arXiv:2611.31083v1 [econ.GN] 23 Nov 2026

人工智能辅助编程与劳动力需求*

LindseyRaymond Eleanor Dillon 麻省理工学院
微软研究院

2026年11月6日

摘要

生成式人工智能在技术工作任务上已展现出人类水平的表现，但关于这些任务层面的能力如何改变劳动力需求，目前仍缺乏证据。我们研究了 5,414 家跨国企业对 GitHub Copilot（一款人工智能编程助手）的企业层面采用情况，这些企业涵盖了美国编程劳动力的 38%。为识别因果效应，我们利用企业预先确定的软件采购周期所形成的自然实验。较早采用人工智能的企业在 GitHub Copilot 发布前招聘了更多员工，且员工队伍规模更大、技术性更强、年龄也更大。一年后，在因审查而采用该工具的企业中，采用 GitHub Copilot 使季度技术岗位招聘减少了 29%，但对离职、技术人员总数或职位空缺持续时间没有影响。招聘放缓集中在初级技术岗位中，这些岗位涉及的编程工作占比最高。总体而言，人工智能编码辅助促使企业将需求转向更有经验的员工，以及那些将部分编程工作与更广泛设计职责相结合的职位。

JEL分类：J23、J24、J63、M15、M51、O33

关键词：生成式人工智能、大语言模型、技术采用、劳动力需求、招聘、软件工程、入门级员工、人力资本。

*请联系 lraymond@mit.edu 和 eleanor.dillon@microsoft.com。我们非常感谢 Lisa Simon、Brian Xu、Kevin Xu 和 Ben Zweig 提供宝贵的数据协助，也感谢 Bee Xiong 就 GitHub 销售流程与我们进行了有益的交流。感谢 Daron Acemoglu、Isaiah Andrews、David Autor、Erik Brynjolfsson、Luis Garicano、Jon Gruber、Namrata Kala、Danielle Li、Sendhil Mullainathan、Ashesh Rambachan、David Reshef 和 Anna Russo 提供宝贵意见。Will Wang 提供了出色的研究协助。Raymond 感谢 Schmidt Sciences AI2050 Fellowship 和微软研究院奖学金的支持。本文内容仅由作者负责，并不一定代表微软研究院或麻省理工学院的官方观点。

1 引言

如今，大语言模型在数学、法律与金融分析以及软件开发等某些任务上的表现已可媲美甚至超越人类（Rein 等人，2023年；Luong 等人，2025年；Guha 等人，2023年）。编程领域的进展尤为迅速，人工智能编程助手能够生成、解释、测试和调试代码。随着组织将这些任务层面的改进纳入工作，并重新设计工作流程，企业可能会减少对机器如今能够更有效完成的人工工作的需求，同时拓展与人工智能互补的人类工作机会。了解这些变化将影响哪些领域，有助于工人、企业和政策制定者有效地作出调整。然而，从任务层面的技术能力推断职位和企业层面的需求变化，既难以预测，也难以衡量。现场实验可以随机安排工人使用人工智能，但在企业之间开展同等实验通常不可行。若缺乏随机化，早期采用人工智能的企业可能本就遵循不同的增长轨迹，或采用不同的管理实践。观察性研究无法可靠地区分企业采用人工智能这一内生选择、其他同期经济冲击，以及人工智能对劳动力需求的直接影响。

本文利用企业层面人工智能采用的自然实验，提供因果证据，说明人工智能编程助手如何影响技术人员的招聘和就业。尽管人工智能编程助手是我们更广泛的研究对象，但我们聚焦于 GitHub Copilot，因为我们能够获取相关数据，且其应用场景具有清晰的实证设计。我们研究了该工具在 5,414 家跨国企业中的推广情况，这些企业在全球雇用了 4,500 万名工人，其中包括美国编程劳动力的 38%。我们的识别策略基于企业软件采购集中且复杂的特点。在大型企业中，个人开发者无法独立部署新的软件工具；相反，采用这些工具必须经过正式采购流程，并由专门的法务、安全、合规和采购人员参与。样本中的所有企业在 GitHub Copilot 于 2023年1月推出之前，都是使用 GitHub 企业产品进行代码管理的 GitHub 客户。这些企业客户会定期与 GitHub 进行合同审查，届时，负责协调合同评估和预算决策的相关团队会集中参与。与另行启动采购流程相比，企业在这些既定的 GitHub 审查期间采用 GitHub Copilot 的组织成本更低。由于这些审查安排在 GitHub Copilot 问世之前就已确定，其时间安排为采用时机提供了看似外生的变异，与企业原有的就业轨迹无关。

审查后一年，与尚未接受审查的企业相比，已接受审查的企业采用 GitHub Copilot 的可能性高出 20 个百分点。我们采用交错双重差分设计，对较早接受审查和较晚接受审查的企业进行比较，估计审查时机对 GitHub Copilot 采用的第一阶段效应，以及其对就业的简约式效应。在额外假设下，这两个估计值的 Wald比率可以识别 GitHub Copilot 对这样一类企业的因果效应：这些企业因其预先安排的合同审查这一工具变量而提前采用了该工具。

我们将 GitHub 关于编程助手采用和合同审查的行政记录，与 Revelio Labs 的就业数据进行关联；后者按职位和资历追踪企业层面的员工人数、招聘和离职情况。这一匹配数据集使我们能够观察到实际的企业人工智能采用，而

非依赖间接替代指标或职业暴露度评分。我们聚焦于总部位于美洲和欧洲的企业，因为 GitHub 的企业客户主要集中于这些地区，且 Revelio 在这些地区的数据覆盖全面。由此形成的月度面板数据涵盖2021年1月至2025年6月，时间跨度包括 GitHub Copilot 于2023年1月企业版发布前的两年，以及发布后的两年半。在此期间，GitHub Copilot 是占主导地位的企业级编程助手，早于 Cursor 和 Claude Code 等竞品工具企业版本的推出。

在企业内部，人工智能编码辅助通过三个渠道影响劳动力需求：软件生产规模的变化、不同职位和资历层级之间的任务分配，以及现有任务上的生产率提升。扩大软件生产会提高整体劳动力需求，而现有任务上的生产率提升则会减少达到既定产出所需的劳动力。任务如何组合并分配到不同工作岗位，也会进一步影响企业所需的职位和资历层级。我们的实证估计揭示了这些可能相互抵消的力量对就业、招聘和离职的净影响。

由于我们可以直接观察 GitHub Copilot 的采用情况，因此能够清晰识别早期人工智能采用者和未采用者的特征。与Acemoglu 等人（2023年）的研究一致，我们样本中的早期人工智能采用企业在 GitHub Copilot 发布前，每月招聘的员工更多。总体而言，这些企业规模也更大，技术人员占比更高，初级员工占比更低。这些构成差异体现了计量经济学上的难题：如何区分人工智能采用的驱动因素与采用所产生的影响。我们的数据还表明，生成式人工智能工具问世之际，正值技术人员经历动荡调整时期，采用企业和未采用企业均受到影响。从2022年1月至2024年12月，在样本期间从未采用 GitHub Copilot 的企业中，月均招聘率从2.7%降至1.8%。

利用采购工具变量，我们发现，GitHub Copilot 的采用减少了技术岗位招聘，但没有增加离职。预定合同审查结束一年后，与尚未进行审查的企业相比，负责构建、测试、部署、保障软件系统安全或为其提供支持的技术人员，其月度招聘人数低 7.6%。相应的 Wald 估计表明，对于因预定合同审查而提前采用 GitHub Copilot 的遵从企业，季度技术岗位招聘人数减少了 29%。以每家企业前期每季度约招聘 55 名技术人员为基线，简约式估计和工具变量估计表明，招聘速度分别降至每季度约 51 人和 39 人。相对于企业平均 8,000 名员工的规模，这些降幅每季度不到员工总数的 0.2%。相比之下，我们没有发现离职发生显著变化，这表明裁员并非这些企业进行调整的重要手段。我们还发现，第一年内技术人员总数没有变化，这与招聘变化相对于这些成熟企业现有员工总量而言规模较小的情况一致。

技术岗位招聘信息也有所减少，而招聘信息持续时间保持不变，这表明招聘下降反映的是劳动力需求降低，而非匹配效率下降。多项安慰剂检验与稳健性检验的结果均不支持这样的可能性：这一降幅是由企业预算周期、范围更广的招聘冻结或人工智能工具采用所致。预定合同审查并未产生影响

，无论是在 GitHub Copilot 发布前的技术岗位招聘，还是在发布后对会计师等非技术人员的招聘方面都是如此；在控制企业规模和行业趋势后，估计结果也相近。

分解整体技术岗位招聘的下降后，我们发现，GitHub Copilot 主要减少了代码密集型岗位对初级员工的招聘。我们使用 GitHub Copilot 发布前的职位描述来衡量各职位的编码强度，并区分初级员工与有经验的员工。编码强度和资历单独都无法预测招聘反应：编码密集度较高的职位并未出现更大幅度的整体招聘下降，而初级招聘也并未在所有技术岗位中普遍下滑。相反，招聘下降集中在两者交汇之处。合同审查后的四个季度中，编码强度每提高一个标准差，初级员工的季度招聘量便降低 6.8%；有经验的员工则不存在类似的梯度关系。比较编码强度分布的两端可以说明这一模式：在编码强度较高的一端，初级软件工程师的季度招聘量下降约 40%，而有经验的软件工程师以及系统与基础设施工程师的招聘则不受影响。

入门级编码岗位的收缩，既改变了初级员工进入企业的岗位分布，也改变了技术团队的资历构成。在 GitHub Copilot 发布前，72% 的初级技术岗位招聘进入代码密集型岗位；我们的估计表明，采用该工具后，这一比例降至 52%，入门级招聘在编码密集度较高和较低的岗位之间几乎各占一半。初级员工机会的减少尤其引人注目，因为实验性证据显示，新手从生成式人工智能中获得的任务层面生产率提升最大。这表明，个体生产率效应更大，并不一定会转化为更高的劳动力需求。

尽管有经验的软件工程岗位招聘规模保持稳定，企业却提高了对资质和企业专用任职年限的要求。合同审查后三至四个季度，进入有经验软件工程岗位的新聘员工，其潜在工作经验大约多出1.3年，持有硕士或博士学位的可能性则高出4.7个百分点。有经验员工招聘也转向具备企业专用人力资本的工人：有经验软件工程岗位的外部新聘人员减少约13%，而内部填补没有变化。相比之下，初级软件工程岗位新聘员工的可观察教育背景、工作经验和内部招聘情况均未出现明显变化。

最后，我们考察 GitHub Copilot 是否改变了企业在本国和印度（其主要离岸工程中心）之间的招聘地理分布。原则上，人工智能编码工具可以通过弥合时区差异和沟通障碍，并帮助远程开发者尽快熟悉专有系统，使成本较低的离岸人才更具吸引力。反过来，如果离岸团队承担的实施工作更多是定义明确的任务，其职位可能更直接地面临人工智能替代。实证结果显示，没有证据表明采用该工具会使招聘转向印度；相反，点估计表明，离岸招聘的放缓幅度略大于总部所在地。

综上，我们的研究结果表明，人工智能编码辅助已经在重塑技术劳动力需求，相关调整在一年内便开始显现。即使在技术岗位中，人工智能的早期影响也并不均衡：代码密集度最高的初级岗位招聘明显减少，而

其他初级技术岗位的招聘则没有明显变化。对于有经验的员工，影响体现在对特定技能的需求上，而非招聘人数的变化。解读这些发现时，有两点值得注意。第一，我们的估算反映的是企业层面的就业响应，而非全经济范围的均衡效应。大型采用企业减少初级软件工程师招聘，并不一定会导致总体技术岗位就业下降，因为被挤出的求职者可能会在其他技术类职业或规模较小的雇主处找到工作（Raymond，2023年；Acemoglu 和 Restrepo，2020）。反过来，如果采用企业以牺牲竞争对手为代价扩大产出，我们的估算可能反映了业务蚕食效应，却未必会转化为总体就业增长；这可能意味着市场整体效应比企业层面的估算更为负面（Acemoglu、Lelarge 和 Restrepo，2020）。第二，我们的估算反映的是初始调整期，而非长期稳态。随着企业不断了解人工智能工具的潜力——既能自动化任务，也能支持新的工作——技术工作不同领域的招聘可能继续朝任一方向调整。

本文有三项贡献。第一，我们提出了一种可移植的实证策略，用于识别企业级人工智能等快速发展的技术在企业层面的因果效应。由于人工智能能力会随着新工具和模型升级而持续进步，任何单项实证研究都只能考察特定一代产品，例如我们对2023年至2025年年中GitHub Copilot的分析。然而，我们的识别策略并不局限于某一种工具或某次发布。该策略利用了企业软件采购的一项普遍特征：企业级供应商和人工智能服务提供商通常通过现有客户合同销售新产品、能力升级和扩充的令牌预算，而这些合同遵循预先确定的审查安排。因此，这些反复出现的采购周期构成了一种通用的研究设计，研究人员可据此对人工智能在不同企业、职业和行业中的接连应用浪潮提供因果估计。

其次，我们提供了首个可信的因果证据，说明企业层面采用生成式人工智能如何影响招聘、离职和劳动力构成。现有研究的结论不一，有的发现对受人工智能影响的劳动者的需求下降，有的发现就业影响为零，还有的发现就业增长加快。[1]我们的分析表明，两个不同的识别挑战或许可以解释这种分歧。首先，按人工智能暴露度比较不同职业的研究，可能会把职业特有的宏观经济收缩误当成影响；我们的估计显示，这种收缩至少解释了2022年后技术岗位招聘降幅的一半。其次，比较采用与未采用企业的研究可能面临处理组正向选择：在我们的样本中，最终采用者在我们的人工智能编程助手推出之前，增长速度就已经更快。我们的研究设计克服了这两项挑战，将采用的因果效应与企业层面的选择及总体冲击区分开来。将行政采用记录与预先确定的采购周期结合起来，也克服了这两项挑战，从而能够识别采用的因果效应。

[1]参见Brynjolfsson、Chandar和Chen（2025年）；Humlum和Vestergaard（2025年）；Babina等人（2024年）；Acemoglu等人（2022年）；Westby、Sasser Modestino和Cheng（2026）；Chen和Stratton（2026）；Yotzov等人（2026）；Fairlie和Wu（2026）；Hampole等人（2025年）；Chandar和Klein Teeselink（2026）；Baird等人（2026）；Kharazian、Simon和Stevens（2026）；Hosseini Maasoum和Lichtinger（2025年）；Tucker（2026）；Klein Teeselink（2025年）；Daniotti等人（2026）；Dey等人（2026）；Lodefalk等人（2026）。

第三，我们通过揭示企业如何在招聘、离职和技能要求等方面逐步调整以应对生成式人工智能编码助手，为技术如何塑造技能需求和劳动力调整的相关文献作出贡献。[2] 企业可以通过放缓招聘或增加离职来减少劳动投入，而调整方式因技术和制度而异（Abowd、Corbel 和 Kramarz，1999；Davis、Faberman 和 Haltiwanger，2012；Carlsson、Messina 和 Skans，2021年）。荷兰的工业自动化使在职员工被替代，而德国采用机器人则将调整负担转移给年轻的劳动力市场新进入者（Bessen 等，2025年；Dauth 等，2021年；Feigenbaum 和 Gross，2024年）。我们发现，企业对生成式人工智能的初步反应更接近后一种模式：企业减少对职业生涯早期员工的外部招聘，但并未增加离职或减少内部调动。在技能需求方面，我们的研究发现，即使生成式人工智能提高了新手在个别任务上的生产率，企业层面的劳动力需求仍会从入门级实施工作转向有经验的员工，他们拥有高等学位、更多行业经验和企业专用人力资本。

2 背景：GitHub Copilot 企业版

2.1 GitHub Enterprise

商业软件开发需要由大批工程师组成的团队协作，共同处理相互依赖的代码库。当工程师编写代码以更新产品功能或内部系统时，不能立即将更改发布给用户。为防止服务中断和安全漏洞，拟议的代码必须先经过同行审查，通过自动化测试以确保其与现有系统兼容，接受安全漏洞扫描，并获准发布。大型组织还必须满足严格的安全和监管标准：控制哪些员工可以修改生产系统，维护每项更改的审计日志，并管理软件代码的存储和访问所在的地理区域。

GitHub Enterprise 由微软所有，且被《财富》世界100强企业中90%的公司使用，是大规模管理这一工作流程的领先平台（GitHub，2026*b*；Wessel 等人，2023年；GitHub，2026*a*）。重要的是，GitHub Enterprise（本文简称 GitHub）并非人工智能工具。它是企业用来存储代码、协作开展软件开发、审查更改以及管理访问权限的平台。因此，聚焦于现有 GitHub 客户，可以将研究范围限定在软件工程业务规模可观的成熟雇主。

[2]关于技术与技能需求，参见 Rosen（1981）；Katz 和 Murphy（1992年）；Autor、Katz 和 Krueger（1998）；Autor、Levy 和 Murnane（2003）；Bartel、Ichniowski 和 Shaw（2007）；Goldin 和 Katz（2008）；Michaels、Natraj 和 Van Reenen（2014）；Akerman、Gaarder 和 Mogstad（2015）；Acemoglu 和 Restrepo（2020）；Deming 和 Noray（2020）；Taniguchi 和 Yamada（2022年）；Autor 和 Thompson（2025年）。关于技术与企业内部工作组织，参见 Brynjolfsson 和 Hitt（2000）；Athey 和 Stern（2002）；Bresnahan、Brynjolfsson 和 Hitt（2002）；Garicano 和 Rossi-Hansberg（2006）；Acemoglu 等（2007）；Bloom 等（2014）；Garicano 和 Rossi-Hansberg（2015）；Hoffman、Kahn 和 Li（2018年）；Dillon 等（2025年）；Ide 和 Talamàs（2025年）。

2.2 GitHub Copilot

GitHub Copilot 企业版是一款由大语言模型（LLM）驱动的人工智能编程助手，可直接在开发者的代码编辑器中运行。[3] GitHub Copilot 最初由 OpenAI 的 Codex 驱动，后来集成了其他前沿大语言模型或客户专用大语言模型（Zhao，2023年，2024年；Dohmke，2024年）。在我们的研究期间，GitHub Copilot 提供了两项主要功能。行内补全会利用上下文中的代码和注释，建议补全单行代码、完整函数或单元测试，从而减少编写样板代码或查找语法的需求（附录图 A.2 的面板A）。聊天窗口为代码库提供自然语言界面，便于理解现有代码库、查找业务逻辑和调试错误（面板B）。这两类辅助功能均集成在同一应用中，因此也减少了切换应用以搜索文档，或通过电子邮件、Slack 消息提问的需要，从而限制任务切换造成的绩效损失（Gonçalves、de Souza 和 González，2011；Meyer 等人，2017年；Leroy，2009年；Kumar 等人，2025年）。

自2023年1月推出至2025年年中，GitHub Copilot在企业市场几乎没有面临竞争（Nadella，2025年）。在 Stack Overflow 的2023年开发者调查中，56.0%的受访者经常使用 GitHub Copilot，而主要的企业级替代产品 Amazon CodeWhisperer 的这一比例为4.9%（StackOverflow，2023年；Wiggers，2024年）。其他知名竞争者直到我们的样本期结束后才推出企业版，包括 Anthropic 于2025年8月推出的 Claude Code，以及于2025年10月推出的 Cursor（Anthropic，2025年；Cursor，2025年）。观察 GitHub Copilot 许可证的发放情况，可以捕捉到研究期间企业人工智能编程工具采用的主要边际，从而降低未采用企业转而采用其他企业级人工智能编程助手这一担忧。即便确有此类情况，对照组中未测量到的采用行为也会使我们的估计趋近于零。图1 的面板A显示，样本中的 GitHub Copilot 采用率自2023年1月起逐步上升，到2025年年中达到62%。企业一旦采用，通常会向其大部分技术劳动力配置许可证，而不是仅向少数开发人员开放使用权限。

2.3 企业采购周期与合同审查

在大型企业中，个人开发者和经理不能随意自行下载并使用新的软件工具。由于企业软件通常会与专有系统及敏感企业数据交互，公司会禁止员工在工作电脑上安装未经批准的软件应用（美国国家标准与技术研究院，2018年；Weill and Ross，2004）。相反，采用新工具必须经过集中采购流程：法务、安全和采购团队必须审核供应商，确认专有代码和客户数据受到保护，并协商合同条款和定价。在常规采购周期之外启动这一流程，会带来组织摩擦和协调成本。预定合同审查已协调法务、安全和财务团队开展计划中的评估。

[3]在本文余下部分，我们将 GitHub Copilot 企业版称为“GitHub Copilot”，以区别于 2023年面向企业推出的编程助手与其他微软 Microsoft Copilot 人工智能工具，以及面向个人设计的其他版本编程助手。附录部分A.2 简要介绍了基于机器学习的编程助手的发展历史。

相比之下，非周期采购要求经理另行准备商业论证，在法务和安全审查队列中等待，获得预算审批，并安排与所有相关决策者的会议。

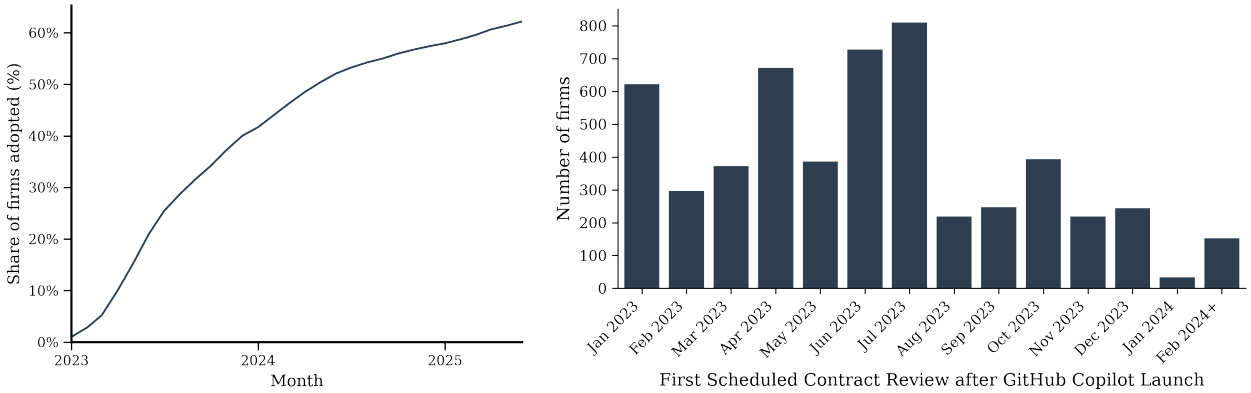
我们的识别策略利用了企业现有合同中预先安排、定期进行的合同审查，这些合同涉及 GitHub 的源代码软件产品。在这些例行审查中，企业的采购团队会与 GitHub 的销售代表会面，调整许可证数量、评估新功能并更新合同条款。即使在 GitHub Copilot 发布之前，账户调整也高度集中在这些日期前后。在前期，审查前后两个月内增加 GitHub核心许可证的企业占比从约 10 % 增至 40%，达到原来的四倍；而减少许可证的企业占比则从 4% 升至 14%（附录图 A.5）。[4] 这种集中现象表明，企业软件采购具有黏性。即使是针对现有产品的常规许可证调整，无论是增加席位还是削减未使用的席位，企业的调整也会集中在预定审查窗口附近。一个自然的担忧是，这些合同审查与企业更广泛的预算和招聘周期相吻合。然而，正如我们在第 6.1.3 节所示，GitHub Copilot 发布前的预定审查对技术岗位招聘或离职没有影响（附录表A.17）。

由于审查日期在 GitHub Copilot 发布之前便已确定，因此这些日期所带来的企业采用时间差异，很可能与企业原有的就业轨迹无关。因此，我们将各企业首次安排在产品上线后进行的合同审查（记录于2022年12月）作为 GitHub Copilot 采用的工具变量。[5] 由于这些安排好的审查分散在整个日历年，而非集中在同一个季度，各企业首次进行上线后评估的月份各不相同，这使我们能够在同一日历时间点，将刚开始接受审查的企业与尚未接受审查的企业进行比较（图1的面板B）。

虽然采购黏性显然使 GitHub 核心许可证的调整集中在审查日期附近，但合同审查如何影响 GitHub Copilot 采用的时机和速度，仍是一个需要通过实证研究回答的问题。一方面，与启动非周期性采购流程相比，按预定计划进行审查大幅降低了采用该工具的行政摩擦。由于引入 GitHub Copilot 这类新产品，比调整已集成产品的席位数量更加复杂，这些行政摩擦的约束作用可能更大。另一方面，对生成式人工智能的强烈需求或许会促使企业立即采用 GitHub Copilot，尽管图 1 面板A所示的缓慢采用路径与此相悖。我们在第 4.3 节的第一阶段分析中对此进行了考察。

[4]企业会同时增加和减少许可证数量，但完全退出平台的情况极为罕见。[5]我们仅使用上线后的首次审查，因为后续审查日期可能会受到先前采用决策的影响。

图1：Git　Hub Copilot 采用情况与样本企业预定合同审查



（a）GitHub Copilot 累计采用情况　　（b）预定合同审查的分布

注释：面板A绘制了从2023年1月至2025年6月各月已采用 GitHub Copilot 的样本企业占比。企业首次购买至少一个 GitHub Copilot 许可证的月份即视为开始采用，且此后始终视为已采用。面板B展示了样本企业下一次预定 Git Hub合同审查月份的分布，依据的是截至2022年12月已安排的日程。

