

数据要素、生成式人工智能与居民就业

李晓龙 陈 宠 廖乐伟 秦雪征*

内容提要 随着数字经济的不断发展，人工智能技术和数据要素在促进经济增长的同时，也对劳动力市场带来了深远的影响。本文通过构建内生增长模型，探讨生成式人工智能对经济增长和劳动力市场的影响，尤其是数据要素在人工智能训练过程中的重要性。研究表明，在分散经济体达到稳态时，尽管人工智能会替代一部分劳动力，但居民能够在减少数据要素和劳动供给的情况下，提升消费水平和自身福利。然而，当数据要素的使用受到限制时，不仅经济增长会放缓，居民福利也将下降。进一步分析发现，劳动供给弹性的引入限制了数据要素的积累时间，导致经济体达到稳态的时间出现滞后。基于此，本文提出了关于数据要素市场建设的政策建议。

关键词 人工智能 数据要素 劳动力市场 经济增长

一 引言

随着大数据、云计算等信息技术的深入应用与加速创新，以生成式人工智能为代表的人工智能技术正深刻改变着经济活动中的各个环节。如图 1 所示，近年来中国人工智能企业的市值不断增长，呈现出快速发展的趋势。生成式人工智能的基本模式是

* 李晓龙，北京邮电大学经济管理学院，电子邮箱：xiaolongli@bupt.edu.cn；陈宠，中国社会科学院大学应用经济学院，电子邮箱：s2022116060@ucass.edu.cn；廖乐伟（通讯作者），北京大学光华管理学院，电子邮箱：2001211208@stu.pku.edu.cn；秦雪征，北京大学经济学院，电子邮箱：xqin@pku.edu.cn。本文系国家自然科学基金应急管理项目（批准号：72241419）的成果之一。本文还得到北京大学中央高校基本科研业务费专项资金项目、北京大学数量经济与数理金融教育部重点实验室的支持。

使用数据要素训练模型以提升性能^①。因此，随着人工智能的持续发展，数据要素的需求量将进一步增加。根据国家互联网信息办公室发布的《数字中国发展报告（2022 年）》，中国的数据产量已经达到 8.1 泽字节，位居世界前列^②。这意味着中国在发展生成式人工智能方面拥有得天独厚的优势。以上事实表明，中国人工智能行业正处于上行周期，且潜力巨大。同时，中国政府高度重视数字经济的发展。2024 年《政府工作报告》明确提出深化大数据、人工智能等研发应用，开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。此外，国家数据局等部门还印发了《“数据要素×”三年行动计划（2024-2026 年）》（国数政策〔2023〕11 号），旨在充分发挥数据要素的乘数效应，推动经济社会发展。因此，如何正确实施“人工智能+”与“数据要素×”两项行动，促进数字经济与实体经济的深度融合，成为加快发展新质生产力、推动中国经济高质量发展的关键。

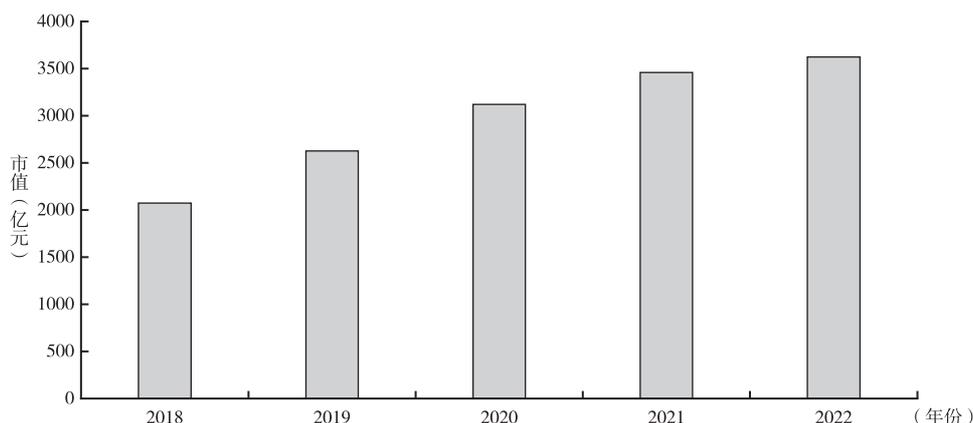


图 1 人工智能主要厂商市值（2018-2022 年）

注：样本企业选取自《互联网周刊》发布的 2023 年人工智能企业百强榜单。

资料来源：根据 A 股上市公司数据中样本企业的市值计算得到。

① 数据要素的核心特征体现在以下几个方面：非竞争性（同一数据可以被多个主体共享使用而不减少其价值，突破了传统要素的独占性）、隐私敏感性（需要通过脱敏、加密等技术手段在挖掘价值与保护个人信息之间取得平衡）以及规模报酬递增性（随着数据量的增加和维度的丰富，数据的整合与分析所产生的价值呈现指数级增长）。与资本、劳动力等传统要素的竞争性、物理性以及边际收益递减规律不同，数据以虚拟形态存在，依靠算法的协同作用创造非线性价值，同时也面临更高的治理复杂度。

② 参见 <http://www.cac.gov.cn/rootimages/uploadimg/1686402331296991/1686402331296991.pdf>。

目前已有大量文献探讨了自动化、机器人等人工智能产品对经济发展的影响 (Acemoglu & Restrepo, 2018b; Aghion et al., 2019; Berg et al., 2018)。这些研究认为, 作为新的生产要素, 人工智能不仅为经济增长提供了新动能, 同时也对劳动力市场产生了较大的冲击。根据中国企业-劳动力匹配调查数据的研究 (程虹等, 2018), 目前机器人的应用对中国劳动力市场的整体替代效应约为千分之三, 这表明自动化、机器人等人工智能产品确实对劳动力产生了一定的替代效应。然而, 以生成式人工智能为代表的其他人工智能是否也会对劳动力产生类似的替代效应, 尚没有文献进行深入探讨。

根据自主学习能力的强弱, 人工智能可分为四种类型 (Schossau & Hintze, 2023): 反应机器 (reactive machines)、有限记忆 (limited memory)、心智理论 (theory of mind) 和自我意识 (consciousness)。既往研究主要集中在自动化、机器人等人工智能产品, 这类人工智能应归类为“反应机器”。而本文的研究对象是基于数据要素训练的生成式人工智能, 属于后三种类型^①。这些类型的主要区别在于其研发过程中对数据要素的依赖程度^②。目前已有部分文献 (Neapolitan & Jiang, 2018; Russell, 2019; Russell & Norvig, 2021) 指出数据在生成式人工智能中的重要性, 这意味着不能将此类人工智能的投资 (研发) 简单视为传统的资本积累过程。因此, 本文研究的人工智能可视为一种提高生产率的通用技术, 与劳动力存在互补性关系。那么, 生成式人工智能将如何影响居民就业? 数据要素作为其研发投入要素, 市场的完备性又如何影响经济增长与劳动力市场? 这些都是值得深入思考的问题。因此, 充分探讨生成式人工智能对劳动力市场的影响, 以及数据要素在其中的作用, 对于推动中国经济的高质量发展和提升居民福利具有重要意义。

本文基于 Romer (1990) 模型, 结合数据要素、人工智能以及弹性劳动供给, 构建了一个分散经济体的基准模型, 探讨了生成式人工智能在促进经济增长的同时, 对劳动力市场产生的影响。已有文献 (Cong et al., 2021; Jones & Tonetti, 2020) 假设劳动供给是无弹性的, 并将数据要素视为独立于劳动决策的外生变量, 主要关注数据要

