

经济学研究“过度模型化”之辨：行为公设、制度框架与经验校准

赵博*

南开大学金融学院

第一版：2026年6月10日。此版：2026年7月3日

摘要

近年来，经济学研究的方法论问题重新受到关注。陆铭教授将经济学研究中的一种倾向概括为“过度模型化”：模型、识别和形式化工具本应服务于对现实问题的解释，但在一些研究中，它们本身却成为衡量论文质量和学术贡献的主要标准(陆铭, 2026)。本文以这一讨论为入口，考察经济学研究如何判断模型、识别和形式化工具的解释边界。模型化是经济学理解现实的必要方式；所谓“过度”，是指模型越过了它在特定问题中能够负责的解释范围。本文提出，经济学研究若要避免这种过度，通常需要把三项要素连接起来：关于行动者如何形成目标、信念和选择方式的行为公设，规定其选择集合和收益后果的制度框架，以及对机制量级和适用边界的经验证据校准。本文首先解释这一三元框架，并说明三者失衡如何构成“过度”；随后回顾李嘉图与熊彼得、门格尔与历史学派、米塞斯与弗里德曼等经济思想史上的方法论争论，指出这一问题并非中国经济学或当代经济学的新现象，而是经济学长期面对的基本方法论问题。接着，本文讨论现代因果识别进入历史研究时引发的争议，借此说明经验工具进入既有命题时，需要面对机制解释、制度语境和知识增量的问题。最后，本文以拍卖理论和期权定价为例，说明模型复杂本身并不一定导致过度；关键在于模型是否有清楚的行为逻辑、制度框架和持续校准机制。本文最后提出，纠正“过度模型化”需要恢复行为公设、制度框架和经验校准之间的连接。

关键词：过度模型化；经济学方法论；因果识别；经济史

1 引言

经济学关于模型的争论，常常被简化为“数学化”与“现实感”的对立。一方强调模型提供清晰假设、逻辑一致性和可检验命题；另一方则批评模型脱离现实、压缩制度

*感谢高凤龙教授、李泽广教授、马晓军教授、张庆元教授和徐良平博士就相关问题进行的讨论以及对本文的帮助。感谢许成钢教授长期以来在研究方法论方面给予的启发。文责自负。

复杂性，并在学术评价中形成技术崇拜。陆铭教授近期把这种倾向概括为“过度模型化”：模型本应服务于解释现实问题，却在一些研究中成为衡量论文质量和学术贡献的主要标准(陆铭, 2026)。¹这一问题值得认真讨论，因为它不仅关系到经济学论文如何写，也关系到经济学如何从现实中获得知识。

类似担忧也出现在国际经济学方法论讨论中。Romer对“mathiness”的批评(Romer, 2015)和Pfleiderer对“chameleon models”的讨论(Pfleiderer, 2014, 2020)，关注的都是模型结论被赋予过多解释分量的问题。前者指出，一些理论写作借助数学符号获得严谨外观，却没有让符号设定、文字解释和经验含义保持清楚对应；后者则强调，模型在原有假设环境中得到的结论，常被用于讨论假设并不成立的现实场景。Romer(2016)对宏观经济学的批评，也延续了这一关切：模型内部推导如果缺少现实约束，就可能在形式上自洽，却难以承担对现实经济的解释。

本文沿着这一问题意识提出一个判断框架。这里首先需要区分“模型化”与“过度模型化”。经济学不能没有模型。只要经济学试图解释价格、成本、激励、约束、信息、产权和预期，它就已经在从复杂现实中抽取相对稳定的行为逻辑。没有抽象，事实仍然存在，但事实之间的机制关系、反事实判断和一般性含义就会变得模糊。因此，问题的关键在于何为“过度”：模型的边界在哪里，对模型的运用何时越过了模型能够负责的范围。

为了说明这种边界，本文把经济学研究拆解为三个相互连接的环节：行为公设、制度框架和经验证据校准。行为公设，是关于行动者如何形成目标、信念、预期和选择方式的基本判断，例如偏好满足何种性质、行动者是否以及如何最大化、如何处理不确定性和信息不足；制度框架，是行动者在什么规则和组织环境中行动，它规定权利、责任、行动后果、信息披露和可行选择；经验证据校准，则回答现实中这种行为反应有多强、在什么条件下成立、是否被其他机制抵消。数据不仅为模型提供支持或反驳，也可以反过来提示模型设定、适用边界和机制理解是否需要调整。更有解释力的经济学研究，通常需要从行为公设出发，在制度框架中展开机制推理，并用经验证据校准其量级、边界和适用条件。相应地，所谓“过度”并不是三项要素中的任何一项本身过多，而是模型在这三者之间失去有效连接：或者只剩形式化的行为假设，或者忽略制度框架，或者把局部经验结果外推成完整解释。

这一表述也有助于更准确地理解陆铭教授所说的“应把是否遗漏了关键性的制度和变量，作为评判中国经济模型优劣的标准”(陆铭, 2026, 200页)。这些“关键性的制度和变量”不仅是回归方程中的控制项，也常常是让行为公设进入具体问题的规则环境。人的行为发生在特定制度规则之下；行为公设和制度框架叠加在一起，才形成具体历史情境中的行为模式。若模型只保留形式上的最优化，却遗漏改变行动者选择集合、信息环境和收益后果的制度框架，抽象就会偏离问题本身；同样，若研究只围绕可识别的数据变化展开，却没有说明何种行为公设和制度规则使这种变化具有因果含义，也会使识

¹围绕这一问题，江艇(2026)从经验研究实践出发，将“过度模型化”理解为经验研究中的模型功能过载，强调模型作为识别装置、事实筛选器和叙事约束时可能越过适用边界。该文与陆铭教授的讨论都主要关注经验实证研究中的模型使用和识别实践。本文初稿形成于该文发表前，两者切入点有所不同：本文更侧重于一般经济学研究中的方法论问题，不限于经验研究，也不限于是否使用数学或计量形式。

别与解释任务之间发生错位。

本文安排如下。第二节解释行为公设、制度框架和经验校准构成的三元框架。第三节说明“过度”如何表现为三元框架的失衡。第四节回顾经济学史中的几次重要方法论争论，说明对这种失衡的讨论是经济学长期面对的基本问题。第五节讨论因果识别在三元框架中的位置，并以量化经济史争议作为案例。第六节以拍卖理论和期权定价两个例子，说明模型复杂并不必然等同于“过度模型化”。第七节提出回应“过度模型化”的几点方法论建议。

