

# 我国CFPS应用研究——基于CiteSpace知识图谱的可视化分析

谢 芳

江西理工大学 江西南昌 330100

**摘 要：**基于2015—2025年中国知网收录的1777篇中国家庭追踪调查（CFPS）应用文献，本文采用文献计量与知识图谱方法，运用CiteSpace从时空分布、研究主体、知识聚类与主题演化四个维度进行分析。结果表明，CFPS研究自2017年起进入快速增长期，学术网络逐步形成；研究主题以家庭微观行为为核心，涵盖数字经济、收入分配、人力资本与农村发展等方向。研究前沿经历“方法建构—机制拓展—政策响应”三阶段演化，与国家战略高度契合，为理解中国社会经济变迁提供了新的数据支撑与研究视角。

**关键词：**中国家庭追踪调查；知识图谱；CiteSpace；研究热点

## 引言

中国家庭追踪调查（China Family Panel Studies, CFPS）由北京大学于2010年启动，是目前覆盖面最广、追踪时间最长的家庭纵向调查项目之一<sup>[1]</sup>。凭借城乡全覆盖与多维结构的数据优势，CFPS已成为研究中国社会经济变迁的重要基础数据库，为探讨家庭经济行为、社会流动与政策效果提供了可靠的微观证据<sup>[2]</sup>。随着数据积累与研究方法的不断进步，CFPS的应用研究迅速拓展，成为社会科学领域的重要知识来源。现有成果虽已涉及经济、教育、社会等多个方面，但整体性的知识结构分析仍然有限，尤其在研究主题的时间演化、学术网络关系与研究前沿识别上缺乏系统梳理。知识图谱方法为弥补这一不足提供有效工具，CiteSpace等软件能够揭示文献间的共现关系与聚类特征，从可视化角度呈现研究热点与演化趋势。基于此，本文以2015—2025年中国知网收录的1777篇CFPS应用研究文献为样本，运用CiteSpace从时空分布、研究主体、热点聚类与前沿演化四个维度系统分析该领域的发展特征与学术格局。

## 一、数据来源与研究方法

### 1. 数据来源与研究方法

本研究聚焦自CFPS数据库创立以来相关期刊，数据来源于中国知网。检索使用“CFPS”或“中国家庭追踪调查”作为主题词，在“篇关摘”设置“家庭追踪调查”、“中国家庭追踪调查”或“CFPS”提高检索精

度，时间区间2015年1月1日至2025年8月31日。限定支持基金为国家自然科学基金和国家社会科学基金，剔除与主题无关文献，获得有效文献1777篇。样本文献覆盖多个学科领域，较为全面反映CFPS应用研究知识体系与学术格局。

### 2. 研究方法

为揭示CFPS应用研究热点与知识结构，本文采用文献计量与知识图谱分析相结合的研究方法。以CiteSpace 6.1.R6为主要工具，对样本文献进行关键词共现、聚类与突现检测，通过可视化方式刻画研究主题的演化路径与结构特征。该方法能够在文献层面呈现学术主题的关联模式与发展趋势，为识别研究重点与领域特征提供量化依据<sup>[3]</sup>。参数设置方面，时间切片设为1年，节点类型选取“关键词”，阈值参数Top N取50，采用Pathfinder算法与Pruning sliced networks对网络进行优化，以确保结构清晰、聚类稳定。聚类质量通过模块化系数Q与平均轮廓值S进行评估， $Q>0.3$ 表示网络划分显著， $S>0.7$ 说明聚类内部一致性较高。研究设计包括三个环节。首先，通过关键词共现分析识别主要研究主题及其相互关系；其次，运用聚类分析提炼知识群落与核心方向；最后，结合突现检测识别近年学术关注度提升的主题，以刻画研究热点的集中与扩散趋势，有助于系统呈现CFPS研究的知识格局及演化特征。

## 二、我国CFPS应用研究基本情况与进展

### 1. 文献的时间分布特征

2015—2025年我国CFPS应用研究发文量总体呈增长趋势。2015年仅27篇，至2024年增至280篇，增长显著。以2017年为界，前期处于起步阶段，2018年后进入

