

基于共词分析的国内互联网金融领域研究热点分析

任国强, 王福珍, 潘永明

(天津理工大学 管理学院, 天津 300384)

摘要: 进入21世纪以来,互联网金融一直受到社会的广泛关注,也是金融学领域中一个研究热点。文章以中国学术文献网络出版总库(CNKI)中2002~2014年互联网金融相关的核心期刊为基础数据源,在确定国内互联网金融领域53个高频关键词的基础上,采用共词分析法及其可视化技术,利用书目共现分析系统(BICOMB)及SPSS统计工具对数据进行分析。建立在多维尺度分析结果上并参考因子分析及聚类分析结果可以得出,现阶段我国互联网金融研究有五大热点主题:互联网小微金融、支付方式与交易、金融改革创新、互联网金融征信和互联网金融风险与监管。

关键词: 共词分析法;互联网金融;因子分析;聚类分析;多维尺度分析

中图分类号: F83

文献标识码: A

文章编号: 1008-407X(2015)03-0074-06

Hotspot of Internet Finance Research Based on the Co-Word Analysis

REN Guoqiang, WANG Fuzhen, PAN Yongming

(School of Management, Tianjin University of Technology, Tianjin 300384, China)

Abstract: Since the 21st century, internet finance not only has received extensive attention from the whole society, but also been a hot research topic in the field of finance. In this paper we conduct an analysis of the data taking internet finance related core journals from CNKI database between 2002 and 2014 as the underlying data source, on the basis of determining the 53 high frequency keywords in domestic Internet financial field, using co-word analysis method and its visualization technology, and using BICOMB and SPSS. Based on the result of multidimensional scaling analysis and reference for the factor analysis and cluster analysis results, we can conclude that the present stage of China's internet financial research has five hotspots: Internet small micro-finance, payment and trading, financial reform and innovation, the internet financial reporting, internet financial risk and regulation.

Key words: co-word analysis method; Internet finance; factor analysis; cluster analysis; multidimensional scaling

一、引言

互联网金融是互联网与金融的结合,是借助互联网和移动通信技术实现资金融通、支付和信息中介功能的新兴金融模式^[1]。互联网金融模式打破了资本市

场长期垄断的格局,弱化传统银行、保险、证券、交易所等金融中介作用,达到接近无中介“瓦尔拉斯一般均衡”的状态,是既不同于商业银行间接融资,也不同于资本市场直接融资的第三种金融融资模式^[2]。

当前互联网金融已经迅速引起社会各界的密切关注,并冲击着现有的金融体系和服务。这一切使2013

收稿日期:2015-01-08; 修回日期:2015-05-19

基金项目:国家社会科学基金项目:“中小企业团体融资路径与机制研究”(11BGL023)

作者简介:任国强(1967-),男,河北文安人,教授,博士,主要从事产业政策与实践研究;王福珍(1993-),女,江西鹰潭人,天津理工大学管理学院硕士研究生,研究方向为产业政策与实践,(E-mail:WFZ1113@163.com);潘永明(1963-),男,江苏南通人,教授,主要从事公司金融与区域经济研究。

年被称为互联网金融元年,2014年3月召开的十二届全国政协二次会议上,国务院总理李克强在2014年政府工作报告中更提到,“促进互联网金融健康发展,建立金融监管协调机制”,说明互联网金融进入最高决策层的视野^{[3][4]}。国内外已有不少学者对该领域进行研究,相当一部分文献集中在互联网金融与传统金融间的关系上,然而对互联网金融领域进行全面文献计量研究的文献则十分稀缺。对我国互联网金融领域的研究热点进行分析,能为我国金融学领域相关学者进行专门而深入的研究提供重要参考。

二、研究方法 with 数据预处理

1. 研究方法

本文的研究方法是共词分析法。共词分析法最早在20世纪70年代由法国文献计量学家最早提出,1986年法国国家科学研究中心 CNRS 的 M. Callon、J. Law & A. Rip 出版的专著《Mapping the Dynamics of Science and Technology》成为共词分析方法的重要里程碑^[5]。共词分析法 (co-word analysis) 的思想来源于文献计量学的引文耦合与共被引概念,即当两个能够表达某一学科领域研究主题或研究动向的专业术语在同一文献中出现时,说明二者存在一定关联性,同时出现的次数与二者之间的关系紧密程度成正比^[6]。

