

# 数字经济发展与农民工城市留居意愿

## ——人力资本异质性视角的分析

曹静 息晨 纪承名\*

**摘要：**将 2012-2018 年中国流动人口动态监测数据与北京大学数字普惠金融指数相匹配形成混合截面数据，本文在县域层面分析流入城市数字经济发展对农民工长期留城意愿的影响。工具变量估计结果表明，城市数字经济发展对农民工长期留居意愿具有正向影响。数字经济指数每提升一个标准差，农民工留城意愿增加 10.3%。数字经济对留城意愿的促进作用存在人力资本异质性，对高人力资本农民工的影响较低人力资本农民工高 1.3%。机制分析表明，支付便利性是人力资本异质性的主要原因。尽管数字经济发展引致城市劳动力需求结构偏向高人力资本农民工，但学习效应下低人力资本农民工更快的劳动生产率提升缩小了两类农民工的工资差距。本文为数字经济时代下农民工福祉问题与新型城镇化建设的讨论提供了新视角。

**关键词：**数字经济；农民工；支付便利性；工资差距

### Digital Economy Development and Rural Migrant Workers' Willingness to Settle in Cities

#### —Analysis from the Perspective of Human Capital Heterogeneity

**Abstract** This paper analyzes the impact of digital economy development of inflowing cities on rural migrant workers' long-term willingness to settle in cities at the county level using cross-sectional data constructed by the China Migrants Dynamic Survey (CMDS) and the Peking University digital finance index over 2012-2018. The instrumental variable estimation results show that cities' digital economy development positively affects migrants' settle willingness. For one standard deviation increase in the digital economy index, migrants' settle willingness increased by 10.3%. We further find that the effect exhibits human capital heterogeneity. The effect of digital economy development on migrant workers with high human capital is 1.3% higher than that on low human capital migrants. Mechanism analysis shows that human capital heterogeneity is mainly explained by the difference of payment facilitation in two migrant groups. Although digital economy development leads the structure of urban labor demand biases toward high human capital migrant workers, faster productivity improvement of low human capital migrant workers through social learning narrows down the wage gap. This paper provides a new perspective for discussing the well-being of migrant workers and the construction of new-type urbanization in the digital economy era.

**Keywords** digital economy; rural migrant workers; payment facilitation; wage gaps

**JEL Classification** G10, J31, J61

---

\*曹静，清华大学经济管理学院；息晨，清华大学经济管理学院；纪承名，中国农业大学经济管理学院。  
通信作者及地址：息晨，清华大学经济管理学院经济系，E-mail: xic20@mails.tsinghua.edu.cn。作者感谢国家社科青年基金（项目编号：21CJY039, 22CJY065）的资助。感谢冷晨昕、祝仲坤、张沛康对本文提出的宝贵意见。感谢第 29 期香樟经济学 Seminar（广州）、北大国发院第四届国家发展青年论坛、第八届清北人三校研究生学术论坛与会学者对本文的建设性意见。感谢匿名审稿人的宝贵意见。作者文责自负。

## 一、引言

中国是流动人口大国，农民工是流动人口的重要组成部分。庞大的流动人口数量为中国的城市劳动力及经济增长做出了巨大的贡献，根据国家统计局与人社部数据显示，2020年中国农民工总量达 2.856 亿人，占全国就业人口约 38.5%。然而，由于长期以来户籍制度的存在，农民工无法享受与城镇职工均等的公共服务，这阻碍了农民工融入城市与实现市民化的进程。农民工市民化关乎经济持续健康发展，是推进新型城镇化与建设经济双循环发展的重要组成部分。自党的十八大以来，我国政府围绕农民工市民化问题进行了一系列改革，力图让农民工“进得来、留得住、增强归属感”。据国家统计局发布的农民工动态监测数据显示，2016 年认为农民工自己是所在城市“城市人”的农民工比例为 35.6%，至 2020 年此比例已增长至 41.4%，推进农民工融入社会的工作取得了积极进展。然而，现阶段包括农民工在内的流动人口社会融入水平仍然不高，距离理想状态差距仍然较大（肖子华等，2018）。

与此同时，伴随着数字技术的发展和数字基础设施的普及，中国正在经历数字经济的快速发展期。数字经济是以人工智能、大数据、云计算、5G 技术等为代表的新型社会经济形态，逐渐成为信息时代下拉动经济发展的“新引擎”。许宪春和张美慧（2020）的测算表明，2017 年中国数字经济增加值与总产出达到 53028.85 亿元与 147574.05 亿元，分别占国内生产总值与总产出的 6.46%与 6.53%，体现了数字经济在经济新常态下对中国经济增长的重要意义。数字经济的发展深刻地改变了国民经济的生产、消费与分配方式，现有研究从居民消费（易行健和周利，2018；何宗樾和宋旭光，2020；张勋等，2020）、创业（谢绚丽等，2018；赵涛等，2020）、就业（戚聿东等，2020）与产业结构（陈晓东和杨晓霞，2021）等角度证实了数字经济对经济高质量发展的促进作用。那么，随着数字经济的发展，农民工群体福祉受到怎样的影响？在农民工融入城市进程中，数字经济起到怎样的作用？其中的影响机制为何？对这些重要的现实问题，尚无完整、细致的研究给出解答。

现有从数字经济与数字金融出发，对象为农村居民与农民工的研究主要涉及如下几方面。其一为数字经济对城乡收入差距的缩小作用。数字金融具有普惠性特点，可以惠及被传统金融服务排除在外的农村低物质资本与低社会资本家庭，促进农村低收入家庭实现收入增长（张勋等，2019）。由于数字经济的地理穿透性，数字经济的发展可弥补乡村信息设施的不足，通过减小城乡间“信息鸿沟”实现城乡收入差距的缩小（李晓钟和李俊雨，2021）。其二，使用农村居民样本，分析数字经济发展对非农就业决策的影响。Zhang et al.（2020）结合 CFPS 数据与北京大学数字普惠金融指数的研究表明，数字经济的发展显著提高了农村居民的非农部门就业概率，这将有助于城镇化的实现。数字经济在此过程中体现出“涓滴效应（trickle-down effect）”，即之前无法接触互联网的家庭也可通过数字技术的发展实现非农就业概率的提升。其三为使用流动人口数据，考察城市数字经济水平对流动人口迁移城市选择的影响。马述忠和胡增玺（2022）将北京大学数字普惠金融指数与流动人口动态监测数据相结合，发现城市数字经济程度的提高将显著促进流动人口流入，数字经济发展为流动人口提供更为丰富的就业机会以及预期收入的增加是主要机制。综合上述研究，可以发现，现有数字经济发展下与农民工相关的研究仍局限于农民工“进得来”城市的角度，对于城市数字经济发展下农村移民福祉的分析仍显不足。本文将 2012-2018 年流动人口动态监测数据与北京大学数字普惠金融指数相结合构建混合截面数据，从微观视角出发，分析农民工长期留城意愿如何受到城市数字经济发展的影响。本文着重数字经济发展对农民工“留得住”的影响，是对现有数字经济背景下农民工福祉研究的后半程补充。随

着“数字中国”的进一步推进，农民工群体在数字化浪潮下的长期留城意愿事关以人为本的高质量城镇化的实现，本文具有鲜明的现实意义与政策意涵。

农民工进城务工获得更高的收入是其流动的初衷，而实现福利提升、增进对城市的认同则有赖于通过消费享受城市提供的多样性服务。生活服务与公共服务更丰富的城市对流动人口吸引力更强，流动人口的长期留居意愿也更高（夏怡然和陆铭，2015；张文武和余泳泽，2021；祝仲坤，2021）。伴随着城市数字经济的发展，消费者可以获得的城服务迅速扩张。以生活服务、社交媒体、网络游戏等为代表的消费互联网的发展打破了传统生活、社交、娱乐等服务的时间与空间限制，深刻地改变了城市居民的生活日常。在消费数字化的转变下，城镇居民得以借助数字平台降低其消费过程中的搜寻成本，通过增加消费以增进个人福祉（张勋等，2020）。考虑到数字普惠金融的包容性特征，流动至城市的农民工群体也应当会受益于城市数字经济的发展，以更低的消费成本实现更为便捷的消费。数字化转型对受众的数字生活技能提出了要求，但模糊了以往生活服务中的身份界限。农民工学习、掌握数字化生活技能的过程会加深其自身对于城市生活方式的理解与“城市人”身份的认同，通过线上方式实现的数字化服务也有助于降低其感知到的被歧视程度，这将有助于实现农民工城市融入意愿的提升。

城市数字经济发展同样可能通过影响农民工群体的就业结构与劳动生产率影响其工资收入，进而影响农民工的留城意愿，即数字经济的就业收入机制。从广延边际（*extensive margin*）角度，数字经济促进了城市创业活跃度的提高（谢绚丽等，2018；赵涛等，2020）。初创企业为流动人口提供了丰富的就业机会，促使流动人口向数字经济更为发达的城市流动。同时，这些诞生于数字时代的企业数字化程度往往更高，这对雇员的数字化生产技能提出了要求。此外，随着数字经济的发展，企业为提高全要素生产率，也将进一步优化其人力资本结构，增加高人力资本员工比例（赵宸宇等，2021）。这意味着，数字经济的发展会在城市或城市-行业的宏观层面表现为对农民工的劳动力需求向高人力资本群体偏移（Acemoglu & Autor, 2011；卢晶亮，2017）。对于人力资本较低的农民工，在这个过程中可能被挤出城市的就业部门，表现为数字经济发展对低技能农民工群体收入的损害（柏培文和张云，2021）。而在企业内部的集约边际（*intensive margin*）上，不同人力资本农民工劳动生产率随数字化的提升速度同样将对相对工资收入产生影响（Sin et al., 2020）。值得说明的是，农民工相对生产率的变化并不一定呈现出类似于广延边际下的高人力资本偏向。数字化生产技能并不完全取决于由受教育程度所代表的人力资本水平，低人力资本农民工可以通过社会学习提升其数字化生产率（Roca & Puga, 2017）。由于社会学习需要相应的时间积累，早进城的低学历农民工反而更可能通过社会学习实现劳动表现的提升（魏东霞和陆铭，2021）。如果由于学习效应的存在，伴随着数字经济发展，低人力资本农民工相对生产率提升更快，则数字经济将通过集约边际缩小两类农民工群体的工资差异。数字经济发展通过就业收入机制对农民工留居意愿的影响将是广延边际效应与集约边际效应的综合。

以上分析表明，城市数字经济的发展将通过改变生活方式与就业收入的两方面机制对农民工留城意愿产生影响。由于数字经济对于高/低人力资本农民工留城意愿的影响可能存在结构差异，有必要从农民工人力资本异质性的视角开展分析，以丰富对农民工群体内数字经济影响不平等的探讨。本文将建立一个理论框架，厘清上述机制的作用途径。依据理论框架的指导，使用详细的农民工个人层面微观数据，对农民工长期留城意愿如何受到城市数字经济发展的影响进行实证解答。

本文的创新与边际贡献体现在以下三方面。第一，不同于以往数字经济发展影响城乡融合研究，本文着眼于已经流动至城市的农民工群体，分析城市数字经济发展对其长期留城意愿的影响。本文补充了数字化下农村劳动力转移研究中的“后半程”，丰富了新发展时

代下与农民工福祉相关的讨论。第二，不同于将农民或农民工视为整体的单一维度的分析，本文进一步区分数字经济对不同人力资本农民工留城意愿的异质性影响，并从生活方式与就业收入效应两方面开展机制分析，突出了农民工群体内部受数字经济发展影响的不平等性。第三，在数据分辨率上，本文使用县域层面流动人口监测数据及北京大学普惠金融指数数据进行实证分析，区别于通常的地级市与省级层面研究，这有助于实现对数字经济发展效应更为准确的识别。县域层面的分析也强化了这两类微观数据的使用深度。

