

## 【劳动经济】

## 产业数字化对性别工资差距的影响研究

张明斗 郭 瑞

**【摘 要】**产业数字化会促进劳动者工资收入增长,并改变工资收入分配格局。根据“生产任务”模型和“肌肉—大脑”假说,产业数字化通过改变脑力劳动与体力劳动的相对价格、促进女性劳动力就业竞争力提升和相对优势发挥等,产生女性偏向的工资增长效应,从而可以缩小性别工资差距。将地区(省份)的产业数字化水平与2014年、2016年和2018年CLDS数据库的微观个体混合截面数据相匹配,分析表明:在样本期间女性劳动者与男性劳动者之间存在显著的工资收入差距;产业数字化能够显著促进劳动者的工资收入增长,该作用对女性劳动者更为显著,从而缩小了性别工资收入差距;产业数字化可以显著缩小低技能劳动群体和经济发展水平较高地区的性别工资差距,但对高技能劳动群体和经济发展水平较低地区的影响不显著;产业数字化对中低工资群体具有显著的促进工资增长和缩小性别工资差距作用,但对高工资群体的影响不显著。因此,各地区(尤其是欠发达地区)应加快推进产业数字化,并优化女性就业结构,充分发挥产业数字化对工资收入分配的改善作用。

**【关键词】**性别工资差距;产业数字化;“生产任务”模型;“肌肉—大脑”假说;劳动技能溢价;脑力劳动;数字技术

**【作者简介】**张明斗(1983—),男,山东济宁人,东北财经大学经济学院教授,博士,主要从事城市经济研究,E-mail:zhangmingdou0537@126.com;郭瑞(1996—),男,黑龙江绥化人,东北财经大学经济学院硕士研究生,主要从事产业经济研究(辽宁 大连 116025)。

**【原文出处】**《西部论坛》(重庆),2023.5.1~17

**【基金项目】**国家社会科学基金重大项目(21ZDA099)。

## 一、引言

随着大数据、人工智能等现代信息技术不断实现突破,数字化已成为驱动经济增长和创造社会价值的重要引擎。产业数字化是指以新一代数字技术为支撑,以数据赋能为主线,以数据为关键要素,对产业链上下游全要素进行数字化转型升级和价值再造的过程。产业的数字化转型和智能化升级不仅带来企业生产方式和经营模式的转变,还会对整个经济结构及市场环境产生重要影响,比如改变劳动力市场的分配格局,在提高劳动者就业质量的同时优化就业结构(蔡昉,2017;王文,2020)<sup>[1-2]</sup>。在传统经济形态下,由于性别歧视

以及不同性别劳动者在劳动能力禀赋上的天然差异,劳动力市场中的性别工资差距(即不同性别劳动者的工资收入存在显著差异)普遍存在。过高的性别工资差距不仅有损整体劳动效率,而且不利于社会公平的实现。那么,产业数字化会对性别工资差距产生怎样的影响?这是值得深入探究的重大课题。

关于性别工资差距的形成及其影响因素,现有文献基于不同的研究视角(如人力资本水平、家庭分工、行业和岗位差异、心理特征以及性别歧视等)进行了较为广泛的探讨,并形成了较为统一的观点。部分学者基于人力资本理论认为,性别工

资差距的产生主要是因为不同性别劳动者在人力资本水平上(如受教育水平、工作经验等)存在差异(Becker, 1965; Mincer, 1974)<sup>[3-4]</sup>;从家庭分工来看,出于兼顾的原因,女性可能会选择薪资报酬相对较低的工作(Becker, 1985)<sup>[5]</sup>;从行业差异来看,性别的行业分割会导致性别工资差距(Katz, 1986)<sup>[6]</sup>;从就业岗位来看,女性就业普遍处于级别较低、晋升难度较高的职位(Albrecht et al., 2003)<sup>[7]</sup>;从心理特征来看,不同性别劳动力在非认知能力上的差异也是造成性别工资差距的重要原因(Heckman et al., 2006)<sup>[8]</sup>。国内学者更多的是从“歧视”角度来分析性别工资差距的成因,认为劳动力市场中的性别歧视越严重,则性别工资差距越大。相关研究发现:我国劳动力市场上存在明显的性别歧视(张丹丹, 2004; 黄志岭等 2008)<sup>[9-10]</sup>,其中,学历低、年纪轻的女性劳动者受到的性别歧视更为严重(葛玉好等, 2011)<sup>[11]</sup>,性别歧视在非国有部门中更严重(亓寿伟等, 2009)<sup>[12]</sup>,不同行业中的性别歧视也存在显著差异(王湘红等, 2016; 罗楚亮等, 2019)<sup>[13-14]</sup>。此外,职位隔离和职位晋升歧视是性别收入差距形成的重要机制,而职位层级内部的性别收入差距主要来源于工资歧视(卿石松等, 2013)<sup>[15]</sup>。

相比之下,由于数字经济发展的时间较短,关于产业数字化对性别工资差距的影响研究目前基本上还停留在理论分析层面。产业数字化主要是利用数字技术对传统产业进行升级改造,提升生产数量和生产效率的过程(肖旭等, 2019)<sup>[16]</sup>,本质上是一种由技术进步驱动的生产力革命,而技术进步通常既具有就业替代效应,又具有就业创造效应(Autor et al., 2003; Acemoglu et al., 2018, 2019)<sup>[17-19]</sup>。随着数字技术的不断发展和广泛应用,一方面中低技能劳动者容易被机器替代(蔡跃洲等, 2019; 张新春等, 2019)<sup>[20-21]</sup>,另一方面也能够通过扩大生产规模、产生新业态等路径创造新的职业类型和就业岗位(杨光等, 2020; 王永钦等,

2020)<sup>[22-23]</sup>。与之类似,技术进步对性别工资差距的影响也可能具有两面性,一些学者对此进行了实证检验。魏下海等(2018)研究发现,生产线升级使性别工资差距得以收敛,但主要是缩小高技能工人的性别工资差距,对低技能工人性别工资差距的缩小作用有限<sup>[24]</sup>;孙早和韩颖(2022)分析表明,人工智能发展可以缩小低技术工业部门的性别工资水平差距,但会扩大高技术工业部门的性别工资差距<sup>[25]</sup>;许健等(2022)研究显示,工业机器人应用有助于缩小性别工资差距,该作用在初始性别工资差距较大、劳动力技能水平较低时以及制造业部门中更为明显<sup>[26]</sup>。然而,对于产业数字化对性别工资差距的影响,还缺乏相关经验证据。

当前,我国正处于产业数字化加速和人口老龄化加剧的发展阶段,明确产业数字化对性别工资差距的影响,进而采取相应措施缩小性别工资差距,既可以改善劳动力市场结构,提高劳动生产率,促进社会公平,也有利于加快产业数字化转型,推动经济高质量可持续发展。有鉴于此,本文在已有研究的基础上,建立一个产业数字化影响性别工资差距的理论模型,并基于工业数字化和服务业数字化两个维度测度地区产业数字化水平,与中国劳动力动态调查(China Labor-force Dynamics Survey, CLDS)2014年、2016年和2018年的微观数据进行匹配,实证检验产业数字化对性别工资差距的影响。

本文的边际贡献主要在于:第一,结合“生产任务”模型和“肌肉一大脑”假说,构建了一个分析产业数字化影响性别工资差距的理论模型;第二,通过引入产业数字化与性别虚拟变量的交乘项进行实证检验,为相关研究提供了思路借鉴和方法参考,并为产业数字化缩小性别工资差距提供了经验证据;第三,从劳动者技能水平、地区经济发展水平两个方面分析产业数字化影响性别工资差距的异质性,并利用无条件分位数回归和分解探究在不同工资水平下产业数字化影响性别工资差

