文选择金融业和制造业进行比较,分析二者规模、结构等的差异,以此进一步确定金融业的发展特点。如表1所示,制造业企业注册量占比是金融业的十几倍,

但是金融业的注册资金量占比将近三分之一。这个数据基本证实了金融业和制造业在国民经济中的不同地位,制造业作为我国经济结构中的基础支柱型产业,规模庞大;而金融业的企业注册资金占全部行业的三分之一,说明金融业的可支配资金充裕,如能够在资源合理配置方面发挥作用,则可保障实体经济企业的长期稳定发展。

表1 金融业和制造业的企业注册情况(单位:%)

企业注册情况	行业	2008	2009	2010	2011	2012	2013	2014
注册量	金融业	2.305	2.533	2.550	2.540	2.475	2.468	2.479
	制造业	37.436	33.505	32.237	31.220	30.244	30.133	30.202
注资金量	金融业	22.202	19.420	18.238	18.410	29.575	29.095	29.104
	制造业	37.880	34.944	52.767	56.252	42.517	41.091	41.088

数据来源:作者根据全国组织机构代码数据服务中心相关数据整理

3.2 金融业分布特征

金融业的发展和扩大带来人才、机构、资本和资源的日益集中,也对我国区域经济的协调发展起着越来越重要的作用。金融集聚不仅显著促进城市经济增长,在邻近城市还存在空间溢出效应^[16]。大多数学者选择城市面板数据如金融业增加值、金融机构年末贷款总额、年末金融机构各项存款余额等验证金融发展的空间效应^[17-18],本文则依据企业经济活动发生前的企业注册信息,分析金融行业空间分布特征。本研究对采集的数据进行脱敏处理以确保分布特征尽量符合正态分布,并引入企业指数这一概念,处理公式为:

$$p = \frac{\ln(x_i)}{\ln(\max(x_j))} \times 100 \quad (\text{公式1})$$

其中 p 为计算对象的企业指数, $\ln(x_i)$ 为计算对象中样本值的自然对数值, $\ln(\max(x_j))$ 为

计算对象样本最大值的自然对数值。依据空间统计学方法,通过等高线地形图进行指数梯度和热点分析,可视化直观描述金融业的城市和区域分布特点。

本文选择2014年度的金融企业和金融外资企业注册量和注册资金量为样本,利用克里格法,按照 200×200 的网格进行时空插值分析得到如图2所示的分布结果。可以看出,金融机构的地域分布条块化趋势显著,基本形成了:以北京、天津为核心的京津冀区带;以哈尔滨和长春为核心的东北区带;以重庆、成都为核心的西南区带;以广州、深圳为核心的珠三角区带;以上海、苏州、杭州、南京为核心的长三角区带。

纪玉俊等基于金融业产值和各地区生产总值计算,认为金融行业在区域空间上大体呈橄榄形分布。本研究根据企业注册信息得出金融

业地理分布的条块化趋势显著。由于统计数据的不同,必然有所差别,但是两项研究在金融业聚集水平上结论比较一致:我国已经形成一定规模的金融聚集,并且以北京为核心的京津冀地区、以上海为核心的长三角地区、以广州、深圳为核心的珠三角地区的金融业集聚水平明显高于其他地区。金融业的高度和发展与当地发

达的经济水平是分不开的,本研究也证明,在经济活动开始前,金融作为一种资源,会流向经济发达,资源利用率和回报率高,风险小的地区^[19]。同时,金融业的发展受益于国家宏观经济战略的支持。由于国家政策对上海自贸区、天津自贸区、广东自贸区的支持,在一定范围内使其对周边城市产生了辐射和渗透效果。