① 为了突出数据训练在生成式人工智能中的重要性, 后文中将不依赖数据训练、仅根据既定规则做出反应的人工智能类型称为规则型人工智能。基于数据要素进行训练的人工智能则称为生成式人工智能。

② 反应机器只需赋予机器相应的知识即可运行, 主要依赖于算法和算力。而生成式人工智能则需要通过大量数据要素的训练自主学习知识, 算法、算力和数据要素构成了生成式人工智能研发的核心三要素 (Jordan & Mitchell, 2015)。

素如何通过直接或间接进入生产函数以促进经济增长。本文的创新在于构建了一个将劳动供给内生化的框架，突破了传统研究中数据要素与劳动供给割裂的视角，揭示了二者协同作用对居民就业的动态影响。基准模型的研究结果表明，在平衡增长路径上，尽管人均数据供给量和劳动供给量减少，居民依然能够提升自身的消费水平和福利。这意味着，当数据要素的所有权归居民所有时，居民不仅可以选择出售数据要素以平衡隐私泄露带来的损失，还能够弥补因劳动供给减少导致的收入损失，即合理的数据要素产权制度能够在数字经济环境中保障居民的利益。通过比较静态分析，本文还发现劳动供给弹性对各经济变量的影响较小，而居民的隐私敏感系数、人工智能产品存量的溢出效应以及数据要素在研发过程中的贡献度，却能在很大程度上影响人均产出与劳动供给。这表明，在数字经济时代，数据要素与人工智能不仅在促进经济增长方面发挥着重要作用，同时也对劳动力市场产生了深远影响，而劳动供给弹性对经济增长与居民就业的影响相对有限。

为进一步探讨限制研发生成式人工智能所需要素供给对经济体带来的影响^①，本文在基准模型的基础上，进一步构建了中央计划者模型，并分析了其在平衡增长路径与转移动态上的表现。转移动态的结果表明，当数据要素的供给受到限制时，生成式人工智能厂商所需的研发投入不足，这不仅会推迟经济体增长的起步时间，还会使经济体进入稳态的时间延后。此外，在经济体增长起步后，居民的人均数据要素供给量将在短期内快速增长，进而引发严重的隐私泄露问题。随着数据要素供给量不再受限，生成式人工智能厂商的研发将不再受阻，经济体将更早开始增长，并快速进入稳态。同时，居民的劳动供给也将更早呈现下降趋势，并且下降速率较低。引入弹性劳动供给的设定后，本文发现，相比于 Cong et al. (2021) 的研究，经济体向稳态收敛的时间推迟了 500 期，这表明劳动供给弹性显著拉长了数据要素积累的时间。

本文剩余部分的结构安排如下：第二部分对相关文献进行简要评述；第三部分介绍基准模型的设定，并对平衡增长路径上的主要变量进行比较静态分析；第四部分构建了中央计划者模型，并对比其与基准模型在平衡增长路径上的性质，同时分析了其在转移动态中的表现；第五部分总结全文并提出政策建议。

① 由于数据要素作为经济活动的附属产物，其在市场上流通的数量不仅受到数据要素市场供求双方的影响，还受到消费量的制约。

二 文献综述

与本文密切相关的文献共有三支，分别涉及反应机器对劳动力市场的影响、生成式人工智能对劳动力市场的影响以及数据要素领域的相关研究。首先，关于反应机器对劳动力市场的影响，现有研究（王永钦、董雯，2020；Acemoglu & Restrepo, 2019；Agrawal et al., 2019；Bessen et al., 2025）认为，反应机器在就业层面表现出创造效应与替代效应，并且两者效应大小基本相等，因此不会导致就业总量出现显著波动（蔡跃洲、陈楠，2019）。从劳动力市场结构的角度看，反应机器的引入对劳动力市场的结构和稳定性产生了较大冲击。例如，Mann & Püttmann（2023）的研究发现，反应机器对制造业部门的就业与工资水平产生负向影响，而服务业部门则能够从中获益。周广肃等（2021）指出，反应机器式人工智能对低学历劳动者等劳动力市场中相对脆弱群体的冲击更为显著。王林辉等（2023a）则指出，在机器人应用的冲击下，女性群体受到的影响更大，并且存在“母职惩罚”现象（即随着女性成为母亲，其工资收入和职业地位受到负面影响）。吴立元等（2023）指出，人工智能投资改变了央行最优货币政策规则，使过度关注产出缺口的政策规则加剧了就业波动。王林辉等（2023b）基于国内数据的实证研究发现，人工智能可能降低劳动者签订劳动合同的概率，并缩短合同期限，从而影响劳动力市场的稳定性。尽管上述文献已探讨了反应机器式人工智能对劳动力市场的影响，但反应机器仅是当代人工智能技术的一种应用。本文关注的重点是生成式人工智能，因此后文将重点探讨生成式人工智能对劳动力市场的影响。

其次，关于生成式人工智能对劳动力市场影响的研究，现有文献认为，生成式人工智能同样具有创造作用（Acemoglu & Restrepo, 2024）和替代作用（Autor & Dorn, 2013；Böhm, 2020；Böhm et al., 2024）。然而，生成式人工智能的独特性使得这一领域的研究具有特殊之处。首先，生成式人工智能具有更强的任务互补性，能够更好地与劳动要素结合，提高劳动力在其他任务中的生产效率（Acemoglu & Restrepo, 2018a；Acemoglu & Restrepo, 2019）。事实上，一些基于自然实验的研究已经为生成式人工智能在提高工人生产率方面提供了证据（Brynjolfsson et al., 2025；Noy & Zhang, 2023）。其次，生成式人工智能能够与反应机器相结合，提高自动化的表现水平。例如，生成式人工智能算法能够使仓储库存管理更加合理（Acemoglu & Restrepo, 2019）。总体而言，当前关于生成式人工智能对劳动力市场影响的研究主要集中在生成式人工智能与劳动者工作中的交互以及基于指标构建的微观实证研究，鲜有研究探讨生成式人工智

能的数据训练特征如何在长期经济增长中影响居民就业的机制。本文将在这方面作出补充。

最后，关于数据要素的相关研究。目前，数据要素在宏观领域的研究主要集中在经济增长方面（Cong et al., 2021; Cong et al., 2022; Jones & Tonetti, 2020）。这些研究强调了数据要素的非竞争性，认为其非竞争性带来的规模报酬递增将产生巨大的社会收益。然而，现有研究通常设定居民的劳动供给弹性为零，因此未能直接探讨数据要素对劳动力供给量的影响。此外，部分学者（陈晓红等，2022；吴梦涛等，2023；Arrieta-Ibarra et al., 2018；Liao et al., 2024）则强调数据要素需要与劳动力相结合才能发挥相应的价值，即数据要素通过赋能人力资本提升劳动力质量。总体而言，现有文献认为数据要素需要与劳动力相结合，通过转化为人力资本促进经济增长，但直接探讨数据要素对劳动力市场影响的研究较少。因此，本文将在数字经济框架下，通过设定居民劳动供给弹性，探究数据要素作为人工智能研发投入如何影响劳动力市场。

与现有文献相比，本文的主要贡献包括三方面。第一，本文将生成式人工智能与数据要素这两个关键生产要素同时纳入内生增长模型，探讨它们如何协同作用，影响经济增长与劳动力市场，同时为生成式人工智能相关的研究提供了基础性理论框架。第二，本文发现当基准模型处于稳态时，若数据要素的所有权归居民所有，尽管居民的人均数据与劳动供给呈递减趋势，他们的消费与福利水平仍然能够持续提升。这表明，在此发展模式下，不仅可以兼顾经济增长、就业和隐私问题，还强调了数据产权制度的重要性。第三，通过系统性分析与比较分散经济体和中央计划经济体，本文发现中央计划经济体在平衡增长路径上的表现优于分散经济体，基于此本文提出了一系列政策建议。