**作者简介：**谢芳（2000—），女，汉族，江西赣州，硕士研究生，江西理工大学，主要从事家庭金融研究。

快速增长期，研究热度持续上升。2025年数据因统计期未满全年而阶段性回落，若以全年维度评估，其研究热度仍具备较强韧性。

## 2. 文献的空间分布特征

(1) 作者分布情况。CFPS研究逐步形成了以高校科研人员为核心的学术共同体。近年，高产作者主要集中在经济学、社会学与相关交叉学科领域，尤其是周广肃、张永奇、周云波等学者，其研究成果对该领域影响深远，研究主题涵盖数字金融、收入分配、家庭行为与社会流动等热点问题。这些高产作者不仅在数字经济、社会保障与贫困研究中占据主导地位，通过不断推出跨领域创新研究，推动CFPS数据应用的多元化发展。CFPS的作者网络呈现出典型的“核心学者—合作团队—边缘研究者”多层结构。核心学者主导该领域的研究方向和学术讨论，而边缘学者则在某些细分领域中贡献了新的视角或数据分析。研究力量呈现出集聚与分化并存的局面，学者们的研究主题趋于集中，但跨领域和跨机构的合作仍相对有限。这种现象不仅反映了学术资源的集中化趋势，也揭示了学科间合作和知识整合的潜力。值得注意，在政策、教育与社会问题的交叉领域，仍存在较大的跨学科合作空白，例如，如何将社会学的视角与经济学的模型相结合，以便更好地解释社会变迁中的家庭行为。

(2) 机构分布情况。CFPS研究已形成以高校为主体的学术格局，经济学与社会学领域的顶尖高校主导研究，南开大学、山东大学、湘潭大学等高校为核心力量。综合性大学与地方高校在人力资本、教育公平与农村发展等领域也有所贡献，形成了“核心集中、外围扩展”的结构。CFPS研究的主力集中于顶尖高校，但跨区域与跨学科合作仍存在障碍。高校间的合作大多局限于地理邻近，缺乏跨学科、跨区域的深度协作。尽管CFPS数据丰富，为跨学科研究提供了基础，但目前仍由经济学与社会学主导，教育学、心理学等学科的融合较少。未来，推动数据资源共享和跨学科合作平台建设，将有助于提升研究深度和政策影响力。

(3) 高被引文献分析。高被引文献反映了研究领域的核心主题与学术影响力，选取2015—2025年间被引频次排名前15的CFPS应用研究文献，其内容分布可较好地揭示该领域的研究焦点与演化趋势。从主题分布看，超过半数文献集中于数字金融、互联网使用与家庭消费、创业和储蓄等议题，表明数字经济背景下的家庭行为研究已成为主流方向。研究内容呈现由现象验证向机制挖掘的转向，早期关注数字经济的影响方向，近年则侧重

作用路径与中介效应分析，反映对微观机制的持续深化。从方法演进看，早期研究主要采用OLS与Probit等传统计量模型，近年来逐渐引入断点回归、双重差分及机器学习方法，显示出方法体系的多元化与实证识别的强化趋势。从时间脉络与学科视角看，2012—2016年文献多聚焦社会保障与家庭金融，2017年后研究主题转向数字经济、共同富裕与多维贫困，议题演进紧随国家战略与技术变革方向。同时，部分文献融入社会学与心理学视角，体现出跨学科融合趋势。核心作者与机构分布进一步印证研究的集群化特征，张勋、周广肃等学者的高被引成果主要集中于数字经济与家庭行为领域，且多依托南开大学经济学院等核心机构产出，形成“核心作者—核心机构—核心议题”的联动格局，巩固了该领域的知识中心与影响力。

## 三、我国CFPS应用研究热点的可视化分析

### 1. 关键词共现图谱分析

关键词共现分析有助于揭示研究主题的核心结构与关联特征。基于CiteSpace的统计结果，CFPS研究领域的关键词网络密度为0.007，模块化系数 $Q=0.8352$ ，加权平均轮廓值 $S=0.9543$ 。依据聚类有效性标准，可见聚类结构显著且内部一致性高，说明该领域研究主题边界清晰、知识体系稳定。

