

人工智能技术进步有助于缓解劳动力市场 教育错配

——来自“岗位-求职者”匹配数据的证据

于航 李泓宇 张丹丹 张润博 李强*

摘要: 本文基于线上招聘平台 2021—2025 年“岗位-求职者”匹配数据,测度了中国劳动力市场的纵向(学历)与横向(专业)错配,并以 2022 年 12 月 ChatGPT 发布为外生冲击,采用双重差分与事件研究法识别人工智能大语言模型(AI-LLM)技术进步的影响。研究发现,2021 年以来错配整体加剧,但高 AI-LLM 暴露职业的纵向错配显著缓解,横向错配则未受显著影响。机制上,AI-LLM 强化了岗位需求信号的清晰度,提升了技能复杂性与学历门槛,从而改善了劳动力配置效率。

关键词: 技术进步;劳动力市场错配;人工智能大语言模型技术暴露指数

DOI: 10.13821/j.cnki.ceq.2026.03.11

一、引言

近年来,人工智能(Artificial Intelligence, AI)技术的快速发展与广泛普及,尤其是大语言模型(Large Language Model, LLM)的持续迭代与规模化应用,正在突破传统生产的边界。与此同时,人工智能驱动的技术变革也对劳动力市场产生了深刻而复杂的结构性冲击:一方面,岗位任务内容与技能需求加速重构,用人单位对劳动者能力的要求不断提高并趋于细化;另一方面,这些变化通过需求侧信号传导至供给侧,进而影响个体的教育投资与职业选择决策,从而重塑劳动力市场的匹配过程与配置效率。在这一背景下,人工智能与就业之间的关系已从单纯“替代”和“创造”的数量问题,转向更为复杂的匹配问题

* 于航,北京大学中国经济研究中心、北京大学国家发展研究院、南南合作与发展学院;李泓宇,北京大学国家发展研究院;张丹丹,北京大学中国经济研究中心、北京大学国家发展研究院;张润博,北京大学国家发展研究院;李强,北京网聘信息技术有限公司。通信作者及地址:张丹丹,北京市海淀区北京大学国家发展研究院,100871;电话:010-62759779;E-mail:ddzhang@nsd.pku.edu.cn。本文得到国家自然科学基金青年项目(72303009)的支持。

(Autor et al., 2003; Acemoglu and Autor, 2011; Goos and Manning, 2007)。由此,一个尚未得到充分回答的问题是:人工智能技术进步究竟如何影响劳动力市场中的教育错配,而不仅仅是影响就业数量或岗位结构本身?在经验识别上,本文以2022年12月ChatGPT发布作为新一轮AI-LLM技术冲击的起点,评估AI-LLM技术冲击对劳动力市场匹配的影响并进行机制的探讨。

劳动力市场错配是指劳动者人力资本特征与岗位需求之间不一致的现象。根据OECD(2017)的划分,劳动力市场错配具有多元性,主要包括学历错配(qualification mismatch)、专业错配(field-of-study mismatch)和技能错配(skill mismatch)三类。受在线招聘数据可得性限制,本文重点关注学历(纵向)错配与专业(横向)错配两类,暂不涉及技能错配。围绕这一主题,既有研究主要沿三条线索展开:教育错配的成因与后果、技术进步的劳动力市场效应,以及人工智能暴露度的量化测度。

首先,大量文献从教育错配的成因与后果角度揭示了劳动力市场中的结构性矛盾,也为本文考察人工智能技术冲击下教育错配的变化提供了直接参照。研究普遍发现,无论是纵向(学历)还是横向(专业)错配,均具有较高发生率,并伴随着显著的“工资惩罚”(wage penalty)与“疤痕效应”(scarring effect)(Duncan and Hoffman, 1981; Verdugo and Verdugo, 1989; Hartog, 2000; McGuinness, 2006; Robst, 2007; Tsai, 2010; Somers et al., 2019; 李骏, 2016; 李晓光, 2024)。在测度方法上,既有研究形成了包括“工作分析法”(job analysis approach)(Eckaus, 1964)、“实际匹配法”(realized matches approach)以及“主客观评估法”(subjective and objective assessment)(Verhaest and Omey, 2006; Bender and Heywood, 2011; Bender and Roche, 2013; Wolbers, 2003)在内的多元体系。在中国背景下,高等教育扩招在短期内加剧了高学历劳动力供给的集中释放(李晓光等, 2023),叠加年龄、经验与性别等个体特征差异(Frank, 1978; Green et al., 2007),以及用人单位对“过度资质”(over-qualification)可能引致更高离职率的预期(Kuhn and Shen, 2013),共同加剧了劳动力市场的匹配摩擦与结构性错配。

其次,关于技术进步的劳动力市场效应,经典文献普遍认为信息技术革命呈现出典型的“技能偏向型技术进步”特征(Autor et al., 2003; Acemoglu and Autor, 2011; Autor and Dorn, 2013; Goos and Manning, 2007),通过替代可程序化的常规性中等技能劳动,推动劳动力市场出现“就业极化”。在中国情境下,工业机器人的广泛应用亦被证实对低技能劳动者产生显著替代效应,并重塑岗位结构(王永钦和董雯, 2020; 李磊等, 2021; 闫雪凌等, 2020; 陈媛媛等, 2022; 王林辉等, 2023)。在信息技术影响匹配效率方面,也有研究表明互联网使用有助于改善教育-工作的匹配状况、降低教育错配风险(谢尚等, 2024)。然

而,与以往主要作用于体力或简单认知任务的自动化技术以及作用较为一般化的互联网技术不同,以大语言模型为代表的人工智能具备复杂语义理解与跨任务迁移能力,其影响范围正由“常规任务”扩展至高度专业化与创造性较强的非程序化认知活动(Brynjolfsson and McAfee, 2014; Brynjolfsson and Mitchell, 2017; Frey and Osborne, 2017; Eloundou et al., 2024)。因此,人工智能对就业的影响不再仅体现为岗位数量与结构的调整,更可能通过改变任务内容与需求表达方式,进一步作用于劳动力市场的匹配过程。总体而言,现有研究普遍认为人工智能对就业具有替代与创造的“双重效应”,其净影响取决于两者的相对强度(Acemoglu and Restrepo, 2019; Bessen, 2019; Deming and Noray, 2020; Acemoglu et al., 2022)。

最后,人工智能暴露度的量化为识别技术冲击提供了关键工具。现有研究多基于“任务视角”(task-based approach)构建度量框架。沿着 Felten et al. (2021)与 Webb (2020)的思路,Eloundou et al. (2024)利用招聘信息与文本语义分析,结合大模型评估,构建了任务层面的人工智能暴露度指标。国内研究中,张丹丹等(2025)将该方法拓展至中国劳动力市场,发现人工智能技术进步显著提高了高技能与高学历岗位的需求。^①

综上所述,既有文献虽已从多个维度揭示了教育错配的成因与后果、技术进步的劳动力市场效应以及人工智能暴露度的测度方法,但仍存在以下几方面不足,有待进一步拓展。第一,研究视角上,现有关于人工智能就业效应的文献多聚焦于宏观层面的岗位数量变化、职业结构调整或技能需求演化,较少从微观“岗位-求职者”匹配过程切入,难以刻画技术冲击下教育错配这一结构性问题的动态演变及其内在机制;而关于教育错配的文献虽积累丰富,但较少将其置于新一轮人工智能技术冲击的分析框架中加以考察,两类文献之间存在明显的研究割裂。第二,测度方法上,现有人工智能暴露度指标多基于职业层面的任务清单进行较为粗粒度的匹配,在任务描述的结构化处理、语义对齐的精确性以及指标构建的透明度方面仍有改进空间,难以充分捕捉 AI-LLM 技术冲击在不同职业之间的异质性影响。第三,政策含义上,既有研究多从就业替代或创造的总量视角提出政策建议,较少从劳动力市场结构性错配的视角出发,为就业引导、高等教育专业结构优化以及求职者匹配策略调整提供系统性的经验证据。

