

# 中国国际关系定量研究十年发展与进路<sup>\*</sup>

庞 珣 漆海霞<sup>①</sup>

**【内容提要】** 中国国际关系定量研究经过近十年的持续发展已经形成了具有集体特征、共享规范和对学科产生整体影响的研究类型,并在学界持续的关注、讨论和争论中被界定为非严格意义上的“研究领域”。对于任何一个研究领域而言,定期进行回顾总结和进路探讨是一项必要而重要的学术工作,有助于凝聚共同努力、形成和强化学术共识、及时发现和纠正偏误、保证可持续创新。清华大学国际关系数据与计算实验室和北京大学全球风险政治分析实验室组建联合研究团队,搜集并分析近十年在中文学术期刊上发表的国际关系定量研究论文,并结合定性观察和扩展性思考,形成本文,供学界同仁进一步讨论。本项研究旨在发现问题、寻找改进方案、避免误区、发挥潜力,推动我国国际关系定量研究自身的健康发展,促进其对我国国际关系学科建设作出更加积极的贡献。

**【关键词】** 学科发展 学术创新 定量研究 数据库 国际关系研究 国际学术传播

<sup>\*</sup> 本文由清华大学国际关系数据与计算实验室和北京大学全球风险政治分析实验室联合组建的研究团队进行资料搜集、分析和撰写,尤其感谢研究生王侯嘉遇、王淇、杜普、黄昱茗、邵志成等出色的研究助理工作和团队组织协调工作。本文系国家自然科学基金重大项目“地缘政治风险预测的理论与方法研究”(项目批准号:17ZDA110)的阶段性成果。需要特别说明的是,本文无论是对成果的总结还是对问题的分析,都是将中国国际关系定量研究作为一种研究类型,而并非针对任何具体的文章或作者,因此在文中着意避免了从团队搜集的文章中进行具体举例和引用,也尽量不提及作者姓名。需要参考本文的分析样本详细数据,可至北京大学开放研究数据平台(<https://doi.org/10.18170/DVN/JCZDZM>)进行下载。

<sup>①</sup> 作者排名不分先后,按姓氏汉语拼音排序。

《国际政治科学》2024年第9卷第1期(总第33期),第1—40页。

*Quarterly Journal of International Politics*

**【作者简介】** 庞珣,北京大学国际关系学院教授。

电子邮箱: xpang@pku.edu.cn

漆海霞,清华大学国际关系研究院副教授。

电子邮箱: qihaixia@tsinghua.edu.cn

## 一、概述

### (一) 国际关系定量研究: 学科发展中的“过渡性”领域

对特定研究领域的发展、问题和进路进行阶段性的回顾、总结和探讨,是一项常见而重要的学术工作,对于该学术领域凝聚共同努力、形成和强化学术共识、保证可持续创新具有不可或缺的作用。从本质上看,一个学术领域是由围绕某个共同研究主题产生、形成和不断发展的一套连贯的知识体系所定义的。此外,要成为学术领域还需具备数个有形要素,包括专业训练和认证专业人士的公认标准、大学和研究所中设置的专门职位、进行领域自律和促进教研交流的学术团体,以及专门期刊等学术成果发布的集中渠道等。<sup>①</sup>

按照这些标准,无论在中国还是国外,国际关系定量研究都并非一个学术领域。国际关系定量研究不同于定量研究方法。定量研究方法的研究主题是系统性数据分析的科学原理和技术体系,在国外已是一个成熟的学术领域。但国际关系定量研究并非围绕一个共享的研究主题,其被视为一种研究“类型”,归类标准是研究资料(量化数据)和研究方法(系统性批量处理数据)。国际关系定量研究旨在向各自选定的国际关系研究主题和问题作出贡献,产生的知识也归于各自的主体所形成的知识体系。例如,三个研究的主题分别为军事结盟、国际贸易政策的国内分配政治、国际组织中的规范形成和传播,它们只要运用了量化数据和量化方法则都被称为国际关系定量研究,但产生的知识分别归属于国际安全、国际政治经济学、全球治理等不同领域中的知识体系,并主要由这些领域中的同行来评估研究的价值。

---

<sup>①</sup> Benjamin J. Cohen, “The Transatlantic Divide: Why are American and British IPE so Different?,” *Review of International Political Economy* 14, 2 May, 2007: 197-219.