共同富裕、互联网使用、数字金融、乡村振兴、家庭消费等构成核心节点，呈现多中心、辐射型结构，反映出CFPS研究的主题多元与学科交叉特征。结合节点联系与聚类关系，可将研究热点概括为三大核心议题及一个跨学科拓展方向。一是政策导向下的贫富差距机制研究。以共同富裕、收入不平等、多维贫困为核心，聚焦收入分配、政策响应与代际传递机制。代表性研究如新农保储蓄效应与社会网络借贷分析，揭示了家庭层面政策效应的微观逻辑<sup>[4]</sup>。二是数字经济背景下的家庭经济行为研究。以数字金融、互联网使用、数字鸿沟为核心，探讨数字化进程对家庭创业与消费升级的影响，显示出数字技术在重塑家庭资源配置与金融决策中的作用。三是“三农”问题的家庭微观视角研究。以乡村振兴、土地流转、非农就业为关键词，分析农村家庭在发展政策下的行为响应与配置逻辑。此外，研究逐渐拓展至幸福感、社会资本、心理健康等社会与心理维度，呈现经济学与社会学的交叉融合趋势。CFPS研究正由单一经济视角向多维融合演进，从现象描述转向机制识别，构建出以家庭微观行为为核心、涵盖经济、社会与数字转型的多维知识体系。

## 2. 关键词聚类图谱分析

利用CiteSpace对关键词进行聚类分析,可将主题关联度较高的词汇归为一类,从而揭示研究领域的知识结构与内在逻辑。CFPS研究的关键词聚类共形成16个主题群,分别包括幸福感、数字经济、收入分配、消费结构、认知能力、劳动参与、家庭金融、倾向得分匹配、非农就业、多维贫困、人力资本、个人所得税、精准扶贫、中介效应、收入不平等与农户。整合相近主题后,结合高频与高中心性关键词特征,可归纳出六个核心知识群。

(1) 数字经济与家庭金融。“数字经济与家庭金融”是CFPS研究中最活跃的主题之一,核心问题在于数字技术如何改变家庭金融约束与财富积累机制。相关研究表明,数字普惠金融扩展在一定程度上缓解信贷约束,促进家庭创业和投资<sup>[5]</sup>。这类研究揭示数字金融通过降低信息不对称、交易成本,提升资源配置效率的作用路径。不同数据口径和群体背景下的研究结论存在分歧。部分学者发现,数字金融的收益主要集中于高教育、城市和高收入群体,可能放大了“数字鸿沟”<sup>[6]</sup>。

(2) 收入分配与共同富裕。该主题聚焦收入差距、多维贫困和代际流动等核心议题。多数研究基于CFPS数据发现,收入分配的不平等兼具结构性与机会性特征,教育、职业与地区差异是主要影响因素<sup>[7]</sup>。社会保障、税收与转移支付政策在缓解不平等中发挥了重要作用,但减贫效应仍具有区域差异。与此同时,“共同富裕”的研究正在从政策口号转向可量化的经济学分析。学者尝试从可行能力、收入结构和社会流动性维度构建测度框架,但在指标体系、时间跨度与因果识别上仍缺统一标准<sup>[8]</sup>。

(3) 家庭消费与经济行为。“家庭消费与经济行为”以家庭消费决策为核心,关注收入结构、支付方式与信贷可得性对消费模式与福利的影响。已有研究发现,移动支付和数字信贷的发展显著推动消费升级和服务型消费扩张<sup>[9]</sup>,尤其在农村地区表现出更高的边际效应<sup>[6]</sup>。但消费扩张的福利效果仍存在争议,一方面,消费升级提升了生活质量;另一方面,消费信贷扩张可能引发过度借贷与债务风险。

(4) 人力资本与劳动参与。该主题聚焦教育投资、认知与非认知能力对劳动市场绩效及社会流动的影响。大量实证研究表明,非认知能力在解释劳动收入与职业流动中具有显著作用,且其边际效应在农村劳动力群体中更为突出<sup>[10]</sup>。此外,数字技能已成为新的劳动市场回报要素,然而,非认知能力测量缺乏跨期一致性,其与认知能力、教育环境间存在双向因果关系<sup>[11]</sup>。