针对上述文献中的不足,本文的贡献如下。第一,在研究视角上,本文基于线上招聘平台“岗位-求职者”匹配数据,从微观匹配过程出发,系统识别了 AI-LLM 技术冲击对劳动力市场教育错配的影响,并对其作用机制进行了深入分

^① 附录 I 进行了更详尽的文献梳理总结。限于篇幅,附录未在正文列示,感兴趣的读者可在《经济学》(季刊)官网(<https://ceq.ccer.pku.edu.cn>)下载。

析,将人工智能就业效应研究与教育错配研究两条线索加以衔接,拓展了既有以宏观岗位或职业结构为主的研究范式。第二,在测度方法上,本文对既有 AI-LLM 暴露度指标进行了细化与优化:引入 Sentence-BERT 等语义嵌入模型,对岗位任务描述进行结构化拆分与标准化匹配,并结合招聘广告抽样数据进行验证,在指标构建的透明性与测度严谨性方面均有所提升,也更有助于捕捉 AI-LLM 冲击在不同职业之间的异质性影响。第三,在政策含义上,本文从劳动力市场结构性错配的视角出发,为政策制定提供了多层面的经验证据:一方面为就业引导政策的精准设计提供依据,另一方面为高等教育专业结构优化与人才培养体系改革提供参考,同时亦为求职者在人工智能冲击下调整职业选择与匹配策略提供了实证支持。

具体而言,本文使用 2021 年 1 月至 2025 年 7 月约 162 万条线上平台“岗位-求职者”匹配数据,从纵向(学历)与横向(专业)两个维度对劳动力市场错配进行测度,并在此基础上使用双重差分与事件研究法评估 AI-LLM 技术进步对两类劳动力市场错配的分别影响。具体而言,本文以 2022 年 12 月 ChatGPT 发布为代表的新一轮人工智能技术冲击作为区分 AI-LLM 产生影响的时点,并用职业层面的 AI-LLM 暴露指数作为对实验组(高暴露指数)和控制组(低暴露指数)的划分,并根据事前不同组间平行趋势的检验论证了对估计结果的因果判断。

本文的实证分析发现,中国劳动力市场基于学历的纵向错配与基于专业的横向错配比例分别平均为 59.9% 和 45.1%,且均呈显著上升趋势;同时,错配程度在不同职业之间存在显著差异,并在求职者年龄、学历等维度上表现出明显异质性。值得注意的是,尽管整体错配呈上升趋势,但以 2022 年 12 月 ChatGPT 发布为代表的新一轮人工智能技术冲击发生后,相较于低暴露度职业,高暴露度职业的向下投递行为及纵向错配比例显著下降,而对横向(专业)错配未见显著影响。机制分析表明,这一变化主要源于人工智能技术强化了岗位需求信号的清晰度,并提高了岗位所需技能的复杂性与学历门槛,从而在一定程度上改善了劳动力配置效率并降低了纵向错配。

二、研究数据与关键指标的测算

本文数据来自智联招聘线上平台。在线招聘广告具有实时性强的特点,能够及时追踪劳动力市场中需求端的变化,对于测算暴露指数有着天然优势。此外,招聘大数据还能够捕捉到求职者的投递记录与入职意愿,有助于研究市场均衡下的匹配情况。需要注意的是,已有研究表明(Kuhn and Shen, 2013;张丹丹等,2025),在线招聘广告通常更集中于城市地区、相对年轻且高学历人群,以

及中高薪酬岗位,因此在将结果推广至整体劳动力市场时需谨慎解释。

(一) 招聘广告数据与 AI-LLM 暴露指数的构建

我们使用来自智联招聘平台 2018—2020 年间的约 60 万条招聘广告随机样本,构造了中国劳动力市场职业层面的“AI-LLM 暴露指数”。指数构造所用数据的时间窗口设定在本文核心“岗位-求职者”匹配样本起始年份(2021 年)之前,有效避免了指数的内生性问题。本文的指数构造在张丹丹等(2025)的基础上加以改进,并借鉴 Hampole et al.(2025)的思路,结合中文招聘广告的具体特征,将任务提取工作细化为“信息过滤—任务提炼—质量清理—标准匹配”四个递进步骤,显著提高了构造过程的透明度与可追踪性。^①

具体而言,第一步为工作任务信息的过滤与抽取。招聘广告通常包含工作任务、任职要求、福利待遇及礼貌性表述等多类信息,仅前两类与指数构建直接相关。本文借助大语言模型结合提示词工程,从中文招聘广告中抽取与工作任务相关的有效信息,并通过包含任务描述、详细说明、逻辑指引、输出模板、参考示例等五个步骤的提示词控制,确保输出的逻辑性与稳定性。第二步为具体工作任务的提炼与翻译。在第一步过滤结果的基础上,参照 O*NET 数据库中任务(Task)的语义表述,进一步提炼出具体工作任务并翻译为英文,在翻译过程中尽可能保留中文广告的原语语义与岗位需求细节。第三步为任务表述的清理与质量审查。为剔除过于笼统(如“按要求完成其他任务”)或与岗位职责无关的低质量内容,本文设计了更严格的筛选与提取规则,最终约 10% 的无效任务表述被剔除。第四步为招聘任务与 O*NET 标准任务的语义匹配。本文采用 Sentence-BERT 框架(使用开源模型 paraphrase-mpnet-base-v2),将招聘任务与 O*NET 标准任务分别编码为 768 维语义向量,通过余弦相似度为每条招聘任务匹配最相近的 O*NET 标准任务;最终匹配相似度均值约为 0.724,表明整体匹配质量较高。四个步骤对应的提示词模板与技术细节详见附录 II。

(二) “岗位-求职者”匹配数据介绍

本文使用 2021 年 1 月至 2025 年 7 月的智联招聘平台数据构建“岗位-求职者”匹配样本。具体而言,研究以需求端岗位为抽样基础,每年随机抽取约 2 万个招聘广告形成“岗位样本”,并提取这些岗位在发布当年收到的全部投递记录,构成“投递样本”,同时识别对应的“求职者样本”。样本记录了岗位的发布

^① 详细的指标构造方法,以及与张丹丹等(2025)的对比,详见北京大学国家发展研究院 AI 与经济学实验室、智联招聘,“中国‘人工智能-大语言模型技术’暴露指数构建方法与数据说明”,v2025.09。我们在附录 II 中亦做了简要总结,并在附录 III 中列出了各职业的暴露度。在本部分抽样数据中,2018—2020 年的数据样本与张丹丹等(2025)完全一致,保证了指数的可比性。

时间、职业、行业、地域等信息,以及求职者的年龄、性别、学历、专业和上一份工作履历等特征。平台根据企业与求职者之间的互动行为识别“正向回复”,本文将将其视为接近达成雇佣的代理变量。为保证数据完整性,研究仅保留每年1—10月(2025年为1—5月)发布的招聘广告,并匹配发布后两个月内的投递记录,同时限定岗位类型为全职。经筛选后,最终样本包含97 447个招聘广告、1 623 430次投递以及1 439 854名求职者,其中大专及以上学历样本占比超过九成。描述性统计显示,样本期间平均正向回复率约为19.4%。此外,约一半招聘广告未收到投递,有投递样本在AI-LLM暴露度、薪资和学历要求等方面显著更高。更为详细的数据构建过程与描述性统计结果见附录IV。有投递岗位与无投递岗位的对比见附录V。

