

区块链激励机制对开源公共品供给业态的影响研究

——基于知识生产的视角

摘要：在知识生产向 Mode 2 演进的背景下，针对传统双轨激励体系在开源公共品供给中的制度错配，本文基于公共品理论探讨区块链代币化激励作为算法契约的治理效能。研究采用配对抽样法，构建了涵盖 244 个区块链底层项目与 730 个传统开源项目的非平衡面板数据集。实证发现：在项目产出层面，代币激励显著增强了项目的生存韧性，使其在扩散效率、粘性效率及开发强度上表现出稳健的制度优势；在微观行为层面，该机制通过所有权绑定实现激励相容，显著提升了开发者留存率与外部协作参与；在治理结构层面，代币化有效降低了把关人集中度，在提升评审多样性的同时缩短了交付周期，实现了去中心化治理与复杂知识生产的深度适配。本研究论证了代币激励能有效缓解集体行动困境，实现开发者目标与项目长效供给的深度锁定，为数字基础设施的持续治理提供了实证支撑。

关键词：知识生产、公共品理论、开源软件、激励机制

第一章 引言

在知识经济的早期阶段，知识生产遵循着严格的科层制与学科分工逻辑。这一阶段被定义为“知识生产模式 1”（Mode 1），其核心是建立在传统的线性创新模型之上，将知识产出严格划分为象牙塔式的基础研究与封闭的企业应用研究（Gibbons et al., 1994; Hessels & van Lente, 2008）。在这一范式下，大学负责生产具有非排他性的基础理论，而企业则通过专利墙和商业秘密将知识内化为私有资产，以保障商业投资的回报（Mazzucato, 2013）。这种模式下的知识生产被限制在特定的物理与组织边界内部，形成了明显的学科孤岛与组织隔阂（Becher & Trowler, 2001; Klein, 2021）。进入 20 世纪后期，随着数字化工具对信息搜索与协作成本的极度压缩以及复杂应用场景的出现（Baldwin & von Hippel, 2011; Goldfarb & Tucker, 2019），知识生产方式产生了深刻的范式跃迁，即向“知识生产模式 2”（Mode 2）的演进。Mode 2 标志着知识生产从封闭的学科导向转向了复杂应用情境导向，生产过程不再局限于特定的科研机构（Gibbons et al., 1994），而是嵌入到了更广泛的、甚至是分布式的协作网络之中（Rikap & Lundvall, 2021）。这种变革催生了基于公地策略的对等生产，使得协作颗粒度从机构下沉至个体（Benkler, 2006）。

在这一演进过程中，开源软件（Open Source Software）逐渐从极客群体的兴趣产物演变为支撑全球运行的数字化基础设施（Eghbal, 2016），其本质上是一个全球性的社会化协同知识生产平台（Turek, 2023），汇聚了高密度的跨国界高级人力资本（Pardo et al., 2022; Liu et al., 2025）。作为数字时代 Mode 2 知识生产的最典型载体，开源代码呈现出纯粹的数字公共品特质：一方面，其边际复制成本为零，具有极强的非竞争性；另一方面，基于 Copyleft、MIT 等许可协议，它主动放弃了排他性收益，允许全球开发者自由复用与再生产。

在开源软件这类分布式协作（Mode 2）的知识生产情境下，在传统上，应对知识公共品生产难题的双轨激励体系，即以财政拨款与学术声誉为核心的公共科学激励和以知识产权为核心的市场专利激励，显现出系统性的适应性难题（Dasgupta & David, 1994; Eghbal, 2020）。

首先，以财政拨款与学术声誉为核心的公共科学激励逻辑与开源知识生产的分布式协作形态脱节。传统的科研资助体系高度依赖于主权边界内的机构-学者身份识别与优先权补偿。然而，开源生态中的贡献者呈现出高度的匿名化、跨国界及非雇佣特质，这导致维系数字基础设施运转的大量碎片化、非职业化贡献难以被现有的学术评价体系与财政资助体系所识别（Hussain et al., 2024）。Turek（2023）指出，在复杂的社会化协作中，传统的作者署名制等名誉激励机制已无

法精准适配颗粒化的贡献，导致人力资本在公共科学领域的投入面临价值被低估的困境，产生了大量的隐形劳动。

其次，以产权保护与垄断租金为核心的市场激励逻辑与开源契约存在天然冲突。根据产权理论，企业通过排他性权利（专利）获取租金以回收研发成本（Hart & Moore, 1990）。但在开源知识生产范式下，开发者基于开放协议主动放弃了排他性收益。在这种语境下，传统的市场逻辑介入往往表现为智力垄断（Intellectual Monopoly）对数字公地的侵蚀（Rikap, 2023）。这种基于租金榨取的逻辑不仅破坏了社区的互惠契约，还引发了严重的敲竹杠风险，使得独立的个体开发者因担心其专用性人力资本投资被商业意志俘获，从而产生系统性的事前投入不足（Wen et al., 2023）。

双轨激励体系的失效在微观层面引发了激励机制的扭曲。早期的信号理论曾认为，开发者可以通过贡献代码积累声誉以换取未来的工资溢价（Lerner & Tirole, 2002）。然而，在项目进入基础设施化的维护阶段后，生产活动转向了高压且低可见度的底层修补。这种隐形劳动难以产生有效的声誉信号，导致核心维护者极易产生职业疲劳（Eghbal, 2020）。

上述激励机制失灵最终外化为全球开源公共品供给的脆弱性。在缺乏稳定激励机制的环境下，社会化协作表现出极高的不协调性。最新的实证研究显示，GitHub 平台的开源项目生命周期呈现出显著的爆发式启动与断裂式衰亡特征：一方面，开源项目以每年两位数的增长速度不断增加；另一方面，超过 80% 的项目在创建后的一年内迅速陷入更新停滞（Coelho et al., 2020）。Arakelyan et al. (2024) 的数据进一步证明，由于缺乏有效的补偿机制，核心人力资本的流失已成为导致开源公共品供给侧坍塌的首要原因。

面对上述开源公共品供给中的激励制度空缺，区块链技术驱动的代币化（Tokenization）不再仅仅被视为一种金融工具，而可以被重新界定为一种面向分布式知识生产的可编程激励契约（Voshmgir, 2019; Catalini & Gans, 2020）。其核心逻辑在于，通过算法化的制度安排，在缺乏中心化权威与资助的环境下重建人力资本的激励相容机制，从而缓解开源项目在知识生产中的结构性困境。

代币化机制对前述激励失灵的补全作用，主要体现在两个维度的激励制度重构：剩余索取权的分配与承诺机制。

针对前文提到的隐形劳动难以衡量及搭便车问题，代币化通过构建一套微观的产权系统，将原本抽象的知识贡献量化为可验证、可交易的数字资产凭证。这种安排实质上是对 Alchian & Demsetz (1972) 团队生产理论的新应用。代币作为项目生态的价值体现和治理凭证，与项目可用性和活跃度强绑定。它打破了必须依靠中心化监督者持有剩余索取权的局限，通过代币形态和算法契约将项目的剩

余索取权直接赋予分布式的贡献者（Bayern, 2021; Reineke & Katila, 2025）。这种贡献绑定所有权收益的契约设计，不仅消解了搭便车动机，有效对冲了因产权模糊引发的敲竹杠风险，实现了开发者从单纯贡献者向合伙人的角色改变。

针对开源项目高夭折率背后的动态不一致性难题，即开发者在早期由于热情承诺投入、但在项目维护与再研究阶段因机会成本增加而选择退出，代币化机制引入了锁定与解锁协议。这种机制充当了 Kydland & Prescott (1977) 所定义的承诺机制：通过算法契约硬性约束代币的释放节奏，项目方能够自动筛选出对项目长期价值具有信心的核心人力资本，并有可能排除追求短期套利的投机者（Fos & Holderness, 2023）。这种算法化的跨期约束，将开发者的长效利益与公共品的持续存续深度锁定，显著降低了由于开发者流失导致的供给侧风险，为开源公共品的可持续供给提供了稳健的激励机制（Catalini & Gans, 2020）。

综上所述，代币化激励通过在分布式情境下重塑生产关系，补全了传统激励体系在 Mode 2 知识生产中的功能缺失，为破解开源公共品的供给危机提供了一套基于算法契约的潜在解决方案。

在前述开源公共品供给危机与代币化激励机制补全的理论背景下，本研究旨在通过严谨的实证分析，验证代币激励机制在真实知识生产环境中的治理效能。本研究不仅关注宏观层面的产出绩效提升，更试图深入微观层面，刻画代币激励机制如何影响开发者的协作行为与社区权力结构，进而实现对传统开源治理困境的突破。具体而言，本研究将聚焦于以下两个核心问题：一是项目绩效（对应第四章），在剔除项目异质性干扰后，代币化激励机制是否显著提升了开源公共品的生存韧性 with 产出绩效；二是微观机理，代币化激励机制如何改变个体开发者的行为并影响整体的微观治理效能。

尽管已有理论对代币化激励机制进行了诸多前瞻性讨论，但既有研究大多集中于规范性分析或单一案例描述，缺乏基于大规模数据的实证检验。为科学回答上述问题，本研究不再局限于单一的案例描述，而是基于开源项目数据进行大规模的数据对比。在数据获取层面，本研究依托于全球最大的开源协作平台 GitHub，通过其公开接口（API）与归档数据库（如 GHTorrent 或 GH Archive），系统性地采集了实验组（区块链项目）与对照组（传统开源项目）两类项目的全维度数据，包括项目维度的产出数据与开发者维度的行为数据，涵盖了项目整体的生命周期。通过构建项目层面与开发者层面的数据集，本研究有助于客观评估代币化激励机制在提升知识公共品供给与优化治理结构方面的真实贡献。这不仅为解析数字时代社会化知识生产的内在规律提供了新的经验证据，也为未来数字基础设施的可持续治理与人力资本的制度化激励提供了参考。

第二章 文献综述

开源社区作为典型的分布式知识生产组织，其激励机制一直是学术界关注的焦点。本章将系统梳理从传统动机逻辑到代币化激励机制的理论演进，并识别现有研究在解析“Mode 2”范式下的研究缺口。

2.1. 激励的起点：内在动机、直接效用与身份认同

开源协作的早期成功，首先被学术界界定为一种文化与组织范式的胜利。Raymond（1999）在其经典著作《大教堂与集市》中，首次将开源社区比作一个去中心化的集市。他指出开源生产遵循的是礼物经济的学理逻辑（源自莫斯的《礼物》理论），即在物质极大丰富的环境中，个体的社会地位不取决于他占有了多少资源，而取决于他向社区捐赠了多少，从而获得直接的满足。Bergquist & Ljungberg（2001）进一步从组织学视角论证，开源生态中的赠予行为构建了一种强有力的社会互惠契约，是维持早期开源组织运行的核心动力。Benkler（2006）将这种模式理论化为“基于公地的对等生产（CBPP）”，在这一模式下，数字化工具使得个体可以绕过科层制，在去中心化的集市中进行大规模协同。这种初期的激励体系主要依赖于开发者的内在动机（Hars & Ou, 2002; Roberts et al., 2006）、极客文化的认同感（Hertel et al., 2003; Weber, 2004）以及解决个人技术问题的直接效用（Lakhani & von Hippel, 2003）。

尽管上述研究合力解释了开源生态初期的增长奇迹，但也共同揭示了其局限性。当项目由于直接效用和纯粹乐趣被建立起来并成为需要不断维护以及迭代的数字基础设施后，繁琐的维护工作往往无法再提供相同的效用满足。动机的边际效用递减，直接导致了前文所述的知识公共品供给萎靡。此时，单靠礼物经济带来的社会名望已不足以抵消沉重的维护成本。这便引出了后续 Lerner & Tirole（2002）关于职业声誉的讨论，即开发者开始需要将这种礼物转化为未来的经济回报。

2.2 激励的演进：声誉信号与互补品理论

2.2.1 声誉信号

当开源软件的生产规模跨越了个人兴趣的临界点后，参与动机开始呈现出显著的职业化特征。开发者通过无偿的代码贡献，实质上是在劳动力市场上进行信号发射行为。Lerner & Tirole（2002）首次系统性地将信号理论引入开源激励机制

研究。他们认为，由于软件代码是完全公开且可追溯的，它为开发者提供了一个极低成本的信息发射平台。因此，开源知识生产本质上即是一种可获得劳动力市场回报的长期投资行为，开发者通过在公开透明的社区中锻炼并展示解决复杂问题的能力，同时也能够有效降低与雇主之间的信息不对称。