2 行为公设、制度框架与经验校准

2.1 行为公设：经济学抽象的必要性

经济学从一开始就具有理论抽象的一面。它研究的对象不是一堆彼此孤立的事实，而是行动者如何形成目标、判断处境、产生预期，并在约束下作出选择。只要经济学试图说明“为什么某个变化会导致另一个变化”，它就已经在使用理论抽象。没有抽象，现实当然仍然丰富，但事实之间的关系很难被组织起来；没有理论，历史材料和经验数据也很难告诉我们哪些事实是关键事实，哪些关系只是表面同时发生。

经济学中的一些基础命题，往往在计量识别之前就已具有可信度。例如，“人在约束下行动”“成本上升会改变选择”“行动者会根据预期收益和风险调整行为”，这类命题的可信度很大程度上来自对人类行动逻辑的反思。需求定律就是典型例子。它讨论的是其他条件不变时，价格沿着同一条需求曲线变化对需求量的影响；而在现实数据中，观测到的价格和数量往往同时受到需求、供给和其他条件变化的影响。例如，春节前后机票价格上升而购票量也上升，可能反映的是节假日出行需求的整体上移；研究生学费上涨而报名人数仍然增加，也可能与教育回报预期、招生规模、就业环境或考生结构变化有关。经济学家不会因为价格与购买量在某些样本中呈现正相关，就立即放弃需求定律，而会追问这种相关性来自价格变化本身，还是来自收入、偏好、预期、供给约束或样本选择等因素。

这说明行为公设在经济学中具有不可替代的地位。它来源于对行动者如何形成目标、理解信息、产生信念并在约束下选择的系统反思，因此能够先把某个机制的可信性讲清楚。这里的行为公设不等同于某一种固定的完全理性最优化假设；研究者可以假定行动者完全理性、有限理性，或者采用其他关于主观价值、判断、预期和选择方式的假设，但必须把这些关于人的行为的前提说清楚。另一方面，若要进一步判断某个机制有多大、在何种制度条件下更强、是否被其他机制抵消、以及相对于替代解释是否更重要，就还需要把它放回制度框架并接受经验证据的校准。

2.2 制度框架：最优化发生在规则之中

行为公设本身通常很抽象。它告诉我们行动者会根据目标、信念和预期作出选择，却不能告诉我们在某个具体社会中，选择集合是什么，收益和成本由什么规则规定，信

息如何披露，权利如何界定，违规成本由谁承担。制度框架正是把一般行为逻辑落实到具体场景的环节。陆铭教授强调的“关键性的制度和变量”，在这个意义上指向行动者的游戏规则和可行选择集合。

中国经济研究中土地财政、户籍制度、中央地方关系、国有企业治理和金融管制等问题，之所以不能被轻易省略，正是因为它们改变了行动者的收益后果、信息环境和选择集合。制度框架不宜被当作普通“背景”或事后加入的“控制变量”；在许多应用研究中，它是行为公设能够具体展开的规则环境。一个看似标准的最优化问题，若放在不同产权制度、监管结构和组织激励下，行动者的可行策略和行为含义都会发生变化。

2.3 经验校准：佐证、量级与边界

行为公设和制度框架可以给出机制方向和反事实想象，但通常不能自动给出量级、边界和不同机制的相对重要性。产权保护影响投资，价格影响需求，规则改变行为，这些命题具有很强的理论可信度；但影响有多大、在什么条件下更强、是否被其他因素抵消、相对于其他解释是否更重要，都需要经验、历史和案例证据来校准。

因此，经验校准的任务在于把理论放回现实中称重，同时检验量级、边界和机制是否与材料相容。它包括估计量级、比较机制、发现边界条件、解释例外情形，并在结果异常时促使研究者重新审视理论、制度描述和测量方式。

总而言之，三元框架的核心，在于把行为逻辑、制度规则和经验事实纳入同一个解释过程。行为公设提供机制方向，制度框架规定机制发生的条件，经验材料则检验其量级、边界和例外，并反过来修正理论设定与制度理解。

3 “过度”如何判断：三元框架的失衡

“过度模型化”是一个有解释力但也容易含混的概念。若不加区分，它可能同时指向数学复杂、形式主义、脱离现实、识别崇拜、评价扭曲等多个问题。本文认为，讨论这一概念时应当把“模型化”和“过度”分开。模型化只是经济学组织现实、表达机制和构造反事实的方式。真正需要判断的是：在某个具体研究中，模型化是否超过了它能够解释、能够识别、能够校准、能够贡献的范围。

据此，“过度模型化”可以被理解为三元框架的失衡：

- 第一种失衡，是行为公设脱离制度框架，研究者只保留最优化或均衡形式，却没有说明行动者实际面对的规则、信息环境和选择集合。
- 第二种失衡，是制度叙述缺乏行为机制，现实被写得很复杂，但读者并不清楚哪些制度通过何种收益后果、信息环境或选择约束影响行动者。
- 第三种失衡，是经验校准替代了理论理解，研究者因为某个问题容易识别、数据可得或方法漂亮而展开研究，却没有说明这个因果关系的重要性，以及它相对于既有知识增加了什么。

- 第四种失衡，是理论公设压倒经验校准，研究者把逻辑上成立的局部机制外推为无需条件的一般结论。

因此，“过度”的实质并不是使用了太多模型，而是模型承担了超出其解释能力的认识任务。一个模型可以用来隔离并澄清某一机制，但若研究者把这种局部机制当成对现实整体的完整描述，便容易越过其适用范围；用于局部校准的识别，如果被当作完整解释，也需要额外说明其外推依据；用于简化制度的假设，如果遮蔽了会改变行为模式的关键制度，就可能削弱模型的解释力；用于表达理论的数学形式，如果替代了对问题重要性的论证，也会使形式清晰与实质贡献之间发生错位。所谓“过度”，总是相对于研究问题和解释任务而言的。

这也解释了为什么“数学太多”并不是过度模型化的本质。一个高度数学化的模型，如果清楚说明了行为假设、规则环境和可校准对象，未必过度；一个完全用文字写成的政策评论，如果把局部经验外推为普遍规律，也可能过度。形式不是问题的核心，“过度”的判断标准才是问题的核心；而本文建议的标准，就是三元连接是否成立。

