

# 经济学研究中“数学滥用”现象及反思

■ 上海财经大学金融学院 陆蓉 邓鸣茂

▶▶▶ [上接12月1日03版]

## （二）数学模型过度运用

数学模型在经济学研究中只是方法和手段,核心依然是经济问题或者故事本身。如果结论本来已经很清楚,引入数学模型并没有得到更有价值的结论或增进逻辑联系,只是为了追求形式上的“科学化”,就属于数学模型过度运用。数学模型过度运用的原因在于模糊了度和手段之间的关系。模型本身并没有错,但研究者忽视了模型的适用边界;或与研究问题相关的数学模型并不存在,硬套模型产生问题。例如,某些经济史学的研究,生搬硬套数学模型,虽然数理分析也是认识历史的途径之一,但是最近几年非历史主义的形式化倾向,研究看似很有科学性,实质并没有在史料发掘以及“以史为鉴”上的边际贡献。

国内外许多学者已经开始反思理论模型的过度运用问题。Caballero(2010)认为,越来越多的学者在研究宏观经济问题时,过度沉迷于动态随机一般均衡模型(DSGE)的内部逻辑,混淆了这一模型在经济现实中的准确性和适用性。Krugman(2009)认为,经济学家有所迷迭的重要原因在于,宏观经济学错误地把披着数学外衣的完美主义理论当作真理和现实。例如,一些学者沉迷于理论模型,得出了诸如经济大萧条的原因在于假期过多这样的极端结论。冯俊新等(2011)指出,经济学者在研究宏观经济问题时,必须防止过度沉迷于数学模型的内部逻辑。尹世杰(2005)认为,本身的逻辑和在联系。尹世杰(2005)认为,本身的逻辑可以而且应该运用数学这个重要的工具来分析经济问题,但是决不能、也不应该“数字化”经济学这门极其重要的学科。贾根良和徐尚(2005)认为,经济学研究应该反思过分沉溺于数学技巧的弊端,倡导回归经济问题本身,拒绝将数学术语作为科学标志与科学主义观念。

## （三）实证研究与理论相脱节

2013年诺贝尔经济学奖授予美国经济学家 Eugene Fama、Lars Hansen 和 Robert Shiller,以表彰他们在“资产价格的实证分析”中做出的贡献。说明实证研究方法在经济学领域越来越重要。余广源和范子英(2017)统计了中国所有“海归”经济学者 1984-2015 年发表的美文论文,发现 2005 年以后,经济学的实证研究已经超过了理论研究,成为主流的研究方法。韩德瑞(Hendry, 1998)指出,现实经济活动中存在一些规律,经济机制的这种规律可以观测并度量,这种可测的机制称作数据生成过程(DGP, Data Generating Process)。洪永淼(2007)认为,现代计量经济学建立在两个公理基础上:(1)任何经济系统都可以看作是服从一定概率分布的随机过程(Stochastic process);(2)任何经济现象(经济数据)都可以看作是这个随机数据生成过程。虽然这两个公理是无法验证的,但是大多数计量经济学家都比较认同,计量经济学模型的设立必须满足数据生成过程。通俗的理解就是实证研究要与经济理论紧密联系。

目前我国经济学研究中一个较普遍的问题是,实证研究没有经济理论的支撑。例如,研究互联网对国际贸易的影响,仅仅将互联网作为解释变量来构造回归模型,就得出了结论,忽视了具体传导机制的研究。再如,研究金融市场变量之间的动态相关性,没有经济理论阐释,也没有事先对数据特征进行分析,就直接对数据进行单位根检验、格兰杰(Granger)因果关系检验、协整(Co-integration)检验、拟合误差修正模型(ECM),最后得出变量 A 与 B 之间存在因果关系、协整关系以及变量之间的动态修正机制。然而,这些模型只能证明动态相关性可能存在,但具体原因是什么、影响机制是什么,都没有理论支持。还有一些研究用向量自回归(VAR)以及结构向量自回归(SVAR)模型的估计系数来解释变量之间的因果关系,但由于内生性的存在,估计的系数实际上并无经济意义。实证模型总会产生结果,但并不意味着经济变量间真的存在这些关系。这些问题在于,只用实证模型回答了“是什么”的问题,忽视了由经济理论所揭示的“为什么”的机制。这些与经济理论相脱节的实证模型结果在经济学原理上难以解释,很难让人相信,也对理解社会经济运行没有什么帮助。

