

数据要素规模效应、产业结构转型与生产率提升

郭凯明, 王钰冰, 杭静

[摘要] 数据成为新型生产要素,是数字经济与工业经济的关键区别之一。数据要素是数字经济深化发展的核心引擎,将对经济增长方式和生产力发展路径产生重大影响,为产业深度转型升级和新质生产力发展带来新的机遇。本文基于数据要素的非竞争性和正外部性特征,从数据要素改变产业规模报酬属性这一新的理论视角出发,在宏观经济层面系统研究了数字经济时代数据要素对产业结构转型、分配结构演化和生产率提升的变革性影响。本文提出,数据要素在达到一定规模后会对不同要素密集程度的产业产生不同程度的规模效应,由此推动产业结构和分配结构转型并提升劳动生产率,甚至可以促使结构转型路径发生方向性转变,形成效率和公平更加统一的新经济形态。对数据要素局部外部性越强的产业,政府应实施更大规模的财政补贴或减税降费,在常态化监管中注重引导创新发展;对数据要素局部外部性越弱的产业,政府应实施更大力度的数据权属保护或数据隐私保护,并适度加强再分配收入调节。本文创新性地提出规模效应是推动产业结构转型的又一重要经济力量,这为理解数字经济时代结构转型趋势提供了新的理论视角,也为政府发挥“有为”作用推动产业转型升级和新质生产力发展提供了政策启示。

[关键词] 数据要素; 产业结构转型; 新质生产力; 劳动收入份额; 数字经济
[中图分类号] F124 **[文献标识码]** A **[文章编号]** 1006-480X(2024)08-0005-19

一、引言

数字经济深化发展由数字技术变革和数据要素供给推动,呈现数字产业化和产业数字化的转型趋势,是新质生产力发展的重要特征与核心要素,有望成为加快产业深度转型升级、变革生产力发展路径和转变经济增长方式的关键力量。党的二十大报告提出,加快发展数字经济,促进数字经济和实体经济深度融合,打造具有国际竞争力的数字产业集群。党的二十届三中全会提出,加快构建促进数字经济发展体制机制,完善促进数字产业化和产业数字化政策体系。值得注意的是,数字技术革命性突破推动数据成为新型生产要素,是数字经济与工业经济的关键区别之一。而与传统要素相比,数据要素在推动产业结构转型和生产率提升上会有哪些新的影响机制,又会促使数字经

[收稿日期] 2024-04-30

[基金项目] 国家社会科学基金重大项目“人口结构重大转变对中国收入分配格局的影响研究”(批准号23&ZD044)。

[作者简介] 郭凯明,中山大学岭南学院教授,博士生导师,经济学博士;王钰冰,中山大学粤港澳发展研究院博士后,经济学博士;杭静,中山大学岭南学院副教授,经济学博士。通讯作者:王钰冰,电子邮箱:wangyb87@mail.sysu.edu.cn。感谢匿名评审专家和编辑部的宝贵意见,文责自负。

济发展和结构转型呈现什么新特征和新趋势?本文研究了数字经济时代数据要素对产业结构转型、分配结构演化和生产率提升的变革性影响,为把握宏观经济结构转型趋势提供新的理论视角。党的二十届三中全会提出,要建设和运营国家数据基础设施,促进数据共享。加快建立数据产权归属认定、市场交易、权益分配、利益保护制度,提升数据安全治理监管能力,建立高效便利安全的数据跨境流动机制。本文围绕这些政策,分析了政府如何发挥“有为”作用推动产业转型升级和新质生产力发展,为进一步深化改革提供理论参考。

本文的理论出发点是:数据成为新型生产要素,是数字经济与工业经济的关键区别之一。数据要素具有非竞争性和正外部性等传统要素并不具备的重要特征,对生产方式变革和经济结构转型将产生重大影响。现有结构转型文献提出,要素结构变化会通过改变具有不同要素密集程度的产业之间的产出相对价格或不同要素的相对使用成本,推动产业之间或要素之间相互替代,对产业结构和分配结构有着重要影响。如果把数据要素等同于资本或劳动等传统要素,那么这些理论同样有助于分析其形成过程对结构转型的影响。但是数据要素还具有非竞争性和正外部性等特征,这些特征并没有被充分研究,应用主流理论分析其结构性影响有一定局限性。本文从数据要素改变规模报酬属性这一新的理论视角出发,在标准的包含要素结构的多部门动态一般均衡模型中引入数据要素。模型中数据要素具有非竞争性和正外部性特征,推动形成了规模效应,且在不同产业中数据要素规模效应大小存在差异,由此纳入了“数据要素通过产生规模效应推动结构转型”的理论机制。模型中数据要素还具有部分排他性特征,且其形成需要支付投资和研发成本,这与资本等传统要素也是有一定共通性的,由此纳入了“数据要素通过改变要素结构推动结构转型”的经典理论机制。

本文的研究与三类文献密切相关。第一类:在产业结构转型研究领域,相关经典理论提出,需求侧非位似消费偏好产生的收入效应和供给侧非平衡技术进步产生的价格效应是推动产业结构转型的重要经济力量(Kongsamut et al., 2001; Ngai and Pissarides, 2007)。之后一些文献提出,在供给侧要素结构及其相关技术变化推动下不同产业之间和不同要素之间相互替代,对产业结构转型也有重要影响(Acemoglu and Guerrieri, 2008; Alvarez-Cuadrado et al., 2017; 郭凯明等, 2020)。这些研究都把推动产业结构转型的经济力量归结为技术进步和要素结构变化,并没有考虑规模效应和生产外部性的重要作用。第二类:在数字经济研究领域,大量文献通过引入数据要素在开发、创新、使用和交易上的重要特征,提出数据要素形成和数字经济发展对企业创新、经济增长、收入分配和社会福利等都会产生重要影响(谢康等, 2020; 徐翔和赵墨非, 2020; Jones and Tonetti, 2020; Cong et al., 2021; 柏培文和张云, 2021; Farboodi and Veldkamp, 2021; 杨俊等, 2022; 徐翔等, 2023)。但是,这些研究并没有直接从宏观经济结构转型层面研究数字要素的作用。Acemoglu and Restrepo (2018)提出,人工智能等技术带来的生产自动化改变了生产结构。郭凯明(2019)、郭凯明和王钰冰(2022)、郭凯明等(2023)指出,人工智能通过影响市场内部产业结构和市场与家庭之间的生产结构,改变了要素配置和分配结构。这些文献的核心理论机制是把人工智能看作要素或要素偏向型技术,由此推动了要素之间或产业之间相互替代,并没有直接关注数据要素推动结构转型的特殊作用。第三类:在经济增长研究领域,发展相对成熟的内生增长理论从知识外部性或人力资本外部性等视角深入解释了持续的技术进步和经济增长现象(Romer, 1986, 1990; Lucas, 1988; Jones, 1995),规模效应也是其理论框架中的核心因素之一。但是,内生增长理论主要聚焦经济增长本身,经济增长过程中产业结构转型的动因并非其关注的主要问题。Acemoglu (2002)基于内生增长理论框架指出,有偏技术进步对技术结构和工资结构有着重要影响,但没有专门关注产业结构转型。

Herrendorf and Valentinyi (2022)在内生增长理论框架中引入产业结构转型过程,郭凯明和刘冲(2023)运用此框架进一步分析了平台企业反垄断政策的影响,但是均没有考虑要素密集程度特别是规模效应存在产业差异的理论机制。

与上述三类文献相比,本文提出,作为新型生产要素,数据要素不仅改变了要素结构,而且改变了不同产业的规模报酬属性。即使技术水平和要素结构保持不变,且偏好是位似的,要素和产出的规模扩张也可以推动产业结构转型和分配结构演化。因此,数据要素规模效应是推动结构转型新的经济力量,这为解释数字经济时代宏观经济结构转型趋势提供了新的理论视角,体现了一定的创新性和边际贡献。

二、数据要素特征

这部分介绍数据要素主要特征,这些特征对本文理论分析有着重要作用。作为新型生产要素,数据要素既有特殊之处,体现为数据要素具有非竞争性和正外部性特征,也有与传统要素相似的共性特征,体现为数据要素具有部分排他性特征,其形成也需要支付投资和研发成本。

(1)数据要素具有非竞争性和正外部性特征,体现为数据要素达到一定规模后,对所有企业生产率提升都有促进作用,而企业并不需要对这一正向作用支付成本(Jones and Tonetti, 2020; Cong et al., 2021; Farboodi and Veldkamp, 2021; Prüfer and Schottmüller, 2021; Acemoglu et al., 2022; 王超贤等, 2022; 李三希等, 2023)。这是因为:一方面,数字技术的发展使得数据要素一旦形成,其传输、复制和使用产生的可变成本大幅降低,甚至接近于零成本,同一数据可被不同企业或不同主体同时使用,一个企业使用数据要素并不会阻碍其他企业使用同样的数据要素。另一方面,内生增长理论指出,知识和信息具有典型的正外部性特征,而作为知识和信息的一种形态,数据要素规模增长、类型扩大,其形成的知识和信息会随之持续积累,正外部性也会快速增强,进而发挥规模效应。例如,大量公共数据、公开报告和统计资料是免费的,网络爬虫、文本分析等数据搜集与分析技术也可以促使一些数据要素的生成成本大幅降低;平台企业由于掌握了丰富的数据资源,更易于运用大数据技术识别、提炼出更多有价值信息,提升企业效率,即使是限于该平台企业内部的数据,也可以被其多个部门共享使用。即使数据由企业购买,该数据的形成过程也不是完全排他的,部分数据也可以通过其他途径由其他企业获得,或者基于该数据形成的新知识或新技术也会或多或少启发其他企业的研发创新活动,提升其生产率。^①

(2)数据要素具有部分排他性特征,体现为企业必须在数据要素市场从数据所有者(数据所有者可以同时是企业所有者)中购买数据,数据与其他要素一同为企业生产提供服务,而这些数据只在企业内部使用(Jones and Tonetti, 2020; Acemoglu et al., 2022; Gaessler and Wagner, 2022)。这是因为,尽管部分数据技术和数据要素可能是完全公开的,可以被免费共享使用,但大量数据要素在形成过程中会被限定在企业内部,不会对外公开,新形成的数据要素也可能会作为知识产权或无

