

· 综论 ·

人工智能时代的社会保障：理念更新 与制度建设

蔡 昉

[摘 要] 人工智能技术的颠覆性在经济上表现为创造性破坏，具有显著提升生产率、冲击就业和扩大收入不均等的双刃剑效应，以及加重结构性就业矛盾和过大收入差距的风险。这相应提出制度需求，迫切要求公共就业服务更加积极、社会保障制度更加普惠。鉴于传统社会保障制度面临对象识别、制度适应和技术红利共享等方面的难题，完善社会保障制度的方向和路径，应是从更接近于“剩余型”向更接近于“制度型”的转变。相应地，公共财政特别是社保筹资模式也需要创新理念，探索以“先分配，助增长”为特征的跨代收支平衡模式，以借助人工智能所达到的更高生产率支撑民生建设。这包括从“对物投资”转向更多“对人投资”，提高政府社会性支出占 GDP 的比重；应对结构性就业矛盾，以普惠保障推动实现老有所养；通过社保兜底消除居民消费的后顾之忧，释放强大内需潜力。

[关键词] 人工智能；创造性破坏；制度需求；普惠性社会保障

一、引言

人们惊叹于近年来人工智能发展和应用的速度，并且越来越多地加入到相关的讨论中：一方面，尝试回答人工智能将为人类创造怎样的美好未来；另一方面，警示人工智能可能给人类社会带来怎样的风险。人工智能作为一门科学，其起点通常追溯至 20 世纪 50 年代，以艾伦·图灵在 1950 年发表该领域开创性论文《计算机器与智能》^① 以及约翰·麦卡锡等科学家在 1956 年的“达特茅斯会议”上首次提出“人工智能”（artificial intelligence）这个概念为标志。然而，按照李开复的说法，在此后的半个多世纪里，人工智能只是部分科学家圈内的工作和话题，长期未见标志性的突破，因而也鲜为大众所关注，以致重大的转折直到 2016 年才发生，以人工智能公司 DeepMind 开发的 AlphaGo 在围棋对局中击败世界冠军李世石为标志。^② 随后，新的转

[作者简介] 蔡昉，中国社会科学院原副院长、学部委员、高端智库首席专家。主要研究方向：人口与劳动经济、收入分配与社会保障等。

① Alan M. Turing, "Computing Machinery and Intelligence," *Mind*, 1950, LIX(236).

② 参见李开复、陈楸帆：《AI 未来进行式》，浙江人民出版社，2022 年。

折性事件便接连不断出现，其中最当之无愧的便是 2022 年 OpenAI 推出 ChatGPT，以及随后包括 DeepSeek-R1 在内的其他大语言模型如雨后春笋般涌现，由此人们第一次直观感受到人工智能技术的颠覆性。

与此同时，作为人工智能发展条件或应用领域的信息、计算、互联网、数据等技术和基础设施也获得加速发展，与人工智能形成相互促进和相互赋能的关系，使几乎所有的科技领域都预期将迎来颠覆性突破。首先，人工智能现有各种模型的应用前景，正在通过各种智能体形式呈现出来，并且，随着通用型人工智能（artificial general intelligence 或简称 AGI）的到来，人工智能技术及其应用将从目前的水平进一步跃升。其次，在现有的数字技术与产业相结合模式的基础上，人工智能的新赋能将把数字经济推到一个新的高度，并创造和塑造新的产业模式和形态。再次，这场科技革命不仅关乎国家现代化的进程，还关乎经济整体和企业的竞争力，也同家庭和个人息息相关，其影响正在渗透于全社会诸多方面。其中，诸如人工智能对生产率、就业、民生的影响都面临着“奇点来临”，且具有双刃剑性质，必须予以高度关注和积极应对。最后，中国在全球人工智能发展和应用的竞赛中处于第一梯队，在人工智能革命中，既能获得巨大收益，也负有大国的责任，应该在自身探索和国际合作的基础上，成为人工智能治理中公共品的供给者。

科学家和思想领袖们对人工智能发展进行了多种展望，分别从不同视角试图对大众予以启发、对发展进程产生影响，并且就人工智能的重要方面和关键领域提出政策建议。例如，以当下为起点，前引李开复的著作以 2042 年之前为参照期，展望了人工智能的主要突破方向和应用领域；凯文·凯利的著作以 2049 年之前为参照期，畅想人类如何迎来一个由技术与创造力驱动的高科技社会；^①戴维·奥托等人则基于以往对技术与就业关系的研究，围绕如何避免人工智能技术对就业的不利影响，提出关于政策及其实施路径的建议。^②布林约尔松及合作者则为人工智能产生影响的三个重要方面，即生产率增长、劳动力市场和产业集中度分别指出了岔路口，由此可以走向正道或者歧途，从而产生大相径庭的结果。^③

不过，任何对人工智能未来的展望，都只是站在一个已知的知识基点上。鉴于人工智能技术演变之快，任何知识基点都在不断发生意想不到的变化，随后的技术发展方向和路径必然是变化多端的。此外，不同国家的技术起点不一样，人工智能发展所处的外部环境也有差别，因而也可能导致不同的演进路径。特别是，体现于各种模型、平台和智能体的人工智能技术和系统，都可能给人类的生产、生活乃至生存带来巨大风险，其中有一条可能的路径，会通过破坏就业对民生福祉造成伤害。鉴于这类风险不可能完全做到事前和过程中控制，旨在引起关注并尝试

① 参见凯文·凯利：《2049：未来 10000 天的可能》，中信出版集团，2025 年。

② 参见戴维·奥托等：《AI 时代的工作》，中信出版集团，2025 年。

③ Erik Brynjolfsson, Gabriel Unger, "The Macroeconomics of Artificial Intelligence," *Finance and Development*, 2023, 60(4).

