

宏观深度报告

报告日期：2026 年 05 月 26 日

如何看硅基通胀和碳基通缩？

核心观点

碳基和硅基二元分类法原本仅存在理论概念维度，随着计算机与人工智能技术迭代，逐步演变为研判经济新旧动能转换的重要视角。碳基生产力依托生物有机体、实体材料、油煤碳基能源与人类创意经验，是实体经济发展的核心根基；硅基生产力以硅基芯片、算力算法、数字软硬件为核心，具备规模化、高效率、可复制优势。

本轮生成式人工智能大规模落地后，下游通用式人工智能发展驱动上游算力、算法、芯片价格上涨，并带动中国 PPI 科技分项上行、进口增速中半导体拉动较大，美国人工智能发展也通过全球半导体价值链带动相关国家地区产业价格上涨。与此同时，中国大模型调用量超过美国，自主可控和进口替代的必要性进一步加强，硅基产业链景气具备更强支撑。

与之相对，碳基通缩部分源自就业替代：AI 提升硅基资本回报和企业生产效率，但对劳动任务的替代会压低部分劳动者边际议价能力，并通过劳动参与率下行、居民接受社会保障、消费降级与居民消费品价格下降，形成碳基通缩压力。

展望硅基与碳基的互动，2026 年硅基快速发展总体偏挤压碳基，科技企业盈利增速跑赢消费企业；但 2027 年碳基可能反噬，并引发 AI 泡沫破裂风险。中长期来看，碳基仍需有效利用硅基，重述社会分配和稳定机制，使硅基智能最终服务于人。

□ 碳基和硅基两分法的由来？

1891 年，德国天体物理学家儒略·申纳在讨论其他类型生命存在的可能时，将以碳元素为有机物质基础的生物称为“碳基生物”，相对地提出“硅基生命”的概念。碳基和硅基二元分类法在相当长一段时间内没有产业属性定义，仅存在理论概念维度，随着计算机与人工智能技术迭代，这两分法成为研判经济新旧动能转换的重要视角。

从产业维度来看，碳基生产力依托生物有机体、实体材料、油煤碳基能源与人类创意经验，覆盖生物医药、现代农业、传统能源、传统制造等实体产业，是实体经济发展的核心根基。硅基生产力以硅基芯片、算力算法、数字软硬件为核心，赋能 AI、数字基建、智能装备等前沿赛道，具备规模化、高效率、可复制优势，契合未来产业布局。

展望看，硅基智能或许正在复刻碳基生命的演化逻辑，从零散孤立的技术单元，逐步走向组网协同、分工细化的体系化发展，终极形态或将是多模块融合的硅基智能生态系统，持续重塑产业生产效率与经济增长格局。

□ 硅基产业受益于本轮人工智能的发展

硅基产业是指围绕硅元素开发、加工形成的产业集群，以半导体（高纯度单晶硅）为核心。产业链上游供应原材料，以工业硅为源头，进一步提纯得到单晶硅、多晶硅等关键原料；中游制造超纯硅材料、高性能陶瓷；下游广泛应用于电子与信息、新能源、新材料等领域。

生成式人工智能成为大规模落地技术后，催生了一种几乎完全集中在硅基经济领域的通胀脉冲，即生产或支撑 AI 算力的产业集群。AI 应用厂商竞相扩大推理算力规模，将需求冲击传递至芯片、高端存储、先进封装等上游环节，价格传导异常迅速且高度集中。

分析师：李超

执业证书号：S1230520030002

lichao1@stocke.com.cn

分析师：林成炜

执业证书号：S1230522080006

linchengwei@stocke.com.cn

分析师：费瑾

执业证书号：S1230524070007

feijin@stocke.com.cn

分析师：王瑞明

执业证书号：S1230525080002

wangruiming@stocke.com.cn

研究助理：汤子玉

tangziyu@stocke.com.cn

相关报告

- 1 《财政收支差额改善，为相机抉择留出空间》 2026.05.23
- 2 《经济走 K 型，资产举杠铃》 2026.05.19
- 3 《“NACHO”交易是什么？——兼论如何看待近期美债利率走高》 2026.05.17



人工智能领域应用提速、算力需求快速增长是 PPI 超预期的重要原因，算力周期与设备更新共振，驱动智算产业链价格将全面提速；未来半导体与 AI 硬件相关进口，或逐步成为中国进口体系中的“新支柱”。

美国作为 AI 大模型和算力需求的源头，通过全球半导体高度集中的供应链，将通胀压力和定价权向东亚主要科技制造经济体传导，形成区域性产业景气共振，中国大陆 AI 产业发展强劲。

□ 碳基通缩部分源自于就业替代

碳基通缩并非单纯来自总需求周期性下行，而是来自技术替代、收入分配和消费结构再平衡的共同作用。AI 提升硅基资本回报和企业生产效率，但其对劳动任务的替代会压低部分劳动者的边际议价能力。若技术进步带来的供给效率提升快于劳动收入改善，宏观经济便容易形成硅基通胀、碳基通缩的结构，即上游算力、芯片、能源、数据中心和科技资本品价格具备韧性，居民消费品、可选消费、传统服务和劳动密集部门价格弹性则趋于下降。

本轮 AI 技术的替代主力直奔知识密集型与行政办公的“白领”领域。企业在采纳人工智能后，其人力策略核心是停止招聘，继而导致初级岗位的直接消失。当职场入口被 AI 截流，部分低年资劳动力在长期的挫败中可能会选择放弃寻找工作，直接导致全社会劳动参与率的结构性承压。

本轮生成式 AI 的核心特征在于其精准替代或重构了高学历、高收入的“中层白领”。这一群体的隐性降薪与失业，正在成为消费降级的核心推手。在收入预期下降和通胀压力的双重挤压下，消费呈现显著的 K 型分化。为争夺有限的购买力，零售商不得不采取惠民定价和折扣策略，从而在实体经济中形成“碳基通缩”的负反馈闭环。

□ 硅基与碳基的互动发展过程推演

2026 年受 AI 数字基建投资拉动，硅基经济和碳基经济呈现明显 K 型分化，传统碳基指数被压缩，AI 相关产业链明显扩张；但主要国家硅基经济仍处在资本开支阶段，当前硅基经济繁荣的驱动力来自于下游 AI 企业融资推进数字基础设施建设，AI 下游尚未形成稳定盈利，企业利润更多用于再投资而非形成大规模应税收益，尚不能为财政带来增量，美债的压力并没有因 AI 发展而缓解。

2024 年以来“硅基”半导体指数显著跑赢碳基消费企业，“硅基”软件+硬件盈利显著跑赢碳基。当前 AI 科技革命、美股、美国经济深度绑定，负面风险在于任何一环被破坏都有形成负反馈的风险；**2027 年科技股 AI 泡沫可能破裂，关注技术变革、技术阻碍的影响。**

AI 是特朗普强化美元美债信用的重要一环，包括美联储在内的美国政策层有较强金融稳定诉求。若科技股泡沫破裂，可能导致美元美债信用恶化，继而引发主权债务危机。

□ 风险提示

人工智能技术迭代及应用落地不及预期风险；硅基产业链供需格局变化超预期风险；就业替代与碳基通缩演绎不及预期风险。

正文目录

1 碳基和硅基两分法的由来	5
2 硅基产业受益于本轮人工智能的发展	7
2.1 下游通用式人工智能发展驱动上游算力算法芯片价格上涨	9
2.2 中国 PPI 科技分项的上涨	9
2.3 进口增速中半导体拉动较大	10
2.4 美国人工智能发展带动相关国家地区产业的价格上涨	10
2.4.1 美国人工智能资本支出持续升高、材料价格上涨	11
2.4.2 日本：半导体设备与材料价格上涨	11
2.4.3 韩国：HBM 供给紧张推动存储价格上行	12
2.4.4 中国台湾地区：先进制程与先进封装	12
2.4.5 中国大陆：AI 产业链与国产替代同步推进	13
2.5 中国大模型调用量超过美国	14
2.5.1 中国大模型调用量快速增长	14
2.5.2 低成本推动 AI 普及和开辟开源模型市场是需求快速增长的主要驱动因素	14
2.6 中国自主可控和进口替代的必要性进一步加强	15
3 碳基通缩部分源自于就业替代	16
3.1 美国 AI 对就业的替代：快速“白领冲击”	16
3.2 劳动参与率下行：社会保障依赖度上升背景下的劳动力市场困境	17
3.3 消费降级与通缩传导：K 型分化后的负反馈循环	18
3.4 沃什截尾平均 PCE 框架，为中期通胀降温 and 降息打开空间	19
3.5 中国结构转型重于总量扩张，K 型分化强化居民消费品通缩压力	20
4 硅基与碳基的互动发展过程推演	22
4.1 2026 年硅基挤压碳基，但主要由 AI 资本支出带动，未缓解债务压力	22
4.2 24 年以来硅基盈利大幅跑赢碳基	23
4.3 2027 年碳基经济极致挤压可能带来泡沫破裂风险	25
4.4 关注技术路线的变革	26
4.4.1 上游涨价方向有可能被新技术路线替代	26
4.4.2 下游盈利兑现存在风险	26
4.5 关注技术路线的阻碍	27
4.5.1 技术失控风险	27
4.5.2 就业替代风险	28
4.6 中长期来看碳基有效利用硅基，重述社会分配和稳定机制	29
4.7 本轮 AI 泡沫破裂不能过度关注美联储	30
4.8 AI 泡沫破裂可能引发美元美债贬值出现主权债务危机	30
4.9 国际货币体系主导货币债券违约将引发百年一遇金融危机和经济危机	33
5 风险提示	34

图表目录

图 1: 第一台通用电子计算机 ENIAC.....	5
图 2: 碳基/硅基产业分类结构.....	6
图 3: 硅基产业链上游、中游与下游.....	9
图 4: 当前美国消费者指数已走弱至历史最低水平.....	19
图 5: 美国家具汽车等耐用品消费当前正持续负向拉动消费总额.....	19
图 6: 核心 PCE 一季度上行, 而截尾 PCE 维持较低读数.....	19
图 7: 25 年以来 PCE 和截尾 PCE 之差由负转正是核心商品贡献.....	19
图 8: 现行核心 PCE 指标泰勒公式政策利率 (%).....	20
图 9: 截尾 PCE 指标泰勒公式政策利率 (%).....	20
图 10: 百年未有至大变局下的“四层次”发展框架.....	20
图 11: 高技术制造业 PMI 与传统制造业 PMI 裂口扩大 (%).....	21
图 12: 高技术制造业增加值持续高于整体工业.....	21
图 13: 消费者信心指数与居民人均可支配收入增速仍处于低位.....	21
图 14: 2023 年至今居民人均消费支出服务类同比增速高于商品.....	21
图 15: 汽车及地产后周期相关品类对社零拉动逐步减弱.....	22
图 16: AI 相关投资 26 年 Q1 贡献非住宅投资增长的 71%.....	23
图 17: 2023 年以来半导体、信息行业经济度远超传统行业.....	23
图 18: A 股 AI 科技板块和消费板块 EBIT 同比增速.....	24
图 19: 标普 500 内 AI 科技板块和消费板块 EBIT 同比增速.....	24
图 20: 美股 AI 链市值占比走势.....	24
图 21: A 股 AI 链市值占比走势.....	24
图 22: 港股 AI 链市值占比走势.....	25
图 23: 新增就业趋势性放缓.....	26
图 24: 信息业就业加速下滑, 但份额提升.....	26
图 25: A 股 AIGC、AI 医疗、AI 教育、智能机器概念研发投入与盈利变化.....	27
图 26: 二战后两轮化债均与科技革命有关.....	33
表 1: 硅基产业核心产品与产业链供应链分布.....	8
表 2: 美国劳工局对选定人工智能相关职业的就业情况及 2023-2033 年预测 (千人).....	17
表 3: AI 风险汇总.....	28

1 碳基和硅基两分法的由来

1891年，德国天体物理学家儒略·申纳在讨论其他类型生命存在的可能时，将以碳元素为有机物质基础的生物称为“碳基生物”，相对地提出“硅基生命”的概念。自然界中碳元素具备独特的四价成键特性，可形成稳定、复杂、多样的长链有机分子，是构成地球所有生命体蛋白质、核酸、有机物的核心基础，故可以将依托碳元素构建有机体，在适应地球常温常压环境下完成新陈代谢、意识形成、遗传信息传递等生命活动的生命形态定义为“碳基生物”。同时，申纳结合硅元素与碳元素同族、化学性质相似的特征，创新性提出了“硅基生命”的假想概念，推测宇宙中或存在以硅元素为有机骨架的另类生命形态。在诞生后的100余年时间里，碳基、硅基的二元划分始终局限于天文物理、天体生物学的纯学术研究范畴，仅用于宇宙生命探索、地外文明形态推演。这两种概念始终停留在纯理论猜想层面，直至通用计算机体系逐步成熟和人工智能技术快速发展，才为硅基这一概念提供了新内涵。

1936年，英国数学家艾伦·图灵发表《论可计算数及其在判定问题中的应用》，首次提出图灵机抽象计算模型，从数学层面证明了所有可计算问题均可通过通用机器自动化求解，奠定了现代计算机的理论根基，首次从科学层面定义了“机器可模拟人类逻辑思维”的可能性。后续依托图灵机理论、布尔代数逻辑与冯·诺依曼体系架构，1946年第一台通用电子计算机ENIAC在美国问世，标志着人类正式迈入机器计算时代，打破了传统碳基计算、推理、决策的单一格局。

图1：第一台通用电子计算机 ENIAC



资料来源：宾夕法尼亚大学官网 www.seas.upenn.edu，浙商证券研究所

随着计算机硬件迭代、数字化技术普及，人工智能开启初步探索进程。1950年，图灵进一步提出“机器能否思考”的核心命题，推出经典图灵测试，正式开启机器智能的研究探索；1956年达特茅斯会议正式确立“人工智能”学科概念，标志着AI成为独立研究领域。早期人工智能以模拟人类逻辑推理、符号运算为核心，依托硅基硬件载体复刻碳基人类的基础思维与算力行为。这一系列技术革命让人类清晰看到全新的生产力形态：碳基人类依托生物躯体、有机禀赋、经验创意输出原创价值，而以硅基芯片、算力算法为核心的智能体系，可高效复刻、放大、替代人。本世纪以来，随着算力技术、大模型算法、硅基芯片产业的飞速发展，以硅材料为核心载体的人工智能体系，展现出区别于人类自然生命的全新生产力形态。产业界发现，传统碳基人类依托生物躯体、有机物质、经验创意创造

价值，而人工智能依托硅基芯片、数字算力、智能算法创造价值，二者契合经典的“碳基-硅基”二元逻辑。

从产业视角看，“碳基”核心指代以碳基物质、生物有机体、实体材料、人工工艺经验为核心的生产力要素，对应生物医药、现代农业、传统能源（石化、煤炭）等实体产业赛道。具体涵盖三大核心维度：一是生命与生物产业，包括合成生物学、基因技术、医药医疗制造等，依托碳基生物本体实现技术落地与成果转化；二是实体材料与高端制造工艺，涵盖各类石化、煤炭等碳基能源、传统化学化工品等，是制造业的核心底座；三是人类创意、工艺经验、复杂决策与场景定义能力。

