

大数据背景下社科实证研究的形式主义陷阱

周靖祥

摘要: 随着数据可获得性的增强,原本不可实现的社科实证研究也变得相对容易,这将改变实证、经验范式主义者依据方法而界定的研究可信性标准。源于统计分析工具的广泛运用,社会科学研究经验主义趋向早已显现。洞悉中国社会科学研究未来走向,梳理中国式实证研究文本意在谋变本土化发展之路。社会科学研究存在数据化倾向、经验主义、低水平重复和经验分析方法论万能认知等形式主义陷阱。为避免深陷经验主义困局而难以自拔,消除“经验实证唯一论”认识误区,坚持实事求是并“以人为本”,提升历史纵向与宏观叙事分析的重要性认知,借助经验实证刻画事实的意义远超于因果关系发现和机制、机理解释。为提升中国社会科学研究水平,需要重拾本土化学术自信,重塑价值理性。

关键词: 数据; 因果关系; 可信性; 可行性; 经验研究

居处大数据时代,社会科学研究发展可谓日新月异。伴随着逼近于总体的数据可得性日趋增强,倚赖概率发生假设的统计、经验分析方法遭遇到前所未有的挑战,新形势下社会科学研究工作者仍在为做出“精确的错误”和“模糊的正确”两难选择而努力着。在大数据时代之前,一个极具说服力的事例是:从15世纪到19世纪,欧洲人绘制的非洲大陆地图不断变化,欧洲人对非洲海岸沿线的探险经历使得这些沿海区域的地图绘制已经十分精确;与此同时,非洲内陆地区地域版图却从地图上消失,海岸沿线之外的非洲大陆成为一片空白。究其原因,随着地图绘制水平和信息质量的快速提高,关于“可靠信息”的标准也在不断提高;借助旅行者的口头描述和估测不再被视为可靠信息,因而被弃之不用。追求精确的结果是“非洲内陆由此无迹可循了”。19世纪末,随着欧洲人的探险活动遍及非洲各地,非洲内陆才重新出现在非洲大陆地图上。这个经典案例的意义在于:科学研究方法的进步及

其规范化、严谨化反而可能会限制我们的视野,导致我们的知识萎缩退步。^①社会科学经验、实证研究陷阱类似于此。研究者一旦过度热衷于追求数据化、公理化、精确化,乱象丛生的错误被形式的完美掩盖。随着数据的可得性增强,实证主义者极力推崇的经验研究偏爱展示通过数据展开统计分析抽取而得到现象和事实。经验研究结果的检验和解释还得依靠事实,倚赖数据的经验研究离不开问题导向思维和实地调查行动,解构现象终归需要思辨性的规范或理论研究。

一 社科实证研究的数据化陷阱

2012年3月29日,美国政府宣布推出“大数据的研究和发展计划”(Big Data Research and Development Initiative)。从收集的相关报告和资料获悉,该计划旨在提高从海量数字数据中提取知识和观点的能力,实现教育与学习的转变。^②据公开信息显示,主管部门指出“过去在信息技术研发方面

作者简介:周靖祥,湘潭大学中国农村与发展研究中心讲师,上海交通大学安泰经济与管理学院博士后。(湖南湘潭,411105)

①周雪光《方法·思想·社会科学研究》,《读书》2001年第7期。

②郎杨琴、孔丽华《美国发布“大数据的研究和发展计划”》,《科研信息化技术与应用》2012年第2期。

的联合投资推动了超级计算机和互联网的创建,而“大数据研究与开发计划”有望使我们利用大数据进行科学发现、环境和生物医学研究、教育以及保护国家安全的能力发生变革。”通过各方面的资料查阅,结合相关的研究分析得知,联邦机构对该计划的首轮资助包括以下机构:国防部(DOD)、NSF与美国国立卫生研究院(NIH)、能源部(DOE)和地质调查局(USGS)等。大数据广泛存在于各个核心领域,包括网络、通信、金融、物流、制造、政府。在中国,随着各类微观数据的可得性增强,盛行拿来主义的社会科学经验研究发展趋向正在变轨。由于中国式经验或实证研究遭遇了“可信性”的质疑,“搞对问题”和“识辨数据”才是解救研究发展受阻的良策。