^① 例如,电子商务平台企业可以获得消费交易数据,但其中部分数据会被物流企业和批发零售类企业获得。旅行平台企业的用户出行数据也不能完全排他,因为航空公司、连锁酒店等也会获得这些数据。对于汽车行驶和个人出行等活动相关数据,汽车公司可以获得一部分,出行服务平台企业同样可以获得一部分,而且这些数据可能是完全公开的。即使汽车公司使用自己获得的行驶数据本身是排他的,但基于这些数据生成的、用于自动驾驶的机器学习算法可能也会产生正外部性,因为企业使用某类相关算法或技术的活动会被其他企业观察到。

形资产进行交易,未获得产权的主体不能任意使用。因此,获得这些数据要素的企业可以从技术上(内部网络、加密算法等)或法律上(知识产权、专利等)排他性地使用这些数据要素,而其他企业无法使用。例如,虽然部分交易数据可以依靠数字技术为公众获取,但平台企业相对而言可以获得更加丰富的交易和用户数据,这些数据及其衍生的数据资源都是企业独有的,不会对外公布;机器学习、大数据、云计算等技术虽然被部分公开,但这些数字技术所使用的数据可能来自企业经营生产活动。

(3)数据要素形成需要支付投资与研发成本,体现为虽然数据资源或信息可以是生产、流通、消费和分配环节的副产品,但把数据资源开发为数据要素则需要投入一定的资源和成本。也就是说,数据要素形成过程需要在数据采集、挖掘、研发和配置等环节进行投资(Farboodi and Veldkamp, 2021; 柏培文和张云,2021;龚强等,2022;Veldkamp and Chung, 2024)。这是因为,人类的各类活动本身会自然产生大量数据资源或信息,但数据资源真正能够成为对生产效率有积极作用的数据要素,需要依赖于发展成熟的数字技术(广义而言,文字、印刷、书籍等都是较为初级的数据要素形成技术)。利用数字技术,把数据资源转化为新的数据要素并不是没有成本的(虽然数据要素易于传输和复制),如果把数据要素看作一种无形资产或知识信息,那么和物质资本类似,数据要素也需要投入一定的资源促进其持续积累。例如,即使是对公众免费的公共数据,相关统计部门在搜集整理这些数据中也需要投入成本;数据市场中企业投入成本获得数据资源再进行出售,数据要素定价本身也包括最初的投资成本。

三、模型框架

下面建立本文理论分析的模型框架。用 $t = \{0, 1, 2, \dots\}$ 代表时期。生产方面分为一个最终品生产部门和两个中间品生产部门,均由完全竞争市场中一个代表性企业进行模型化。用 $j \in \{1, 2\}$ 区分两个中间品生产部门,代表要素密集程度和规模效应大小存在差别的两个产业。

产业 j 的代表性企业使用数据 D_{jt} 、资本 K_{jt} 、劳动 L_{jt} 等要素生产,其产出为 Y_{jt} 。企业可能采用两种生产技术,其差别是是否存在数据要素规模效应。一种生产技术下不存在数据要素规模效应,产出为 $Y_{jt} = A_{jt}(D_{jt}^\eta K_{jt})^\alpha (L_{jt})^\beta$ 。其中,变量 A_{jt} 为全要素生产率。参数 $\eta > 0, 0 < \alpha_j, \beta_j < 1$ 为常数,满足 $\eta\alpha_j + \alpha_j + \beta_j = 1$ 。另一种生产技术下存在数据要素规模效应,产出为 $Y_{jt} = B_{jt}(D_{jt}^\eta K_{jt})^\alpha (L_{jt})^\beta$ 。其中,变量 $B_{jt} = A_{jt}(D_t)^\gamma$ 为全要素生产率,变量 D_t 为数据要素总量,参数 $\gamma_j > 0$ 为常数。由于数据要素在生产过程中的正外部性,企业获得所有数据要素 D_t 带来的全要素生产率提升的正向影响,但无须为此额外支付成本,参数 γ_j 衡量了数据要素正外部性所产生的规模效应大小。^①数据要素数量较低时,只有部分企业或部分生产环节才使用数据要素,其外部性可能难以显现,只有数据要素积累到一定数量,才可能发挥规模效应。因此,不失一般性地,假设只有当数据要素数量超过一定阈值 ($D_t \geq 1$) 时,数据要素才会产生外部性和发挥规模效应,而当数据要素数量低于这一阈值时,数据要素还未形成规模效应,企业只使用购买的数据要素生产。基于此,设定生产函数为:

^① 注意此时生产技术和规模报酬递增的,即使数据要素全部为一个企业所有,企业也可以享受到数据要素规模效应。所有数据要素中,可能只有其中一部分数据要素会产生外部性,这一情形也可以纳入本文模型进行分析,不会改变主要结论。

$$Y_{jt} = \begin{cases} A_{jt}(D_{jt}^\eta K_{jt}^{\alpha_j})(L_{jt})^{\beta_j}, & \text{当 } D_{jt} < 1 \text{ 时, 无数据要素规模效应} \\ B_{jt}(D_{jt}^\eta K_{jt}^{\alpha_j})(L_{jt})^{\beta_j}, & \text{当 } D_{jt} \geq 1 \text{ 时, 有数据要素规模效应} \end{cases} \quad (1)$$

由(1)式可知,在数据要素达到一定规模后,经济中数据要素总量会对两个产业全要素生产率产生影响,这种外部性被称为一般外部性。此时随着数据要素总量 D_t 增长,所有企业全要素生产率和产出都会提高。由于两个产业中参数 η 是相同的,企业购买的数据要素产出弹性和资本产出弹性在两个产业中的相对大小均由参数 α_j 决定。但在存在数据要素外部性和规模效应时,参数 α_j 越高(即资本产出弹性越高)的产业中,数据要素的产出弹性(即 $\eta\alpha_j + \gamma_j$)未必更大,还取决于数据要素外部性产生多高的规模效应(即参数 γ_j)。^① 设定数据价格、资本租金和劳动工资分别为 R_t^D, R_t^K, W_t , 产出价格为 P_{jt} 。求解企业利润最大化问题,得到最优性条件:

$$R_t^D D_{jt} = \eta\alpha_j P_{jt} Y_{jt}, R_t^K K_{jt} = \alpha_j P_{jt} Y_{jt}, W_t L_{jt} = \beta_j P_{jt} Y_{jt} \quad (2)$$

由(1)式和(2)式可知,由于数据要素的部分排他性特征,企业需要为使用数据支付成本 $R_t^D D_{jt}$ 。但同时由于数据要素具有非竞争性和正外部性,数据要素总量也会对企业生产率产生正向影响,企业无需为数据要素带来的这一正外部性支付成本。因此,(2)式中数据要素需求函数体现为企业只考虑其购买的数据要素的影响(由 $\eta\alpha_j$ 决定),而不考虑数据要素总量的影响(由 γ_j 决定)。

最终品生产部门代表性企业购买两个产业的中间品,以常替代弹性技术生产最终品 Q_t :

$$Q_t = \left[(\omega_1)^{1/\sigma} (Y_{1t})^{(\sigma-1)/\sigma} + (\omega_2)^{1/\sigma} (Y_{2t})^{(\sigma-1)/\sigma} \right]^{\sigma/(\sigma-1)} \quad (3)$$

其中, $\omega_j > 0$ 为常数; $\sigma > 0$ 为常数,衡量了两个产业的产出替代弹性。把最终品作为计价物,价格标准化为1。求解企业利润最大化问题,得到最优性条件:

$$\frac{P_{1t}}{P_{2t}} = \left(\frac{\omega_1}{\omega_2} \right)^{1/\sigma} \left(\frac{Y_{1t}}{Y_{2t}} \right)^{-1/\sigma} \quad (4)$$

模型中,两个产业在要素密集程度、规模效应大小和全要素生产率水平等方面都可以有所差别,其产出又复合形成了用于投资或消费的最终品。这两个产业可以从广义上指代经济中按照任意标准划分的两个产业、两种业态、两种模式或两种生产方式等生产结构(只要二者之间在要素密集程度、规模效应大小或全要素生产率水平上存在显著差别)。后文重点关注新兴产业和传统产业(既可以用四分位数行业分类区分,也可以用不同生产方式区分,分别对应 $j = 1$ 和 $j = 2$),其背后的经济机制同样适用于其他标准划分的生产结构。

生产、流通、消费和分配活动产生数据资源,作为这些活动的副产品,数据资源形成过程并不耗费成本。假设每一期新形成的数据资源 \tilde{D}_t 与总产出 Q_t 正相关,形式上满足: $\tilde{D}_t = \kappa Q_t$ 。其中,参数 $\kappa > 0$ 为常数。

需求方面由一个代表性家庭进行模型化。该家庭拥有生产所使用的数据 D_t 和资本 K_t ,同时提供劳动 L_t , 获得所有要素收入 $R_t^K K_t + R_t^D D_t + W_t L_t$ 。家庭将部分收入用于当期消费 C_t 。每一期消费加总后形成家庭一生效用,满足: $\sum_{t=0}^{\infty} \rho^t (C_t^{1-\theta} - 1)/(1-\theta)$ 。其中,参数 $0 < \rho < 1$ 表示时间偏好因子,参数 $\theta > 0$ 为常数。家庭把剩余收入用于储蓄,而储蓄要么被用于投资 I_t^K , 进而增加家庭拥有的资本,要么被用于开发数据 I_t^D , 进而增加家庭拥有的数据要素。于是,家庭面临如下约束方程:

$$C_t + I_t^K + I_t^D = R_t^K K_t + R_t^D D_t + W_t L_t \quad (5)$$

^① 更一般性的参数设定下的理论分析参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

$$K_{t+1} = (1 - \delta^k)K_t + I_t^k \quad (6)$$

$$D_{t+1} = (1 - \delta^D)D_t + \mu_t(I_t^D)^{1-\varepsilon}(\tilde{D}_t)^\varepsilon \quad (7)$$

其中,参数 $0 < \delta^k, \delta^D < 1$, 分别表示资本和数据的折旧率; $0 < \varepsilon < 1$ 为常数。由(7)式,数据要素形成过程既受新的数据资源 \tilde{D}_t 影响,也受用于开发数据的投资 I_t^D 影响,其影响的相对大小由参数 ε 决定。变量 $\mu_t > 0$ 衡量数据要素形成效率。可以看到,数据资源无法被直接用于生产,必须首先成为可以提高产出的生产要素,而数据资源转化为数据要素的过程需要投入一定成本。考虑到部分数据要素不会在长期持续对产出有直接影响,数据要素积累过程也引入数据要素折旧。求解家庭效用最大化问题,得到:

$$\left(\frac{C_{t+1}}{C_t}\right)^\theta = \rho [R_{t+1}^k + (1 - \delta^k)] \quad (8)$$

$$R_{t+1}^k + (1 - \delta^k) = \mu_t(\tilde{D}_t)^\varepsilon \left[\frac{1 - \varepsilon}{(I_t^D)^\varepsilon} R_{t+1}^D + \left(\frac{I_{t+1}^D}{I_t^D}\right)^\varepsilon \frac{1 - \delta^D}{\mu_{t+1}(\tilde{D}_{t+1})^\varepsilon} \right] \quad (9)$$

