

数智时代的政策学习方法探索^{*}

——以最低生活保障为例

方悦¹ 胡诗云² 苏良军³ 解海天⁴

摘要 政策学习作为人工智能技术与因果推断理论的深度融合,为大数据时代的精准社会治理提供了科学依据。本文立足于中国治理规模巨大、政策约束条件复杂的现实背景,系统梳理了政策学习领域的国际前沿文献,并提出一种基于凸化处理理论的政策学习算法。该方法有效解决了在复杂高维政策空间中的优化计算难题,显著提升了大数据及多约束场景下的政策学习效率。文章进一步以中国最低生活保障制度为例,利用中国家庭金融调查数据进行实证分析。结果表明,该方法能够生成具备高度可解释性的精准低保分配方案,有效扩大社会福利增益并提升治理效能。本研究为运用前沿数智技术推动中国社会治理现代化提供了重要参考。

关键词 数字治理;政策学习;因果推断;大数据方法;最低生活保障

0 引言

习近平总书记指出:“治理体系和治理能力现代化是中国式现代化的应有之义。”^①随着数字化时代的全面到来,数字治理作为一种新型治理范式,正通过大数据、人工智能等技术的深度应用,推动政策制定从经验驱动向数据驱动转变,提升决策的科学化与精准化水平。近年来,随着数字政府建设的不断深化与政务数字化的全面普及,各级政府部门在社会治理、公共服务、经济调控等领域积累了海量的微观数据。然而,如何有效盘活这些数据资产,使其跨越统计监测的初级阶段,真正转化为驱动精准治理的科学决策能力,是当前深化数智技术应用、提升国家治理效能所面临的核心挑战。

在此背景下,政策学习(Policy Learning)作为一种新兴的数据驱动型经济政策研究范

^{*} 本研究受国家自然科学基金青年项目“政策学习中的理论创新与实践应用”(72503208)、国家自然科学基金重点基金项目“高维计量模型的机器学习方法及其在经济管理中的应用”(72133002)、国家自然科学基金青年项目“基于机器学习的因果政策制定——政策学习”(72403008)和重大项目“大规模商务场景下的统计学习与管理实践”(72495123)的资助。作者按照拼音首字母排序。

1 方悦,香港中文大学(深圳)经管学院助理教授,E-mail:fangyue@cuhk.edu.cn。

2 胡诗云,北京大学国家发展研究院博雅博士后,E-mail:hushiyun@pku.edu.cn。

3 苏良军,清华大学经济管理学院 C. V. Starr 讲席教授,E-mail:sulj@sem.tsinghua.edu.cn。

4 解海天(通信作者),北京大学光华管理学院助理教授,E-mail:xht@gsm.pku.edu.cn。

① 出自习近平总书记 2024 年 4 月 22 日至 24 日在重庆考察时的讲话,参见:https://www.gov.cn/yaowen/liebiao/202404/content_6947266.htm。

式应运而生。一方面,传统的机器学习工具虽长于在静态环境下的精准预测,却难以直接回答“改变政策干预将如何影响社会福利”这一核心经济学问题;另一方面,传统计量经济学的静态政策评估对处理效应高维度异质性的刻画极其有限,难以为设计精准化的干预政策提供现实指导(陶旭辉和郭峰,2023)。政策学习则创造性地将机器学习在处理高维非结构化大数据方面的计算优势,与计量经济学中因果推断在识别政策干预效应方面的严谨逻辑深度融合。它实现了从单纯结果预测向决策优化的跨越,能够以社会治理目标最大化为导向,为政府制定个性化、精准化的政策分配规则提供坚实的微观计量基础(陈强,2025;张征宇等,2026)。

尽管政策学习在理论界展现出巨大的应用潜力,但将其直接移植于中国社会治理的现实场景时,仍面临严峻的本土化挑战与方法论瓶颈。一方面,中国超大规模的经济体量与复杂的社会结构,决定了政策实施对象的基数极其庞大,数据维度极为丰富;另一方面,中国公共政策的制定绝非简单的无约束寻优,而是往往受到多重现实条件的严格约束,如政策规则的易解释性、区域与群体的公平性、财政预算约束,以及必须体现对弱势群体倾斜的特定函数形态(如单调性约束)等。这些约束条件使政策空间变得复杂,让最优政策的搜寻过程在数学上变得极其困难。以传统的经验福利最大化为代表的政策学习算法,在处理此类问题时通常会转化为混合整数规划问题。当面对中国的微观样本和高维约束空间时,传统算法遭遇了严重的维数灾难(Curse of Dimensionality)与计算瓶颈,运算量呈指数级暴增,甚至难以在有效时间内收敛至全局最优解。这极大地限制了政策学习算法在大数据环境和中国复杂治理场景中的实际应用效能。

针对上述现实痛点与技术局限,本文立足中国治理实践,致力于在理论方法与经验研究两方面作出实质性拓展,探索形成一套适用于中国复杂政策环境的精准决策方法论。

首先,在方法论层面,本文系统整合了国际前沿的政策学习理论,提出了一种基于凸化处理(Convexification)理论的改进型政策学习算法。具体而言,我们基于 Kitagawa et al. (2023)提出的铰链损失(Hinge Loss)凸化方法,将原始非凸的福利最大化问题转化为可高效求解的凸优化问题,显著降低了大数据场景下的计算复杂度。同时,我们融合 Athey and Wager(2021)的双重稳健估计框架,确保在观察性数据下政策效应的一致估计。另外,我们还引入 Mbakop and Tabord-Meehan(2021)的带惩罚项福利最大化(PWM)方法,实现对高复杂度政策空间的合理筛选与过拟合控制。这一套整合方法在保持理论最优收敛速率的前提下,极大提升了政策学习的计算效率和实际可操作性。

其次,在经验应用层面,本文选取中国社会保障体系中的核心制度——最低生活保障(简称“低保”)制度作为切入点,利用中国家庭金融调查(CHFS)的大样本微观数据展开实证分析。低保制度在保障弱势群体与维护社会公平中发挥着兜底作用,但长期面临保障标准难统筹、保障对象难精准识别等痛点。已经有一些文献讨论了低保制度的瞄准偏差,并利用双重机器学习等技术进行了初步的估计(韩华为,2021;赵明华和田北海,2024)。我们分别以降低家庭恩格尔系数和提高家庭净收入为政策目标,展示了如何通过数据驱动的方式生成具有高度可解释性的精准低保分配方案。进一步,本文阐明了如何运用上述改进的政策学习算法,在满足政策可解释性与防范道德风险(如单调性约束)的前提下,依据家庭多维经济特征(如收入来源多样性、历史消费结构等)精准识别低保需求群体。

研究不仅构建了具备高度实操性的精准低保分配规则,还有效量化了在考虑财政机会成本情况下的社会福利增益。

综上所述,本文的边际贡献在于:不仅为微观计量经济学中的政策学习文献提供了更加高效的算法实现路径,更重要的在于,以中国低保政策为典型案例,打通了从海量微观数据到科学因果推断,再到政策精准制定的完整逻辑闭环。本研究既为运用前沿数智技术推动中国社会治理现代化提供了重要的方法论参考,也为各级政府在多重约束下优化公共资源配置、全面提升数字治理效能提供了科学且可落地的分析工具。

1 政策学习方法的回顾、改进与整合

1.1 政策学习的定义与目标

从数据中学习最优的政策分配规则,是微观计量经济学与政策评估领域面临的核心任务之一。无论是政府向不同特征的失业者提供就业培训(Flores et al., 2012),还是企业向消费者发放不同额度的价格补贴(Dubé and Misra, 2023; Liu, 2023),抑或是为失业者提供不同时长的就业培训计划(Flores et al., 2012),在小额信贷项目中发放现金补助(Kaboski and Townsend, 2011)或对商品和服务制定精准税率(Chen et al., 2023),决策者都希望在有限的预算中,依据可观测的个体特征决定政策的分配方式,从而实现资源精准配置与整体效用最大化。

本文采用因果推断中经典的潜在结果框架(Potential Outcome Framework)来定义政策学习的目标。从形式上看,政策学习本质上是一个从个体特征空间到政策干预空间的映射函数的求解过程。假设观测样本包含 n 个独立同分布的个体数据 $\{(X_i, D_i, Y_i)\}_{i=1}^n$ 。其中 $X_i \in \mathcal{X}$ 为个体的多维协变量特征; $D_i \in \{0, 1\}$ 为二元政策处理变量($D_i = 1$ 表示受到干预, $D_i = 0$ 表示未受干预); Y_i 为决策者关注的结果变量。每个个体存在两个潜在结果: $Y_i(1)$ 与 $Y_i(0)$, 分别代表接受和未接受政策干预时的潜在表现。