Hann 等人（2004）通过对 Apache 基金会开发者的实证研究发现，拥有开源贡献经历的开发者在劳动力市场中能获得显著的职位晋升和薪资溢价，其在开源社区中的贡献度（尤其是获得核心提交权限）与劳动力市场的薪资水平正相关。Bitzer 等人（2007）进一步明确指出，开源知识生产的参与不仅是信号发放，更是一种干中学的人力资本积累过程。开发者在协作中获得的技能提升，构成了其影子工资的重要组成部分。Belenzon & Schankerman（2015）也发现开发者通过在具有高技术挑战的项目中贡献代码，能够向潜在雇主展示其难以被标准化考试测评的高级解决问题能力。这种信号在校内或初级工程师中表现得尤为强烈。

但 Lakhani & Wolf（2005）、Eghbal（2020）都认为信号理论能够解释外在激励的一个关键前提是信号的可见性，由此解释了为何开发者热衷于创建新项目（高可见度信号），但也导致了开发者不愿意承担枯燥的后期维护工作和持续的知识生产。Schweik & English（2012）通过对数千个开源项目的对比研究发现，绝大多数项目在度过爆发式增长后会陷入长期的沉寂。Ghapanchi & Aurum（2011）进一步验证了当项目进入成熟期，由于信号强度衰减，开发者的未来工资预期会显著下降。这种信号衰减现象正是导致开源项目在维护期面临人力资本流失的深层原因之一。

2.2.2 互补品逻辑与成本分摊

随着开源软件成为主流，企业的角色从消费者转变为开源知识生产者。这种转型基于企业战略的理性选择。Lerner & Tirole（2005）以及 Bonaccorsi & Rossi（2006）指出，企业资助开源、成为开源知识生产者的核心动机在于增加其自有互补品的价值。初期的研究焦点主要集中在软硬件互补，目的在于通过软件操作系统的传播以扩大硬件的市场份额。West（2003）提出了服务与软件互补的混合模型，即企业通过开源知识生产来扩大市场份额，同时通过提供定制化服务、技术支持或附加的专用插件，进一步获取利润。

除了有利于企业扩大市场份额以外，企业加入开源知识生产还能够降低自身研发成本，Henkel（2006）在对嵌入式 Linux 行业的研究中发现，企业会主动开放那些不具有竞争优势的基础模块，以吸引外部开发者共同维护，进而降低内部研发成本，但会严格封锁涉及核心竞争力的专有代码。Alexy et al.（2013）也指出，企业通过战略性地放弃部分知识产权，实质上是将其非核心模块的研发工

作外包给了全球开源社区，利用社区的人力资本来替代内部的研发投入。因此，企业在开源知识生产中的激励特点表现为选择性开源，这与传统社区的全量开放有本质区别。

这种基于企业战略的开源知识生产激励模式，也引发了开源知识生产中一定程度的激励失灵。Dahlander & Wallin (2006) 的研究发现，当企业过度干预开源项目时，会产生挤出效应，使得一部分原本基于内在动机参与的极客会感到自主性受损，从而撤出社区。另一方面，O'Mahony (2003) 指出，企业倾向于通过各种法律工具（如 CLA，贡献者许可协议）获取对社区代码的控制权，这让独立贡献者面临严重的资产专用性投资风险，使其担心自己的无偿贡献最终会沦为企业获取租金的免费投入。

2.3 代币化激励机制

代币化激励 (Tokenized Incentives) 被视为一种解决开源公共品供给困境的算法化制度方案。其核心逻辑在于通过对产权和权利束的重构，在分布式环境下重建激励相容机制。在制度经济学视角下，代币并非单纯的金融资产，而是一种可编程的权利束 (Voshmgir, 2023)。不同于传统货币，代币可以在同一载体上同时赋予持有者收益权、治理权和使用权。这种设计使得开发者持有的代币凭证不仅代表了过往贡献所获得的市场价值，更锚定了项目未来的治理权限 (Werbach, 2018)。

Cong & He (2019) 以及 Allen et al. (2020) 进一步认为，通过算法契约，代币赋予了持有者对项目的剩余控制权与剩余索取权。在传统模式中，独立开发者的资产专用性投资（针对特定项目的技能投入）极易在项目商业化过程中被企业通过敲竹杠行为剥夺。而代币化的产权结构确保了这种投入被固化为不可篡改的资产凭证。Malinova & Park (2023) 强调，这种由算法保障的剩余权益分配，实现了开发者从纯粹贡献者向合伙人的角色跃迁，从而在根本上解决了由于所有权缺失引发的激励缺乏与人力资本撤出。根据 Alchian & Demsetz (1972) 的团队生产理论，最优激励来源于剩余索取权与贡献的对等。代币化激励机制通过算法将个体的私人收益（代币的市场价值）与项目的整体生态价值（软件的易用性与采用率）直接挂钩。这种利益锁定效应使得开发者从外部志愿者转变为内部股东，从而实现了开发者个人目标与项目长效供给目标的激励相容 (Catalini & Gans, 2020)。

2.4 文献述评：为何比较两类项目的创新业态

通过对既有文献的系统梳理，本研究发现，尽管学术界对开源激励的演进过程（从礼物经济到信号理论，再到互补品战略）已有深刻洞察，但仍尚缺失对于

市场化激励开源知识生产的相关研究，这也构成了本研究开展对比实证的逻辑起点。具体到本研究所涉及的代币化激励机制，目前尚缺乏对代币化作为一种新型知识生产激励制度的绩效实证。现有关于代币化的研究，鲜有文献将其回归到知识生产模式的本质，去考察其是否真的改变了知识公共品的产出效率。通过将区块链项目（代币化激励）与传统高质量开源项目（非代币化激励）进行对比，以此观察传统开源项目表现出的高夭折率与人力资本撤出现象，是否在代币化项目中得到了根本性逆转，进而可以剥离出新的激励制度对知识生产产生的影响。

第三章 实证设计与数据说明

3.1 样本选择与对比设计

本研究旨在探究区块链开源项目与传统开源项目在知识生产业态上的差异。为了实现这一目标，本文构建了一个涵盖项目产出绩效（静态结果）与开发者微观协作行为（动态过程）的双层分析框架。在样本选择上，本文采取基于准实验设计的配对抽样法，构建了包含区块链底层基础设施项目（实验组）与传统高质量开源项目（对照组）的基准数据集。

3.1.1 实验组：区块链底层公链项目

本文实验组聚焦于区块链技术的底层基础设施——公链的开源代码仓库。在样本选择与覆盖率方面，本研究初始识别出 108 条 Layer 1 公链和 156 条 Layer 2 公链。经过对 GitHub 仓库可用性及代码开源协议的严格核实，最终确定 252 条具有可比性且数据完整的公链仓库作为实验组样本。其中，Layer 2 项目实现了全样本追踪，Layer 1 则涵盖了市场绝大部分主流及技术代表性项目。为缓解实证研究中常见的幸存者偏差问题，样本中不仅包含当前生态繁荣的活跃项目，还涵盖了处于停滞及已停止维护的项目。这一特征为后续开展生存分析提供了必要的保障。

3.1.2 对照组：传统开源项目与匹配策略

为了深度剥离宏观技术周期与 GitHub 平台用户规模增长对实证结果的潜在干扰，本文采用了严谨的时间队列匹配抽样策略。具体而言，研究以实验组公链仓库的建仓月份作为时间基准锚点，按照 1:3 的比例在同期创建的非区块链开源项目中进行随机匹配。

在构建抽样框时，考虑到公链项目通常具备较高的初始关注度与技术复杂度，

为了确保对照组具备对等的创新潜力和技术声誉,本文并非进行全样本的无差别抽样,而是利用 GitHub Search API 特意筛选出该月历史热度排名前 1000 的活跃仓库作为基础样本池。这种基于质量对齐的准实验设计,有效排除了海量低质量噪声仓库的干扰,确保了 756 个 对照组项目与实验组在社区认可度及潜在产出效率上具备较强的统计可比性。

3.2 数据来源与预处理

3.2.1 多源异构数据的抽取与内涵解析

GitHub 不仅是代码托管平台,更是一个高度数字化的社会化协作系统,其产生的数据流能够真实刻画开源创新的微观动态。本研究利用 Python 自动化脚本,通过 GitHub REST API 与 Search API 构建了一个涵盖“项目-事件-开发者”三维特征的异构数据集。

(1) 仓库级元数据 (Metadata) 抽取

在数据采集的第一阶段,本文利用 GitHub Search API 执行了自动化采样,旨在构建一个支撑后续深度追踪的高价值抽样框。为确保对照组项目具备与公链项目同等的技术产出潜力,本文在各匹配月份创建的仓库中,通过 API 接口筛选出该月历史热度排名前 1000 名的活跃仓库作为基准索引。通过此接口,本研究首先成功获取了 974 个目标仓库的仓库级元数据 (Metadata),包括仓库唯一标识、主要编程语言、创建时间、最终星标数 (Stars)、分叉数 (Forks) 以及项目状态 (如是否归档) 等。这一步骤实现了技术层面的质量对齐,即通过构建这一高质量抽样框,本文成功排除了 GitHub 平台上数以亿计的随机噪声仓库,确保了对照组样本均具备一定的社区关注度与潜在创新价值。这部分元数据构成了本文第四章项目层面产出分析的数据基础。

(2) 细粒度行为事件流 (Event Stream) 抓取

在锁定 974 个目标仓库构成的抽样框后,本文进一步利用 GitHub REST API,以仓库为单位,回溯并抓取了各项目自建仓起至观测终点 (2025 年) 的全量行为日志。不同于静态的元数据,这些细粒度的事件流能够动态还原公共品生产过程中的微观互动,能够动态地刻画开发者在开源生态中的互动过程。具体而言,研究重点抓取了以下三类核心行为维度,以揭示开源公共品生产过程中的动态特征:

i. 代码提交 (Commits)

代码提交是开源协作中最基础的知识生产行为。它代表了开发者对软件代码仓库进行的一次原子化修改,是 Github 公共知识生产中的流量与存量数据的基本提交单位。通过这一数据,通过不断的代码提交,散布于全球的开发者将零散

的知识碎片转化为连续的软件资产。

在数据抓取层面，本研究抓取了对应开源仓库的仓库名（repo）、代码提交唯一识别码（sha）、提交者账号（author_login）、作者姓名/邮箱（author_name/email）、提交日期（author_date）、提交的内容说明（message）。

利用上述字段，本研究可以将零散的提交记录串联起来，精准捕捉开发者在项目生命周期中的每一次活跃，并完整地重构出每一个项目的知识贡献事件流，还原出项目在微观层面上的知识生产脉络。基于此，进一步能够揭示开发者是如何通过持续的微贡献逐步构建起庞大的公共产品，从而观察区块链激励机制是否改变了公共品生产的强度，以及它如何影响底层技术资产的积累效率。

ii. 拉取请求（Pull Requests）

拉取请求是开发者请求将自己编写的代码合并到主项目中的行为，是开源社区处理复杂协作与集体决策的核心机制。它是一个完整的协作流程，涵盖了从发起申请、社区讨论到最终执行合并的全过程，代表了从个人知识生产向公共采纳转化的提案过程。通过 Pull Requests 流程，项目维护者得以在外部贡献正式并入公共知识资产库之前，对其进行评估与协调。

在数据抓取层面，本研究获取了拉取请求的所属仓库（repo）、编号（number）、发起者账号（author_login）、发起者身份属性（author_association）、创建时间（createdAt）、合并时间（mergedAt）、执行合并者账号（mergedBy）以及参与讨论人（participants）等核心字段。

利用上述字段，本研究可以完整地重构出每一个项目的协同治理事件流，可以精准测量社区的决策响应速度与代码交付周期，并识别识别出项目权力结构的把关人及其对外部创新的接纳程度。基于此，进一步能够揭示公共品生产过程中的协调成本与准入门槛。通过观察不同项目在治理效率上的差异，本文旨在探讨区块链代币激励是否通过优化治理规则，降低了大规模协作中的交易成本，从而提升了公共产品演进的整体效率。

iii. 代码评审（PR Reviews）

代码评审（Review）是在 Pull Requests 合并之前，由其他成员对代码质量进行的审核与反馈，是开源公共产品在正式发布前的最后一道质量防线。它是由社区成员对他人贡献进行审核、纠偏或批准的过程，体现了开源生态中特有的集体监督机制与去中心化质量控制。通过评审，社区得以确保公共知识资产的安全性、稳健性与一致性。

在数据抓取层面，本研究详细记录了评审所属仓库（repo）、对应的 PR 编号（pr_number）、评审人账号（reviewer_login）、评审结论状态（state）以及评审时间（createdAt）等关键字段。