本文接下来的结构安排如下：第二部分建立理论模型，通过对理论模型的分析厘清数字经济的影响机制，并提出相应的研究假说。第三部分介绍本文的研究数据及实证设定。第四部分实证解答数字经济对农民工长期留城意愿的影响，并进行相应的稳健性检验与机制分析。最后一部分简要总结本文发现，提出政策建议并对未来的研究方向进行展望。

## 二、理论模型与实证假说

本部分通过建立一个简单的家庭-厂商静态一般均衡模型，从理论上阐明数字经济发展对高/低两类人力资本农民工留城意愿影响的差异及影响机制。通过理论模型的分析，提出相应的待检验假说以指导后续的实证分析。

### （一）家庭

城市中存在高、低两类人力资本的农民工家庭， $j \in \{H, L\}$ 。两类家庭的效用取决于消费量及消费中的搜寻成本，家庭的效用函数为：

$$U(c_j) = (c_j/\varepsilon_j)^\gamma, \quad \gamma \in (0, 1). \quad (1)$$

其中 $c_j$ 与 $\varepsilon_j$ 分别代表家庭消费量与为实现消费而付出的搜寻成本。消费量越大，家庭效用水平越高。同等消费量下，搜寻成本的降低也会提升家庭的效用水平。

代表性家庭具有 1 单位劳动禀赋，其被用于两方面：劳动 $\varepsilon_j$ 在产品市场进行搜寻以实现消费，其余劳动 $L_j$ 投入就业市场以获得工资收入。显然， $L_j + \varepsilon_j = 1$ 。家庭的预算约束来自工资收入，将消费品价格标准化为 1，即得到：

$$c_j = w_j L_j = w_j (1 - \varepsilon_j), \quad (2)$$

其中 $w_j$ 为 $j$ 类型人力资本家庭的工资率。

与张勋等（2020）类似，在上述设定中， $\varepsilon_j$ 刻画了家庭消费过程中的搜寻成本。搜寻成本的变化会通过消费与就业两种渠道对农民工家庭效用产生综合影响。在消费端，搜寻成本的降低意味着消费便捷性的提升，对家庭效用具有直接的促进作用。而在就业端，更低的搜寻成本使家庭可将更多的劳动投入就业市场，通过就业投入与均衡工资率对家庭消费预算约束产生影响，从而间接影响家庭的效用水平。

数字经济的发展使得更多的线上搜寻成本可能，这将极大降低消费中的搜寻成本，促进支付便捷性与消费的提升（张勋等，2020）。对于农民工家庭，假定其所在城市数字经济水平的提升会降低其用于消费搜寻中的劳动投入，即 $\partial \varepsilon_j / \partial DE < 0$ 。此外，假定数字经济对搜寻成本的降低效应是边际递减的，即 $\partial^2 \varepsilon_j / \partial DE^2 > 0$ 。

### （二）厂商

本文旨在从农民工角度考察数字经济发展对其福利及留城意愿的影响，故使用简约形

式刻画数字经济下厂商的生产行为<sup>②</sup>。依据 Card & Lemieux (2001) 的设定, 企业仅使用两类人力资本的劳动力生产消费品, 其生产函数为 CES 函数:

$$Q = \left[ \mu (\varphi_L L_L)^{\frac{\sigma-1}{\sigma}} + (1-\mu) (\varphi_H L_H)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (3)$$

其中为  $Q$  消费品产量。 $\mu$  与  $(1-\mu)$  分别代表代表性企业生产中低人力资本与高人力资本劳动力投入比重, 设定  $\mu$  为不随时间变化的常数。 $L_L$  与  $L_H$  代表对低人力资本与高人力资本劳动力的需求。 $\varphi_L$  与  $\varphi_H$  代表两类劳动力的生产率水平。 $\sigma$  为两类劳动力间的替代弹性, 根据以往实证研究的发现, 设定  $\sigma$  为大于 1 的常数。

代表性厂商处于一个完全竞争市场, 以工资  $w_L$  与  $w_H$  雇佣两类劳动力进行生产, 厂商的利润最大化问题为:

$$\max_{\{L_L, L_H\}} \pi = \left[ \mu (\varphi_L L_L)^{\frac{\sigma-1}{\sigma}} + (1-\mu) (\varphi_H L_H)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} - w_L L_L - w_H L_H, \quad (4)$$

求解上述问题的一阶条件, 可得到:

$$\begin{aligned} (1-\mu) \cdot Q^{\frac{1}{\sigma}} \varphi_H^{\frac{\sigma-1}{\sigma}} L_H^{-\frac{1}{\sigma}} &= w_H \\ \mu \cdot Q^{\frac{1}{\sigma}} \varphi_L^{\frac{\sigma-1}{\sigma}} L_L^{-\frac{1}{\sigma}} &= w_L \end{aligned}, \quad (5)$$

(5) 式即为就业市场局部均衡的结果。

### (三) 一般均衡

当达到一般均衡时, 经济中的产品市场与劳动力市场出清。产品市场出清意味着厂商雇佣两类农民工, 使用其有效劳动时间  $(1-\varepsilon_j)$  进行生产, 产出的最终品被经济中两类农民工家庭消费。家庭的单位禀赋劳动一部分用于在产品市场进行消费搜寻, 另一部分则参与就业市场进行生产, 投入就业市场的劳动力与厂商需求相等, 实现劳动力市场出清。当市场出清时, 由 (2) 与 (5) 式, 可得到:

$$\ln\left(\frac{c_H}{c_L}\right) = \ln\left(\frac{1-\mu}{\mu}\right) + \frac{\sigma-1}{\sigma} \ln\left(\frac{L_H}{L_L}\right) + \frac{\sigma-1}{\sigma} \ln\left(\frac{\varphi_H}{\varphi_L}\right). \quad (6)$$

进一步结合 (6) 式与 (1) 式, 两类农民工家庭效用差异可被分解为:

$$\ln\left(\frac{u_H}{u_L}\right) = \gamma \ln\left(\frac{1-\mu}{\mu}\right) + \gamma \frac{\sigma-1}{\sigma} \ln\left(\frac{L_H}{L_L}\right) + \gamma \frac{\sigma-1}{\sigma} \ln\left(\frac{\varphi_H}{\varphi_L}\right) - \gamma \ln\left(\frac{\varepsilon_H}{\varepsilon_L}\right), \quad (7)$$

其中  $(1-\mu)/\mu$  为常数, 不会受到数字经济发展的影响。(7) 式揭示了数字经济水平  $DE$  的发展对两类人力资本农民工家庭效用差距的三方面影响, 即:

$$\frac{\partial \ln(u_H/u_L)}{\partial DE} = \gamma \frac{\sigma-1}{\sigma} \frac{\partial \ln(L_H/L_L)}{\partial DE} + \gamma \frac{\sigma-1}{\sigma} \frac{\partial \ln(\varphi_H/\varphi_L)}{\partial DE} - \gamma \frac{\partial \ln(\varepsilon_H/\varepsilon_L)}{\partial DE}. \quad (8)$$

(8) 式的一般均衡结果表明, 数字经济发展对两类农民工家庭效用的异质性影响可被分解为对劳动力需求结构的相对影响、对劳动生产率的相对影响以及对搜寻成本的相对影响。我们称前两项为数字经济发展的就业收入效应, 后一项为数字经济发展的生活便利性效应。在就业收入机制中, 数字经济发展通过影响劳动力需求结构与农民工相对生产率进而影响两类农民工的相对工资收入。数字经济促进了创业活跃度 (谢绚丽等, 2018; 赵涛

<sup>②</sup> 从生产角度详细建模数字经济、人工智能技术发展对产业结构、劳动力工资差距等影响的理论研究可参见 Acemoglu & Restrepo(2018)、郭凯明 (2019) 及王林辉等 (2020)。

等，2020），这些新创企业为农民工提供了大量的就业机会，从而促进了城市经济的高质量发展。然而，数字经济背景下催生的企业可能是更为数字化的，其通常要求劳动力具有相应的数字化技能，能够与企业的数字化需求相匹配。数字经济的发展在此层面上可视为一种技能偏向型的技术进步（Acemoglu & Autor, 2011；卢晶亮，2017）。高人力资本农民工通常能更快适应企业的数字化技能要求并与企业的数字化程度互补，因而劳动力需求结构将更偏向高人力资本农民工，这意味着在宏观的城市或城市-行业层面将有  $\frac{\partial \ln(L_H/L_L)}{\partial DE} > 0$ 。由于替代弹性  $\sigma > 1$ ，表明数字经济发展引致的技能偏向型技术进步将拉大高/低人力资本农民工家庭间的工资收入及效用差异。此外，城市数字经济发展也将通过劳动力生产率的改变对异质性人力资本农民工的均衡收入产生影响。通常而言，数字经济发展对劳动生产率的相对影响  $\frac{\partial \ln(\varphi_H/\varphi_L)}{\partial DE}$  难以直接观测，且劳动生产率的相对提升不一定偏向高人力资本群体。区别于普通劳动者，农民工群体具有明显的聚集居住现象，同乡、文化相近的农民工自发居住于同一或相近社区（Zhang & Xie, 2013）。在数字经济发展下，聚居的农民工间可能通过社会学习、信息共享等方式实现数字生产技能的提高（Battisti et al., 2016）。数字生产技能并不完全取决于以受教育程度为代表的人力资本水平，非认知能力等不通过学校教育而习得的技能对于生产率的提升同样非常重要。农民工可以通过向其周围的农民工学习增进数字化生产技能、提升劳动生产率，由于社会学习需要一定的时间积累，早进城、受教育程度较低的低人力资本农民工反而可能从中收益更多（魏东霞和陆铭，2021）。学习效应的存在可能缩小两类人力资本农民工间的生产率差距，数字经济对该项影响的符号并不确定。对于生活便利性机制， $-\gamma \frac{\partial \ln(\varepsilon_H/\varepsilon_L)}{\partial DE}$  体现了数字经济发展对两类农民工群体消费端的直接影响。由于  $\partial \varepsilon_j / \partial DE < 0$ ，若消费搜寻成本的降低效应对于高人力资本农民工更强，则城市数字经济水平的提高将拉大高/低人力资本农民工间的效用差距。张勋等（2020）的研究表明，数字金融具有普惠性，可以服务难以被传统金融覆盖的低人力资本人群，帮助其通过数字支付实现便捷消费（张勋等，2020）。数字经济发展对两类农民工搜寻成本的相对影响方向同样不确定，有待于通过实证分析进行解答。

#### （四）农民工效用与留城意愿

上述分析阐明了数字经济发展对两类人力资本农民工效用影响的差异来源，为考察其对农民工留城意愿的影响，仍然需要农民工效用与留城意愿间的关系。为契合后续的实证分析，我们简单假定农民工效用与其留城意愿间关系是单调递增的。将农民工主观幸福感作为其效用的近似，我们使用 2012 年流动人口动态监测数据对上述假设进行了检验，幸福感与城市留居意愿间的单调关系非常显著<sup>③</sup>。

结合上述分析，提出本文的待检验假说：

**假说：**城市数字经济发展对高人力资本与低人力资本农民工留城意愿的影响由数字经济的就业收入效应与生活便利性效应综合决定。若数字经济引致城市劳动力需求偏向高人力资本农民工、高人力资本农民工相对生产率提升更快、高人力资本农民工消费成本降低更多，则高/低人力资本农民工间的留城意愿差距将被拉大。

<sup>③</sup> 仅 2012 年的流动人口动态监测数据询问了农民工主观幸福感问题，具体的检验细节请参见附录中表 A1。

### 三、研究数据与实证策略

#### （一）数据来源

本文所关注的农民工留城意愿数据及其个人、家庭信息来自 2012-2018 年流动人口动态监测调查（China Migrants Dynamic Survey, CMDS）。CMDS 由国家卫生与健康委员会组织，每年 5 月在全国 31 个省（区、市）及新疆生产建设兵团，采取分层、多阶段、与规模成比例的 PPS 方法对流动人口进行抽样调查，数据具有极好的微观基础与全国代表性。

