DOI:10.19361/j.er.2026.04.04

熵减与增长

——数据要素和人工智能如何重塑经济增长路径

徐 翔 李 涛*

摘要: 本文构建内生化信息过载现象的动态增长模型,将人工智能刻画为执行熵减功能的特殊资本,揭示其将高熵原始数据提炼为低熵有效信息以驱动创新的微观机制。研究发现:第一,原始数据对创新效率的贡献呈倒 U 型特征,信息过载会显著抑制知识生产;第二,在数据积累与人工智能投资的动态演化中,经济体可能因投资不足陷入低水平增长陷阱,亦可通过充足投资跃迁至高生产率路径;第三,市场存在两类新型失灵,即源于隐私成本的数据生成负外部性,以及人工智能投资状态依存的双重外部性。据此,本文为数字经济时代纠正新型市场失灵、引导宏观经济向社会最优增长路径跃迁提供了坚实的理论依据。

关键词: 数据要素;人工智能;信息熵;内生增长;市场失灵

中图分类号: F061.2;F062.5

一、引言

当前,大数据与人工智能正作为新型通用目的技术,被视为重塑宏观经济增长路径的核心动能(Crafts, 2021;刘涛雄等,2024)。主流理论框架中,数据要素的非竞争性使其能够被多主体无成本复用,从而在算法的驱动下打破传统物理要素的报酬递减约束,释放显著的规模报酬递增效应(Jones and Tonetti, 2020)。然而,这一宏观层面的理论预期在实践中正遭遇严峻的"数据悖论"。现实中的原始数据并非即插即用的同质化要素,而是高度非结构化、充满噪音的复杂信息集合。当数据积聚的规模超越微观主体的有效处理边界时,企业普遍陷入"数据丰富而信息贫乏"的困境:过量的原始数据不仅难以直接转化为有价值的经济洞察,反而会增加搜寻成本、引发决策摩擦。此外,从预测经济学的视角来看,数据作为一种旨在降低不确定性的要素投入,必然受制于固有的边际收益递减规律。随着数据量级的增加,其对预测误差的边际缩减效应存在客观下限,这为数据驱动的效率提升设定了自然边界(Veldkamp and Chung, 2024)。这意味着,若缺乏相匹配的智能提炼能力,数据规模的盲目扩张不仅无法自然转化为生产率的跃升,反而可能成为阻碍创新活动的结构性摩擦。

*徐翔,中央财经大学经济学院,邮政编码:100081,电子邮箱:xuxiang@cufe.edu.cn;李涛(通讯作者),中央财经大学经济学院,邮政编码:100081,电子邮箱:litao@cufe.edu.cn。

本文得到国家自然科学基金项目"数据要素与技术创新:理论机制与实证分析"(72473167)的资助。感谢第十一届中国经济增长与发展青年学者论坛暨《经济评论》第 11 期审稿快线审稿专家和编辑部的宝贵意见,文责自负。

上述微观层面的双重摩擦——信息过载导致的处理能力瓶颈与收益递减带来的自然边界——共同导致宏观层面的经济增长未能达到预期。尽管对数据要素驱动生产率高速增长的乐观预测屡见不鲜(徐翔、赵墨非,2020;Cong et al., 2022;刘涛雄等, 2023),但这一新兴虚拟要素在实际应用中所暴露出的复杂性,使其对宏观经济的贡献至今仍显温和,尚未充分释放理论上的巨大潜力。一系列实证研究也已揭示,尽管企业在数据与人工智能领域进行了巨额投资,其对全要素生产率的提升作用在多数情况下依然有限且不显著(Acemoglu, 2025;Babina et al., 2024)。这一现象,便被归纳为大数据与人工智能时代的"新生产率悖论"(Brynjolfsson et al., 2017)。微观"数据悖论"与宏观"新生产率悖论"的交织共存,表明现有文献在刻画数据要素驱动经济增长的内在机制时仍存在局限,亟待引入新的理论视角加以深化。

本文提出,人工智能对数据的处理过程并非仅仅是消耗燃料,而是作为一种精炼技术,将高信息熵的、嘈杂的原始数据提炼为低信息熵的、用于知识生产的有效信息。"熵"的概念最早源于热力学,后由克劳德·香农借鉴并引入信息论,确立了"信息熵"这一概念,用以度量系统信息的不确定性或随机性(Shannon, 1948;Cabrales et al., 2013)。各种经济社会活动产生的原始数据,无论是来自消费者的交易记录、物联网传感器的实时读数,还是社交媒体上的非结构化文本,都充满了噪音、冗余、不一致和不相关的信息,从而具备较高的信息熵。人工智能——特别是机器学习与生成式大模型——具备强大的熵减能力:其通过模式识别、预测分类与异常检测,能从多模态的原始海量数据中精准提取高信噪比的有效信息。由此可见,人工智能的宏观经济价值绝非简单的资本深化,而是从根本上破解了知识生产过程中的信息过载瓶颈,进而重塑了经济的内生增长轨迹。

为深入剖析上述机制,本文构建了一个包含数据要素、人工智能与信息熵机制的动态一般均衡模型。该模型以扩展产品种类的内生增长理论为基础,将数据界定为一种伴随消费活动生成且存在折旧的虚拟要素。本文的核心理论设定在于,将信息熵函数嵌入知识生产过程,使有效数据量服从于原始数据与人工智能资本存量之比的倒 U 型分布。这一设定刻画了信息过载的微观摩擦:当原始数据规模越过由人工智能处理能力决定的临界阈值时,数据对创新效率的边际贡献将由正转负;而人工智能作为执行熵减功能的专有资本,能内生地拓展经济体的信息承载边界。此外,模型显性引入了数据生成衍生的隐私成本,将其作为一种负外部性纳入家庭效用函数。在此基础上,本文系统求解了去中心化市场均衡的平衡增长路径,并通过与社会计划者最优配置的比较,识别了数字经济特有的新型市场失灵,进而提出旨在纠正此类微观摩擦的针对性干预机制,为构建适应数字经济特征的宏观治理体系提供了理论依据。

本文的研究贡献主要体现在以下四个方面。第一,将信息过载这一在多学科中被广泛认可的现象,通过信息熵函数的形式内生化于宏观增长模型,为理解数据驱动型经济的内在摩擦提供了新的理论视角。第二,重新定义了人工智能在经济增长中的角色,揭示了对人工智能的投资如何通过提升数据处理能力,帮助经济体克服数据信息过载,实现向更高水平平衡增长路径的跃迁。第三,识别出一种新的市场失灵——人工智能投资的双重外部性。单个企业投资人工智能时主要考虑其自身收益,但这种投资通过影响全社会的信息熵水平,对所有其他创新企业产生影响。由于企业无法将这一外部性内部化,将导致市场的人工智能投资系统性地偏离社会最优水平。第四,基于对市场去中心化均衡与社会最优配置的对比

剖析,揭示了纠正上述新型市场失灵、引导资源向社会最优水平配置的内在逻辑,为构建适应数字经济特征的宏观治理体系提供了理论支撑。

二、理论基础与典型事实

(一)文献述评

本文的理论构建主要交汇于数据要素的宏观建模、信息过载机制以及人工智能经济学分析这三大领域。近年来,将数据作为新型资产纳入宏观经济模型已成为前沿研究的热点。Jones 和 Tonetti(2020)的开创性工作基于数据的非竞争性,探讨了不同产权安排与隐私成本下的福利效应。随后,Cong 等(2021)、刘涛雄等(2024)进一步从时间非竞争性等维度剖析了数据驱动经济增长的潜在路径。Farboodi 和 Veldkamp(2021)等学者揭示了数据反馈循环对市场结构的影响,指出缺乏技术基础设施的企业在处理海量数据集时面临严峻挑战。然而,现有宏观增长模型大多隐含了一个强假设,即数据对生产率的贡献是单调递增的,这显然未能充分刻画由于数据规模激增而带来的微观负面效应。