一项“政策”可以被严格定义为一个指标函数 $\pi: \mathcal{X} \rightarrow 0, 1$, 其根据个体的特征 X 决定是否对其施加干预。政策学习的根本目标,在于寻找一个能够最大化全社会预期福利(Social Welfare)的最优政策分配规则。给定政策 π , 其对应的社会福利函数 $W(\pi)$ 定义为在该政策下潜在结果的数学期望:

$$W(\pi) = \mathbb{E}(Y(1)\pi(X) + Y(0)(1 - \pi(X))) \quad (1)$$

举例而言,假设政府计划实施一项旨在提高劳动者收入的“就业技能培训政策”。其中,协变量 X 代表劳动者的年龄、受教育程度、过往行业等特征;干预变量 $D = 1$ 代表参加培训;结果变量 Y 代表其未来的收入水平。此时, $Y(1)$ 和 $Y(0)$ 分别是该劳动者参加与不参加培训的潜在收入。政策 $\pi(X) = 1$ 意味着政府规定具备特征 X 的群体应当参加培训。上述福利函数 $W(\pi)$ 衡量的正是在政策 π 的分配规则下,全社会劳动者所能实现的平均总收入。

1.2 经验福利最大化(EWM)的三大局限与文献演进

Kitagawa and Tetenov(2018)在上述框架基础上,开创性地提出了“经验福利最大化”(Empirical Welfare Maximization, EWM)方法。

该方法将最优政策的搜索范围限定于一个预先设定的备选政策空间 Π 中,基于可忽略性假设,通过最大化样本经验福利[即式(1)的经验估计]来逼近真实的社会福利最大化。然而,传统的 EWM 方法在实际经济数据应用中面临三大严峻的理论与计算局限,后续前沿文献分别针对这三个维度进行了突破。

一是观察性数据下的倾向得分(Propensity Score)未知问题。原始 EWM 方法要求政策分配的倾向得分已知,方能实现样本量平方根级别的最优福利收敛速率。但在经济学常见的观察性研究中,个体的政策参与往往存在自选择偏误,倾向得分未知且必须通过数据估计,这将导致传统 EWM 方法失去最优收敛性。针对这一局限,Athey and Wager(2021)引入了因果推断中的双重稳健估计(Doubly Robust Estimation)。他们证明,通过结合倾向得分与结果条件均值的正交矩条件构建双重稳健得分,即使在倾向得分未知的情境下,依然能够保持最优的福利收敛速率,从而夯实了观察性数据下政策学习的计量基础。

二是政策空间复杂度的约束问题。传统 EWM 仅适用于复杂度较低的简单政策类别。但在真实的社会治理中,政策往往需要兼顾弱势群体保护、特定单调性等复杂的伦理与公平约束,其函数形态极其灵活。在高度复杂的政策空间中,传统方法极易陷入过拟合。针对这一问题,Mbakop and Tabord-Meehan(2021)提出了带惩罚项的福利最大化(Penalized Welfare Maximization, PWM)框架。该方法采用筛分法(Sieve Method),构建一系列复杂度递增的局部政策空间($\Pi_1 \subset \Pi_2 \subset \dots \subset \Pi_\infty$),通过引入衡量复杂度的惩罚项,在样本内拟合度与泛化能力之间寻求平衡,从而能够逼近任意复杂的全局目标政策。

三是非凸优化带来的计算瓶颈。从数学本质上看,寻找最优政策 $\pi(X) \in \{0,1\}$ 包含离散的示性函数,其对应的损失函数为 0-1 损失。这是一个典型的非凸优化(Non-convex Optimization)问题。当样本量达到数千或面临多维约束时,采用混合整数规划进行求解会导致计算时间呈指数级增长。针对这一计算瓶颈,Kitagawa et al. (2023)揭示了政策学习与机器学习分类风险最小化问题之间的等价联系。他们提出,可以通过引入铰链损失函数(Hinge Loss)对 0-1 损失进行凸松弛(Convexification)。这一处理在不改变原福利函数最优解的前提下,将非凸的组合优化问题转化为多项式时间内可解的凸优化问题,极大地提升了算法在大数据场景下的运行效率。

除了以上三方面的算法改进,最新的文献还将政策学习的设定进行了多方面拓展。Zhou et al. (2023)、Fang et al. (2025)以及 Ai et al. (2026)讨论了存在多种处理政策选项时的估计算法和模型选择;Zhan et al. (2024)将框架拓展到了自适应实验的情形中,从而使算法能够适应政策的不断变化与迭代;Kallus and Zhou(2021)讨论了存在不可观测干扰变量时的稳健政策学习方案;Ida et al. (2026)则将政策学习扩展到了局部平均处理效应(LATE)框架当中,使得政策学习能够适应并利用来自自我选择(Self-selection)的信息。

1.3 本文的一般化政策学习数学框架

结合上述三大前沿进展,本文提出并构建了一个融合“双重稳健估计—复杂空间筛分—凸化优化处理”的一般化政策学习框架,以应对高维复杂约束下的大数据决策问题。

第一步是对福利函数的双重稳健转化。首先,我们估计倾向得分 $e(X) = P(D=1|X)$ 与结果方程 $m(d, X) = E[Y|X, D=d]$ ($d \in \{0,1\}$)。为每个个体构建双重稳健的异质性

处理效应得分 w_i ：

$$w_i = m(1, X_i) - m(0, X_i) + \frac{D_i(Y_i - m(1, X_i))}{e(X_i)} - \frac{(1 - D_i)(Y_i - m(0, X_i))}{1 - e(X_i)} \quad (2)$$

双重机器学习的系列文献表明,只要倾向得分或结果方程中任意一个设定正确, w_i 的期望就等于平均处理效应(胡诗云等, 2026)。 w_i 在文献中也被称作伪结果(Pseudo-outcome), 可以视为对条件平均处理效应的一个粗略代理。可以证明, 最大化经验社会福利在数学上等价于最大化 $\frac{1}{n} \sum_{i=1}^n w_i \pi(X_i)$ 。

第二步则是映射为加权分类问题的 0-1 损失。我们将指标政策函数转化为连续的决策函数形式 $\pi(X) = 1\{f(X) \geq 0\}$ 。不难看出, 若 $w_i \geq 0$, 表明个体接受干预能产生正向福利, 最优策略应为其分配干预(即 $f(X_i) > 0$); 反之亦然。如果政策分配出错, 其带来的福利损失幅度恰好为 $|w_i|$ 。因此, 最大化福利等价于最小化以下以 $|w_i|$ 为权重的 0-1 分类损失函数：

$$\min_{f \in \mathcal{F}} W(f) = \min_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n |w_i| \mathbf{1}(\text{sign}(w_i) f(X_i) \leq 0) \quad (3)$$

其中, $\mathbf{1}(\cdot)$ 是示性函数(Indicator Function)。在这个意义下, 政策学习问题可以理解为对个体处理效应符号的分类问题。

第三步, 引入铰链损失的凸化松弛(Convexification)。为了解决上述 0-1 分类损失函数不可导、非凸导致的计算困难, 我们引入机器学习中具备优良数值计算特性的铰链损失函数 $\phi(v) = \max\{0, 1 - v\}$ 替代原有的示性函数。定义 $Z_i = \text{sign}(w_i) \in \{-1, 1\}$ 为目标标签。凸化处理后的经验风险目标函数变为：

$$\hat{L}^\phi(f) = \frac{1}{n} \sum_{i=1}^n |w_i| \max\{0, 1 - Z_i f(X_i)\} \quad (4)$$

这一转化使得原问题变为严格的凸优化问题, 可使用成熟的内点法或梯度下降法高效求解。

第四步是基于 PWM 的局部政策空间筛分与选择。面对包含复杂约束(如单调性约束、公平性约束)的原政策空间 $\mathcal{F}_\infty$, 我们构造一系列复杂度逐渐增加的逼近空间(如不同阶数的三角核函数空间) $\mathcal{F}_1 \subset \mathcal{F}_2 \cdots \subset \mathcal{F}_k$ 。