来自《2010年年度信息流通普查报告》的数据显示,信息社会未被使用的信息比例高达99.4%。在大数据时代背景下,企业微观主体遭遇了如何应对数据资源利用率低的现实困惑,社会科学研究的可信性终将受到挑战。让我们以金融行业为例,将信用卡消费者行为资料转化为研究文本,则需要通过长期记录获取消费行为数据。无论是技术突破还是观察对象,现实社会经济运行已经积淀了这样的研究素材。统计数据显示,2011年信用卡新增发行量5500万张,交易笔数达到28.5亿笔,交易金额达7.56万亿元;同年中国信用卡消费额达4.1万亿元,占全社会消费品零售额的22.6%。来自央行的数据显示,中国信用卡累计发卡量已从2008年末的1.4亿张,增长到2011年末的2.85亿张,3年翻了一倍。2012年底,中国累计发行信用卡的数量已经高达3.3亿张,同比增长13.8%;累计激活卡量为1.86亿张,同比增长22.3%;活卡率为56.1%;全年信用卡交易金额达10万亿元,同比增长31.6%;信用卡未偿信贷余额新增3257.1亿元,累计达11386.7亿元,同比增长40.1%。由于金融部门掌控着这样的数据,他们最具微观消费者行为研究优势,此集成数据的商业性价值更是不可估量。此举例意在阐释理论建构和现实社会的鸿沟已被跨过,取材于大数据的深度研究或许会改变对中国正在发生或已经发生过的事件的既成认识和判断。

在大数据时代,社会科学研究面临重大机遇和严峻挑战。相比之下,中国产生的数据量极为庞大,但是真正存储下来的数据仅仅是北美的7%、日本的60%。值得注意,数据获取渠道增加和信

息量级数级增长特征:随着互联网、智能手机及传感器等的普及,信息流量有了爆发性的增长。根据GreenIT协议会预测,2025年,社会的信息量将猛增至2006年的200倍。考虑到全样本数据信息能够镜像出真实世界,社会科学研究的可信性终将获得来自“大数据”所能够描摹真实图景检验。具体而言,基于概率和可能性的实证和经验研究将会被大量的数据分析报告和专项预测研究取代,此时“历史的分析方法”和“科学预知未来”的重要性就会凸显。2005年后,中国式经验研究发文数量呈现级数级增长,可谓经验主义盛行。然而,在社会科学界,可获取数据的渠道多元化、实现方法论复杂化以及可研究的学术问题多样化,反倒使得众多研究者陷入迷茫,深陷经验主义的中国社会科学研究备受可信性和可实现性质疑。在当前数据可获得性增强情况下,社会科学界更应该思考的问题是,怎样更新范式、方法和开发工具利用大数据资源,以能够更好地展开深度研究。

当然,借助数据讲“事实”,原因是经验事实对于研究尤其重要,这又是实事求是讲道理的前提。然而,由于经验、实证研究容易模仿、借鉴的研究范本众多,当然这也为“思维偷懒”的研究者提供了便利,有思想性的研究文本俨然成了少数人的专利。进入21世纪,中国掀起经验研究的追捧热潮,量化研究也就备受学术刊物的“青睐”,研究者更是热衷于迎合。一种解释唾手可得,数据可得性增强,“拿来主义”思潮滋生。在这个躁动的、信息流“云化”的时代,奉行经验主义的研究者自以为有“知识”,实际上已经抛弃了常识,社会科学物化某种程度上能够用以刻画社会科学研究独立之精神,理论创新速度正在减慢。进入21世纪,国内期刊刊发的“庸俗”和“肤浅”经验研究日渐增多,高校、科研院所激励不相容的学术奖励如出一辙,各种形式的论文短平快生产行动可以视为“工具理性”扭曲的学术认知和评价更是催生了论文的流水线制造风潮。借此能够很好地解释为什么国内盛行经验主义研究,即便是缺乏基础知识储备和学术经验积累的研究者仍然能够玩弄“数据”快速实现发文的目的,无知又无畏戴帽子的研究者比比皆是。社会科学研究结果及解释的合理性检验在于研究者对经验事实的熟悉和了解,一旦研究者习惯了闭门造车、热衷于数据挖掘、公式化的文本写作,只会让研究之路越发狭窄,结果使得社科研究沦为“数据游戏”。