利用上述字段，本研究可以量化测量项目内部的质量审查与评审互动事件流，可以计算出评审多样性，用以观察项目的技术评估权是广泛散布于社区，还是被少数精英所垄断。基于此，进一步能够揭示公共品生产中的权力分配与民主化程度。在分布式知识生产协作的叙事背景下，评审数据的透明度与多样性是验证其治理真实度的关键证据。通过这一分析，本文旨在衡量区块链激励机制是否强化了社区的监督强度，从而有效降低了公共产品演化过程中的寻租风险。

综上，通过对上述代码提交（Commits）、拉取请求（Pull Requests）、代码评审（PR Reviews）等行为数据的串联，本研究不仅能获取仓库层面的静态元数据，更能追踪开发者在项目生命周期中的动态行为轨迹

3.2.2 样本清洗规则

为了确保实证分析的高质量与严谨性，本文对从 GitHub 采集的原始数据进行了精细化清洗。由于 GitHub 生态中存在大量碎片化、非工程化的干扰数据，本研究制定了以下五项核心过滤规则，旨在将研究视角聚焦于具有实质性创新意义的软件项目：

（1）确保创新的原创性：排除派生项目

在 GitHub 生态中，Fork（派生）是指开发者将他人的代码仓库完整复制到自己账户下的行为。这一操作通常用于对原项目进行临时的代码修改、提交拉取请求（Pull Request）或作为个人知识生产的起点。Fork 仓库的本质是已有创新的副本，而非独立的创新源头。排除此类项目可确保本研究的观察对象均为具有独立创新意图、拥有完整生命周期的原生项目，从而真实反映区块链与传统开源在原创性产出上的差异。

（2）确保数据的动态有效性：排除归档项目

Archived（归档）是项目所有者对仓库设定的一种状态，标志着该项目正式停止维护。一旦项目被归档，它将变为只读模式，社区成员无法再提交代码、发起讨论或反馈问题。本研究涉及项目的生存分析与开发者留存研究，由于归档项目已丧失了持续演化的动力，无法提供观测区间内的动态交互数据，因此排除归档项目能确保样本在观测窗口内具备活跃更新的可能性，避免静态死项目对生存概率估计的偏误。

（3）确立软件工程基准：设定活跃度门槛

Stars（星标）是 GitHub 用户对仓库的点赞或收藏行为，是衡量开源项目受欢迎程度与社区认可度的核心指标。由于 GitHub 平台上存在数以亿计的僵尸仓库或个人练习项目，设定 $\text{Stars} > 5$ 的低限阈值，旨在过滤掉缺乏社区关注、不具备软件工程质量的个人习作。这一门槛确保了入选样本均已跨过了可被观测的

初始门槛，具备了进行社会化协作创新的基础规模。

（4）维护技术环境的可比性：编程语言对齐

区块链底层设施（如一层公链与二层网络）的开发具有较高的技术壁垒，通常采用 Rust、Go、C++ 等高性能语言或 Solidity 等专有合约语言。由于不同编程语言的社区文化、开发难度与协作模式存在显著差异，通过指定主流编程语言对齐，本研究排除了因技术栈差异导致的干扰因素，确保实验组与对照组在技术复杂性与开发者准入门槛上处于对等地位。

（5）保证研究对象的纯净度：内容类型校验

GitHub 不仅是代码托管平台，也常被用于发布学习笔记、教程、资源列表（如 Awesome List）或项目模板。这些仓库虽有高热度，但其本质是内容汇编而非知识生产。通过关键词检索与人工复核，本研究剔除了对照组中可能含有的区块链相关项目及非代码类项目，确保研究结论能够精准推导至软件创新产出这一核心范畴。

由于 GitHub 生态中存在海量的自动化工具操作与冗余记录，经过以上五项核心过滤规则后，为了确保实证结果能够真实反映开发者的创新业态，本文在构建面板数据前进行了数据清洗与质量控制：（1）识别并去除机器人账号。在现代开源协作中，大量重复性任务（如代码格式化、依赖包更新、自动化测试报告）由机器人账号自动完成。本文通过多重过滤机制识别机器人。首先，利用 GitHub API 自带的账号类型标签进行筛选；其次，通过关键词匹配（如账号后缀包含 [bot]、-helper 等）锁定潜在机器人；最后，针对高频次、模式化提交的账号进行人工抽样核实。（2）开发者身份对齐与去重。同一开发者在不同时期可能使用不同的 GitHub Login 账号或多个电子邮箱进行提交，这会导致对社区规模和活跃度的误判。本文采用了身份对齐算法。通过对 author_name 和 author_email 字段进行交叉映射，将具有相同姓名或邮箱的记录归集至同一自然人 ID 之下。

（3）跨仓库冗余事件剔除。在 GitHub 生态中，经常存在将一个上游仓库的代码批量同步至下游仓库的行为，这会产生大量的重复事件。本文利用代码提交的唯一识别码（sha）进行全局去重。在计算特定项目的知识增量时，确保每一个 sha 标识的知识贡献事件在统计中只被计入一次。（4）异常值处理与缩尾处理。开源指标（如 Stars、Forks、Commits）通常呈现长尾分布，极端异常值可能会扭曲回归系数。在进行指标计算后，本文对所有连续型变量进行了 1% 水平上的双侧缩尾处理。

经过上述严格的样本筛选与行为去噪处理，本研究最终构建了一个底层行为数据库。该数据库涵盖了实验组与对照组共 974 个目标仓库自建仓起至观测终点的全量交互记录，具体数据规模如下：共计 3121213 条开发者代码提交

（Commits）的有效提交记录，构成了研究知识生产强度的核心；共计 533069 条拉取请求（Pull Requests）的协作提案记录，是分析社区治理效率与开放度的基础；共计 151759 条代码评审（PR Reviews）的审核记录，为测度社区权力结构与流动性提供了详实证据。

3.2.3 面板数据构建

在完成原始数据的抽取与深度清洗后，本研究通过对代码提交（Commits）、拉取请求（Pull Requests）与代码评审（PR Reviews）三类动态事件流的串联，实现了从底层字段到行为序列的转化，构建了支撑全文实证分析的非平衡面板数据集。

本研究以项目一月份作为基本观测单元。通过对样本仓库全量行为数据的追踪，利用向量化计算方法将数百万行的事件记录按月度进行聚合。每一个项目都被重构为一个从建仓月份开始，到观测终点为止的连续月度时间序列。最终生成的面板数据集涵盖了 974 个仓库（含 244 个实验组与 730 个对照组）在多个维度上的月度演化指标。通过对开发者 ID 的跨月度追踪，面板数据成功捕捉到了开发者从进入项目到活跃、再到晋升核心层或流失的完整行为轨迹。

3.3 指标体系构建

为了全面刻画区块链公链项目与传统开源项目与创新产出、市场认可及生态演化上的本质差异，本研究构建了一个多维度的指标体系。该体系由仓库层面的宏观绩效指标（结果视图）与开发者层面的微观协作指标（过程视图）共同组成。

3.3.1 项目层面指标

项目层面指标反映了开源公共产品在技术资产密度与社区影响力方面的静态存量特征。由于两类项目在激励机制上存在本质差异，本研究通过构建比率型衍生指标，旨在消除名气效应造成的统计偏误。

表 3.1 项目层面指标定义及说明

指标名称	计算方式	意义解读
关注数	Stars	开发者对仓库点击“收藏”的累计次数，反映了项目在开源生态中被观测到的难易程度，可作为项目热度的核心代理变量。
派生数	Forks	开发者将仓库代码复制到个人空间以便进行修改或二次开发的次数。Forks 是强度更高的参与信号，每一个 Fork 都代表了一次潜在的创新尝试或技术采纳，体现了知识从被观察到再生产的转化潜力。
订阅数	Watchers	主动选择接收该仓库所有动态通知的用户总数。Watchers 代表了用户愿意承受信息干扰来保持对项目技术演进的同步。这反映

未解决问题数	Open_Issues	了开发者对于一个项目的深度关注与持续跟踪意愿。 仓库中尚未关闭的任务、缺陷报告、功能建议或讨论贴的总数。 该指标既代表了项目的维护压力，也反映了社区的反馈密度。大量的未解决问题数意味着社区正处于活跃的博弈与迭代中，体现更强的协作负荷与反馈循环。
仓库体量	Size	仓库所有文件占用磁盘空间的累计大小。该指标是项目资产规模的缩影，反映了项目在时间维度上的技术积累与历史积淀。体量越大往往意味着项目的技术积累更为雄厚，但相应的维护成本也越高。
扩散效率	$E_{diff} = \frac{Forks}{Stars+1}$	衡量从被动关注（Stars）向主动采纳/二次开发（Forks）的转化比例，它反映了知识溢出和新增开发的速度。若 E_{diff} 显著更高，说明有效地降低了外部开发者从观察者转变为应用者与生产者的转化门槛。
粘性效率	$E_{stick} = \frac{Watchers}{Stars+1}$	衡量用户从点赞关注（Stars）到持续关注动态（Watchers）的转化比例，它反映了社区的运营粘性。该指标反映了社区治理的利益相关者化程度。更高的 E_{stick} 说明了能够有效缓解开源社区中常见的注意力涣散现象，将流量固化为长期关注者。
开发强度	$I_{dev} = \frac{Open_Issues}{\sqrt{Stars+1}}$	规模归一化协作负荷，衡量在单位影响力下，开源项目社区处理创新诉求或反馈的强度。该指标体现了去中心化治理的反馈韧性。如果 I_{dev} 显著更高，说明其社区协作更活跃，存在更密集的创新反馈循环。

3.3.2 开发者行为指标

项目层面指标揭示了创新的产出结果，而开发者行为指标则旨在通过微观行为数据，解构开源公共品生产过程中的协作逻辑与权力结构。利用构建的非平衡面板数据，本研究从动员、协作治理和所有权三个维度刻画开发者行为特征。

（1）社区动员与留存指标

该维度衡量项目吸纳人力资本及维持社区活力的能力，反映了激励机制对开发者参与动机的持续影响。

表 3.2 社区动员与留存指标定义及说明

指标名称	计算方式	意义解读
活跃开发者数（每月）	当月产生贡献 (Commit/PR/Review)的去重用户数	衡量社区公共品供给的微观人力规模与瞬时协作动能，反映激励机制的动员范围。
次月留存率	t 月活跃且在 t+1 月持续活跃的用户比例	衡量社区的短期组织粘性，体现代币化激励对开发者行为的初步锁定效应。
季留存率	t 月活跃且在 t+3 月持续活跃的用户比例	衡量社区的中长期忠诚度与公共品生产的持续稳定性，反映制度环境的稳态能力。
外部贡献率	非核心成员提交的拉取请求(Pull Requests)比例	衡量社区的开放程度与边界渗透性，体现激励机制在克服公共品搭便车问题上的效能。

(2) 协作治理与效率指标

该维度揭示了项目内部的协调机制、审核门槛及决策反馈速度。

表 3.3 协作治理与效率指标定义及说明

指标名称	计算方式	意义解读
合并效率	当月已合并拉取请求（Pull Requests）数/总申请数	反映治理团队对外部贡献的接纳程度与审核门槛。
交付周期	拉取请求（Pull Requests）从创建到合并的时间间隔中位数	衡量协作响应速度及社区内部的决策链路长度。
评审多样性	$1 - \sum_{i=1}^n \left(\frac{Review_Count_i}{Total_Reviews} \right)^2$ n 代表当月参与代码评审的所有开发者总数， $Review_Count_i$ 代表开发者 i 在当月执行的评审次数。 $Total_Reviews$ 为项目当月总的评审活动次数	衡量代码审核权力的分散程度，值越高表示治理越民主。
把关人集中度	第一大拉取请求（Pull Requests）合并者执行合并的占比	衡量项目是否存在权力过于集中的治理倾向。

(3) 所有权与稳定性指标

该维度采用 6 个月滚动窗口计算，用以捕捉知识生产的深度特征与社区阶层的流动性。

表 3.3 知识所有权与稳定性指标定义及说明

指标名称	计算方式	意义解读
代码所有权集中度	基于 6 个月累计代码提交量的 HHI 指数	衡量代码所有权是否长期由少数核心开发者把持。
精英占比	$1 - (\text{贡献累计占比 } 50\% \text{ 代码的开发者数} / \text{月活跃开发者总数})$	识别支撑项目核心功能的实质性核心团队规模占比。
核心稳定性	t 月与 t-1 月窗口核心成员集的 Jaccard 相似度	衡量权力核心的更替频率，反映社区结构的稳定性。
阶层流动性	从非核心成员晋升为核心成员的比例	衡量项目的社会流动性，即普通参与者晋升的难易程度。