为避免过拟合, 我们将样本随机分为独立的估计样本(样本量为 n_E)与检验样本(样本量为 n_T)。首先在估计样本上, 针对每一个复杂度 k , 求解上述凸优化问题, 得到局部最优策略。随后, 在检验样本上评估这些策略的福利表现, 并减去衡量复杂度的技术惩罚项 $\tau(k, n_T)$ (例如取 $\tau \propto \sqrt{\log(k)/n_T}$), 从而选择使得惩罚后检验福利(Penalized Out-of-sample Welfare)最大化的最优复杂度 k^* ：

$$k^* = \underset{k \geq 1}{\operatorname{argmax}} \{-\hat{L}_{n_T}^\phi(\hat{f}_{n_E, k}^\phi) - \tau(k, n_T)\} \quad (5)$$

通过这一严格的交叉验证与复杂性惩罚机制, 本框架实现了在非实验数据、高维复杂约束及海量样本环境下的稳健、高效求解, 为后续具体经济政策的机制设计提供了坚实的算法支撑。

2 中国低保政策现状

中国的最低生活保障制度(简称“低保”)作为社会保障体系的重要组成部分,已成为帮助城乡困难群体维持基本生活、促进社会公平的核心工具。该制度最早起步于 20 世纪 90 年代,经过不断优化与完善,现已覆盖城乡居民,形成了以县级为单位实施的基本保障机制。低保政策通过提供现金补助,确保家庭人均收入低于当地最低生活保障标准的居民能够维持最基本的生活水平。这一制度不仅是中国社会保障体系的关键环节,也是政府应对贫困、推动社会公平的重要政策抓手。

目前,中国的低保政策主要分为城市和农村两类。城市低保制度于 1997 年率先实施,旨在保障城市贫困居民的基本生活。周文明和谢圣远(2016)对城镇居民最低生活保障制度进行了系统回顾,指出该政策在保障城市困难群体生活方面发挥了关键作用。农村低保制度则于 2007 年起全面推行,专门面向农村贫困人口,以确保其基本生活。符华平和顾海(2009)对农村低保的现状 & 政策进行了相关探讨,认为农村低保政策的实施在缓解农村贫困方面取得了显著成效。

近年来,随着新型城镇化的推进和城乡一体化的发展,低保制度在城乡之间逐步趋于统一,农村与城市的老保政策差距正在逐步缩小。根据民政部发布的 2023 年《民政事业发展统计公报》,截至 2023 年底,全国共有城市低保对象 663.6 万人,平均保障标准为 785.9 元/(人·月),全年支出城市低保资金 464.7 亿元;农村低保对象为 3399.7 万人,平均保障标准为 621.3 元/(人·月),全年支出农村低保资金 1483.9 亿元(民政部,2023)。上述数据充分反映了中国低保制度在覆盖面和资金投入方面取得的显著成效。

低保制度在贫困救助方面取得了显著成效。作为社会保障体系的重要组成部分,它有效填补了传统扶贫政策的空白,尤其是在农村贫困地区,通过现金补助及其他社会救助措施,帮助低收入家庭缓解生活困难、逐步摆脱贫困。低保制度不仅在物质层面为贫困家庭提供了基本保障,也在促进社会公平方面发挥了积极作用。通过覆盖城乡困难群体,低保政策为社会弱势群体提供了必要支持,有助于缩小收入差距、维护社会正义,从而为社会的稳定与和谐提供了保障。

尽管低保制度取得了显著成绩,但在当前地方财政持续承压的背景下,如何精准高效地做好“六稳”工作、落实“六保”任务,兜牢基本民生底线,已成为政策制定的关键议题。一是对象认定的精准度不足。受基层信息不对称与主观判断因素的影响,“应保未保”与“错保”现象依然存在(李棉管,2017;朱梦冰和李实,2017;宋锦等,2020)。二是低保标准的动态调整机制尚不完善,难以灵敏地反映不同地区物价与生活成本的实时变化。这些问题的叠加,使得低保资金的使用效率面临挑战。为此,低保管理有必要借助大数据与数字化手段,将居民收入、消费等多维信息整合至社会保障数据平台,推动认定过程向更透明、更个性化的方向发展,确保最需要救助的群体能够及时获得支持,从而在保障基本民生的同时,减轻财政无效支出的压力,推动社会保障的可持续发展与社会和谐稳定。

3 精准低保政策学习

在考虑精准低保政策学习时,我们可以设定不同的政策目标。本节首先以降低家庭

平均恩格尔系数作为政策目标,随后将最大化家庭净收入作为备选目标。

3.1 数据、政策目标及处理变量

本研究使用的数据来自西南财经大学中国家庭金融调查与研究中心组织管理的“中国家庭金融调查”项目(甘犁等,2013)。基于CHFS 2017年与2019年的调查数据,本文为如何依据数据精准学习低保政策提供了一个示例性分析。我们通过计算2019年家庭食品消费支出占家庭消费总支出的比重,得到家庭恩格尔系数 Y_i 。恩格尔系数是衡量家庭及国家贫困程度的重要指标,其数值越低,通常意味着家庭和国家越富裕。本研究的政策目标设定为降低家庭的平均恩格尔系数,从而引导更多收入释放至教育、医疗保健等领域。

然而,从理论上,恩格尔系数应反映基本食品支出的占比,而CHFS统计的食品消费中,除基本食品外还包含烟酒支出,后者恰恰是政策制定者并不希望鼓励的消费类别。已有实证文献表明,家庭在获得转移收入后,未必会将其用于政策制定者所期望的必要食品、生活用品、教育和医疗等方面,反而可能增加烟酒消费。受此影响,并非所有家庭在领取最低生活保障后都会出现恩格尔系数的下降;将低保收入用于烟酒消费,显然有违最低生活保障制度的初衷。更为关键的是,烟酒支出因被计入食品消费,会直接推高CHFS统计口径下的恩格尔系数,这便在事实上构成了政策学习中的一个潜在成本项。

正是由于这一成本项的存在,即使未显式引入财政约束,政策学习也不会得出“人人获得低保”的平凡解。出于最大限度降低平均恩格尔系数的目标,我们会在异质性家庭中有选择地覆盖那些能够真正降低该系数的家庭,而非简单地倾向于覆盖所有家庭。换言之,烟酒支出作为降低恩格尔系数过程中的事实成本项,使得政策分配呈现出非平凡(Non-trivial)的结果。这一选择体现了基于异质性作用的指向性政策制定思路,旨在确保资源的有效配置,使最低生活保障切实用于改善家庭的基本生活水平。

尽管本实例分析并未显式考虑财政限制,但如前所述,烟酒支出通过提高CHFS食品消费统计数据进而推高恩格尔系数,本身就已充当了潜在的成本项。本分析框架并不排除将政策成本直接纳入考量的可能性,只是在实际应用中需要因地制宜:既要结合各地政府的具体低保支付标准,也取决于政府对目标变量与成本变量之间关系连接方式的设定,例如政府愿意为恩格尔系数每降低一个百分点支付多少成本。

本文的政策处理变量 D_i 根据问卷中“过去一年是否获得过最低生活保障”这一问题构建。我们假设,低保获得状态在家庭总消费发生之前即已确定,即低保资格的确认至少发生在入户调查的前一年。但必须承认这一假设的局限性:仅凭该问题,部分家庭可能在距离调查较近的时期才获得低保资格,其对消费结构的影响或许尚未充分显现。对此,我们进一步假设这类测量误差在不同特征的群体中是随机的,不会造成选择性偏误。鉴于本研究关注的是针对不同协变量群体的政策分配方案,研究问题可界定为:基于已观测到的低保对不同特征群体消费结构的影响,如何最优化地分配低保资格。

在识别策略上,我们假设,在控制2017年家庭是否为低保户以及其他一系列协变量后,2019年的低保状态 D_i 与当年恩格尔系数的潜在结果 $Y_i(1)$ 、 $Y_i(0)$ 相互独立。所有控制变量均采用2017年取值,以保证它们在时间上先于处理变量,避免反向因果。这些协