国外也有学者对实证研究与经济理论相脱离问题进行了批判。例如,Desen(2008)认为,用动态随机一般均衡模型(DSGE)来研究金融危机的成因或者短期宏观经济的剧烈波动并没有太大用处,因为这些模型是在封闭、完美状态下的,没有考虑摩擦因素。此外,这些模型都把 GDP 增长率波动归因于全要素生产率(TFP)的波动,却没有解释 TFP 波动的原因,因而这些研究并不能揭示经济衰退的深层次原因。

## （四）实证过程不规范

实证研究主要是为了验证现实是否符合理论预期,揭示经济变量之间的关系或探寻经济运行的机理。无论采用哪种实证方法,研究内容的重要性仍然大于形式。高质量的经济学实证研究,不仅强调研究内容的重要性,而且注重论证过程的规范性,一般由引言、文献回顾与研究设计、数据描述、实证分析、稳健性检验以及结论等几大部分构成,并且论证过程严谨、规范。然而,一些低质量的实证研究,虽然看上去包含实证研究过程,但是仔细阅读论文,实证过程不规范、实证结果阐述不严谨的现象还是经常可见。



验以及结论等几大部分构成,并且论证过程严谨、规范。然而,一些低质量的实证研究,虽然看上去包含实证研究过程,但是仔细阅读论文,实证过程不规范、实证结果阐述不严谨的现象还是经常可见。

举几个比较常见的例子:(1) 回归分析结果的表述错误。在回归分析中,如果变量 X 的回归系数为正,但是不显著,这种情况下如何表述结果? 是表述成 X 对 Y 没有影响,还是表述成有正向影响但不显著呢? 原假设一般设定为没有影响,统计学上不显著,说明不能拒绝原假设,而不是接受原假设,因此第一种表述不正确。既然统计意义不显著,那么是正向影响还是负向影响已经没有意义,因此第二种表述也不正确。正确的表述方式是, X 对 Y 的影响在统计意义上不显著。(2) 协整关系等同于均衡关系。王美今和林建浩(2012)指出,这种等同存在三方面的问题,因此协整关系并不等同于均衡关系⑥。(3) 内生性问题与工具变量(IV)。内生性问题存在的原因包括解释变量与被解释变量之间互为因果、遗漏解释变量以及解释变量的测量误差。寻找工具变量是解决内生性问题的一种较为常见的做法。好的工具变量必须与随机扰动项正交,且与内生解释变量具有较强的相关性。经济变量之间的关系错综复杂,多数经济学研究都会存在内生性问题。然而,多数研究却缺少内生性问题的排查,没有对工具变量的适用性进行评估,更有甚者,所选用的工具变量不仅没有解决原来的内生性,还带来了更加严重的内生性。

实际上,“数学滥用”的现象不仅在经济学研究中,在其他人文社会科学,如管理学、法学、史学等研究中也在一定程度上存在。现实中,管理实践有很强的独立性,套用数学模型或大样本的统计方法,往往会掩盖成功管理特质的独特性,得到的结论反而没有借鉴意义。所以,案例分析在管理学领域的很多研究中可能更加适用。黄宗智和高原(2015)认为,法学研究应该拒绝数学形式主义的方法,应立足于经验证据,首先从归纳出发,然后再演繹逻辑导出可靠的推断和假说,最后再返回到经验世界中去检验。徐祥运(2005)认为,哲学社会科学有其固有的特征和界限,是一个多参数多变量的动态体系,难以用数学形式表达,反对在哲学社会科学研究中不切实际的“滥用数学”。