二 社科实证研究的经验化陷阱

在中国社会科学界(尤其是社会学与经济学)经验研究(Empirical Research)通常也被张贴或称其为“实证研究”,持此方法论者试图与纯理论(规范)研究划清界限,依靠搜集的“第一手资料”而展开。由此产生实证研究的两种取向:其一,力图透过抽象概念探讨建立一套概念体系(社会学);其二,透过抽象理论分析以解释某种现象,但未经实地验证(经济学)。当然,实证研究被视为一种透过客观的自然科学方法,通过观察、实验来检验的有效方法。在方法论层面上,为找到自然科学方法“实证化”如何嵌入社会科学研究,理当追溯到孔德的“实证主义”。20世纪以后,统计、概率分析方法不断改进,计量方法得到精湛应用,但是,在实现思路上仍然难以解决真正意义上的“证实”难题的困扰,更难以做到“证伪”。对理论假设或经验事实进行概率分析虽然能够揭示出经验事实提取学术问题,但很难展开较为完美的解释;实证研究的发展已经在不确定性、可能性的话语体系道路上止步不前或者徘徊。

在宏大历史进程中,研究者时常受制于有限思想认知,系统性难题及其交互影响往往会成为未知区域。总体而言,社会科学研究运用的数量分析方法在快速革新。此时,研究者更需要了解基本事实,经验研究主义发展趋向及实现方法选择本身就嵌入方法论内容范畴。在中国社会科学研究文本呈现中,研究展开和论题阐述的实现途径早已经陷入对实证、经验研究的“追捧”陷阱。在经济学、心理学、社会学甚至法学领域研究中,标贴实证研究早已成为一种时髦。从国内外各类社科期刊所发表的实证或经验研究论文数量来看,实证主义趋向越发明显,2000年后国内期刊更是趋之若鹜。在此情景之下,描摹社会科学研究需要静而观之和谨慎对待,展望未来、面对质疑才能够使得事实阐述“有的放矢”。

(一) 因偏离“人性”假设而迷失方向

社会科学研究对象离不开人及人的行为选择,经济人和社会人假设是基本前提。在中国经济学研究追随国际经济学界,日趋公理化、数理化的现实中,我们不禁要问,经济学研究究竟能做什么?

这并非实用主义使然,反倒是追本溯源。从亚当·斯密的《国富论》(被誉为“第一部系统的伟大的经济学著作”)说起,原因是这一部著作将经济学、政治理论、哲学、历史和实践计划奇妙地结合在一起,不同阅读目的的读者都会有意外收获。殊不知,《国富论》哲学基础是《道德情操论》(类似于凯恩斯的《通论》的哲学基础是《概率论》),在《道德情操论》(通常被认为是伦理学著作)的阐述中,相关论调早已触及到道德哲学,中文版的内容介绍中还提及“而后的《通论》更是用同情的基本原理来阐释正义、仁慈、克己等一切道德情操产生的根源,以此说明道德评价的性质、原则以及各种美德的特征,揭示出人类社会赖以维系、和谐发展的基础,以及人的行为应遵循的一般道德准则”^①。关于发展要义可以有不同的解读,可以从目的和过程两个维度进行阐释:就研究目的而言,体现人本主义关怀和人文情怀;可以将发展过程视为秩序建构(社会和经济)。经济学研究最重要的工作是场景给定,然后探讨人类行为选择的规律性。一旦植入了“人性”和“秩序”的发展或是增长研究解读,必定会对主流经济学公理化和形式化的科学性“权威”形成挑战,而所谓的现代主流经济学研究更像是“皇帝的新装”,裹挟在形式完美、结构严谨及“理性假设”范式危机之下,并未彰显人的社会性和自然性(原始动物性)特质才是行为选择的决定变量。