每一期产品市场和要素市场出清,满足:

$$Q_t = C_t + I_t^k + I_t^D \quad (10)$$

$$K_{1t} + K_{2t} = K, D_{1t} + D_{2t} = D, L_{1t} + L_{2t} = L_t \quad (11)$$

四、理论分析

1. 比较静态分析

将每一期产业比重和要素收入份额对数据要素进行比较静态分析,可以评估数据要素如何影响产业结构和分配结构。为便于表示,这里变量省略下标 t 。引入变量 $x^L = L_V / (L_1 + L_2)$ 、 $x^D = D_V / (D_1 + D_2)$ 、 $x^K = K_V / (K_1 + K_2)$ 、 $x^Y = P_1 Y_V / (P_1 Y_1 + P_2 Y_2)$, 分别衡量新兴产业的就业比重、数据要素比重、资本比重和产出比重,反映产业结构转型过程。用劳动收入与资本收入之比 $WL/R^K K$ 衡量劳动要素和资本要素的分配结构,该比值越高,劳动收入份额越大,反映分配结构演化过程。^①由(1)式、(2)式和(4)式可进一步得到:

$$\frac{\alpha_2}{\alpha_1} \frac{x^K}{1 - x^K} = \frac{\alpha_2}{\alpha_1} \frac{x^D}{1 - x^D} = \frac{\beta_2}{\beta_1} \frac{x^L}{1 - x^L} = \frac{x^Y}{1 - x^Y} \quad (12)$$

$$\frac{WL}{R^K K} = \frac{x^K \beta_1}{x^L \alpha_1} \quad (13)$$

两个产业的资本产出弹性 α_j 和劳动产出弹性 β_j 可以有所差别,数据要素正外部性 γ_j 在两个产业中也可以不相等。变量 x^K 、 x^L 、 x^D 衡量了各类要素在不同产业中的配置比例。为便于分析,用 x^K 、 x^L 的相对大小区分资本密集型和劳动密集型产业,而 x^K 、 x^L 的相对大小又取决于资本产出弹性 α_j 和劳动产出弹性 β_j 。当 $x^L > x^K$ 时,新兴产业的劳动比重高于资本比重,劳动资本比也会高于传统产

^① 引入数据要素后,劳动收入份额应为劳动收入除以劳动收入、资本收入和数据要素收入之和,但国民收入核算和资金流量表中还未纳入关于数据要素的收入统计,国民收入初次分配仍然只分为资本收入和劳动收入两类,因此,这里只关心劳动收入占劳动收入和资本收入之和的比重,隐含假设了数据要素收入按同比例最终核算进劳动收入和资本收入之中。

业,此时新兴产业是劳动密集型产业,传统产业则为资本密集型产业,反之亦然。

当数据要素还未形成规模效应(即数据要素数量 $D < 1$)时,有:

$$\frac{\alpha_2}{\alpha_1} \frac{(x^D)^{1-(\eta+1)\alpha_1 \frac{\sigma-1}{\sigma}}}{(1-x^D)^{1-(\eta+1)\alpha_2 \frac{\sigma-1}{\sigma}}} = \left(\frac{\omega_1}{\omega_2} \right)^{\frac{1}{\sigma}} \left(\frac{A_1}{A_2} \right)^{\frac{\sigma-1}{\sigma}} \frac{(x^L)^{\beta_1 \frac{\sigma-1}{\sigma}}}{(1-x^L)^{\beta_2 \frac{\sigma-1}{\sigma}}} D^{\eta(\alpha_1-\alpha_2) \frac{\sigma-1}{\sigma}} K^{(\alpha_1-\alpha_2) \frac{\sigma-1}{\sigma}} L^{(\beta_1-\beta_2) \frac{\sigma-1}{\sigma}} \quad (14)$$

当数据要素形成规模效应(即数据要素数量 $D \geq 1$)时,有:

$$\frac{\alpha_2}{\alpha_1} \frac{(x^D)^{1-(\eta+1)\alpha_1 \frac{\sigma-1}{\sigma}}}{(1-x^D)^{1-(\eta+1)\alpha_2 \frac{\sigma-1}{\sigma}}} = \left(\frac{\omega_1}{\omega_2} \right)^{\frac{1}{\sigma}} \left(\frac{A_1}{A_2} \right)^{\frac{\sigma-1}{\sigma}} \frac{(x^L)^{\beta_1 \frac{\sigma-1}{\sigma}}}{(1-x^L)^{\beta_2 \frac{\sigma-1}{\sigma}}} D^{[\eta(\alpha_1-\alpha_2)+(\gamma_1-\gamma_2)] \frac{\sigma-1}{\sigma}} K^{(\alpha_1-\alpha_2) \frac{\sigma-1}{\sigma}} L^{(\beta_1-\beta_2) \frac{\sigma-1}{\sigma}} \quad (15)$$

将(12)式—(15)式对数据要素 D 进行比较静态分析,可以得到:

$$\frac{d \log x^D / (1-x^D)}{d \log D} = \frac{d \log x^K / (1-x^K)}{d \log D} = \frac{d \log x^L / (1-x^L)}{d \log D} = \frac{(\sigma-1)[\eta(\alpha_1-\alpha_2)+(\gamma_1-\gamma_2)]}{1+(\sigma-1)(\eta+1)(\alpha_1-\alpha_2)(x^D-x^L)} \quad (16)$$

$$\frac{d \log (WL/R^K K)}{d \log D} = (x^L - x^K) \frac{(\sigma-1)[\eta(\alpha_1-\alpha_2)+(\gamma_1-\gamma_2)]}{1+(\sigma-1)(\eta+1)(\alpha_1-\alpha_2)(x^D-x^L)} \quad (17)$$

结论 1: 当 $\gamma_1 = \gamma_2 = 0$, 即数据要素还未形成规模效应时, 有:

$$\frac{d \log x^D}{d \log D} > 0 \Leftrightarrow \frac{d \log x^K}{d \log D} > 0 \Leftrightarrow \frac{d \log x^L}{d \log D} > 0 \Leftrightarrow \frac{d \log x^Y}{d \log D} > 0 \Leftrightarrow (\sigma-1)(\alpha_1-\alpha_2) > 0 \quad (18)$$

$$\frac{d \log (WL/R^K K)}{d \log D} > 0 \Leftrightarrow (x^L - x^K)(\sigma-1)(\alpha_1-\alpha_2) > 0 \quad (19)$$

此时,数据要素形成对产业结构和分配结构的影响方向取决于不同产业产出替代弹性的大小和数据要素产出弹性的差别。在新兴产业和传统产业之间产出替代弹性较高($\sigma > 1$)时,数据要素产出弹性越高的产业所占比重将会提高;在新兴产业和传统产业之间产出替代弹性较低($\sigma < 1$)时,数据要素产出弹性越高的产业所占比重将会降低。如果比重提高的产业是劳动密集型产业,那么劳动收入份额就会上升,反之亦然。

结论 1 的经济含义与 Acemoglu and Guerrieri (2008) 是一致的,核心机制是,数据形成改变了要素相对结构,通过价格效应推动了产业结构转型。具体地,数据作为新型要素,其形成过程提高了数据要素相对其他生产要素的数量,进而降低了数据的相对成本。于是,数据要素产出弹性越高的产业(如当 $\alpha_1 < \alpha_2$ 时,即为传统产业),其生产成本相对下降,进而其产出相对价格降低。在不同产业之间产出替代弹性较高时,该产业的产出会大幅替代其他产业产出,其实际数量的上升幅度会超过相对价格的下降幅度,于是名义产出的相对比重就会提高(x^Y 下降),进而促使各类要素更多配置到该产业,提高其各类要素比重(x^K, x^L, x^D 下降)。如果该产业同时是资本密集型产业($x^L < x^K$),那么相对于劳动,该产业对资本的相对需求会更大幅度提高,进而劳动收入份额减少。如果上述条件相反,那么各个理论机制的影响方向也会相反。

结论 2: 当 $\gamma_1, \gamma_2 \neq 0$, 即数据要素形成规模效应时, 有:

$$\frac{d \log x^D}{d \log D} > 0 \Leftrightarrow \frac{d \log x^K}{d \log D} > 0 \Leftrightarrow \frac{d \log x^L}{d \log D} > 0 \Leftrightarrow \frac{d \log x^Y}{d \log D} > 0 \Leftrightarrow (\sigma-1)[\eta(\alpha_1-\alpha_2)+(\gamma_1-\gamma_2)] > 0 \quad (20)$$

$$\frac{d \log (WL/R^K K)}{d \log D} > 0 \Leftrightarrow (x^L - x^K)(\sigma-1)[\eta(\alpha_1-\alpha_2)+(\gamma_1-\gamma_2)] > 0 \quad (21)$$

此时,数据要素形成对产业结构和分配结构的影响方向不仅取决于不同产业之间产出替代弹

性的大小和数据要素产出弹性的差别,还取决于数据要素的正外部性在不同产业中产生的规模效应大小。在新兴产业和传统产业之间产出替代弹性较高($\sigma > 1$)时,数据要素产出弹性越高、规模效应越强的产业所占比重将会提高;在新兴产业和传统产业之间产出替代弹性较低($\sigma < 1$)时,数据要素产出弹性越高、规模效应越强的产业所占比重将会降低。如果比重提高的产业是劳动密集型产业,那么劳动收入份额就会上升,反之亦然。

相对于结论1而言,结论2提出了推动产业结构转型的又一个重要的经济力量,即数据要素的正外部性产生的规模效应大小在不同产业中存在差异时也会产生价格效应。具体地,如果数据要素数量达到一定程度产生正外部性和规模效应,那么即使不同产业数据要素产出弹性相同(如 $\alpha_1 = \alpha_2$),数据要素规模效应较为显著的产业(如当 $\gamma_1 > \gamma_2$ 时,即为新兴产业)的产出相对价格也会随着数据要素的形成而下降。因为数据要素在该产业形成了更为显著的规模效应,会降低其单位产出的生产成本,进而降低其产出相对价格,产生价格效应。在不同产业之间产出替代弹性较高时,数据要素规模效应越显著的产业的产出就会大幅替代其他产业,其名义产出的相对比重就会提高(x^y 上升),进而提高各类要素比重(x^k, x^l, x^d 上升)。如果该产业同时是劳动密集型产业($x^l > x^k$),那么劳动收入份额会随之上升。反之,在不同产业之间产出替代弹性较低时,这些比重的变化方向就会相反。