第四章：区块链开源项目的创新产出表现（项目层面）

本章旨在通过定量分析，评估区块链开源项目与传统高质量开源项目在宏观产出绩效上的异质性。首先，通过生存分析方法探讨两类项目的生命周期稳健性；其次，对比分析其在技术规模、市场认可及生态转化效率上的静态分布规律。本章研究结论将为理解代币化激励机制在公共品生产中的宏观绩效提供实证证据。

4.1 项目生存能力对比分析

为了衡量两类项目在动态演化过程中的持续创新动力，本文首先引入 Kaplan-Meier 生存分析法对项目的生命周期进行建模。在开源软件工程语境下，项目的“死亡”并非简单表现为代码库的注销，而往往体现为社区活跃度的彻底停滞。因此，本文将死亡事件严格定义为：项目出现连续 365 天无任何代码提交（Commits）更新的状态。这一标准旨在剔除因短期休假或版本迭代间隙导致的伪停滞，从而精准捕捉项目真正丧失维护动力的时刻。通过以建仓月份为起点，计算各样本至死亡事件发生或观测终点（右删失）的持续月数，本文构建了支撑生存分析的时间序列基础。

实证结果显示，拥有代币激励机制的区块链项目组在全生命周期内的生存表现显著优于对照组。如图 4.1 所示，实验组的生存曲线始终位于传统高质量开源项目组的上方，这意味着在同等观测时长下，区块链项目维持活跃开发的概率更高，表现出更强的生命韧性。特别是在项目建立的前 75 个月内，实验组的存活概率下降斜率显著平缓，而对照组在此阶段面临更剧烈的流失压力。此外，实验组在跨越初期生存危机后表现出更为明显的平台期特征，反映出区块链项目一旦建立起初始生态，往往能比传统项目形成更为稳固且持久的开发群体。

从公共品理论视角审视，上述生存溢价揭示了代币化作为一种内生激励机制，对缓解开源软件开发中集体行动困境的显著作用。传统开源模式常因缺乏直接经济回报而导致贡献者倾向于搭便车，使得公共产品的长期维护面临资源枯竭的风险。而区块链项目通过代币化手段将未来的生态价值与代币绑定并分配给贡献者，不仅提供了必要的资源补偿，更通过利益绑定机制将开发者转化为社区的共同所有者。这种经济与治理的双重约束为公共产品的早期演化提供了必要的韧性，使其在跨越死亡谷时展现出比传统模式更强的制度稳健性。

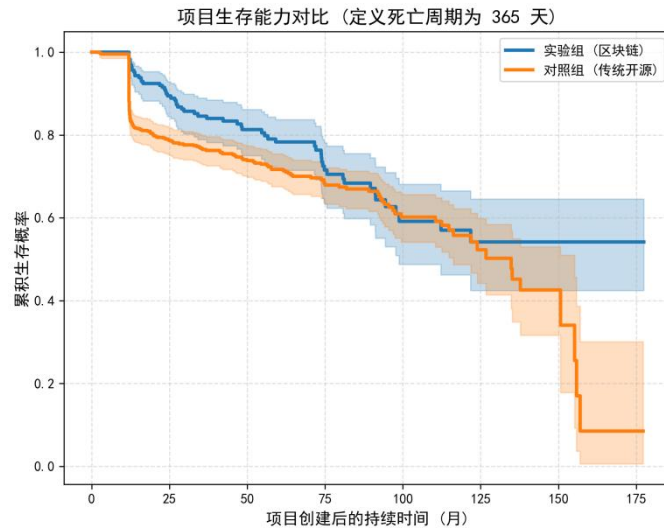


图 4.1 项目生存能力对比分析图

4.2 项目热度与转化效率分析

在证实了区块链项目具备更强的生存韧性后，本节进一步通过组合图对比两类项目在创新业态上的异质性。如图 4.2 所示，本文从关注度规模与扩散效率两个维度进行了对比分析。图 4.2 (a) 显示，实验组（区块链项目）的星标数（Stars）中位数显著高于对照组，且分布区间更为集中，呈现明显的关注下限更高的特征，而控制组（传统开源项目）的分布离散度更高。这有力地证明了实验组在吸引社区注意力方面并不比优秀的传统开源项目逊色。其关注下限更高的特征也反映了区块链项目特有的资产属性：由于代币的存在，投资者、矿工及生态参与者构成了天然的监测群体，使其在初创期即能动员起远超普通项目的社会化资本。

为了排除区块链项目的知识生产繁荣仅为投机泡沫的可能性，图 4.2 (b) 引入了归一化后的扩散效率指标（ E_{diff} ）。结果显示，在一定程度剔除了项目热度的干扰后，区块链项目的扩散效率呈现出对对照组的压倒性优势。在公共品理论视域下，分叉数（Forks）衡量了技术的可扩散性和二次生产。实证结果揭示，区块链项目每吸引一个单位的外部关注，能激发出更多的开发者进行代码派生与二次创作。这种从围观到构建的高转化率，证明了其创新并非仅靠热度驱动，而是通过高效的知识溢出和知识生产协作实现了实质性的知识产出。然而，这种由于资产属性带来的扩散效率是否能自发转化为持续的产出，仍需在控制关注度干扰后进行进一步的实证检验。

区块链项目与传统开源项目的关注数与扩散效率对比

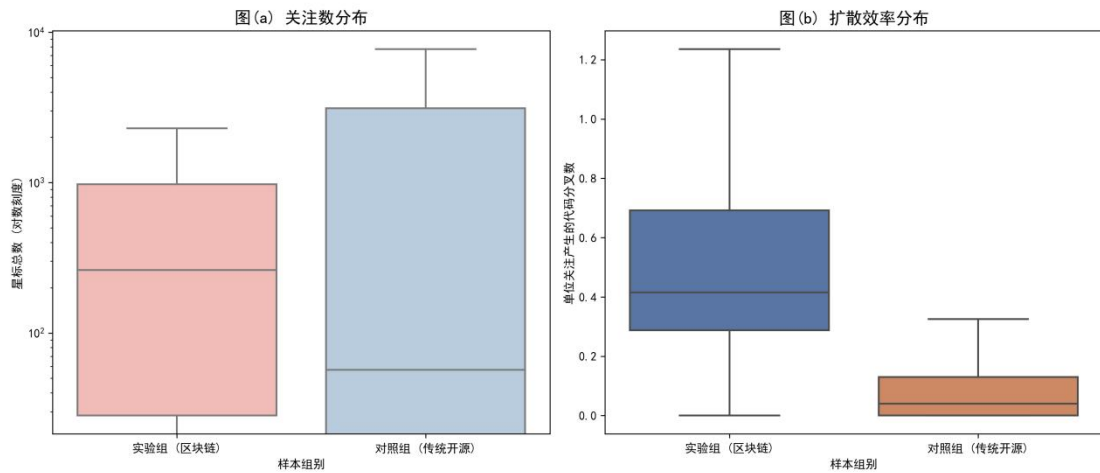


图 4.2 项目层面关注数与扩散效率对比图

4.3 项目热度与开发者粘性分析

针对代币激励可能引发的投机性关注问题，本文通过对比关注规模与粘性效率来识别社区的忠诚度结构。如图 4.3 所示，虽然实验组与对照组在关注规模（图 4.3a）上均展现出较高的关注规模，但两组在粘性效率（图 4.3b）上的表现却呈现出显著的异质化特征。实证发现，实验组单位关注转化为持续订阅（Watchers）的能力不仅在中位数上超过对照组，且其分布密度高度集中在相对稳定的高数值区间。相比之下，对照组的粘性效率表现出极高的离散性，大量样本在吸引关注后未能有效实现用户的深度留存。

这种高热度、高留存的共存状态，有力地回应了关于区块链项目仅存在投机式创新的假设。订阅行为不仅是简单的信息获取，更代表了开发者对项目技术演进进行长期跟踪的直接表征。较高的粘性效率证明，代币激励机制很有可能实现了对搭便车行为的有效过滤，通过利益锚定，区块链项目成功将广泛的、瞬时的外部关注固化为愿意长期关注与投入的参与者，构建起了一个高密度的核心利益共同体。这种社区结构的沉淀，不仅为公共产品的持续生产提供了稳定的人力资本支撑，也为后文 4.5 节中关于区块链机制对冲注意力稀释效应的回归发现提供了直观的分佈证据。

区块链项目与传统开源项目的关注数与粘性效率对比

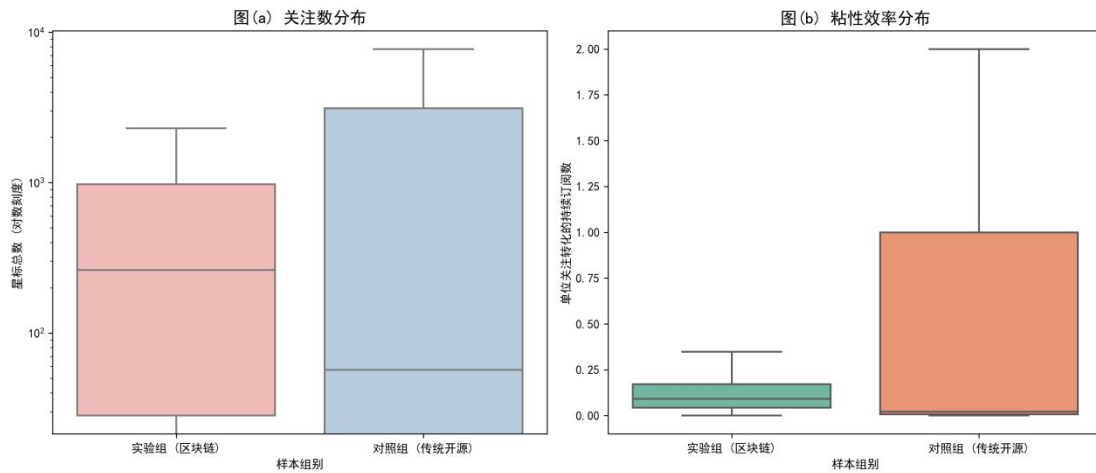


图 4.3 项目层面关注数与粘性效率对比图

4.4 项目复杂度与开发强度分析

在剖析了项目的扩散效率与粘性效率后，本节进一步探讨高技术复杂度的公链项目能否与其庞大的社会化协作需求相匹配，即考察技术复杂度与开发强度的状态。由图 4.4（a）的对比可见，实验组与对照组在技术资产规模上表现出显著的异质性。实验组（区块链项目）的仓库体量（Size）中位数显著高于对照组，且其实验组的下四分位数已高于对照组的中位数水平。这一分布特征直观地反映了区块链底层协议重型化的技术特征：由于涉及共识机制、加密算法及分布式账本等多个高门槛模块，区块链项目作为数字公共基础设施，具备相当高的技术密度和较大的技术资产规模。

令人关注的是图 4.4（b）所示的开发强度对比结果。在控制了热度规模的干扰后，实验组在单位关注下的 Issue 响应负荷依然显著领先于对照组。这意味着，即便在面临超大规模技术资产带来的维护压力下，区块链社区依然保持了远超传统项目的高频需求反馈与问题响应。这种高复杂度与强互动性的共存，有力地论证了区块链创新业态的治理韧性。从知识生产的角度审视，这种重运营、强互动的模式是典型的知识生产模式 2（Mode 2）的具体体现，即生产过程不仅是技术代码的堆砌，更是一个多方高频互动的动态演进体系。这种社会化协作范式揭示了代币激励在驱动治理效能方面的核心作用。代币激励不仅缓解了公共品维护中的资源匮乏，更构建了一套能够消化高复杂度需求的动态维护系统，实现了技术产出与社会化协作的深度耦合。