随后有许多国内外学者以共词分析法对各个领域进行研究探索,该方法应用的广泛性已经得到验证^{[7][8][9]}。但是纵观国内外文献发现,目前在金融学领域进行文献计量分析的研究少之又少。考虑到共词

分析法对新兴领域研究的焦点具有很大的适用性,本文拟在共词分析法的基础上,对我国互联网金融领域的关键词进行多元统计分析,得出现阶段该领域研究热点。

2. 数据预处理

由于核心期刊在学科研究中具有代表性和广泛的学术影响性,且国内首篇关于互联网金融的核心期刊发表于2002年,故本文把检索范围限定为2002~2014年核心期刊,这有利于更好地选取互联网金融领域的代表性关键词。按上述条件以中国学术文献网络出版总库(CNKI)为检索源,以互联网金融为主题进行检索共得到440条记录,剔除7条重复登载、征文通知和没有标识关键词的文献,剩余333条记录,将其作为研究的样本文献库并进行高频关键词提取。

(1) 构建关键词共词矩阵

关键词是一篇论文的精华浓缩,能够展现文献的核心要义。本文利用中国医科大学研发的书目共现分析系统(BICOMB)抽取关键词,共得到2622个关键词,再统计关键词频数。由于国内关键词的拟定具有一定程度的主观性和不规范性,对关键词进行统计时采取了手动筛选和合并,例如,将“互联网金融”、“启示”、“美国”、“建议”等无益研究的高频关键词剔除,将关键词“P2P”并入“P2P网络借贷”等。然后对关键词按词频高低进行排序,综合考虑之后将频次阈值设为5,得到53个关键词作为高频关键词。随后通过该系统对53个高频关键词进行两两检索进而得到 53×53 的共词矩阵,见表1。由于篇幅限制只截取前9个关键词,后续内容亦作类似处理不再一一说明。

表1 共词矩阵

关键词	商业银行	第三方支付	P2P网络借贷	社交网络	移动支付	金融监管	金融创新	支付宝	余额宝
商业银行	40	0	4	0	0	1	0	1	2
第三方支付	0	37	3	11	6	2	0	10	1
P2P网络借贷	4	3	27	0	1	1	2	0	1
社交网络	0	11	0	22	9	0	0	2	0
移动支付	0	6	1	9	20	0	0	2	0
金融监管	1	2	1	0	0	19	4	2	1
金融创新	0	0	2	0	0	4	18	0	1
支付宝	1	10	0	2	2	2	0	18	1
余额宝	2	1	1	0	0	1	1	1	16

共词矩阵是一个对称矩阵,主对角线上的数值表示该关键词出现的频次,非主对角线上的数值则为两个相异关键词的共现频次。其中,“商业银行”出现40

次,初步说明目前对互联网金融领域的研究主要集中在商业银行与互联网金融之间的关系上。“社交网络”和“第三方支付”同时出现11次,初步说明二者有很大

的关联性。

(2) 相关矩阵与相异矩阵的转化

为了进行后续的聚类分析及多维尺度分析,消除

词频差距的影响,先使用 $ochiia$ 等值系数对互联网金融的共词矩阵进行相关矩阵的转化,构造出的相关矩阵如表 2 所示。

表 2 相关矩阵

关键词	商业银行	第三方支付	P2P 网络借贷	社交网络	移动支付	金融监管	金融创新	支付宝
商业银行	1	0	0.122	0	0	0.036	0	0.037
第三方支付	0	1	0.095	0.386	0.221	0.075	0	0.387
P2P 网络借贷	0.122	0.095	1	0	0.043	0.044	0.091	0
社交网络	0	0.386	0	1	0.429	0	0	0.101
移动支付	0	0.221	0.043	0.429	1	0	0	0.105
金融监管	0.036	0.075	0.044	0	0	1	0.216	0.108
金融创新	0	0	0.091	0	0	0.216	1	0
支付宝	0.037	0.387	0	0.101	0.105	0.108	0	1

表 2 中的数值表示对应两个关键词之间的距离,数值越大说明对应关键词之间距离越近、相似度越高。对角线上的数据表示该关键词自身的相关程度,经转化计算后均为 1。由于该矩阵中数值 0 过多,为减少统计误差且便于后续数据分析,用 1 与相关矩阵各数值相减得到表示两词间相异程度的相异矩阵,由于篇幅所限,此处不列出相异矩阵。在相关矩阵和相异矩阵的基础上,我们对数据进行多元统计分析,以便揭示互联网金融研究的热点。

