

技术变革下的共同富裕： 数字素养与农民工收入分配*

张锦华 游葭露

内容提要：提升数字素养是数字时代的要求，也是弥合数字鸿沟、扎实推进共同富裕的关键举措。本文基于中国劳动力动态调查(CLDS)数据，测算数字素养对不同行业农民工收入差距的贡献度，系统研究了数字素养对农民工收入分配的影响。研究发现，第一，制造业中数字素养与文化素养呈现“互补效应”，而服务业从“替代效应”逐渐过渡到“生产率效应”；第二，数字素养对制造业农民工收入差距的贡献度趋于提高，但对流通性服务业和生活性服务业的作用呈弱化趋势；第三，城市包容性有助于缩小收入差距，而传统社会网络的影响正在减弱。本文为扎实推进共同富裕提供了微观依据。

关键词：收入差距 农民工 数字素养 文化素养

作者简介：张锦华，上海财经大学财经研究所、城乡发展研究院研究员，200433；

游葭露(通讯作者)，海德堡大学在站博士后，69117。

中图分类号：F062.5 **文献标识码：**A **文章编号：**1002-8102(2024)07-0094-18

一、引言

党的二十大报告指出：“我们要实现好、维护好、发展好最广大人民根本利益，紧紧抓住人民最关心最直接最现实的利益问题……增强均衡性和可及性，扎实推进共同富裕。”该论断表明，居民收入差距已成为当前中国社会面临的重要问题。2022年，中国农民工群体规模达2.9亿人，近年来其收入水平虽不断提升，但与城市居民相比，农民工群体工作强度大，在就业、医疗、居住、生活等方面受到不平等待遇，福利水平严重不足，收入增长速度缓慢(袁方、史清华，2013)。如何提升农民工收入一直是中国社会发展面临的难题。

以互联网为代表的信息与通信技术(ICT)开启了信息时代新纪元。这不仅是一场技术革命，而且影响着人们生活的方方面面。数字素养是居民学习、工作、生活应具备的数字获取、制作、使

* 基金项目：国家自然科学基金面上项目“教育准入、职业选址与农民工市民化”(72173080)；上海市教育委员会科研创新重大项目“解决相对贫困问题的长效机制研究”(2023SKZD13)；国家社会科学基金重大项目“新发展格局下我国粮食安全风险防范研究”(23&ZD118)。感谢匿名审稿专家的宝贵意见，文责自负。游葭露邮箱：lunajialu@163.com。

用、交互、分享、创新等一系列素质与能力的集合,在“互联网+”“智慧城市”“数字中国”等一系列国家信息化发展战略背景下研究数字素养对农民工群体收入的影响无疑具有重要的现实意义。已有文献大多认为使用互联网可以提高微观个体居民的收入(DiMaggio和Bonikowski,2008),但目前国内的相关研究多数集中于单一的互联网使用对农村居民收入的影响(曾亿武等,2018),鲜有涉及综合的数字素养是否有助于提升农民工群体收入。一方面,互联网使用会使个体获取更优质的工作信息,降低个体工作搜寻过程中的信息成本,使农民工与雇主更快、更精准地完成劳动市场匹配,打破农民工务工过程中的时间、空间限制,提高就业概率(张卫东等,2021)。另一方面,Bonfadelli(2002)发现受教育程度和收入水平高的用户一般将互联网用于获取经济收益,而社会经济地位低的用户则更多地将互联网用于娱乐,农民工更可能将闲暇时间用于刷抖音、追剧等消遣而非利用便捷的网络学习知识技能。此外,农民工群体进入城市务工,未能从电商等ICT中获益,促进收入增加。因此,数字素养究竟是否有利于农民工增收尚未有明确答案,这是本文要探究的第一个问题。进一步,已有研究发现,我国农民工常规教育回报率很低,且呈现下降趋势,农民工会在继续接受教育和进城务工之间选择后者(王春超、叶琴,2014),该决策带来的后果是农民工文化素养的缺失,那么农民工群体能否通过提升数字素养弥补其文化素养回报的不足?究竟是提高农民工的数字素养还是文化素养,才能更加匹配当前劳动力市场的需求?这是本文要探究的第二个问题。

习近平总书记在《习近平谈治国理政》第二卷中强调,要着力解决发展失衡、治理困境、数字鸿沟、分配差距等问题。技术变革在优化产业结构、提升经济运行效率的同时,往往也会对社会公平和收入分配提出更大挑战。ICT的快速发展与应用带来了全球生产模式、管理模式和营销模式以及产业链、供应链、价值链的深刻变革,随之可能对居民收入分配产生重要影响。部分学者认为,ICT会在贫富阶层之间画下“数字鸿沟”,扩大收入差距(Britz,2004)。但也有学者认为,如果信息是低成本且完全的,那么低成本提供信息的ICT会使市场价格离散进一步缩小,市场的产品配置会更加合理,使得低收入群体也享受到“信息红利”(Stigler,1961)。一些学者研究发现手机信号的覆盖和互联网的使用对提升农产品市场效率、提高农产品销售价格、改善中间商和农民的福利有着显著的积极作用(Jensen,2007;Andrianarison等,2010;许竹青等,2013)。而近期的研究直接强调了数字素养在收入差距中的作用。例如,Acemoglu(2020)发现在美国移民群体中信息技术造成的劳动力市场技能溢价是收入不平等加剧的原因之一。Naik(2022)认为数字素养较高的个体利用“互联网社会网络”迅速获取市场信息并转换为资本,数字素养较低的个体只能依赖传统社会网络获取滞后信息,信息机会不平等间接引致收入不平等。综观以上研究,关于ICT对收入分配影响的讨论主要集中在高-低收入群体的差异方面,尚未有文献研究ICT带来的行业变革对不同行业农民工群体收入差异的影响。

《中国劳动统计年鉴》数据显示,2016—2019年,行业间收入差距逐渐扩大,由2.7081上升到3.2406。已有文献表明行业工资差距与技术偏向性进步(Plasman等,2007;Acemoglu和Author,2011)、行业垄断特征(罗楚亮等,2019)密切相关,但不同行业劳动力市场对个体技能需求变化、行业间工资差距的影响逐渐引起学界关注。当前,数字化改变了产业结构和劳动力市场需求。世界银行发布的《2019世界发展报告:工作性质的变革》指出,发达经济体对非重复性认知技能和基本数字技能密集型就业比例从33%上升到41%。这意味着企业对表征沟通和融入能力的文化素养技能需求以及全新的数字素养技能需求增长不对称是引致行业间工资差距扩大的重要因素。有文献认为数字技术增加了制造业对劳动力数字素养的需求,减少了对劳动力传统文化素

养的需求(Gupta等,2019;周云波等,2017;余玲铮等,2021),造成行业间收入差距的“马太效应”。也有文献认为,依托“互联网”快速发展的新兴经济体在劳动力市场上为传统技能劳动者创造了更多岗位,实现了行业数字素养需求与传统文化素养需求的“技能互补效应”,从而缩小了行业间的相对收入差距(Hodgson,2014;Koning,2019)。在当前数字素养重要性日益凸显的背景下,比较农民工由沟通、团队协作和城市融入能力产生的文化素养与其数字素养技能孰能更加匹配因数字化而改变的劳动力市场需求,探析二者对不同行业间农民工收入差距的作用,是本文要研究的第三个问题。

基于以上三个问题,本文首先建立包含数字素养和文化素养两种投入要素的两部门生产函数理论模型,阐明数字素养和文化素养在行业间的替代弹性对农民工收入差距的影响,随后利用中国劳动力动态调查(CLDS)数据对理论假说进行实证检验,进而使用FFL-OB分解测度数字素养和文化素养对农民工行业间收入差距的贡献度,并通过随机森林模型预测不同行业农民工收入差距变动趋势,最后探讨数字素养和文化素养影响农民工收入分配的潜在机制。

本文的边际贡献可能包括以下三个方面。第一,从数字素养和文化素养的全新视角,引入二者的替代弹性,探究技术变革对农民工收入及收入差距的影响,丰富了数字经济、技能互补和收入不平等领域的文献。第二,将制造业、生产性服务业和生活性服务业纳入统一的研究框架,使用FFL-OB分解和随机森林模型探析数字素养和文化素养对各行业农民工收入差距不平等的贡献度以及收入差距的变动趋势,弥补了以往文献中静态收入分解存在过度噪声和拟合不足的缺陷。第三,讨论了包括行业间技能结构、社会网络和城市包容度在内的数字素养影响农民工收入的直接机制、间接机制和调节因素,对如何提高农民工收入、缩小农民工行业收入差距和实现共同富裕提供了政策建议。