变量包括:2017 年家庭是否为低保户、家庭人均收入、家庭收入来源多样性指数、家庭总消费、食品消费、医疗消费、住宿消费、户主受教育水平、户主性别以及地区等。其中,家庭收入来源多样性指数基于赫芬达尔-赫希曼指数构建:

$$h_i = 1 - \sum_{j=1}^{J_i} (I_{ij}/I_i)^2$$

其中, I_{ij} 为家庭 i 第 j 种收入来源的金额, J_i 为收入来源总数, h_i 越大表示收入来源越多样。该指标反映了家庭收入的多元化程度,间接衡量其抵御经济风险的能力。在不确定的经济环境中,收入多样性较高的家庭通常能更好地应对突发冲击,从而降低陷入贫困的可能性。户主受教育水平被分为三类:小学及以下、初中、高中及以上。本文选取 2017 年家庭总收入低于 100000 元的家庭作为分析样本,主要变量的描述性统计见表 1。

表 1 主要变量的描述性统计

变量	样本量	均值	标准差	最小值	最大值
恩格尔系数	11603	0.430	0.199	0.000	0.994
2017 年恩格尔系数	11603	0.525	0.202	0.000	0.997
2019 年低保	11603	0.087	0.281	0.000	1.000
2017 年低保	11603	0.075	0.264	0.000	1.000
收入来源多样性	11603	0.279	0.248	0.000	1.000
户主受教育程度					
小学及以下	11603	0.434	0.496	0.000	1.000
初中	11603	0.356	0.479	0.000	1.000
高中及以上	11603	0.210	0.407	0.000	1.000
农村	11603	0.492	0.500	0.000	1.000
户主为女性	11603	0.181	0.385	0.000	1.000
总收入	11603	38773.366	27985.390	0.175	99990.000
教育支出	11603	3503.632	7709.701	0.000	150360.000
医疗支出	11603	7218.356	19258.990	0.000	615000.000

分析样本共包含 11603 个家庭。其中,2019 年低保户占 8.7%,2017 年低保户占 7.5%;2017 年家庭总收入均值为 38773.366 元,收入来源多样性指数均值为 0.279;户主为女性的家庭占 18.1%;户主受教育程度为小学及以下的家庭占 43.4%,初中占 35.6%,高中及以上占 21.0%;农村地区家庭占比 49.2%。2017 年平均恩格尔系数为 0.525,2019 年降至 0.463。

3.2 基于降低家庭的平均恩格尔系数的政策目标的精准政策学习

我们首先选取家庭收入来源多样性与过去恩格尔系数两个变量作为政策变量,这两个变量均为衡量家庭抗风险能力的变量,家庭收入来源多样性越高,则抗风险能力越强;过去的恩格尔系数越低,则家庭经济状况越好,抗风险能力越强。我们认为,从政策透明

度及可解释性的角度出发,选取这两个变量作为政策变量是合理的。

为便于后文政策阐述,我们对2017年恩格尔系数进行如下变换:取 $\widetilde{PE}_i = 1 - PE_i$ 。在此变换下,单调性约束体现为: $\widetilde{PE}_i$ 和 h_i 越小的家庭,越应当获得低保资格。我们认为,抗风险能力更低的家庭更应获得最低生活保障。例如,在收入来源多样性相同的情况下,若2017年恩格尔系数为0.5($\widetilde{PE}_i = 0.5$)的家庭根据政策可获得最低生活保障,那么过去恩格尔系数为0.6($\widetilde{PE}_i = 0.4$)的家庭也应该获得该保障;同理,在过去恩格尔系数相同的情况下,收入来源多样性较低的家庭相较于多样性较高的家庭,也更应获得低保资格。

记 $X_i = (\widetilde{PE}_i, h_i)$,我们考虑的组成原政策空间的政策函数为:

$$\pi(X_i) = 1\{f_1(\widetilde{PE}_i) + f_2(h_i) \leq 0\}$$

其中, f_1, f_2 均为单调递增函数。参考Mbapok and Tabord-Meehan(2021),利用不同阶数的三角核函数对具有单调性约束的政策空间进行逼近。关于低保情境下政策优化的具体数学细节,请参见本文附录1、2部分。

在政策学习过程中,我们首先考虑未纳入低保政策成本的情形,随后再将低保政策的成本纳入分析框架。

我们在图1中展示了当 k 取1至9时,通过数据驱动学习得到的低保资格分配政策,以及依据所学政策,检验样本中家庭是否获得低保资格的情况。对比不同政策空间下的最优政策可以发现, k 越大,政策越复杂。在 $k=7,8,9$ 的政策空间下,数据学习得到的政策极度偏好收入来源单一的家庭;而对于收入来源更为多样化的家庭,是否给予低保资格则对过去恩格尔系数提出了较为不同的要求。这一结果也印证了我们对数据在复杂政策空间下可能存在过拟合的担忧是合理的。

通过惩罚福利最大化(PWM)方法,最终选择了 $k=2$ 的局部政策空间。在该最优政策下,23.4%的检验样本应获得低保资格,估计得到的平均恩格尔系数为0.25。举例而言,如果一个家庭的收入来源多样性指数为0.5,过去恩格尔系数为0.5,则该家庭可获得低保资格。原样本中,实际拥有低保资格的家庭占比为8.65%,其平均恩格尔系数为0.43。由于我们并未显性考虑低保政策的成本,因此估计得到的低保政策给予更多家庭以低保资格,是合理的结果。

我们进一步将低保政策的成本纳入研究范围。为了将用于低保的财政支出与恩格尔系数相结合,我们计算得到低保政策关于恩格尔系数的机会成本。基于2004—2019年国内人均生产总值与恩格尔系数的数据,可以发现人均GDP与恩格尔系数呈现负相关关系,即人均GDP的提高伴随着恩格尔系数的下降。我们认为,如果将用于低保政策的支出转投于改善民生的其他领域,恩格尔系数潜在降低的程度可由人均GDP与恩格尔系数之间的关系来刻画。因此,我们通过简单的线性回归估计恩格尔系数关于人均GDP的系数,并据此计算低保支出关于恩格尔系数的机会成本。图2回归得到的相关系数为 -1.12×10^{-6} ($se = 1.41 \times 10^{-7}$)。

根据中华人民共和国国家发展和改革委员会发布的数据,“截至2023年1月底,全国共有城市低保对象679万人,低保平均标准达到752元/(人·月);农村低保对象3338万人,