直接参与风险控制的第三方行动，通常也不能左右人工智能发展的主流路径，因而都难以起到杜绝风险的作用。因此，体现民生本位的社会保障新理念和新体系，应该成为因应人工智能发展的一项重要制度建设任务。

给定诸多的不确定性，对未来的预想与其拘泥于细节，不如更加粗犷一些，针对更大的颗粒度进行观察，反而能够在更宏观的层面、以更体现事在人为的原则对未来方向做出展望。况且，对于非人工智能专家而言，更适合于着眼于所熟悉的特定领域，对人工智能最可能带来的积极影响，以及发生概率较大的风险做出分析和判断。本文将主要关注人工智能对就业的挑战，以及社会保障制度的可能应对路径。人工智能的发展和应用，将把人类对未来经济可能性的预期无数倍地放大。与此同时，像以往任何一次科技革命一样，人工智能革命也是双刃剑，不仅将伴随着创造性破坏，更有可能的是，这个新科技革命将具有史无前例的“破坏”力度，同时也能为解决问题提供前所未有的强有力手段。人类有史以来赖以谋生的工作，将在这场创造性破坏中首当其冲。一方面，就业创造和就业破坏之间的平衡关系究竟如何，取决于政策选择及其所决定的技术发展方向和路径。另一方面，就业创造和就业破坏不仅必然同时发生，并且往往因不可预期而难以避免。从底线思维着眼，从宏观政策入手，应该立足于在三个方面筑起“防波堤”，即人工智能与社会规范对标、劳动者能力建设以及社会保障的保护，以系统性、建设性的方式应对人工智能的就业挑战。

作为一种事前的制度保障，应该着眼于通过规制、激励和引导等方式规范人工智能的开发与应用，最大限度使就业创造既不少于也不慢于就业破坏。达龙·阿西莫格鲁把人工智能界定为两类，分别是增强劳动者能力的“正确的人工智能”，以及替代劳动者工作的“错误的人工智能”。^① 诸如此类在政策设计中需要进行方向引导的问题，在人工智能研究中被称为“对齐问题”（alignment problem）。或许，将其翻译为“对标”更符合问题的性质，因为这里的“对标”并非仅仅针对人工智能本身，让其去适应某种抽象标准，而是通过各种制度安排，将规制、激励和劝导机制作用于这项技术的开发者、投资人和使用者。^② 例如，一位研究者指出，“技术正在改变工作场所”这个说法，实际上掩盖了一个事实，是首席执行官而不是技术本身，不断做出自动化和减少岗位的决定。^③ 不过，即便在这个方面的努力差强人意，也还不足以保障人工智能必然走在“正确的”轨道上。

因此，我们还需要建造另外两个重要的防波堤。在个人层面，通过教育和培训培养人类劳动者迄今尚有相对优势的人力资本，以抵御人工智能替代就业的冲击，并与智能体建立起协作

① Daron Acemoglu, Pascual Restrepo, "The Wrong Kind of AI? Artificial Intelligence and the Future of Labour Demand," *Cambridge Journal of Regions, Economy and Society*, 2020, 13.

② 蔡昉：《引领人工智能创造更多更高质量就业》，《劳动经济研究》2025年第3期。

③ Guru Madhavan, "It's Time to Retire the Word 'Technology'," *Financial Times*, 30 July 2025.

关系。这需要教育和职业培训体系做出相应调整，形成人力资本培养新模式。^①在社会层面，必然需要做好社会兜底保障，以之作为最后的制度屏障，确保民生不受到实质性冲击。人工智能的发展及其在提升生产率乃至局部替代人类劳动方面的表现，终究要服从于人的全面发展这个目标。我们面临着一个千载难逢的历史机遇，发展的资源制约最终可以被打破，然而，只有建立起分享的制度和机制，才能使社会均等地享受人工智能红利。就以上三个方面来说，我们仍然面临着很多研究课题和广阔的探索空间。本文将把分析重点放在如何发展和完善社会保障，以有力应对人工智能挑战这一问题上。

二、从人工智能视角重新认识社会保障

诚然，以规制、激励和劝善这样的方式引导人工智能对标人类意图，是政策制定和实务推动所不可或缺的。然而，从理论和经验来看，也有诸多理由使我们不能有足够的把握认为，只需做出这些努力便能如愿以偿地影响人工智能发展，从而可以一劳永逸地解决所关注的问题——譬如避免人工智能带来就业破坏大于就业创造的结果，在数字鸿沟的基础上进一步形成智能鸿沟，以至造成劳动力市场两极化，进而扩大收入差距等等。与此同时，人工智能发展方向也不是可以预先决定的，劳动者的何种技能会在何时、何种程度上被替代也是不确定的。鉴于人工智能在经济社会发展中的积极作用，既不能指望任何“涓流效应”自动开启，也不能完全依赖任何单一调控手段予以发动，因此，我们需要建立一个作为终极社会保障的制度体系，确保生产率得到分享、劳动者受到保护。从以下几个方面，我们将揭示人工智能表现出的特性，并据此论证为什么建立健全普惠性社会保障制度，是应对人工智能就业冲击的题中应有之义。

首先，创造性破坏是一个整体，没有“破坏”就没有“创造”。如同任何先进技术一样，人工智能被赋予了增进民生福祉的终极期望，但技术和福祉终究不是同一个东西，两者之间存在着动机差别和功能边界。例如，人工智能提高生产率和促进经济增长的功能，归根结底在于其创造性破坏性质，创造性与破坏性本质上是一枚硬币的两面。对于理解这种相反相成的性质，以下两项研究成果具有一定的启发性。其一，发表在《科学》杂志的一篇论文介绍的研究结果，揭示出在自动化与增强就业能力之间，并非一定是对立的关系，反而常常呈现互为条件的关系。也就是说，在该研究所考察的领域中，实现自动化通常是提高低技能群体劳动生产率，进而缩小整体收入差距的前提条件。^②其二，基于在50个国家的微观调查数据，普华永道在一份以《希望与忧虑》为题的报告中展示：一方面，“最受人工智能影响的产业”与“最不受人工智能影

① 蔡昉：《面向未来的人力资本培养模式》，《学习时报》，2025年6月6日第1版。

② Ajay Agrawal, et al., "Do We Want Less Automation? AI May Provide a Path to Decrease Inequality," *Science*, 2023, 381(6654).