区块链项目与传统开源项目的项目复杂度与开发强度对比

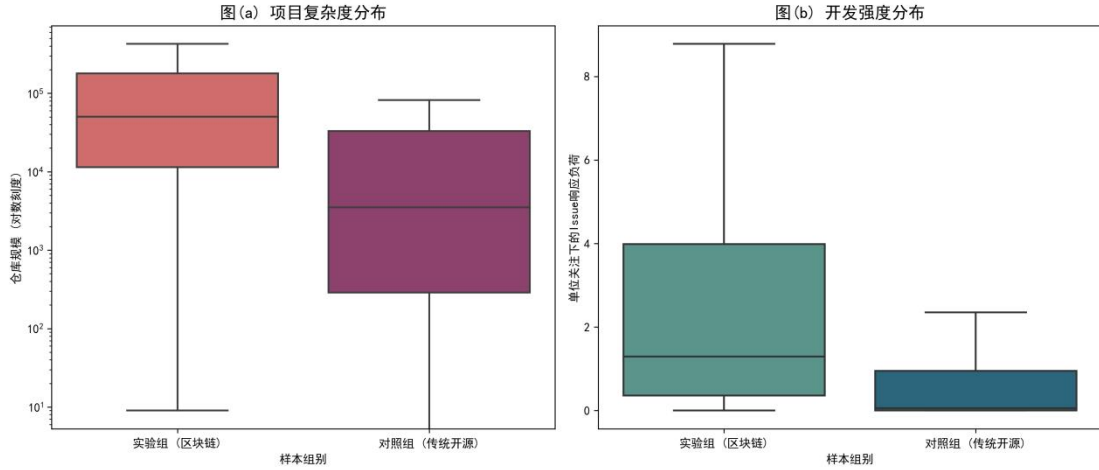


图 4.2 项目层面复杂度与开发强度对比图

本节前述生存分析与效率指标对比初步勾勒出区块链项目长效存续、高转化且有粘性、高复杂度但强互动性的知识生产特征。然而，这些直观差异是否仅由区块链项目的规模优势、创建年份或特定技术栈（如语言差异）所驱动？为剥离上述混淆因素，本章接下来将通过构建计量回归模型，在相对严苛的控制变量环境下，对区块链激励机制的促进效果进行进一步的实证检验。

4.5 项目层面回归检验

为了进一步验证前文观察到的区块链项目与传统开源项目在生存韧性、转化效率及治理强度上的异质性分布，本节构建了多元线性回归模型。通过引入一系列控制变量并固定编程与时间效应，本研究旨在剥离外部环境干扰，识别代币化激励机制对开源公共品产出绩效的净影响。

4.5.1 计量模型设定

为进一步考察区块链激励机制对开源项目创新转化效率的净效应，本研究构建了如下基准回归模型：

$$\ln(Efficiency_i + \epsilon) = \alpha + \beta \cdot is_blockchain_i + \sum_{j=1}^n \gamma_j \cdot Controls_{i,j} + \lambda_{lang} + \delta_{year} + \varepsilon_i$$

模型设定说明：

1. 因变量 ***Efficiency_i***：代表项目 *i* 的三个衍生效率指标，即扩散效率 (*E_{diff}*)、粘性效率 (*E_{stck}*) 及开发强度 (*I_{dev}*)。考虑到开源数据普遍存在的长尾分布特征，本研究对因变量执行了对数转换，并添加偏移量 $\epsilon = 0.0001$ 以处理零值。

2. 核心解释变量 (***is_blockchain_i***)：虚拟变量，当项目属于区块链项目类

别时取 1，否则取 0。其系数 β 反映了在同等外部条件下，激励机制对产出效率的百分比边际贡献。

3. **固定效应项：**模型加入了编程语言固定效应 (λ_{lang}) 以控制不同技术栈（如 Rust、Go、Solidity 等）在开发者基数及工具链成熟度上的异质性，以及项目创建年份固定效应 (δ_{year}) 以吸收不同出生队列的宏观环境干扰。

4. **估算方法：**采用普通最小二乘法 (OLS) 进行参数估计，并引入 HC3 异方差稳健标准误，以确保在数据分布不均时的推断稳健性。

4.5.2 控制变量选取与理论逻辑

为了缓解遗漏变量导致的偏误，并精准剥离出区块链激励机制对比传统开源模式的差异，模型从以下三个维度引入了关键控制变量：

(1) 基础规模与生命周期

项目知名度 (ln Stars)：控制项目的社区吸引力与社会资本存量。由于知名度越高的项目天然拥有更高的协作基数，控制该变量有助于识别在同等关注水平下的转化效率，排除因名气效应导致的统计偏误。

技术资产规模 (ln Size)：以代码仓库体量衡量项目的技术厚度与复杂度。复杂的公共产品（如底层公链）往往面临更高的维护门槛与历史积淀，控制此变量可排除因技术难度或资产冗余导致的效率差异。

项目生命周期 (Age_month)：以项目自建仓至观测点的月数衡量。旨在捕捉开源项目随时间推移而产生的声誉积累与社区成熟度效应，消除项目所处不同演化阶段对产出效率的影响。

(2) 社区治理与领域跨度

管理规范程度 (Has_Wiki / Has_Projects)：通过项目是否启用 Wiki 文档或 Projects 管理面板作为虚拟变量，控制项目在组织协调与知识库建设层面的正规化程度。完善的治理基础设施能有效降低外部贡献者的进入门槛。

标签丰富度 (Topic_Counts)：通过计算仓库技术标签的数量，衡量项目的跨领域特征、技术定位清晰度及潜在受众的覆盖广度。

(3) 组织背景与环境效应

所有者类型 (Owner_Type)：模型以个人开发者为基准组，引入组织/机构虚拟变量。旨在控制机构化运作在资源调度、资本投入与信用背书上相对于个人开发者的天然优势。

4.5.3 实证结果分析

据表 4-5 的基准回归结果，我们可以对区块链激励机制在开源创新中的效能溢价进行深度的逻辑解构。实证结果显示，区块链激励机制在扩散效率、粘性

效率及开发强度三个维度上均表现出极强的正向显著性，这不仅验证了代币化创新对开源生态的重塑作用，也揭示了其在解决公共品供给困境中的独特性。

首先，在衡量知识溢出效率的扩散效率模型中，区块链激励机制展现出了高达 2.4710 的回归系数，这在三个效率指标中最为显著。这意味着在控制了项目知名度、技术体量及时间等变量后，区块链项目将外部关注转化为技术采纳和二次开发的能力显著优于传统模式。从公共品理论视角分析，这意味着代币化激励有效地降低了外部开发者参与再创作的门槛。通过将未来的生态价值提前锚定并进行透明化分配，区块链项目成功引导了更多开发者从被动观察者转变为主动贡献者，加速了知识公共品从社会资本向生产资本的转化进程。

进一步观察可以发现，区块链机制在粘性效率与开发强度上也表现出稳健的促进作用。尤其在粘性效率模型中，一个有趣的现象是项目知名度的系数显著为负（系数值位-0.2965），这揭示了开源生态中普遍存在的注意力稀释悖论，即随着项目关注度的扩张，单位关注度所能带来的深度跟踪比例反而会因信息噪音的增加而下降。然而，区块链机制的正向显著性反衬出其在对冲注意力流失方面的制度优势。这种利益绑定效应建立了一种强烈的契约关系，促使持份者为了维护自身数字资产价值而愿意承担更高的长期跟踪成本，从而在流量扩张的过程中维持了社区的深度参与。在组织背景的对比分析中，所有者类型表现出明显的异质化逻辑。机构背景显著提升了项目的扩散效率与开发强度，但在粘性效率上则并不显著且系数偏负。这一对比说明，组织背书虽然能凭借其信用溢价和资源堆砌带来早期的流量红利与产出压力，但难以自发产生深层的社区忠诚。相比之下，区块链的激励红利是全维度的，尤其在建立长期承诺方面，其作为一种去中心化的制度设计，表现出了超越传统层级化组织背书的治理效能。

最后，治理基础设施的控制变量表现也为上述逻辑提供了微观支撑。文档规范化程度（Has_Wiki）对三项效率指标均有显著的正向贡献，证明了知识库建设在降低进入门槛方面的普适价值。然而，协作管理工具（Has_Projects）在提升开发强度的同时，却对粘性效率产生了显著的负向抑制。这暗示了过于严苛或结构化的管理流程虽然能驱动即期产出，但可能对松散型开发者产生一定的挤出效应。综合来看，区块链模式通过一种“低门槛、强激励、高频互动”的协作范式，在无需过度依赖层级化管理工具的背景下，实现了对开源公共品供给效能的优化，这为其在数字公共品领域的长期生命力提供了实证证据。

表 4.1 项目层面效率指标回归表

变量	(1) 扩散效率	(2) 粘性效率	(3) 开发强度
区块链激励机制	2.4710*** (10.99)	0.7968*** (4.30)	1.9293*** (6.60)
项目知名度	0.5770*** (19.72)	-0.2965*** (-9.6288)	0.7606*** (19.37)
技术资产规模	0.0968*** (3.26)	-0.0387 (-1.28)	0.1139*** (2.90)
项目月龄	-0.0034 (-0.14)	0.0027 (-0.12)	-0.0201 (-0.63)
文档规范化程度	0.3915** (2.35)	0.3275* (1.82)	0.4258* (1.81)
协作管理工具应用	-0.1733 (-0.90)	-0.3154** (-2.11)	0.4944* (1.80)
技术标签丰富度	0.0181 (0.90)	-0.0172 (-1.21)	0.0327 (0.96)
所有者类型	0.6586*** (3.77)	-0.0916 (-0.66)	1.0932*** (4.40)
编程语言固定效应	Yes	Yes	Yes
时间固定效应	Yes	Yes	Yes
R ²	0.6774	0.3422	0.6454
Observations	971	971	971

注：a.圆括号里为 t 值；
b.* p<0.10, ** p<0.05, *** p<0.01；
c.均采用聚类稳健标准误

第五章：代币化激励下开发者协作行为的微观演化分析

5.1 开发者行为实证研究设计

在证实了区块链激励机制对项目宏观绩效的正向影响后，本章旨在进一步解构其微观动力系统。本节将详细说明开发者行为数据的构建过程、计量模型设定以及核心控制变量的定义。

5.1.1 数据来源与非平衡面板构建

为了精准捕捉开发者在动态协作过程中的行为变迁，本研究利用 GitHub 事件流数据，本研究提取了包括 Pull Request（PR）申请、代码评审（Review）、以及跨窗口的贡献权重变化等数据，构建了一个微观开发者-月度-项目层面的非

平衡面板数据。不同于第四章的项目截面数据，本章的数据颗粒度下沉至开发者行为维度。每一行观测值代表开发者们在特定月份对特定项目的贡献实态（包括提交次数、评审行为、角色晋升等）。考虑到开发者进入与退出社区的动态过程，非平衡面板能够真实还原社区的演化逻辑。数据集涵盖了样本项目自建仓起至观测终点的全生命周期活跃记录，为后续动员、治理和权力结构三个模块的分析提供了高频微观证据。

5.1.2 计量模型设定与因果识别策略

本章采用带有多重固定效应的线性回归模型进行实证检验。考虑到开发者层面的行为指标（如留存率、晋升概率、贡献占比等）普遍存在分布不均特征，本研究统一采取如下对数化计量方程：

$$\ln(Y_{it} + \epsilon) = \alpha + \beta \cdot is_blockchain_i + \gamma \cdot Controls_{it} + \lambda_{lang} + \delta_t + \epsilon_i$$

模型设定说明：

- Y_{it} ：代表项目 i 在 t 月份的微观行为指标（涵盖动员、治理和权力结构三个模块）。
- $is_blockchain_i$ ：核心解释变量。用于识别代币化激励机制对开发者行为的净效应。
- ϵ ：平滑项（取值为 0.0001），用于处理包含零值的微观比例指标，确保对数转换的有效性。
- λ_{lang} ：编程语言固定效应，控制不同技术社区特有的协作习惯。
- δ_t ：月份-年份固定效应。这是微观面板分析的关键，用以吸收宏观市场波动（如加密货币价格周期）。且由于实验组与控制组的抽样对比是基于建仓月份的随机抽样，该效应的控制同时完成了对于实验组与控制组基于出生月份的匹配。

5.1.3 控制变量库定义与分类逻辑

为了在开发者微观层面精准识别代币化激励机制产生的制度净效应，本研究基于动员与留存、治理与协作、所有权与流动性三个维度的理论特性，构建了差异化的控制变量组合。

（1）全局基础控制变量

无论在何种微观行为回归模型中，以下变量均被纳入以控制样本间最基础的禀赋差异：①月份-年份固定效应（ δ_t ），锁定宏观技术周期与外部市场行情带来的整体冲击以及月份匹配；②编程语言固定效应（ λ_{lang} ），控制不同技术栈

对于开发者生态的影响；③项目月龄，控制项目在不同生命周期阶段（如爆发期、平稳期）表现出的自然协作规律。

（2）模块化差异化控制逻辑

本研究针对不同模块的识别偏误，分别引入了特定的控制变量组，其逻辑阐述如下：

在动员与留存方面，一个重要的识别陷阱在于区块链项目是否仅因为其天生的高热度而吸引人留存或贡献？因此，控制项目初始时刻的项目热度以及活跃开发者数量，有利于避免此类问题。故此，本研究在动员与留存的回归中，控制了**初始项目知名度**和**初始人力资源禀赋**，旨在剔除因社会资本带来的引流效应，并控制项目创建初期的参与热度，从而识别代币激励机制在同等知名度下的转化红利。

