

数字经济下互联网使用能否缓解过度教育的收入惩罚？*

李玉茹 张杉杉

【摘 要】从互联网使用视角探究缓解过度教育的收入惩罚问题,是数字经济时代提高劳动力市场效率的新思路。文章使用 2014~2018 年中国家庭追踪调查数据,探讨互联网使用对劳动者的过度教育风险及其收入惩罚的缓解效应,并分析了其作用机制。研究发现,互联网使用有助于降低过度教育的发生风险(即“数量效应”),对过度教育者而言,互联网使用有助于增加其收入(即“收入效应”),由此缓解劳动者的过度教育风险及其收入惩罚。细分研究显示,使用互联网进行工作、学习和社交活动均有助于降低过度教育风险,并缓解过度教育者的收入惩罚。异质性分析表明,互联网使用显著降低了非垄断行业劳动者的过度教育风险,对女性和农村户口过度教育劳动者的收入惩罚缓解效应更大。据此,文章建议充分利用数字经济发展契机,通过互联网赋能提升劳动者数字技能,促进更高质量的充分就业。

【关键词】互联网使用 过度教育 收入惩罚 缓解作用

【作 者】李玉茹 首都经济贸易大学劳动经济学院,博士研究生;张杉杉 首都经济贸易大学劳动经济学院,教授。

一、引 言

2020 年 10 月 29 日,《中共中央关于制定国民经济和社会发展第十四个五年规划和二〇三五年远景目标的建议》提出“实现更加充分更高质量就业”的目标,并强调“更加注重缓解结构性就业矛盾”。过度教育是就业结构性矛盾的一种表现形式,主要指劳动者的教育水平高于工作所需的教育水平的状态。过度教育意味着劳动者所在岗位的知识能力要求低于劳动者自身水平,这可能会限制高学历劳动者生产力的释放,从而给劳动者带来收入损失,产生“收入惩罚”效应(傅琼,2023),降低个体的就业质量。可见,缓

* 本文为首都经济贸易大学 2022 年校级研究生教改项目“研究生培养质量评价新范式的研究——基于过度教育的视角”(编号:00292254301003)的阶段性研究成果。

解中国劳动力市场结构性矛盾、提高劳动者就业质量,关键之一在于提升个体教育水平与职业岗位要求之间的匹配度(Wang, 2021; 汪鲸、罗楚亮, 2023)。

学界对过度教育形成的影响因素进行了广泛讨论,部分学者从劳动者的性别、年龄、学历、户口类型等人口学特征(方超、黄斌, 2018)进行探讨,也有学者从工作特征视角切入,研究岗位需求、行业类型、企业所有制等因素的影响(周敏丹, 2021; 李勇、段诗宁, 2021)。随着中国的互联网普及率的快速提升^①,众多研究者也将互联网使用纳入过度教育影响因素的分析框架。已有研究发现,互联网使用极大提高了劳动者的信息搜寻能力,提升个人工作效率并带来人力资本的快速积累,通过信息改善效应显著减少教育错配(刘诗洋等, 2022; 丁述磊、刘翠花, 2022)。然而,既有研究普遍倾向于从互联网接入这一较为宏观的视角出发,往往忽略了互联网使用的细节,包括劳动者使用互联网的频率差异及不同活动场景下的具体作用,从而未能充分探讨互联网使用对过度教育收入惩罚的影响机制。

鉴于此,本文利用中国家庭追踪调查(CFPS)数据,从互联网使用的微观视角切入,研究互联网使用对缓解过度教育收入惩罚效应的影响。本文的边际贡献如下。第一,从工作、学习、社交和商业活动等多个视角构建互联网使用变量,且细化分析使用互联网进行具体活动对劳动者过度教育及其收入惩罚的不同影响,为探索缓解当前普遍存在过度教育问题提供了新的经验证据和启示。第二,深入剖析互联网在降低过度教育风险和缓解收入惩罚方面所发挥的双重作用机制,即“数量效应”和“收入效应”。研究发现为互联网使用缓解收入惩罚效应的作用机制提供了微观证据,丰富了互联网劳动力市场效应领域的文献。第三,从性别、户口、行业等多角度考察互联网使用对过度教育及其收入惩罚的异质性影响,为政府部门制定促进高质量充分就业的政策提供理论支撑。

二、文献综述与研究假设

(一) 文献综述

1. 过度教育的产生及其收入效应

现有研究对于过度教育的关注集中在两点。其一是过度教育的内在发生机制研究。基于劳动力市场完全竞争假设的理论视角,人力资本理论认为在完全竞争市场上,劳动力供过于求导致的教育错配是短暂现象,长期来看,过度教育状况有望通过工作转换、人员流动以及工资调整等实现平衡状态。随后,信息经济学派认识到信息完备假设的局限性,转而从信息不对称视角剖析过度教育问题。有学者提出,在城乡、区域、行业、技术、岗位等多重分割下,劳动力市场上供求双方的信息不对称使得劳动者难以实现就业

^① 资料来源:中国政府网,《关于数字经济发展情况的报告》, https://www.gov.cn/xinwen/2022-11/28/content_5729249.htm?eqid=b34cc16b000456ec00000006645f7ff7。

流动和匹配(盛斌、郝碧榕,2021),学历成为劳动者求职的重要信号。为了提升自身就业竞争力、获取优质工作机会,劳动者会主动增加额外的教育投资。在就业竞争激烈的情况下,高学历劳动者也会降低求职预期,开始向较低学历要求的工作流动(沈煜等,2023),导致过度教育愈加严重。其二是过度教育的收入惩罚效应研究。工作竞争理论认为,工资水平的主要决定因素是工作而不是个人,超过工作所需的教育投资几乎无利可图(Wu等,2018)。当过度教育发生时,劳动者在没有获得相对优势的工作岗位上,其技能无法得到充分发挥,这会阻碍生产效率的提升,进而产生收入惩罚(李根丽、尤亮,2020)。此外,一些学者进一步探讨了影响过度教育收入惩罚效应的因素,如周敏丹(2021)的研究显示,年龄的增长和工作经验的积累会加剧劳动者过度教育状态对收入的负面影响,而认知能力的提升能有效缓解这种影响;蒋文莉和赵昕(2022)指出,过度教育会对农民工的收入产生惩罚效应,但这种效应会随着单位规模的扩大而减轻。