CMDS 的调查对象为在流入地居住一个月以上，非本区（县、市）户口的 15 周岁以上的流动人口。本文的关注对象为在城市居住、工作的农民工，参照朱明宝和杨云彦（2016）以及魏东霞和陆铭（2021）的研究，对样本进行如下筛选与剔除：（1）保留在城市流动的农民工样本，剔除样本点类型为村委会的样本；（2）保留户口为农业户口的样本；（3）保留男性年龄在 15-59 周岁，女性年龄在 15-55 周岁的劳动人口样本，剔除在此年龄范围内已退休的样本；（4）保留流动原因为务工经商或随迁的样本，剔除因婚嫁、拆迁、投亲、学习、参军等其他原因流动的样本；（5）剔除户籍地为港澳台及跨国迁移的样本。此外，由于无法与数字经济指数匹配，分析中未纳入新疆生产建设兵团、黑龙江农垦总局与森工总局的调查样本。除 2013 年 CMDS 的 A 卷调查中未包括农民工社会融入信息，使用 C 卷 8 城市的小样本进行分析外，其他年份均使用 A 卷进行分析<sup>④</sup>。经剔除异常值及不合逻辑的样本后，最终形成 2012-2018 年 CMDS 混合截面数据用于实证分析，逐年样本量分别为 80221、8493、102750、105993、86003、81974、63365。

本文的核心解释变量，县域层面数字经济发展水平来自北京大学数字金融研究中心发布的北京大学数字普惠金融指数。此指数包含了全国 31 个省、337 个地级以上城市和约 2800 个县域的数字普惠金融总指数、覆盖广度、使用深度、数字化程度及更为丰富的二级指标（郭峰等，2020）。此指数具有较好的地区间横向可比性与时间上纵向可比性，可满足本研究的需要，此指数也被广泛应用于与数字经济、数字普惠金融相关的研究中（谢绚丽等，2018；张勋等，2019；Zhang et al., 2020）。

值得说明的是，尽管此套数字普惠金融指数的底层数据为来自蚂蚁金服的支付宝使用信息，更多反映其服务于小微企业与中低收入人群的金融普惠性。但基于以下原因，本文仍将其作为农民工群体可感知到的数字经济的表征，从而不将数字金融与数字经济的表述在本文中进行区分。其一，基于上文的理论分析，数字经济对农民工留城意愿的影响主要体现为生活便利性与就业收入效应两方面。生活便利性上，通过数字技术惠及以往被传统金融服务覆盖较少的农民工群体，促进其消费及福祉的提升本就是数字普惠金融的意涵。普惠金融对居民消费的促进作用已得到广泛的证实（易行健和周利，2018；张勋等，2020）。对于就业收入效应，尽管上市公司、规模以上企业不是支付宝的直接客群，使用普惠金融指数进行此类大企业劳动力需求的分析并不合适（郭峰和熊云军，2021）。但本文所关心的农民工群体通常也难以进入到此类大企业中，而是集中于普惠金融所服务的小微企业。现有研究证实，城市数字普惠金融的发展会促进创业（谢绚丽等，2018），为流动人口提供就业机会从而促进劳动力流动（马述忠和胡增玺，2022）。本文并非意在对数字经济的城市就业收入效应进行系统分析，而是着眼于较为特殊的农民工群体。从这个角度上，使用普惠金融数据进行分析是较为合适的。其二，为全面反映城市数字经济发展水平，部分

<sup>④</sup> 在 2013 年 CMDS 的 A 卷中不包含受访者留城意愿的相关信息。上海市松江区、江苏省苏州市、无锡市、福建省泉州市、湖北省武汉市、湖南省长沙市、陕西省西安市、陕西省咸阳市的受访者在回答 A 卷问题之后继续回答 C 卷的社会融合专题调查问题。

研究选取多个子指标以合成数字经济综合指标（如赵涛等，2020；潘为华，2021）。但在指标构建上，数字普惠金融指数通常是输入指标之一，这些指标与普惠金融指数间存在较高的相关性。此外，从数据分辨率上，限于县域数据的可获得性，此类指标通常仅能在地级市层面进行构建。农民工流入地县域数量丰富，县域层面的数字经济指数更能捕捉地区与时间上的数字经济水平波动，从而进行更为准确的因果识别，北京大学普惠金融指数使之成为可能。

由于县域层面普惠金融指数仅在 2014 年后可得，对于 2012-2014 年 CMD5 数据，根据农民工流入地地级市与普惠金融指数进行合并。对于 2015-2018 年 CMD5 数据，根据流入地县级市与普惠金融指数进行合并，最终得到 2012-2018 年 528799 条观测用于实证分析。

（二）变量选取与描述性统计

本文的被解释变量为农民工留城意愿，此变量从心理认同角度反映了农民工的城市融入意愿（杨菊华，2015）。根据受访者对其未来 5 年是否打算在本地长期居住的回答构建二元变量。在 CMD5 各年调查中，对受访者长期留居意愿问题的设计稍有不同，我们对其进行处理以保证该问题跨年间的纵向可比性。在基准分析中，若受访者回答未来 5 年“打算”留在流入城市，则将留城意愿变量赋值为 1。若回答“不打算”或“没想好”，则将留城意愿赋值为 0。55.12%的农民工回答未来 5 年打算继续留在所在城市。

核心解释变量数字经济发展水平来自北京大学普惠金融指数。在基准分析中，选取代表城市数字经济综合水平的总指数进行分析。2012-2014 年数字经济指标使用城市指数，2015-2018 年数字经济指标来自县域指数。在样本期间，农民工所在城市平均经济指数由 2012 年的 64.91 增长至 2018 年的 119.17，体现了在此间中国城市数字经济水平的迅速增长。

本文的控制变量包括农民工个人与家庭特征及城市经济特征。农民工个人特征包括其性别、年龄、民族、受教育年限、婚姻状况等，家庭特征包括与受访农民工家庭随迁人员数量、迁移类型、流动时长等，上述信息均可从 CMD5 问卷中获得。考虑到城市的数字经济水平与城市经济水平可能高度相关，进一步控制了流入地县级市层面人均 GDP。县级市层面人均 GDP 来自于 EPS 数据平台的《中国区域经济数据库》中分县数据<sup>⑤</sup>。对于其中缺失的信息，根据各城市各年的《国民经济和社会发展统计公报》等信息进行手工填补。对于手工填补后仍缺失的县级市数据，使用其所在地级市的人均 GDP 进行补充，地级市数据来自《中国城市统计年鉴》。

表 1 汇报了变量定义及其描述性统计结果。

表 1 变量定义与描述性统计

变量	变量定义	均值	方差	最小值	最大值
留居意愿	未来 5 年打算留在所在城市=1，不打算或没想好=0	0.551	0.497	0	1
数字经济总指数	2012-2014 年为地级市层面指数，2015-2018 年为县级市层面指数	125.918	46.654	11.3	272.72
性别	男=1，女=0	0.535	0.499	0	1
年龄	根据出生年月与当年调查时间计算，岁	34.362	9.190	16	60
受教育年限	未上过学=0 年，小学=6 年，初中=9 年，高中/中专=12 年，大学专科=15 年，大学本科=16 年，研究生=19 年	9.784	2.840	0	19
民族	汉族=1，其他=0	0.927	0.261	0	1
婚姻状况	在婚=1（包括初婚、再婚），非在婚	0.806	0.395	0	1

<sup>⑤</sup> EPS 数据平台的网址为：www.epsnet.com.cn。感谢 EPS 数据平台对本研究提供的数据支持。



	=0（包括未婚、离婚、丧偶）				
家庭随迁人数	与受访者在流入地同住的家庭人员数量，人	1.590	0.395	0	9
迁移类型：市内跨县	流动范围为市内跨县=1，其他=0	0.195	0.396	0	1
迁移类型：省内跨市	流动范围为省内跨市=1，其他=0	0.333	0.471	0	1
迁移类型：跨省	流动范围为跨省=1，其他=0	0.472	0.499	0	1
流动时长	本次流动时长，根据本次流动开始时间与当年调查时间计算，年	4.976	5.079	0.833	41
县级市人均GDP对数	县级市人均GDP（元）取对数	11.068	0.606	8.515	13.056

注：由于数据缺失，迁移类型变量样本量为 528704，人均 GDP 变量样本量为 527623，其他变量样本量为 528799。

### （三）实证设定

首先使用如下基准设定考察城市数字经济发展对农民工总体留城意愿的影响：

$$plan\_to\_stay_{ict} = \alpha + \beta DE_{c,t-1} + \mathbf{X}_{i,t} \gamma + \theta_{c,p} + \lambda_t + \Omega_o + \varepsilon_{ict} \quad (9)$$

其中  $plan\_to\_stay_{ict}$  代表  $t$  年位于流入地城市  $c$  的  $i$  农民工的长期留城意愿<sup>⑥</sup>。 $DE_{c,t-1}$  为  $c$  城市的数字经济水平，为了缓解潜在的反向因果问题，参照文献中的一般做法（张勋等，2019），对数字经济变量取滞后一年处理。本文使用的是 2012-2018 年混合截面数据， $\beta$  借助城市与年份间的数字经济水平的波动识别了其对农民工留居意愿的影响。 $\mathbf{X}_{i,t}$  代表农民工个人、家庭与城市层面的控制变量，相应变量在表 1 中列出。根据上文的理论分析，数字经济水平可能通过支付便利性与就业等机制影响农民工留居意愿，此处未控制相应的机制变量，如农民工就业状况、收入、城市劳动力结构等变量， $\beta$  实际上体现了通过各机制实现的总体影响。由于 2012-2014 年数字经济变量仅在地级市层面可得，控制了  $c$  城市所在地级市  $p$  的城市固定效应  $\theta_{c,p}$ 。 $\lambda_t$  为年份固定效应， $\Omega_o$  为农民工流出地省份固定效应<sup>⑦</sup>。将标准误聚类在县级市层面，以允许同县域内农民工之间的相关性。被解释变量留居意愿为二元变量，通常可使用 Probit 或 Logit 模型进行估计。但在控制年份固定效应或交互项时，非线性概率模型的估计结果是不一致的。本文使用更为简洁的线性概率模型（linear probability model, LPM）进行估计，以方便系数解释与内生性处理。

尽管在上述设定中，使用滞后一年数字经济变量的设定缓解了潜在的反向因果问题，但模型仍可能面临因遗漏变量导致的内生性问题。若随机误差项  $\varepsilon_{ict}$  中遗漏了与城市数字经济水平及农民工留城意愿同时相关的变量，则  $\beta$  的估计将是不一致的。为缓解潜在的内生性问题，借鉴 Nunn & Qian（2014）以及柏培文和张云（2021）的想法，使用上一年度全国城市平均数字经济水平（ $\overline{DE}_{t-1}$ ） $\times$  县域平均坡度作为数字经济指数的工具变量。地形特征是外生的地理要素，会通过影响城市数字基础设施的建设影响数字经济的发展，但不会直接影响农民工的留城意愿。具体而言，一个城市的地势越为平坦，越有利于数字基站等基础设施的建设，进而有助于城市数字经济的发展。全国平均数字经济水平是相较于城市更为宏观的变量，与城市的数字经济水平相关，但不会影响农民工对于特定城市的留居意愿。二者的乘积体现了由外生的地理条件引致的城市数字经济水平的逐年扩散，在控制

<sup>⑥</sup> 2012-2014 年  $c$  为地级市指标，2015-2018 年  $c$  为县级市指标。为表述方便，下文均简称城市。

<sup>⑦</sup> 2018 年 CMDS 数据中未报告农民工流出地省份信息，后续在控制流出地省份固定效应时无法使用 2018 年 CMDS 样本。

了相应的控制变量与固定效应后，工具变量估计给出了数字经济对农民工留居意愿因果影响的局部平均处理效应（local average treatment effect, LATE; Imbens & Angrist, 1994）。在后续的实证结果与分析部分，我们进一步论述并检验了工具变量的相关性与其排他性。