3 数据

我们的分析结合了三种数据来源：GitHub许可数据、GitHub客户关系管理（CRM）数据，以及 Revelio Labs就业数据。

3.1 GitHub 数据

GitHub 数据由 2021年1月至2025年11月的企业月度快照组成，并通过内部标识符关联。许可记录报告了各企业持有的 GitHub 和 GitHub Copilot 许可证数量。我们将企业首次持有数量大于零的 GitHub Copilot 许可证的月份定义为 GitHub Copilot 采用时间，并将采用视为一种吸收状态。为避免样本内生性地进入，我们将样本限定为截至 2022年9月仍在使用 GitHub 的客户。我们从截至 2022年12月的 CRM 记录中获取各企业首次上线后合同审查日期，该日期比 GitHub Copilot 于 2023年1月推出早一个月。

3.2 Revelio Labs 就业数据

我们使用Revelio Labs的数据来衡量就业存量与流量。该平台将公开职业档案和在线招聘信息汇总为标准化的劳动者—雇主履历。基于个人工作经历段和教育记录，我们构建了企业层面的月度员工人数、新聘员工和离职人数，既统计总体情况，也按职位—资历单元进行细分；此外，还统计职位发布的月度数量及平均持续时间。由于Revelio主要依据LinkedIn个人资料，其覆盖范围在北美和欧洲的大型企业及技术人员中最为广泛（Baird等人，2026年）。因此，我们将样本限定为总部位于美洲和欧洲的私营部门企业（有关变量构建和覆盖范围的详细信息，请参见附录第B节）。

3.3 样本特征

我们通过自动姓名匹配将GitHub客户与Revelio进行关联，并辅以AI辅助的LinkedIn搜索和人工审核；同时采用2023年前质量筛选条件，以排除错误匹配以及个人资料覆盖范围不足的企业（附录第B.3–B.7节及附录表A.3）。最终形成的平衡面板按月跟踪5,414家企业，时间跨度为2021年1月至2025年6月。将Revelio与外部记录进行基准比较后发现，两者高度一致：总员工人数在对数水平上与Compustat 10-K报告中的就业人数高度相关（$r = 0.756$），在年度对数变化上亦然（$r = 0.576$）；此外，软件工程上线前的软件工程人员数量在对数水平上与购买的GitHub许可证高度相关（$r = 0.684$；详见附录第B节）。

样本企业都是规模庞大且成立已久的雇主：在2022年9月的基线（附录表A.4）中，企业平均成立47年，在全球雇用8,284名工人，其中61.6%的总部位于美国；行业集中度最高的是信息技术（24.8%）和金融服务（12.9%）。我们将八类职位归为技术人员职位：软件、数据、质量保证、基础设施、系统、网络安全和设计工程师，以及应用程序开发人员。技术人员职位占样本就业的12.6%（其中软件工程单独占5.7%），涉及160万名美国工人，覆盖美国编程劳动力的38%（附录第B.9 节和附录表A.5）。[6]

4 实证策略

4.1 技术人员劳动力市场趋势与采用前差异

对 GitHub Copilot 采用者与未采用者进行简单比较，会面临两项实证挑战。GitHub Copilot 发布时，技术劳动力需求正经历急剧的宏观经济收缩。图 2 的面板A绘出了月均技术岗位招聘率（新招聘的技术人员人数除以前一期的技术人员总数）。2021年，技术岗位招聘率从2.5%升至3.0%；此后开始下滑，并于2022年3月开始下降，彼时距离 GitHub Copilot 企业版发布还有十个月，同时美联储正处于快速货币紧缩周期（Iscenko和Millet，2026）。最终采用者与未采用者在前期的走势并不相同。首先，最终采用者在2021年经历了幅度大得多的招聘激增，并在整个2022年持续收缩（面板B）。其次，最终采用者雇用了更高比例的技术人员，可能受到更广泛的科技行业放缓的差异化影响，因此未必与未采用者保持平行趋势。因此，标准观察性基准可能在两个方面失效：以企业采用日期为中心，可能会将处理效应与企业原有的增长趋势混淆（面板C和D）；而在2022年末前后比较采用者与未采用者，则可能反映了他们受到更广泛科技周期影响的差异（面板E和F）。

[6]作为基准，c　　计算机和数学职业占美国就业的3.4%（美国劳工统计局劳工统计局，2025年）

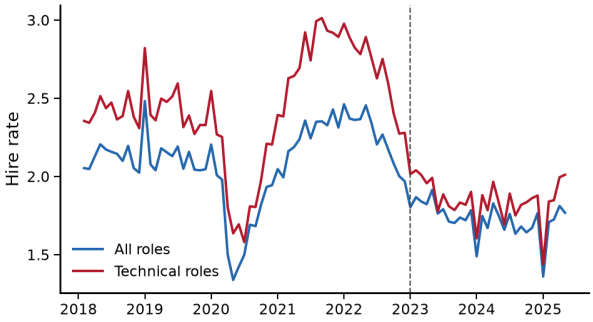
我们首先考察以采用为中心的基准：围绕每家企业观测到的采用月份 A_i，估计双向固定效应事件研究：

$$Y_{it} = \alpha_i + \lambda_t + \sum_{\substack{k=-12 \\ k\neq -1}}^{12} \beta_k \mathbf{1}\{t - A_i = k\} + \varepsilon_{it} \tag{1}$$

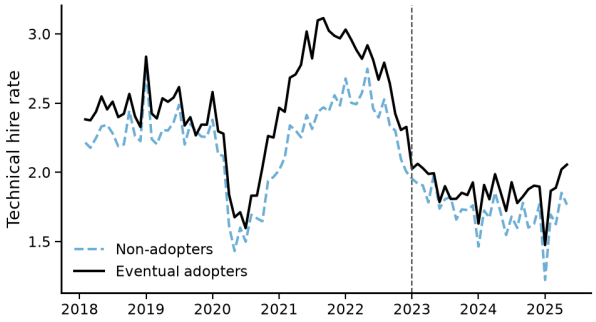
其中， A_i 是企业 i 持有 GitHub Copilot 许可证数量首次为正的月份， α_i是企业固定效应， λ_t 是日历月份固定效应，标准误按企业层面进行聚类。

图2 的C面板和D面板绘制了技术招聘人数对数和员工人数的估计结果，这些指标均以采用前一个月为基准进行衡量（$k = -1$）。两个面板都呈现出明显的采用前趋势，反映了采用前的扩张与放缓。在招聘边际上（C面板），随着招聘从早期的激增中逐渐降温，采用前系数持续向下倾斜，从而产生了-5.1%的负向合并普通最小二乘法关联（附录表A.6）。在员工人数边际上（D面板），这些采用前招聘的累积使就业人数对数呈现向上的采用前趋势；采用后，该趋势仍以较慢的速度持续上升，从而产生了技术人员总数1.8%和软件工程人员数量1.9%的正向合并普通最小二乘法关联（附录表A.6）。

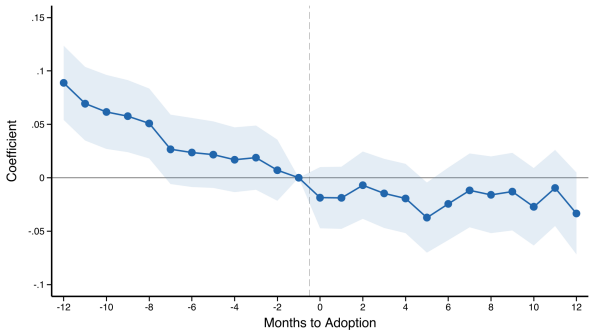
图2：GitHub Copilot 采用前后技术劳动力市场结果 n



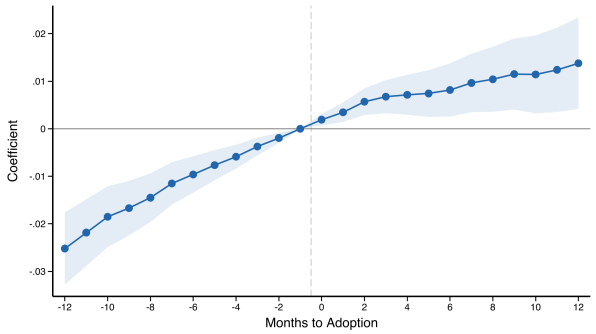
(a) 技术岗位招聘总量



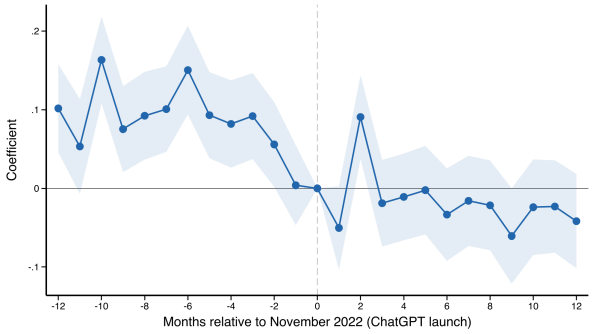
(b) 按最终采用情况划分的技术岗位招聘



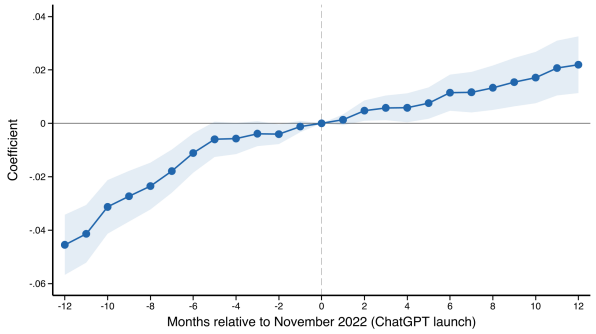
(c) 技术招聘人数对数（以采用为中心）



(d) 技术就业人数对数（以采用为中心）



(e) 技术招聘人数对数（ChatGPT 发布时点）



(f) 技术就业人数的对数（ChatGPT发布）

注释：A面板和B面板绘制了2018年1月至2025年6月经季节性调整的月度招聘率，计算方式为招聘人数除以滞后一期的就业人数。A面板比较所有职位与技术岗位；B面板比较最终采用者与未采用者。虚线标示 GitHub Copilot 于2023年1月发布企业版。C面板和D面板绘制了根据采用时间居中处理的事件研究系数及95%置信区间，依据公式（1）。E面板和F面板绘制了根据2022年11月ChatGPT发布时点估算的事件研究结果，依据公式（2）。标准误在企业层面聚类。

接下来，我们考察第二个基准，该基准通过比较最终采用者与从未采用者在ChatGPT于2022年11月发布前后的情况，避免了内生采用日期问题

(t_0)：[7]

$$Y_{it} = \alpha_i + \lambda_t + \sum_{\substack{k=-12 \\ k\neq 0}}^{12} \delta_k \mathbf{1}\{t - t_0 = k\} \times \mathbf{1}\{\text{Adopter}_i\} + \varepsilon_{it} \tag{2}$$

其中 $\mathbf{1}\{\text{采用者}_i\}$ 对于在2025年6月之前采用 GitHub Copilot 的企业取值为 1，且 t_0 为2022年11月（关于2022年11月之后的双重差分汇总估计，见附录表 A.7 ）。

固定日历日期并不能消除采用者对2021—2022年科技周期的差异性暴露期。面板E和F重现了相同的相反前趋势以及符号相反的汇总估计（附录表 A.7）。由于采用者的招聘在2021—2022年繁荣期之前大幅攀升，随后才趋近未采用者水平（面板E），将2022年之后的招聘与前期相比，会得出相对下降10.1%的结果（第1列）。反过来，由于此前招聘激增扩大了采用者的员工规模，且之后没有发生裁员（第2列），其2022年之后的技术人员总数仍比2022年之前的基线高出6.7%（面板F和第3列）。因此，标准的观察性回归可能得出人工智能减少就业（以招聘流量衡量）或增加就业（以员工数量存量衡量）的结论。

4.2 工具变量 法：具有错开的采用时点的双重差分

为分离采用时点中的外生变异，我们使用 GitHub Copilot 发布后各企业首次预定的 GitHub合同审查时间作为工具变量，以识别 GitHub Copilot 的采用（R_i）。令 $\ell = t - R_i$ 表示相对于首次预定审查的事件时间。我们使用以下事件研究设定，估计进行合同审查对 GitHub Copilot 采用的第一阶段效应，以及对就业结果的简约式效应：

$$D_{it} = \alpha_i + \lambda_t + \sum_{\ell \in \mathcal{L}} \pi_\ell \mathbf{1}\{t - R_i = \ell\} + \varepsilon_{it} \tag{3}$$

$$Y_{it} = \alpha_i + \lambda_t + \sum_{\ell \in \mathcal{L}} \rho_\ell \mathbf{1}\{t - R_i = \ell\} + u_{it} \tag{4}$$

其中， D_{it} 是一个吸收性指标，当企业 i 持有的 GitHub Copilot 许可证数量为正时取值为一；Y_{it} 是就业结果，例如员工人数、招聘或离职。对于在企业-月份层面可能为零的招聘和离职，我们分别考察广延边际（取值为正的指标变量）和密集边际（在取值为正的条件下，结果的对数）（Chen 和 Roth，2024年）。我们使用平衡月度面板数据估计企业整体和宽泛职业类别的设定；对于更细分的职位-资历分析和企业内部分析，由于单个单元中的月度招聘数据较为稀疏，我们将数据汇总至季度层面。

在两个方程中，企业固定效应 α_i 吸收企业间不随时间变化的差异，例如基线规模和行业，而日历期固定效应 λ_t 吸收共同的宏观经济冲击和季节性冲击。系数 π_ℓ 和 ρ_ℓ 衡量合同审查对 GitHub Copilot 采用的影响

[7]这是企业层面与之对应的回归：这些回归以任务暴露度或职业暴露度作为采用的代理变量，比较 ChatGPT 发布前后的就业动态。

以及时间跨度 ℓ 相对于尚未到达预定审查日期的企业的就业影响。由于合同审查在时间上是交错进行的，当处理效应在不同队列之间或随时间变化时，传统的双向固定效应估计量可能会产生偏差（Miyaji、2024年*b*）。因此，我们使用Borusyak、Jaravel 和 Spiess的插补估计量来估计方程（3）和（4），该估计量对处理效应异质性具有稳健性。合同审查安排在 2023年之后的队列作为我们的最后处理对照组，仅在其审查日期之前进入样本（R_i）。所有回归均未加权，因此代表企业层面的影响，而非按就业人数加权的影响。标准误在企业层面进行聚类。

在就业结果满足平行趋势和无预期效应的前提下，简化式系数ρ_ℓ 识别了达到预定 GitHub合同审查对就业产生的意向处理效应（ITT）。平行趋势要求，在不存在合同审查的情况下，就业轨迹不会随企业预定的审查日期而系统性变化。无预期效应要求企业不会在临近合同审查之前提前调整就业。即使不将合同审查解释为工具变量，简化式仍具有信息价值：它捕捉了达到预定的评估和采购 GitHub Copilot 机会对就业结果产生的意向处理效应。

要从意向处理效应转向采用 GitHub Copilot 本身的影响，我们使用特定时间跨度的 Wald比率，根据第一阶段采纳反应对简化式估计进行缩放：

$$\widehat{\theta}_\ell = \frac{\widehat{\rho}_\ell}{\widehat{\pi}_\ell} \tag{5}$$

我们使用包含 1,000 次抽样的企业区块自助法构造 Wald 估计的 95% 百分位数置信区间，在同一子样本上重新估计两部分及其比率。

$\widehat{}$将 θ_ℓ 解释为因果处理效应，需要将在错期面板设定下调整后的标准工具变量假设（Miyaji, 2024a、b）。采用也必须满足平行趋势和无预期效应。[8] 合同审查必须在平均意义上提高采用率（相关性），不得延迟任何企业的采用（单调性），并且只能通过 GitHub Copilot 采用影响就业（排除限制）。最后，时间跨度为 ℓ时的就业结果取决于当前的采用状态，而不是此前暴露期的具体持续时间（无延续效应）（Miyaji，2024a）。在这些$\widehat{}$条件下，θ_ℓ 识别了时间跨度为 ℓ 时，服从者中对已处理者的局部平均处理效应（LATET）（Miyaji，2024a、b）。在我们的设定中，服从者是指这样的企业：预定审查降低了行政障碍后采用了 GitHub Copilot，但如果没有审查，即使处于非周期时点，也不会在当时采用。对于这些企业，θ_ℓ 捕捉了截至时间跨度为 ℓ 时已经采用 GitHub Copilot 相对于尚未采用所产生的影响。我们在第 4.3 、 4.4 和 6.1.3 节评估这些假设的合理性。

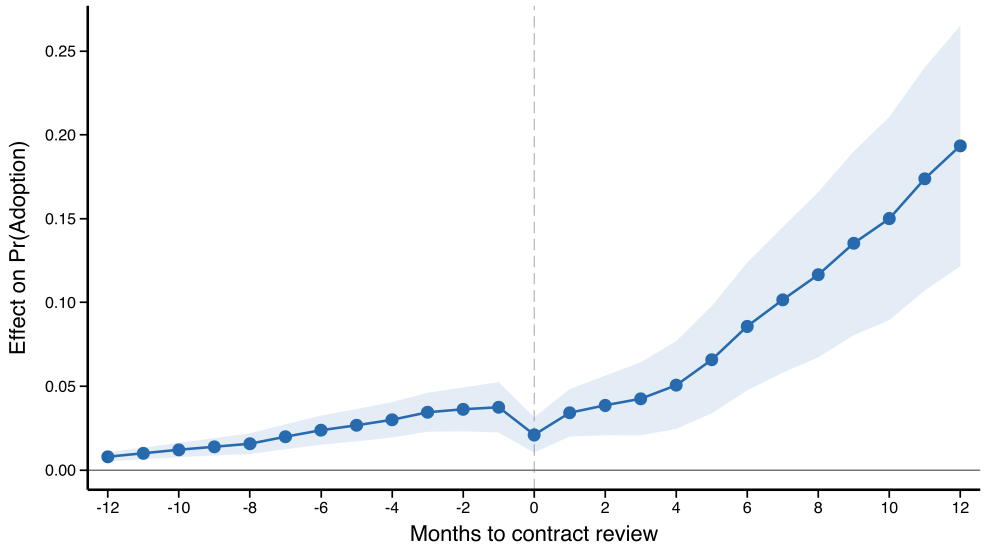
4.3 第一阶段：合同审查与 GitHub Copilot 采用

图3 绘制了方程（3）中的第一阶段事件研究（季度估计见附录表A.9）。预定合同审查持续提高 GitHub Copilot 的

[8]无预期效应并不排除企业在审查前采用；它要求企业不得因临近的审查而安排采用时间。

采用率；六个月后，采用概率上升 8.6 个百分点，十二个月后上升 19.3 个百分点。这两项显著增幅均高于样本期末 62% 的采用率。在四个季度的时间跨度上，第一阶段 F统计量为 30.7，远高于弱工具变量的常规阈值。[9] 现有 GitHub Enterprise 许可证数量的调整集中在合同审查月之后的两个月内，而 GitHub Copilot 采用的第一阶段效应则逐步增强。更新现有协议中的席位数量属于常规行政变更，而预定审查则会启动对 GitHub Copilot 这一新人工智能产品的评估；这需要安装软件并向员工推出。

图3：合同审查与 GitHub Copilot 采用



注：该图报告了方程 3中的事件时间系数。估计采用 Borusyak、Jaravel 和 Spiess (2024年) 的插补估计量，基于包含企业固定效应和日历月份固定效应的平衡企业-月份面板数据。阴影带表示95%置信区间；标准误按企业聚类。

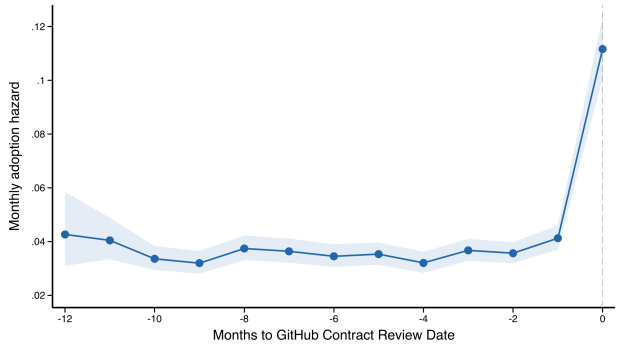
在标准的事件研究中，前期系数（$\ell < 0$）可用于检验预期效应。然而，在我们的情境中，图 3中预定合同审查前的上升趋势可能由两种不同的因素造成。首先，GitHub Copilot 于 2023年1月推出带来了机械性的样本构成效应。2023年1月之前，GitHub Copilot 尚未推出，因此不可能在非周期时间采用，采用率恒为零。然而，2023年1月之后，始终接受者可以在预定审查之前开始于非周期时间采用。因此，随着事件时间从 $\ell = -12$移至 $\ell = -1$，审查前的企业-月份观测值中，发生在 2023年1月之后的比例不断上升：在 $\ell = -12$时，几乎所有观测值都发生在 2022年，此时采用率在机械意义上为零；而到了 $\ell = -2$时，大多数观测值发生在 2023年，此时非周期采用可能不断增加。[10] 这种样本构成的变化本身就可能造成审查前提前项呈现上升趋势。其次，企业可能会

[9]在我们恰好识别且采用按企业聚类的标准误的设定中，第一阶段平方 t统计量等于 Montiel Olea 和 Pflueger (2013年) 的有效 F统计量；对于最坏情况下10%的偏差，其临界值为23.1。[10]这意味着，尽管样本中“始终接受者”的比例可能保持不变，但能够在周期外采用的比例会发生变化。

确实因为预期合同审查临近而以更高的比率采用 GitHub Copilot。由于基线的审查前提前项混合了推出前的零值（2022年12月及之前）与推出后的观测值（2023年1月及之后），前期系数无法直接区分预期效应。

为直接识别并检验预期效应，我们将样本限制在2023年1月之后的月份，此时所有企业都可以在合同审查周期之外采用。由于采用是一种吸收状态，图3 中的累计采用率会随时间累积此前的所有采用。因此，即使合同审查周期之外的采用率保持不变，随着审查日期临近，采用者的累计占比仍会逐渐上升。相比之下，月度采用风险率衡量的是某个月采用人数的流量：对于进入第 t个月时尚未采用的企业，该指标表示其中有多少企业在第 t个月首次采用 GitHub Copilot。若企业因预期合同审查临近而安排采用时间，那么审查前几个月的月度风险率应会上升。然而，在尚未采用的企业中，预定合同审查前各月的月度采用风险率持平在约4%，并在审查当月大幅跃升（$\ell = 0$）。这一模式表明没有预期效应的证据。使用控制了 GitHub Copilot 发布以来月份数的互补对数-对数模型估算风险比，也显示审查前的个人资料同样平稳（附录图A.6）。

图4：距离预定合同审查月份与 GitHub Copilot 采用风险率



注释：该图描绘了发布后期间尚未采用 GitHub Copilot 的企业风险集中，首次采用 GitHub Copilot 的月度风险率。第 0 月为企业预定合同审查所在月份，负值表示审查前的月份。阴影区域为 95% 威尔逊置信区间。

4.4 合同审查时间安排之间的协变量平衡

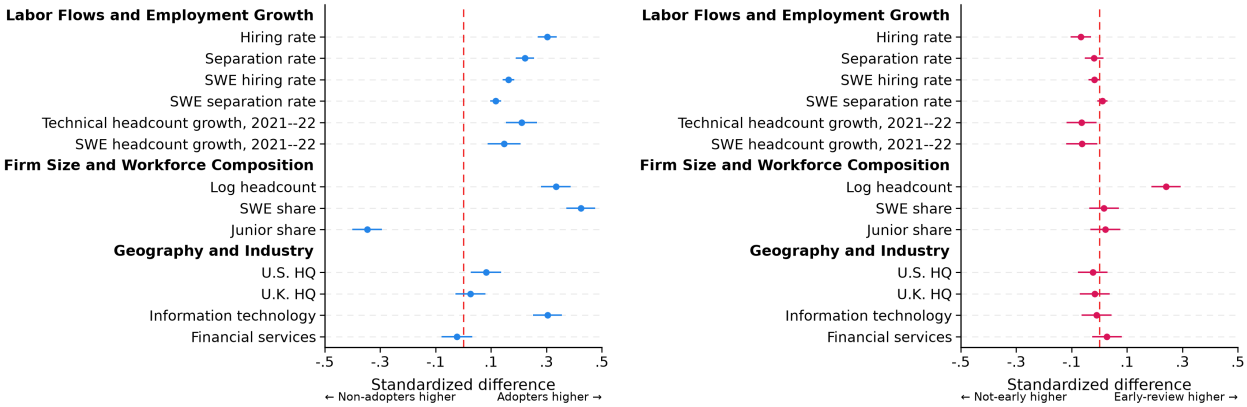
我们的识别策略要求：若没有合同审查，安排在不同日期接受审查的企业，其就业轨迹应当平行。虽然基线协变量平衡无法直接验证平行趋势，但有助于评估审查时间安排是否与发布前的企业特征或就业增长存在系统性关联。

图 5的面板A展示了不同采用状态之间的差异。与未采用者相比，最终采用者的规模明显更大，基线招聘率和离职率更高，雇佣的软件工程师占比更高、初级员工占比更低，并且更集中于信息技术行业。其

中，面板B将较早进行合同审查的企业与较晚进行审查的企业进行比较，并以2023年6月的合同审查月份中位数为界进行划分。在劳动力构成、招聘、离职以及行业指标方面，标准化差异均较小，处于协变量平衡通常采用的0.10阈值以内（Austin，2009年）。