2. 互联网使用对个体劳动力市场表现的影响

根据已有研究文献,互联网使用主要从两个方面影响个体劳动者的劳动力市场表现。一方面,互联网极大地提升了劳动者获取高质量就业信息的能力(王军、韩悦,2023)。它不仅为劳动者创造了多样化的灵活就业与创业机会,还拓宽了职业搜寻渠道,使得劳动者能够更好地匹配个人技能与市场需求,缩短失业周期,增强就业稳定性(戚聿东等,2021)。另一方面,互联网有助于劳动者提升人力资本和社会资本(宁光杰、杨馥萍,2021),提高工作或生产效率,进而提高工资水平和工作满意度(周烁、张文韬,2021)。

整体来看,现有文献关于过度教育的形成机理及互联网使用对就业或收入影响的研究发现为本研究奠定了重要基础,但关于互联网使用与过度教育及其收入惩罚之间关系的探讨尚显不足,不仅缺乏互联网使用如何缓解劳动者过度教育收入惩罚的理论机制框架建构,也欠缺对互联网使用的具体活动场景的影响后果分析与实证检验。鉴于此,本文从互联网使用的具体活动行为视角出发,系统研究其对过度教育及其收入惩罚的影响与作用机制,并进一步分析这些影响在不同互联网使用活动中和个体特征上的异质性。

(二) 研究假设

理论上,互联网使用主要通过双重作用机制缓解劳动者过度教育的收入惩罚,如图所示。首先,互联网使用通过降低过度教育风险来减轻收入惩罚。这本质上强调了过度教育的数量效应。换句话说,过度教育可能导致收入惩罚,而互联网使用降低了过度教育

的程度,从而减少了劳动者面临的收入惩罚。其次,互联网使用通过增加收入来缓解过度教育者的收入惩罚。这主要强调了过度教育的收入回报,即已经发生过度教育的劳动者可以通过

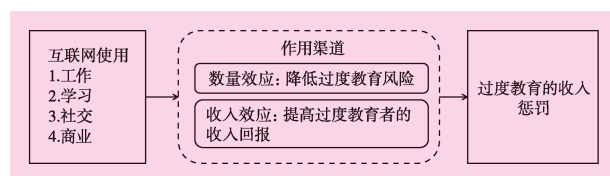


图 互联网使用缓解过度教育的收入惩罚机制

互联网使用提升或稳定工资水平,从而在一定程度上缓解收入惩罚效应。

1. 互联网使用对过度教育的影响

工作搜寻匹配理论指出,劳动力市场信息摩擦是造成劳动者就业匹配失衡的重要因素。互联网使用可以拓宽信息渠道,有效缓解信息不对称现象,扩大劳动者在社会网络的人际交往,提供更多在线学习机会,减少劳动者过度教育的发生(赵放、蒋国梁,2023)。随着互联网使用下过度教育情况的改善,工资因过度教育受损的程度降低,收入惩罚效应得到缓解。

在具体活动场景中,互联网使用主要从以下几个方面降低过度教育风险:(1)互联网技术的发展促成了点对点的新工作模式,这种工作模式使工作地点和工作时间不再成为劳动者就业的限制因素,有助于提升劳动者的工作满意度(Bloom等,2015),从主观上减少过度教育的发生风险;(2)近年来,许多学习类网站及平台得到了广泛应用,它们可以减少传统教育的时间和空间限制,使得个体学习更为灵活。劳动者可以借此以较低成本提升个人的就业能力、积累职业技能,增强在劳动市场中的竞争力(赵建国、周德水,2019),进而获得与自身匹配和满意的职位;(3)网络社交平台提供的即时沟通功能,一方面促使劳动者维持与亲朋好友等原有亲密关系的强联系,另一方面也为劳动者结交陌生人、发展弱联系提供了更多的可能性。强弱联系均有助于劳动者获得与教育水平更为匹配的工作(Lu等,2016);(4)互联网商业使用有助于提高求职者的职业搜寻效率,增加劳动者找到与自身受教育水平相匹配工作的概率。例如,即时支付减少劳动者在交通和通讯上的精力消耗,互联网购物节省用于家庭琐事的时间(魏浩然等,2023)。由此,本文提出假设1:互联网使用有助于降低过度教育风险;假设1a:使用互联网进行工作、学习、社交和商业活动有助于降低过度教育风险。

2. 互联网使用对过度教育者收入惩罚效应的影响

在数字经济时代,互联网使用是劳动者提高生产率和收入的有效途径,劳动者可以通过互联网使用有效缓解过度教育的收入惩罚。从能力角度来看,过度教育的劳动者相比于教育匹配者更有能力有效利用互联网资源,获得更高的收入,从而减少过度教育的收入惩罚。从动机角度讲,过度教育可能会降低工作满意度(尤亮、李根丽,2023),过度教育的劳动者往往会不满足于现状,更有动力通过流动或者利用外部资源如互联网使用等方式改善就业环境,获取更高收入,以弥补过度教育带来的收入惩罚。

细化到具体的活动场景中,过度教育的劳动者主要从以下几个方面通过互联网使用缓解收入惩罚:(1)互联网提高了信息获取的便捷性和内外部沟通效率,使劳动者能够通过检索联系信息、查找参考资料等方式提高工作效率,最终获得更高的工资收入(王军、韩悦,2023);(2)互联网提供了更多的学习机会,打破了地域限制,使劳动者能够通过在线直播、录制视频等形式获得职业技能培训(王少国、匡国静,2022),这可以弥补其