在治理与协作方面，为了回应治理延迟是由于协作的人群数量抑或是激励制度低效这一问题。本研究控制了**活跃开发者数**（即当月的协作人数）和**技术资产规模**（即仓库代码体量），以此验证控制了相同协作规模和协作难度下，区块链项目是否比传统模式更具效率。并且，同时控制了**文档规范化程度**（Has_Wiki），旨在控制常规管理手段（如完善的文档）对进入门槛及评审效率的稀释作用。

在所有权与流动性方面，为了排除项目发起方背景对社区层级的先天塑造作用，本研究控制了**所有者类型**和**初始代码所有权集中度**，尽可能排除机构以及项目早期领袖魅力对于激励机制在所有权与阶层流动性方面的影响。

5.1 社区动员与留存的动态演化

从活跃开发者数（MAD）的全生命周期演化轨迹来看（见图 5.1 左上），实验组（区块链项目）展现出显著的冷启动优势。在项目建立初期，其动员规模即优于对照组，且随月龄增加呈现出更为陡峭的上升斜率，在观测期内保持了持续的扩张态势。这种直观的规模差距验证了激励机制在缓解公共品生产早期协调失灵方面的核心作用：通过将未来生态价值与代币绑定，有效补偿了早期开发者的投入不确定性与机会成本。然而，值得深思的是表 5.1 的回归结果显示，在控制了初始人力资源禀赋与初始知名度后，区块链激励对活跃开发者数（MAD）的净影响并不显著（T 值为-0.18）。这一统计显著性的消失揭示了动员规模的深层逻辑：区块链项目的规模优势很大程度上源于其更强的初始动员能力（初始人力禀赋系数 0.5545，T 值大于 4）和项目热度（初始知名度系数 1.0992，T 值大于 2）。一旦在同一配对层级内拉齐，代币机制对参与者总数的扩张作用便让位于对参与质量的塑造。

次月留存率与季留存率是观察激励机制能否将偶然贡献转化为持续参与的关键指标。描述性统计显示，尽管两组项目均随月龄增加表现出自然下降的疲劳曲线，但实验组的两项留存指标始终稳定在对照组上方，且两者间的差距随时间演进展现出稳健的持续性。回归分析进一步证实了这种持续的留存溢价：区块链激励机制显著提升了次月留存率（系数 0.0607，T 值 3.39）与季留存率（系数 0.0560，T 值 2.76）。这种溢价揭示了代币激励下的所有权锁定机制，代币不仅是经济补偿，更是社区治理权的表征。在这种经济—治理耦合体系下，开发者从单纯的劳动力提供者转变为利益相关的治理者，这种身份重塑有效对冲了开源社区常见的注意力漂移。特别是在初始知名度显著抑制留存率的背景下，代币机制的正向显著性有力回应了激励仅带来短期投机的观点，证明其在流量冲击下依然能保护社区的核心忠诚度。

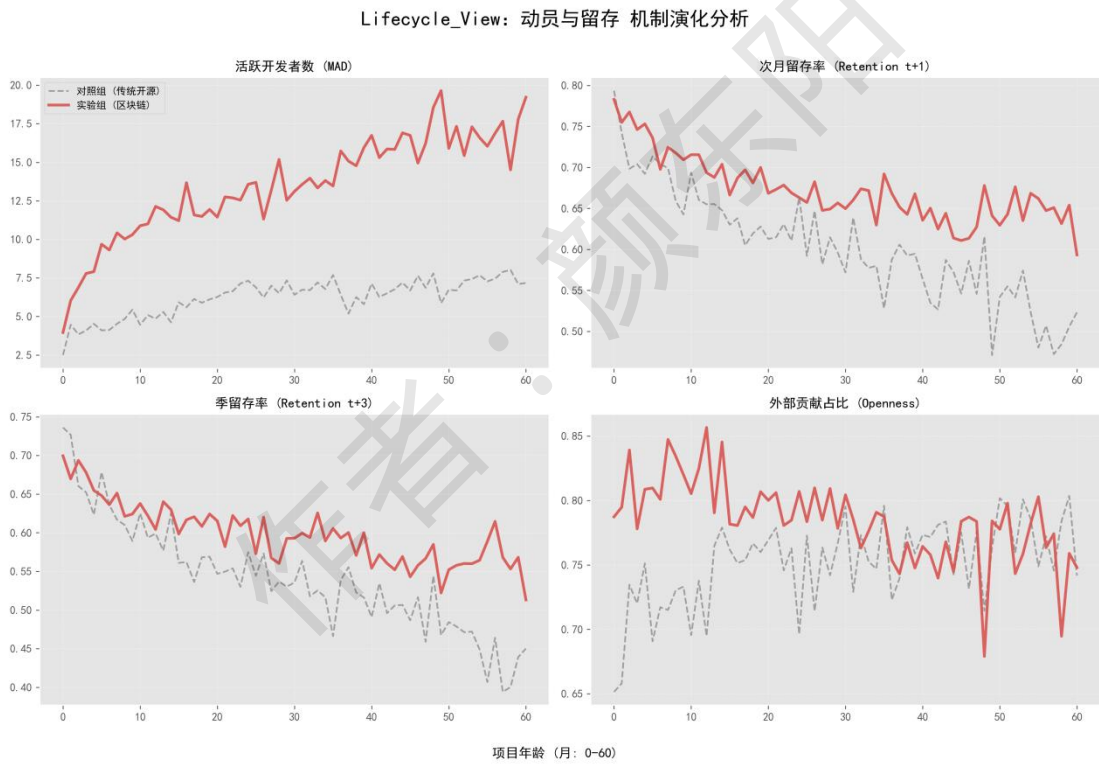


图 5.1 生命周期视角下两类开源社区动员与留存演化图

外部贡献率反映了项目核心层之外的社会化协作深度。实证观察发现，区块链项目的外部参与始终维持在高位，且在回归模型中表现出显著的增量效应（系数 0.0710，T 值 2.37）。这意味着代币化机制并未导致社区闭锁，反而通过所有权动员实现了更广域的社会化动员。这种无许可的参与特质，使得区块链创新生态更接近于知识生产模式 2（Mode 2）下的开放协作范式，即知识生产能够跨越组织边界、在高度异质化的群体间协同实现。控制变量中初始人力禀赋对外部贡

献的正向拉动说明了早期的核心积累能产生更强的外部吸引，而区块链机制通过制度化的分配规则，确保了这种吸引力能够转化为持续的社会化协作，实现了技术产出与外部劳动力供给的深度耦合。

表 5.1 社区动员与留存回归表

变量	(1) 活跃开发者数	(2) 次月留存率	(3) 季留存率	(4) 外部贡献率
区块链激励机制	-0.9499 (-0.18)	0.0607*** (3.39)	0.0560*** (2.76)	0.0710** (2.37)
项目月龄	0.6314 (1.62)	-0.0024*** (-3.76)	-0.0029*** (-3.91)	-0.0013 (-1.57)
初始人力资源禀赋	0.5545*** (4.19)	-0.0046** (-2.52)	-0.0057*** (-3.16)	0.0054*** (3.37)
初始项目知名度	1.0992** (2.18)	-0.0311*** (-7.48)	-0.0329*** (-6.60)	0.0335*** (3.74)
配对固定效应	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
编程语言固定效应	YES	YES	YES	YES
R ²	0.149	0.195	0.203	0.160
Observations	30419	26775	25439	20555

注：a.圆括号里为 t 值；
b.* p<0.10, ** p<0.05, *** p<0.01;
c.均采用聚类稳健标准误

5.2 协作治理与效率

在开源治理中，代码合并的准入控制是权力分配的核心体现。由图 5.2 所示，实验组（区块链项目）的把关人集中度在全生命周期内始终维持在 0.6 至 0.75 之间，显著低于对照组（约 0.85 至 0.90）的垄断水平。这一发现得到了表 5.2 回归结果的有力支撑：在控制了固定效应与协作负荷后，区块链激励机制显著降低了把关人集中度（系数-0.1381,T 值为-4.99）。这意味着代币化机制打破了传统开源项目中常见的集中式治理，实现了评审权力结构的相对扁平化。与此同时，激励机制对评审多样性展现出极强的正向推动力（系数 0.1084,T 值为 3.97）。这种权力的社会化扩散揭示了区块链项目独特的治理逻辑，通过所有

权激励引入多方利益相关者参与决策，将代码审核从少数核心成员的私人领地转化为多中心驱动力的社区共识过程，从微观行为层面落实了去中心化的制度设计。

一个长期的学术争论是去中心化是否会以牺牲协作效率为代价。然而，表 5.2 的实证数据给出了反直觉的证据：区块链激励机制显著缩短了合并耗时（系数-4.6292,T 值 2.11），且对代码合并率没有显著的负向干扰。结合图 5.2 观察可以发现，尽管区块链项目由于底层技术资产规模较大且开发强度更高，导致初期交付周期波动剧烈，但在进入成长期后（约 10 个月后），其响应速度往往能与对照组持平甚至在特定阶段展现出更高的柔性。这种去中心化且高效率的协作范式，证明了代币激励下的异步协作机制能够通过利益耦合有效补偿沟通成本。回归结果中，技术资产规模对耗时的显著负向影响进一步说明，在高复杂度的项目环境下，区块链机制所激发的评审韧性能够更有效地分担维护压力，实现了知识生产模式 2（Mode 2）下的高效治理。

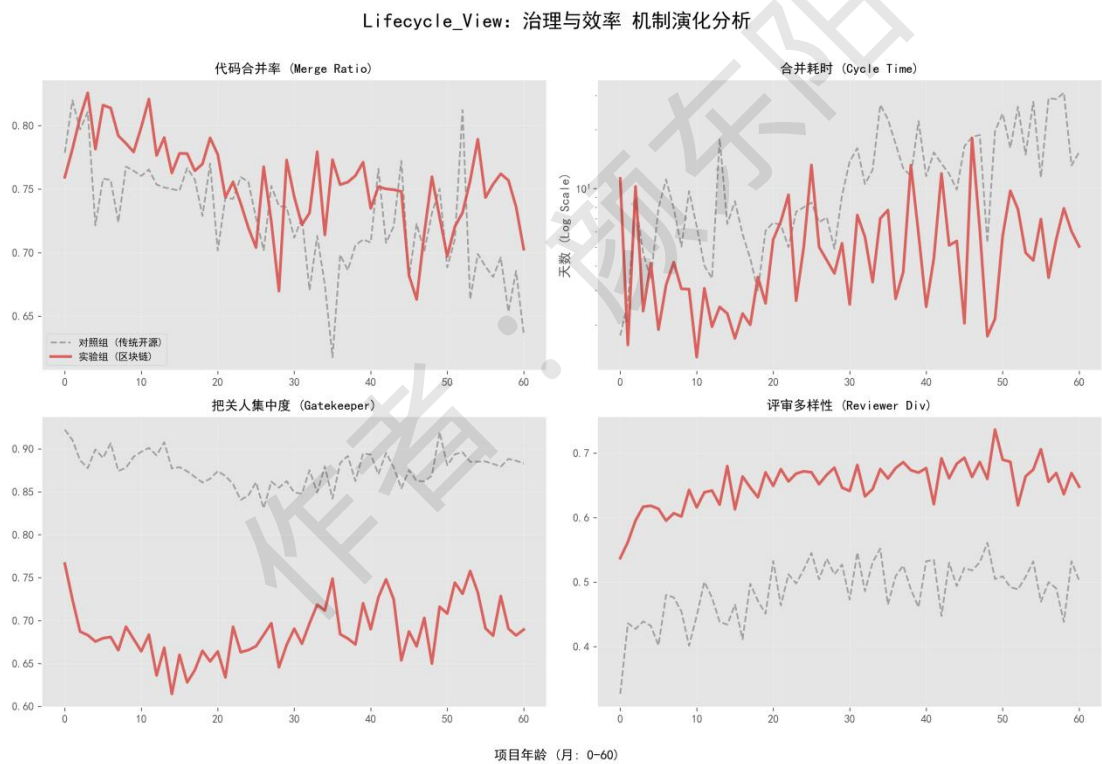


图 5.2 生命周期视角下两类开源社区治理与协作演化图

本小节的实证发现进一步揭示了治理结构与技术负荷之间的耦合逻辑。控制变量显示，活跃开发者数的增加会显著降低把关人集中度（-0.0035）并加速合并反馈速度，这说明协作规模的扩张在区块链背景下是正向的治理资产。结合前文发现的区块链项目技术资产规模大、开发强度高的特征，这种高评审多样性不仅没有导致系统性的决策瘫痪，反而成为了消化复杂技术需求、应对高压协作环境的必要手段。文档规范化程度（Has_Wiki）对合并率的负向抑制暗示了在严格治