响的产业”相比,每个雇员创造的收入高出三倍;另一方面,将更多受到影响的岗位与较少受到影响的岗位进行比较,前者的增长率比后者低27个百分点。^①

从宏观经济层面,特别是结合中国的情况来看,包括节约劳动力在内的减少要素投入功能、缩减低生产率行业份额、推动低生产率企业退出乃至破产、通过市场结清消除过剩和无效产能,都是作为可持续驱动力的生产率得以提高的必由之路。人工智能的强大赋能性,也恰恰体现在将这些过程成倍地加快。与此同时,这些提高生产率的途径也是创造性和破坏性的统一,不可避免地造成对就业的冲击和对部分劳动者的伤害。因此,合理的制度安排,就是要确保创造性和破坏性之间不会成为一种零和博弈,不会表现为“不是东风压倒西风,就是西风压倒东风”的非此即彼关系。除了通过促进人工智能发展和应用的对标、增强劳动者应对挑战的人力资本禀赋,以最大限度保持创造性和破坏性之间的合理平衡外,还需要筑牢制度性屏障,防止创造性破坏造成对人本身的“破坏”。

其次,天底下没有所谓“涓流效应”,而是往往产生“合成谬误”。基于科技进步与经济社会发展之间关系的经验,我们可以做出这样的判断,人工智能发展与民生福祉提高之间,隐含着某种微妙的关系。一方面,人工智能发展的最突出特征,是不会自然而然带来程度上相等、节奏上同步、时空上并存的民生福祉效应。经济学家们对所谓“涓流经济学”的批评已形成越来越多的共识。^②另一方面,一种常见的经济社会现象很可能将在这里大行其道,即当事人看似理性的分散行为造成“合成谬误”。人工智能前所未有的进步速度和颠覆性质并非免费午餐,而是伴随着巨大投资,包括人才、资金、能源、基础设施的大规模投入,这使技术演进成为一场激烈的竞争。

对主要国家政府来说,人工智能发展成为占领制高点的国家间竞赛,以及形成经济增长点的关键抓手,以致给予大量补贴和其他政策优惠。对投资者来说,大规模投入该领域毕竟不是出于慈善或公益,而是预期高回报的风险投资。对技术创新企业来说,能否在这个领域占据先机,不仅关乎竞争力的位次,而且似乎成为生死存亡分水岭。对应用技术的经营主体来说,其出发点自然是减员增效、降本增盈。上述所有当事人的动机和追求不仅具有极大的力度,还表现为:无论出于任何理由选择观望,皆不会得到伙伴企业和竞争对手的响应,而只会让自己落在后面。既然市场主体不可能主动去控制人工智能潜在风险和冲击效应,政府规制和调控举措则不能缺位或滞后。

最后,人工智能的就业冲击性质表明,识别社会保障受益人的难度日益凸显。在理查德·蒂特马斯所框定的剩余型和制度型社会福利模式两极之间,或者在哥斯塔·埃斯平-安德森概括的三种福利国家模式之中,不同社会保障制度的独特之处,通常表现在保障和覆盖原则的普惠

① Anjali Raval, "Disrupted or Displaced? How AI Is Shaking Up Jobs," *Financial Times*, 9 June 2025.

② 参见蔡昉:《成长的烦恼:中国迈向现代化进程中的挑战及应对》,中国社会科学出版社,2021年。

性差异上。^①也就是说，在那些更靠近剩余型模式的社会保障体制下，制度的实施需要建立在对保障对象的资格进行严格识别的基础上。一旦涉及与就业相关的社会保险和其他保障或福利待遇，头等重要的就是确认对象是自愿放弃就业，还是缺乏获得或接受可能就业机会的必要条件。换句话说，社会保障制度力图保障真正陷入困境的人群，而尽量避免“养懒汉”。然而，在人工智能时代，这种模式的运行条件将发生根本的变化。

在人工智能赋能的自动化技术作用下，劳动者与工作岗位、社会保障之间的关系，正在并将继续发生颠覆传统观念的变化，这种变化要求对认识范式和政策导向做出相应转变。一个日益显现出来的现象是，社会作为一个整体越来越难以预先判断：何种劳动者技能可以同人工智能形成互补关系，从而可以预期由新技术予以增强；何种劳动者技能与人工智能恰好形成竞争关系，从而使掌握该技能的劳动者成为岗位被人工智能替代的对象。可见，劳动力市场上的人力资源错配，不是任何劳动者个人的过错。并且，劳动者一旦从遭遇就业冲击的经历中懂得这个道理，理性的选择便是不断接受培训和再教育，有时甚至需要一些时间观察和等待。对社会而言，更好的方式则是以更加普惠的方式对此给予兜底保障，允许受冲击劳动者试错和等待。概而言之，如果劳动者个人的就业意愿、失业原因和劳动参与状态均无法得到确定性的识别，社会保障制度则应该更具有普惠性，以便为劳动力市场行为留出更充分的余地。

三、中国劳动力市场表现及其制度需求

伴随着技术进步和自动化而来的创造性破坏过程，应该放在中国劳动力市场演变的大背景下，予以具体和动态的阐释。改革开放初期后的前30年的中国经济增长，如果借助发展经济学家阿瑟·刘易斯的理论模型予以刻画，恰好处于二元经济发展阶段；如果用人口红利理论和相关经验研究来定位，则可以说处于人口机会窗口期。如图1所示，在1978年改革起步之前，生育率和人口出生率都很高，也不存在老龄化问题，因此，16—59岁劳动年龄人口与其他年龄人口的增长率大致相当，中国虽然具有劳动力丰富的特征，人口抚养比（非劳动年龄人口与劳动年龄人口的比率）却并不低，人口结构的生產性也不十分显著。在1978—2010年期间，劳动年龄人口始终保持着快速增长，年均增长率为1.90%，同期非劳动年龄人口规模增长几乎停滞，年均增长率仅为-0.12%。两个人口年龄组的不同变化趋势构成一个剪刀状，对应着人口抚养比的下降，形成人口机会窗口期的典型特征——“生之者众、食之者寡”。

生育率下降和人口出生率、增长率减慢的效应，以老龄化的方式逐渐显现出来，并导致劳动年龄人口峰值的到来。2004年前后，中国经济以遭遇民工荒为标志，跨过了标志性的“刘易

