

政策研究工作论文

11263

生成式人工智能时代的劳动力需求

美国招聘数据中的早期证据

严刘何王舒宇



WORLD BANK GROUP

数字化转型垂直 & 繁荣垂直 20
25年11月



Reproducible Research Repository

本文的一个已验证的可重复性软件包可在 <http://reproducibility.worldbank.org> 获取，点击[此处](#)直接访问。 [这里](#)

摘要

本文使用在线招聘数据考察了生成式人工智能对美国劳动需求产生的因果影响。利用2022年11月ChatGPT的发布作为一个外生冲击，本文采用双重差分和事件研究设计方案来估计生成式人工智能导致的岗位替代效应。识别策略通过比较ChatGPT发布后高风险与低风险人工智能替代脆弱性职业的劳动需求（同时控制相似的生成式人工智能暴露水平），以将替代效应与协同使用效应区分开来。分析使用了Lightcast从2018年第一季度到2025年第二季度第二季度收集的2.85亿条招聘信息。研究结果表明，招聘数量

对于人工智能替代得分高于中位数的职业，其平均降幅为12%，而得分低于中位数的职业则有所上升。该效应从发布后的第一年6%增加到第三年的18%。对于入门级职位（既不需要高等学位也不需要丰富经验，分别损失18%和20%），以及行政支持（40%）和专业服务（30%）等职位，损失尤为严重。尽管生成式人工智能会创造新的职业并提高生产力，从而可能增加劳动力需求，但早期证据表明，某些职业可能比其他职业不太可能被生成式人工智能所补充。

本文是数字化转型与首席经济学家办公室、繁荣垂直的产品。它是世界银行提供其研究公开访问并就全球发展政策讨论做出贡献的一项更大努力的一部分。政策研究工作论文也发布在<http://www.worldbank.org/prwp>上。作者可通过yanliu@worldbank.org联系。此论文的验证可重复性包可在<http://reproducibility.worldbank.org>获取，点击[这里](#)用于直接访问。



政策研究工作论文系列发布研究进展中的成果，以鼓励就发展问题进行思想交流。该系列的目标是尽快发布研究成果，即使展示的内容尚未完全完善。论文会署名作者，并应相应引用。本文中表达的观点、解释和结论完全是作者的个人观点。它们不一定代表国际复兴开发银行/世界银行及其附属机构的观点，也不一定代表世界银行的执行董事或他们所代表的政府。

生成式人工智能时代的劳动力需求： 美国招聘数据中的早期证据

*

刘岩
†¹
，他王
‡¹
，以及舒宇 1 世界银行
§¹

Authorized for distribution by Kamal M. Siblini, Acting Manager, Digital Transformation Vertical, World Bank Group

JEL代码：O33, J23, J21

关键词：生成式人工智能，技术采纳，劳动力需求，在线

职位发布

^{*} 我们谨感谢Aart Kraay、Daniel Lederman、Franziska Lieselotte Ohnsorge、Jonah Matthew Rexer以及各研讨会参与者提出的有益评论和建议。本文中表达的研究发现、解释和结论完全是作者的观点。它们不一定代表世界银行及其附属机构、或世界银行的执行董事或其代表政府的观点。

[†] 通讯作者。邮箱：yanliu@worldbank.org

[‡] 邮件：hwang21@worldbank.org

[§] 邮件：syu2@worldbank.org

1 引言

生成式人工智能（GenAI）是否已经开始取代工人？什么样的工作面临更大的替代风险？自2022年底ChatGPT发布以来，生成式人工智能的采用已经迅速传播到各国、各行业和各职业（刘和王）[2024](#)；比克，布兰丁和戴明[2024](#)；邦尼等人[2024](#)；刘，黄和王[2025](#)）。随着GenAI能力的不断扩展及其加速应用，人们再次关注由技术驱动的工作岗位流失问题，美国各行各业的商界领袖都发出了警告——从IT到汽车制造业。

¹

人工智能驱动的工作岗位流失的担忧并非空穴来风，新出现的证据表明自由职业者正面临最严重的影响。

²

然而，关于GenAI在总体层面的工作替代效应，目前有限且缺乏定论的证据。

本文利用涵盖2018年第一季度至2025年第二季度的近乎全体的在线招聘数据，研究了GenAI对美国劳动力需求的影响。

³

我们利用2022年11月ChatGPT的公开发布作为外生冲击，并采用双重差分（DiD）和事件研究方法来识别GenAI的劳动力替代效应。我们的分析追踪了随着GenAI能力的提高和采用深化，这些效应如何随时间演变。它还检验了GenAI在不同教育水平、经验要求和行业中的异质性影响。该分析基于Lightcast提供的2.85亿个工作职位信息，这些信息被汇总成一个包含680万个州-行业-职业-季度单元的平衡面板，从而全面展示了GenAI的采用如何重塑美国劳动力需求。虽然本研究侧重于美国数据，但其关于GenAI对工作替代影响的发现对面临GenAI所造成相同技术冲击的其他国家的劳动力市场具有更广泛的启示。

¹ 行业领袖的警告非常明确。2025年5月，Anthropic首席执行官Dario Amodei预测，一半的入门级工作可能在1到5年内消失，可能导致美国失业率上升到10-20%（[链接](#) 一个月后，亚马逊CEO安迪·贾西宣布，随着公司增加对GenAI和“AI代理”的采用，它“在某些岗位上需要更少的人”，并预计由于AI驱动的效率提升，未来几年公司总员工数将缩减（[链接](#) 福特汽车的首席执行官表示人工智能将取代美国一半的白领工作[链接](#)）。

² 参见乔瑞、熊（[2023](#)），会，试剂，和周[2024](#)），布林约夫森，查南达，和陈（[2025](#)），德米里奇，汉娜内，和朱（[2025](#)），和Teutlof等人（[2025](#)）。

³ 美国为这项分析提供了一个尤其强有力的实证环境。首先，它引领全球通用人工智能发展，主要模型（ChatGPT、Claude、Gemini）源自美国公司。这种邻近性促进了更早、更快的国内劳动力市场扩散。其次，美国劳动力市场规模庞大、充满活力且资料详尽，为新技术效应在更广泛扩散前的测试提供了理想场所。第三，美国Lightcast数据已被证明能提供更优的覆盖范围和一致性。它已广泛用于先前关于自动化和人工智能的研究（例如，Acemoglu和Restrepo（[2020](#)），阿西莫格鲁等人（[2022](#)），和Bonfiglioli等人（[2025](#)）），使其特别适合识别新兴的由GenAI驱动的劳动力需求变化。

我们的实证框架结合了两个互补的维度，共同决定通用人工智能如何影响不同职业的劳动需求：通用人工智能接触度，它衡量通用人工智能的理论技术适用性，以及人工智能替代脆弱性，它捕捉了雇主用人工智能替代工人的实际可能性。通用人工智能接触度反映了人工智能对特定任务和职业的有用性和适用程度，但仅凭它本身并不能预测实际的劳动需求效应，因为它将自动化和增强混为一谈。为了解决这个问题，我们纳入了从Pizzinelli等人（2023）—其AI互补性指标的倒数—沿着六个职业维度构建：沟通、责任、物理条件、关键性、常规性和技能要求。该指标考虑了社会偏好和监管限制等情境和非技术因素，这些因素塑造了实际位移风险。与以往仅依赖暴露指数来定义处理组的研究不同，我们的策略比较了在有高于中位数替代脆弱性职业与低于中位数替代脆弱性职业之间，在相似的GenAI暴露水平条件下，职位发布的变动。这个二维框架使我们能够通过保持技术适用性差异和现有自动化潜力的差异恒定，从而分离出劳动力位移渠道。

⁴

我们的结果表明，通用人工智能对可替代性较强的职业的职位发布产生了巨大、统计上显著且加剧的负面影响。在美国，我们估计到2025年年中，在具有高于中位数人工智能替代得分的职业中的职位发布平均比具有低于中位数得分的职业下降了12%，在通用人工智能暴露水平相当的情况下。事件研究分析进一步证实，两组之间的职位发布趋势没有显著差异，这与条件平行趋势（CPT）假设一致（Baker等人）。2025）。发散现象仅在ChatGPT引入后才出现，突显其因果关系。替代效应随着时间的推移而不断加剧，从推出后第一年的6%增长至第三年（截至2025年6月）的18%。这些发现在对职业进行随机分配处理、替代AI指标和不同AI暴露指标作为控制变量的安慰剂检验中依然稳健。异质性分析揭示，入门级职位受到的负面效应尤为显著，这些职位既不需要高级学位（硕士、专业和博士学位），也不需要大量的工作经验（6年及以上），以及专业服务和管理支持行业的职位。

⁴一种类似的二维方法也应用于世界银行（2025a）和世界银行（2025b）。

本文对 GenAI 与劳动力需求的新兴文献做出了几项贡献。

⁵首先，据我们所知，这是首次使用大规模在线招聘数据识别GenAI对长期、稳定就业岗位劳动需求因果影响的研究。现有关于该主题的文献主要集中在对零工经济的探讨上，自由职业平台的研究表明，GenAI降低了写作和翻译等高度可自动化任务的劳动需求与工资。[2025](#)；胡伊，雷谢夫，和周 [2024](#)；特乌特洛夫等 [2025](#)发现顶尖自由职业者受GenAI影响不如其他从业者，可能是因为他们的工资水平较高。相反，我们的研究将重点转向正式的全职工作，这些工作代表了劳动力市场的绝大多数，并且是大多数工人的主要收入来源。由Brynjolfsson、Chandar和Chen进行的密切相关研究([2025](#))利用就业数据发现，高度暴露职业中初级工人的就业率相对下降了13%。我们使用在线招聘数据，因为它们更具前瞻性，并且比就业数据更能捕捉劳动力需求变化的早期信号。

⁶

其次，我们的扩展样本使我们能够追踪GenAI随时间的动态影响。先前使用职位发布数据的研究，例如Ham pole等人 ([2025](#))和leisfeldt等人 ([2023](#))，受限于有限的时间覆盖范围或数据截止到2023年，无法捕捉到不断发展的技术进步和扩散效应。我们的2018-2025数据集能够追踪GenAI的负面效应如何随着能力的提升和采用加速而加剧。

第三，我们的识别策略通过结合职业的GenAI适用性（GenAI暴露度）及其对AI替代的实践脆弱性，隔离了GenAI的位移效应。具体来说，我们在GenAI暴露度相同的情况下，比较了具有高于或低于中位数替代潜力的职业的劳动需求如何演变。这一设计使我们能够区分GenAI的因果影响与仅仅由于技术适用性变化而产生的差异化雇佣趋势。尽管一些论文试图区分自动化

⁵关于生成式人工智能（GenAI）对劳动力市场影响的实证文献正迅速扩展。虽然一些实验研究已记录了从GenAI采纳中获得的生产力大幅提升（Brynjolfsson, Li和Raymond [2025](#)；崔等人。[2024](#)；诺伊和张 [2023](#) 等人；[2023](#)），其他研究发现对工资和工作时间的影响混合或微乎其微（Jiang等人 [2025](#)；humlum和vestergaard [2025](#)；查南达 [2025](#)；哈莫尔等 [2025](#)）。

⁶此外，Brynjolfsson、Chandar 和 Chen ([2025](#)) 覆盖不到美国劳动力的15%，而我们的在线职位发布数据更具代表性，自2010年以来始终捕捉到美国60%的职位空缺。

以及使用任务级聚合的增强效果 (Eisfeldt等人 2023 ;布林约夫森，钱德拉，和陈 2025), 我们的方法 - 基于 Pizzinelli 等人 (2023) - 依赖于工作环境特征和职业技能要求来捕捉现实中的替代风险。现有的依赖任务暴露得分、任务级别得分的标准差或来自用户与 GenAI 工具交互的主观评价的方法，并未完全捕捉到任务与监管和社会因素之间的联系。因此，这些方法可能高估了 GenAI 的自动化潜力。我们的替代指标包含了影响 GenAI 采用的监管、社会和情境因素，从而对 GenAI 在短期到中期内可能替代劳动力的地点进行更现实的评估。

第四，我们的异质性分析为GenAI采用的分布性后果提供了新的见解。我们表明，负面影响不成比例地集中在入门级和技能较低的工作者中，而经验更丰富的工人受影响较小。这些负面影响在行政和支持服务、专业服务和公司管理中也更明显。这些结果与最近的研究证据 (Bryn-jofsson, Chandar, and Chen) 一致。 2 025 ;特乌特洛夫等 2025), 这表明人工智能不成比例地减少了入门级和低技能工人的需求。我们的研究强调了考虑有针对性的支持以减轻技术冲击对弱势群体潜在影响的必要性。

本文的其余部分安排如下。第 2 描述了分析中使用的數據，澄清了人工智能暴露、增強和自動化等關鍵概念；並呈現了描述性證據。部分 3 概述了经验策略。节 4 报告了基线结果。第 5 执行了一系列稳健性检验。第 6 考察跨工人特征和行业的异质性位移效应。第 7 总结，讨论政策启示，并阐明未来研究的潜在途径。

2 数据和描述性证据

在本节中，我们首先介绍了用于我们分析的Lightcast职位发布数据。然后，我们概述了AI暴露度与替代指标，并阐明与我们框架相关的关键概念区别。最后，我们呈现了关于劳动力需求如何分化的描述性证据。

具有高和低AI替代脆弱性的职业。

2.1 来自 LightCast 的职位发布数据

在这篇论文中，我们使用 Lightcast（前身为 Burning Glass）提供的实时职位发布数据来捕捉美国劳动力需求。Lightcast 从全球的公司网站和在线招聘板上收集数据。⁷

它使用去重算法处理原始数据，并进一步使用算法从详细的职位描述中清理和提取关键信息。它报告发布日期、工作地点、公司名称、职位名称、招聘公司的行业（北美产业分类2位代码水平）、职业（国际标准职业分类08年4位代码水平）以及教育和经验要求。现有研究表明，职位发布数据比标准劳动力调查或职位空缺调查更能提供及时和更细粒度的洞察力，了解雇主的需求侧行为，使其更适用于分析对外部冲击的劳动力需求。

