

数字经济背景下的人力资本 积累新模式探究

吴梦涛 张龙天 武康平

【摘要】新兴数字技术正重新定义着包括人力资本积累在内的经济生活各个方面,但目前还未有文献从理论层面探讨数字经济核心要素数据的人力资本效应。本文将数据作为一种新的关键要素引入人力资本的积累过程,在动态一般均衡模型框架内探究数字经济背景下人力资本的周期性以及数据要素对人力资本积累的影响,并讨论了目前文献中所提及的不同来源类型的数据在该过程中发挥的不同作用。研究发现,数据的引入提升了人力资本积累的效率,使其出现了从传统的逆周期变动向顺周期变动逐渐转换的趋势。进一步地,相较于来自消费的数据会受到隐私问题的制约,来自生产过程的数据则更能推动人力资本周期性的转变,且后者能够使经济达到更高的稳态福利水平。

【关键词】数字经济;数据要素;人力资本积累;动态一般均衡模型

【作者简介】吴梦涛,清华大学经济管理学院博士研究生,E-mail:wumt.18@sem.tsinghua.edu.cn;张龙天(通讯作者),中央财经大学国际经济与贸易学院助理教授,E-mail:zhanglongtian@cufe.edu.cn;武康平,清华大学经济管理学院教授,E-mail:wukp@sem.tsinghua.edu.cn。

【原文出处】《经济学报》(京),2023.2.1~27

【基金项目】本研究获得中央财经大学国际经济与贸易学院发展基金的资助。

0.引言

人力资本是一国综合实力和核心竞争力的体现,在推动经济社会实现可持续发展的过程中具有重要意义和巨大潜力。自改革开放以来,中国的“人才强国战略”取得了显著成效,以受教育年限和人口预期寿命为代表的人力资本水平得到了明显提升^①。中国是世界第一人口大国^②,然而不同于过去几十年的是,当前的经济发展形势正面临“人口红利”消失的重大挑战与由“人口红利”转向“人力资本红利”的历史转折点,劳动力市场开始呈现出数量供给减少但质量不断提升的新趋势(张同斌,2016;柏培文和张云,2021)。人力资本投资是提供人才支撑、释放人才红利、挖掘增长潜能的关键切入点,为回应新时代经济高质量发展的要求,推动人力资本积累势在必行。

与此同时,数字经济在中国经济生活中的作用日益凸显,以大数据技术为代表的一系列新兴技术逐步登上时代舞台。2017年,“数字经济”“人工智能”等新词汇首次被写入国务院政府工作报告,并在随后几年时间里,“数据”一词被中央文件多次提及,最终被正式纳入生产要素范畴^③。2022年12月,中共中央、国务院对外发布《关于构建数据基础制度更好发挥数据要素作用的意见》(简称“数据二十条”),旨在构建数据基础制度,以更好地发挥中国海量数据规模的优势、充分激活并释放数据要素潜能、做强做优做大数字经济,为中国经济增长注入澎湃动力。中国的数字经济发展迅猛^④,成为推动经济增长的新的引擎(蔡跃洲和牛新星,2021),此外,中国还拥有巨大的数据资源潜在储量^⑤。在这样的时代背景下,如何利用数据要素赋能经济高质量发展,已

经成为当前经济领域的重要研究课题之一。高质量发展意味着兼顾生产效率与“人民日益增长的美好生活需要”,数据作为数字经济的核心要素在提升生产效率、满足消费需求、实现供需匹配等多方面发挥着关键作用,有望与经济高质量发展的要求协调统一(蔡跃洲和马文君,2021)。

数字技术和海量数据固然重要,但人力资本才是数字经济发展的真正决定因素。新兴数字技术的创新,以及全球数据量的激增,为教育等众多行业带来了向数字化和信息化转型的新的变革机遇。在此背景下,各行各业对数字人才的需求急剧增加,人力资本的概念逐渐扩展到数字思维和特殊技能^⑥。为进一步引导和培养数字方向的新型人才,为中国经济高质量发展提供有力的人力资本支撑,促进数字经济与人力资本的融合至关重要。

有鉴于此,本文将致力于通过建立理论模型,针对数字经济背景下人力资本积累的新模式进行研究,并就数字经济时代经济建设中可能面临的一系列问题,提出相应的政策建议。具体而言,本文将主要参考数据经济长期增长理论领域中的两篇开创性文献Jones and Tonetti(2020)和Cong et al.(2021b)(以下称为CXZ模型),并沿用最新的工作论文Xie and Zhang(2023)中所提出的针对数据生产函数的建模方法,依托人力资本投资理论将数据要素纳入包含教育投资的人力资本积累方程,从而在动态一般均衡模型的分析框架下,探究人力资本的周期性以及数据要素参与人力资本积累过程的新模式。研究发现,当经济中发生正向全要素生产率冲击时,产出增加且数据产生量也随之增加,数据要素在经济中的使用水平明显提升,积极推动人力资本积累进程,使得人力资本存量由传统的逆周期变动趋势转变为顺周期趋势。因此,数据要素的加入能够大幅提升人力资本积累的效率,促使人力资本顺周期变动。

此外,本文还通过研究和对比文献中关于数据要素的不同解读,通过不同的组合形式描绘出人力资本积累由传统的逆周期性向顺周期性发展的变化趋势。以往有关人力资本和经济周期的大部分文献都指出人力资本具有逆周期性,然而,本文在动态一般均衡模型框架下通过将数据要素加入人力资本的

积累过程,发现数据要素水平上升的同时能够驱使人力资本顺周期变动,由此为人力资本的周期性以及数字经济的相关研究做出了一定的补充。最后,本文从推动数据共享平台的建立、发挥数据要素的关键作用,以及鼓励教育投资和人力资本积累方式的创新三个方面提出相关的政策建议,为数据要素推动人力资本积累、促进经济高质量发展提供参考。

由于近年来数字经济的迅速发展以及人力资本在经济中的重要地位,国内一些学者已经开始关注数字经济(包括数据要素)对人力资本的影响,这其中大致可从就业结构、劳动力技能结构和人力资本水平三个视角洞悉:首先,数字经济催生了新业态、新产业和新职业,在优化就业结构的同时能够改善劳动者就业环境并提升就业能力,有助于增强知识技能、促进工作效率和提升人力资本水平(戚聿东等,2020)。具体而言,数字经济推动就业结构向第三产业服务业、知识与技术密集型行业和高技能职业方向转变,人力资本存量在该过程中发挥着紧要的中介作用,赋能数字经济实现对劳动力市场和就业结构影响的飞跃式提升(叶胥等,2021)。其次,具有高边际产出特征的信息和通信技术一方面通过“替代效应”淘汰过时技能^⑦,另一方面通过“提升效应”督促劳动者掌握新型技能,让技能结构得以吐故纳新(Michaels et al., 2014)。李梦娜和周云波(2022)的研究同样发现,数字经济发展与人力资本结构高级化显著相关,为实现人工智能领域的人机互动和合作提供便利。数字经济增加了对劳动者数字化思维、数据处理分析和数字技术应用技能的需求,从而不断推动人力资本结构转向高级化。再次,数据要素激发了人力资本潜力,赋能人才要素助力经济高质量发展。当前时代正不断从知识经济转型为数字经济,而实现人力资本从“量”到“质”转变的核心要义是把握数据要素,通过数据要素培养数字思维人才,深化人力资本时代价值,推动经济高质量发展(陶小龙和肖培,2021)。数据要素有助于充分挖掘人力资本价值,在现有人力资本存量的基础上提高人力资本质量。在中国经济发展过程中,人才要素和数据要素之间应当形成互为促进、共同提升的良好态势,数字经济发展同人力资本优势相结合能够真正发挥

最大效用(俞伯阳和丛屹,2021)。

由此可见,数字经济以及数据要素对人力资本产生了不可估量的影响,使其成为最具潜力和竞争力的独特资源。然而,以上文献对数字经济的研究只停留在定性分析和实证分析的层面,对数据要素作用与价值的探索还不够深入。面对数字经济的兴起,国外一些研究则开始尝试从理论模型层面刻画数据要素的作用机制,这些研究大致可分为信息摩擦和经济增长两个主要领域。对于前一类研究,数据被看作一种信息载体,能够减少未知结果的不确定性;而对于后一类研究,数据则被视为一种生产要素或技术并加入经济运行过程,这类要素的收益递增性质为经济增长提供了关键的驱动力量。

数据本质上是信息的一种存在形式,为经济活动提供了记录和证据,并有效提高了企业分析预测的科学性和准确性。Farboodi and Veldkamp(2021)参照索洛模型中资本积累的思想构建了企业数据积累模型,从而发现数据能够降低预测误差,为企业选择生产技术和优化运营提供必要信息,提高了企业的生产效率和盈利能力。数据引导完美的预见、远见使企业更有利可图。互联网和移动设备的普遍应用将个人行为在数字和物理环境中细粒化、数据化,形成了涵盖丰富信息的“数字足迹”[®]。数字足迹为金融从业人员监控、了解和优化个体投资决策提供了新的视角,有助于优化服务流程、预测未来业绩、制定发展战略,从而提高金融机构的盈利能力和经济效益(Cong et al.,2021a)。然而从长期来看,数据的这一作用同其他要素类似也会遵循边际收益递减规律,由此对微观运行效率的提升作用逐渐减弱而无法维持长期增长(Veldkamp and Chung,2019)。其他学者如Ichihashi(2020,2021a,b)还从数据的非竞争性和数据平台或中介竞争的角度对数据要素的特性展开了研究。此外,国内学者如徐翔和赵墨非(2020)采用了产业组织的创新模型为数据资本构建了进入经济运行过程的微观基础,也是对于数据要素和数字经济建模的一个重要尝试。