这种由海量数据引发的负面效应,在跨学科研究中被广泛定义为信息过载。自 Simmel(1950)和 Miller(1960)的早期观察以来,过量信息导致决策质量下降、创新受阻的现象,已在管理学与心理学等领域得到详尽证实(Eppler and Mengis, 2004;Anderson and De Palma, 2012)。近期的经济学前沿文献也开始关注这一摩擦,如 Acemoglu 等(2022)强调了海量原始数据引发的经济低效,Dugast 和 Foucault(2025)发现数据丰裕反而会恶化金融市场的价格发现功能。尽管微观层面的实证证据十分丰富,但信息过载这一深刻的现实摩擦,至今尚未被正式、内生地整合到主流宏观经济增长框架中,导致理论模型与现实中的"数据悖论"存在脱节。

面对海量高熵数据带来的增长摩擦,人工智能理应在宏观经济中扮演化解信息过载的破局角色,但现有关于人工智能与经济增长的文献尚未充分捕捉到这一机制。当前的主流研究多将人工智能视为一种高级自动化技术,其主要增长路径被设定为替代劳动力、提升资本效率或创造新任务(Acemoglu and Restrepo, 2018;陈彦斌等,2019;李健斌、周浩,2025)。Aghion 等(2017)、Agrawal 等(2019)以及 Trammell 和 Korinek(2023)探讨了人工智能如何通过自动化研发任务加速思想生产;Acemoglu(2025)则测算了人工智能在任务层面对未来十年经济增长的潜在拉动作用,并指出其单纯的生产力增强效果可能相对有限。有鉴于此,本文的创新之处在于跳出"人工智能等于自动化生产力"的传统设定,将其重新定位为解决信息过载的关键熵减工具,从而在统一的宏观增长框架内,为理解数据积累与人工智能投资的互补联动提供了一个全新的微观理论机制。

(二)人工智能的熵减机制

本文的核心观点是,人工智能——特别是机器学习与生成式人工智能——的核心价值正在于其强大的熵减能力。将未经处理的高熵数据直接作为研发投入,不仅会导致搜寻与试错成本激增,更易因信息过载引发边际创新报酬为负,这构成了微观层面"数据悖论"的成因。相反,人工智能凭借模式识别与高维映射等算法,能够从海量多模态数据中精准剥离噪音,提炼出高信噪比的有效信息,从而有效降低信息熵。宏观视角下,对人工智能资本的投资等价于扩张经济体的信息承载与提炼边界。AI 资本的丰裕度直接决定了经济体有效消化原始数据的阈值,从而保障海量数据能够持续转化为创新动能(Cabrales et al., 2013)。

数据积累规模与数据处理能力是否匹配,内生决定了要素投入与创新产出之间的非线性关系。在数据相对稀缺的经济起步阶段,原始数据的扩容能带来显著的边际创新收益,表现为效率曲线的单调递增。然而,随着数据总量的指数级膨胀,经济体不可避免地逼近由既有 AI 资本存量决定的信息处理能力边界。一旦越过临界点,高熵的原始数据流便开始造成拥堵,其处理成本将超过其所能带来的边际收益。创新效率曲线随之呈现倒 U 型回落,经济体陷入信息过载状态,更多的原始数据反而导致了更低的创新效率。由此可见,人工智能投资的宏观机制,在于不断向右上方拓展该倒 U 型曲线的效率峰值(见图 1),从而确保经济体在更高的数据规模下维持创新的可持续性。

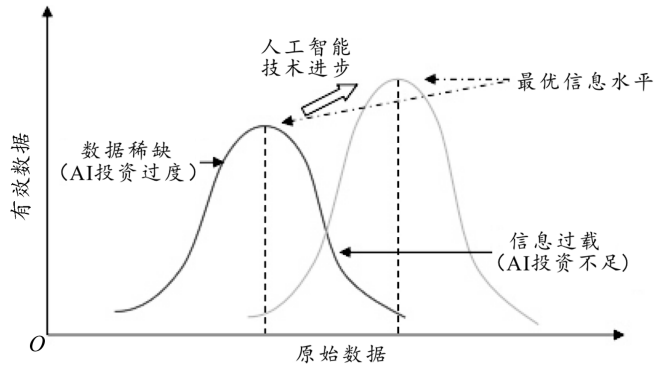

图 1　数据、人工智能与信息熵曲线示意图

在现实中,人工智能可以通过多种技术渠道提升数据要素质量,降低数据中存在的噪声。主要包括:(1)模式识别与分类,从非结构化数据(如图像、文本)中提取结构化特征;(2)预测与异常检测,在充满噪音的时间序列数据中识别趋势和异常信号;(3)数据整合与清洗,将来自多源、多模态的数据进行标准化和对齐,剔除冗余和错误信息。这些过程共同将企业面临的信息过载困境,转化为真正的决策优势。上述一系列微观应用场景中的熵减功能,构成了本文宏观建模的经验基础。接下来,本文将把这一数据提炼过程抽象为信息熵函数,并将其嵌入动态一般均衡框架中进行严格求解。

三、模型设定

为将前述理论机制形式化,本部分构建一个动态一般均衡模型。设定一个包含代表性家庭、厂商和研发部门的经济环境,并在此基础上引入数据要素、人工智能与信息熵的核心互动机制。

(一)家庭与偏好

假设经济中存在一个无限期限的、追求效用最大化的代表性家庭。总劳动力(人口)L_t以恒定的外生速率 n 增长,即:

$$\dot{L}_t = nL_t \tag{1}$$

家庭的效用取决于人均消费$c_t = C_t/L_t$和由人均原始数据存量$d_t = D_t/L_t$带来的隐私成本,本文采用连续时间动态系统表示,变量上方的点号表示其时间导数。家庭的目标是选择人均消费路径c_t以最大化其总贴现效用:

$$U=\int_0^{\infty} e^{-(\rho-n)t}\left(\frac{c_t^{1-\sigma}-1}{1-\sigma}-\chi\frac{d_t^{1+\psi}}{1+\psi}\right)dt \tag{2}$$

(2)式中:ρ 是时际贴现率,因此有效贴现率为$\rho-n$。$\sigma>0$ 是相对风险厌恶系数。本文模型的一个关键特征是引入了非货币性的隐私成本项$\chi\frac{d_t^{1+\psi}}{1+\psi}$。该成本被建模为对家庭效用的直接减损,其大小取决于经济中的人均原始数据存量d_t。参数$\chi(\chi>0)$表示隐私成本强度,衡量了社会对隐私的重视程度;参数 ψ 表示隐私成本的曲率,$\psi>0$ 则确保了边际隐私成本是递增的,与实证研究中关于隐私担忧随数据收集量增加而加剧的发现相符。由于单个消费者对人均数据存量的影响微乎其微,他们不会将自己的消费行为对社会整体隐私风险的贡献内部化,这使得隐私成本成为一种典型的负外部性。

家庭的人均财富a_t来源于劳动报酬$w_{labor,t}$和资产回报$r_t a_t$(r_t 表示经济中的资产实际回报率),其动态演化遵循标准的预算约束:

$$\dot{a_t}=(r_t-n)a_t+w_{labor,t}-c_t \tag{3}$$

(二)生产部门

经济中包含两个产品生产部门:最终品部门和中间品部门。

首先是最终品生产部门。遵循经典文献设定,假设一个代表性的、完全竞争的企业使用劳动力$L_{Y,t}$和A_t种类的差异化中间品$x_{i,t}$来生产最终品Y_t。其生产函数采用标准的 Dixit-Stiglitz 形式:

$$Y_t=L_{Y,t}^{1-\alpha}\int_0^{A_t}x_{i,t}^{\alpha}di \tag{4}$$

(4)式中:A_t是经济中可用的中间品种类的测度,代表知识或技术的现有存量。$\alpha(0<\alpha<1)$代表中间品在最终品生产中的份额。

其次是中间品生产部门。每一种中间品$x_{i,t}$由一个拥有该产品设计专利的垄断厂商生产,其中 $i\in[0,A_t]$。为简化模型,假设生产一单位的中间品$x_{i,t}$需要一单位的资本。因此,中间品部门的总资本需求为 $K_{Y,t}=\int_0^{A_t}x_{i,t}di$。