“硅基”核心则是指代以硅基芯片、光伏新材料、大模型算法、算力集群、数字软硬件、具身智能等为核心的数字化、智能化生产力工具，是新质生产力的效率底座与赋能载体。硅基核心价值体现在规模化、可复制、高效率的智能赋能：以 GPU、ASIC、先进封装、HBM 存储等硅基硬件为算力底座，以大模型、工业软件、数字孪生、智能调度等算法软件为核心引擎，覆盖人工智能、数字基建、智能机器人、脑机接口、智慧电网、新能源等前沿领域。硅基产业的核心作用是替代重复性劳动、放大碳基创新价值、实现产业规模化升级、重构生产效率边界。全面匹配我国《“十五五”规划》前瞻布局的具身智能、大模型、量子科技、6G 通信、算力网络、智能装备等未来产业赛道。

图2：碳基/硅基产业分类结构



资料来源：浙商证券研究所

展望“硅基生命”的终点，从生命演化进程看，如果未来硅基智能真的可以取得长足发展，或许与地球“碳基生命”演化规律高度契合。地球现存复杂的多细胞“碳基生命”和以“碳基生命”为核心的生态系统，均起源于最原始的单细胞生物，经过漫长的自然演化，“碳基生命”逐步从零散独立的个体，迭代为分工明确、协同共生、体系完备的复杂生命系统。当前，以人工智能为核心的硅基智能体系，仍处在完全对应的初级演化阶段。未来以各类硅基基础单元或告别零散、孤立的发展状态，持续开展组网连接、功能分工与体系演化：以 GPU、ASIC、HBM、超算集群为核心的算力单元，承担核心智能运算与数据处理；以光伏、储能、智能供电系统为支撑的能源单元，提供持续稳定的动力供给；以算力网络、6G 通信、物联网、光纤链路为载体的通信链接单元，搭建全域高速传输通道；以工业软件、大模型算法、数字孪生为核心的智能大脑单元，负责逻辑决策与智能调度；以智能机器人、具身智能设备、各类智能终端为载体的执行落地单元。各类功能模块互补协同，逐步从单一硬件、孤立算法，演化成规模化、组织化、体系化的硅基智能生产力集群。市场普遍将人形机器人视作硅基的终极形态，但可能有些片面，人形机器人或仅是其中的一个重要组成部分，硅基智能的高级演化形态有可能并不是单一的智能个体设备，而是依托算力、网络、算法、智能装备等产业的深度融合形成的硅基生态系统。

2 硅基产业受益于本轮人工智能的发展

硅基产业是指围绕硅元素开发、加工形成的产业集群，以半导体（高纯度单晶硅）为核心。

产业链上游供应原材料，以工业硅（金属硅）为源头，进一步提纯得到单晶硅、多晶硅等关键原料。1工业硅由硅石（ SiO_2 ）和碳质还原剂在电炉中冶炼而成，是所有硅基新材料的“粮食”。其的纯度通常在98%以上，但还未达到电子或光伏级别，其生产能耗高，是典型的资源型产业。硅的开采属于B10非金属矿采选业，冶炼属于C32常用有色金属冶炼和压延加工业。硅基产业链的源头特指工业硅和冶金级硅的冶炼生产。

产业链中游制造各类新材料，如超纯硅材料（单晶硅、多晶硅、晶圆/硅片/衬底）、高性能陶瓷（碳化硅粉体、陶瓷制品）。

单晶硅是整个微电子产业的物理基石，它是在严格控制条件下，从一个籽晶生长而成的、原子排列高度有序的完整晶体。近乎完美的原子结构为电子提供了稳定、高效的移动通道，确保了器件性能的一致性，因此在半导体领域，所有用作衬底的硅晶圆都必须是单晶硅。全球超过95%的半导体器件和99%的集成电路都采用单晶硅作为衬底材料。

晶圆/硅片/衬底是芯片的物理载体，它是将单晶硅棒经过切割、研磨和精密抛光等一系列工序后，制成的厚度在0.2至0.75毫米之间、表面光滑如镜的圆形薄片。在芯片制造过程中，所有的光刻、刻蚀、沉积等几百道工艺，都将在这些薄薄的晶圆上进行。它也是整个半导体材料市场中占比最高的品类，其成本约占芯片总成本的30%-40%。在九大类晶圆制造材料中，硅片自身的销售额占比就高达30%，是晶圆制造中耗用最大的材料。

碳化硅粉体纯度要求极高，它本身并不直接成为芯片，而是作为制造碳化硅晶圆的源头，为高端半导体材料生产打下基础。粉体也被用于制备化学机械抛光液，对晶圆表面进行精密平坦化处理。同时，它还可以作为电子封装的导热填料，帮助高性能芯片散热。碳化硅陶瓷制品具有其极高的硬度、优异的导热性、耐高温和抗化学腐蚀等特性，是制造芯片所必需的高端生产设备中的核心耗材。

产业链下游广泛应用于电子与信息、新能源、新材料等领域。其中，电子与信息产业包括生产集成电路（芯片），应用于电脑CPU、手机处理器等；新能源产业，如制造光伏电池板，将太阳能转化为电能；新材料产业，如生产有机硅（如硅胶、硅油等），应用于建筑、汽车、医疗等领域；建筑与基础材料，如生产玻璃、水泥等，其基础成分是硅酸盐；其他高端领域，如航空航天、新能源汽车等。

总体而言，硅基产业是一个典型的横跨多个国民经济行业的产业生态。根据国家统计局发布的《国民经济行业分类》（GB/T 4754-2017），根据产业链的不同环节对应着特定的行业：上游硅的开采属于B10非金属矿采选业，冶炼属于C32常用有色金属冶炼和压延加工业；中游超纯硅材料属于C39计算机、通信和其他电子设备制造业，高性能陶瓷碳化硅的制造环节归入C26化学原料和化学制品制造业，陶瓷制品的成型与烧结环节则归入C30非金属矿物制品业；下游集成电路（芯片）属于C39计算机、通信和其他电子设备制造业。

1 我国是全球最大的工业硅生产国，产量占全球50%以上。

表1：硅基产业核心产品与产业链供应链分布

产业链环节	核心产品及说明	国标行业分类	细分行业名称及代码	应用场景
上游（原材料供应）	工业硅（又名金属硅） ，它由硅石（SiO ₂ ）和碳质还原剂在电炉中冶炼而成，是所有硅基新材料的“粮食”。其的纯度通常在98%以上，但还未达到电子或光伏级别，其生产能耗高，是典型的资源型产业。	硅的开采属于 B10 非金属矿采选业 ，冶炼属于 C32 有色金属冶炼和压延加工业 。硅基产业链的源头，特指工业硅和冶金级硅的冶炼生产。	B1019 粘土及其他土砂石开采 C3218 硅冶炼	为中游新材料、下游各类产品提供基础原料
中游（新材料制造）	超纯硅材料：单晶硅、多晶硅、晶圆 （硅片/衬底）。半导体和光伏产业的核心，技术和资本壁垒极高，产品形态包括太阳能级和电子级多晶硅、单晶硅棒/片及外延片等。	属于 C39 计算机、通信和其他电子设备制造业 大类，如 C398 电子元件及电子专用材料制造中类。	C3985 电子专用材料制造	支撑半导体和光伏产业下游产品生产
	高性能陶瓷 ：碳化硅（SiC）粉体、陶瓷制品。具有高熔点、超高硬度（莫氏硬度9.5）与耐磨性。半导体与高纯级陶瓷通常要求高纯（>99.9%）乃至超高纯，关键杂质控制在极低水平。	碳化硅的制造环节归入 C26 化学原料和化学制品制造业 ，其陶瓷制品的成型与烧结环节则归入 C30 非金属矿物制品业 。	C2613 无机盐制造（粉体） C3073 特种陶瓷制品制造（制品）	支撑第三代半导体与光电、新能源汽车、航空航天与国防、核能与新能源、工业与化工
	新型锂电池负极材料 ：硅碳、硅氧。新型锂电池负极材料，通常由纳米硅与碳材料复合而成，是新能源产业的关键材料。	属于 C30 非金属矿物制品业，C39 计算机、通信和其他电子设备制造业。	C3091 石墨及碳素制品制造 C3985 电子专用材料制造	新型锂电池负极材料
	化工新材料 ：有机硅单体（D4/DMC）、硅烷偶联剂、硅橡胶、硅树脂等。分别用于工业大宗通用原料，“分子桥”、弹性体、高性能涂层骨架。	属于 C26 化学原料和化学制品制造业的重要分支，有机硅产品线极长，从上游单体到下游聚合物，涵盖多个行业小类。	C2614 有机化学原料制造（单体） C2651 初级形态塑料及合成树脂制造（聚合物） C2652 合成橡胶制造（硅橡胶） 2661 化学试剂和助剂制造（如硅烷偶联剂）	以用于建筑/汽车/电子/新能源/医疗领域
下游（应用领域）	集成电路（芯片）	C39 计算机、通信和其他电子设备制造业	C2659 其他合成材料制造	电脑 CPU、手机处理器等
	光伏电池板	C38 电气机械和器材制造业	C3862 太阳能器具制造	-
	有机硅（硅胶、硅油等）	C26 化学原料和化学制品制造业	有机硅（硅胶、硅油等）	汽车、医疗器械、建筑材料等
	各类硅基配套器件/材料	同上述三行		航空航天、新能源汽车等

资料来源：《国民经济行业分类》（GB/T 4754-2017），浙商证券研究所

图3：硅基产业链上游、中游与下游



资料来源：浙商证券研究所

2.1 下游通用式人工智能发展驱动上游算力算法芯片价格上涨

生成式人工智能（GenAI）成为大规模落地技术后，催生了一种几乎完全集中在硅基经济领域的通胀脉冲，即生产或支撑 AI 算力的产业集群。AI 应用厂商（下游）竞相扩大推理算力规模，将需求冲击传递至芯片、高端存储、先进封装等上游环节。由于半导体供应链“卡脖子”环节（先进制程晶圆代工、高带宽内存 HBM、先进封装）由少数企业掌控、扩产周期长达数年，价格传导异常迅速且高度集中。

一是推理芯片成为新增长极。英伟达（NVIDIA）2025 年一季度财报显示，数据中心业务收入同比增长 427%，达 226 亿美元；2025 年三季度进一步增至 512 亿美元，同比增速维持 66% 高位；2025 年全年营收 1305 亿美元，较 2024 财年 609 亿美元实现翻倍。有研究机构预测，2026 年全球生成式 AI 芯片整体收入逼近 5000 亿美元，约占全球半导体总销售额的一半。AMD 预计，到 2030 年全球数据中心 AI 加速芯片潜在市场规模将达 1 万亿美元。

二是内存市场通胀机制尤为典型。AI 算力对高带宽内存（HBM）、DRAM 需求激增，产能从消费端向企业端倾斜。2026 年初起，特斯拉、苹果等十余家巨头披露全球 DRAM 短缺制约生产、挤压利润；苹果 CEO 库克预内存短缺将拖累 iPhone 利润率。2026 年一季度 DRAM、NANDFlash 均价环比分别暴涨 60%、70%。

三是算法端定价权向稀缺环节集中。2026 年 5 月，CME 集团与 SiliconData 推出算力期货市场——基于 GPU 价格指数的衍生品，帮助 AI 企业对冲租赁成本。GPU 算力已成为可交易金融标的，这在 AI 稀缺前难以想象。

2.2 中国 PPI 科技分项的上涨

2026 年 4 月 PPI 同比增速录得 2.8%（前值 0.5%），显著高于市场预期，为近 45 个月以来的最大涨幅。PPI 环比 1.7%（前值 1.0%），连续 7 个月上涨，涨幅比上月扩大 0.7 个百分点，为 54 个月以来最大涨幅。石油产业链上游价格传导外，人工智能领域应用提速，算力需求快速增长是 PPI 超预期的重要原因。

算力产业链方面，4 月份有色金属矿采选业同比上涨 38.9%，有色金属冶炼和压延加工业同比上涨 22.5%，电气机械和器材制造业同比上涨 3.6%，计算机通信和其他电子设备制造业同比上涨 1.5%，合计影响 PPI 同比上涨共约 2.04 个百分点。受“人工智能+”加快拓展、算力需求快速增长影响，3、4 月份光纤制造价格分别环比上涨 76.1%、22.5%。

我们认为，算力周期与设备更新共振，驱动智算产业链价格将全面提速。2026 年国家正式确立“超大规模智算集群”与“算力电力协同”机制，标志着推理需求正式反超训练需求，触发行业“智算节点”替换潮。在大规模设备更新政策的强力引导下，高附加值的智算服务器对传统通用服务器的结构性替代，将对计算机、通信和其他电子设备制造业的价格中枢建立上行支撑。

我们预计，计算机、通信和其他电子设备制造业有望对 PPI 逐月贡献 0.2-0.25% 的新涨价因素。同时，算力基础设施的爆发式增长将传导至上游有色金属行业，形成“算力+电力”双轮驱动的产业链繁荣，成为下半年 PPI 内部 K 形分化中的重要上行牵引力。

2.3 进口增速中半导体拉动较大

2026 年 1-4 月我国进口同比增长 20%，持续保持强劲增长。在经济总需求仍偏弱运行的背景下，AI 链是我国进口的重要支撑。2026 年 4 月，集成电路、自动数据处理设备分别拉动我国进口增速的 7.8%、3.3%。集成电路进出口双增体现 AI 链条下“大进大出”的特征：

一方面我国在全球半导体产业体系中处于后端环节，封测、集成组装占比较高，是进出口高增的产业逻辑。我国半导体产业链优势集中于后端封测环节，形成了“大进大出”加工贸易格局：从韩国、中国台湾地区进口 AI 芯片、高端存储、光模块等核心元器件，依托后端优势完成封装集成与整机组装后出口。2025 年我国集成电路进口额达 4243.3 亿美元，同比增长 10.1%；出口额达 2019 亿美元，同比增长 26.8%，进出口双向均实现强劲增长。

另一方面国内 AI 算力中心建设加速推进。“十五五”规划第十二章明确提出要加强算力基础设施支撑，“加强高性能高质量智算资源供给，论证建设超大规模智算集群”。根据新华社报道，2026 年中国 AI 核心产业规模预计突破 1.2 万亿元人民币，较 25 年 9000 亿水平同比增长率接近 30%。

我们认为，未来半导体与 AI 硬件相关进口，或逐步成为中国进口体系中的“新支柱”。国内 AI 算力中心建设正在加速推进，“十五五”规划明确提出加强超大规模智算集群建设，意味着 AI 主线对半导体进口的趋势性拉动仍将持续。

2.4 美国人工智能发展带动相关国家地区产业的价格上涨

人工智能的发展和应用大致可以分为以下五个主要阶段：

第一阶段基础理论突破期（2006-2011）：深度学习的奠基与算力瓶颈的打破。标志性事件：2006 年 Geoffrey Hinton 发表在 Science 上的文章《通过神经网络实现数据降维》(Reducing the Dimensionality of Data with Neural Networks)。