数据产生于消费和生产过程,是数字经济的关键生产要素,也是经济增长的新型驱动力。数据在经济中具有多种用途,除了以上提到的提高预测准

确度之外,还可作为投入要素进入生产或创新过程。Jones and Tonetti(2020)在Romer(1990)内生增长模型的基础上开创性地构建了一个数据经济理论模型,指出具有非竞争性的数据是消费活动产生的副产品(比如驾驶汽车产生的行驶数据),数据作为可提高想法质量的一种中间投入要素直接参与最终品的生产过程,分析发现被广泛使用的数据将会带来巨大的社会收益。Cong et al.(2021b)则把消费者产生的数据看作知识积累的关键因素,允许创新型中间品生产商在研发过程中利用数据,并由此分析数据在推动经济长期增长的过程中所涉及的隐私保护等问题。在后续研究Cong et al.(2022)中,作者们还在一个完全内生增长模型的框架下进一步分析和对比了数据在生产和创新过程中的不同作用,成功地将以上两个模型联系在一起。产生于消费过程的数据是目前数据经济研究中的主流设定,然而最新的工作论文Xie and Zhang(2023)指出,数据也可来自工业企业的生产过程(如工业物联网[®]),此类数据不会对消费者隐私造成侵害,能够将经济增长推至更高的水平。现实中“数字孪生”技术的应用,正是生产者数据影响人力资本积累的典型案例[®]。本文根据人力资本积累的独特性质,将主要参考此研究的设定方式,同时对以上所述两项研究中的设定进行对比和讨论。

本文首次从理论层面探讨数字经济的人力资本效应,通过将数据引入人力资本投资过程,研究数据要素在人力资本积累过程中的作用与机制。本文的主要贡献体现在以下三个方面:第一,创造性地提出了包含数据要素的人力资本积累方程,并通过数值模拟展现数据要素对人力资本积累的重要性,同时也为其他相关研究提供了基础性框架。第二,基于数据的产生渠道设定了同时包含生产者数据和消费者数据的数据组合公式,并通过讨论这两种数据的不同替代方式以体现它们的区别以及各自对经济运行的效应。第三,通过构建动态一般均衡模型并引入各类外生冲击进行政策实验,从而揭示了数据要素在推动人力资本积累方面所能够发挥的更突出于教育投入的作用,体现了发展数字经济、激发数据潜能的实践意义。

文章的剩余部分安排如下:第二部分描述模型

设定,其中包括只含有来自生产过程的数据基准模型、只含有来自消费的数据对照模型,以及同时含有两种数据的综合模型。第三部分在参数设定的基础上,对单一数据模型的稳态效用、综合模型中各变量的稳态值,以及各类外生冲击的脉冲响应模式进行数值模拟和分析。最后,第四部分进行总结并提出相应的政策建议。

1. 模型设定

本节建立一个包含人力资本的动态一般均衡模型,并将数据作为除教育以外的重要因素引入人力资本积累方程。根据现有文献中使用的两种不同来源的数据类型,即来自生产过程的数据和来自消费的数据设定只含有来自生产过程数据的模型,并将此看作基准模型。紧接着考虑只含有来自消费数据的模型以及同时包含两种数据的综合模型,与基准模型形成对照,便于体现数据对人力资本积累的作用。在现有的数据经济理论研究中,Jones and Tonetti(2020)将数据视作生产要素,Cong et al.(2021b)将数据视作创新投入要素。而在本文的模型设定中,数据是人力资本的投入要素,直接参与人力资本积累进而间接参与商品生产过程,这是本文与已有文献的不同之处。

在构建模型之前,我们有必要对数据与数据要素的含义进行界定和区分。数据是一种在经济活动中不断产生、反映一定事实或概念的符号总称,是无序的、未经加工处理的原始素材,表现形式包括数字、文字、文件、图像、声音和视频等。只有经过处理分析并以要素身份投入到经济运行过程当中的数据,才转变为具有经济价值的数据要素。数据要素是一种参与社会生产经营活动、为使用者或所有者带来经济效益的基础性数据资源,拥有丰富的表现形式和应用场景,并且通过与其他要素协调配合,数据要素提升了整体要素效益和资源配置效率。数据要素属于经济学范畴,在后续研究中,我们将重点关注数据要素,对于“数据”一词的表述,则均为数据要素的含义。在本文的设定下,数据是一种能够扩展人力资本的要素。这是因为个体为了积累人力资本,除了投入时间和精力在学校或培训机构接受教育和掌握技能,还会购入学习资料以辅助知识水平

的提升。学习资料来自经济活动的各个方面且表现形式多样,可以由本文所研究的数据整理而成。

1.1 基准模型:只含有来自生产过程的数据

1.1.1 代表性家庭

假设经济中有离散时间无限期存在的代表性家庭,该家庭在时期 t 的效用水平由消费 C_t 和闲暇时间 O_t 共同决定:

$$U(C_t, O_t) \quad (1)$$

家庭将非闲暇时间分配到工作 L_t (包括商品生产与数据生产活动,将在下文详述)和教育 E_t 两种活动当中。将家庭拥有的总有效时间标准化为1,由此家庭的时间约束可表示为:

$$L_t + E_t + O_t = 1 \quad (2)$$

具体地,代表性家庭的效用最大化问题可写成如下形式:

$$\max_{C_t, L_t, E_t} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma} - 1}{1-\sigma} - \Omega \frac{(L_t + E_t)^{1+\eta}}{1+\eta} \right] \quad (3)$$

在上式中, β 为效用的主观贴现因子, σ 为相对风险厌恶系数(也可视为跨期消费替代弹性的倒数), η 为Frisch劳动力供给弹性的倒数, Ω 表示闲暇在效用中的重要性。将消费品 C_t 的价格标准化为1,消费以 C_t 表示。

每个家庭都以投资物质资本的形式进行储蓄,即 $S_t = I_{k,t}$,并在每期收到资本的利息收入 $R_t K_t$,其中, R_t 为资本回报率, K_t 为物质资本存量。除了来自资本的收入,家庭还通过参加工作获得劳动报酬 $W_t H_t L_t$,其中 W_t 为工资率, H_t 为家庭拥有的人力资本存量,因此 $W_t H_t$ 表示家庭实际收到的效率工资。注意,家庭只有在参加工作的时候才能获得劳动报酬,而花费在教育 E_t 上的非闲暇时间将不产生直接收益。除了教育投入,代表性家庭在人力资本积累的过程中还用到了数据要素,该种数据来源于生产过程,可以称之为“生产者数据”,并记为 $D_{p,t}$,其产生过程的建模将在下文中予以详述。关于这类数据的更多信息,可参考Xie and Zhang(2023)中的描述。生产者数据需要家庭从厂商处购买,相应的价格可记为 $p_{p,t}$ 。此外,家庭拥有厂商的股份,由此厂商利润(记为 Π_t)将成为家庭收入的一部分。因此,代表性家庭的预算约束方程为:

$$C_t + I_{K,t} + p_{P,t} D_{P,t} = R_t K_t + W_t H_t L_t + \Pi_t \quad (4)$$

此外,物质资本的运动方程为:

$$K_{t+1} = (1 - \delta_K) K_t + I_{K,t} \quad (5)$$

其中, δ_K 为物质资本折旧率, $I_{K,t}$ 表示物质资本投资。同理,人力资本的运动方程为:

$$H_{t+1} = (1 - \delta_H) H_t + I_{H,t} \quad (6)$$

其中, δ_H 为人力资本折旧率,反映了个体知识和能力的更新换代,有 $0 < \delta_H < 1$ 。 $I_{H,t}$ 表示人力资本投资,在数字经济背景下,此项投资将在传统的受教育时间投入的基础上,添加生产者数据作为另一项投入要素,数据与教育一同提高人力资本水平和劳动效率。由此,人力资本积累方程为:

$$I_{H,t} = B_t D_{P,t}^\theta E_t^{1-\theta} \quad (7)$$

在上式中, E_t 表示受教育时间,可指代个体通过接受学校教育而获得的知识和能力, $D_{P,t}$ 表示生产者数据,可指代个体通过使用数据要素而获得的数字技能和素养。参数 θ 表示数据在人力资本积累过程中的贡献程度,已有人力资本研究更多地强调教育和培训等传统途径,而本文通过纳入数据要素扩展了一种新途径。 B_t 为人力资本积累的效率项,这是一个随机过程,假设其以 AR(1) 的形式进行演进:

$$\ln B_t - \ln B = \rho_B (\ln B_{t-1} - \ln B) + \sigma_B \varepsilon_B \quad (8)$$

其中, B 为稳态下的效率项, ρ_B 和 σ_B 分别为此过程对应的参数, ε_B 服从正态分布 $N(0, \sigma_B^2)$ 。由以上设定,求解家庭效用最优化问题,可得到如下所示的一阶条件:

$$\beta(R_t + 1 - \delta_K) = \left(\frac{C_t}{C_{t-1}} \right)^\sigma \quad (9)$$

$$\Omega(L_t + E_t)^\eta C_t^\sigma = W_t H_t \quad (10)$$

$$\frac{\beta W_t L_{t+1}}{C_{t+1}^\sigma} + \frac{\beta \Omega(1 - \delta_H)(L_{t+1} + E_{t+1})^\eta}{(1 - \theta) B_{t+1} D_{P,t+1}^\theta E_{t+1}^{-\theta}} =$$

$$\frac{\Omega(L_t + E_t)^\eta}{(1 - \theta) B_t D_{P,t}^\theta E_t^{-\theta}} \quad (11)$$

$$\frac{\beta W_t L_{t+1}}{C_{t+1}^\sigma} + \frac{\beta(1 - \delta_H) p_{P,t+1}}{C_{t+1}^\sigma \theta B_{t+1} D_{P,t+1}^{0-1} E_{t+1}^{-1-0}} =$$

$$\frac{p_{P,t}}{C_t^\sigma \theta B_t D_{P,t}^{0-1} E_t^{-1-0}} \quad (12)$$

以上各式子的含义为:式(9)为家庭跨期消费的最优条件;式(10)为家庭最优劳动供给方程,也是额

外一单位休闲的机会成本;式(11)为家庭最优教育决策方程,可见家庭将教育投入的当期成本与投入工作的收益以及通过教育投入积累人力资本的未来收益进行比较;式(12)为家庭对于生产者数据的最优决策,可见家庭将数据投入的当期成本与投入工作的收益以及通过数据投入积累人力资本的未来收益进行比较。