三、多元统计分析

多元统计分析是对若干相关随机变量观测值的分析。大部分的共词分析研究都使用多种统计分析方法进行的研究。本文使用共词分析中常用的三种多元统计方法,即因子分析(factor analysis)、聚类分析(cluster analysis)和多维尺度分析(multidimensional scaling)进行具体研究。

1. 因子分析(factor analysis)

因子分析是对多变量的数据进行整理和简化,在保证数据信息较为完整的情况下,对高维变量空间进行降维处理的一种多元统计分析方法^[10]。其目的是把多个观测变量转换为几个不相关的综合指标,使得这些综合指标能够反映原数据的主要信息,然后根据因子成分矩阵对研究对象进行分类。由于旋转之后的成分矩阵中第 12 类与 13 类中各关键词因子载荷系数都低于 0.3,无因子参与分析,且旋转前各因子比旋转后具有更好的解释能力,故本文不采取旋转。

由 SPSS18.0 软件进行因子分析得到总方差解释表,共抽取出 14 个因子,这 14 个因子的累计方差解释

贡献率达 80.891%,即将 53 个关键词分为 14 类便可揭示互联网金融领域 80.891% 的研究热点。经观测后发现,前 6 个因子解释的方差比例较高,分别达到 24.676%、9.682%、6.975%、6.267%、5.269%、4.386% 累计方差解释贡献率 57.254%,表明这是近年来互联网金融研究中最重要 6 个领域。

遵照因子负载超过 0.3 的因子才能被提取的原则,对因子分析结果进行分析并根据关键词的属性及含义对 14 个因子进行命名,可以确定国内互联网金融领域热点分布,见表 3。从表 3 可以看出,因子分析结果中的关键词分类情况有下列特点:

(1) 大部分关键词都能归至相应的因子。然而有 2 个关键词由于载荷系数小于 0.3,故无法参与分类。它们是“监管协调”、“民间借贷”。这些关键词的关联程度比较小,一方面说明有些关键词虽然具有较高频次,但其所代表的研究主题范围较广,在因子的相关度分析中未形成鲜明特色;另一方面则说明学者对这些关键词的研究没达成统一共识,或者是其代表的主题较新,还没有很好地同其他研究方向相结合。

(2) 一些关键词因子载荷系数较低。由于关键词之间在概念上具有一定包含性,有些关键词可以被细划为更小的领域,因而在因子分析中以不同的方式参与分类。例如,“监管”与“金融监管”,“风险管控”与“风险控制”、“风险管理”等,因此“监管”(0.341)和“风险管控”(0.501)具有较小因子载荷。

(3) 有 3 个关键词跨区分布。其中,关键词“网络支付”同时出现在第 1 个因子和第 3 个因子当中,且二者因子载荷系数均处于 0.4~0.7 之间,体现出这两个因子即“支付方式与交易”、“电子商务与支付”间具有一定的关联性。

表3 研究热点

热点分类	关键词构成
支付方式与交易	(1)第三方支付 0.949,(2)社交网络 0.907,(3)移动支付 0.788,(4)支付宝 0.853,(5)基金公司 0.707,(6)腾讯 0.715,(7)网络支付 0.682,(8)支付结算 0.781,(9)人民银行 0.607,(10)筹融资 0.782,(11)风险管控 0.501,(12)手机银行 0.727,(13)金融支付 0.747,(14)金融信息 0.736,(15)交易规模 0.913,(16)电子商务平台 0.733,(17)客户体验 0.627
借贷与交易风险	(1)商业银行 0.578,(2)P2P网络借贷 0.549,(3)大数据 0.526,(4)小微企业 0.481,(5)风险控制 0.467,(6)交易成本 0.486,(7)融资 0.796,(8)征信系统—0.659
电子商务与支付 金融创新	(1)金融模式 0.372,(2)支付清算 0.402,(3)电子商务 0.559,(4)互联网经济 0.496,(5)网络支付 0.408
征信与监管	(1)金融创新 0.514,(2)金融改革 0.402,(3)金融监管 0.416,(4)金融机构—0.642
互联网普惠金融	(1)金融监管 0.602,(2)资本市场 0.321,(3)征信体系 0.387,(4)传统金融 0.361
金融风险	(1)互联网 0.523,(2)普惠金融 0.408
银行创新	(1)金融脱媒 0.580,(2)风险管理 0.389
小微企业筹融资	(1)影子银行 0.512,(2)民营银行 0.512,(3)利率市场化 0.478
信用信息	(1)融资 0.796,(2)众筹 0.339,(3)小微企业 0.404
监管	(1)信用信息 0.400
余额宝与金融机构	(1)监管 0.341
风险	(1)余额宝 0.368,(2)金融机构 0.378
资源配置	(1)风险 0.308
	(1)资源配置 0.512