4 三元失衡在经济学方法史中的反复出现

本节回顾经济思想史上几个经典方法论争论，以此说明陆铭教授所说的“过度模型化”并非当代经济学才出现的新问题。经济学从形成之初，就不断在一般理论、制度语境和经验校准之间做权衡。不同学派和作者的分歧，常常可以理解为三元结构中某一环节被过度强调，或者某一环节被放在了过弱的位置。

4.1 李嘉图式抽象与长链条推理

从经济分析方法演变的历史看，李嘉图可能是最早系统使用抽象演绎方法的经济学家之一。工资、利润、地租、比较优势和分配关系，在李嘉图那里被置于一套高度抽象的逻辑体系中。这样的处理方式使经济学能够从表面现象背后寻找结构性关系，也让经济学获得了不同于历史描述和政策评论的理论形态。

李嘉图式推理的特点，是先把现实经济压缩成少数关键关系，再沿着这些关系推导出分配和政策含义。以地租和利润理论为例，李嘉图把土地质量差异、人口增长、粮食价格、工资和利润率放入一条逻辑链条：随着人口增加，耕作扩展到较差土地，或者在同一土地上追加劳动和资本而边际回报下降；较好土地因相对于边际土地的生产率差异获得地租；粮食成本上升推高维持工人工资所需的生活资料成本；工资压力又压缩利润。这个推理常以数值例子展开，类似在一个“模型农场”中改变劳动投入、土地质量和产出分配，再观察地主、工资劳动者和资本家的收入份额如何变化。Morgan (2012) 对李嘉图“模型农场”的讨论说明，这种抽象与当时农业实验、经营账簿和关于农业改良的讨论有密切关系。这个推理的力量在于，它把看似分散的地租、工资和利润变化，组织成一个关于经济增长与收入分配的结构性解释。

但这一链条也清楚显示，长链条推理每前进一步，都需要额外的现实条件。第一，从人口增长推到边际土地扩张或农业边际收益递减，需要假定粮食供给主要受本国土

地约束，并且农业改良、固定资本投入和经营方式变化不足以抵消边际递减。Gehrke (2015) 考察了 Richard Jones 对李嘉图地租理论的批评，其核心正是农业固定资本和土地改良在理论中被忽视；这也说明，李嘉图式地租推导能否成立，取决于农业生产的技术和资本结构是否可以被简化为边际土地差异。第二，从粮价推到工资，需要假定工资主要受生活资料成本约束；若劳动力市场制度、济贫法、产业结构和工资谈判方式改变工资形成过程，这一环节就不能自动成立。第三，从工资上升推到利润率下降，还需要说明生产率提高、资本深化、产业转移或粮食贸易是否会改变工资和利润之间的剩余分配关系。英国工业革命时期农业制度和生产率变化的经济史研究表明，18 世纪末至 19 世纪中叶的农业并不是一个技术和制度静止的背景 (Allen, 2004)；长期地租资料也说明，地租变化本身需要通过实际账簿、地产记录和区域差异来加以校准，而不能只由理论链条推出 (Turner et al., 1997)。

因此，熊彼得 (Schumpeter, 1954) 用“李嘉图恶习” (Ricardian Vice) 概括一种典型的理论风险：研究者从少数高度简化的假设出发，经过一连串逻辑推演，便直接得到关于复杂现实的结论。这个批评强调的是，抽象前提和现实结论之间的距离，可能被推理的连贯性遮蔽。事实上，Reich (1980) 对李嘉图经济学衰落的讨论表明，关于地租份额和边际递减是否符合当时英国经验，历史文献中并不存在简单的一边倒结论。因此，一旦结论从模型世界进入现实世界，被放在模型外部的制度条件、技术变化、市场结构和历史事实，就必须重新进入论证。尤其当理论结论被用于现实政策判断时，这些模型外部条件就不再只是背景，而会直接影响政策推论能否成立、能否外推以及可能产生何种后果。逻辑越顺畅，越需要说明每一环节的现实约束；形式越简洁，越需要交代哪些细节被牺牲、哪些细节不能牺牲。

用本文的语言说，李嘉图式抽象所提示的失衡，是行为公设和演绎推理压倒制度框架与经验校准。研究者当然可以从少数前提出发推导机制；问题在于，推理链条越长，就越需要说明中间环节在何种制度条件下成立，以及经验事实是否支持这些环节的连接。否则，模型看上去很严密，实际承担的却可能是超出其前提的解释任务。

4.2 一般理论与历史语境的分裂

门格尔与德国历史学派的争论，是经济学史上关于理论与历史关系的经典分歧。门格尔在《社会科学方法论探究》中反对把经济学完全变成历史描述，主张经济学需要研究经济现象中的一般关系和“精确”规律 (Menger, 1985, 1883)。他的关切是可以理解的：如果经济学只记录不同国家、不同时期、不同制度下的事实差异，而缺少关于人的选择、交换、价格和制度后果的一般理论，经济学就很难从历史材料中提炼机制，也很难解释为什么相似的制度安排会产生相似或不同的后果。

德国历史学派强调经济现象的历史具体性，主张从制度、法律、社会结构和经验材料中理解经济行为，而不能把经济学化约为一套普遍抽象法则。以施穆勒为代表的历史学派面对的是 19 世纪德国的国家建构、工业化、社会问题和政策改革。他们关心的是，经济行为总是在特定法律制度、行政体系、社会阶层、组织形式和道德习俗中展开；价格、工资、企业组织、家庭结构和社会政策并不是抽象市场关系的外在附属品，而是会

改变行动者选择空间和行为模式的制度条件。Grimmer-Solem (2003) 对德国历史经济学与社会改革关系的研究说明, 历史学派的方法论主张与当时德国关于社会问题、国家能力和政策改革的具体语境密切相关。Hodgson (2001) 对经济学中“历史特殊性”问题的讨论, 也可以帮助理解这种批评的持久意义: 同一套不变的抽象命题未必能够直接覆盖不同经济制度。

因此, 这场争论指向一般行为逻辑如何进入具体制度语境。门格尔强调一般理论, 是为了避免经济学退回到事实堆叠; 历史学派强调历史语境, 是为了防止抽象理论把制度差异看成可有可无的细节。从三元框架看, 二者各自抓住了经济学研究中的一个真实环节。一般理论提供关于目标、信念、预期和选择方式的机制语言, 使研究者能够提出“在某类规则和约束下, 行动者可能如何反应”的解释; 历史语境则说明这些规则的具体内容, 以及行动者面对的组织、权力、信息和约束条件如何构成。