距的作用效果,有利于深入认识产业数字化的就业效应和性别工资差距的形成原因,有助于采取针对性措施有效缩小性别工资差距。

## 二、理论模型与研究假设

本文借鉴 Autor 等(2003)的“生产任务”模型<sup>[17]</sup>,假定最终品由一系列生产任务  $y(i)$  加总而成,设定总体生产函数如式(1)所示:

$$Y = \left[ \int_0^1 y(i)^{\frac{\sigma-1}{\sigma}} di \right]^{\frac{\sigma}{\sigma-1}} \quad (1)$$

其中,  $\sigma$  代表不同生产任务间的替代弹性,  $\sigma \in (0, \infty)$ 。假定生产存在阈值  $I$  和  $S$ , 满足  $0 < I < S < 1$ , 生产任务  $y(i)$  的生产情况如下: 当  $i \leq I$  时, 生产任务实现数字化, 要素投入为数字化资本  $Q$ ; 当  $I < i \leq S$  时, 生产任务为低技能生产任务, 要素投入为传统资本  $K$  和体力劳动  $L$ ; 当  $S < i$  时, 生产任务为高技能生产任务, 要素投入为脑力劳动  $H$ 。设定任务  $y(i)$  的生产函数如式(2)所示:

$$y(i) = \begin{cases} Q(i), & i \leq I \\ K(i)^\alpha L(i)^{1-\alpha}, & I < i \leq S \\ AH(i), & S < i \end{cases} \quad (2)$$

式(2)中,  $A$  为偏向技术进步率, 代表产业数字化发展对脑力劳动的效率改进, 设定  $A$  是数字化资本  $Q$  的增函数,  $\partial A / \partial Q > 0$ , 即数字化资本存量越大, 产业数字化对脑力劳动的效率改进就越强。同时, 为简化运算, 借鉴孙早和韩颖(2022)的做法<sup>[25]</sup>, 假定数字化资本投入在数字化生产任务中平均分配, 传统资本投入和体力劳动投入在低技能生产任务中平均分配, 脑力劳动投入在高技能生产任务中平均分配, 则式(1)可以改写为式(3):

$$Y = [I^{\frac{1}{\sigma}} Q^{\frac{\sigma-1}{\sigma}} + (S-I)^{\frac{1}{\sigma}} (K^\alpha L^{1-\alpha})^{\frac{\sigma-1}{\sigma}} + (1-S)^{\frac{1}{\sigma}} (AH)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \quad (3)$$

将最终品单位价格标准化为 1, 在完全竞争的市场下, 市场出清时体力劳动和脑力劳动的要素价格等价于每单位劳动投入的边际产品, 如式(4)和式(5)所示。其中,  $\omega_L$  和  $\omega_H$  分别代表体力劳动和

脑力劳动的要素价格。

$$\omega_L = (1-\alpha)(S-I)^{\frac{1}{\sigma}} K^{\frac{\alpha(\sigma-1)}{\sigma}} L^{\frac{\alpha-\alpha\sigma-1}{\sigma}} Y^{\frac{1}{\sigma}} \quad (4)$$

$$\omega_H = (1-S)^{\frac{1}{\sigma}} A^{\frac{\sigma-1}{\sigma}} H^{-\frac{1}{\sigma}} Y^{\frac{1}{\sigma}} \quad (5)$$

产业数字化对生产的影响可以分为两个方面: 一方面, 数字化技术与传统产业融合将促使更多低技能生产任务转向数字化生产模式, 生产阈值  $I$  升高; 另一方面, 产业数字化会提高市场的数字化资本水平  $Q$ 。从式(4)和式(5)来看, 数字化生产对低技能生产任务的替代会对体力劳动要素价格产生负面冲击, 而由数字化资本积累带来的效率改进则会进一步提高脑力劳动的要素价格。定义脑力劳动技能溢价函数  $\varphi$  如式(6)所示:

$$\varphi = \frac{\omega_H}{\omega_L} = \frac{(1-S)^{\frac{1}{\sigma}} A^{\frac{\sigma-1}{\sigma}} H^{-\frac{1}{\sigma}}}{(1-\alpha)(S-I)^{\frac{1}{\sigma}} K^{\frac{\alpha(\sigma-1)}{\sigma}} L^{\frac{\alpha-\alpha\sigma-1}{\sigma}}} \quad (6)$$

生产阈值  $I$  升高将导致低技能生产任务数量  $(S-I)$  变少, 同时随着数字化资本  $Q$  不断增长, 偏向技术进步率  $A$  也随之提高。将技能溢价对生产阈值  $I$  和数字化资本  $Q$  分别求偏导可得式(7)和式(8):

$$\frac{\partial \varphi}{\partial I} = \frac{(1-S)^{\frac{1}{\sigma}} A^{\frac{\sigma-1}{\sigma}} H^{-\frac{1}{\sigma}}}{\sigma(1-\alpha)(S-I)^{\frac{1+\sigma}{\sigma}} K^{\frac{\alpha(\sigma-1)}{\sigma}} L^{\frac{\alpha-\alpha\sigma-1}{\sigma}}} > 0 \quad (7)$$

$$\frac{\partial \varphi}{\partial Q} = \frac{(\sigma-1)(1-S)^{\frac{1}{\sigma}} H^{-\frac{1}{\sigma}}}{\sigma A^{\frac{1}{\sigma}} (1-\alpha)(S-I)^{\frac{1}{\sigma}} K^{\frac{\alpha(\sigma-1)}{\sigma}} L^{\frac{\alpha-\alpha\sigma-1}{\sigma}}} \cdot \frac{\partial A}{\partial Q} > 0 \quad (8)$$

由式(7)和式(8)可知, 无论是数字化生产范围扩大还是数字化资本存量增加, 都将提高脑力劳动相对体力劳动的技能溢价, 进而提高劳动力市场对高技能劳动力的相对需求。

为进一步分析产业数字化对性别工资差距的影响, 本文基于 Welch(2000)提出的“肌肉—大脑”假说<sup>[27]</sup>, 并借鉴 Ge 等(2020)的研究思路进行如下设定<sup>[28]</sup>: 在其他条件一致的情况下, 男性劳动力和女性劳动力在脑力劳动方面具有相同的能力, 但是相对女性劳动力而言, 男性劳动力可以从事更多的体力劳动, 如式(9)和式(10)所示。其中,  $L'$  和



$L^f$ 分别代表由女性劳动力和男性劳动力提供的体力劳动, $H^f$ 和 $H^m$ 分别代表由女性劳动力和男性劳动力提供的脑力劳动。

$$L^f < L^m \quad (9)$$

$$H^f = H^m \quad (10)$$

不同性别劳动力的工资水平由劳动要素价格和个体的劳动禀赋共同决定,决定方程如式(11)和式(12)所示,其中, $W_f$ 和 $W_m$ 分别代表女性劳动力和男性劳动力的工资收入。

$$W_f = \omega_L L^f + \omega_H H^f \quad (11)$$

$$W_m = \omega_L L^m + \omega_H H^m \quad (12)$$

定义性别工资差距 $\pi$ 为式(13)所示的函数:

$$\pi = \frac{W_m}{W_f} = \frac{\omega_L L^m + \omega_H H^m}{\omega_L L^f + \omega_H H^f} = \frac{L^m + \varphi H^m}{L^f + \varphi H^f} \quad (13)$$

结合式(7)和式(8),由性别工资差距 $\pi$ 对生产阈值 $I$ 和数字化资本 $Q$ 分别求偏导可得:

$$\frac{\partial \pi}{\partial I} = \frac{H^m L^f - H^f L^m}{(L^f + \varphi H^f)^2} \cdot \frac{\partial \varphi}{\partial I} < 0 \quad (14)$$

$$\frac{\partial \pi}{\partial Q} = \frac{H^m L^f - H^f L^m}{(L^f + \varphi H^f)^2} \cdot \frac{\partial \varphi}{\partial Q} < 0 \quad (15)$$