1.1.2 生产厂商

假设经济中存在一个代表性生产厂商,在完全竞争的环境下租赁资本并雇佣劳动力生产供家庭消费的商品。这里将生产函数设定为 Cobb-Douglas 的形式:

$$Y_t = A_t K_t^\alpha (H_t L_{E,t})^{1-\alpha} \quad (13)$$

其中, $L_{E,t}$ 表示从事生产活动的劳动力,参数 α 表示资本在生产过程中的贡献程度, A_t 表示全要素生产率。这是一个随机过程,同样假设其以 AR(1) 的形式演进:

$$\ln A_t - \ln A = \rho_A (\ln A_{t-1} - \ln A) + \sigma_A \varepsilon_A \quad (14)$$

其中, A 为稳态下的全要素生产率, ρ_A 和 σ_A 分别为此过程对应的参数, ε_A 服从正态分布 $N(0, \sigma_A^2)$ 。注意到此处厂商在雇佣劳动力时,还能获得来自人力资本 H_t 的加成效应。

除了供给消费的商品外,生产过程中还会生成数据副产品。这种数据从产生获得到投入实践需要经过一部分劳动力的收集、清洗、处理和整合,使其充分要素化然后以生产者数据的形式被消费者用于人力资本积累过程。参考 Xie and Zhang(2023)文章中源于生产过程的数据产生方式,将生产者数据的生产函数设定为如下形式:

$$D_{P,t} = v_t Y_t H_t L_{D,t} \quad (15)$$

其中, $L_{D,t}$ 为从事数据收集和清洗活动的劳动力,有 $L_{E,t} + L_{D,t} = L_t$ 。 v_t 为数据生产过程的效率项,同样假设其服从如下 AR(1) 过程:

$$\ln v_t - \ln v = \rho_v (\ln v_{t-1} - \ln v) + \sigma_v \varepsilon_v \quad (16)$$

其中, v 为稳态下的效率项, ρ_v 和 σ_v 分别为此过程对应的参数, ε_v 服从正态分布 $N(0, \sigma_v^2)$ 。注意到在数据的生产过程中,产出 Y_t 并不是投入要素,而只是作为数据收集清洗活动的参照。厂商每生产一单位的商品,就会产生一定单位的生产者数据,在额外投入一

定的劳动力之后,这些生产者数据提供给消费者以提升人力资本水平。类似地,人力资本在数据生产过程中也发挥了加成效应,数据要素与人力资本的有机结合和相互促进,能够更有效地推动数字经济发展。另外,鉴于数据的产生由产出和人力资本共同决定,数据与其他投入要素如资本和劳动力并不相同。

生产厂商通过出售商品和数据获得收入,最大化利润函数如下:

$$\Pi_t = (1 + p_{p,t} V_t H_t L_{D,t}) A_t K_t^\alpha (H_t L_{E,t})^{1-\alpha} - R_t K_t - W_t H_t (L_{E,t} + L_{D,t}) \quad (17)$$

求解生产者的利润最大化问题,分别得到资本、劳动力和生产者数据的价格,以及厂商的均衡利润:

$$R_t = \alpha (1 + p_{p,t} V_t H_t L_{D,t}) A_t K_t^{\alpha-1} (H_t L_{E,t})^{1-\alpha} \quad (18)$$

$$W_t = (1 - \alpha) (1 + p_{p,t} V_t H_t L_{D,t}) A_t K_t^\alpha (H_t L_{E,t})^{-\alpha} \quad (19)$$

$$p_{p,t} = \frac{1 - \alpha}{V_t H_t [L_{E,t} - (1 - \alpha) L_{D,t}]} \quad (20)$$

$$\Pi_t = \frac{-(1 - \alpha) L_{D,t}}{L_{E,t} - (1 - \alpha) L_{D,t}} A_t K_t^\alpha (H_t L_{E,t})^{1-\alpha} \quad (21)$$

1.1.3 定义竞争性均衡

模型的竞争均衡由9个内生变量 C_t 、 K_t 、 $I_{k,t}$ 、 H_t 、 $I_{h,t}$ 、 $D_{p,t}$ 、 $L_{E,t}$ 、 $L_{D,t}$ 、 E_t 和3个生产要素价格 R_t 、 W_t 、 $p_{p,t}$ 组成,使得:

●将资本回报率 R_t 、工资率 W_t 、生产者数据价格 $p_{p,t}$ 视为给定,消费 C_t 、物质资本投资 $I_{k,t}$ 、人力资本投资 $I_{h,t}$ 、生产者数据 $D_{p,t}$ 、工作 L_t 、教育 E_t 是消费者效用最大化问题的解;

●将资本回报率 R_t 、工资率 W_t 和生产者数据价格 $p_{p,t}$ 视为给定,物质资本 K_t 、人力资本 H_t 、劳动力投入 $L_{E,t}$ 和 $L_{D,t}$ 是生产厂商利润最大化问题的解;

●物质资本积累遵循 $K_{t+1} = (1 - \delta_k) K_t + I_{k,t}$, 人力资本积累遵循 $H_{t+1} = (1 - \delta_h) H_t + I_{h,t}$;

●所有市场出清,即产品市场出清 $C_t + I_{k,t} = A_t K_t^\alpha (H_t L_{E,t})^{1-\alpha}$, 劳动力市场出清 $L_{E,t} + L_{D,t} = L_t$, 生产者数据市场出清 $D_{p,t}^D = D_{p,t}^S$ (需求等于供给)。

1.2 对照模型:只含有来自消费的数据

上一节构建了一个以生产者数据作为人力资本积累过程关键投入要素的基准模型。然而,目前大多数文献都将数据设定为来自个体的消费行为。因

此,本文另外给出了以下只含有来自消费的数据模型作为对照,以便更清晰地展示数据在人力资本积累中所起的作用。此处针对数据的设定方式来自 Jones and Tonetti(2020)和 CXZ 模型中所使用的形式, Chen et al.(2021)也从国际规制的角度对这一数据类型进行了讨论。不同于这两篇文献的是,在本文的模型中,数据被用于消费者自身的人力资本积累。对于此类数据与消费者个人隐私受到潜在侵害之间关系的理解,可以认为是由于不同消费者个体在使用数据进行学习(积累人力资本)时可能从中得知他人信息而造成的。为了叙述方便,在以下模型的设定中,混用了部分与基准模型含义类似的变量。

假设经济中有离散时间无限期存在的代表性家庭,与基准模型不同的是,该家庭在时期 t 的效用水平由消费 C_t 、闲暇时间 O_t 以及数据使用量 $D_{c,t}$ 三部分组成:

$$U(C_t, O_t, D_{c,t}) \quad (22)$$

代表性家庭的效用最大化问题可写成如下形式:

$$\max_{C_t, I_{k,t}, E_t, D_{c,t}} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma} - 1}{1-\sigma} - \Omega \frac{(L_t + E_t)^{1+\eta}}{1+\eta} - z_t \tau D_{c,t}^2 \right] \quad (23)$$

在上式中, $D_{c,t}$ 表示家庭在时期 t 使用的数据,该种数据经由消费过程产生,可以称之为“消费者数据”。消费者数据和生产者数据的主要区别在于它们的成本形式不同:前者来源于消费行为并包含消费者个人信息,在提供和使用这些数据的过程中,家庭将会面临数据隐私泄露的风险,上式中的最后一项负效用即为成本的体现,这里用 τ 表示消费者对数据隐私泄露问题的厌恶程度。另一方面,在生产过程中产生的数据虽然与某个消费者的隐私问题没有直接关联,但需要生产者承担数据收集和清理的劳动力投入成本。另外, z_t 表示消费者数据风险偏好,这是一个随机过程,假设以 AR(1) 的形式进行演进^②:

$$\ln z_t - \ln z = \rho_z (\ln z_{t-1} - \ln z) + \sigma_z \varepsilon_z \quad (24)$$

其中, z 为稳态下的风险偏好, ρ_z 和 σ_z 分别为此过程对应的参数, ε_z 服从正态分布 $N(0, \varepsilon_z^2)$ 。

消费过程产生的数据并不会被直接用于人力资本积累,参考 Jones and Tonetti(2020)的设定思路,消

费者数据由数据中介收集,随后中介将数据卖给需要使用数据来积累人力资本的家庭。数据中介的优化问题为:

$$\max_{p_{C,t}, p_{S,t}, \varphi_{C,t}} p_{C,t} D_{C,t} - p_{S,t} \varphi_{C,t} \quad (25)$$

其中, $\varphi_{C,t}$ 表示中介从消费者手中收集的数据, $D_{C,t}$ 表示中介卖给家庭的数据(对应上文中家庭使用的数据), $p_{S,t}$ 和 $p_{C,t}$ 分别表示收集数据的价格和出售数据的价格。在自由进出市场的条件下,垄断中介的利润将为零,因此,家庭的预算约束为:

$$C_t + I_{K,t} = R_t K_t + W_t H_t L_t \quad (26)$$

需要注意的是,尽管家庭和中介之间进行数据的交易,但该对照模型中的数据市场并非有效,这里无法求解得到消费者数据的价格。此外,在上一节的基准模型中,数据交易也可交由数据中介处理,即中介从厂商手中购买生产者数据,然后再卖给家庭用于人力资本积累。但为了简单起见,本文将其设定为家庭和厂商直接进行数据交易。