(三)知识、数据与人工智能的动态演化

本小节阐述模型的核心动态机制。遵循相关文献做法,将原始数据D_t视为一种可积累的资本存量(徐翔、赵墨非,2020;刘涛雄等,2023),其积累过程取决于总消费C_t(作为数据"投资"的来源)和数据自身的折旧。这一做法捕捉了数据作为经济活动副产品的特性,同时也反映了旧数据的时效性会逐渐降低。据此,数据积累的跨期公式为:

$$\dot{D}_t=\eta C_t-\delta_D D_t \tag{5}$$

(5)式中:η 是数据的生成效率参数,反映了每次消费行为能转化为多少原始数据的能力。δ_D为数据的折旧率或过时率,反映了旧数据可能随着时间推移而失去其预测或分析价值,现有文献对此有从 0.05 到 0.5 的多个预测值。公式(5)捕捉了"交易即生成数据"的核心思想:经济活动越活跃,产生的数据就越多。由于D_t是非竞争性的,假设数据一旦产生,就可以被所有研发部门无成本地使用。

新思想的产生,即知识存量A_t的增长,发生在研发(R&D)部门。其生产函数遵循现代半内生增长理论的设定,引入了此类文献常见的"淘金效应"和"拥挤效应":

$$\dot{A}_t = \delta L_{A,t}^{\lambda} D_{eff,t} A_t^{\varphi} \tag{6}$$

(6)式中:δ 是基础研发效率参数,$L_{A,t}$是投入研发部门的劳动力。参数 φ 代表知识的跨期溢出弹性,即"淘金效应",$0<\varphi<1$ 反映了随着知识的积累,新的突破性思想变得越来越难发现。新引入的参数 λ 表示研发劳动力的产出弹性,当 $0<\lambda<1$ 时,意味着研发投入存在"拥挤效应"或收益递减。与标准模型不同,设定研发活动的生产率还受到有效数据$D_{eff,t}$的直接影响。下面对$D_{eff,t}$的生成进行讨论。

信息熵函数与人工智能资本是本模型的核心机制,用于捕捉信息过载现象。有效数据$D_{eff,t}$是原始数据存量D_t和人工智能资本(AI capital)存量$K_{AI,t}$的函数:

$$D_{eff,t} = \left(\frac{D_t}{K_{AI,t}^{\gamma}}\right) exp\left(1-\frac{D_t}{K_{AI,t}^{\gamma}}\right) \tag{7}$$

生成有效数据的函数形式具有以下关键特性。一是原始数据投入的倒 U 型关系。对于给定的人工智能资本,当原始数据D_t从零开始增加时,$D_{eff,t}$先增加后降低。这一设定充分捕捉了信息过载的效应:在数据稀缺的初始阶段,更多的数据有助于创新;数据规模超过某个阈值后,如果分析数据的人工智能资本未能随之增加,过多的噪音和冗余信息会淹没研发人员,降低研发效率。二是人工智能的熵减作用。人工智能资本$K_{AI,t}$代表了经济体处理和分析数据的能力。参数 $\gamma>0$ 衡量了人工智能资本处理数据的效率。$K_{AI,t}$的增加会将倒 U 型曲线的峰值点(即$D_t=K_{AI,t}^{\gamma}$)向右移动,这意味着拥有更高人工智能水平的经济体能够有效利用更大规模的原始数据,从而推迟或缓解信息过载的出现。

在信息论中,高熵代表着系统的高度不确定性、随机性和无序性。原始数据D_t因其庞杂、充满噪音和冗余,可被视为一个高熵系统。有效数据$D_{eff,t}$则是经过处理、提炼后的低熵信息,可直接用于知识生产。需要强调的是,这里的人工智能资本指代广义的计算能力——即企业使用硬件、软件流程或预训练模型处理信息的能力,而并非特指生成式模型或大语言模型,尽管它们是计算密集型发展轨迹的例证。这种更广阔的框架反映出本文关心的是从数据中学习的基础设施,而非所使用的具体架构①。

设定比率 $z \equiv D_t/K_{AI,t}^{\gamma}$可以被理解为经济体面临的"信息负荷"或未经处理的熵的程度。函数$D_{eff,z}=ze^{1-z}$的结构巧妙地捕捉了熵减过程的两个对立效应:线性项 z 代表了从更多原始数据中提取潜在有用信号的收益,而指数衰减项e^{1-z}则代表了随着信息负荷加重,筛选和处理成本呈指数级增长的负面效应。因此,该函数形式为信息过载这一复杂的经济现象提供了一个简洁、可解且具有坚实理论基础的数学表达。

该经济作为一个整体,面临以下三条资源约束,分别是劳动力($L_t=L_{Y,t}+L_{A,t}$)、最终品产出($Y_t=C_t+I_t$)与投资约束($I_t=\dot{K}_t+\dot{K}_{AI,t}$)。在本文模型中,总投资并非用于积累单一的同质化资本,而是需要在两种具有根本不同功能的资本之间进行分配。这一设定引入了一个核心的经济权衡,即社会需要在扩大现有生产能力与提升信息处理能力之间作出选择。

四、均衡分析与人工智能的作用

本部分首先求解模型在市场机制下的去中心化均衡,并刻画其长期动态,即平衡增长路

①此处的信息熵函数是在借鉴信息论核心思想的基础上,为捕捉信息过载现象而构造的一个具有理想经济学性质的简化函数,而非对香农熵的直接数学应用。

径(BGP)。随后将深入分析模型的核心机制,探讨在不同的人工智能投资水平下,经济体可能经历的两种截然不同的增长路径。

(一)去中心化均衡

在去中心化均衡中,家庭和厂商都根据市场价格进行最优化决策,并且所有市场(最终品、中间品、劳动力和资产市场)都出清。

1.家庭与厂商的最优化

代表性家庭通过构建现值汉密尔顿函数求解其最优化问题。其一阶条件给出了标准的人均消费欧拉方程:

$$\frac{\dot{c}_t}{c_t}=\frac{1}{\sigma}(r_t-\rho) \tag{8}$$

(8)式表明,消费的增长率由资本的净回报率r_t和时间偏好ρ之间的差异决定。

家庭最大化其终身效用,受制于其财富积累的动态预算约束,即远期人均资产的贴现值趋于0,同时满足标准的横截性条件$\lim_{t\to\infty}e^{-\int_0^t(r_s-n)ds}a_t=0$。

下面讨论生产厂商。最终品部门是完全竞争的,其利润最大化行为决定了对每种中间品$x_{i,t}$的需求函数。给定中间品价格$p_{i,t}$,最终品厂商选择$x_{i,t}$以最大化利润:

$$\max_{\{L_{Y,t},x_{i,t}\}}\Pi_{Y,t}=Y_t-\int_0^{A_t}p_{i,t}\,x_{i,t}di-w_{labor,t}\,L_{Y,t} \tag{9}$$

利润最大化的一阶条件给出了对劳动力和每种中间品的需求函数。设$p_{i,t}$为中间品价格,对中间品$x_{i,t}$的需求为:

$$p_{i,t}=\alpha\,L_{Y,t}^{1-\alpha}x_{i,t}^{\alpha-1} \tag{10}$$

基于对称性假设,所有中间品厂商面临相同的需求曲线,并设定相同的价格和产量,即$p_{i,t}=p_t$和$x_{i,t}=x_t$。每个拥有专利的中间品厂商都是其产品领域的垄断者,它选择价格p_t以最大化其利润$\pi_t=(p_t-1)x_t$,其中边际成本为1(因为生产一单位中间品需要一单位最终品)。结合需求函数,利润最大化的一阶条件得出,中间品价格是其边际成本的一个固定加成:$p_t=p=1/\alpha$。由于价格是固定的,每个厂商的瞬时利润与其产量成正比:$\pi_{i,t}=(p-1)x_t=\frac{1-\alpha}{\alpha}x_t$。