注:表中数值为各关键词在相应类别中的因子载荷系数。

(4)有2个关键词的因子载荷系数为负,同各自的因子形成负相关关系。包括第2个因子(借贷与交易风险)中的“征信系统”(因子载荷系数为-0.659),第4个因子(金融创新)中的“金融机构”(-0.642),观察关键词“征信系统”和第2个因子“借贷与交易风险”可以知道,该关键词与各自的因子之间的关联度很低。

2. 聚类分析(cluster analysis)

聚类分析是根据事物“物以类聚”的思想将相似度较大的变量聚类的一种多元统计方法。聚类分析的方法有很多种,基于数据适用情况本文采用在实际操作中应用最为广泛的系统聚类方法进行聚类分析。本文在相异矩阵的基础上进行互联网金融高频关键词聚类分析。

根据聚类结果并结合因子分析的分类结果及实际情况,本文将53个高频关键词划分成以下7类:

第一类:商业银行、金融创新、互联网、余额宝、金融脱媒、风险、普惠金融、监管、影子银行、金融机构、金融模式、民营银行、利率市场化、风险管控、传统金融、征信系统、金融改革、信用信息、民间借贷。第二类:支付宝、移动支付、第三方支付、社交网络、支付结算、筹融资、手机银行、交易规模、客户体验。第三类:P2P网络借贷、小微企业、众筹、融资。第四类:金融监管、基

金公司、腾讯、人民银行、支付清算、备付金、电子商务、金融支付、互联网经济。第五类:大数据、风险控制、资本市场、资源配置、征信体系、信息不对称、交易成本。第六类:风险管理、监管协调。第七类:金融信息、电子商务平台、网络支付。

3. 多维尺度分析

聚类分析的结果将高频关键词分为7类,但是无法反映出各关键词之间的关联程度,而多维尺度分析(multidimensional scaling,简称MDS)通过低维空间中的点结构(configuration of points)展现的相似性来研究客体,可以更加直观地挖掘出数据的潜在结构关系。与因子分析和聚类分析不同,该方法对数据的要求较低,并且由于其确定了所有变量的位置,对之前的因子分析和聚类分析有一定的矫正调整作用,能更形象地展示分析结果^[11]。

在相异矩阵的基础上利用SPSS18.0得到多维尺度图,如图1所示。图中有相似性高的点被划分成一类;位置越为居中则该关键词与其他关键词关联度越大,在该领域中的地位越核心。

现根据图1拟将关键词分为3类:

(1)筹融资风险与监管。该类关键词主要分布在第二象限,且共有4个关键词分别散落于坐标轴上限

右端及坐标轴左限下端。主要包括:筹融资、风险、监管、小微企业、风险管理、风险管控、风险控制、P2P 网络借贷、监管协调等。根据该类的关键词方位观察可知,其主要分布于第二象限中,分布较为分散,于原点远处聚集成类,中心方位不易确定。

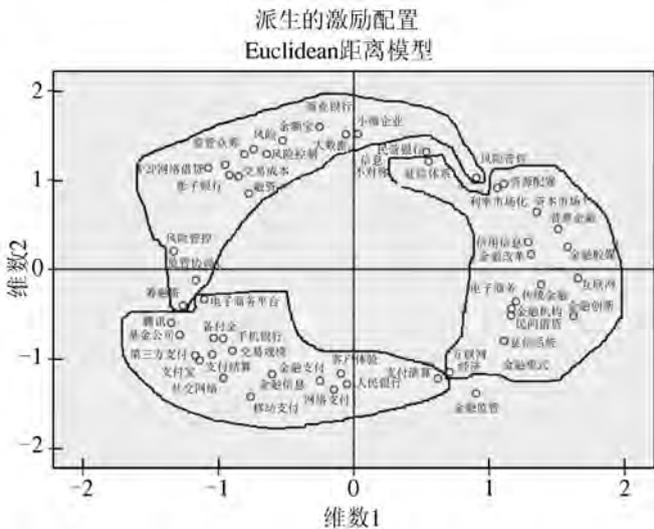


图 1 多维尺度图

(2)支付方式与交易。该类关键词主要分布在第三象限,且有一个关键词位于坐标轴下限右端。主要包括:第三方支付、支付宝、支付结算、金融支付、移动支付、网络支付、支付清算、备付金、手机银行等。观察该类关键词方位可知,其距离两条中心轴较远且分布较为分散,中心方位亦不易确定。