对比结论1和结论2,一个引申结论是,数据要素产生的规模效应可以促使经济结构转型路径发生方向性转变。在数据要素还未形成规模效应时,传统产业所占比重可能在数据形成和资本深化过程中逐渐上升(如 $\alpha_1 < \alpha_2$ 时),在该产业为资本密集型产业时还会同时降低劳动收入份额。然而,一旦数据要素达到一定数量形成规模效应后,新兴产业所占比重就可能由于形成了更为显著的规模效应(如 γ_1 显著大于 γ_2 ,使 $\gamma_1 + \eta\alpha_1 > \gamma_2 + \eta\alpha_2$ 时)而由降转升,在该产业为劳动密集型产业时还会提高劳动收入份额,从而促使产业结构和分配结构的转型方向发生转变。

2. 数值模拟演示

这里通过数值模拟方式展示理论模型中的经济结构转型过程。模型中每一期代表每一年,模拟300期经济,即假设经济在第300期处于稳态。首先求解稳态,然后设定第1期资本和数据要素取值,使第1期新兴产业产出比重与劳动收入份额基本在50%左右,由此再求解转移动态路径。

考虑到要素产出弹性通常小于1,这里先设定影响数据要素的产出弹性的参数 $\eta = 0.5$ 。根据世界投入产出数据库的社会经济账户,2000—2014年中国计算机程序设计、咨询、信息服务活动业、科学研究与发展业等与数字经济直接相关的新兴行业劳动收入份额均值分别是0.647和0.729,显著高于整体经济的均值0.498。这里假设新兴产业是劳动密集型,传统产业是资本密集型,由此设定 $\alpha_1 = 0.25, \alpha_2 = 0.5, \beta_1 = 0.625, \beta_2 = 0.25$ 。新兴产业和传统产业的产出替代弹性的取值可以借鉴不同产业、行业或产品之间替代弹性的相关估计进行校准。根据Simonovska and Waugh (2014),同一细分行业不同品类产品之间的替代弹性的范围是2.79—4.46,基准估计结果为4;关于家庭服务和市场服务之间的替代弹性的大量文献表明,二者之间的替代弹性基本位于1.5—2.3的范围内(Aguiar et al., 2012; Rogerson and Wallenius, 2016);颜色等(2018)估计中国三次产业产出在消费和投资中的替代弹性基本上接近于1。考虑到新兴产业和传统产业之间的产出同质性较高,其替代弹性应会高于三次产业之间的产出替代弹性,但应低于同一细分行业不同品类产品之间的替代弹性或者家庭服务和市场服务之间的替代弹性,这里设定其产出替代弹性高于1,即 $\sigma = 1.5$ 。决定两个产业相对比重的权重参数设为 $\omega_1 = \omega_2 = 0.5$ 。

在内生增长理论中,关于技术或知识的外部性的相关估计文献非常丰富,几乎所有估计结果都在1以内,多数在0.5左右(Griliches, 1992);关于人力资本外部性的相关估计表明,人力资本外部性远低于技术或知识的外部性,且不超过人力资本的私人回报率,通常在0.1左右(Acemoglu and Angrist, 2000)。借鉴这些研究,设定决定数据要素外部性的参数 $\gamma_1 = 0.5$ 、 $\gamma_2 = 0$,新兴产业数据要素规模效应高于传统产业。Jones and Tonetti (2020)设定最终产出中数据外部性为0.33或0.4, Cong et al. (2021)设定研发中知识外部性为0.85,考虑到数据要素外部性类似于数字技术外部性但低于知识外部性,这里的取值与这些研究是相符的。

与多数研究类似,时间偏好因子 ρ 设定为0.96,跨期替代弹性的倒数 θ 设定为1。资本和数据的折旧率设定为0.1。劳动总量固定为1,由此设定两个产业的技术参数 A_1 和 A_2 均为0.7,数据要素形成效率 μ 为2,参数 κ 为0.01,数据资源对数据要素的影响程度 ε 为0.1,从而形成数据要素数量从小于1增长到大于1的转移动态路径。把上述参数取值的模型作为基准模型。

图1给出了前150期模型转移动态路径下产业结构转型、分配结构演化和生产率提升的趋势。由图1可知,在基准模型中,由于数据要素数量在第1期低于形成规模效应的阈值,数据形成和资本深化过程使新兴产业所占比重逐渐下降,配置于该产业的就业比重和数据要素比重会随之下降。由于新兴产业的劳动收入份额更高,劳动收入份额也呈现下降趋势。但数据要素积累到第10期就超过了阈值,之后产生了规模效应。由于此时数据要素规模效应在新兴产业中更加显著,新兴产业所占比重及各类要素在其中的配置比重均会由降转升,进而也会促使劳动收入份额转为上升。数据要素规模效应形成直接提升了新兴产业全要素生产率,与此同时,新兴产业比重转为提高,由此带来总体经济劳动生产率的持续提升。这些结果很好地展示了结论1和结论2所描述的数据要素的影响方向和机制。

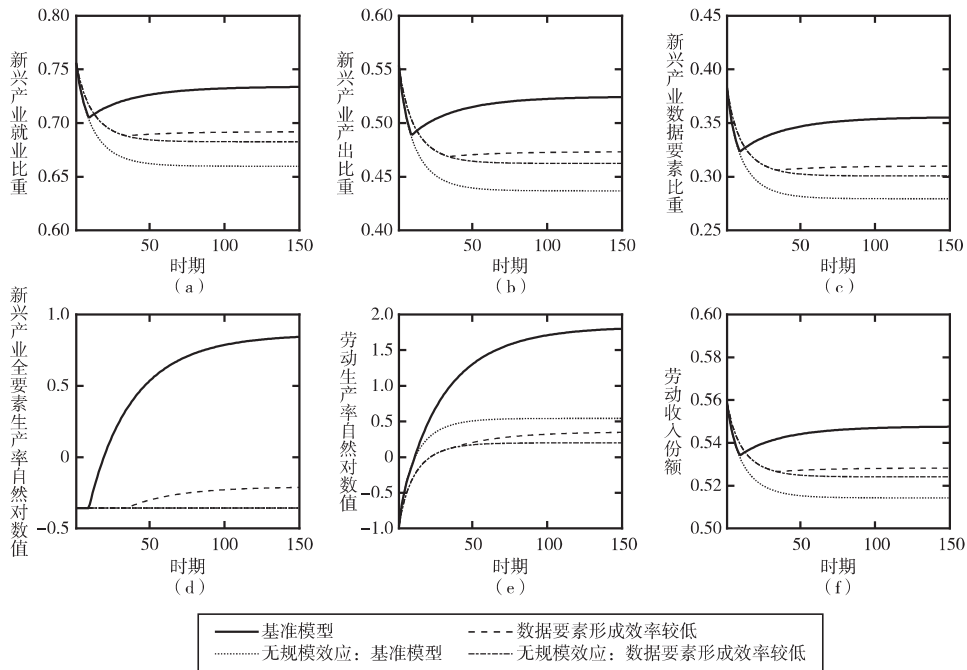


图1 数据要素推动产业结构转型、分配结构演化和生产率提升趋势

图1通过设定 γ_j 取值始终为零,模拟了不存在数据要素规模效应的情形。与之对比,数据要素规模效应改变了产业结构和分配结构的转型方向,带来了生产率提升,使100期后新兴产业所占比重及各类要素在其中的配置比重大约提高了8个百分点,劳动收入份额大约提高了3个百分点,新兴产业全要素生产率和总体经济劳动生产率大约提高了2倍。图1通过设定数据要素形成效率 μ 取值为1,模拟了数据要素形成效率较低的情形。此时,由于数据要素形成更慢,更晚达到产生规模效应的阈值,结构转型路径也会更晚发生方向性转换,在第35期后新兴产业所占比重才会由降转升,进而提高劳动收入份额和生产率。与不存在数据要素规模效应的情形相比,此时数据要素规模效应的影响程度相对有限,100期后新兴产业所占比重及各类要素在其中的配置比重的提高幅度均在1个百分点左右,劳动收入份额变化幅度更小,同时新兴产业全要素生产率和总体经济劳动生产率的提高幅度也均在0.5倍以内。本文还对模型主要参数做了敏感性分析,即在一定范围内改变相关参数取值,评估模拟结果是否会显著变化。结果表明,只要参数相对大小不发生变化,其取值变化只会影响产业结构、分配结构和生产率的大小,并不会改变转型路径,因此,这里的定量结果具有较高稳健性。^①

3. 局部外部性引入

前文基准模型设定了一般外部性,即经济中数据要素总量会对两个产业的生产率产生影响,但数据要素带来的外部性可能只会局限在特定产业中,形成局部外部性。这里拓展基准模型,考察局部外部性和一般外部性同时存在的经济环境。为了突出理论机制,这里只聚焦数据要素已达到一定规模而产生外部性的阶段。假设存在局部外部性,体现为某一产业所使用的数据要素只会提升该产业所有企业的全要素生产率,但不会对其他产业的企业产生影响。把基准模型中两个产业的代表性企业的生产函数变换为:

$$Y_j = B_j (D_j^\eta K_j)^\alpha (L_j)^\beta = A_j (D_j)^{v_j} (D_j)^{\gamma_j} (D_j^\eta K_j)^\alpha (L_j)^\beta \quad (22)$$