二、理论机制

基于Van de Klundert(1989)的两部门经济理论模型,构建“制造业-服务业”两部门劳动力市场模型,刻画文化素养和数字素养与行业间收入的关系。假设劳动力市场存在服务业(S)和制造业(M)两个封闭经济部门,劳动者数字素养和文化素养水平各异。每个部门需要三种投入进行生产:文化素养 U 、数字素养 T 两种劳动要素投入以及部门固定资本。一般化生产函数为:

$$Y_{it} = K_{it}^{1-\alpha_{it}} e_{it} L_{it}^{\alpha_{it}}; i = M, S \quad (1)$$

假设生产函数是规模报酬不变的, $K_{it}^{1-\alpha_{it}}$ 是外生给定的部门固定资本投入。其中, Y_{it} 是部门 i 在 t 年的总产出; e_{it} 和 L_{it} 分别是部门劳动要素生产率和劳动要素投入; α_{it} 是劳动要素的产出弹性,每个部门劳动要素投入是文化素养和数字素养之和。则劳动要素投入CES函数为:

$$L_{it} = C_{it} \left[\beta_{it} (\delta_{it}^U U_{it})^{\frac{\rho_{it}-1}{\rho_{it}}} + (1 - \beta_{it}) (\delta_{it}^T T_{it})^{\frac{\rho_{it}-1}{\rho_{it}}} \right]^{\frac{\rho_{it}}{\rho_{it}-1}} \quad (2)$$

其中, C_{it} 是部门 i 在 t 年的每单位劳动要素投入; β_{it} 和 $1 - \beta_{it}$ 表示CES函数的分配系数; δ_{it}^U 和 δ_{it}^T 分别表示文化素养和数字素养的质量变化; U_{it} 和 T_{it} 分别是部门 i 在 t 年的文化素养和数字素养需求数量; ρ_{it} 为文化素养和数字素养劳动要素替代参数。根据式(1)和式(2),将部门生产函数设定为:

$$Y_{it} = K_{it}^{1-\alpha_{it}} \theta_{it} \left[\beta_{it} (\delta_{it}^U U_{it})^{\frac{\rho_{it}-1}{\rho_{it}}} + (1-\beta_{it}) (\delta_{it}^T T_{it})^{\frac{\rho_{it}-1}{\rho_{it}}} \right]^{\frac{\alpha_{it} \rho_{it}}{\rho_{it}-1}}; \theta_{it} = e l_{it} C_{it} \quad (3)$$

其中, θ_{it} 表示部门 i 的希克斯中性技术进步水平; δ_{it}^U 和 δ_{it}^T 分别是劳动力市场中文化素养和数字素养技术进步偏向。

(一) 不存在比较优势两部门模型

假设劳动力市场是完全竞争市场, 服务业和制造业对数字素养和文化素养的需求相同, 且各部门不存在剩余劳动力。文化素养和数字素养的分配系数相等, 即 $\beta_{it} = 1 - \beta_{it}$; 文化素养和数字素养的质量无变化, 即 $\delta_{it}^U = \delta_{it}^T = 1$; ρ_{it} 外生给定。劳动投入要素一阶条件是劳动力工资等于企业预期劳动力价值, 劳动力工资(劳动边际产出)为:

$$w_{it} = \frac{1}{2} \alpha_{it} \delta_{it}^{\frac{\rho_{it}-1}{\rho_{it}}} Y_{it}^{[\rho_{it}(\alpha_{it}-1)+1]/\alpha_{it} \rho_{it}} / I_{it}^{\frac{1}{\rho_{it}}} \theta_{it}^{\frac{\alpha_{it} \rho_{it}}{\rho_{it}-1}} \quad (4)$$

其中, $I_{it}^{1/\rho_{it}} = U_{it}^{1/\rho_{it}} + T_{it}^{1/\rho_{it}}$, 劳动力工资由可观察到的文化素养 $U_{it}^{1/\rho_{it}}$ 和数字素养 $T_{it}^{1/\rho_{it}}$ 决定。劳动力工资 w_{it} 是可观察到的文化素养 $U_{it}^{1/\rho_{it}}$ 和数字素养 $T_{it}^{1/\rho_{it}}$ 的增函数, 行业间劳动力技能需求结构变化决定劳动者技能回报。在数字素养和文化素养需求比例相同的劳动力市场, 劳动力文化素养 $U_{it}^{1/\rho_{it}} \rightarrow 1$, 数字素养 $T_{it}^{1/\rho_{it}} \rightarrow 1$, 企业支付劳动力工资 $w_{it} \rightarrow 1$ 。

根据式(4), 可得到研究假说1。

研究假说1: 数字素养和文化素养的需求变化决定劳动力工资水平。

(二) 比较优势两部门模型

假设劳动力市场是完全竞争市场, 服务业是数字素养密集型行业, 其经济效益高于制造业。数字素养劳动生产率大于文化素养劳动生产率, 即 $e l_{T_{it}} > e l_{U_{it}}$; ρ_{it} 外生给定。根据式(2)和式(3), 文化素养和数字素养(劳动边际产出)工资率分别为:

$$w_{U_{it}} = \frac{\alpha_{it} \beta_{it} \delta_{it}^{\frac{\rho_{it}-1}{\rho_{it}}} Y(U_{it})^{\frac{\rho_{it}(\alpha_{it}-1)+1}{\alpha_{it} \rho_{it}}}}{\frac{1}{U_{it}^{\rho_{it}}} \theta_{it}^{\frac{\alpha_{it} \rho_{it}}{\rho_{it}-1}}}; i = M, S \quad (5)$$

$$w_{T_{it}} = \frac{\alpha_{it} (1-\beta_{it}) \delta_{it}^{\frac{\rho_{it}-1}{\rho_{it}}} Y(T_{it})^{\frac{\rho_{it}(\alpha_{it}-1)+1}{\alpha_{it} \rho_{it}}}}{\frac{1}{T_{it}^{\rho_{it}}} \theta_{it}^{\frac{\alpha_{it} \rho_{it}}{\rho_{it}-1}}}; i = M, S \quad (6)$$

假设每个行业的流动要素边际劳动生产率等于工资比, 根据式(5)和式(6), 劳动力市场出清条件为 $\sum_i U_{it} = \bar{U}_t$, $\sum_i T_{it} = \bar{T}_t$, 决定了文化素养和数字素养均衡工资。假设服务业对数字素养的需求大于对文化素养的需求, 因此 $Y(M_t) = \partial_{it} \delta_{it}^U U_{it}$, $Y(S_t) = (1 - \partial_{it}) (\delta_{it}^T T_{it})$, 行业收入差距与劳动要素相对价格成正比, 代入式(3)可得行业收入差距为:

$$W(S_0) = \frac{\bar{U}_t}{\bar{T}_t} = \left(\frac{\theta_{M_t}}{\theta_{S_t}} \right)^{\frac{\rho_{it}}{1-\alpha_{it} \rho_{it}}} \left(\frac{\delta_{it}^U}{\delta_{it}^T} \right)^{\frac{(1-\alpha_{it}) \rho_{it}}{1-\alpha_{it} \rho_{it}}} \quad (7)$$

式(7)表示工资差距取决于各行业文化素养与数字素养的替代关系 $(1 - \alpha_{it}) \rho_{it} / (1 - \alpha_{it} \rho_{it})$, 当 $0 < \rho_{it} < 1$ 时, 文化素养与数字素养在该行业是替代效应; 当 $\rho_{it} < 0$ 时, 文化素养与数字素养在该行

业是技能互补效应。假设在某行业中,数字素养与文化素养是替代效应,导致那些只具备文化素养的农民工因缺乏数字素养而面临收入的负面冲击,从而形成了明显的“收入鸿沟”。而当劳动力市场上具备数字素养的农民工供应量超过行业的实际需求,且数字素养与文化素养在该行业中呈现技能互补关系时,那些仅具备数字素养的农民工的收入有可能会受到压制,这有助于缩小“收入鸿沟”。各行业数字素养和文化素养的替代弹性不同,引致行业间农民工收入发生变化。基于上述理论模型推导,得到如下研究假说。