在满足上述假设的条件下,由式(14)和式(15)可知,生产阈值 $I$ 和数字化资本 $Q$ 提高会缩小不同性别劳动力的工资收入差距。以上推导证明,产业数字化可以缩小性别工资差距。

一般来讲,生产力的提高伴随着劳动者收入的增长。产业数字化在带来生产方式革新和生产进步的同时,也会促进劳动者的工资收入增长,这种促进作用的性别差异则会改变性别工资差距。本文认为,产业数字化将提高女性劳动力的就业竞争力,对女性劳动者工资收入增长的促进作用比男性更大,进而有利于缩小性别工资差距。具体来说,产业数字化促进劳动者工资收入增长的女性偏向性主要体现在两个方面:

一方面,产业数字化改变了脑力劳动与体力劳动之间的相对价格(Autor et al., 2003)<sup>[17]</sup>。数字技术的应用使得许多传统的体力劳动过程转变为

自动化、数字化的生产过程,人工智能和工业机器人等新技术能够取代部分传统体力劳动,进而降低了对体力劳动的需求。与此同时,产业数字化显著提高了生产效率,企业通过数字技术能够更有效地利用资源和人力,降低信息传递和获取的成本,有助于推动人与人之间、人与机器之间更紧密地合作,从而增加对脑力劳动的需求。因此,随着产业数字化的推进,脑力劳动的相对价格上升,这会弱化女性劳动者在体力劳动能力禀赋上的天然相对弱势对其劳动报酬的不利影响(Goldin et al., 2006)<sup>[29]</sup>,进而缩小性别工资差距。

另一方面,产业数字化的就业创造效应为女性劳动力提供了更多的就业岗位和更便捷高效包容的就业环境,并有助于充分发挥女性劳动力的相对优势(戚聿东等,2020)<sup>[30]</sup>。数字技术的应用推动了弹性工作时间和远程工作的普及,许多工作岗位允许员工利用碎片化时间远程办公,这使得那些因为照料家庭或其他原因而受到限制的女性劳动力可以根据家庭和自身的需要,在特定的时间段内自由选择办公地点,从而更好地平衡家庭和职业生活(陈璐等,2016)<sup>[31]</sup>。产业数字化催生了许多在线平台经济,为女性劳动力提供了更多的灵活就业机会,如网上销售、社交媒体营销和远程咨询等(李建奇,2022)<sup>[32]</sup>。同时,人工智能、大数据等数字技术在招聘网站的应用改变了求职方式,有效降低了女性劳动力获取就业信息和就业机

会的搜寻成本(仇化等,2023)<sup>[33]</sup>。综上所述,产业数字化主要通过提高脑力劳动的相对价格和强化女性劳动力的就业优势等,对女性劳动者产生比男性劳动者更强的工资收入增长效应,进而缩小性别工资差距。当然,这种性别工资差距缩小效应在不同的情形下可能具有异质性表现。这里主要从劳动者技能水平和地区经济发展水平两个维度进行简要探讨:从劳动者的技能水平来看,产业数字化会改变劳动力市场的技能需求结构。随着传统产业与数字技术深度融

合,低技能劳动力更容易被取代,而高技能劳动力则可以从数字化带来的生产效率改进中受益,因而产业数字化对不同技能水平劳动群体性别工资差距的影响可能不同。高技能劳动力主要提供脑力劳动,低技能劳动力主要提供体力劳动,而不同性别劳动者的劳动能力禀赋差异主要表现在体力劳动上,因此相较于高技能劳动群体,产业数字化对低技能劳动群体性别工资差距的缩小作用更为明显。此外,数字化生产主要替代的是低技能体力劳动,而这部分劳动主要由男性劳动力提供,由此造成对男性低技能劳动力的挤出,并通过降低男性劳动力的技能溢价缩小低技能劳动群体的性别工资差距。从劳动者所在地区的经济发展水平来看,在经济发展水平较高的地区,产业数字化的程度较高、范围较广,对社会经济的影响也较大;而在经济发展水平较低的地区,产业数字化发展相对滞后,所产生的影响也相对有限。

基于以上分析,本文提出如下假说:地区产业数字化水平提高有利于缩小性别工资差距(H1);相比于高技能劳动群体和经济发展水平较低地区,产业数字化的性别工资差距缩小效应在低技能劳动群体中和经济发展水平较高地区更为显著(H2)。

### 三、实证研究设计

#### 1. 模型设定与变量选择

为检验产业数字化对性别工资差距的影响,本文基于Mincer(1974)提出的工资收入方程<sup>[4]</sup>,借鉴魏下海等(2018)、孙早和韩颖(2022)的研究思路<sup>[24-25]</sup>,构建如下基准计量模型:

$$\text{模型(1): } \ln W_{i,p} = \alpha_0 + \alpha_1 \text{dig}_p + \alpha_2 \text{gender}_i + \alpha_3 Z_{i,p} + v_i + \zeta_p + \mu_d + \varepsilon$$

$$\text{模型(2): } \ln W_{i,p} = \beta_0 + \beta_1 \text{dig}_p + \beta_2 \text{gender}_i + \beta_3 \text{gender}_i \times \text{dig}_p + \beta_4 Z_{i,p} + v_i + \zeta_p + \mu_d + \varepsilon$$

其中,下标*i*、*p*、*t*、*d*分别代表劳动者个体、地区(省份)、年份、行业,*v<sub>i</sub>*、*ζ<sub>p</sub>*、*μ<sub>d</sub>*和*ε*分别表示年份固定效应、省份固定效应、行业固定效应和随机误差项,*Z<sub>i,p</sub>*为控制变量。被解释变量为“工资收入”(ln*W<sub>i,p</sub>*),采用地区*p*中劳动者*i*的年工资收入(取自然对数)来衡量。模型(1)中核心解释变量有2个:一是“产业数字化”(dig<sub>*p*</sub>),为劳动者*i*所在地区*p*的产业数字化水平,若其回归系数α<sub>1</sub>显著为正,则表明地区产业数字化促进了劳动者的工资收入增长;二是“性别”(gender<sub>*i*</sub>),为劳动者*i*是否为女性的虚拟变量(男性赋值为0,女性赋值为1),若其回归系数α<sub>2</sub>显著为负,则表明女性劳动者的工资收入显著低于男性劳动者,即存在性别工资差距。模型(2)在模型(1)的基础上加入“产业数字化”与“性别”的交乘项(gender<sub>*i*</sub>×dig<sub>*p*</sub>),若其回归系数β<sub>3</sub>显著为正,则表明“产业数字化”可以弱化“性别”对“工资收入”的负向作用,即地区产业数字化水平的提高能够缩小性别工资差距。

对“产业数字化”的测度借鉴王军等(2021)的方法<sup>[34]</sup>,并考虑到本文所用微观调查数据对应年份的产业数字化主要集中在工业和服务业,从工业数字化和服务业数字化两个维度构建地区产业数字化水平的综合评价指标体系(见表1),采用熵值法计算各项指标的权重(郭显光,1998)<sup>[35]</sup>,从而计算得到样本地区(省份)的产业数字化水平。

表1 地区产业数字化水平评价指标体系

| 一级指标  | 二级指标   | 变量             | 指标单位 | 指标属性 | 权重    |
|-------|--------|----------------|------|------|-------|
| 产业数字化 | 工业数字化  | 工业机器人安装密度      | 台/万人 | +    | 0.400 |
|       |        | 两化融合发展指数       | /    | +    | 0.090 |
|       | 服务业数字化 | 电子商务销售额占生产总值比重 | %    | +    | 0.214 |
|       |        | 数字普惠金融指数       | /    | +    | 0.166 |
|       |        | 电子商务交易活动企业比重   | %    | +    | 0.130 |