例外是企业规模：较早接受合同审查的企业，其员工人数对数比接受较晚审查的企业高出0.24个标准差。这一模式反映了最初的 GitHub企业合同早在 GitHub Copilot 发布数年前就已签订。由于销售佣金随合同价值增加，企业销售团队在日历年初优先跟进规模最大的客户，而后续年度审查仍以最初的签约月份为准。在我们的基准设定中，企业固定效应吸收了企业间不随时间变化的规模差异，但无法排除不同企业规模对应的就业趋势存在差异。在第6.1.3节中，我们加入基准企业规模与季度交互固定效应以及行业与季度交互固定效应，并表明结果并未改变。

图5：基准企业特征的平衡性



（a）最终采用者与未采用者　　（b）较早与较晚接受预定合同审查

注释：面板A比较最终采用者与未采用者；面板B比较在2023年1月至6月接受审查的企业与其余样本。条形表示95%置信区间；标准误按企业聚类。附录表A.8 报告了底层回归结果。

5 概念框架

为了指导我们的实证分析，我们引入一个简单的会计框架，将 GitHub Copilot 对企业就业的影响分解为三个渠道：软件活动规模、不同职位和资历层级之间的任务分配，以及员工生产率。[11]

令 T_i 表示企业 i的软件相关任务总量。工人按职位 r 和资历 s分组为单元 (r,s) 。令 $m_{i,rs}$ 表示分配给单元 (r,s)的软件相关工作份额，其中 $\sum_{r,s} m_{i,rs} = 1$。最后，$a_{i,rs}$ 表示任务层面的生产率，定义为

[11]由于我们的实证设计捕捉的是企业短期调整，而非工资和就业的全经济均衡反应，因此我们采用会计恒等式，而非均衡模型。

该单元每单位劳动完成的任务数。根据这一结构，企业在特定职位—资历单元所需的劳动量为：

$$L_{i,rs} = \frac{m_{i,rs} T_i}{a_{i,rs}}. \tag{6}$$

对企业层面的 GitHub Copilot 采用 (Δ) 取对数并进行差分，可得：

$$\Delta \ln L_{i,rs} = \underbrace{\Delta \ln T_i}_{\text{scale}} + \underbrace{\Delta \ln m_{i,rs}}_{\text{task reallocation}} - \underbrace{\Delta \ln a_{i,rs}}_{\text{productivity}}. \tag{7}$$

方程（7）凸显了三种可能相互抵消的力量。首先，通过降低软件开发成本，GitHub Copilot 可能促使企业开展更多项目，扩大任务总量 ($\Delta \ln T_i > 0$)，并使所有单元的劳动力需求按比例增加。其次，GitHub Copilot 可能改变任务分配 ($\Delta \ln m_{i,rs}$)。例如，资深工程师可能使用 GitHub Copilot 承担此前委派给初级工程师的实施工作，或者开发人员可能接手此前分配给质量保证团队的测试任务。第三，GitHub Copilot 会改变任务生产率 ($\Delta \ln a_{i,rs}$)。[12] 更快完成分配的任务会减少完成给定工作量所需的劳动力，而花费额外时间审查和调试人工智能生成的代码，则会降低有效生产率并增加劳动力需求。

这一框架阐明了我们的实证估计能够识别和不能识别的内容。在数据中，我们观察到每个单元中的就业和招聘 ($L_{i,rs}$)，但无法观察企业软件工作的总量 (T_i)、任务如何在各单元之间分配 ($m_{i,rs}$)，或任务层面的生产率 ($a_{i,rs}$)。因此，对于任何单一群体，我们的估计都反映了三种渠道的综合净效应。然而，比较同一企业内各单元的就业响应，例如同一职位中初级与有经验的工程师之间的就业响应，可以差分掉企业层面的规模效应 ($\Delta \ln T_i$)，因为该效应对企业 i 的所有单元都是共同的。因此，企业内部不同群体之间的任何差异，都必须反映任务分配 ($\Delta \ln m_{i,rs}$) 或任务层面生产率 ($\Delta \ln a_{i,rs}$) 的差异化变化。最后，由于招聘决策具有前瞻性，我们的短期估计无法区分对已实现生产率提升的响应，与经理人因对人工智能能力的看法和预期而做出的调整。随着企业逐渐了解该工具的实际表现，以及项目规模和任务分配以不同速度进行调整，这些净就业影响可能会在更长的时间跨度内发生变化。

6 结果

6.1 对技术劳动力的总体影响

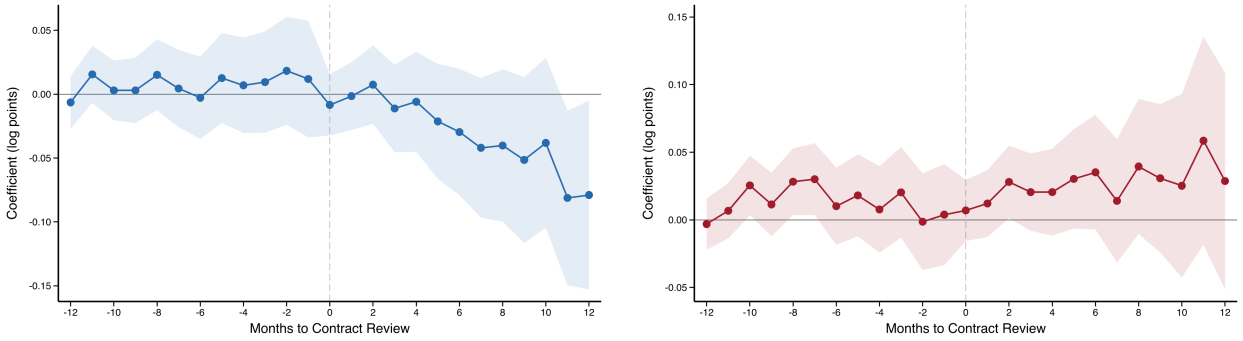
6.1.1 招聘、离职与员工人数

我们首先考察预定合同审查如何影响整体技术就业流动和存量。图 6展示了技术岗位招聘和离职的月度简约形式系数。面板A显示，预定合同审查后约五个月，招聘的简化式系数转为负值

[12]原则上，对任务层面的生产率的影响可能是正面的，也可能是负面的。虽然 GitHub Copilot 可能会加快代码生成，但有效生产率可能会下降（$\Delta \ln a_{i,rs} < 0$）。

。到第十二个月，与尚未到达预定审查时间的企业相比，技术岗位招聘减少了 7.6%（−0.079 对数点）（表 1，第1列）。面板B显示，离职并未出现相应增加；离职包括员工主动和非主动离职：到第十二个月，点估计为 2.9%，在统计上不显著（0。029对数点；第3列）。95%置信区间排除了幅度超过约12%的离职增加。完整的月度系数见附录表 A.10 。综上，这些结果表明，企业在合同审查后会通过放缓新招聘，而非增加离职，来调整其技术劳动力。

图6：对技术岗位招聘和离职的简化式效应



(a) 招聘　　(b) 离职

注释：该图报告了技术人员招聘人数对数和技术人员离职人数对数的简化式事件研究系数及95%置信区间。估计采用 Borusyak、Jaravel 和 Spiess (2024年) 的平衡面板插补估计量，并纳入企业和日历月份固定效应。标准误按企业聚类。

表1 的第（2）列和第（4）列显示，没有证据表明招聘或离职的广延边际发生相应变化（附录图A.7 绘制了完整的广延边际事件研究图）。平均而言，企业每月仍会招聘技术人员，但招聘人数有所减少。

我们发现，合同审查十二个月后，技术人员总数没有出现统计上可检测到的变化，点估计为 −0.016 对数点，即下降1.6%（表1第5列）。员工人数的点估计恰好与招聘放缓所隐含的结果一致。样本中的企业平均雇用823名技术人员，每月招聘17.7名技术人员。举例来说，将估计的招聘降幅7.6%应用于这一均值，同时假设招聘概率和离职人数保持不变，这意味着每月约少招聘1.35人，一年约少招聘16人。这一累计缺口将在一年后使技术人员总数减少约2%，与估计的1.6%降幅接近。如果这一招聘缺口持续下去，约两年后员工人数将出现约3.5%的差距；考虑到当前标准误，这一差距大致达到5%显著性水平所需的幅度，不过在更长的时间跨度上，效应及其精确度都可能发生变化。这些结果证实，由于每月招聘流量仅占员工总数的一小部分，即使招聘大幅下降，在一年时间跨度内对劳动力总规模的累计影响仍然有限。

表1：简化式　　合同审查十二个月后对技术劳动力的影响

因变量	技术岗位招聘		技术人员离职		技术人员总数的对数
	Log	Any	Log	Any	
	(1)	(2)	(3)	(4)	(5)
合同审查后第12个月	-0.079**	-0.010	0.029	0.011	-0.016
	(0.038)	(0.017)	(0.041)	(0.015)	(0.018)
企业固定效应	✓	✓	✓	✓	✓
月份固定效应	✓	✓	✓	✓	✓
趋势前联合 p	0.849	0.244	0.120	0.014	0.289
因变量均值	1.722	0.707	1.533	0.661	4.812
企业	5,339	5,414	5,357	5,414	5,414
观测值	149,565	221,346	143,324	221,346	221,298

注释：本表报告了附录表 A.10中月度事件研究的第12个月简化式系数，该系数采用 Borusyak、Jaravel 和 Spiess (2024年) 的插补估计量，在包含企业和日历月份固定效应的平衡企业-月份面板上进行估计。第（1）列和第（3）列分别为企业-月份正向流量样本中的技术岗位招聘人数对数和离职人数对数；第（2）列和第（4）列为是否存在上述流量的指标变量；第（5）列为技术人员总数的对数。趋势前联合 p值检验提前期系数是否联合为零。回归未加权。标准误按企业聚类，括号内为标准误。* $p < 0.10$，** $p < 0.05$，*** $p < 0.01$。

将这些约化形式估计值除以第一阶段估计值，可得出GitHub Copilot 采用对服从者招聘流量的隐含影响。第十二个月的Wald估计值表明，对于服从者而言，截至第十二个月已采用GitHub Copilot，会使月度技术岗位招聘减少约34%。在下文的细分分析中，我们采用季度设定；其第四季度Wald估计值为 -0.340 对数点，即季度技术岗位招聘下降29%，企业区块自助法得出的95%置信区间为 $[-0.948,\ 0.088]$ 对数点（附录表A.9 报告了完整的季度变化轨迹）。虽然各季度的特定时间跨度Wald比率大小相近，但其自助法区间较宽，因此无法根据这些估计判断效应是立即出现且保持恒定，还是随着暴露期延长而逐渐显现。我们的研究设计也无法区分这两种情况，因为要识别暴露期的持续时间，需要合同审查时间之外的额外工具变量。总体而言，虽然Wald比率的估计精确度低于约化形式估计值，因为它们叠加了两个阶段的抽样变异，但两组估计都表明，劳动力调整是通过减少招聘逐步实现的，而非通过员工离职或裁员。

6.1.2 招聘活动、职位空缺持续时间和发生率

实际招聘情况既反映企业希望填补的职位数量，也反映其填补空缺职位的速度。因此，招聘减少可能是劳动力需求下降所致，也可能是匹配过程变慢所致，例如人工智能采用促使企业寻找更稀缺、具备人工智能专门技能的候选人。为区分这两种解释，附录表A.11 报告了预定合同审查对技术岗位招聘活动三项季度指标的简化式效应：招聘信息、职位发布保持开放的时长，以及是否有任何职位发布活动。预定合同审查后四个季度，开放中的技术岗位发布的密集边际下降了 14%（-0.148 对数点）。相比之下，合同审查对是否存在任何

开放中的技术岗位发布，或职位发布开放的时长，均没有统计上可检测的影响，尽管职位发布时长存在一些前趋势。技术岗位招聘放缓似乎是由企业减少职位发布所致，而非填补现有空缺所需时间变长。

综合来看，就业流动和职位发布的结果表明，GitHub Copilot 的采用最初通过减少新招聘而非取代在职员工，减缓了技术劳动力的增长。尽管这一调整的直接影响落在潜在新员工身上，在职员工仍可能受到间接影响。随着采用企业的招聘放缓，外部职位空缺减少可能会缩小在职员工的外部选择，降低工作间流动，并削弱企业内部的工资议价能力（Caldwell 和 Danieli，2024年；Jäger 等人，2024年）。将调整集中在招聘边际上，可能符合包含准固定劳动成本的劳动力需求模型，也可能反映人工智能编码辅助降低了新手技能的价值（Oi，1962；Abowd、Corbel 和 Kramarz，1999）。

6.1.3 安慰剂与稳健性检验

我们检验三种替代性解释，以确认技术岗位招聘下降反映的是 GitHub Copilot 的采用，而非反复出现的企业预算周期、同期企业范围的招聘冻结或一般性人工智能采购，亦非企业规模和行业的差异化趋势。首先，预定合同审查在理论上可能与企业通常重设招聘目标的年度企业预算编制或财政规划日期重合。为直接检验这一点，我们进行发布前安慰剂检验，将每家企业的预定合同审查日期设定在 GitHub Copilot 的 2023 年发布之前。尽管这使发布后窗口仅涵盖 GitHub Copilot 发布前的两个季度，但若由预算驱动的周期性收缩是在合同审查时决定的，那么在此时间范围内就应开始体现在实际招聘中。然而，我们没有发现其对总体技术岗位招聘、离职，或软件工程初级及有经验岗位招聘的影响（附录表A.17）。此外，即使在 GitHub Copilot 发布之后，技术岗位招聘的下降也并未如年度预算重置所暗示的那样，在审查日期立即出现，而是在审查后的第三和第四季度逐渐显现，与 GitHub Copilot 许可证的推广进程相吻合。

其次，即使发布前的审查并未影响就业，企业在2023年进行审查，也可能恰逢2022年后更广泛的白领招聘冻结、全 IT 范围预算削减，或同时采购通用办公人工智能工具（尽管 GitHub Copilot 是 GitHub 的专属销售团队销售的唯一人工智能产品）。为检验企业整体冲击，我们考察了会计师的招聘情况。会计师是我们数据中最常见的非技术专业职业（98% 的样本企业雇用了会计师，发布前每家企业的中位数为34人），且既受到通用人工智能工具的影响，也受到更广泛的白领招聘放缓影响。附录图A.12显示，合同审查后，无论是会计师招聘人数的对数，还是招聘会计师的概率，均未下降。审查四个季度后，会计师招聘人数对数的简约式估计在统计上并不显著，为0.035个对数点。此外，正如我们在第6.2节所示，即使在技术劳动力内部，招聘放缓也并非普遍现象：代码密集度较低的技术岗位和资深工程师的招聘均未改变，这不支持存在广泛的技术岗位或 IT 部门招聘冻结。

第三，由于合同审查日期较早的企业发布前劳动力规模更大，如果大企业和小企业，或不同行业，在发布后的招聘轨迹各异，我们的估计就可能受到混杂因素影响。在附录表A.18中，我们重新估计各项设定，加入日历季度固定效应与发布前三分位企业规模的交互项、信息行业和专业服务业的指标变量，或同时加入两者。在所有设定中，技术岗位总招聘和离职的四季度简约式估计最多变动0.008个对数点（第1列和第2列），这证实企业规模或行业的差异化趋势并未驱动我们的研究发现。

6.2 按代码强度和资历划分的招聘构成 y

前述结果表明，GitHub Copilot 采用会减少整体技术岗位招聘。接下来，我们将考察这一降幅是均匀分布于技术劳动力各处，还是集中在特定职位和资历层级，从而改变新聘员工的构成。

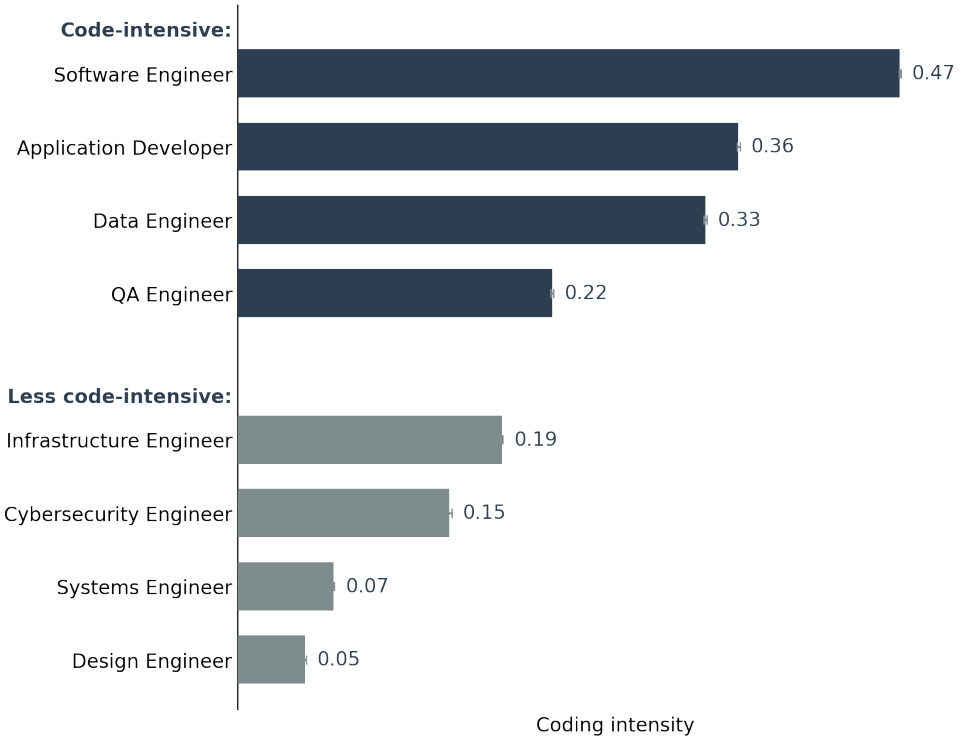
6.2.1 衡量技术工作的编码内容

我们将某个职位的代码强度定义为：在研究期间，该职位所描述的职责中，用于 GitHub Copilot 核心领域活动的比例，包括编写、补全、测试、记录、解释和调试代码。为衡量不同职位和资历层级的代码强度，我们采用样本企业在 2018年至2022年间、GitHub Copilot 发布前招聘的技术岗位的 LinkedIn 职位描述文本。[13] 不同于为宽泛职业代码指定单一平均值的职业层面任务度量，职位描述能够体现职责如何随更细分的技术岗位以及同一职位内部的资历层级而变化。附录图A.8 通过样本之外的案例展示了这种差异：一名软件工程师描述了如何构建作业调度平台并实现 API（面板A）；而一名测试自动化工程师则将 Python 和 LabVIEW 编程与液氮热交换器的实体设计和测试结合起来（面板B）。同样，在基础设施工程领域，一名技术负责人描述了制定团队战略和招聘工作，而一名技术负责人经理则兼顾工程管理和机器学习基础设施建设（C面板和D面板）。

我们使用 Claude Sonnet 4.6 对每段描述在这一维度上的得分进行评估（提示词和实现细节见附录第 B.10节）。模型仅查看自由文本描述，不会看到职位名称、职业、资历、地点、企业或入职日期。在我们的基线跨职位分析中，我们计算八类技术岗位中每类岗位描述得分的平均值。因此，代码强度是对任务与 GitHub Copilot 能力重叠程度的事前衡量，而非对实际编程时间或实际 GitHub Copilot 使用情况的衡量。

[13]由于 LinkedIn 用户可以追溯编辑职位描述，因此部分文本可能是在 GitHub Copilot 发布后撰写或修订的；我们的样本限制依据的是该职位的开始日期。

图 7：技术岗位的代码强度



注：该图报告了开始日期介于2018年和2022年之间的职位描述的平均代码强度得分。每份职位描述的权重相同。区间为基于职位描述层面变异计算的正态95%置信区间。得分最高的四个职位构成代码密集型群体。

图 7 展示了各技术岗位的平均代码强度。主要产出为软件代码的岗位得分最高，其中软件工程（47%）、应用开发（36%）、数据工程（33%）和质量保证工程（22%）位居前列。我们将这四类高于中位数的岗位称为代码密集型岗位。编程在基础设施工程（19%）、网络安全工程（15%）、系统工程（7%）和设计工程（5%）中的核心程度较低；在我们的样本中，这些岗位主要涵盖硬件、硅和产品设计领域，编程在其中辅助承担更广泛的运营、安全或实体工程职责。

6.2.2 按职位和资历划分的招聘反应

大量实证研究通过评估某一职业的任务与人工智能能力重叠的比例，来衡量职业层面的人工智能暴露度（Eloundou 等人，2024年；Felten、Raj 和 Seamans，2023年）。如果职业层面的任务重叠直接决定人工智能如何改变劳动力需求，那么招聘反应应随职位发布前的代码强度呈现系统性差异。然而，将某一职位的所有招聘合并分析，可能会掩盖不同资历层级之间的异质性：企业可能像应对周期性衰退那样，在各类技术岗位中一律减少初级员工招聘（Forsythe，2022年）；也可能是因为相较于资深职位所包含的编码任务，GitHub Copilot 更容易替代初级员工承担的编码任务。为评估职位编码强度和员工资历如何影响招聘，我们构建了按公司、岗位、资历和季度划分的面板数据，并估计：

$$Y_{irst} = \alpha_{irs} + \lambda_{rst} + \sum_{\ell} \left[\rho_{\ell} + \gamma_{\ell}\text{Junior}_s + \theta_{\ell} x_r + \beta_{\ell} \left(\text{Junior}_s \times x_r\right)\right] \mathbf{1}\{t - R_i = \ell\} + u_{irst} \tag{8}$$

其中， Y_{irst} 表示在招聘人数为正的单元中，公司 i、职位 r、资历组 s和季度 t的季度招聘人数对数。资历分为两组：如果属于初级职位（Revelio 资历层级 1 和 2，涵盖实习生、受训人员和职业生涯早期员工），则 Junior_s 取值为 1；如果属于资深职位（资历层级 3 及以上），则取值为 0。该设定纳入公司—职务—资历固定效应 α_{irs}以及职位—资历—季度固定效应 λ_{rst}。

表2 报告了这些设定下第四季度的估计值。第（1）列首先给出职位层级基准，该基准在公司—职位—季度面板数据上汇总不同资历的员工，并控制公司—职位和职位—季度固定效应。合同审查后四个季度，职位编码强度每增加一个标准差，招聘人数的变化在统计上并不显著 -0.023 个对数点（标准误 $= 0.022$）。如果与 GitHub Copilot 的能力在事前任务重叠足以概括劳动替代性，那么招聘降幅应该在编码密集型程度最高的职位中最为明显。然而，仅凭职位层级编码强度，无法可靠预测招聘会在哪些职位中下降。第（2）列检验所有技术岗位的初级员工招聘是否普遍减少，结果同样显示，初级员工与有经验员工之间的差异较小且在统计上不显著（-0.013 个对数点，标准误 $= 0.040$）。无论是职业层级编码强度还是员工资历，都无法单独预测招聘会在哪些职位中下降。

第（3）列将职位编码强度与资历进行交互，呈现出更明显的规律。在有经验员工的资深职位中，估计的编码强度梯度 θ_4 为0.015个对数点，在统计上并不显著。然而，在初级职位中，编码强度的斜率比有经验员工的斜率低0.086个对数点（$\beta_4 = -0.086$）。合并这两项后，职位编码强度每增加一个标准差，初级员工招聘降幅就比有经验员工大0.071个对数点。第（4）列用一个指标变量取代连续评分 x_r，该变量表示高于中位数的四个代码密集型职位：与代码密集度较低的职位相比，代码密集型职位中初级员工与有经验员工的招聘差异低0.227个对数点，由此推算的初级员工招聘高低组差异为 -0.180 个对数点。由于两个职位组的评分相差1.6个标准差，第（3）列的斜率意味着高低组差异约为0.14个对数点，与连续估计值非常接近。

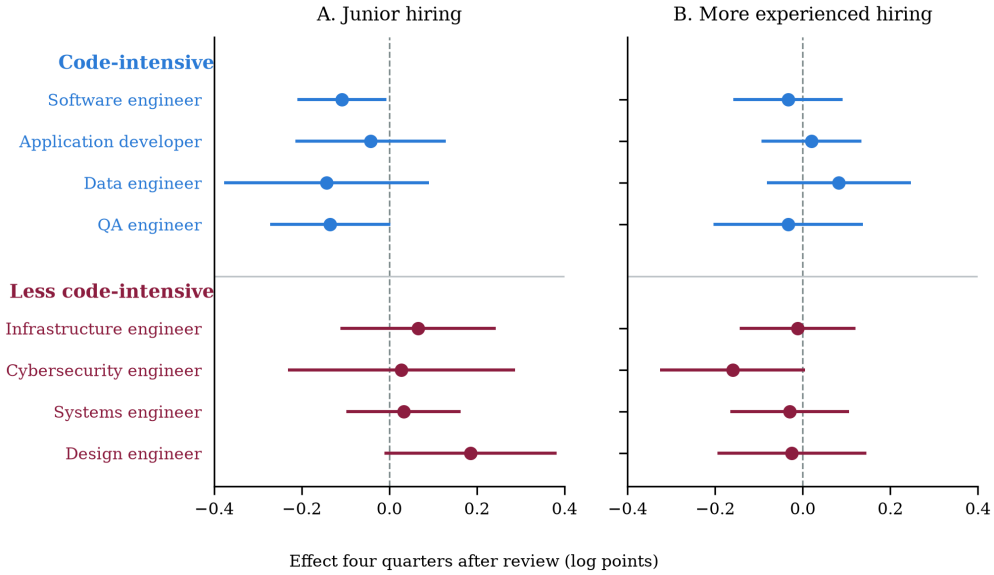
图 8 按职位代码强度排序，展示了16个职位—资历单元各自四个季度的简化式招聘估计。虽然各单元的估计值比合并回归更为嘈杂，但它们显示出一种贯穿整个职位分布的一致模式，而非由某个异常职业所致。面板A中，初级员工招聘的点估计在四个代码密集型职位中均为负，在四个代码密集度较低的职位中均为正；而面板B中，有经验员工招聘在不同职位间并未呈现类似的排序。