三 分散经济模型与平衡增长路径

在 Romer（1990）的模型基础上，本文构建了一个同时考虑数据要素、生成式人工智能与弹性劳动供给的内生增长模型。与传统资本通过居民投资不断积累的方式（Acemoglu & Restrepo, 2018b）不同，本文中的生成式人工智能是通过数据要素在研发部门的投入提升性能。根据现有文献（Cong et al., 2021; Jones & Tonetti, 2020; Liao et al., 2024）的讨论，本文假设数据要素的所有权归居民，并以此为基础构建分散经济模型作为基准模型，用于探讨经济体在平衡增长路径上的性质。

(一) 代表性居民

t 时期分散经济体存在 N_t 个可存活无穷期的代表性居民，他们的数量以恒定速度 n 增长，即 $\dot{N}_t/N_t = n$ 。借鉴 Jones & Tonetti (2020) 的研究，本文认为代表性居民出售数据要素会导致隐私泄露。因此，本文设定隐私泄露带来的效用损失为 d_t^ϵ 。此外，参考 de Hek (1998) 的研究，代表性居民的劳动供给不再是无弹性的。居民因为劳动供给减少闲暇造成了效用损失。据此，本文设定劳动供给引致的效用损失为 $\theta l_t^{1+\eta}/(1+\eta)$ 。综上所述， t 时期代表性居民的效用函数 U_t 表达式为：

$$U_t = \frac{c_t^{1-\sigma} - 1}{1-\sigma} - \theta \frac{l_t^{1+\eta}}{1+\eta} - d_t^\epsilon \quad (1)$$

其中， c 为消费， l 为劳动供给， σ 为居民的相对风险厌恶系数（跨期消费替代弹性的倒数）， η 为居民的弗里希（Frisch）劳动供给弹性的倒数， θ 衡量了闲暇的重要性， ϵ 代表居民的隐私敏感系数。 t 时期代表性居民面临的预算约束为：

$$\dot{a}_t = \underbrace{(r_t - n) a_t}_{\text{投资回报}} + \underbrace{w_t l_t}_{\text{劳动收入}} + \underbrace{P_{d,t} d_t}_{\text{出售数据要素的收入}} - \underbrace{c_t}_{\text{消费支出}} \quad (2)$$

其中， a_t 为 t 时期居民的资本存量。 r_t 、 w_t 和 $P_{d,t}$ 分别表示 t 时期的利率、工资率和数据要素的价格，最终产品的价格标准化为 1。式 (2) 表明，当数据要素的所有权属于居民时，出售数据要素获得的收入 $P_{d,t} d_t$ ，与资本回报 $(r_t - n) a_t$ ，以及劳动收入 $w_t l_t$ 共同构成了 t 时期居民总收入。总收入减去消费支出 c_t 等于资产的变化量 \dot{a}_t 。鉴于数据要素来源于消费行为，因此居民的数据要素拥有量受限于其消费量。具体约束形式如下^①：

$$d_t \leq \Gamma(c_t) \quad (3)$$

为了突显数据要素量来源于消费的属性，函数 $\Gamma(\cdot)$ 需要满足以下性质： $\Gamma'(\cdot) > 0$ ， $\Gamma(0) = 0$ 且 $\lim_{x \rightarrow +\infty} \Gamma(x) \rightarrow +\infty$ 。以上函数性质表明，数据要素量将随着消费量的增加而增加。当消费量为 0 时，所产生的数据要素数量也将趋近于 0。代表性居民通过选择消费 c_t 、劳动供给 l_t 和数据供给 d_t 以最大化效用：

$$\begin{aligned} \max \int_0^{+\infty} e^{-(\rho-n)t} \left(\frac{c_t^{1-\sigma} - 1}{1-\sigma} - \theta \frac{l_t^{1+\eta}}{1+\eta} - d_t^\epsilon \right) dt \\ \text{s. t. } \dot{a}_t = (r_t - n) a_t + w_t l_t + P_{d,t} d_t - c_t \end{aligned}$$

① 基准模型中，本文暂不讨论消费量对数据要素供应量的限制。本文将在转移动态部分详细探讨这一约束条件。

其中, ρ 为主观贴现因子。求解效用最大化问题可得到以下一阶条件:

$$\frac{\dot{c}_t}{c_t} = \frac{r_t - \rho}{\sigma} \quad (4)$$

$$\frac{\dot{w}_t}{w_t} - \eta \frac{\dot{l}_t}{l_t} = r_t - \rho \quad (5)$$

$$\frac{\dot{P}_{d,t}}{P_{d,t}} - (\epsilon - 1) \frac{\dot{d}_t}{d_t} = r_t - \rho \quad (6)$$

(二) 代表性最终产品厂商

为了刻画数据要素在训练生成式人工智能中的作用, 以及生成式人工智能产品对最终产品生产的影响, 本文在 Romer (1990) 模型的基础上, 构建了一个包含代表性最终产品厂商和负责研发与生产生成式人工智能产品厂商的经济模型。

代表性最终产品厂商处于完全竞争市场环境。 t 时期, 最终产品厂商会挑选不同种类 $v \in [0, A_t]$ 的生成式人工智能产品, 并确定每种生成式人工智能产品的需求量 $x(v, t)$ 。此外, 最终产品厂商还需要雇佣一定数量的劳动力 L_t 用于最终产品生产。最终产品厂商的生产函数表达式为:

$$Y_t = L_t^\beta \int_0^{A_t} x(v, t)^{1-\beta} dv \quad (7)$$

其中, A_t 为 t 时期生成式人工智能产品的种类总数。 L_t 为最终产品厂商对劳动力的总需求量, 根据劳动力市场出清条件可知 $L_t = l_t N_t$ 。 β 衡量了劳动力在最终产品生产过程中的贡献率。最终产品厂商通过选择劳动力数量 L_t 和各种类生成式人工智能产品的数量 $x(v, t)$ 以最大化利润:

$$\max (l_t N_t)^\beta \int_0^{A_t} x(v, t)^{1-\beta} dv - w_t l_t N_t - \int_0^{A_t} P(v, t) x(v, t) dv$$

其中, $P(v, t)$ 为 t 时期第 v 种生成式人工智能产品的价格。求解利润最大化问题可得以下一阶条件:

$$w_t = \beta l_t^{\beta-1} N_t^{\beta-1} \left[\int_0^{A_t} x(v, t)^{1-\beta} dv \right] \quad (8)$$

$$P(v, t) = (1 - \beta) (l_t N_t)^\beta x(v, t)^{-\beta} \quad (9)$$

(三) 生成式人工智能产品厂商

经济体中存在无限多的潜在生成式人工智能产品厂商, 他们各自决定是否参与生成式人工智能产品的研发。一旦研发成功, 生成式人工智能产品厂商将垄断其成功研发的产品生产。本文首先讨论生产部门的利润最大化问题, 评估生成式人工智能产品