新思想的创造者(即新厂商或研发机构)可以自由进入研发部门。在均衡状态下,创造一项新专利的价值$V_{A,t}$必须等于其成本。假设研发成本完全由劳动力构成,则创造新专利的边际成本($p_{A,t}$)为$p_{A,t}=w_t\Big/\frac{\partial\dot{A}}{\partial L_A}$。同时,专利的价值$V_{A,t}$等于其带来的未来利润流的现值。在均衡中,满足标准的资产定价方程(无套利条件):

$$r_tV_{A,t}=\pi_t+\dot{V}_{A,t} \tag{11}$$

这一条件确保了持有专利的回报率$r_tV_{A,t}$等于其带来的利润流(股息)和资本利得之和。

2.平衡增长路径

在平衡增长路径(BGP)上,所有总量宏观变量(C,Y,K,K_{AI},D,A)以一个共同的常数增长率增长,而所有人均变量(c,y,k,k_{AI},d)也以一个共同的常数增长率增长。劳动力在生产和研发部门的分配比例L_Y/L和L_A/L保持不变。为了使"信息负荷"$z\equiv D_t/K_{AI,t}^{\gamma}$在BGP上保

持为一个常数,D_t和$K_{AI,t}^{\gamma}$必须以相同的增速增长。考虑到K_{AI}和 D 的相同增长率,在此后设定$\gamma=1$。这一假设简化了稳态分析,也不影响模型关于信息过载和人工智能熵减作用的核心机制与结论。

从知识生产函数出发,可以推导出各变量的长期增长率。知识存量A_t的增长率为:

$$g_A=\frac{\dot{A}_t}{A_t}=\delta\frac{L_{A,t}^{\lambda}}{A_t^{1-\varphi}}D_{eff,t} \tag{12}$$

为了使g_A在 BGP 上保持为一个常数,等式右侧也必须为一个常数。在 BGP 上,$L_{A,t}$以人口增长率 n 增长(因为L_A/L 为常数),因此分子中的$L_{A,t}^{\lambda}$以 λn 的速率增长。同时,$D_{eff,t}$是一个常数,因为它取决于人均数据与人均人工智能资本的比率,而这些比率在 BGP 上是稳定的。因此,为了使整个表达式为常数,分母$A_t^{1-\varphi}$的增长率必须与分子$L_{A,t}^{\lambda}$的增长率相等。A_t的增长率为g_A,所以$A_t^{1-\varphi}$的增长率为$(1-\varphi)g_A$。

令分子与分母的增长率相等,得到 $\lambda n=(1-\varphi)g_A$。整理后,得到知识的长期增长率为$\lambda n/(1-\varphi)$。在对称均衡下,人均产出 y 与知识存量 A 成正比,因此人均产出的长期增长率g_y等于知识增长率。这说明,在一个由人口增长驱动的创新经济中,人均 GDP 的长期增长率由人口增长的速度 n、研发的拥挤程度 λ 以及创新的难易程度(由 $1-\varphi$ 体现)这三个外生参数共同决定。

上述结果也与近年来被重新强调的"新生产率悖论"相一致:尽管人工智能等颠覆性技术正以前所未有的速度发展,但宏观层面测算的全要素生产率(TFP)增长却持续乏力(Acemoglu, 2025)。模型显示,数据要素的积累与人工智能资本的投资水平,尽管能够通过提升有效数据的水平以显著影响经济的产出水平,但它们并不直接改变由人口和创新活动基础规律决定的长期增长率。换言之,人工智能的革命性作用可能更多地体现在帮助我们更高效地利用现有知识和数据,而非直接加速新知识边界的扩张速度。因此,即使能够观察到人工智能在微观层面带来了巨大的效率提升,宏观增长率仍然可能受制于更为基础且变化缓慢的因素。

(二)人工智能的变革性作用

尽管长期人均产出增长率g_y由外生参数决定,但这并不意味着数据和人工智能并不重要。相反,它们通过影响研发部门的生产率,共同决定了经济体能够达到的平衡增长路径的水平。不同经济体可能具备相同的长期增长率,但其人均收入水平可能因其数据处理能力的不同而存在巨大差异。

考虑一个经济体,其对人工智能资本$K_{AI,t}$的投资倾向较低。高昂的固定成本、投资与监管不确定性、金融摩擦与互补性要素的缺失都可能是其深层次原因(Acemoglu and Restrepo, 2018)。在经济增长过程中,总消费C_t和数据存量D_t随之增加。如果$K_{AI,t}$的增长跟不上D_t的增长,比率$D_t/K_{AI,t}$将会持续上升,最终将经济体推向信息熵函数倒 U 型曲线的右侧下降区域。此时,一个与微观层面的"数据正反馈循环"相反的负反馈循环将可能被触发。

伴随着经济发展与非竞争的数字技术进步,数据要素规模迅速增长。初始的经济增长导致数据量D_t上升。由于人工智能资本$K_{AI,t}$不足,$D_t/K_{AI,t}^{\gamma}$超过了最优阈值 1,经济进入信息熵曲线的右侧下降区域,此时$\frac{\partial D_{eff,t}}{\partial D_t}<0$。此外,原始数据的进一步增加反而导致有效数据$D_{eff,t}$

下降,这直接降低了研发部门的生产率。研发效率的下降意味着创造一项新技术的预期回报降低。由于创新激励减弱,厂商将减少对研发的投入,将劳动力从研发部门$L_{A,t}$转移到生产部门$L_{Y,t}$。虽然长期人均产出增长率g_y由外生参数n、φ、λ决定,但研发效率的下降会降低经济的整体生产力水平。此时,经济体仍会增长,但会沿着一条人均收入水平较低的平衡增长路径前进。这个过程形成了一个恶性循环,即增长的副产品(数据)最终压制了经济所能达到的繁荣高度。经济体可能因此陷入一个长期低收入水平的"信息过载陷阱"。由于作为$K_{AI,t}$互补品的有效数据$D_{eff,t}$正在减少,投资新人工智能资本的预期回报也随之下降。这进一步抑制了对$K_{AI,t}$的投资,确保了$D_t/K_{AI,t}$的比率持续处于高位。

现在考虑一个正向冲击,例如技术突破使得人工智能资本的生产成本大幅下降,或一国政府意识到人工智能技术的重要性,通过一系列政策鼓励对$K_{AI,t}$的投资。这一冲击会激励经济体增加对$K_{AI,t}$的投资,使其增长速度与数据生成的速度相匹配,从而将比率$D_t/K_{AI,t}$维持在最优阈值附近。结果是,经济体能够持续、有效地利用其不断增长的原始数据存量D_t。在这种情况下,人工智能不仅仅是另一种形式的物质资本,它的核心作用是提升了创新过程中另一个关键要素——数据要素的生产率。因此,人工智能投资的激增可以将整个经济体从一个低水平的平衡增长路径,跃迁到一个高水平的平衡增长路径。这与将人工智能视为一种通用目的技术的观点相一致,支持人工智能通过从根本上改变知识创造的过程,从而可能带来变革性的长期增长效应的观点(Kurzweil, 2005; Aghion et al., 2017)。例如在生命科学等数据密集型领域,人工智能资本的外生冲击已成功释放了海量存量数据的潜在价值,清晰验证了这一向高生产力前沿跃迁的理论机制(Nussinov et al., 2022)。