表2：按代码强度和资历划分的招聘效应

因变量	季度招聘人数对数			
	(1)	(2)	(3)	(4)
合同审查后的第4季度（ρ_4）	-0.012	-0.018	-0.022	-0.023
	(0.036)	(0.042)	(0.042)	(0.041)
代码强度（θ_4）	-0.023		0.015	0.048
	(0.022)		(0.023)	(0.041)
初级（γ_4）		-0.013	0.008	0.007
		(0.039)	(0.040)	(0.039)
初级 ×代码强度（β_4）			-0.086***	-0.227***
			(0.030)	(0.061)
观测单位	Role	职位 × 资历	职位 × 资历	职位 ×资历
企业 × 职位固定效应	✓			
职位 × 季度固定效应	✓			
企业 × 职位 × 资历固定效应		✓	✓	✓
职位 × 资历 ×季度固定效应		✓	✓	✓
推断的初级员工梯度（$\theta_4 + \beta_4$）			-0.071***	-0.180***
			(0.024)	(0.053)
企业	5,048	5,048	5,048	5,048
观测值	231,880	357,346	357,346	357,346

注：表中报告了预定合同审查四个季度后，对数季度招聘人数的简化式效应。第（1）列在职位—公司样本上估计方程 8 。第（2）至（4）列在共同的公司—职位—资历样本上估计方程 8 。第（4）列将其替换为一个指标变量，该变量表示高于中位数的四个职位，并以0.5为中心，因此其 θ_4 和 β_4 是高组减低组对比，其初级组主效应则在两组的中点处进行评估。初级职位为 Revelio 级别1–2，资历较深的职位为级别3–7。如果某企业在某职位的初级组和有经验组中，发布前平均员工人数均至少为1，则该企业纳入该职位的样本。标准误按企业聚类，并列于括号内。推断的初级员工梯度是上述估计值的线性组合 $\theta_4 + \beta_4$ ，其标准误根据估计协方差计算。 * $p < 0.10$， ** $p < 0.05$，*** $p < 0.01$ 。

图8：按职位和资历划分的招聘效应



注：该图报告了预定合同审查四个季度后各职位的简化式效应。面板 A 和 B 展示了初级职位（级别 1–2）和经验更丰富职位（级别 3–7）的招聘时段对数，估计样本为有正向招聘的公司—职位—季度。各职位按 GitHub Copilot 发布前的代码强度排序。水平误差线为基于按企业聚类的标准误计算的 95%置信区间。

这种差异在不同替代指标和公司内设定下均稳健。尽管代码密集型岗位中的初级职位所涉及的编码工作略多于同一职位中的资深职位（附录图A.9），但将职位层级得分 x_r 替换为按单元格划分的发布前编码强度得分 x_{rs} 后，估计结果几乎没有变化：初级员工梯度为 -0.062 对数点，资深员工梯度为 0.020 对数点，两者之差为 -0.082 对数点（附录表 A.12第 1 和第 2 列）。

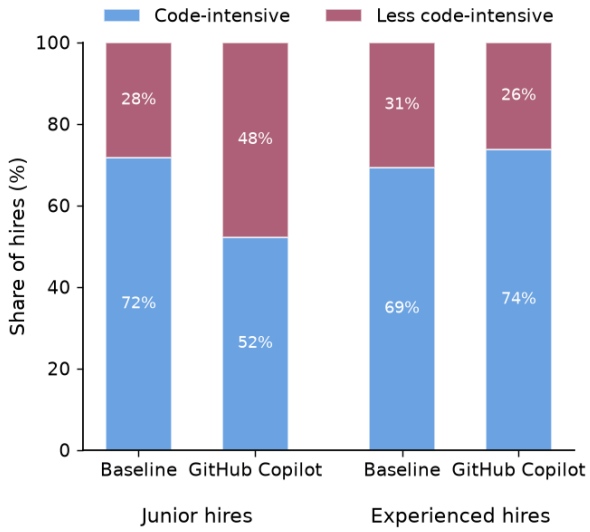
为剥离单个企业在某一职位内调整资历构成的方式，我们还使用公司内初级与资深员工招聘比率对数重新估计模型：对数(H^J_{irt}/H^E_{irt})，针对在职位 r 和季度 t同时招聘两类员工的企业，该指标消除了公司—职位—季度冲击的影响；这种冲击会在不同资历层级之间按比例调整招聘人数。合同审查后四个季度，职位编码强度增加一个标准差，会使企业的初级员工与资深员工招聘比率降低7.7%（-0.080 对数点，$p < 0.05$；附录表A.12第3列），而四个代码密集型岗位的指标变量会使该比率降低15.5%（-0.169 对数点，$p < 0.05$；附录图A.10）。

在各岗位组中，合同审查不会影响季度招聘为正的概率，也不会影响同一季度同时招聘初级和有经验员工的概率，这证实我们的集约边际和比率估计并非由广延边际选择所驱动（附录表A.13）。最后，与第 6.1.3 节一样，这些按职位和资历划分的估计没有显示出发布前安慰剂效应；加入企业规模与季度交互固定效应以及行业与季度交互固定效应后，估计结果也没有变化（附录表 A.17 和A.18）。

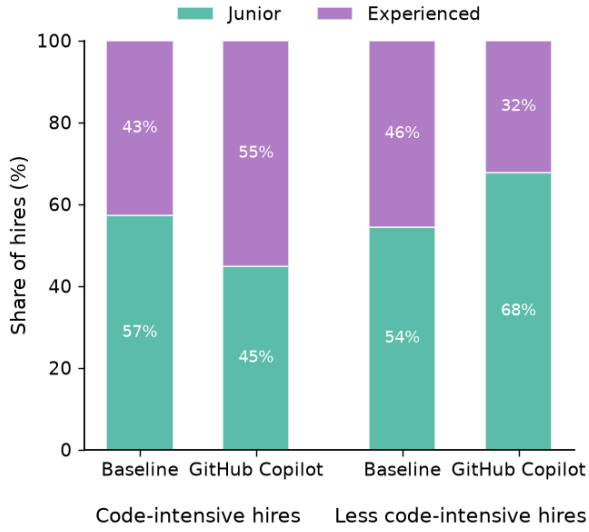
6.2.3 隐含幅度与构成变化

由于招聘收缩集中在代码密集型岗位中的初级职位，GitHub Copilot 采用重塑了初级员工的入职路径和技术团队的资历结构（图9）。在基线情况下，所有初级技术岗位招聘中有72%来自代码密集型岗位；随着 GitHub Copilot 采用，这一比例降至52%，使入门级招聘在代码密集度较高和较低的岗位之间几乎均分，而有经验员工招聘的岗位构成基本不变（代码密集型岗位占比从69%升至74%；面板A）。在各岗位组内部观察，代码密集型岗位招聘中的初级员工占比从57%降至45%，而代码密集度较低岗位中的初级员工占比从54%升至68%（面板B）。从招聘人数看（面板C），代码密集型岗位的初级招聘每企业季度减少约7个工作经历段，而这些岗位的有经验员工招聘基本持平。直接比较编码强度分布的两端，可清楚看到这种分化（附录表 A.14；月度简化式事件研究见附录图 A.11 ）。GitHub Copilot 采用使季度初级软件工程师招聘减少约40%（-0.513对数点），而有经验的软件工程师（-0.160对数点）或系统与基础设施工程师（0.212对数点）的招聘均未出现可检测的下降。

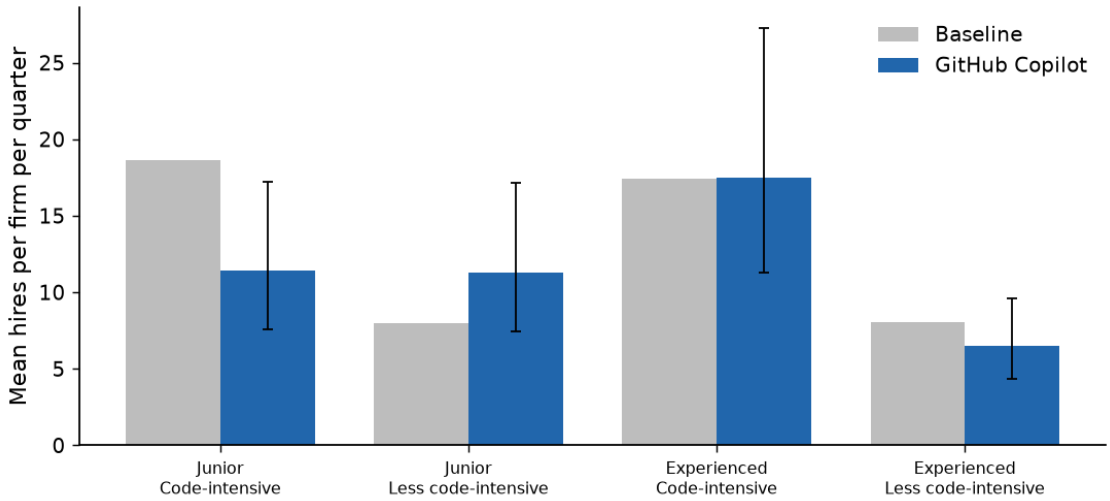
Fi图9：按代码强度和资历划分的推算招聘人数与构成



（a）按代码强度划分的构成（资历组内）



（b）按资历划分的构成（代码强度组内）



（c）按代码强度和资历划分的每家企业季度招聘人数

注释：基线值为2022年第四季度经过筛选的企业—职位配对中，每家样本企业平均每季度招聘任期数。带“GitHub Copilot”的数值是将各基线均值乘以$\exp(\hat{\rho}_g/\hat{\pi})$得到的，其中$\hat{\rho}_g$是组别$g$在表2第（4）列中评估后四个季度的简化式估计，而$\hat{\pi}$是同一时间跨度下的全样本第一阶段估计。面板A报告初级员工招聘和有经验的员工招聘中代码密集型职位所占比例；面板B报告代码密集型职位招聘和代码密集度较低职位招聘中的初级份额。面板C中的误差线表示$\hat{\rho}_g$的95%置信区间，并按相同方式缩放；这些区间未反映第一阶段的不确定性。面板A和面板B未显示区间。

从我们的概念框架来看，软件生产规模的普遍变化无法解释为何同一职位中初级员工与有经验员工的招聘趋势出现分化。相反，这种分化反映了不同资历层级在任务层面的生产率提升、工作捆绑或任务分配方面存在差异。首先，GitHub Copilot 可能在通常分配给初级员工的编码任务上，带来更大的直接节省劳动力的生产率提升。其次，初级职位与资深职位

在编码与其他职责的捆绑方式上有所不同，这也解释了为何即使在职位—资历单元层面衡量编码强度，资历差距依然存在（Garicano、Li 和 Wu，2026）。在代码密集型岗位中，初级职位的工作重心是离散的实施任务，这些任务可以独立地以相对较低的成本进行说明、委派和评估；而资深职位则将代码生产与系统架构、领域判断以及跨团队问责结合起来。由于在初级职位中，编码与工作的其余部分更容易分离，即使 GitHub Copilot 提高了两组员工的生产率，它也可能更容易替代初级员工的招聘。

与此同时，有经验员工招聘的编码强度梯度呈平缓趋势，并不意味着 GitHub Copilot 对有经验员工的工作毫无影响，或只是对有经验劳动力起到补充作用。即使 GitHub Copilot 能帮助有经验的工程师节省完成现有编码任务的时间，企业也可能扩大分配给他们的工作范围，或调整他们的时间分配。有经验的工程师可能会借助 GitHub Copilot 完成过去交由初级队友负责的实施工作，也可能在核心架构和协调职责之外，花更多时间审查、调试并整合人工智能生成的代码。这些相互抵消的力量——既有编码任务上的劳动力节省、从初级职位转移而来的任务，以及新增的验证需求——可能使有经验员工的净招聘量在不同编码强度分布中大致保持平稳，即便 GitHub Copilot 正在重塑他们的日常工作。

6.3 软件工程师中的经验、资历与内部招聘

尽管企业在采用 GitHub Copilot 后维持了有经验的软件工程岗位招聘规模，但稳定的招聘量可能掩盖企业招聘的工程师类型发生的变化。接下来，我们将考察在软件工程这一样本中规模最大、编码强度最高的职业中，企业是否会调整所招聘工人的教育资历、既往经验和内部招聘情况。基线时期，有经验的软件工程师获得学士学位后的平均年限为 8.7 年，而初级工程师为 4.8 年；前者拥有学术型硕士或博士学位的可能性也显著更高（附录表 A.15）。

原则上，人工智能编码工具既可能替代，也可能补充积累的一般人力资本和企业专用人力资本。如果 GitHub Copilot 能让员工在工作中轻松获取编程知识和代码库规范，企业在填补职位空缺时可能会放宽对学位、经验或内部任职年限的要求。反过来，如果更深入的技术培训、行业经验或对企业架构的熟悉程度，能让工程师更有效地指导人工智能工具、评估生成的代码并协调系统设计，那么企业可能会提高对新聘员工这些资质的要求。

Tabl表 3：合同审查、资质与软件工程师招聘来源 es

因变量	初级软件工程师 (1)	资深软件工程师 (2)
A. 资历与经验		
研究生学位占比	-1.83	4.69**
	(2.23)	(2.21)
占比 5+ 获得学士学位后的年数	-3.32	4.29*
	(3.05)	(2.39)
获得学士学位后的年数	-0.01	1.28**
	(0.40)	(0.51)
B. 按招聘来源划分		
内部填补	-0.103	0.111
	(0.085)	(0.085)
外部填补	-0.096*	-0.140**
	(0.058)	(0.067)
企业固定效应	✓	✓
季度固定效应	✓	✓

注释：表中报告了合同审查影响的约化形式估计值，并合并了第3和第4个事件季度。初级软件工程师是第1列中的1–2级，有经验的软件工程师是第2列中的3–7级。占比结果以百分点表示，平均经验以年表示，招聘结果以对数点表示。研究生学位包括学术型硕士学位和博士学位。内部填补是指上一年在同一企业任职的工人新填补的职位任期。估计采用 Borusyak、Jaravel 和 Spiess (2024年) 插补估计量；括号内为按企业聚类的标准误。 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ 。

表3 的面板A通过估计合同审查的简化式效应来检验这些预测，分析中将审查后的第三和第四季度合并，考察软件工程岗位招聘人员的教育背景和经验。对于初级软件工程岗位（第2列），尽管招聘总量大幅下降，我们并未发现受聘工人的教育资历或既往经验出现统计上显著的变化。相比之下，在有经验的职位（第2列）中，总招聘量保持稳定，却掩盖了招聘对象向一般人力资本更高的候选人转变这一趋势。拥有研究生学位的有经验员工招聘占比上升了4.7个百分点，而获得学士学位至少五年的人才占比也以相近幅度上升。平均而言，新聘的有经验软件工程师从获得学士学位起的潜在工作经验多出1.28年。企业并未降低资格门槛，反而提高了对有经验的软件工程岗位的要求。

面板B考察企业是否也会转向招聘具备企业专用人力资本、或其能力为企业所熟知的内部候选人（Becker，1962；Li、Raymond 和 Bergman，2026）。若员工在前一年曾在同一家企业担任某个职位（例如实习或此前的全职岗位），我们便将该职位填补界定为内部填补；否则则界定为外部填补。在有经验的软件工程岗位中，外部新聘人员减少约13%，而内部填补没有出现统计显著变化。在初级职位中，内部和外部填补均以相近幅度下降，没有统计上可区分的转向内部招聘的变化。

综合面板A和面板B可见，企业调整软件工程劳动力资历分布的方式并不对称。企业减少初级软件工程岗位招聘人数，同时不改变其可观察到的资质；而有经验的员工招聘规模保持不变，其教育背景、既往经验和内部任职年限则有所提升。尽管在实验环境中，人工智能编程助手可能大幅提升新手的任务层面生产率，但在实践中，它们可能与积累的一般人力资本和企业专用人力资本相辅相成。

6.4 技术岗位招聘的地域构成

除了改变新聘员工的职位、资历和学历之外，人工智能编程助手还可能重塑企业安排技术工作的地点。我们样本中的企业既雇用国内技术人员，也雇用离岸技术人员，其中印度是最常见的离岸地点。2021—2022年，约58%的企业至少雇用了一名在印度工作的技术人员，而83%的企业在印度至少雇用了一名从事任何职业的工人。这种集中反映了印度作为全球领先离岸软件开发中心的崛起。自21世纪初起，印度发展出了规模庞大的出口导向型软件服务业，其发展得益于数量可观的英语流利工程师、实力雄厚的理工大学，以及远低于美国和欧洲的劳动力成本（Aroraand Gambardella，2005年；Parthasarathy，2013年）。

原则上，人工智能编程助手可以通过多种渠道使离岸工人获得不成比例的更大收益。首先，大语言模型能够促进跨语言障碍的沟通，降低地理分散的工程团队所产生的协调成本（Brynjolfsson, Li andRaymond，2025年）。其次，编程助手可以帮助工人解读技术规范、浏览不熟悉的代码库，并理解内部系统。对于离岸工程师而言，这些能力可能尤其宝贵，因为与企业总部的同事相比，他们往往较难接触到非正式知识和可供同事利用的制度背景。第三，鉴于国内工程师与离岸工程师之间存在显著的工资差异，能够减少这些生产率劣势和协调劣势的人工智能工具，可能会使离岸劳动力成为国内劳动力更具吸引力的替代品。

另一方面，可以在境外完成的任务可能更容易被编纂，因此也更容易受到自动化的影响。此外，企业可能更愿意减少总部所在国之外地区的招聘，因为在这些地区，声誉顾虑、社会压力和信息摩擦对就业调整的约束可能更少（Landier、Nair 和 Wulf，2009年；Bassanini、Brunello 和 Caroli，2017年）。

为考察这一边际，附录表A.16 估计了合同审查对企业总部所在国（“国内”）和印度技术岗位招聘的影响，样本限定为前期在印度设有业务的4,509家企业。简约式估计的精确度较低，部分原因在于样本规模较小且海外招聘更加稀疏，因此无法揭示不同地点之间在招聘反应上的统计显著差异。合同审查四个季度后，印度本地技术岗位招聘的点估计为 -0.060 对数点（标准误 $= 0.189$），相比之下，国内技术岗位招聘为-0.029 对数点。尽管这些估计的精确度不足，无法可靠地检验不同地点之间的差异，但没有证据表明在采用 GitHub Copilot 后，技术岗位招聘转向离岸招聘

。如果一定要说有什么趋势，点估计反而表明印度的招聘放缓幅度更大。

6.5 GitHub Copilot 对技术岗位招聘总体下降的贡献

最后，我们将处理效应估计与 2022 年之后技术人才招聘的整体收缩幅度进行比较。2022 年至 2024 年（数据中最后一个完整日历年）期间，样本企业的平均季度技术岗位招聘人数下降了 29%，即每个企业每季度减少 16.1 名（从 55.3 名降至 39.2 名）。为评估这一总体放缓有多少可归因于采用 GitHub Copilot，我们利用处理效应估计来推算反事实后期招聘。设 H_{i0} 和 H_{i1} 分别表示企业 i 在 2022 年和 2024 年的平均季度技术岗位招聘人数，$H_{i1}(0)$ 表示企业 i 在未采用 GitHub Copilot 的情况下于 2024 年本会招聘的人数。我们将每家企业平均季度招聘人数的实际降幅 Δ 分解为两部分：一部分归因于采用 GitHub Copilot，另一部分为反映其他因素的剩余部分：

$$\Delta = \frac{1}{N}\sum_i \big(H_{i0} - H_{i1}\big) = \underbrace{\frac{1}{N}\sum_i \big(H_{i1}(0) - H_{i1}\big)}_{\Delta_{\text{GitHub Copilot}}} + \underbrace{\frac{1}{N}\sum_i \big(H_{i0} - H_{i1}(0)\big)}_{\Delta_{\text{Other}}}. \tag{9}$$

对于未采用者，$H_{i1}(0) = H_{i1}$，因此只有采用者会计入 $\Delta_{\text{GitHub Copilot}}$。对于估计对数招聘效应为 β_{i1}的企业，若未采用，反事实招聘为 $H_{i1}(0) = H_{i1}e^{-\beta_{i1}}$，而其放弃的招聘量为 $H_{i1}(e^{-\beta_{i1}} - 1)$。我们评估各企业2023年合同审查后各事件季度中放弃的招聘量（$\ell = 0, \ldots\ldots, 4$），并对季度和企业取平均，得到$\hat{}\Delta_{\text{GitHub Copilot}}$。更多细节和注意事项，见附录第F 节。

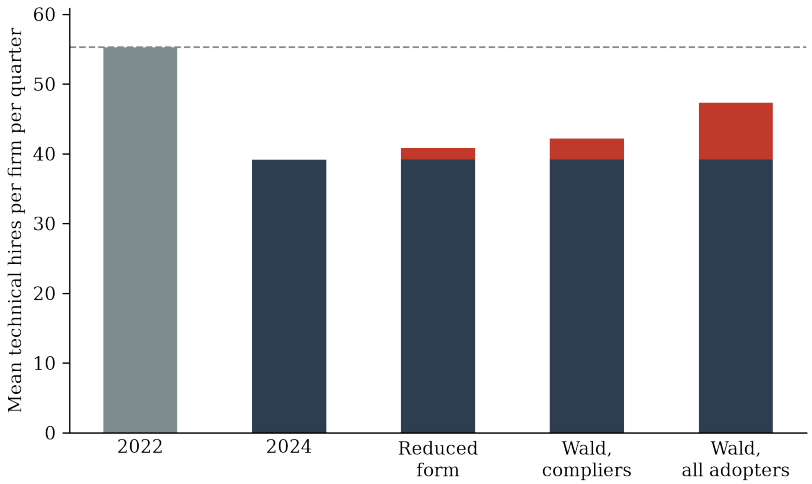
图10 展示了 $\Delta_{\text{GitHub Copilot}}$ 在三种逐步增强的假设下的结果，这些假设涉及估计效应适用于哪些企业：

1. 简化式反事实。 将简约式系数$\hat{}\rho$ 应用于所有企业，反映了接受预定合同审查的意向治疗效应。由此估算，每家企业每季度放弃招聘1.7人，占总体招聘总降幅的10%。

$\hat{}$ 2. *Wald*估计值、服从者反事实。将Wald估计值 θ 应用于采用企业，并按采用者中的服从者占比进行缩放，可将每家企业每季度少招聘的3名工人（占总体降幅的19%）归因于 GitHub Copilot。由于采用企业规模更大，且基线招聘的工人多于未采用者，因此这些估计值高于简化式基准。

3. *Wald*，所有采用者的反事实。 这采用了更强的假设，即所有采用者的招聘比例降幅及其变化路径都与服从者相同。由此估算，每家企业每季度减少招聘 8.1 人，相当于总体降幅的 51%。

图10：反事实招聘



注释：本图展示了一项反事实分析的结果，用于估算2022年与2024年之间季度平均技术岗位招聘总体下降的幅度中，有多少可归因于GitHub Copilot 采用。计算详情请参见附录第F节。

一个简单的粗略计算为这一第三项基准提供了依据：由于第四季度的处理效应与 2022年 至 2024年 总体招聘降幅（29%）的幅度相近，且到 2024年 时样本中约有一半已采用 GitHub Copilot，因此将服从者效应应用于所有采用者，可以解释总体招聘放缓的大约一半，另一半则由其他宏观经济因素驱动。

我们强调，这些计算只是说明性基准，并非精确的反事实，原因有四。第一，处理效应是按事件时间估算的，时间范围为各企业2023年合同审查后的一年（不同队列涵盖2023年第一季度至2024年第四季度）；而总量降幅 Δ 则以2022年和2024年日历年的差值衡量。因此，这一分解将审查后第一年的平均效应视为2024年后期的平均效应。^第二，$\hat{\rho}_\ell$ 和 θ_ℓ 是基于有正招聘的企业-季度观测，以未加权对数形式估算的，因此将其应用于企业招聘人数时，假设各企业的处理效应是同质的。第三，全体采用者基准（附录方程（12））没有明确的上界。由于始终接受者平均比服从者更早采用，他们可能最初部署的是较早、功能较弱的工具版本，也可能从人工智能采用的规模效应中获益更多，因而招聘降幅较小。另一方面，他们的招聘放缓幅度也可能更大。第四，这一计算属于局部均衡分析：它假设未采用企业的招聘不受其他企业采用人工智能的影响，而在均衡状态下，未被采用企业录用的部分候选人可能会被样本中的其他雇主吸纳。最后，我们的研究设计无法区分导致剩余二分之一至五分之四招聘放缓的具体因素，例如货币紧缩、疫情期间过度招聘的调整、软件研发税务处理的变化，或人工智能采用通过其他渠道产生的影响。

7 结论

随着人工智能系统能力不断增强，一个核心问题是企业将如何调整其员工队伍。哪些职位的需求会增加、减少或保持不变？企业会通过招聘边际进行调整、重新配置或裁撤在职员工，还是三者兼有？我们研究了采用 GitHub Copilot（早期的企业级人工智能编程助手）的大型技术密集型跨国企业在这些方面的调整，并利用既有企业软件合同审查时间的差异来识别因果效应。我们发现，劳动力市场最初的调整发生在招聘边际，而非通过离职实现。采用企业并未普遍削减招聘，而是减少了代码密集型岗位的入门级招聘，同时维持有经验员工的招聘规模，并提高了对其教育背景、既往经验和企业内部任职年限的要求。因此，人工智能编码辅助的早期影响与其说是广泛的岗位替代，不如说是入门级人才输送渠道收窄，同时对积累的人力资本需求增加。由此可见，即使生成式人工智能为新手带来了最大的任务层面生产率提升，这些个体收益也未必会转化为更大的劳动力需求。