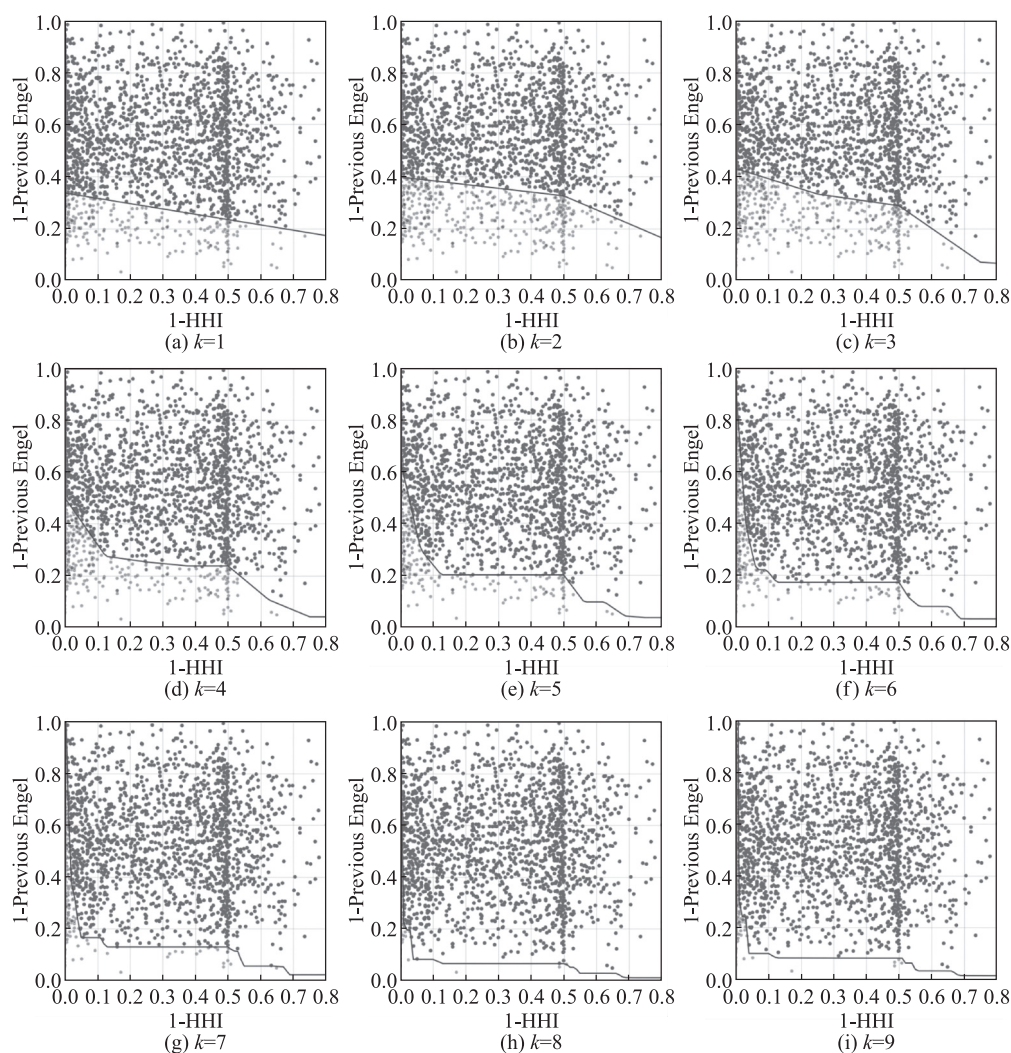


图 1 数据驱动下的低保政策(恩格尔系数):不考虑财政成本

注:该图以降低平均恩格尔系数为政策目标,选择家庭收入来源多样性(1-HHI)和过去恩格尔系数(1-previous Engel)为政策变量。图中实线为通过数据学习得到的该政策空间中的最优低保政策。浅色点为在该政策下检验样本中应得到低保资格的家庭,深色点为该政策下检验样本中不拥有低保资格的家庭。 $k=2$ 的政策空间被最终选择。在该最优政策下,检验样本中应获得低保的家庭占比为 23.4%。

低保平均标准达到 582 元/(人·月)”^①。我们将城镇低保与农村低保进行简单的加权平均,得到 610.7 元/(人·月)。根据姚建平(2018),低保家庭的平均家庭成员人数为 2.28 人。据此计算,最低生活保障的户均年支出为 16708 元,其对应的恩格尔系数机会成本为 0.0286。

我们将该成本纳入福利函数,对上述问题进行重新分析,不同政策空间下的最优政策如图 3 所示。经过上述分析步骤,我们最终选择了 $k=3$ 的局部政策空间。在该最优政策

① 参见: https://www.ndrc.gov.cn/fggz/jyysr/jysrsbxf/202304/t20230428_1355158.html。

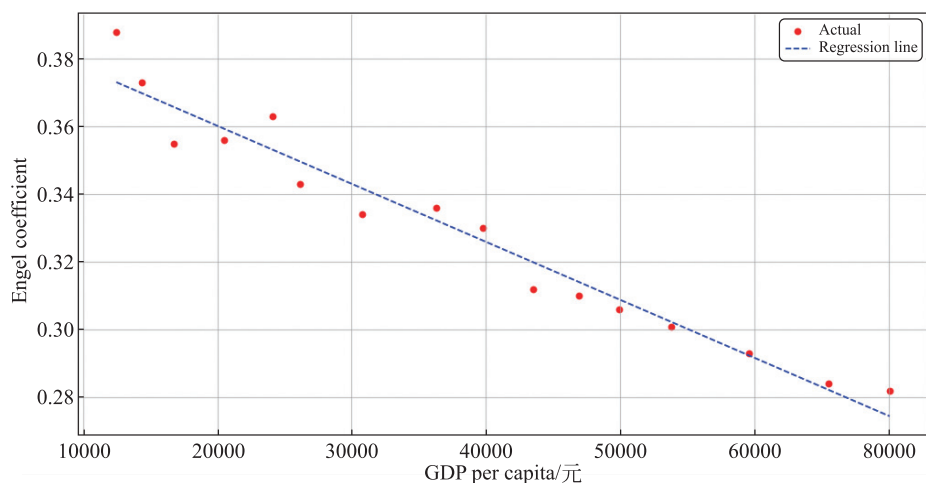


图2 恩格尔系数与人均国内生产总值的关系

下,检验样本中 16.7% 的家庭应获得低保资格。相比不考虑成本的情况,该比例降低了约 28.6%。

3.3 基于最大化家庭净收入的政策目标的精准政策学习

在实际的政策制定中,可以考虑不同的政策目标。在本节我们将最优化的目标变为最大化家庭净收入 NI_i ,即家庭总收入 I_i 减去家庭总支出 E_i ,在政策变量组合上,我们选取过去家庭总收入与家庭收入来源多样性指标。具体的数学细节,请参见本文附录 3 部分。

在图 4 和 5 中,我们分别展示了不考虑成本与考虑成本两种情形下,以最大化家庭净收入为目标的最优低保政策。在两种情形中,均为 $k=6$ 的局部政策空间被选择。

在不考虑成本时,检验样本中有 27.7% 的家庭可获得低保资格。从图 4 中可见,与之前得到的低保政策极为不同的是,当家庭的过去总收入超过 56000 元时,无论该家庭收入来源多样性的取值,其均不会获得低保资格;而当家庭过去总收入在 40000 元及以下时,则家庭收入来源多样性成为该家庭是否可以获得低保资格的重要指标。这种政策分配机制体现了我们的单调性约束,也是在低保政策制定中的一种理性原则,即优先保障那些人均收入较低且抗风险能力相对较弱的家庭。

在考虑成本的情况下,应获得低保资格的家庭比例略有下降,为 26.2%,且最优低保政策与不考虑成本时基本类似。

4 政策建议与结论

本文系统构建了一个融合双重稳健估计、复杂空间筛分与凸优化理论的数据驱动型政策学习框架,并以中国最低生活保障制度为典型应用场景,利用大样本微观调查数据进行了实证检验。研究表明,相较于传统的经验决策,基于机器学习与因果推断深度融合的政策学习方法,能够在满足政策可解释性、防范道德风险等多重现实约束的前提下,有效