正如陆铭所言“当代中国,政治、经济、社会全面发展所面临的深刻挑战即在于,人们对于建设理想社会的目标缺乏共识,局部和短期的利益损害了全局和长期的利益。”^②将经济学纳入社会科学研究学科范畴,那么经济学研究对人性的关注和人性观的改变的作用不容忽视。当然,政治学、社会学领域的研究也毫不逊色,而且总是与制度规则和道德联系在一起。^③社会科学研究需要多一点人文关怀,少一点形式主义;多一点人性探讨,少一点物性套用。经济学研究尤其如此,人的复杂性决定了并非几个假设就可以概括,恰恰需要“以人为本”。事实上,凯恩斯之警告、克鲁格曼之观察都早已有过对社会各界发问“经济学家为什么错了?”的尖锐回应。

(二) 经验数据质量引发可信性深思

在中国社会科学研究领域,标贴实证研究早已

^①亚当·斯密《国富论》孙善春、李春长译,中国华侨出版社2010年。

^②陆铭《为了公共利益——一个经济学家的理想社会建设论纲》,《社会》2013年第3期。

^③陈涛《人造社会,亦或自然社会》,《社会学研究》2013年第3期。

成为一种时髦,心理学、管理学和经济学最为典型。在此情景之下,描摹社会科学研究需要我们静而观之和谨慎对待,展望未来、面对质疑才能够使得论述“有的放矢”。从国内外各类社科期刊所发表的实证或经验研究论文数量来看,实证主义趋向越发明显,2000年后国内期刊更是趋之若鹜。回过头来,我们自然要问实证研究方法是否有值得推敲和斟酌的地方?余国杰对实证研究的哲学逻辑进行过深度阐释^①:假定和假说是有严格区分的,验证则是搜集经验证据并与假设相对照。如果经验证据与假设一致,则假说或理论命题被证实;如果不一致,则被证伪。实证主义方法论的“实证”两字不仅包括“证实”,而且最主要和可实现的工作却是“证伪”。社会经济各种现象之间存在着极为复杂的相互作用机制,运用数学、计量工具可以将有关影响因素予以控制,从而解析复杂现象之间的内在逻辑联系或作用机理,并力求消除变量内生性,通过反复推演总能够让计算机运行出来的经验结果表现得更为“干净(Clean)”。然而,一旦使用的数据源存在“缺陷”,分析结果越是“精确”越应该受到质疑,必须承认没有终结的归因研究。在现实中数理统计意义上的相关关系未必存在,相关关系和因果联系的检验和判断已成为考释研究者识辨能力的重要内容,还可以将此类研究升格为应用统计分析技术的“规范性”讨论。实证或经验研究对数据质量的要求甚高,采集的数据源和中间加工过程的任一疏漏和错误都可能会导致结果或判断的致命性错误。

通常而言,“证实性”经验研究结论是或然性的。而当检验的直接结果是否定时,由于逻辑上存在着“否定后件推理”的原理,很容易找寻用以证伪理论的经验证据,而且在逻辑上则是必然的。基于文献的研究能够识别既有社会科学因果理论的命题结构及其类型,进而可以保持既有机制陈述不变,引入其他辅助条件,发展新的因果关联命题,形成基于因果关联命题意义上的新的理论观点与研究假设。此时经验数据质量起决定性作用。在研究论证过程中,使用不同结构和来源的数据得到的结论大相径庭,国内学界(尤其是经济学界),已经由质疑使用统计数据的研究转向质疑“数据”本

身。正当国内经济学发展在形式上的“国际接轨”提速之时,经验研究对数据的依赖最具有代表性。于是,有媒体指出“中国经济学研究的背景资料质量存在较大缺陷,……资料不全、资料掺水、资料有偏见等多种问题同时存在,许多方面的标准处于空白状态。”^②深究起来,追求逻辑自洽和形式完美的社会科学研究广受统计数据的可得性与可信性双重困扰,“缺失有质量的数据就必定缺乏有解释力的研究”俨然成了共识。^③谨记:其一,数据质量与经验研究可信性的关系论证;其二,如何展开能够取得最大“真实性”资料的社会科学研究,熟知历史、立足现实、贴近理论和提取出超前性认识。