技能经验的不足,提高其职业适应力,摆脱收入惩罚的困境;(3)互联网社交使用可以扩大个人的社交网络,有助于过度教育者与客户快速建立联系、与企业达成合作以提高工作绩效,同时也可以通过线上联系获得亲友等主要社交群体的支持,得到优质工作机会和丰厚的收入回报(Meliciani等,2016);(4)网上购物、订餐及网银支付等互联网商业活动也可以帮助过度教育的劳动者减轻生活琐事负担(Dettling,2017),专注于工作,从而提高工资收入。由此,本文提出假设2:过度教育者互联网使用强度越高,得到的收入回报越多,因过度教育面临的收入惩罚越小;假设2a:过度教育者在工作、学习、社交和商业活动中使用互联网程度越强,其面临的收入惩罚越小。

三、实证模型

(一)数据来源与变量选择

本文使用的数据来自中国家庭追踪调查(China Family Panel Studies,简称CFPS)。该调查采用内隐分层的、多阶段、与人口规模成比例的概率抽样方式,调查对象涵盖25个省(市、自治区)的家庭户及家庭成员,具有良好的代表性,调查内容包括被访者的基本人口特征、教育、就业等详细信息。在CFPS目前已公布的调查结果中,2010、2012、2020年的调查问卷中收集的互联网使用信息有限,因此本文主要使用2014~2018年的3期数据。根据研究对象的特征,本文将分析样本限定为劳动年龄人口,保留16~60岁的男性和16~55岁的女性样本中有工资收入的被访者,在剔除了离退休人员、在校大学生及教育或职业等核心变量信息缺失的样本后,2014年剩余4569个样本、2016年剩余3236个样本、2018年剩余6297个样本。在进一步剔除工资收入、健康状况、行业类型、户口类型等变量缺失的样本后,保留三年全部追踪个体,最终形成由1639位被访者的4917人次有效观测记录构成的平衡面板数据。

本文的被解释变量是过度教育和工资收入。过度教育在一定程度上反映了劳动力供需匹配状况。本文通过比较受访者的实际受教育年限与胜任工作所需受教育年限来衡量过度教育,若实际受教育年限高于胜任工作所需受教育年限,视为过度教育,其差值为过度教育年限。考虑到胜任工作所需的受教育年限难以直接衡量,本文采用实际匹配法进行测算,以众数法为主要测量方法,以标准差法为辅助测量方法^①。众数法的测算步骤如下:根据国家统计局《职业分类标准》,将受访者所从事的职业划分为七大职业类型(不包括军人);再依据省份一职业分组,求出某省份某一职业中频率最高的教育年限作为工作所需教育年限;最后将受访者实际受教育年限与工作所需年限相减,超过的年

① 标准差法的步骤:首先根据省份职业利用平均值法得到某省份某一职业要求的平均教育年限,在此均值上加减标准差得到工作所需的教育年限区间;当个人受教育年限位于区间右侧,则为过度教育,得到的值为过度教育年限。

限即为过度教育年限。调查样本中部分受访者的实际受教育年限低于胜任工作所需受教育年限,该情况被视为教育不足,不在本研究的关注范畴内,因此本文在数据分析时剔除了此类样本。工资收入是教育—职业匹配结果的直接体现,能够有效反映过度教育的收入惩罚效应。考虑到数据的可得性,本文参考高志鹏等(2023)的研究,将收入定义为过去一年内所有工资、奖金、现金福利和实物补贴的总和,扣除税和五险一金后的净收入,分析时取其自然对数。

本文的核心解释变量为互联网使用。基于职业发展需求和终身学习理念,劳动者在使用数字技术时存在主动性和目的偏向性(Bonfadelli, 2002),其活动范围扩展到工作、学习和社交等方面,甚至包括金融领域(许冰、胡俊, 2022)。因此,本文将互联网使用范围界定为包括工作、学习、社交和商业(侧重于使用网银和网上购物等)在内的多种活动。本文通过被访者“使用互联网工作、学习、社交和商业的频率”来衡量互联网使用,并将其从“从不”到“几乎每天”进行1~7的数值赋值,值越大表明个体互联网使用的程度越高。为了反映综合的互联网使用状况,本文采用变异系数法进行客观权重的赋值测算。具体步骤为:(1)将工作、学习、社交和商业活动中的互联网使用指标进行标准化处理;(2)计算每个指标的变异系数 $V_m = \sigma_m / \mu_m$, 其中 σ_m 为第 m 项指标的标准差, μ_m 为第 m 项指标的均值;(3)计算各个指标的权重值 $W_m = V_m / \sum_{m=1}^4 V_m$; (4)将标准化处理的4个指标数值按线性加权法加总合成互联网使用的得分值 $Digital = \sum_{m=1}^4 (W_m X_m)$ 。

本文参考已有文献的研究发现,在数据可得的基础上,尽可能地控制了相关影响变量,包括年龄、受教育年限、性别、户口类型、婚姻状况、健康状况等个体特征变量,以及行业类型等工作特征变量。性别变量中,男性赋值为1,女性赋值为0。年龄为连续变量。受教育年限根据学历水平进行赋值:文盲/半文盲为0年、小学程度为6年、初中程度为9年、高中程度为12年、中专程度为13年、大专程度为15年、大学本科为16年、研究生程度为19年。户口类型中,农村户口赋值为1,城镇户口赋值为0。婚姻状况中,已婚赋值为1,其他情况赋值为0。健康状况中,非常健康、很健康、比较健康赋值为1,一般、不健康赋值为0。对于行业类型,本文参照岳希明和蔡萌(2015)的做法,将19个行业划分为垄断行业和非垄断行业两类。此外,本文还加入了时间固定效应和省份固定效应,以减少对省份—职业中的过度教育及收入影响效应的估计偏误。主要变量的描述性统计结果如表1所示。具体来看,分析样本中被访者年总收入均值为27 553元,取对数后为8.375,过度教育年限平均为1.249年,互联网使用标准化指标均值为0.328。在个体和工作特征方面,男性占比(63.4%)高于女性,被访者的平均年龄为35.419岁,平均受教育年限为10.265年,已婚比例为79.3%,82.9%的被访者自评健康状况良好,农村户口的被访者占70.6%,81.8%的被访者在非垄断行业工作。