厂商研发成功后可能获得的垄断利润。然后，本文通过求解研发部门的利润最大化问题，确定成功训练生成式人工智能产品所需的数据要素使用量。

1. 生产过程

对于任意类型 $v \in [0, A_t]$ 的生成式人工智能产品，生成式人工智能产品厂商每销售一单位的生成式人工智能产品将获得 $P(v, t)$ 的收入，并需要支付 $\phi(v)$ 的成本费用。由于本文关注重点是数据要素在生成式人工智能研发过程中发挥的作用，为了简化分析，本文设定生成式人工智能厂商生产部门的边际成本为恒定常数，即 $\phi(v) \equiv \phi$ 。生产部门选择生成式人工智能产品的产量 $x(v, t)$ 以利润最大化：

$$\max P(v, t)x(v, t) - \phi x(v, t)$$

将式 (9) 代入并求解利润最大化问题可得一阶条件：

$$x(v, t) = \left[\frac{(1 - \beta)^2}{\phi} \right]^{\frac{1}{\beta}} l_t N_t = x_t \quad (10)$$

$$P(v, t) = \frac{\phi}{1 - \beta} = P_t \quad (11)$$

结合式 (8)、式 (10) 和式 (11) 以及最终产品厂商生产函数式 (7)，可得生成式人工智能产品厂商的垄断利润 $\Pi(v, t)$ 、居民工资率 w_t 和最终产品的产量 Y_t 的表达式：

$$\Pi(v, t) = P(v, t)x(v, t) - \phi x(v, t) = \beta \frac{(1 - \beta)^{\frac{2}{\beta}-1}}{\phi^{\frac{1}{\beta}-1}} l_t N_t = \Pi_t \quad (12)$$

$$w_t = \beta \frac{(1 - \beta)^{\frac{2}{\beta}-2}}{\phi^{\frac{1}{\beta}-1}} A_t \quad (13)$$

$$Y_t = \left[\frac{(1 - \beta)^2}{\phi} \right]^{\frac{1}{\beta}-1} A_t l_t N_t \quad (14)$$

2. 研发过程

在过去的知识创新文献中（如 Cong et al., 2021; Funke & Strulik, 2000; Jones, 1995; Romer, 1990），传统知识的生成与劳动力密切相关。然而，在生成式人工智能研发过程中，劳动力的作用相对有限^①。参考 Jordan & Mitchell (2015) 的研究，由于人工智能具有类似人类的自主学习能力，数据要素与当前生成式人工智能技术对新一

^① 在人工智能训练的三元要素框架（数据、算力、算法）中，本研究在核心研究目标的限定性下进行建模。本文将算力与算法视为资本性投入，并通过外生化处理（即式 (15) 中的参数 γ ），这一设定符合无监督知识生产框架。通过这一设定，算力与算法的作用被转化为技术参数，从而使得本文能够聚焦于劳动供给与数据要素的交互机制研究。

代生成式人工智能的研发具有重要影响。因此，不仅数据要素在生成式人工智能研发中扮演着关键角色，过去的生成式人工智能产品种类存量也对下一代产品研发产生溢出效应。基于这一点，本文构建了生成式人工智能产品种类拓展的模型。综上所述，生成式人工智能产品种类的增量 \dot{A}_t 受到 t 时期的生成式人工智能产品种类存量 A_t 的溢出效应和数据要素使用量 $d_t N_t$ 的影响：

$$\dot{A}_t = \gamma A_t^\zeta (d_t N_t)^\alpha \quad (15)$$

其中， $\gamma \in (0,1)$ 为生成式人工智能产品研发的效率， $\zeta \in (0,1)$ 为生成式人工智能产品存量对下一代产品研发的溢出效应， $\alpha \in (0,1)$ 为研发过程中数据要素的贡献度。 D_t 为 t 时期生成式人工智能产品厂商研发过程中数据要素的总需求量。根据数据要素市场出清条件，数据要素总需求等于总供给，即 $D_t = d_t N_t$ 。

研发部门主要关注研发成功后能否实现生成式人工智能产品厂商垄断利润的最大化。定义成功研发的产品未来可实现的总价值 $V(v,t)$ ，即将未来的持续性利润贴现到 t 时期的总价值：

$$V(v,t) = \int_t^{+\infty} e^{-\int_t^s r(\tau) d\tau} \Pi(v,s) ds = \int_t^{+\infty} e^{-\int_t^s r(\tau) d\tau} \Pi_s ds := V_t \quad (16)$$

研发部门选择研发所需的数据要素数量 $d_t N_t$ 以最大化利润：

$$P_{d,t} = \alpha \gamma A_t^\zeta (d_t N_t)^{\alpha-1} V_t \quad (17)$$

(四) 均衡定义

均衡是指生成式人工智能产品厂商选择产品价格与数据要素的需求量 $\{P(v,t), d_t\}_{v \in [0, A_t], t=0}^\infty$ 以最大化利润。自由出入市场原则决定了生成式人工智能种数的变化 $\{A_t\}_{t=0}^\infty$ 。利率、工资和数据要素价格 $\{r_t, w_t, P_{d,t}\}_{t=0}^\infty$ 由市场出清条件决定。最终产品厂商根据利润最大化原则选择劳动力和各类生成式人工智能产品的需求量 $\{l_t, x(v,t)\}_{v \in [0, A_t], t=0}^\infty$ 。代表性居民选择消费、数据供给和劳动供给 $\{c_t, d_t, l_t\}_{t=0}^\infty$ 以最大化效用。

(五) 参数校准

基准模型的参数取值主要参考既往文献的设定（表 1）。参考梅冬州和龚六堂（2011）的设定，相对风险规避系数 $\sigma = 2$ 。Chetty et al. (2011) 估算的 Frisch 劳动供给弹性为 0.75，据此本文将 η 设为 1.33。参考 Cong et al. (2021) 的研究，本文对影响生成式人工智能产品种类增量的参数设定如下：出售数据要素导致效用损失的隐私敏感系数 $\epsilon = 1.5$ ，生成式人工智能产品存量的溢出效应 $\zeta = 0.85$ ，数据在研发过程中的贡献度 $\alpha = 0.5$ 。参考 Jones & Tonetti (2020)，代表性居民人口的增长速度 $n = 0.02$ 。参考 Jones (2016) 的研究，设定代表性居民效用的贴现因子 $\rho = 0.03$ 。

表 1 参数基准值

参数含义	参数名称	数值	来源
相对风险规避系数	σ	2	梅冬州和龚六堂 (2011)
劳动供给弹性的倒数	η	1.33	Chetty et al. (2011)
隐私敏感系数	ϵ	1.5	Cong et al. (2021)
生成式人工智能产品存量的溢出效应	ζ	0.85	Cong et al. (2021)
数据要素在研发过程中的贡献度	α	0.5	Cong et al. (2021)
代表性居民人口的增长速度	n	0.02	Jones & Tonetti (2020)
代表性居民效用的贴现因子	ρ	0.03	Jones (2016)

资料来源：根据相关文献计算得到。

(六) 分散经济的平衡增长路径

在平衡增长路径上，各变量增速均达到恒定值（包括利率 $r_t = r^*$ ）。定义分散经济的稳态增速表达式为 $(\dot{\cdot})/(\cdot) := g_{(\cdot)}$ 。结合式 (4) 至式 (6)、式 (8) 至式 (12)、式 (14) 至式 (17)，以及最终产品市场出清条件 $Y_t = N_t y_t = N_t c_t$ ，可得平衡增长路径上劳动供给增速 g_l 、人均数据供给增速 g_d 、人均消费增速 g_c （人均产出增速 g_y ）以及生成式人工智能产品种类增速 g_A 的表达式：

$$g_l = - \frac{\alpha \epsilon (\sigma - 1) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (18)$$

$$g_d = - \frac{\alpha (\sigma - 1) (1 + \eta) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (19)$$

$$g_c = g_y = \frac{\alpha \epsilon (1 + \eta) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (20)$$

$$g_A = \frac{\alpha \epsilon (\sigma + \eta) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (21)$$