数字经济对高人力资本与低人力资本农民工留城意愿的影响可能不同，体现为数字经济对农民工群体内城市发展的不平等影响。参照卢晶亮（2017）与李天成等（2022）的设定，并根据 CMDS 数据中农民工受教育年限的分布，设定高中及以上受教育程度农民工为高人力资本农民工，其余受教育程度农民工为低人力资本农民工。依此设定，32.95%的样本为高人力资本农民工。使用如下的交互项设定分析数字经济发展对两类人群的异质性影响：

$$plan\_to\_stay_{ict} = \alpha + \beta_1 high\_hc_{ict} \times DE_{c,t-1} + \beta_2 low\_hc_{ict} \times DE_{c,t-1} + \mathbf{X}_{i,t} \gamma + \theta_{c,p} + \lambda_t + \Omega_o + \varepsilon_{ict} \quad (10)$$

其中  $\beta_1$  与  $\beta_2$  分别体现了数字经济对高人力资本农民工与低人力资本农民工留城意愿的影响。在（10）中， $DE_{c,t-1}$  仍可能是内生的，仍使用全国城市平均数字经济水平  $\times$  城市平均坡度作为工具变量进行估计。 $\beta_1$  与  $\beta_2$  间的差异可通过 Delta 法进行检验，或进行如下更为简洁的回归：

$$plan\_to\_stay_{ict} = \alpha + \beta_3 high\_hc_{ict} \times DE_{c,t-1} + \beta_4 DE_{c,t-1} + \mathbf{X}_{i,t} \gamma + \theta_{c,p} + \lambda_t + \Omega_o + \varepsilon_{ict} \quad (11)$$

其中  $\beta_3$  的系数与显著性体现了数字经济留居效应的高/低人力资本差异。

## 四、实证结果与分析

### （一）基本结果：数字经济的留城意愿效应

首先分析城市数字经济发展对农民工总体留城意愿的影响，结果报告于表 2 中。表 2 的（1）与（3）列报告了使用 2012-2018 年数据的分析结果。由于 2018 年 CMDS 未报告农民工流出地省份信息，（2）与（4）列在控制流出地省份固定效应后未包含 2018 年 CMDS 数据。考虑到核心解释变量城市数字经济水平的指数特征，采用其标准差变化进行效应解释。（1）与（2）列未加入控制变量，流入地城市 1 单位数字经济标准差的增加使农民工留城意愿提高 2.7%-2.9%。（3）与（4）列进一步控制农民工个人、家庭与城市层面控制变量，单位数字经济标准差效应有所减小，对农民工留城意愿的促进作用在 2% 左右。可以发现，是否控制流出地省份固定效应对数字经济系数估计的影响非常微弱，表明缺失 2018 年农民工数据不会影响基本结果的稳健性。此外，（3）与（4）列估计中，城市人均 GDP 增加 1%，农民工留城概率将显著增加 4.2%-4.5%。由于城市经济水平与数字经济水平正向相关，缺失城市人均 GDP 变量将导致对数字经济效应的高估。其他控制变量的结果均符合预期。

表 2 数字经济的留城意愿效应

变量	被解释变量：农民工留城意愿			
	(1)	(2)	(3)	(4)
数字经济总指数	0.00058*** (0.00009)	0.00060*** (0.00009)	0.00042*** (0.00007)	0.00043*** (0.00007)
性别			-0.00125 (0.00169)	-0.00487*** (0.00173)
年龄			0.02064***	0.01983***

			(0.00080)	(0.00085)
年龄 <sup>2</sup>			-0.00027***	-0.00026***
			(0.00001)	(0.00001)
受教育年限			0.01825***	0.01694***
			(0.00062)	(0.00052)
民族			-0.01423	-0.00389
			(0.00972)	(0.00588)
婚姻状况			0.06694***	0.05959***
			(0.00425)	(0.00418)
家庭随迁人数			0.06328***	0.06714***
			(0.00119)	(0.00120)
流动时长			0.01567***	0.01589***
			(0.00025)	(0.00028)
迁移类型：省内 跨市			-0.03368***	-0.03344***
(市内跨县为参 照组)			(0.00388)	(0.00384)
迁移类型：跨省			-0.11138***	-0.09945***
			(0.00419)	(0.00433)
人均 GDP 对数			0.04242***	0.04517***
			(0.00673)	(0.00706)
数字经济总指数 标准差	46.654	48.713	46.631	48.688
年份固定效应	是	是	是	是
流入地地级市固 定效应	是	是	是	是
流出地省份固定 效应	否	是	否	是
观测值	528,799	465,433	527,528	464,327
R <sup>2</sup>	0.070	0.075	0.153	0.157

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。

表 2 中的基本结果初步验证了数字经济发展对农民工总体留城意愿的促进作用，但上述估计仍可能面临因遗漏变量导致的内生性问题。依据实证分析部分中的阐述，通过构建全国城市平均数字经济水平与外生的城市坡度起伏作为工具变量，以进行两阶段工具变量估计。我们使用流入地县域平均坡度代表地形特征，其会通过影响数字基础设施的建设影响城市数字经济发展。为检验以县域坡度均值构建工具变量的稳健性，我们进一步以县域坡度标准差构建工具变量进行分析。表 3 汇报了工具变量估计的结果。

在第一阶段中，两类工具变量与县域数字经济水平均显著负相关，这意味着平均坡度较高、地形起伏较大的城市内数字化基础设施的建设更为困难，城市数字经济发展较缓，工具变量的相关性得到验证。进一步地，借助 CMDS 数据的混合截面特征，我们对 2011-2017 年样本县域地形起伏与数字经济水平间的相关性进行逐年检验。附表 A4 报告了逐年截面检验的结果，可以发现，无论以坡度的均值还是标准差作为城市地形起伏的代表，样本期各年内城市坡度与数字经济水平均显著负相关，这意味着工具变量识别的外生来源并非由特定年份所驱动，进一步验证了工具变量的相关性条件。在第一阶段中，F 值大于弱工具变量经验法则下的临界值 10，且 Cragg-Donald Wald F 值远大于 10%显著性水平下 Stock-Yogo 弱识别检验的临界值，表明使用坡度均值与坡度标准差构建工具变量进行估计不存在弱工具变量问题。在第二阶段中，城市数字经济发展对农民工留城意愿的促进效应仍是显著的。单位数字经济标准差的增加引致农民工留城意愿提高 8.7%-10.3%，数字经济发展对农民工留居意愿的促进作用非常可观。表 3 中工具变量估计结果是表 2 中 LPM 估计的约 2 倍，表明潜在的遗漏变量问题将导致对数字经济效应的低估。

工具变量的排他性条件要求在控制了相应的控制变量与固定效应后，其仅通过影响县域数字经济发展对农民工留居意愿产生影响。我们对两个潜在的竞争性影响渠道进行检

验，以考察表 3 中的工具变量分析是否违背排他性条件<sup>®</sup>。第一为城市环境质量渠道。尽管城市坡度为外生的地理条件，但其可能影响污染物的扩散条件，从而直接影响城市的空气质量。此外，坡度可能会影响城市交通基础设施建设，间接影响通勤出行过程中的污染排放（孙传旺等，2019）。由于空气污染会损害居民的心理健康与主观幸福感（Zhang et al., 2017; Chen et al., 2018），这可能进一步影响农民工的留居意愿。使用  $0.01^{\circ} \times 0.01^{\circ}$  分辨率的栅格层面地表 PM2.5 数据，并根据县域行政边界进行解析，我们得到了县域-年维度地表 PM2.5 数据。附表 A5 的（1）与（2）列表明，无论以坡度均值还是标准差构建工具变量，工具变量均与县域 PM2.5 不显著相关，这排除了工具变量通过空气质量影响农民工留居意愿的可能。第二个竞争性渠道为城市内逐年变化的移民社会网络。全国层面的数字经济发展促进农村居民为更好的就业机会与更高的预期收入流入城市（马述忠和胡增玺，2022），若县域特定的地形条件影响了农村居民对流入地的偏好，则工具变量将与城市内的移民社会网络相关。以“老乡会”为代表的移民社会组织密度对农村移民延续乡俗惯性、提升生活质量紧密相关（李树和邓睿，2021），进而会影响农民工的长期留居意愿。我们以县域人口密度的年增长率作为城市内移民社会网络变化的近似，以检验工具变量与移民社会网络的相关性。附表 A5 的（3）与（4）列表明，工具变量与县域人口密度增长率的相关性并不显著，这排除了潜在的移民社会网络渠道。上述分析表明，工具变量不会通过影响城市环境质量与移民网络密度影响农民工的留居意愿，工具变量的排他性条件较好地得以满足。

此外，我们借鉴黄群慧等（2019）与赵涛等（2020）的研究，使用历史上 1984 年地级市人均固定电话数与邮电业务量重新构造工具变量进行分析。一方面，以固定电话与邮电业务为代表的历史电信基础设施会通过使用习惯与技术基础影响现代以互联网为代表的数字技术发展，满足工具变量的相关性条件。另一方面，随着信息技术的变革，历史固定电话及邮电业务对当前农民工长期留居意愿影响甚微，满足工具变量的排他性条件。我们以“全国城市数字经济指数均值 $\times$ 1984 年流入地地级市每百人固定电话数（或每万人邮电业务量）”作为数字经济发展的工具变量，估计结果汇报于附录表 A6 中。第一阶段结果表明，两类工具变量与数字经济发展显著正向相关，且统计检验表明不存在弱工具变量问题。在第二阶段中，数字经济发展促进了农民工的城市留居意愿，且效应非常显著。上述分析证实了以坡度构造工具变量分析的稳健性，意味着在克服相应的内生性问题后，城市数字经济的发展有助于提升农民工融入城市的意愿。

表 3 数字经济的留城意愿效应：工具变量估计

变量	第一阶段被解释变量：数字经济总指数	
	(1)	(2)
全国城市平均数字经济水平 $\times$ 流入地县域坡度均值	-0.00651*** (0.00119)	
全国城市平均数字经济水平 $\times$ 流入地县域坡度标准差		-0.00823*** (0.00153)
第一阶段 F 值	29.81	28.87
Cragg-Donald Wald F 值	9206.105	5614.015
	第二阶段被解释变量：农民工留城意愿	
	(1)	(2)
数字经济总指数	0.00211** (0.00092)	0.00179** (0.00091)
数字经济总指数标准差	48.688	48.688
控制变量	是	是
年份固定效应	是	是
流入地地级市固定效应	是	是

<sup>®</sup> 篇幅所限，未在正文中报告排他性检验的结果，具体结果参见附录表 A5。

流出地省份固定效应	是	是
观测值	464,327	464,327
R <sup>2</sup>	0.082	0.085

注：\*\*\*、\*\*、\*分别代表系数在 1%、5% 与 10% 水平上显著，标准误聚类在流入地县域层面。控制变量包括表 2 中所有控制变量，下同。

## （二）稳健性检验

此部分从多方面检验上述基本结果的稳健性。首先，在基准分析中，为保留尽可能多的样本进行分析，对于未来长期留城意愿回答“没想好”的样本，将其留城意愿赋值为 0。此处，若农民工对未来 5 年留在所在城市的回答为“没想好”，则剔除此样本。重新定义留城意愿变量后的工具变量估计结果见表 4 中（1）列。由于样本量的减少，数字经济系数的显著性有所下降，但其仍在 10% 水平上显著。重新定义农民工留城意愿变量后，数字经济的留居意愿促进作用仍然稳健。其次，考虑到模型设定中数字经济变量的滞后一年处理，上一年的数字经济指数可能无法作用于流动时间较短的农民工群体，导致潜在的结果偏差。在表 4 的（2）列中，剔除了流动时长不足 6 个月的样本进行重新估计。结果表明，剔除流动时间较短的样本后，数字经济的促进作用仍非常显著，系数大小与表 3 的（1）列中结果非常接近。此外，由于 CMDS 问卷的设置，在 2012-2015 年调查中，我们仅能获知农民工是否愿意长期留城。但对于不愿意留城的农民工样本，无法得知其打算回乡或继续流动。城市生活、就业的数字化发展可能增进农民工适应数字化的能力，从而促进其融入城市或继续流动以寻求更高的经济收益。数字经济的发展也可能使部分农民工感到难以适应，从而将他们“挤出”城市而寻求返乡。由于数字经济对两种不愿留城结果的作用机制并不相同，基准分析结果是两种效应的综合。2016 与 2017 年 CMDS 问卷中进一步区分了农民工不愿留城时的下一步打算，仅保不愿留城样本中下一步打算返乡的样本进行分析，剔除了下一步打算继续流动或没想好的样本。表 4 中（3）列的估计结果与基准结果仍非常接近<sup>⑨</sup>。在基准分析中，使用的数字经济变量来自北京大学数字金融研究中心发布的数字普惠金融指数，尽管我们论述了此指数是农民工所感知到的城市数字经济发展程度的主要表征，但仍可能担心此指数侧重反应数字普惠金融的发展，区别于城市数字经济整体发展。借鉴赵涛等（2020）的设定，从互联网发展与数字普惠金融角度对地级市数字经济发展进行综合测度。其中选取的变量包括互联网用户数、移动电话用户数、电信业务量、信息产业从业人数与城市普惠金融指数，通过对上述变量进行主成分分析，得到数字经济水平的综合度量。表 4 的（4）列表明，数字经济发展对农民工留居意愿仍具有显著的正向影响，基准分析的发现是稳健的。