其中,引入参数 $v_j > 0$,为常数。与前文模型相比,此时每个产业所使用的数据 D_j 也会产生正外部性和规模效应,体现在(22)式等号右边 $(D_j)^{v_j}$ 部分。这一正外部性只限于每个产业中,并不会影响其他产业,因此,参数 v_j 衡量了局部外部性所产生的规模效应大小。注意到局部外部性在不同产业的相对大小是可以有所差别的。局部外部性仍然是外部性的一种,企业在决定数据要素投入数量时不会考虑增加数据要素也会通过局部外部性影响企业收入,于是其利润最大化的最优性条件(2)式在形式上仍然不变。其他设定沿用基准模型设定。按照前文方法重新求解模型,可知(12)式和(13)式仍然成立,但(15)式变为:

$$\frac{\alpha_2}{\alpha_1} \frac{(x^D)^{1-\frac{[(\eta+1)\alpha_1+v_1]\frac{\sigma-1}{\sigma}}}{(1-x^D)^{1-\frac{[(\eta+1)\alpha_2+v_2]\frac{\sigma-1}{\sigma}}}} = \left(\frac{\omega_1}{\omega_2}\right)^{\frac{1}{\sigma}} \left(\frac{A_1}{A_2}\right)^{\frac{\sigma-1}{\sigma}} \frac{(x^L)^{\beta_1 \frac{\sigma-1}{\sigma}}}{(1-x^L)^{\beta_2 \frac{\sigma-1}{\sigma}}} D^{\left[\eta(\alpha_1-\alpha_2)+(\gamma_1-\gamma_2)+(v_1-v_2)\right] \frac{\sigma-1}{\sigma}} K^{(\alpha_1-\alpha_2) \frac{\sigma-1}{\sigma}} L^{(\beta_1-\beta_2) \frac{\sigma-1}{\sigma}} \quad (23)$$

把(12)式、(13)式和(23)式对数据要素 D 进行比较静态分析,可以得出:

$$\begin{aligned} \frac{d \log x^D / (1-x^D)}{d \log D} &= \frac{d \log x^K / (1-x^K)}{d \log D} = \frac{d \log x^L / (1-x^L)}{d \log D} = \frac{d \log x^Y / (1-x^Y)}{d \log D} \\ &= \frac{(\sigma-1) \left[\eta(\alpha_1-\alpha_2) + (\gamma_1-\gamma_2) + (v_1-v_2) \right]}{1 + (\sigma-1) \left\{ (\alpha_1-\alpha_2)(\eta+1)(x^D-x^L) - [v_1(1-x^D) + v_2 x^D] \right\}} \end{aligned} \quad (24)$$

^① 敏感性分析的更多模拟结果参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

$$\frac{d\log(WL/R^kK)}{d\log D} = (x^l - x^k) \frac{(\sigma - 1)[\eta(\alpha_1 - \alpha_2) + (\gamma_1 - \gamma_2) + (v_1 - v_2)]}{1 + (\sigma - 1)\{(\alpha_1 - \alpha_2)(\eta + 1)(x^d - x^l) - [v_1(1 - x^d) + v_2x^d]\}} \quad (25)$$

当 $(\alpha_1 - \alpha_2)(\eta + 1)(x^d - x^l) > v_1(1 - x^d) + v_2x^d$ 时, 可知:

$$\begin{aligned} \frac{d\log x^d}{d\log D} > 0 &\Leftrightarrow \frac{d\log x^k}{d\log D} > 0 \Leftrightarrow \frac{d\log x^l}{d\log D} > 0 \Leftrightarrow \frac{d\log x^y}{d\log D} > 0 \\ &\Leftrightarrow (\sigma - 1)[\eta(\alpha_1 - \alpha_2) + (\gamma_1 - \gamma_2) + (v_1 - v_2)] > 0 \end{aligned} \quad (26)$$

$$\frac{d\log(WL/R^kK)}{d\log D} > 0 \Leftrightarrow (x^l - x^k)(\sigma - 1)[\eta(\alpha_1 - \alpha_2) + (\gamma_1 - \gamma_2) + (v_1 - v_2)] > 0 \quad (27)$$

虽然(26)式、(27)式与(20)式、(21)式相比,形式上有所不同,但结论2及其经济机制没有发生本质变化。结论2强调了数据要素外部性所产生的规模效应在不同产业之间存在差异,由此改变产业结构与分配结构的转型路径。结论2是在一般外部性的情形下展开讨论,这里进一步把外部性分为一般外部性(由 γ_1 和 γ_2 反映)和局部外部性(由 v_1 和 v_2 反映),但无论外部性来自何处,最终决定结构转型方向的还是数据要素规模效应的相对大小。因此,形式上只不过从关注 $(\gamma_1 - \gamma_2)$ 的正负号变为关注 $(\gamma_1 - \gamma_2) + (v_1 - v_2)$ 的正负号,规模效应的影响仍然成立。

4. 进一步讨论

与现有文献对比,本文理论模型纳入了数据要素规模效应,发现即使技术没有进步、要素结构不变且偏好是位似的,要素和产出的规模扩张所产生的规模效应也可以推动产业结构转型,这为解释产业结构转型提供了一个新的理论视角和经济机制。也就是说,在技术结构和要素结构不变时,依靠市场规模扩大也会推动产业结构转型,进而改变分配结构,这有别于产业结构转型动因分析的现有文献的结论。上述理论分析也为产业转型升级和新质生产力发展提供了一个重要理论视角。一方面,不同于三次产业之间产出替代弹性通常较低的情况,新兴产业和传统产业的产出同质性可能较高,其产出替代弹性也较高(新兴产业和传统产业所生产的产出可以覆盖同一细分行业或同一种类的产品和服务,如新能源汽车制造业和传统汽车制造业生产的汽车、直播电商和传统零售提供的销售服务);另一方面,相比传统产业,新兴产业对新技术和新模式的应用程度可能更高,其数据要素规模效应也会更强。因此,数据要素形成规模效应不仅直接提升了新兴产业和传统产业的全要素生产率,而且会加快新兴产业对传统产业的替代,推动产业结构转型升级,进一步增强新质生产力。

尽管前文是以要素密集程度区分新兴产业和传统产业进行分析,但数据要素推动结构转型的理论机制是普遍成立的,有助于解释或预测中国现实经济中各种形式的结构转型过程。根据前文分析,如果三次产业之间产出替代弹性较低,且数据要素在服务业产生的外部性较强(考虑到大量服务行业与数字产业相关,互联网经济也多集中在服务业等行业),那么数据要素形成会倾向于降低服务业比重和提高制造业比重。这就可能改变之前中国制造业比重下降的趋势,为解释近年来制造业比重稳中有升的变化提供一些理论视角,也可以为下一步稳定制造业比重提供一定的前瞻性指引。常见的产业类别或生产方式类别不但可以区分为资本密集型和劳动密集型,也可以区分为技能密集型和技能非密集型、技术密集型和要素密集型等。根据前文结论,在这些不同产业或生产方式之间替代弹性较高(如果不同产业或不同生产方式的产出同质性较高)时,只要数据要素在两类产业或两类生产方式中产生的外部性强弱有一定差别,数据要素形成便既可能加快之前生产结构和要素结构转型的趋势(如果之前比重扩大的产业或生产方式中数据要素规模效应较强),也

可能延缓甚至逆转之前生产结构和要素结构转型的趋势(如果之前比重扩大的产业或生产方式中数据要素规模效应较弱)。这些动态变化特征都可以对应到中国现实经济中进行解释或检验。考虑到产业内部细分行业之间产出替代弹性通常较高,上述机制在解释产业内部结构转型时也是适用的。此外,数字经济和传统经济的结构变化(如其所体现的数字产业化和产业数字化发展、数字经济和实体经济融合发展等趋势)也会受数据要素所产生的规模效应的影响,影响方向取决于数字经济和传统经济的替代弹性和规模效应的差别。这对于把握和预测数字经济发展趋势具有重要参考价值,也可以在未来数字经济发展实践中得到进一步检验。

五、政策评估

这部分在前文模型中分别引入数字基础设施建设、结构性税费政策、数据权属保护与数据隐私保护等三类政策,评估这些相关政策的影响效应。这三类政策分别对应政府的支出政策、税费政策和产业政策,是中国当前数字经济发展和治理领域的重要政策实践。

1. 数字基础设施的影响分析

在开发数据资源形成数据要素过程中,信息网络基础设施、大数据中心、东数西算工程、绿色数字中心、智能计算中心等数字基础设施起着积极作用。政府建设数字基础设施,把数据资源转化为数据要素,加快了数据要素积累过程。为了评估这一机制影响,这里建立引入数字基础设施的理论模型作拓展分析。假设数据要素形成过程既取决于私人部门数据开发投入 I_t^D 和数据资源 \tilde{D}_t ,还取决于数字基础设施规模 G_t ,为此修改(7)式为:

$$D_{t+1} = (1 - \delta^D) D_t + \mu_t (I_t^D)^{1-\varepsilon-\chi} (\tilde{D}_t)^\varepsilon (\zeta G_t)^\chi \quad (28)$$

其中,参数 $\zeta > 0$ 为常数,衡量数字基础设施的数据生成效率;参数 $0 \leq \chi < 1$ 为常数,衡量数字基础设施对数据要素形成的影响。政府通过向家庭部门收取总量税收为数字基础设施建设筹资,即家庭收入在缴纳总量税收 G_t 后再用于消费和储蓄,(5)式变为:

$$C_t + I_t^K + I_t^D = R_t^K K_t + R_t^D D_t + W_t L_t - G_t \quad (29)$$

重新求解家庭效用最大化问题,最优性条件(8)式在形式上保持不变,但(9)式变为:

$$R_{t+1}^K + (1 - \delta^K) = \mu_t (\tilde{D}_t)^\varepsilon (\zeta G_t)^\chi \left[\frac{1 - \chi - \varepsilon}{(I_t^D)^{\chi+\varepsilon}} R_{t+1}^D + \left(\frac{I_{t+1}^D}{I_t^D} \right)^{\chi+\varepsilon} \frac{1 - \delta^D}{\mu_{t+1} (\tilde{D}_{t+1})^\varepsilon (\zeta G_{t+1})^\chi} \right] \quad (30)$$

产品市场出清条件为最终品用于消费、投资、开发数据或数字基础设施建设:

$$Q_t = C_t + I_t^K + I_t^D + G_t \quad (31)$$

数字基础设施建设规模由政府决定,设定为总产出的 $0 < g < 1$ 比例。因此,参数 g 衡量数字基础设施投资率,是外生变量,满足:

$$G_t = gQ_t \quad (32)$$

由(28)式得到:

结论3:如果政府加强数字基础设施建设显著提升了数据要素生成效率,那么数据要素产生规模效应、转换结构转型路径(如存在路径转换)和提升生产率便会更早发生。

结论3意味着数字基础设施建设有助于加快数据要素形成,但在政策实践中不同类型数字基础设施对数据要素形成和外部性发挥可能会产生差异化影响。例如,信息网络基础设施建设促进物联网发展,其影响可能更多集中于工业制造、农业生产、公共服务和应急管理等领域;建设大数据