(三) 基于“岗位-求职者”匹配数据构建教育错配指标

1. 职业、专业的标准化对应

为构建岗位-求职者错配指标,本文首先对招聘广告中的职业类别与求职者专业信息进行标准化处理。参照国际通行的职业与专业分类体系,本文分别将岗位职业与求职者专业对应到SOC(Standard Occupational Classification)职业分类标准与CIP(Classification of Instructional Programs)学科专业分类标准。在具体操作上,针对“岗位样本”中近千种职业类别,本文结合大语言模型辅助识别与专家筛选评估,将其映射至SOC的具体职业层级;针对“求职者样本”中约4.4万类自填专业,则基于Sentence-BERT语义相似度匹配并结合专家审核,将其对应至CIP具体专业分类。在专业维度分析中,本文主要关注大专及以上学历群体,该群体在样本中占比超过90%,且专业信息更具识别度。关于SOC与CIP分类体系及具体匹配流程的更为详细说明,见附录VI。

2. 纵向错配与横向错配的测算

本文从投递视角出发,定义“向下投递”与“跨专业投递”。“向下投递”指“投递样本”中,求职者的学历高于申请岗位的学历要求的投递。本文将学历划分为五个层级:不限学历或初中及以下为1级;高中、中专、中技为2级;大专为3级;本科为4级;研究生为5级。根据学历层级的相对关系判断是否为“向下投递”。

“跨专业投递”指“投递样本”中,求职者所学专业与职位所要求的专业不一致的投递。其识别依赖于求职者的专业信息与职位所需专业之间的匹配关系,现有文献(如Manuel and Plesca, 2020; Manuel, 2024)大多将美国劳工统计局(Bureau of Labor Statistics, BLS)和美国国家教育统计中心(National Center for Education Statistics, NCES)协作构建的CIP-SOC对照表作为参考依据。本研究依据大语言模型辅助与Sentence-BERT语义相似度匹配,将智联招聘的

职业分类与求职者专业分别对应到 SOC 职业标准与 CIP 专业标准中,将不在 CIP-SOC 对照表中的投递行为视作跨专业投递。为减少模糊界定的影响,我们将 CIP-SOC 对照表中的 SOC 职业聚合到职组小类(Minor Group, 98 个),将 CIP 职业聚合到大类(Major, 48 个)^①。

由于大专及以上学历的申请者专业性较强且该人群数量占求职者样本的比例超过 90%,因此在研究“跨专业投递”和“横向错配”问题时,我们将把样本分别局限在由大专及以上学历的申请者发出的投递和获得的正向回复中(以下简称大专及以上学历子样本),需要注意的是,对岗位要求学历不进行限制。

错配一般指实际入职匹配的情况,与投递行为仍有概念上的差距。由于招聘数据中无法追踪到实际匹配情况,本文用“正向回复”这一表示招聘双方意愿的指标进行间接反映。在有正向回复的“投递样本”子样本中,本文将“向下投递”视为“纵向错配”^②,将“跨专业投递”视为“横向错配”。

在“投递样本”中,存在 2.0%的投递记录缺乏求职者学历信息,无法直接判断是否为“向下投递”;在“大专及以上学历子样本”中,存在 1.6%的投递记录缺乏原始专业信息,或是与 CIP 标准专业的语义相似度略低于阈值 0.5^③,无法有效判断是否为“跨专业投递”。经分析,这两部分投递记录的其他特征分布与全样本特征分布存在显著差异,直接去除可能会对结果产生一定影响。本文采用机器学习算法,基于投递记录的其余信息,对关键信息缺失的投递样本进行“向下投递”与“跨专业投递”的预测,详细信息参见附录 IX。

三、人工智能大语言模型技术职业暴露度对劳动力市场错配的影响

(一) 劳动力市场错配趋势分析

应用“岗位-求职者”匹配数据,采用本文第二部分的测算方法,中国投递行为和劳动力错配的趋势如图 1 所示。

图 1 刻画了 2021—2025 年劳动力市场投递行为与错配的动态变化。样本期内,约 40%—45%的投递属于高学历求职者向低学历岗位的“向下投递”,约 40%的投递为大专及以上学历申请者的“跨专业投递”,两类行为整体较为稳定。然而,以“正向回复”作为匹配成功的代理指标,错配程度显著上升:纵向错配比例由 2021 年的 52.0%上升至 2025 年的 64.9%,横向错配比例也经历了从最低 2022 年的 40.7%上升至高值 49.3%的过程。需要指出的是,本文测算的错配比

① 聚合层级会影响横向错配的测算,在附录 VII 中对不同聚合层级下的平均横向错配比例予以呈现。

② 理论上,“向上投递”也属于“纵向错配”,但由于正向回复样本中几乎不存在求职者学历低于岗位要求的情况,因此本文不单独对“向上投递”进行考虑。

③ 关于这一阈值的选择问题,参见附录 VIII。

例高于既有基于劳动力调查数据的结果。既有研究多基于就业存量样本,通常得到20%—40%的错配率区间(Duncan and Hoffman, 1981; Verdugo and Verdugo, 1989; 李晓光等, 2023);而本文基于招聘筛选阶段的“岗位-求职者”匹配数据,刻画的是流量型错配,数值不宜直接对比,但在时序上具有良好内部可比性。

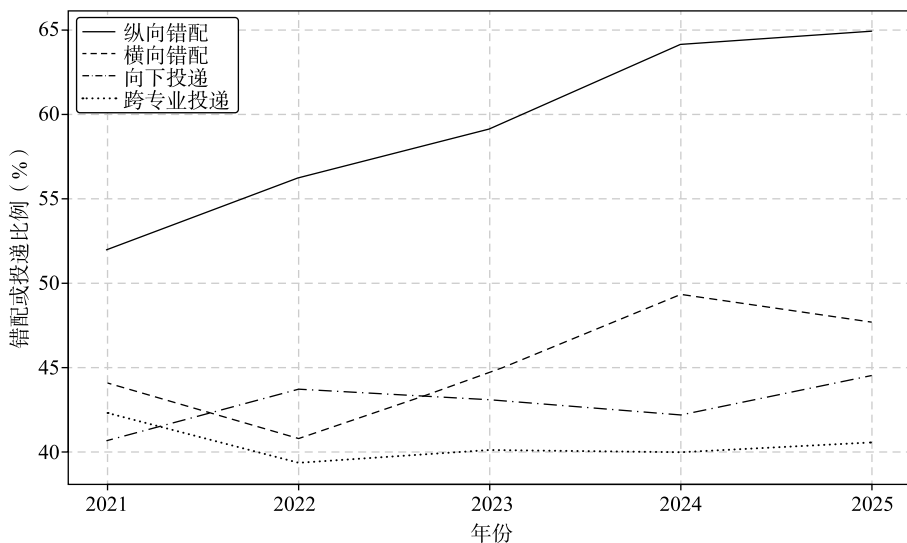


图1 基于“岗位-求职者”匹配数据的投递行为和劳动力市场错配时间趋势

进一步比较发现,纵向错配显著高于向下投递,横向错配亦高于跨专业投递,表明用人单位在筛选中对高学历及跨专业求职者存在偏好,且这一趋势有所强化。其中,对向下投递的偏好尤为突出,反映出日益强化的“高学历优势”:高学历求职者更易获得用人单位认可,但也可能挤出原本匹配的申请者,从而推动纵向错配上升。纵向错配和横向错配反映了当下劳动力市场配置中的结构性问题。^①不断增长的纵向错配表明,部分高学历群体面临“学历过剩”的困境,难以充分发挥其学历优势,可能造成人力资源的浪费,也可能加剧这部分群体的就业压力与职业满意度偏低的情况。不断增长的横向错配表明,教育专业设置与市场行业需求之间的衔接不够紧密,产业结构调整与人才培养体系之间存在一定的滞后性,难以快速响应市场对特定专业人才的需求变化。关于不同职组及各类求职者群体错配分布的详细图表与具体分析,请参见附录X。