表 1 样本的描述性统计(N=4917)

变 量	变量说明及赋值	均值	最小值	最大值
工资收入	年收入,取对数	8.375	0	13.816
过度教育	实际受教育年限-工作所需受教育年限(年)	1.249	0	10
互联网使用	变异系数法计算的互联网使用综合得分值	0.328	0	1
年龄	年龄(岁)	35.419	16	60
性别	男性=1;女性=0	0.634	0	1
受教育年限	学历水平转化为受教育年限(年)	10.265	0	19
婚姻状况	已婚=1;未婚=0	0.793	0	1
健康状况	健康=1;不健康和一般=0	0.829	0	1
户口类型	农村户口=1;城镇户口=0	0.706	0	1
行业类型	垄断行业=1;非垄断行业=0	0.182	0	1

(二) 模型选择

1. 互联网使用对过度教育的影响模型

本文的被解释变量“过度教育”为连续变量,且存在过度教育年限零值数据即实际受教育年限等于胜任工作所需受教育年限的情况(约占总样本的 56%),该被解释变量的分布断尾会使概率密度随之发生变化。在这种情况下,如果直接使用普通最小二乘法(OLS)对整个样本进行线性回归,得到的结果是有偏的,因而本文选择使用面板 Tobit 模型。考虑到样本量大但时期数较少时,Tobit 面板模型的固定效应估计不一致,并且还会导致其他解释变量估计的非一致性,因此,本文将模型设定为随机效应模型。计量模型如下:

$$Overedu_{ijt}^* = \alpha_0 + \alpha_1 Digital_{ijt} + \alpha_2 CV_{ijt} + \theta_j + \gamma_t + \varepsilon_{ijt} \tag{1}$$

$$Overedu_{ijt} = \begin{cases} Overedu_{ijt}^*, Overedu_{ijt}^* > 0 \\ 0, Overedu_{ijt}^* \leq 0 \end{cases} \tag{2}$$

其中,被解释变量 $Overedu_{ijt}$ 表示过度教育, $Overedu_{ijt}^*$ 为不可观测的潜变量,核心解释变量 $Digital_{ijt}$ 表示互联网使用,控制变量 CV_{ijt} 包括年龄、受教育年限、性别、户口、婚姻、健康状况等个体特征变量以及行业类型等工作特征变量。待估系数 α_1 反映互联网使用对劳动者过度教育的效应, α_2 为控制变量的待估计系数, θ_j 为省份固定效应, γ_t 为时间固定效应, ε_{ijt} 为随机扰动项。

2. 互联网使用对过度教育者收入惩罚效应的影响模型

基于明瑟收入方程,本文重点考察过度教育的收入惩罚效应以及互联网使用对过度教育者收入惩罚的影响,以互联网使用、过度教育及其交互项作为解释变量,纳入模型进行分析。考虑到工资收入存在零值数据,属于受限被解释变量,本文采用随机效应的面板 Tobit 模型,该模型如下:

$$Lnwage_{ijt}^* = \beta_0 + \beta_1 Overedu_{ijt} + \beta_2 Digital_{ijt} + \beta_3 (Overedu_{ijt} \times Digital_{ijt}) + \beta_4 CV_{ijt} + \theta_j + \gamma_t + \varepsilon_{ijt} \tag{3}$$

$$Lnwage_{ijt} = \begin{cases} Lnwage_{ijt}^*, & Lnwage_{ijt}^* > 0 \\ 0, & Lnwage_{ijt}^* \leq 0 \end{cases} \quad (4)$$

其中, $Lnwage_{ijt}$ 表示劳动者工资收入的自然对数, $Lnwage_{ijt}^*$ 为不可观测的潜变量, $Overedu_{ijt}$ 表示过度教育, $Digital_{ijt}$ 表示互联网使用, $Overedu_{ijt} \times Digital_{ijt}$ 为过度教育与互联网使用的交互项, CV_{ijt} 为控制变量, 包括年龄、受教育年限、性别、户口、婚姻、健康状况等个体特征变量以及行业类型等工作特征变量, θ_i 为省份固定效应, γ_t 为时间固定效应, ε_{ijt} 为随机扰动项。此外, 为了避免交互项与主效应之间可能存在的严重共线性问题, 本文参考已有文献的研究方法, 对过度教育和互联网使用进行中心化处理。

四、实证结果分析

(一) 互联网使用对过度教育及过度教育者收入惩罚的影响

根据前述的分析策略, 本文运用 Tobit 模型估计互联网使用对过度教育及过度教育者收入惩罚的影响, 结果如表 2 所示。本文首先分析了互联网使用对过度教育的影响, 模型 1 控制了个体和工作特征变量, 模型 2 在此基础上加入了省份和年份固定效应。模型结果均显示, 互联网使用对劳动者的过度教育具有显著的负向影响, 这表明劳动者互联网使用程度越高, 越有助于降低过度教育风险。假设 1 得到验证。

接下来, 本文探讨了互联网使用对过度教育者收入惩罚的影响。模型 3 未加入过度教育与互联网使用的交互项, 回归结果显示互联网使用的系数显著为正, 说明互联网使用为劳动者带来更多的收入回报。同时, 过度教育对劳动者的收入产生了显著的负面影响, 其系数值为 -0.093, 即过度教育带来了收入惩罚效应, 这与既有研究的结论一致。

表 2 互联网使用对过度教育及其收入惩罚的影响(N=4917)