研究假说2:数字素养对农民工行业间的“收入鸿沟”是先扩大后缩小。

研究假说3:各行业对数字素养需求增加的同时也增加对文化素养的相对需求。

三、实证设计

(一)模型设定

1.基准模型:数字素养与收入

本部分的实证模型主要检验研究假说1。使用标准工资方程(Mincer, 1974):

$$\ln wage_i = \alpha_0 + \beta_1 literacy_{1i} + \beta_2 literacy_{2i} + \gamma_0 X_i + \varepsilon_i \quad (8)$$

其中,因变量 $\ln wage_i$ 为第 i 个农民工的时工资对数,以年工资对数检验其稳健性。自变量 $literacy_{1i}$ 和 $literacy_{2i}$ 分别为第 i 个农民工的文化素养和数字素养; X_i 是可能影响农民工收入的个体特征变量和企业特征变量,控制个体特征和企业特征可以尽量缓解遗漏变量偏误,选取的控制变量包括年龄、性别、受教育年限、工作经验、社会保障、工作年限、婚姻状况等个人特征以及企业规模、所属行业等企业特征; γ_0 为农民工个体特征向量对收入的影响; β_1 、 β_2 分别为文化素养和数字素养的系数; ε_i 为随机误差项; α_0 为常数项。

2.行业间收入差距:FFL-OB分解

对工资差距进行分解,关注的焦点是所有行业样本及全部人群,较少关注行业间工资差距及农民工群体。为进一步讨论文化素养和数字素养对不同分位数上行业间工资差距的影响,解决遗漏变量造成回归结果不一致的问题,使用FFL-OB分解^①(Firpo等, 2018)检验研究假说2和研究假说3,刻画2014年、2016年文化素养和数字素养对行业间收入差距的影响。具体模型为:

$$\ln E(wage_{it}) = \beta_{1it} E(literacy_{1i}) + \beta_{2it} E(literacy_{2i}) \quad (9)$$

其中, $\ln E(wage_{it})$ 是 t 年不同行业间第 i 个农民工的时工资对数样本均值, $E(literacy_{1i})$ 、 $E(literacy_{2i})$ 分别是 t 年文化素养和数字素养的样本均值, β_{1it} 、 β_{2it} 分别是2014年和2016年工资回归方程系数,关注点是不同行业间农民工工资平均差异,即 $D[\ln E(wage)] = \ln E(wage_{2014i}) - \ln E(wage_{2016i})$ 受到文化素养和数字素养的影响,代入式(9)可得:

$$D[\ln E(wage)] = \beta_{2016} E(X_{2016}) - \beta_{2014} E(X_{2014}) = E(X_{2016})(\beta_{2016} - \beta') + \beta_{2014}[E(X_{2016} - X_{2014})] + E(X_{2016})(\beta' - \beta_{2014}) \quad (10)$$

其中, $X_{2016} = literacy_{12016i} + literacy_{22016i}$, $X_{2014} = literacy_{12014i} + literacy_{22014i}$; β' 是均衡要素回报率;

① 采用样本权重回归以解决无条件分位数回归样本限制问题。

$E(X_{2016})(\beta_{2016} - \beta')$ 是要素回报率差异,即数字素养和文化素养回报率; $\beta_{2014}[E(X_{2016} - X_{2014})]$ 是要素禀赋差异,即收入结构效应; $E(X_{2016})(\beta' - \beta_{2014})$ 是不可解释部分,避免遗漏变量造成回归结果不一致。利用再中心化影响函数(RIF)对2014年和2016年样本分别进行回归后再进行FFL-OB分解。FFL-OB分解结果为:

$$D[\ln E(wage)] = E(X_{2016}^T)(\beta_{2016}^v - \beta_c^v) + [E(X_{2016}^T) - E(X_{2014}^T)]\beta_{2016}^v + E(X_{2014}^T)(\beta_c^v - \beta_{2014}^v) \quad (11)$$

其中, $E(X_{2016}^T)(\beta_{2016}^v - \beta_c^v)$ 表示 $T=2016$ 样本和反事实样本信息回报效应; $[E(X_{2016}^T) - E(X_{2014}^T)]\beta_{2016}^v$ 表示 $T=2014$ 样本和反事实样本收入结构效应; $E(X_{2014}^T)(\beta_c^v - \beta_{2014}^v)$ 表示不可解释部分,由技术误差^①和再权重误差^②两部分构成。选取2014年和2016年作为反事实分析的原因如下:一方面,受CLDS数据可得性限制;另一方面,2016年《国家信息化发展战略纲要》颁布以及中国电子商务稳定赶超美国,这意味着在2016年后,中国大力发展信息化技术与信息化平台,选取2016年作为反事实分析是能够对信息化技术发展前后进行对比的。

3. 随机森林模型的不平等贡献度

FFL-OB分解只能静态判别文化素养与数字素养对不同行业间农民工工资差距的影响程度和方向,无法确定文化素养与数字素养对工资差距的贡献度,故采用随机森林模型度量变量的重要性(Raschka, 2015)。具体步骤如下:首先,将影响收入的解释变量如文化素养、数字素养和其他特征变量共同代入随机森林收入决策函数,估算GINI指数 R' ;其次,将解释变量和特征变量依次代入每个决策树,估算特征变量GINI指数差距 R ;最后,计算每棵决策树GINI指数差距的平均值,估算每个特征变量的贡献度 D_K 。具体模式为:

$$D_K = \sum_r \sum R \varepsilon_{\theta}(\gamma, R) \quad (12)$$

其中, θ 表示所有特征变量随机特征的随机子集; γ 和 R 是随机森林收入决策函数的两个决策树; $\varepsilon_{\theta}(\gamma, R)$ 表示数字素养和文化素养导致的分裂准则函数工资差距。

(二) 数据来源与指标构建

1. 数据来源

本文数据来源于2014年和2016年CLDS数据。该调查问卷由个人、家庭和社区三部分构成,个人与家庭问卷调查涉及务工家庭的人口基本特征、收入与消费、就业与幸福感、职业与工资福利等信息;社区问卷调查涉及村委基本特征、土地与经济、文化与习俗、环境与生态建设等信息。该调查采用多阶段、多层次、与劳动力规模成比例的概率抽样方法,并与轮换样本追踪方式相结合。2014年共401个村庄14214个家庭,其中包括7884个追踪调查家庭;2016年共404个村庄14191个家庭,其中包括8412个追踪调查家庭。未将2014年和2016年个人层面的截面数据合并为非平衡面板的原因如下:首先,CLDS数据中个人样本包括新增和追踪两部分,非平衡面板会导致样本不具有随机性;其次,如果匹配成平衡面板,必然会损失样本容量进而降低估计效率。因此,在第四

① 技术误差是由偏离线性估计造成FFL-OB分解无法解释的部分,技术误差越小,说明分解结果越优。

② 再权重误差是加权最小二乘法“再权重过程”中无法解释的部分,再权重误差越小,说明反事实估计效果越好。

部分提取追踪样本合并成两年短期面板,检验模型的稳健性。为尽可能准确地反映全国的情况,本文对所有数据进行权重回归。^①

数据处理过程如下:首先,剔除缺少样本值的省份以及个体特征变量为“不知道”的样本;其次,由于研究主体是农民工工资,因此只保留农村、非失业^②和非务农人口,并删除自我雇佣者、私营企业主以及从事农、林、牧、渔业和其他行业^③的样本;再次,将农民工年龄限制在16~70周岁^④;最后,对收入变量进行缩尾处理。根据上述筛选方法,2014年最终样本量为5461个,2016年最终样本量为4643个。