(二) 职业错配比例的变动与职业暴露度的关系

图2呈现了职业错配比例的变动与职业暴露度的关系。具体而言,图中每

^① 值得注意的是,线上招聘数据不足以覆盖整个劳动力市场的情况。就智联招聘平台而言,求职者以青壮年、中学历居多,这部分群体往往更容易发生错配,所以图1中测算的比例可能是对中国总体劳动力市场情况的高估。

个点代表 SOC 广义职业,横坐标为该职业的暴露指数,纵坐标为该职业在 ChatGPT 发布后的平均错配比例与 ChatGPT 发布前的平均错配比例的差值。为了稳健性,图中删除了包含样本较少的职业点。图 2(a)呈现了纵向错配的变动情况与职业暴露度的关系,图 2(b)呈现了横向错配的变动情况与职业暴露度的关系。可以初步发现,高暴露度职业的平均错配比例的增长幅度显著低于低暴露度职业,甚至在较高暴露度的职业中有所下降;而在横向错配的比例变动中,高低暴露度职业并没有呈现出明显的异质性。下一小节将通过更严谨的实证方法提供证据。

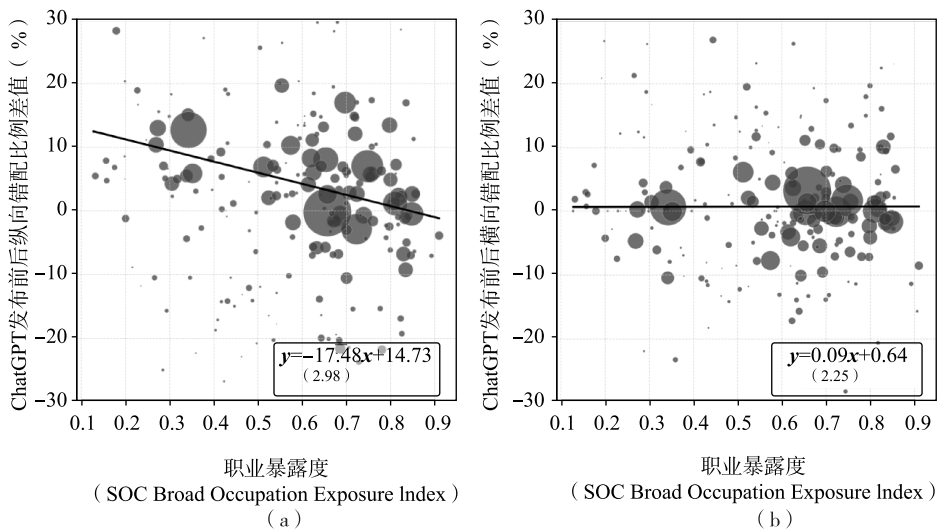


图 2 职业错配比例的变动与职业暴露度的散点图

注:图 2(a)代表纵向错配的情况,图 2(b)代表横向错配的情况。图中每一个点代表职业(SOC 广义职业,6 位编码),点的大小代表在 ChatGPT 发布前的职业份额占比大小,横坐标为该职业的暴露度,纵坐标为该职业在 ChatGPT 发布后(2022 年 12 月至 2025 年 7 月)与 ChatGPT 发布前(2021 年 1 月至 2022 年 11 月)的平均错配比例的差值。每个图的右下角标注了采用加权最小二乘估计的拟合曲线,括号中为对应系数的标准误。

(三) 双重差分实证方法

本文以招聘广告为主要研究对象,采用双重差分法,识别新一轮人工智能技术冲击对中国劳动力市场错配的因果影响。我们以 ChatGPT 的发布作为技术冲击到来的标志,识别其对潜在“AI-LLM 暴露指数”不同的职业的不同影响。具体而言,我们使用以下回归方程:

1. 双重差分(Difference-in-Differences, DiD)回归

$$Y_{ijct} = \alpha + \beta ExpOccu_j \times Post_t + \gamma X_i + \mu_{ct} + \lambda_j + \epsilon_{ijct}, \quad (1)$$

其中, i 表示招聘广告, j 表示招聘广告所在的 SOC 广义职业, c 表示城市, t 表示时间, 具体到月份。 Y_{ijct} 为本文关注的被解释变量, 本文使用了两类特征展开分析。第一类特征聚焦市场的投递与正向回复(接近雇佣)情况: ①招聘广告是否有求职者投递; ②招聘广告收到的投递量; ③招聘广告的正向回复率。第二类特征聚焦市场的错配情况: ①招聘广告收到的向下投递比例; ②招聘广告的纵向错配比例; ③招聘广告收到的跨专业投递比例; ④招聘广告的横向错配比例。

$ExpOccu_j$ 表示该招聘广告所在的职业 j 的暴露指数, 为方便解释回归结果, 回归中使用的暴露指数进行了 Z-Score 标准化处理。本文以 2022 年 12 月 ChatGPT 发布作为主要事件冲击的时点。 $Post_t$ 是关于冲击时期的指示变量, 在 2022 年 12 月及之后取值为 1, 此前取值为 0; X_i 表示招聘广告层面的一系列控制变量, 包括该岗位的拟招聘数量、要求学历、月薪、岗位所在企业所有制类型、岗位所在企业规模等。 μ_{ct} 表示“时间-城市”层面的固定效应, 它可以控制由于招聘平台的经营策略引起的地区间劳动力市场趋势不一致的情况; λ_j 表示职业层面的固定效应。 ϵ_{ijct} 表示误差项。研究将标准误差聚类到 SOC 广义职业层面, 与核心解释变量的层面保持一致。

2. 事件研究法(Event Study)

$$Y_{ijct} = \sum_{\tau \neq -1} \beta_{\tau} I(t = \tau) \times ExpOccu_j + \gamma X_i + \mu_{ct} + \lambda_j + \epsilon_{ijct}. \quad (2)$$

为进一步研究 ChatGPT 发布前后, 不同暴露度职业错配情况的动态变化, 本文采用事件研究法进行分析。为保证结果的稳健性, 本部分将时间加总在季度层面, 将 2022 年第 4 季度(即 2022 年 12 月 ChatGPT 发布所在季度)作为事件冲击的第 0 期。 $I(t = \tau)$ 为时期的指示函数, 当且仅当第 τ 期时取值为 1, 其余情况取值为 0。其余变量与式(1)保持一致。

(四) 主要实证结果

表 1 基于双重差分法, 呈现了人工智能大语言模型技术应用对市场投递与正向回复(接近雇佣)情况的影响。第(1)列展示了技术冲击并不显著影响招聘投递的外扩边际(extensive margin), 即招聘广告收到投递的概率。第(2)列则显示技术冲击使高暴露度职业获得了更多投递, 这一增加主要来自集约边际(intensive margin): 与冲击前相比, 职业暴露度每增加 1 个标准差, 在有投递岗位中的投递数量平均增加 11.30 人次。投递人次的增加意味着竞争更加激烈, 如第(3)列所示, 高暴露度职业在技术冲击后的正向回复率显著下降。