表 4 稳健性检验：工具变量估计

变量	被解释变量：农民工留城意愿			
	（1） 剔除对长期留城“没想好”的样本	（2） 剔除流动时长不足 6 个月的样本	（3） 剔除未来打算流动到其他城市样本	（4） 依据赵涛等（2020）改变数字经济变量测度
数字经济指数	0.00156* (0.00087)	0.00196** (0.00085)	0.00186** (0.00080)	0.00180** (0.00090)
数字经济指数标准差	49.262	48.795	48.609	123.316
Cragg-Donald Wald F 值	8778.266	8532.895	9211.125	2266.724
控制变量	是	是	是	是

<sup>⑨</sup> 2016 年 CMDS 问卷中，此问题为 Q305。2017 问卷中，此问题为 Q317。2016 年回答继续流动的样本量为 2963，2017 年回答继续流动或没想好的样本量为 647。

年份固定效应	是	是	是	是
流入地地级市固定效应	是	是	是	是
流出地省份固定效应	是	是	是	是
观测值	308,443	413,097	460,717	417,447
R <sup>2</sup>	0.067	0.071	0.083	0.072

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。(1) -

(3) 列数字经济变量为流入地县域北大普惠金融总指数，(4) 数字经济变量根据赵涛等（2020）对各组成维度进行主成份分析测算得到。数字经济指数的工具变量为全国城市数字经济指数均值×流入地县域平均坡度。

### (三) 数字经济留城意愿效应：人力资本异质性

在数字经济背景下，不同人力资本农民工适应数字化生活的能力存在差异，高人力资本人群更能借助数字技术降低消费过程中的搜寻成本。此外，数字经济发展下不同人力资本农民工相对生产率的变化，以及城市层面劳动力需求结构改变将通过影响均衡工资进而影响两类农民工的留城意愿。此部分中，将使用（10）式的设定，分析数字经济发展对两类人力资本农民工留城意愿的异质性影响。表 5 的（1）列汇报了数字经济影响的 LPM 估计结果。数字经济发展对高人力资本农民工留城意愿的促进作用更强，单位数字经济标准差增加使高人力资本农民工的留城意愿提高 2.6%。欣喜的是，数字经济的发展没有降低低人力资本农民工的留城意愿，将其“挤出”城市。单位数字经济标准差对于低人力资本农民工留城意愿的促进作用为 1.9%，在 1%水平上显著，这体现了数字经济对低人力资本农民工群体融入城市的包容性作用。本文设定高中及以上农民工为高人力资本农民工，出于对人力资本标准的设定可能影响结果的担心，进一步放松相应设定，构建数字经济指数与七类受教育程度的交互项进行异质性分析<sup>⑩</sup>。结果显示，数字经济对留城意愿的促进效应随农民工受教育程度增加而逐渐增强，对于受教育程度最低的“未上过学”与“小学”农民工估计系数不显著。数字经济对不同人力资本农民工融入城市的包容性作用再次得到证实。

出于对 LPM 估计结果存在内生性问题的担心，使用全国城市平均数字经济水平×流入地县域平均坡度作为数字经济指数的工具变量再次进行估计，结果汇报于表 5 的（2）列。可以发现，使用工具变量估计后，数字经济对两类农民工留城意愿的促进效应较 LPM 估计均增强，单位数字经济标准差的增加使高人力资本农民工留居意愿提高 10.8%。对影响进行人力资本异质性检验，结果表明其对高人力资本农民工促进作用显著更强。表 5 的（1）列中数字经济留城意愿效应的差异为 0.7%，差异在 1%水平上显著。使用工具变量估计，差异进一步扩大至 1.3%，且仍是统计显著的。上述结果表明，随着城市数字经济的发展，人力资本禀赋更高的农民工群体更愿意留在城市、融入城市，成为数字化背景下城市组成的一份子，但其背后的影响机制仍有待于进一步分析。

表 5 数字经济留城意愿影响：农民工人力资本异质性

变量	被解释变量：农民工留城意愿	
	(1)	(2)
数字经济总指数 ×高人力资本农民工	0.00053*** (0.00007)	0.00221** (0.00091)
数字经济总指数 ×低人力资本农民工	0.00038*** (0.00007)	0.00195** (0.00091)
数字经济效应人力资本差异	0.00015***	0.00026***

<sup>⑩</sup> 篇幅所限，没有在正文中展示相应结果，请参见附录中表 A2。

数字经济总指数标准差	48.688	48.688
Cragg-Donald Wald F 值		4576.170
控制变量	是	是
年份固定效应	是	是
流入地地级市固定效应	是	是
流出地省份固定效应	是	是
观测值	464,327	464,327
R <sup>2</sup>	0.157	0.083

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。控制变量中包括了农民工的受教育年限，以允许人力资本对留城意愿的直接影响，后续分析同。数字经济效应人力资本差异检验依据（11）式的设定，下同。

#### （四）机制分析

根据第二部分的理论分析，数字经济对留城意愿影响的人力资本异质性可被分解为生活方式变化与就业收入效应两类机制，本部分将对此进行实证检验。

##### 1. 生活方式变化

在理论分析中，家庭效用由消费实现，数字经济引致的消费便利性差异体现在两类农民工消费时搜寻成本的相对差异。禀赋更高的高人力资本农民工通常能更快地学习、掌握数字化的线上消费技能，借助数字金融降低消费中成本以实现福祉的提升。根据（8）式，相较低人力资本农民工，高人力资本农民工在数字化下更低的消费成本将导致两类群体间留城意愿差异的扩大。由于 CMDS 数据库中缺乏衡量农民工消费成本变量，我们参照张勋等（2020），使用覆盖广度指数、使用深度指数中支付业务子指数与数字化程度指数三类可以反映支付便利性的指数，考察支付便利性对两类农民工留城意愿的影响<sup>11</sup>。

表 6 汇报了三类支付便利性指数对高/低人力资本农民工留城意愿影响的估计结果。

（1）-（3）列中，反映支付便利性的三类数字经济指数均对两类人力资本农民工留城意愿具有显著促进作用。数字金融具有普惠性特征，被传统金融服务排除在外的低人力资本农民工群体同样可以受益于数字金融的发展享受到便利的消费，最终体现为三类指数对其留居意愿的促进。分两类农民工来看，支付便利性的留居意愿效应对于高人力资本农民工从量级和显著性均更强。组间差异检验表明，两类群体间的差异在 1%水平上显著。根据子指数的不同，单位支付便利性指数标准差对高/低人力资本留城意愿的影响差异在 1.3%-1.7%之间，人力资本异质性背后的生活便利性机制得到证实。

表 6 机制分析：支付便利性

变量	被解释变量：农民工留城意愿		
	（1）	（2）	（3）
覆盖广度指数 ×高人力资本农民工	0.00274** (0.00122)		
覆盖广度指数 ×低人力资本农民工	0.00249** (0.00123)		
使用深度：支付业务指数 ×高人力资本农民工		0.00163** (0.00072)	
使用深度：支付业务指数 ×低人力资本农民工		0.00134* (0.00072)	

<sup>11</sup> 覆盖广度体现了城市电子账户的覆盖程度，支付业务指数体现了使用电子账户进行支付的用户总量与活跃度，数字化程度指数计算中使用以移动支付占比等指标，三类指标均可反映支付便利性。更为细致的论证请参见郭峰等（2020）与张勋等（2020）。

数字化程度指数 ×高人力资本农民工			0.00122*** (0.00043)
数字化程度指数 ×低人力资本农民工			0.00100** (0.00043)
数字经济效应人力资本差异	0.00025***	0.00029***	0.00021***
支付便利性指数标准差	50.992	56.964	74.524
Cragg-Donald Wald F 值	2798.283	7093.702	5321.421
控制变量	是	是	是
年份固定效应	是	是	是
流入地地级市固定效应	是	是	是
流出地省份固定效应	是	是	是
观测值	464,327	464,327	464,327
R <sup>2</sup>	0.077	0.085	0.084

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。数字经济总指数的工具变量为全国城市平均覆盖广度/使用深度中支付业务/数字化程度指数×流入地县级市平均坡度。

数字经济发展引致生活方式变化的一个延伸是其对农民工主观被接纳感受的影响。数字生活下，农民工日常消费、事务办理、休闲等活动更多通过线上方式实现，这将减少农民工日常生活中感知到的歧视。此外，掌握数字生活技能，也将增进农民工同质于城市人的感知。被歧视感知下降、被城市接纳感知上升有助于提升农民工的幸福感，进而提升农民工的留城意愿（卢海阳和张敏，2020）。我们使用 2012、2013、2014 与 2017 年 CMDS 数据对此机制进行检验<sup>12</sup>。若农民工对“我感觉本地人看不起外地人”的回答为完全同意或基本同意，则将感知被歧视变量赋值为 1，若回答为完全不同意或者不同意，将感知被歧视变量赋值为 0。若农民工对“我觉得本地人愿意接受我成为其中一员”的回答为完全同意或基本同意，将感知被接纳变量赋值为 1，若回答为完全不同意或不同意，将感知被接纳变量赋值为 0。分析结果表明，数字经济对两类农民工被歧视感知的估计系数为负，对两类农民工被接纳感知的估计系数为正，但系数均是统计不显著的<sup>13</sup>。这意味着，从定性角度，城市数字经济发展确实有助于通过影响农民工的心理感知促进其接纳城市、融入城市。然而，数字经济降低歧视的效应在本文中并未完全显现，仍有待于后续的进一步探索。

## 2. 就业收入机制

数字经济将通过就业收入机制影响两类人力资本农民工的工资收入，进而影响农民工的留城意愿。首先考察数字经济对农民工个体月工资收入的影响及影响的人力资本差异。为剔除价格因素对工资收入的影响，以 2011 年为基期对各年的工资进行了 CPI 平减处理。此外，为了剔除极端工资的影响，将工资收入进行 1%缩尾。在农民工个体月工资收入之外，进一步选取了家庭月收入与家庭月消费进行稳健性检验，对此变量同样进行 CPI 平减与 1%缩尾处理。表 7 汇报了数字经济发展对农民工个体收入及对家庭收入与消费的影响。表 7 中（1）列对有就业农民工样本的分析表明，数字经济发展带来农民工工资收入的大幅增长，单位数字经济标准差的增加使农民工月工资收入显著提高 19.7%。（2）列人力资本异质性分析结果表明，数字经济的增收作用对两类农民工均成立，且对低人力资本农民工效应略强。单位数字经济标准差的提高将使低人力资本农民工工资收入相对增加 1.1%，此效应相对总体增收效应在量级上较小。（3）列与（4）列使用农民工家庭月收入对数作为被解释变量，数字经济对家庭收入的促进作用仍然显著。单位数字经济标准差对农民工家庭