(三)稳态结果

如前文所述,模型的长期人均产出增长率g_y完全由人口增长和创新技术的结构性参数决定:

$$g_y=\frac{\lambda n}{1-\varphi} \tag{13}$$

利率r必须是一个常数。根据消费欧拉方程(公式(8)),可以得到稳态利率:

$$r=\rho+\sigma\, g_y=\rho+\sigma\,\frac{\lambda n}{1-\varphi} \tag{14}$$

这两个结果独立于人工智能的投资水平,为求解其他变量提供了锚点。

在对称均衡下,所有中间品厂商生产相同的数量$x=K/A$。根据中间品厂商的利润最大化条件与最终品部门对中间品的需求,可以推导出在平衡增长路径上,"生产性资本-产出"比是一个仅由技术参数α决定的常数:

$$\left(\frac{K}{Y}\right)^{*}=\alpha^2 \tag{15}$$

劳动力的分配比例$s_A=L_A/L$和$s_Y=L_Y/L$由研发部门和最终品生产部门的自由进入条件决定。在BGP上,专利的价值V_A等于其未来利润流的现值,也等于创造该专利的成本,因此$V_A=\dfrac{\pi}{r-n}$。经过将工资、利润、利率和资本产出比等变量代入$V_A=p_A$,可以推导出研发劳动力与生产劳动力的比例。结合劳动力约束$s_A+s_Y=1$,可以解出BGP上劳动力的分配比例:

$$s_A^{*}=\frac{\alpha\,\lambda g_y}{\rho+(\sigma+\alpha\lambda)\,g_y-n},\ s_Y^{*}=1-s_A^{*} \tag{16}$$

值得注意的是,劳动力的长期分配比例同样仅由模型的深层参数决定,在 $\gamma=1$ 的假设下,与人工智能的投资水平无关。

接下来,计算人工智能投资的稳态水平。人工智能投资的作用在于决定经济体所能达到的增长路径的水平。首先,从知识增长率的 BGP 条件中,可以反解出在 BGP 上必须维持的有效数据水平:

$$D_{eff}^{*}=\frac{g_A}{\delta}\left(\frac{A_t^{1-\varphi}}{L_{A,t}^{\lambda}}\right)=\frac{g_A}{\delta\ (s_A^{*})^{\lambda}}\left(\frac{A_t}{L_t^{\lambda/(1-\varphi)}}\right)^{1-\varphi} \tag{17}$$

由于括号内的项(人均有效知识存量$\frac{A_t}{L_t^{\lambda/(1-\varphi)}}$)在 BGP 上是一个常数,因此$D_{eff}^{*}$也是一个由所有基础参数决定的常数。为了维持这一必需的研发效率,经济体必须通过投资人工智能资本$K_{AI,t}$,将其面临的信息负荷 $z\equiv D_t/K_{AI,t}$稳定在某个最优水平z^{*}上,该水平满足:

$$D_{eff}^{*}=z^{*}e^{1-z^{*}} \tag{18}$$

经济体对人工智能的投资强度可以用 BGP 上的人均人工智能资本与人均数据存量的比率来衡量,定义为$\kappa_{AI}\equiv\left(\frac{k_{AI}}{d}\right)^{*}=\left(\frac{K_{AI}}{D}\right)^{*}$。因此,一个更高的人工智能投资强度意味着一个更低的z^{*},这使得经济体能够在D_{eff}曲线更有效率的区间运行($\kappa_{AI}^{*}=1/z^{*}$)。

利用数据积累方程与宏观资源约束来推导最终的宏观指标比率。在 BGP 上,所有总量以$g_Y=g_y+n$ 的速度增长。数据积累方程变成$g_Y D=\eta C-\delta_D D$,可得$\left(\frac{D}{C}\right)^{*}=\frac{\eta}{g_Y+\delta_D}$。这一比例表明 BGP 上的数据存量与消费流量成正比。现在,就可以将人工智能资本与居民消费联系起来:

$$K_{AI}=\kappa_{AI}^{*}D=\kappa_{AI}^{*}\left(\frac{\eta}{g_Y+\delta_D}\right)C\equiv\Omega C \tag{19}$$

(19)式中,$\Omega\equiv\kappa_{AI}^{*}\frac{\eta}{g_Y+\delta_D}=\frac{1}{z^{*}}\frac{\eta}{g_Y+\delta_D}$。$\Omega$ 是一个综合参数,由创新需求决定的人工智能投资强度κ_{AI}转化为一个宏观的人工智能资本-消费比率。最后,将此关系代入最终品产出约束 $Y=C+I=C+\dot{K}+\dot{K}_{AI}=C+g_Y(K+K_{AI})$,可以得到"消费-产出"比的最终表达式:

$$\left(\frac{C}{Y}\right)^{*}=\frac{1-g_Y\alpha^2}{1+g_Y\Omega} \tag{20}$$

以及"人工智能-产出"比的表达式:

$$\left(\frac{K_{AI}}{Y}\right)^{*}=\Omega\left(\frac{C}{Y}\right)^{*}=\frac{\Omega(1-g_Y\alpha^2)}{1+g_Y\Omega} \tag{21}$$

上述稳态方程组是本模型的核心结果。它们清晰地表明,尽管长期人均增长率g_y外生,但决定经济长期福利水平的消费和资本的相对规模,内生地取决于由创新效率所要求的人工智能投资强度。一个需要更高研发效率(更高的D_{eff}^{*})的经济体,需要维持一个更优的信息负荷z^{*},这决定了其必需的人工智能投资强度κ_{AI}^{*},并最终体现在更高的"人工智能-产出"比与更低的"消费-产出"比上,从而将经济体推向一条人均产出与消费水平都更高的平衡路径。

(四)模型拓展

在基准模型中,假设经济体的人工智能投资强度κ_{AI}给定,从而决定了经济所处的平衡增

长路径的水平。然而在现实中，κ_{AI}是无数企业进行利润最大化决策的加总结果。本拓展旨在将这一决策过程内生化，从而解出市场自发形成的人工智能投资水平。

本文假设，企业投资人工智能资本k_{AI}能够获得直接的私人回报，同时也会产生外部性。具体而言，修改知识生产函数，使其同时包含企业自身的人工智能资本$k_{AI,i}$和全社会的总人工智能资本K_{AI}。参考关于企业人工智能投资的近期研究（Babina et al., 2024），本文将企业的知识生产函数设定为：

$$\dot{A}_{i,t}=\delta L_{A,t}^{\lambda}k_{AI,i,t}^{\varepsilon}D_{eff,t}A_t^{\varphi} \tag{22}$$

（22）式中：$D_{eff,t}$仍然由全社会的信息负荷$z\equiv D_t/K_{AI,t}$决定（为保持BGP的存在性，继续假设$\gamma=1$）。参数$\varepsilon>0$代表企业人工智能投资的私人产出弹性。企业在决策时，会选择最优的$k_{AI,i}$以最大化其创新价值，但由于其规模相对于整个经济是没有影响的，它会将D_{eff}视为外生给定，不会考虑其自身投资$k_{AI,i}$通过增加K_{AI}对D_{eff}产生的影响，从而造成企业人工智能投资的外部性。在这一设定下，企业投资人工智能资本的边际收益（$MPK_{AI,private}$）必须等于其边际成本。在均衡中，人工智能资本的租金率等于真实利率r。因此，企业的最优决策满足：

$$r=MPK_{AI,private}=\varepsilon\frac{V_{A,t}\dot{A}_{i,t}}{k_{AI,i,t}^{*}} \tag{23}$$

（23）式中：$V_{A,t}$代表创新的价值，一般可以理解为专利价值。在对称均衡下，所有企业都相同（$k_{AI,i}=k_{AI}$，$L_{A,i}=\frac{L_A}{A}$），且$\dot{A}_i=\frac{\dot{A}}{A}=g_A$。可以解出均衡中单个企业的人工智能资产存量$k_{AI,DE}^{*}=\frac{\varepsilon V_A^{*}g_A}{r}$。其中$V_A^{*}$是BGP上以人均产出标准化的创新价值。由于$K_{AI}=A\cdot k_{AI}$，可以得到市场自发形成的“人工智能-产出”比：

$$\left(\frac{K_{AI}}{Y}\right)^{*}_{DE}=\frac{A\cdot k_{AI,DE}^{*}}{Y}=\frac{A}{Y}\frac{\varepsilon V_A^{*}g_A}{r} \tag{24}$$