## 四、经济学研究中“数学滥用”的负面影响

如果“数学滥用”只是少数情况,不影响最终达成科学的、一致的结论,那么其危害也仅仅是局部的、暂时的。不幸的是,“柠檬市场”理论表明,随着“数学滥用”的不断增多,区分“数学滥用”和传统(严谨的)数理经济学将会十分困难。

“数学滥用”问题会随着研究者的学习、引用而自我加强,并通过教学活动产生代际影响,经由学术期刊等传播平台进一步放大这种不利影响,最终影响政策制定、经济决策,影响经济生活中的每个人。Romer(2015)也指出,“数学滥用”的影响可能是普遍的、永久的。

### （一）对经济学学术研究的负面影响

经济学研究中引入数学的一个重要原因是数学能够准确、简洁且有效地表达观点,使逻辑推理更加严谨。然而数学工具的引入也存在方法论的导向性问题,能否运用数学模型来分析经济关系逐渐成为判断一项研究学术价值的重要依据(林毅夫,2005)。这种导向的负面作用,导致了经济学研究者对“数学滥用”现象。“数学滥用”会使学者忽视对经济学问题本身的深入思考,扭曲经济学者思考问题的思维模式。比如经济问题首先应该思考的是:这个因素的真实内涵是什么,存在何种影响机制?带着这样的思考去调查研究,才可能得到有意义的结论。而沉迷于数学工具者,首先不是关心经济运行的真实情况如何,而是首先思考:这个因素可以量化吗,需要采用什么模型,改变哪些条件可以实现自己想要的结果。如果不

能量化,就放弃了此问题的研究。更有甚者,则通过玩弄数学模型达到想要的结论。久而久之,学者会失去对问题的创新性见解和敏锐的洞察力,研究越来越脱离实际。因此,“数学滥用”会阻碍思想的创新,导致的结果就是产生出有大量包含冗余数理模型的经济学论文,而真正有思想的文章不多。正如保罗·罗默(2015)批判的那样,“经济学研究中的‘数学滥用’,忽视了紧密的逻辑演绎,往往导致逻辑滑坡。如果这种不严谨甚至是学术不端持续下去,数学模型就会丧失解释力和说服力。”

经济学期刊放大了“数学滥用”对于经济学研究的不利影响。学术期刊在某种程度上发挥着研究成果的鉴定作用。什么样的研究是好的研究,由同行评议决定。学者会通过学术期刊最新发表的论文来学习、模仿他们认为正确的研究模式。所以,一旦研究模式出现了某种错误倾向,例如“数学滥用”,同行评议这种自我强化的机制很难自发地进行纠正。

### （二）对经济学教学的负面影响

2000 年 6 月,法国一批经济学专业的学生在互联网上发表请愿书,掀开了经济学改革国际运动的序幕(陆夏,2011)。请愿书认为,目前的经济学教学存在缺陷,新古典经济学占压倒性支配地位,过于迷恋数学形式主义且严重脱离现实,数学本身成为追求的目标。这场经济学改革的国际运动,说明了经济学中“数学滥用”、过度数学化现象普遍存在的一个原因在于,经济学教学中过分强调研究的一个原因,忽视了模型背后的经济学思想训练。

学者在经济学研究中的“数学滥用”会影响其经济学教学。特别是《宏观经济学》、《微观经济学》、《计量经济学》等课程,教师十分注重数学模型的推导,缺乏对模型背后大量经济现实的剖析和引导。教师上课忙于推导公式,学生片面地将经济学课程理解为掌握各种数学模型。毕业后去业界工作的学生,无法有效应用所学的经济学模型,认为经济学“无用”,实践与理论无法互补,两者越来越脱节;继续从事科学研究的学学生,又不断向新的学生重复着老师们的教学方法。经济学的“数学滥用”在经济学的教学过程中被不断延续。