(5) 研究方法与时区分析。这一主题反映出CFPS研究方法论的持续演进。倾向得分匹配、双重差分、工具变量与中介效应模型成为主流实证工具,显著提高了因果识别的可靠性<sup>[12]</sup>。近年来,机器学习与因果推断的结合被引入CFPS研究中,用于高维变量选择与非线性机制探测<sup>[13]</sup>。然而,部分研究仍存在工具变量选择不稳健、平行趋势假设未验证等问题。

(6) 农村发展与农户行为。以乡村振兴为背景,该主题聚焦土地流转、要素禀赋与非农就业对农业效率与家庭收入的影响。研究普遍发现,适度规模经营与农业效率正相关,土地流转的禀赋效应显著<sup>[14]</sup>。数字乡村建设通过改善信息环境与市场接入,提升了农户生产率和家庭韧性<sup>[15]</sup>。但现有研究多集中于经济因素,对村级治理与社会网络等社会资本变量考虑不足。

总体来看,CFPS研究的聚类结构呈现以“家庭微观行为”为枢纽、连接“方法—行为—政策”的三维互动格局。研究方法的创新强化了机制识别,微观证据反向支撑政策评估与社会治理,主题演进则体现了经济与社会议题的融合趋势。CFPS数据的纵向性与多维性,使其成为贯通理论研究与实践政策的重要平台,也为多学科综合研究提供了坚实基础。

## 3. 关键词突现图与时区图分析

对CFPS研究领域的关键词进行突现与时区分析,有助于揭示研究热点的历史演进与主题变迁。结合分析结果,并参考侯海燕等提出的“突现强度—持续时间”判定方法,研究演化可大致分为三个阶段。整体上,2015—2025年间CFPS研究呈现出“基础传承—衍生拓展—前沿创新”的渐进格局,与知识聚类结果相呼应。

(1) 2015—2018年:奠基阶段。突现关键词主要包括消费、倾向得分匹配等,反映出研究早期对家庭经济行为与实证方法的集中探索。学者以消费为切入点构建家庭经济行为分析框架,同时引入倾向得分匹配等因果推断工具,为微观机制研究奠定方法基础。该阶段以数据验证与方法建构为主,开启了CFPS研究的实证化路径。

(2) 2019—2022年:深化阶段。这一时期的突现词如多维贫困、精准扶贫、贫困脆弱性等频次显著上升,研究内容与国家脱贫攻坚和乡村振兴战略紧密相关。CFPS数据广泛应用于贫困治理、教育投资与社会政策评估,研究重心逐渐由经济行为转向民生与社会公平议题。同时,人力资、认知能力等关键词的出现,显示学界关注点从结果变量拓展到能力形成与代际传递,研究视角更加多维。

(3) 2023年至今: 拓展阶段。近年来, 共同富裕、数字经济、数字普惠金融等关键词突现频繁, 研究焦点进一步转向数字化背景下的收入分配与社会公平。学者从数字赋能视角探讨技术变革对家庭行为与财富分配的重塑机制, 体现出CFPS研究在数字社会经济转型中的前沿性与政策关联性。

CFPS研究主题的演化体现出明显的层次递进与结构特征, 早期聚焦方法建构与理论验证, 中期转向机制深化与社会议题拓展, 近年则形成政策导向下的创新融合。研究热点与国家战略形成同频共振, 显示出强烈的现实回应性与学术开放性。时间脉络上, CFPS研究已实现从经济行为到社会治理、从静态描述到动态机制的跃迁, 逐步构建起贯通“理论—数据—政策”的知识网络, 为理解中国家庭行为与社会变迁提供了新的实证视角与研究范式。

### 结论

本文基于2015—2025年中国知网收录的1777篇CFPS应用研究文献, 运用CiteSpace知识图谱方法, 从时空分布、研究主体、知识聚类与主题演化等维度系统梳理我国CFPS应用研究的整体格局与发展趋势。研究发现, 过去十年间, CFPS研究已由零散探索迈向体系化发展, 其学术结构、研究主题与方法范式均呈现出持续深化与多维融合的特征。