参照许健等(2022)的研究<sup>[26]</sup>,选取以下控制变量:一是反映劳动者个体的人口学基本特征的“年龄”“婚姻状态”“健康状况”,二是反映劳动者个体的社会经济特征的“政治面貌”“户口性质”“专业证书”“受教育水平”“单位性质”等,三是反映劳动者所在城市经济发展水平的“人均GDP”和“平均工资”。此外,考虑到劳动者的年龄可能对其劳动收入具有非线性影响,进一步控制了年龄的平方项。各变量的测度方法见表2。

2. 样本选择与数据处理

本文所使用的劳动者个体数据来自2014年、2016年和2018年的“中国劳动力动态调查”(CLDS),样本覆盖29个省份(不包括西藏、海南和港澳台地区)。将三个年份的数据组合成为混合截面数据<sup>①</sup>,并进行如下样本筛选:女性年龄限定为18-55岁,男性年龄限定为18-60岁,保留非农就业样本,删除在校生的失去劳动能力人员,同时对主要变量存在缺失值和异常值的样本予以剔除。基于CLDS的数据特征,个体的年工资额为所有工资、奖金和补贴的总和,工业机器人安装密度

根据国际机器人联合会(IFR)公布的中国各行业工业机器人安装量进行计算,两化融合发展指数来自相应年度的《中国信息化与工业化融合发展水平评估蓝皮书》,其余宏观数据来源于《中国统计年鉴》、北京大学数字普惠金融指数、国泰安数据库以及部分省级统计年鉴。对所有连续变量进行上下1%的缩尾处理,所有以货币单位衡量的变量均使用地区生产总值平减指数折算为2013年不变价格。表2报告了主要变量的定义和描述性统计结果。

四、实证检验结果分析

1. 基准模型回归与内生性处理

表3的(1)(3)(5)列为基准模型(1)的回归结果,(2)(4)(6)列为基准模型(2)的回归结果;(1)(2)列仅控制了年份、省份和行业固定效应,(3)(4)列加入劳动者人口学基本特征控制变量,(5)(6)列进一步加入其他控制变量。逐步回归结果显示,在模型(1)中,“产业数字化”的回归系数显著为正,表明劳动者所在地区的产业数字化水平提高具有显著的工资增长效应;“性别”的回归系数显著为负,表明在样

表2 主要变量定义和描述性统计结果

| 变量    | 变量定义          | 总体    |        |        | 男性   |        |        | 女性    |        |        |
|-------|---------------|-------|--------|--------|------|--------|--------|-------|--------|--------|
|       |               | 样本量   | 均值     | 标准差    | 样本量  | 均值     | 标准差    | 样本量   | 均值     | 标准差    |
| 产业数字化 | 省份产业数字化水平     | 20474 | 0.287  | 0.169  | 9022 | 0.286  | 0.170  | 11452 | 0.287  | 0.168  |
| 工资收入  | 年工资额的自然对数值    | 14495 | 10.332 | 0.802  | 7973 | 10.466 | 0.781  | 6522  | 10.167 | 0.796  |
| 年龄    | 受访者年龄         | 20474 | 40.403 | 11.220 | 9022 | 39.211 | 10.350 | 11452 | 41.342 | 11.775 |
| 受教育水平 | 受访者受教育年限      | 20474 | 10.389 | 3.951  | 9022 | 10.905 | 3.521  | 11452 | 9.982  | 4.215  |
| 政治面貌  | 党员=1;其他=0     | 20474 | 0.106  | 0.308  | 9022 | 0.139  | 0.346  | 11452 | 0.080  | 0.271  |
| 健康状况  | 1~5,数值越大越健康   | 20474 | 2.760  | 0.931  | 9022 | 2.843  | 0.907  | 11452 | 2.694  | 0.944  |
| 婚姻状态  | 已婚=1;未婚、丧偶=0  | 20474 | 0.836  | 0.371  | 9022 | 0.802  | 0.398  | 11452 | 0.862  | 0.345  |
| 单位性质  | 国营、集体=1;其他=0  | 14495 | 0.315  | 0.464  | 7973 | 0.322  | 0.467  | 6522  | 0.306  | 0.461  |
| 户口性质  | 城市户口=1;农村户口=0 | 20474 | 0.449  | 0.497  | 9022 | 0.433  | 0.495  | 11452 | 0.462  | 0.499  |
| 专业证书  | 有专业证书=1;没有=0  | 20474 | 0.225  | 0.417  | 9022 | 0.264  | 0.441  | 11452 | 0.194  | 0.395  |
| 平均工资  | 城市年平均工资的自然对数值 | 20474 | 10.931 | 0.245  | 9022 | 10.922 | 0.237  | 11452 | 10.938 | 0.251  |
| 人均GDP | 城市人均GDP的自然对数值 | 20474 | 10.916 | 0.593  | 9022 | 10.897 | 0.598  | 11452 | 10.931 | 0.589  |

表3 基准模型回归结果

| 变量             | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 产业数字化×性别       |                      | 0.194*<br>(0.102)    |                      | 0.209**<br>(0.093)   |                      | 0.223***<br>(0.078)  |
| 产业数字化          | 0.714***<br>(0.109)  | 0.622***<br>(0.120)  | 0.631***<br>(0.127)  | 0.532***<br>(0.138)  | 0.588***<br>(0.109)  | 0.482***<br>(0.118)  |
| 性别             | -0.324***<br>(0.022) | -0.382***<br>(0.029) | -0.334***<br>(0.023) | -0.396***<br>(0.030) | -0.304***<br>(0.017) | -0.371***<br>(0.025) |
| 年龄             |                      |                      | 0.106***<br>(0.007)  | 0.106***<br>(0.007)  | 0.089***<br>(0.007)  | 0.089***<br>(0.007)  |
| 年龄平方           |                      |                      | -0.144***<br>(0.011) | -0.144***<br>(0.011) | -0.116***<br>(0.010) | -0.116***<br>(0.010) |
| 婚姻状态           |                      |                      | 0.075***<br>(0.019)  | 0.075***<br>(0.019)  | 0.127***<br>(0.022)  | 0.127***<br>(0.022)  |
| 健康水平           |                      |                      | 0.080***<br>(0.011)  | 0.080***<br>(0.011)  | 0.075***<br>(0.008)  | 0.075***<br>(0.008)  |
| 政治面貌           |                      |                      |                      |                      | 0.095***<br>(0.031)  | 0.094***<br>(0.031)  |
| 户口性质           |                      |                      |                      |                      | 0.140***<br>(0.029)  | 0.141***<br>(0.029)  |
| 专业证书           |                      |                      |                      |                      | 0.104***<br>(0.022)  | 0.105***<br>(0.022)  |
| 受教育水平          |                      |                      |                      |                      | 0.053***<br>(0.003)  | 0.053***<br>(0.003)  |
| 单位性质           |                      |                      |                      |                      | -0.001<br>(0.019)    | -0.001<br>(0.019)    |
| 人均GDP          |                      |                      |                      |                      | 0.174***<br>(0.031)  | 0.173***<br>(0.031)  |
| 平均工资           |                      |                      |                      |                      | 0.190**<br>(0.091)   | 0.191**<br>(0.091)   |
| 年份固定效应         | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| 省份固定效应         | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| 行业固定效应         | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| R <sup>2</sup> | 0.184                | 0.185                | 0.229                | 0.229                | 0.333                | 0.333                |
| 样本量            | 14495                | 14495                | 14495                | 14495                | 14495                | 14495                |