（24）式反映由市场微观决策内生决定的人工智能投资水平。然而，社会计划者在进行决策时，会考虑人工智能投资的全部社会回报，即私人部门回报加上其通过降低信息负荷z而产生的外部性。在人工智能投资内生的设定下，对于人工智能投资进行补贴就具备重要的实践意义。人工智能资本的社会边际产出（$MPK_{AI,social}$）为：

$$MPK_{AI,social}=\underbrace{\varepsilon\frac{V_{A,t}\dot{A}_t}{K_{AI,t}}}_{\text{私人回报}}+\underbrace{\frac{V_{A,t}\dot{A}_t}{K_{AI,t}}(z-1)}_{\text{外部性}} \tag{25}$$

由（25）式可知，当经济处于信息过载区域（$z>1$）时，外部性项为正。此时，每个企业的人工智能投资通过降低全社会的z值，为所有其他企业创造正外部性。由于企业忽略了这一社会收益，市场自发形成的投资水平将低于社会最优水平。当经济处于数据稀缺区域（$z<1$）时，外部性项为负。此时，每个企业的人工智能投资反而会因过度降低z值（使其偏离最优峰值1），而对其他企业造成负外部性。在这种情况下，市场自发形成的投资水平将高于社会最优水平。其经济学含义是，人工智能的产业政策设计取决于经济体当前的信息负荷水平。

五、福利分析与数值模拟

家庭与厂商在进行最优化决策时不会考虑其行为对社会整体福利的全部影响,因此市场自发形成的均衡配置通常并非社会最优。本部分将识别导致这种无效率的四种市场失灵来源,通过数值模拟探讨市场去中心化均衡与社会最优配置的潜在偏离。

(一)市场失灵的主要来源

通过将去中心化均衡中个体决策的驱动因素与社会总福利的决定因素进行对比,我们可以在模型中识别出四种主要的市场失灵。

首先是数据生成的负外部性,即隐私成本。市场失灵的第一个来源在于数据的生成过程。根据模型设定,原始数据是消费活动的副产品,而人均数据存量的增加会给代表性家庭带来直接的负效用,即隐私成本。然而,单个家庭在决定其消费水平时,仅考虑其自身的边际效用与成本,并将社会人均数据存量d_t视为外生给定。他们不会将自己的消费行为通过增加数据总量D_t而对所有其他家庭造成的边际隐私成本内部化。因此,数据的社会成本高于其私人成本,导致市场均衡下的消费水平和数据生成量系统性地高于社会最优水平。

其次是研发投入的负外部性,即拥挤效应。研发过程本身也存在一种负外部性,体现在知识生产函数中研发劳动力的产出弹性$\lambda<1$。这就意味着,当一个企业决定多用研发人员时,它只考虑这个新员工对其自身创新产出的边际贡献。由于研发投入存在收益递减,新员工的加入会轻微地降低经济中所有其他研发人员的平均生产率,如雷同的研发活动或对有限科研资源的竞争。由于企业不会考虑其用人决策对其他企业研发效率造成的这种负面影响,市场均衡下的研发投入可能会高于社会最优水平。

再者是知识创造的正外部性,即知识溢出。新思想的产生依赖于现有知识存量,当一个企业成功创新时,它虽然可以通过专利获得暂时的垄断利润,但却无法捕获其创新所产生的全部社会收益。这项新知识会溢出到整个经济体,成为后续所有创新活动可以利用的公共知识基础,从而提升了未来所有企业的研发效率。由于私人创新者无法将这部分正的外部性收益内部化,市场自发形成的研发投入水平将低于社会最优水平。

最后是人工智能投资的熵减效应。这一效应的正负方向取决于当前市场的信息负荷水平z。当一个企业投资于人工智能资本K_{AI}时,其主要动机是提升自身处理数据、降低信息熵的能力,从而获得更高的创新效率和利润。然而,单个企业的人工智能投资在增加自身$k_{AI,i}$的同时,也增加了全社会的人工智能资本总量K_{AI}。根据信息熵函数,K_{AI}的增加会降低整个经济面临的信息负荷$z\equiv D_t/K_{AI,t}$,从而提升所有企业都从中受益的公共创新资源——有效数据$D_{eff,t}$的水平。由于单个企业无法捕获其人工智能投资通过降低全社会平均信息熵所造成的影响,导致其私人回报与社会回报之间存在差异。因此,市场自发形成的人工智能投资水平将系统性地偏离社会最优水平。在经济体处于信息过载区域(即$z>1$)时,更多的人工智能投资将带来正外部性,增加有效数据供给;在经济体处于数据稀缺区域(即$z<1$)时,更多的人工智能投资将带来负外部性,导致市场有效数据水平更低。

综上所述,本文模型对市场失灵的分析在继承经典增长理论的同时,提供了两个针对数字经济时代的核心拓展,通过对数据和人工智能的显性建模,识别出两种全新的、与之交织的市场失灵来源。其一,是源于隐私成本的数据生成负外部性,导致市场产生了过量的、未经提炼的原始数据;其二,是人工智能投资所具有的双重外部性,导致市场对这一资本的投

资难以达到最优水平。

(二)参数校准与数值模拟

数值模拟所需的参数可划分为四大类:家庭偏好参数、生产技术参数、知识创新参数、数据与人工智能参数。参数的选择遵循以下原则:对于宏观增长模型中的标准参数,如时际贴现率(ρ)、跨期替代弹性倒数(σ)、资本份额(α)、人口增长率(n)、知识溢出弹性(φ)和研发拥挤弹性(λ),参考 Jones 和 Tonetti(2020)与 Bloom 等(2020)等相关领域的经典文献。鉴于本文新引入的数据与隐私参数(χ, ψ, η, δ_D)尚无公认的宏观估算值,本文为其设定了合理的基准取值,并通过后续的比较静态分析检验核心结论的稳健性。表 1 汇总了数值模拟的核心参数及其经济学含义、基准取值与校准依据。

表 1　　模型参数设定

类型	符号	定义	取值	依据与说明
家庭偏好	ρ	时际贴现率	0.04	满足收敛性要求
	σ	相对风险厌恶系数	2.5	参考 Havranek(2015)
	χ	隐私成本强度参数	0.1	作者设定
	ψ	隐私成本的曲率	1	作者设定
生产技术	α	中间品在最终品生产中的份额	0.4	基于宏观经济指标设定
	n	劳动力(人口)增长率	0.02	参考 Jones 和 Tonetti(2020)
创新过程	δ	基础研发效率参数	0.1	作者设定,避免知识生产贡献过强
	φ	知识的跨期溢出弹性	0.5	参考 Bloom 等(2020)
	λ	研发劳动投入产出弹性	0.5	参考 Jones 和 Tonetti(2020),满足求解市场均衡所需的参数约束
数据与 AI	η	数据生成效率参数	0.1	作者设定
	δ_D	数据折旧率	0.2	参考徐翔和赵墨非(2020)、刘涛雄等(2023)
	γ	人工智能资本数据处理效率	1	形成平衡增长路径的要求

基于表 1 的参数设定与平衡增长路径条件,可推导出该经济体在稳态下的各项核心宏观指标。其中人均经济增长速度为 2%,总体经济增速为 4%。研发部门的劳动力占比为 5.41%,最终品生产部门的劳动力占比为 94.59%,生产性资本与总产出的比例为 0.16。在这一参数下,经济体必需的有效数据水平D^*_{eff}为 0.86。通过数值求解,我们得到两个z^*的取值:0.547和 1.63。经严格证明,$z^*=0.547$ 构成稳定均衡,而$z^*=1.63$ 为不稳定均衡,经济由此将滑入持续恶化的动态发散路径①。由此,市场均衡分化为两种典型的次优形态:人工智能投资过度的低水平稳态($z^*<1$),以及投资不足引发的低效发散路径($z^*>1$)。这表明,若无政策干预,市场机制无法自发收敛于社会最优的人工智能投资水平($z^*=1$)。