第二阶段从理论可行变为工程可用（2012-2015）：感知智能的工程化与商业落地。标志性事件：2012 年 AlexNet 在 ImageNet 竞赛中夺冠。

第三阶段风险投资涌入（2016-2019）：认知智能的突破与 AI 技术的扩散，全球市场持续扩张。标志性事件：2016 年 AlphaGo 击败李世石。

第四阶段算法算力与市场的同步跃迁（2020-2024）：大模型的集中爆发与生成式 AI 的产业化。标志性事件：2022 年 11 月 ChatGPT 发布。

第五阶段推理能力的跃迁与 AI 效率革命（2025 至今）：推理能力的跃迁与 AI 效率革命。标志性事件：2025 年 DeepSeek 引发算法效率变革。

总体上讲，美国是本轮 AI 投资周期的需求核心，在基础研究和技术应用方面处于领先地位。微软、谷歌、亚马逊、Meta、OpenAI、xAI 等云巨头与前沿模型厂商，年投入数千亿美元建设 AI 基础设施。并且资金并未局限于美国，而是沿全球半导体价值链传导，有些经济体受益、定价权提升，其中韩国绝对收益较大，中国台湾地区成为全球逻辑代工枢纽，日本加速半导体产业复兴，中国大陆 AI 产业发展强劲。

2.4.1 美国人工智能资本支出持续升高、材料价格上涨

2023 年以来，以生成式人工智能为代表的新一轮技术周期推动全球科技资本支出持续扩张，美国大型科技企业进入以 AI 基础设施为核心的投资阶段。谷歌、亚马逊、微软和 Meta 四大科技企业计划在 2026 年将 AI 领域的资本开支提升至 7250 亿美元，较 2025 年的 4100 亿美元增长 77%。与过去的互联网周期相比，本轮 AI 投资呈现更强的资本密集型特征，其核心投入不再局限于软件研发，而是进一步延伸至 GPU、HBM、高速互联、先进封装以及数据中心等硬件基础设施领域。

人工智能正通过资本开支成本、算力资源挤占等渠道推高生产者价。2026 年 4 月，美国最终需求 PPI 同比上涨 6.0%（创 2022 年 12 月以来新高），环比上涨 1.4%（创 2022 年 3 月以来最大单月涨幅）；核心 PPI（剔除食品、能源和贸易服务）同比上涨 4.4%，环比上涨 0.6%。其中，计算机和电子产品制造同比增长 12.3%，计算机存储设备制造同比增长 19.2%，半导体及其他电子元件制造同比增长 26%，多项 AI 产业链价格创有史以来新高。

美国作为 AI 大模型和算力需求的源头，通过全球半导体高度集中的供应链，将通胀压力和定价权向东亚主要科技制造经济体传导，形成区域性产业景气共振，日本、韩国、中国台湾地区以及中国大陆均在本轮 AI 产业扩张中受到不同程度的影响。与传统需求驱动型通胀不同，AI 带来的价格上涨更多集中于半导体、数据中心、电力设备、先进制造以及算力基础设施等硅基产业链。

2.4.2 日本：半导体设备与材料价格上涨

受 AI 发展的驱动，高性能计算和数据中心内存芯片出货量激增，全球晶圆厂投资和设备采购持续景气。根据美国半导体行业协会（SIA）发布最新研报称，2025 年全球半导体销售额将达到 7917 亿美元，较 2024 年的 6305 亿美元增长 25.6%。WSTS 预期 2026 年全球半导体营收有望同比增长 26.3%，达到 9750 亿美元，存储和逻辑 IC 仍是主要成长动能，两者增长率都超过三成，分别增长 39.4% 及 32.1%。

日本是全球半导体制造设备与核心材料的重要供应方，处于硅基产业链的中游核心位置。在 AI 发展的第二阶段，日本企业就开始受益于半导体材料行业的复苏。根据贸促会的报告，日东京电子在晶圆涂布光刻胶（感光材料）、呈现电子电路“涂布显影设备”领域占全球份额的 90%；信越化学工业和 SUMCO 合计掌握约全球 60% 的市场份额的硅片供应；在光刻胶市场占有 90% 的份额。

2026 年以来，多家半导体产业大厂宣布产品涨价。日本化学材料大厂信越化学针对其主力产品有机硅实施国内外全面性的价格调涨，全线产品的调幅高达 10% 以上。日本半导体与电子材料巨头 Resonac 控股公司对旗下铜箔基板、黏合胶片等印刷电路板核心上游材料进行全面调价，整体涨幅高达 30%。Resonac 的客户涵盖村田制作所、TDK、三星电机、台积电封测厂等国际头部电子企业，此次调价将传导至 MLCC、HDI 板、IC 载板、高频高速 PCB 等高端制造环节。

两家公司均表示价格上涨源于原材料价格上涨、中东局势导致的原油及轻油价格大幅高涨、以及劳动力成本和运输费用显著增加，使产品总成本增加。在政策方面，日本在 2023 年 3 月 31 日就已宣布对包括光刻机在内的多种半导体设备加强出口限制以控制高端芯片制造能力和关键材料供应。2026 年初，日本进一步发布公告并对部分半导体材料发起反倾销调查。半导体材料多为高纯度、高标准产品，生产线建设周期长达 2-3 年，且技术壁垒高、认证严格，短期内难以快速扩产。高端材料领域，信越、JSR、东京应化等国际厂商长期垄断，供给高度集中导致议价能力倾斜。此轮由 AI 需求驱动、供给刚性与地缘风险推动的涨价，将在未来 2-3 年持续影响全球电子产业。

日本 2026 年 4 月生产者物价指数（PPI）同比上涨 4.9%，为 12 年来最高涨幅，其中电子元件及设备、半导体器件、集成电路等 PPI 分项分别上涨 2.42%、4.25%、1.8%。本轮日本 PPI 分项没有跟风明显上涨，主要原因是全球涨价的集成电路主要是 HBM、DDR5、AI GPU 等高端逻辑/存储芯片，而日本厂商优势集中在 MCU（单片机）、功率半导体、模拟芯片，这些领域竞争激烈、替代性强，难以跟风涨价。

2.4.3 韩国：HBM 供给紧张推动存储价格上行

长期以来，韩国半导体产业深受传统大宗存储芯片通缩周期的结构性困扰。然而，AI 时代改变了存储产业的定价逻辑。2022 年，ChatGPT 引爆大模型浪潮后，AI 训练对算力与内存带宽的需求骤然激增，高端 GPU 对高带宽内存（HBM）的需求快速增长。这个原本服务于高性能计算、应用于高端显卡的技术，突然成为了 AI 芯片的核心刚需，导致 HBM 供给持续紧张，成为当前 AI 服务器产业链中的关键瓶颈之一。

SK 海力士与 AMD 联合开发了全球首款硅通孔 HBM，至今已经迭代升级了 4 代 HBM 产品，性能和容量持续提升。目前，全球 HBM 被三星、SK 海力士与美光科技垄断，产能严重不足，DRAM 价格正不断飙升。据市场调研机构 DRAM Exchange 统计，截至今年 2 月底，DDR4 8Gb 产品的单价为 13 美元，较去年同期的 1.35 美元上涨了 9.6 倍。从第二季度开始，三星电子处于最领先地位的 HBM4 产品将正式开始出货。预计第一季度达到 3 万亿韩元的 HBM 销售额将逐步增加，到下半年时，季度销售额有望接近 10 万亿韩元。随着 HBM4 供应的增加，三星电子抢占价格更高的最新工艺——第六代（1c）DRAM 市场的可能性也随之增大。三星电子的 HBM4 采用了 1c DRAM 进行堆叠，这与使用第五代（1b）DRAM 的竞争对手形成对比。

韩国半导体产业的数据表明硅基通胀具备强大的去周期属性。即使在全球传统消费电子面临低迷和局部通缩风险时，AI 所急需的硅基硬件升级依然能够推动整个国家级优势产业实现逆周期的价格拉升。HBM 将存储芯片从标准化大宗商品转为高门槛、定制化的短缺资源。韩国厂商在此领域构建了极高的垄断壁垒，从而在这一轮硅基通胀中获得了极强的定价自主权。全球 HBM 产能仍然严重不足，无法满足 AI 服务器需求，未来价格走向或取决于国产 HBM 能否实现技术突破，成为破局关键。

2.4.4 中国台湾地区：先进制程与先进封装

台湾地区的核心产品包括先进制程晶圆代工（3nm/4nm）、CoWoS 先进封装，处于硅基产业链的中游制造与封装核心环节，是连接美国设计与全球应用的关键枢纽。在 AI 芯片产业链中，台湾地区（尤其是台积电）几乎垄断了全球尖端 AI 芯片的代工与先进封装产能。当前英伟达高端 AI GPU 高度依赖台积电的先进制程与封装产能，推动中国台湾地区半导体产业景气度明显提升。随着 AI 服务器需求快速增长，先进封装成为新的产能瓶颈，价格与订单持续增长。台积电 2026 年一季度营运报告显示，7nm 及以下的先进制程占总收入的 74%，并且 76% 的收入来自北美的客户，包括英伟达、AMD、苹果等高科技企业。

由于高端 GPU 制造门槛较高，短期内全球供给难以快速扩张，从而导致相关产业链价格持续上涨。台积电(TSM.US)已开始通知苹果(AAPL.US)等主要客户，其先进制程芯片将实施价格上调，此次涨价包含 2nm、3nm、5nm 制程产品，这些工艺广泛应用于智能手机、GPU 等高端电子设备，涨价幅度最高 10%。代工与先进封装具备强技术垄断，短期内全球无替代方案，虽然台积电等厂商的产能预计在 2026 年逐步释放，供需缺口有所收窄，但随着 2nm 制程的量产落地，晶圆的物理制造难度和成本再次飙升，预计台湾地区的尖端代工价格难以下跌，将维持在高位。

2.4.5 中国大陆：AI 产业链与国产替代同步推进

芯片的封装与测试是中国大陆硅基产业的核心支柱，是产业链中营收规模最大、最稳定的环节之一。委外封测是芯片从“原料”走向“产品”的最后两道关卡，也是保证芯片质量的关键。芯思想研究院（ChipInsights）数据显示，2025 年全球委外封测营收达 3,332 亿元人民币，创历史新高。根据总部所在地划分，前十大委外封测公司中，中国内地有五家（长电科技、通富微电、华天科技、智路封测、盛合晶微），市占率为 32.6%，略低于中国台湾的 33.42%。其中长电科技继续位列全球第三位，中国大陆第一。

由于 AI 半导体需求极为强劲，长电科技发布的业绩暨现金分红说明会记录公告内容称，“价格方面，从去年开始在原材料价格上涨趋势下，联动机制已逐步成熟，得到越来越多客户理解和积极响应。在产能供不应求、产能相对紧缺情况下，产线资源宝贵，公司不断优选客户、优化产品结构，提升单位价格”，但并未公布涨价幅度。可供参考的是，中国台湾的封测龙头日月光因产能不足、订单量激增的调价涨幅预期落在 5%至 20%之间。

与其他国家地区不同的是，中国正在逐步建立自主的完整硅基产业链，因此涨价热潮并不局限与人工智能直接相关的产业。头部集成电路代工厂商中芯国际在 2026 年一季度业绩说明会上表示，对于供不应求的产品类别，公司已与客户协商上调代工定价。中芯国际不制造最顶尖的 AI 算力芯片，但 AI 产业的爆发带动了 AI 服务器、数据中心以及端侧 AI 设备所使用的成熟制程配套芯片，使得中芯国际在这些细分领域的产能出现了明显的供不应求，其涨价效应是通过“产能挤压”和“需求虹吸”引发的连锁反应。全球顶级的晶圆代工厂（如台积电、三星）为了满足 AI 主芯片的高利润需求，纷纷将产能资源向先进制程倾斜，主动收缩了成熟制程的产能。原本在海外代工的常规芯片（如手机、电脑、物联网芯片）拿不到足够的产能，只能转向中国大陆寻找产能。中芯国际作为国内晶圆代工龙头，承接了大量回流的订单。手握充足的订单储备，也让公司获得了更大的定价权。

美国对高端 AI 芯片出口限制持续升级，使中国大陆进一步强化自主可控与国产替代的发展。当前，中国在 AI 大模型、国产 GPU、AI 服务器以及先进制造领域持续加大投入，推动本土产业链快速发展。因此中国大陆的硅基产业发展逻辑并非完全跟随美国的 AI 周期，而是叠加了自主可控的独立周期，在 2019 年（中美科技摩擦加剧）后进入加速期，促使中国企业加快本地化部署与国产算力适配。近期随着国产大模型调用量超越美国，对国产算力的需求倒逼上游材料国产化，带动了国产硅片、光刻胶等产品的价格重估。华为、寒武纪、

中芯国际等企业在国产 AI 生态中的重要性明显提升，进一步强化了国内人工智能、半导体等相关产业景气度。

2.5 中国大模型调用量超过美国

2.5.1 中国大模型调用量快速增长

根据全球最大的 AI 模型 API 聚合平台 OpenRouter 披露的数据，2026 年 2 月，中国大模型的 Token 调用量激增，首次超过了美国模型。这一增长并非偶然，5 月 18 日至 5 月 24 日，全球 AI 大模型总调用量为 28.9 万亿 Token，较此前一周增长 7.4%，已连续五周上涨。其中，中国 AI 大模型周调用量达 9.22 万亿 Token，环比增长 19.89%；同期美国 AI 大模型周调用量为 4.93 万亿 Token，环比增长 16.27%。中国大模型周调用量连续四周超过美国并稳居全球首位。截至目前，DeepSeek-V4-Flash 已登顶 OpenRouter 全球 AI 大模型调用榜，以 3.43 万亿 Token 位居第一。

中国模型调用量增长不只是源于自国内用户需求，也来自海外开发者与企业使用。根据该平台于 2025 年 12 月发布的统计数据，其用户主要由海外开发者构成，其中美国用户使用的 Token 数量占比高达 47.17%，而中国用户仅占 6.01%。因此该榜单数据能够客观反映中国 AI 模型在全球范围内的真实吸引力。这一现象表明，中国 AI 产业已经从追赶模型能力、扩大国内用户群体的阶段，逐步进入向海外扩散与输出的阶段。不再局限于本土市场，而开始参与全球 AI 生态竞争。

2.5.2 低成本推动 AI 普及和开辟开源模型市场是需求快速增长的主要驱动因素

一、低成本推动 AI 普及

与美国大型闭源模型相比，中国 AI 企业更强调低成本部署与开源生态，在推理成本与部署效率方面具备一定优势，降低了企业与开发者使用门槛，从而推动了国产模型的普及与 Token 调用量的快速增长。

OpenRouter 平台 5 月 18 日至 5 月 24 日期间调用量前十的 AI 模型及其每百万 Token 价格，国产模型中价格最高的是 Kimi K2.6，输出价格为 3.49 美元每百万 Token，其他国产模型的输出价格则均在 1 美元每百万 Token 以下。然而从宏观层面来看，AI 模型的需求弹性较低，价格变化对使用量影响较小，如调用量前二的 DeepSeek V4 Flash 和 Hy3 preview 输出价格均在 0.2 美元每百万 Token 左右，而位于第三和第四的 Claude Opus 4.7 与 Claude Sonnet 4.6 的输出价格分别高达 25 和 15 美元每百万 Token。这表明 AI 模型的客户群体差异较大，专业的企业用户愿为高难度、高精度的任务支付更昂贵的成本，使在用率高的同时，指令难度也导致 Token 调用量大；而个人或业余用户则对成本更加敏感，倾向于选择更便宜的模型，导致了高效模型的大量使用。在被广泛集成到更多任务后，由于人们运行更长的上下文，更多迭代等，总 Token 消耗量同样很高。