⁸

在美国，Lightcast数据自2010年以来涵盖了超过4.35亿个工作职位。如图所示¹，光速始终捕捉到就业岗位和劳动力流动调查（JOLTS）中记录的60%左右的职位空缺，并紧密追踪职位空缺的整体演变。

⁹

此外，以往研究表明，Lightcast数据中的职业和行业构成与职业就业统计（OES）中发现的数据高度一致，这使得数据足以满足我们研究的目的。为了捕捉ChatGPT于2022年11月发布前后的劳动力需求动态，我们使用了从2018年1月1日到2025年6月30日的高频每日招聘数据，这使我们能够建立长期的前处理趋势并观察发布后的发展；这些数据被汇总到季度层面进行分析。

⁷总共，数据集自2010年以来涵盖了超过25亿个职位，分布在超过150个经济体。

⁸参见赫什贝因和卡恩（2018），戴明和诺雷（2020），阿西莫格鲁等人（2022），福赛特等（2022），和 Braxton 和 Taska（2023）。

⁹这与以往的研究发现一致，例如赫什贝因和卡恩（2018）和Acemoglu等人（2022）。

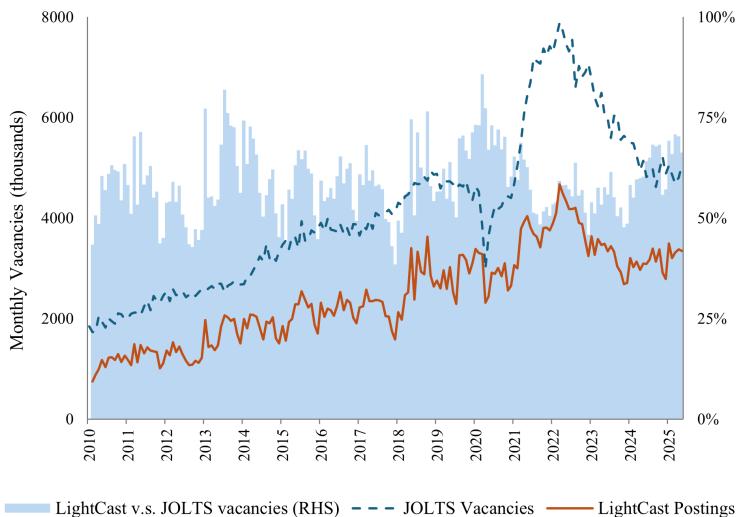


图1：比较Lightcast数据与职位空缺和劳动力流动调查

注意： 该图展示了从2010年到2025年6月Lightcast和美国职位空缺与劳动力流动调查 (JOLTS) 记录的职位空缺数量。Lightcast密切跟踪美国整体空缺的演变，并捕捉了大约60%的总空缺。

表 A1 展示了按职业、行业划分的工作岗位数量，以及教育和经验要求。在我们的样本期间，一半的工作岗位是为专业人士 (31%) 和技术人员 (19%)，经理 (11%) 和销售人员 (16%) 合计占超过四分之一。在包含行业信息的80%的岗位中，岗位主要集中在医疗和社会援助 (12.4%)、行政和支持以及废物管理 (12.2%)、零售贸易 (9.0%)、专业、科学和技术服务 (9.0%)，以及制造业 (6.1%)。同时，发布岗位最少的四个行业——农业、采矿、公司管理服务和公用事业——合计仅占所有岗位的1.1%。

绝大多数 (59%) 的职位招聘不指定教育要求。在这些指定了要求的职位中，四分之一要求至少拥有高中文凭，而10%和6.4%分别要求拥有大学学历和高等学位（包括硕士、专业学位和博士学位）。类似地，大多数职位 (55%) 也不声明经验要求。对于这些指定了经验要求的职位，24%寻找具有0-2年经验的候选人，15%寻找具有3-5年经验的人才，而6%要求至少6年经验。这些详细的职位要求使我们能够在第几节中考察GenAI在不同行业、教育和经验群体中的异质性效应 [6](#)。

2.2 基因人工智能暴露和人工智能替代测量

如前所述，两个互补维度共同决定了GenAI如何影响各职业的劳动力需求：GenAI暴露反映了GenAI可能影响工作的领域，而AI替代脆弱性则基于固有的职业特征、社会偏好和监管框架，捕捉了某职业被AI现实中取代的概率。

研究人员已经开发出各种职业级别的AI暴露指数，以估计哪些任务和职业最有可能受到AI的影响（Frey和Osborne）[2017](#)；费尔滕，拉杰，和希曼斯[2018](#)；费尔滕，拉杰，和希曼斯[2019](#)；韦伯[2019](#)；费尔滕，拉杰，和希曼斯[2021](#)；皮齐内利等人。[2023](#)），基于在Acemoglu和Autor（[2011](#)），作者（[2015](#)），以及阿西莫格鲁和雷斯特雷波（[2019](#)）。

¹⁰近期研究将这些框架调整以更具体地关注生成式人工智能的接触。

¹¹

研究人员通常从O*NET数据库中获取信息

¹²并要求人类专家或GenAI系统判断每个任务可以在多大程度上由AI执行或辅助。然后他们汇总任务级分数，以生成职业级暴露分数（Felten、Raj和Seamans [2023](#)；埃隆杜等 [2024](#)；艾斯费尔德特等 [2023](#) 贝内特斯-鲁埃达和帕拉多 [2024](#)；Gmyrek、伯格和贝斯孔德 [2023](#)；格雷雷克等 [2025](#)）。一些方法进一步使用现实世界用户交互数据。例如，Anthropic 经济指数（Handa 等人）[2025](#)）和微软的AI适用性评分（Tomlinson等人）[2025](#)）分析数百万与 GenAI 工具的对话，以识别使用最集中的领域。一些指标尝试按任务和职业分解自动化和增强分数。

本文主要使用 Gmyrek、Berg 和 Bescond 开发的 GenAI 暴露评分 [2023](#)）（GBB）由于以下原因。首先，GBB 分数是首个广泛使用的 GenAI 暴露指标之一。一些其他常用的 AI 暴露度量，如 Felten、Raj 和 Seamans ([2021](#)) 和 Pizzinelli 等人 ([2023](#))，都不特指生成式人工智能。其次，GBB

¹⁰基于任务的模型被广泛用于分析技术如何影响劳动力需求。该框架强调，人工智能等技术并非直接在职业层面影响劳动力需求，而是通过它们对构成职业的任务的影响来发挥作用。每种职业都由一系列任务组成。技术可以自动化或增强现有任务，并创造新任务。这些力量之间的平衡决定了技术是降低还是提高总劳动力需求。

¹¹参见，例如，Felten、Raj 和 Seamans ([2023](#))，埃尔朗杜等人 ([2024](#))，benitez-rueda 和 parrado ([2024](#))，gmyrek、berg 和 bescond ([2023](#))，gmyrek 等人 ([2025](#))，手合等人 ([2025](#))，和汤姆森等人 ([2025](#))。

¹²该数据库包含一套丰富的变量，描述了与约1000个美国职业相关的工作和工人特征（包括技能要求）

分数在ChatGPT公开发布后不久发布，并且没有受到模型功能扩展和GenAI日益增长的采用的污染，这与它最近更新的版本（即Gmyrek等人 2025）。第三，大多数现有的AI风险敞口衡量指标之间存在高度正相关关系。我们在第 5 节中使用这些替代的AI风险敞口指标作为控制变量。 5.2 为了测试我们基线结果的鲁棒性。

我们使用由Pizzinelli等人反向推导的AI替代分数来识别治疗职业(2023)。他们的衡量指标由五个工作-环境特征——沟通、责任、物理条件、关键性和常规性——以及基于O*NET的复合技能维度构成。在这些维度（除“常规性”外）得分高的职业更有可能被人工智能补充，而不是被取代。逆转这个量表使我们能够捕捉雇主实际用人工智能替代工人的可能性，考虑到不仅技术可行性，还包括职业特征、工作场所环境和社会偏好。虽然原始衡量指标并非专门为生成式人工智能设计，但它关注与任务无关的工作属性，因此非常适合我们的分析。在 5，我们进一步分解这个分数，以考察哪些职业特征可以抵御替代风险，并将其与人际交往技能强度等替代性替代措施进行比较。

我们更倾向于这种整体替代指标，而非基于任务的自动化或增强指标，原因有几点。首先，职业并非固定任务包，而是动态发展的活动集合，在不同公司、行业和国家间存在差异。像 O*NET 这样的任务分类体系往往遗漏工作中关键但无形的一面——例如，O*NET 详细描述了卡车司机的物理活动，但并未完全捕捉协调与调度员或应对中断等任务。此外，当任务相互依存时，自动化或增强某个环节并不必然意味着整个职业都可以被自动化或增强。新任务如何出现、如何分布到现有职业中，或是否会催生全新职业，这些都仍高度不确定。

其次，除了定义和测量方面的挑战之外，将任务层面的影响聚合到职业层面的劳动力需求需要通过更广泛的经济机制进行追溯。劳动力需求是由一个职业所生产的商品和服务的需求得出的。产品需求本身受到多种因素的影响，包括人工智能的出现。任务层面的技术冲击的影响受到复杂的总体均衡效应的影响，

依次通过各行各业塑造涟漪效应。

第三，位移的速度和程度取决于非技术因素。容忍度高（例如，数据录入员）的职业比责任重大或几乎没有出错空间（例如，外科医生或司机）的职业更容易被取代。需要社交互动、同理心和解决问题的职业——如教学——也存在更大的抵抗力。此外，监管限制、社会信任和消费者期望对采纳产生了强烈影响，而这些因素均未在任务层面的暴露分数中体现。这些局限性解释了为什么基于任务的自动化和增强分数通常无法预测职业层面的结果。相比之下，我们的替代分数更准确地衡量了GenAI可能取代劳动力的领域。

图 2 通过将职业的 GenAI 暴露分数（技术适用性）与 AI 替代分数（基于职业特征的脆弱性）绘制在一起，图 1 说明了这两个维度的互补性。具有相同暴露水平的职业表现出广泛不同的替代潜力。在右上象限，具有高暴露和高替代分数的职业——包括一般秘书、应用程序员以及统计、财务和保险文员——面临着巨大的自动化风险。相反，右下象限包含暴露度高但替代潜力低的职业，例如生命科学技术员和大专院校教师，这些职业更有可能被 AI 增强而不是直接取代。左上象限揭示了像建筑结构清洁工和织纺机操作员这类职业，尽管 GenAI 暴露度低，但其替代潜力很高。最后，左下象限包括像警察和牙医这类职业，它们表现出较低的 GenAI 暴露度和最小的替代风险。

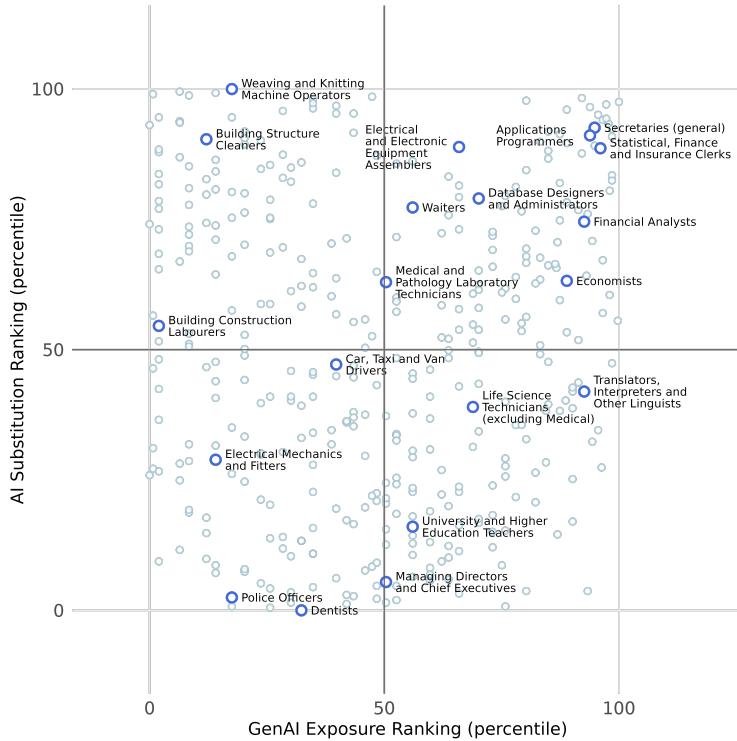


图2：各职业的AI替代和GenAI接触

注意：x轴展示了Gmyrek、Berg和Bescond（的研究）中GenAI暴露指数的百分位数排名（2023）(GBB)。y轴展示了AI替代得分的中位数排名，该得分被定义为Pizzinelli等人（中的AI互补性措施的倒数。2023）。人工智能互补性度量沿六个职业维度构建：沟通、责任、物理条件、关键性、常规和技能要求。典型职业被突出显示。

一个主要问题是，即使在ChatGPT推出之前，具有高和低AI替代脆弱性的职业可能已经遵循了不同的雇佣轨迹，这可能会削弱我们DiD设计中所需的平行趋势假设。如果将高AI暴露的职业与低暴露的职业进行比较，这种风险会更加突出，因为它们的潜在采用方式不同。为了解决这个问题，我们基于具有相似GenAI暴露分数的职业来进行我们的分析。通过保持暴露度不变，我们确保处理组和控制组有同等可能采用GenAI，从而将取代效应与更广泛的技术暴露差异分离开来（参见第3（详细信息））。

2.3 描述性证据

图3绘制了在线职位发布趋势图，涵盖高于和低于中位数AI替代分数的职业，数值在ChatGPT发布季度（2022Q4）归一化到1。在那之前

点，两组的Posting平行移动，但之后它们的轨迹急剧分歧。

¹³

自ChatGPT推出以来，高替代率职业的职位发布量下降幅度比低替代率职业更陡峭。尽管归一化削弱了测量幅度，但差异仍然很明显。

¹⁴

仅此分化时间本身并不能建立与生成式人工智能的因果关系。其他因素，如2022年的利率上升以及疫情后白领招聘的常态化，也可能导致这种分化。为区分生成式人工智能的影响与这些混杂趋势，下一节提出了我们的实证策略，概述了双重差分法和事件研究规范。