时下,中国社会科学领域的经验、规范研究质量提升得益于数据可得性增强,数据源的多元化。即便是纯粹理论研究(包括方法论)形成的经验事实描述依然受到数据质量的约束,问题提出和背景阐述离不开直观的数据支持。理论建构依赖于现实,检验结果又必须回归现实。重返社会科学研究目标追求,赵刚早已给出警告“越是追求‘结果的完美’,越是偏离社会事实。”再就是“提倡小题大做。”但小题大做必须具备两个条件:第一,在细致研究上所投下的精力必须与研究结果的功能相称,过分细致的追索其实没有必要,甚至是一种人力之浪费,原因是“学术研究不是手工艺品,只见功力,没有功用”。第二,小题经过大做之后,仍需有人来综合处理,才能看出全貌。形象化的概述则是“把森林中许多叶子做了十分详尽的化验与分析,但对整个森林的全相却一无所知,并不是理想的林业研究工作者。”^④为提高经验结果的解释力,研究需要建立在坚实的理论基础之上,更需要有历史的、全局和系统性思维,以防缺乏纵向时间维度的比较而忽视了约束条件的重要性。

三 社科实证研究的“低水平重复”陷阱

社会科学经验研究主要依赖于概率,具体方法不外乎“归纳”和“演绎”两种研究实现路径。概率的通俗化表述皆为“可能性”,重提拉姆齐和凯恩斯的有关阐述对于我们对社会科学研究的“模糊性判断”做出客观描述尤其重要。由于个人的经验不同,导致研究者对客观概率的测度存在偏差。从逻辑

^①余国杰《析实证研究的哲学逻辑》,《光明日报》2005年8月16日。

^②郑良中《中国经济学家离诺贝尔奖有多远》,《长江商报》2013年10月15日。

^③郑良中《中国经济学家离诺贝尔奖有多远》,《长江商报》2013年10月15日。

^④赵冈《中国城市发展史论集》序言部分,北京:新星出版社,2006年。

辑学的角度来说,凯恩斯的演绎方法较之拉姆齐的归纳方法更加安全或者说更趋稳健。通过简单的比较会发现,其实都是在类似的角度下考虑概率,即从概率如何产生来考虑概率。社会科学研究的实证主义尝试和哲学讨论早已被卷入到研究可信性大讨论之中,证伪主义除了受到自然科学哲学的批评外,更不适合于基于概率性因果关系的社会科学。社会科学的统计检验自有其意义,并不需要冠以“证伪”的名号,证伪主义不是社会科学赖以立基的逻辑基础。^①在不可试验的研究过程中,因果关系的讨论总会深陷“鸡生蛋还是蛋生鸡”的逻辑困扰。假以时日,科学界若能够搞清楚“第一只鸡从哪里来?”,一切问题都将迎刃而解。研究问题决定方法,而不是方法套用问题。卡尔的认识极为深刻,其中一句话虽然通俗易懂但却道破全部真理:“历史就是人与时间的对话。”^②

在此,重述“历史总是在不断重演”的古训和判断,当代中国所处的大转型时代,各种应运而生的社会、经济问题,英国、法国和美国、日本都曾经遭遇过,在总结发展、研究此类现象时就已经在向人类发展规律总结逼近。因此,社会科学研究主题、内容和对象是动态的,而唯有规律是恒定的。早有这方面经典著述,《马克思恩格斯全集》及《双城记》描摹的英国和法国历史场景,和加拿大记者道格·桑德斯《落脚城市》给出的世界各地场景所展现出的实质就是人的活动和发展的规律。社会科学研究问题是全球性的、非时代性的,切入点却大不相同。在描述和解释社会经济现象(学术问题解释)时,不仅需要知道事实形成的过去,阐述现实表现(终将成为过去)以及未来发展演变的可能状态,某种意义上讲不同维度、不同层面的全面深度解析是构成社会科学研究规范的重要内容体系。福山早已论及道:“将理论放在历史之后,我认为是正确的分析方法。……社会科学往往以高雅理论出发,再搜寻可确认该理论的实例,我希望这不是我的态度。”^③在此,给出严密因果联系的精确判断和演绎缜密逻辑的实例。2012年11月29日,凤凰网有一篇“‘千城一面’——城市规划的滥觞”的文章,指出“城市规划在先,城市建设按照规划进行”。历史就像一面镜子,对于社会科学研究而言,重现历史就是在寻找实证证据。问题“起源”是研