我们的研究结果也引出了若干尚待解答的问题：人工智能对劳动力市场的影响将如何在不同企业之间以及随时间推移而演变。首先，我们的估计反映的是大型、成熟的跨国企业，它们运营着复杂的软件系统。面临较少协调约束的小型或年轻企业，或许能够更迅速地重组生产，或更灵活地通过降低软件开发成本来扩大产出和技术岗位招聘。其次，我们的样本期涵盖了人工智能编程助手的第一波发展，这类助手主要通过在开发者工作流程中自动补全代码来发挥作用。新一代能够自主执行多步骤工程任务的智能体系统，可能会加剧替代，或将替代效应进一步扩展至资历分布的更高层级。反过来，如果企业早期的调整部分源于超出实际生产率提升的预期，或者缩减初级员工群体最终导致资深人才供给出现瓶颈，那么企业的招聘策略可能会在更长的时间跨度内重新平衡。

最后，我们的企业层面估计并未涵盖工资调整或总就业。即使员工人数保持稳定，招聘放缓和技能需求的变化也可能改变薪酬以及企业与工人之间的剩余分配。此外，大型采用企业特定入职渠道的收缩，并不一定会一对一地转化为技术人员总就业的下降（Raymond，2023；Acemoglu、Lelarge 和 Restrepo，2020年）。原本可能加入这些企业担任初级软件工程师的工人，可能会转而进入系统工程或基础设施工程等相邻职位，或在规模较小或较晚采用技术的企业中担任软件工程岗位。另一方面，采用人工智能的企业可能会从其他企业手中夺取市场份额。确定这些企业层面的招聘和职业晋升阶梯变化如何重塑全市场就业、工资和长期劳动力需求，仍是未来研究的核心问题。

参考文献

Abowd, John M.、Patrick Corbel 和 Francis Kramarz。1999。《工人的进入与退出及就业增长：法国企业分析》。《经济学与统计学评论》，81(2)：170–187。Acemoglu, Daron 和 Pascual Restrepo。2020年。《机器人与工作岗位：来自美国劳动力市场的证据》。《政治经济学杂志》，128(6)：2188–2244。Acemoglu, Daron、Claire Lelarge 和 Pascual Restrepo。2020年。《与机器人竞争：来自法国企业层面的证据》。《美国经济学会论文与会议录》，110：383–388。Acemoglu, Daron、David Autor、Jonathon Hazell 和 Pascual Restrepo。2022年。《人工智能与工作岗位：来自在线职位空缺的证据》。《劳动经济学杂志》，40(S1)：S293–S340。Acemoglu, Daron、Gary Anderson、David Beede、Catherine Buffington、Eric Childress、Emin Dinlersoz、Lucia Foster、Nathan Goldschlag、John Haltiwanger、Zachary Kroff、Pascual Restrepo 和 Nikolas Zolas。2023年。《先进技术采用：选择效应还是因果效应？》。《美国经济学会论文与会议录》，113：210–214。Acemoglu, Daron、Philippe Aghion、Claire Lelarge、John Van Reenen 和 Fabrizio Zili-botti。2007。《技术、信息与企业分权》。《经济学季刊》，122(4)：1759–1799。Akerman, Anders、Ingvil Gaarder 和 Magne Mogstad。2015。《宽带互联网的技能互补性》。《经济学季刊》，130(4)：1781–1824。Ani, Zhamri Che、Zauridah Abdul Hamid 和 Nur Nazifa Zhamri。2024年。《GitHub Copilot 研究的近期趋势：系统综述》。355–366。新加坡：施普林格·自然。Anthropic。2025年。《Claude Code 及面向商业计划的新管理控制功能》。《Anthropic 新闻》，2025年8月，https: // www. anthropic. com/ news/ claude-code-on-team-and-enterprise 。Arora, Ashish 和 Alfonso Gambardella。2005年。《从弱者到猛虎：巴西、中国、印度、爱尔兰和以色列软件行业的崛起与发展》。牛津：牛津大学出版社。Athey, Susan 和 Scott Stern。2002。《信息技术对急诊医疗结果的影响》。《兰德经济学杂志》，33(3)：399–432。Austin, Peter C.。2009。《比较倾向得分匹配样本中处理组基线协变量分布的平衡性诊断》。《医学统计学》，28(25)：3083–3107。Autor, David 和 Neil Thompson。2025年。《专业知识》。《欧洲经济学会杂志》，23(4)：1203–1271。Autor, David H.、Frank Levy 和 Richard J. Murnane。2003。《近期技术变革的技能构成：一项实证探索》。《经济学季刊》，118(4)：1279–1333。Autor, David H.、Lawrence F. Katz 和 Alan B. Krueger。1998。《计算机与不平等：计算机改变了劳动力市场吗？》。《经济学季刊》，113(4)：1169–1213。Babina, Tania、Anastassia Fedyk、Alex He 和 J ames Hodson。2024年。《人工智能、企业增长与产品创新》。《金融经济学杂志》，151：103745。Baird, Matthew、Mar Carpanelli、Brian Xu 和 Kevin Xu。2026。《企业采用 GitHub Copilot 与软件工程师的劳动力市场结果》。《当代经济政策》，1–22。

Bartel, Ann、Casey Ichniowski 和 Kathryn Shaw。2007年。“信息技术如何影响生产率？产品创新、流程改进与工人技能的工厂层面比较。”《经济学季刊》122(4)：1721–1758。Bassanini, Andrea、Giorgio Brunello 和 Eve Caroli。2017年。“不在我的社区：社会压力与解雇地理分布。”《劳动经济学杂志》35(2)：429–483。Becker, Gary S. 1962年。“人力资本投资：理论分析。”《政治经济学杂志》70(5，第2部分)：9–49。Bessen, James、Maarten Goos、Anna Salomons 和 Wiljan van den Berge。2025年。“自动化企业中的工人会怎样？”《经济学与统计学评论》107(1)：125–141。Bloom, Nicholas、Luis Garicano、Raffaella Sadun 和 John Van Reenen。2014年。“信息技术与通信技术对企业组织的不同影响。”《管理科学》60(12)：2859–2885。Borusyak, Kirill、Xavier Jaravel 和 Jann Spiess。2024年。“重新审视事件研究设计：稳健且有效的估计。”《经济学研究评论》91(6)：3253–3285。Bresnahan, Timothy F.、Erik Brynjolfsson 和 Lorin M. Hitt。2002年。“信息技术、工作场所组织与熟练劳动力需求：企业层面证据。”《经济学季刊》117(1)：339–376。Brynjolfsson, Erik 和 Lorin M. Hitt。2000年。“超越计算：信息技术、组织转型与企业绩效。”《经济学展望杂志》14(4)：23–48。Brynjolfsson, Erik、Bharat Chandar 和 Ruyu Chen。2025年。“煤矿中的金丝雀？关于近期人工智能就业影响的六个事实。”斯坦福数字经济实验室工作论文。Brynjolfsson, Erik、Danielle Li 和 Lindsey Raymond。2025年。“工作中的生成式人工智能。”《经济学季刊》140(2)：889–942。Caldwell, Sydnee 和 Oren Danieli。2024年。“劳动力市场中的外部选择。”《经济学研究评论》91(6)：3286–3315。Carlsson, Mikael、Julián Messina 和 Oskar Nordström Skans。2021年。“企业层面冲击与劳动力流动。”《经济学杂志》131(634)：598–623。Chandar, Bharat 和 Bouke Klein Teeselink。2026年。“人工智能如何改变劳动力需求？来自41个国家的证据。”斯坦福数字经济实验室工作论文。Chen, Fiona 和 James Stratton。2026年。“企业中的人工智能：软件生产瓶颈。”求职市场论文。Chen, Jiafeng 和 Jonathan Roth。2024年。“含零值的对数？一些问题与解决方案。”《经济学季刊》139(2)：891–936。Cui, Kevin Zheyuan、Mert Demirer、Sonia Jaffe、Leon Musolff、Sida Peng 和 Tobias Salz。2026年。“生成式人工智能对高技能工作的影响：来自软件开发人员三项现场实验的证据。”《管理科学》。Cursor。2025年。“推出企业版 Cursor。”Cursor 博客，2025年10月，https: // cursor. com/ blog/ enterprise 。Daniotti, Simone、Johannes Wachs、Xiangnan Feng 和 Frank Neffke。2026年。“谁在使用人工智能编程？生成式人工智能的全球扩散与影响。”《科学》391(6787)：831–835。Dauth, Wolfgang、Sebastian Findeisen、Jens Suedekum 和 Nicole Woessner。2021年。“劳动力市场对机器人的调整。”《欧洲经济学会杂志》19(6)：3104–3153。

Davis, Steven J.、R. Jason Faberman 和 John Haltiwanger。2012年。“横截面与时间维度上的劳动力市场流动。”《货币经济学杂志》59(1)：1–18。Deming, David J. 和 Kadeem Noray。2020年。“收入动态、工作技能变化与科学、技术、工程和数学职业。”《经济学季刊》135(4)：1965–2005。Dey, Matthew、Emin Dinlersoz、Lucia Foster、Cheryl Grim、John Haltiwanger、Peter B. Meyer、Sabrina Wulff Pabilonia、Aditya Pande、Martha Stinson、Sean Wang 和 Zoltan Wolf。2026年。“美国企业的人工智能暴露与采用。”美国人口普查局、经济研究中心CES工作论文26-61。Dillon, Eleanor W.、Sonia Jaffe、Nicole Immorlica 和 Christopher T. Stanton。2025年。“生成式人工智能引发的工作模式转变。”美国国家经济研究局工作论文33795。Dohmke, Thomas。2024年。“通过 Anthropic 的 Claude 3.5 Sonnet、Google 的 Gemini 1.5 Pro 和 OpenAI 的 o1-preview，为 Copilot 带来开发者选择。”GitHub 博客，2024年10月29日。Eloundou, Tyna、Sam Manning、Pamela Mishkin 和 Daniel Rock。2024年。“GPT即GPT：大语言模型对劳动力市场的潜在影响。”《科学》384(6702)：1306–1308。Fairlie, Robert W. 和 Jane Wu。2026年。“人工智能对近期大学毕业生就业的早期影响。”美国国家经济研究局工作论文35796。Feigenbaum, James 和 Daniel P. Gross。2024年。“回应自动化的召唤：劳动力市场如何适应电话接线机械化。”《经济学季刊》139(3)：1879–1939。Felten, Edward W.、Manav Raj 和 Robert Seamans。2023年。“生成式人工智能暴露度的职业异质性。”SSRN工作论文4414065。Forsythe, Eliza。2022年。“经济衰退期间企业为何不雇用年轻工人？”《经济学杂志》132(645)：1765–1789。Friedman, Nat。2021年。“推出 GitHub Copilot：你的人工智能结对程序员。”GitHub 博客，2021年6月，https: // github. blog/ news-insights/ product-news/ introducing-github-copilot-ai-pair-programmer/ 。Fu, Yujia、Peng Liang、Amjed Tahir、Zengyang Li、Mojtaba Shahin、Jiaxin Yu 和 Jinfu Chen。2025年。“GitHub 项目中 Copilot生成的代码的安全弱点：一项实证研究。”《ACM软件工程与方法学汇刊》34(8)：1–34。Garicano, Luis 和 Esteban Rossi-Hansberg。2006年。“知识经济中的组织与不平等。”《经济学季刊》121(4)：1383–1435。Garicano, Luis 和 Esteban Rossi-Hansberg。2015年。“知识型层级：利用组织理解经济。”《经济学年度评论》7(1)：1–30。Garicano, Luis、Jin Li 和 Yanhui Wu。2026年。《繁杂工作：人工智能无法触及的工作》。Upriver出版社。GitHub。2026a。“关于 GitHub企业版。”GitHub Enterprise Cloud文档，https: // docs. github. com/ en/ enterprise-cloud@ latest/ admin/ overview/ about-github-for-enterprises 。GitHub。2026b。“客户案例。”GitHub，https: // github. com/ customer-stories 。Goldin, Claudia 和 Lawrence F. Katz。2008年。《教育与技术之间的竞赛》。马萨诸塞州剑桥：哈佛大学出版社。Gonçalves, Márcio Kuroki、Cleidson R. B. de Souza 和 Victor M. González。2011年。“协作、信息搜寻与沟通：软件开发人员工作实践的观察性研究。”《通用计算机科学杂志》17(14)：1913–1930。

Guha, Neel, Julian Nyarko, Daniel E. Ho, Christopher Ré, Adam Chilton, Aditya Narayana, Alex Chohlas-Wood, Austin Peters, Brandon Waldon, Daniel N. Rock- more 等。2023年。《LegalBench：一个协作构建、用于衡量大语言模型法律推理能力的基准》。*arXiv* 预印本 *arXiv:2308.11462*。**Hampole, Menaka, Dimitris Papanikolaou, Lawrence D. W. Schmidt 和 Bryan Seeg- miller**。2025年。《人工智能与劳动力市场》。美国国家经济研究局工作论文 33509。**Hoffman, Mitchell, Lisa B. Kahn 和 Danielle Li**。2018年。《招聘裁量权》。《经济学季刊》，133(2)：765–800。**Hosseini Maasoum, Seyed Mahdi 和 Guy Lichtinger**。2025年。《生成式人工智能作为资历偏向型技术变革：来自美国简历和职位发布数据的证据》。SSRN 工作论文 5425555。**Humlum, Anders 和 Emilie Vestergaard**。2025年。《静水之下，暗流涌动：生成式人工智能引发的早期劳动力市场转型》。美国国家经济研究局工作论文 33777。**Ide, Enrique 和 Eduard Talamàs**。2025年。《知识经济中的人工智能》。《政治经济学杂志》。**Iscenko, Zanna 和 Fabien Curto Millet**。2026年。《寻找晋升阶梯：人工智能是否正在影响初级职位？》经济创新集团工作论文。**Jäger, Simon, Christopher Roth, Nina Roussille 和 Benjamin Schoefer**。2024年。《工人对外部选择的信念》。《经济学季刊》，139(3)：1505–1556。**Jiang, Shaokang 和 Michael Coblenz**。2024年。《对集成开发环境中自动补全的成本与效益分析》。*ACM* 软件工程会议论文集，1(FSE)：1284–1306。**Jimenez, Carlos E., John Yang, Alexander Wettig, Shunyu Yao, Kexin Pei, Ofir Press 和 Karthik Narasimhan**。2024年。《SWE-bench：语言模型能否解决真实世界的 GitHub 问题？》*arXiv* 预印本 *arXiv:2310.06770*。**Katz, Lawrence F. 和 Kevin M. Murphy**。1992年。《相对工资的变化，1963–1987：供需因素》。《经济学季刊》，107(1)：35–78。**Kharazian, Ara, Lisa Simon 和 Ryan Stevens**。2026。《重新审视人工智能对就业的影响：企业层面的人工智能支出与劳动力调整》。Ramp经济学实验室工作论文。**Klein Teeselink, Bouke**。2025年。《生成式人工智能与劳动力市场结果：来自英国的证据》。SSRN 工作论文 5516798。**Kumar, Sukrit, Drishti Goel, Thomas Zimmermann, Brian Houck, B. Ashok 和 Chetan Bansal**。2025年。《Time Warp：人工智能驱动时代开发人员理想工作周与实际工作周之间的差距》。*arXiv* 预印本 *arXiv:2502.15287*。**Landier, Augustin, Vinay B. Nair 和 Julie Wulf**。2009年。《亲近办公地点的权衡：企业决策与地理分散》。《金融研究评论》，22(3)：1119–1148。**Leroy, Sophie**。2009年。《为什么我的工作这么难做？在不同工作任务间切换时注意力残留带来的挑战》。《组织行为与人类决策过程》，109(2)：168–181。**Li, Danielle, Lindsey Raymond 和 Peter Bergman**。2026年。《探索性招聘》。《经济学研究评论》，93(2)：1200–1240。**Lodefalk, Magnus, Lydia Löthman, Michael Koch 和 Erik Engberg**。2026年。《同一场风暴，不同的航船：生成式人工智能与招聘中的年龄梯度》。厄勒布鲁大学商学院工作论文 2026:2。

Luong, Thang, Dawsen Hwang, Hoang H. Nguyen, Golnaz Ghiasi, Yuri Chervonyi, Insuk Seo, Junsu Kim, Garrett Bingham, Jonathan Lee, Swaroop Mishra 等。2025年。《迈向稳健的数学推理》。arXiv 预印本 arXiv:2511.01846。Maslej, Nestor, Loredana Fattorini, Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik Brynjolfsson, John Etchemendy 等。2025年。《2025年人工智能指数报告》。arXiv 预印本 arXiv:2504.07139。McKenna, Nick 和 Bartek Perz。2026年。《GitHub Copilot CLI 结合多个模型系列，提供第二意见》。GitHub 博客，2026年4月，https: // github. blog/ ai-and-ml/ github-copilot/ github-copilot-cli-combines-model-families-for-a-second-opinion/。Meyer, Andre N., Laura E. Barton, Gail C. Murphy, Thomas Zimmermann 和 Thomas Fritz。2017年。《开发人员的工作生活：活动、切换与感知生产力》。《IEEE软件工程汇刊》，43(12)：1178–1193。Michaels, Guy, Ashwini Natraj 和 John Van Reenen。2014年。《信息与通信技术是否使技能需求两极分化？来自11个国家、25年间的证据》。《经济学与统计学评论》，96(1)：60–77。Miyaji, Sho。2024a年。《具有异质处理效应的工具变量双重差分》. arXiv 预印本 arXiv:2405.12083。Miyaji, Sho。2024b年。《交错式双重差分工具变量设计中的双向固定效应工具变量回归》。arXiv 预印本 arXiv:2405.16467。Montiel Olea, José Luis 和 Carolin Pflueger。2013年。《弱工具变量的稳健检验》。《商业与经济统计学杂志》，31(3)：358–369。Murphy, Gail C., Mik Kersten 和 Leah Findlater。2006年。《Java软件开发人员如何使用 Eclipse 集成开发环境？》《IEEE 软件》，23(4)：76–83。Nadella, Satya。2025年。《微软2025财年第四季度财报电话会议》。微软投资者关系部，2025年7月。Nagle, Frank, Shane Greenstein, Maria P. Roche, Nataliya Langburd Wright 和 Sarah Mehta。2023年。《Copilot：微软与 GitHub 的生成式人工智能》。哈佛商学院案例 624-010。美国国家标准与技术研究院。2018年。《改进关键基础设施网络安全框架》。美国商务部，1.1版。Oi, Walter Y. 1962年。《作为准固定要素的劳动力》。《政治经济学杂志》，70(6)：538–555。Parthasarathy, Balaji。2013年。《印度离岸信息与通信技术服务供给特征的变化，1985–2010》。载于《牛津离岸业务与全球就业手册》，Ashok Bardhan、Dwight Jaffee 和 Cynthia Kroll 编，第380–403页。牛津大学出版社。Raymond, Lindsey。2023年。《算法的市场效应》。麻省理工学院工作论文。Rein, David, Betty Li Hou, Asa Cooper Stickland, Jackson Petty, Richard Yuanzhe Pang, Julien Dirani, Julian Michael 和 Samuel R. Bowman。2023年。《GPQA：研究生水平的谷歌免疫问答基准》。arXiv 预印本 arXiv:2311.12022。Revelio Labs。2026a年。《Revelio Labs 数据字典》。https: // www. data-dictionary. reveliolabs. com/ data. html。Revelio Labs。2026b年。《Revelio Labs 方法论》。Revelio Labs 数据字典，https: // www. data-dictionary. reveliolabs. com/ methodology. html。Rosen, Sherwin。1981年。《超级明星经济学》。《美国经济评论》，71(5)：845–858。

Shah、Anshul、Anya Chernova、Elena Tomson、Leo Porter、William G. Griswold 和 Adalbert Gerald Soosai Raj。2025 年。“ 学生使用 GitHub Copilot 处理大型代码库。” 第 1 卷，1050–1056。**Stack Overflow**。2023年。“Stack Overflow 开发者调查 2023。” *https://survey.stackoverflow.co/2023/*。**Taniguchi、Hiroya 和 Ken Yamada**。2022年。“信息通信技术资本—技能互补性与工资不平等：来自经合组织国家的证据。”《劳动经济学》，76：102151。**Tucker，Lee C**。2026。“你‘（不）被录用’：季度劳动力指标中的人工智能与职业生涯早期招聘。”美国人口普查局经济研究中心 CES工作论文 26-27。**美国劳工统计局**。2025年。“职业就业与工资——2024年5月。”美国劳工部经济新闻稿 USDL-25-0451，华盛顿特区。**Weill、Peter 和 Jeanne W. Ross**。2004年。《*IT*治理：顶尖企业如何管理*IT*决策权以取得卓越成果》。马萨诸塞州波士顿：哈佛商学院出版社。**Wessel、Mairieli、Tom Mens、Alexandre Decan 和 Pooya Rostami Mazrae**。2023年。“GitHub开发工作流自动化生态系统。” *arXiv* 预印本 *arXiv:2305.04772*。**Westby、Samuel、Alicia Sasser Modestino 和 Peiran Cheng**。2026。“生成式人工智能与入门级软件工作的重新定义。”IZA劳动经济研究所讨论论文 18723。**Wiggers，Kyle**。2024。“Amazon CodeWhisperer 现更名为 Q Developer，并正在扩展其功能。”*TechCrunch，2024*年*4*月*30*日。**Yotzov、Ivan、Jose Maria Barrero、Nicholas Bloom、Philip Bunn、Steven J. Davis、Kevin M. Foster、Aaron Jalca、Brent H. Meyer、Paul Mizen、Michael A. Navarrete、Pawel Smietanka、Gregory Thwaites 和 Ben Zhe Wang**。2026。“企业人工智能数据。”美国国家经济研究局工作论文 34836。**Zhao，Shuyin**。2023。“GitHub Copilot Chat 现已面向组织和个人正式推出。” *GitHub* 博客，*2023*年*12*月。**Zhao，Shuyin**。2024。“微调模型现已在 GitHub Copilot Enterprise 中限量公开测试。”GitHub 博客，2024年9月10日。

附录材料——在线发表

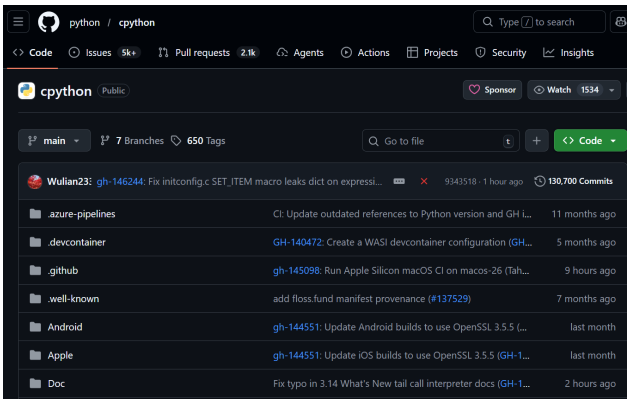
GitHub 与人工智能编码工具的背景

A.1 GitHub

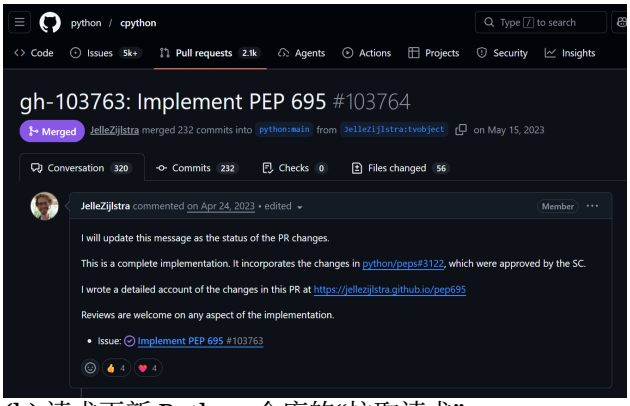
基于拉取的软件开发起源于开源社区。2005年，林纳斯·托瓦兹发布了 Git，这是一种分布式版本控制系统，使开发人员能够维护代码库的独立副本，并提出更改以供整合。三年后，汤姆·普雷斯顿-沃纳、克里斯·万斯特拉斯和 P.J. Hyett 在 Git 之上搭建了基于网络的协作层，推出了结构化的拉取请求工作流，开发人员通过该工作流提交、审查并合并提议的代码更改（附录图 A.1）。该平台的用户数量从2009年的10万人增长到2018年的逾3100万人。

在创立初期，GitHub 主要作为开源协作平台，为公共代码仓库提供免费托管服务。商业领域的采用源于开发人员自下而上的需求：随着软件工程师越来越多地使用个人 GitHub 账户处理工作项目，取代传统的内部工具，企业客户开始要求推出具备集中管理和安全控制功能的企业版，GitHub Enterprise 因此于 2011 年推出。在随后的十年间，拉取请求工作流成为大型组织进行商业软件开发的标准模式。微软于 2018年收购了 GitHub，但 GitHub 仍作为独立业务部门运营，拥有自己的企业销售团队和签约系统（Nagle 等人，2023年）。

图 A.1：GitHub 仓库与拉取请求工作流程



(a) Python 语言的公共仓库



(b) 请求更新 Python 仓库的“拉取请求”

注释：截图 展示了 Python 语言公共仓库的截图，以及一项提议修改的拉取请求 it.