从 AI 企业的角度来看，会出现两种战略模式。美国头部 AI 企业普遍采用较为明显的模型分层策略，根据模型性能、推理能力、响应速度以及上下文长度实施差异化定价。例如，Google 推出 Gemini Flash 与 Gemini Pro 等不同型号产品，分别面向高吞吐低成本场景与高复杂度推理场景，轻量级任务被分配到更便宜、更快速的模型；高级型号则支持复杂或耐延迟的工作负载。本质上，此类模式与传统云计算服务类似，通过高端模型维持利润率，并利用轻量化模型覆盖高频应用场景，其核心目标在于提升单用户收入（ARPU）与商业化效率。

相比之下，中国 AI 企业更强调降低 Token 调用成本与扩大开发者生态。以 DeepSeek、Qwen、Kimi 等国产模型为代表，中国 AI 厂商普遍采用更低价格策略，并通过免费开放、低价 API 等方式推动 AI 应用快速普及。部分国内模型 API 价格在过去一年中出现大幅下调，

部分企业甚至开放免费商用权限。其核心目标并非短期利润最大化，而是通过降低使用门槛扩大用户规模与模型调用量，注重“低成本、高调用量、快速普及”的发展路径。

二、开辟开源模型市场

海外 AI 模型，如 GPT、Gemini、Claude 等，当前更加倾向于构建闭源商业生态，通过 API 授权、企业订阅以及高端模型服务强化定价权，并逐步形成以技术壁垒为核心的商业模式。然而自 2025 年以来，中国开源大模型正在成为全球 AI 需求增长的重要驱动力，向用户公开源代码、模型训练等内容，重视开源生态扩张，通过开放模型权重、降低调用成本以及加快版本迭代推动开发者与企业快速接入 AI 能力。随着 DeepSeek V3、Kimi K2、Qwen 等中国开源模型陆续发布，中国模型使用量快速提升，开源模型整体使用占比到 2025 年底已接近三分之一。其中，中国开源模型的增长尤为明显，其周度 Token 使用份额从 2024 年底不足 1.2%，逐步提升至部分时期接近 30%，全年平均占周度 Token 成交量约 13%。

中国开源模型需求增长的核心原因，在于其快速迭代能力、较低使用成本以及对技术类场景的适配能力不断增强。与传统闭源模型不同，以 Qwen、DeepSeek 为代表的中国模型持续保持高频版本更新，能够更快适应代码生成、Agent、基础设施运维以及企业自动化等新兴工作负载，从而推动开发者与企业侧调用量快速增长。

从使用结构来看，中国开源模型的需求已经逐渐从娱乐与创意场景，转向更具生产属性的技术场景。OpenRouter 数据显示，编程与技术相关任务的总和占中国开源模型使用量的 39%，超过传统创意类需求（33%），表示中国 AI 模型正在逐步进入代码开发、数据处理、云计算以及企业生产流程等核心应用环节，逐渐缩小与国外高端闭源模型在技术上的差距。逐步转向全球 AI 生态的重要供给方。AI 需求开始由消费型需求向生产型需求扩张。随着开源模型持续推动开发者与企业 AI 化，中国 AI 需求增长正从单纯用户规模扩张，转向以 Token 消耗和推理需求为核心的新阶段。

从宏观层面来看，海外 AI 产业模式更接近于高利润软件服务业，其核心在于维持模型能力领先与商业化利润率；而中国 AI 产业模式则更接近于基础设施普及型扩张，其重点在于通过低成本与高渗透率推动 AI 大规模应用。后者虽然短期盈利能力相对有限，但更容易形成高频 Token 消耗、高推理需求以及大规模云端调用。

三、AI 普及驱动下的 K 型分化

在传统的碳基经济中，价格下跌往往伴随着产能过剩与需求萎缩。在硅基经济中，单位 Token 调用成本的下降，反而刺激了总调用量上升。通过降低 AI 使用门槛，推动 AI 从少数科技企业内部工具，逐步扩展为全社会范围的基础性生产工具。Token 的消耗即算力的消耗，代表推理需求与算力需求的指数级增长，从而带动 GPU、数据中心、电力以及云计算资源需求持续扩张。

但 AI 技术扩散并未消除资源约束，而是进一步强化了经济中的结构性分化。一方面，拥有 GPU、先进制程、云计算与电力资源的硅基部门持续受益于 AI 需求扩张，价格与资本开支维持高景气；另一方面，传统制造业、劳动密集型行业以及需求偏弱的碳基部门仍面临需求不足、价格下行与盈利压力等问题。

2.6 中国自主可控和进口替代的必要性进一步加强

中国 AI 半导体自主化早于本轮 AI 潮，源于 2018-2022 年美国出口管制升级。2025 年初 DeepSeek 技术突破及后续国内模型爆发，显著加速自主化紧迫性与可行性，核心逻辑为能力建设、成本下降、供应链安全三重叠加。

硬件以华为昇腾成为国产算力核心。2026 年 4 月，DeepSeek 发布专为华为昇腾芯片优化的 V4 模型，系中国前沿大模型首次基于国产芯片联合开发。华为确认全系列昇腾超节点支持 DeepSeek-V4，V4-Flash 模型训练亦采用昇腾芯片。业内认为，华为昇腾是英伟达最佳国产替代，DeepSeekV4 适配证明中国顶尖模型可跑在国产硬件上，能够降低企业全栈国产应用门槛。有研究机构预测，2026 年中国高端 AI 芯片市场增速超 60%，国产芯片有望占据半壁江山。

国产替代有条不紊。虽然华为昇腾 910C 推理性能约为英伟达 H100 的 60%（DeepSeek 基准），下一代 950PR 支持 FP4 精度（对标英伟达 Blackwell），但 2026 年规划产能仅 75 万片（远低于英伟达），但是国产替代仍有条不紊进行，华为正牵头 2000 余家中国企业，目标 2028 年半导体全链条自主率超 70%。预计 2030 年前中国晶圆代工产能占比超 20%，以 28nm 及以上成熟制程为主。

算法替代效率突破。DeepSeek 用更少算力实现前沿性能（混合专家 MoE、低成本训练），证明硬件约束下中国仍能打造竞争力 AI，并可向全球输出低成本方案——算法层面进口替代，与硬件替代形成互补。

实现进口替代具有重要的经济意义。国产替代使中国企业采购从英伟达等高溢价进口芯片转向华为昇腾，技术通胀收益留在国内，而非流向美—台—韩供应链。国内 PPI 也将更多反映本土供需，而非单纯进口价格信号。

3 碳基通缩部分源自于就业替代

碳基通缩并非单纯来自总需求周期性下行，而是来自技术替代、收入分配和消费结构再平衡的共同作用。AI 提升硅基资本回报和企业生产效率，但其对劳动任务的替代会压低部分劳动者的边际议价能力。若技术进步带来的供给效率提升快于劳动收入改善，宏观经济便容易形成硅基通胀、碳基通缩的结构，即上游算力、芯片、能源、数据中心和科技资本品价格具备韧性，居民消费品、可选消费、传统服务和劳动密集部门价格弹性则趋于下降。

3.1 美国 AI 对就业的替代：快速“白领冲击”

当前，美国在生成式人工智能的应用与落地速度上领跑全球，其对劳动力市场的“替代效应”正在以远超以往技术革命的速度推进。这种“快”不仅体现在局部行业的直接裁员上，更体现在对特定白领岗位、职场初级入门岗位的精准定向渗透，以及企业的招聘策略逆转。

与传统工业革命主要冲击低技术蓝领工人不同，本轮 AI 技术的替代主力直奔知识密集型与行政办公的“白领”领域。美国作为高度依赖服务业和智力资产的经济体，其白领岗位暴露出的技术可替代性极高。根据美国橡树岭国家实验室、北卡罗来纳州议会与麻省理工学院等多机构联合发布的最新“冰山指数项目”表明，除科技行业外，AI 的技术能力已实质性延伸至认知自动化领域，覆盖了美国劳动力市场中多达 11.7% 的工作内容，涉及薪酬总额高达 1.2 万亿美元，其中法律助理、行政管理、金融初级分析和合规审查成为了被替代的第一前线。这与我们在 2025 年 10 月 23 日发布的前期报告《人工智能“创造性破坏”的宏观推演》中推演一致，即人工智能的就业替代呈现出显著的行业异质性，尤其是在

信息、金融和商业服务行业中，超半数受访企业已将人工智能作为其业务流程的一部分。虽然不同行业的人工智能可替代性虽呈现出较大差异，但技能可替代性总体正逐步上升。

表2：美国劳工局对选定人工智能相关职业的就业情况及 2023-2033 年预测（千人）

职业名称	2023 年就业人数	2033 年就业人数	2023-2033 年人数变化	2023-2033 年变化率	全行业中总排名
总就业人数	167,849.80	174,589.00	6,739.20	4%	NA
法律类职业	1,394.40	1,446.20	51.8	3.70%	15/22
律师助理与法律辅助人员	366.2	370.5	4.3	1.20%	590/832
艺术、设计、娱乐、体育与媒体类职业	3,073.10	3,203.20	130.1	4.20%	14/22
平面设计师	267.2	273.8	6.6	2.50%	511/832
广播播音员与电台 DJ	26.5	25.4	-1.1	-4%	713/832
技术文档工程师	50.1	52.1	2	4%	394/832
口笔译人员	78.3	80.1	1.8	2.30%	521/832
销售及相关职业	14,609.50	14,319.50	-290	-2%	21/22
保险销售代理人	547.6	581	33.4	6.10%	219/832
服务类销售代表（不含广告、保险、金融服务与旅游）	1,170.30	1,214.50	44.2	3.80%	415/832
批发与制造业销售代表（技术与科学产品方向）	325.3	335.3	10	3.10%	476/832
批发与制造业销售代表（不含技术与科学产品方向）	1,356.10	1,366.10	10	0.70%	608/832
模特	3.6	3.6	0	-0.40%	633/832
销售工程师	59.7	63.4	3.7	6.30%	214/832
办公室与行政支持类职业	19,671.90	18,975.60	-696.3	-3.50%	22/22
采购文员	63.3	59.1	-4.2	-6.70%	747/832
信用审核员、核对员与文员	14.5	13.7	-0.8	-5.20%	732/832
客户服务代表	2,954.60	2,805.80	-148.8	-5%	728/832
行政秘书与高级行政助理	516.1	517.2	1.2	0.20%	618/832
法律秘书与行政助理	154.2	146.3	-7.9	-5.10%	730/832
医疗秘书与行政助理	768.9	810.1	41.3	5.40%	275/832
秘书与行政助理（不含法律、医疗与高级行政方向）	1,982.30	1,976.20	-6.1	-0.30%	630/832

资料来源：BLS，浙商证券研究所

此外，我们先前在报告《人工智能“创造性破坏”的宏观推演》中也提示，企业在采纳人工智能后，其人力策略核心是停止招聘，继而导致初级岗位的直接消失。核心原因在于，人工智能模型更擅长替代可通过正规教育获得的“程式化知识”（即课本中学来的知识），而这正是年轻劳动者的主要技能储备；相比之下，人工智能难以复制依赖长期实践积累的“默会知识”（即工作经验中的知识），这正是资深员工的核心优势。相关数据曾揭示，在高 AI 相关职业中，22-25 岁职场新人的就业率自 2022 年底至 2025 年 7 月下降了 6%。以软件开发岗位为例，22-25 岁群体的就业人数相比 2022 年底的峰值下降了近 20%。因入门级任务被人工智能自动化，年轻人的职业起步路径正面临断裂风险。

3.2 劳动参与率下行：社会保障依赖度上升背景下的劳动力市场困境

从政治经济学的现实博弈角度，随着通用人工智能对白领与行政任务的解构，全社会劳动参与率的下降，让失业人口依赖社会保障正在逐步成为当前美国部分极端思维蔓延下，后续处理大规模技术性失业的主流模式。由于 AI 替代人类劳动的速度显著快于劳动者技能重塑周期，政府主导的长期社会保障支持，可能形成劳动力市场退出激励。当大量劳动力议价能力下降、重返职场的薪酬预期低于福利保障水平时，部分人群会在统计上转为非劳动力人口，劳动技能与市场参与度逐步弱化。

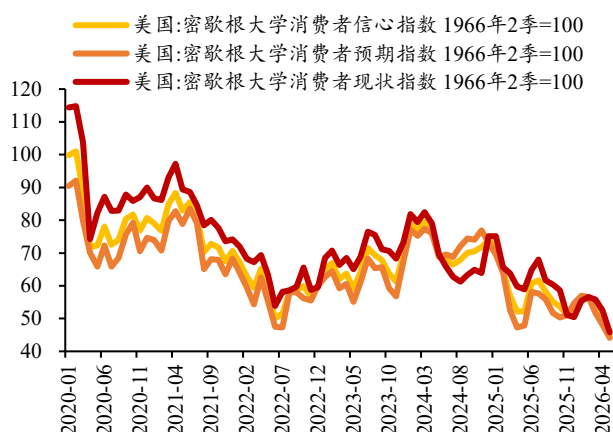
“以社会保障维持社会稳定”的模式，可能成为一种收入再分配机制，加剧美国贫富分化。在资本与技术高度集中的财富分配格局中，科技企业与资本方通过 AI 应用降低人力成本，利润集中度提升；而依赖社会保障的技术性失业群体，其生活水平主要由救济标准决定，难以分享技术红利，形成上层技术精英、下层福利依赖群体的分层结构，社会流动性受到约束。

当前美国移民政策呈现高技能优先、低技能配额收紧、边境管控强化的特征。政策意图在于减少外部劳动力供给，提升本土劳动力参与就业的必要性，推动失业人口重返劳动市场。白领与传统蓝领在收缩的就业市场中竞争加剧，社会保障依赖度上升，社会阶层向上流动通道收窄。

3.3 消费降级与通缩传导：K 型分化后的负反馈循环

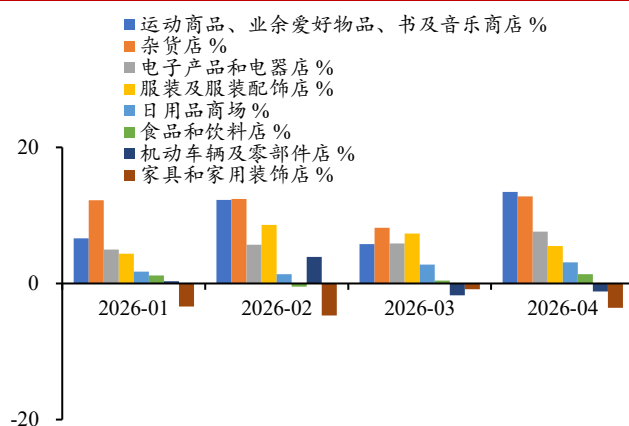
本轮生成式 AI 的核心特征在于其精准替代与重构高学历、高收入的“中层白领”岗位。这一群体的收入下调与失业，正在成为消费降级的核心推手。传统的工业革命或自动化主要冲击低收入蓝领，而蓝领的消费边际倾向虽高，但其绝对消费总额有限。当 AI 智能体在企业后台实现大规模替代后，被裁撤或被迫转入低薪岗位的白领阶层，其可支配收入出现断崖式下跌。不仅如此，由于 AI 技术替代的边界具有高度的不确定性，即使尚未被裁员的在职白领，在“失业焦虑”的驱动下也会大幅提高预防性储蓄。这种由技能错配与职业恐慌交织构成的心理防御，正在促使居民消费行为全面向“刚需化、低客单价”的降级路径靠拢，导致宏观总需求长期低迷。