2. 指标构建

核心解释变量。核心解释变量为数字素养与文化素养。按照Tuominen等(2005)、Majid等(2013)对数字素养的界定,在信息技术出现前,数字素养是指阅读与运用语言的能力,即文化素养;在信息技术出现后,数字素养是指通过电脑、手机等电子设备获取信息,实现资本积累的能力。数字素养的主要测评方法是问卷调查和职业测评,其中问卷调查是国内外学者经常采用的方法(Dobson和Willinsky,2009)。参照Lanciotti(2019)评估文化素养和数字素养的方法,结合CLDS相关问题,基于社会构建理论以信息价值观、个人通用技能、社会通用技能三个指标衡量文化素养(Oakleaf,2009)。Hodgson(2014)定义的数字素养是数字运用能力或者数字技术转换能力,因此以数字积累能力衡量数字素养(Abdi等,2019)。^⑤鉴于被采访者的回答带有主观意愿,易掩盖真实性,故结合采访员评价,综合测评文化素养和数字素养。^⑥

行业分类。依照问卷中的行业划分标准,即参照《国民经济行业分类代码表》(2009)将行业分成十六个大类。^⑦为探究行业间的收入差距,将十六大类行业分为四类^⑧,分别是农业、制造业、流通性服务业和生活性服务业。由于其他行业不属于四大行业分类,加上该行业样本量较少,所以本文只对十五个行业进行划分。^⑨

控制变量。在企业层面和个人层面均添加控制变量。企业层面控制变量包括单位规模(以平均数作为参照组构建虚拟变量,平均数以上为0,平均数以下为1)、所属行业。^⑩个人层面控制变量包括性别、婚姻状况、年龄、年龄平方、受教育年限、受教育水平(将未上过小学群体作为参照组构建虚拟变量)、职工养老保险、工伤险、生育险、合同签订、技能培训、工作时长和工作经验。^⑪此外,还控制地区(城市层面)对收入的影响。

工具变量。鉴于核心解释变量与被解释变量存在反向因果关系,一方面,文化素养和数字素养通过提高劳动者生产率,直接提高劳动者收入水平;另一方面,收入改善会进一步加大劳动者对

① 2014年权重回归为 svy ,2016年权重回归为 wfd ,基准回归和收入方程分解皆为权重回归,以体现样本具有全国代表性。

② 失业一年以上为失业人口。

③ 行业分为制造业、流通性服务业和生活性服务业,其他行业不属于三大分类,故删除。

④ 70周岁及以下为低能老年人,虽然已返乡,但还会去邻近的镇和县打工(贺雪峰,2015)。

⑤ 培养文化素养的主要目的是扫盲;数字素养是数字资本的一部分,培养数字素养的主要目的是提高数字积累能力。

⑥ 限于篇幅,解释变量定义和赋值说明未列示,留存备案。

⑦ 这十六个大类分别为:农、林、牧、渔业,采掘业,制造业,电力、煤气及水的生产和供给业,建筑业,地质勘察业、水利管理业,交通运输、仓储及邮电通信业,批发和零售贸易、餐饮业,金融、保险业,房地产业,社会服务业,卫生、体育和社会福利业,教育、文化艺术及广播电视业,科学研究和综合技术服务业,国家机关、党政机关和社会团体,其他行业。

⑧ 主要依据2018年国家统计局颁布的《关于修订〈三次产业划分规定(2012)〉的通知》。

⑨ 本文关注的重点是农民工工资,农、林、牧、渔业属于经营性活动,故删除。

⑩ 全样本回归时控制行业变量,考察行业间工资差距时没有控制行业变量。

⑪ 选取控制变量的依据是经过LASSO回归筛选,以减少遗漏变量偏误。

自身文化素养和数字素养的投资。对于遗漏变量,采用LASSO回归以尽可能筛选出影响被解释变量的控制变量,但不排除依旧存在一些不可观察的遗漏变量(Legarra等,2011)。此外,借鉴周云波等(2017)、柏培文和王亚文(2023)的做法,选取1984年农民工户籍所在地的地级市教育从业人数与地区的地形起伏度分别作为文化素养与数字素养的工具变量。选择这两个工具变量的主要原因在于,地形起伏度会影响光纤电缆、通信基站等与数字技术发展相关的设备设施的普及度,但不会影响农民工收入。此外,城市教育水平对文化素养具有外部溢出效应,但与30年后个人收入相独立。

(三)描述性统计

数字素养描述性统计。通过被采访者自评与采访者他评两部分综合评分测量文化素养和数字素养。被采访者自评赋值为1~5,采访者他评赋值为1~4。为确保数据的归一性,首先,对文化素养和数字素养指标进行标准化处理并加总平均,分别求文化素养和数字素养指标数值;其次,考虑到指标相关性,采用主成分分析(PCA)法对指标提取主成分,生成文化素养和数字素养总数值;最后,考虑到文化素养和数字素养可能存在多重共线性问题,对变量进行方差膨胀系数(VIF)检验,检验值为1.33,远小于经验临界值10,说明不存在多重共线性问题。^①

行业间收入分布。2014年和2016年最高工资行业与最低工资行业间的工资差距分别为3785.14元和3900.44元。对行业时均工资的收入数据统计描述表明,2014年批发和零售贸易、餐饮业的时均工资最低,为13.41元;2016年社会服务业的时均工资最低,为13.88元。2014年和2016年金融、保险业的时均工资最高,分别为31.16元和32.96元。由此可以看出,无论是年均工资还是时均工资,各行业间的工资差距均在扩大。高工资行业集中在金融、保险业,房地产业,科学研究和综合技术服务业等知识密集型产业;相反,低工资行业集中在批发和零售贸易、餐饮业,制造业,建筑业等劳动密集型产业。对比2014年和2016年数据可以看出,劳动力供给逐渐从制造业向生活性服务业转移。^②

主要变量描述性统计。农民工工作时长普遍较长,以年均工资作为被解释变量易出现测量误差,故选取时工资为被解释变量,用年工资验证时工资的稳健性。^③

四、实证分析

(一)基准模型回归:数字素养与农民工收入

行业特性差异会造成劳动要素禀赋差异。这意味着,制造业固定资产投资比例大,资产流动性差,企业面临的数字产业升级成本相对较高,因此从事该行业的劳动者大多是文化素养熟练工;与之相反,流通性服务业和生活性服务业资产流动性较强,固定资产成本低,实现数字转型的成本低,因此从事服务业的劳动力大多需要具备基本数字素养。本文按照制造业、流通性服务业和生活性服务业分组探究文化素养与数字素养对不同行业间收入的差异性影响。表1报告了文化素养与数字素养对行业间工资差异的影响。模型1为2014年文化素养与数字素养对不同行业工资影响的基准回归。结果显示,在控制一系列变量后,文化素养对流通性服务业和生活性服务业的影

① 限于篇幅,碎石图、PCA结果、数字素养和文化素养单独回归结果(回归结果再次证明不会受到多重共线性影响)未列示,留存备案。

② 限于篇幅,行业分布与收入差距未列示,留存备案。

③ 限于篇幅,主回归方程中各变量说明、样本数、均值、标准差未列示,留存备案。

响系数为负且不显著,对制造业的影响系数为正且不显著,这意味着从事该行业的劳动者需要具备较高的社交技能。从数字素养的回归结果看,数字素养对所有行业至少在5%的水平下能够促进农民工提高时工资,其中对生活性服务业的影响最大,是制造业的1.40倍。我国生活性服务业拥有庞大的消费群体,企业须不断进行数字消费升级以满足消费者的数字化需求。在模型2中,基于2016年的数据分析结果显示,文化素养对制造业工资的影响并不显著,而在流通性服务业和生活性服务业中,文化素养对农民工的时工资有显著的负向影响,这意味着随着数字技术的不断进步,服务业对数字素养技能的需求越来越高,而文化素养的重要性逐渐被数字素养所替代,从而使得文化素养对工资产生负向影响。这一趋势暗示了在数字经济的推动下,就业结构正在经历转型,其中以数字素养为核心的岗位逐步成为服务业的主导。此外,数字素养对制造业工资的影响系数为0.679,对流通性服务业工资的影响系数上升至0.871,而数字素养对生活性服务业的收入效应由2014年的0.838下降至2016年的0.565。^①

表 1 行业间基准回归

变量	2014年			2016年		
	ln时工资			ln时工资		
	模型 1			模型 2		
	制造业	流通性服务业	生活性服务业	制造业	流通性服务业	生活性服务业
文化素养	0.261 (0.175)	-0.210 (0.150)	-0.189 (0.174)	0.212 (0.714)	-0.084*** (0.045)	-0.011*** (0.004)
数字素养	0.598*** (0.024)	0.634** (0.241)	0.838*** (0.058)	0.679*** (0.006)	0.871*** (0.000)	0.565*** (0.005)
个体特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
企业特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
地区固定效应	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.340	0.131	0.286	0.508	0.386	0.343
N	2325	924	1699	2096	784	1453