中国的经济学教学尤其需要注意“数学滥用”的负面影响。经济学作为社会科学“皇冠上的明珠”,课程的核心应该在于传授经济学模型背后的原理,培养经济学直觉,训练经济学分析思路,让学生具备理解和解决现实经济问题的能力。中国学生的数学功底普遍较好,这使得学生在经济学的本土和国际学习过程中,会强化对于数理模型的偏好。一个经济思想的产生通常并不依赖于数学推导,而是通过细心观察经济现象、阅读大量文献并经过长期思考才能产生出来(洪永淼,2014)。学生的经济学直觉训练普遍不足,发现问题的能力不够。这些学生成长之后的研究也会剑走偏锋,发表论文的国际学术影响力不强。这可能也是余广源和范子英(2017)研究发现,中国经济学科的国际影响力较弱,全国仅两所大学(北京大学和清华大学)的“经济学·商学”学科进入 ESI 全球前 1%,离国家制订的世界“一流学科”的目标还相差较大的原因所在。

《计量经济学》是经济学实证研究方法的基础,也是重要的经济学课程。但是,借用和滥用计量经济学模型的现象也不断发生,甚至普遍存在(李子奈,2007)。计量经济学不是经济统计学,也不是数学在经济学研究中的应用,而是属于经济学的范畴。它是使用经济观测数据,采用适合经济数据特点的统计方法,估计、验证经济理论或经济模型,解释现实经济现象的一门学科,是连接经济学理论与现实的桥梁(洪永淼,2014)。因此《计量经济学》除了要学习“桥梁”本身的构造外,判断“桥梁”是否合适才是重要的。要判断“桥梁”是否合适,需重点关注“桥梁”的两端——理论与现实,明确“桥梁”只是起到联结(工具)的作用。但是不少《计量经济学》的课程直接上成了计算机软件的课程,只教了“桥梁”的构造,学生并不具备判断“桥梁”是否合适的能力。在教学阶段埋下了错

用、滥用数学的种子,可能是学生今后在经济学研究中滥用数学的一个重要原因。

### （三）对经济决策和社会生活的负面影响

2011 年 11 月 7 日,70 名哈佛大学学生退选格雷文里·曼昆(Gregory Mankiw)的经济学课程,称其课程是引爆金融危机的原因之一。曼昆是美国哈佛大学著名的经济学教授,29 岁便已成为哈佛大学的终身教授,其所著的《经济学原理》和《宏观经济学》是国际上最有影响力的经济学教科书之一。哈佛大学学生此举并非针对曼昆教授,而是借此反映学生对当今保守的经济学教育的不满。他们认为,经济学教育太教条,是有偏的。这种偏见会影响哈佛大学的学生,他们中的许多人已成为当今政策的决策者。缺乏系统性、批判性思维使他们的决策出现失误,或许是造成本次金融危机的原因之一。

这一事件说明,经济学教学和研究中存在的问题,会直接影响政策决策和经济社会生活,与每个人都息息相关。“数学滥用”现象表明,部分经济学者已经从坚持科学精神转向了坚持某种学术政治立场。如果事实果真如此,那么经济增长理论的停滞不前和两极分化,并不是科学方法出了错,而是学术政治导致的结果(Romer,2015)。罗默认为,正是“数学滥用”导致了经济增长理论中的严谨数学模型消失,取而代之的是不严谨、随意的“数学滥用”行为。“数学滥用”可能会导致经济学某些领域的研究长期停滞,缺乏创新性见解。当 2008 年全球金融危机全面爆发,主流宏观经济学无法对危机做出准确的判断并提供有效的解决方法。金融危机爆发后,一些经济学家开始反思,金融危机真正的原因在于人性的贪婪和金融监管的失灵,显然这是 DSGE 等模型无法刻画的(Caballero,2010)。