在研究格局方面, 发文规模稳步增长, 2017年成为领域发展的关键拐点。以南开大学、山东大学、湘潭大学等为核心节点的学术网络逐步形成, 研究力量由个体分散转向团队协作, 显示出数据共享与学术合作的显著增强。研究主题方面, CFPS已形成以家庭微观行为为中心的知识体系, 主要涵盖数字经济与家庭金融、收入分配与共同富裕、人力资本与劳动参与、农村发展与农户行为等方向。主题间呈现明显的交叉与融合, 推动了经济学、社会学与政策研究的协同发展。在研究前沿方面, 关键词突现与时区演化表明, CFPS研究经历了由方法建构到机制拓展、再到政策响应的三阶段跃迁。早期聚焦实证方法与数据验证, 中期深化至贫困治理与教育公平, 近年则转向数字经济与共同富裕等新兴议题。这一演化路径反映出研究方向与国家战略的高契合度, 体现了社会科学研究服务现实决策的学术功能。

CFPS数据在连接微观家庭行为与宏观政策评估方面展现出独特优势。其城乡覆盖、纵向追踪与多维结构, 为收入分配、社会流动与数字化转型研究提供了坚实的

实证基础。未来研究应进一步强化跨学科合作与方法创新, 深化微观机制与政策效果的交互识别, 探索家庭行为研究与社会治理的有机融合, 以推动形成具有中国特色的社会经济研究体系与数据驱动型政策分析范式。

### 参考文献

- [1] 孙妍, 王赫, 潘修明, 等. 大型追踪调查样本流失规律及模式——以中国家庭追踪调查为例[J]. 调研世界, 2025, (03): 63-76.
- [2] 谢宇, 胡婧炜, 张春泥. 中国家庭追踪调查: 理念与实践[J]. 社会, 2014, 34(02): 1-32.
- [3] 陈悦, 陈超美, 刘则渊, 等. CiteSpace知识图谱的方法论功能[J]. 科学学研究, 2015, 33(02): 242-253.
- [4] 马光荣, 周广肃. 新型农村养老保险对家庭储蓄的影响: 基于CFPS数据的研究[J]. 经济研究, 2014, 49(11): 116-129.
- [5] 周广肃, 丁相元, 张维昊. 最低工资标准、居民收入不平等与共同富裕——基于CFPS面板数据的分析[J]. 经济问题探索, 2023, (10): 31-47.
- [6] 张勋, 万广华, 张佳佳, 等. 数字经济、普惠金融与包容性增长[J]. 经济研究, 2019, 54(08): 71-86.
- [7] 李燕, 周云波. 数字金融发展与家庭消费升级——基于多源微观数据的分析[J]. 财贸经济, 2023, (9): 112-128.
- [8] 刘润芳. 家庭资产与代际收入传递研究——基于CFPS数据[J]. 经济科学, 2022, (3): 54-67.
- [9] 王一鸣, 刘世锦. 共同富裕的经济学逻辑与实现路径[J]. 中国社会科学, 2023, (10): 15-33.
- [10] 王春超, 张承莎. 非认知能力与劳动参与——基于CFPS数据的分析[J]. 劳动经济研究, 2021, (2): 65-78.
- [11] 刘梦瑶, 张永奇. 数字技能、人力资本与劳动市场回报[J]. 中国劳动经济, 2022, (5): 90-104.
- [12] 郑春荣. 双重差分法在社会科学研究中的应用综述[J]. 统计研究, 2020, 37(6): 3-15.
- [13] 王震, 陈悦. 社会科学可视化研究的演进与趋势[J]. 情报理论与实践, 2023, 46(5): 47-58.
- [14] 钱龙, 洪名勇. 非农就业、土地流转与农业生产效率变化——基于CFPS的实证分析[J]. 中国农业经济, 2016, (10): 45-60.
- [15] 李春梅, 赵立. 数字乡村建设与农户生产效率提升研究——基于CFPS数据[J]. 农业技术经济, 2024, (2): 89-105.