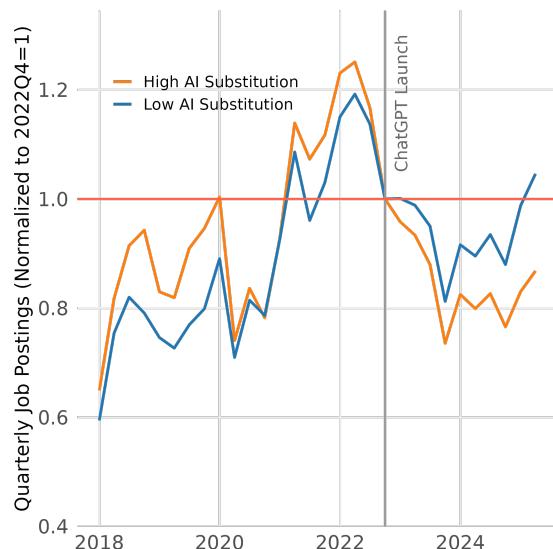


图3：人工智能替代分数高低的不同职业的职位发布趋势标准化图

注意：这张图展示了LightCast中高于和低于中位数AI替代得分的职业的在线招聘趋势。招聘数值已在ChatGPT发布的季度（2022Q4）标准化为1。AI替代得分定义为Pizzinelli等人（中）中AI互补性指标的倒数。[2023](#)，沿着六个维度构建：沟通、责任、物理条件、关键性、常规性和技能要求。工作职位的原始趋势报告在附录表格中。[A1](#)。

¹³

图A1展示了原始的季度发布计数，并证实两组在2022Q4之前处于相似水平。

¹⁴

图A2突出了两种职业的对比：秘书（高暴露、高替代）和牙医（低暴露、低替代）。两者从2018年到2022年都经历了稳定的需求增长，但在ChatGPT发布后，它们的轨迹分叉了：牙医的需求增长了近50%，而秘书的职位发布则下降了大约50%。

3 经验设定

我们的分析聚焦于职业层面的处理差异，评估了GenAI在2022年11月ChatGPT公开发布前后的影响。我们按职业和季度汇总了地方劳动力市场的职位发布数量，这些市场被定义为51个省级单位的横截面、2位数的NAICS层面20个行业以及近400个4位数的ISCO层面的职业。我们的分析单位选择与既定的理论模型（Topel 1986；莫雷蒂 2010），这表明劳动力需求受到当地市场状况的影响，并且与考察当地劳动力市场的实证研究密切相关（例如，Hershbein和Stuart（2024））。与在企业层面考察人工智能影响的研究（Babina等人）相比（2023；阿西莫格鲁等人 2022），我们的方法有两个关键优势。首先，它避免了样本期内企业进入或退出市场可能造成的潜在选择偏差。其次，它解决了过多的零值问题，因为企业可能不会每年都为每个职业招聘。通常，将职位发布汇总到地方市场层面可以平滑这种企业特定的波动，并提供更可靠的系统性劳动力需求变化估计。

基线规范。 我们的基准识别策略采用双重差分（DiD）框架，利用职业对人工智能替代的脆弱性差异来定义处理组，并将2022年11月ChatGPT的发布视为一个外生冲击。虽然发布时间很大程度上出乎意料，但ChatGPT的发布随后被快速采用，因此有必要在我们的模型设定中控制GenAI暴露水平。因此，我们的基准模型设定形式如下：

$$\text{arsinh} (Y) = \alpha D \times 1_{Post} + \beta X + \theta + \gamma + u_{sot o \text{ 坐卧站 } i(o) t t} \quad (1)$$

其中 Y 表示该州的职位发布数量 s ，行业 i 职业 o ，以及四分之一

t 这里，我们包含没有观察到招聘信息的行业-职业-季度单元格中的零计数。¹⁵

我们应用反双曲正弦（ arsinh ）转换来处理零值的出现，同时保持可解释性。该转换对大值近似自然对数，但在零处保持良好定义，允许系数被解释为

¹⁵

我们保留至少有一个职位的行业-职业单元，并对其余季度填充零，从而丢弃样本期内没有职位的单元。

职位发布量的百分比变化，类似于对数转换。

¹⁶

关键解释变量是交互项 $D \times 1$

P_{ost}

. 治疗指标，

o_t

D ，是一个二元变量，对于具有高于中位数的AI替代分数的职业等于一

o

(i.e.,治疗组)且其他情况为零。事后处理期指标，1

P_{ost}

是

t

也是一个虚拟变量，在ChatGPT发布后的季度等于1（包括2022Q4，ChatGPT发布的季度）以及在更早的时期等于0。感兴趣系数， α ，识别了ChatGPT发布对具有高与低人工智能替代脆弱性职业的职位发布所带来的差异影响。

我们包含三组固定效应作为控制。首先，州行业季度固定效应(θ)

θ 被添加以吸收所有可能影响州-行业层面地方劳动力市场劳动力需求的时变因素。

¹⁷

其次，我们控制了州-行业-职业固定效应(γ)，

sio

它们捕获了特定职业-地点-行业三元组的所有时间不变特征。这些固定效应捕捉了基准职位发布水平的持续差异、地方比较优势以及其他影响工作地理和部门分布的未观测因素。

最后，我们在职业层面控制“按季度分组的GenAI暴露十位数固定效应”（即， χ ，其中 $j(o)$ 表示职业的十分位数排名 o 基于我们的首选 GenAI

$j(o)_t$

暴露指数）。GenAI 对劳动力需求的影响可能因其 GenAI 暴露分数而异。如果控制不当，这种差异会导致对迪尔方法 (DiD) 所必需的平行趋势假设的违反。因此，我们依赖于贝克等人 (Baker et al. 提出的条件平行趋势 (CPT) 假设。[2025](#))，其中我们比较了在同一代人工智能 (GenAI) 暴露指数（即，GBB分数）的同一十分位数排名中具有高和低AI替代得分的工作职业。这些控制因素吸收了影响所有具有相似总体GenAI 暴露水平的职业的普遍趋势，例如人工智能采用浪潮和对人工智能暴露型工作的需求变化。在这种设定下，参数 α 识别AI替代效应

¹⁶

具体而言 $arsinh(X) = \ln X + X^2 + 1$ $arsinh$ 为零时 X 为零，且它接近于 $\ln(X)$ ，当 X 很大。这种方法广泛用于计数数据，包括像Acemoglu等人([2022](#))。

$arsinh$ 函数，类似于 $\ln(1+X)$ ，是一个良好定义的函数，它在零处可微，其导数在 X 较大时近似于 $\ln(X)$ 。尽管如此，最近的研究工作（例如，陈和罗特 ([2024](#))))显示了对数型变换，包括 $arsinh$ ，处理零时可能导致依赖于单位且难以解释的效果。尽管如此，在我们的设定中，具有较大的样本量和高维固定效应 $arsinh$ 仍然是一种近似处理效应的可行方法，作为捕捉因变量中比例变化的合适转换。

¹⁷

这类因素可能包括地方性产业特有的商业周期、国家政策变化以及季节性就业模式

基于 GenAI 并将其与整体曝光效果分离。

动态效果。 随着 GenAI 模型能力与采用随时间演进，我们修改了基准规范，并使用以下事件研究设计来估计 GenAI 的动态处理效应：

$$\begin{aligned} & X \\ & \text{arsinh}(Y) = \alpha D \times 1 \\ & k \\ & + X + \theta + \gamma + u \\ & \text{slot } o \text{ 坐卧站 } k j(o) t tk = -1 \end{aligned} \tag{2}$$

哪里 1

k 是 ChatGPT 发布后第 k 个季度的虚拟指示（负值表示

t 在 ChatGPT 之前时期），以发布前的季度 ($k = -1$) 作为省略的周期作为参考。系数 $\{\alpha\}$ 追踪动态治疗效果。为了评估

k 对于年度效应，我们还估计了一个模型，该模型将 ChatGPT 发布后的每个季度合并为一个单元，同时将所有 ChatGPT 之前的季度作为参考。

¹⁸

在整个分析过程中，我们根据 2017 年各州-行业-职业单元中的职位发布数量对观察值进行加权，以确保我们的估计反映有意义的劳动力市场细分。标准误差在职业-季度层面聚类，以解释职业类别内部随时间推移的潜在序列相关性和共同冲击。

¹⁹

4 实证结果

本节分为两部分介绍我们的实证发现。首先，我们报告了 GenAI 在劳动力替代效应方面的基准双重差分 (DiD) 结果。其次，我们呈现了事件研究的结果，这些结果既支持了我们的条件平行趋势 (CPT) 假设，又揭示了 GenAI 影响的季度动态。

4.1 基准结果

表 1 报告了我们主要的双重差分估计值，即生成式人工智能对劳动力需求的影响。我们报告了跨越三个逐步严格的规范的结果，这些规范添加了固定效应来解决潜在的偏误问题。

¹⁸这种方法采取以下形式：

$$\begin{aligned} & + \alpha D \text{ arsinh}(Y) = \alpha D \ 1 \times \times 1 \times \\ & \text{Post 3 yr Post 2 yr Post 1 yr} \\ & + X + \theta + \gamma + u + \alpha D \ 1 \\ & \text{坐卧站 3 o 2 o slot 1 o j(o) t t t t} \end{aligned} \tag{3}$$

¹⁹

作为鲁棒性检验，附录图 A3 和 A4 分别展示无加权以及职业层面聚类的结果。它们与基线结果基本一致。

创立因素。前三列遵循公式(1)并展示ChatGPT发布后(即从2022年最后一个季度到2025年中期)治疗组和控制组之间职位发布的平均差异。治疗组包括AI替代得分高于中位数的职业，而控制组包括AI替代得分低于中位数的职业。最后三列显示ChatGPT发布后第一、第二和第三年的动态影响，如公式(3)。

第(1)列和第(4)列仅呈现控制了州-行业-职业固定效应和季度固定效应的估计值。前者固定效应吸收了特定职业-地点-行业组合的所有时间不变特征，而后的固定效应控制了随时间变化的因素(如宏观经济冲击、疫情效应和利率上升)。根据这一设定，ChatGPT的推出导致从2022年末至2025年6月，那些具有相对较高替代脆弱性的职位发布平均减少了23%(第(1)列)。第(4)列的动态估计显示，这种替代效应随着时间的推移而加剧，从推出后的第一年16%增长到第二年25%，并在第三年达到31%。

列(2)和(5)用州-行业-年份-季度固定效应替代了总时间固定效应。这种更严格的设定控制了所有在特定州-行业中均匀影响劳动力需求的时变因素。虽然这些额外的控制在一定程度上削弱了我们的估计，但对高替代性行业的负面影响在经济和统计上仍然显著，平均为18%。动态模式持续存在，效应从第一年的12%增长到第二年的20%和第三年的25%(列(5))。

我们首选的规格，如第(3)列和第(6)所示，增加了“按季度划分的GenAI暴露度十位数固定效应”。这种方法使我们能够比较AI替代得分高于中位数和低于中位数的职业，且均在相同的整体GenAI暴露度水平下进行比较。此方法使用Gmyrek、Berg和Bescond的GenAI暴露度得分(2023)，提供了对位移对劳动力需求影响更精确的估计。

在此首选规格下，ChatGPT的推出显著降低了具有高于中位数的AI替代分数的职位的招聘信息，平均降低了12%。同样，这种效应随时间增强，从第一年的6%上升到第二年的14%和第三年的18%。我们使用第(3)列和第(6)列的规格作为后续所有分析的基础

分析。

表1：ChatGPT对劳动力需求的影响

因变量 arsinh(# 职位发布)	(1)	(2)	(3)	(4)	(5)	(6)
高取代率 occ. × 喜帖GPT后 -0.232*** -0.183*** -0.122*** (0.066) (0.05) (0.036)						
高取代率占位符 × 术后第1年 -0.155* -0.121* -0.064* (0.064) (0.047) (0.031)						
高代取代率×后续第二年 -0.252*** -0.197*** -0.135** (0.072) (0.055) (0.041)						
高替代率从业 × 服第三年 -0.310*** -0.247*** -0.182*** (0.064) (0.051) (0.04)						
季度FE 是 是 国家×行业×职业 FE 是 是 是 是 是 国家×行业×季度固定效应 是 是 是 GenAI经验十级×季度效应 Yes Yes 数.观察值 6,753,930 6,753,930 6,753,930 6,753,930 6,753,930 6,753,930 R平方 0.959 0.970 0.971 0.959 0.970 0.971						

注意：以职业-季度层面聚类稳健标准误括号内显示。

* $p < 0.1$,

** $p < 0.05$,

*** $p < 0.01$.

回归单元处于行业-职业-区域层面，包括零计数的单元格。因变量是职位发布的反正弦 (arsinh) 变换，该变换近似对数变换，在职位发布为零时等于零，允许系数被解释为职位发布近似对数变化。“高替代率职业”是表示处理组的虚拟变量，定义为具有中位数以上AI替代分数（或Pizzinelli等人(2023)）。“ChatGPT发布后”是一个虚拟变量，在ChatGPT发布后的季度等于1，包括2022年第四季度，即发布季度。“第一年后”、“第二年后”和“第三年后”分别指2022年第四季度至2023年第三季度、2023年第四季度至2024年第三季度和2024年第四季度至2025年第二季度。“GenAI经验十级数”是基于Gmyrek、Berg和Bescond的GenAI风险暴露指数排名的一系列虚拟变量 (2023).