变量	2014年			2016年		
	ln年工资			ln年工资		
	模型 3			模型 4		
	制造业	流通性服务业	生活性服务业	制造业	流通性服务业	生活性服务业
文化素养	0.261 (0.175)	-0.210 (0.151)	-0.189 (0.173)	0.246 (0.511)	-0.082*** (0.045)	-0.011*** (0.004)
数字素养	0.598*** (0.024)	0.634** (0.242)	0.838*** (0.058)	0.679*** (0.000)	0.871*** (0.000)	0.565*** (0.005)
个体特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
企业特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
地区固定效应	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.340	0.131	0.286	0.310	0.303	0.343
N	2325	924	1699	2096	784	1453

注:*,**和***分别表示在10%、5%和1%的水平下显著。下同。括号内为基于个人层面的聚类标准误。

(二)稳健性检验

工具变量法。本文的内生性问题可能来源于遗漏变量及反向因果,即无法控制所有与数字素

① 限于篇幅,全样本回归结果、城镇就业群体和农民工群体对比结果未列示,留存备索。

养、文化素养相关且会影响农民工收入的变量,收入高的农民工个体也可能具有更好地提升数字素养与文化素养的机会。针对以上两个潜在的内生性来源,本文采用工具变量法予以解决。具体而言,本文借鉴周云波等(2017)、柏培文和王亚文(2023)的做法,选取1984年农民工户籍所在地地级市教育从业人数与地区的地形起伏度分别作为文化素养和数字素养的工具变量。工具变量回归结果与基准回归结果基本一致。^①

基于追踪样本的面板回归。为进一步考察基准回归结果的可靠性,使用2014年和2016年追踪样本构建面板模型验证数字素养与行业间农民工收入的关系,解决不可观测个体异质性导致遗漏变量的问题。基于追踪样本的面板回归结果显示,文化素养和数字素养对全样本农民工收入的影响显著为正,但数字素养的收入效应大于文化素养;基于行业分类的分析发现,文化素养对流通性服务业和生活性服务业工资效应的影响显著为负,而对制造业工资效应的影响显著为正。这说明流通性服务业与生活性服务业可能处于数字技术带动下的就业结构转变中,从而增大了对具备数字素养技能劳动力的需求,导致传统文化素养在这些行业中的需求减少;而制造业对传统文化素养仍具有依赖性,其中文化素养与数字素养之间存在技能互补效应,因此文化素养在该行业对农民工收入产生的影响为正。^②

(三)进一步分析

信息作为数字经济的关键生产要素,在提升劳动力市场运行效率的同时,也拉大了行业间、个体间的收入差距。^③文化素养和数字素养对行业间农民工收入差距的影响趋势如何变化?本部分基于FFL-OB分解方法,以2014年为参照年,将2014年和2016年农民工行业间收入差距分解为信息回报效应和收入结构效应两部分。此外,采用随机森林决策树模型进一步探析收入回归方程中各变量对行业间收入差距的贡献度变化趋势。

1. 行业工资差距分解^④

为考察行业间农民工收入差距的变化趋势,本文采用Firpo等(2018)、Rios-Avila(2020)关于基尼系数和方差总体不平等统计量再中心化影响函数的构造方法,以2014年为参照年,基于收入回归方程对各行业收入差距分位数、基尼系数和方差分布进行分解。对比来看,各行业无论是以时工资对数还是以年工资对数为分解对象,两者分解结果都比较相似。^⑤在基尼系数为被解释变量的分解中,信息回报效应反映劳动力市场上企业对具备数字素养技能农民工需求的强度,即信息回报效应越大,具备该技能劳动力的工资水平越高;收入结构效应反映数字化进程中产业结构的变化对具备文化素养和数字素养技能农民工工资差异的影响。制造业时工资差距缩小1.36%是信息回报效应与收入结构效应共同作用的结果,但信息回报效应大于收入结构效应。从信息回报效应看,文化素养和数字素养信息回报效应的系数皆为负,数字素养信息回报效应不显著,即文化素养与数字素养是互补效应。从信息回报特征的工资分布看,数字素养信息回报效应对90~10分位数、50~10分位数、90~50分位数的分解系数分别为0.3742、0.1537、-0.2205,这说明劳动力平均数字素养提升会缩小中高收入群体差距,扩大高低收入群体差距。这意味着在制造业中,数字经济时代下的“收入鸿沟”现象在低收入群体中最为突出,但在中高收入群体中则趋

① 限于篇幅,工具变量回归结果未列示,留存备案。

② 限于篇幅,基于追踪样本的面板回归结果未列示,留存备案。

③ 中国信息通信研究院:《中国数字经济发展与就业白皮书(2019年)》,2019年4月。

④ 限于篇幅,行业工资差距分解表和分解图未列示,留存备案。

⑤ 限于篇幅,本文以ln时工资分解为例。

于缩小。

以流通性服务业时工资对数为分解对象的结果显示,文化素养弥合 90~10 分位数和 50~10 分位数工资差距,但对整体工资差距的影响不显著,因此 2014 年和 2016 年行业收入差距扩大主要是由常数项导致的。从信息回报效应看,数字素养对总回报特征具有正贡献,即数字素养能够扩大行业内工资差距,但在高低端和中低端工资分布系数为负。总体而言,流通性服务业更偏向数字素养劳动力,数字素养和文化素养相对工资差距是造成行业内工资差距的主要原因。

以生活性服务业时工资对数为分解对象的结果显示,数字素养对基尼系数的贡献值为正,即数字素养回报特征能够扩大 0.137% 的工资差距。从收入结构效应看,文化素养对高低端和中低端工资分布的系数分别为 1.8164 和 2.522,这意味着在工资函数中文化素养能够扩大工资差距。此外,数字素养能够缩小 78.7% 的中低端工资差距,即低收入群体跨入中部收入群体的关键是提升数字素养。

从不同行业农民工工资差距的分解结果可以看出,以 50 分位数为分界点,制造业信息回报效应和收入结构效应近乎呈“倒 S 型”;相反,在流通性服务业和生活性服务业中,数字素养在高分位数群体中对收入的解释作用越来越弱。从收入结构效应看,在制造业和流通性服务业中,数字素养收入结构效应在中高收入群体中呈上升趋势,但在生活性服务业中则呈下降趋势。值得注意的是,制造业信息回报效应和收入结构效应曲线具有几乎完全相同的形状,这说明制造业收入差距是信息回报效应与收入结构效应共同作用的结果。此外,流通性服务业收入差距是信息回报效应大于收入结构效应,生活性服务业收入差距是收入结构效应大于信息回报效应。

2. 行业工资差距贡献度

FFL-OB 分解考察了数字素养与文化素养在 2014 年和 2016 年行业间工资差距的静态变化趋势,但无法计算工资方程中各变量行业收入差距的贡献度。本部分基于随机森林机器学习的方法度量各变量的贡献度,^①优化静态分解过度噪声和减少遗漏特征值引起的测量误差。表 2 的变量重要性分解(贡献度)结果显示,2016 年,数字素养贡献度比文化素养贡献度高 0.1525,地区综合指标对制造业的贡献度为 0.1406。横向比较看,2014—2016 年,制造业数字素养对收入不平等的贡献度上升了 0.0115;从流通性服务业随机森林决策树变量重要性分解结果看,文化素养贡献度下降了 0.1612,数字素养贡献度上升了 0.0311。分解结果再次证明流通性服务业处于“替代效应”阶段,而生活性服务业农民工收入差距呈现缩小趋势,数字素养贡献度从 2014 年的 0.0970 下降至 2016 年的 0.0459。此外,地区对各行业的收入不平等贡献度显著上升,说明行业区域溢价现象趋于显现。

表 2 随机森林决策树变量重要性分解

变量		2014 年			2016 年		
		ln 时工资			ln 时工资		
		制造业	流通性服务业	生活性服务业	制造业	流通性服务业	生活性服务业
素养指标	文化素养	0.1480	0.1820	0.1235	0.0446	0.0208	0.0629
	数字素养	0.0913	0.1010	0.0970	0.1028	0.1321	0.0459