物质资本和人力资本的运动方程分别与式(5)和式(6)相同。不同的是,人力资本投资过程使用的是消费者数据 $D_{C,t}$,由此人力资本积累方程为:

$$I_{H,t} = B_t D_{C,t}^0 E_t^{1-\theta} \quad (27)$$

求解家庭的效用最大化问题,得到如下阶段条件:

$$\frac{\beta W_t L_{t+1}}{C_{t+1}^\sigma} + \frac{\beta \Omega (1 - \delta_H) (L_{t+1} + E_{t+1})^\eta}{(1 - \theta) B_{t+1} D_{C,t+1}^0 E_{t+1}^{1-\theta}} = \frac{\Omega (L_t + E_t)^\eta}{(1 - \theta) B_t D_{C,t}^0 E_t^{1-\theta}} \quad (28)$$

$$\frac{\beta W_t L_{t+1}}{C_{t+1}^\sigma} + \frac{2\beta \tau (1 - \delta_H) z_{t+1} D_{C,t+1}}{C_{t+1}^\sigma \theta B_{t+1} D_{C,t+1}^{0-1} E_{t+1}^{1-\theta}} = \frac{2\tau z_t D_{C,t}}{C_t^\sigma \theta B_t D_{C,t}^{0-1} E_t^{1-\theta}} \quad (29)$$

其中,式(28)为家庭最优教育决策方程,式(29)为家庭对于消费者数据的最优决策。

与基准模型的设定类似,代表性生产厂商租赁资本和雇佣劳动力以生产商品供家庭消费。但此时的生产过程没有产生数据副产品,家庭只参与生产活动而不再参与数据收集和清洗活动,即 $L_{D,t} = 0$,厂商可利用的劳动力为 $L_{E,t}$ 且 $L_{E,t} = L_t$ 。生产厂商最大化如下利润函数:

$$\Pi_t = A_t K_t^\alpha (H_t L_{E,t})^{1-\alpha} - R_t K_t - W_t H_t L_{E,t} \quad (30)$$

求解厂商的利润最大化问题,得到资本和劳动力价格,即资本回报率和工资率:

$$R_t = \alpha A_t K_t^{\alpha-1} (H_t L_{E,t})^{1-\alpha} \quad (31)$$

$$W_t = (1 - \alpha) A_t K_t^\alpha (H_t L_{E,t})^{-\alpha} \quad (32)$$

关于均衡的定义,本模型与基准模型较为相似,此处将不再赘述。

1.3 综合模型:两类数据的组合

以上两个模型都只包含一种数据,体现了单一类型数据的人力资本积累效应。接下来,本文给出一个同时包含生产者数据和消费者数据的综合模型,探究两种数据对人力资本积累的复合影响以及各自在该过程中所发挥的作用。具体地,消费者一方面需要从厂商手中购买数据,另一方面还需要面临由于消费者数据被使用所产生的隐私问题。此时,代表性家庭的效用最大化问题为:

$$\max_{C_t, I_t, E_t, D_{C,t}, D_{P,t}} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t^{1-\sigma} - 1}{1 - \sigma} - \Omega \frac{(L_t + E_t)^{1+\eta}}{1 + \eta} - z_t \tau D_{C,t}^2 \right] \quad (33)$$

预算约束方程为:

$$C_t + I_{K,t} + p_{P,t} D_{P,t} = R_t K_t + W_t H_t L_t + \Pi_t \quad (34)$$

物质资本和人力资本的运动方程分别与式(5)和式(6)相同。然而不同的是,此处人力资本投资过程中使用的数据为消费者数据和生产者数据的组合。前者可被个体用于提升对自我行为和能力的认知,后者(相当于一种数据产品)可被个体用于提升数字技能和素养。这两类数据常以替代弹性系数的形式(即 CES 函数)一同构成复合的数据集,与教育投入共同构成人力资本积累方程。由此,人力资本积累方程变为:

$$I_{H,t} = B_t D_t^0 E_t^{1-\theta} = B_t \left[\gamma D_{C,t}^{\frac{\rho-1}{\rho}} + (1 - \gamma) D_{P,t}^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho\theta}{\rho-1}} E_t^{1-\theta} \quad (35)$$

在上式中, D_t 表示复合数据集, γ 和 $1 - \gamma$ 分别为消费者数据 $D_{C,t}$ 和生产者数据 $D_{P,t}$ 在整个复合数据集的权重, ρ 为消费者数据和生产者数据之间的替代弹性。数据于消费过程和生产过程自然生成,虽然数量庞大、种类多样,但仍难以发挥作用,本文将其置于人力资本投资公式,为人力资本积累和最终产

品生产作出贡献,使其真正成为一种投入要素。

求解家庭效用最优化问题,得到如下—阶条件:

$$\frac{\Omega(L_t + E_t)^\eta}{(1-\theta)B_t L_t^{1-\theta} D_t^\theta} = \frac{\beta W_{t+1} L_{t+1}}{C_{t+1}^\sigma} + \frac{\beta(1-\delta_H)\Omega(L_{t+1} + E_{t+1})^\eta}{(1-\theta)B_{t+1} E_{t+1}^{1-\theta} D_{t+1}^\theta} \quad (36)$$

$$\frac{2\tau z_t D_{C,t}}{\theta \gamma B_t E_t^{1-\theta} D_t^{1-\theta} D_{C,t}^{\frac{\theta p - p + 1}{\rho}} D_{P,t}^{-\frac{1}{\rho}}} = \frac{\beta W_{t+1} L_{t+1}}{C_{t+1}^\sigma} + \frac{\beta(1-\delta_H)2\tau z_{t+1} D_{C,t+1}}{\theta \gamma B_{t+1} E_{t+1}^{1-\theta} D_{t+1}^{1-\theta} D_{C,t+1}^{\frac{\theta p - p + 1}{\rho}} D_{P,t+1}^{-\frac{1}{\rho}}} \quad (37)$$

$$\frac{\beta W_{t+1} L_{t+1}}{C_{t+1}^\sigma} + \frac{\beta(1-\delta_H)p_{P,t+1}}{\theta(1-\gamma)C_{t+1}^\sigma B_{t+1} E_{t+1}^{1-\theta} D_{t+1}^{1-\theta} D_{P,t+1}^{\frac{\theta p - p + 1}{\rho}} D_{P,t+1}^{-\frac{1}{\rho}}} = \frac{p_{P,t}}{\theta(1-\gamma)C_t^\sigma B_t E_t^{1-\theta} D_t^{1-\theta} D_{P,t}^{\frac{\theta p - p + 1}{\rho}} D_{P,t}^{-\frac{1}{\rho}}} \quad (38)$$

以上式子中,式(36)为家庭最优教育决策方程,式(37)表示家庭对于消费者数据的最优决策,式(38)表示家庭对于生产者数据的最优决策。代表性生产厂商的设定以及利润最大化问题的求解与基准模型中的式(18)到(21)相同。该模型的竞争性均衡定义如下。

综合模型的竞争均衡由 10 个内生变量 C_t 、 K_t 、 $I_{K,t}$ 、 H_t 、 $I_{H,t}$ 、 $D_{C,t}$ 、 $D_{P,t}$ 、 $L_{E,t}$ 、 $L_{D,t}$ 、 E_t 和 3 个生产要素价格 R_t 、 W_t 、 $p_{P,t}$ 组成,使得:

●将资本回报率 R_t 、工资率 W_t 、生产者数据价格 $p_{P,t}$ 视为给定,消费 C_t 、物质资本投资 $I_{K,t}$ 、人力资本投资 $I_{H,t}$ 、生产者数据 $D_{P,t}$ 、消费者数据 $D_{C,t}$ 、工作 L_t 、教育 E_t 是消费者效用最大化问题的解;

●将资本回报率 R_t 、工资率 W_t 和生产者数据价格 $p_{P,t}$ 视为给定,物质资本 K_t 、人力资本 H_t 、劳动力投入 $L_{E,t}$ 和 $L_{D,t}$ 是生产厂商利润最大化问题的解;

●物质资本积累遵循 $K_{t+1} = (1-\delta_K)K_t + I_{K,t}$,人力资本积累遵循 $H_{t+1} = (1-\delta_H)H_t + I_{H,t}$;

●所有市场出清,即产品市场出清 $C_t + I_{K,t} = A_t K_t^\alpha (H_t L_{E,t})^{1-\alpha}$,劳动力市场出清 $L_{E,t} + L_{D,t} = L_t$,生产者数据市场出清 $D_{P,t}^D = D_{P,t}^S$ (需求等于供给)。

1.4 参数设定

本文模型中的参数取值主要通过借鉴其他文献加以设定。首先,采用相关文献中的数值(梅冬州和龚六堂,2011;朱超等,2018):家庭效用的主观贴现

因子 β 为 0.99 (对应于 4% 的年利率),相对风险厌恶系数 σ 为 2,劳动力供给弹性的倒数 η 为 0.3,闲暇的重要性 Ω 为 1.315。其次,借鉴人力资本研究的相关文献 (DeJong and Ingram, 2001; Alessandrini et al., 2015),设定物质资本折旧率 δ_K 为 0.06,人力资本折旧率 δ_H 为 0.01,教育的弹性系数 $(1-\theta)$ 为 0.8 (即数据的贡献度 θ 为 0.2,还将在下一小节对该参数的取值进行讨论)。为了进一步校准参数 θ ,使用 WDI 数据库中的人力资本指数 (Human Capital Index),该指数较为客观准确地衡量了教育和健康对劳动生产力的贡献。根据本文的设定,数据是除教育以外的影响人力资本积累的重要因素,故而这里关注数字经济发展相对迅速的经济体,以估算数据的贡献程度。参考联合国贸发会发布的《2019 年数字经济报告》^③,选取 ICT 行业增加值排名前 10 的国家与地区并计算人力资本指数的三年平均值^④,由此得到数据的贡献程度约为 0.26。考虑到影响人力资本积累的因素不只教育和数据,这里将参数 θ 的数值设定为 0.2。最后,关于消费者对数据隐私泄露问题的厌恶程度,在 CXZ 模型中参数 τ 位于指数位置,作者们将其设定为 1.5,并且还讨论了其他取值的情况。但在本文的模型设定中,参数 τ 位于系数位置,对此,这里先给参数 τ 赋值 0.3,然后在数值模拟部分给出了不同取值的对比 (分别为 0.4 和 0.5),发现不同取值的 τ 将会给效用稳态值带来不同的影响。具体地,参数设定结果见表 1。