经济学研究最终需要服务于社会。好的经济学研究应该能够解答复杂的问题简单化,而不是反过来。晦涩的数学语言或者专业术语如果作为终极形式,只能起到相反的效果。

## 五、对经济学研究中“数学滥用”现象的反思

经济学研究中使用数学工具与模型,大多数学者并不反对,他们反对的是经济学研究中“数学滥用”的现象。罗默的批判引起诸多学者的讨论。Joshua Gans (2015)、Brad DeLong (2015)、Simon Wren-Lewis(2015)等对罗默的观点表示赞同,认为罗默并不是反对经济假设微观基础的简单化,更不是反对数学在经济学中的应用,而是反对“数学滥用”。他们认为,罗默通过期刊(AER)和会议(美国经济学会年会)的广泛影响来表达对经济学研究整体方向的批判,对学术发展是有益的。当然也有一些学者对罗默的批判表示不同看法,比如 Dani Rodrik (2015)、Noah Smith(2015),他们主要对罗默文章中认为经济学研究应该达成一致的论点提出了反对。Dani Rodrik 对罗默观点的回应是:经济学并不是自然科学,能够在所有情况下找出一个最优的模型。经济学研究没有必要达成某种共识,应保持经济学中理论模型的多样性,提倡研究方法的多元化。

罗默对经济学中的“数学滥用”现象的批判及在经济学界引发的这场大讨论值得所有经济学家反思。我们可以对“数学滥用”问题所引发的两个重大问题进行思辨,以图找出解决“数学滥用”的思路。首先,“数学滥用”混淆了经济学研究的目的与方法,这一问题即经济学研究的思想性与技术性之辨;其次,“数学滥用”不仅没有揭示出经济变量之间的关系,反而会使人们对经济运行规律误读,这一问题有关经济学研究的原因关系与相关关系之辨。

### （一）经济学研究:重思想还是重技术?

面对国内经济学术界的思想性和技术性之争,全国哲学社会科学规划办曾牵头,在国内开展了三轮较为广泛的讨论。经过充分讨论,经济学学术界、期刊界达成了一些共识。

第一次讨论的代表性观点刊登于《光明日报》2012 年 10 月 28 日第 5 版。中国社会科学院数量经济与技术经济研究所李金华发表的代表性文章《经济学论文:重思想还是重模型》(李金华,2012),剖析了经济学论文数学模型泛化的成因,并认为期刊界有责任把好用稿关、守好入门关。国内经济学论文中数学模型泛化可能有 3 个方面的原因:第一,“海归”学者带回国外学术研究范式的影响;第二,国内主流学术期刊模仿国外期刊的数理偏好;第三,国内高校经济学科的专业和课程设置中重数理而轻思想,导致研究者基本功不扎实的后果。经济学期刊应担负起“让科学、有效的数学模型见诸于世,应用于实际,服务于社会;不让错误的模型、伪模型见诸于刊,防止数学模型应用的泛滥和混乱”的作用。

第二次讨论开始于 2013 年 8 月 23 日,由《经济理论与经济管理》发起,国家社科基金资助的几家经济学期刊联合召开了“经济学论文的思想性与技术性的关系”专题学术研讨会,并最终由 5 家经济学期刊发布了“坚持‘思想性优先’的选稿原则”的联合声明:(1)经济学研究应当坚持“问题导向”,而不是“技术导向”;(2)坚持“思想性优先”的选稿用稿原则;

(3)正确处理好思想性与技术性的关系。会议还邀请了国内多位著名经济学家,与会专家一致认为,经济学研究应将思想性放在优先位置,数学模型应当服务于新思想、新观点的发现,而不能片面追求数学模型本身的复杂性和形式化。此次联合倡议可看作是经济学界对“数学滥用”问题的一次主动出击。

第三次讨论是由全国哲学社会科学规划办发起,国家社科基金资助的全部 26 家中国经济学、管理学期刊共同发起的对中国经济学研究选题、研究范式的导向讨论。《经济研究》《管理世界》《世界经济》《中国工业经济》《金融研究》《会计研究》《财经研究》等经济学、管理学期权威期刊明确表示,应发挥经济学期刊的引领作用,对中国的经济学研究进行选题引领、研究范式引领。经济学研究中使用数学模型本身没有问题,但目前出现了过分依赖实证分析、重模型轻思想的倾向,这是需要纠正的。经济学期刊选题应该更关注中国经济的现实问题和热点问题,注重理论创新,既要坚守学术为本、也要服务政府决策;经济学研究范式应坚持科学标准,避免对任何单一范式的迷信,方法必须服务于思想(樊丽明等,2015)。本次讨论可以看作是主流经济学期刊对经济学研究中“数学滥用”现象的集体抵制和纠偏。