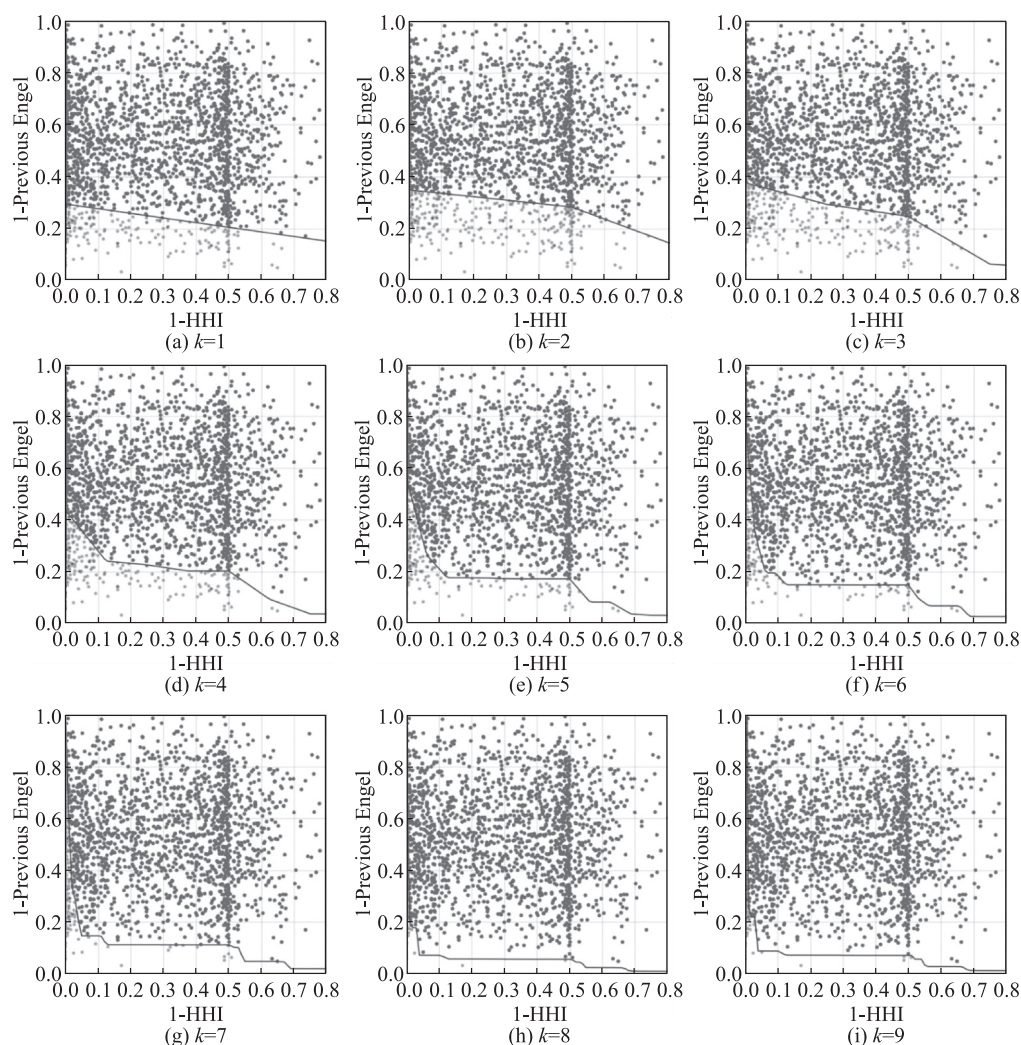


图 3 数据驱动下的低保政策(恩格尔系数):考虑财政成本

注:该图以降低平均恩格尔系数为政策目标,选择家庭收入来源多样性(1-HHI)和过去恩格尔系数(1-previous Engel)为政策变量。图中实线为通过数据学习得到的该政策空间中的最优低保政策。浅色点为在该政策下检验样本中应得到低保资格的家庭,深色点为该政策下检验样本中不拥有低保资格的家庭。 $k=3$ 的政策空间被最终选择。在该最优政策下,检验样本中应获得低保的家庭占比为 16.7%。

识别出干预效果最佳的目标群体。这不仅显著提升了复杂高维政策空间中的计算效率,更为实现社会福利最大化提供了坚实的微观计量支撑透过低保政策的实证镜像,本文的研究结论对数智时代一般性公共政策(如就业补贴、教育资助、税收减免等)的精准制定与效能提升,具有广泛而深刻的启示。具体而言,可得出以下五个方面的政策建议。

4.1 重塑干预机制:从“经验阈值划分”向“数据驱动寻优”转变

传统的公共政策往往依赖单一的静态指标(如以当地平均收入设定门槛)进行资格认定与资源分配。这种“一刀切”的经验阈值划分虽然易于操作,却难以捕捉不同群体在面

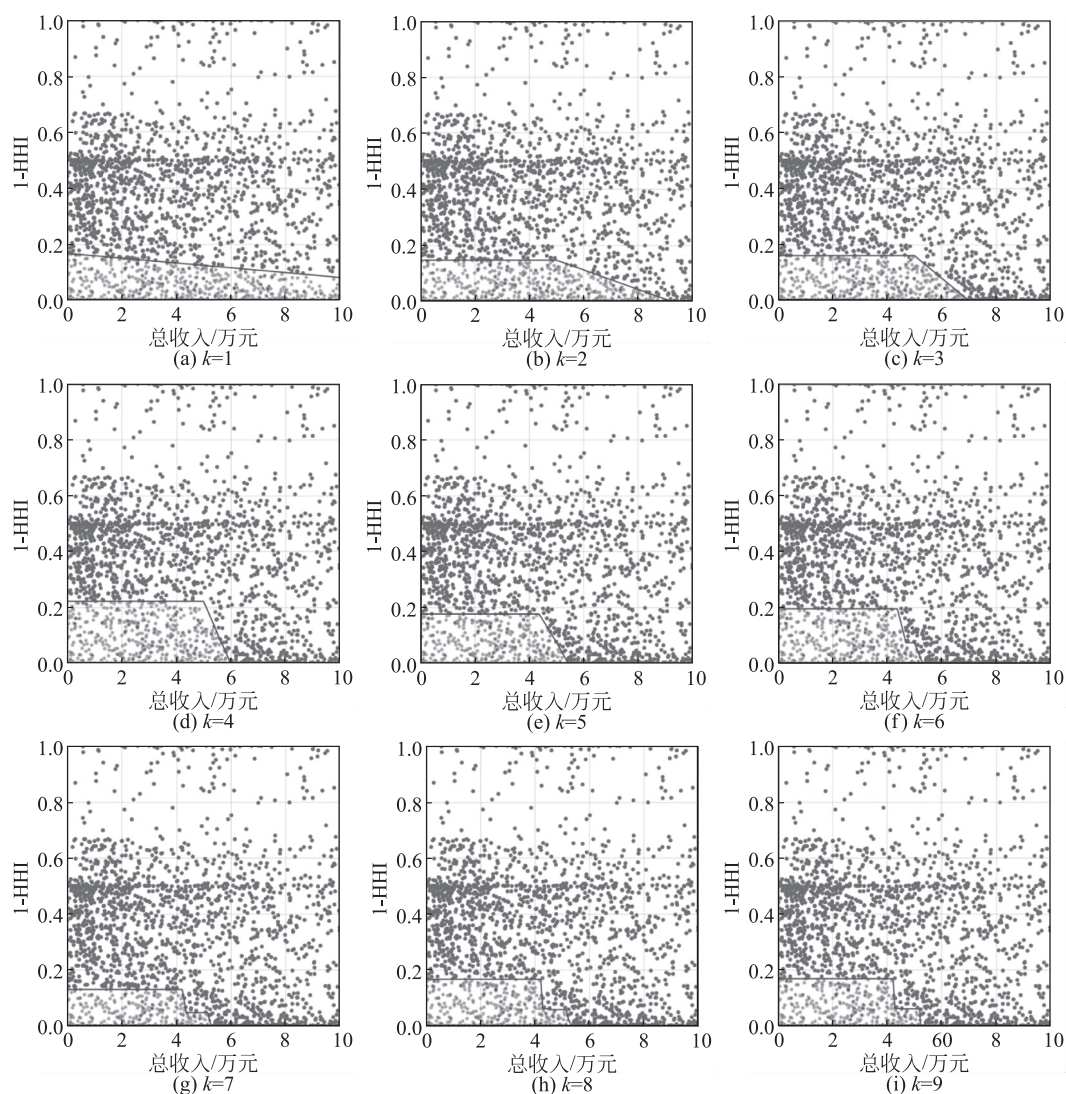


图4 数据驱动下的低保政策(家庭净收入): 不考虑成本

注:该图以提高家庭净收入为政策目标,选择过去家庭总收入(previous income)和家庭收入来源多样性(1-HHI)为政策变量。图中实线为通过数据学习得到的该政策空间中的最优低保政策。浅色点为在该政策下检验样本中应得到低保资格的家庭,深色点为该政策下检验样本中不拥有低保资格的家庭。 $k=6$ 的政策空间被最终选择。在该最优政策下,检验样本中应获得低保的家庭占比为27.7%。

对政策干预时所表现出的复杂异质性,更容易引发潜在的道德风险(Tan-Soo et al., 2025)。本文的实证结果表明,引入数据驱动的政策学习技术,可以全方位评估不同特征群体从政策中获得的潜在收益(即异质性处理效应)。对于一般性政策制定而言,政府应逐步打破依赖单一经验阈值的决策惯性,在较为宽泛的政策空间内,利用算法求解多约束条件下的最优分配规则。这不仅能大幅提升稀缺公共资源的配置效率,还能在数据框架下实现政策灵活性与规则透明度的有机平衡。