为进一步剖析深层结构参数对宏观均衡的影响机制,并兼顾数值结论的稳健性检验,对核心参数展开比较静态分析。

首先是知识生产的基础效率参数 δ。参数 δ 捕捉了创新过程本身的基础效率——即在给定研发劳动力和有效数据的情况下,将它们转化为新知识的效率。我们考察当 δ 从 0.1(基准值)逐步提高到 0.2 时,对经济稳态的影响,结果见表 2。知识生产基础效率的提高,对经济的资本结构产生了强大且与直觉相反的影响。由于长期人均增长率g_y由其他参数锁定,一个更高的 δ 意味着经济体可以用更少的有效数据投入(更低的D^*_{eff})就实现其必需的创

①证明过程请参见《经济评论》网站(http://jer.whu.edu.cn)附件。

新产出。为了达到这个更低的D^*_{eff}水平,经济体必须在一个离信息熵曲线效率峰值更远的点上运行。稳定的均衡要求这个点位于曲线的左侧,这意味着信息负荷z^*必须显著降低。表2结果显示,当δ从0.1翻倍至0.2时,均衡的z^*从0.547急剧下降至0.192。更低的信息负荷意味着,对于每一单位的数据,需要一个远大于以往的人工智能资本存量来进行处理。这是因为经济体的基础创新效率已经非常高,使其对数据质量变得极为敏感。这导致了"人工智能资本-产出"比的爆炸性增长,从0.735飙升至1.987。这项对人工智能资本的巨大投资是以牺牲消费为代价的,表现为"消费-产出"比从96.4%下降至91.4%。这一分析揭示了基础科研效率与数据处理能力之间强烈的替代关系:一个基础研发部门效率极高的经济体,可能悖论性地需要变得在人工智能资本的使用上更加密集,以维持其创新优势。

表2　　知识生产参数δ的比较静态分析

变量	数值($\delta\in[0.1,0.2]$)										
δ	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19	0.20
z^*	0.547	0.452	0.389	0.343	0.307	0.279	0.255	0.236	0.219	0.204	0.192
$\left(\frac{K_{AI}}{Y}\right)^*$	0.735	0.883	1.020	1.151	1.278	1.401	1.522	1.641	1.758	1.873	1.987
$\left(\frac{C}{Y}\right)^*$	0.964	0.958	0.953	0.948	0.943	0.938	0.933	0.928	0.923	0.919	0.914

其次是数据生成效率参数η。数据生成效率η的提高,意味着在相同的消费水平下,经济会产生更多的原始数据,这与大数据技术、物联网等重要技术的进步方向一致。由于维持长期增长所必需的有效数据水平D^*_{eff}和最优信息负荷z^*是由其他深层参数决定的(在此分析中保持不变),经济体为了避免陷入信息过载,必须积累更多的人工智能资本来处理增加的数据流。表3的结果清晰地验证了这一机制:随着η从0.1上升至0.2,稳态下的"人工智能资本-产出"比$\left(\frac{K_{AI}}{Y}\right)^*$从0.735几乎翻倍至1.427。这一结果揭示了一个重要的权衡关系:一个数据生成能力更强的经济体,也必须是一个在数据处理能力上投资更多的经济体。这种对人工智能资本的额外投资需求会挤出消费,表现为"消费-产出"比从96.4%下降至93.7%。

表3　　数据生成效率参数η的比较静态分析

变量	数值($\eta\in[0.1,0.2]$)										
η	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19	0.20
$\left(\frac{K_{AI}}{Y}\right)^*$	0.735	0.806	0.877	0.947	1.017	1.086	1.144	1.224	1.292	1.360	1.427
$\left(\frac{C}{Y}\right)^*$	0.964	0.961	0.959	0.955	0.953	0.950	0.947	0.945	0.942	0.939	0.937

最后是数据资本折旧率δ_D。对于数据资本的一系列研究以及各国统计局的核算工作为这一参数提供了多种估计结果,本节关注δ_D从0.1逐步提高到0.2(基准值)的经济效应。数据折旧率δ_D的提高,意味着数据的时效性更短,过时得更快。这导致在相同的消费水平下,经济体在稳态下维持的原始数据存量D更低。由于需要处理的数据总量减少了,管理信息熵所需的人工智能资本K_{AI}也相应减少。表4结果与此判断一致:随着δ_D从0.1上升至0.2,稳态下的"人工智能资本-产出"比从1.234平稳下降至0.735。这一发现具有重要的政策和产业启示:在一个数据折旧速度极快的行业(高δ_D),如即时零售或实时金融交易,对人工智

能资本的长期积累需求相对较低,因为旧数据的价值迅速消失;相反,在一个数据价值持久的行业(低δ_D),如基因组学或地质勘探,为了有效利用不断积累的历史数据,对人工智能资本的长期投资需求会大得多。因此,最优的人工智能投资策略与数据本身的生命周期特性密切相关。

表 4　　数据资本折旧率δ_D的比较静态分析

变量	数值($\delta_D \in [0.1,0.2]$)										
δ_D	0.10	0.11	0.12	0.13	0.14	0.15	0.16	0.17	0.18	0.19	0.20
$\left(\frac{K_{AI}}{Y}\right)^*$	1.234	1.155	1.086	1.025	0.970	0.921	0.877	0.836	0.800	0.766	0.735
$\left(\frac{C}{Y}\right)^*$	0.944	0.947	0.950	0.953	0.955	0.957	0.959	0.960	0.962	0.963	0.964

根据本文的数值模拟结果,市场形成的稳定均衡点为$z^*=0.547$,处于$z<1$的区间。然而在现实中,诸多证据表明,我们仍处于人工智能投资不足的区间,无论是从全球视角还是中国视角,对于原始数据的使用可能远未达到拐点(Babina et al., 2024; Acemoglu, 2025; 刘涛雄等,2024)。因此,当前最优政策并非对人工智能投资征税,而是通过适度的人工智能投资补贴抑制低效率的资本错配,同时将信息负荷推向$z=1$的最优值。

六、结论与启示

本文构建一个包含数据要素、人工智能与信息熵的内生增长模型,为理解数字经济时代的增长动力提供了新的理论框架。本文的主要创新在于,将信息过载这一关键的现实摩擦,通过信息熵函数内生化到宏观增长模型中。在该框架下,原始数据对创新的边际贡献呈现非线性的倒 U 型特征,人工智能的宏观角色被重新定义为一种有效降低信息熵、提炼有效数据的特殊资本。模型动态刻画了数据积累与人工智能投资的协同演化轨迹,揭示了经济体滑入“信息过载陷阱”或跃迁至高生产率路径的内生条件。此外,通过对比市场去中心化均衡与社会最优配置,本文精准识别出两类数字经济特有的新型市场失灵,即源于隐私成本的数据生成负外部性,以及人工智能投资状态依存的双重外部性。

基于上述发现,本文提出如下三条政策建议。

第一,实施数据与隐私的源头治理。为纠正数据生成的负外部性,应考虑对与大规模数据收集直接相关的消费行为或企业活动进行征税。这不仅有助于保护消费者隐私,更能从源头上控制低质量、高熵数据的泛滥,为后续的人工智能应用降低成本、提高效率。除征税以外,应严格落实数据最小化和目的限制原则,要求企业证明其数据收集的绝对必要性,而非默认全量收集。对于违规收集、超范围使用或长期囤积数据的行为,应建立阶梯式惩罚性税收或罚款机制,将数据泛滥的外部成本内生化。还应主导建立和推广行业性、区域性的公共数据质量标准与认证体系,引导市场从数据规模竞争转向数据精炼能力的比拼。