图4：当前美国消费者指数已走弱至历史最低水平



资料来源：Wind，浙商证券研究所

图5：美国家具汽车等耐用品消费当前正持续负向拉动消费总额



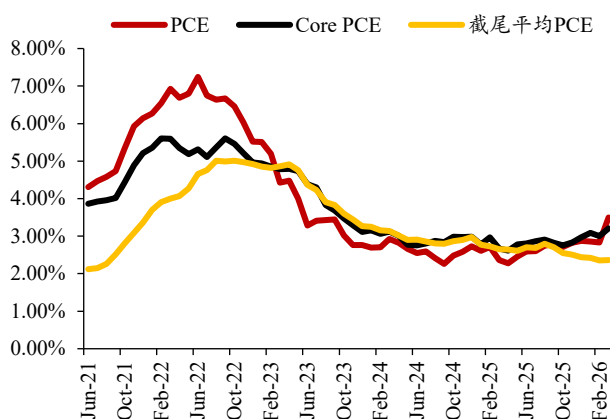
资料来源：Wind，浙商证券研究所

3.4 沃什截尾平均 PCE 框架，为中期通胀降温 and 降息打开空间

美国通胀读数短期可能受到能源、关税和商品价格扰动推升，但截尾平均 PCE 更偏向观察内生通胀粘性。2026 年美国 CPI 读数可能偏高，但新锚点截尾 PCE 下半年可能维持低位；预计 2026 年全年美国 CPI 同比增速为 3%，而截尾 PCE 全年同比增速为 2.4%。

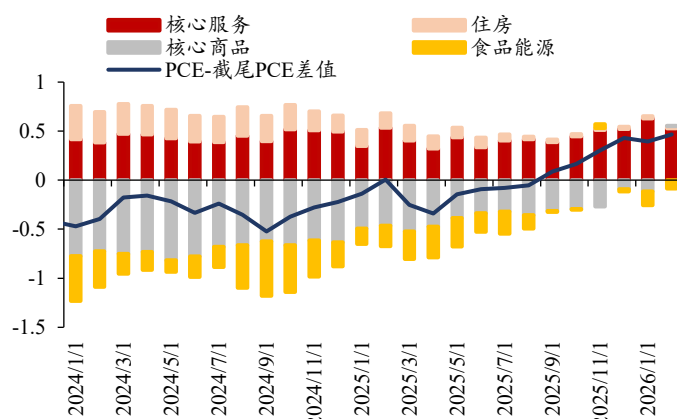
沃什主张以截尾平均 PCE 替代现行核心 PCE，政策含义在于淡化短期外生冲击，更加重视就业市场风险。截尾平均 PCE 每月将 177 个 PCE 分项按价格变动幅度排序，剔除支出权重低于第 24 百分位和高于第 69 百分位的分项，对剩余约 45% 的分项取加权平均。2026 年 3 月，美国截尾平均 PCE 同比为 2.4%，同期 PCE 和核心 PCE 分别为 3.5% 和 3.2%；截尾 PCE 与 PCE 之间 1.1 个百分点的差距，可理解为短期冲击对整体通胀的阶段性拉动。

图6：核心 PCE 一季度上行，而截尾 PCE 维持较低读数



资料来源：Bloomberg，达拉斯联储，浙商证券研究所

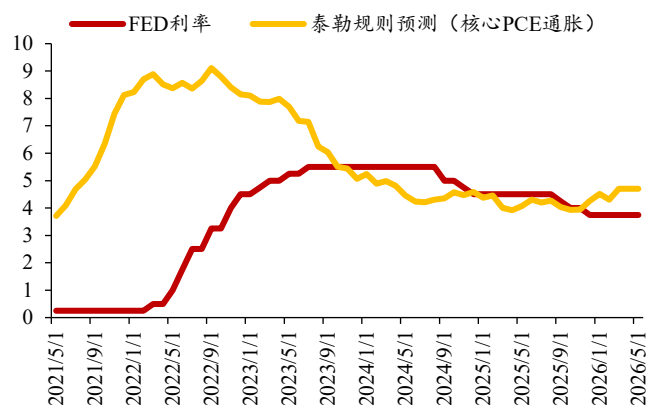
图7：25 年以来 PCE 和截尾 PCE 之差由负转正正是核心商品贡献



资料来源：达拉斯联储，浙商证券研究所

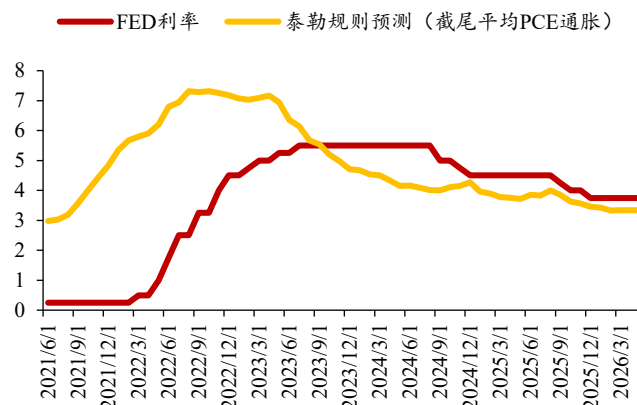
这一指标框架与碳基通缩逻辑高度一致。若 AI 替代和 K 型分化压制劳动收入预期，居民消费意愿和价格承受力下降，则美国内生通胀压力会弱于整体通胀所显示的水平。能源和关税可以阶段性推高 CPI、PCE 或核心 PCE，但截尾平均 PCE 更能反映中期需求放缓和广义价格压力回落。泰勒规则视角下，若以截尾 PCE 衡量通胀，截至 2026 年 3 月隐含的联邦基金利率上限为 3.34%，较当前联邦基金利率仍有约 41BP 降息空间。

图8：现行核心 PCE 指标泰勒公式政策利率（%）



资料来源：Bloomberg，达拉斯联储，浙商证券研究所

图9：截尾 PCE 指标泰勒公式政策利率（%）



资料来源：达拉斯联储，Bloomberg，浙商证券研究所

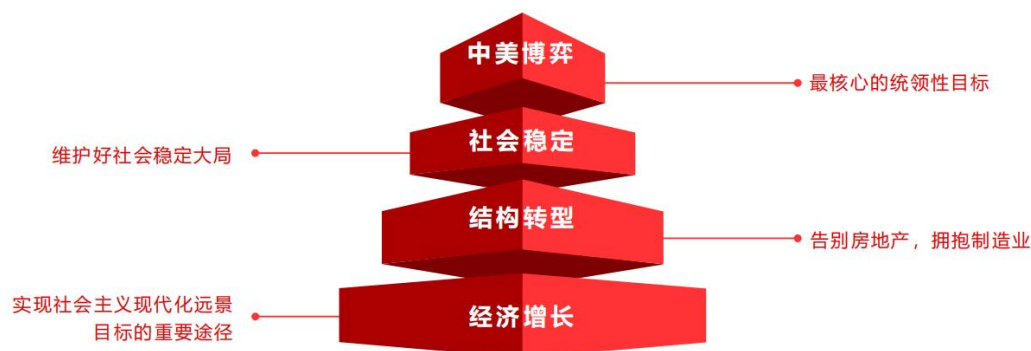
因此，沃什截尾平均 PCE 框架本质上是在高通胀读数中识别低内生通胀。只要就业端和消费端压力继续累积，货币政策就存在从抗通胀转向防就业和需求过度走弱的空间。碳基通缩越明显，截尾平均 PCE 相较核心 PCE 的解释力越强，降息的宏观叙事也越容易成立。

3.5 中国结构转型重于总量扩张，K 型分化强化居民消费品通缩压力

中国宏观决策正在形成中美博弈—社会稳定—结构转型—经济增长的四层次框架。经济增长仍是实现现代化远景目标的重要路径，但已不再是孤立目标；结构转型决定增长质量，社会稳定约束转型节奏，中美博弈则决定长期战略方向。由此看，经济增长要服务于结构转型，结构转型要服务于科技自立自强、供应链安全和战略竞争，社会稳定则是转型持续推进的底线。

这一框架意味着，在全年增长目标总体可实现的前提下，政策不会重新依赖房地产和传统基建来换取短期高增，而是将资源更多投向 AI、半导体、高端制造、数字经济、绿色低碳和新型基础设施。传统地产链、低附加值制造业和部分产能过剩行业的压力，更多被视为结构升级过程中需要承受的转型成本，政策取向也更接近托底而非强刺激。

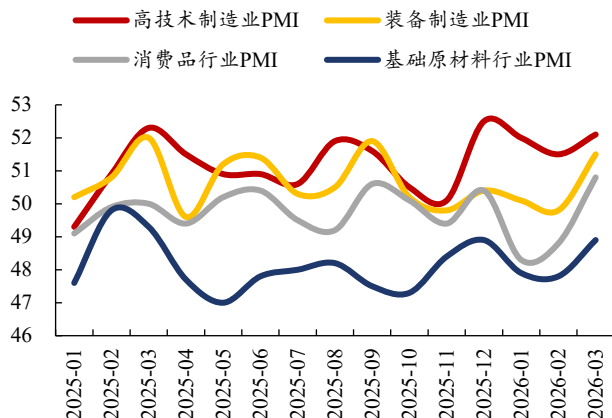
图10：百年未有至大变局下的“四层次”发展框架



资料来源：浙商证券研究所

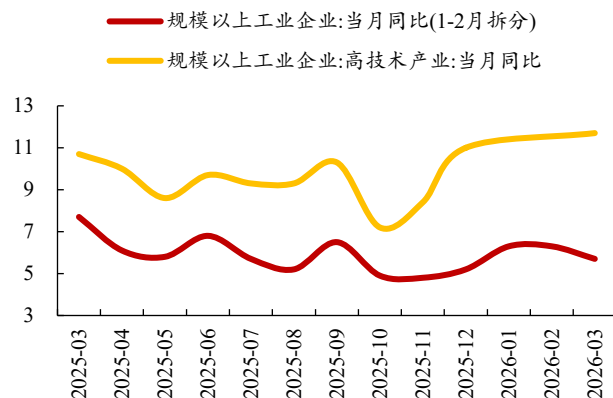
经济数据已经体现出明显 K 型分化。一端是新经济维持高景气，高技术制造业 PMI、战略新兴产业 EPMI 与传统制造业 PMI 裂口扩大，工业用电量和计算机等新兴产业用电量增速持续背离，新产业呈现物价、利润同步改善。2026 年一季度，中国 GDP 同比增长 5.0%，规模以上工业增加值同比增长 6.1%，其中高技术制造业增加值同比增长 12.5%，明显高于整体工业。

图11：高技术制造业 PMI 与传统制造业 PMI 裂口扩大（%）



资料来源：Wind，浙商证券研究所

图12：高技术制造业增加值持续高于整体工业

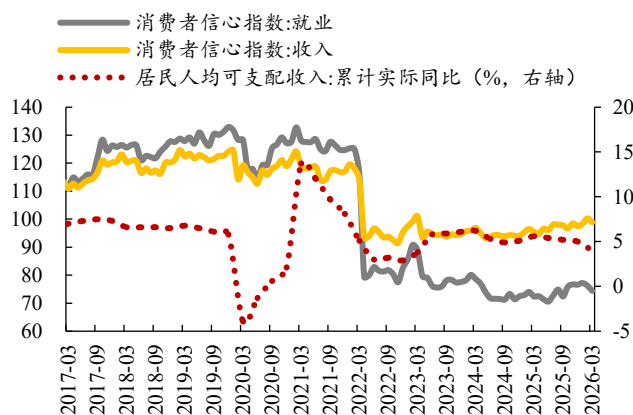


资料来源：Wind，浙商证券研究所

另一端是传统内需仍偏弱。居民消费仍处弱修复阶段，收入预期、就业压力和资产价格调整后的财富效应减弱，是制约消费弹性的核心因素。2022 年以来，消费者信心指数明显下台阶，其中就业和收入分项长期处于低位；居民人均可支配收入实际同比虽维持正增长，但中枢较疫情前并未显著抬升，显示消费能力和消费意愿均处缓慢修复状态。

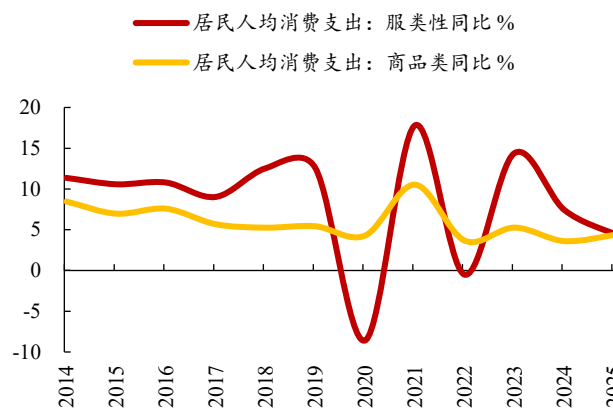
投资结构同样分化。2026 年固定资产投资预计增长 1.4%，其中制造业投资预计增长 5.2%；基建投资预计增长 5.6%；房地产开发投资预计下降 11.2%。这说明投资并非全面扩张，而是由高技术制造、新型基础设施和政策性项目拉动，地产链条仍是主要拖累项。

图13：消费者信心指数与居民人均可支配收入增速仍处于低位



资料来源：Wind，浙商证券研究所

图14：2023 年至今居民人均消费支出服务类同比增速高于商品



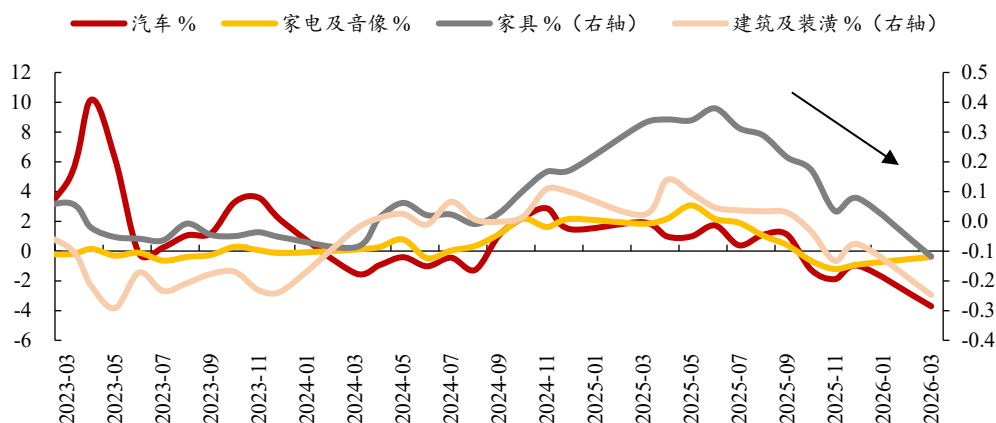
资料来源：Wind，浙商证券研究所

K 型分化对通缩的影响，核心在于价格结构和定价权分化。AI、高技术制造、上游资源和新型基础设施相关部门具备价格和利润弹性，而传统消费、地产后周期、低附加值服务和居民消费品更容易受到需求偏弱和竞争加剧约束。即便 2026 年下半年 CPI、PPI 和

GDP 平减指数可能温和修复，若价格回升主要来自能源、原材料和上游成本推动，而居民收入预期和消费倾向改善有限，则碳基部门仍面临结构性通缩压力。

居民消费品价格下行压力由三条链条共同形成：1) 地产财富效应下降压制汽车、家电、家具、建材等地产后周期消费；2) 就业和收入预期偏弱压制居民消费倾向；3) 供给效率提升和行业竞争加剧压低企业定价权。名义社零读数回升若主要来自油价和部分价格因素抬升，更多体现为收入端托底，而非实际消费量和利润率同步改善；汽车、家电、家具、建材等大宗和地产后周期消费仍受到价格竞争、地产销售偏弱和居民加杠杆意愿不足制约。