2. 数值模拟与政策实验

2.1 基准模型与对照模型的稳态效用对比

在确认各参数取值后,本文首先从代表性家庭的角度比较了基准模型与对照模型的稳态效用水平。具体来说,本文计算了数据贡献程度 θ 从 0 到 1 不同取值下的效用稳态值 (图 1)。从图中可以看出,随着数据对人力资本积累的贡献程度不断增大,生产者数据模型的效用水平逐渐上升,其上升趋势相对于消费者数据模型的变动较为缓慢,而消费者数据模型的效用水平不断下降,两种模型的效用差距越来越大。消费者数据隐私问题造成的负面影响注定使其处于弱势地位,而生产者数据成功避开了此类问题,故而生产者数据模型能够实现更高的效用

表 1 模型参数设定

参数含义	参数名称	校准值	来源
效用贴现因子	β	0.99	朱超等(2018)
相对风险厌恶系数	σ	2	梅冬州和龚六堂(2011)
劳动力供给弹性的倒数	η	0.3	梅冬州和龚六堂(2011)
闲暇的重要性	Ω	1.315	朱超等(2018)
资本产出弹性	α	1/3	标准常用设定
物质资本折旧率	δ_k	0.06	Alessandrini et al.(2015)
人力资本折旧率	δ_H	0.01	DeJong and Ingram(2001)
教育的贡献度	$1-\theta$	0.8	DeJong and Ingram(2001)
消费者对数据隐私泄露问题的厌恶程度	τ	0.3	待讨论

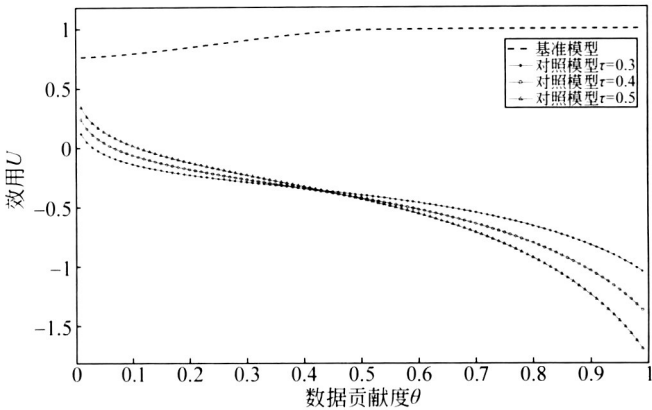


图 1 数据贡献程度 θ 在不同取值下的基准模型和对照模型稳态效用的比较

注:基准模型为只有生产者数据的模型,对照模型为只有消费者数据的模型。

水平。因此,从稳态效用水平来看,生产者数据具有更明显的效用促进价值。

另外,本文还考虑了数据风险偏好 τ 的不同取值(分别为 0.3、0.4 和 0.5)对家庭效用的影响。总体来看,随着数据风险偏好的不断增加,效用水平值逐渐递减,体现了隐私泄露风险给消费者带来的负面影响。分开来看,当数据贡献程度 θ 数值较小时,家庭选择使用更多的消费来弥补数据隐私问题引致的效用损失,因此 τ 取值 0.5 的曲线位于最上方,此时数据风险偏好数值越大,效用稳态值越高。随后,当数据贡献程度逐渐增加,即数据占比逐渐提高后,数据使用带来的负面影响也更加显著,并超过了消费带来的正面影响,此时数据风险偏好 τ 数值越大,效用稳态值越低。换句话说,隐私损失越少,福利水平越高。由此可见,数据使用对家庭效用的影响随着数据贡献程度 θ 和数据风险偏好 τ 取值呈现出非线性的关系。

2.2 综合模型中主要变量的稳态值

接下来,本文将生产者数据与消费者数据结合起来,重点关注同时包含两种数据的综合模型。具体而言,计算了替代弹性 ρ 从 0 到 10 不同取值下的主要变量的稳态值(图 2)。生产者数据和消费者数据之间的替代弹性 ρ 决定了哪种数据占主导的问题, ρ 取值小于 1 代表数据的替代缺乏弹性,即两类数据难以相互替代,复合的数据集由数量较少的一类数据所主导。 ρ 取值大于 1 则代表数据的替代富有弹性,即两类数据比较容易相互替代,复合的数据集则由数量较多的一类数据所主导。参数 ρ 取值从 0 到 10 正是两种数据从缺乏弹性向富有弹性的替代的过渡,针对此参数的变化进行比较便于观察每种数据主导情况下各变量稳态值的变动趋势。

这里比较了与产出相关的变量(产出 Y 、物质资本 K 和人力资本 H)、两类数据及由它们组合而成的复合数据集(消费者数据 D_c 、生产者数据 D_p 和数据集

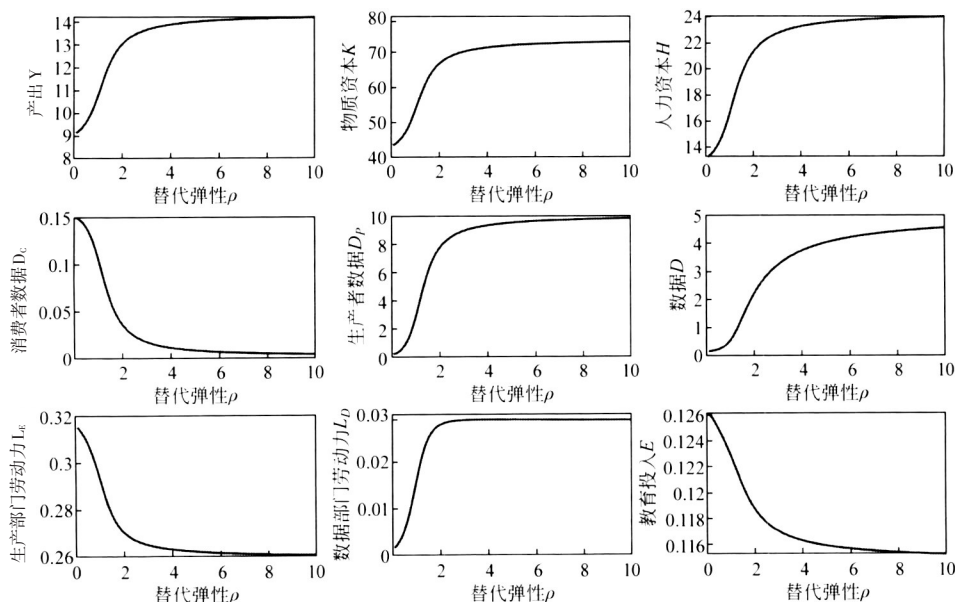


图2 参数 ρ 在不同取值下综合模型中各主要变量的稳态值

注:综合模型为同时包含消费者数据和生产者数据的模型。

D),以及三类劳动力水平(产品生产部门劳动力 L_E 、数据生产部门劳动力 L_D 和教育投入 E)。首先,对于第一组变量,可以观察到它们随着参数 ρ 的增加而增加,并在 ρ 约为4之后进入平台区域。其次,由第二组变量的变化趋势可以看出,随着参数 ρ 不断增加,消费者数据减少而生产者数据增加,而从消费者数据到生产者数据转变的过程中,数据总量不断增加。生产者数据占据主导地位,在推动人力资本积累的过程中担负主要职责,人力资本存量的变动与生产者数据的关系更加密切。再结合第一组变量中产出和物质资本水平也不断上升的变动趋势,说明无论是从生产者角度还是消费者角度看,生产者数据都具有更重要的意义。最后,由第三组变量的变化趋势可知,更多的劳动力从产品生产部门流向了数据生产部门,生产者数据提高了数据副产品的生产效率。教育投入在较小的范围内波动,替代弹性对其影响较小。结合前一小节分析,无论是从稳态效用水平角度还是人力资本积累角度出发,生产者数据相对于消费者来说都具有更重要的作用与价值。

2.3 脉冲响应分析

在讨论完综合模型处于稳态时的各类重要情形之后,本文转而对脉冲响应分析。具体来说,选取四个主要的外生冲击:全要素生产率冲击、人力资本

积累效率冲击、生产者数据生产效率冲击,以及消费者数据风险偏好冲击。这四类冲击涵盖了本模型从生产到人力资本积累,再到两类数据各自的产生过程,对理解模型具有重要意义。

2.3.1 全要素生产率冲击

首先在模型中引入正向全要素生产率冲击(即式(13)中的 A_t),得到脉冲响应函数来描述在受到冲击后至经济重新进入稳态的动力学过程,如图3所示,该图展示了一些重要变量在冲击后的百分比变动情况。每个小图包含6条曲线,分别对应替代弹性 ρ 取值为0.1、0.5、1.5和10的综合模型,以及只包含生产者数据的基准模型和只包含消费者数据的对照模型。

从图3可以看出,全要素生产率正向冲击提升了产出、消费和物质资本水平,参数 ρ 不同取值下的综合模型连同基准模型和对照模型的结果都较为一致,变动幅度差别不大。然而,人力资本的变动情况稍显复杂。当 ρ 数值较小时,消费者数据和生产者数据之间的替代弹性较低,总数据更偏向于消费者数据,而此时消费者数据的使用量处于较低水平,故而综合模型的脉冲响应函数呈现出不完全的渐进变化趋势。来自消费过程的数据受隐私问题的干扰,难以带动总数据水平提升,故而数据要素对人力资本