第二,构建以有效熵减为目标的新型研发补贴机制。传统研发补贴往往按人员或资金投入规模进行“一刀切”式激励,但在本文模型揭示的信息过载环境中,单纯扩张研发规模极易加剧拥挤效应,导致资源错配。因此,创新激励政策必须从补贴研发投入规模转向补贴熵减能力提升。政府的财政补贴与税收抵免应精准定向于能够产生显著正向溢出效应的基础人工智能项目,如致力于多模态数据清洗、开源高质量语料库建设以及研发基础性大规模预

训练模型的企业或科研机构。通过此类强针对性的激励措施,放大核心技术对全社会的知识溢出效应,从而在源头上化解数据驱动型创新中的内卷竞争与拥挤摩擦。

第三,建立动态、精准的人工智能产业政策体系。政府对人工智能的干预,应是一种高度状态依存的精细化治理。政策制定者需要对经济体和各个产业的信息负荷进行动态评估与诊断。为克服决策滞后与误判风险问题,政府可以通过一系列代理指标进行判断,如追踪数据密集型行业的全要素生产率变化、调研企业在人工智能应用上面临的主要瓶颈是数据不足还是算力或算法不足,以及分析数据中心投资与算力增长的相对速度等。当识别出行业或经济整体出现信息过载时,应果断实施投资补贴、税收抵免等政策;当评估发现人工智能投资存在过热倾向,导致资本与数据要素失衡时,则应考虑减少甚至取消补贴,或通过征税引导资本更有效地配置,确保人工智能的发展与实体经济的数据承载能力相匹配。

当然,本文的研究亦存在一些局限性。为保持模型的简洁与可解性,本文将人工智能的作用限定于降低信息熵上。这一设定虽抓住了数字经济当前发展中的一个核心摩擦,但也搁置了人工智能作为通用目的技术对劳动力供需的直接影响,以及人工智能对真实世界的熵减效果。此外,如何将本文的核心变量信息负荷转化为现实中可观测、可度量的宏观指标,以对模型推论展开严谨的计量检验,将是未来量化宏观研究的重要拓展方向。

参考文献:

1.陈彦斌、林晨、陈小亮,2019:《人工智能、老龄化与经济增长》,《经济研究》第 7 期。
2.李健斌、周浩,2025:《人工智能、资本-技能互补与企业全要素生产率》,《经济评论》第 1 期。
3.刘涛雄、戎珂、张亚迪,2023:《数据资本估算及对中国经济增长的贡献——基于数据价值链的视角》,《中国社会科学》第 10 期。
4.刘涛雄、张亚迪、戎珂、周迪,2024:《数据要素成为中国经济增长新动能的机制探析》,《经济研究》第 10 期。
5.徐翔、赵墨非,2020:《数据资本与经济增长路径》,《经济研究》第 10 期。
6.Acemoglu, D. 2025. "The Simple Macroeconomics of AI." *Economic Policy* 40(121): 13-58.
7.Acemoglu, D., A. Makhdoumi, A. Malekian, and A. Ozdaglar. 2022. "Too Much Data: Prices and Inefficiencies in Data Markets." *American Economic Journal: Microeconomics* 14(4): 218-256.
8.Acemoglu, D., and P. Restrepo. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review* 108(6): 1488-1542.
9.Aghion, P., B.F. Jones, and C.I. Jones. 2017. "Artificial Intelligence and Economic Growth." NBER Working Paper 23928.
10.Agrawal, A., J. Gans, and A. Goldfarb. 2019. *The Economics of Artificial Intelligence: An Agenda*. Chicago: University of Chicago Press.
11.Anderson, S. P., and A. De Palma. 2012. "Competition for Attention in the Information (Overload) Age." *The RAND Journal of Economics* 43(1): 1-25.
12.Babina, T., A. Fedyk, A. He, and J. Hodson. 2024. "Artificial Intelligence, Firm Growth, and Product Innovation." *Journal of Financial Economics* 151, 103745.
13.Bloom, N., C.I. Jones, J. Van Reenen, and M. Webb. 2020. "Are Ideas Getting Harder to Find?" *American Economic Review* 110(4): 1104-1144.
14.Brynjolfsson, E., D. Rock, and C. Syverson. 2017. "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics." NBER Working Paper 24001.
15.Cabrales, A., O. Gossner, and R. Serrano. 2013. "Entropy and the Value of Information for Investors." *American Economic Review* 103(1): 360-377.
16.Cong, L.W., D. Xie, and L. Zhang. 2021. "Knowledge Accumulation, Privacy, and Growth in a Data Economy." *Management Science* 67(10): 6480-6492.
17.Cong, L.W., W. Wei, D. Xie, and L. Zhang. 2022. "Endogenous Growth under Multiple Uses of Data." *Journal of Economic Dynamics and Control* 141(1), 104395.

18. Crafts, N. 2021. "Artificial Intelligence as a General-purpose Technology: An Historical Perspective." *Oxford Review of Economic Policy* 37(3): 521-536.
19. Dugast, J., and T. Foucault. 2025. "Equilibrium Data Mining and Data Abundance." *The Journal of Finance* 80(1): 211-258.
20. Eppler, M.J., and J. Mengis. 2004. "The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines." *The Information Society* 20(5): 325-344.
21. Farboodi, M., and L. Veldkamp. 2021. "A Growth Model of the Data Economy." NBER Working Paper 28427.
22. Havranek, T. 2015. "Measuring Intertemporal Substitution: The Importance of Method Choices and Selective Reporting." *Journal of the European Economic Association* 13(6): 1180-1204.
23. Jones, C.I., and C. Tonetti. 2020. "Nonrivalry and the Economics of Data." *American Economic Review* 110(9): 2819-2858.
24. Kurzweil, R. 2005. *The Singularity is Near*. London: Palgrave Macmillan.
25. Miller, J.G. 1960. "Information Input Overload and Psychopathology." *American Journal of Psychiatry* 116(8): 695-704.
26. Nussinov, R., M. Zhang, Y. Liu, and H. Jang. 2022. "AlphaFold, Artificial Intelligence (AI), and Allostery." *The Journal of Physical Chemistry B* 126(34):6372-6383.
27. Shannon, C.E. 1948. "A Mathematical Theory of Communication." *The Bell System Technical Journal* 27(3): 379-423.
28. Simmel, G. 1950. *The Sociology of Georg Simmel*. New York: Simon and Schuster.
29. Trammell, P., and A. Korinek. 2023. "Economic Growth under Transformative AI." NBER Working Paper 31815.
30. Veldkamp, L., and C. Chung. 2024. "Data and the Aggregate Economy." *Journal of Economic Literature* 62(2):458-484.

Entropy Reduction and Economic Growth: How Data Factors and Artificial Intelligence Reshape the Path of Economic Growth

Xu Xiang and Li Tao
(School of Economics, Central University of Finance and Economics)

Abstract: This paper constructs a dynamic growth model that endogenizes the phenomenon of information overload. It characterizes artificial intelligence as a special type of capital performing an "entropy reduction" function, revealing the underlying micro-mechanism through which high-entropy raw data is refined into low-entropy effective information to drive innovation. The findings are threefold. First, the contribution of raw data to innovation efficiency exhibits an inverted U-shaped pattern, wherein information overload significantly inhibits knowledge production. Second, in the dynamic interplay between data accumulation and AI investment, an economy may fall into a low-level growth trap due to AI underinvestment, or it may achieve a transition to a high-productivity path through adequate investment. Third, the paper identifies two novel market failures: a negative externality of data generation stemming from privacy costs, and a state-dependent dual externality associated with AI investment. Accordingly, this paper provides a solid theoretical basis for correcting novel market failures in the digital economy and steering macroeconomic growth toward the socially optimal path.

Keywords: Data Factors, Artificial Intelligence, Information Entropy, Endogenous Growth, Market Failure
JEL Classification: O33, O41, D62

(责任编辑:彭爽)