也正因为如此, 这里存在两种相反的失衡风险。若一般理论完全压倒历史语境, 制度差异就会被处理成无关紧要的背景, 模型中的行动者似乎可以不进入具体组织、规则和权力关系。反过来, 若历史描述完全拒绝一般行为逻辑, 研究就可能停留在事实堆叠和语境叙述中, 难以说明哪些因素真正通过何种机制产生作用。门格尔与历史学派之争的持续意义, 正在于它提醒我们: 经济学既不能把制度历史化约为抽象前提的附属品, 也不能取消关于人如何行动的一般理论。好的理论抽象, 应当能够说明自己保留了哪些一般行为逻辑, 又是在何种制度和历史条件下使用这些逻辑。

4.3 先验逻辑与经验校准的断裂

上述一般理论的重要性在米塞斯那里被推到更彻底的位置。米塞斯把经济学理解为人的行动学 (praxeology) 的一部分, 认为经济学应当从“人有目的地行动”这一先验范畴出发, 通过演绎推理建立其理论体系, 而不是依赖经验归纳来发现经济规律 (Mises, 1949)。这个立场有其力量, 因为它提醒我们, 许多经济学命题来自对人类行动逻辑的反思。例如, “自愿交换的双方之所以愿意交换, 是因为他们各自预期交换后处境改善”这类命题, 首先是包含在“有目的地行动”和“自愿交换”概念之中的逻辑判断, 而不是先由统计检验发现的经验规律。

不过, 本文认为先验逻辑有其边界: 有些底层经济命题可以通过对行动逻辑的反思确认; 但许多经济命题涉及制度结构、市场组织和经验量级, 必须由理论、制度分析与证据共同判断。先验逻辑可以说明某个机制为何可信, 却不能自动说明现实中的交换条件是否真正自愿、产权边界如何划定、规则如何执行、信息是否对称, 以及这一机制在具体历史环境中有多大作用。例如, 对于“信贷扩张在什么条件下会导致经济波动”这类复杂问题,²需要结合金融制度、政策规则和经验证据来判断其适用条件。弗里德曼在批评这种奥地利学派式方法论时, 强调经济学必须接受经验事实和预测表现的检验。Friedman (1953) 的实证经济学方法论常被理解为带有工具主义色彩, 但其核心关切

²Hayek 与 Myrdal 共同获得 1974 年诺贝尔经济学奖, 获奖理由包括他们在货币理论和经济波动理论方面的开创性工作, 以及对经济、社会和制度现象相互依赖关系的分析 (Nobel Prize Committee, 1974)。这里并不否认奥地利学派相关解释的重要性, 而是强调这类宏观命题本身具有较强的制度和历史条件依赖性。

之一，正是避免经济学变成纯粹的逻辑体系。³Friedman (1991) 在回忆与米塞斯及其追随者的分歧时，曾用一种尖锐的方式指出：如果两个坚持米塞斯式先验方法的人发生争论，而他们又拒绝看经验证据，那么他们除了打架 (fight) 之外，没有其他办法解决分歧。

本节的三个例子共同说明，经济学史上的几次重要方法论辩论，从某种程度上说都可以归结于关于行为公设、制度框架和经验校准之间的关系的讨论。李嘉图式抽象提示我们，长链条推理可能忽略现实中的其他条件约束；门格尔与历史学派之争提示我们，一般理论和历史语境可能彼此脱节；米塞斯与弗里德曼之争提示我们，先验逻辑若缺乏经验裁判，就会削弱经济学作为社会科学的可争辩性。

在当代经济学研究中，经验研究，特别是以研究设计和因果识别为核心的经验研究，在主流经济学发表结构和学术影响中占据了越来越中心的位置 (例如，Hamermesh, 2013; Angrist et al., 2017)。这一变化提高了经济学面对现实问题的能力，也使与因果识别相关的失衡变得更加突出：有时识别设计本身被放在了过于中心的位置，而研究问题的重要性、行为机制和制度语境反而说明不足。陆铭教授原文第三节专门讨论了这一问题。在此基础之上，本文将这个问题也置于三元失衡的框架之内进行进一步讨论。

5 因果识别在三元框架中的位置

5.1 识别是经验校准的一种形式

现代实证经济学的“可信性革命”极大改变了经济学论文的研究范式。Angrist and Pischke (2009) 强调，经济学需要更可信的研究设计，减少对随意控制变量相关性分析的依赖。自然实验、工具变量、断点回归、双重差分等方法，使经济学家能够更严肃地面对内生性问题。与此相配套，潜在结果框架为“处理”“反事实”和“平均处理效应”等概念提供了更清楚的语言 (Rubin, 1974; Holland, 1986; Imbens and Rubin, 2015)。

这一进步不可低估。没有因果推断，许多经验研究就会停留在描述性相关层面；没有对反事实的认真构造，政策评价很容易被选择偏误、共同趋势和遗漏变量误导。现代因果推断迫使研究者说明：处理是什么，结果是什么，反事实是什么，识别假设是什么，哪些样本支撑了估计。这些要求实质性提高了经验研究的论证标准。

不过，因果识别的可信性并不由估计量本身自动保证，而取决于识别设定是否可信。在线性回归中，由于扰动项不可观测，解释变量与扰动项不相关不能由回归方程自己保证；在工具变量法中，工具变量的外生性不是一个可以由估计结果直接验证的统计结论，而是关于工具变量来源、作用路径和未观测因素关系的概念性假设；在双重差分中，平行趋势假设常常需要通过事前趋势、制度背景和反事实想象来增强可信度；在断点回归中，临界点附近个体是否不可操纵、是否局部可比，也需要结合制度规则和行为

³本文在此只取其关于现实反馈的最低限度主张，而不把预测表现视为评价经济理论的唯一标准。由于经济现象存在归因困难和内生性问题，机制与制度条件的讨论仍然不可省略。Coase (1994, 16-17) 在《经济学家应当如何选择?》中指出，理论不能像航空或公交时刻表那样只按预测准确性评价；理论还应帮助我们组织思考，理解系统如何运转。