^① 参见理查德·蒂特马斯：《社会政策十讲》，吉林出版集团，2011年；哥斯塔·埃斯平-安德森：《福利资本主义的三个世界》，商务印书馆，2010年。

斯转折点”，及至2010年之后，以各种统计口径反映的劳动力资源陆续达到峰值并开始负增长。这标志着中国经济发生阶段性变化：在不同地区和各类行业中，劳动力短缺渐渐成为常态，普通劳动者工资也随之上涨，意味着劳动力无限供给特征的式微，以及人口红利期的结束。^① 这些变化相应驱动了更快的自动化即资本替代劳动的过程，集中表现在工业特别是制造业资本劳动比的提高，以及就业增长速度明显减慢。例如，在2008年第二次全国经济普查和2023年第五次全国经济普查之间，全部工业企业每个从业人员对应的名义总资产额，从39.4万元提高到175.9万元，提高了3.5倍；在全国法人单位从业人员增长52.4%的同时，制造业法人单位就业人数却减少了0.5%，因而制造业就业比重从38.4%降低到25.1%。这与制造业产值比重的变化趋势也是一致的。例如，世界银行数据表明，中国制造业增加值占GDP的比重，从2006年峰值时的32.0%显著下降到2024年的24.9%。

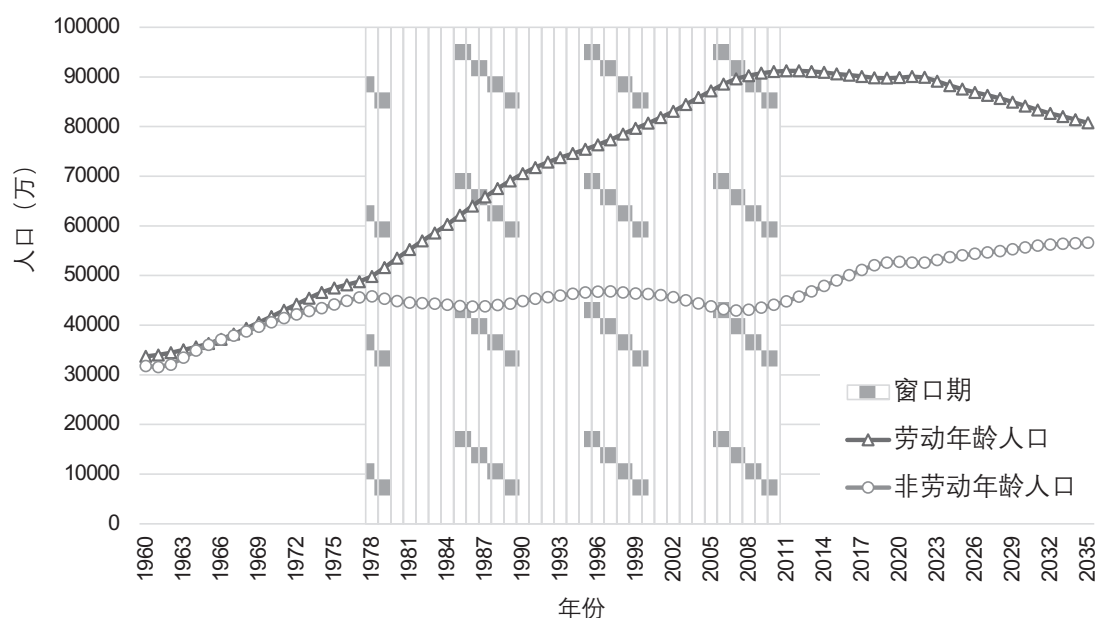


图1 人口机会窗口的开启和关闭

资料来源：United Nations, Department of Economic and Social Affairs, Population Division, *World Population Prospects 2024*, Online Edition, 2024.

在农村劳动力仍在向城镇非农产业转移，从而城镇就业整体上继续扩大的背景下，中国就业结构发生了巨大变化。在2005—2024年期间，第一产业就业比重从44.8%降低到22.2%，第三产业就业比重从31.4%提高到48.8%，第二产业就业比重则从2012年30.5%的峰值水平，小幅降低到2024年的29.0%。对于就业的这种相对份额变化，人们习惯于从结构变化的分析范式和观察角度来认识。然而，在观察技术变化的就业影响时，一个同样有用且更有针对性的框架，则是从就业创造和就业破坏的关系，对就业结构展开阐释。经济学家始终认为，技术进

① 蔡昉：《刘易斯转折点——中国经济发展阶段的标识性变化》，《经济研究》2022年第1期。

步既通过机器和自动化替代人的工作，造成就业破坏的结果，也通过新需求诱致出新岗位，导致就业创造的结果。真正的问题在于，由于发展阶段、技术导向、政策取向及举措等诸多不同，在就业创造与就业破坏之间，岗位的相对规模、时间契合程度以及岗位类型也分别有所差异。以最简化的方式来概括，包括人工智能在内的技术应用通常表现为三种类型或阶段，分别经历程度不同的就业冲击：即就业创造大于、等于或小于就业破坏的情形。

中国城镇就业的宏观数据，可供我们进行这样的观察与分析。政府每年正式公布的“城镇新增就业”是一个城镇就业的毛增量，可以被作为就业创造的代理指标。同时，我们还可以从年度统计数据得到“当年净增就业”，相当于是就业净增量。净增量与毛增量之间的差，即可作为就业破坏的代理指标，换句话说，如果这个差为负数，就意味着就业创造的一部分被就业破坏抵消掉了。鉴于在我们所考察的期间，劳动力市场发生了巨大变化，就业数据的统计方法及含义在不同年份未必相同，所以这里并无必要拘泥于具体数值及其含义，只是借助相关数据鸟瞰总体趋势，以支持本文对劳动力市场变化的基本判断。