图15：汽车及地产后周期相关品类对社零拉动逐步减弱



资料来源：Wind, 浙商证券研究所

因此，碳基通缩不是简单的总量物价下行，而是居民部门购买力、消费意愿和企业定价权相对下降。结构转型越快，硅基部门资本开支和利润改善越明显；但若就业吸纳、工资增长和收入分配调整滞后，传统消费品和服务价格仍将承压，形成硅基部门通胀、碳基部门通缩的典型 K 型价格格局。

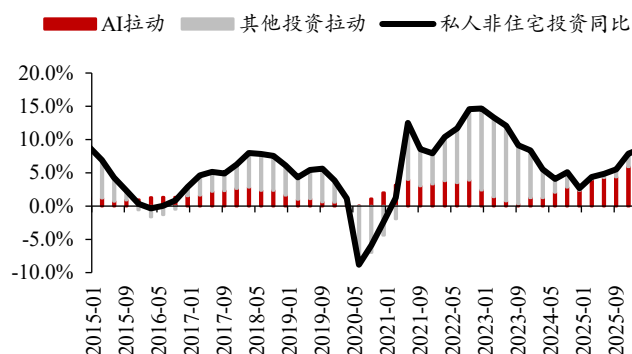
4 硅基与碳基的互动发展过程推演

4.1 2026 年硅基挤压碳基，但主要由 AI 资本支出带动，未缓解债务压力

2026 年受 AI 数字基建投资拉动，硅基经济和碳基经济呈现明显 K 型分化。从投资分项来看，AI（私人信息处理设备和软件固定投资）对非住宅投资增长的贡献率从 2024 年 Q1 的 23% 持续上升至 26 年 Q1 的 71%，传统碳基指数被压缩。

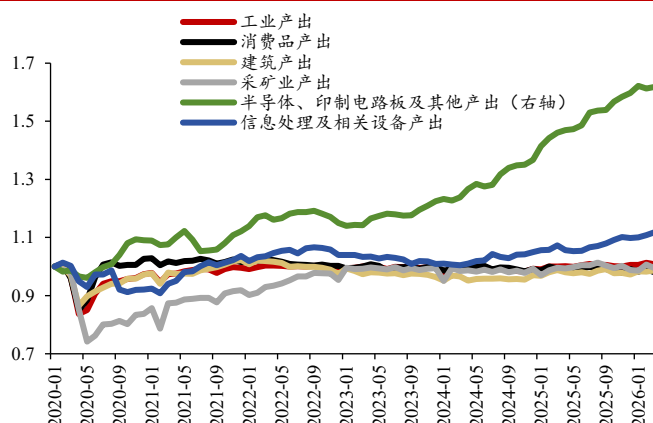
工业产出指数方面：截至 2026 年 3 月消费品产出指数降至 0.980，较 2020 年初下降约 2.0%；建筑产出指数降至 0.987，下降约 1.3%。相较之下，AI 相关产业链明显扩张，半导体、印制电路板及其他产出指数升至 1.618，累计增长约 61.8%；信息处理及相关设备产出指数升至 1.118，累计增长约 11.8%。

图16：AI 相关投资 26 年 Q1 贡献非住宅投资增长的 71%



资料来源：Wind，浙商证券研究所

图17：2023 年以来半导体、信息行业经济度远超传统行业



资料来源：Wind，浙商证券研究所

财政视角来看，韩国等少数局部科技实力强，人口少，可以对 AI 征税补贴全民。韩国央行 26 年 Q1 数据显示，在 AI 需求拉动出口回暖以及内需回暖的双重驱动下，Q1 韩国 GDP 同比增 3.6%，创下自 2021 年第 Q4 以来的最快增速。其中，受 AI 投资扩大带动，第一季度芯片出口出货量同比暴涨 139.1%，增幅较上一季度扩大两倍多。2026 年 5 月 11 日，韩国总统府政策室长金容范公开提议设立“公民红利”，认为 AI、半导体产业是国家数十年产业积累的结果，超额税收应全民共享，应当使用 AI/ 半导体产业贡献的超额税收来进行普惠性质的补贴。

主要国家硅基经济仍处在资本开支阶段。当前硅基经济繁荣的驱动力来自于下游 AI 企业融资推进数字基础设施建设，AI 下游尚未形成稳定盈利。据 Gartner 预测，2026 年全球人工智能支出将达到 2.52 万亿美元，同比增长 44%。预计五家头部企业 2026 年资本支出将接近运营现金流，而过去十年平均约为 40%。

企业利润更多用于再投资而非形成大规模应税收益，尚不能为财政带来增量，美债的压力并没有因 AI 发展而缓解。当前硅基经济的产业链重心集中在算力基建、先进制程与设备采购，企业利润更多用于再投资而非应税收益。CBO 在 2026 年 2 月上调 2026 财年美国赤字率预期由 5.5% 升至 5.8%。在美伊冲突之后，更慢的降息、更多的军费支出还可能导致 26 年、27 年美国财政赤字率进一步上修。截至 2026 年 5 月 24 日，美国 10 债、30 债纷纷升破 4.5%、5% 的关键心理关口，债务压力依然紧迫。

4.2 24 年以来硅基盈利大幅跑赢碳基

我们使用半导体与软件开发行业观测硅基股票表现（硅基=软件+半导体硬件），使用一般零售、服装家纺、饮料乳品、白色家电、酒店餐饮、化妆品观测碳基消费股票的表现，美股使用标普镜像指数。

2024 年以来“硅基”半导体指数显著跑赢碳基消费企业，2024 年至今 A 股消费指数累计上涨 0.4%，同期半导体指数累计上涨 159%。

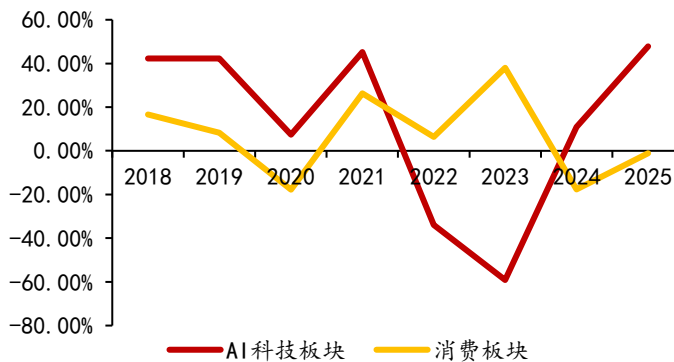
从盈利增速来看，2024 年以来“硅基”软件+硬件盈利显著跑赢碳基。2025 年，A 股 AI 科技板块 EBIT 同比增速为 47.81%，消费板块为 -1.16%；标普 500“硅基”板块 EBIT 同比增速为 37.27%，消费板块为 7.34%。

市值占比方面，美股、港股、A股中半导体硬件快速上升，软件疲软。美股“硅基”市值占标普500总市值比重从2020年末的13.32%升至2025年末的24.05%，截至2026年5月18日进一步升至25.58%，其中半导体市值17.20%；软件8.38%。

A股“硅基”市值占比从2020年末的3.12%升至2025年末的6.95%，截至2026年5月21日进一步升至9.22%。分项看，半导体市值占比从2020年末的1.54%升至2026年5月21日的7.84%；软件市值占比从1.57%降至1.37%。

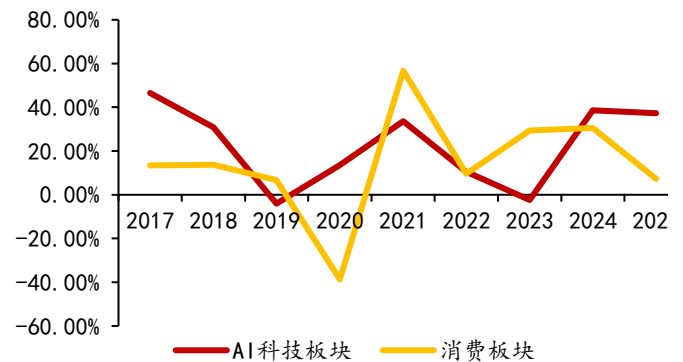
港股“硅基”市值占比从2021年末的13.48%降至2024年末的10.08%，2025年末回升至10.87%，截至2026年5月22日为11.81%。分项看，软件市值占比从2021年末的13.01%降至2026年5月22日的8.43%；半导体/硬件市值占比从0.47%升至3.38%。

图18：A股AI科技板块和消费板块EBIT同比增速



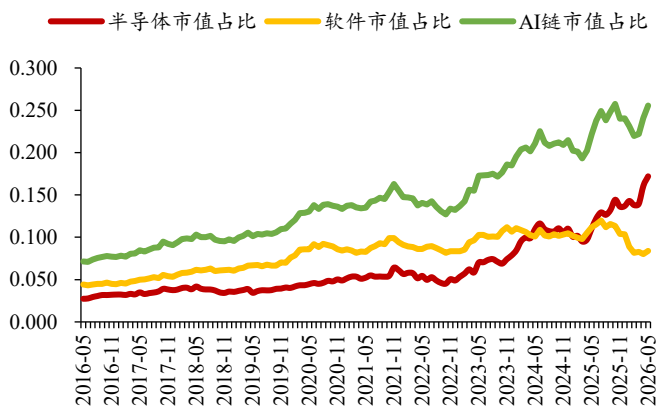
资料来源：浙商证券研究所

图19：标普500内AI科技板块和消费板块EBIT同比增速



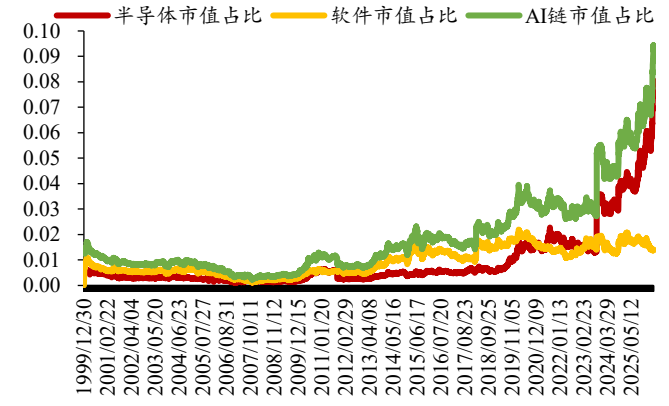
资料来源：浙商证券研究所

图20：美股AI链市值占比走势



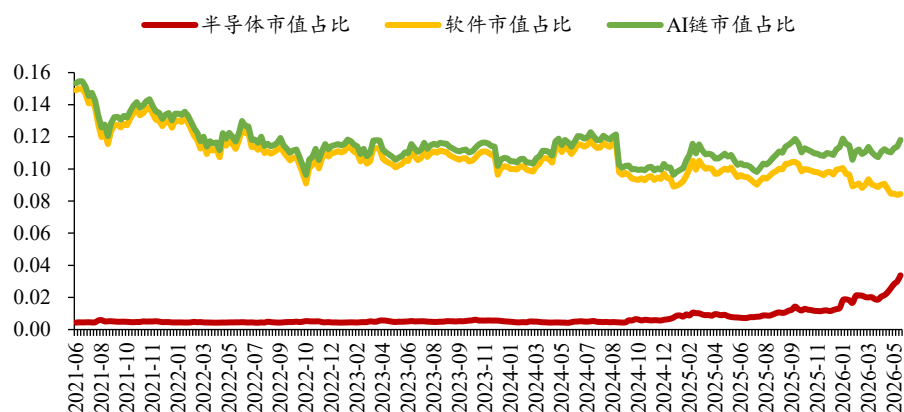
资料来源：Bloomberg，浙商证券研究所

图21：A股AI链市值占比走势



资料来源：Bloomberg，浙商证券研究所

图22：港股 AI 链市值占比走势



资料来源：Wind，浙商证券研究所

4.3 2027 年碳基经济极致挤压可能带来泡沫破裂风险

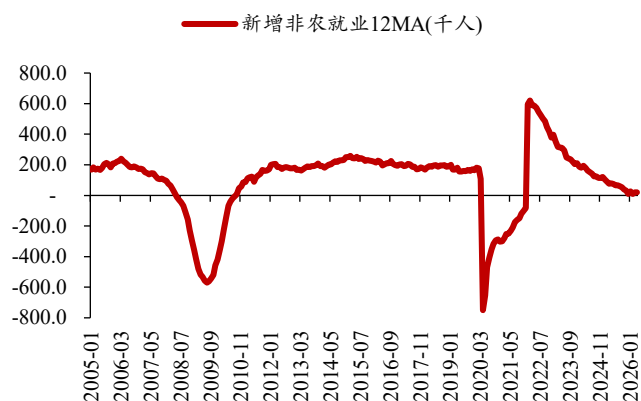
当前 AI 科技革命、美股（美股绑定科技巨头 MAG7）、美国经济（转向以 AI 为代表的投资驱动）深度绑定。一是 AI 科技革命带动美股科技巨头股价上行，AI 基础设施投入为美国经济增长提供韧性；二是美股上行为 AI 科技革命提供更好的投、融资环境；通过财富效应为美国经济增长提供动能（当前美国居民权益资产比例为历史极值）；三是美国经济增长催生 AI 需求（生产率瓶颈），并通过盈利增长支撑美股。

深度绑定的负面风险在于任何一环被破坏都有形成负反馈的风险，我们认为硅基经济极致挤压碳基经济的后果是 2027 年宏观增长可能出现下行风险，可能导致 AI 泡沫破裂。

一方面硅基经济就业拉动弱，美国劳动力市场持续放缓，美国新增非农（12 个月移动平均）已经由 22 年 7 月 52.6 万人降至 2026 年 4 月 2.9 万人，在硅基经济内部就业强度持续走弱，2023 年至今标普信息技术板块市值占总市值比重由 26.2% 上升至 34.1%，而信息技术就业占总就业比重反而从 2% 降至 1.7%。微观视角来看，在 AI 提高效率的背景下，多家大厂积极裁员，科技企业裁员人数 26 年 1 月再创新高。

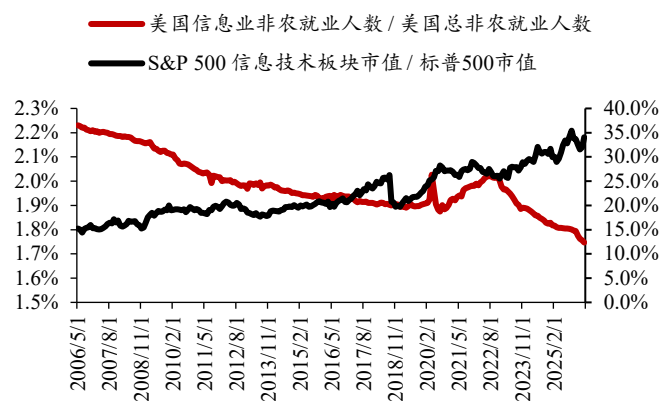
另一方面资本市场给 AI 下游软件的定价出现松动，资金向利润更为确定的硅基经济“上游”集中。硅基经济依赖下游 AI 大厂融资进行资本支出，2025 年 10 月后美国私募信贷问题爆发（是 AI 初创企业的重要融资来源）、26 年 2 月路透社报道 OPEN AI 用 17.5% 的利率融资、ORCL 等美国 AI 投资大厂的信用违约互换 CDS 快速上升，均指向下游融资压力升高。在这样的背景下，2025 年 10 月后同属于硅基经济的上游“半导体硬件”和下游“AI 软件”出现明显分化，市场切换至利润更为确定的 HALO（重资产、低淘汰风险）主线。