逻辑来判断。也就是说，因果识别越强调设计，越需要说明制度规则如何运行、行动者会如何反应、样本可比性从何而来，以及反事实为何可信。

由此可见，在本文的框架中，因果识别更适合作为经验证据校准的一种形式来理解。它可以帮助我们估计某种机制的量级，比较不同解释的相对重要性，并说明某个制度变化在特定样本、时期和规则下产生了什么影响；这些估计之所以可信，则取决于识别设定能否得到行为公设和制度框架的支撑。换言之，反事实是否成立、样本是否可比、冲击是否可以视为外生，往往要回到行动者如何反应、制度规则如何运行以及相关变量如何产生这些问题上说明。正是在这个意义上，“因果”计量经济学进入历史研究时，会更集中地反映识别设定与历史语境之间的关系。

5.2 量化经济史争议：识别进入其他学科旧命题

计量经济学的因果识别处理的是社会科学共同面对的反事实和内生性问题，因此原则上可以进入历史学、政治学、社会学等研究对象。量化经济史把历史事件、制度变迁、战争冲击和地方差异转化为可度量、可比较、可估计的经验对象，也因此成为现代因果识别进入其他学科传统问题的典型场景。这里的特殊之处在于，许多问题原本已有深厚的史学研究传统；因果识别若作为经验校准进入这些问题，就应该说明自己校准了什么，改变了对问题的什么理解。量化经济史争议之所以适合作为方法论案例，正是因为它把这一点暴露得很清楚：经济学者进入历史问题时，通常会强调历史研究不够“正式”（formal），或者难以识别因果；历史学者则常常认为，经济学为了干净识别裁剪掉了真正重要的语境，而且最终得到的结论不过是史学界早已知道的常识。

本文认为，这个分歧不应被理解为“经济学有因果，历史学没有因果”，或者“经济学研究‘正式’，历史学研究‘不正式’”。计量经济学中谈论的“因果分析”，并不是社会科学中唯一的因果分析。历史研究的核心，是确认史料的真实性与完整性，从史实出发重建事件脉络，并在多源材料、时间顺序和行动者语境中判断解释是否成立。其背后同样有关于人的行为的公设，例如行动者会利己、避害、追求权力、维护声誉，并会在其信息位置和组织位置下行动；有时还会通过具体史料确认某些行动者的偏好、信念和行为模式，再据此推断他们在关键事件中的选择。换言之，历史学也有“行为公设-演绎分析”，只是这种分析通常不以方程或显式公理的形式出现，而是嵌入史料批判、过程追踪和语境叙述之中。Mahoney (2012) 和 Ermakoff (2019) 对历史社会科学中因果推断方式的讨论说明，这类判断未必表现为回归方程，但并非无规则的主观判断。因此，经济学识别进入历史研究时，若只是把一个既有命题重新估计一遍，却没有说明自己在量级、边界、其他候选解释或反常识发现上增加了什么，历史学者自然会觉得这种贡献有限。

Acemoglu 等关于第一次世界大战后意大利政治变化的研究，是一个典型案例。Acemoglu et al. (2022) 研究战后政治动员、地方精英反应与意大利法西斯主义兴起之间的关系，文章提供了清楚的实证设计，也把一个历史命题转化为可识别、可估计的实证问题。该研究发表后，引发了史学界和跨学科公共讨论中的批评。⁴ 批评者关心的并不只是识

⁴这类批评较多表现为公共学术讨论和跨学科评论，而不是正式期刊论文或者评论。公共讨论中，研究

别技术，而是“贡献归属”和“知识地图”：如果地方精英、地主组织、战后政治动员和法西斯主义崛起之间的关系，在意大利史、法西斯主义史和历史社会科学中已经长期被讨论，那么经济学论文需要说明自己新增的是量级估计、机制比较、反常识发现，还是对既有解释的修正。

这种争议并不是孤立个案。Bai et al. (2023) 研究曾国藩、湘军动员与战后政治权力分布。文章把个人精英网络如何影响国家层面政治经济结果的问题，转化为可度量网络与县级结果之间的实证关系。围绕该文再次产生了历史学和经济学的争论。耶鲁大学张泰苏教授提问：在史学界已有大量关于湘军、团练、地方军事化、督抚政治、厘金财政和晚清权力结构研究的情况下，经济学论文相对于这些研究的知识增量何在？若论文结论大体上接近史学界已有判断，经济学贡献就不能只说“我们做了识别”，而要进一步说明识别提供了什么新的数量关系、反事实比较或理论修正。在这个案例中，历史学者很可能并不否认精英网络会影响政治权力，也不否认曾国藩和湘军网络在晚清政治中重要；这些判断本来就是史学叙述的一部分。争议的核心在于，经济学论文通过可识别的经验设计，提供了何种知识增量：新的校准、旧争议的澄清，还是只是把已有的史学共识结论转化为经济学可发表的形式。⁵

因此，本文认为，经济学进入历史学、政治学或社会学的传统问题时，不能仅仅以“我们从经济学角度做了可识别的估计”作为充分辩护。量化历史研究之所以不只是普通经济研究，正是因为它声称要用现代经济学方法重新理解历史命题；既然如此，它就需要把自己的发现放回既有知识地图中说明边际贡献。这也把讨论带回因果识别本身：经济学可能带来的新知识，不只来自估计量，更来自识别假设如何被说明、检验和放回历史语境。

5.3 识别假设与历史语境的对接

由此来看，量化历史研究或许可以把识别设计更明确地表述为一组条件：哪些个体或地区可以比较，哪些冲击可以视为外生，哪些机制需要被排除，何种反事实相对可信。若这些条件能够同既有历史研究中的史料判断、过程叙述和机制解释逐项对照，经济学

一战时期意大利史的 Vanda Wilcox 曾以讽刺口吻指出，仿佛“经济学家阅读史学是违法的”，否则很难解释为何把陈旧的史学解释呈现为新的经济学发现；欧洲当代史学者 Holger Nehring 则建议回到 Paul Corner 的经典研究，并对类似论点被呈现为新颖贡献感到惊讶。Anup Malani 认为，若在缺少充分引用的情况下检验一个旧史学假设，并使其看起来像新的经济学贡献，就会给人以经济学傲慢的印象；Claire Adida 则质疑，摘要中的“我们认为”是否应更准确地写成“我们检验一个长期存在的争论”。这些评论主要来自社交媒体和公共学术讨论；计量经济圈 (2023) 对这场讨论作了整理，本文据此转述其大意，只把它们作为跨学科争议的材料，而不作为正式学术反驳。Acemoglu et al. (2020) 在介绍该研究时也承认，已有若干史学和历史社会科学研究强调了战后政治恐惧与法西斯兴起之间的联系；其所列文献包括 Snowden (1972)、Comer (1975)、Cardoza (1982) 和 Lyttelton (2003) 等。一战意大利背景可参见 Wilcox (2021)，法西斯主义解释传统另见 De Felice (1977)。