注:括号内的数值为省份层面的聚类稳健标准误,\*\*\*、\*\*、\*分别表示在1%、5%、10%的水平上显著,下表同。

本期间样本地区存在性别工资差距。在模型(2)中,“产业数字化”和“性别”的回归系数依然分别显著为正和负,且“产业数字化×性别”的回归系数显著为正,表明地区产业数字化水平的提高有利于缩小性别工资差距,假说H1得到验证。

基准模型可能存在反向因果关系和遗漏变量等内生性问题,参考黄群慧等(2019)的思路<sup>[36]</sup>,采用2SLS工具变量法来缓解内生性问题。基于1984年各省份邮电业务量构造“产业数字化”的工具变量,由于本文有三个年度的数据,为赋予工具



变量年度变异性,参照赵涛等(2020)的做法<sup>[37]</sup>,将上一年全国互联网用户数与1984年邮电业务总量的交乘项作为该年该省份产业数字化水平的工具变量。第一阶段的检验结果显示(限于篇幅,具体结果略,备索),工具变量与“产业数字化”显著正相关;第二阶段的估计结果见表4,工具变量通过了识别不足检验和弱工具变量检验,说明工具变量有效,“产业数字化×性别”的回归系数仍显著为正,表明在缓解内生性问题后,产业数字化能够缩小性别工资差距的结论依然成立。

2. 稳健性检验

为进一步验证基准模型回归结果的可靠性,进行以下稳健性检验:

(1)Heckman 两步法。本文采用微观调查数据和宏观数据相结合的方法进行计量回归,而在微观调查中可能存在个体工资水平不可观测的情况,造成样本选择偏误和回归结果有偏(李宏兵等,2014)<sup>[38]</sup>。对此,采用 Heckman 两步法进行样本选择纠正,并与 OLS 回归结果进行对比。首先,对

所有样本进行 Probit 估计: $P_{it}=\delta Q+\varepsilon$ 。其中, $P_{it}$ 为个体工资是否可观测变量(个体工资可观测赋值为1,个体工资不可观测赋值为0),协变量  $Q$  包括“性别”“年龄”“年龄平方”“婚姻状态”“健康状况”“政治面貌”“户口性质”“受教育水平”“人均GDP”“平均工资”等, $\varepsilon$ 为随机误差项。然后,根据估计结果求得逆米尔斯率(imr)。最后,将“逆米尔斯率”作为控制变量加入基准模型进行检验,回归结果见表5。在加入全部控制变量的(5)(6)列中,“逆米尔斯率”的回归系数不显著,且核心解释变量的回归系数和显著性变化不大,表明基准模型回归中不存在严重的样本选择性偏误问题(孙早等,2022)<sup>[25]</sup>,分析结果是可信的。

(2)更换计量模型。基准模型通过引入性别与产业数字化水平的交乘项来验证产业数字化是否缩小了性别工资差距,该模型允许产业数字化水平在不同组别存在差异,但假定其他控制变量的系数不随性别发生变化。为避免其他控制变量系数也随性别发生变化对回归结果产生影响,将样

表 4 工具变量法第二阶段回归结果

| 变量                                                    | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
|-------------------------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 产业数字化×性别                                              |                      | 0.622***<br>(0.157)  |                      | 0.664***<br>(0.154)  |                      | 0.669***<br>(0.141)  |
| 产业数字化                                                 | 1.562***<br>(0.334)  | 1.156***<br>(0.340)  | 1.507***<br>(0.326)  | 1.073***<br>(0.333)  | 1.153***<br>(0.304)  | 0.714**<br>(0.310)   |
| 性别                                                    | -0.324***<br>(0.013) | -0.508***<br>(0.049) | -0.333***<br>(0.013) | -0.531***<br>(0.048) | -0.304***<br>(0.012) | -0.503***<br>(0.044) |
| 人口学特征控制变量                                             | 未控制                  | 未控制                  | 控制                   | 控制                   | 控制                   | 控制                   |
| 其他控制变量                                                | 未控制                  | 未控制                  | 未控制                  | 未控制                  | 控制                   | 控制                   |
| 固定效应                                                  | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| R <sup>2</sup>                                        | 0.041                | 0.040                | 0.093                | 0.092                | 0.216                | 0.215                |
| 样本量                                                   | 14495                | 14495                | 14495                | 14495                | 14495                | 14495                |
| Kleibergen-Paap rk LM 统计量(中括号内为 p 值)                  | 2379.938<br>[0.000]  | 2219.334<br>[0.000]  | 2380.666<br>[0.000]  | 2218.890<br>[0.000]  | 2318.451<br>[0.000]  | 2184.574<br>[0.000]  |
| Kleibergen-Paap rk Wald F 统计量(中括号内为 Stock-Yogo 检验临界值) | 1478.834<br>[16.380] | 695.055<br>[7.030]   | 1476.388<br>[16.380] | 694.097<br>[7.030]   | 1420.295<br>[16.380] | 673.303<br>[7.030]   |

注:限于篇幅,控制变量估计结果略,固定效应包括年份、省份和行业3种固定效应,下表同。



表 5

Heckman 两步法检验结果

| 变量             | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 产业数字化×性别       |                      | 0.228**<br>(0.084)   |                      | 0.180**<br>(0.074)   |                      | 0.223***<br>(0.078)  |
| 产业数字化          | 0.689***<br>(0.125)  | 0.581***<br>(0.131)  | 0.643***<br>(0.135)  | 0.558***<br>(0.141)  | 0.588***<br>(0.109)  | 0.482***<br>(0.118)  |
| 性别             | -0.040*<br>(0.022)   | -0.107***<br>(0.032) | 0.051*<br>(0.029)    | -0.003<br>(0.037)    | -0.310***<br>(0.033) | -0.378***<br>(0.035) |
| 逆米尔斯率          | -0.853***<br>(0.049) | -0.855***<br>(0.049) | -1.129***<br>(0.058) | -1.127***<br>(0.058) | 0.018<br>(0.073)     | 0.022<br>(0.071)     |
| 人口学特征控制变量      | 未控制                  | 未控制                  | 控制                   | 控制                   | 控制                   | 控制                   |
| 其他控制变量         | 未控制                  | 未控制                  | 未控制                  | 未控制                  | 控制                   | 控制                   |
| 固定效应           | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| R <sup>2</sup> | 0.245                | 0.246                | 0.267                | 0.267                | 0.333                | 0.333                |
| 样本量            | 14495                | 14495                | 14495                | 14495                | 14495                | 14495                |

本划分为“女性”和“男性”两个子样本,分别检验“产业数字化”对“工资收入”的影响,比较地区产业数字化水平提高对不同性别劳动者工资水平的提升效果,检验结果如表6所示。“产业数字化”的回归系数在“女性”子样本中显著大于“男性”子样本,表明地区产业数字化水平提高对女性劳动者工资收入增长的促进作用显著大于对男性劳动者工资收入增长的促进作用,从而有利于缩小性别工资差距。

(3)更换被解释变量。基准模型中的被解释变

量“工资收入”采用劳动者的年工资额进行衡量,CLDS数据库中还调查了劳动者每周平均工作小时数,据此可以计算得到劳动者的小时工资收入,将其自然对数值作为被解释变量重新进行模型检验,回归结果见表7。逐步加入控制变量后,“产业数字化×性别”的回归系数显著为正,进一步表明基准模型的回归结果具有稳健性。