如图 2 所示，在 2000—2012 年期间，中国城镇劳动力市场恰好处于就业创造的高峰期，全部就业创造都转换为历年净增城镇就业，累计数量高达 1.49 亿，可以说不存在净就业破坏。2012 年以后，虽然净就业破坏开始出现且规模呈逐年扩大的趋势，但总体来说就业创造仍然大于就业破坏。在 2012—2024 年期间，劳动力市场上累计创造了 1.67 亿个城镇就业岗位，仅有 5344 万个岗位被破坏，每年净增就业的累计数为 1.13 亿。由此可见，中国城镇劳动力市场已经不再处于就业创造大于就业破坏的阶段，大体上处于就业创造与就业破坏相对均衡的状态。当然，从此时点出发，随着人工智能就业影响的加剧，就业破坏占主导的阶段不可避免会来到。

作为一种劳动力市场结果，就业破坏现象的出现和凸显，必然对应着不利的收入分配状况。包括戴维·奥托和达龙·阿西莫格鲁在内的经济学家，都通过研究表明，自动化对工人的替代，通过资本与劳动之间的报酬差距，以及高端人力资本回报与低端岗位工资之间的差距，导致劳动力市场两极化、工资中位数增长缓慢，以及总体收入差距扩大等结果。更具体而言，阿西莫格鲁提出了就业替代的相应假说，即由于中间年龄劳动者相对短缺，自动化最先针对这一群体，形成自动化影响程度的倒 U 形曲线。^① 奥托则提出了工资水平及其增长的 U 形曲线，即由于劳动者从制造业被分流，低技能劳动力向服务业岗位转移造成就业极化，可以同自动化相协作的高技能劳动者收入呈现增长态势，造成工资极化，而中间层次被替代的劳动者则经历工资停滞。^② 两个假说具有内在的逻辑一致性，两种现象也具有警示意义。这些结论也适用于分析人工智能

① Daron Acemoglu, Pascual Restrepo, "A Task-based Approach to Inequality," *Oxford Open Economics*, 2024, (3).

② David H. Autor, David Dorn, "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 2013, 103(5).

的就业影响,或者说,作为一种“特征化事实”,^①无论由何种技术赋能的自动化,并不会因生产率提高而自然惠及所有人群,反而可能导致收入差距的扩大。

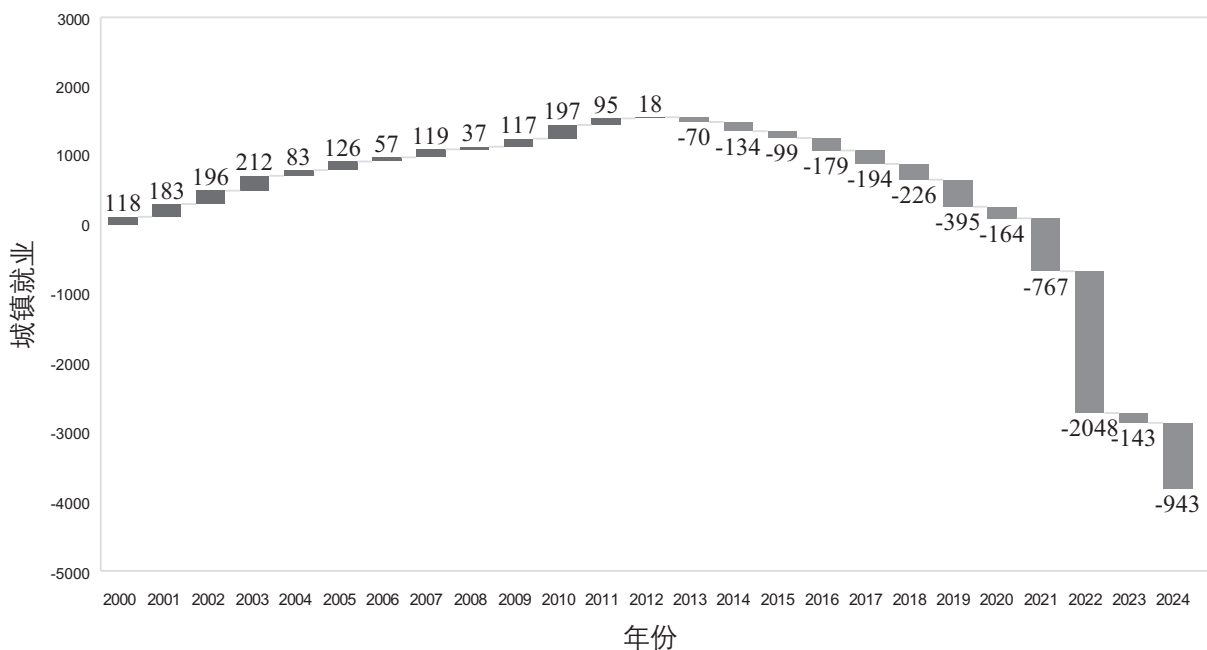


图2 就业创造与就业破坏

资料来源:作为计算依据的“城镇净增就业”(城镇就业的年度之差)来自国家统计局“国家数据”: <https://data.stats.gov.cn/easyquery.htm?cn=C01>;“城镇新增就业”分别来自历年《政府工作报告》或统计公报。

注:图中标出数字的框图为岗位创造(正值)和岗位破坏(负值)的年度增量,图形在纵向上显示的绝对幅度为岗位创造或岗位破坏的累计水平。

参照上述理论逻辑和实证经验,我们现在来观察伴随着劳动力市场演变过程,中国收入分配状况的对应轨迹及可能趋势。在基尼系数于2009年、城乡收入差距于2007年达到峰值之前,这两个反映收入差距的指标均呈现出增加的趋势,并且由于城乡收入差距是基尼系数的重要组成部分,两个指标始终保持着同步变化(图3)。在2015年基尼系数降低到0.462的谷底水平后,一方面,基尼系数处于不进不退的徘徊状态,另一方面,基尼系数与城乡收入差距的变化不再同步。换句话说,城乡收入差距的继续降低未能带动基尼系数缩小,这个事实说明在基尼系数这一收入不均等指标的三个组成部分(即农村内部收入不平等、城镇内部收入不平等和城乡收入差距)中,前两项的相对贡献分量显著增大。根据联合国大学发展经济学研究院的数据,^②在分别以基尼系数和帕尔马指数反映的总体收入不平等程度、农村收入不平等程度和城镇收入不平等程度中,总体不平等程度显著大于农村或城镇内部不平等,说明总体收入不平等在很大

① 我把与此相关的几个研究结论称为“阿西莫格鲁事实”。参见蔡昉:《人工智能时代的就业挑战》第二章“理解就业挑战的性质”,中信出版集团,2025年,第57-63页。

② 参见 UNU-WIDER, *World Income Inequality Database (WIID)*, Version 30 June 2022, <https://doi.org/10.35188/UNU-WIDER/WIID-300622>.