<sup>12</sup> 2012 年 CMDS 相应问题为 Q405，2013 年相应问题为 Q530，2014 年相应问题为 Q518，2017 年相应问题为 Q503。

<sup>13</sup> 篇幅所限，没有展示具体结果，请参见附录中表 A3。



月收入的促进作用为 12.8%，且效应在两类农民工家庭中几乎完全一致。（5）列与（6）列进一步验证了数字经济水平对家庭消费的促进作用，单位数字经济标准差的增加使农民工家庭月消费显著提升 16.1%，且数字经济对消费的促进作用在不同人力资本家庭中不存在显著差异。

表 7 数字经济发展与农民工收入及消费

变量	解释变量：农民工个体月工资收入对数		被解释变量：农民工家庭月收入对数		被解释变量：农民工家庭月消费对数	
	(1)	(2)	(3)	(4)	(5)	(6)
数字经济总指数	0.00404*** (0.00081)		0.00262*** (0.00081)		0.00331*** (0.00118)	
数字经济总指数×高人力资本农民工		0.00394*** (0.00081)		0.00263*** (0.00080)		0.00333*** (0.00118)
数字经济总指数×低人力资本农民工		0.00417*** (0.00083)		0.00261*** (0.00082)		0.00327*** (0.00120)
数字经济效应人力资本差异		-0.00022***		0.00002		0.00006
数字经济总指数标准差	48.651	48.651	48.667	48.667	48.672	48.672
Cragg-Donald Wald F 值	9127.278	4537.130	9218.563	4582.420	9229.732	4587.859
控制变量	是	是	是	是	是	是
年份固定效应	是	是	是	是	是	是
流入地地级市固定效应	是	是	是	是	是	是
流出地省份固定效应	是	是	是	是	是	是
观测值	402,295	402,295	463,935	463,935	463,988	463,988
R <sup>2</sup>	0.073	0.072	0.173	0.173	0.210	0.211

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。数字经济总指数的工具变量为全国城市平均数字经济指数×流入地县级市平均坡度。

我们进一步从广延边际与集约边际角度分解数字经济发展对两类人力资本农民工就业收入的影响

**广延边际角度。**数字经济发展会通过替代效应、创造效应与补偿效应综合影响农民工在流入城市中的就业与否<sup>14</sup>。数字经济发展所催生的数字化生产资本（如机器人等）将替代传统生产模式下的劳动力需求，引致企业对农民工需求的总体降低。而在农民工内部，由于操作数字化设备需要一定的数字生产技能，低人力资本农民工受到的就业冲击较高人力资本农民工更大，这可能导致低人力资本农民工退出数字化程度较高的行业，甚至被“挤出”城市。其次，数字经济进步带来的生产效率提升将引致总体生产规模的扩大，弥补数字经济发展对劳动力的替代作用，表现为数字经济的补偿效应。此外，数字经济发展也将催生一系列全新的就业岗位，这些岗位通常需要就业者具有更加专业化的数字生产能力，对农民工的人力资本也提出了更高的要求。考虑到 CMDS 数据的混合截面特征，我们在更为宏观的地级市及地级市-行业层面检验数字经济发展对农民工需求的人力资本结构的影响<sup>15</sup>，表 8 汇报了 LPM 与工具变量估计的结果。结果表明，数字经济发展引致就业需求结构向高人力资本农民工偏移，即（8）式中  $L_H/L_L$  的增加，这将拉大高人力资本农民工与

<sup>14</sup> 感谢审稿专家指出数字经济的广延边际影响是多种效应的综合结果，应对此进行更为充分的讨论。

<sup>15</sup> 如果按县级市进行加总，县级市及县级市-行业观测数量过少，影响分析的统计效率。

低人力资本农民工间的就业收入差距<sup>16</sup>。数字经济发展对农民工就业收入的广延边际影响是上述替代效应、创造效应与补偿效应的综合，限于本文的数据结构，我们无法对各效应进行详细分解。较多研究从理论与实证角度对数字经济发展的就业效应进行了探讨（赵昱名与黄少卿，2020；陈富贵等，2022；胡拥军与关乐宁，2022；郭东杰等，2022），对影响效应的具体分析仍有待后续的进一步研究。值得说明的是，由于农民工通常在小企业中就业，进入大企业比较困难，此处的分析着重于吸纳农民工的小微企业，可能区别于城市中的企业整体。此外，我们的分析限定于农民工样本，城市中其他居民的就业需求结构受数字经济发展的影响可能与此存在差异。

表 8 就业收入的广延边际：人力资本需求结构

变量	被解释变量：地级市层面高人力资本农民工/低人力资本农民工比例		被解释变量：地级市-行业层面高人力资本农民工/低人力资本农民工比例	
	(1)	(2)	(3)	(4)
地级市数字经济总指数	0.00397** (0.00171)	0.04094** (0.01831)	0.00647*** (0.00207)	0.61591 (3.94432)
地级市人均 GDP 对数	-0.10908* (0.06515)	-0.27802*** (0.10715)	-0.08716 (0.08453)	-3.69606 (23.43029)
数字经济总指数标准差	57.351	57.351	39.556	39.556
Cragg-Donald Wald F 值		32.587		0.289
年份固定效应	是	是	是	是
流入地地级市固定效应	是	是	是	是
行业固定效应			是	是
观测值	1,895	1,895	13,978	13,978
R <sup>2</sup>	0.394	-0.270	0.273	-3.750

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地地级市层面。地级市数字经济总指数的工具变量为全国城市平均数字经济指数×流入地地级市平均坡度。（1）与（2）列观测单位为地级市-年，（3）与（4）列观测单位为地级市-行业-年。

进一步分析数字经济发展对不同行业农民工人力资本需求结构的变化。根据 CMDS 中对农民工就业行业的分类，构建 19 个行业与数字经济指数的交互项进行行业异质性分析<sup>17</sup>，图 1 展示了相应结果。可以发现，数字经济发展引致劳动力需求结构偏向高人力资本农民工的行业主要为金融、信息技术服务、教育、科研与技术服务、卫生和社会工作等现代服务业行业（魏东霞和陆铭，2021），此类行业对于总体农民工进入壁垒较高，就业人数相对较少。农民工就业较为集中的第一产业、第二产业与传统服务业行业在数字经济发展下的高人力资本偏向并不明显，这从一定程度上解释了表 7 中数字经济对两类农民工收入影响差异微弱的发现。

根据魏东霞和陆铭（2021）的行业分类，我们将 19 个农民工就业行业划分为第一产业、第二产业、传统服务业与现代服务业，分析数字经济发展对两类人力资本农民工就业行业选择的影响，表 A7 汇报了具体的分析结果。结果表明，在数字经济的就业创造与补偿作用下，其显著提升了农民工就业于现代服务业的概率。然而，此种影响存在人力资本异质性，数字经济对高人力资本农民工的高数字技能就业促进作用显著更强。尽管低人力资本农民工进入现代服务业的概率有所提升，部分缺乏数字化生产技能的低人力资本农民工将被分流至第二产业中。随着数字经济的发展对制造业低端岗位的毁灭（赵昱名和黄少卿，2020），这些低人力资本农民工的福祉将受到损害，甚至被进一步“挤出”城市。限于

<sup>16</sup> 表 8 的（4）列城市-行业层面工具变量分析未通过弱工具变量检验，对其结果应谨慎解读。

<sup>17</sup> 由于行业 20-国际组织的就业人数过少，不对此行业进行分析。

CMDS 数据的混合截面特征，此处的就业行业分析仅提供了支持性证据，但其发现与上述就业收入的广延边际分析一致，并提供了数字经济发展对低技能工作岗位存在替代效应的证据。

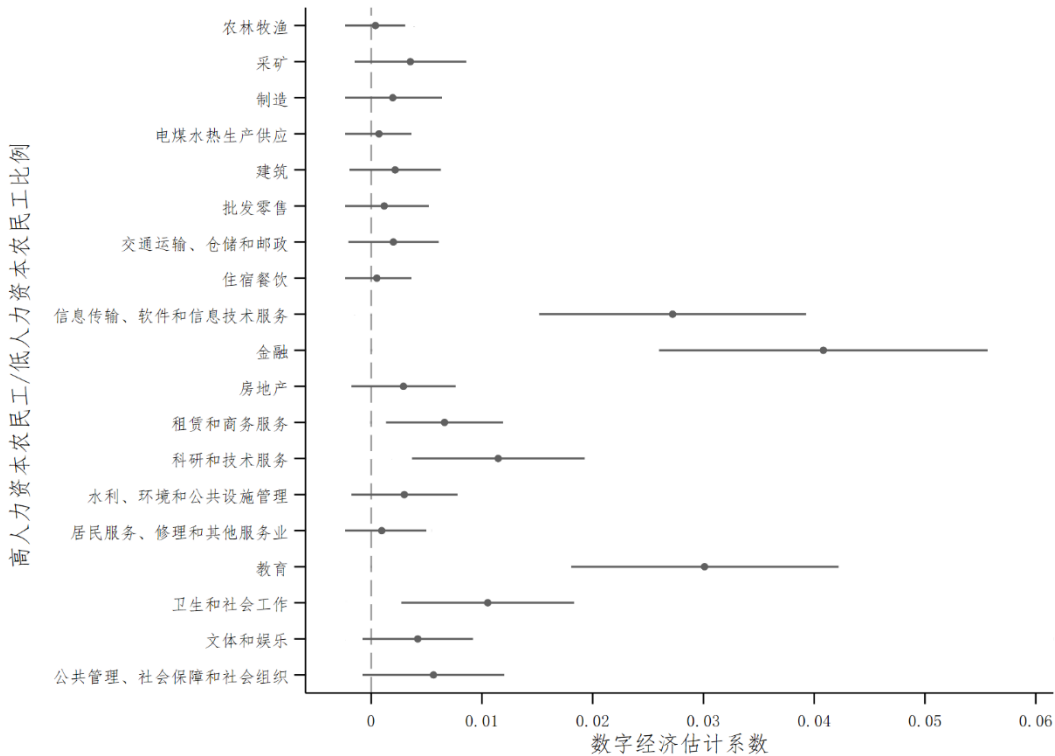


图 1 人力资本需求结构影响：行业异质性

**集约边际角度。**随着数字经济发展，两类人力资本农民工间相对生产率的变化同样对工资差异产生影响。以受教育年限为代表的人力资本水平一定程度上反映了农民工的认知能力，但对于数字化生产，非认知能力的习得同样非常重要。与通过在学校学习获得的人力资本不同，适应数字化生产的非认知能力可以通过社会学习实现，且不完全取决于受教育程度高低。魏东霞和陆铭（2021）的研究表明，受教育年限较低但进城较早的农民工，可以在城市生活、工作过程中通过社会学习实现收入的增长。在本文中，学习效应将通过调节数字经济对农民工间相对生产率的影响从而影响工资收入。若低人力资本农民工在学习效应下生产率提升相对更快，则数字经济发展将通过相对生产率机制缩小高/低人力资本农民工间的收入差异。遗憾的是，CMDS 数据库中无直接衡量个体生产率的指标，无法直接检验相对生产率机制。作为间接检验，我们使用农民工流动时长作为其学习效应的度量，通过流动时长将有就业农民工样本均分为三组，观察在不同组别中数字经济发展对工资收入影响的人力资本差异。在三组样本中，流动时间最短组流动时间均值为 0.86 年，流动时间中等组均值为 3.28 年，最长组均值为 10.37 年。不同组别之间流动时长差异较大，表明学习效应的离散程度较大。在分析中，控制了城市层面就业农民工人力资本结构变量，以切断数字经济通过广延边际机制对工资的影响。根据理论部分的分析，此时数字经济将通过劳动生产率的集约边际机制影响农民工工资差异。表 9 汇报了依流动时间分组回归的结果。可以发现，在控制了城市层面劳动力需求结构后，随着流动时间的增加，低人力资本农民工相对工资收入增长逐渐提升。单位数字经济指数的相对增收效应由 0.9%（表 9（1）列）逐渐增加至 1.3%（表 9（3）列），且统计显著性也逐渐增强。这意味着，随着流动时间的增加，农民工逐渐通过社会学习增进其数字化生产技能，低人力资本农民工实