(3)金融创新与信息征集。该类关键词主要分布在坐标轴的右限上下两侧。主要包括:金融改革、金融创新、信用信息、征信系统、征信体系、利率市场化、普惠金融、金融脱媒、互联网经济等。观察该类关键词方位可知,其分布于中心横轴上下两端,于距离原点较远处聚集成类。其中,“金融改革”、“传统金融”和“互联网”等关键词位于主轴附近,表明这些关键词居于该类的中心地位。从内容上看,尽管以金融创新特性为主,但是出现了“民间借贷”“金融脱媒”等与第一类有内涵交叉的关键词。

全面观测多维尺度图可看出,关键词“民营银行”和“金融监管”没有归类。考量词义后将二者与相邻的类团合并有失偏颇,鉴于二维平面的视觉缺陷并参考因子分析与聚类分析结果决定将“民营银行”并入金融创新与信息征集类,“金融监管”则被划入支付方式与交易类。

四、国内互联网金融研究热点透视

互联网金融在我国是个新的研究领域,国内引入时间较短,尚不能对其脉络及流派进行十分精准的把握。但综合上述的因子分析、聚类分析和多维尺度分析仍可得出我国互联网金融领域的研究热点分类。由于因子分析在计算因子得分时采用的是最小二乘法,而该方法在分析过程中有可能失效,故可能存在一定的偏差。而聚类分析虽能定量地反映出各关键词间的亲疏关系,但是由于该方法通过词间距离的传递性将词间距离最短的关键词进行聚集,而聚类过程中又不能有效地将该关系取出,导致一些相关性不大的关键词形成聚集。通过前文三种多元统计方法的热点分类结果我们可以发现,因子分析和聚类分析各自分类过细,而多维尺度分析分类则略广,考虑到后者与另两种方法相比的优越性,建立在多维尺度分析结果上并参考其他两种分类方法可以得出我国互联网金融领域的研究热点主要集中在以下五个方面:

1. 互联网小微金融

该热点主要研究互联网小微金融。小微金融是银行对小微企业提供的金融服务,互联网小微金融以P2P和众筹为代表,是互联网与小微金融的结合。互联网小微金融业务正在填补目前小微金融需求的巨大缺口,为传统融资模式中处于不利地位的小微企业提供广阔的发展空间,有利于解决小微企业个人贷款难等问题^[12]。表征这一类热点的关键词包括小微企业、筹融资、P2P网络借贷、众筹、融资、余额宝、民间借贷等。

2. 支付方式与交易

该热点主要研究互联网金融业态下相对传统金融下支付方式的革新与多元化。互联网金融模式下的支付方式以移动支付为基础,通过移动通信设备、利用无线通信技术可以转移货币价值以清偿债权债务关系^[13]。截至2014年6月,我国互联网支付用户总数达到2.92亿^[14]。随着互联网支付的更加普及,由互联网金融的支付方式为代表的引起的支付革命对传统金融体系是强有力的冲击,并将对整体宏观金融业态产生深远的影响。表征这一类热点的关键词包括第三方支付、支付宝、支付结算、金融支付、移动支付、支付清算、备付金、基金公司、交易规模等。

3. 金融改革创新

该热点主要研究如何对传统金融进行改造,提高

资本市场运行效率,改变我国长期的“金融压抑”现象。《2014 第二季度中国货币政策执行报告》指出,要深化金融机构改革,通过增加供给和竞争改善金融服务^[15]。互联网金融作为一种新型金融模式提供了一个更有效率且低成本的新平台,填补了我国传统金融市场的金融缺口,优化了经济主体间的资源配置,从而推进经济增长。表征这一类热点的关键词包括金融改革、金融创新、金融脱媒、利率市场化、普惠金融、民营银行、传统金融等。