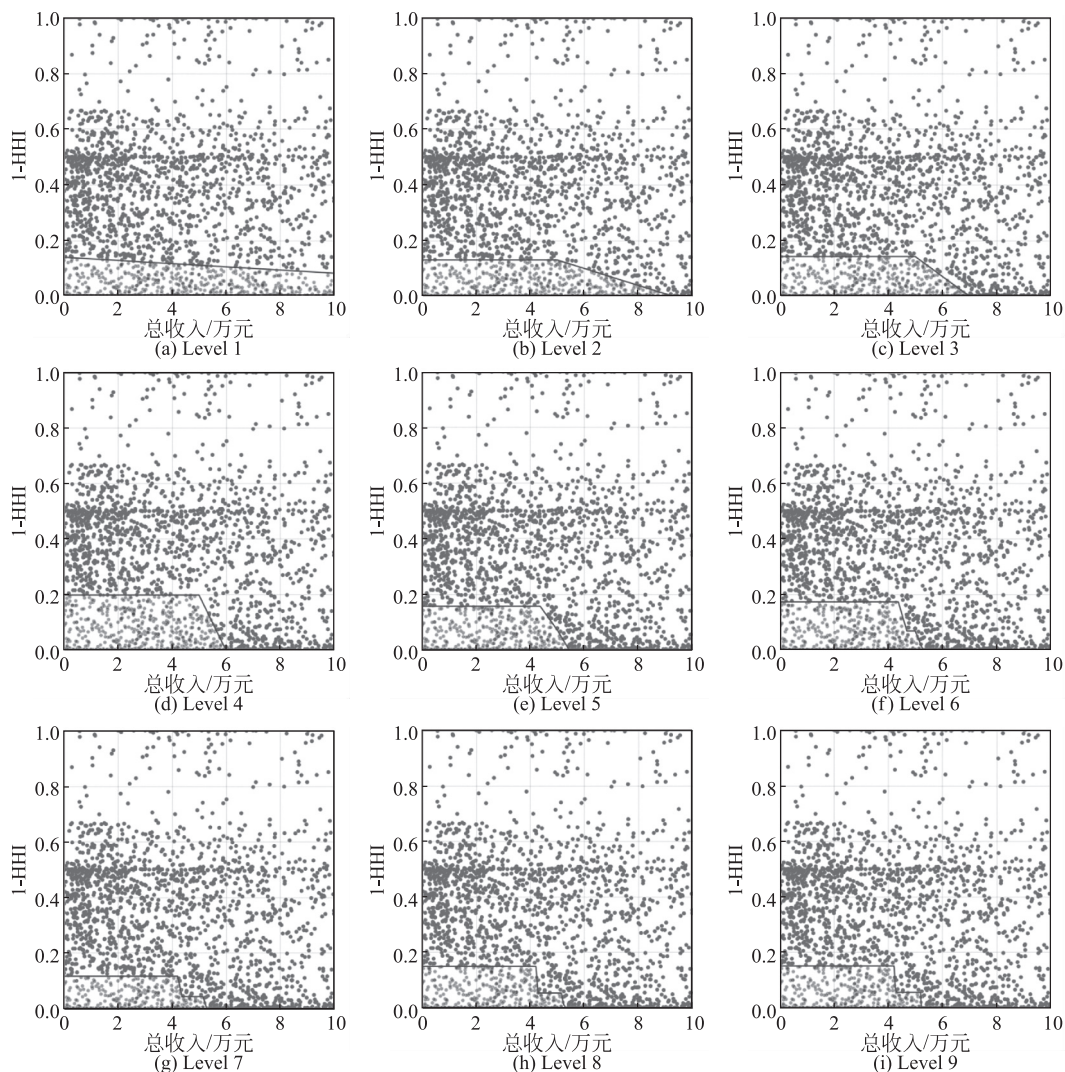


图 5 数据驱动下的低保政策(家庭净收入): 加入成本

注: 该图以提高家庭净收入为政策目标, 选择过去家庭总收入(previous income)和家庭收入来源多样性(1-HHI)为政策变量。图中实线为通过数据学习得到的该政策空间中的最优低保政策。浅色点为在该政策下检验样本中应得到低保资格的家庭, 深色点为该政策下检验样本中不拥有低保资格的家庭。 $k=6$ 的政策空间被最终选择。在最优政策下, 检验样本中应获得低保的家庭占比为 26.2%。

4.2 拓宽考量维度: 构建多维特征与多目标的协同决策体系

在本文的分析中, 当我们将家庭收入来源的多样性、历史支出结构(如恩格尔系数)引入决策变量后, 政策的精准度得到了显著优化。这启示我们在更广泛的社会经济政策设计中, 不应局限于当前的静态财务指标, 而应引入反映微观主体抗风险能力与经济脆弱性的多维动态指标。此外, 现实中的政策往往面临财政预算约束或需要兼顾多元目标(如促进公平与提升效率)(Haushofer et al., 2025)。因此, 政策制定者应构建多目标协同的寻优框架, 将干预成本、辐射范围以及多维政策目标(如改善生活质量、促进教育公平、减少

医疗负担等)显性地纳入政策学习的成本效益分析模型中,通过构建反映多目标权衡的效用函数,得出更加契合宏观战略导向的综合性分配方案。

4.3 打破空间桎梏：推动政策分配的因地制宜与动态迭代

中国区域经济发展极不平衡,各地的生活成本、产业结构与社会资源存在显著差异,这意味着最优的政策分配规则绝不应是全局统一的静态映射。政策学习方法赋予了政策设计极强的时空适应能力。一方面,各级地方政府应基于本地的微观特征数据,针对性地训练和调整政策学习模型,生成符合当地地区情的最优干预规则,实现真正意义上的因地制宜。另一方面,政策制定是一个随宏观经济周期动态演化的过程。政府应依托现代数字治理平台,建立政策实施的实时反馈与数据更新机制,根据经济形势波动及时重估政策效应,推动政策规则的动态迭代与持续优化。

4.4 夯实底层基建：完善高质量微观数据的采集与治理机制

数据驱动型决策的上限,取决于底层数据的广度与质量。政策学习中机器学习算法的准确性与因果推断的稳健性,高度依赖于全面、及时且无偏的微观数据。当前,部分地区(尤其是偏远农村)在数据采集的时效性与深度上仍存在短板,数据孤岛现象依然严重。为此,国家应进一步加强数字政务基础设施建设,建立跨部门(如税务、民政、医疗、教育)的数据互通与动态维护机制,全面绘制微观主体的经济画像。同时,在数据的深度应用过程中,必须建立严格的数据脱敏与隐私保护机制,以透明、安全的数据治理体系赢得公众信任,为精准社会治理夯实底层基础。

4.5 向纵深拓展：探索从“二元干预”到“分层精准施策”的升级

本文重点探讨了二元干预(即是否给予低保资格)的政策学习。然而,真实的社会治理场景远比二元选择复杂。正如低保制度包含不同的保障档次,就业支持涵盖不同时长的培训计划,税收优惠具备多级累进税率。未来的政策学习与公共决策应积极向多元处理变量(Multi-action)与连续型干预拓展(Zhou et al., 2023; Fang et al., 2025; Ai et al., 2026)。通过评估不同强度、不同类型的干预工具组合对微观主体的实际效果,政府可以从“给与不给”的粗放式决策,走向“给多少、怎么给”的精细化分层施策。除此之外,政府还可以将以基尼系数为代表的社会平等指标作为干预目标,提高底层收入群体的福利权重(Kitagawa and Tetenov, 2021)。这将为构建全方位、多层次的现代公共服务体系提供强大的方法论支撑。

综上所述,政策学习不仅仅是一种新兴的计量经济学工具,更是推动国家治理体系与治理能力现代化、实现数字政府高质量发展的重要技术路径。随着底层数据的不断丰富与算法算力的持续跃升,基于大数据的因果推断与政策学习必将在深化中国经济体制改革、优化资源配置、促进社会公平等方面释放出更大的潜能。

参考文献

陈强. 2025. 计量经济学中的因果推断：过去、现在与未来[J]. 中山大学学报(社会科学版), 65(1): 43-64.