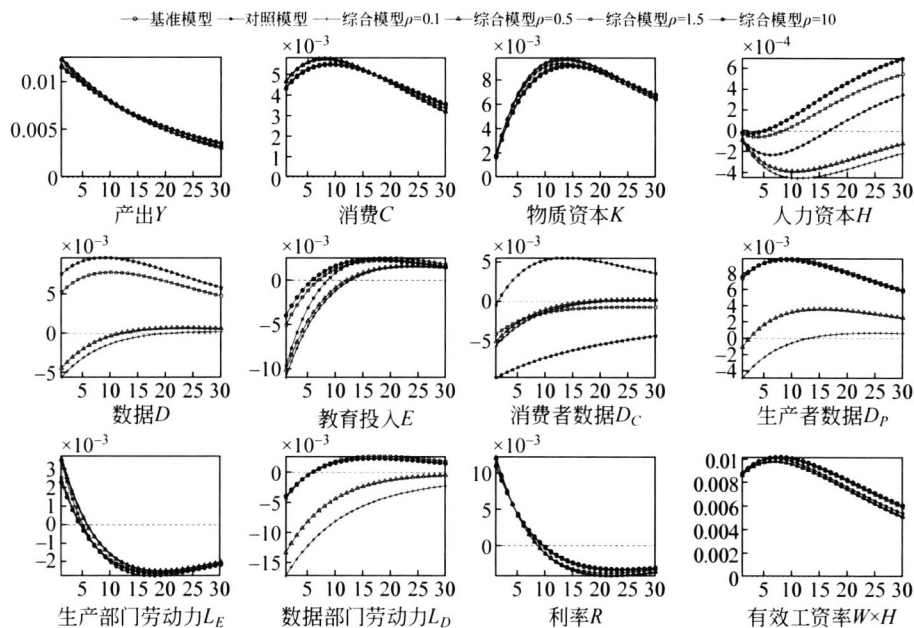


图3 全要素生产率正向冲击的脉冲响应函数

的促进作用受限,人力资本存量在初期出现下降,与其他变量的变动趋势相反。与之相对地,当 ρ 数值较大时,消费者数据和生产者数据之间富有替代弹性,总数据更偏向于生产者数据,且此时生产数据的使用量处于较高水平,故而综合模型接近于生产者数据模型,从 ρ 等于10时综合模型与基准模型几乎重合的脉冲响应函数也可看出。作为生产过程的一种副产品,生产者数据与生产率密切相关,于是在正向全要素生产率冲击下脱颖而出并占据主导地位,由此总数据水平明显提升,能够有效发挥人力资本积累促进效应。因此得到如下结论,当替代弹性较小时,人力资本存量逆周期变动,然而随着替代弹性不断变大,生产者数据逐渐占据主导地位,数据使用水平显著提升,随之而来的数据价格下降使得人力资本积累更加合算。因此,数据(尤其是生产者数据)的引入能够显著提升人力资本积累的效率,促使人力资本呈现出顺周期变动的趋势。

以往研究人力资本周期性的文献表明(DeJong and Ingram, 2001; Méndez and Sepúlveda, 2012; Alessandrini et al., 2015),正向TFP冲击在提高工作收入的同时增加了教育的机会成本,导致人们减少学习时间而增加工作时间,因此与人力资本积累相关的技能获得活动和学校入学率具有逆周期性。从图3同样可

以看出,正向TFP冲击使代表性家庭减少了教育时间而增加了工作时间,与之前文献的结果相似。常见的人力资本方程主要由教育投入或教育和健康投入组成(杨建芳等, 2006; 余长林, 2006),人力资本对产出的影响存在延迟效应,当期积累的人力资本等到下期才能决定生产,因此面临生产率正向冲击时,人们倾向于将更多的时间分配到工作生产而不是学习教育中。然而在本文的模型设定中,人力资本积累方程由教育投入和数据要素共同组成,人力资本存量在教育投入减少时仍然能够提升,正是数据要素(尤其是生产者数据)在背后驱动的结果。生产率冲击在刺激产出的同时增加了数据副产品的生成,使其充分发挥对人力资本积累的推动作用,最终人力资本存量顺周期变动,这是与之前研究不同的结果。产生于生产和消费过程的数据被家庭用来提升人力资本水平,其活力与价值被更大程度地激发了出来。

教育投入首先出现下降但随后逐渐上升,表明数据提高了人力资本投资回报,使其在正向TFP冲击下依然保有一定的效益,故而人们愿意将更多的时间投入到学习教育中,最终数据要素与教育投入共同推动人力资本积累。Lu(2021)在探究人工智能的经济增长效应时,提出了人工智能可以改善人力资本积累的观点,于是在标准的Lucas(1988)人力资

本积累模型中加入了人工智能因素。研究发现随着人工智能自我积累效率的提升或人工智能在CES生产函数中替代弹性的提高,家庭将在教育方面投入更多时间,人工智能通过多途径促进人力资本积累,这与本文的结论不谋而合。

2.3.2 人力资本积累效率冲击

在对全要素生产率冲击进行分析之后,本文转而分析人力资本积累冲击(即式(7)中的 B_t)以更直接地反映数据在此过程中的作用,相应的脉冲响应函数见图4。在受到冲击后,人力资本存量、数据要素(包括复合数据集、消费者数据和生产者数据)以及教育投入均明显增加,产品生产部门劳动力减少而数据生产部门劳动力增加。产出、消费和物质资本先下降后上升,因为劳动力从产品生产部门集中到了数据生产部门,资源从物质资本领域流向了人力资本领域,导致在冲击发生初始时产出、消费和物质资本出现下降。但随着人力资本积累的作用逐渐显现,转而会促进产出并刺激物质资本积累,于是产出、消费和物质资本水平又逐渐回升。以上结果表明,在人力资本积累效率的正向冲击下,数据要素与教育投入共同推动人力资本积累,人力资本存量呈现出先于产出水平变化的变动趋势。

2.3.3 数据生产效率冲击

针对生产者数据,本节还对其生产过程中可能产生的冲击进行分析(即式(15)中的 v_t),相应的脉冲响应函数见图5。在本节的分析中,由于在只含有消费者数据的对照模型中并不涉及此种冲击,故不将其纳入讨论范围。从图中大致可以看出,替代弹性 ρ 等于10的综合模型脉冲响应函数和生产者数据模型的几乎完全重合,再次印证了生产者数据占据主导地位时发挥的重要作用,并且参数 ρ 不同取值下的综合模型脉冲响应函数连同只含有生产者数据的基准模型脉冲响应函数均体现出明显的渐变趋势。人力资本、数据要素和教育投入明显增加,生产者数据相比消费者数据受到更大的影响。产品部门劳动力减少而数据部门劳动力增加,这意味着代表性家庭缩短了工作时间,于是产出、消费和物质资本短期内出现下降。但从长期来看,人力资本的加速积累最终推动产出和物质资本水平不断回升。此外,从综合模型的脉冲响应函数能够发现,替代弹性 ρ 数值较小时,数据生产效率的正向冲击缩减了数据生产部门的劳动力规模;但随着 ρ 数值变大,生产者数据产量逐渐攀升,数据生产部门的劳动力随之也不断增多。以上结果表明,在数据生产效率的正向冲击下,数据要素(尤其是生产