A.2 人工智能编码工具

自动化代码辅助工具的出现早于大语言模型。微软的 IntelliSense 等基于规则的工具于 1997 年在 Visual Studio 中推出，可根据语法补全标识符并标记错误。到 2006 年，Java 开发人员已广泛使用代码补全工具，不过早期实证研究发现，这些工具主要减少了查阅文档的需求，并未显著加快代码编写速度（Murphy、Kersten 和 Findlater，2006；Jiang 和 Coblenz，2024年）。之后，Tabnine 等工具引入统计模型和早期深度学习模型，以提供简短的逐行补全。

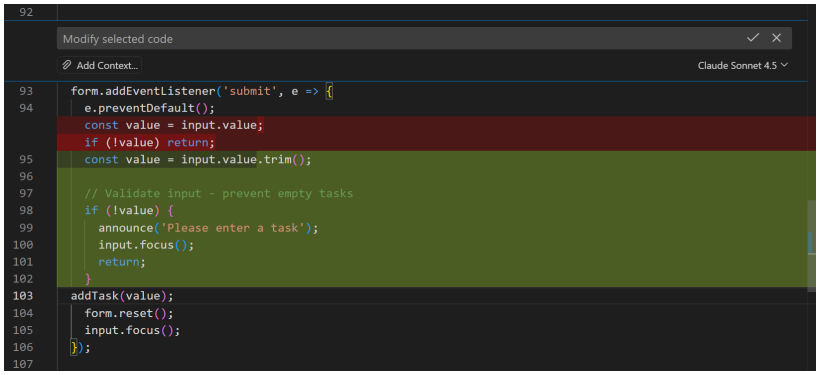
大语言模型的出现，标志着代码生成能力的一次重大飞跃。由于大语言模型会根据周围上下文预测后续词元，编程语言尤其适合其架构：代码遵循严格的语法和语义规则；大量源代码与自然语言文档配对的语料可公开用于训练；编译器和测试套件则提供了客观的验证信号。这些特点推动了能力的快速提升。在 SWE-bench（一项于2023年推出的评测基准）中，模型需要

解决12个 Python代码库中真实的 GitHub问题，前沿模型的解决率从发布时的不到5%上升至2026年初的近80%（Jimenez 等人，2024年；Maslej 等人，2025年）。

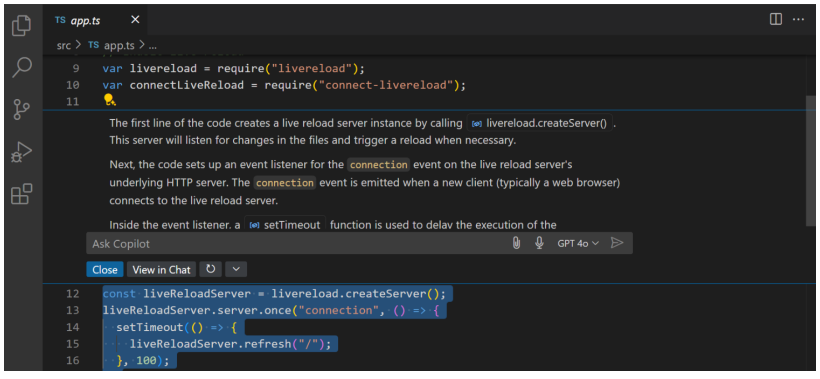
GitHub Copilot 于2021年6月首次以技术预览版形式发布（Friedman，2021年）。GitHub Copilot 个人版于2022年6月正式推出，随后在2023年初面向企业客户推出了 GitHub Copilot 商业版。开发人员迅速采用了这款工具：在2023年的一项行业调查中，56%的专业开发人员表示正在使用 GitHub Copilot，使用比例是排名第二的人工智能编程工具的四倍多（Stack Overflow，2023年）。GitHub Copilot 的初始版本由 OpenAI Codex 提供支持，该模型使用数十亿行公开源代码进行训练。后续版本扩展了产品功能，不再局限于行内自动补全，还加入了交互式聊天界面（附录图A.2），允许用户在多个前沿模型中进行选择（McKennaand Perz，2026），并允许企业客户为私有代码库建立索引，以便根据其内部代码库生成回答。

越来越多的研究开始考察 GitHub Copilot 对个体开发人员的影响；Ani、Hamid 和 Zhamri（2026）提供了早期综述。受控实验和实地研究评估了它对开发人员任务层面的速度和产出（Cui 等人，2026）、生成代码的安全性和质量（Fu 等人，2025年）以及计算机科学教育（Shah 等人，2025年）的影响。

图 A.2：GitHub Copilot 的功能



(a) 代码生成



(b) 聊天

注释：GitHub Copilot 行内代码生成和聊天界面的屏幕截图。

B Da构建、测量与验证 on

B.1 Revelio职业与资历测量

LinkedIn 上的工人常用不同的职位名称描述相似的工作，而相同的职位名称在不同企业和行业中可能对应不同的职业。Revelio 使用层次聚类算法对自由文本职位名称进行标准化，该算法综合考虑职位名称、工人和雇主的职位描述文本、工人的技能、同事以及既往经验（RevelioLabs, 2026*a*）。例如，“软件开发人员”“后端工程师”和“全栈开发人员”等职位名称会被归入共同的“软件工程”类别。我们采用 Revelio 的 150 类职位分类体系，这是公司年度分析所推荐的最细粒度分类（Revelio Labs，2026*b*）。

我们还使用Revelio对员工资历的测量。由于职位名称在不同企业和职业中所代表的资历层级各不相同，Revelio采用集成模型，根据个人的职位名称、公司和行业，以及过往就业经历的持续时间与资历和年龄，构建资历评分（Revelio Labs，2026*b*）。该测量结果涵盖从入门级职位到高级高管职位的范围。[14]

基于底层个人职位任期数据，我们构建了企业层面的就业存量、新入职和离职月度面板数据，以及公司—职务层面和公司—职务—资历层面的细分面板数据。

B.2 Revelio覆盖范围与测量

线上职业档案过度代表白领和技术类职业、较年轻的队列以及规模更大且更成熟的企业。在北美和欧洲，信息技术行业以及专业、科学与技术从业者的覆盖尤其全面（Baird等人，2026）。这些特征使Revelio非常适合追踪大型GitHub客户的技术劳动力。由于其他地区的覆盖较为有限，且覆盖范围随着平台采用而不断扩大，我们将样本限定为总部位于美洲和欧洲的企业。

除原始任职记录数量外，Revelio 还会调整数据，以纠正两种测量误差来源（Revelio Labs, 2026*b*）。首先，横截面抽样权重用于调整不同职业、行业和国家之间档案代表性的差异。其次，时间序列调整用于处理就业转换发生与员工更新其在线个人资料之间的报告滞后。这两项调整均依据监管申报文件中的企业层面员工人数基准进行校准。在整个分析中，我们使用这些经过调整的就业与流动序列。

[14]资历类别从入门级（例如，会计实习生、律师助理、软件工程师培训生）、初级（例如，初级软件质量保证工程师、法律顾问）、副职（例如，律师、高级税务会计师）、经理、总监、高管到高级高管（例如，首席执行官、首席财务官、首席运营官）（Revelio Labs，2026*b*）。

B.3 GitHub–Revelio 公司匹配流程

我们使用公司名称、国家、行业和规模指标，将 GitHub 行政记录与 Revelio Labs 的数据关联起来。首先，我们应用 Revelio 内部名称匹配算法：该算法会剔除常见的法律后缀（如“LLC”、“Inc.”和“GmbH”）及行政标签（如“母公司”），然后将名称与 Revelio 数据库中的标准化公司名称进行精确字符串匹配。该算法会排除记录员工人数为零的公司简介页面；如果匹配对象是企业子公司，则将其就业数据汇总至最终母公司。

这一初始字符串匹配步骤为约 64% 的 GitHub 账户生成了高置信度匹配结果。其余账户未被自动匹配，因为它们记录的名称较为通用、经过缩写，或在两个数据库中的格式不同。当企业在职业社交平台上使用不同的品牌名称或部门名称运营（例如“ABC Inc.”与“ABC Science”）时，当非英语字符或音译在不同系统中存在差异时，或者当 GitHub 许可证由一家子公司持有，而该子公司的法定名称在 Revelio 中并未作为独立实体出现时，通常会产生不一致。为解决这些情况，我们将人工智能辅助的匹配流程与人工核验相结合，利用企业所在国家、行业和员工规模来识别并验证相应的公司简介页面，具体内容见下一小节。

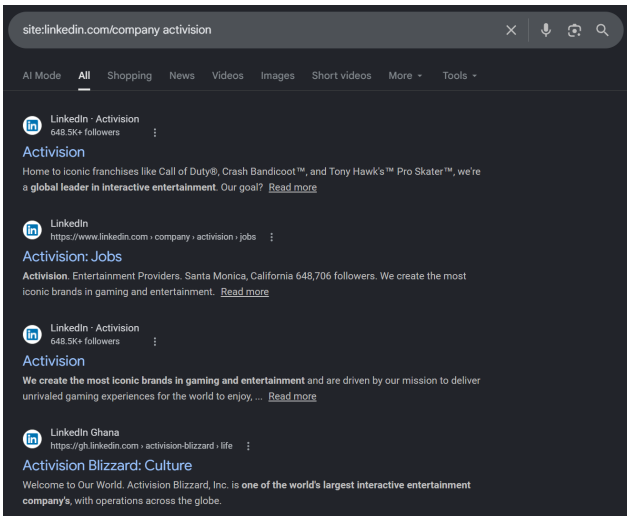
B.4 LinkedIn 搜索算法

1. **搜索**。 使用 Google 的站内定向搜索，根据 Salesforce 中的公司名称和国家在 LinkedIn 上搜索该企业：site:linkedin.com/company “COMPANY_NAME” COUNTRY。如果搜索结果不理想，则使用限制较少的查询方式，包括取消精确搜索、移除公司后缀、重新排列词语顺序，以及使用名称片段而非完整名称进行搜索。

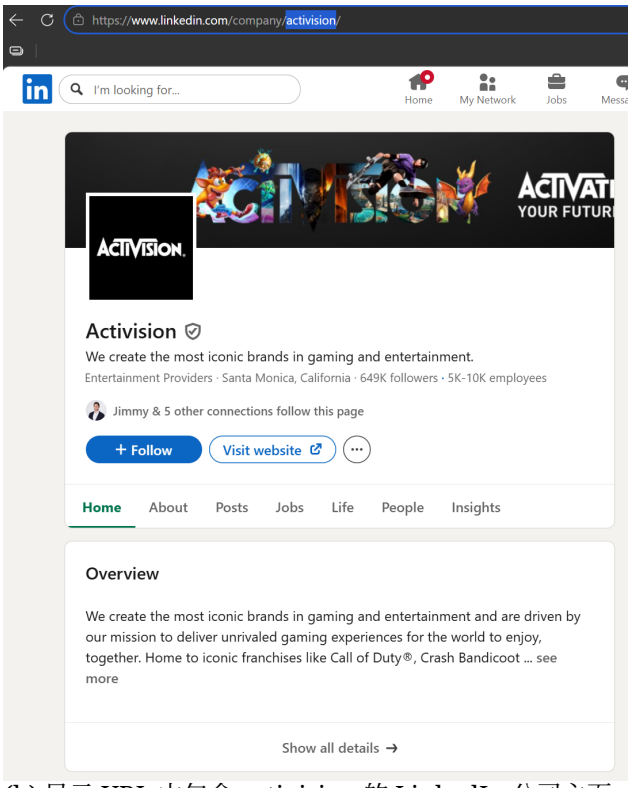
2. **筛选**。在搜索结果中，找出名称与 Salesforce 账户名称完整匹配且 LinkedIn 标注的位置与企业账户所在国家一致的公司主页。其他可供参考的条件包括：公司主页是否已由所有者“认领”（因此并非自动生成），以及是否拥有更多关注者和员工人数。

3. **记录**。 记录网址中的页面 LinkedIn 标识符（例如 `linkedin.com/company/activision`；图 A.3 ），并注明匹配中可能存在的问题，例如公司更名、母子公司关系和国家不匹配。

图 A.3：智能体辅助的企业匹配：搜索与验证工作流



(a) 使用站内定向搜索查找“activision”



(b) 显示 URL 中包含 activision 的 LinkedIn 公司主页

注释：第 B.4 节企业匹配流程示例：通过站内定向搜索查找企业的领英页面，并使用搜索结果中的个人资料记录其公司标识符。

对于未被 Revelio 的算法匹配的企业，我们让一个 AI 智能体（Claude Sonnet 4.6）遵循该算法来构建候选匹配。随后，另一个智能体通过评估该企业的相似性，将每个拟议匹配的置信度评分为高、中、低或无_匹配，以该企业的

Salesforce 名称与 LinkedIn 页面名称是否匹配、国家是否一致，以及 LinkedIn 页面显示的员工人数是否合理。我们仅保留置信度为高或中的匹配结果，纳入最终样本。

RCID 查找。 通过在 WRDS 上查询 Revelio 的公司_映射表，将每个匹配的 LinkedIn 标识符转换为 Revelio 公司标识符（RCID）。该映射表将 LinkedIn 网址与 RCID 对应起来。在 2,803 个置信度为高或中的代理匹配结果中，有 2,679 个（96%）获得了 RCID，这证实匹配到的 LinkedIn 页面对应于 Revelio 就业数据库中的企业。

重复项处理。 在某些情况下，多个 GitHub 企业（通常是母公司及其子公司）会匹配到同一个 LinkedIn 页面。我们发现了 82 个此类重复匹配，涉及母子公司关系，并对其进行了合并处理。

网址修复。 另一位人工审核员分四批核查了 LinkedIn 标识符失效、发生重定向或因其他原因无法访问的企业。审核员在可能的情况下提供了更正后的标识符，为 59 家网址失效的企业中的 46 家恢复了匹配。此网址修复步骤解决了链接失效问题，但不会独立重新评估代理对企业的识别结果。

B.5 匹配质量筛选

我们采用两项发布前筛选，以剔除 GitHub 与 Revelio 之间的错误链接，并确保匹配企业的个人资料覆盖范围充足。首先，我们排除在 GitHub Copilot 发布前，Revelio 中记录的总就业人数或软件工程师人数为零的匹配企业。其次，我们排除这样一些匹配对象：Revelio 中企业发布前的平均总就业人数少于其购买的 GitHub 许可证数量。由于每个 GitHub 许可证对应一个个人用户许可证，企业的总员工人数必须至少与其许可证数量相当；若不满足这一条件，则表明实体匹配有误，或个人资料覆盖范围严重不足。

这两项标准剔除了 1,215 个匹配对象，留下 6,959 家企业（占匹配账户的 80%），其链接已核实，且就业覆盖范围充足。被排除的账户中，公共部门或政府组织，以及总部位于北美和欧洲以外地区的企业占比偏高；这些地区的线上个人资料覆盖范围较为稀疏。由于两项筛选均严格依据 GitHub Copilot 发布前的 2023 年以前数据制定，因此不会根据后续的 GitHub Copilot 采用情况或发布后的就业轨迹来筛选企业。

B.6 验证样本与匹配准确率

我们通过三项验证实验评估公司匹配流程的准确性。前两项分别从算法流程和人工智能辅助流程中随机抽取100家企业，并将分配的匹配结果与人工审核员根据公司名称、国家和行业独立识别的公司简介页面进行比较。第三项则使用已知真实标识符的企业，对 Revelio 名称匹配算法开展更大规模的压力测试。

1. **Revelio 算法匹配（精确率审计，N = 100）**。审核员将 Revelio 分配的公司（RCID）与其独立识别的企业进行比较，并记录匹配是否正确、是否错误（错误公司），或者是否本就不应存在正确匹配。由于审计仅抽取 Revelio 已匹配的企业，因此衡量的是精确率，而非召回率。

2. **AI 智能体辅助匹配（精确率和召回率，N = 100）**。审阅者将智能体提出的 LinkedIn 标识符与其独立识别出的标识符进行了比较。抽样也包括智能体拒绝匹配的企业，因此该样本可用于评估精确率和召回率。

3. **名称扰动压力测试（N = 995）**。对于 995 家真实 RCID 已知的企业，我们先在公司名称后添加法律后缀（例如“Ltd.”、“有限责任公司”）和母公司限定词（例如“– 母公司”），再将其输入 Revelio 的匹配算法。通过固定目标企业，该测试能够单独检验算法对名称规范度变化的敏感性。

表 A.1 报告了各项测试的精确率、召回率和 F1分数。Revelio 的算法匹配结果中，抽样的 100 家企业有 83 家匹配正确（精确率为 0.83）。AI 智能体取得更高精确率（0.93），且召回率接近完整（0.99），F1分数为 0.96：大多数错误案例属于企业层级错误（通常是将子公司匹配到其母公司），而非漏匹配。名称扰动测试的精确率（0.93）与智能体流程相同，但召回率较低（0.76）：对于 22.7% 的受扰动名称，算法未返回匹配结果。这反映出该算法依赖清理后精确匹配，也说明对于名称表示不规范的企业，采用智能体辅助流程具有必要性。

表 A.1：匹配验证摘要：精确率、召回率和 F1 分数

Test	N	精确率	召回率	F1
算法审计	100	0.83	–	–
智能体辅助审计	100	0.93	0.99	0.96
算法扰动	995	0.93	0.76	0.84

注释：该表报告了第B.6节中三项验证实验的精确率、召回率和F1分数：Revelio算法匹配结果的精确率审计（$N = 100$）、AI智能体匹配结果的精确率和召回率审计（$N = 100$），以及针对具有已知标识符的企业开展的Revelio算法名称扰动压力测试（$N = 995$）。人工审核员在每项实验中独立确定了正确的LinkedIn公司主页；精确率是候选匹配结果中正确结果所占的比例，召回率是具有正确匹配结果的企业中被成功匹配的比例。

附录表 A.2 将匹配错误分解为五个互斥类别，每类均以相应验证样本的百分比表示。在算法审计中，错误地分配到其他企业的情况占绝大多数错误（占抽样企业的16.0%），但在AI辅助审计（7.0%）和扰动测试（5.3%）中，这一比例大幅下降。相反，漏匹配（假阴性）在AI辅助审计中很少见（1.0%），但在扰动测试中占22.7%，这证实格式噪声主要导致基于规则的算法未能匹配有效企业，而非建立错误链接。

表 A.2：匹配验证：按验证测试划分的误差分解（占 N)

误差类型	算法	代理辅助	算法扰动
精确匹配	83.0%	88.0%	72.0%
错误公司	16.0%	7.0%	5.3%
漏匹配	–	1.0%	22.7%
正确地不可匹配	–	4.0%	–
本不应匹配	1.0%	–	–

注释：该表将表 A.1 中每项验证练习的匹配错误分为五类，每类占该练习样本量的百分比（$N = 100$、100 和 995）。

B.7 E估计样本构建与汇总统计 s

附录表 A.3概述了我们如何从现有 GitHub 账户总体中构建估计样本。为确保预定合同审查时间相对于 GitHub Copilot 的推出是预先确定的，我们首先选取截至2022年9月持有有效 GitHub 合同，且预定的上线后合同审查日期在 GitHub Copilot 推出前已确定的企业。通过算法和AI辅助匹配程序将这些账户与 Revelio Labs 关联，并应用附录第B.5节所述的上线前匹配质量筛选条件后，得到6,959个就业覆盖范围经核实的匹配组织。

在这个匹配样本中，我们施加两项最终样本限制。首先，剔除 599 个公共部门、政府和非营利实体，以聚焦商业企业，剩余 6,360 家私营部门企业。[15] 其次，为确保各国个人资料覆盖范围一致且具有代表性，我们将样本限制为总部位于美洲或欧洲的企业，这两个地区是 Revelio 覆盖最全面的地区。[16] 最终的平衡面板包含 5,414 家企业，观测期为 2021年1月 至 2025年6月。附录表 A.4报告了完整样本以及截至 2025年6月 是否采用 GitHub Copilot 的企业在 2022年9月 的基线特征。

[15]我们通过两位数北美行业分类系统代码 92（“公共管理”）、在组织名称中搜索政府和非商业相关关键词，以及人工核实，识别公共部门和非营利实体。[16]具体而言，我们保留总部位于 Revelio 划分的六个地理区域内的企业：北美洲、中美洲和加勒比地区、南美洲、北欧、西欧和南欧。这一限制排除了 946 家总部位于亚洲、非洲、中东或大洋洲的企业。

表 A.3：估计样本构建

	企业
初始样本	8,756
与 Revelio (RCID) 匹配	8,174
算法匹配	5,558
人工匹配	2,644
匹配质量筛选	6,959
保留私营企业	6,360
将总部限制在美洲 + 欧洲	5,414

注释：该表报告了经过附录第 B.7 节所述的每项样本限制后剩余的企业数量；最后一行是估计样本。缩进的行按匹配方法划分匹配企业，但有 28 家企业通过两种方法均匹配成功。“私营企业”一行剔除了根据北美行业分类系统代码或组织名称识别出的公共部门组织和非营利组织，以及经人工审查后移除的企业。

表 A.4：汇总统计

	完整样本 (1)	最终采用者 (2)	未采用者 (3)
员工人数	8,284	10,754	4,218
	(28,515)	(34,777)	(11,594)
企业年龄	47	45	50
	(45)	(45)	(45)
技术人员占比（%）	12.6	14.2	10.1
	(11.2)	(11.1)	(10.8)
SWE占比（%）	5.7	6.9	3.7
	(7.1)	(7.2)	(6.5)
初级份额（%）	59.6	57.7	62.8
	(14.5)	(14.4)	(14.1)
招聘率（%）	2.55	2.80	2.15
	(1.45)	(1.52)	(1.23)
离职率（%）	1.92	2.03	1.73
	(0.83)	(0.82)	(0.82)
员工人数增长，2021—2022年（%）	8.6	10.5	5.4
	(14.5)	(15.4)	(12.2)
SWE员工人数增长，2021—2022年（%）	13.4	15.6	9.6
	(31.0)	(29.0)	(33.8)
合同审查周期（月）	12.1	11.9	12.3
	(2.0)	(1.4)	(2.4)
总部位于美国（%）	61.6	63.2	58.8
信息技术（%）	24.8	29.7	16.6
金融服务（%）	12.9	12.6	13.3
企业	5,414	3,368	2,046

注释：该表报告了第 3 节所述分析样本的公司层面汇总统计。员工人数、企业年龄、就业占比、合同审查周期、所在地和行业均以2022年9月为测量时点；招聘率和离职率为上线前期间的月均值；员工人数增长是2021年9月至2022年9月四舍五入后员工人数的变化，并在第1和第99百分位数处进行缩尾处理。括号内为标准差。第（1）列涵盖完整样本；第（2）列和第（3）列根据企业是否在面板结束前采用GitHub Copilot对企业进行划分。

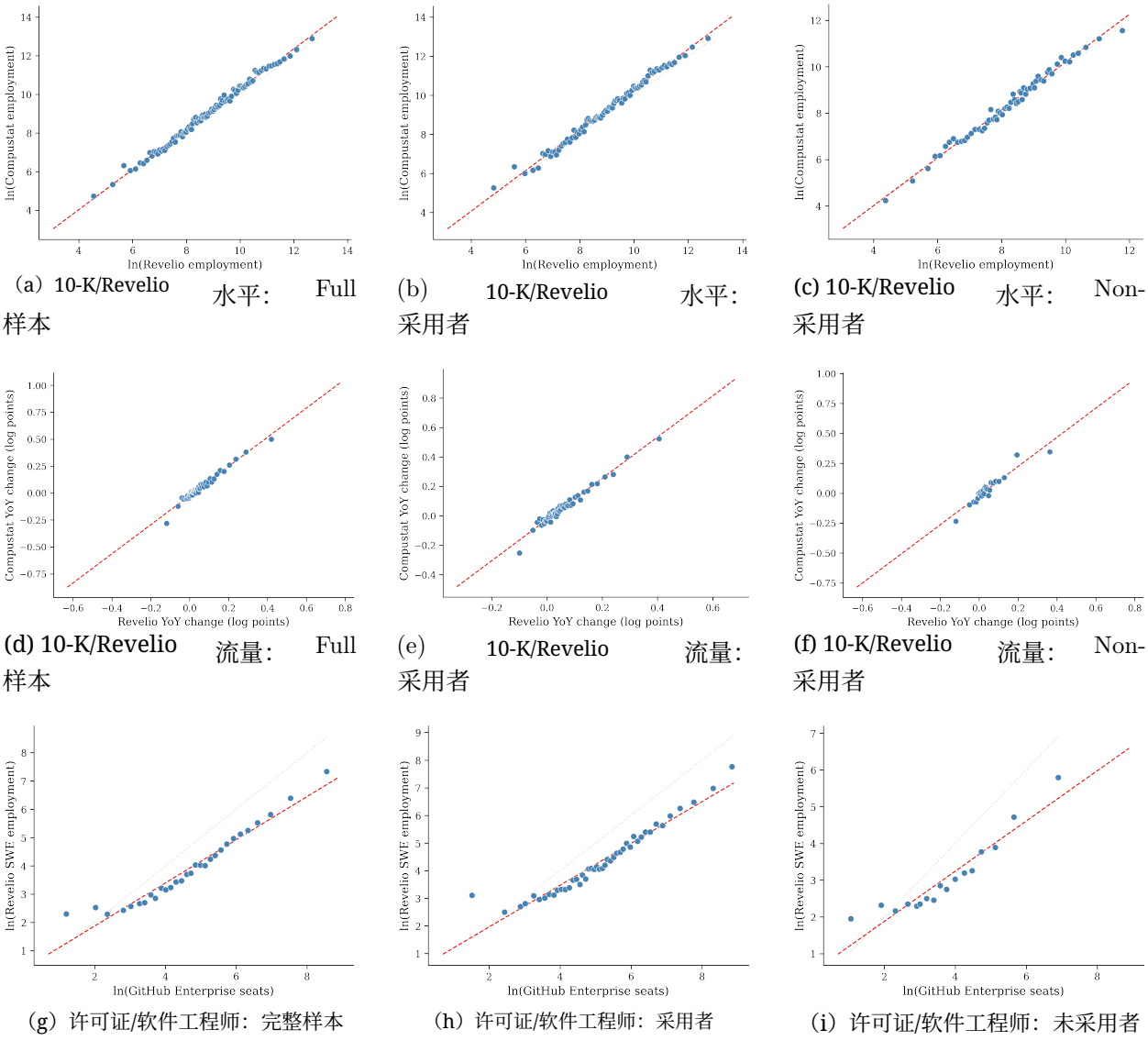
B.8 就业数据验证

我们的实证分析依赖 Revelio Labs 数据来衡量技术岗位中的公司层面就业存量、招聘与离职流量以及劳动力构成。我们将这些指标与两个独立的外部基准进行验证：监管申报文件中经审计的员工总人数，以及 GitHub Enterprise 许可证的行政记录。