程度上由城乡收入差距导致。在总体收入差距越过峰值之后，该指数与城镇不平等指数的差异趋于缩小，意味着城镇的收入不平等程度在加大，成为总体收入不平等的更主要驱动力。^①造成这种情况的原因，越来越同就业破坏、结构性就业矛盾、就业非正规化，以及社会保护不完善有着密切的关系。



图3 收入差距变化趋势

资料来源：国家统计局“国家数据”：<https://data.stats.gov.cn/easyquery.htm?cn=C01>；2003年之前的基尼系数的数据取自 Fang Cai, *China's Economic Growth Prospects: From Demographic Dividend to Reform Dividend*, Cheltenham, UK, Edward Elgar, 2016, 180.

中国劳动力市场和收入分配格局发生的变化，诱致出新的制度需求，由此所要解决的问题具有必然性，既迫在眉睫又将长期存在。这些问题包括：在劳动力市场主要矛盾转化之后，如何更有针对性地应对结构性就业矛盾；如何推动全体劳动者均等分享技术进步带来的更高生产率；在充满风险的技术变革环境中，如何切实保障劳动者权益及民生福祉等。回应这些问题，需要在新发展理念的统领下，加深对人工智能发展规律的认识，恰当地确定制度建设的出发点和落脚点。首先，营造创造性破坏与社会保护相统一的环境。既要通过充分进入、自由退出和优胜劣汰机制，让市场主体有充分的激励，确保技术被用来提高整体生产率，又要通过社会保障、劳动力市场制度、再分配制度和公共就业服务，确保在创造性破坏过程中劳动者不会被置于边缘化地位，使最终发展结果得到充分的分享。其次，制度建设先立后破、改革驱动。旨在促进基本公共服务均等化和人口、劳动力合理分布的户籍制度改革，应该成为这一系列制度建设的

① 蔡昉：《破除城乡二元结构》，浙江人民出版社，2025年，第203页。

突破口。最后，社会保障制度体现新理念。人工智能时代的最大机遇，是生产率以前所未有的幅度提高，最大挑战则是如何实现生产率的均等分享。应对挑战的制度建设，将体现在一个更具普惠性的社会保障制度中。

四、如何增强社会保障的普惠性？

社会保障制度的演进路径，就是一个保障水平持续提高、普惠性不断增强的过程。对于普惠性增强这个重点，学术界和政策研究界在很长时间里未予足够的重视，以致在一定程度上使制度完善缺失了一个重要的维度。从中国的社会保障体系，特别是很多社会保险项目的建设实践来看，迄今取得的非凡成就，都是经历从无到有、由小到大跨越式发展的结果。实际上，社会保障水平和均等化程度的提高，都是以增强普惠性的方式得以实现。以城乡居民养老保险制度为例，虽然这个项目并不是一种无需缴费或覆盖全民的社会养老金制度，保障水平也有待显著提高，但是，该项目所具有的制度特点颇值得关注，譬如说很低的缴费门槛、符合年龄要求人群的制度性全覆盖、改革决策和制度的建立健全均属于中央政府事权等等，恰恰都体现了较强的普惠性。此外，与此类似的城乡居民医疗保险、覆盖城乡的最低生活保障等制度，以及失业保险金通过用于公共就业服务和“稳岗援企”的方式惠及更广泛劳动者群体等做法，也都是社会保障普惠性增强的体现。

上述同一个命题，往往还表现为社会保障与社会福利两种制度的异同。关于这两种制度体系，从以下两个层面来理解，可以获得一定的启发。纯粹从理论上讲，社会保障更接近于剩余型模式，社会福利更接近于制度型模式，因而两者可以被看作不同（层次）的制度形式。从制度运行和政策实施的角度来看，实践中也形成了两种制度模式彼此相异的一些做法。不过，在制度的实施和演进中，两者之间的差别通常只是剩余型和制度型性质的程度差别，或者只是叙事意义上的差别，犹如在“七分剩余、三分制度”模式与“三分剩余、七分制度”模式之间的差别。中国的社会保障是广义的制度体系，包括社会保险和社会救助等诸多内容，也蕴含着社会福利的诸多基因。不过，从方法论的意义上，把社会保障与社会福利这种目标大致相同、功能有所差异、手段常有不同的模式，作为制度光谱的两极和参照系，可以帮助我们获得一个坐标系，以便在比较中看清社会保障制度演进中的普惠性方向和对应的途径。

从这个意义上说，虽然社会保障与社会福利两种制度具有根本的共同之处，即以社会共济的方式达到社会保护的目标、以社会统筹的方式分享发展成果，我们也可以看到两者之间的一些明显不同。首先，依据的制度理念和原则存在一定差异。社会保障仅对界定的人群最基本需求做出保障承诺，社会福利则根据社会必要水平，确定一个关于人的基本权利的普适标准，无差别地做出社会福利安排。所以，社会福利提供的公共服务范围要显著大于社会保障，旨在满

足生命周期各个阶段的需求。同时，社会福利的保障水平也更高，且伴随着生产率不断提高而水涨船高。其次，提供公共服务的方式有原则性差别。很多社会保障项目，通常需要以生计调查或类似的方式识别对象，据以保障“符合条件”的人群。社会福利制度则以普惠的方式，以全体国民或具有某一共同特征的人群整体为对象提供公共服务。最后，财政理念和预算纪律观不同。社会保障制度更倾向于奉行传统财政理念，追求增收节支、以收定支和当期平衡，每逢遭遇外生性冲击，往往还会主动提出或被迫接受紧缩要求。社会福利体系则从更长期的维度考量财政“平衡”，既看重支出对于社会保护的意义，又看重支出对人力资本培养的作用，着眼于更好的经济表现和水涨船高的财政收支。