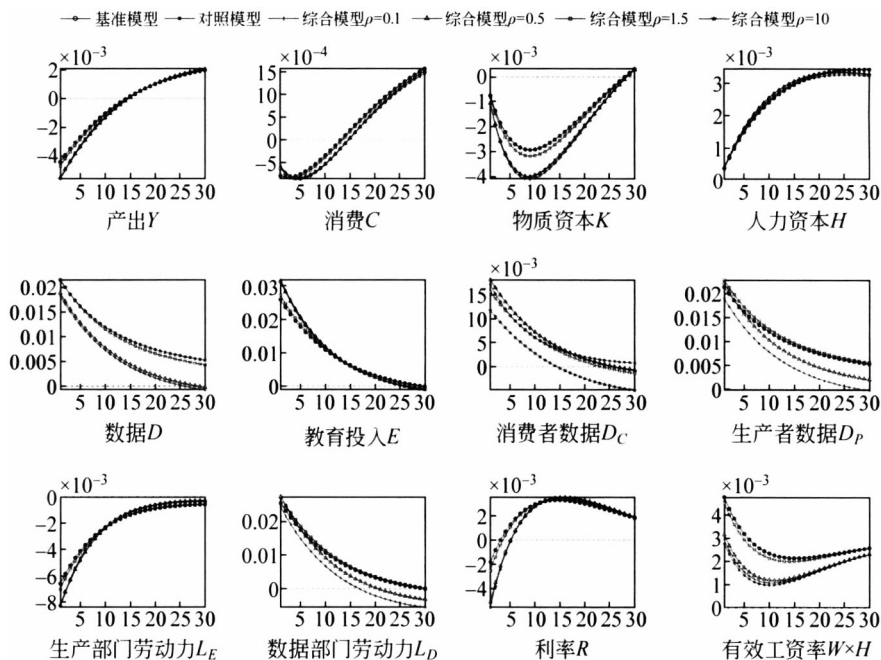


图4 人力资本积累效率冲击的脉冲响应函数

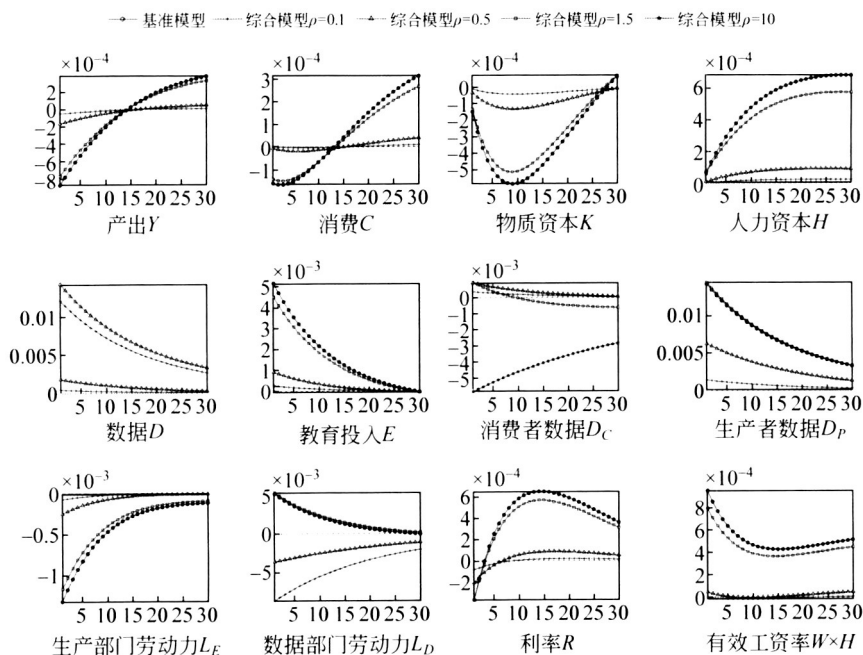


图5 数据生产效率冲击的脉冲响应函数

者数据)通过刺激人力资本加速积累,进而使得产出在经历短暂的下降之后回升到比原来更高的水平。

2.3.4 消费者数据风险偏好冲击

最后,我们还将针对消费者数据对其产生过程中涉及的风险偏好冲击进行分析(即式(23)中的 z_t),相应的脉冲响应函数见图6。由于此处的偏好冲击

属于成本增加冲击而非前三小节中所涉及的效率提升冲击^⑤,由图可知,人力资本、数据要素和教育投入明显减少,消费者数据相比生产者数据受到更大的负面影响,这些变化都与消费者数据面临的隐私泄露问题直接相关。劳动力停留在产品生产部门而未出现明显的流动情况,代表性家庭增加了工作时间

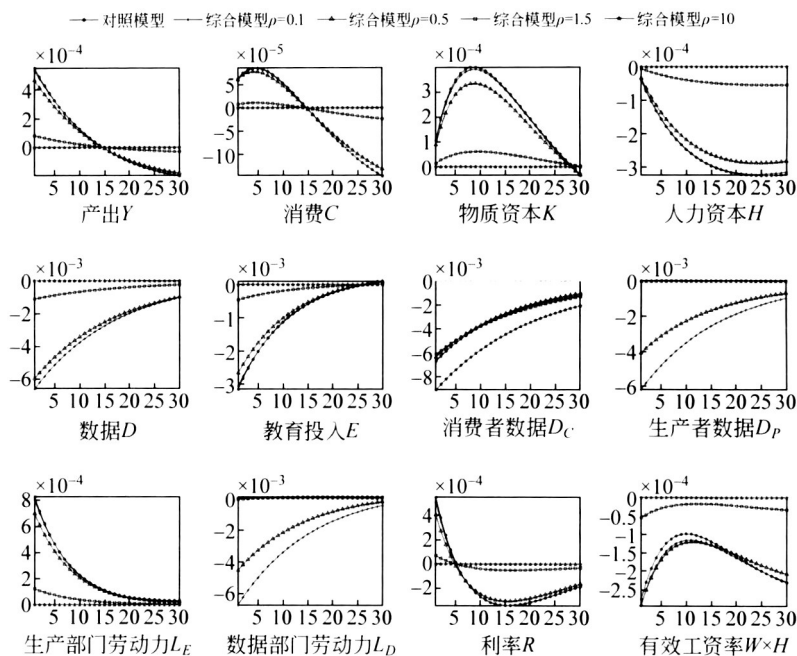


图6 消费者数据风险偏好冲击的脉冲响应函数

而缩短了学习时间,于是产出、消费和物质资本短期内明显上升。一方面,从综合模型的脉冲响应函数能够发现,随着 ρ 数值变大,消费者数据不断减少而生产者数据不断增加,有一部分劳动力从产品生产部门流向了数据生产部门,更关键的是人力资本的逆周期性逐渐减弱。另一方面,当 ρ 较大时,图中所示的各变量在冲击下均变得几乎无变化,这也说明了此种情形下消费者数据对经济的影响已几乎被生产者数据所主导。由此可见,消费者数据的隐私问题注定使其处于弱势地位,而能够占据主导地位的生产者数据更具有人力资本效应,促使人力资本存量顺周期变动。

3. 结论与启示

3.1 结论

提升人力资本质量,发挥人力资本红利,是推动中国经济实现高质量发展的关键环节。数字经济下的人力资本,其内涵已经从受教育程度和基本工作能力扩展到数字化战略思维和数据处理分析能力。与此同时,数字人才的培养和需求也达到了前所未有的高度。在这样的时代背景下,本文尝试将数字经济与人力资本相结合,把数字经济的核心要素——数据,纳入人力资本积累方程,在动态一般均衡模型框架下讨论数据要素对人力资本积累的影响,探究数据对人才培养和经济发展的作用。数值模拟结果表明,作为经济活动的副产品,数据量在正向技术冲击下显著提升,不仅能够带动教育投资的增加,还提升了人力资本积累的效率,促使人力资本存量呈现出由逆周期逐渐向顺周期转变的趋势。因此,数据的引入改变了人力资本的周期性,有利于人力资本在经济形势向好的阶段仍不断积累。

在本文的模型设定中,数据要素的来源有两个渠道,分别是基于 Jones and Tonetti(2020)和 Cong et al.(2021b)所描述的来自消费的数据(即“消费者数据”),以及基于最新研究 Xie and Zhang(2023)所提出的来自生产的数据(即“生产者数据”)。对于前者来说,数据产生于消费活动,使用这类数据将会给消费者带来潜在的隐私风险;而对于后者来说,数据产生于生产过程,其使用虽与消费者个人隐私无关,但需要企业投入资源进行收集和清洗工作。通过设定只

含有生产者数据的基准模型、只含有消费者数据的对照模型,以及同时包含两类数据的综合模型,本文对数据要素的构成方式以及两种数据各自的作用机制予以进一步分析。研究发现,在提升数据总体使用量和人力资本积累效率方面,来自生产过程的数据能够发挥出更大的作用。随着生产者数据与消费者数据之间替代弹性的不断增加,生产者数据的水平明显超过了后者并逐渐占据主导地位,在提升数据总量和刺激教育投资的同时,促使人力资本存量从逆周期不断向顺周期转变。

本文通过将数据纳入人力资本投资方程,把数字经济与人力资本有机地结合了起来,为之后的研究提供了简明的建模思路。将数据视作人力资本积累的投入要素,剖析不同来源类型的数据对人力资本周期性的影响,得到了与以往研究不同的结论,为数字经济时代的教育发展新模式提供了理论基础,具有重要的实践意义和政策含义。

3.2 政策启示

基于本文的研究结论,可提出如下三点政策建议:第一,应当推动建立多类型数据的共享平台。数据具有非竞争性,同样的数据可以供不同的场景使用而不会耗尽,总量数据产生的价值总是大于单个数据价值的加总。这种特征对于释放数据的全部价值潜力至关重要。从用户的应用程序到企业的数字平台,从消费者数据到工业生产数据,随着数据来源类型的增多,从数据中获取价值的范围也逐渐扩大。因此,应当在实现高质量发展目标的相关领域,以尊重公民个人隐私的方式,建立一个广泛的、包含生产者和消费者多方利益相关的以汇集多种类型数据、共享数据公共产品为目标的平台,从而最大限度地发挥数据的价值,推动人力资本积累,促进经济高质量发展。

第二,发挥数据要素在数字技术教育和企业技能培训中的关键作用。数字基础设施是数据要素塑造数字人才的必要条件,与之对应的充分条件则是增加数据与数字技术教育和企业技能培训的融合度。具体而言,在学校非数字专业中增设数据科学综合课程,对于数字专业则更新与现阶段需求相适配的课程内容并增加实训教学。让学生在夯实数字

技术知识和计算机基本操作的基础上,不断培养数据处理分析技能,真正掌握数字技术开发原理。另外,在企业技能培训中增加数据分析的应用场景和实践机会,促进劳动者传统技能数字化转型,鼓励技能应用与数字技术紧密结合。

第三,应当鼓励教育投资和人力资本积累方式的创新。教育投入的逆周期性意味着国民收入上升时教育投资反而出现下降,将会带动人力资本积累逆周期变动,不利于教育行业的长期发展以及人才队伍的有效供给。因此,应当在经济形势向好阶段也同样加大财政教育支出力度,鼓励家庭和个人增加教育投资,充分释放教育和知识红利,发挥人口大国的要素禀赋优势,促进人口红利转向人力资本红利。除此之外,在数字经济的背景下不断探索和创新人力资本的培养与积累方式,将数字投资纳入人力资本发展战略,鼓励数字技术与教育投资深度融合,推动人力资本结构走向多元化与高级化,为经济高质量发展增添动力。

本文曾于中国居民收入与财富分配论坛(2022)、第四届互联网与数字经济论坛(2022)进行报告。感谢谢丹夏、刘盼、张熠等专家提出的宝贵意见,以及感谢匿名评审专家和编辑部的宝贵意见,当然文责自负。

注释:

①全国劳动年龄人口平均受教育年限从1985年的6.24年增至2020年的10.7年(高于世界平均水平8.6年),同时人口预期寿命也从1978年的65.9岁增至2020年的77.9岁(高于世界平均水平72.6岁)。数据来源:中央财经大学中国人力资本与劳动经济研究中心发布的《中国人力资本报告2022》,联合国开发计划署发布的《人类发展报告2021-2022》(Human Development Report)。