首先，对于样本企业中提交了 Compustat 所涵盖的 10-K 表格报告的 19%，我们将 Revelio 中的年均就业人数与 Compustat 报告的员工总人数进行比较。这两组数据在若干制度层面存在差异：10-K 文件报告的是财年末正式薪资册员工人数，而 Revelio 根据在线个人资料衡量年度平均活跃任职期，其中可能包括承包人员和自由职业者。尽管存在这些概念上的差异，两项指标在对数水平上仍高度相关（皮尔逊相关系数为 $r = 0.756$，斯皮尔曼相关系数为 $\rho = 0.705$）。对于依赖公司内部随时间变化的纵向研究设计而言，更重要的是，Revelio 与 Compustat 中就业人数对数的同比变化也高度一致，皮尔逊相关系数为 0。576，斯皮尔曼秩相关系数为 0。598（最终采用 GitHub Copilot 的企业中为 0.593，未采用者中为 0.458）。附录图A.4的前两行绘制了水平值和增长率的比较。

其次，我们使用企业发布前的 GitHub 许可证数量，验证 Revelio 对技术劳动力的测量。行政记录中的许可证数量与 Revelio 记录的职位人数并非一一对应：企业合同通常按批次分配许可证，而非与实际在职人数精确匹配；并非所有软件工程师都需要使用特定企业实例的许可证；而且企业经常会向核心软件工程以外相邻职位的技术人员分配许可证。即便如此，发布前的 GitHub Enterprise 许可证数量与 Revelio 记录的软件工程人员数量在对数水平上仍呈现强正相关（皮尔逊 $r = 0.684$）。在中位数企业中，GitHub Enterprise 许可证数量与 Revelio 记录的软件工程师人数之比约为 2.4。若将分母扩大至包括经常与代码库交互的相邻技术岗位（具体包括应用开发、数据分析、质量保证以及系统与基础设施工程），对数水平上的相关性仍同样强（$r = 0.673$），而 GitHub Enterprise 许可证数量与技术人员人数的中位数之比为 0。9，这证实 Revelio 捕捉到了样本企业中相关的技术人员群体。附录图A.4 的第三行绘出了这些比较结果。

图 A.4：Revelio 就业指标的验证



（a）10-K/Revelio 水平：Full 样本　(b) 10-K/Revelio 水平：采用者　(c) 10-K/Revelio 水平：Non-采用者

(d) 10-K/Revelio 流量：Full 样本　(e) 10-K/Revelio 流量：采用者　(f) 10-K/Revelio 流量：Non-采用者

（g）许可证/软件工程师：完整样本　（h）许可证/软件工程师：采用者　（i）许可证/软件工程师：未采用者

注释：第 1 行以对数水平绘制 Revelio 就业人数与 Compustat 10-K报告就业人数的关系；第 2 行绘制两项指标同比对数变化的关系；第 3 行以对数水平绘制前期 GitHub 许可证与 Revelio 软件工程就业人数的关系。每行均按最终是否采用 GitHub Copilot 对比较进行划分。样本说明见第 3 节。

B.9 美国技术劳动力的覆盖范围

尽管我们的估计样本并不能代表美国雇主总体，因为 GitHub 客户中大型企业和技术密集型企业所占比例过高，但这些企业在美国技术人员就业总量中占有相当大的份额。截至 2025年6月，样本企业在美国雇用了 1,646,968 名技术人员。作为外部基准，美国劳工统计局职业就业与工资统计（OEWS）中相应编程与计算机职业的美国就业总人数为 4,347,980（美国劳工统计局，

2025年），这意味着我们的样本覆盖了约38%的美国编程劳动力。由于我们的估计样本仅包含私营部门企业，而OEWS基准涵盖了私营部门和公共部门就业，因此我们的样本占美国私营部门技术就业的比例更高。附录表A.5 列出了基准中各职业的全国就业人数。

表A.5：2025年美国编程相关职业就业人数

SOC 代码	职业	就业人数（2025年）
11-3021	计算机和信息系统经理	670,570
15-1211	计算机系统分析师	519,530
15-1212	信息安全分析师	190,650
15-1243	数据库架构师	67,140
15-1244	网络和计算机系统管理员	314,340
15-1251	计算机程序员	92,230
15-1252	软件开发人员	1,687,890
15-1253	软件质量保证分析师和测试人员	186,740
15-1254	网络开发人员	70,190
15-1255	网络和数字界面设计师	113,330
15-1299	其他所有计算机职业	435,370
	总计	**4,347,980**

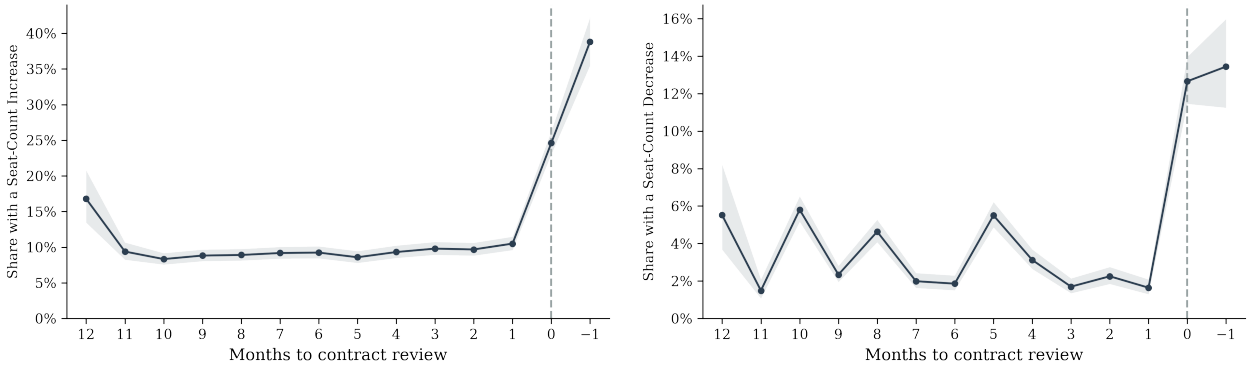
注释：美国劳工统计局职业就业与工资统计中的美国全国就业数据，2025年职业就业与工资统计。不包括15-2051数据科学家（262,440名工人）和15-1241计算机网络架构师（179,740名工人），这两个职业同样不在我们对技术人员的定义之内。样本企业在2025年6月雇用了1,646,968名位于美国的技术人员，这意味着 ≈38%覆盖率。

B.10 代码强度提示

我们使用 Claude Sonnet 4.6，在温度为零的设置下，对每份职位描述评分一次。模型只能看到职位描述文本，该文本插入提示词中 <职位描述文本>所示的位置；它看不到职位名称、职业、资历、雇主、地点或日期。下方提示词逐字重现了提示词中的指令、代码强度维度的定义以及所要求的输出格式。

```
You are scoring the work described in an anonymized job-position description on five dimensions. Score ONLY
what the text describes; do not infer from writing style, language, or any place names what kind of
person wrote it.The dimension is a SHARE of the described work, scored as a continuous number in [0, 1]:
[...]
copilot_domain: the share of the described work inside GitHub Copilot's 2023 capability surface: writing,
completing, testing, documenting, explaining, or debugging code (1.0 = all inside; 0.0 = none)[...]
Position description:___<position description text>___
```

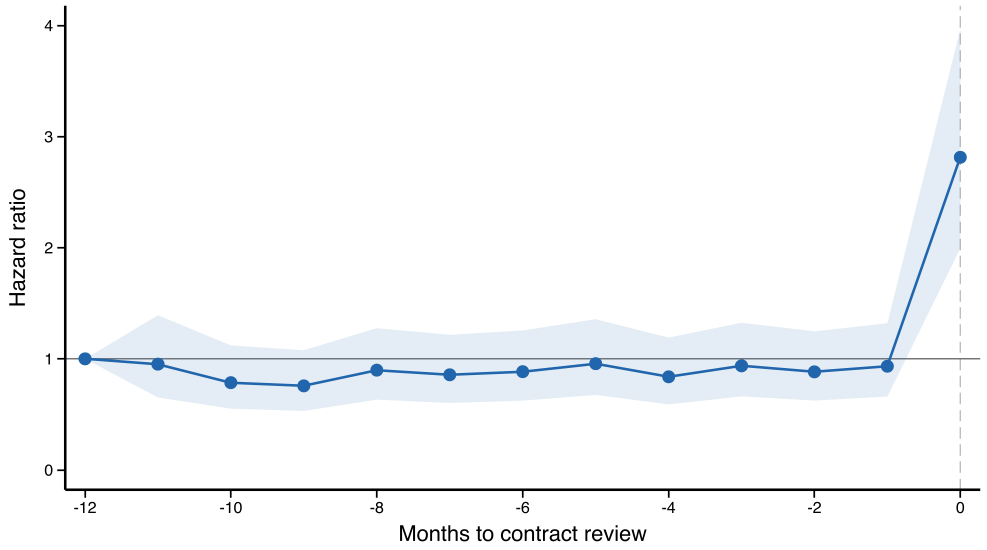
图 A.5：预定合同审查前后各月 GitHub Enterprise 许可证数量的增加与减少



(a) GitHub 许可证数量增加　　(b) GitHub 许可证数量减少

注释：在 GitHub Copilot 发布前，面板A绘制的是企业-月份中，企业缩尾处理后的 GitHub Enterprise 许可证数量高于上月的比例；面板B绘制的是低于上月的比例。第 0 月为审查月份，第 -1 月为审查后的月份。

图 A.6：G　　距预定合同审查月数对应的 GitHub Copilot 采用风险比　　w



注释：此图报告了互补对数-对数模型得出的风险比，该模型包含距审查的月数和自 GitHub Copilot 发布以来月数的指标变量；距合同审查 12 个月为参照类别。标准误按企业聚类。

表 A.6：GitHub Copilot 采用与就业：普通最小二乘法估计

因变量	员工人数对数			技术人员招聘人数对数	技术人员离职人数对数
	All	技术人员	SWE		
	(1)	(2)	(3)	(4)	(5)
企业固定效应 ＋月份固定效应	0.011***	0.018***	0.019***	-0.051***	-0.012
	(0.002)	(0.004)	(0.005)	(0.009)	(0.008)
＋行业 × 月份固定效应	0.017***	0.028***	0.029***	-0.037***	-0.005
	(0.003)	(0.004)	(0.005)	(0.009)	(0.008)
＋ 规模四分位数 ×月份固定效应	0.015***	0.023***	0.023***	-0.029***	0.002
	(0.003)	(0.004)	(0.005)	(0.009)	(0.008)
因变量均值	7.316	4.664	3.464	1.563	1.428
企业	5,414	5,414	5,282	5,192	5,208
观测值	191,242	191,147	182,162	114,934	113,632

注释：每个条目均为双向固定效应回归中一个指标变量的系数。该指标变量从企业首次采用 GitHub Copilot 的月份起及之后均取值为 1；该回归以单一的采用后指标变量替代方程（1）中的事件时间指标。第一行纳入企业和日历月份固定效应；第二行进一步纳入两位数北美行业分类系统行业 × 日历月份固定效应；第三行再进一步纳入发布前员工人数四分位组 × 日历月份固定效应。标准误按企业聚类，并列于括号内。效应以对数点表示。样本包括采用后十二个月内的采用企业月度观测值，以及所有未采用企业月度观测值。 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ 。

表 A.7：ChatGPT 发布前后最终采用者的双重差分估计

因变量	技术人员招聘人数对数	技术人员离职人数对数	技术人员总数的对数
	(1)	(2)	(3)
11月之后的 2022 × 最终采用者	-0.101***	-0.006	0.067***
	(0.014)	(0.011)	(0.007)
企业固定效应	✓	✓	✓
月份固定效应	✓	✓	✓
因变量均值	1.691	1.522	4.868
企业	5,272	5,278	5,414
观测值	192,425	187,526	292,261

注释：每一项均为双向固定效应回归中“11月后 2022 × 最终采用者”指标变量的系数；该回归以单一的“11月后 2022 × 最终采用者”指标变量替代方程（2）中的事件时间指标变量。标准误按企业聚类，并列于括号内。效应以对数点表示。数据为平衡的企业—月份面板数据，时间范围为2021年1月至2025年6月。 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$。

表 A.8：按采用情况和合同审查时间划分的平衡性

	最终采用者与未采用者	较早与较晚的合同审查
招聘率	0.303	-0.067
	（0.018）	（0.019）
离职率	0.222	-0.019
	(0.017)	(0.017)
软件工程师招聘率	0.163	-0.019
	(0.011)	(0.010)
软件工程师离职率	0.117	0.011
	(0.010)	(0.009)
科技行业增长率 2021年-2022年	0.210	-0.064
	(0.028)	(0.028)
软件工程师增长率 2021-2022	0.147	-0.063
	(0.030)	(0.029)
员工人数的自然对数	0.334	0.241
	(0.027)	(0.027)
软件工程师就业占比	0.424	0.017
	(0.027)	(0.027)
初级份额	-0.347	0.021
	(0.027)	(0.027)
总部位于美国	0.082	-0.024
	（0.028）	（0.027）
总部位于英国	0.025	-0.017
	（0.028）	（0.028）
信息技术	0.304	-0.010
	(0.026)	(0.027)
金融服务	-0.023	0.027
	(0.028)	(0.027)

注释：表中报告了图5所依据的系数。每个单元格列出一个组别指标的系数，下方为标准误；该指标来自对某项标准化处理前特征进行的独立回归。第（1）列比较最终采用者与未采用者；第（2）列比较预定合同审查时间较早和较晚的企业，以2023年6月为界。对于时变特征，使用发布前公司月度数据并纳入日历月份固定效应；对于时不变特征，使用发布月份截面；2021—2022年增长各行表示2021年12月至2022年12月员工人数取整后的变化。括号内为按企业聚类的标准误。

D 补充就业与招聘结果

表 A.9：技术岗位招聘与离职的季度沃尔德估计

距合同审查的季度数（ℓ）	简化式 (1)	第一阶段 (2)	LATET (3)	95% 区间 (4)
面板*A*。技术岗位招聘				
$+1$	-0.021	0.057^{***}	-0.365	[-1.351, 0.354]
	(0.020)	(0.015)		
$+2$	-0.050	0.107^{***}	-0.467	[-1.205, 0.083]
	(0.031)	(0.024)		
$+3$	-0.063	0.161^{***}	-0.391	[-1.020, 0.046]
	(0.039)	(0.032)		
$+4$	-0.073	0.215^{***}	-0.340	[-0.948, 0.088]
	(0.051)	(0.039)		
面板*B*。技术人员离职				
$+1$	0.031^{*}	0.057^{***}	0.556	[-0.108, 1.583]
	(0.018)	(0.015)		
$+2$	0.018	0.107^{***}	0.163	[-0.348, 0.740]
	(0.025)	(0.024)		
$+3$	0.026	0.161^{***}	0.161	[-0.285, 0.623]
	(0.033)	(0.032)		
$+4$	0.034	0.215^{***}	0.160	[-0.323, 0.635]
	(0.049)	(0.039)		
自助法抽样次数		1,000		

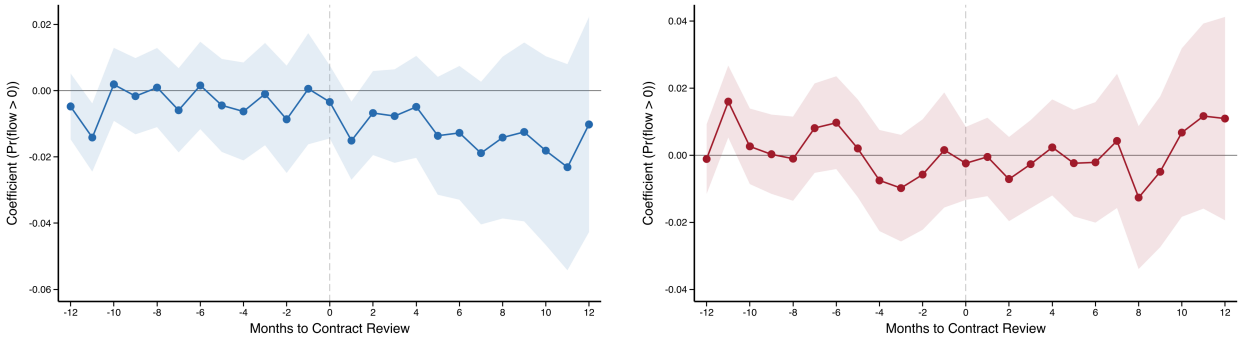
注：第（1）、第（3）和第（4）列的单位为对数点；第（2）列的单位为概率单位。第（1）列为各时间跨度下的简约式系数；第（2）列为在相同时间跨度下，合同审查对采用 GitHub Copilot 概率的第一阶段效应；第（3）列为二者的比率（方程 5）。第（4）列报告了基于 1,000 次企业区块自助法抽样得到的 95% 百分位区间。由于审查季度的第一阶段效应较小，使比率不稳定，因此省略第 0 季度。括号内为按企业聚类的标准误。 * $p < 0.10$, ** $p < 0.05$ *** $p < 0.01$ 。

表 A.10：合同审查对技术劳动力结果的影响

因变量	技术岗位招聘		技术人员离职		技术人员总数的对数
	Log	Any	Log	Any	
距合同审查的月数（ℓ）	(1)	(2)	(3)	(4)	(5)
−12	-0.006	-0.005	-0.003	-0.001	-0.001
	(0.011)	(0.005)	(0.010)	(0.005)	(0.003)
−11	0.015	-0.014***	0.007	0.016***	-0.001
	(0.011)	(0.005)	(0.010)	(0.005)	(0.003)
−10	0.003	0.002	0.025**	0.003	-0.003
	(0.012)	(0.006)	(0.011)	(0.006)	(0.004)
−9	0.003	-0.002	0.011	0.000	-0.003
	(0.013)	(0.006)	(0.012)	(0.006)	(0.004)
−8	0.015	0.001	0.028**	-0.001	-0.005
	(0.014)	(0.006)	(0.013)	(0.006)	(0.005)
−7	0.004	-0.006	0.030**	0.008	-0.006
	(0.015)	(0.006)	(0.014)	(0.007)	(0.006)
−6	-0.003	0.002	0.010	0.010	-0.008
	(0.016)	(0.007)	(0.015)	(0.007)	(0.006)
−5	0.013	-0.004	0.018	0.002	-0.010
	(0.018)	(0.007)	(0.015)	(0.007)	(0.007)
−4	0.007	-0.006	0.008	-0.008	-0.011
	(0.019)	(0.008)	(0.016)	(0.008)	(0.008)
−3	0.009	-0.001	0.020	-0.010	-0.012
	(0.020)	(0.008)	(0.017)	(0.008)	(0.008)
−2	0.018	-0.009	-0.001	-0.006	-0.013
	(0.022)	(0.008)	(0.018)	(0.008)	(0.009)
−1	0.012	0.001	0.004	0.002	-0.015
	(0.023)	(0.009)	(0.019)	(0.009)	(0.010)
0	-0.008	-0.003	0.007	-0.002	-0.004
	(0.012)	(0.006)	(0.012)	(0.006)	(0.003)
+1	-0.001	-0.015**	0.012	-0.001	-0.007*
	(0.014)	(0.006)	(0.013)	(0.006)	(0.004)
+2	0.007	-0.007	0.028**	-0.007	-0.008*
	(0.016)	(0.006)	(0.014)	(0.006)	(0.005)
+3	-0.011	-0.008	0.020	-0.003	-0.007
	(0.017)	(0.007)	(0.015)	(0.007)	(0.005)
+4	-0.006	-0.005	0.020	0.002	-0.008
	(0.020)	(0.008)	(0.016)	(0.007)	(0.006)
+5	-0.021	-0.014	0.030	-0.002	-0.010
	(0.023)	(0.009)	(0.019)	(0.008)	(0.007)
+6	-0.030	-0.013	0.035	-0.002	-0.009
	(0.025)	(0.010)	(0.022)	(0.009)	(0.009)
+7	-0.042	-0.019*	0.014	0.004	-0.010
	(0.028)	(0.011)	(0.023)	(0.010)	(0.010)
+8	-0.040	-0.014	0.039	-0.013	-0.009
	(0.030)	(0.012)	(0.025)	(0.011)	(0.012)
+9	-0.052	-0.012	0.031	-0.005	-0.010
	(0.033)	(0.014)	(0.028)	(0.011)	(0.013)
+10	-0.038	-0.018	0.025	0.007	-0.011
	(0.034)	(0.015)	(0.035)	(0.013)	(0.015)
+11	-0.081**	-0.023	0.058	0.012	-0.013
	(0.035)	(0.016)	(0.039)	(0.014)	(0.017)
+12	-0.079**	-0.010	0.029	0.011	-0.016
	(0.038)	(0.017)	(0.041)	(0.015)	(0.018)
企业固定效应	✓	✓	✓	✓	✓
月份固定效应	✓	✓	✓	✓	✓
前趋势联合 p	0.849	0.244	0.120	0.014	0.289
因变量均值	1.722	0.707	1.533	0.661	4.812
企业	5,339	5,414	5,357	5,414	5,414
观测值	149,565	221,346	143,324	221,346	221,298

注释：表中报告了使用 Borusyak、Jaravel 和 Spiess (2024年) 的插补估计量，在包含企业和日历月份固定效应的平衡企业-月份面板数据上得出的简约式事件研究估计值。（1）列和（3）列的因变量分别为技术岗位招聘人数对数和离职人数；（2）列和（4）列分别为是否发生此类流动的指标变量；（5）列的因变量为技术人员总数的对数。括号内为按企业聚类的标准误。前趋势联合 p 值报告了 p值，表示 Wald 检验的结果，用于检验审查前系数是否联合为零。 * $p < 0.10$， ** $p < 0.05$， *** $p < 0.01$。

图A.7: 合同审查对技术岗位招聘和离职的广延边际效应 s



(a) 招聘 (b) 离职

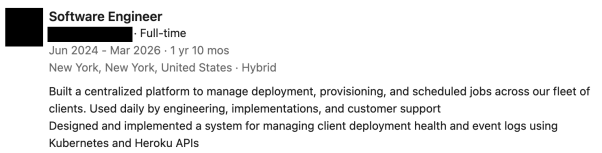
注释：图中报告了技术岗位招聘指标（面板A）和技术岗位离职指标（面板B）的简约式系数及95%置信区间。估计使用Borusyak、Jaravel 和 Spiess (2024年的插补估计量，基于包含企业和日历月份固定效应的平衡企业-月份面板数据。标准误按企业聚类。

表 A.11：对技术岗位招聘信息的简约式效应

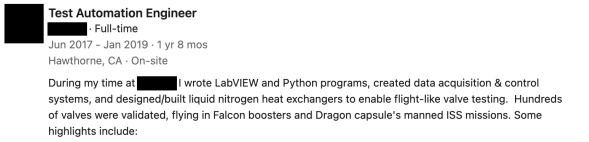
因变量	招聘信息对数	招聘信息发布时长对数	任何招聘信息
距合同审查的季度数（ℓ）	(1)	(2)	(3)
-4	-0.042^{*}	0.026^{*}	-0.008
	(0.022)	(0.014)	(0.006)
-3	-0.050	0.045^{**}	-0.002
	(0.034)	(0.020)	(0.010)
-2	-0.063	0.085^{***}	-0.004
	(0.045)	(0.027)	(0.013)
-1	-0.072	0.098^{***}	0.006
	(0.056)	(0.033)	(0.016)
0	0.017	-0.005	0.003
	(0.018)	(0.013)	(0.005)
$+1$	0.012	0.012	0.003
	(0.030)	(0.020)	(0.009)
$+2$	-0.011	0.021	0.000
	(0.047)	(0.033)	(0.014)
$+3$	-0.080	0.042	-0.013
	(0.067)	(0.041)	(0.018)
$+4$	-0.148^{*}	0.042	-0.021
	(0.088)	(0.051)	(0.023)
企业固定效应	✓	✓	✓
季度固定效应	✓	✓	✓
预趋势联合 p	0.386	0.011	0.069
因变量均值	3.677	3.266	0.857
企业	4,822	4,661	5,414
观测值	43,297	40,782	53,363

注释：本表报告了基于Borusyak、Jaravel 和 Spiess （2024年）插补估计量得出的季度简化式估计，回归纳入企业和日历季度固定效应，使用2022年第一季度至2025年第二季度企业—季度面板数据。结果变量包括：（1）开放中的技术岗位发布数的对数，样本为2022年四个季度中至少有两个季度存在开放技术岗位发布的企业；（2）从发布到撤除的平均天数对数，样本为2022年四个季度中至少有两个季度存在已撤除岗位发布的企业；（3）是否存在任何开放技术岗位发布的指标变量。删除具有插补的撤除日期或在两天内撤除的岗位发布。回归未加权。标准误按企业聚类，并列于括号内。* $p < 0.10$， ** $p < 0.05$， *** $p < 0.01$。