鉴于很多国家因削弱福利体系的实践而付出惨痛代价，经济学界和政策研究领域开始反思乃至重建福利国家理论。约瑟夫·斯蒂格里茨做过权威性的综述，在一定程度上可以代表主流经济学家的共识，也可谓经济理论在这方面的最新发展。^①首先，福利国家以制度化的形式，向国民提供具有共济、互惠、诚信和团结性质的公共品，弥补市场失灵，维护社会公平正义。其次，福利国家通过提供必要的保障和激励，鼓励个体承担风险和从事创新，并为每一个人享有社会必要的生活质量提供全生命周期支持，以制度化的方式协调代际合作。最后，信仰、偏好和习惯是由超越市场范围的诸多因素，譬如社会规范和非正规制度所决定的，因此，社会福利并不会削弱受益人的责任感和工作努力程度。福利国家自诞生以来，历时已久并经历种种变化和起落反复，各国在不同时期分别积累了成功经验，也付出了失败的代价。总体来说，当今世界的社会保障或社会福利，已经超越剩余型模式，在不同程度上接近制度型模式。

对中国来说，面对现实存在的和预期出现的挑战，这方面的制度建设仍然任重道远。首先，社会保险的给付水平和均等化程度，仍然不能充分满足日益增长的城乡居民需要。一方面，不同人群参加的社会保险项目之间待遇差距过大。另一方面，很多社会保险项目尚未实现应保尽保。其次，社会保险的普惠性仍然不强，仍然基于各种既定制度方式识别受益对象，提供排他性或差别性待遇。这种做法不仅在人工智能时代日益变得没有意义，也不符合全体人民共同富裕的现代化的理念。最后，虽然社会保障制度建设的成效显著，但现行的推进速度仍然属于常规态，因而与人工智能非常规演进的速度不甚匹配，从而不足以应对相应的就业和民生冲击。人工智能预期带来的巨大生产率提升，终究可以打破资源的绝对约束，为构建更高水平福利国家提供坚实的物质基础。在此基础上，中国的社会保障制度，将接受更多的社会福利理念和做法，从增强普惠性和覆盖全生命周期着眼，为全体居民提供更为广泛、充分和均等的基本公共服务。

无论是采取低调的表达方式和渐进的过渡方式，实质性转向更加普惠性的基本公共服务体

① 约瑟夫·斯蒂格里茨：《21 世纪的福利国家》，《比较》2023 年第 1 期。

系,还是高调宣示以建设中国特色福利国家为目标,对于国家财力能否支撑、如何筹资等问题,都不能避而不答。在进一步分析之前,在此有必要预先强调一个两点论:一方面,增强社会保障制度的普惠性,或以更多的社会福利特征代替社会保障特征,无疑是制度建设的方向,也是目前应该探讨的重要话题;另一方面,这并不意味着可以淡化“尽力而为、量力而行”的原则,只不过在更高的经济发展水平上,以及在人工智能显著提高生产率的预期下,政府工作的“尽力”应该与日俱增,财政能力的“量力”也到了水涨船高的时刻。

更大规模的投入,并非总是表现为资金总量的增加,在更多的情况下更应该是支出结构的调整。这里涉及的做法便是从“对物投资”更多转向“对人投资”。从民生需要和国际比较来看,中国政府的财力和支出结构仍然有调整的余地。根据中国近期人均 GDP 跨过世界银行定义的高收入国家门槛(14005 美元),进而在 2035 年成为中等发达国家(27000 美元)的发展目标,我们可以把人均 GDP 处于 12000—30000 美元的国家作为参照系,就政府总支出、社会性支出占 GDP 比重进行比较。2020 年,中国政府支出占 GDP 的比重为 33.9%,相当于参照国家算术平均值的 83.9%;同年中国政府社会性支出占 GDP 的比重为 17.8%,仅相当于参照国家平均水平的 70.9%。^① 研究还表明,对中国经济的物质资本投入已经面临持续的报酬递减,而对人的投资将在长期内保持很高的回报率。^② 更多、更高质量的社会福利供给,本质上相当于国家以制度形式,将面临报酬递减的物质资本转化为具有报酬递增性质的人力资本。

以全民分享生产率提高收益为目标的社会福利制度,如同公共财政制度一样,处理好收入与支出的平衡始终面临着两难的抉择,也是关于增长与分配关系的永恒课题。发展经济学家曾经概括出两种增长与分配关系模式。一是“先增长,后分配”模式,将资本积累和投资激励置于优先地位,主张社会忍受一段时间的收入不均等,在社会变得足够富裕时再实施再分配。二是“边增长,边分配”模式,主张在经济发展创造出社会财富增量的同时,即通过社会福利制度进行再分配,以随时随地缩小收入和基本公共服务可得性之间的差距。后一种模式因其增长与分配的同步性,被认为有利于实现公平与效率的统一。与此同时,长期以来公共政策理念都采取谨慎小心的态度,防止“边增长,边分配”模式超越应有的限度,以致演变成为一种“先分配,助增长”模式,陷入寅吃卯粮境地,甚至陷入民粹主义政策陷阱。可以说,这种谨慎态度,特别是对“先分配,助增长”模式的否定,长期以来都有着充分的理由。

然而,至少由于两个重要的变化以及由此发生的新情况,对于这类问题日益需要重新认识,换句话说,我们亟待进行一个基于范式转换的理念更新。第一个变化是,中国经济增长的主要制约因素已经从供给侧转到需求侧,居民消费日益成为增长的决定性拉动力,而后者

① 参见蔡昉、贾朋:《构建中国式福利国家的理论和实践依据》,《比较》2022 年第 3 辑。

② 吴舒钰等:《教育回报率显著高于资本回报率的实证分析——基于 1978 年以来我国宏观教育回报率的测算》,《教育研究》2024 年第 3 期。

的提振有赖于更具普惠性的社会保障。第二个变化是，人工智能的发展不啻为我们带来生产率大幅提高的极大可能性，预先分配这个更大生产率蛋糕的必要性，可以得到经济理论的支持。对预期未来将出现的极高生产率和充分涌流的物质财富进行预先分配，在逻辑上相当于对后代的负债。正如瑞·达里欧所指出的那样，“一个实体的债务是另一个实体的资产”，^①因此，认为政府长期负债意味着把债务留给子孙后代的看法并不准确，对此忧心忡忡更无必要。也就是说，如果以这笔“借款”保障和改善了民生福祉、增强了人力资本、提高了个人和社会的生产率，并且建立起生产率的分享制度，我们这一代人无疑为后代留下了一笔优质资产。