究讲述的重要背景,历史是实证主义的“灵魂”。

进入21世纪,中国经济学界对于实证和经验的“追捧”超前了其他学科,单从形式上讲可谓跑步实现了“国际接轨”。有英文作为辅助,“走出去”发表也更多。中国发展是否适合用拿来的知识加以概括至今无人能解答,甚至于马克思主义政治经济学也被赶时髦地贴上了“实证”的标签。事实上,国内报刊媒体刊发社会科学研究“短论”的思想性、简明性让我们更多地体味到研究成果所能起到的“意识”导向和对现实的深度解剖作用。重新回到“可信性”问题的讨论,中国式经验或实证研究遭遇的困境不仅仅局限于可信性,更深层次的问题则是“如何搞对问题”,又如何识辨方法“是否适宜”?在统计分析技术在服务于数据分析的过程中,数据质量就比什么都重要。计量经济学作为应用经济学的分支方法论学科,基础是“数据”。数据的类型通常有三类T-C-P(依序为时序数据、截面数据和面板数据),这也是数据二次采用和研究的基础。诚然,研究过程中内生性的讨论固然重要,但需要建立在“数据”可靠性基础之上。经济学研究中,问题意识和历史感极为重要。使用数据展开经验研究是研究者的拿手工作,依赖于数据开展经验研究时会出现“差之毫厘,谬以千里”,由于社会科学描摹的各国社会现实。当然,由于社会科学研究结论偏误很难做出“检验”,需要社会科学研究者自律、自省,储备历史的知识,多开展一些调查研究工作。

四 社科实证研究的方法论“万能”陷阱

在历史长河中,融合、统一社会科学和自然科学的尝试与努力总需要人们在“科学的目的与方法究竟是什么”这一问题上先达成某种程度的共识。尽管“解释”不是恰当的科学语言,在生物学、物理学和化学中也并非特别重要,但解释、阐释、文本却是社会科学最为普遍的基础,在社会科学领域种种充满主观色彩的人性化解释不可避免。融合自然科学与社会科学研究的研究方法,社会科学研究各领域交叉已经在改变着我们的研究去向。原因是,作为研究者无法也无力去回避社会科学研究融合,被我们所熟知的社会科学家本难以做出严格的“学

①张杨《再论证伪主义的问题:对尹海洁、庞文商榷的答复》,《社会》2009年第4期。

②[英]卡尔《历史是什么》陈恒译,北京:商务印书馆,2008年。

③毛俊杰《政治秩序的起源》桂林:广西师范大学出版社,2012年。

科领域”归属判断。问题选择虽有区别,实现方法不尽相同,实证尤其是证伪才是科学性建立的根基。当然,实证研究并非万能,理论基础或理论建构决定了社会科学研究的“水平”。作为一门显学和最接近科学的经济学研究却总是遭到批判,某种程度上讲经济学研究实现方法论走向科学的同时理论发展几近停滞,两者的此消彼长对于这一领域的研究者而言是“短板”而非“长处”。现代经济学者滥用经济学理论去解释现实并不构成经济学理论本身被“证伪”的理由,这就好比一个蹩脚的建筑师误用了平面几何定理造成了设计失误,但这并不构成平面几何定理被“证伪”的证据一样。^①自斯密以来,经济学研究“大三角”(由问题意识、理论建构和实现方法论构筑)关系互动已经成为学科发展以及决定未来方向的内核,数以万计的经济学家还努力在新古典经济学的形式模型体系内工作,而大量的研究文本选择“理性选择理论”、“公共选择理论”作为基础(社会科学研究理论建构离不开“人性”的讨论)。为解答这个问题,在一定程度上取决于站在何种立场上看待自然科学和社会科学在方法论上的整合。在社会科学研究之文本呈现中,借助数据可以将问题具体化,甚至增强了“可视化”程度,使得研究范式从思辨性论战的批评中解放出来。