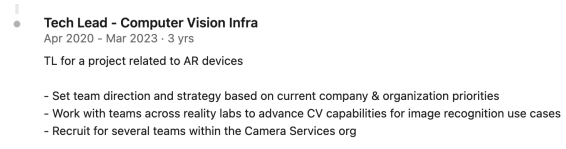
图 A.8：技术岗位的 LinkedIn 职位描述示例



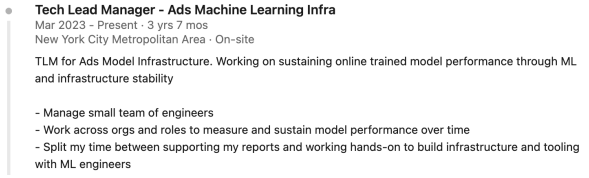
（a）软件工程师



（b）质量保证工程师



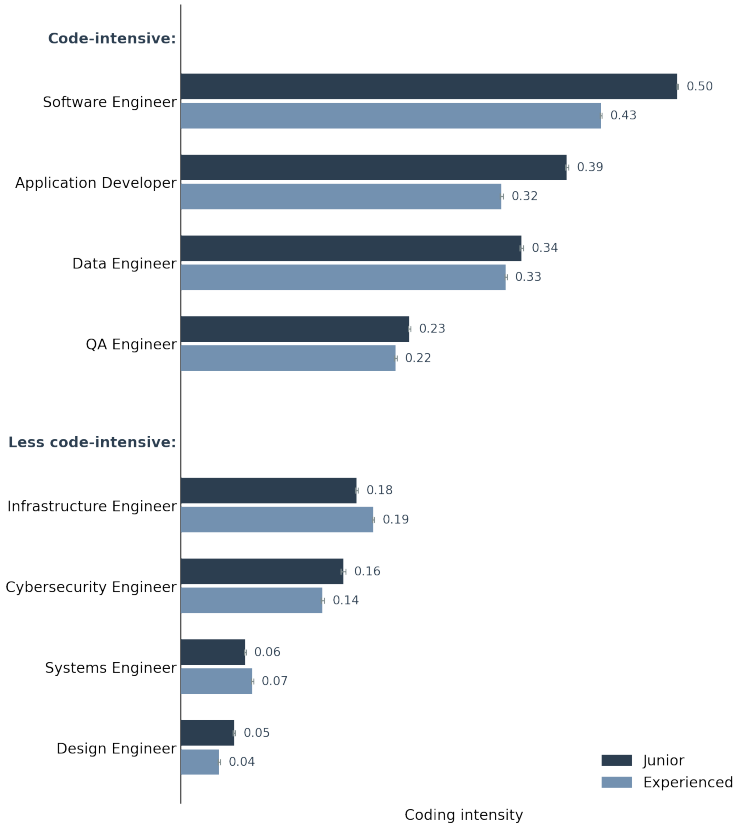
（c）基础设施工程师，计算机视觉



（d）基础设施工程师，机器学习

注释：示例取自样本之外的企业，以保护匿名性；雇主名称和徽标已涂黑。代码强度指标仅使用自由文本描述；在评分前，会删除标题、雇主、日期、地点和资历信息。

图 A.9：按技术岗位和资历划分的编码强度



注：该图报告了按职位和资历划分的、始于2018年–2022年的职位描述的未加权平均代码强度得分（初级职位为Revelio 级别1–2；有经验的职位为级别3–7）。区间为基于职位描述层面变异计算的正态95%置信区间。

表 A.12：按代码强度和资历划分的招聘效应：单元特定得分与企业内部比率

因变量	季度招聘人数对数		初级员工与有经验员工招聘比率的对数 招聘比率	
代码强度指标	职位 × 资历单元		岗位得分	代码密集型岗位
	(1)	(2)	(3)	(4)
合同审查后的第4季度（ρ_4）	-0.017	-0.021	-0.019	-0.026
	(0.035)	(0.042)	(0.050)	(0.049)
代码强度（θ_4）	-0.029*	0.020	-0.080**	-0.169**
	(0.018)	(0.025)	(0.038)	(0.077)
初级（γ_4）		0.011		
		(0.040)		
初级 ×代码强度（β_4）		-0.082***		
		(0.030)		
企业 × 职位 × 资历固定效应	✓	✓		
职位 × 资历 ×季度固定效应	✓	✓		
企业 × 职位固定效应			✓	✓
职位 × 季度固定效应			✓	✓
推断的初级员工梯度（$\theta_4 + \beta_4$）		-0.062***		
		(0.021)		
企业	5,048	5,048	4,490	4,490
观测值	357,346	357,346	125,856	125,856

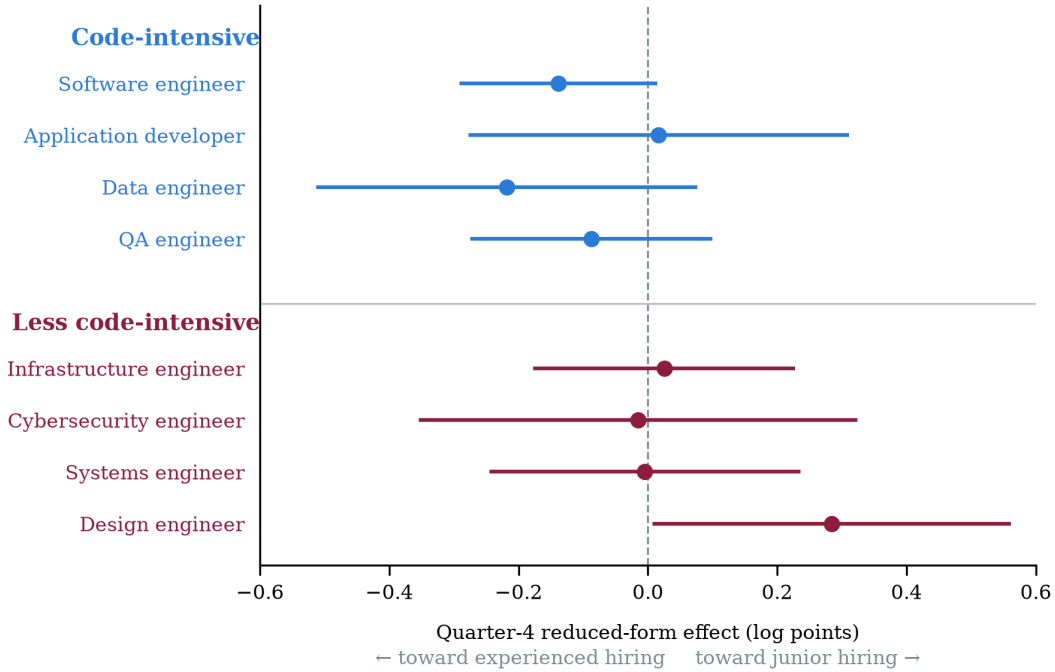
注释：本表报告预定合同审查后四个季度的简约式效应。第（1）列和第（2）列使用 16 个职位—资历单元各自的启动前编码强度得分，并在各单元间进行标准化（图 A.9），样本为表 2中的企业—岗位—资历—季度样本。推断的初级员工梯度为 $\theta_4 + \beta_4$，其标准误根据估计协方差计算。第（3）列和第（4）列报告职位内初级员工与有经验员工招聘比率对数的效应。第（3）列中，ρ_4 表示平均代码强度下的效应，θ_4 表示该效应随启动前岗位得分每增加一个标准差而发生的变化。第（4）列中，代码强度是一个指标变量，用于标识上述四个代码密集型岗位（得分高于中位数）。标准误按企业聚类，并列于括号内。 * $p < 0.10$， ** $p < 0.05$， *** $p < 0.01$。

表 A.13：按职位组划分的招聘广延边际估计

因变量	任何招聘				两组均招聘	
	软件工程	系统与基础设施	代码密集度较高	代码密集度较低	初级与资深软件工程师	软件工程与系统/基础设施
	(1)	(2)	(3)	(4)	(5)	(6)
合同审查后的第4季度	0.002	-0.002	-0.016	0.009	-0.001	-0.009
	(0.024)	(0.022)	(0.019)	(0.021)	(0.027)	(0.027)
企业固定效应	✓	✓	✓	✓	✓	✓
季度固定效应	✓	✓	✓	✓	✓	✓
趋势前联合 p	0.475	0.687	0.094	0.806	0.189	0.594
因变量均值	0.712	0.647	0.792	0.697	0.540	0.551
企业	5,182	5,130	5,360	5,202	4,561	4,932
观测值	71,741	70,980	74,230	71,995	63,029	68,189

注释：该表报告了预定合同审查四个季度后，对职位组中任何招聘（第1–4列）以及同一季度一对职位组中两个组均有招聘（第5–6列）这些指标的简约式效应；该效应是在季度企业面板上，采用 Borusyak、Jaravel 和 Spiess (2024年) 的插补估计量，并控制企业固定效应和日历季度固定效应进行估计的。回归未加权。趋势前 p-值对四个提前项进行联合检验。标准误按企业聚类，并列于括号内。 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$。

图 A.10：按职位划分的初级员工与资深员工招聘比率



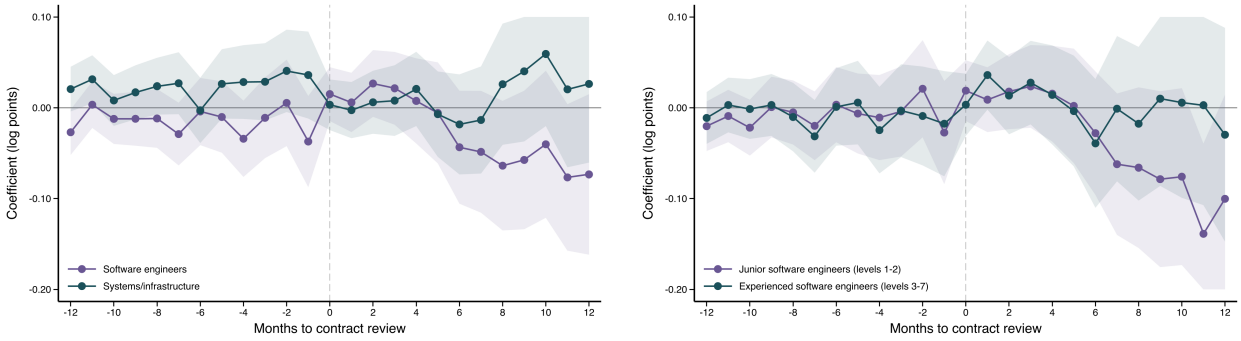
注释：图中报告了预定合同审查后四个季度时，针对各职位的初级员工与有经验员工招聘比率对数的简约式效应。估计值为负表示招聘转向经验更丰富的员工；估计值为正表示招聘转向初级员工。职位按 GitHub Copilot 发布前的代码强度排序。点表示各职位的估计值，水平误差线表示基于按企业聚类的标准误计算的 95%置信区间。

表 A.14：软件工程师与系统/基础设施招聘的 Wald 估计

	季度招聘人数对数			
	按职位		按资历划分的软件工程	
	软件工程学	系统/基础设施	初级	有经验的
	(1)	(2)	(3)	(4)
面板*A*。简化式				
合同审查（第四季度）	−0.064	0.050	−0.108**	−0.034
	(0.048)	(0.055)	(0.052)	(0.064)
面板*B*。第一阶段				
采用的第一阶段效应	0.206***	0.238***	0.211***	0.211***
	(0.042)	(0.039)	(0.047)	(0.047)
*C*面板。*Wald*估计值				
GitHub Copilot 采用	−0.310	0.212	−0.513	−0.160
	[−0.884, 0.186]	[−0.247, 0.798]	[−1.262, 0.008]	[−0.830, 0.492]
企业固定效应	✓	✓	✓	✓
季度固定效应	✓	✓	✓	✓
企业	4,901	4,825	4,407	4,182
观测值	51,022	45,843	44,191	38,895

注释：每列报告预定合同审查后四个季度，按职位或按职位和资历划分的组的季度招聘人数对数估计值。面板A报告合同审查的简约式效应，单位为对数点。面板B报告第一阶段估计值，单位为概率单位；C面板报告Wald估计值，单位为对数点（方程 5）；方括号内为95%百分位数公司区块自助法置信区间。两阶段均采用 Borusyak、Jaravel 和 Spiess (2024年) 的插补估计量，并纳入企业和日历季度固定效应。回归未加权。标准误按企业聚类，括号内为标准误。
* $p < 0.10$， ** $p < 0.05$， *** $p < 0.01$。

图A.11：简　　化式软件工程和系统/基础设施招聘估计



（a）软件工程与系统/基础设施　　（b）软件工程内部比较

注释：该图报告了在存在正招聘的企业-月份样本中，招聘时段对数的月度简化式事件研究系数。面板A比较所有软件工程职位与系统和基础设施工程职位；面板B比较初级（级别1–2）和有经验的（级别3–7）软件工程职位。估计采用Borusyak、Jaravel 和 Spiess （2024年）的插补估计量，并纳入企业和日历月份固定效应。带状区域表示95%置信区间；标准误按企业聚类。

Ta表 A.15：不同编码强度下技术工人的教育程度与工作经验

		获得学士学位后的年数	学术型硕士（%）	博士学位（%）	工程学位（%）	工作经历段
SWE	初级	4.8	26.9	1.6	93.6	1,518,836
	有经验的	8.7	32.7	2.5	92.7	874,670
代码密集型	初级	5.0	25.9	1.6	92.6	2,361,788
	有经验的	8.9	31.6	2.5	90.6	1,824,329
代码密集度较低	初级	6.0	26.0	1.8	90.2	1,067,491
	有经验的	10.4	32.4	3.0	86.2	916,679
所有技术岗位	初级	5.2	26.0	1.6	91.9	3,429,279
	有经验的	9.3	31.9	2.6	89.3	2,741,008

注释：样本为 2021年和2022年样本企业中软件工程师的职位任期，未加权；涵盖所有代码密集型和代码密集度较低的技术岗位，以及按资历划分的所有技术岗位。代码密集型岗位包括软件工程师、应用程序开发人员、数据工程师和质量保证工程师；代码密集度较低的岗位包括基础设施、系统、网络安全和设计工程师。获得学士学位后的年数是指员工获得学士学位至职位开始之间的平均年数。学位占比是指具有相关教育记录的工作经历段所占百分比：学术型硕士（不包括工商管理硕士学位和博士学位）、博士学位，或工程学或信息技术领域的任何学位。44%至56%的工作经历段缺少学位年份，35%至44%缺少学位层级，45%至56%缺少专业领域信息。工作经历段指职位任期的数量。

表 A.16：印度本地与国内技术岗位招聘

因变量	技术人员招聘人数对数		任何技术岗位招聘	
距合同审查的季度数（ℓ）	印度本地 (1)	国内 (2)	印度本地 (3)	国内 (4)
-4	0.023	-0.009	0.000	0.000
	(0.026)	(0.012)	(0.005)	(0.005)
-3	0.020	0.007	-0.004	-0.006
	(0.034)	(0.015)	(0.006)	(0.006)
-2	0.007	0.011	-0.008	-0.007
	(0.045)	(0.020)	(0.008)	(0.008)
-1	0.004	0.005	-0.012	-0.006
	(0.056)	(0.024)	(0.009)	(0.009)
0	0.005	0.009	-0.001	0.004
	(0.034)	(0.015)	(0.005)	(0.006)
$+1$	-0.015	-0.001	-0.005	0.011
	(0.055)	(0.022)	(0.008)	(0.009)
$+2$	-0.075	-0.022	0.012	0.016
	(0.106)	(0.034)	(0.013)	(0.016)
$+3$	-0.079	-0.042	0.009	0.014
	(0.128)	(0.045)	(0.017)	(0.020)
$+4$	-0.060	-0.029	0.029	0.016
	(0.189)	(0.054)	(0.019)	(0.025)
企业固定效应	✓	✓	✓	✓
季度固定效应	✓	✓	✓	✓
前趋势联合 p	0.869	0.568	0.643	0.744
因变量均值	2.367	2.331	0.275	0.872
企业	2,368	4,434	4,509	4,509
观测值	16,633	52,821	62,336	62,336

注释：样本为 2021年 至 2022年期间某个季度中至少有一名担任任何职位的印度本地员工的企业。印度驻地招聘是指在印度境内技术职位上的招聘任期；国内招聘是指在企业总部所在国境内的招聘任期。第 (1) 和 (2) 列是在该单元格中有正招聘的企业-季度观测上以对数点表示的结果；第 (3) 和 (4) 列是该单元格中是否有任何招聘的指标变量。所有列均在季度面板数据上使用Borusyak、Jaravel 和 Spiess (2024年)的插补估计量，并纳入企业和日历季度固定效应。趋势前联合 p-值检验四个提前期系数是否联合为零。回归未加权。标准误按企业聚类，并列于括号内。 * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$。

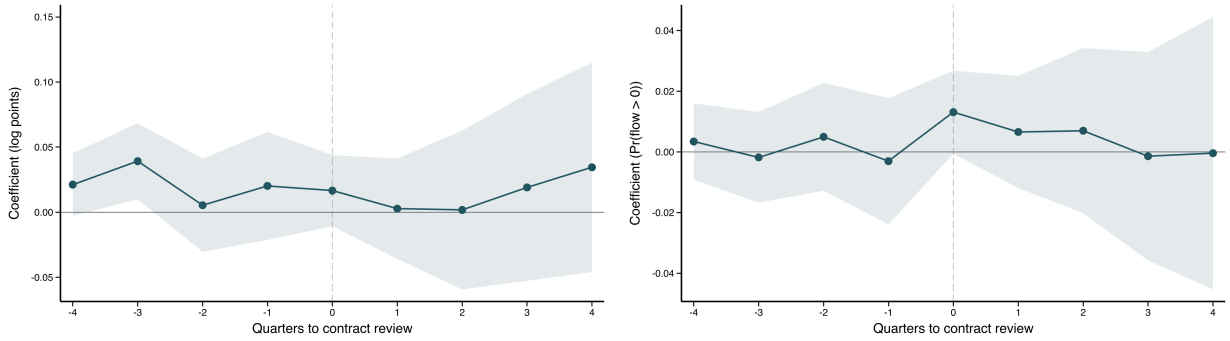
E 安慰剂检验与稳健性检验

表 A.17：一年期合同审查安慰剂

因变量 距离安慰剂审查的季度数（ℓ）	记录技术人员 人数 (1)	记录技术人员 离职 (2)	初级 软件工程师招聘 (3)	有经验的 软件工程师招聘 (4)
-2	-0.042^{**}	0.009	-0.018	-0.023
	(0.018)	(0.018)	(0.022)	(0.024)
-1	-0.049^{**}	0.004	-0.021	-0.063^{**}
	(0.023)	(0.023)	(0.028)	(0.031)
0	-0.016	0.006	-0.001	0.001
	(0.012)	(0.012)	(0.016)	(0.017)
$+1$	-0.006	0.026	0.030	-0.017
	(0.018)	(0.018)	(0.022)	(0.024)
$+2$	-0.033	0.045	0.005	-0.058
	(0.028)	(0.028)	(0.037)	(0.042)
企业固定效应	✓	✓	✓	✓
季度固定效应	✓	✓	✓	✓
预趋势联合 p	0.104	0.416	0.867	0.038
企业	5,212	5,180	4,249	3,934
观测值	36,896	35,621	27,171	22,703

注释：样本包括从2021年第一季度到2022年第四季度的企业-季度观测。安慰剂日期将预定合同审查提前四个季度，同时保留基线处理组和对照组。结果变量包括技术岗位招聘人数对数、离职人数，以及初级和有经验的软件工程师（SWE）招聘任期。估计采用Borusyak、Jaravel 和 Spiess (2024) 的插补估计量，并纳入企业和日历季度固定效应。Q0 表示安慰剂审查所在季度；由于结果数据截至2022年第四季度，后续时间跨度涵盖的队列较少，因此样本并非完全平衡。前趋势联合 p值检验三个估计的提前项。标准误按企业聚类，并列于括号内。效应以对数点表示。 * $p < 0.10$， ** $p < 0.05$， *** $p < 0.01$。

图 A.12：会计师招聘安慰剂：事件研究估计值



(a) 招聘人数对数

(b) 任何招聘

注：该图报告了正向会计师招聘的企业-季度观测中，会计师招聘时段对数的简约式效应（面板A），以及任何会计师招聘的指标变量的简约式效应（面板B）。估计采用 Borusyak、Jaravel 和 Spiess （2024年）的插补估计量，并控制企业和日历季度固定效应。阴影区域表示95%置信区间；标准误按企业聚类。

表 A.18：对规模和行业趋势的稳健性

因变量	技术人员		SWE			按代码密集度划分的招聘	
	招聘人数对数	离职人数对数	初级与资深员工比例的对数	初级员工招聘人数对数	资深员工招聘人数对数	代码密集度较高	代码密集度较低
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
基线							
合同审查后的第4季度	−0.073	0.034	−0.139*	−0.108**	−0.034	−0.081*	0.027
	(0.051)	(0.049)	(0.078)	(0.052)	(0.064)	(0.047)	(0.057)
企业规模三分位组 × 季度							
合同审查后的第4季度	−0.077	0.037	−0.139*	−0.094*	−0.048	−0.090*	0.025
	(0.051)	(0.046)	(0.075)	(0.054)	(0.062)	(0.046)	(0.056)
行业 × 季度							
合同审查后的第4季度	−0.072	0.026	−0.116	−0.106**	−0.053	−0.087*	0.026
	(0.050)	(0.050)	(0.080)	(0.052)	(0.065)	(0.046)	(0.057)
企业规模三分位组 × 季度和行业 × 季度							
合同审查后的第4季度	−0.072	0.031	−0.148**	−0.092*	−0.058	−0.093**	0.026
	(0.049)	(0.048)	(0.074)	(0.054)	(0.063)	(0.044)	(0.056)

注：表中报告了基于季度面板数据、采用插补估计量得到的简约式估计，方法见Borusyak、Jaravel 和 Spiess (2024年）。基线模型纳入企业和日历季度固定效应；后续面板则加入按启动前规模三分位组、行业组或两者划分的季度效应。第（1）–（2）列分别为技术人员招聘人数对数和离职人数。第（3）–（5）列涵盖软件工程岗位招聘期，并要求两个资历组的启动前平均就业人数均至少为 1；第（3）列为初级与资深员工对数招聘比。第（6）–（7）列为代码密集度较高的四类职位和代码密集度较低的四类职位的招聘时段对数。标准误按企业聚类，并列于括号内。* $p < 0.10$， ** $p < 0.05$， *** $p < 0.01$。

F GitHub Copilot 对技术人员招聘下降的贡献

本节详细介绍第 6.5 节和图 10中报告的反事实计算。样本由首次预定的上线后合同审查发生在2023年的企业组成。我们将上线前基准招聘 H_{i0} 定义为企业 i在2022年期间的季度技术岗位招聘人数均值，将事后期间招聘 H_{i1} 定义为2024年期间的季度技术岗位招聘人数均值，由此得出每个企业—季度观测到的招聘人数下降 $\Delta = 16.1$。

为计算 $\Delta_{\text{GitHub Copilot}}$，我们按企业—季度计算各事件时间跨度 $\ell = 0$至 4 内相对于各企业计划合同审查季度（方程 ^的事件时钟4）所放弃的招聘人数，并取平均值。在审查季度本身（$\ell = 0$），我们在点估计和每次自助法抽样中都设定 $\hat{\rho}_0 = \theta_0 = 0$，因为该季度的第一阶段较小，Wald比率不稳定；平均值涵盖五个季度。令 $H_{i\ell}$ 表示企业 i在计划审查季度之后 ℓ 个季度的观测到的技术岗位招聘人数，令 $D_{i\ell}$ 表示该季度的采用指标。图 10 中的三个基准计算如下：

$$\hat{\Delta}^{RF}_{\text{GitHub Copilot}} = \frac{1}{5}\sum_{\ell=0}^{4}\frac{1}{N}\sum_{i} H_{i\ell}\left(e^{-\hat{\rho}_\ell} - 1\right) \tag{10}$$

$$\hat{\Delta}^{C}_{\text{GitHub Copilot}} = \frac{1}{5}\sum_{\ell=0}^{4}\frac{1}{N}\sum_{i} D_{i\ell}\, H_{i\ell}\left(e^{-\hat{\theta}_\ell} - 1\right)\frac{\hat{\pi}_\ell}{\hat{P}(D_{i\ell}=1)} \tag{11}$$

$$\hat{\Delta}^{A}_{\text{GitHub Copilot}} = \frac{1}{5}\sum_{\ell=0}^{4}\frac{1}{N}\sum_{i} D_{i\ell}\, H_{i\ell}\left(e^{-\hat{\theta}_{k_{i\ell}}} - 1\right) \tag{12}$$

在方程（10）中，特定时间跨度的简约式系数 $\hat{\rho}_\ell$ 适用于所有企业。在方程（^ 11）中，特定时间跨度的Wald估计值 $\theta_\ell = \hat{}\rho_\ell/\hat{}\pi_\ell$ 适用于采用企业，并乘以^时间跨度 ℓ时采用者中的服从者占比，$\hat{\pi}_\ell/P(D_{i\ell}=1)$, 其中 $\hat{\pi}_\ell$ 为第一阶段效应，$\sim P(D_{i\ell}=1)$ 为时间跨度 ℓ时的样本采用率。由于采用企业的平均规模大于未采用者，按采用者的招聘水平对服从者效应重新加权，所得招聘降幅大于未加权的简约式估计（每企业每季度减少3.1个招聘岗位，而非1.7个）。在方程（12）中，Wald估计值适用于所有采用企业，不按服从者占比缩减。这一假设认为，始终接受者在自身采用后经过 k 个季度所经历的比例招聘效应，与服从者在其评估后经过 k 个季度所经历的效应相同。