图23：新增就业趋势性放缓



资料来源：布鲁金斯学会，浙商证券研究所

图24：信息业就业加速下滑，但份额提升



资料来源：FED，浙商证券研究所

4.4 关注技术路线的变革

技术演进并非线性坦途，而是螺旋式上升，当前高资本开支推动硅基经济繁荣发展的背景下，技术路线变革带来两个方面的风险，可能引发科技股泡沫的破裂。

4.4.1 上游涨价方向有可能被新技术路线替代

2026年AI产业链上下游出现明显分化，上游无论是GPU、存储、光模块还是数据中心电力资源，均处于供给偏紧状态持续涨价，但下游应用端竞争激烈，用户付费意愿存疑，这一背景下AI企业有动力通过技术迭代化解成本压力：

一是算法迭代，可能颠覆当前数字基建的投资逻辑。当前中美主流AI大模型均基于Transformer架构，该架构下存在“键值缓存”概念（即大模型逐词生成文本时，把过往所有token的计算特征临时存起来，避免重复运算），使得内存占用随文本长度平方级暴涨。近来Mamba为代表的新架构凭借不同技术逻辑，舍弃KV缓存，可以形象的理解为Transformer每次回答问题前，需要将过往的文本历史完整翻阅，因而需要大量内存来保存数据；而Mamba边回答边做摘要，在回答前不再需要完整翻阅历史文本。当前Mamba技术还存在训练困难、推理能力不如Transformer等问题，但不排除短期内技术非线性突破，Mamba生态升值新的模型架构突破颠覆当前算力基建逻辑。

二是模型压缩。不同大模型1个词元(token)消耗的算力和每次计算调用的模型参数相关。当前大模型已经积极采用剪枝、量化、蒸馏、低秩分解等压缩技术，减少模型参数量，降低计算复杂度，同时降低芯片计算和存储需求。

三是参数调用模式的优化。例如当前主流大模型已经积极从稠密模型(Dense，每token所有参数都参与计算)转向MoE混合专家模型(部分参数参与计算)，也是同时降低计算和存储需求。

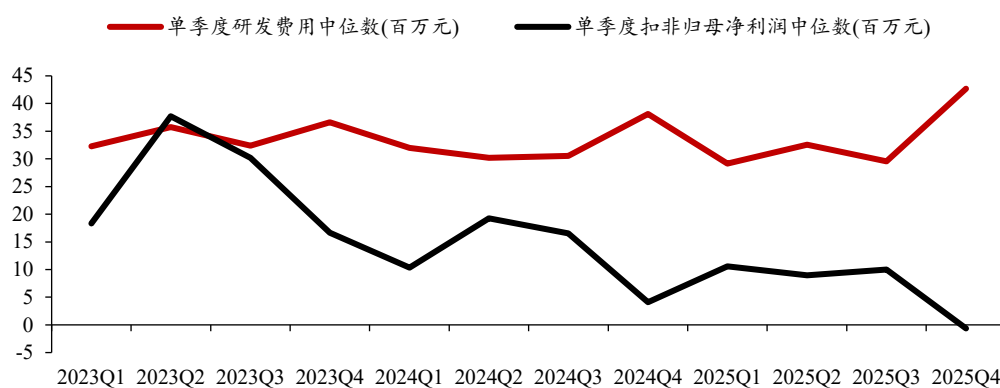
4.4.2 下游盈利兑现存在风险

二是下游盈利兑现风险。一方面当前AI投入尚未形成显著的利润增长，AI应用商业化面临技术和经营的不确定性。我们使用AIGC、AI医疗、AI教育、智能机器概念进行筛选来统计AI应用盈利情况，数据显示2023Q1至2025Q4期间，我们选取的AI下游板块单

季度研发费用中位数呈上升趋势，而扣非归母净利润中位数却波动下行，至 25 年 Q4 转为负数，直观反映了高研发投入与低盈利的背离。

另一方面头部 AI 存在同质化问题，有且已经出现价格战问题。当前主流 AI 大厂技术路径趋同（Transformer 架构等），大模型能力虽有区别但普通消费者感知差异有限，中美均是多头竞争格局，例如美国头部三家 OpenAI（ChatGPT）、Anthropic（Claude）、Google（Gemini）竞争激烈，这种行业生态存在价格战风险。5 月 22 日，DeepSeek 宣布旗下 V4-Pro 模型 API 永久降价至原价的四分之一，缓存输入成本低至 0.025 元/百万 Token，而 Gemini 3 pro 最低为 1.45 元/百万 Token，OpenAI GPT-5.5 则为 3.63 元/百万 Token。一旦 AI 行业价格战冲击盈利预期，可能引发硅基经济的重估。

图25：A 股 AIGC、AI 医疗、AI 教育、智能机器概念研发投入与盈利变化



资料来源：wind，浙商证券研究所

4.5 关注技术路线的阻碍

硅基经济快速发展，OECD 2026 年的报告显示 2023 年仅 8.7% 的 OECD 企业使用 AI，到 2025 年翻倍至 20.2%，AI “存在感”越来越强的同时也使其自身风险和弊端暴露出来。

我们认为，AI 风险可能因安全、监管、社会公平等议题导致监管政策收紧，迫使其渗透率提升放缓，制约硅基经济的发展速度。我们认为最需要关注的是技术失控风险和就业替代风险：

4.5.1 技术失控风险

截至 2026 年，AI 技术已经涉及多类刑事犯罪，包括 AI 教唆杀人（25 年佛罗里达州校园枪击案）、AI 电信诈骗（26 年 2 月最高法发布典型案例，使用 AI 拟声诈骗老人等）；AI 涉黄、AI 虚假信息、AI 抄袭等。由于 AI 技术快速发展，监管很难对未来风险进行管控。

人工智能的潜在负外部性已经推动主要司法辖区加速构建监管框架。根据英国人工智能安全研究所《2026 年国际人工智能安全报告》，通用人工智能系统的发展可能带来三个大类的风险，即恶意使用、故障、系统性风险。2024 年欧盟出台了《人工智能法案》，采用基于风险的分级监管，对于高风险 AI 序满足严格合规要求，不可接受风险如实时远程生物识别责备直接禁止。

监管强度的提升与技术进步之间存在两难权衡：前置审批与多辖区合规抬高市场准入壁垒，认证成本对中小企业形成挤出效应，产品迭代周期被迫拉长，对行业增长动能构成约束。根据 IMF 2025 年的报告，欧盟 AI 法案、数据隐私法和职业监管叠加，可能使 AI 生产率增益降低超过 30%。同时欧盟的 AI 法案合规成本过高（达 50 万欧元），碎片化监管一定程度上遏制了技术的创新，阻碍了新产品的上市，例如 Meta 的 AI 智能眼镜因监管障碍被迫推迟在欧洲的上市。

技术失控可能导致监管收紧。2026 年以“龙虾”等产品为代表，AI 正从“回答问题”进化为“自主行动”并深度嵌入日常生活工作，短期内其涉及风险还将从信息失真转向行为失控，对监管提出更高的要求。虽然当前中美正处于人工智能竞争白热化的阶段，但如果 AI 的发展引发了系统性的技术失控或某类刑事违法案件的趋势性增长，可能引发监管层面的超预期收紧，使市场对 AI 发展的预期发生重大调整，最终冲击到资本市场和硅基经济。

表3：AI 风险汇总

类别。	风险。	解释。
恶意使用风险。	AI 生成内容与犯罪活动。	通用 AI 系统能生成逼真的文本、音频、图像和视频，可用于诈骗、勒索、诽谤、非自愿亲密影像和儿童性虐待材料等犯罪目的。
	影响与操纵。	AI 系统可通过生成内容影响人们的信仰和行为。
	网络攻击。	通用 AI 系统可执行或协助网络攻击的多个任务。
故障风险。	可靠性挑战。	通用 AI 系统故障可造成伤害，包括产生虚假信息（“幻觉”）、编写有缺陷的代码、提供误导性医疗建议。
	失控。	一个或多个通用 AI 系统在任何人控制之外运行，重新获得控制极其昂贵或不可能。
系统性风险。	劳动力市场影响。	通用 AI 系统可自动化或协助与全球许多工作相关的任务。
	对人类自主性的风险。	通用 AI 系统可通过多种方式影响人们的自主性，包括认知技能、信仰和偏好形成、决策和行动方式。

资料来源：英国人工智能安全研究所，浙商证券研究所

4.5.2 就业替代风险

目前尚处于人工智能革命的早期，AI 对就业市场的影响尚处可控。美联储 2026 年报告显示 AI 仅对就业总体影响较小，估算自 2022 年 11 月至 2026 年仅导致整体失业率仅上升 0.1%。我们研究同样显示 AI 对就业的负面影响主要集中在科技领域内部，占整体就业比重较小（详见前期报告《经济走 K 型，资产举杠铃——2026 下半年宏观展望》）

人工智能对就业市场的影响，既延续了历次技术革命“创造性破坏”的共性，又在冲击的速度、范围和模式上展现出前所未有的独特性，未来其对就业市场的冲击可能超过历次科技革命：

在速度上，人工智能技术扩散极快，从“iPhone 时刻”到 10%渗透率仅用 1 年，远快于电力的 20 年和互联网的 7 年。

在范围上，IMF 2024 年的报告显示全球约 40%的工人处于高暴露职业，其中发达经济体约为 60%，新兴市场经济体约为 40%。

在模式上，人工智能率先冲击年轻和资浅员工。NBER 2025 年论文显示，在软件开发、客户服务等高 AI 暴露职业中，22-25 岁职场新人的就业率自 2022 年底至 2025 年 7 月

下降了6%，而同期35-49岁年长劳动者的就业率反而增长了6%至9%。以软件开发岗位为例，22-25岁群体的就业人数相比2022年底的峰值下降了近20%。颠覆了以往技术变革中年轻劳动力更具适应优势的传统认知。

人工智能背景下就业市场面临的核心挑战，是技术变革的指数级速度与社会适应能力之间的“剪刀差”。这种不匹配体现在教育体系改革滞后、劳动力技能转型缓慢、以及社会保障与收入分配政策未能及时跟进等多个方面。若无法有效弥合这一差距，可能导致结构性失业加剧、技能鸿沟扩大和收入不平等恶化，从而引发社会问题，进一步可能导致监管收紧或社会动荡。

4.6 中长期来看碳基有效利用硅基，重述社会分配和稳定机制

历次科技革命均发生“创造性破坏”，初期造成结构性失业、贫富分化问题，后期新技术诞生新岗位，社会制度逐步匹配，生产率得到提升。

第一次工业革命（1760—1840年）。工业革命前英国的纺织业主要依靠家庭作坊式的手工生产，工匠需多年训练才能掌握技艺。1760年蒸汽机诞生，珍妮纺纱机等机器大幅提高纺织业劳动生产率，导致英国织工的数量、工资下降。社会矛盾激化的背景下，1811年爆发“卢德运动”，有组织的破坏机器，最终被军事镇压。机器大工业对传统手工岗位的排斥，导致劳动力结构性失业加剧，红利在初期高度向工厂主集中，在缺乏制度保障的三十年间，英国工人实际工资停滞不前。

虽然传统织工失业，但操作火车、机械的需要催生了机械师、火车司机等新型职业；大规模生产也催生了工厂会计、管理等“白领”阶层，最终科技革命使得英国国力跃升。政府方面也通过《工厂法》限制工时及《普及教育法》推动劳动力转型等方式，逐步促使社会矛盾缓解。

第二次科技革命（19世纪末—20世纪初）。该革命以电力与内燃机为基础，直至1882年发电站与1885年变压器完备后才具备广泛应用基础。电力、内燃机、化工技术普及，一方面农业机械化使得大量农民被迫进城务工；另一方面流水线取代旧式作坊。但在结构性失业的同时城镇化快速发展，制造业、服务业工作比例上升，美国中产阶级逐步形成。为缓解社会矛盾，以美国为代表通过《反垄断法》拆解垄断巨头、德国率先建立现代社保体系，将垄断溢价转化为社会公共保障，缓解了社会压力。

第三次科技革命（20世纪中叶—21世纪初）。以计算机技术、生物技术、航空航天的发展为代表。科技革命的发展并非普惠，产业高端化和全球化的背景下美国传统铁锈地带出现大规模下岗潮，美国底特律、德国鲁尔区均出现严重衰退，旧的打字员、档案管理员等工作岗位逐步消失，结构性失业加重。但长期来看，第三次科技革命加速了就业结构中高端服务业的占比，电商、程序员、数据分析师等新的工作岗位出现。在1996至2000年间，互联网与计算机的普及直接拉动了美国潜在经济增速的提升，实现了“低通胀繁荣”。

人工智能与历次工业革命在技术替代的速度、范围、冲击群体及替代模式上存在根本性差异，其变革的烈度与广度可能前所未有。人工智能的广泛应用对就业的替代效应主要呈现三大显著特征：一是行业异质性突出，各行业人工智能渗透率差异较大，且岗位技能可替代性分化明显，当前就业市场中，近半数岗位技能具备人工智能可替代性，不同行业因此采取直接替代或员工再培训的差异化应对策略，且完全可被人工智能替代的技能数量呈现突破性趋势；二是初级岗位面临消失风险，企业采纳人工智能后更倾向通过停止招聘而非直接裁员调整人力，明显降低初级员工招聘率，悄然重塑人力金字塔；三是年龄冲击

非对称，年轻和资浅员工因技能知识更易被人工智能替代，在高人工智能暴露职业中就业率显著下降，而资深员工凭借难复制的经验知识实现就业增加，经验溢价大幅提升。

长期来看，人工智能部门物质财富贡献能力极大时，生产关系从按劳分配向按需分配过渡，最终服务于人。以大模型、算力基础设施与机器人为核心的广义 AI 资本或将重塑要素结构与分配机制。一方面，算法与数据的可复制性使知识类产出呈现递增规模报酬，推动大量数字化与标准化环节的边际成本下降，即推动供给曲线右移，价格下降；另一方面，通过“AI——机器人——实物产出”的链条，把近零边际成本向制造业和服务业传导，尤其是标准化环节，或将在制造、物流、能源、医疗与公共服务中形成自动化与自主演进的复合供给能力。

在此背景下，居民获得基本生活（如衣食住行等）所需要的成本占比将出现结构性下行，伴随其成本降至社会可承受的财政供给阈值之下，且私营部门在 AI 红利下有稳定的超额收益，社会即可通过税收与公共平台，将基本生活所需要的产品或服务以低价甚至零价的方式实现按需供给，以劳动收入作为重要购买力来源的地位将被削弱，居民的基础购买力由劳动所得转化为劳动+公共供给的复合结构，届时生产关系存在从按劳分配为主有序过渡到基本生活按需、超额需求市场化的可能。