中心、东数西算工程、智能计算中心等基础设施可能更有利于数据要素外部性较强的产业,更快推动与算力、数据和智能服务直接相关的行业发展。数字基础设施也可以包括政府提供免费的公开数据。考虑数据要素还可以分为排他性数据要素(如企业不公开的数据)和公共数据要素(如公开数据)两类,可能二者之间并不如前文模型设定的那样是重合的,而是有所区别。分析这一环境,可以沿用类似这里引入数字基础设施的模型设定。具体地,用 \hat{D}_i 表示公开数据,公开数据也有利于积累数据要素,即(28)式变为:

$$D_{i+1} = (1 - \delta^D) D_i + \mu_i (I_i^D)^{1-\varepsilon} (\tilde{D}_i)^\varepsilon (\zeta \hat{D}_i)^x \quad (33)$$

由上式可知,企业在使用数据要素进行生产时,由于非竞争性和正外部性而可免费使用的数据要素中也包括了公开数据。当然,公开数据形成也需要耗费一定的社会资源,体现为家庭预算约束满足: $C_i + I_i^K + I_i^D = R_i^K K_i + R_i^D D_i + W_i L_i - \hat{D}_i$,社会资源约束满足: $Q_i = C_i + I_i^K + I_i^D + \hat{D}_i$ 。此时重新求解模型,(30)式变为:

$$R_{i+1}^K + (1 - \delta^K) = \mu_i (\tilde{D}_i)^\varepsilon (\zeta \hat{D}_i)^x \left[\frac{1 - \chi - \varepsilon}{(I_i^D)^{\chi+\varepsilon}} R_{i+1}^D + \left(\frac{I_{i+1}^D}{I_i^D} \right)^{\chi+\varepsilon} \frac{1 - \delta^D}{\mu_{i+1} (\tilde{D}_{i+1})^\varepsilon (\zeta \hat{D}_{i+1})^x} \right] \quad (34)$$

此时易知,公开数据也有助于提升数据要素生成效率,于是数据要素更易于产生规模效应、转换结构转型路径并提升生产率。因此,可以把公开数据视为数字基础设施的一类,这里关于数字基础设施影响的分析也适用于理解公开数据的影响。

2. 结构性税费政策的影响分析

当局部外部性大小在不同产业之间存在差异时,完全竞争市场所决定的产业结构可能不再是社会最优,这就为政府税费政策通过调节要素配置进而改变结构转型提供了有为空间。为了展示这一结果,首先定义每一期的社会最优产业结构为一个可以自由调配要素配置、产业结构和需求结构的社会计划者最大化当期产出下的产业结构,即给定资本、劳动和数据等要素数量及其相关技术变量,在(11)式和(22)式约束下最大化总产出(3)式。求解这一最优化问题,(12)式和(23)式变为:

$$\frac{\alpha_2}{\alpha_1} \frac{x^K}{1 - x^K} = \frac{\eta \alpha_2 + v_2}{\eta \alpha_1 + v_1} \frac{x^D}{1 - x^D} = \frac{\beta_2}{\beta_1} \frac{x^L}{1 - x^L} = \frac{x^Y}{1 - x^Y} \quad (35)$$

$$\frac{\eta \alpha_2 + v_2}{\eta \alpha_1 + v_1} \frac{(x^D)^{1-(\eta \alpha_2 + v_2) \frac{\sigma-1}{\sigma}}}{(1 - x^D)^{1-(\eta \alpha_2 + v_2) \frac{\sigma-1}{\sigma}}} = \left(\frac{\omega_1}{\omega_2} \right)^{\frac{1}{\sigma}} \left(\frac{A_1}{A_2} \right)^{\frac{\sigma-1}{\sigma}} \times \quad (36)$$

$$\frac{(x^K)^{\alpha_1 \frac{\sigma-1}{\sigma}} (x^L)^{\beta_1 \frac{\sigma-1}{\sigma}}}{(1 - x^K)^{\alpha_2 \frac{\sigma-1}{\sigma}} (1 - x^L)^{\beta_2 \frac{\sigma-1}{\sigma}}} D^{\left[\eta(\alpha_1 - \alpha_2) + (\gamma_1 - \gamma_2) + (v_1 - v_2) \right] \frac{\sigma-1}{\sigma}} K^{(\alpha_1 - \alpha_2) \frac{\sigma-1}{\sigma}} L^{(\beta_1 - \beta_2) \frac{\sigma-1}{\sigma}}$$

把(35)式、(36)式分别与(12)式、(23)式对比,可知如果存在局部外部性且 $v_2/v_1 \neq \alpha_2/\alpha_1$,那么完全竞争市场的产业结构与社会最优的产业结构并不相同。

此时,在完全竞争市场中引入具有结构性特征的税费政策,考察通过税费政策调节,完全竞争市场下产业结构是否可以与社会最优相同。具体地,假设政府对数据、资本和劳动等要素收入征税,企业支付的要素价格为 $(1 + \tau_j^D) R_i^D, (1 + \tau_j^K) R_i^K, (1 + \tau_j^L) W_i$ 。其中,参数 $\tau_j^D, \tau_j^K, \tau_j^L$ 分别衡量了数据、资本、劳动的收入税率。如果 $\tau_j^D, \tau_j^K, \tau_j^L < 0$,政府税费政策实际上变为政府财政补贴,因此,这里的分析可以直接应用于分析财政补贴的影响。政府也会对不同产业的企业收入征税,使企业获得税后收入 $(1 - \tau_j^Y) P_j Y_j$,其中,参数 τ_j^Y 衡量企业收入税。政府把所有税费收入 $T_i = \sum_j (\tau_j^D R_i^D D_{ji} + \tau_j^K R_i^K K_{ji} + \tau_j^L W_i L_{ji} + \tau_j^Y P_j Y_j)$ 转移

给家庭部门,因此,家庭预算约束变为:

$$C_t + I_t^K + I_t^D = R_t^K K_t + R_t^D D_t + W_t L_t + T_t \quad (37)$$

重新求解模型,(12)式和(23)式变为:

$$\frac{\alpha_2}{\alpha_1} \frac{1 + \tau_1^K}{1 + \tau_2^K} \frac{x^K}{1 - x^K} = \frac{\alpha_2}{\alpha_1} \frac{1 + \tau_1^D}{1 + \tau_2^D} \frac{x^D}{1 - x^D} = \frac{\beta_2}{\beta_1} \frac{1 + \tau_1^L}{1 + \tau_2^L} \frac{x^L}{1 - x^L} = \frac{1 - \tau_1^Y}{1 - \tau_2^Y} \frac{x^Y}{1 - x^Y} \quad (38)$$

$$\frac{\alpha_2}{\alpha_1} \frac{1 + \tau_1^D}{1 + \tau_2^D} \frac{1 - \tau_2^Y}{1 - \tau_1^Y} \frac{(x^D)^{1 - (\eta\alpha_1 + v_1)\frac{\sigma-1}{\sigma}}}{(1 - x^D)^{1 - (\eta\alpha_2 + v_2)\frac{\sigma-1}{\sigma}}} = \left(\frac{\omega_1}{\omega_2} \right)^{\frac{1}{\sigma}} \left(\frac{A_1}{A_2} \right)^{\frac{\sigma-1}{\sigma}} \frac{(x^K)^{\alpha_1 \frac{\sigma-1}{\sigma}}}{(1 - x^K)^{\alpha_2 \frac{\sigma-1}{\sigma}}} \frac{(x^L)^{\beta_1 \frac{\sigma-1}{\sigma}}}{(1 - x^L)^{\beta_2 \frac{\sigma-1}{\sigma}}} \times \quad (39)$$

$$D^{\left[\eta(\alpha_1 - \alpha_2) + (\gamma_1 - \gamma_2) + (v_1 - v_2) \right] \frac{\sigma-1}{\sigma}} K^{(\alpha_1 - \alpha_2) \frac{\sigma-1}{\sigma}} L^{(\beta_1 - \beta_2) \frac{\sigma-1}{\sigma}}$$

把(35)式、(36)式分别与(38)式、(39)式对比,可以得到:

结论4:在存在数据要素局部外部性的环境中,对于数据要素局部外部性越强的产业的企业收入和数据使用,政府最优税费应当越低,或者说,政府应实施更大幅度的减税降费或财政补贴,此时政府最优税费满足:

$$\frac{1 + \tau_1^D}{1 + \tau_2^D} \frac{1 - \tau_2^Y}{1 - \tau_1^Y} = \frac{\alpha_1}{\alpha_2} \frac{\eta\alpha_2 + v_2}{\eta\alpha_1 + v_1}, \quad \frac{1 + \tau_1^K}{1 + \tau_2^K} \frac{1 - \tau_2^Y}{1 - \tau_1^Y} = \frac{1 + \tau_1^L}{1 + \tau_2^L} \frac{1 - \tau_2^Y}{1 - \tau_1^Y} = 1 \quad (40)$$

结论4的经济含义如下:当存在数据要素局部外部性,使不同产业数据要素规模效应存在差别时,随着数据要素形成,数据要素局部外部性越强的产业全要素生产率提高更快,生产成本会更快下降,从而该产业所占比重增大。但是在完全竞争市场中,企业决策时并没有考虑局部外部性对其生产成本的影响,这就使该产业的企业产出和要素配置低于社会最优。此时,政府降低该产业的企业收入和数据使用的相关税费,或者对其采取财政补贴,就可以促使要素更多配置到该产业,从而发挥局部外部性的规模效应,达到社会最优。