- Chen Q. 2025. Causal inference in econometrics: The past, the present and the future[J]. *Journal of Sun Yat-Sen University (Social Science Edition)*, 65(1): 43-64. (in Chinese)
- 符华平, 顾海. 2009. 我国农村最低生活保障制度的现状、问题及对策[J]. *南京社会科学*, (1): 101-104.
- Fu H P, Gu H. 2009. Current situation, problems and countermeasures of China's rural minimum living security system[J]. *Social Sciences in Nanjing*, (1): 101-104. (in Chinese)
- 甘犁, 尹志超, 贾男, 等. 2013. 中国家庭资产状况及住房需求分析[J]. *金融研究*, (4): 1-14.
- Gan L, Yin Z C, Jia N, et al. 2013. Assets and residential demand of Chinese households[J]. *Journal of Financial Research*, (4): 1-14. (in Chinese)
- 韩华为. 2021. 代理家计调查、农村低保瞄准精度和减贫效应——基于中国家庭金融调查的实证研究[J]. *社会保障评论*, 5(2): 93-109.
- Han H W. 2021. Proxy means tests analysis of rural Dibao's targeting performance and anti-poverty effectiveness: empirical study based on China household finance survey[J]. *Chinese Social Security Review*, 5(2): 93-109. (in Chinese)
- 胡诗云, 江弘毅, 解海天. 2026. 双重机器学习的理论与应用——从“黑箱”到“工具箱”的实践指南[J]. *数量经济技术经济研究*, 43(3): 154-179.
- Hu S Y, Jiang H Y, Xie H T. 2026. The theory and practices of double machine learning: From “black box” to “tool box”[J]. *Journal of Quantitative & Technological Economics*, 43(3): 154-179. (in Chinese)
- 李棉管. 2017. 技术难题、政治过程与文化结果——“瞄准偏差”的三种研究视角及其对中国“精准扶贫”的启示[J]. *社会学研究*, 32(1): 217-241.
- Li M G. 2017. Technological barriers, political process and cultural consequence: Three research perspectives on targeting error and the implications for “targeting poverty” in China[J]. *Sociological Studies*, 32(1): 217-241. (in Chinese)
- 民政部. 2023. 2023 年民政事业发展统计公报[R/OL]. <https://www.mca.gov.cn/n156/n2679/c1662004999980001204/attr/355717.pdf>.
- Ministry of Civil Affairs. 2023. Statistical bulletin on the development of civil affairs in 2023[R/OL]. <https://www.mca.gov.cn/n156/n2679/c1662004999980001204/attr/355717.pdf>. (in Chinese)
- 宋锦, 李实, 王德文. 2020. 中国城市低保制度的瞄准度分析[J]. *管理世界*, 36(6): 37-48.
- Song J, Li S, Wang D W. 2020. Analysis on the targeting of urban Dibao program in China[J]. *Journal of Management World*, 36(6): 37-48. (in Chinese)
- 陶旭辉, 郭峰. 2023. 异质性政策效应评估与机器学习方法: 研究进展与未来方向[J]. *管理世界*, 39(11): 216-235.
- Tao X H, Guo F. 2023. Policy effect heterogeneity evaluation and machine learning methods: Research progress and future orientation[J]. *Journal of Management World*, 39(11): 216-235. (in Chinese)
- 姚建平. 2018. 中国城市低保瞄准困境: 资格障碍、技术难题, 还是政治影响? [J]. *社会科学*, (3): 61-72.
- Yao J P. 2018. The targeting dilemma of urban Dibao in China: Eligibility obstruction, technological problem, or political influences? [J]. *Journal of Social Sciences*, (3): 61-72. (in Chinese)
- 张征宇, 陈浩文, 郎旭华. 2026. 面向共同富裕问题的因果推断[J]. *经济评论*, (1): 3-19.
- Zhang Z Y, Chen H W, Lang X H. 2026. Causal inference on the issue of common prosperity[J]. *Economic Review*, (1): 3-19. (in Chinese)
- 赵明华, 田北海. 2024. 农村低保瞄准偏差估计及乡村治理手段创新的改进效应——基于双重机器学习的因果推断[J]. *华中农业大学学报(社会科学版)*, (5): 167-180.
- Zhao M H, Tian B H. 2024. Estimation of targeting bias for rural Dibao and the improvement effect of rural governance innovation: Causal inference based on double machine learning[J]. *Journal of Huazhong Agricultural University (Social Sciences Edition)*, (5): 167-180. (in Chinese)
- 周文明, 谢圣远. 2016. 中国城镇居民最低生活保障制度的发展演进及政策评估[J]. *广东社会科学*, (2): 206-212.

- Zhou W M, Xie S Y. 2016. The development and policy evaluation of China's urban minimum living security system[J]. *Social Sciences in Guangdong*, (2): 206-212. (in Chinese)
- 朱梦冰, 李实. 2017. 精准扶贫重在精准识别贫困人口——农村低保政策的瞄准效果分析[J]. *中国社会科学*, (9): 90-112.
- Zhu M B, Li S. 2017. The key to precise poverty alleviation rests in the precise identification of impoverished populations: An analysis of the targeting effect of the rural subsistence allowance policy[J]. *Social Sciences in China*, (9): 90-112. (in Chinese)
- Ai C R, Fang Y, Xie H T. 2026. Data-driven policy learning for continuous treatments[J]. *Journal of Econometrics*, 253: 106170.
- Athey S, Wager S. 2021. Policy learning with observational data[J]. *Econometrica*, 89(1): 133-161.
- Chen Z, Jiang X, Liu Z K, et al. 2023. Tax policy and lumpy investment behaviour: Evidence from China's VAT reform[J]. *The Review of Economic Studies*, 90(2): 634-674.
- Dubé J P, Misra S. 2023. Personalized pricing and consumer welfare[J]. *Journal of Political Economy*, 131(1): 131-189.
- Fang Y, Xi J, Xie H T. 2025. Model selection for multivalued-treatment policy learning in observational studies[J]. *Journal of Business & Economic Statistics*, 43(4): 897-909.
- Flores C A, Flores-Lagunes A, Gonzalez A, et al. 2012. Estimating the effects of length of exposure to instruction in a training program: The case of job corps[J]. *Review of Economics and Statistics*, 94(1): 153-171.
- Haushofer J, Niehaus P, Paramo C, et al. 2025. Targeting impact versus deprivation[J]. *American Economic Review*, 115(6): 1936-1974.
- Ida T, Ishihara T, Ito K, et al. 2026. Choosing who chooses: Selection-driven targeting in energy rebate programs[J]. *Econometrica*, 94(1): 225-247.
- Kaboski J P, Townsend R M. 2011. A structural evaluation of a large-scale quasi-experimental microfinance initiative [J]. *Econometrica*, 79(5): 1357-1406.
- Kallus N, Zhou A. 2021. Minimax-optimal policy learning under unobserved confounding[J]. *Management Science*, 67(5): 2870-2890.
- Kitagawa T, Tetenov A. 2018. Who should be treated? Empirical welfare maximization methods for treatment choice [J]. *Econometrica*, 86(2): 591-616.
- Kitagawa T, Tetenov A. 2021. Equality-minded treatment choice[J]. *Journal of Business & Economic Statistics*, 39(2): 561-574.
- Kitagawa T, Sakaguchi S, Tetenov A. 2023. Constrained classification and policy learning [Z]. arXiv preprint arXiv: 2106.12886.
- Mbakop E, Tabord-Meehan M. 2021. Model selection for treatment choice: Penalized welfare maximization[J]. *Econometrica*, 89(2): 825-848.
- Tan-Soo J S, Cheng M D, Chen S. 2025. Evaluating the labor supply implications of a cash transfer program: Evidence from China[J]. *Journal of Development Economics*, 176: 103513.
- Zhan R H, Ren Z M, Athey S, et al. 2024. Policy learning with adaptively collected data[J]. *Management Science*, 70(8): 5270-5297.
- Zhou Z Y, Athey S, Wager S. 2023. Offline multi-action policy learning: Generalization and optimization[J]. *Operations Research*, 71(1): 148-183.

Policy Learning in the Digital Age: A Case Study of Precise Policy-Making in the China's Minimum Livelihood Guarantee

Yue Fang¹ Shiyun Hu² Liangjun Su³ Haitian Xie⁴

- (1. School of Management and Economics, The Chinese University of Hong Kong, Shenzhen;
2. National School of Development, Peking University;
3. School of Economics and Management, Tsinghua University;
4. Guanghua School of Management, Peking University)

Abstract Policy learning, as a deep integration of artificial intelligence technologies with causal inference theory, provides a scientific foundation for precision social governance in the era of big data. Grounded in the practical context of China's large-scale governance and complex policy constraints, this paper systematically reviews the international frontier literature on policy learning and proposes a policy learning algorithm based on convexification theory. This approach effectively addresses the computational optimization challenges inherent in complex, high-dimensional policy spaces, thereby significantly improving policy learning efficiency under big data and multiple-constraint settings. The paper further conducts an empirical analysis using China's Minimum Livelihood Guarantee (Dibao) program as a case study, drawing on data from the China Household Finance Survey (CHFS). The results demonstrate that the proposed method can generate highly interpretable and precisely targeted Dibao allocation schemes, effectively expanding social welfare gains and enhancing governance efficiency. This study provides an important reference for leveraging cutting-edge digital and intelligent technologies to advance the modernization of social governance in China.

JEL Classification C18, H53