根据参数基准值可得各变量在平衡增长路径上的增速值如表 2 所示。

表 2 分散经济中的稳态增速值

g_A	g_l	g_d	$g_c (g_y)$
0.0391	-0.0118	-0.0183	0.0274

资料来源：根据表 1 及式 (18) 至式 (21) 计算得到。

结论 1：在分散经济的平衡增长路径上，生成式人工智能种类的增加降低了单位产出所需的劳动需求。当数据要素的所有权属于居民时，居民能够在减少劳动供给和数

据供给的同时增加消费。

当分散经济体进入平衡增长路径后，稳态增速由风险规避系数 σ 、劳动供给弹性的倒数 η 、隐私敏感系数 ϵ 、技术溢出效应 ζ 和创新过程中的数据需求弹性 α 共同决定。由表 2 可知在参数基准值设定下，生成式人工智能产品种类的增加将促进长期的经济增长。

结论 1 的具体分析如下：由于生成式人工智能产品的种类可以不断叠加，从最终产品厂商的角度看，他们可选择的生成式人工智能产品范围将不断扩大。根据式 (13)，随着生成式人工智能技术水平的提升，工资率将持续增长，即劳动力的雇佣成本将不断上升。然而，生成式人工智能产品的价格却是恒定的（参见式 (11)）。根据成本最小化原则，最终产品厂商倾向于选择更多的生成式人工智能产品参与生产。对于生成式人工智能厂商而言，由于生成式人工智能产品存量具有溢出效应，当存量足够多时，当前研发所需的数据要素需求量将逐渐减少。从居民的角度看，当他们拥有数据要素的所有权时，出售数据所得的补贴不仅可以弥补隐私泄露带来的效用损失，还能弥补由于劳动供给减少而导致的收入损失。综上所述，在分散经济体的平衡增长路径上，居民的隐私泄露与闲暇不足等问题得到了有效解决，同时他们的福利水平也得到了提升，形成了因人工智能发展带来的“技术红利共享”机制。在这种经济模式下，数据要素在居民端与企业端都发挥着重要作用。然而，这一良性循环机制的前提是居民必须拥有数据要素的所有权。

（七）比较静态分析

本部分通过变动居民特征参数 $\{\eta, \epsilon\}$ 和生成式人工智能产品的特征参数 $\{\zeta, \alpha\}$ ，探究居民特征和生成式人工智能产品特征对平衡增长路径的影响。

结论 2：在分散经济体的平衡增长路径上，人均数据要素供给增速随劳动供给弹性倒数的增加而下降，放缓了人均产出和消费增速；而隐私敏感系数、生成式人工智能产品存量的溢出效应与数据要素在研发过程中的贡献度提升，都能够显著促进生成式人工智能产品种类增速，增加人均产出和消费增速，同时降低劳动供给增速。

具体的分析如下：如图 2 所示，在分散经济体平衡增长路径上，居民的劳动供给弹性对人均产出的影响并不明显。这是因为，当居民更愿意通过增加劳动供给而非出售数据获得收入时，系统中核心的增长动力“数据要素驱动生成式人工智能发展”受到抑制。此外，隐私敏感系数的提升能够加快生成式人工智能产品种类拓展速度，进而提高人均产出增速。这与 Cong et al. (2021) 中的观点相似，即居民需要通过更高的经济增速弥补更多因隐私泄露带来的负外部性。敏感系数越大（数据供给弹性越小），

居民将出售更多的数据获得数据补贴，通过增加人均消费和减少劳动供给提升自身福利。

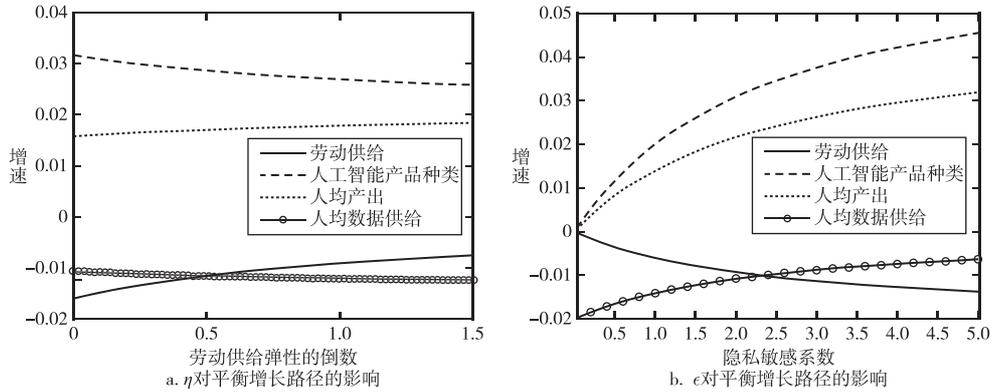


图2 居民特征参数对平衡增长路径的影响

资料来源：根据表1及式(18)至式(21)计算得到。

如图3所示，数据要素在研发过程中的贡献度越大，生成式人工智能厂商研发部门对于当期数据要素的使用效率越高。生成式人工智能产品存量的溢出效应越高，研发部门对过去数据要素的使用效率越高。总而言之，以上两个参数的变化都能够有效提升生成式人工智能产品的研发效率，促进人均产出增长，同时降低劳动供给与数据供给增速。这表明，增加以上参数取值不仅可以减少居民对隐私泄露的顾虑，增加居民的闲暇时间，还能提升居民的福利水平。

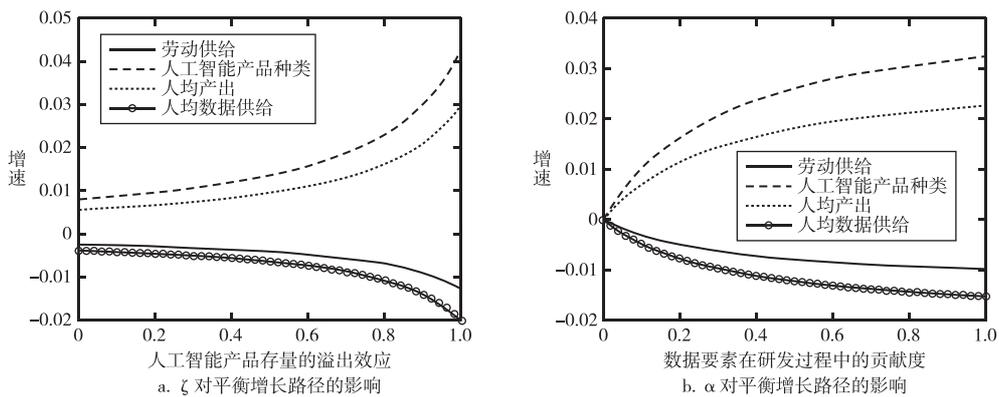


图3 生成式人工智能产品特征参数对平衡增长路径的影响

资料来源：根据表1及式(18)至式(21)计算得到。

四 中央计划者经济模型与转移动态

(一) 中央计划者经济

在前文中，本文已经探讨了分散经济体的平衡增长路径及其性质。参考 Cong et al. (2021)，本文希望在分散经济体模型的基础上，进一步研究中央计划者经济模型，主要原因包括两方面。第一，在分散经济体中，由于生成式人工智能产品厂商的垄断行为，市场势力的存在必然导致稳态结果无法达到帕累托最优。因此，通过对这两种经济体进行比较，本文能够揭示基准模型中的市场扭曲现象。第二，数据要素作为生成式人工智能研发过程中的关键投入，其在市场中的流通量不仅受到供需双方的影响，还会受到消费量的显著影响。因此，通过构建中央计划者经济模型，本文可以进一步探讨当数据供应量受到限制时，这种限制将如何影响经济增长的转移动态过程。

在资源约束下，中央计划者最大化代表性居民效用。根据最终产品市场出清原则，居民总消费等于净产出：

$$N_t c_t = (l_t N_t)^\beta \int_0^{A_t} x(v, t)^{1-\beta} dv - \int_0^{A_t} \phi x(v, t) dv$$