2016 年 5 月 17 日,习近平总书记主持召开哲学社会科学工作座谈会并发表了重要讲话,在讲话中指出:“创新是哲学社会科学发展的永恒主题,也是社会发展、实践深化、历史前进对哲学社会科学的必然要求。”理论创新只能从问题开始,从某种意义上说,理论创新的过程就是发现问题、筛选问题、研究问题、解决问题的过程。”习近平总书记的讲话强调了哲学社会科学研究问题的重要性。习近平总书记在讲话中还指出“哲学社会科学研究范畴很广,不同学科有自己的知识体系和研究方法。对一切有益的知识体系和研究方法,我们都要研究借鉴……对现代社会科学积累的有益知识体系,运用的模型推演、数量分析等有效手段,我们也可以,而且应该好好用。需要注意的是,在采用这些知识和方法时不要忘了老祖宗,不要失去了科学判断力……如果用国外的方法得出与国外同样的结论,那也就没有独创性可言了。”习近平总书记的讲话为解决经济学研究中思想性与技术性的矛盾指明了方向。只有原创性的思想和理论才能构建经济学科的知识体系、学术体系和话语体系,模型推演、数量分析只是证明思想和理论的手段。

### （二）经济学研究:因果关系还是相关关系?

大数据应用于经济学研究领域,就产生了经济学研究的方法之争。“经济学研究的是相关关系还是因果关系?”大数据倡导者维克托·迈尔·舍恩伯格(2013)指出,大数据时代最为重要的转变就是从因果关系的研究转向相关关系的研究(谢志刚,2015)。Noah Smith 认为,随着当今信息和计算机技术,特别是大数据技术的发展,未来(基于海量数据的)实证研究方法必然会成为主流。这种潮流的一个典型例子,是机器学习(Machine Learning)在经济学中的应用。机器学习不需要依赖于经济学理论,通过对大量数据进行分析,识别数据本身的主要特征。换句话说,就是“让数据说话”。他们认为,未来经济学将从完全依赖于理论,逐渐转向数据驱动(Data-driven)的研究模式。

然而,这种观点是否正确呢?凯恩斯在《就业、利息和货币通论》中,认为基于概率上的相关关系推导出的经济结果值得商榷,因为仅凭借理论推演或个人经验来预测经济事件或得出结论,这些预测和结论事后几乎都被证明是错误的(凯恩斯,1997)。洪永淼(2007)指出,计量经济学更关注的是经济变量之间的因果关系,以揭示经济运行规律。数学模型的推演,并不一定能揭示经济变量之间的因果关系。

经济变量的因果关系,存在于经济学现象之中,由经济学理论所揭示,数学模型可能正确也可能错误验证出因果关系。综观国内外的经济学学术研究,对于因果关系的揭示已经越来越重视。不少经济学研究都包含内生性检验,包括巧妙地应用外生冲击、特殊样本以及克服内生性的各种计量方法应用,研究结果一般都会包含稳健性检验,以各种计量手段保证因果关系的成立。这是一个可喜的现象,说明经济学与数学并不对立,数学如果应用恰当,对于经济学研究因果关系的揭示是有帮助的。

更为有趣的是,Athey 和 Imbens (2015)的最新研究指出,最新的机器学习技术可以用来分析数据间的因果联系,将这一技术运用于经济学领域,可以分析经济学家们最为关注的一类问题——“政策对经济的影响(因果关系)”。该研究在美国国民经济研究局(NBER)会议上引起了学者的广泛兴趣。与传统机器学习主要研究数据预测,即主要考察数据间的因果关系不同,新的大数据技术开始注重经济学研究最根本的关切——因果关系。可见,在大数据时代,经济学研究注重变量之间因果关系的宗旨还是没有改变。技术只要服务于目的,两者就可以协调。因果关系的确立建立在经济理论的基础之上,因此未来的经济学研究需要学者们不断创新经济理论,同时也应发挥大数据技术在揭示经济变量因果关系上的作用。

(下转 04 版)