社会科学研究总是在努力使得结果和解释逼近于真实世界或是揭露运行规律,也因此结论的大相径庭和说理对抗(或者研究向前迈进一小步的过程)总是在不断上演。为对此作出说明,我们以中国基尼系数的“大讨论”故事为例,原因是此话题不仅是经济学问题,更是社会和政治问题,与现实联系得最为紧密,社会各界对此都给以高度关注。2013年1月18日,国家统计局公布2003年至2012年中国居民收入的基尼系数。据此可知,全国的基尼系数在2003年为0.479,自那之后逐年上升,到2008年曾达到历史最高点的0.491,其后的几年有所回落,到2012年降为0.474。在此数据公布之前,西南财经大学中国家庭金融调查中心根据其住户调查计算并且“抢先”公布了2010年全国居民收入差距的基尼系数数值为0.61。正如有论者指出“有可能因为西南财大的基尼系数‘先入为主’的缘故”,国家统计局公布的基尼系数受到了一定的质疑。”查阅资料发现,批评西财基尼系数

数值的一方指出:正是由于样本偏差和收入指标统计上的问题,让该调查对很多家庭给出了不现实的低收入,居民收入差距(0.61的基尼系数)被高估的原因之一也在于此。^②从社会公众的心理反应来看,基尼系数不仅是一个数,被倾注了更多理性的,非理性的主观情绪,也有“历史的记忆”,此次基尼系数的公布还引发了统计制度的讨论。从研究过程来看,基尼系数只是测算数,但取得该数背后的过程是极其复杂的,对估计结果加以详细说明是完全必要的。

超越数值之争,作者根据公开的CGSS2008年的数据测算“农村户口”个人和家庭收入的基尼系数,不同分组都在0.5左右,农村户口家庭基尼系数值是0.491,与国家统计局公布的中国全国居民收入的基尼系数2008年0.491在数值上一致,绝非巧合。依据户口类型的测算将“不知道”和“拒绝回答”两类少量样本用已知的样本数值的“平均值”来替代,作此处理主要理由是不改变样本的分布情况,但补足缺失数值以保留“有效样本”。作此处理也是为了说明,其实未剔除“最高收入组”和“最低收入组”,基尼系数测算依据样本总会收敛于0.3~0.6,原因是大样本情况下的“正态分布”假设总能够得以满足。作此处理排除居住地、职业的城乡差异,控制个人和家庭特征变量后,对个人收入不平等做分解结果发现受教育年限和家庭人数变量(和被调查者居住的人口数)两项贡献最大,数值分别为34.8%和12.2%,借此可以解释差距的50%左右。接着,由于2013年国家统计局公布了2003年以来的基尼系数,通过描述其与经济增长率的关系发现,两者存在着高度的相关性(考虑滞后影响),自2003年以来两者都呈现出“倒V型变化”,也即经济增速经历了最高点、不平等程度也达到最高点,但不能以此短期波动变化来判断不平等已经得以改善。如果延长时限此变化趋势仍然成立,那么由此数据的简单拟合我们似乎可以推知在一定发展阶段,高增长必然导致不平等的因果判断,经济增长是不平等的根源,前提条件是:起点差距甚小,但由于经济的快速发展结果使得不平等加剧。