3. 数据权属保护与数据隐私保护的影响分析

在数据权属和数据隐私保护不足的环境中,如果政府加强数据权属保护与数据隐私保护,也可能使产业结构达到社会最优水平。

(1)考虑数据权属保护不足的环境。具体地,假设由于数据权属保护不足,企业使用数据 D_j 时只需为 $\xi_j D_j$ 的数据进行支付,其支付成本为 $R_j^D(\xi_j D_j)$ 。其中,引入参数 $\xi_j < 1$,衡量数据权属保护程度, ξ_j 越大,越接近于1,数据权属保护程度越高。这可以理解为政府加强数据权属保护后,企业必须支付购买成本的数据要素所占比例更高,由参数 ξ_j 的上升刻画。重新求解模型,可知社会最优产业结构仍由(35)式和(36)式决定,但完全竞争市场中产业结构满足:

$$\frac{\alpha_2}{\alpha_1} \frac{x^K}{1 - x^K} = \frac{\xi_1}{\xi_2} \frac{\alpha_2}{\alpha_1} \frac{x^D}{1 - x^D} = \frac{\beta_2}{\beta_1} \frac{x^L}{1 - x^L} = \frac{x^Y}{1 - x^Y} \quad (41)$$

$$\frac{\xi_1}{\xi_2} \frac{\alpha_2}{\alpha_1} \frac{(x^D)^{1 - (\eta\alpha_1 + v_1)\frac{\sigma-1}{\sigma}}}{(1 - x^D)^{1 - (\eta\alpha_2 + v_2)\frac{\sigma-1}{\sigma}}} = \left(\frac{\omega_1}{\omega_2} \right)^{\frac{1}{\sigma}} \left(\frac{A_1}{A_2} \right)^{\frac{\sigma-1}{\sigma}} \frac{(x^K)^{\alpha_1 \frac{\sigma-1}{\sigma}}}{(1 - x^K)^{\alpha_2 \frac{\sigma-1}{\sigma}}} \frac{(x^L)^{\beta_1 \frac{\sigma-1}{\sigma}}}{(1 - x^L)^{\beta_2 \frac{\sigma-1}{\sigma}}} \times \quad (42)$$

$$D^{\left[\eta(\alpha_1 - \alpha_2) + (\gamma_1 - \gamma_2) + (v_1 - v_2) \right] \frac{\sigma-1}{\sigma}} K^{(\alpha_1 - \alpha_2) \frac{\sigma-1}{\sigma}} L^{(\beta_1 - \beta_2) \frac{\sigma-1}{\sigma}}$$

把(35)式、(36)式分别与(41)式、(42)式对比,可以得到:

结论5:在存在数据要素局部外部性和数据权属保护不足的环境中,政府加强数据权属保护使下式成立,可以达到社会最优的产业结构:

$$\frac{\xi_1}{\xi_2} = \frac{\alpha_1}{\alpha_2} \frac{\eta\alpha_2 + v_2}{\eta\alpha_1 + v_1} \quad (43)$$

(2)考虑数据隐私保护不足的环境。具体地,假设由于存在数据隐私,企业购买数据的支付须为 $\varphi_j R_j^D D_j$ 。其中,引入参数 $\varphi_j > 1$,衡量数据隐私保护程度, φ_j 越大,数据隐私保护程度越高。这可以理解为,如果政府保护了数据隐私,那么企业必须多支付 $(\varphi_j - 1) R_j^D D_j$,即承担更多数据使用成本,使隐私受侵犯的数据所有者得到货币补偿。此时,参数 $\varphi_j = 1$ 可以刻画数据隐私保护不足的情形,而 φ_j 取值上升反映了政府加强数据隐私保护。重新求解模型,可知社会最优产业结构仍由(35)式和(36)式决定,但政府加强数据隐私保护时,产业结构满足:

$$\frac{\alpha_2}{\alpha_1} \frac{x^k}{1-x^k} = \frac{\varphi_1 \alpha_2}{\varphi_2 \alpha_1} \frac{x^D}{1-x^D} = \frac{\beta_2}{\beta_1} \frac{x^L}{1-x^L} = \frac{x^Y}{1-x^Y} \quad (44)$$

$$\frac{\varphi_1 \alpha_2}{\varphi_2 \alpha_1} \frac{(x^D)^{1-(\eta\alpha_1+v_1)\frac{\sigma-1}{\sigma}}}{(1-x^D)^{1-(\eta\alpha_2+v_2)\frac{\sigma-1}{\sigma}}} = \left(\frac{\omega_1}{\omega_2}\right)^{\frac{1}{\sigma}} \left(\frac{A_1}{A_2}\right)^{\frac{\sigma-1}{\sigma}} \frac{(x^k)^{\alpha_1 \frac{\sigma-1}{\sigma}}}{(1-x^k)^{\alpha_2 \frac{\sigma-1}{\sigma}}} \frac{(x^L)^{\beta_1 \frac{\sigma-1}{\sigma}}}{(1-x^L)^{\beta_2 \frac{\sigma-1}{\sigma}}} \times D^{[\eta(\alpha_1-\alpha_2)+(\gamma_1-\gamma_2)+(\nu_1-\nu_2)]\frac{\sigma-1}{\sigma}} K^{(\alpha_1-\alpha_2)\frac{\sigma-1}{\sigma}} L^{(\beta_1-\beta_2)\frac{\sigma-1}{\sigma}} \quad (45)$$

把(35)式、(36)式分别与(44)式、(45)式对比,可以得到:

结论6:在存在数据要素局部外部性和数据隐私保护不足的环境中,政府加强数据隐私保护使下式成立,可以达到社会最优的产业结构:

$$\frac{\varphi_1}{\varphi_2} = \frac{\alpha_1 \eta \alpha_2 + v_2}{\alpha_2 \eta \alpha_1 + v_1} \quad (46)$$

根据结论5和结论6,对于数据要素局部外部性越弱的产业(如 v_1 取值更大),政府应当实施更大力度的数据权属保护(如 ξ_2 取值更大)或数据隐私保护(如 φ_2 取值更大)。这背后的经济含义如下:由于局部外部性提升整个产业全要素生产率的正向作用并没有在单个企业的决策中被考虑,对于局部外部性越强的产业,数据等要素的配置比例低于社会最优,而局部外部性越弱的产业,数据等要素配置比例高于社会最优。如果政府对局部外部性更弱的产业实施更大力度的数据权属保护或数据隐私保护,那么就会提高该产业的企业使用数据的成本,推动更高比例的要素流向局部外部性更强的产业,从而充分发挥数据要素规模效应,达到社会最优产业结构。

结论4—结论6意味着,政府在数字经济时代应坚持发展和规范并重的政策取向,在政策权衡中应充分考虑数据要素的外部性和规模效应在不同产业的差异性。实践中,结构性减税降费和产业规制政策的实施同样应当重视数据要素外部性的异质性特点。对数据要素局部外部性越强的产业,政府应实施更大规模的减税降费或财政补贴政策,在常态化监管中允许一定程度的“野蛮生长”,引导数字技术和数据应用的探索与创新,充分发挥正外部性效应;对数据要素局部外部性越弱的产业,政府应实施更大力度的数据权属保护或数据隐私保护,并适度加强数据要素的再分配收入调节,保障小微企业、工商个体户和消费者能够分享数字技术红利。在数据要素形成和数字技术创新所催生的新产业、新业态和新模式中,数据要素的外部性和规模效应大小可能会动态变化,传统产业在使用数据要素和应用数字技术过程中也可能呈现曲线型发展路径。因此,在加大结构性财税调节力度、加强数据权属保护和数据隐私保护等政策实践中,应准确把握数据要素的差异化影响,在重点领域和关键环节建立相关政策动态调整机制,以规模效应大小为重要原则适时优化政策,推动产业结构转型和生产率提升。^①

① 本文对数据要素市场化建设的相关政策做了模拟分析和讨论,具体内容参见《中国工业经济》网站(ciejournal.ajcass.com)附件。

六、结论与启示

不同于工业经济时代中的传统生产要素,作为数字经济时代中的新型生产要素,数据要素具有非竞争性和正外部性等特征,将深刻改变产业规模报酬属性。本文从这一理论视角系统研究了数据要素对产业结构转型、分配结构演化和生产率提升的变革性影响,结论如下:①数据要素在达到一定规模后会针对不同要素密集程度的产业产生不同程度的规模效应,由此推动产业结构与分配结构转型,并可能提升生产率。其影响方向不仅取决于不同产业之间产出替代弹性的大小和数据要素产出弹性的差别,还取决于数据要素的正外部性在不同产业中产生的规模效应大小。在不同产业之间产出替代弹性较高时,数据要素产出弹性越高和数据要素规模效应越强的产业所占比重将会提高;如果比重提高的产业是劳动密集型产业,那么劳动收入份额会随之上升,反之亦然。②数据要素产生的规模效应可以促使结构转型路径发生方向性转变。特别地,如果新兴产业相对传统产业的数据要素规模效应更强、劳动密集程度更高,且新兴产业和传统产业之间的产出替代弹性较高,那么工业经济时代资本收入份额和传统产业所占比重随着资本深化呈现逐渐提高趋势,就会转变为数字经济时代劳动收入份额和新兴产业所占比重随着数据要素形成呈现逐渐提高趋势,并加快提升生产率,形成效率和公平更加统一的新经济形态。③数据要素产生规模效应为更好发挥政府作用提供了有为空间,政府在数字经济时代应坚持发展和规范并重的政策取向。政府可以加强数字基础设施建设,提升数据要素生成效率和促进数据要素积累,进而加快形成规模效应和转换结构转型路径。如果数据要素产生的规模效应只限于特定产业中,那么就会形成局部外部性。对数据要素局部外部性越强的产业,政府应实施更大规模的财政补贴或减税降费,在常态化监管中注重引导创新发展;对数据要素局部外部性越弱的产业,政府应实施更大力度的数据权属保护或数据隐私保护,并适度加强再分配收入调节。

以上结论为理解数字经济时代结构转型趋势提供了新的理论视角和经济机制,也为政府发挥“有为”作用推动产业深度转型升级和新质生产力发展提供了现实依据和政策建议:

(1)把握数字技术革命机遇加快发展新质生产力,推动形成效率和公平更加统一的数字经济形态。本文研究表明,数据成为数字经济时代新型生产要素,其规模效应在不同产业中存在差别,有望促使产业结构和分配结构的转型趋势发生方向性转变,实现更高效率和更加公平的双重目标。中国经济市场规模巨大,加快释放数据资源丰富的发展潜力,以数字技术创新加快数据要素形成,有望为产业结构转型升级和发展新质生产力提供有力支撑。建议政府加大数字经济通用技术和基础技术的研发支持力度,完善人工智能、大数据、云计算等数据核心技术发展政策和治理体系,引导数字技术和数据要素优先应用于新产业、新模式和新业态,以数字技术红利释放对冲人口红利收缩对生产效率和收入分配的不利影响。

(2)全面加强数字基础设施和产业升级类基础设施建设,优化高质量数据要素供给模式。本文研究表明,在对私人部门数据要素供给的挤出效应有限的前提下,数字基础设施建设将会加快数据要素有效供给,推动数据资源向数据要素转换,通过形成规模效应增强产业结构转型动力。中国大规模基础设施建设是集中力量办大事制度优势的重要体现,而当前基础设施的内涵和范畴也在发生快速转变。全面加强数字基础设施、创新基础设施和产业升级类基础设施建设,形成现代化基础设施体系,有望在数据要素外部性和规模效应较为显著的领域率先实现深度转型升级。建议政府加快构建全国一体化大数据中心体系,布局建设新一代超算、云计算、人工智能平台、宽带基础网络