表 2 是主要结果, 呈现了在以 2022 年 12 月 ChatGPT 发布为代表的新一轮 AI-LLM 技术冲击对市场错配情况的影响。第(1)列的结果显示, 在技术冲击之后, 更高暴露度的职业吸引了更少比例的向下投递, 职业暴露度每增加 1 个标准差, 向下投递比例将减少 2.6 个百分点。第(2)列使用了所有形成了“近

似雇佣”(正向回复)的岗位样本,被解释变量是岗位中所有形成了正向回复的匹配中纵向错配的比例。回归显示,技术冲击使得高暴露度职业的纵向错配比例同样下降。这意味着,与低暴露度职业相比,高暴露度岗位所吸引的向下投递与最终形成的纵向错配比例在冲击后明显降低,本文的机制部分将具体对这一现象展开分析。表2的第(3)、(4)列相应地分析了跨专业投递与横向错配,没有发现技术冲击的显著影响。

图3对事件研究法中交互项系数 β_x 的趋势变化进行呈现。其中子图(a)至(d)分别对应表2的第(1)–(4)列回归的事件研究版本。技术冲击到来之前,暴露度不同的岗位在向下投递和纵向错配方面没有显著差异(系数接近0)和变化趋势;2022年第4季度之后,交互项系数呈现出明显的降低,且在2024年年底至2025年年初尤为突出。

表1 人工智能大语言模型技术应用对市场投递与回复情况的影响

样本 结果变量	全部岗位		有投递岗位	
	有投递 (1)	投递人次 (2)	正向回复率 (3)	
ChatGPT发布后×职业暴露指数	0.015 (0.0103)	11.30*** (2.069)	-0.0300*** (0.00613)	
岗位控制变量	是	是	是	
年月-城市固定效应	是	是	是	
职业固定效应	是	是	是	
观测数	91 897	46 462	46 462	
调整后R ²	0.482	0.157	0.334	
被解释变量均值	0.517	33.274	0.343	

注:“岗位控制变量”包括某岗位的拟招聘数量、要求学历、月薪、岗位所在企业所有制类型、岗位所在企业规模,第(3)列还控制了该岗位的投递人次。括号中为标准误(标准误聚类到职业层面,SOC广义职业,6位编码);***代表在1%的水平上统计显著,**代表在5%的水平上统计显著,*代表在10%的水平上统计显著,下同。

表2 人工智能大语言模型技术应用对市场错配情况的影响

岗位样本 结果变量	有投递	有正向回复	有投递#	有正向回复#
	向下投递比例 (1)	纵向错配比例 (2)	跨专业投递比例 (3)	横向错配比例 (4)
ChatGPT发布后×职业暴露指数	-0.0263*** (0.00589)	-0.0230*** (0.00572)	-0.00214 (0.00344)	-0.000123 (0.00487)
岗位控制变量	是	是	是	是
年月-城市固定效应	是	是	是	是

(续表)

岗位样本 结果变量	有投递	有正向回复	有投递#	有正向回复#
	向下投递比例	纵向错配比例	跨专业投递比例	横向错配比例
	(1)	(2)	(3)	(4)
职业固定效应	是	是	是	是
观测值	46 462	31 096	42 262	26 999
调整后 R ²	0.721	0.63	0.603	0.524
被解释变量均值	0.545	0.592	0.496	0.505

注：#表示第(3)、(4)列仅包含岗位要求为大专及以上学历的样本。“岗位控制变量”包括拟招聘人数、要求学历、月薪、岗位所在企业所有制类型、岗位所在企业规模、投递人次。

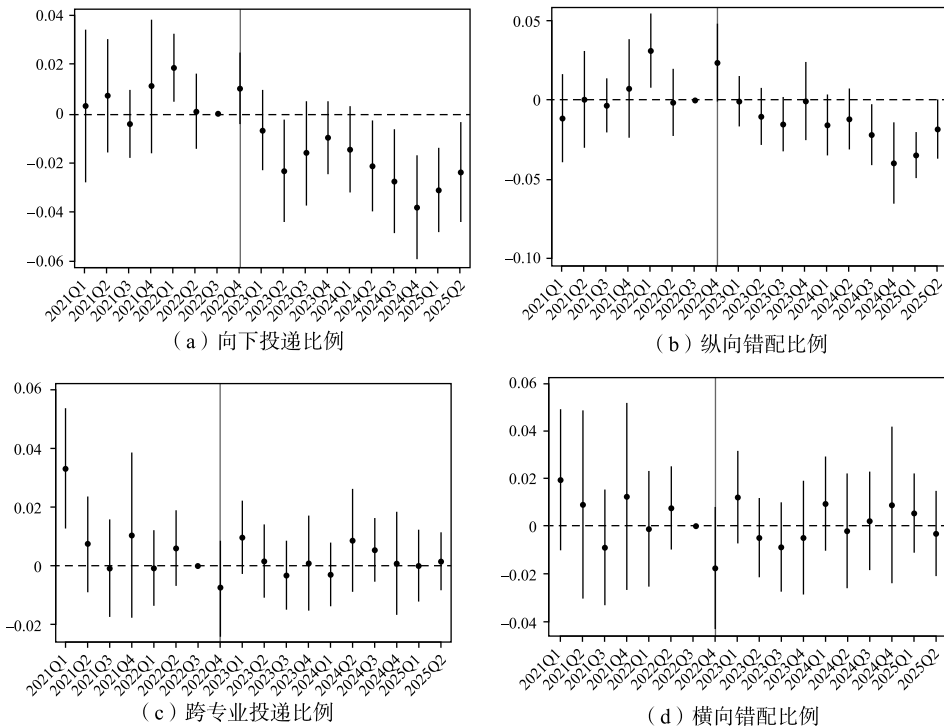


图 3 人工智能大语言模型技术冲击对市场错配情况的动态影响

注：图中呈现了事件研究中交互项的系数。回归包含职业和年-季度的固定效应，同时控制了岗位要求学历、薪资水平等招聘广告层面的特征。标准误在职业层面聚类，置信区间为 95% 水平。

四、技术冲击下的错配机制与异质性分析

人工智能技术的快速发展，给企业带来了显著的任务重组与技能替代压力。对于高度暴露的职业而言，技术冲击不仅改变可自动化的任务边界，也改变岗位内部的技能结构，从而影响企业的岗位发布方式和求职者的投递行为。

本部分旨在探讨总体样本中错配持续上升而高暴露职业的纵向错配显著下降的现象,并揭示支撑这一变化的微观机制。

(一) 需求端:高暴露岗位的任务与技能专业化

技术冲击首先作用于企业端的生产组织方式与岗位需求结构。随着生成式 AI 的广泛应用,高暴露职业的任务边界与技能需求发生系统性重构,使岗位内容呈现出更高复杂度与专业化水平。本文基于“岗位-求职者”匹配数据,从招聘广告中提取任务并映射至 O*NET 任务体系,同时利用大语言模型提取岗位技能(提示词见附录 XI),并通过语义匹配对应至 Lightcast 分类^①;进一步参考 Tacchella et al.(2012)与 Aufiero et al.(2024),采用 Economic Fitness and Complexity(EFC)算法^②刻画技能复杂度与专业性。表 3 结果表明,技术冲击后高暴露职业的岗位内容发生显著调整:一方面,任务数量与技能数量均显著增加,表明岗位涉及的工作内容更加多样,所需技能组合更为广泛;另一方面,技能复杂度显著提升,反映出不同技能之间的互补性增强,对综合处理能力的要求提高。同时,技能专业性显著上升,说明岗位技能由通用型向特定领域聚焦,岗位“画像”更加清晰。