4.2 事件研究结果

图 4 提出基于公式 (方程) 规范的事件研究估计值 (2). 该图展示了在控制三组固定效应 (如第 (3) 列和第 (6) 列所示) 后，具有高于中位数AI替代分数的职业与低于中位数分数的职业之间职位发布的季度差距，蓝色点是我们样本期间估计的差距，尖峰捕捉相应的95%置信区间。结果证实了表第 (3) 列和第 (6) 列所示的影响。 1 是因果关系，并详细展示了GenAI对劳动力市场的动态演变。

预处理验证。 图 4 显示在ChatGPT推出之前，估计值在零附近波动，没有系统趋势，表明高低替代职业表现出相似的劳动力需求轨迹。这一视觉证据由正式的预趋势检验 (Wald统计量=0.83, p值=0.66) 证实，该检验未能拒绝共同预趋势的零假设。这验证了我们的实证识别策略，该策略涉及

在每十分位数人工智能接触水平的范围内，将具有高于中位数和低于中位数人工智能替代分数的 occupation 分别作为处理组和控制组。

处理后动态。 从图可知⁴ 在chatgpt推出后的第一季度，治疗效果立即转为负面。在2023年第一季度，人工智能替代分数高于中位数的职业招聘数量相对于对照组下降了约5%。这种初始影响随着时间的推移而系统性地加剧，到2023年年中（Q=3）加深到约10%，到2024年底（Q=8）达到15%，到2025年第二季度达到约20%。

²⁰

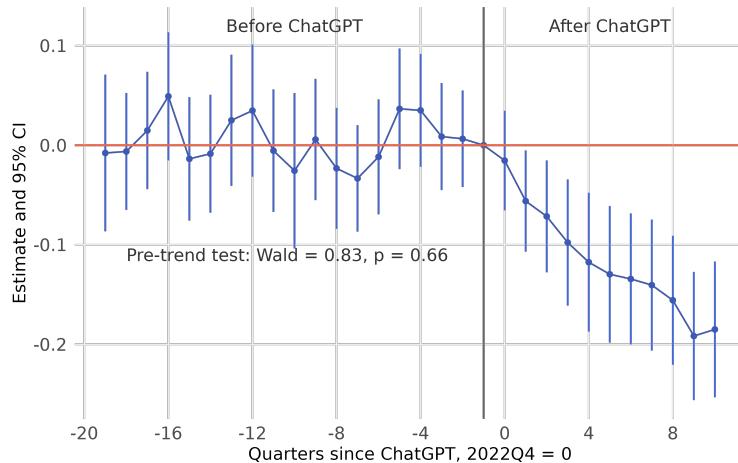


图 4：事件研究：基于 GenAI 暴露的“替代”效应

注意： 该图描绘了基于方程中事件研究规范的点估计和95%置信区间² 参考期是ChatGPT发布之前的季度（Q=-1）。点代表治疗组和控制组在职位发布上的季度差距。治疗组被定义为具有高于中位数AI替代分数（或AI互补分数，如Pizzinelli等人（[2023](#)））。回归单元处于省-行业-职业-季度层面，包括零计数的单元格。该设定控制了省-行业-职业固定效应、省-行业-季度固定效应，以及一系列基于Gmyrek、Berg和Bescond的Gen AI暴露指数组十位数排名的虚拟变量。[2023](#)），与季度虚拟变量交互。

动态效应显示了GenAI对劳动力需求的三种显著特征。首先，负面效应的即时显现表明劳动力市场对新技术的可用性做出了快速反应，这与延迟调整的说法相矛盾（Brynjolfsson, Rock, and Syverson [2017](#)；Hap аянан и Kapur [2025](#)）。chatgpt的快速采用可能会导致

²⁰

为了检验治疗强度效应，我们进一步使用第50百分位数和第90百分位数作为阈值，将高度替代性职业细分为两组。图A5将其对中等偏上替代职业（50-90百分位数）和最高替代职业（90-100百分位数）相对于中位数以下职业的影响进行比较。作为稳健性检验，图A6 使用上述中位数AI暴露分数来识别治疗组。

归功于立竿见影的效果。

²¹

其次，快速且加剧的响应也可以归因于招聘趋势更具灵活性和前瞻性，因为新增招聘通常比现有就业更容易调整，而现有就业则需要裁员或组织结构调整。第三，持续向下的轨迹可能反映了人工智能工具的逐步组织采用以及模型能力的持续改进。

负面影响的程度凸显了生成式人工智能的潜在破坏力。对于人工智能替代得分高于中位数的职业的职位发布数量下降了近20%，这在经济上具有重要意义，与近期研究中的估计一致（Brynjolfsson、Chandar和Chen²²；胡伊、雷谢夫、周²⁰²⁴；艾斯费尔德特等²⁰²³）。值得注意的是，这一收缩发生在Chat GPT发布后的短短三年内，凸显了技术颠覆的快速步伐。为了进行比较，Acemoglu和Restrepo^(2020)发现，每千名工人中增加一个机器人，在1990年至2007年间，仅使美国的就业人口比例降低了0.18-0.34个百分点，而Acemoglu等人^(2022)报告称，在2010年至2018年间，没有发现早期人工智能接触对就业或工资有任何可辨别的冲击。与机器人技术或外包不同，就业人数的下降是逐渐积累了几十年，而生成式人工智能的影响是快速、集中，并且不成比例地打击白领工作。这种影响的速度和规模表明，在短时间内，生成式人工智能可能会深刻地重塑劳动力市场，强调快速政策响应和教育培训体系及时适应的必要性。

5 鲁棒性检验

为确保我们主要发现的可靠性，我们进行了两组稳健性检验。首先，我们通过运行一系列安慰剂检验并使用²³我们的识别策略依赖于对GenAI暴露水平的适当控制（参见第 5.2）。

²¹

人工智能正以历史上任何其他技术都比不上的速度传播，ChatGPT 成为仅用两个月就最快达到1亿用户的 ứng dụng。到2025年年中，美国成年人中超过三分之一的人已经使用过ChatGPT（[链接](#)），并且该工具在2025年8月达到了7亿周活跃用户，占全球劳动力人口的近20%（[链接](#)）。

5.1 治疗组的选取

5.1.1 安慰剂测试

为了验证我们的治疗组别是否识别得当，我们通过随机分配职业给对照组和治疗组来实施安慰剂检验。我们进行了400次安慰剂试验，在每次迭代中随机将所有职业中的一半分配到一个“伪治疗组”，忽略其实际的AI替代得分。对于每次随机试验，我们重新估计了表6中(6)列的优选规格。[1](#)。

[图 5](#)展示了三个治疗后时期的分析结果。每个面板显示一个治疗后年份（灰色直方图）的安慰剂系数分布，以及我们的基准估计值（红色垂直线）。安慰剂分布近似正态分布，并围绕零值中心，这与随机分配下无系统性治疗效应的零假设一致。相比之下，我们的实际估计值远落在这些分布的左侧尾部。对于第1年，我们的系数-0.06代表安慰剂分布的6.5分位数。对于第2年，基准系数-0.14位于0.25分位数处，而对于第3年，我们的系数-0.18明显超出了随机分配下产生的系数范围。

我们将稳健性检验扩展到我们的事件研究结果上。我们使用了相同的随机化程序进行了400次安慰剂试验，并重新估计了方程([2](#))中详细说明的模型设定。[2](#)。[图 6](#)将我们的基线事件研究结果表明（粗蓝线）与安慰剂估计的完整分布（浅蓝线）叠加在一起。与我们基线结果不同，安慰剂估计在治疗前和治疗后时期大多集中在零附近，表明没有系统性模式或趋势。与此同时，我们的基线估计在治疗前时期围绕零波动，这与平行趋势假设一致。更重要的是，它们在ChatGPT发布后（以灰垂直线标记）表现出明显的下降趋势，低于零并且低于安慰剂测试估计的下限。

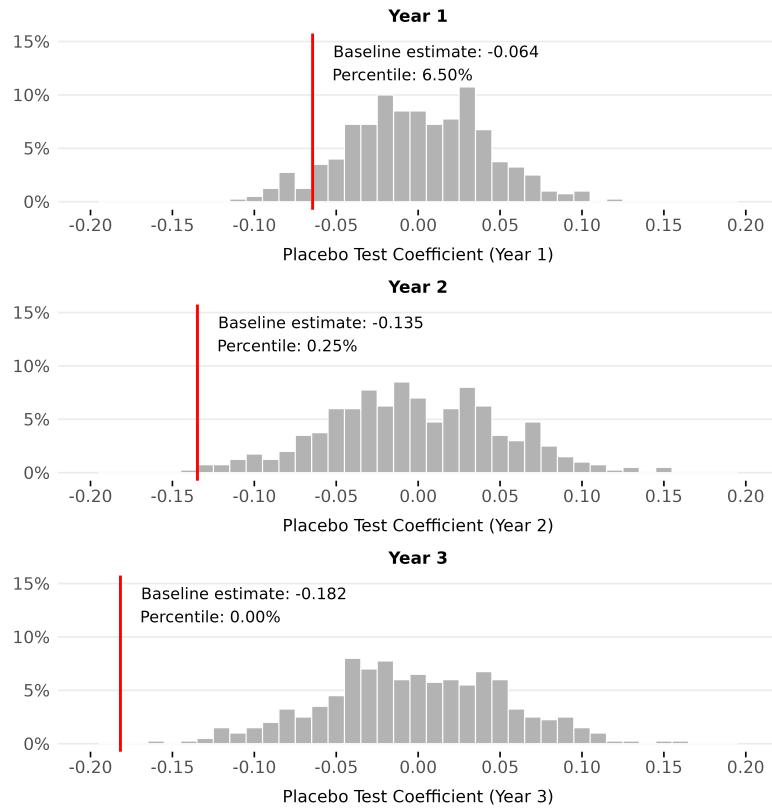


图5：安慰剂检验：双重差分模型设定

注意： 在每个面板中，红色垂直线表示基线系数，如表6第6列所示 1 对于每个治疗后年份，灰色条形图显示了从400次安慰剂测试中，随机将一半所有职业分配到“伪治疗”组，并根据公式估计的每个治疗后年份的系数分布 3 .

这些安慰剂检验的结果在很大程度上验证了我们对处理组的识别，因为随机处理分配在相同的模型设定下产生了零效应。最重要的是，安慰剂检验证实了ChatGPT的发布对美国的那些具有高于中位数的AI替代分数的职业的需求产生了不成比例的影响。处理效应随时间的增强——这在我们的基准结果中可见，但在安慰剂检验中不存在——为我们在捕捉GenAI采用对劳动力需求的因素影响提供了强有力的证据。

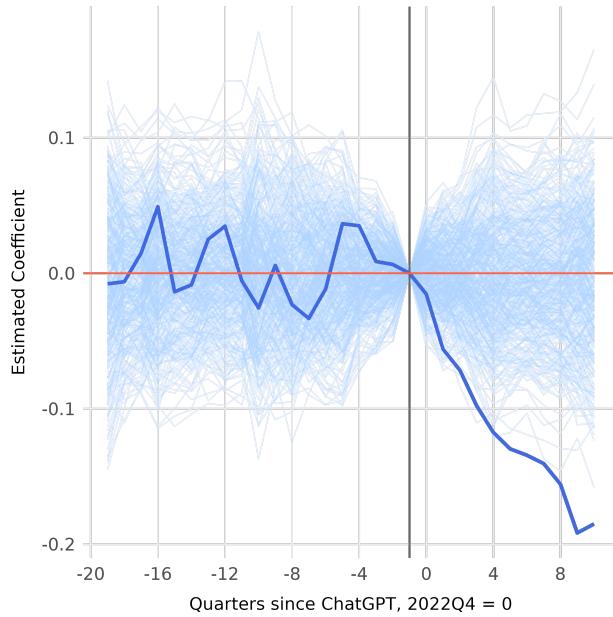


图6：安慰剂测试：事件研究

注意：深蓝色线条表示基线系数，如图 4，并且浅蓝色线条显示了使用随机分配一半所有职业到“伪处理”组进行400次安慰剂检验所得到的系数，这些系数是使用方程中的规格进行估计的 2.

5.1.2 替代人工智能替代措施

我们检验了在使用替代的AI替代指标时，我们的基线结果是否仍然稳健。我们首先使用构成总体AI替代分数的六个单个组成部分，然后采用O*NET提供的另一个相关职业特征：人际交往能力要求。预期人际交往能力要求较高的职业在短期内不太可能被GenAI替代。

如前所述，Pizzinelli等人（2023）识别了六个与人工智能取代的可能性相关的职业维度：沟通、责任、身体条件、关键性、常规性和技能要求。其中五个维度与职业的工作环境相关，最后一个关于技能区（参见 Pizzinelli 等人）。2023 为详细信息）。需要更高水平的沟通、责任、身体在场和专门技能的工作，虽然不那么常规，但被ai取代的风险可能较低。我们使用每个维度的中位数分数将职业分为处理组和控制组。高于中位数的职业被放在处理组，而那些