等基础设施,有序促进公共数据和政府数据开放共享,加强社会数据资源优化整合和有效利用,提高数字基础设施、公共数据和企业数据的协同化水平。

(3)优化调整财政政策、货币政策和产业政策,统筹促进数字产业化和产业数字化发展。本文研究表明,在数据要素规模效应体现在特定生产环节和领域时,政府加强对这些生产环节和领域的政策支持力度,将会优化数据要素和其他生产要素的配置结构,提升要素配置效率,加快产业结构转型升级和生产率提升。中国财政政策和货币政策具有鲜明的结构性特征,产业政策的指向性和执行力也较强。应当立足数字产业化发展中数据要素规模效应可能较为明显而在部分产业数字化转型中又相对有限的特点,充分发挥财政政策、货币政策和产业结构调整的结构性优势。建议政府依据数据要素的外部性和规模效应的相对大小,有效识别一批加快数字化转型的重点行业,在数字设备更新改造、数字技术研发应用和数字产业园区建设等领域,实施更大力度的减税降费、再贷款和产业发展政策,鼓励和支持头部平台企业带动中小企业数字化转型,以数字产业化发展促进传统产业数字化和智能化改造。

(4)健全数据要素市场化配置体制机制,加强数据权属界定、交易流通和安全保护。本文研究提出,数据要素的权属保护和隐私保护应当充分评估其不同产业和不同领域产生的规模效应差异性,考虑到允许一定程度的野蛮生长有利于发挥规模效应,数据要素市场规制应循序渐进,优先应用于外部性较弱的生产环节和领域。中国数据要素市场化进程正在加快,数据要素流通和交易规则也在探索建立中,在发展和规范中加快完善要素市场化配置体制机制,有望强化数据要素的结构性调整功能,推动改革取得积极进展。建议政府加快培育统一规范的全国一体化数据要素市场,制定数据权属保护、隐私保护和安全审查的制度和分级分类措施,优先在传统产业使用数据要素和数字技术过程中完善数据要素权属界定、开放共享、交易流通等制度和措施,依法保护企业和个人的数据信息和产权。

〔参考文献〕

- [1]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,(5):91-108.
- [2]龚强,班铭媛,刘冲.数据交易之悖论与突破:不完全契约视角[J].经济研究,2022,(7):172-188.
- [3]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,(7):60-77.
- [4]郭凯明,刘冲.平台企业反垄断、数字经济创新与产业结构升级[J].中国工业经济,2023,(10):61-79.
- [5]郭凯明,王钰冰.人工智能技术方向、时间配置结构转型与人类劳动变革远景[J].中国工业经济,2022,(12):33-51.
- [6]郭凯明,王钰冰,龚六堂.劳动供给转变、有为政府作用与人工智能时代开启[J].管理世界,2023,(6):1-15.
- [7]郭凯明,颜色,杭静.生产要素禀赋变化对产业结构转型的影响[J].经济学(季刊),2020,(4):1213-1236.
- [8]李三希,王泰茗,刘小鲁.数据投资、数据共享与数据产权分配[J].经济研究,2023,(7):139-155.
- [9]王超贤,张伟东,颜蒙.数据越多越好吗——对数据要素报酬性质的跨学科分析[J].中国工业经济,2022,(7):44-64.
- [10]谢康,夏正豪,肖静华.大数据成为现实生产要素的企业实现机制:产品创新视角[J].中国工业经济,2020,(5):42-60.
- [11]徐翔,赵墨非.数据资本与经济增长路径[J].经济研究,2020,(10):38-54.
- [12]徐翔,赵墨非,李涛,李帅臻.数据要素与企业创新:基于研发竞争的视角[J].经济研究,2023,(2):39-56.
- [13]颜色,郭凯明,杭静.需求结构变迁、产业结构转型和生产率提高[J].经济研究,2018,(12):83-96.
- [14]杨俊,李小明,黄守军.大数据、技术进步与经济增长——大数据作为生产要素的一个内生增长理论[J].经济研

- 究, 2022, (4): 103-119.
- [15] Acemoglu, D. Directed Technical Change[J]. *Review of Economic Studies*, 2002, 69(4): 781-809.
- [16] Acemoglu, D., and J. Angrist. How Large Are Human-Capital Externalities? Evidence from Compulsory Schooling Laws[J]. *NBER Macroeconomics Annual*, 2000, 15: 9-59.
- [17] Acemoglu, D., and V. Guerrieri. Capital Deepening and Non-balanced Economic Growth[J]. *Journal of Political Economy*, 2008, 116(3): 467-498.
- [18] Acemoglu, D., A. Makhdoui, A. Malekian, and A. Ozdaglar. Too Much Data: Prices and Inefficiencies in Data Markets[J]. *American Economic Journal: Microeconomics*, 2022, 14(4): 218-256.
- [19] Acemoglu, D., and P. Restrepo. The Race between Man and Machine: Implications of Technology for Growth, Factor Shares and Employment[J]. *American Economic Review*, 2018, 108(6): 1488-1542.
- [20] Aguiar, M., E. Hurst, and L. Karabarbounis. Recent Developments in the Economics of Time Use[J]. *Annual Review of Economics*, 2012, 4: 373-397.
- [21] Alvarez-Cuadrado, F., N. Van Long, and M. Poschke. Capital-Labor Substitution, Structural Change, and Growth[J]. *Theoretical Economics*, 2017, 12(3): 1229-1266.
- [22] Cong, L. W., D. Xie, and L. Zhang. Knowledge Accumulation, Privacy, and Growth in a Data Economy [J]. *Management Science*, 2021, 67(10): 6480-6492.
- [23] Farboodi, M., and L. Veldkamp. A Growth Model of the Data Economy[R]. NBER Working Paper, 2021.
- [24] Gaessler, F., and S. Wagner. Patents, Data Exclusivity, and the Development of New Drugs[J]. *Review of Economics and Statistics*, 2022, 104(3): 571-586.
- [25] Griliches, Z. The Search for R&D Spillovers[J]. *Scandinavian Journal of Economics*, 1992, 94(S): 29-47.
- [26] Herrendorf, B., and Á. Valentinyi. Endogenous Sector-Biased Technological Change and Industrial Policy [J]. *Economic Modelling*, <https://doi.org/10.1016/j.econmod.2022.105875>, 2022.
- [27] Jones, C. I. R&D-Based Models of Economic Growth[J]. *Journal of Political Economy*, 1995, 103(4): 759-784.
- [28] Jones, C. I., and C. Tonetti. Nonrivalry and the Economics of Data[J]. *American Economic Review*, 2020, 110(9): 2819-2858.
- [29] Kongsamut, P., S. Rebelo, and D. Xie. Beyond Balanced Growth[J]. *Review of Economic Studies*, 2001, 68(4): 869-882.
- [30] Lucas, R. E. On the Mechanics of Economic Development[J]. *Journal of Monetary Economics*, 1988, 22(1): 3-42.
- [31] Ngai, L. R., and C. A. Pissarides. Structural Change in a Multisector Model of Growth [J]. *American Economic Review*, 2007, 97(1): 429-443.
- [32] Prüfer, J., and C. Schottmüller. Competing with Big Data [J]. *Journal of Industrial Economics*, 2021, 69(4): 967-1008.
- [33] Rogerson, R., and J. Wallenius. Retirement, Home Production and Labor Supply Elasticities[J]. *Journal of Monetary Economics*, 2016, 78: 23-34.
- [34] Romer, P. M. Increasing Returns and Long-Run Growth[J]. *Journal of Political Economy*, 1986, 94(5): 1002-1037.
- [35] Romer, P. M. Endogenous Technological Change[J]. *Journal of Political Economy*, 1990, 98(5): S71-S102.
- [36] Simonovska, I., and M. E. Waugh. The Elasticity of Trade: Estimates and Evidence [J]. *Journal of International Economics*, 2014, 92(1): 34-50.
- [37] Veldkamp, L., and C. Chung. Data and the Aggregate Economy[J]. *Journal of Economic Literature*, 2024, 62(2): 458-484.

Scale Effect of Data, Structural Change and Productivity Growth

GUO Kai-ming¹, WANG Yu-bing², HANG Jing¹

(1. Lingnan College, Sun Yat-sen University;

2. Institute of Guangdong, Hong Kong and Macao Development Studies, Sun Yat-sen University)

Abstract: One of the key distinctions between the digital economy and the industrial economy is that data has become a new type of production factor. Data serves as the core engine that drives the deep development of the digital economy and significantly transforms the economic growth model and productivity development path, which brings new opportunities for profound industrial upgrades and the development of new quality productive forces. Based on the non-rivalry and positive externality of data, this paper systematically studies the macroeconomic effects of data on structural change, income distribution and productivity growth from a novel theoretical perspective that data changes the properties of returns to scale.

Existing studies emphasize that factor structure may change industrial structure and income distribution through the mechanism of substitutions between different industries or between different factors, induced by changes in the relative output prices of industries with different levels of factor intensity or in the relative costs of different factors. Data, as a new type of factor, has similar effects. However, different from traditional factors, data is non-rivalry and shapes positive externality. Existing literature has not fully studied the importance of these characteristics. This paper incorporates data into a standard multi-sector dynamic general equilibrium model with factor structure. In the model, data forms different extents of scale effect in different industries, and through this novel mechanism, it can drive structural change.

This paper finds that once the amount of data exceeds a certain level, it exhibits different degrees of scale effect on industries with varying levels of factor intensity, which can cause changes in industrial structure and income distribution, and promote productivity growth. The mechanism may even change the direction of structural change and present a new economic form that unifies efficiency and equity, if the new industry has larger scale effect of data and higher labor intensity than the traditional industry, and the elasticity of substitution between outputs of both industries is large. For industries with strong local scale effect of data, the government should implement more active fiscal subsidies or tax cuts, and promote innovation and development under normalized supervision. For industries with weak local scale effect of data, the government should enforce stronger protection of data property right and data privacy and moderately strengthen income redistribution.

This paper proposes that data changes not only the factor structure but also the return to scale in different industries. Even if the technology and factor structure stay invariant and the preference is homothetic, the growth of the scale of factor and output can still drive the transformation of the industrial structure and the evolution of income distribution. This paper contributes to existing literature by firstly showing that the scale effect of data can be a new driving force of structural change, which offers novel perspectives to understand the trends of structural transformation in the era of the digital economy, and derives policy implications for the government to promote industrial upgrades and the development of new quality productive forces.

Keywords: data factor; structural change; new quality productive forces; labor income share; digital economy

JEL Classification: O11 O14 O41

[责任编辑:覃毅]