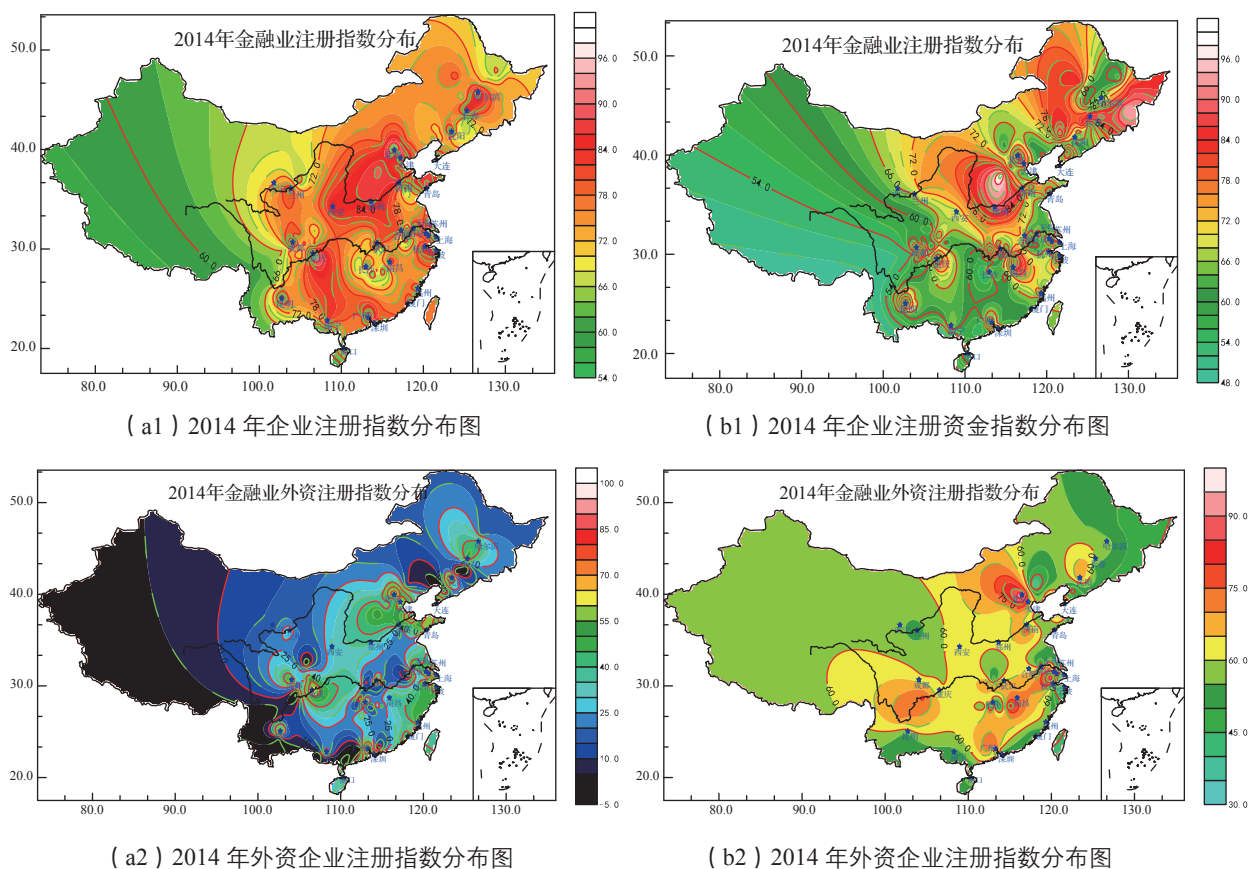


图2 金融业企业注册分布图

部分城市缺少外资情况的统计数据,做了适当的数据整理

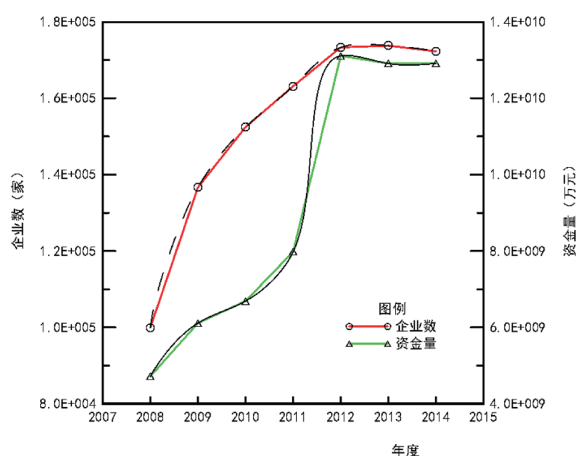
从外资企业注册数量和注册资金量的分布情况看,外资发达地区主要集中在北京、天津、广州、深圳、上海等核心城市,重庆、成都等城市有一定的优势。外资金融机构在华注册往往会受到地区经济水平、制度因素、政策因素、集聚因素等影响。深圳、

北京、上海等城市作为市场化程度较高的地区,最受外资的青睐。由于地区经济发展增长快、国家政策导向支持影响,天津、成都、重庆等城市在吸引外资上也具备优势。但是外资企业整体规模上并不突出,未形成典型的区域或条块特色。