得分低于中位数的被分配到对照组。

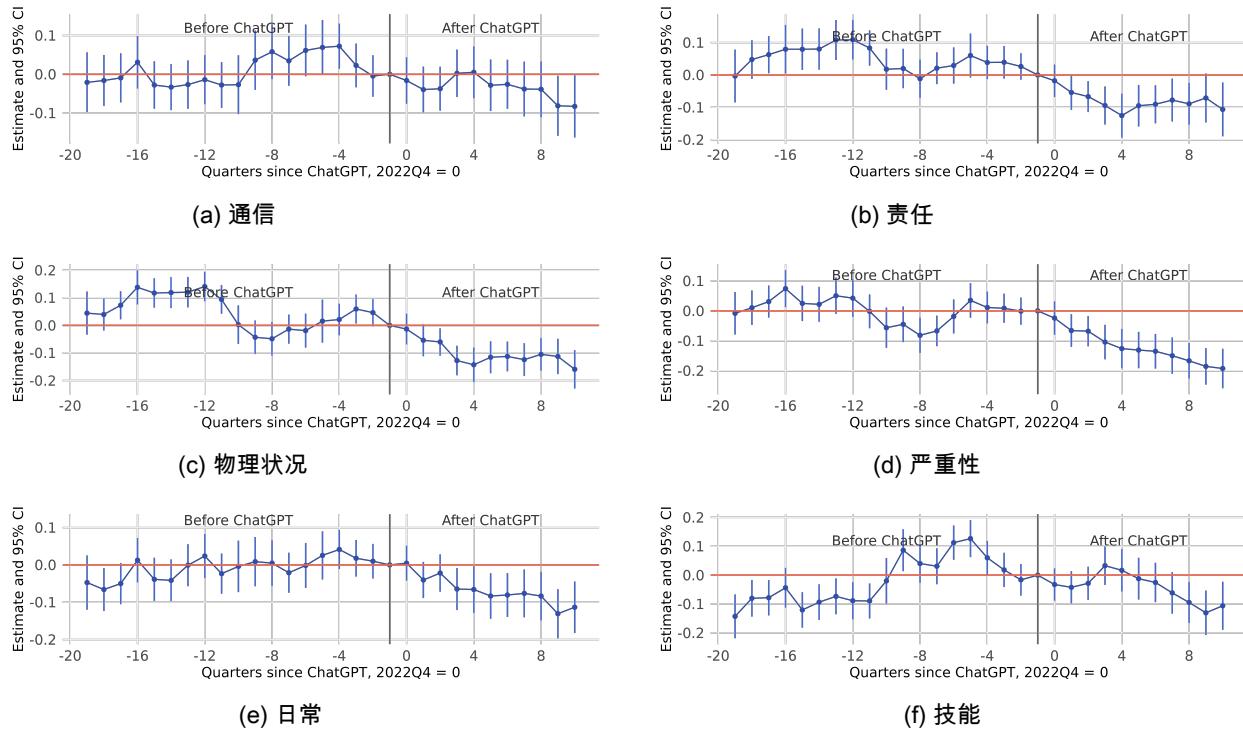


图7：鲁棒性检验：使用AI替代分数的子组件定义处理

注意： 这些图形绘制了基于方程中事件研究规范的点估计和95%置信区间 [2](#) 参考期间为ChatGPT首次推出前的季度 ($Q=-1$)。点代表治疗组和控制组之间职位发布中的季度差距。治疗组定义如下：在沟通、责任、物理条件、关键性或技能要求上低于中位数，或在O*NET上的常规任务强度高于中位数。回归单元是州-行业-职业-季度级别，包含零计数。该规格控制了州-行业-职业固定效应、州-行业-季度固定效应，以及一系列基于Gmerek、Berg和Bescond (的GenAI暴露指数 decile 排名生成的虚拟变量。[2023](#))，与季度虚拟变量交互。

图 [7](#) 展示了跨所有六个维度的结果。ChatGPT的岗位替代效应在AI替代分数的几乎所有维度上都有体现。然而，当考虑以下因素来识别处理组时，即身体状况、关键性、责任和工作常规性，这种效应最为显著。这一发现表明，这四个工作环境特征似乎与确定一个职业对AI替代的脆弱性更为相关。沟通需求在用于分类处理组时产生较弱且大体上不显著的效果。由于许多早期采用者利用ChatGPT来增强沟通(Bick, Blandin, and Deming [2024](#))，这方面可能无法有效识别最初预期的治疗组。此外，ChatGPT有可能被用来弥补技能差距，从而导致

技能评分用于分组时存在延迟效果。

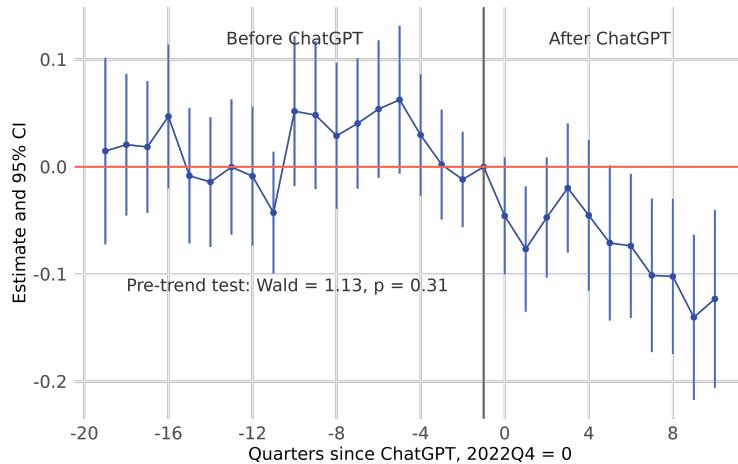


图 8：鲁棒性检验：人际交往能力要求

注意：这些图形绘制了基于方程中事件研究规范的点估计和95%置信区间²。参考期间是ChatGPT首次发布之前的季度($Q=-1$)。点代表治疗组和控制组之间职位发布中的季度差距。治疗组根据O*NET对人际交往能力的要求低于中位数定义。回归单元为州-行业-职业-季度级别，包含零计数。该模型控制了州-行业-职业固定效应、州-行业-季度固定效应，以及一系列基于Gmyrek、Berg和Bescond (的GenAI暴露指数十位数排名的虚拟变量。[2023](#))，与季度虚拟变量交互。

我们进一步使用人际技能要求作为人工智能替代脆弱性的替代衡量标准来测试我们的发现。借鉴gmyrek, b erg, 和 bescond ([2023](#)), 我们采用O*NET的人际交往技能工作环境得分，该得分衡量职业需要面对面互动、建立关系和社会协调能力的程度。预计需要极少人际交往的职业更有可能被人工智能取代。因此，我们将需要人际交往技能得分低于中位数的职业定义为治疗组，其中人类社交互动对绩效不那么重要。图 8 证实了我们的基线发现：需要有限人际交往技能的职业招聘信息在ChatGPT发布后显著下降，且这种效应随着时间的推移持续并加剧。总而言之，这些稳健性检验支持我们的主要结论，同时强调了某些职业特征（例如责任和关键性）在决定职业被AI替代的风险方面起着更重要的作用。

5.2 作为对照的替代AI暴露测量

我们的基线规范使用来自gmyrek、berg和bescond的GenAI暴露指数(2023)来构建 GenAI 暴露十分位-季度固定效应。为确保我们的发现不是这种特定控制选择产生的伪影，我们使用替代的暴露指标和模型设定进行了额外的稳健性检验。具体而言，我们首先使用了两种替代的 AI 暴露分数作为控制：由 Gmyrek 等人提供的 GenAI 暴露分数 (2025) 和 Felten、Raj、Seamans 的 AI 职业暴露指数 (AIOE 2021)。前者是 Gmyrek、Berg 和 Bescond 的更新版本。 2023)，其中工人在自动化中面临的风险被整合到新的暴露指标中。后者是首批人工智能暴露指标之一，并在许多实证研究中得到应用 (Acemoglu 等。 2022 ; Fossen 和 Sorgner 2021 ; 戈德法布, 特斯卡, 和 提奥多里迪斯 2023)。这个 AIOE 指数通过在亚马逊机械土耳其上调查 2000 名工人，将 10 个 AI 应用与 52 个 O*NET 工作能力联系起来，创建了一个应用-能力相关关系矩阵，该矩阵被聚合为职业级别的接触分数。我们还展示了在不控制 GenAI 接触十分位-四分位固定效应的结果，这使我们能够评估我们的条件策略对于识别是否是必要的。

图 9 基于公式(呈现了结果 2 在三种不同的控制规范下，)。这些模式大体上与我们的基线发现一致。在每个情况下，我们观察到在 ChatGPT 发布之后，具有高于中位数的 AI 替代分数的职业的职位发布数量显著下降，并且随着时间的推移，这种效应加剧。当使用 Gmyrek 等人 (2025) 的暴露分数来构建控制虚拟变量，负向效应的幅度略微减小，且在统计上从第二年变得显著。这主要是因为更新的 GenAI 暴露指数将 ChatGPT 发布后工人感受到的自动化潜力纳入了暴露分数。因此，它可能会稀释在条件 Gmyrek 等人 (2025) 的得分。

在使用 AIOE 或没有曝光控制的情况下，职位发布数量的下降似乎早在 ChatGPT 发布日期之前就开始了，并且显示出整体的长期下降趋势。因此，这两个规格未能通过趋势前测试。这可能是因为 AIOE 早于 ChatGPT 的发布，并且不专门针对 GenAI。这些结果证实，针对 GenAI 特定的曝光进行适当的条件化对于因果识别至关重要。否则，

由相应的双重差分法 (DiD) 规范识别的治疗效应，如表中第(2)和第(5)列所示¹，可能会捕捉到治疗组和对照组之间长期分化趋势，有可能高估了GenAI对劳动力需求的影响。相比之下，我们的基本模型规范通过区分GenAI影响和一般自动化效应，比替代模型更合适。

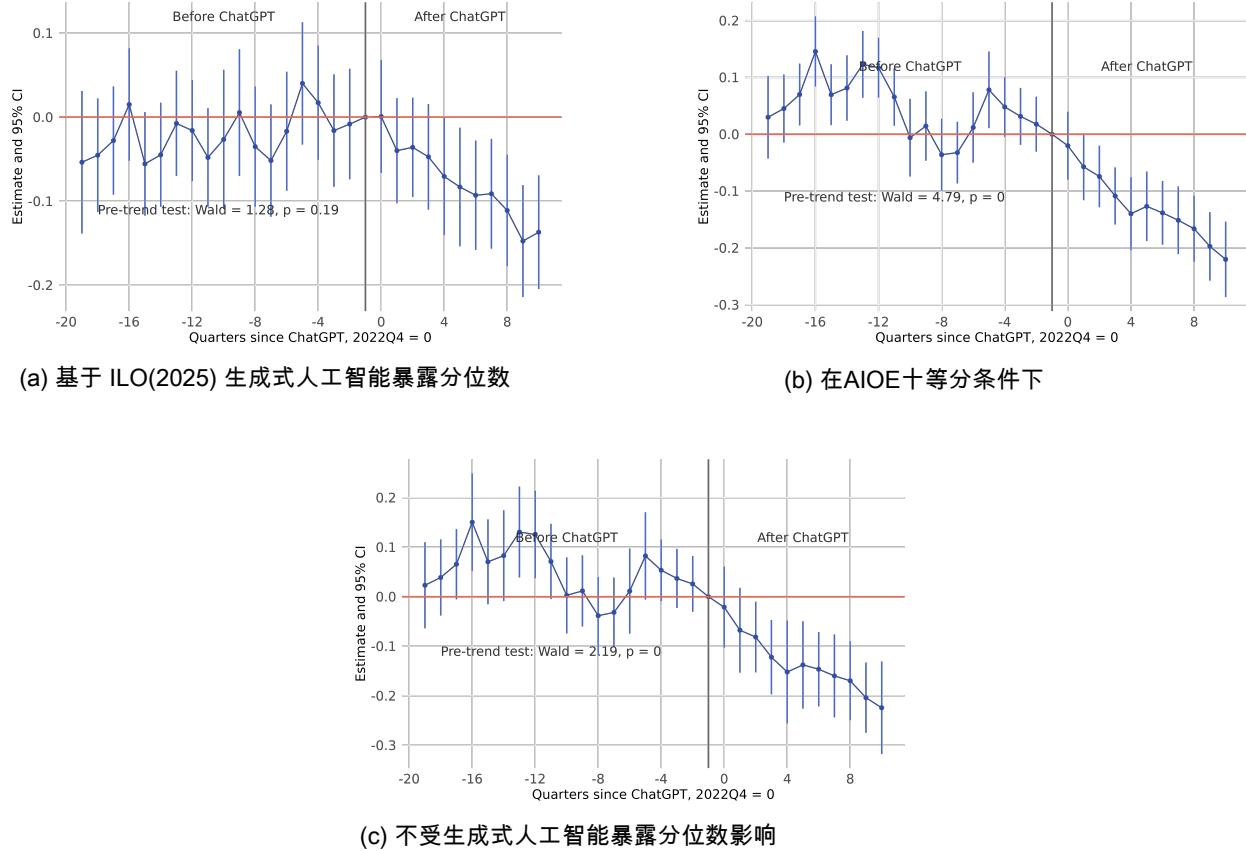


图 9：鲁棒性检验：使用替代的人工智能敞口分数作为对照或不使用对照

注意：这些图形绘制了基于方程中事件研究规范的点估计和95%置信区间²。参考时期是ChatGPT首次推出的季度之前($Q=-1$)。点表示治疗组和控制组之间职位发布中的季度差距。治疗组根据O'NET要求的社交技能中位数以下定义。回归单位是州-行业-职业-季度级别，包含零计数。该规定控制了州-行业-职业固定效应、州-行业-季度固定效应。在(a)和(b)面板中，包括额外的控制变量，即来自Gmyrek等人的一系列虚拟变量，表示GenAI暴露指数的十分位数排名()。²⁰²⁵对于面板(a)，以及来自Felten、Raj和Seamans的AI暴露指数(²⁰²¹)，与季度虚拟变量交互。

6 异质性

由教育和工作经验要求造成的异质性。 本节首先从两个关键维度考察了GenAI对不同职位的影响：1) 教育要求，2) 经验要求。在同一职业中，GenAI可能会以系统的方式对不同教育水平的工人产生影响。

²² 同样地，其影响可能在经验较少和经验较丰富的工人之间有所不同。一方面，生成式人工智能可以让经验较少的工人执行以前需要大量培训或专业知识的任务，有可能减少公司对经验更丰富工人的需求。另一方面，经验丰富的工人可以利用生成式人工智能来自动化通常分配给初级员工的常规任务，有可能取代入门级职位。最近的研究表明，后一种影响占主导地位，入门级和经验较少的工人承担了由生成式人工智能引起的不成比例的替代效应 (Beane 2024 ; 布林约夫森，钱德拉，和陈 2025 ; Lichtinger和Hosseini Maasoum 2025).