在 t 时期，给定生成式人工智能产品种类 $\{A_t\}_{t=0}^\infty$ ，中央计划者选择第 v 种生成式人工智能产品的数量 $x(v, t)$ 以最大化净产出。求解上述最优化问题可得一阶条件：

$$x(v, t) = \beta \left(\frac{1-\beta}{\phi} \right)^{\frac{1}{\beta}} A_t N_t \quad (22)$$

将式 (22) 代入净产出表达式可得资源约束表达式：

$$c_t = \beta \left(\frac{1-\beta}{\phi} \right)^{\frac{1}{\beta}-1} A_t l_t \quad (23)$$

给定资源约束式 (23) 与生成式人工智能产品种类增量式 (15)，中央计划者最大化代表性居民效用问题可表示为：

$$\begin{aligned} \max \int_0^{+\infty} e^{-(\rho-n)t} \left(\frac{c_t^{1-\sigma} - 1}{1-\sigma} - \theta \frac{l_t^{1+\eta}}{1+\eta} - d_t^\epsilon \right) dt \\ \text{s. t. } c_t = \beta \left(\frac{1-\beta}{\phi} \right)^{\frac{1}{\beta}-1} A_t l_t \\ \dot{A}_t = \gamma A_t^\xi (d_t N_t)^\alpha \end{aligned}$$

其中，人均最优消费 c_t 、人均数据供给 d_t 与劳动供给 l_t 均为选择变量，生成式人工智能产品种类 A_t 为状态变量。构建现值汉密尔顿函数，分别对 c_t 、 l_t 、 d_t 、 A_t 求偏导可得一阶条件：

$$c_t^{-\sigma} = \lambda_t \quad (24)$$

$$\theta l_t^\eta = \beta \left(\frac{1-\beta}{\phi} \right)^{\frac{1}{\beta}-1} \lambda_t A_t \quad (25)$$

$$\epsilon d_t^{\epsilon-1} = \alpha \gamma A_t^\zeta d_t^{\alpha-1} N_t^\alpha \quad (26)$$

$$\lambda_t \beta \left(\frac{1-\beta}{\phi} \right)^{\frac{1}{\beta}-1} l_t + \gamma \zeta \mu_t A_t^{\zeta-1} (d_t N_t)^\alpha = -\dot{\mu}_t + (\rho - n) \mu_t \quad (27)$$

其中， λ_t 和 μ_t 分别对应约束条件式(23)和式(15)的影子价格。

(二) 中央计划者经济的平衡增长路径

定义中央计划者经济体中主要变量的稳态增速表达： $(\dot{\cdot})/(\cdot) := \tilde{g}_{(\cdot)}$ 。结合式(15)、式(23)至式(27)，可得平衡增长路径上稳态增速表达式：

$$\tilde{g}_l = - \frac{\alpha \epsilon (\sigma - 1) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (28)$$

$$\tilde{g}_d = - \frac{\alpha (\sigma - 1) (1 + \eta) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (29)$$

$$\tilde{g}_c = \frac{\alpha \epsilon (1 + \eta) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (30)$$

$$\tilde{g}_A = \frac{\alpha \epsilon (\sigma + \eta) n}{(1 - \zeta) (\sigma + \eta) \epsilon + \alpha (\sigma - 1) (1 + \eta)} \quad (31)$$

结论3：中央计划者经济的平衡增长路径上可求得 $\tilde{g}_A > 0$ ， $\tilde{g}_l < 0$ ， $\tilde{g}_d < 0$ ， $\tilde{g}_c > 0$ 。

对比分散经济和中央计划者经济可以发现，由于分散经济下生成式人工智能产品厂商的垄断加价行为，分散经济的人均产出 $(\beta [(1-\beta)^2/\phi]^{1/\beta-1} A_t l_t)$ 小于中央计划者经济的人均产出 $(\beta ((1-\beta)/\phi)^{1/\beta-1} A_t l_t)$ 。此外，因为生成式人工智能产品厂商的垄断定价策略是固定比例的成本加成定价，生成式人工智能产品厂商的垄断加价行为只会影响人均产出水平，并不影响平衡增长路径上人均产出增速。这一结果与 Cong et al. (2021) 的研究结论一致。

(三) 中央计划者经济的转移动态

为了进一步探究数据要素供给受到限制对经济增长动态过程的影响，本部分进一步分析中央计划者经济体的转移动态过程。式(3)表明，市场中数据要素的流通数量不仅受到供需双方的影响，还会受到消费量的限制。参考 Jones (2016) 的研究，本文

通过中央计划者经济的转移动态过程模拟限制数据要素供给的影响。根据式 (15)、式 (23) 至式 (27) 可得动态系统:

$$-\sigma g_c(t) = g_\lambda(t) \quad (32)$$

$$\eta g_l(t) = g_\lambda(t) + g_A(t) \quad (33)$$

$$(\epsilon - \alpha) g_d(t) = g_\mu(t) + \zeta g_A(t) + \alpha n \quad (34)$$

$$\dot{g}_\mu(t) = [(1 + \eta) g_l(t) - g_A(t) - g_\mu(t)] [\zeta g_A(t) \quad (35)$$

$$+ g_\mu(t) - \rho + n] - \zeta \dot{g}_A(t)$$

$$g_c(t) = g_A(t) + g_l(t) \quad (36)$$

$$\dot{g}_A(t) = g_A(t) [(\zeta - 1) g_A(t) + \alpha g_d(t) + \alpha n] \quad (37)$$

其中, $\dot{g}_{(\cdot)}(t) := (\dot{\cdot})_t / (\cdot)_t$ 表示变量或影子价格在时期 t 的增速。参考 Judd (1998) 的分析, 本文从中央计划者经济体的稳态值出发, 模拟经济体的转移动态过程。

结论 4: 在数据要素训练生成式人工智能的生产模式下, 当数据要素供给量不受限制时, 经济体将更早开始增长, 并快速进入稳态。居民劳动供给也会更早呈现出下降趋势, 且下降的速率更快。当数据要素供给受到限制时, 经济体增长起步的时间将推迟, 进入稳态的时间更晚。居民需要在较长时间内提供更多的劳动力, 直至经济体增长起步后劳动供给增速才开始下降。

在数据要素驱动生成式人工智能的经济体中, 当数据要素的供应量受到限制时, 生成式人工智能厂商的研发部门将缺乏关键投入要素, 从而导致生成式人工智能产品种类的拓展速度放缓, 这使得经济体难以实现快速发展。相反, 当数据要素供应充足时, 生成式人工智能产品的研发将顺利进行, 经济体能够迅速起步并较早进入稳态阶段 (图 4)。

此外, 当数据要素供应受限时, 最终产品厂商面临的挑战是无法获得丰富的生成式人工智能产品, 因此它们必须依赖更多的劳动力进行生产。对于居民而言, 如果他们无法通过出售数据获得收入补贴, 唯一的选择便是增加劳动供给以弥补福利损失。同时, 在经济体发展初期, 人均数据供给量通常会呈现爆炸式增长, 以确保生成式人工智能厂商的研发工作能够顺利进行。

结论 5: 引入弹性劳动供给设定后, 动态系统的收敛速度显著降低。

通过对比本文和 Cong et al. (2021) 的转移动态结果可以发现, 不论数据要素是否存在限制, Cong et al. (2021) 的动态系统都会在 400 期左右收敛。但引入弹性劳动供给设定后, 动态系统会在 1000 期左右收敛。出现该结果的原因是在本文设定下, 代表