① 由于解释变量过多,本文只选取基准收入方程中显著性在 10% 水平下的变量,并合并含义相似变量,以提高分解效率(年龄平方和年龄运用降维的方法合并)。

续表 2

变量		2014年			2016年		
		ln时工资			ln时工资		
		制造业	流通性服务业	生活性服务业	制造业	流通性服务业	生活性服务业
个人综合指标	年龄	0.1657	0.2300	0.1879	0.0991	0.0925	0.1660
	受教育年限	0.0701	0.1720	0.0934	0.0907	0.1201	0.0890
	性别	0.0463	0.2628	0.0378	0.1132	0.0840	0.0303
	工作经验	0.1246	0.0210	0.1326	0.0067	0.1200	0.0179
	工伤险	0.0524	0.0130	0.0478	0.0113	0.0312	0.0345
企业综合指标	企业规模×企业类型	0.0211	0.0088	0.0216	0.0991	0.0701	0.0270
地区综合指标	地区类型	0.0498	0.0121	0.0379	0.1406	0.1346	0.2014
拟合优度	—	0.5725	0.9309	0.5529	0.5274	0.4293	0.4177

注：以市辖区人口密度对城市进行地区类型划分。

图 1 为 2014 年和 2016 年各行业文化素养与数字素养偏依赖。^①对比 2014 年和 2016 年分解结果可以发现,文化素养对行业间农民工收入差距的贡献度有所下降,而数字素养的贡献度上升很快。这说明数字经济的确造成了“收入鸿沟”现象。此外,2014—2016 年,数字素养对制造业收入差距的贡献度提升 12.60%,对流通性服务业收入差距的贡献度提升 30.79%,而对生活性服务业的贡献度趋于稳定,并呈下降趋势。这也再次说明数字经济时代下各行业农民工收入差距具有差异性,这种差异性可能是由行业要素禀赋、平台依赖程度和劳动要素结构等行业特性因素引起的。

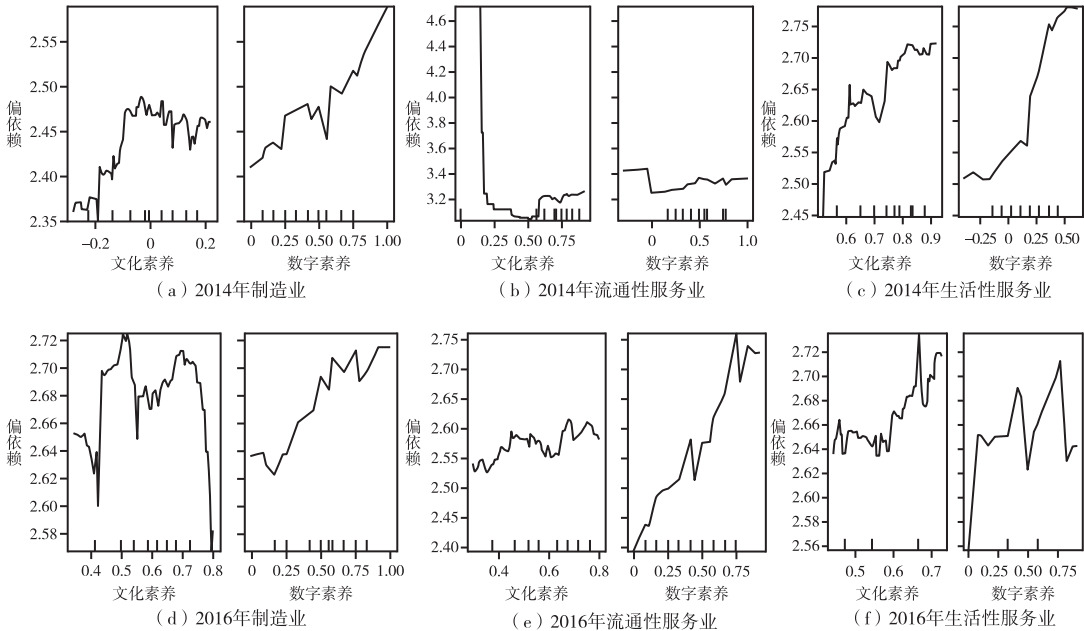


图 1 2014 年和 2016 年文化素养与数字素养偏依赖

① 特征变量文化素养与数字素养对时工资的边际效应为： $\phi(x_1, x_2) = \frac{1}{n} \sum_{i=1}^n f(x_1, x_2)$ 。

五、机制讨论

(一)行业间技能结构

根据行业特性,服务业消费总量大于制造业,能够创造更多低技能岗位,因此数字素养对行业收入差距的差异受技能结构影响。本文根据调查问卷中“为掌握这份工作所需技能花费的时间”构造技能虚拟变量^①,探究技能结构对农民工收入的作用机制。表3结果显示,2014年人均数字素养每提高1个单位,制造业高技能劳动力收入和低技能劳动力收入分别上升0.7367个和0.3587个单位,这说明技能结构的异质性能够增加学习交流机会,带来人力资本外部溢出效应,提升企业生产效率。从文化素养看,2014—2016年,文化素养对制造业高技能劳动力的收入效应下降0.1892个单位,对低技能劳动力由收入效应转为抑制效应,主要原因是制造业产业结构升级诱发企业要素节约偏向(蒲艳萍、顾冉,2019),弱化技能互补的学习交流互动机制,抑制收入增长。从流通性服务业看,技能偏向性技术进步使流通性服务业对高技能劳动力的需求增长,从而使高技能劳动力的供给量与工资水平呈现同步提升的趋势,数字素养对技能溢价的影响系数从1.0382扩大到3.6970。文化素养对生活性服务业高技能劳动力工资产生抑制作用,2014年和2016年人均文化素养每提高1个单位,生活性服务业高技能劳动力收入分别下降0.1206个和0.3908个单位,“收入鸿沟”在生活性服务业中呈现缩小趋势。生活性服务业不可贸易和劳动密集型的属性,一方面,决定了生活性服务业能够吸纳更多低技能劳动力,改善低技能劳动力的收入;另一方面,生活性服务业更需要人与人面对面交流,而人力资本外部性主要是通过面对面交流产生的(梁文泉、陆铭,2016),故“收入鸿沟”在生活性服务业中呈现缩小趋势。

表 3 行业间技能工资差异

变量	2014 年					
	ln 时工资		ln 时工资		ln 时工资	
	制造业		流通性服务业		生活性服务业	
	高技能劳动力	低技能劳动力	高技能劳动力	低技能劳动力	高技能劳动力	低技能劳动力
文化素养	0.7179*** (0.001)	0.1368*** (0.001)	1.3720*** (0.002)	0.9536*** (0.002)	-0.1206*** (0.001)	0.5729*** (0.002)
数字素养	0.7367*** (0.001)	0.3587*** (0.001)	0.8121*** (0.001)	0.7822*** (0.001)	0.7847*** (0.001)	0.29184*** (0.001)
个体特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
企业特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
地区固定效应	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.2425	0.1619	0.0264	1.3770	1.0670	2.0917
N	706	578	269	263	526	327

变量	2016 年					
	ln 时工资		ln 时工资		ln 时工资	
	制造业		流通性服务业		生活性服务业	
	高技能劳动力	低技能劳动力	高技能劳动力	低技能劳动力	高技能劳动力	低技能劳动力
文化素养	0.5287*** (0.001)	-0.2550*** (0.001)	-0.7322*** (0.002)	-0.5270*** (0.002)	-0.3908*** (0.001)	0.1390*** (0.002)
数字素养	0.9185*** (0.001)	0.7727*** (0.001)	0.1830*** (0.001)	0.0495*** (0.001)	0.6781*** (0.001)	0.6067*** (0.001)

① 花费的时间在平均时间以上为低技能,赋值为 0;花费的时间在平均时间以下为高技能,赋值为 1。

续表 3

变量	2016年					
	ln时工资		ln时工资		ln时工资	
	制造业		流通性服务业		生活性服务业	
	高技能劳动力	低技能劳动力	高技能劳动力	低技能劳动力	高技能劳动力	低技能劳动力
个体特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
企业特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
地区固定效应	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.4760	0.0286	0.2925	0.3779	0.3080	0.3080
N	1274	822	467	317	766	687