变 量	过度教育		工资收入	
	模型 1	模型 2	模型 3	模型 4
互联网使用	-1.011*** (0.212)	-0.858*** (0.217)	0.519** (0.245)	0.274 (0.285)
过度教育			-0.093** (0.041)	-0.180*** (0.062)
过度教育 × 互联网使用				0.168* (0.090)
年龄	-0.016 (0.010)	-0.009 (0.009)	-0.010 (0.008)	-0.011 (0.008)
性别	0.438** (0.180)	0.423** (0.162)	0.636*** (0.133)	0.643*** (0.133)
受教育年限	1.349*** (0.038)	1.401*** (0.036)	0.141*** (0.032)	0.152*** (0.033)
婚姻状况	0.194 (0.189)	0.261 (0.177)	0.022 (0.180)	0.033 (0.179)
健康状况	-0.141 (0.171)	-0.053 (0.162)	0.651*** (0.163)	0.659*** (0.163)
户口类型	0.408** (0.168)	0.231 (0.158)	-0.681*** (0.151)	-0.679*** (0.151)
行业类型	-2.097*** (0.163)	-2.227*** (0.156)	0.319* (0.168)	0.315* (0.168)
常数项	-15.436*** (0.659)	-22.765*** (1.086)	5.239*** (0.848)	5.275*** (0.847)

注: 模型 1 未控制省份固定效应、年份固定效应, 模型 2 至模型 4 均控制了省份固定效应、年份固定效应。括号中数值为个体层面聚类稳健标准误。*、** 和 *** 分别表示在 10%、5% 和 1% 的水平上显著。下表同。

模型 4 在模型 3 的基础上加入了过度教育与互联网使用的交互项,结果显示交互项对劳动者的工资收入起到了显著正向作用,说明互联网使用有助于缓解过度教育产生的收入惩罚,假设 2 得到验证。

(二) 稳健性检验

为了检验前述结论的稳健性,本文从以下 4 个方面进行检验。(1)为排除变量和模型设定方法对结果的影响,本文进行混合 OLS 回归估计。(2)考虑到过度教育的测量误差问题,本文采用“标准差法测算的过度教育”代替“众数法测算的过度教育”作为解释变量,对模型进行重新估计。(3)为进一步考察互联网使用指标衡量方法的稳健性,本文借鉴夏璋煦和丁守海(2023)的研究,利用因子分析将互联网使用的原始测量指标降维得到互联网使用的综合指标,并使用新生成的解释变量对模型进行重新估计。(4)为了避免因个体编码匹配形成追踪面板数据带来的样本缺失,本文扩大样本量,利用匹配前数据重新进行回归。上述稳健性检验结果呈现在表 3 中。检验结果表明,互联网使用降低过度教育风险并缓解过度教育者的收入惩罚的结论稳健。

表 3 更换模型、变量和样本的稳健性检验结果

变 量	混合 OLS 回归		标准差法		因子分析降维		扩大样本量	
	模型 5	模型 6	模型 7	模型 8	模型 9	模型 10	模型 11	模型 12
	过度教育	工资收入	过度教育	工资收入	过度教育	工资收入	过度教育	工资收入
互联网使用	-0.437*** (0.105)	0.233 (0.241)	-0.803*** (0.103)	0.271 (0.287)	-0.240*** (0.072)	0.069 (0.092)	-0.907*** (0.189)	-0.039 (0.249)
过度教育		-0.167*** (0.056)		-0.176** (0.084)		-0.125*** (0.044)		-0.364*** (0.039)
过度教育 × 互联网使用		0.170** (0.074)		0.213* (0.122)		0.052* (0.029)		0.287** (0.063)
观测值	4917	4917	4917	4917	4917	4917	9700	9700

(三) 工具变量回归

过度教育的劳动者更可能借助互联网来获取更匹配的工作,这会导致式(1)存在内生性问题。为克服该估计偏误,本文参考已有文献(蒋文莉、赵昕,2022),选取“社区或村居互联网使用率”作为互联网使用的工具变量。选取该工具变量的依据如下:第一,社区或者村居层面的互联网普及率高意味该区域较低的互联网使用成本,加之地域范围内的同伴效应,会促使劳动者倾向于选择使用互联网,满足工具变量所需的相关性要求;第二,社会或者村居层面的互联网使用率较难通过其他途径影响劳动者的就业表现,满足工具变量所需的外生性要求。

表 4 展示了工具变量的回归结果。第一阶段的回归以社区或村居互联网使用率作为解释变量,以个体互联网使用作为被解释变量,得到的 F 统计值大于 10,表明不存在

弱工具变量问题。第二阶段使用工具变量法得到的互联网使用系数为负且在1%的水平上显著。综上所述,解决内生性后,互联网使用对过度教育的影响显著为负的结论成立。

表 4 互联网使用影响过度教育的内生性检验(工具变量法)(N=4917)

变 量	第一阶段 互联网使用	第二阶段 过度教育
互联网使用		-0.725**(0.295)
社区或村居互联网使用率	0.360*** (0.015)	
控制变量	是	是
第一阶段的 F 值	572.144	
DWH 检验		0.279

针对互联网使用与过度教育和工资收入存在的较强内生性问题,本文进一步尝试寻找过度教育变量的工具变量,对内生性问题进行处理。借鉴颜敏和王维国(2018)的研究方法,本文选取“标准差法的过度教育”作为众数法过度教育的工具变量。理由如下:假设 $overedu$ 是真实值 $overedu^*$ 的一种测量值,则 $overedu=overedu^*+e_1$,其中, e_1 为测量误差。 $overedu^2$ 是真实值 $overedu^*$ 的第二种测量值,则 $overedu^2=overedu^*+e_2$,其中, e_2 为测量误差。根据对数工资公式 $lnwage=\gamma_0+\gamma_1 overedu^*+\gamma_2 cv+\varepsilon$,其中 $lnwage$ 表示工资收入的对数形式, cv 是其他可观测控制变量, ε 是随机扰动项。将 $overedu$ 代入公式得: $lnwage=\gamma_0+\gamma_1 overedu+\gamma_2 cv+(\varepsilon-\gamma_1 e_1)$ 。一方面, $overedu^2$ 和 $overedu$ 是对同一种概念的两种测度方式,显然相关,满足工具变量所要求的相关性条件;另一方面,假定不同方法测量误差不相关,测量误差与观测不到的变量不相关,则 e_1 与 $overedu^*$ 、 e_2 、 ε 不相关,由 $overedu^2=overedu^*+e_2$ 可知 $overedu^2$ 与 e_1 不相关,满足工具变量所需的外生性条件。