### 3.异质性分析

根据前文理论分析,进行劳动者技能水平和地区经济发展水平两个方面的异质性检验:一是

表 6

按性别分组回归结果

| 变量             | 女性                  | 男性                  | 女性                  | 男性                  | 女性                  | 男性                  |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| 产业数字化          | 0.759***<br>(0.146) | 0.668***<br>(0.127) | 0.708***<br>(0.153) | 0.537***<br>(0.145) | 0.648***<br>(0.147) | 0.520***<br>(0.121) |
| 人口学特征控制变量      | 未控制                 | 未控制                 | 控制                  | 控制                  | 控制                  | 控制                  |
| 其他控制变量         | 未控制                 | 未控制                 | 未控制                 | 未控制                 | 控制                  | 控制                  |
| 固定效应           | 控制                  | 控制                  | 控制                  | 控制                  | 控制                  | 控制                  |
| R <sup>2</sup> | 0.203               | 0.132               | 0.245               | 0.191               | 0.362               | 0.282               |
| 样本量            | 6522                | 7973                | 6522                | 7973                | 6522                | 7973                |
| 组间系数差异 P 值     | 0.067               |                     | 0.022               |                     | 0.024               |                     |

注:组间系数差异 P 值根据 Chow 检验结果计算得到。

表 7

更换被解释变量回归结果

| 变量             | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 产业数字化×性别       |                      | 0.206<br>(0.123)     |                      | 0.218*<br>(0.115)    |                      | 0.238**<br>(0.098)   |
| 产业数字化          | 0.744***<br>(0.184)  | 0.647***<br>(0.202)  | 0.663***<br>(0.200)  | 0.561**<br>(0.215)   | 0.583***<br>(0.198)  | 0.471**<br>(0.212)   |
| 性别             | -0.265***<br>(0.024) | -0.326***<br>(0.034) | -0.272***<br>(0.024) | -0.336***<br>(0.035) | -0.239***<br>(0.017) | -0.310***<br>(0.030) |
| 人口学特征控制变量      | 未控制                  | 未控制                  | 控制                   | 控制                   | 控制                   | 控制                   |
| 其他控制变量         | 未控制                  | 未控制                  | 未控制                  | 未控制                  | 控制                   | 控制                   |
| 固定效应           | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| R <sup>2</sup> | 0.151                | 0.152                | 0.175                | 0.175                | 0.271                | 0.272                |
| 样本量            | 14378                | 14378                | 14378                | 14378                | 14378                | 14378                |

借鉴李红阳和邵敏(2017)的方法<sup>[39]</sup>,采用最高学历来衡量劳动者的技能水平,将具有大专及以上学历的个体划为“高技能劳动群体”子样本,高中及以下学历的个体划为“低技能劳动群体”子样本,分别进行模型检验,回归结果见表8。二是参照张军扩等(2019)、王思博和庄贵阳(2023)的研究<sup>[40-41]</sup>,基于人均GDP进行区域划分,将人均GDP在7000美元以下、7000~10000美元、10000美元以上的省份分别划归为“较低发展水平地区”“中等发展水平地区”“较高发展水平地区”3个子样本<sup>②</sup>,分别进行模型检验,回归结果见表9。在“低技能劳动群体”和“高发展水平地区”中,“产业数字化×性别”的回归系数显著为正;在“高技能劳动群体”和“中等发展水平地区”中,“产业数字化×性别”的回归系数为正但不显著;在“较低发展水平地区”中“产业数字化×性别”的回归系数为负但不显著。

上述结果表明,产业数字化水平提高可以显著缩小低技能劳动群体和经济发展水平较高地区的性别工资差距,但对高技能劳动群体和经济发展水平较低地区的性别工资差距没有显著影响,假说H2得到验证。

4. 进一步分析:工资水平的影响

本文验证了产业数字化的性别工资差距缩小效应及其在劳动者技能水平和地区经济发展水平方面的异质性,但忽略了劳动者自身工资水平的影响。在不同的工资水平下,产业数字化对劳动者工资收入以及性别工资差距的影响程度是否存在差异?对此,本文进一步采用再中心化影响函数(RIF)的无条件分位数回归和分解方法进行分析。首先对样本工资水平分布的不同分位点进行回归,然后在经典Oaxaca-Blinder分解的基础上分解出不同分位点下单个协变量对总体效应的具体

表 8

劳动者技能水平异质性分析结果

| 变量             | 低技能劳动群体              |                      | 高技能劳动群体              |                      |
|----------------|----------------------|----------------------|----------------------|----------------------|
|                | 0.245***             |                      | 0.191                |                      |
| 产业数字化×性别       |                      | (0.067)              |                      | (0.132)              |
| 产业数字化          | 0.620***<br>(0.117)  | 0.508***<br>(0.127)  | 0.535***<br>(0.194)  | 0.438**<br>(0.192)   |
| 性别             | -0.352***<br>(0.017) | -0.426***<br>(0.033) | -0.207***<br>(0.027) | -0.262***<br>(0.042) |
| 控制变量           | 控制                   | 控制                   | 控制                   | 控制                   |
| 固定效应           | 控制                   | 控制                   | 控制                   | 控制                   |
| R <sup>2</sup> | 0.240                | 0.241                | 0.333                | 0.333                |
| 样本量            | 10153                | 10153                | 4342                 | 4342                 |

表9 地区经济发展水平异质性分析结果

| 变量             | 较高发展水平地区             |                      | 中等发展水平地区             |                      | 较低发展水平地区             |                      |
|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 产业数字化×性别       |                      | 0.157**<br>(0.056)   |                      | 0.331<br>(0.276)     |                      | -0.163<br>(0.379)    |
| 产业数字化          | 0.338***<br>(0.084)  | 0.267***<br>(0.084)  | 0.881<br>(0.596)     | 0.776<br>(0.640)     | -1.696<br>(1.276)    | -1.629<br>(1.396)    |
| 性别             | -0.293***<br>(0.024) | -0.357***<br>(0.036) | -0.324***<br>(0.035) | -0.377***<br>(0.069) | -0.332***<br>(0.020) | -0.310***<br>(0.061) |
| 控制变量           | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| 固定效应           | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| R <sup>2</sup> | 0.352                | 0.353                | 0.272                | 0.272                | 0.313                | 0.313                |
| 样本量            | 8720                 | 8720                 | 4066                 | 4066                 | 1709                 | 1709                 |

贡献,并扩展到所有可定义RIF的统计量中(Firpo et al.,2007)<sup>[42]</sup>。

对不同工资水平分位点进行无条件分位数回归的结果见表10。在10%和50%分位点,“产业数字化×性别”和“产业数字化”的回归系数均显著为正,说明当劳动者工资处于中低水平时,产业数字化可以显著提高劳动者的工资水平,并有效缩小性别工资差距;在90%分位点,“产业数字化×性别”和“产业数字化”的回归系数均不显著,表明对于工资水平处于顶端的劳动者,产业数字化并不能显著提高其工资收入,对性别工资差距的影响也不显著。进一步进行无条件分位数分解,通过构建“被视为男性的女性”的反事实组,将性别工资差距分解为由个体、地区等特征禀赋差异带来

的禀赋效应(可解释的部分)和由工资决定结构差异带来的结构效应(不可解释的部分),其中结构效应可归因为性别歧视(Oaxaca,1973)<sup>[43]</sup>。分位数分解结果见表11。从禀赋效应来看,中低工资群体在产业数字化禀赋上的差异能够有效缩小性别工资差距;从结构效应来看,产业数字化显著减轻了中等工资群体在劳动力市场中面对的歧视;而对于高工资群体,产业数字化的禀赋效应和结构效应均不显著。上述分析表明,产业数字化能显著提高中低工资群体的工资水平,并显著缩小中低工资群体的性别工资差距。