②2021年人口数达14.13亿,占全球总人口的五分之一(约18%)。

③2019年10月31日党的十九届四中全会通过的《决定》中指出:“要鼓励勤劳致富,健全劳动、资本、土地、知识、技术、管理和数据等生产要素按贡献参与分配的机制”。2020年3月30日由中共中央、国务院共同发布的《中共中央国务院关于构建更加完善的要素市场化配置体制机制的意见》中提出:“加快培育数据要素市场……提升社会数据资源价值”,正式将数据要素纳入为市场化配置改革的第五大基础生产要素。

④2020年,中国数字经济规模便已达到19.14万亿元,占国内生产总值的比重从1993年的3.1%上升至18.8%,且增长

速度持续强势(平均增速16.3%)。

⑤2018年,中国共产生了大约7.6泽字节数据(占全球数据总量的23.4%),到2025年数据产生量预计将增至48.6泽字节(占比将达到27.8%),成为全世界数据储量最丰富的国家。数据来源:国际数据公司希捷(Seagate)2019年白皮书《2025年中国将拥有全球最大的数据圈》。

⑥欧盟委员会构建的数字经济与社会指数(Digital Economy and Society Index)便是将人力资本与数字经济相结合的典型应用。该指数由人力资本、网络连接、技术商用和数字公共服务四个要素组成,其中,人力资本指代新型数字技能和拥有数字技能的人口,衡量指标具体包括:掌握基本数字技能的人口数量、ICT劳动力数量、提供ICT培训的企业数量以及ICT专业的毕业生数量。详细内容可见:<https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2021>。

⑦信息和通信技术(Information and Communication Technologies, ICT):由信息技术与通信技术融合而成,信息技术强调信息编码或解码,通信技术强调信息传输以及传送方式。

⑧数字足迹:是指用户在使用互联网时留下的数据痕迹,比如使用网上银行或手机银行进行投资理财时留下的操作记录。可分为主动和被动两种,主动数字足迹如“接受网站的Cookies”(一种由网站发送到设备的小文件,用来监视活动和记录信息),被动数字足迹如广告商根据用户过去的浏览内容提供个性化广告服务。

⑨工业物联网(Industrial Internet of Things, IIoT):是一种将计算机与具有感知功能的传感器和射频识别设备以及具有监控功能的控制器和仪表等设备连接起来,通过收集、传输和分析数据并将其应用于工业生产过程的数字技术。

⑩数字孪生(Digital Twin):是一种对真实世界实体对象、过程或服务的数字表示,也是一种利用数字技术对大型物体(如喷气发动机、风力发电场、建筑物甚至整个城市)和生产制造流程的数字复制,以便收集数据、创建模拟,并预测执行情况和产品或过程的性能。鉴于大型建筑项目对实时数据的较高需求,数字孪生技术广泛应用于基础设施的建设过程。通过汇集来自平面设计、建筑工地和生产设备的实时数据,数字孪生平台充当了建筑资产虚拟生产副本的角色,一方面让建筑公司更加高效合理地生产和管理建筑,另一方面让工程师和专家们获得实时数据以便及时发现问题并实行改进措施。

⑪除了本文展示的柯布-道格拉斯形式,我们还构建了常替代弹性函数(CES)形式的人力资本积累方程,并得到了与柯布-道格拉斯形式相似的脉冲响应结果,再次印证了本文的结论。读者如对此有兴趣,可向作者索取详细过程。

⑫这里假设数据风险偏好是AR(1)过程的依据是具有时变性的风险规避系数。Brandt and Wang(2003)建立了一个基于消费的资产定价模型,研究发现总风险厌恶程度随总消费和通货膨胀率的变化而不断变化;Donadelli and Proserpi(2012)同样认为风险规避系数随时间而不断变化。进一步地,Fernández-Villaverde(2009)使用贝叶斯方法分析了风险规避冲击在产出、通货膨胀和利率动态中的作用,Benchimol(2014)在New Keynesian DSGE的模型框架中探究了风险偏好冲击对产出的影响。

⑬《2019年数字经济报告》的详细内容可见：https://unctad.org/system/files/official-document/der2019_en.pdf。

⑭ICT行业增加值排名前10的国家与地区分别是：美国、中国、日本、德国、韩国、印度、英国、法国、意大利、中国香港。

⑮在模型中，消费者数据的使用会给代表性消费者带来负效用，冲击 z_t 发生正向变化实际上提升了此类数据使用所产生的效用成本，这与前三小节中加在产品生产、人力资本投资品生产和生产者数据生产上的冲击是不同的。

参考文献：

[1]柏培文,张云.2021.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,56(5):91-108.

[2]蔡跃洲,马文君.2021.数据要素对高质量发展影响与数据流动制约[J].数量经济技术经济研究,38(3):64-83.

[3]蔡跃洲,牛新星.2021.中国数字经济增加值规模测算及结构分析[J].中国社会科学,(11):4-30.

[4]李梦娜,周云波.2022.数字经济发展的“人力资本”结构效应研究[J].经济与管理研究,43(1):23-38.

[5]梅冬州,龚六堂.2011.新兴市场经济国家的汇率制度选择[J].经济研究,46(11):73-88.

[6]戚聿东,刘翠花,丁述磊.2020.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,(11):17-35.

[7]陶小龙,肖培.2021.数字经济时代人力资本的价值特点与开发策略[J].信息与管理研究,6(6):74-84.

[8]徐翔,赵墨非.2020.数据资本与经济增长路径[J].经济研究,55(10):38-54.

[9]杨建芳,龚六堂,张庆华.2006.人力资本形成及其对经济增长的影响——一个包含教育和健康投入的内生增长模型及其检验[J].管理世界,(5):10-18,34.

[10]叶胥,杜云哈,何文军.2021.数字经济发展的就业结构效应[J].财贸研究,32(4):1-13.

[11]俞伯阳,丛屹.2021.数字经济、人力资本红利与产业结构高级化[J].财经理论与实践,42(3):124-131.

[12]余长林.2006.人力资本投资结构与经济增长——基于包含教育资本、健康资本的内生增长模型理论研究[J].财经研究,32(10):102-112.

[13]张同斌.2016.从数量型“人口红利”到质量型“人力资本红利”——兼论中国经济增长的动力转换机制[J].经济科学,(5):5-17.

[14]朱超,余颖丰,易祯.2018.人口结构与经常账户：开放DSGE模拟与经验证据[J].世界经济,41(9):26-50.

[15]Alessandrini D, Kosempel S, Stengos T. 2015. The business cycle human capital accumulation nexus and its effect on hours worked volatility[J]. Journal of Economic Dynamics and Control, 51: 356-377.

[16]Benchimol J. 2014. Risk aversion in the Eurozone[J]. Research in Economics, 68(1): 39-56.

[17]Brandt M W, Wang K Q. 2003. Time-varying risk aversion and unexpected inflation[J]. Journal of Monetary Economics, 50(7): 1457-1498.

[18]Chen Y M, Hua X Y, Maskus K E. 2021. International protection of consumer data[J]. Journal of International Economics, 132: 103517.

[19]Cong L W, Li B, Zhang Q T. 2021a. Alternative data in FinTech and business intelligence[M]//Pompella M, Matousek R. The Palgrave Handbook of FinTech and Blockchain. London: Palgrave Macmillan. Switzerland: Springer, 217-242.

[20]Cong L W, Xie D X, Zhang L T. 2021b. Knowledge accumulation, privacy, and growth in a data economy[J]. Management Science, 67(10): 6480-6492.

[21]Cong L W, Wei W S, Xie D X, et al. 2022. Endogenous growth under multiple uses of data[J]. Journal of Economic Dynamics and Control, 141: 104395.

[22]DeJong D N, Ingram B F. 2001. The cyclical behavior of skill acquisition[J]. Review of Economic Dynamics, 4(3): 536-561.

[23]Donadelli M, Proserpi L. 2012. The equity premium puzzle: Pitfalls in estimating the coefficient of relative risk aversion[J]. Journal of Applied Finance and Banking, 2(2): 177-213.

[24]Farboodi M, Veldkamp L. 2021. A model of the data economy[R]. NBER Working Paper No. 28427.

[25]Fernández-Villaverde J. 2009. The econometrics of DSGE models[R]. NBER Working Paper No. 14677.

[26]Ichihashi S. 2020. Online privacy and information disclosure by consumers[J]. American Economic Review, 110(2): 569-595.

[27]Ichihashi S. 2021a. Competing data intermediaries[J]. The RAND Journal of Economics, 52(3): 515-537.

[28]Ichihashi S. 2021b. The economics of data externalities[J]. Journal of Economic Theory, 196: 105316.

[29]Jones C I, Tonetti C. 2020. Nonrivalry and the economics of data[J]. American Economic Review, 110(9): 2819-2858.

[30]Lu C H. 2021. The impact of artificial intelligence on economic growth and welfare[J]. Journal of Macroeconomics, 69: 103342.

[31]Lucas R E Jr. 1988. On the mechanics of economic development[J]. Journal of Monetary Economics, 22(1): 3-42.

[32]Méndez F, Sepúlveda F. 2012. The cyclical nature of skill acquisition: Evidence from panel data[J]. American Economic Journal: Macroeconomics, 4(3): 128-152.

[33]Michaels G, Natraj A, van Reenen J. 2014. Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years[J]. The Review of Economics and Statistics, 96(1): 60-77.

[34]Veldkamp L, Chung C. 2019. Data and the aggregate economy[J]. Journal of Economic Literature, Forthcoming.

[35]Xie D X, Zhang L. 2023. A generalized model of growth in the data economy[R]. Working Paper 4033576.

[36]Zhu C. 2019. Big data as a governance mechanism[J]. The Review of Financial Studies, 32(5): 2021-2061.