对于当下中国的社会科学研究而言,急需营造坚守“顶天立地”之学术追求大环境:何为“顶天”,

^①张杨《再论证伪主义的问题:对尹海洁、庞文商榷的答复》,《社会》2009年第4期。

^②相关资料参见廉薇、周慧兰《基尼系数助推收入分配改革破题》,《21世纪经济报道》2013年1月28日。

也即专注社会科学领域的尖端、前瞻性理论问题;“立地”就是立足于中国社会经济发展大转型现实,捕捉中国问题、建构中国理论和寻求中国解释。问题是研究的起点,也是学科发展的生长点。^①结合学术与研究的发展认识可以做出判断,问题导向抑或知识(文献)导向决定着中国未来社会科学研究发展的命运,归属于工具理性范畴的实现方法论(尤其是数理化、公理化)很可能加速研究者远离学术(学术生命短暂),大多数情况下的研究工作只是服务于“唯成果发表”的研究认知。目前看来,数据的可得性增强对于社会科学研究而言是一把“双刃剑”:一方面,为社会科学研究注入了活力;另一方面,正在挑起社会科学研究的新的“可信性”革命(这一次或将超越实现方法论的可信性),也不排除中国社会科学研究会再来一次“跟风”。纵观而论,中国研究方法革新(奉行拿来主义)与学术“大跃进”(盛行发展主义)齐头并进,以至于最终脱离方法论研究的初衷,其对于研究所能起到的作用只是“规范”。研究的起点是“问题”,人文社会科学研究势必需要建立起问题意识、学理意识和方法意识,方法论特征则可以用“法无定法,道有常道”来概括。^②在大数据时代,中国社会科学研究的传统范式或将被改变,中国实践和社会运转甚至会颠覆传统经验、实证方法。倡导和展开社会科学研究的可复制性和可检验性实际行动也可视为研究的可信性革命,然而公布、公开社会科学定量分析原材料亟待建立一个透明和开源的学术机制,机构或是个人在积极倡导“可复制性研究”(Replication Study)。在具体行动上,经济学研究开创了先河,这就是《美国经济评论》(AER)自1999年改版后开始论文上网,2003年以后经验和实证研究论文的数据及方法公布在期刊主页上,可供“复制”和检验。并于2011年第1期再现100年来的TOP20论文,从中能够捕捉到期刊的成长和经济学研究发展经验化之路。经济学、社会学、政治学等社会科学研究领域已经有大量的学者积极倡导,并付诸于具体研究行动,在可知的范围内,《美国经济评论》公布数据、程序开创了典范。

五 结语

进入21世纪,中国社会科学研究步入了飞速发展快轨,尤其是2003年各大主流期刊改版后拉开了经验主义盛行序幕。伴随着可获取数据渠道的多元化、实现方法论的复杂化以及可研究的学术问题的多样化,发展也因此而陷入迷茫。就中国社会科学研究现状而言,作者和期刊的偏好选择促成了经验主义趋向,深陷可信性陷阱。

在大转折时代,大数据生成很可能会诱发社会科学研究嬗变;更进一步,展开问题选择、方法论认知和可信性三维度深度论述都会是宏大的研究工程,本文的论述意在指出如何提防形式主义陷阱。重塑中国社会科学研究新格局,研究者的洁身自好是准入的前提,学术期刊的平台管束和引导可谓最后的“把关”,当然更需要社会各界重新认识社会科学研究的真实面。以思辨性思维探讨和反思中国社会科学研究之发展,以期建构起数据与经验研究之间的辩证联系,需要反思的恰恰是社会科学研究者自身。

知晓常识,尊重知识,重温杰文斯在《科学不涉及终极关怀》一文中的经典论述“任何科学研究中,逻辑法则和方法的有效性,即我们在这个世界上确定方向的一般基础,都是有前提的。……社会科学各学科知识教给我们如何从源头上理解政治、艺术、文学和社会现象。”^③社会科学研究是一种知识生产活动,常言道“一切皆有可能。”在信息时代,过去不能实现的研究也将会变成现实。从理论上讲,实现数据的集成已经是一项极其容易的工作,包括政府机构和统计部门及大型企业或网络平台数据、监控视频等已经构成了很丰富的大数据资源库。社会科学研究本该是创新知识和服务社会的实践活动,有良知的社会科学研究工作者不应当被昙花一现的形式化浪潮吞没,勇于担当引领正道,不拘一格再建中国本土化学术体系,重获学术自信。

(责任编辑:向志柱)

①王永斌《人文社会科学研究中的问题意识》,《光明日报》2005年4月25日。

②劳凯声《人文社会科学研究的问题意识、学理意识和方法意识》,《北京师范大学学报(社会科学版)》2009年第1期。

③斯坦利·杰文斯《政治经济学理论》,郭大力译,北京:商务印书馆,1984年,第200页。