## 五、结论与启示

产业数字化在提高生产效率和就业质量的同时,也在重塑劳动力市场的分配机制,使得不同性

表10 分位数回归结果

| 变量             | 10%分位数               |                      | 50%分位数               |                      | 90%分位数               |                      |
|----------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 产业数字化×性别       |                      | 0.517***<br>(0.140)  |                      | 0.258**<br>(0.121)   |                      | -0.109<br>(0.128)    |
| 产业数字化          | 0.921***<br>(0.218)  | 0.676**<br>(0.246)   | 0.576***<br>(0.113)  | 0.453***<br>(0.134)  | 0.313<br>(0.231)     | 0.364<br>(0.224)     |
| 性别             | -0.235***<br>(0.044) | -0.389***<br>(0.056) | -0.330***<br>(0.026) | -0.407***<br>(0.032) | -0.269***<br>(0.030) | -0.236***<br>(0.044) |
| 控制变量           | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| 固定效应           | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   | 控制                   |
| R <sup>2</sup> | 0.087                | 0.088                | 0.261                | 0.262                | 0.177                | 0.177                |
| 样本量            | 14495                | 14495                | 14495                | 14495                | 14495                | 14495                |

表 11 分位数分解结果

| 变量    | 10%分位数               |                       | 50%分位数               |                      | 90%分位数               |                      |
|-------|----------------------|-----------------------|----------------------|----------------------|----------------------|----------------------|
|       | 禀赋效应                 | 结构效应                  | 禀赋效应                 | 结构效应                 | 禀赋效应                 | 结构效应                 |
| 产业数字化 | -0.071***<br>(0.027) | 0.754<br>(0.583)      | -0.039***<br>(0.012) | -0.537**<br>(0.252)  | 0.008<br>(0.011)     | -0.178<br>(0.454)    |
| 年龄    | 3.528***<br>(0.590)  | 31.909***<br>(6.769)  | 1.051***<br>(0.162)  | 7.941***<br>(2.129)  | 1.111***<br>(0.201)  | 2.222<br>(3.359)     |
| 年龄平方  | -3.591***<br>(0.584) | -24.793***<br>(4.998) | -1.119***<br>(0.166) | -6.392***<br>(1.589) | -1.141***<br>(0.202) | -2.649<br>(2.532)    |
| 婚姻状态  | -0.094**<br>(0.037)  | 4.278***<br>(0.823)   | -0.042***<br>(0.016) | 1.687***<br>(0.318)  | -0.035**<br>(0.015)  | 1.311**<br>(0.599)   |
| 健康状况  | 0.075***<br>(0.028)  | 0.691<br>(1.183)      | 0.027***<br>(0.009)  | -0.433<br>(0.439)    | 0.005<br>(0.008)     | -0.839<br>(0.750)    |
| 政治面貌  | -0.075**<br>(0.031)  | -0.185*<br>(0.096)    | 0.020**<br>(0.010)   | -0.055<br>(0.044)    | 0.076***<br>(0.024)  | -0.183*<br>(0.111)   |
| 户口性质  | -0.127***<br>(0.039) | 0.677*<br>(0.368)     | -0.033***<br>(0.011) | -0.161<br>(0.153)    | -0.060***<br>(0.019) | -0.125<br>(0.254)    |
| 专业证书  | 0.022<br>(0.014)     | 0.156<br>(0.172)      | 0.018*<br>(0.010)    | -0.186**<br>(0.084)  | 0.008<br>(0.007)     | -0.312*<br>(0.178)   |
| 受教育水平 | -0.017<br>(0.031)    | 2.580*<br>(1.383)     | -0.017<br>(0.030)    | -0.689<br>(0.524)    | -0.031<br>(0.057)    | 1.355<br>(0.968)     |
| 单位性质  | 0.013<br>(0.012)     | 0.072<br>(0.220)      | -0.000<br>(0.003)    | -0.113<br>(0.096)    | -0.039*<br>(0.020)   | -0.363*<br>(0.193)   |
| 人均GDP | -0.435***<br>(0.075) | 3.825***<br>(1.188)   | -0.078***<br>(0.018) | -0.172<br>(0.469)    | -0.048*<br>(0.026)   | 1.961***<br>(0.761)  |
| 平均工资  | 0.074<br>(0.052)     | -0.220<br>(0.868)     | -0.150***<br>(0.025) | -0.355<br>(0.349)    | -0.318***<br>(0.051) | -2.424***<br>(0.753) |

别劳动力的工资收入差距呈现出新的特征。科学识别产业数字化对性别工资差距的影响,有助于破解新时代性别工资差距难题,促进经济可持续发展。产业数字化通过增加脑力劳动需求改变了脑力劳动与体力劳动之间的相对价格,还提供了更多的灵活就业机会和更便捷包容的就业环境,促使女性在劳动力市场中的就业竞争力得以提高,并有助于女性劳动力相对优势的充分发挥,从而对女性劳动者具有更强的工资增长效应,由此带来性别工资差距的缩小。本文基于“生产任务”模型和“肌肉—大脑”假说构建理论模型,探究产

业数字化与性别工资差距之间的关系,同时将地区(省份)产业数字化水平与2014年、2016年和2018年CLDS数据库的劳动者个体调查数据相匹配,实证检验产业数字化对性别工资差距的影响及其异质性,结果发现:(1)地区产业数字化水平的提高能够显著促进劳动者的工资收入增长,这种工资增长效应对女性劳动者更为显著,从而缩小了女性劳动者与男性劳动者之间的工资收入差距,该结论经过稳健性检验和内生性处理后依然成立。(2)对于低技能劳动群体和经济发展水平较高地区,产业数字化可以显著缩小性别工资差距;



对于高技能劳动群体和经济发展水平较低地区,产业数字化对性别工资差距的影响不显著。(3)对于中低工资群体,产业数字化具有显著的工资增长促进效应和性别工资差距缩小效应;对于高工资群体,产业数字化对工资收入和性别工资差距没有显著影响。

基于上述结论,本文提出如下启示:第一,加强数字基础设施建设,持续推动产业数字化提速升级,充分发挥产业数字化的就业结构优化效应。进一步投资数字基础设施建设,扩大数据流通范围,提高数据流通效率。完善数据中心、云计算平台等关键数字基础设施,加强对物联网、人工智能等数字技术的研发与应用,增强数字基础设施服务能力,着力打造数字基础设施共享平台,为产业数字化提供有力支撑。第二,因地制宜,推进各地区产业数字化多路径发展。对于经济发展水平较高的地区,要以全球领先技术为目标,优先布局核心数字技术领域,抢占科技创新高地,充分发挥数字技术创新的引领性作用,推动互联网、大数据分析等各类数字技术与各产业深度融合;对于经济发展相对滞后的地区,应加大对产业数字化的政策倾斜和财政支持,缩小与发达地区在产业数字化方面的差距,结合当地的要素禀赋和资源环境制定适宜的数字化转型方案,实现区域间数字经济的均衡发展。第三,促进两性教育公平,优化女性就业结构。为更好发挥产业数字化对女性收入增长的偏向性促进作用,应积极引导女性提高自身人力资本水平,落实教育平等政策,鼓励女性进入科学、技术、工程和数学领域,提供面向女性的数字化培训课程,不断促进女性就业质量提升。同时,还应不断完善社会保障制度和就业创业服务体系,支持新型灵活就业方式,促进女性多元化就业,优化女性就业结构。

本文探讨了产业数字化对性别工资差距的影响,为通过产业数字化缩小性别工资差距提供了政策启示,但仍存在改进和拓展的空间,比如:囿

于数据限制,本文仅检验了省份层面产业数字化对性别工资差距的影响,需要进一步基于更为微观的城市层面进行实证分析;产业数字化对性别工资差距的影响可能存在复杂的路径和多样化的异质性,有待进行更为全面细致深入地研究。