⁵ 搜狐转载整理了张泰苏教授与作者之一贾瑞雪教授围绕该文的公开讨论。张泰苏教授的追问集中在：文章的“量化结论和传统史学中的共识有何不同”，以及这一套实证究竟“刷新了我们哪些传统共识”；贾瑞雪教授则回应说，论文试图把微观网络与宏观权力结构联系起来，并且“需要使用经济学的语言与同行对话”（搜狐，2023）。聂辉华 (2023) 对这个争议进行了概述，把分歧放在学科评价标准上：经济学界是否发表一篇论文，关键在于它对经济学界是否有贡献；经济学论文通常需要借助数学工具或数据证明因果关系，并揭示具体机制。

研究便有机会说明既有结论在什么范围内更可能成立，在什么边界上可能需要修正，以及哪些候选解释较难与材料相容。

从这个角度看，量化经济史的贡献可以不只来自估计量本身，也来自识别假设、历史语境和经验结果之间的相互校准。当相关条件较为可信时，已有历史判断可以获得更清楚的量级估计和边界说明；当条件与史料脉络存在冲突时，争论也可以更具体地落到若干识别假设上。这种改进也许有助于展示量化经济史如何生成新的知识，并减少“用复杂方法证明常识”的观感。

6 复杂模型何以不必然“过度”

前文的讨论强调，“过度”的问题不在于数学复杂本身，而在于模型是否越过了自己的解释边界。反之，不容易走向过度的模型，通常具有几个共同特征：它面对的是边界相对清楚的问题，而不是直接承担对整个社会或整个经济系统的完整解释；它的行为公设和制度规则能够被明确写出来，关键假设在现实中有可理解的含义；它的结果能够在实践、数据或反复使用中接受反馈，并促使研究者根据偏差重新修正模型设定、制度描述或适用范围。

在这种意义上，复杂性本身可以成为一种解释资源。数学形式如果服务于明确的问题，就可以帮助研究者更精确地表达策略、约束、价格关系和反事实条件。拍卖理论和 Black-Scholes-Merton 期权定价公式，分别提供了制度设计和金融定价中的两个例子。

6.1 拍卖理论

拍卖模型通常面对一个较窄的问题：在给定物品、潜在买家、信息结构和拍卖规则时，参与者会怎样出价，物品会配置给谁，卖方能获得多少收入，规则是否会诱发策略性行为。我们可从行为公设、制度框架和经验校准观察这类模型的建模方式。

行为公设很简洁：投标人会根据自己的估值和对获胜可能性的判断，选择对自己更有利的出价策略。这一假设在进入具体规则之后有了分析内容：什么信息被公开，投标人是否知道其他人的估值分布，报价是否可以被观察和修改，获胜者如何支付，是否存在保留价，投标人能否串通，都会改变“最大化收益”的具体含义。

在这个例子中，拍卖者的行为不一定只是私人收入最大化：政府出售频谱、矿产或国有资产时，可能同时关心财政收入、配置效率、竞争结构、公共服务覆盖和后续市场集中度。同时，“制度框架”也成为模型要分析的对象本身，而不是模型外部的背景介绍。信息披露规则、报价能否修改、支付规则、进入门槛、保留价、反串通安排等，都是制度设计的一部分。Vickrey (1961) 对拍卖规则与真实报价关系的分析，Myerson (1981) 对最优拍卖设计的讨论，以及 Milgrom and Weber (1982) 对共同价值、关联价值和信息结构的研究，反映了不同的最大化收益假设在不同信息环境和规则设计下的不同结果。形式化在这里的作用，是把规则、信息和策略之间的关系说清楚。

拍卖理论的经验证据校准，来自真实拍卖留下的可观察材料：有多少潜在买家进入，谁选择退出，各轮报价如何变化，成交价格是否充分反映竞争，报价中是否出现默

契分割或信号传递，赢家事后是否出现支付困难、经营失败或资产转售，最终配置是否符合效率、竞争和公共服务目标。这些材料会反过来检验模型中的估值分布、信息结构、进入成本、共同价值成分、互补性和合谋可能性。若参与不足，模型需要重新考虑进入门槛和资格规则；若价格路径显示需求缩减或默契分割，就需要调整公开信息、报价格和活动规则；若赢家事后经营困难，赢家诅咒和共同价值问题就不能被忽略；若互补资产被拆散，则需要重新考虑标的划分或组合出价。20世纪90年代以来，频谱拍卖成为机制设计进入公共政策的重要场景。McMillan (1994) 讨论了美国频谱权利出售中的经济学设计问题；2020年诺贝尔经济学奖官方介绍也强调，Milgrom 和 Wilson 的拍卖理论改进被用于无线电频谱等复杂资产拍卖，并影响了全球拍卖设计实践 (Nobel Prize Committee, 2020)。这些应用说明，复杂模型可以在一个制度边界清楚的场景中通过出价数据、成交结果和事后续效不断接受反馈，并据此修正规则设计。

6.2 期权定价：无套利原则与持续市场校准

期权定价模型提供了另一个不同类型的例子。期权收益取决于未来标的资产价格，其定价涉及未来价格、波动、利率、到期时间和交易条件。Black and Scholes (1973) 和 Merton (1973) 用无套利假设把定价问题转化为局部价格约束：如果一种期权的收益可以通过标的资产和无风险资产的动态组合复制，那么期权价格就不能任意偏离复制组合的成本。

这里的行为公设是明确的：如果存在可以被利用的无风险套利机会，市场参与者会试图利用它，买入被低估资产、卖出被高估资产，直到价格关系被压回无套利范围之内。这个公设只要求在特定交易条件下，明显的无风险套利机会难以长期存在。正因如此，Black-Scholes-Merton 公式的解释力量来自一个很窄但很强的机制：套利约束使不同金融资产价格之间形成特定关系。