以社区或村居互联网使用率作为互联网使用的工具变量,将标准差法的过度教育与其生成交互项纳入模型进行检验,结果如表 5 所示。第一阶段不可识别检验的统计量为 216.840,在 1%的水平上拒绝不可识别的原假设;弱工具变量检验的 F 统计量均超过了经验标准值 10,表明工具变量和内生变量具有较强的相关性,不存在弱工具变量问题。第二阶段的结果显示,在工具变量法回归中,过度教育与互联网使用的交互项对工资收入的影响仍然在 1%的水平上显著为正,即互联网使用有效缓解了过度教育带来的收入惩罚。

(四) 不同活动场景下的互联网使用对过度教育收入惩罚的影响

考虑到劳动者通常会利用互联网进行工作、学习、社交和商业等不同活动,这些活动产生的影响可能也不尽相同,本文进一步探讨在不同活动场景下的互联网使用对劳动者过度教育和收入惩罚效应的影响,并采用前文的工具变量,使用两阶段最小二乘法(2SLS)解决不同互联网使用、过度教育与收入惩罚之间存在的内生性问题。

表 6 结果显示,使用互联网进行工作和学习活动均对过度教育产生了显著的负向影响,而使用互联网进行社交和商业活动的结果并不显著。在考虑了内生性问题后,互联网社交活动对过度教育的影响在 10%的水平上显著,这说明使用互联网进行工作、学习

表 5 互联网使用影响过度教育者工资效应的内生性检验(N=4917)

变 量	第一阶段			第二阶段
	互联网使用	过度教育	过度教育 × 互联网使用	工资收入
互联网使用				-0.290*** (0.106)
过度教育				-0.036 (0.677)
过度教育 × 互联网使用				0.447** (0.182)
社区或村居互联网使用率	0.323*** (0.016)	0.055 (0.062)	-0.133** (0.058)	
标准差法过度教育	-0.046*** (0.007)	1.098*** (0.029)	-0.040 (0.027)	
过度教育 II × 社区 / 村居互联网使用率	0.033*** (0.009)	0.003 (0.035)	0.775*** (0.032)	
控制变量	是	是	是	是
第一阶段的 F 值		216.840		
不可识别检验		578.878***		

和社交活动均有助于降低过度教育风险。假设 1a 得到部分验证。在使用工具变量法处理了内生性问题后,使用互联网进行工作、学习、社交和商业活动均对过度教育者的收入惩罚产生了显著的正向影响(见表 7),说明不同活动场景下的互联网使用都有助于缓解过度教育者的收入惩罚效应。假设 2a 得到验证。产生这一结果的可能原因如下:使用互联网进行工作、学习和社交活动均有助于提升劳动者的人力资本和社交资本,对劳动者获取就业信息和提高就业能力进而缓解过度教育和提升收入有着重要意义。而劳动者使用互联网进行的商业活动,如网银支付和网上购物等,主要属于消费体验。尽管这些活动可以减少劳动者在家务劳动等活动上的时间和精力耗费,从而增加其工作时间并带来更高收入,但这种影响对于直接改善就业环境的效果可能较为有限。

(五) 异质性分析

如前文所述,中国劳动力市场深受城乡、行业等多重细化分割的影响。此外,人口属

表 6 不同活动场景下的互联网使用影响过度教育的估计结果

变 量	模型 13	模型 14	模型 15	模型 16	模型 17	模型 18	模型 19	模型 20
	Tobit 回归	2SLS 回归	Tobit 回归	2SLS 回归	Tobit 回归	2SLS 回归	Tobit 回归	2SLS 回归
互联网使用	-0.161*** (0.024)	-0.135** (0.055)	-0.069*** (0.025)	-0.125** (0.051)	0.009 (0.025)	-0.071* (0.029)	-0.006 (0.033)	-0.174 (0.172)

注:模型 13 和模型 14 自变量为互联网工作;模型 15 和模型 16 自变量为互联网学习;模型 17 和模型 18 自变量为互联网社交;模型 19 和模型 20 自变量为互联网商业。

表 7 不同活动场景下的互联网使用影响过度教育者收入惩罚的估计结果

变 量	模型 21	模型 22	模型 23	模型 24	模型 25	模型 26	模型 27	模型 28
	Tobit 回归	2SLS 回归	Tobit 回归	2SLS 回归	Tobit 回归	2SLS 回归	Tobit 回归	2SLS 回归
互联网使用	0.063*	0.018	0.001	-0.014	0.027	-0.002	-0.051	0.037
	(0.034)	(0.136)	(0.035)	(0.117)	(0.029)	(0.061)	(0.044)	(0.168)
过度教育 ×	0.020*	0.069**	0.021*	0.085**	0.004	0.066***	0.023	0.092**
互联网使用	(0.011)	(0.031)	(0.012)	(0.033)	(0.012)	(0.023)	(0.014)	(0.037)

注：模型 21 和模型 22 自变量为互联网工作；模型 23 和模型 24 自变量为互联网学习；模型 25 和模型 26 自变量为互联网社交；模型 27 和模型 28 自变量为互联网商业。

性的多元化和区域发展的差异化均可能在不同程度上影响个人的互联网使用及其效应。为全面理解劳动者互联网使用及其认知回报率的人群差异,本文从性别、户口类型、行业类型等多角度进行异质性分析,结果如表 8 和表 9 所示。