从劳动力市场机制看,岗位任务与技能要求的重构提高了招聘信号的清晰度与筛选标准。在信息不完全程度较高时,不匹配的求职者仍可能进入申请序列;而在技能要求更加明确且专业化程度提高后,这类求职者的投递行为将减少,从而直接导致高暴露职业中的向下投递与纵向错配显著下降。

表 3 企业发布招聘广告的主动调整

被解释变量	工作任务数量	技能数量	最高技能复杂度排名
	(1)	(2)	(3)
ChatGPT 发布后 × 职业暴露指数	0.0710** (0.0344)	0.0812* (0.0468)	-2.013** (1.016)
年月-城市固定效应	是	是	是
职业固定效应	是	是	是
观测值	46 462	45 390	45 053
调整后 R ²	0.256	0.506	0.346
被解释变量均值	3.699	9.532	271.862

注:回归中控制了该岗位的拟招聘人数、要求学历、月薪、岗位所在企业所有制类型、岗位所在企业规模、投递人次等变量,同时分别控制了年月和城市层面的固定效应。

① Lightcast 技能分类体系由美国劳动力市场数据服务商 Lightcast 构建,整合机器学习与专家人工审核,依托海量招聘信息和职业档案动态更新,可快速捕捉新兴技能。体系涵盖超 32 000 项独立技能,采用三层层级结构(大类-子类-具体技能),为劳动力市场技能需求分析提供标准化框架。

② EFC 算法基于职业-技能网络,迭代计算职业适应性与技能复杂度;目的是量化职业技能组合复杂度和技能稀有度,揭示技能结构与薪资的内在关联。

（二）岗位层级的异质性

企业端岗位内容的专业化不仅改变了高暴露职业内部的技能需求,也系统性地影响了不同层级岗位的可进入性和吸引力。为此,本文从岗位学历要求与薪酬水平两个维度考察技术冲击在不同职业层级上的影响,并分析其与向下投递及纵向错配变化之间的关系。表4结果表明,在学历维度上,本科岗位的向下投递与纵向错配在技术冲击后整体保持稳定,反映出岗位专业化提升强化了技能信号,使高技能岗位的匹配结构相对稳定;大专岗位的向下投递与错配显著上升,表明随着本科岗位准入门槛上移,高学历求职者更多向中等技能岗位集中,形成“中段挤占”;而大专以下岗位的错配则明显下降,高学历求职者对低技能岗位的兴趣减弱。薪酬维度的结果与之相一致:低薪岗位的向下投递与错配显著下降,而高薪岗位变化较小且不显著。总体而言,AI技术冲击通过提升岗位技能要求与专业化程度,推动高技能门槛上移、中技能岗位竞争加剧、低技能岗位吸引力下降,从而在不同岗位层级上形成差异化的错配调整。

（三）供给端的异质性反应:技能比较优势和可迁移性

本节从求职者专业背景与工作经历两个维度出发,检验不同类型人力资本在新技术冲击中的相对表现。我们沿用双重差分模型的方法,在“投递”层面进行回归分析,以求更精细地描述求职者行为。具体来说,我们使用了以下回归方程:

$$Y_{pict} = \alpha + \beta ExpOccu_j \times Post_t + \gamma X_i + \theta Z_p + \lambda_j + \mu_{ct} + \epsilon_{pict}. \quad (3)$$

相对于式(1),式(3)以每个岗位-求职者投递为分析单元。下标 p 代表求职者,其他字母含义保持与式(1)相同, Z_p 为求职者层面的控制变量。我们关心的被解释变量 Y_{pict} 为该投递是否为向下投递;如果使用所有正向回复的投递样本,被解释变量则可解读为是否形成了纵向错配。

表5从求职者专业背景与工作经历两个维度考察技术冲击下求职行为的异质性。首先,从专业背景看,科学、技术、工程、数学类专业(Science、Technology、Engineering、Mathematics,以下简称STEM专业)求职者在高暴露职业中的向下投递与纵向错配均未出现显著上升,部分群体甚至表现出更低的错配水平。这表明,在岗位专业化与技能复杂度提升的背景下,具备科学技术相关背景的求职者更容易适应岗位技能结构的变化,从而在高技能职业轨道中保持相对优势。相比之下,非STEM专业求职者在高暴露职业中的向下投递与纵向错配显著上升,说明随着岗位技能要求提高,这类求职者在高技能岗位中的竞争力下降,被迫向中等技能岗位转移,从而加剧了中段职业梯度的竞争。

表4 分岗位学历要求、岗位薪资的回归结果

分组依据	分岗位学历要求						分岗位薪资					
	本科岗位		大专岗位		大专以下岗位		高薪岗位		中薪岗位		低薪岗位	
	向下投递 比例	纵向错配 比例	向下投递 比例	纵向错配 比例	向下投递 比例	纵向错配 比例	向下投递 比例	纵向错配 比例	向下投递 比例	纵向错配 比例	向下投递 比例	纵向错配 比例
样本	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		
结果变量												
ChatGPT发布后×职业暴露指数	0.00595 (0.00773)	-0.00990 (0.0162)	0.0239*** (0.00919)	0.0242* (0.0131)	-0.0332*** (0.00630)	-0.0358*** (0.00557)	-0.0124* (0.00714)	-0.00549 (0.0101)	-0.0230*** (0.00623)	-0.0231*** (0.00666)		
岗位控制变量	是	是	是	是	是	是	是	是	是	是	是	是
年月-城市固定效应	是	是	是	是	是	是	是	是	是	是	是	是
职业固定效应	是	是	是	是	是	是	是	是	是	是	是	是
观测值	9 197	4 844	14 352	9 040	18 923	14 300	14 673	8 836	29 913	20 804		
R ²	0.337	0.309	0.306	0.311	0.243	0.238	0.757	0.676	0.708	0.627		
被解释变量均值	0.111	0.114	0.452	0.467	0.860	0.866	0.453	0.514	0.590	0.624		

注：奇数列的样本限定在有投递的招聘广告中，偶数列的样本限定在有正向回复的招聘广告中。“岗位控制变量”包括拟招聘人数、要求学历、月薪、岗位所在企业所有制类型、岗位所在企业规模、投递人次。

表 5 分求职者专业、工作经历的回归结果(投递层面)

分组依据	分求职者专业					分求职者工作经历				
	STEM 专业求职者		其他专业求职者		有高暴露度工作经历		有低暴露度工作经历		无工作经历	
	是否为 向下投递	是否为 纵向错配	是否为 向下投递	是否为 纵向错配	是否为 向下投递	是否为 纵向错配	是否为 向下投递	是否为 纵向错配	是否为 向下投递	是否为 纵向错配
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
样本	0.000576	-0.00363	0.00865**	0.0139***	0.00775**	0.00996***	0.00263	0.00541	-0.000607	0.00525
结果变量	(0.00315)	(0.00380)	(0.00343)	(0.00353)	(0.00383)	(0.00320)	(0.00312)	(0.00388)	(0.00421)	(0.00590)
ChatGPT 发布后 ×	是	是	是	是	是	是	是	是	是	是
职业暴露指数	是	是	是	是	是	是	是	是	是	是
岗位控制变量	是	是	是	是	是	是	是	是	是	是
求职者控制变量	是	是	是	是	是	是	是	是	是	是
年月-城市固定效应	是	是	是	是	是	是	是	是	是	是
职业固定效应	是	是	是	是	是	是	是	是	是	是
观测值	459 380	78 756	967 768	167 777	1 235 891	208 671	232 448	61 944	94 735	22 690
R ²	0.779	0.792	0.784	0.768	0.760	0.751	0.751	0.772	0.792	0.767
被解释变量均值	0.392	0.541	0.430	0.599	0.400	0.552	0.515	0.646	0.501	0.676