现了更快的劳动生产率进步，并以此获得了更快的工资收入增长。

表 9 就业收入的集约边际：生产率相对变化

变量	被解释变量：农民工个体月工资收入对数		
	(1) 流动时长：最短	(2) 流动时长：中等	(3) 流动时长：最长
数字经济总指数×高人力资本农民工	0.00288*** (0.00082)	0.00370*** (0.00090)	0.00476*** (0.00123)
数字经济总指数×低人力资本农民工	0.00307*** (0.00082)	0.00393*** (0.00091)	0.00503*** (0.00124)
地级市层面高人力资本农民工/低人力资本农民工比例	0.02298 (0.02599)	0.06254*** (0.02383)	0.06730** (0.02736)
数字经济效应人力资本差异	-0.00019**	-0.00024**	-0.00027***
数字经济总指数标准差	47.512	48.919	49.323
Cragg-Donald Wald F 值	1465.844	1441.265	1480.544
控制变量	是	是	是
年份固定效应	是	是	是
流入地地级市固定效应	是	是	是
流出地省份固定效应	是	是	是
观测值	132,909	134,453	134,565
R <sup>2</sup>	0.09656	0.07528	0.04837

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地地级市层面。数字经济总指数的工具变量为全国城市平均数字经济指数×流入地县域平均坡度。

对就业收入机制的分析表明，尽管数字经济发展引致宏观层面对农民工劳动力需求偏向高人力资本，但此种偏向主要集中于农民工就业较少的现代服务业中。此外，低人力资本农民工通过学习效应实现学校教育以外非认知技能的增进，相对高人力资本农民工具有更高的生产率进步，这缩小了两类人群间工资收入差异，数字经济对农民工群体的包容性发展得以实现。考虑到数字经济对两类农民工工资影响几乎无差异，前文中发现的数字经济对农民工留城意愿的人力资本异质性结果主要被支付便利性机制所解释。

五、结论与展望

本文将 2012-2018 年流动人口动态监测数据与北京大学数字普惠金融指数匹配形成多年混合截面数据，首次在县域层面分析了城市数字经济发展对流动农民工长期留城意愿的影响。使用全国城市平均数字经济指数×流入地县域平均坡度作为数字经济的工具变量，估计结果表明城市数字经济发展对农民工总体的留居意愿具有显著的促进作用，单位数字经济指数标准差增加的促进效应为 10.3%。数字经济的促进作用对受教育程度在高中及以上的高人力资本农民工更强，但对低人力资本农民工的促进效应同样显著，从城市融入角度印证了数字经济对两类农民工群体城市发展的包容性。进一步的机制分析表明，数字经济时代下，高人力资本农民工拥有的更强的数字生活技能，能更好地利用数字经济带来的支付便利性是人力资本异质性的主要成因。在就业收入机制上，数字经济发展使城市与城市-行业宏观层面农民工劳动力需求向高人力资本农民工偏移，拉大了高/低人力资本农民工间的工资差异。然而在就业群体内部，低人力资本农民工通过社会学习获得了生产率的更快提高，这弥补了两类农民工间的工资差异。综合就业收入效应的两种机制，数字经济对两类农民工收入影响的差异并不明显。

本文从农村移民角度为数字经济时代下城乡融合的探讨提供了“后半程”视角，具有鲜明的现实意义。加强农民工对城市的认同感、归属感，使农民工在数字时代下分享城市数字经济建设的成果是实现“以人为本”的新型城镇化建设的关键。本文的政策启示在于，应

采用宣传、培训等方式增强农民工的数字技能，使农民工能充分适应城市中数字化生活与生产方式。同时，数字经济的发展使越来越多的生活服务与公共服务由线下转移至线上。应降低相应服务的使用成本与准入成本，借助数字方式减缓农民工群体与城市居民间公共服务不均等问题，进而增强农民工融入城市的心理认同。

本文针对已流动至城市的农民工群体开展分析，得到了城市数字经济发展有利于农民工留居城市的积极发现，这意味着数字城市的建设可实现现代化与农民工福祉提升的双赢。然而，已经决定迁移的农民工可能本身适应数字化城市的能力就较强。由于选择效应的存在，本文的结论难以直接推广至所有农村居民。随着数字经济的进一步发展，未来城乡融合走向何处，仍是后续研究应关注的问题。此外，本文通过对高/低人力资本农民工的异质性分析，揭示了数字经济发展对农民工群体内部潜在的不平等影响，但未对农民工与城市居民间进行对比。由于城乡二元户籍制度仍在多地存在，即使是人力资本水平较高的农民工，可能也无法享受与城市居民对等的数字服务资源。思考现有户籍制度的改革方向，加快流动人口的市民化进程，使其“流向城市、认同城市、融入城市”，仍有待未来的进一步探索。

## 参考文献

- 柏培文,张云.数字经济、人口红利下降与中低技能劳动者权益[J].经济研究,2021,56(05):91-108.
- 陈晓东,杨晓霞.数字经济发展对产业结构升级的影响——基于灰关联熵与耗散结构理论的研究[J].改革,2021(03):26-39.
- 陈贵富,韩静,韩恺明.城市数字经济发展、技能偏向型技术进步与劳动力不充分就业[J].中国工业经济,2022(08):118-136.
- 郭凯明.人工智能发展、产业结构转型升级与劳动收入份额变动[J].管理世界,2019,35(07):60-77+202-203.
- 郭峰,王靖一,王芳,孔涛,张勋,程志云.测度中国数字普惠金融发展:指数编制与空间特征[J].经济学(季刊),2020,19(04):1401-1418.
- 郭峰,熊云军.中国数字普惠金融的测度及其影响研究:一个文献综述[J].金融评论,2021,13(06):12-23+117-118.
- 郭东杰,周立宏,陈林.数字经济对产业升级与就业调整的影响[J].中国人口科学,2022(03):99-110+128.
- 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,(08):5-23.
- 何宗樾,宋旭光.数字金融发展如何影响居民消费[J].财贸经济,2020,41(08):65-79.
- 胡拥军,关乐宁.数字经济的就业创造效应与就业替代效应探究[J].改革,2022(04):42-54.
- 卢晶亮.资本积累与技能工资差距——来自中国的经验证据[J].经济学(季刊),2017,16(02):577-598.
- 卢海阳,张敏.融合策略、歧视感知与农民工幸福感——基于福建省 2393 个农民工的调查数据[J].社会发展研究,2020,7(02):90-109+243-244.
- 李树,邓睿.乡俗惯性与农民工主观生活质量——基于流动人口动态监测的经验研究[J].财经研究,2021,47(11):109-123.
- 李晓钟,李俊雨.数字经济发展对城乡收入差距的影响研究[J].农业技术经济,2022(02):77-93.
- 李天成,孟繁邨,李世杰,高健.技术进步影响农民工就业和收入了吗——来自劳动力异质性视角下的微观证据[J].农业技术经济,2022(03):100-116.
- 马述忠,胡增玺.数字金融是否影响劳动力流动?——基于中国流动人口的微观视角[J].经济学(季刊),2022,22(01):303-322.
- 潘为华,贺正楚,潘红玉.中国数字经济发展的时空演化和分布动态[J].中国软科学,2021(10):137-147.
- 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(11):17-35.
- 孙传旺,罗源,姚昕.交通基础设施与城市空气污染——来自中国的经验证据[J].经济研究,2019,54(08):136-151.
- 王林辉,胡晟明,董直庆.人工智能技术会诱致劳动收入不平等吗——模型推演与分类评估[J].中国工业经济,2020(04):97-115.
- 魏东霞,陆铭.早进城的回报:农村移民的城市经历和就业表现[J].经济研究,2021,56(12):168-186.
- 夏怡然,陆铭.城市间的“孟母三迁”——公共服务影响劳动力流向的经验研究[J].管理世界,2015(10):78-90.
- 肖子华,徐水源,刘金伟.《中国城市流动人口社会融合年度报告》,2018,载肖子华主编《中国城市流动人口社会融合评估报告 No.1》社会科学文献出版社.
- 谢绚丽,沈艳,张皓星,郭峰.数字金融能促进创业吗?——来自中国的证据[J].经济学(季刊),2018,17(04):1557-1580.
- 许宪春,张美慧.中国数字经济规模测算研究——基于国际比较的视角[J].中国工业经济,2020(05):23-41.

- 杨菊华.中国流动人口的社会融入研究[J].中国社会科学,2015(02):61-79+203-204.
- 易行健,周利.数字普惠金融发展是否显著影响了居民消费——来自中国家庭的微观证据[J].金融研究,2018(11):47-67.
- 朱明宝,杨云彦.城市规模与农民工的城市融入——基于全国 248 个地级及以上城市的经验研究[J].经济学动态,2016(04):48-58.
- 张勋,万广华,张佳佳,何宗樾.数字经济、普惠金融与包容性增长[J].经济研究,2019,54(08):71-86.
- 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.
- 张勋,杨桐,汪晨,万广华.数字金融发展与居民消费增长:理论与中国实践[J].管理世界,2020,36(11):48-63.
- 赵昱名,黄少卿.创造抑或毁灭:数字技术对服务业就业的双向影响[J].探索与争鸣,2020(11):160-168+180.
- 赵宸宇,王文春,李雪松.数字化转型如何影响企业全要素生产率[J].财贸经济,2021,42(07):114-129.
- 张文武,余泳泽.城市服务多样性与劳动力流动——基于“美团网”大数据和流动人口微观调查的分析[J].金融研究,2021(09):91-110.
- 祝仲坤.公共卫生服务如何影响农民工留城意愿——基于中国流动人口动态监测调查的分析[J].中国农村经济,2021(10):125-144.
- Acemoglu D, Autor D. Skills, tasks and technologies: Implications for employment and earnings[M] Handbook of labor economics. Elsevier, 2011, 4: 1043-1171.
- 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.
- Battisti M, Peri G, Romiti A. Dynamic effects of Co-Ethnic networks on immigrants' economic success[R]. National Bureau of Economic Research, 2016.
- Card D, Lemieux T. Can falling supply explain the rising return to college for younger men? A cohort-based analysis[J]. Quarterly Journal of Economics, 2001, 116(2): 705-746.
- Chen S., Oliva P, Zhang P. Air pollution and mental health: Evidence from China. NBER working paper, 2018, No. w24686.
- Imbens G, Angrist J. Identification and estimation of local average treatment effects. [J] Econometrica. 1994, 62(2): 467-475.
- Nunn N, Qian N. US food aid and civil conflict [J]. American Economic Review, 2014,104(6): 1630-1666.
- Roca J D L, Puga D. Learning by working in big cities[J]. Review of Economic Studies, 2017, 84(1): 106-142.
- Sin I, Stillman S, Fabling R. What drives the gender wage gap? Examining the roles of sorting, productivity differences, bargaining and discrimination[J]. Review of Economics and Statistics, 2020: 1-44.
- Zhang C, Xie Y. Place of origin and labour market outcomes among migrant workers in urban China[J]. Urban Studies, 2013, 50(14): 3011-3026.
- Zhang X, Zhang X, Chen X. Happiness in the air: How does a dirty sky affect mental health and subjective well-being? [J] Journal of Environmental Economics and Management, 2017, 85: 81-94.
- Zhang X, Tan Y, Hu Z, et al. The trickle-down effect of fintech development: From the perspective of urbanization[J]. China and World Economy, 2020, 28(1): 23-40.