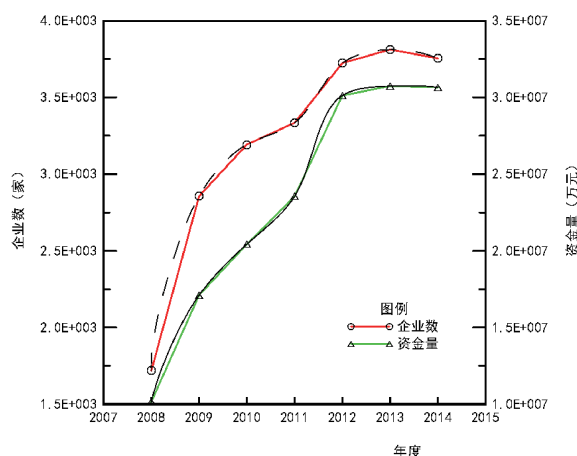
3.3 金融业发展水平

关于金融业发展水平的测度,国内外不同学者基于统计数据提出了许多不同的指标,构建了相应的金融发展评价指标体系。

从前向观测角度分析金融业发展水平,以144个城市的金融业企业注册量、注册资金量以及外资注册数据进行统计分析,其结果见图3。



(a) 金融业企业注册变化图



(b) 金融业外资企业注册变化图

图3 金融业企业变化图

注: 多个城市在统计年限内无任何外资利用情况的数据, 做了适当清洗

可以看出, 金融业企业注册量自2008年起, 一直都在以一个相对稳定的速度快速增长, 2012年以后基本趋于稳定; 从企业注册资金量上看, 随着企业注册量的快速增长, 企业注册资金也在同步增长, 于2012年到达顶峰, 在2012年后逐步趋于稳定。综合来看, 2008年国家实行“四万亿计划”后, 金融业企业注册量和注册资金量大幅增长。2012年以后, 经济刺激的效果逐渐减弱, 2013年受宏观经济下行影响, 金融业保持适度增长。为了防范金融风险, 2013年国务院出台了“金融十条”, 提出要切实解决经济运行中的结构性矛盾和资金分布不合理的问题, 同时要守住不发生区域性系统性金融风险。正是在一系列国家政策的调控下,

金融业的发展转为平稳。

从外资统计数据 and 变化趋势看, 2008年外资金金融企业注册数量为1719家, 2014年的数量为3755家, 注册资金量也从2008年的1020亿元上升到3070亿元, 外资金金融企业增长达到一倍多, 一直在稳定增长。说明国家逐步取消对外资金融机构的限制, 加强金融业对外开放政策取得初现成效。

4 注册数据与增加值的关联度分析

金融行业的产出受到多种条件的影响, 已有学者证实国家政策的制定、金融人才的多少、整体经济水平等因素都和金融业增加值的

变化相关^[20-22]。本文在收集整理金融业增加值数据和企业注册数据时发现,二者的增长趋势十分相似,具体数据如表2所示。鉴于此,本研究认为金融业企业注册数据与金融行业增加值存在一定关联关系,依据企业注册量、企业注册资金和金融业增加值数据,验证其关联程度大小,并选择灰色关联法进行测度。运用灰色关联法选取2008年到2014年的金融行业增加值数据为参考序列,金融业企业注册量数据和企业注册资金数据为比较序列。测度的结果显示,金融业的增加值与其企业注册量关联度 $r_{01}=0.698$,与其企业注册资金量关联度 $r_{02}=0.748$ 。 r 的值就是比较数列与参考数列间的关联程度, r 值越大则关联程度越大,反之,关

联程度越小。为进一步验证企业注册量和注册资金的影响程度,本文选择双对数回归模型分析并确定数值,方程如公式2:

$$\ln Y=c+\alpha\ln x_1+\beta\ln x_2 \quad (\text{公式2})$$

其中 Y 是金融业增加值, x_1 、 x_2 分别代表企业注册数量和注册资金量, α 、 β 分别为二者的弹性系数, c 为常数。将统计数据导入SPSS中进行回归分析,结果为: $\ln Y=0.299\ln x_1+0.652\ln x_2-2.166$ 。即金融业增加值每增加1%,企业注册数量就增加0.299%,企业注册资金随之增加0.652%。综上分析结果得出金融业增加值与企业注册资金的关联最显著,企业注册数量是影响金融业增加值的重要因素之一。

表2 金融业增加值和企业注册情况统计

年度	金融业增加值(单位:亿元)	企业注册数量(单位:家)	企业注册资金(单位:亿元)
2008	1.83E+04	1.00E+05	4.72E+05
2009	2.18E+04	1.37E+05	6.11E+05
2010	2.57E+04	1.52E+05	6.69E+05
2011	3.07E+04	1.63E+05	7.98E+05
2012	3.52E+04	1.73E+05	1.31E+06
2013	4.12E+04	1.74E+05	1.29E+06
2014	4.67E+04	1.72E+05	1.29E+06

数据来源:作者根据相关统计资料整理

5 结论

本文以统一社会信用数据中的企业注册数据为基础,以国内144个地区的金融行业为分析对象,运用时空统计学和灰色关联法,对2008年到2014年期间国内金融行业的规模、特征和趋势进行统计分析,初步得出如

下结论:

在行业发展规模上,金融业在国内所有行业中的数量规模偏小,但可支配资金数额庞大,能够较好的支持实体经济发展。在一定程度上,金融业受到政策举措的影响较为显著。

在空间分布上,金融行业基本形成了条块化格局,其中京津冀、长三角和珠三角地区

的金融行业发展最为迅速,集聚规模和潜力也最大。

金融行业外资企业规模虽小,但呈现稳定增长趋势;随着国家层面宏观政策的进一步放宽,预计外资企业将会有更大的发展空间。

结合数据分析结果和国家相关政策判断,国内金融行业度过了一段快速增长期,在国家宏观调控以及国际经济形势的影响下,金融业的发展已转为平稳状态。

利用全样本数据可以清晰地分析金融业的时空分布特征和发展趋势。后续研究可统计银行业、证券业和保险业的企业注册数据,以便分析金融业的行业构成特征。本研究仅对金融业注册数据及其增加值进行关联度分析,未来将进一步研究金融业注册数据和区域以及国家 GDP 增加值的关系。注册大数据是经济前向数据,目前对其挖掘和利用的研究还不足,未来将从经济理论、大数据模型等方面深入研究。

参考文献

- [1] 石影. 新时代金融业“营改增”面临的挑战与风险应对策略探讨[J]. 财会学习, 2018(25): 166.
- [2] 卓贤. 金融膨胀与中国 经济转型[EB/OL]. [2018-11-09]. <http://www.drc.gov.cn/xsyzcfx/20180611/4-4-2896327.htm>.
- [3] 项俊波. 经济新常态下我国现代保险业的发展[J]. 保险研究, 2015(2): 3-13.
- [4] 徐洪才. 我国金融业对外开放: 回顾及展望[J]. 金融发展研究, 2018(9): 1-5.
- [5] 上海自贸区推出金融开放 25 条举措[EB/OL]. [2018-11-09]. <https://finance.ifeng.com/a/20180623/163505640.shtml>.
- [6] 高万东, 吕鹰飞. 我国金融业发展水平测度及区域比较[J]. 东北师大学报(哲学社会科学版), 2013(2): 22-28.
- [7] Zouaoui H, Mazioud M, Ziedi N. A Semi-Parametric Panel Data Analysis on Financial Development-Economic Volatility Nexus in Developing Countries[J]. Social Science Electronic Publishing, 2018(8): 50-55.
- [8] Xiong L X, Qi S Z. Financial development and carbon emissions in chinese provinces: a spatial panel data analysis[J]. Singapore Economic Review, 2016, 62(2): 1-18.
- [9] 唐雨虹. 西藏金融发展的实证研究[D]. 北京: 对外经济贸易大学, 2015.
- [10] 纪玉俊, 周素娟. 我国金融业空间集聚的区域分异及其特征——基于非线性傅立叶单位根检验的分析[J]. 经济与管理, 2014, 28(1): 78-83.
- [11] Blazquez D, Domenech J. Big Data sources and methods for social and economic analyses[J]. Technological Forecasting and Social Change, 2017, 27(7): 99-113.
- [12] 杨英锐. 整合科学之美[EB/OL]. [2017-07-01] <http://news.sciencenet.cn/htmlnews/2017/7/383264.shtm>.
- [13] 李军利. 河南省县域经济空间信息系统的设计与实现[D]. 郑州: 河南理工大学, 2010.
- [14] 纪玉俊, 周素娟. 我国金融业空间集聚的区域分异及其特征——基于非线性傅立叶单位根检验的分析[J]. 经济与管理, 2014, 28(1): 78-83.
- [15] 刘军, 程中华, 李廉水. 中国制造业发展: 现状、困境与趋势[J]. 阅江学刊, 2015, 7(4): 15-21.
- [16] 李红, 王彦晓. 金融集聚、空间溢出与城市经济增长——基于中国 286 个城市空间面板杜宾模型的经验研究[J]. 国际金融研究, 2014(2): 89-96.
- [17] 茹乐峰, 苗长虹, 王海江. 我国中心城市金融集聚水平与空间格局研究[J]. 经济地理, 2014, 34(2):

- 58-66.
- [18] 周丽丽, 杨刚强, 江洪. 中国金融发展速度与经济增长可持续性——基于区域差异的视角 [J]. 中国软科学, 2014(2): 58-69.
- [19] 纪玉俊, 周素娟. 我国金融业空间集聚的区域分异及其特征——基于非线性傅立叶单位根检验的分析 [J]. 经济与管理, 2014, 28(1): 78-83.
- [20] 刘文革, 周文召, 仲深, 等. 金融发展中的政府干预、资本化进程与经济增长质量 [J]. 经济学家, 2014(3): 64-73.
- [21] 聂宇贤. 我国金融人才集聚对金融业发展贡献的实证研究 [D]. 重庆: 重庆理工大学, 2018.
- [22] 徐江. 金融业发展与实体经济增长关系: 基于河南省的实证分析 [J]. 生态经济, 2013(9): 151-154.