注：偶数列的样本限定在有正向回复的投递记录中。“岗位控制变量”包括拟招聘人数、要求学历、月薪、岗位所在企业所有制类型、岗位所在企业规模、投递人次。
“求职者控制变量”包括求职者的年龄、性别、学历层级、毕业院校层级。

其次,从工作经历维度看,是否具备高暴露职业经历同样影响求职者的调整方式。曾在高暴露职业中工作的求职者,在技术冲击后表现出更高比例的向下投递和更明显的纵向错配,表明岗位任务结构的变化对其既有技能形成了直接冲击,使其在职业梯度上出现一定程度的“向下迁移”。相比之下,不具备高暴露经历的求职者受岗位技能重构的约束较弱,其投递行为与错配变化相对温和。

总体而言,技术冲击通过岗位专业化和技能结构重构,在企业需求与求职者供给两端共同作用。一方面,高暴露岗位的技能要求提升提高了招聘信号的清晰度,从而降低了职业内部的错配;另一方面,不同能力与经验的求职者在技术冲击下面临不同的调整路径,进一步强化了劳动力市场错配的结构化变化。

五、政策建议与研究局限

基于本文结论,人工智能技术进步并未简单地加剧所有岗位的错配,而是通过重塑岗位任务与技能结构,使高暴露职业内部的招聘信号更加清晰、纵向错配相对下降;与此同时,整体劳动力市场错配仍在上升,并表现出显著的层级分化与群体异质性。据此,政策重点不宜停留于笼统的稳就业或促培训,而应转向围绕技术冲击下的匹配效率提升,构建更具针对性的调节机制。

首先,应建立以岗位需求变化为基础的动态监测与预警体系。本文发现,AI冲击后高暴露职业的技能要求、学历门槛和专业化程度明显提升,中技能岗位则成为高学历劳动者“向下流动”的主要承接层。因此,政策部门应依托招聘平台、职业分类和岗位文本数据,持续监测不同职业的AI暴露度、技能要求变化和错配走势,及时识别错配高发职业、高风险群体和技能门槛快速抬升的领域,并将其纳入就业公共服务和职业培训资源配置之中。与其在错配形成后被干预,不如在岗位需求结构变化初期提前预警,从源头上缓解结构性挤压。

其次,应推动教育供给体系从规模扩张转向适配优化。本文显示,高暴露岗位内部错配下降,并非因为市场自动出清,而是因为岗位要求更清晰、筛选标准更高;相应地,缺乏相关技能和专业背景的劳动者更容易被挤出并向中技能岗位集中。这意味着,当前教育体系面临的核心问题已不只是“是否培养足够多的高学历劳动者”,而是“能否培养与新岗位要求相匹配的人力资本”。因此,高等教育和职业教育应加强与岗位技能需求的动态衔接,适度扩大与AI互补性较强领域的人才培养,强化通用数字能力、任务迁移能力和跨学科应用能力训练,同时避免单纯以学历提升替代技能提升。对于大专层次教育,应更加突出应用型、职业特定型能力建设,减少中技能岗位被动承接高学历挤压所带来的错配累积。

再次,就业政策应更加关注技术冲击下的脆弱群体。本文发现,STEM背景求职者在高暴露职业中并未出现显著错配上升,而非STEM群体以及具有高暴露职业经历的劳动者更容易发生向下投递和纵向错配。这表明,技术冲击并非均匀作用于所有劳动者,而是通过技能可迁移性和技术互补性形成差异化影响。基于此,公共政策应将就业帮扶和再培训重点更多投向非科技专业背景劳动者、专用人力资本较强但可迁移性较弱的从业者,以及处于职业转换阶段的中年劳动者。培训设计也不应停留于一般性数字技能普及,而应强调与具体职业场景结合的转岗能力和适应能力提升。

最后,应引导企业在招聘端发布更清晰的岗位信息。本文表明,岗位任务与技能要求越清晰,越有助于减少高暴露职业内部的错配。因此,企业在招聘中应进一步提高岗位描述、技能要求与职业路径信息的透明度,减少模糊性筛选和过度学历偏好。同时,可探索建立更强调技能识别而非单一学历门槛的招聘机制,为具备可迁移能力但学历不占优的劳动者保留进入机会。只有在政策、教育与企业三端协同推进的条件下,才能将AI带来的岗位重构转化为匹配效率提升,而不是演变为更大范围的人力资本错置。

本文存在以下两点局限。一方面,本文尝试从需求与供给两个角度,识别大语言模型等人工智能技术进步对劳动力市场错配的影响机制,但从一般均衡视角来看,二者的调整过程必然相互作用,因此难以完全分离出两者的独立影响。例如,从供给侧来看,求职者针对人工智能大语言模型技术冲击所做出的投递行为调整,本身就会受到新技术冲击引发的岗位需求结构变动的影响。另一方面,人工智能大语言模型技术冲击不仅会影响求职者的短期投递行为,还可能通过更为缓慢的人力资本投资与技能结构重塑产生长期作用。但受限于本文仅使用流量数据(新增岗位招聘信息),无法系统分析人工智能大语言模型技术冲击对求职者的长期影响。未来研究可结合更长时间跨度的数据与更严谨的识别策略,弥补本文的不足。

参考文献

- [1] Acemoglu, D., and D. H. Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings", In: Card, D., and O. Ashenfelter (eds.), *Handbook of Labor Economics* (Vol. 4). Elsevier, 2011, 1043-1171.
- [2] Acemoglu, D., D. Autor, J. Hazell, and P. Restrepo, "Artificial Intelligence and Jobs: Evidence from Online Vacancies", *Journal of Labor Economics*, 2022, 40(S1), S293-S340.
- [3] Acemoglu, D., and P. Restrepo, "Automation and New Tasks: How Technology Displaces and Reinstates Labor", *Journal of Economic Perspectives*, 2019, 33(2), 3-30.
- [4] Aufiero, S., G. De Marzo, A. Sbardella, and A. Zaccaria, "Mapping Job Fitness and Skill Coherence into Wages: An Economic Complexity Analysis", *Scientific Reports*, 2024, 14(1), 11752.