## 附录

在附录中，我们将报告受篇幅限制，正文中未报告的相应结果，以方便审稿专家及读者审阅。

理论分析部分假设了农民工效用与留城意愿间单调递增关系，此部分使用 2012 年 CMDS 数据对此假设进行检验。2012 年 CMDS 问卷中 Q406 为“与流出地相比，您现在是否感觉幸福？1、很幸福；2、幸福；3、一般 4、不幸福；5、很不幸福”，使用此变量作为农民工效用的近似。将主观幸福感分别视为基数与序数进行分析。在基数分析中，以幸福感取值 1-5 分别代表很不幸福至很幸福。在序数分析中，将 Q406 中 4（不幸福）与 5（很不幸福）的回答作为参照组，考察其他类别回答与留城意愿间的关系。表 A1 展示了分析结果，（1）与（2）列为基数分析结果，其中（2）列控制了正文表 2 基准回归中所有的控制变量。（3）与（4）列为序数分析结果，（4）列同样加入控制变量。表 A1 的结果显示，农民工效用与留城意愿间关系是单调递增的，理论分析中的假定得到了数据的实证支持。

表 A1 农民工效用与留城意愿关系检验：LPM 估计

变量	被解释变量：农民工留城意愿			
	(1)	(2)	(3)	(4)
幸福感	0.18504***	0.16123***		
(1=很不幸福, 2=不幸福, 3=一般, 4=幸福, 5=很幸福)	(0.00408)	(0.00403)		
幸福感：一般 (参照组为不幸福)			0.09366*** (0.01683)	0.08117*** (0.01508)
幸福感：幸福			0.32504*** (0.01647)	0.28379*** (0.01508)
幸福感：很幸福			0.45355*** (0.01742)	0.39495*** (0.01636)
控制变量	否	是	否	是
年份固定效应	是	是	是	是
流入地地级市固定效应	是	是	是	是
流出地省份固定效应	是	是	是	是
观测值	80,104	79,957	80,104	79,957
R <sup>2</sup>	0.14541	0.19243	0.14874	0.19509

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。



此部分中，放松表 5 中对于高/低人力资本标准的划分，分别将七类农民工受教育程度与数字经济指数构建交互项，以进行数字经济对农民工留城意愿影响的教育程度异质性分析。表 A2 中（1）列为 LPM 估计结果，（2）列为工具变量估计结果，数字经济指数的工具变量为全国城市平均数字经济指数×流入地县域平均坡度。

表 A2 数字经济留城意愿影响异质性：受教育程度异质性

变量	被解释变量：农民工留城意愿	
	(1)	(2)
数字经济总指数 ×未上过学	0.00001 (0.00009)	0.00124 (0.00083)
数字经济总指数 ×小学	0.00010 (0.00008)	0.00135 (0.00084)
数字经济总指数 ×初中	0.00028*** (0.00007)	0.00161* (0.00087)
数字经济总指数 ×高中/中专	0.00062*** (0.00007)	0.00202** (0.00086)
数字经济总指数 ×大学专科	0.00118*** (0.00008)	0.00266*** (0.00085)
数字经济总指数 ×大学本科	0.00150*** (0.00008)	0.00309*** (0.00086)
数字经济总指数 ×研究生	0.00174*** (0.00014)	0.00320*** (0.00089)
数字经济总指数标准差	48.688	48.688
Cragg-Donald Wald F 值		1306.293
控制变量	是	是
年份固定效应	是	是
流入地地级市固定效应	是	是
流出地省份固定效应	是	是
观测值	464,327	464,327
R <sup>2</sup>	0.15753	0.08473

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。

表 A3 汇报了数字经济对高人力资本农民工与低人力资本农民工被歧视/被接纳感知的影响，所用数据为 CMDS 2012、2013、2014 与 2017 调查。对于被歧视感知变量，若农民工对“我感觉本地人看不起外地人”的回答为完全同意或基本同意，则将其赋值为 1。若回答为完全不同意或不同意，则赋值为 0。对于被接纳感知变量，若农民工对“我觉得本地人愿意接受我成为其中一员”的回答为完全同意或基本同意，则将其赋值为 1。若回答为完全不同意或不同意，则赋值为 0。表 A3 展示了使用全国城市平均数字经济指数×流入地县域平均坡度作为数字经济指数的工具变量估计结果。

表 A3 机制分析：被歧视与被接纳感知

变量	被解释变量： 被歧视感知	被解释变量： 被接纳感知	被解释变量：被 歧视感知	被解释变量：被接 纳感知
	(1)	(2)	(3)	(4)
数字经济总指数	-0.00725	0.00245		
	(0.00953)	(0.00290)		
数字经济总指数 ×高人力资本农民工			-0.00724	0.00245
			(0.00936)	(0.00281)
数字经济总指数 ×低人力资本农民工			-0.00693	0.00227
			(0.00924)	(0.00280)
数字经济总指数标准 差	32.704	32.704	32.704	32.704
Cragg-Donald Wald F 值	423.939	423.964	215.155	215.151
控制变量	是	是	是	是
年份固定效应	是	是	是	是
流入地省份固定效应	是	是	是	是
流出地省份固定效应	是	是	是	是
观测值	179,639	179,706	179,639	179,706
R <sup>2</sup>	-0.08187	-0.01773	-0.07786	-0.01591

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。数字经济总指数的工具变量为全国城市平均数字经济指数×流入地县级市平均坡度。

表 A4 汇报了 2011-2017 年间样本县域的坡度均值与标准差与数字经济水平间的相关性，以验证以坡度构建工具变量进行估计的结果不是由特定年份的偶然相关引致的。

表 A4 工具变量相关性逐年检验：县域坡度与数字经济水平

被解释变量：城市数字经济水平				
<b>Panel A</b>				
年份	2011	2012	2013	2014
县域平均坡度	-1.12153*** (0.08778)	-0.80923*** (0.27888)	-1.56891*** (0.10600)	-3.43705*** (0.23304)
观测值	908	59	1,103	1,149
R <sup>2</sup>	0.14693	0.10443	0.15997	0.14754
年份	2015	2016	2017	
县域平均坡度	-4.14797*** (0.25772)	-0.65639*** (0.06723)	-0.74992*** (0.08144)	
观测值	1,040	1,076	953	
R <sup>2</sup>	0.18851	0.05061	0.04897	
<b>Panel B</b>				
年份	2011	2012	2013	2014
县域坡度标准差	-1.60574*** (0.17467)	-1.15470** (0.50230)	-2.15063*** (0.21902)	-4.97777*** (0.45763)
观测值	908	59	1,103	1,149
R <sup>2</sup>	0.09043	0.08824	0.08640	0.09017
年份	2015	2016	2017	
县域坡度标准差	-6.47328*** (0.48958)	-0.82572*** (0.14343)	-0.99517*** (0.16599)	
观测值	1,040	1,076	953	
R <sup>2</sup>	0.13735	0.02300	0.02579	

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。Panel A 汇报了流入地县域平均坡度与城市数字经济水平间的相关性，Panel B 汇报了流入地坡度标准差与城市数字经济水平间的相关性，分析单元为县域城市。

此部分中，我们检验了工具变量“县域坡度×全国平均数字经济指数均值”是否与本文关心的消费便利性及就业收入机制外的其他因素相关，进而违反工具变量的排他性要求。我们关注了两类潜在的竞争性要素：以 PM2.5 为代表的城市环境质量，以及以城市人口密度增速为代表的移民网络。“县域-年”层面地表 PM2.5 数据来自华盛顿大学圣路易斯分校提供的 0.01° × 0.01° 分辨率栅格数据产品<sup>18</sup>，根据县域行政区划进行解析。“县域-年”人口密度数据来自 WorldPop 提供的栅格数据<sup>19</sup>，根据县域区划进行解析并计算年人口密度增速。

表 A5 工具变量排他性检验：县域-年层面

变量	被解释变量：地表 PM2.5（微克/立方 米）	被解释变量：地表 PM2.5（微克/立方 米）	被解释变量：人口 密度增长率（%）	被解释变量：人口 密度增长率（%）
	(1)	(2)	(3)	(4)
全国城市平均数字经济 水平 ×流入地县域坡度均值	-0.00009 (0.00031)		-0.00001 (0.00002)	
全国城市平均数字经济 水平 ×流入地县域坡度标准差		0.00004 (0.00074)		-0.00004 (0.00004)
年份固定效应	是	是	是	是
县域固定效应	是	是	是	是
观测值	4,815	4,815	4,780	4,780
R <sup>2</sup>	0.95172	0.95172	0.90372	0.90374

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在县域层面。

<sup>18</sup> 原始数据下载网址：<https://sites.wustl.edu/acag/datasets/surface-pm2-5/>

<sup>19</sup> 原始数据下载网址：<https://hub.worldpop.org/project/categories?id=18>

根据黄群慧（2019）与赵涛等（2020）的想法，重新构造工具变量进行分析。两个工具变量分别为“全国城市数字经济指数均值×1984 年流入地地级市每百人固定电话数量”与“全国城市数字经济指数均值×1984 年流入地地级市每万人邮电业务量”。

表 A6 数字经济的留城意愿效应：重新构造工具变量估计

变量	第一阶段被解释变量：数字经济总指数	
	(1)	(2)
全国城市平均数字经济水平 ×1984 年流入地地级市每百人固 定电话数量	0.02641***	
	(0.00607)	
全国城市平均数字经济水平 ×1984 年流入地地级市每万人邮 电业务量		0.00536***
		(0.00132)
第一阶段 F 值	18.91	16.37
Cragg-Donald Wald F 值	1740.879	2076.832
	第二阶段被解释变量：农民工留城意愿	
数字经济总指数	0.00424***	0.00361**
	(0.00159)	(0.00162)
数字经济总指数标准差	48.267	48.267
控制变量	是	是
年份固定效应	是	是
流入地地级市固定效应	是	是
流出地省份固定效应	是	是
观测值	420,959	420,959
R <sup>2</sup>	0.04207	0.05511

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。控制变量与表 2 相同。

此部分检验数字经济发展对两类人力资本农民工在城市中不同产业就业的广延边际影响。四类产业划分参照魏东霞和陆铭（2021）。具体地，第一产业包括农林牧渔业；第二产业包括采矿，制造，电煤水热生产供应，以及建筑业；传统服务业包括批发零售，住宿餐饮，居民服务、修理及其他服务业；现代服务业包括交通运输、仓储和邮政，信息传输、软件和信息技术服务，房地产，金融，租赁和商务服务，科研和技术服务，水利、环境和公共设施管理，教育，文体和娱乐，公共管理、社会保障和社会组织，卫生和社会工作行业。若有就业农民工在特定产业中就业，则被解释变量为 1，否则为 0。

表 A7 数字经济发展对农民工就业行业选择的影响

变量	被解释变量：在第一产业就业	被解释变量：在第二产业就业	被解释变量：在传统服务业产业就业	被解释变量：在现代服务业产业就业
	(1)	(2)	(3)	(4)
数字经济总指数×高人力资本农民工	-0.00029 (0.00027)	-0.00008 (0.00093)	-0.00161 (0.00102)	0.00197*** (0.00036)
数字经济总指数×低人力资本农民工	-0.00030 (0.00027)	0.00014 (0.00093)	-0.00154 (0.00103)	0.00170*** (0.00038)
数字经济效应人力资本差异	0.00001	-0.00022***	-0.00006	0.00028***
数字经济总指数标准差	48.654	48.654	48.654	48.654
Cragg-Donald Wald F 值	4550.741	4550.741	4550.741	4550.741
控制变量	是	是	是	是
年份固定效应	是	是	是	是
流入地地级市固定效应	是	是	是	是
流出地省份固定效应	是	是	是	是
观测值	403,038	403,038	403,038	403,038
R <sup>2</sup>	0.00034	0.01467	0.02074	0.02335

注：\*\*\*、\*\*、\*分别代表系数在 1%、5%与 10%水平上显著，标准误聚类在流入地县域层面。控制变量与表 2 相同。被解释变量中第一产业、第二产业、传统服务业、现代服务业的划分参照魏东霞和陆铭（2021）。