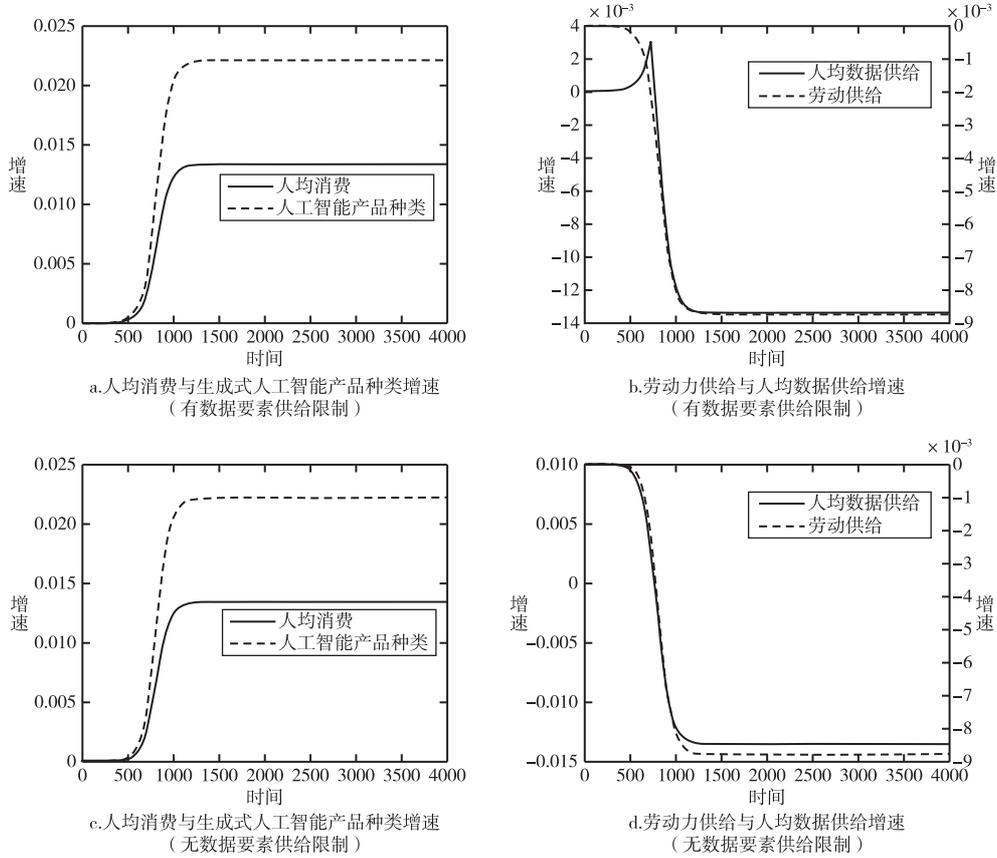


图4 有(无)数据要素供给限制时经济体的转移动态过程

资料来源：根据表1及式(32)至式(37)计算得到。

性居民供给劳动和出售数据是互相替代的：数据要素受限将增加企业用工需求，从而引致居民增加劳动供给、减少数据要素供给，最终降低了生成式人工智能研发企业的技术积累速率。

综上所述，数据要素在数字经济发展中的重要性不言而喻。它不仅是生成式人工智能研发厂商的关键生产要素，也是居民改善自身福利的重要资产，贯穿于整个生产和消费环节。数据要素供给量受到限制，不仅会导致生成式人工智能产品种类增速放缓，从而拖慢经济增长，还会减少居民出售数据所获补贴，增加居民劳动供给，进而导致居民福利水平的整体下降。同时，忽视居民劳动供给弹性将低估数据要素在转移动态过程中的作用。

五 结论与政策建议

现有文献已深入探讨人工智能对劳动力市场的重大影响，但大部分研究集中在反应机器式人工智能领域，这不仅导致生成式人工智能领域的学术空白，也使得结合数据要素与生成式人工智能的相关研究相对匮乏。因此，本文通过构建数据要素训练生成式人工智能的内生增长模型，并结合比较静态和转移动态分析，揭示了这一生产模式下数据要素、劳动力、生成式人工智能与居民福利水平之间的相互关系及其发展规律。

基于上述分析，本文得出了以下主要结论。首先，数据要素确权在数据要素训练生成式人工智能的生产模式中至关重要。如果数据要素归居民所有，居民在减少劳动供给和数据供给的同时，可以提高消费和福利水平。其次，隐私敏感系数、人工智能产品存量溢出效应以及数据要素贡献度都会影响生成式人工智能技术的发展。当居民对数据隐私更为重视、生成式人工智能产品存量溢出效应更大、数据要素在研发中的贡献更高时，生成式人工智能技术的发展速度加快，这将推动经济体在稳态时表现出更高的经济增长率和居民福利水平。再次，宽松的数据限制规则能够促进生成式人工智能产品种类的积累，从而加速经济体进入快速增长阶段。居民能够更早地享受到更多的闲暇时间和更高的福利水平。最后，本文发现劳动供给弹性的加入将显著延长经济体收敛至稳态的时间，这表明数据要素和劳动供给的互动极大影响了经济发展速度。

面对中国人口老龄化加速、经济结构转型压力加大以及外部冲击不断增加的现实，劳动力市场的供求关系日益紧张。然而，充分利用数据要素训练生成式人工智能的生产模式，可以在一定程度上缓解劳动力市场的供需矛盾，并进一步提升居民福利水平。基于此，本文提出以下政策建议。首先，明确数据要素的居民所有权。确保居民能够拥有自身数据的所有权，并通过出售数据要素获得收入补贴。这不仅能够提升生成式人工智能研发部门对数据要素的利用效率，也能成为居民增收的有效途径，从而缓解劳动力市场的供需矛盾。

其次，增强居民对数据隐私的重视程度。中国的数据隐私立法正在逐步完善，但居民隐私意识较弱。因此，在明确数据产权的同时，应加强居民对数据隐私保护的重视。研究表明，居民对隐私问题的关注度越高，数据将得到更高效地利用，生成式人工智能技术的提升速度加快，居民福利与经济增长也会相应提升。

再次，提高数据要素的利用率并加强技术研究。虽然数据要素是当前生产模式中

的关键投入，但其给居民带来负效用的特征不容忽视。尽管在稳态下人均数据供给量会逐步降低，隐私问题不再加剧，但隐私保护仍需得到重视。提升数据要素利用率不仅能促进经济增长和居民福利水平，还能减少当期数据要素的使用量。

最后，鼓励数据要素流通与交易。研究表明，当数据要素的供给没有限制时，数字经济体能够更早进入稳态阶段。为此，本文建议采取以下措施：建立完善的数据交易平台，降低数据交易过程中的机会成本，促进数据需求和供给的高效匹配；在数据要素市场化的前提下，适当放松隐私限制政策，以促进更多数据的利用；加强数据收集技术的研发，提高数据收集的效率和精准度；鼓励发展数据密集型产业（如制造业和互联网服务业），这些产业能够产生和收集大量数据，为经济结构转型提供支持。

参考文献：

- 蔡跃洲、陈楠（2019），《新技术革命下人工智能与高质量增长、高质量就业》，《数量经济技术经济研究》第5期，第3-22页。
- 陈晓红、李杨扬、宋丽洁、汪阳洁（2022），《数字经济理论体系与研究展望》，《管理世界》第2期，第208-224页。
- 程虹、陈文津、李唐（2018），《机器人在中国：现状、未来与影响——来自中国企业-劳动力匹配调查（CEES）的经验证据》，《宏观质量研究》第3期，第1-21页。
- 梅冬州、龚六堂（2011），《新兴市场经济国家的汇率制度选择》，《经济研究》第11期，第73-88页。
- 王林辉、钱圆圆、宋冬林、董直庆（2023a），《机器人应用的岗位转换效应及就业敏感性群体特征——来自微观个体层面的经验证据》，《经济研究》第7期，第69-85页。
- 王林辉、钱圆圆、周慧琳、董直庆（2023b），《人工智能技术冲击和中国职业变迁方向》，《管理世界》第11期，第74-95页。
- 王永钦、董雯（2020），《机器人的兴起如何影响中国劳动力市场？——来自制造业上市公司的证据》，《经济研究》第10期，第159-175页。
- 吴立元、王忬、傅春杨、龚六堂（2023），《人工智能、就业与货币政策目标》，《经济研究》第1期，第56-72页。
- 吴梦涛、张龙天、武康平（2023），《数字经济背景下的人力资本积累新模式探究》，《经济学报》第2期，第1-27页。