²³

利用 Lightcast 提供的详细信息，我们根据职位描述中指定的最低要求对招聘信息进行分类，分为离散类别。教育要求分为高中毕业、大学学位和高等教育（硕士、专业学位或博士学位）。经验要求分为三个等级：入门级职位（0-2年）、中级职位（3-5年）和高级职位（6年及以上）。为估计异质性效应，我们应用以下规格：

$$\begin{aligned} & X \quad X \\ & x \\ & Post\ 1\ yr \\ & arsinh(Y) = \alpha D_{1 \times 1} \end{aligned} \tag{4}$$

^e 下标 $D_{1 \times 1}$ 指示不同的工作特征（即，教育要求和经验水平）。上标 e 用于构造虚拟变量，¹

^{Post\ 2+\ yr}
其值为

² ~~一个观察~~ g_e 属于组 e 当且仅当满足条件时，其值为1，否则为0。这使得我们可以为每个组估计单独的系数 β_{post} 和 β_{job} 提示 GenAI 的取代效应

³ $\epsilon_{e\ 2\ e}$
它们在不同类型的职位上有所差异，分别体现在第一年和长期。

²² 例如，在软件开发中，博士生可能专注于设计新算法，而同一公司的大学毕业生可能被分配任务去实现和部署它们。

²³ 金融时报 (2025). “一个没有初级职位的白领世界？”

图 10 展示了GenAI对按教育和经验要求的工作职位的影响。

²⁴ 结果表明，在GenAI对劳动力市场的影响方面，对于第一年效应存在明显的技能梯度，只有完成高中学位的工作岗位经历了最急剧的下降。与对照组相比，在AI替代脆弱性中位数以上的职位中，这些职位发布量下降了12%。大学学位职位显示出6%的温和下降，而高等学位职位实际上增加了5%，这表明可能存在互补效应或对更熟练的专业人员需求增加。

然而，长期效应揭示了GenAI的取代影响加剧和范围扩大。所有教育类别最终都经历了显著下降，并且幅度转变为与技能要求相关的U形曲线。大学学位职位显示出最严重的长期下降，为18%，其次是高中职位，为15%，而高等教育学位职位的降幅仅为7%。这种模式表明，虽然GenAI最初幸免于高技能职位，但其不断发展的能力最终对中等技能工作造成了最严重的影响。

经验维度表明，初级职业工作者更容易受到生成式人工智能的取代效应影响。这一发现支持了生成式人工智能的知识转移能力特别威胁到传统上通过经验学习提供价值的职位的观点。入门级职位（0-2年）经历了最严重的即时影响，第一年下降了10%。中级职位（3-5年）显示出7%的适度取代，而高级职位（6年以上）则没有出现明显的短期效应。从长远来看，取代效应在所有经验类别中加剧，同时保持了其等级结构。入门级和中级职位的降幅分别为19%和14%，而高级职位则下降了10%。这表明生成式人工智能补充了而不是替代了工作者多年积累的专业知识和判断力。这些发现证实了新兴的证据，表明年轻的初级职业工作者代表着生成式人工智能取代的“矿井中的金丝雀”（布赖恩约夫森、钱达尔和陈）。[2025](#)).

²⁴ 回归结果展示在表中 [A2](#) 和 [A3](#) 为了确保我们的研究结果不受样本选择的影响，我们在前两列包含缺乏明确教育或经验要求的招聘信息，并在后两列排除此类招聘信息。包含无特定要求的招聘信息所得到的结果与前两列（3）和（4）的结果非常相似，这证实了我们的研究结果对样本构成是稳健的。

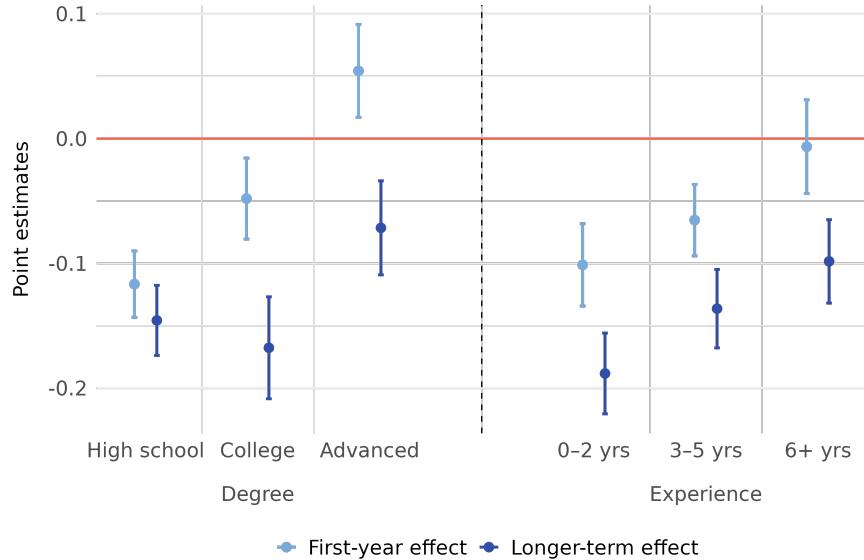


图10：按教育程度和按经验异质性

注意：这个图显示了基于公式回归的估计值 4，分别按教育程度和工作经验展示。高级学位包括硕士、专业学位和博士学位。“yrs”是“年”的缩写。不包括未规定教育程度或工作经验要求的职位发布。点代表估计效应，而尖峰表示相应的95%置信区间。浅蓝色点和尖峰显示ChatGPT发布后第一年的效应，而深蓝色点和尖峰显示截至2025年6月的长期效应。完整回归结果展示在表格的第一列。A2 和 A3，其他列提供了额外的规格作为鲁棒性检验。

行业异质性。接下来，我们分析 GenAI 在 20 个两位数的 NAICS 行业中的劳动力替代效应，以捕捉其部门差异。具体而言，我们将行业虚拟变量与前面的一组 DiD 交互项进行交互，如下所示：

$$X \quad X \\ x \\ \begin{matrix} Post & 1 \\ yr & \end{matrix} \\ arsinh(Y) = \alpha D_1 \times 1 \\ \text{哪里是上标 } i \text{ 用于虚拟变量以指示一个行业，以便 } 1 \\ \text{如果观测值为 } 1 \text{ 属于行业 } i \text{ 并且其余为 } 0. \\ \begin{matrix} 1 & \times & 1 \end{matrix} \quad \quad (5)$$

图 11+展示了 GenAI 对各行各业产生的短期和长期影响。总的来说，六个行业在第一年内经历了显著的负面影响，如行政和支持服务、专业服务、公司管理、房地产、其他服务以及医疗保健。十一个行业没有出现显著的短期影响，涵盖从教育、建筑到信息技术、酒店和制造业的各个领域。零售贸易、金融和运输三个行业最初显示出积极影响。

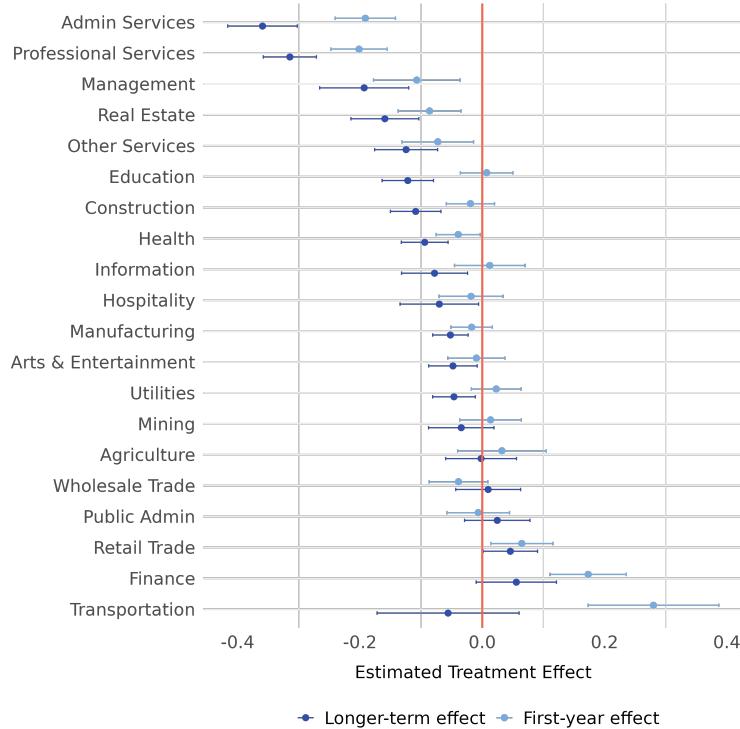


图 11：按行业异质性

注意：这个图显示了基于公式回归的估计值⁵，按行业分别呈现。“yrs”是“年”的缩写。点代表估计效应，而竖线表示相应的95%置信区间。浅蓝色点和竖线显示ChatGPT发布后第一年的效应，而深蓝色点和竖线显示直至2025年6月的长期效应。

动态分析揭示了大多数行业向负面效应的显著趋同。在二十个行业中，最终有十三个行业的工作岗位数量经历了显著下降，这种效应在先前受影响的部门加剧，并在先前具有抵抗性的部门中出现。行政服务、专业服务、公司管理、房地产、其他服务和教育显示出尤为突出的长期效应。这些行业也往往具有很高的通用人工智能（GenAI）采用率和很高的职业暴露率，正如现有研究所示。²⁵

最后，七个行业在整个期间保持了微不足道的影响。它们是交通运输、金融、公共管理、批发贸易、农业、采矿和公用事业。这些部门可能具有限制GenAI适用性的特征，包括体力任务要求、监管约束或专业领域知识要求。结果

²⁵

参见，例如，Felten, Raj, 和 Seamans (2021), gmyrek, berg, and bescond (2023),以及比克、布兰丁和戴明(2024)

支持 Korinek 所提出的论点 2024) 在 GenAI 存在的情况下 , 技术和社交障碍 , 加上劳动的根本以人为本的方面 , 可以在短期内至中期内维持对经济中人力投入的一定需求。

7 结论

本文首次全面分析了生成式人工智能对美国劳动力需求的影响 , 利用实时在线招聘数据。以 2022 年 11 月 ChatGPT 的公开展示作为外生技术冲击 , 我们应用双重差分法和事件研究法来估计生成式人工智能对不同职业招聘广告的影响。预计在相同生成式人工智能暴露水平下 , 人工智能替代脆弱性较高的职业比人工智能替代脆弱性较低的职业更可能受到 ChatGPT 发布带来的负面影响。因此 , 我们比较了在生成式人工智能暴露水平相似的情况下 , 人工智能替代脆弱性高于中位数和低于中位数的职业的劳动力需求。这种模型设置使我们能够在控制生成式人工智能随时间推移的采用效果和其他混杂因素的同时 , 隔离生成式人工智能的替代劳动效应。我们的动态分析追踪了随着生成式人工智能能力的提升和采用深化 , 这些影响是如何随时间演变的。

在控制了 GenAI 暴露水平的情况下 , 我们记录了 GenAI 对具有较高 AI 替代脆弱性的职业岗位发布造成 的巨大、统计上显著且加剧的负面影响。在 ChatGPT 发布后 , 我们发现从 2022 年底到 2025 年 6 月 , 对于 替代得分高于中位数的职业 , 岗位发布平均下降了 12% 。这种替代效应从第一年的约 6% 增长到第三年的 18% , 表明随着技术的成熟和扩散 , GenAI 对劳动力市场的影响会累积。

异质性分析揭示了 GenAI 在工人特征和行业间的系统性效应差异。既不需要高等学位也不需要丰富工作经 验的入门级职位面临着不成比例的失业 , 特别是在专业服务和行政支持行业 , 减少尤为严重。这种模式表明 , GenAI 对专业知识的民主化最直接地威胁了早期职业技能发展的传统途径。这些发现表明 , GenAI 工具 的进步可能有助于解释美国劳动力市场一些观察到的趋势 , 例如白领招聘的放缓。

工人人群和GenAI采用率高的行业。此外，我们的结果也对应了近届大学毕业生就业前景的恶化。

²⁶

早期证据表明，生成式人工智能对劳动力需求的影响与过去技术的影响不同。过去技术的颠覆往往在几十年间逐渐显现，为工人、企业和政策制定者提供了适应的时间。相比之下，生成式人工智能在其应用仅三年内就导致对脆弱职业的需求下降了近20%。这种下降对没有高等学历的求职者和初入职场的工作者更为明显，在行政支持（40%）和专业服务（30%）领域最为严重。此外，与早期主要取代常规体力劳动岗位的技术不同，生成式人工智能似乎对传统白领职业的需求产生了更大的影响。

我们的研究结果具有重要的政策启示。首先，教育和培训体系需要适应，为学生们做好准备以迎接人工智能时代。这可能包括更加重视培养创造性解决问题的能力、情商、人际交往能力，以及数字或人工智能素养。其次，入门级职位中传统“边做边学”机会的减少凸显了支持早期职业发展的必要性。此类支持项目例如学徒计划、实习计划，以及创业或技能培训计划。第三，政策制定者也应考虑扩大社会保障体系，以支持工人在这一转型期间。最后，如果这种工作替代效应变得更加永久（Korinek 2024）。

GenAI的快速采用及其劳动力市场的启示为未来的研究开辟了广阔的途径（Brynjolfsson, Korinek, and Agrawal 2025）。跨国比较研究—尤其是发展中国家之间的比较研究—目前仍然缺乏。它们可以揭示哪些国家特有的因素会影响GenAI对劳动力市场的影响。除了劳动力需求之外，未来的研究还可以评估对工资、收入分配和技能需求的影响。更重要的是，应该投入更多努力来研究GenAI如何创造全新的工作岗位，并以能够增加劳动力需求的方式提高生产力。

²⁶

来自当前人口调查的数据揭示了23-27岁年轻工人群体令人担忧的模式。该群体中的大学毕业生在2025年的失业率平均为4.59%，自2019年以来上升了1.34个百分点。这一增幅远超其他人口群体：同一年龄段的非大学毕业生仅上涨了0.47个百分点，而年长的大学毕业生上涨了0.38个百分点。这些差异表明，GenAI对劳动力市场的影响不仅仅是周期性的，而是反映了经济吸收新教育人才方式的结构性变化（联邦储备委员会2025年（[链接](#)））

参考文献

阿西莫格鲁，达龙和戴维·奥托尔（2011年）。“劳动经济学讲座”。文稿。<http://economics.mit.edu/files/4689>，第22页。

阿西莫格鲁，达龙，大卫·奥托，乔纳森·哈泽尔，和帕斯夸尔·雷斯特雷波（2022年）。“人工智能与就业：来自在线职位的证据”。*劳动经济学杂志* 40.S1, S293–S340。

阿西莫格鲁，达龙和帕斯库亚尔·雷斯特雷波（2019年）。“自动化和新的任务：技术如何取代和重新安置劳动力”。*经济视角杂志* 33.2, 第3-30页。

— (2020). “机器人与工作：来自美国劳动力市场的证据”。*政治经济学杂志* 128.6, 第2188-2244页。

作者，大卫·H·阿特奥 (2015年)。\"为什么还有这么多工作？工作自动化的历史和未来\"。*经济视角杂志* 29.3, 第3-30页。

巴比纳，塔尼娅，阿纳斯塔西娅·费德克，亚历克斯·何，詹姆斯·霍德森 (2023)。 公司在人工智能技术上的投资以及劳动力构成的变化 . 第31325卷。美国国家经济研究局。

贝克，安德鲁，布兰蒂·卡劳，斯科特·坎宁安，安德鲁·古德曼-贝肯，和佩德罗·HC桑塔纳 (2025)。“双重差分设计：从业者指南”。 arXiv预印本 [arXiv:2503.13323](https://arxiv.org/abs/2503.13323) .