理环境下贡献接纳门槛的提升，而区块链激励通过所有权绑定，使得这种高门槛并不会阻碍开发者的参与热情，反而确保了在快速迭代中维持较高的代码质量。综上所述，区块链开源生态并非单纯的无序扩张，而是通过所有权激励构建了一套能够支撑重型公共基础设施持续演进的微观协作体系，实现了治理机制与技术复杂度的深度适配。

表 5.2 社区治理与协作回归表

变量	(5) 把关人集中度	(6) 评审多样性	(7) 代码合并率	(8) 合并耗时
区块链激励机制	-0.1381*** (-4.99)	0.1084*** (3.97)	0.0056 (0.25)	-4.6292** (-2.11)
项目月龄	-0.0009** (-2.21)	0.0025*** (3.7)	0.0005 (0.65)	0.0213 (0.28)
文档规范化程度	-0.0088 (-0.47)	-0.0205 (-0.94)	-0.0382* (-1.85)	3.1689 (1.41)
技术资产规模	-0.0235*** (-5.18)	0.0339*** (5.28)	0.0058 (1.01)	-1.2771** (-2.56)
活跃开发者数	-0.0035** (-2.11)	0.0023 (1.24)	-0.0007*** (-4.66)	-0.1112** (-2.21)
配对固定效应	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
编程语言固定效应	YES	YES	YES	YES
R ²	0.381	0.310	0.124	0.054
Observations	17200	13473	20593	18083

注：a.圆括号里为 t 值；

b.* p<0.10, ** p<0.05, *** p<0.01；

c.均采用聚类稳健标准误

5.3 知识所有权与核心阶层流动性

代码所有权集中度与精英占比是衡量项目权力边界的关键判准。如图 5.3 所示，实验组（区块链项目）展现出极具特征的演化曲线：其代码所有权集中度在项目建立后的前 20 个月内经历了剧烈下降，从初始的 0.7 迅速降低至 0.3 以下，并长期稳定在低位。表 5.3 的回归结果强力证实了这一视觉发现：在控制了初始结构与所有者类型后，区块链激励机制显著降低了代码所有权集中度

（系数-0.0977, T 值-4.39）及精英占比（系数-0.0824, T 值-4.22）。这有力地论证了代币机制对精英俘获现象的打破，代币所有权的分配逻辑导致了技术贡献权的社会化扩散。相比之下，控制变量中的所有者类型表现出极强的中心化倾向，这意味着机构背书往往伴随着更高的权力垄断风险。区块链激励的负向显著性证明了其作为一种制度力量，成功对冲了组织层面的中心化偏误，使知识产出表现出更纯粹的公共品特征。

进一步观察核心团队稳定性可以发现（见图 5.3），实验组的 Jaccard 相似度指标虽然维持在 0.8 以上的高位区间，但在全生命周期内普遍略低于对照组。回归结果显示，区块链激励对稳定性的影响在统计上并不显著，而所有者类型则表现出显著的正向影响。这种对比揭示了区块链项目独特的动态演化逻辑。传统开源项目或机构主导项目由于缺乏利益退出机制，核心层往往在一定时期内形成一种高度封闭的稳定结构；而在区块链激励框架下，代币的流动性为贡献者的更替提供了市场化路径。虽然初始代码所有权集中度对核心团队稳定性有相当强的预示作用（系数 0.1529，T 值大于 10），但区块链机制通过分散化的激励路径，在维持核心协作连贯性的同时，赋予了社区更高的演化柔性，确保了公共产品在不同阶段的创新活力。

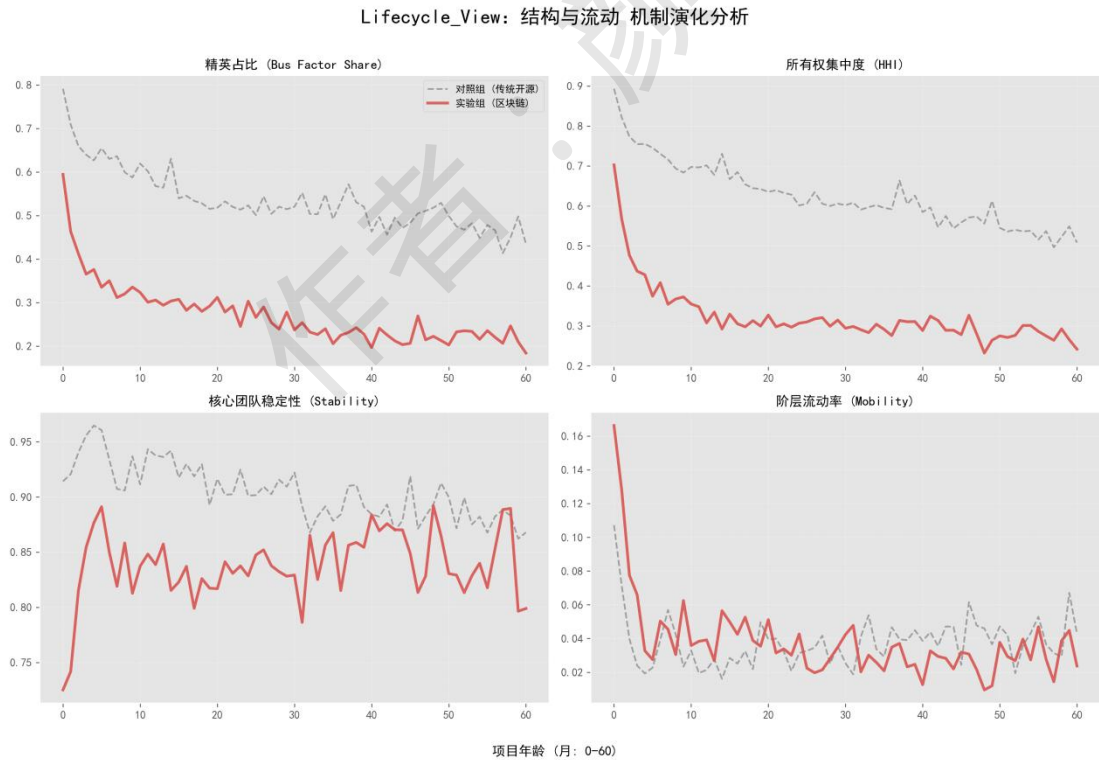


图 5.3 生命周期视角下两类开源社区所有权与流动性演化图

阶层流动率是判定社区是否具备可持续演化能力的微观指标。由图 5.3 可见，区块链项目在生命周期初期展现出极高的波动性，其流动率短时高达 0.16，

远超对照组。尽管回归模型中该指标的整体系数不显著，但这应与其高度波动的脉冲式特征有关。回归显示，项目月龄显著抑制了流动性，这符合项目成熟后结构趋稳的自然规律，但区块链项目在初期的爆发式流动成功打破了初始结构的固化。这种持续的阶层流动潜力与前述低所有权集中度共同构成了一个实证闭环。通过降低准入门槛并明确激励路径，区块链项目成功实现了无许可参与的愿景。虽然活跃开发者数的增加会略微降低单位流动率，但区块链机制通过制度设计确保了知识资产不被特定利益集团长期垄断。这种由代币激励驱动的微观行为模式，从公平性与开放性两个维度落实了公共品生产的民主化，完成了从宏观绩效到微观结构的逻辑统一。

表 5.3 社区所有权与流动性回归表

变量	(9) 代码所有权集 中度	(10) 精英占比	(11) 核心稳定性	(12) 阶层流动性
区块链激励机制	-0.0977*** (-4.39)	-0.0824*** (-4.22)	-0.0083 (-0.85)	-0.0048 (-1.14)
所有者类型	0.1482*** (6.11)	0.2022*** (8.21)	0.0343*** (3.37)	0.0048 (0.74)
项目月龄	-0.0037*** (-5.54)	-0.0039*** (-7.15)	-0.0002 (-0.68)	-0.0007*** (-3.59)
初始代码所有权集 中度	0.4563*** (12.24)	0.2712*** (7.66)	0.1529*** (10.42)	-0.0314*** (-4.15)
活跃开发者数	-0.001 (-1.62)	-0.0009 (-1.55)	-0.0001 (-0.96)	-0.0001** (-2.34)
配对固定效应	YES	YES	YES	YES
时间固定效应	YES	YES	YES	YES
编程语言固定效应	YES	YES	YES	YES
R ²	0.5543	0.4479	0.0864	0.0429
Observations	28084	28084	26898	22910

注：a.圆括号里为 t 值；

b.* p<0.10, ** p<0.05, *** p<0.01；

c.均采用聚类稳健标准误

第六章 结论与展望

本研究通过对区块链开源项目与传统高质量开源项目的深度对比，系统性地揭示了代币激励机制对开源公共品供给效能的促进作用。在宏观产出层面，本研

究发现区块链项目展现出比传统模式更强的生存韧性与产出绩效。代币化机制通过将未来生态价值通过代币分配,有效补偿了开发者在公共品生产早期的机会成本,使其能够顺利跨越开源社区常见的死亡谷。这种制度红利直接转化为宏观上的效率提升,即在同等社会化关注下,区块链项目表现出显著更高的扩散效率、粘性效率及开发强度。这一发现不仅论证了代币激励机制在解决开源知识公共品供给困境中的作用,也为理解市场化激励与社会化协作的融合提供了宏观证据。

更深层次的研究发现,宏观绩效的涌现根源于开发者微观协作逻辑与社区权力结构的转变。通过对开发者行为的追踪分析,本研究证实代币机制实现了从偶然贡献向所有权绑定的飞跃。代币不仅是经济补偿,更作为一种数字化所有权的表征,将开发者从搭便车者转化为利益相关的持份者,从而显著提升了社区的长期留存率与外部协作深度。同时,实证结果也回应了关于去中心化治理是否会导致低效这一问题,揭示了区块链项目能够在降低把关人集中度、提升评审多样性的同时,利用利益耦合机制缩短决策耗时。这种由激励机制改变的权力扩散与阶层流动,成功对冲了机构化运作可能带来的精英俘获,确保了知识资产不被特定所有者垄断,从而在微观层面落实了公共品生产的相对稳定性。

在理论层面,本研究通过构建“激励路径—协作行为—产出绩效”的全链条分析框架,拓展了公共品激励理论在数字经济时代的适用边界。通过论证代币激励制度对集体行动困境的缓解作用,本文为理解分布式知识生产与市场化提供公共品的理论探索提供了实证支撑。此外,本研究对高复杂度与强互动耦合关系的探讨,深化了知识生产模式 2 (Mode 2) 的理论内涵,揭示了在重型数字化基础设施建设中,多中心协作场景也能够通过激励制度红利补偿沟通成本,从而实现传统层级化管理模式的替代。

在实践层面,本研究为全球开源生态的治理优化提供了重要参考。对于开源社区而言,研究建议应辩证看待经济激励的引入,利用代币分配规则构建更具粘性的开发者核心层,以增强社区应对外部流量冲击的抵抗力。主动利用代币激励机制的流动性特性引导权力向社会化贡献者扩散,以维持长期的创新活跃度。此外,政府及相关政策制定者在推动数字公共基础设施建设时,可借鉴代币化治理的分配逻辑,通过数字化所有权确认机制,激发全社会参与公共产品供给的积极性。

尽管本研究利用精准配对与面板回归提供了稳健的实证证据,但受限于数据可获得性与环境复杂性,仍存在进一步深化的空间。首先,目前的研究主要基于 GitHub 的行为痕迹数据,虽然捕捉了协作业态,但尚未深入结合链上真实的代币转移记录 (On-chain Data)。未来的研究可尝试将开发者行为数据与链上财富效应、代币锁定周期进行关联,以更细微地刻画激励参数对协作深度的影响。此

外，治理规则的多样性也是未来值得挖掘的方向。不同区块链生态在所有权设计、及分配算法上的细微差异，可能对开发者阶层流动性产生不同的塑造作用。未来的研究可进一步通过多案例对比或实验研究，识别不同治理参数下的最优激励曲线。