4.7 本轮 AI 泡沫破裂不能过度关注美联储

在中美竞争的大背景下，AI 科技革命是国家战略竞争的关键变量，美国政策层有强烈动机维护金融稳定。当前 AI 泡沫和 2000 年科技股泡沫关键不同在于当前中美 AI 科技领域存在激烈竞争，而 2000 年美国在互联网科技革命上断崖式领先。2000 年是美国单边主导的科技革命，即便美国科技股泡沫破裂，其在全球科技领导权并未受到实质挑战；而当前中国在大模型、AI 应用、机器人、智能制造等方面已经紧紧追赶甚至部分超越。在“国家安全”考虑的背景下，美国对华实施芯片出口限制、半导体设备封锁、AI 人才和资本流动审查，均体现国家战略竞争的特征。当前 AI 已不仅是产业周期变量，更是中美两国竞争的国家战略层面变量。

在 AI 是大国竞争“前线”的宏观视角下，特朗普或联储主席沃什有必要维护金融稳定，在其当前的政策表态上已经多有体现，包括特朗普将股票指数新高作为自己的重要政绩宣传，以及任命新的联储主席沃什来推动其降息目标；沃什多次指出当前正处于生产力革命的初期，AI 是美国相对于竞争对手的巨大领先优势，未来将带来低通胀繁荣。

金融稳定约束下，联储货币政策可能相对审慎，AI 泡沫破裂的风险不应过度关注联储政策。展望下半年，虽然石油冲击下通胀读数可能偏高，但年内降息+扩表+放松金融监管可能是最大预期差（详见前期报告《经济走 K 型，资产举杠铃》）。

4.8 AI 泡沫破裂可能引发美元美债贬值出现主权债务危机

2025 年以来“去美元化”交易的根源在于美债可持续性预期恶化。当前美国债务总量已经达到极高水平。从存量指标看，截至 2026 年，美国政府对公众负债已经达到 31.4 万亿美元，CBO 口径负债率已达 100%，处于二战后高位；到 2035 年该比例预计进一步升至 118%。从流量指标看，CBO 在 2026 年 2 月曾预测 2026 年赤字率可能小幅收缩至 5.8%，但关税退款和战争预算可能推高赤字率约 1.2%。2027 财年特朗普更是提出 1.5 万亿美元的国防支出计划，使赤字收窄难度进一步加大。

从财政结构看，美国财政赤字易增难减：一是美国法定支出占总支出比重约 60%，既难以削减福利也难以大幅增税。二是自主性支出军费、产业投资有进一步扩张需要。三是利率中枢抬升，利息支出处于历史高位。

特朗普上台以来多项政策试图强化美元美债信用，但尚未取得成功。

一是关税手段强化美元美债地位。米兰在 2024 年提出“海湖庄园构想”，并在特朗普胜选后得到重用被任命为白宫政策委员会主席及美联储理事。其“海湖庄园构想”一是“财政贬值”，即在强美元的背景下，美国通过进口关税+工资税削减的组合财政组合去复刻货币贬值的效果，达到强化国内制造业和美元地位的目标。二是汇率干预，由于美国的债务规模较高，因此美元贬值后对债券市场的利率衍生冲击必须予以考虑，需要以关税和取消安全保护为压力工具，1) 要求对手国抛售美元，推动本国货币升值；2) 将美国的安全保护功能与债券购买挂钩，使得利率保持稳定。

各国强硬反制和美债风险超过米兰预期，通过关税强化美元并不实际。4 月特朗普对全球施加“对等关税”，特朗普政策扰动过大，通胀和经济衰退预期下随即 4 月 7 日美国 10 年期国债收益率出现快速飙升，“基差交易”平仓和米兰“对美债加税”的发言进一步放大波动，市场出现“去美元化”交易。

二是稳定币战略强化美元地位，但规模较小影响力有限。稳定币本质是利用加密货币交易+跨境支付需求加持美元美债。2025 年 7 月特朗普签署《GENIUS》法案，一是通过稳定币监管强化美元地位，包括发币机构需配合美国金融监管、不被认可的外国人禁止发币等。二是严格要求抵押资产类型，必须持有等额的美元、美国短债等高流动性资产。美国财政部测算，截至 2024 年稳定币持有约 1200 亿美元短期国债，预计到 2028 年稳定币市场可能扩张到 2 万亿美元规模，短债需求增长至 1 万亿美元。美元通过“购买稳定币—支持加密货币交易—投资美国国债”这一循环，又回到美国金融市场。

尽管美国财政部对稳定币的扩张保持乐观的预测，但稳定币规模依然不足。截至 2026 年 2 月，DefiLlama 显示美国稳定币市值仅有 3089 亿美元，较 2025 年 7 月法案签署之日的 2599 亿美元仅累计多增 490 亿美元，合计每月仅多增约 82 亿美元，而 2025 年至今美债月均增量约 1835 亿美元，稳定币对美债贡献有限。

三是重建石油美元体系，但在美伊战争中遭遇重挫。石油美元的目的并不是盲目追求低油价，而是让美国从适度的高油价中获益。特朗普 2025 年上台以来多举措试图强化石油美元，包括：1) 获取产油国对美国的投资，打通石油美元利润回流的渠道；2) 特朗普主导成立“和平委员会”，中东多国主动捐赠，服务于美国“影子财政”。3) 对委内瑞拉、伊朗动武，特朗普明确表示为了获取石油资源；4) 对格陵兰岛、古巴、加拿大发出主权威胁，意在垄断美洲、北极资源（详见前期报告《北极地区会是下一个热点吗？》）。2026 年以来美伊战争陷入泥潭，伊朗实质控制霍尔木兹海峡威胁石油美元地位。

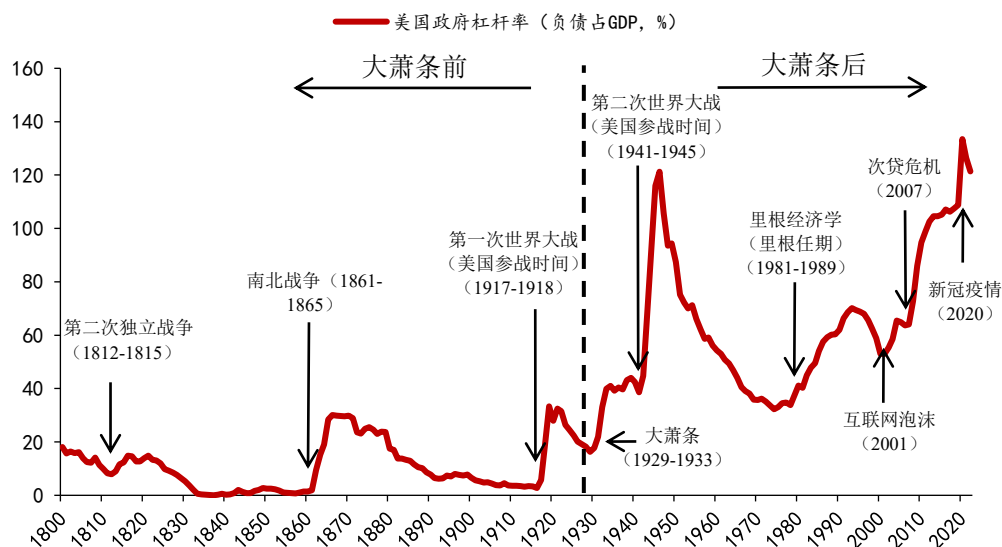
在这样的背景下，AI 重振美元美债的重要性被动提升，特朗普施压科技企业投资、叫停 AI 监管，多次表态要求美国 AI 技术领先世界。2025 年以来特朗普确实多次召集科技公司高管，公开施压科技企业投资 AI，例如 2025 年 1 月星际之门项目，特朗普称中国是这一新兴产业中的主要竞争对手；2025 年 7 月宾州 AI/能源峰会，包括 Google、Blackstone、CoreWeave 等企业现场宣布超过 900 亿美元 AI 和能源投资，白宫措辞是“确保美国成为无可争议的 AI 领导者”；2026 年 5 月，特朗普叫停签署 AI 行政令，明确表态“不想做任何妨碍美国领先地位的事”。

当前 AI 以投资为主，未看到明显税收贡献。当前硅基经济的产业链重心集中在算力基建、先进制程与设备采购，企业利润更多用于再投资而非应税收益。为实现大规模资本

支出，企业债务融资力度加大，2025 年亚马逊、GOOGLE、META、ORCL、微软五家公司发行了 1210 亿美元的企业债券，是 2020 年至 2024 年间的平均年均发行额为 280 亿美元的 4 倍。由于减税影响，2026 年 1-4 月美国财政收入累计同比增速仅 3.3%，低于 2025、2024 年同期 9.2%、11.8% 增速。企业税方面，截至 2025 年 Q4 同比增速 -1.8%，同样低于 2010 年以来均值 7.3%。

长周期视角来看美国国债和战争风险、科技革命两大因素强相关，AI 可能实现分母端化债。二战后两次化债一是 50 至 70 年代黄金时代 + 第三次科技革命；二是 90 年代戈尔巴乔夫上台后冷战逐步走向终结+美国信息技术革命带来的潜在增速和税收提升。若 AI 科技革命顺利进行可能从全要素生产率方面推动美国经济增长，并减少美国福利财政的依赖，实现分母端化债。若 AI 泡沫破裂，则意味着 2025 年以来特朗普强化美元美债的措施纷纷告破，将进一步恶化美元美债信用，主权债务风险大增。

图26：二战后两轮化债均与科技革命有关



资料来源：Fed，浙商证券研究所

4.9 国际货币体系主导货币债券违约将引发百年一遇金融危机和经济危机

主权债务违约有两种形式：一是拒绝兑付直接违约；一方面是拒付外债，美国当前外债占比极低，截至 2025 年美国外债(政府外币债务)占 GDP 比重为 0.37%，外债风险极低；另一方面是拒付本币债，历史来看，从英国到美国，主权债务直接违约的情形均相对有限，多数由技术性因素（79 年美债系统故障）或制裁引起（例如冻结资产）。

二是通过通胀、贬值等手段间接违约，例如 1933 年罗斯福“黄金条款废除”（国债原本以黄金兑付，改为用贬值的纸币偿还）或 1971 年尼克松冲击（停止兑换黄金），均对资本市场造成重大影响。

虽然美国宪法要求国债需得到兑付，但我们提示，全球“百年未有之大变局”背景下，不应忽视美债直接违约的极端尾部风险。若科技股泡沫破裂，中美竞争激化，极端情形下不排除美国政府通过司法解释绕过宪法债务兑付的约束：

一是可能对外部美债直接违约，例如大规模冻结、扣押外国政府美债资产；美国《敌国贸易法》、《国际紧急经济权力法》均许可美国政府在特殊情形下处置外国政府资产。2022 年俄乌冲突后，欧盟与 G7 体系内被冻结的俄罗斯主权资产规模被估计约 2895 亿欧元，G7 冻结俄罗斯主权资产显著强化了各国央行对单一结算体系集中暴露的警惕。

二是也可能通过行政手段违约，例如发行低利率特殊债券，参考战时“战争债券”，以行政手段迫使机构、个人购买折价债务。

一旦出现上述两类美债违约行为，将迅速打破 70 年代至今建立的美元信用体系，引发百年一遇金融危机和经济危机：

一是美债资产价值大幅缩水，截至 2026 年 Q1 外国投资者共持有美债 9.3 万亿美元，美国国内投资者持有 22.1 万亿美元，包括各国央行外汇储备、主权财富基金、商业银行、养老基金在内均有大比例资产以美债形式持有（详见前期报告《“乱世”新纪元，拥抱“新黄金”》），一旦资产负债表受损可能引发全球流动性危机的快速蔓延，从金融信贷收紧、资产负债表收缩等渠道引发金融、经济危机。

二是国际贸易与跨境资本往来秩序紊乱。美元作为国际贸易储备货币发生违约，将导致全球贸易以美元计价、结算、清算的传统模式难以运转；大宗商品价格发生混乱；各国企业跨境收付、贸易融资等业务停滞。

三是地缘博弈可能升级。美元是多国重要外汇储备，美债违约可能导致金融制裁、贸易壁垒手段频繁加码，叠加科技泡沫破裂带来的资本市场重创，最终演化成百年一遇的全球性金融危机与经济萧条。

5 风险提示

1、人工智能技术迭代及应用落地不及预期风险。若 AI 发展引发系统性的技术失控或某类刑事违法案件的趋势性增长，监管层面可能超预期收紧，使市场对 AI 发展的预期发生重大调整，AI 渗透率提升放缓，并冲击资本市场和硅基经济。

2、硅基产业链供需格局变化超预期风险。若先进制程、HBM、先进封装等产能释放快于预期，或算法效率突破显著降低单位算力需求，则上游稀缺性和定价权可能回落，削弱硅基产业链价格上行和 PPI 科技分项上涨的判断。

3、就业替代与碳基通缩演绎不及预期风险。若 AI 更多体现为放大碳基创新价值而非替代劳动，或政策托底、居民收入改善和服务消费修复强于预期，则居民消费品、可选消费、传统服务和劳动密集部门价格弹性可能不降反稳。

股票投资评级说明

以报告日后的 6 个月内，证券相对于沪深 300 指数的涨跌幅为标准，定义如下：

1. 买入：相对于沪深 300 指数表现+20%以上；
2. 增持：相对于沪深 300 指数表现+10%~+20%；
3. 中性：相对于沪深 300 指数表现-10%~+10%之间波动；
4. 减持：相对于沪深 300 指数表现-10%以下。

行业的投资评级：

以报告日后的 6 个月内，行业指数相对于沪深 300 指数的涨跌幅为标准，定义如下：

1. 看好：行业指数相对于沪深 300 指数表现+10%以上；
2. 中性：行业指数相对于沪深 300 指数表现-10%~+10%以上；
3. 看淡：行业指数相对于沪深 300 指数表现-10%以下。

我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重。

建议：投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者不应仅仅依靠投资评级来推断结论。

法律声明及风险提示

本报告由浙商证券股份有限公司（已具备中国证监会批复的证券投资咨询业务资格，经营许可证编号为：Z39833000）制作。本报告中的信息均来源于我们认为可靠的已公开资料，但浙商证券股份有限公司及其关联机构（以下统称“本公司”）对这些信息的真实性、准确性及完整性不作任何保证，也不保证所包含的信息和建议不发生任何变更。本公司没有将变更的信息和建议向报告所有接收者进行更新的义务。

本报告仅供本公司的客户作参考之用。本公司不会因接收人收到本报告而视其为本公司的当然客户。

本报告仅反映报告作者的出具日的观点和判断，在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，投资者应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求。对依据或者使用本报告所造成的一切后果，本公司及/或其关联人员均不承担任何法律责任。

本公司的交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。本公司没有将此意见及建议向报告所有接收者进行更新的义务。本公司的资产管理公司、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

本报告版权均归本公司所有，未经本公司事先书面授权，任何机构或个人不得以任何形式复制、发布、传播本报告的全部或部分内容。经授权刊载、转发本报告或者摘要的，应当注明本报告发布人和发布日期，并提示使用本报告的风险。未经授权或未按要求刊载、转发本报告的，应当承担相应的法律责任。本公司将保留向其追究法律责任的权利。

浙商证券研究所

上海总部地址：杨高南路 729 号陆家嘴世纪金融广场 1 号楼 25 层

北京地址：北京市东城区朝阳门北大街 8 号富华大厦 E 座 4 层

深圳地址：广东省深圳市福田区广电金融中心 33 层

上海总部邮政编码：200127

上海总部电话：(8621) 80108518

上海总部传真：(8621) 80106010

浙商证券研究所：<https://www.stocke.com.cn>