注:括号内为基于个人层面的聚类标准误。

(二)城市包容度与数字素养

数字素养提升农民工收入的路径包括提升农民工工作技能的直接机制以及通过社会网络提高其收入的间接机制。除此之外,本文还研究了影响农民工数字素养对收入作用的另一重要因素——城市包容度。城市包容度越高的城市,流动人口长久居住的意愿越高,流动人口落户的比例越高,流动人口就业的稳定性越高(李叶妍,2017)。此时,农民工更有动力提升数字素养和工作技能,以更好地融入当地生活。Florida和Gates(2003)研究表明城市包容度能够提高城市发展效率和创新水平,形成人才集聚效应,外部人力资本溢出会影响农民工劳动生产率,城市公共服务均等化能够降低劳动力市场信息不对称,保护农民工在劳动力市场的权益。因此,在政府倡导包容性增长、增强内生动力的背景下,探究城市包容度对农民工数字素养发挥收入促增效应的影响具有重要意义。本文参照李叶妍(2017)、周颖刚等(2020)的做法,从社会保障和民生性公共服务^①两个维度衡量城市包容度,借鉴Landgraf和Lee(2020)的Logistics PCA^②方法提取一个主成分后,验证城市包容度的外部人力资本溢出效应机制。表4第(1)至(4)列为城市包容度的作用机制,第(1)列和第(3)列是没有纳入城市包容度变量的估计,结果显示数字素养对农民工时工资具有显著的正向影响;第(2)列和第(4)列是纳入城市包容度变量的估计,结果显示2014年和2016年文化素养与城市包容度交互项的系数分别为0.1930和0.0013,验证了城市包容度通过民生性公共服务均等化和集聚效应间接提高农民工收入。

表 4 城市包容度集聚效应与社会网络机制

变量	2014年		2016年		2014年	2016年
	ln时工资		ln时工资		ln时工资	ln时工资
	(1)	(2)	(3)	(4)	(5)	(6)
文化素养	0.4024*** (0.098)	0.06783*** (0.015)	-0.0117*** (0.001)	-0.0455*** (0.007)	0.0948*** (0.000)	-0.0105*** (0.000)
数字素养	0.6504*** (0.070)	0.6092*** (0.064)	0.7382*** (0.009)	0.7177*** (0.034)	0.5593*** (0.000)	0.2641*** (0.000)
文化素养×城市包容度		0.1930*** (0.010)		0.0013*** (0.003)		
数字素养×城市包容度		0.7360*** (0.004)		0.7437*** (0.000)		

① 将调查问卷中“五险一金”纳入社会保障层面指标,将“是否参加政府提供的技能培训”纳入民生性公共服务指标。

② 限于篇幅,碎石图和代码未列示,留存备案。 $M=4,k=1$ 模型最优。

续表 4

变量	2014 年		2016 年		2014 年	2016 年
	ln 时工资		ln 时工资		ln 时工资	ln 时工资
	(1)	(2)	(3)	(4)	(5)	(6)
文化素养×社会网络					0.0899*** (0.000)	0.0058*** (0.000)
数字素养×社会网络					0.02313*** (0.000)	0.1183*** (0.000)
个体特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
企业特征控制变量	Yes	Yes	Yes	Yes	Yes	Yes
地区固定效应	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.2454	0.0644	0.2706	0.0644	0.2407	0.2725
N	5463	5463	4641	4641	5463	4641

注：括号内为稳健标准误。

(三) 社会网络与数字素养

社会网络可能是数字素养提升农民工收入的另一条间接机制。一方面,农民工提高数字素养可以增强其社会网络关系,进而获得更多的能够提高其收入的信息;另一方面,中国是一个关系型社会(Li 和 Guo, 2022),社会网络在日常生活中的作用非常重要,在劳动力市场上,它能够传递关于就业岗位的信息,从而促进劳动力与就业岗位之间的配置,或帮助求职者找到待遇更好的工作岗位,因此如果一个农民工拥有更广泛的社会网络,就更容易匹配到能够足以发挥其数字素养技能的职位,从而提高其收入。本文参照陈斌开和陈思宇(2018)的做法,选取农民工求职途径^①作为社会网络指标,将社会网络与文化素养和数字素养的交互项作为解释变量纳入方程。对比表 4 第(5)列和第(6)列结果,社会网络收入结构效应有所降低,且阻碍数字素养收入结构效应实现,这说明依赖数字网络的社会关系网络逐渐取代了传统宗族网络、亲友网络而成为农民工获取信息、寻找工作的主要渠道。

六、结论与启示

数字经济在提高经济生产率、形成技术和人才集聚效应的同时,也伴随技能需求结构变化和个人机会不均等现象。未来,数字技术必然影响劳动力市场收入分配格局和行业间劳动力要素分配。然而,关于数字素养与文化素养对行业间农民工收入差距变动趋势的研究较少,本文试图从行业异质性的角度考察文化素养与数字素养对农民工收入的作用机制及收入分配效应。研究发现,文化素养与数字素养在不同行业具有不同的替代弹性,文化素养与数字素养在制造业中由“互补效应”过渡到“替代效应”,在流通性服务业中是“替代效应”,在生活性服务业中由“替代效应”过渡到“生产率效应”和“还原效应”的合力,相对收入差距缩小。通过 FFL-OB 静态分解和随机森林决策树模型进一步证明了行业间数字素养与文化素养偏向是造成农民工中低收入群体、高低收入群体和中高收入群体收入差距变化的重要原因。异质性分析结果表明,数字素养扩大了制造业高

^① 本文的农民工求职途径是指传统求职途径,即同乡途径、朋友或者同学途径、亲戚或者家人途径、报纸或者杂志途径、人才市场途径。

技能劳动力与低技能劳动力收入差距,但缩小了流通性服务业和生活性服务业高技能劳动力与低技能劳动力收入差距。机制检验结果表明,城市包容度效应和社会网络效应是数字素养影响农民工收入差距的重要途径。

基于实证结果,本文提出以下政策建议。第一,对于由“互补效应”过渡到“替代效应”的制造业,要注重同时提升低收入群体的文化素养和数字素养,预防数字技术带来收入差距的进一步扩大;针对具有“替代效应”的流通性服务业,需注重对中低收入群体数字信息技能培训和软件指导,跨越数字素养带来的数字鸿沟,缩小高低收入群体的收入差距;针对具有“生产率效应”和“还原效应”合力的生活性服务业,要提高中低收入群体劳动力质量,储备未来高技能劳动力,弥合相对收入差距。第二,在大规模劳动力流动的趋势下,城市包容度对农民工收入结构优化具有重要作用,要促进农民工在医疗、教育、工作环境等方面的社会福利效应,打破行业间户籍壁垒,确保农民工职业流动性。第三,数字经济时代下的收入差距不是持续性和长期性问题,当前要利用“平台效应”提升农民工“互联网社会网络”存量,通过在各地区新建农民工信息化网站,发布有关就业、政府补贴和技能培训等相关信息,打破劳动力市场信息壁垒,加快农民工技能转型,提高农民工数字资本。

参考文献:

1. 柏培文、王亚文:《中国细分行业技能资本替代弹性与技术偏向性》,《经济研究》2023年第3期。
2. 陈斌开、陈思宇:《流动的社会资本——传统宗族文化是否影响移民就业?》,《经济研究》2018年第3期。
3. 贺雪峰:《新一代农民工能否回到农村去》,《经济导刊》2015年第4期。
4. 李叶妍:《中国城市包容度、流动人口与城市发展研究》,社会科学文献出版社2017年版。
5. 梁文泉、陆铭:《后工业化时代的城市:城市规模影响服务业人力资本外部性的微观证据》,《经济研究》2016年第12期。
6. 罗楚亮、滕阳川、李利英:《行业结构、性别歧视与性别工资差距》,《管理世界》2019年第8期。
7. 蒲艳萍、顾冉:《劳动力工资扭曲如何影响企业创新》,《中国工业经济》2019年第7期。
8. 王春超、叶琴:《中国农民工多维贫困的演进——基于收入与教育维度的考察》,《经济研究》2014年第12期。
9. 许竹青、郑风田、陈洁:《“数字鸿沟”还是“信息红利”?信息的有效供给与农民的销售价格——一个微观角度的实证研究》,《经济学(季刊)》2013年第4期。
10. 余玲铮、魏下海、孙中伟、吴春秀:《工业机器人、工作任务与非常规能力溢价——来自制造业“企业—工人”匹配调查的证据》,《管理世界》2021年第1期。
11. 袁方、史清华:《不平等之再检验:可行能力和收入不平等与农民工福利》,《管理世界》2013年第10期。
12. 曾亿武、郭红东、金松青:《电子商务有益于农民增收吗?——来自江苏沭阳的证据》,《中国农村经济》2018年第2期。
13. 张卫东、卜偲琦、彭旭辉:《互联网技能、信息优势与农民工非农就业》,《财经科学》2021年第1期。
14. 周颖刚、蒙莉娜、林雪萍:《城市包容性与劳动力的创业选择——基于流动人口的微观视角》,《财贸经济》2020年第1期。
15. 周云波、田柳、陈岑:《经济发展中的技术创新、技术溢出与行业收入差距演变——对U型假说的理论解释与实证检验》,《管理世界》2017年第11期。
16. Abdi, E.S., Partridge, H., Bruce, C.S., & Watson, J., Skilled Immigrants: A Resettlement Information Literacy Framework. *Journal of Documentation*, Vol.75, No.4, 2019, pp.110-127.
17. Acemoglu, D., & Author, D., Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, Vol.4, 2011, pp.1043-1171.
18. Acemoglu, D., AI and Jobs: Evidence from Online Vacancies. NBER Working Paper, Vol.15, No. 28, 2020, pp.57-79.
19. Andrianarison, F., Davies, V. A. B., & Dessy, S., The Economic Foundations of Institutional Stagnation in Commodity-Exporting Countries. CIRPEE Working Paper, 2010, pp.10-36.
20. Bonfadelli, H., The Internet and Knowledge Gaps: A Theoretical and Empirical Investigation. *European Journal of Communication*,

Vol.17, No.1,2002, pp.65–84.

21.Britz, J.J., To Know or Not to Know: A Moral Reflection on Information Poverty. *Journal of Information Science*, Vol.30, No. 3, 2004, pp.192–204.

22.DiMaggio, P., & Bonikowski, B., Make Money Surfing the Web? The Impact of Internet Use on the Earnings of U.S. Workers. *American Sociological Review*, Vol.73, No.2,2008, pp.227–250.

23.Dobson, T., & Willinsky, J., Digital Literacy. *The Cambridge Handbook of Literacy*, 2009, pp.286–312.

24.Firpo, S. P., Fortin, N. M., & Lemieux, T., Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, Vol. 6, No.2,2018, pp.28–37.

25.Florida, R., & Gates, G., *Technology and Tolerance: The Importance of Diversity to High-Technology Growth*. Emerald Group Publishing Limited, 2003.

26.Gupta, A., Ponticelli, J., & Tesel, A., Technology Adoption and Access to Credit via Mobile Phones. *SSRN Electronic Journal*, Vol.14, No.5,2019, pp.139–201.

27.Hodgson, G.M., What Is Capital? Economists and Sociologists have Changed Its Meaning: Should It Be Changed Back? . *Cambridge Journal of Economics*, Vol.38, 2014, pp.1063–1086.

28.Jensen, R., The Digital Divide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector. *The Quarterly Journal of Economics*, Vol.122, No.3, 2007, pp.879–924.

29.Koning, R., Industrial Change, the Boundary of the Firm, and Racial Employment Segregation. Harvard Business School Working Paper, No.20, 2019, pp.20–69.

30.Lanciotti, M., Adult Literacy: An Essential Component of the CRRF. *Forced Migration Review*, 2019, pp.31–34.

31.Landgraf, A. J., & Lee, Y., Generalized Principal Component Analysis: Projection of Saturated Model Parameters. *Technometrics*, Vol.62, No.4,2020,pp.459–472.

32.Legarra, A., Robert-Granié, C., Croiseau, P., Guillaume, F., & Fritz, S., Improved Lasso for Genomic Selection. *Genetics Research*, Vol.93, No. 1,2011, pp.77–87.

33.Li, X., & Guo, X., Dynamics of Social Capital in Urban China, 1999 to 2014: An Age-Period-Cohort Analysis. *Social Networks*, Vol.68,2022,pp.394–406.

34.Majid, S., Chang, Y. K., Foo, S., Mokhtar, I. A., Theng, Y. L., & Zhang, X., Strengthening Information Literacy Competencies through Incorporating Personal Information Management Skills. *Communications in Computer and Information Science*, Vol.397, No.2,2013,pp.331–337.

35.Mincer, J., *Schooling, Experience and Earnings*. New York: Columbia University Press for the National Bureau of Economic Research, 1974, pp. 7–22.

36.Naik, L., Usage of Electronic Resources, Internet, and Choices of Resources in College Libraries of India: A Study. IGI Global, 2022.

37.Oakleaf, M., Using Rubrics to Assess Information Literacy: An Examination of Methodology and Interrater Reliability. *Journal of the American Society for Information Science and Technology*, Vol.60, No.5,2009, pp.969–983.

38.Plasman, R., Rusinek, M., & Rycx, F., Wages and the Bargaining Regime under Multi-Level Bargaining: Belgium, Denmark and Spain. *European Journal of Industrial Relations*, Vol.13, No.2, 2007, pp.61–180.

39.Raschka, S., *Python Machine Learning*, Packt Publishing, 2015.

40.Rios-Avila, F., RIF: Stata Module to Compute Recentered Influence Functions (RIF) : RIF-Regression and RIF-Decomposition. *Statistical Software Components*, Vol.9, 2020, pp.117–121.

41.Stigler, G. J., The Economics of Information. *Journal of Political Economy*, Vol.69, No.3,1961, pp.213–225.

42.Tuominen, K., Savolainen, R., & Talja, S., Information Literacy as a Sociotechnical Practice. *The Library Quarterly: Information, Community, Policy*, Vol.75, No.3, 2005, pp.329–345.

43.Van de Klundert, Th., Wage Differentials and Employment in a Two-Sector Model with a Dual Labor. *Metroeconomica*, Vol.40, No.3, 1989, pp.235–256.

Common Prosperity in the Era of Technological Change: Digital Literacy and Income Distribution among Rural-to-Urban Migrant Workers

ZHANG Jinhua (Shanghai University of Finance and Economics, 200433)

YOU Jialu (University of Heidelberg, 69117)

Summary: Enhancing digital literacy is essential in the digital era and a key initiative to bridge the digital divide and promote common prosperity. This paper presents a theoretical model of a two-sector production function that incorporates digital literacy and cultural literacy as input factors. It examines the impact of the elasticity of substitution between digital literacy and cultural literacy across industries on the income gap of rural-to-urban migrant workers. The study conducts empirical tests using both macro-city-level data and micro-survey data and employs a random forest model to predict the trend of the income gap among migrant workers in different industries.

The findings are as follows. Firstly, in the manufacturing industry, digital literacy and cultural literacy exhibit a “complementary effect,” whereas the service industry transitions from a “substitution effect” to a “productivity effect.” Secondly, digital literacy significantly influences the income gap of migrant workers in the manufacturing industry. The contribution of digital literacy to the income gap among migrant workers in the manufacturing industry tends to increase, while its role in the distribution and lifestyle service industries tends to diminish. Thirdly, urban inclusiveness helps reduce the income gap, while the influence of traditional social networks is weakening.

The academic contributions of this paper include three main aspects. First, it enriches the literature on the fields of digital economy, skill complementarity, and income inequality. Second, by using the FFL-OB decomposition technique and random forest model, the paper explores the contribution of digital literacy and cultural literacy to income inequality among migrant workers across industries and the trends in income gaps. This approach addresses the limitations of static income decompositions in existing literature, which often suffer from excessive noise and underfitting. Third, the paper discusses the direct and indirect mechanisms and moderating factors of digital literacy, including inter-industry skill structure, social networks, and urban inclusiveness, and their impact on migrant workers’ income.

This paper provides a micro-basis for the solid advance toward common prosperity. Future research can explore the impact of digital literacy on the consumption gap among migrant workers from a consumption perspective.

Keywords: Income Inequality, Rural-to-Urban Migrant Workers, Digital Literacy, Cultural Literacy

JEL: R2, J6, D8

责任编辑:非 同