- [5] Autor, D. H., and D. Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market", *American Economic Review*, 2013, 103(5), 1553-1597.
- [6] Autor, D. H., F. Levy, and R. J. Murnane, "The Skill Content of Recent Technological Change: An Empirical Exploration", *The Quarterly Journal of Economics*, 2003, 118(4), 1279-1333.
- [7] Bender, K. A., and J. S. Heywood, "Educational Mismatch and the Careers of Scientists", *Education Economics*, 2011, 19(3), 253-274.
- [8] Bender, K. A., and K. Roche, "Educational Mismatch and Self-employment", *Economics of Education Review*, 2013, 34, 85-95.
- [9] Bessen, J., "Automation and Jobs: When Technology Boosts Employment", *Economic Policy*, 2019, 34(100), 589-626.
- [10] Brynjolfsson, E., and A. McAfee, *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: W. W. Norton & Company, 2014.
- [11] Brynjolfsson, E., and T. Mitchell, "What Can Machine Learning Do? Workforce Implications", *Science*, 2017, 358(6370), 1530-1534.
- [12] 陈媛媛、张竞、周亚虹, "工业机器人 与 劳动力的 空间配置", 《经济研究》, 2022 年第 1 期, 第 172—188 页。
- [13] Deming, D. J., and K. Noray, "Earnings Dynamics, Changing Job Skills, and STEM Careers", *The Quarterly Journal of Economics*, 2020, 135(4), 1965-2005.
- [14] Duncan, G. J., and S. D. Hoffman, "The Incidence and Wage Effects of Overeducation", *Economics of Education Review*, 1981, 1(1), 75-86.
- [15] Eckaus, R. S., "Economic Criteria for Education and Training", *The Review of Economics and Statistics*, 1964, 46(2), 181-190.
- [16] Eloundou, T., S. Manning, P. Mishkin, and D. Rock, "GPTs Are GPTs: Labor Market Impact Potential of LLMs", *Science*, 2024, 384(6702), 1306-1308.
- [17] Felten, E., M. Raj, and R. Seamans, "Occupational, Industry, and Geographic Exposure to Artificial Intelligence: A Novel Dataset and Its Potential Uses", *Strategic Management Journal*, 2021, 42(12), 2195-2217.
- [18] Frank, R. H., "Why Women Earn Less: The Theory and Estimation of Differential Overqualification", *American Economic Review*, 1978, 68(3), 360-373.
- [19] Frey, C. B., and M. A. Osborne, "The Future of Employment: How Susceptible Are Jobs to Computerisation?", *Technological Forecasting and Social Change*, 2017, 114, 254-280.
- [20] Goos, M., and A. Manning, "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain", *Review of Economics and Statistics*, 2007, 89(1), 118-133.
- [21] Green, C., P. Kler, and G. Leeves, "Immigrant Overeducation: Evidence from Recent Arrivals to Australia", *Economics of Education Review*, 2007, 26(4), 420-432.
- [22] Hampole, M., D. Papanikolaou, L. D. W. Schmidt, and B. Seegmiller, "Artificial Intelligence and the Labor Market", *NBER Working Paper*, 2025, No. 33509.
- [23] Hartog, J., "Over-education and Earnings: Where Are We, Where Should We Go?", *Economics of Education Review*, 2000, 19(2), 131-147.
- [24] Kuhn, P., and K. Shen, "Gender Discrimination in Job Ads: Evidence from China", *The Quarterly Journal of Economics*, 2013, 128(1), 287-336.
- [25] 李骏, "中国高学历劳动者的教育匹配与收入回报", 《社会》, 2016 年第 3 期, 第 64—85 页。

- [26] 李磊、王小霞、包群,“机器人的就业效应:机制与中国经验”,《管理世界》,2021年第9期,第104—119页。
- [27] 李晓光,“教育失配经历对企业雇佣决策的影响:来自简历投递实验和在线招聘职位的证据”,《管理世界》,2024年第4期,第121—145页。
- [28] 李晓光、陆瑶、吴晓刚,“高学历劳动者的教育失配”,《教育研究》,2023年第6期,第122—137页。
- [29] Manuel, N., “The Use of Major-Related Knowledge by Early Career College Graduates”, *Applied Economics*, 2024, 56(55), 7302-7316.
- [30] Manuel, N., and M. Plesca, “Skill Transferability and the Earnings of Immigrants”, *Canadian Journal of Economics*, 2020, 53(4), 1404-1428.
- [31] McGuinness, S., “Overeducation in the Labour Market”, *Journal of Economic Surveys*, 2006, 20(3), 387-418.
- [32] OECD, *Getting Skills Right: Spain*. Paris: OECD Publishing, 2017.
- [33] Robst, J., “Education and Job Match: The Relatedness of College Major and Work”, *Economics of Education Review*, 2007, 26(4), 397-407.
- [34] Somers, M. A., S. J. Cabus, W. Groot, and H. M. van den Brink, “Horizontal Mismatch between Employment and Field of Education: Evidence from a Systematic Literature Review”, *Journal of Economic Surveys*, 2019, 33(2), 567-603.
- [35] Tacchella, A., M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero, “A New Metrics for Countries' Fitness and Products' Complexity”, *Scientific Reports*, 2012, 2(1), 723.
- [36] Tsai, Y., “Returns to Overeducation: A Longitudinal Analysis of the US Labor Market”, *Economics of Education Review*, 2010, 29(4), 606-617.
- [37] Verdugo, R. R., and N. T. Verdugo, “The Impact of Surplus Schooling on Earnings: Some Additional Findings”, *Journal of Human Resources*, 1989, 24(4), 629-643.
- [38] Verhaest, D., and E. Omey, “The Impact of Overeducation and Its Measurement”, *Social Indicators Research*, 2006, 77(3), 419-448.
- [39] 王林辉、钱圆圆、宋冬林、董直庆,“机器人应用的岗位转换效应及就业敏感性群体特征——来自微观个体层面的经验证据”,《经济研究》,2023年第7期,第69—85页。
- [40] 王永钦、董雯,“机器人的兴起如何影响中国劳动力市场?——来自制造业上市公司的证据”,《经济研究》,2020年第10期,第159—175页。
- [41] Webb, M., “The Impact of Artificial Intelligence on the Labor Market”, SSRN Working Paper, 2020.
- [42] Wolbers, M. H., “Job Mismatches and Their Labour-Market Effects among School-leavers in Europe”, *European Sociological Review*, 2003, 19(3), 249-266.
- [43] 谢尚、韦东明、唐琦,“互联网使用对教育-工作匹配的影响——基于CFPS2016~2020的研究发现”,《中国人口科学》,2024年第4期,第36—51页。
- [44] 闫雪凌、朱博楷、马超,“工业机器人使用与制造业就业:来自中国的证据”,《统计研究》,2020年第1期,第74—87页。
- [45] 张丹丹、于航、李力行、胡佳胤、莫怡青、李泓宇,“中国人工智能技术暴露度的测算及其对劳动需求的影响——基于大语言模型的新证据”,《管理世界》,2025年第7期,第59—75页。

Artificial Intelligence and the Alleviation of Labor Market Mismatch: Evidence from Online Job Matching Data

YU Hang LI Hongbo ZHANG Dandan* ZHANG Runbo

(Peking University)

LI Qiang

(Zhaopin Limited)

Abstract: Using 1.62 million matched vacancy-applicant observations from an online recruitment platform, we refine an occupation-level AI large language model (AI-LLM) exposure index, construct measures of vertical and horizontal mismatch, and apply difference-in-difference and event-study designs to identify the impact of AI-LLM technologies on labor market mismatch. We show that overall mismatch increased after 2021. But following the technological shock associated with ChatGPT, vertical mismatch declined in occupations with higher AI-LLM exposure, accompanied by more precise vacancy signals, higher skill requirements, and educational thresholds. These results provide vacancy-applicant evidence on how AI improves labor market matching efficiency.

Keywords: technological progress; labor market mismatch; AI large language model exposure index

JEL Classification: J24, J62, O33

* Corresponding Author: ZHANG Dandan, National School of Development and China Center for Economic Research, Peking University, Haidian District, Beijing 100871, China; Tel: 86-10-62759779; E-mail: ddzhang@nsd.pku.edu.cn.