五、结语和政策建议

人工智能继承了以往技术进步对经济、社会和就业产生影响的诸多特征，最突出地表现为创造性破坏的过程和后果。与此同时，这一最新且更具颠覆性技术的发展和应用后果，也在两个方面表现出与以往经验大不相同之处，具体来看就是两对矛盾现象。第一对矛盾是：人工智能产生的后果，无论就创造性还是就破坏性而言，其力度均远超以往技术的常规性变迁或颠覆性突破。第二对矛盾是：人工智能影响的趋势和方向具有必然性（即创造性破坏），路径和后果却具有巨大的不确定性。这都启示我们，着眼于未来并与人工智能共舞，必然需要做出更加普惠的社会保障制度安排。借用数学概念来说明的话，以剩余型与制度型两种模式作为两端，一种具体的社会保障（福利）制度模式可处于两者之间的无数点位，通常对应特定的国情和制度需求。可见，与当前中国所处的发展阶段、面临的挑战和预期的发展目标相对应的某种制度形式，必然处在这个光谱的某一点或某一区间，只是我们尚不清楚其具体位置而已。因此，转换认识范式、创新发展理念、深刻认识现实，进而准确判断未来，正是不断趋近制度最适点的探索过程。

从这个意义上来说，已有的各种制度模式和政策举措，譬如说面向全体国民发放的“无条件基本收入（universal basic income）”、无需缴费和资格识别的“全民基本养老金（social pension）”，以及依据社会必要生活质量保证确定的“生活工资（living wage）”等制度，虽然未必适宜全盘照搬，却都有着可供借鉴的有益成分。不过，既然中国经济社会发展已经处在新的阶段，并将继续向更高阶段迈进，或者说处于从必然王国进入自由王国这种过渡的更高形态上，所要借鉴的制度因素无疑应该更接近于制度型模式，而不是剩余型模式。下面，通过对人工智能挑战和发展面临的制约这两个因素进行交叉比照，聚焦几个最迫切的任务领域，提出有针对性的政策建议。

^① 瑞·达里欧：《原则：应对变化中的世界秩序》，中信出版集团，2022年，第79页。

首先,应对结构性就业矛盾,是抵御即将来临的人工智能就业冲击与应对当下劳动力市场挑战的结合点。这种结合让我们既始终坚守制度建设的方向,也善于抓住最紧迫的政策突破口,久久为功与只争朝夕的努力可以相互赋能。具体来说,这要求在制度建设和政策实施中,把兜底性保障和普惠性福利及再分配手段结合运用,加强以培训为核心的公共就业服务,重点扶助劳动力市场重点人群,特别是努力使其免于技术性失业冲击。同时,把普惠性收入保障制度与对工作本质和内涵的重新定义相结合,以生产率的不断提高支撑就业向深度和广度拓展。

其次,应对老龄化和“未富先老”,努力实现老有所养、老有所为、老有所乐,需要以制度建设为中心,充分调动和发挥人工智能作为强大赋能工具的作用,用以破解发展中的难题,消除成长中的烦恼,疏通民生改善中的堵点。针对中国“未富先老”新国情特点,一方面,借助社会保障普惠性制度安排,让老年人和大龄群体同全体居民一起,得以均等地分享生产率提高的结果;另一方面,借力人工智能系统和智能体的全方位应用,破解老龄化带来的经济社会问题,让技术发挥更大的改善民生、造福社会的作用。

最后,消除居民消费的后顾之忧,需要建设更充分、更普惠、更均等的社会保障制度。在老龄化和未富先老的条件下,居民消费能够在多大程度上满足支撑经济增长的需要,归根结底取决于两个方面的提升。一是通过收入水平的持续提高,实现居民消费能力的显著提升。这个条件对劳动年龄人口来说,通常来自劳动报酬的提高;而对于老年人来说,则更加依靠社会养老保险制度的保障水平。二是通过基本公共服务水平的持续提高,实现居民消费意愿的显著提升。这个条件无论对于劳动者群体还是老年人,都有赖于社会保障等公共品的供给,从而为居民消费解除后顾之忧。如果说经济发展的供给侧资源制约,正在日益被人工智能提高生产率的效应所打破,且劳动力市场参与者也能从充分就业中分享生产率提高的成果的话,打破老龄化背景下经济增长需求侧的消费制约,则更需仰仗分享性制度建设,以支撑经济合理增长,形成技术、经济、社会发展相互促进的良性循环。

Social Security in the Era of Artificial Intelligence: Conceptual Renewal and Institutional Construction

Cai Fang

(Chinese Academy of Social Sciences, Beijing 100732, China)

Abstract: The disruptive nature of artificial intelligence (AI) manifests as creative destruction in the economy, with a double-edged impact of significantly improving productivity, reshaping employment, and widening income inequality, as well as heightening the risks of structural employment contradictions and excessive income gaps. This in turn raises institutional demands, calling for more

proactive public employment services and a more inclusive social security system. Traditional social security systems face challenges such as identifying target populations, adapting institutional frameworks, and equitably sharing the dividends of technological progress. Therefore, the direction for improvement should involve a transition from a "residual-type" model toward a more "institutionalized" model. Innovative approaches are also required in public finance, particularly in social security financing, by exploring a cross-generational balance of revenues and expenditures characterized by "pre-distribution to promote growth." Higher productivity enabled by AI should be harnessed to support livelihood improvements. This includes shifting from material investment to human capital investment and increasing the share of government social spending in GDP; addressing structural employment contradictions and promoting elderly care through universal and inclusive security; and eliminating residents' concerns about consumption through social security, thereby unlocking strong domestic demand potential.

Keywords: artificial intelligence; creative destruction; institutional demand; universal social security

(责任编辑：华 颖)