- 周广肃、李力行、孟岭生 (2021), 《智能化对中国劳动力市场的影响——基于就业广度和强度的分析》, 《金融研究》第 6 期, 第 39 - 58 页。
- Acemoglu, Daron & Pascual Restrepo (2018a). Artificial Intelligence, Automation, and Work. *NBER Working Paper*, No. 24196.
- Acemoglu, Daron & Pascual Restrepo (2018b). The Race Between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108 (6), 1488 - 1542.
- Acemoglu, Daron & Pascual Restrepo (2019). Automation and New Tasks: How Technology Displaces and Reinstates Labor. *Journal of Economic Perspectives*, 33 (2), 3 - 30.
- Acemoglu, Daron & Pascual Restrepo (2024). Automation and Rent Dissipation: Implications for Wages, Inequality, and Productivity. *NBER Working Paper*, No. 32536.
- Aghion, Philippe, Benjamin Jones & Charles Jones (2019). Artificial Intelligence and Economic Growth. In Ajay Agrawal, Joshua Gans & Avi Goldfarb (eds.), *The Economics of Artificial Intelligence: An Agenda*. Chicago: University of Chicago Press, pp. 237 - 282.
- Agrawal, Ajay, Joshua Gans & Avi Goldfarb (2019). Artificial Intelligence: The Ambiguous Labor Market Impact of Automating Prediction. *Journal of Economic Perspectives*, 33 (2), 31 - 50.
- Arrieta-Ibarra, Imanol, Leonard Goff, Diego Jiménez-Hernández, Jaron Lanier & Glen Weyl (2018). Should We Treat Data as Labor? Moving Beyond “Free”. *AEA Papers and Proceedings*, 108, 38 - 42.
- Autor, David & David Dorn (2013). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review*, 103 (5), 1553 - 1597.
- Berg, Andrew, Edward Buffie & Luis-Felipe Zanna (2018). Should We Fear the Robot Revolution? (The Correct Answer Is Yes). *Journal of Monetary Economics*, 97, 117 - 148.
- Bessen, James, Maarten Goos, Anna Salomons & Wiljan van den Berge (2025). What Happens to Workers at Firms That Automate? *Review of Economics and Statistics*, 107 (1), 125 - 141.
- Böhm, Michael (2020). The Price of Polarization: Estimating Task Prices Under Routine-Biased Technical Change. *Quantitative Economics*, 11 (2), 761 - 799.
- Böhm, Michael, Hans-Martin von Gaudecker & Felix Schran (2024). Occupation Growth, Skill Prices, and Wage Inequality. *Journal of Labor Economics*, 42 (1), 201 - 243.

- Brynjolfsson, Erik, Danielle Li & Lindsey Raymond (2025). Generative AI at Work. *Quarterly Journal of Economics*, qjae044.
- Chetty, Raj, Adam Guren, Day Manoli & Andrea Weber (2011). Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins. *American Economic Review*, 101 (3), 471 – 475.
- Cong, Lin William, Danxia Xie & Longtian Zhang (2021). Knowledge Accumulation, Privacy, and Growth in a Data Economy. *Management Science*, 67 (10), 6480 – 6492.
- Cong, Lin William, Wenshi Wei, Danxia Xie & Longtian Zhang (2022). Endogenous Growth Under Multiple Uses of Data. *Journal of Economic Dynamics and Control*, 141, 104395.
- de Heek, Paul (1998). An Aggregative Model of Capital Accumulation with Leisure-Dependent Utility. *Journal of Economic Dynamics and Control*, 23 (2), 255 – 276.
- Funke, Michael & Holger Strulik (2000). On Endogenous Growth with Physical Capital, Human Capital and Product Variety. *European Economic Review*, 44 (3), 491 – 515.
- Jones, Charles & Christopher Tonetti (2020). Nonrivalry and the Economics of Data. *American Economic Review*, 110 (9), 2819 – 2858.
- Jones, Charles (1995). R&D-Based Models of Economic Growth. *Journal of Political Economy*, 103 (4), 759 – 784.
- Jones, Charles (2016). Life and Growth. *Journal of Political Economy*, 124 (2), 539 – 578.
- Jordan, Michael & Tom Mitchell (2015). Machine Learning: Trends, Perspectives, and Prospects. *Science*, 349 (6245), 255 – 260.
- Judd, Kenneth (1998). *Numerical Methods in Economics*. Cambridge: The MIT Press.
- Liao, Lewei, Xuezheng Qin, Xiaolong Li & Liutang Gong (2024). Creative Destruction, Human Capital Accumulation, and Growth in a Digital Economy. *Macroeconomic Dynamics*, 28 (5), 1206 – 1230.
- Mann, Katja & Lukas Püttmann (2023). Benign Effects of Automation: New Evidence from Patent Texts. *Review of Economics and Statistics*, 105 (3), 562 – 579.
- Neapolitan, Richard & Xia Jiang (2018). *Artificial Intelligence: With an Introduction to Machine Learning (2nd)*. Boca Raton: CRC Press.
- Noy, Shakked & Whitney Zhang (2023). Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence. *Science*, 381 (6654), 187 – 192.
- Romer, Paul (1990). Endogenous Technological Change. *Journal of Political Economy*, 98

(5, Part 2), S71 – S102.

Russell, Stuart & Peter Norvig (2021). *Artificial Intelligence: A Modern Approach* (4th). Hoboken: Pearson.

Russell, Stuart (2019). *Human Compatible: AI and the Problem of Control*. New York: Viking.

Schossau, Jory & Arend Hintze (2023). Towards a Theory of Mind for Artificial Intelligence Agents. Paper presented at the 2023 Artificial Life Conference ALIFE 2023: Ghost in the Machine, Sapporo, Japan, July 24 – 28.

Data Factors, Generative Artificial Intelligence and Employment

Li Xiaolong¹, Chen Chong², Liao Lewei³ & Qin Xuezheng⁴

(School of Economics and Management, Beijing University of Posts and Telecommunications¹;

Faculty of Applied Economics, University of Chinese Academy of Social Sciences²;

Guanghua School of Management, Peking University³;

School of Economics, Peking University⁴)

Abstract: With the continuous development of the digital economy, artificial intelligence technologies and data factors have profoundly impacted the labor market while promoting economic growth. This paper explores how generative artificial intelligence impacts economic growth and the labor market by constructing an endogenous growth model, focusing on the importance of data factors in the AI training process. The research findings indicate that when decentralized economies reach a steady state, although AI may replace some labor, residents can improve their consumption levels and welfare even with a reduced supply of data factors and labor. However, when the use of data factors is restricted, not only does economic growth slow down, but the welfare of residents also declines. Further analysis reveals that the introduction of an elastic labor supply limits the time available for accumulating data factors, resulting in a delay for the economy to reach a steady state. Based on this, we propose policy recommendations for building the data factors market.

Keywords: artificial intelligence, data factors, labor market, economic growth

JEL Classification: E24, O11, O30

(责任编辑: 马 超)