#### 注释:

①采用混合截面数据主要基于以下两面的考虑:一是CLDS采用轮换追踪方式进行调查,每次调查会去掉部分上次调查的样本,并增加新样本,用面板数据会导致样本量大幅减少,用混合截面数据则可以保持足够的样本量;二是相较于横截面数据,混合截面数据具有更好的样本代表性,可以获得更准确的估计值和更有效的检验统计量。

②其中,“较高发展水平地区”包括北京、天津、内蒙古、上海、江苏、浙江、福建、山东、广东、重庆、湖北,“中等发展水平地区”包括安徽、湖南、辽宁、江西、山西、四川、河南、云南、陕西、新疆、宁夏、青海,“较低发展水平地区”包括河北、贵州、广西、吉林、黑龙江、甘肃。

#### 参考文献:

- [1]蔡昉.中国经济改革效应分析——劳动力重新配置的视角[J].经济研究,2017,52(7):4-17.
- [2]王文.数字经济时代下工业智能化促进了高质量就业吗[J].经济学家,2020,32(4):89-98.
- [3]BECKER G S. A theory of the allocation of time[J]. The Economic Journal, 1965, 75: 493-517.
- [4]MINCER J. Schooling, experience, and earnings[M]. New York: National Bureau of Economic Research, 1974.
- [5]BECKER G S. Human capital, effort, and the sexual division of labor[J]. Journal of Labor Economics, 1985, 3(1): 33-58.
- [6]KATZ L F. Efficiency wage theories: a partial evaluation[J]. NBER Macroeconomics Annual, 1986, 1(1): 235-276.
- [7]ALBRECHT J, BJÖRKLUND A, VROMAN S. Is there a glass ceiling in Sweden?[J]. Journal of Labor Economics, 2003, 21(1): 145-177.
- [8]HECKMAN J J, STIXRUD J, URZUA S. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior[J]. Journal of Labor Economics, 2006, 24(3): 411-482.
- [9]张丹丹.市场化与性别工资差异研究[J].中国人口科学,2004,18(1):34-43+81.
- [10]黄志岭,姚先国.教育回报率的性别差异研究[J].世界

经济, 2009, 22(7): 74-83.

[11]葛玉好, 曾湘泉. 市场歧视对城镇地区性别工资差距的影响[J]. 经济研究, 2011, 46(6): 45-56+92.

[12]亓寿伟, 刘智强. “天花板效应”还是“地板效应”——探讨国有与非国有部门性别工资差异的分布与成因[J]. 数量经济技术经济研究, 2009, 26(11): 63-77.

[13]王湘红, 曾耀, 孙文凯. 行业分割对性别工资差异的影响——基于 CGSS 数据的实证分析[J]. 经济学动态, 2016, 44(1): 44-53.

[14]罗楚亮, 滕阳川, 李利英. 行业结构、性别歧视与性别工资差距[J]. 管理世界, 2019, 35(8): 58-68.

[15]卿石松, 郑加梅. “同酬”还需“同工”: 职位隔离对性别收入差距的作用[J]. 经济学(季刊), 2013, 12(2): 735-756.

[16]肖旭, 戚聿东. 产业数字化转型的价值维度与理论逻辑[J]. 改革, 2019, 32(8): 61-70.

[17]AUTOR D H, LEVY F, MURNANE R J. The skill content of recent technological change: an empirical exploration[J]. The Quarterly Journal of Economics, 2003, 118(4): 1279-1333.

[18]ACEMOGLU D, RESTREPO P. The race between man and machine: implications of technology for growth, factor shares, and employment[J]. American Economic Review, 2018, 108(6): 1488-1542.

[19]ACEMOGLU D, RESTREPO P. Automation and new tasks: how technology displaces and reinstates labor[J]. Journal of Economic Perspectives, 2019, 33(2): 3-30.

[20]蔡跃洲, 陈楠. 新技术革命下人工智能与高质量增长、高质量就业[J]. 数量经济技术经济研究, 2019, 36(5): 3-22.

[21]张新春, 董长瑞. 人工智能技术条件下“人的全面发展”向何处去——兼论新技术下劳动的一般特征[J]. 经济学家, 2019, 31(1): 43-52.

[22]杨光, 侯钰. 工业机器人的使用、技术升级与经济增长[J]. 中国工业经济, 2020, 37(10): 138-156.

[23]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[J]. 经济研究, 2020, 55(10): 159-175.

[24]魏下海, 曹晖, 吴春秀. 生产线升级与企业内性别工资差距的收敛[J]. 经济研究, 2018, 53(2): 156-169.

[25]孙早, 韩颖. 人工智能会加剧性别工资差距吗? ——基于我国工业部门的经验研究[J]. 统计研究, 2022, 39(3): 102-116.

[26]许健, 季康先, 刘晓亭, 等. 工业机器人应用、性别工资差距与共同富裕[J]. 数量经济技术经济研究, 2022, 39(9):

134-156.

[27]WELCH F. Growth in women's relative wages and in inequality among men: one phenomenon or two?[J]. American Economic Review, 2000, 90(2): 444-449.

[28]GE S, ZHOU Y. Robots, computers, and the gender wage gap[J]. Journal of Economic Behavior & Organization, 2020, 178(10): 194-222.

[29]GOLDIN, CLAUDIA. "The quiet revolution that transformed women's employment, education, and family." [J]. American Economic Review 2006, 96(2): 1-21.

[30]戚聿东, 刘翠花, 丁述磊. 数字经济发展、就业结构优化与就业质量提升[J]. 经济学动态, 2020, 48(11): 17-35.

[31]陈璐, 范红丽, 赵娜, 等. 家庭老年照料对女性劳动就业的影响研究[J]. 经济研究, 2016, 51(3): 176-189.

[32]李建奇. 数字化变革、非常规技能溢价与女性就业[J]. 财经研究, 2022, 48(7): 48-62.

[33]仇化, 尹志超. 数字化转型、信息搜寻与女性高质量就业[J]. 财贸经济, 2023, 44(7): 124-141.

[34]王军, 朱杰, 罗茜. 中国数字经济发展水平及演变测度[J]. 数量经济技术经济研究, 2021, 38(7): 26-42.

[35]郭显光. 改进的熵值法及其在经济效益评价中的应用[J]. 系统工程理论与实践, 1998, 18(12): 99-103.

[36]黄群慧, 余泳泽, 张松林. 互联网发展与制造业生产率提升: 内在机制与中国经验[J]. 中国工业经济, 2019, 36(8): 5-23.

[37]赵涛, 张智, 梁上坤. 数字经济、创业活跃度与高质量发展——来自中国城市的经验证据[J]. 管理世界, 2020, 36(10): 65-76.

[38]李宏兵, 蔡宏波, 王永进. 市场潜能加剧了性别工资不平等吗? [J]. 数量经济技术经济研究, 2014, 31(1): 22-41.

[39]李红阳, 邵敏. 城市规模、技能差异与劳动者工资收入[J]. 管理世界, 2017, 33(8): 36-51.

[40]张军扩, 罗雨泽, 宋荃柯. 突破“制度高墙”与跨越“中等收入陷阱”——经验分析与理论研究结合视角[J]. 管理世界, 2019, 35(11): 1-7+71+230.

[41]王思博, 庄贵阳. 中国省域碳达峰与基本现代化“双目标”何以协同实现? [J]. 中国人口·资源与环境, 2023, 33(3): 1-12.

[42]FIRPO S P, FORTIN N M, LEMIEUX T. Unconditional quantile regressions[J]. Econometrica, 2007, 77(3): 953-973.

[43]OAXACA R. Male-female wage differentials in urban labor markets[J]. International Economic Review, 1973, 14(3): 693-709.