比恩，马特 (2024)。 技能代码：如何在智能机器时代拯救人类能力 . 哈珀柯林斯。

贝内特兹-鲁埃达，米格尔和埃里克·帕拉多 (2024)。 镜子，墙上的镜子：哪些工作最终会被人工智能取代？：职业暴露新指数 .Tech.report.IDB.

贝克，亚历山大，亚当·布兰丁，和大卫·J·德明 (2024)。 生成式人工智能的快速采用 科技报告. 美国国家经济研究局.

邦菲格利奥利，亚历山德拉，罗萨里奥·克雷诺，吉诺·甘恰，以及伊奥annis·帕帕达基斯（2025）。“人工智能与就业：来自美国通勤区的证据”。《经济政策》40.121, 第145-194页。

邦尼，凯瑟琳，科尔·布雷奥，凯茜·巴菲特，埃明·丁勒索兹，卢西娅·S·福斯特，内森·戈尔施拉格，约翰·C·哈蒂旺纳，扎卡里·克罗夫，基思·萨维奇（2024）。《实时追踪企业AI使用情况：商业趋势与展望调查的快照》。科技报告。美国国家经济研究局。

布拉克顿、J·卡特和布萊迪·塔斯卡（2023年）。“技术变革与失业的后果”。《美国经济评论》113.2, 第279—316页。

布赖恩约夫森、埃里克·巴拉特·钱达尔，和瑞雨·陈（2025）。《煤炭矿中的金丝雀？六个关于人工智能近期就业效应的事实》。Tech. 报告

布赖恩约夫森、埃里克·安东·科里内克，和阿贾伊·K·阿格拉瓦尔（2025年）。“人工智能经济学的变革性研究议程”。

布赖恩约夫森、埃里克·莉莉安·李，和林德赛·雷蒙德（2025年）。“工作中的生成式人工智能”。《经济学季刊》140.2, pp. 889–942。

布赖恩约夫森、埃里克·丹尼尔·罗克，和查德·西弗森（2017）。《人工智能与现代生产力悖论：预期与统计的冲突》。科技报告。美国国家经济研究局。

钱达尔，巴巴特（2025）。“追踪暴露于人工智能工作中的就业变化”。可在SSRN上找到 5384519。

陈嘉峰和乔纳森·罗特（2024年）。“含有零的日志？一些问题和解决方案”。《经济学季刊》139.2, pp. 891–936。

崔泽远凯文、Mert Demirer、Sonia Jaffe、Leon Musolff、彭思达、Tobias Salz（2024年）。“生成式人工智能对高技能工作的影响：来自三个软件开发者现场实验的证据”。可在ssrn上找到，编号4945566。

德·阿夸, 法布里齐奥, 爱德华·麦克弗洛兰三世, 埃than·R·莫利克, 赫拉·利夫希茨-阿萨夫, 凯瑟琳·凯洛格, 萨兰·拉贾恩德拉, 丽莎·克拉耶, 弗朗索瓦·坎德隆, 和 卡里姆·R·拉克哈尼 (2023). “ navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality ”。 哈佛商学院技术与运营管理单位工作论文 24-013.

戴明 , 大卫·J和卡迪姆·诺雷 (2020年)。“收益动态、变化的职业技能和STEM职业”。 《经济学季刊》 13 5.4 , 第1965-2005页。

铁米尔奇, 奥兹杰, 约纳斯·哈纳内, 和 朱欣荣 (2025). “谁被AI取代了？生成式AI对在线自由职业平台的影响” 。 管理科学 。

爱斯菲尔德特 , 安德烈亚·L , 格雷戈尔·舒伯特 , 张淼 , 以及布雷迪·塔斯卡 (2023)。“生成式人工智能对企业价值的影响”。 可在SSRN上找到 4436627 。

伊隆杜、蒂娜、山姆·曼宁、帕梅拉·米施金和丹尼尔·罗克 (2024年)。“GPTs就是GPTs : 大语言模型对劳动力市场的影响潜力”。 科学 384.6702 , 第1306-1308页。

Felten, Edward, Manav Raj, 和 Robert Seamans (2021)。“职业、行业和地理领域的人工智能暴露 : 一个新型数据集及其潜在应用”。 战略管理杂志 42.12 , 第2195-2217页。

Felten, Edward W, Manav Raj, 和 Robert Seamans (2018)。“一种将人工智能的进步与职业能力联系起来的方法”。 AEA 论文与会议录 108, pp. 54–57.

— (2019). “人工智能对职业的影响 : 劳动力、技能和分化”。 纽约大学斯特恩商学院 。

— (2023)。“生成式人工智能暴露的职业异质性”。 SSRN电子期刊 福赛斯 , 伊丽莎白 , 丽莎·B·卡恩 , 法比安·兰格 , 和大卫·维泽尔 (2022年)。“工人们去哪儿了？新冠疫情恢复中的召回、退休和重新分配”。 劳动经济学 78 , 第102251页。

Fossen, Frank和Alina Sorgner (2021)。“工作的数字化与创业的进入”。 商业研究杂志 125 (C). 出版商 : elsevier , 第548-563页。

弗雷，卡尔·本尼迪克特和迈克尔·A·奥斯本（2017年1月1日）。“就业的未来：工作容易受到计算机化的影响吗？”*技术预测与社会变革* 114, pp. 254–280.

格默雷克，帕维尔，珍妮娜·伯格，和大卫·贝斯孔（2023年）。“生成式人工智能和工作：对工作数量和质量潜在影响的全局分析”。*国际劳工组织工作文件* 96.

格米雷克，帕韦尔，扬娜·伯格，卡罗尔·卡明斯基，菲利普·科诺普琴斯基，阿格涅什卡·拉达娜，巴林特·纳夫拉迪，康拉德·罗萨尼切克，马尔克·特罗什尼斯基（2025）。*生成式人工智能与工作：一个职业暴露的精细化全球指数* . 140. 国际劳工组织工作文件。

Goldfarb , Avi , Bledi Taska和Florenta Teodoridis (2023年) 。“机器学习可以是通用目的技术吗？使用来自在线招聘信息的兴起技术比较”。*研究政策* 52.1 , p. 104653。

哈莫尔，门卡，迪米特里斯·帕帕尼科洛乌，劳伦斯·D·施密特，以及布莱恩·西格米尔（2025）。*人工智能与劳动力市场 科技报告*. 美国国家经济研究局.

Handa , Kunal , Alex Tamkin , Miles McCain , Saffron Huang , Esin Durmus , Sarah Heck , Jared Muelle r , Jerry Hong , Stuart Ritchie , Tim Belonax 等 (2025) 。 “哪些经济任务由 AI 执行？来自数百万次 Claude 对话的证据”。*arXiv预印本 arXiv:2503.04761* .

赫什贝因、布拉德和丽莎·B·卡恩（2018年）。“经济衰退会加速常规偏向的技术变革吗？来自职位空缺信息的证据”。*美国经济评论* 108.7, pp. 1737–1772.

赫什贝因、布拉德和布萊恩·A·斯图尔特（2024年）。“经济衰退后地方劳动力市场的演变”。*美国经济期刊：应用经济学* 16.3 , 第399-435页。

会, 翔, 奥伦·雷谢夫, 和 卓枫 (2024). “生成式人工智能对就业的短期影响：来自一个在线劳动力市场的证据”。*组织科学* 35.6 , 第1977-1989页。

Humlum, anders和Emilie Vestergaard (2025)。 大型语言模型，小劳动力市场效应 科技报告. 美国国家经济研究局.

江, Wei, Junyoung Park, Rachel Jiqiu Xiao, 和 Shen Zhang (2025). 人工智能与工作时长延长：生产力、合同效率与租值分配 科技报告. 美国国家经济研究局.

科里内克 , 安东 (2024)。 人工智能时代的经济政策挑战 科技报告. 美国国家经济研究局.

lichtinger, guy和seyyed mahdi hosseini maasoum (2025)."生成式人工智能作为偏向高龄的技术变革：来自美国简历和招聘数据的证据"。 在SSRN上可用 .

刘, 喻, 景云黄, 和王 (2025). 地球上有谁在使用生成式人工智能？追踪2025年全球趋势和转变 .Tech. 报告

刘岩和何王 (2024年8月)。 地球上到底有谁在使用生成式人工智能？ 政策研究工作论文10870。华盛顿特区：世界银行，第1-56页。

莫雷蒂 , 恩里科 (2010)。本地劳动力市场。技术报告。国家经济研究局。

。”和萨亚什·卡普尔 (2025年)。“人工智能作为常规技术”纳拉扬南、阿维 骑士第一修正案学院 .

诺伊 , 沙克德和张薇薇 (2023)。“生成式人工智能对生产率影响的实验证据”。 科学 381.6654 , 第187-192页。

皮齐内利 , 卡洛 , 奥古斯都·J·潘顿 , 玛琳娜·门德斯·塔瓦雷斯 , 马乌罗·卡扎尼亚加和栗俊 (2023)。 人工智能对劳动力市场的暴露：跨国差异与分配影响 .国际货币基金组织。

乔丹丹、华夏瑞、钱雄 (2023)。“AI与就业：拐点已至？来自一个在线劳动力平台的证据”。 arXiv预印本 arXiv:2312.04180 .

Teutloff, Ole, 约翰娜·艾因赛德勒, 奥托·卡西, 法比安·布雷塞曼, 帕梅拉·米施金, 和 R·玛丽亚·德尔·里奥·查诺纳 (2025). “生成式人工智能的赢家与输家：自由职业者需求变化的早期证据”。 经济学行为与组织 journal , 第 106845 页。

汤林森 , 基兰 , 索尼·贾菲 , 威尔·王 , 斯科特·库恩斯 , 和西迪哈斯·苏里 (2025年)。“使用人工智能：衡量生成式人工智能的职业影响”。 arXiv预印本 arXiv:2507.07935 .

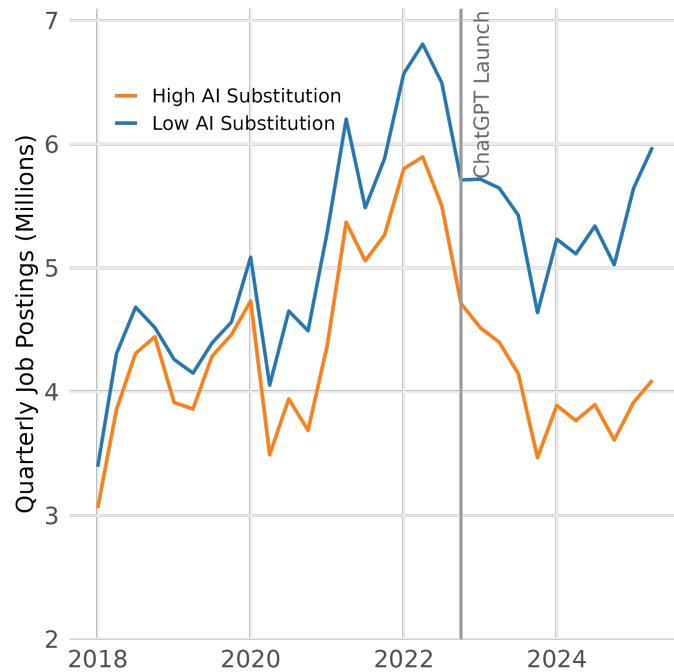
托佩尔，罗伯特·H（1986年）。“地方劳动力市场”。*政治经济学杂志* 94.3，第二部分，S111–S143。