另一方面，无套利行为依赖制度和市场条件。动态复制需要交易者能够买卖标的资产，能够借贷或持有无风险资产，能够进行足够频繁的再平衡，并且交易成本、卖空限制、流动性冲击和价格跳跃不会破坏复制关系。无套利定价是在具体金融产品、交易制度和市场摩擦条件下成立的局部推理。制度条件越接近模型要求，公式越有解释力；制度条件偏离越大，模型越需要修正。

期权市场也会持续校准模型。交易者可以把市场报价倒推出隐含波动率，观察不同执行价和期限上的波动率曲面；也可以用模型给出的 delta、gamma 等风险敏感度构造对冲组合，再用实际对冲损益检验复制是否有效。若隐含波动率随执行价系统变化，常数波动率假设就需要修正；若对冲误差在流动性紧张或价格跳跃时集中出现，连续交易、无交易成本和正态价格过程等假设就受到挑战；若不同期限的期权价格显示出不同风险补偿，利率、波动率期限结构和市场摩擦也需要进入模型。Black-Scholes-Merton 公式并没有完整描摹期权市场，但它进入报价、做市、对冲和风险管理之后，市场数据不断把模型偏差暴露出来。MacKenzie (2006) 用“引擎而非照相机”来概括金融模型与市场实践之间的关系，正说明模型在参与塑造市场的同时，也会被市场实践反复校准。

6.3 小结

这两个例子共同说明，复杂模型之所以不必然过度，是因为复杂性服务于清楚的解释任务。拍卖模型的数学结构用于表达规则、信息环境和策略之间的关系；期权定价模型的数学结构用于表达无套利、复制组合和价格约束之间的关系。它们的力量不在于把现实世界全部装进模型，而在于把一个局部问题的行为机制、制度条件和可校准结果清楚地连接起来。

由此看，“过度”的判断并不取决于模型是否复杂、是否数学，而取决于复杂性是否仍然服务于行为机制、制度条件和经验校准之间的连接。数学可以使模型更精确，也可以使模型更封闭；区别在于模型是否清楚说明自己试图解释什么、在何种条件下解释，以及如何接受现实反馈。一个复杂模型若能把解释任务限定在自己能够负责的范围内，并在使用中根据反馈修正设定和适用范围，就不必然构成过度模型化；真正需要警惕的，是长链条、多假设、弱校准的外推：模型越想解释更宏大的历史、制度和社会结果，越需要说明外推的依据。

7 回应“过度模型化”的几点建议

在上述框架下，陆铭教授关于“过度模型化”的批评可以被理解为对若干方法论风险的提醒。陆铭(2026)所说的“为了让模型成立而忽略现实”和“为了让方法成立而牺牲问题”，分别对应两种典型失衡：前者是行为公设或形式模型脱离制度框架，后者是经验识别或技术方法压倒问题本身。类似风险在中国经济研究中尤其值得重视。中国经济中的转型过程、区域差异、央地关系和金融监管体系，都会影响行动者的选择集合、收益后果和激励结构。若直接套用在其他市场或制度环境中形成的模型，就容易使模型设定与现实规则发生错位，进而影响研究结论的解释。

不过，对“过度模型化”的批评也不应被理解为反模型、反数学或反因果识别。许多有现实解释力的研究，恰恰依赖清楚的模型来组织机制；案例研究、调查研究和历史叙事若缺少理论组织，同样可能选择性呈现事实或误导因果判断。本文最后把上述讨论概括为几点建议。要使研究更有解释力，关键在于让行为公设、制度框架和经验校准形成一个可追问、可争辩、可修正的解释结构。

第一，从问题出发。研究最好从真实经济现象或能够澄清现实机制的理论疑问出发，避免主要由现成方法牵引。一个问题的重要性，不应由数据是否现成、识别是否便利或模型是否容易展开来决定，而应取决于它是否能改变我们对行动者选择、制度规则、市场机制或经验事实的理解。

第二，交代行为假设。研究应当让读者看出自己依赖怎样的行为公设：行动者依据什么目标和信念行事，面对激励、约束、信息或风险时如何产生相应反应。这种公设不一定必须写成数学模型，但需要足够清楚，使读者能够判断从原因到结果之间的行为推理是否成立。

第三，交代制度条件。研究需要说明行动者所处的规则约束、权利安排、信息环境、行动后果和可行选择集合。所谓“关键性的制度和变量”，不能只是事后补充的背景介

绍；它们常常构成行为推理能否成立的前提。

第四，接受经验校准。理论推理和机制解释都需要接受事实和经验证据的反馈。研究者应当尽量说明哪些条件下结论成立，哪些条件下可能失效；当经验结果与理论预期不一致时，也需要回到行为公设、制度描述、数据情况和识别假设中检查原因。

第五，说明边际贡献。研究者应当说明自己的发现相对于既有知识的边际贡献。若研究的是一个已有稳定共识的问题，就需要说明新增的是量级估计、机制比较、边界条件、反直觉异质性，还是对既有解释的修正。尤其在跨学科研究中，不能把其他学科长期积累的解释简单当作未经检验的背景材料。

8 结论

本文从“过度模型化”的讨论出发，试图把问题从是否使用模型、数学或因果识别，转向模型和方法在具体研究中承担何种解释任务。文章提出的三元框架，强调经济学研究需要同时处理行动者如何选择、行动者在何种规则中选择，以及理论机制在现实中具有多大作用、适用边界何在。

由此得到的结论是：经济学需要模型，也需要识别和叙述；问题在于它们能否相互约束。模型应当说明自己的行为公设和制度条件，经验识别应当说明自己校准的是何种机制及其外推范围，历史和制度叙述也应当说明复杂语境如何进入行动者的选择过程。纠正“过度模型化”的方向，不是简单减少模型，而是使模型、证据和制度语境形成更清楚的解释关系。

参考文献

- Daron Acemoglu, Giuseppe De Feo, Giacomo De Luca, and Gianluca Russo. Revisiting the rise of Italian fascism. VoxEU.org, October 2020.
- Daron Acemoglu, Giuseppe De Feo, Giacomo De Luca, and Gianluca Russo. War, socialism, and the rise of fascism: An empirical exploration. *Quarterly Journal of Economics*, 137(2): 1233–1296, 2022. doi: 10.1093/qje/qjac001.
- Robert C. Allen. Agriculture during the industrial revolution, 1700–1850. In Roderick Floud and Paul Johnson, editors, *The Cambridge Economic History of Modern Britain*, volume 1, pages 96–116. Cambridge University Press, Cambridge, 2004. doi: 10.1017/CHOL9780521820363.005.
- Joshua Angrist, Pierre Azoulay, Glenn Ellison, Ryan Hill, and Susan Feng Lu. Economic research evolves: Fields and styles. *American Economic Review*, 107(5):293–297, 2017. doi: 10.1257/aer.p20171117.

- Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, 2009.
- Ying Bai, Ruixue Jia, and Jiaojiao Yang. Web of power: How elite networks shaped war and politics in china. *Quarterly Journal of Economics*, 138(2):1067–1108, 2023. doi: 10.1093/qje/qjac041.
- Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. doi: 10.1086/260062.
- Anthony L. Cardoza. *Agrarian Elites and Italian Fascism: The Province of Bologna, 1901–1926*. Princeton University Press, Princeton, NJ, 1982.
- R. H. Coase. How should economists choose? In *Essays on Economics and Economists*, pages 15–33. University of Chicago Press, Chicago, 1994.
- Paul Corner. *Fascism in Ferrara, 1915–1925*. Oxford University Press, London and New York, 1975.
- Renzo De Felice. *Interpretations of Fascism*. Harvard University Press, Cambridge, MA, 1977.
- Ivan Ermakoff. Causality and history: Modes of causal investigation in historical social sciences. *Annual Review of Sociology*, 45:581–606, 2019. doi: 10.1146/annurev-soc-073117-041140.
- Milton Friedman. The methodology of positive economics. In *Essays in Positive Economics*, pages 3–43. University of Chicago Press, Chicago, 1953.
- Milton Friedman. Say “no” to intolerance. *Liberty*, 4(6):17–20, 1991.
- Christian Gehrke. Fixed capital in agriculture: Richard jones’s critique of ricardo’s theory of rent. *Journal of the History of Economic Thought*, 37(3):411–430, 2015. doi: 10.1017/S1053837215000243.
- Erik Grimmer-Solem. *The Rise of Historical Economics and Social Reform in Germany, 1864–1894*. Clarendon Press, Oxford, 2003. doi: 10.1093/acprof:oso/9780199260416.001.0001.
- Daniel S. Hamermesh. Six decades of top economics publishing: Who and how? *Journal of Economic Literature*, 51(1):162–172, 2013. doi: 10.1257/jel.51.1.162.
- Geoffrey M. Hodgson. *How Economics Forgot History: The Problem of Historical Specificity in Social Science*. Routledge, London, 2001.
- Paul W. Holland. Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960, 1986. doi: 10.1080/01621459.1986.10478354.

- Guido W. Imbens and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge, 2015.
- Adrian Lyttelton. *The Seizure of Power: Fascism in Italy, 1919–1929*. Routledge, London, 3 edition, 2003.
- Donald MacKenzie. *An Engine, Not a Camera: How Financial Models Shape Markets*. MIT Press, Cambridge, MA, 2006.
- James Mahoney. The logic of process tracing tests in the social sciences. *Sociological Methods & Research*, 41(4):570–597, 2012. doi: 10.1177/0049124112437709.
- John McMillan. Selling spectrum rights. *Journal of Economic Perspectives*, 8(3):145–162, 1994. doi: 10.1257/jep.8.3.145.
- Carl Menger. *Investigations into the Method of the Social Sciences with Special Reference to Economics*. New York University Press, New York, 1985. Originally published as *Untersuchungen über die Methode der Socialwissenschaften und der politischen Oekonomie insbesondere*; edited by Louis Schneider; translated by Francis J. Nock.
- Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973. doi: 10.2307/3003143.
- Paul R. Milgrom and Robert J. Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, 1982. doi: 10.2307/1911865.
- Ludwig von Mises. *Human Action: A Treatise on Economics*. Yale University Press, New Haven, 1949.
- Mary S. Morgan. *The World in the Model: How Economists Work and Think*. Cambridge University Press, Cambridge, 2012. doi: 10.1017/CBO9781139026185.
- Roger B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. doi: 10.1287/moor.6.1.58.
- Nobel Prize Committee. The sveriges riksbank prize in economic sciences in memory of alfred nobel 1974, 1974.
- Nobel Prize Committee. The prize in economic sciences 2020: Popular science background, 2020.
- Paul Pfleiderer. Chameleons: The misuse of theoretical models in finance and economics. *SSRN Electronic Journal*, 2014. doi: 10.2139/ssrn.2414731.
- Paul Pfleiderer. Chameleons: The misuse of theoretical models in finance and economics. *Economica*, 87(345):81–107, 2020. doi: 10.1111/ecca.12295.

- Michael Reich. Empirical and ideological elements in the decline of ricardian economics. *Review of Radical Political Economics*, 12(3):1–14, 1980. doi: 10.1177/048661348001200301.
- Paul M. Romer. Mathiness in the theory of economic growth. *American Economic Review*, 105(5):89–93, 2015. doi: 10.1257/aer.p20151066.
- Paul M. Romer. The trouble with macroeconomics. Working paper, Stern School of Business, New York University, 2016.
- Donald B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5):688–701, 1974. doi: 10.1037/h0037350.
- Joseph A. Schumpeter. *History of Economic Analysis*. Oxford University Press, New York, 1954.
- Frank M. Snowden. On the social origins of agrarian fascism in italy. *European Journal of Sociology*, 13(2):268–295, 1972.
- M. E. Turner, J. V. Beckett, and B. Afton. *Agricultural Rent in England, 1690–1914*. Cambridge University Press, Cambridge, 1997. doi: 10.1017/CBO9780511549205.
- William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, 1961. doi: 10.1111/j.1540-6261.1961.tb02789.x.
- Vanda Wilcox. *The Italian Empire and the Great War*. Oxford University Press, Oxford, 2021.
- 搜狐. 耶鲁教授质疑“曾国藩”QJE文章，作者迅速做出学术回应. 搜狐网, 2023.
- 江艇. 对“经济学过度模型化”的再讨论. *中国社会科学*, pages 185–195, 2026. 第4期.
- 聂辉华. 社会科学没有终极真理，只有不同眼镜：对一篇经济史文章争论的评论. 财新博客, 2023.
- 计量经济圈. Acemoglu, 22年QJE文章被史学家爆炸了!. 计量经济圈, January 2023.
- 陆铭. 经济学研究“过度模型化”的误区及其纠正. *中国社会科学*, pages 196–203, 2026. 第2期.