作者：颜东阳

参考文献

- [1] Alchian A A, Demsetz H. Production, information costs, and economic organization[J]. The American economic review, 1972, 62(5): 777-795.
- [2] Alexy O, George G, Salter A J. Surrendering control to gain advantage: Reconciling open controls and the role of the firm[J]. Strategic Management Journal, 2013, 34(10): 1147-1171.
- [3] Allen D W E, Berg C, Markey-Towler B, et al. Tokenomics as a new economic institution[J]. Available at SSRN 3564027, 2020.
- [4] Arakelyan A, Spinellis D, Mitropoulos D. The GitHub repository lifecycle[J]. Journal of Systems and Software, 2024, 208: 111894.
- [5] Baldwin C Y, von Hippel E. Modeling a strategy of open collaborative innovation[J]. Strategic management journal, 2011, 32(13): 1399-1423.
- [6] Bayern S. Autonomous organizations[M]. Cambridge University Press, 2021.
- [7] Becher T, Trowler P R. Academic tribes and territories: Intellectual enquiry and the culture of disciplines[M]. Open University Press, 2001.
- [8] Belenzon S, Schankerman M. Motivation and sorting of open source software developers[J]. Management Science, 2015, 61(1): 79-101.
- [9] Benkler Y. The wealth of networks: How social production transforms markets and freedom[M]. Yale University Press, 2006.
- [10] Bergquist M, Ljungberg J. The power of gifts: organizing social relationships in open source communities[J]. Information Systems Journal, 2001, 11(4): 305-320.
- [11] Bitzer J, Schrettl W, Schröder P J H. Intrinsic motivation in open source software development[J]. Labour Economics, 2007, 14(1): 160-169.
- [12] Bonaccorsi A, Rossi C. Why firms contribute to open source software: Notes on a taxonomy of motives[J]. Research policy, 2006, 35(8): 1077-1098.
- [13] Catalini C, Gans J S. Some simple economics of the blockchain[J]. Journal of Economic Literature, 2020, 58(1): 80-126.
- [14] Coelho J, Valente M T, Milani L L, et al. Why modern open source projects fail[J]. Empirical Software Engineering, 2020, 25: 1581-1607.
- [15] Cong L W, He Z. Blockchain disruption and smart contracts[J]. The Review of Financial Studies, 2019, 32(5): 1754-1797.
- [16] Dahlander L, Wallin M W. A man on the inside: Unlocking communities as complementary assets[J]. Research policy, 2006, 35(8): 1243-1259.
- [17] Dasgupta P, David P A. Toward a new economics of science[J]. Research policy, 1994, 23(5): 487-521.
- [18] Eghbal N. Roads and bridges: The unseen labor behind our digital infrastructure[R]. New York, NY: Ford Foundation, 2016.
- [19] Eghbal N. Working in Public: The Making and Maintenance of Open Source Software[M]. Stripe Press, 2020.
- [20] Fos V, Holderness C G. The supply and demand for shareholder voting rights[J]. The Review of Financial Studies, 2023, 36(11): 4487-4523.
- [21] Ghapanchi A H, Aurum A. Antecedents to IT personnel's intentions to leave open

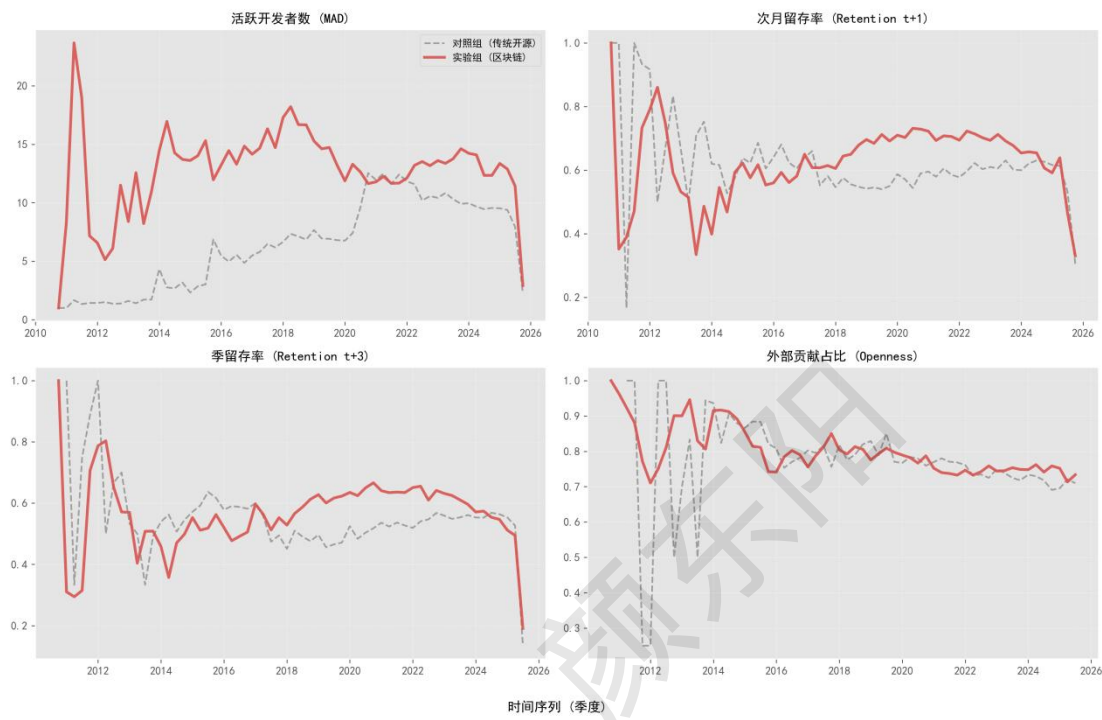
- source software projects[J]. *Journal of Systems and Software*, 2011, 84(6): 951-965.
- [22] Gibbons M, Limoges C, Nowotny H, et al. *The new production of knowledge: The dynamics of science and research in contemporary societies*[M]. Sage, 1994.
 - [23] Goldfarb A, Tucker C. Digital economics[J]. *Journal of Economic Literature*, 2019, 57(1): 3-43.
 - [24] Hann I H, Roberts J, Slaughter S, et al. The returns to open source software development[J]. *Management Science*, 2004, 50(12): 1745-1760.
 - [25] Hars A, Ou S. Working for free? Motivations for participating in open-source projects[J]. *International journal of electronic commerce*, 2002, 6(3): 25-39.
 - [26] Hart O, Moore J. Property rights and the nature of the firm[J]. *Journal of political economy*, 1990, 98(6): 1119-1158.
 - [27] Henkel J. Selective revealing in open innovation processes: The case of embedded Linux[J]. *Research policy*, 2006, 35(7): 953-972.
 - [28] Hertel G, Niedner S, Herrmann S. Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel[J]. *Research policy*, 2003, 32(7): 1159-1177.
 - [29] Hessels L K, van Lente H. Re-thinking new knowledge production: A literature review and a strategic agenda[J]. *Research policy*, 2008, 37(4): 740-760.
 - [30] Hussain W, Murtaza G, Zhou S, et al. Sustainability of open source software: A systematic literature review[J]. *Journal of Systems and Software*, 2024, 211: 111956.
 - [31] Klein J T. *Beyond interdisciplinarity: Boundary work, communication, and collaboration*[M]. Oxford University Press, 2021.
 - [32] Kydland F E, Prescott E C. Rules rather than discretion: The inconsistency of optimal plans[J]. *Journal of political economy*, 1977, 85(3): 473-491.
 - [33] Lakhani K R, von Hippel E. How open source software works: "free" user-to-user assistance[J]. *Research policy*, 2003, 32(6): 923-943.
 - [34] Lerner J, Tirole J. Some simple economics of open source[J]. *The journal of industrial economics*, 2002, 50(2): 197-234.
 - [35] Lerner J, Tirole J. The economics of technology sharing: Open source and beyond[J]. *Journal of Economic Perspectives*, 2005, 19(2): 99-120.
 - [36] Malinova K, Park A. Tokenomics: When do tokens beat equity?[J]. *Management Science*, 2023, 69(12): 7245-7268.
 - [37] Mazzucato M. *The entrepreneurial state: Debunking public vs. private sector myths*[M]. Anthem Press, 2013.
 - [38] O'Mahony S. Guarding the commons: How community managed software projects protect their work[J]. *Research policy*, 2003, 32(7): 1179-1198.
 - [39] Pardo A, Duarte D, Serebrenik A, et al. Global collaboration in open source software: A study of human capital distribution[J]. *Journal of Systems and Software*, 2022, 186: 111183.
 - [40] Raymond E S. *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*[M]. O'Reilly Media, 1999.
 - [41] Reineke P, Katila R. Decentralization in organizations: A revolution or a

- mirage?[J]. Academy of Management Annals, 2025, 19(1): 112-145.
- [42] Rikap C, Lundvall B Å. The digital innovation race: Conceptualizing the new world order[M]. Palgrave Macmillan, 2021.
- [43] Roberts J A, Hann I H, Slaughter S A. Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the Apache projects[J]. Management science, 2006, 52(7): 984-999.
- [44] Schweik C M, English R C. Internet success: A study of open-source software commons[M]. MIT Press, 2012.
- [45] Turek K. Accelerating social science knowledge production with the coordinated open-source model[J]. Nature Human Behaviour, 2023, 7(12): 2056-2066.
- [46] Voshmgir S. Token economy: How the Web3 reinvents the Internet[M]. Token Kitchen, 2019.
- [47] Weber S. The success of open source[M]. Harvard University Press, 2004.
- [48] Wen W, Ceccagnoli M, Forman C. Open source software and the direction of technical change[J]. Management Science, 2023, 69(11): 6712-6734.
- [49] Werbach K. The blockchain and the new architecture of trust[M]. MIT Press, 2018.
- [50] West J. How open is open enough? Melding proprietary and open source platform strategies[J]. Research policy, 2003, 32(7): 1259-1285.

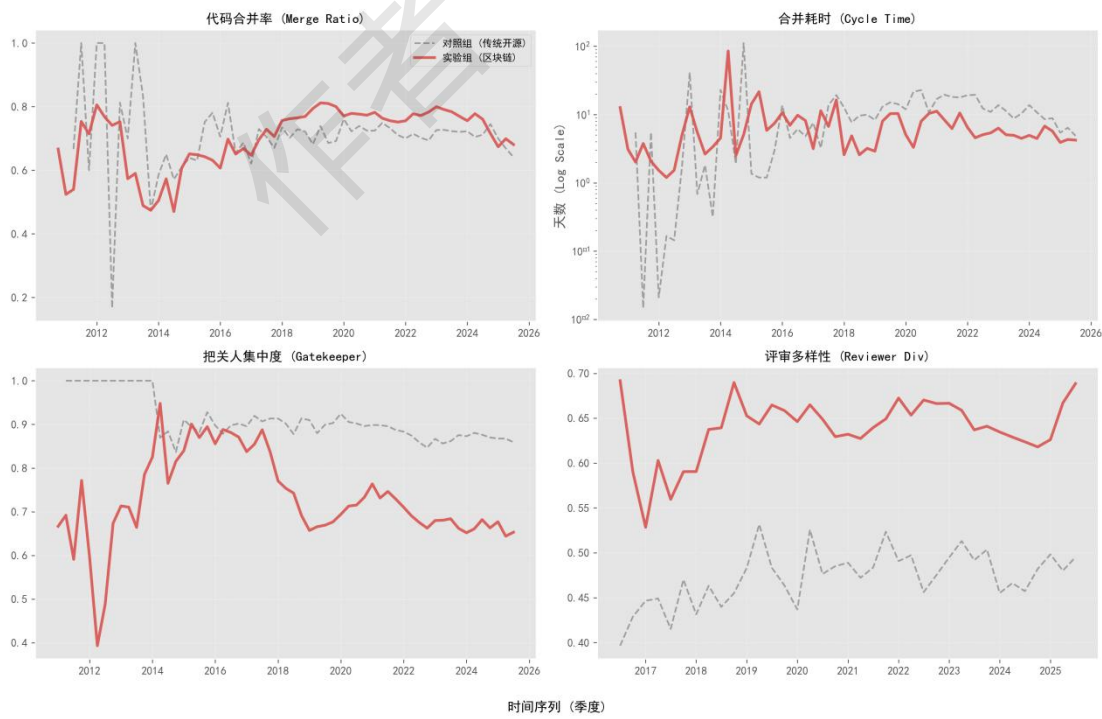
附录

第五章以日历时间视角的“动员”、“治理”、“结构” 3 个模块子图

Evolution_View: 动员与留存 机制演化分析



Evolution_View: 治理与效率 机制演化分析



Evolution_View: 结构与流动 机制演化分析