表 8 性别的异质性分析

变 量	过度教育		工资收入	
	男性	女性	男性	女性
互联网使用	-0.859*** (0.268)	-0.835*** (0.364)	-0.340 (0.357)	-0.202 (0.467)
过度教育 × 互联网使用			0.121 (0.113)	0.319** (0.150)
观测值	3117	1800	3117	1800

表 9 户口类型、行业类型的异质性分析

变 量	过度教育		工资收入		过度教育		工资收入	
	农村户口	城市户口	农村户口	城市户口	垄断行业	非垄断行业	垄断行业	非垄断行业
互联网使用	-0.832***	-0.919***	-0.365	0.037	-0.783	-0.817***	0.171	0.284
	(0.283)	(0.340)	(0.363)	(0.168)	(0.625)	(0.214)	(0.468)	(0.338)
过度教育 ×			0.332***	-0.013			0.051	0.181*
互联网使用			(0.126)	(0.123)			(0.163)	(0.105)
观测值	3471	1446	3471	1446	893	4024	893	4024

首先,无论是男性还是女性劳动者,互联网使用对其过度教育均有显著负向影响,同时,互联网使用能显著缓解女性过度教育者的收入惩罚,但对男性劳动者不显著。可能的原因是,相比于男性,受困于家庭与工作平衡问题的女性更易从互联网使用中突破资源限制和信息隔离,获得更多远程、弹性及智慧工作的机会,从而获得更高收入(Santiago等,2022)。其次,对不同户口类型的劳动者而言,互联网使用均能显著降低其过度教育风险。互联网使用能显著缓解农村户口劳动者过度教育收入惩罚,但对城镇户口劳动者的影响不显著。这可能是因为互联网使用可以打破资源的时空、地域限制,突破传统上只能通过学历教育增加人力资本的局限,减轻农村户口劳动者进行人力资本投资的经济负担,为其提高职业能力提供便利(李晓钟、李俊雨,2022),从而提高农村户口劳

动者的工资收入。最后,互联网使用对非垄断行业劳动者的过度教育有显著负向影响,而对垄断行业劳动者的影响不显著。过度教育与互联网使用的交互项对非垄断行业的劳动者的工资收入产生了显著正向影响,但是在垄断行业样本中并不显著。这说明互联网在市场化环境中更能得到充分的发展和应用(刘诗洋等,2022),非垄断行业的劳动者可以通过互联网使用扩展就业空间,减少就业市场中的摩擦,寻求更匹配的就业岗位,从而提高生产效率,获得更高的收入和更好的福利。

五、结论与讨论

本文使用2014~2018年中国家庭追踪调查数据实证检验了互联网使用对过度教育收入惩罚效应的影响及作用机制,得出以下结论:首先,互联网使用有助于降低过度教育的发生风险,对过度教育者而言,互联网使用有助于增加其收入,由此缓解劳动者的过度教育风险及其收入惩罚。该结论在经过稳健性检验和内生性问题处理后仍然成立。其次,使用互联网进行工作、学习和社交活动有效降低过度教育风险,而使用互联网进行商业活动并未产生显著效果;同时使用互联网进行工作、学习、社交和商业活动均对过度教育者的收入惩罚效应起到了缓解作用。最后,互联网使用在不同群体中的影响效果存在异质性,互联网使用显著降低了非垄断行业劳动者的过度教育风险,对女性、农村户口的过度教育者的收入惩罚缓解效应更大。

基于以上研究发现,本文提出如下政策建议。首先,政府在加大“互联网+”基础建设投入的同时,应鼓励和创造更多互联网使用场景,改善互联网就业服务的可得性和可及性。政府应强化劳动力市场的数字化建设,优化数字技术在公共服务、商务交易等领域的应用,进而拓展劳动者互联网使用的工作生活场景;推广落实“互联网+职业技能培训计划”,倡导社会公益组织开展数字技能培训等活动,积极构建覆盖各类职业发展需求的在线学习平台,帮助劳动者掌握和运用工作所需的数字技术,如信息搜集、数据处理等,实现“授之以渔”;强化劳动者在工作、学习、社交和商业活动中的互联网使用,引导劳动者有针对性地提升数字能力,最大限度地减少教育过度、缓解收入惩罚。其次,重点关注女性、农村户口、非垄断行业就业群体,加强数字技能培训,引导不同群体劳动者跨越数字鸿沟,实现数字就业红利的充分共享。为女性群体在互联网使用方面创造有利条件,重点扶持既能帮助女性脱离工作与生活平衡困境,又能激发女性认知与社交技能等优势的新业态,鼓励女性提高自身的数字素养并积极参与各类互联网使用相关的创收活动;推动数字技能培训、信息服务等资源向农村地区和非垄断行业倾斜,充分利用互联网平台优势,降低农村户口劳动者、非垄断行业就业群体参与市场活动的门槛和成本,最大程度发挥互联网使用的增收作用。

本文中还存在一定的局限性。一是研究仅考虑了教育程度与职业要求之间的匹配,但对于工作领域和技能方面的匹配缺少考量。随着数据的更新和补充,相关研究可纳入

更多的维度进行测量和讨论。二是尽管本文发现“数量效应”和“收入效应”是互联网使用缓解劳动者过度教育收入惩罚的作用机制,但囿于数据资料限制,研究未能进一步探究过度教育发生前后更深入的作用路径,有待日后进一步开展研究。

参考文献:

1. 丁述磊、刘翠花(2022):《数字经济时代互联网使用、收入提升与青年群体技能溢价——基于城乡差异的视角》,《当代经济管理》,第8期。
2. 方超、黄斌(2018):《中国过度教育测量、趋势及其影响因素的实证研究》,《教育科学》,第4期。
3. 傅琼(2023):《教育错配与薪资效应》,《浙江社会科学》,第4期。
4. 高志鹏等(2023):《数字技术对农民工劳动收入和福利的影响——基于产业结构转型视角的分析》,《中国人口科学》,第2期。
5. 蒋文莉、赵昕(2022):《互联网使用能够减缓劳动者工资扭曲吗?——来自CFPS数据的证据》,《华中科技大学学报(社会科学版)》,第2期。
6. 李根丽、尤亮(2020):《教育错配、非正规就业与工资惩罚效应》,《财政研究》,第12期。
7. 李晓钟、李俊雨(2022):《数字经济发展对城乡收入差距的影响研究》,《农业技术经济》,第2期。
8. 李勇、段诗宁(2021):《高校扩招如何影响了人力资本配置?》,《南京财经大学学报》,第4期。
9. 刘诗洋等(2022):《互联网普及能否改善人力资本错配》,《当代财经》,第6期。
10. 宁光杰、杨馥萍(2021):《互联网使用与劳动力产业流动——对低技能劳动者的考察》,《中国人口科学》,第2期。
11. 戚聿东等(2021):《数字经济背景下互联网使用与灵活就业者劳动供给:理论与实证》,《当代财经》,第5期。
12. 沈煜等(2023):《高校扩招、过度教育与职业陷阱》,《财贸经济》,第4期。
13. 盛斌、郝碧榕(2021):《全球价值链嵌入与技能溢价——基于中国微观企业数据的经验分析》,《国际贸易问题》,第2期。
14. 汪鲸、罗楚亮(2023):《社会网络有助于降低过度教育概率吗?——基于中国家庭追踪调查(CFPS)的数据分析》,《财经研究》,第10期。
15. 王军、韩悦(2023):《互联网使用对农村劳动力非农就业的影响研究——理论机制与微观证据》,《经济问题》,第9期。
16. 王少国、匡国静(2022):《互联网使用、技能异质性与工资收入——基于CGSS数据的实证检验》,《求是学刊》,第4期。
17. 魏浩然等(2023):《女性更能在互联网使用中获益吗——基于互联网使用偏好的调节效应》,《财经科学》,第3期。
18. 夏璋煦、丁守海(2023):《人力资本贬值预期对女性生育决策的影响——基于职业生涯动态视角的分析》,《中国人口科学》,第4期。
19. 许冰、胡俊(2022):《数字金融发展与人力资本提升——基于中国地级城市层面的经验证据》,《技术经济与管理研究》,第12期。
20. 颜敏、王维国(2018):《教育错配对工资的惩罚效应——来自中国微观面板数据的证据》,《财经研究》,第3期。
21. 尤亮、李根丽(2023):《过度教育与劳动者离职倾向》,《外国经济与管理》,第1期。
22. 岳希明、蔡萌(2015):《垄断行业高收入不合理程度研究》,《中国工业经济》,第5期。
23. 赵放、蒋国梁(2023):《数字科技赋能劳动力就业:内在机理与微观证据——基于CFPS数据的实证分

析》,《人口学刊》,第2期。

24. 赵建国、周德水(2019):《教育人力资本、互联网使用与新生代农民工职业选择》,《农业经济问题》,第6期。

25. 周敏丹(2021):《人力资本供给、工作技能需求与过度教育》,《世界经济》,第7期。

26. 周烁、张文韬(2021):《互联网使用的主观福利效应分析》,《经济研究》,第9期。

27. Bloom N., Liang J., Roberts J., et al. (2015), Does Working from Home Work? Evidence from a Chinese Experiment. *The Quarterly Journal of Economics*. 130(1):165-218.

28. Bonfadelli H. (2002), The Internet and Knowledge Gaps: A Theoretical and Empirical Investigation. *European Journal of Communication*. 17(1):65-84.

29. Dettling L.J. (2017), Broadband in the Labor Market: The Impact of Residential High-Speed Internet on Married Women's Labor Force Participation. *ILR Review*. 70(2):451-482.

30. Lu Y., Xie H., Xu L.C. (2016), Telecommunication Externality on Migration: Evidence from Chinese Villages. *China Economic Review*. 39(7):77-90.

31. Meliciani V., Radicchia D. (2016), Informal Networks, Spatial Mobility and Overeducation in the Italian Labour Market. *The Annals of Regional Science*. 56:513-535.

32. Santiago-Vela A., Mergener A. (2022), Gender Overeducation Gap in the Digital Age: Can Spatial Flexibility through Working from Home Close the Gap?. *Social Science Research*. 106:102727.

33. Wang Z.Z. (2021), The Incidence of Overeducation among International Students upon Graduation in Ireland: The Effects of Social Capital and Job Searching Methods. *Irish Educational Studies*. 40(1):51-69.

34. Wu N., Wang Q. (2018), Wage Penalty of Overeducation: New Micro-Evidence from China. *China Economic Review*. 50(C):206-217.

Can Internet Use Mitigate the Income Penalty of Overeducation in the Digital Economy?

Li Yuru Zhang Shanshan

Abstract: Exploring the use of the internet to mitigate the income penalty of overeducation offers a new perspective for enhancing labor market efficiency in the digital economy era. Using data from the China Family Panel Studies from 2014 to 2018, this paper investigates the impact of internet use on workers' risk of overeducation and its mitigating effect on income penalties, and analyzes the underlying mechanisms. The study finds that internet use helps reduce the risk of overeducation (the "quantity effect") and increases the income of overeducated workers (the "income effect"), thus mitigating the risk and income penalty associated with overeducation. Detailed research shows that using the internet for work, learning, and social activities all contribute to reducing overeducation risk and mitigating the income penalty for overeducated workers. Heterogeneity analysis reveals that internet use significantly reduces the risk of overeducation in non-monopoly industries and has a greater effect on mitigating the income penalty for overeducated female and rural hukou workers. Based on these findings, the paper suggests leveraging the opportunities provided by the digital economy to enhance workers' digital skills through internet empowerment, promoting higher-quality and fuller employment.

Keywords: Internet Use; Overeducation; Income Penalty; Mitigating Effect

(责任编辑:华 颖)