韦伯，迈克尔（2019）。“人工智能对劳动力市场的影响”。可在SSRN上找到3482150。

世界银行（2025a）*未来工作：东亚和大洋洲的机器人、人工智能和数字平台*。完整报告。华盛顿特区：世界银行。

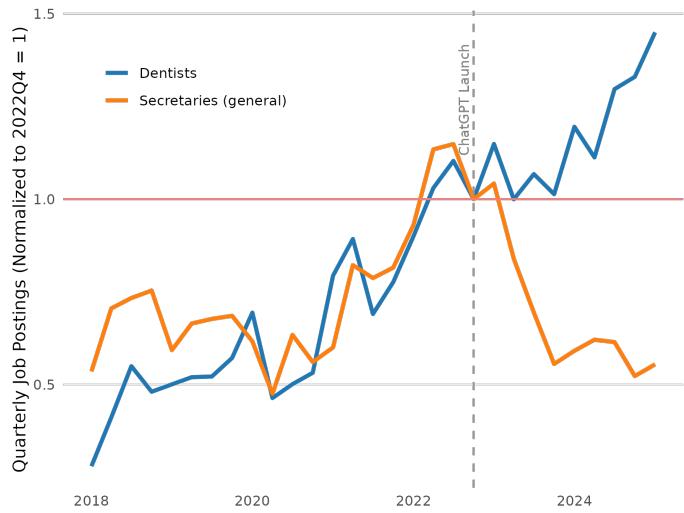
—(2025b). 南亚发展动态，2025年10月。华盛顿特区：世界银行。

附录：补充图表



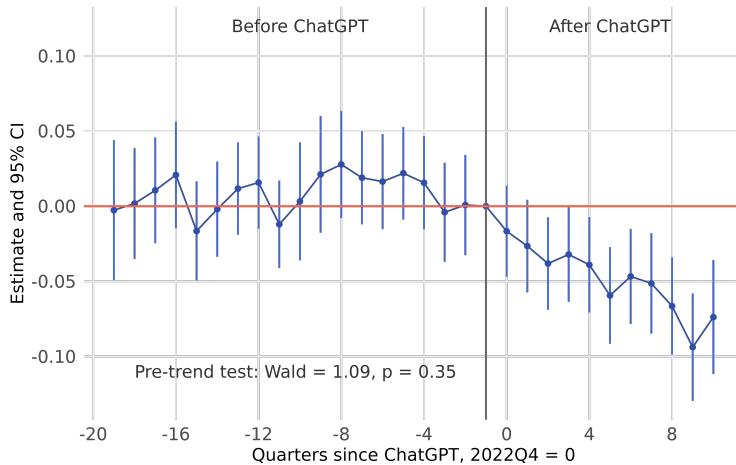
图A1：具有高与低AI替代评分的职业招聘趋势

注意：该图显示了LightCast中高于和低于中位数AI替代分数的职位的在线招聘趋势。AI替代分数被定义为Pizzinelli等人的AI互补性指标的倒数（）。[2023](#)，沿六个职业维度构建：沟通、责任、物理条件、关键性、常规和技能要求。



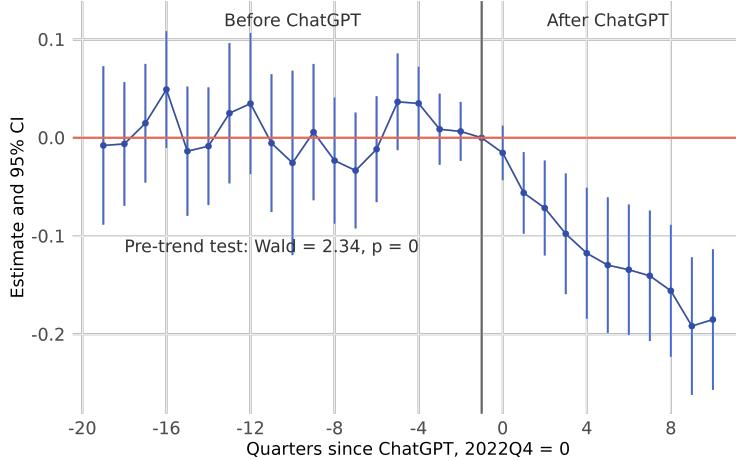
图A2：两个典型职业的职位发布趋势：牙医 对比 秘书（综合）

注意：这张图展示了LightCast中两个职业（牙医和秘书（综合））的在线招聘趋势，分别位于中位数AI替代分数的下方和上方。数值在ChatGPT发布季度（2022Q4）归一化到1。



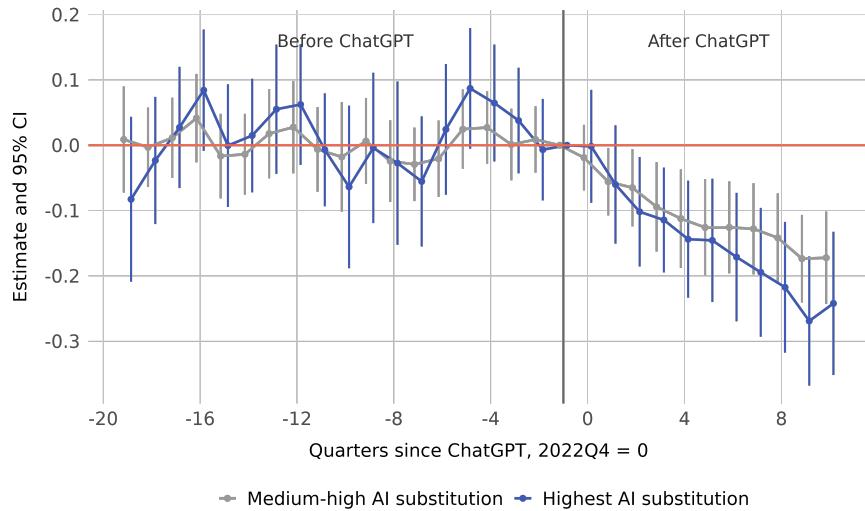
图A3：鲁棒性：非加权回归

注意：该图描绘了基于方程中事件研究规范的点估计和95%置信区间²，不使用2017年的州-行业-职业岗位公告作为样本权重。参考期是Chat GPT首次发布前的季度($Q=-1$)。点代表治疗组和控制组之间岗位公告的季度差距。治疗组定义为AI替代得分高于中位数的职业（或Pizzinelli等人()中AI互补得分的倒数）。²⁰²³）。回归单元处于州-行业-职业-季度级别，包含零计数。该设定控制了州-行业-职业固定效应、州-行业-季度固定效应，以及一系列基于Gmyrek、Berg和Bescond的GenAI暴露指数十位数排名的虚拟变量。²⁰²³），与季度虚拟变量交互。



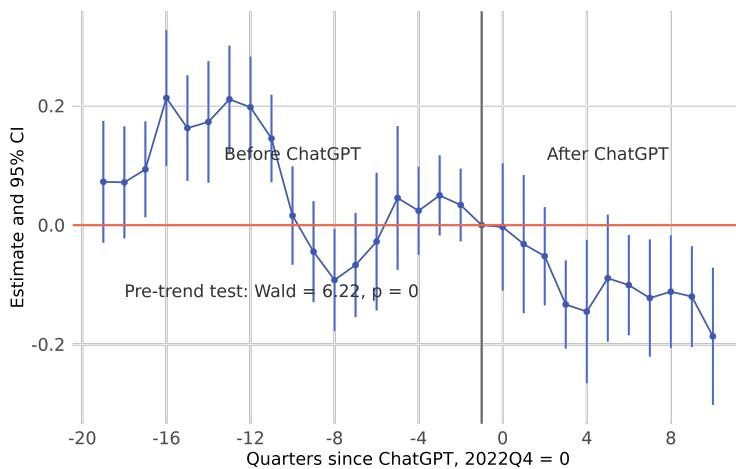
图A4：鲁棒性：职业层面的SE聚类

注意：该图描绘了基于方程中事件研究规范的点估计和95%置信区间²，而标准误在职业层面聚类。参考期为ChatGPT首次发布前的季度（Q=-1）。点代表治疗组和控制组之间职位发布的季度差距。治疗组定义为AI替代分数高于中位数的职业（或Pizzinelli等人()中AI互补分数的倒数）。^{20 23}）。回归单元处于州-行业-职业-季度级别，包含零计数。该设定控制了州-行业-职业固定效应、州-行业-季度固定效应，以及一系列基于Gmyrek、Berg和Bescond的GenAI暴露指数组十位数排名的虚拟变量。²⁰²³），与季度虚拟变量交互。



图A5：稳健性检验：其他定义处理组的截止点

注意：该图描绘了基于方程中事件研究规范的点估计和95%置信区间²，同时我们实施了包含两个处理组的事件研究。这两个处理组是根据AI替代得分定义的：中高分替代组包括50至90百分位数范围，最高分替代组涵盖90至100百分位数范围。参考期间是ChatGPT首次发布前的季度（Q=-1）。点代表处理组与对照组之间职位发布中的季度差距。处理组定义为AI替代得分高于中位数（或Pizzinelli等人（²⁰²³））。回归单元处于州-行业-职业-季度级别，包含零计数。该设定控制了州-行业-职业固定效应、州-行业-季度固定效应，以及一系列基于Gmyrek、Berg和Bescond的GenAI暴露指数十位数排名的虚拟变量。²⁰²³，与季度虚拟变量交互。



图A6：鲁棒性检验：使用高GBB GenAI指数定义处理组

注意： 该图描绘了基于方程中事件研究规范的点估计和95%置信区间 [2](#) 参考期间是ChatGPT首次推出的季度之前 ($Q=-1$)。点代表治疗组和控制组之间职位发布中的季度差距。治疗组定义为来自Gmyrek、Berg和Bescond的GenAI暴露指数高于中位数的职业 ([2023](#)) (GBB)。回归单元处于省-行业-职业-季度层面，包含零计数。该设定控制了省-行业-职业固定效应和省-行业-季度固定效应。

操作维护专员

关联交易机器

和渔业人

和销售员

按职业 (ISCO-98 一位数)

文员类职业员 指定员 Notes-

职业的

离子

表A1：按职业、行业、教育水平和经验要求划分的招聘职位（百万），2018年第一季度至2025年第二季度

表A2 : ChatGPT对劳动力需求的影响 , 按教育要求分类

因变量 arsinh(# 职位发布)	(1)	(2)	(3)	(4)
已处理 \times 短 (高中) -0.117*** -0.118***				(0.014) (0.014)
已处理 \times 短 \times 高中 -0.117*** -0.118***		(0.014)	(0.014)	
已处理 \times 短 \times 大学学历 -0.048** -0.050** 0.069*** 0.069***	(0.017)	(0.017)	(0.017)	(0.017)
已处理 \times 短 \times 高学历 0.054** 0.053** 0.171*** 0.171***	(0.019)	(0.019)	(0.022)	(0.022)
已处理 \times 短 \times 教育要求未提及 -0.056*** 0.063***		(0.014)		(0.015)
已处理 \times 长 (高中基线) -0.146*** -0.155***			(0.014)	(0.014)
已处理 \times 长 \times 高中 -0.146*** -0.155***	(0.014)	(0.014)		
已处理 \times 长 \times 大学学历 -0.168*** -0.177*** -0.022 -0.022	(0.021)	(0.021)	(0.021)	(0.021)
已处理 \times 长 \times 高学历 -0.072*** -0.081*** 0.074*** 0.074***	(0.019)	(0.019)	(0.02)	(0.02)
已处理 \times 长 \times 学历要求未提及 -0.128*** 0.027*		(0.014)		(0.013)
教育 \times 状态 \times 行业 \times 职业 FE 是 是 是				
教育 \times 状态 \times 行业 \times 年季度 FE 是 是 是				
职业密度FE \times 年季度 FE 是 是 是				
观察值数量 20,261,790 27,015,720 20,261,790 27,015,720				
R2 0.943 0.956 0.943 0.956				

注意：以职业-季度层面聚类稳健标准误括号内显示。

* $p < 0 . 1$,

** $p < 0 . 05$,

*** $p < 0 . 01$.

回归单元处于教育-状态-行业-职业-区域层面，包含零计数。因变量是职位发布的反双曲正弦 (arsinh) 变换，该变换近似于对数变换，并且在职位发布为零时等于零，允许系数被解释为职位发布的近似对数变化。处理变量是一个虚拟变量，表示处理组，定义为具有中位数以上AI替代分数（或皮齐内利等人的AI互补分数（ 2023 '））。‘短’是虚拟变量，在ChatGPT发布后第一年（包括2022Q4）等于1，而‘长’是发布后的第二年和第三年（2023Q4-2025Q2）。‘职业十分位数固定效应’是基于Gmyrek、Berg和Bescond的GenAI暴露指数十分位数排名的一系列虚拟变量（ 2023 ）。高等学历包括硕士、专业和博士学历。

表A3 : ChatGPT对劳动力需求的影响，按经验要求分类

因变量 arsinh(# 职位发布)			
	(1)	(2)	(3)
已处理 \times 短 (0-2 年经验作为参考群体) -0.101*** -0.098***			
		(0.017)	(0.017)
已处理 \times 短 \times 0-2年经验 -0.101***-0.098***			
	(0.017)	(0.017)	
已处理 \times 短 \times 3-5年经验 -0.065*** -0.062*** 0.036** 0.036**			
	(0.015)	(0.015)	(0.013)
已处理 \times 短 \times 6年+yrs经验 -0.006-0.0030.095***0.095***			
	(0.019)	(0.019)	(0.021)
已处理 \times 短 \times 经验要求未提及 -0.077*** 0.020			
	(0.015)		(0.012)
已处理 \times 长 (0-2年经验作为参考群体) -0.188*** -0.188***			
	(0.017)	(0.016)	
已处理 \times 长 \times 0-2年经验-0.188*** -0.188***			
	(0.017)	(0.016)	
已处理 \times 长 \times 3-5年经验 -0.136*** -0.136*** 0.052*** 0.052***			
	(0.016)	(0.016)	(0.012)
已处理 \times 长 \times 6年+yrs经验 -0.098*** -0.098*** 0.090*** 0.090***			
	(0.017)	(0.017)	(0.018)
已处理 \times 长 \times 要求经验未提及 -0.139*** 0.048***			
	(0.014)		(0.011)
经验 \times 状态 \times 行业 \times 职业 FE 是 是 是			
经验 \times 状态 \times 行业 \times 年季度 FE 是 是 是			
职业密度 FE \times 年季度 FE 是 是 是			
观察值数量 20,261,790 27,015,720 20,261,790 27,015,720			
R2 0.952 0.96 0.952 0.96			

注意：以职业-季度层面聚类稳健标准误括号内显示。

* $p < 0 . 1$,

**

$p < 0 . 05$,

$p < 0 . 01$. 回归单元处于教育-状态-行业-职业-区域层面，包含零计数。因变量是职位发布的反双曲正弦 (arsinh) 变换，该变换近似于对数变换，并且在职位发布为零时等于零，允许系数被解释为职位发布的近似对数变化。处理变量是一个虚拟变量，表示处理组，定义为具有中位数以上AI替代分数（或皮齐内利等人的AI互补分数（ 2023 ”）。“短”是虚拟变量，在ChatGPT发布后第一年（包括2022Q4）等于1，而“长”是发布后的第二年和第三年（2023Q4-2025Q2）。“职业十分位数固定效应”是基于Gmyrek、Berg和Bescond的GenAI暴露指数组十分位数排名的一系列虚拟变量（ 2023 ）。