4. 互联网金融征信

该热点主要研究如何完善我国互联网金融行业征信体系建设。互联网金融与征信体系互相推动、互相影响,征信系统通过网贷平台实现信息共享从而促进互联网金融业发展,互联网金融业则通过自身平台发布的借贷信息完善征信体系。建立完善的信用体系关乎互联网金融发展的稳定性与长久性,其重要性不容忽视。表征这一类热点的关键词包括征信体系、征信系统、信用信息、金融信息和信息不对称。

5. 互联网金融风险与监管

该热点主要研究互联网金融模式带来的风险及其监管问题。互联网金融通过优化资源配置等对整体经济产生利好作用的同时也带来了新的风险,因此针对互联网机构的监管必不可少。2013年4月,众贷网和城乡贷的相继破产给监管机构敲响了警钟。有鉴于此,监管机构应该明确互联网企业的性质,以各互联网金融企业的性质和特点采取不同的监管策略,实现经济良性健康发展^[16]。表征这一类热点的关键词包括风险、风险管控、风险管理、风险控制、监管、金融监管、监管协调等。

五、结 论

进入21世纪后,国内外学者对互联网金融的研究日益增加,然而以共词方法对其进行热点剖析和审视的基本没有。有鉴于此,本文以互联网金融为主题对2002~2014年的核心期刊进行检索,通过书目共现分析系统建立共词矩阵,并利用SPSS18.0对其进行多元统计分析,得到国内互联网金融研究的研究热点为互联网小微金融、支付方式与交易、金融改革创新、互联网金融征信和互联网金融风险与监管五个方面。由于国内文献期刊的关键词选取存在很大主观性,对关键词的整合及数据处理也有一定偏差,低频关键词的舍弃和文献发表的时滞性等造成的关键词缺失在一定

程度上降低了研究热点的全面性和预见性,不利于对我国互联网金融领域热点的整体把握,这些因素使本文存在一定的局限性,可为以后相关领域学者的深入研究提供新的思路与视角。

参考文献:

- [1] 中国人民银行金融稳定分析小组. 中国金融稳定报告[R]. 北京:中国金融出版社,2014.
- [2] 谢平,邹传伟. 互联网金融模式研究[J]. 金融研究,2012,(12):11-22.
- [3] 莫易娴. 互联网时代金融业的发展格局[J]. 财经科学,2014,(4):1-10.
- [4] 皮天雷,赵铁. 互联网金融:范畴、革新与展望[J]. 财经科学,2014,(6):22-30.
- [5] CALLON M, LAW J, RIP A. Mapping the Dynamics of Science and Technology[M]. London:Macmillan Press, 1986. 102.
- [6] 张勤,徐绪松. 共词分析法与可视化技术的结合:揭示国外知识管理研究结构[J]. 管理工程学报,2008,22(04):30-35.
- [7] QIN H. Knowledge discovery through co-word analysis[J]. Library Trends,1999,48(1):133-159.
- [8] LAW J, BAUIN S, COURTIAL J P, et al. Policy and the mapping of scientific change:a co-word analysis of research into environmental acidification[J]. Scientometrics,1988,14(3-4):251-264.
- [9] LAW J, WHITTAKER J. Mapping acidification research: A test of the co-word method[J]. Scientometrics,1992,23(3):417-461.
- [10] 陈希镇,林俊涛. 用多元统计方法分析浙江省各地区的经济结构[J]. 数理统计与管理,2010,29(6):1043-1051.
- [11] 张勤,马费成. 国外知识管理研究范式——以共词分析为方法[J]. 管理科学学报,2007,12(6):65-75.
- [12] 孙晨辉,李富有. 民间金融与正规金融的比较优势与均衡发展——基于 Logistic 模型的研究[J]. 大连理工大学学报(社会科学版),2014,35(1):81-85.
- [13] 帅青红. 电子支付与结算[M]. 大连:东北财经大学出版社,2011. 56.
- [14] 中国互联网络信息中心(CNNIC). 2014年中国互联网络发展状况统计报告[EB/OL]. <http://www.cnnic.net.cn/hlwfzyj/hlwxyzbg/hlwjtjbg/201407/t20140721-47437.htm>,2015-05-19.
- [15] 中国人民银行货币政策分析小组. 2014年第二季度中国货币政策执行报告[R]. 北京:中国金融出版社,2014.
- [16] 李炳,赵阳. 互联网金融对宏观经济的影响[J]. 财经科学,2014,(8):21-28.