换言之,国际关系定量研究的方向和成果并不在实质(substantive)或方法(methodological)上推进一套连贯知识体系的发展,也就无法构成一个研究领域。

尽管如此,在学科发展特定阶段中,国际关系定量研究展现出践行共享的学术理念、研究方式、行业标准的群体性努力,倡导以系统性数据分析为主要途径和信度保证来回答国际关系实质性问题(substantive questions),在对实证证据进行观察、提取、选择、甄别和解读中采纳通用的科学标准。可以认为,这些群体性的投入界定了一个领域,这就是国际关系定量研究。不过,国际关系定量研究即使视为一个领域,也具有过渡的性质。在定量研究初入一个学科时,学科整体上缺乏对这类方法的原理理解和应用经验,在规范性上的自律和他律也较薄弱。一部分学者开始探索和推广定量方法在国际关系研究中的应用,并确立和内化研究规范性和评判标准,形成知识构建和研究推动的集体性努力。在进一步发展中,定量方法的应用变得越来越普及且常规,这种努力不再是某个群体的行为,国际关系定量研究也就完成了其过渡使命,而后作为一个“领域”自然消散。这个发展规律可以在国外的经验中观察到。例如,亚洲和澳洲地区政治学国际关系研究中的定量研究方法应用时间较短、普及程度较低,在这些地区就活跃着以“定量政治学”等命名的学会团体和学术会议。但在定量方法的应用已变得更为普遍的北美地区,则没有了此类的专门会议或组织。<sup>①</sup>

中国国际关系学界大约从20世纪末开始出现使用定量方法来进行国际关系研究的成果,并有学者和机构持续地致力于定量研究理念和技术的学习和推广工作。但定量研究在很长时间内只是个别现象,直到最近十年才出现较为明显而稳定的发展势头,形成“领域”的雏形。这不仅表现为在参与的科研人员数量、成果所占期刊版面比例、研究体现的学术共识和研究规范、相关教学培训等重要方面的持续进步,还体现在定量研究的发展在中国

---

<sup>①</sup> 美国政治学方法学会不是“定量政治学”(应用定量方法来研究政治学问题),而是“政治方法学”(政治学中的数据问题和数据特征促进方法的创新)。为了促进新方法在政治学研究中得到推广,最近一些年学会的年会会接收一些应用型研究作为研究墙报,也会评选出应用研究的研究墙报奖,但将“应用”(applied)专门区分出来就是对两者之间区别的强调。

国际关系学界成了一个学科级的现象,受到与其当前阶段和规模不相匹配的高度关注和各种争议。这些都表明定量研究作为“过渡性”领域已经形成——它被人为地画出了边界和定义了领地,同时对国际关系学科在理念、规范和实践等方面产生整体的影响。这也凸显了本研究对国际关系定量研究的发展进行回顾、审视、分析和探讨的必要性和重要性。

## (二) 本研究的定位、范围和特点

为承担这一工作,清华大学国际关系数据与计算实验室和北京大学全球风险政治分析实验室组建联合研究团队,搜集和分析近十年在中文学术期刊上发表的国际关系定量研究论文,并结合研究团队的定性观察和扩展性思考,形成了本研究论文。

研究团队的分析资料主要建立在2013—2022年正式发表在中文学术期刊上的使用定量研究方法对国际关系问题进行研究的学术文章。我国国际关系定量研究成果的主要呈现形式为期刊论文,此外也有部分散见于著作章节、智库文章、咨政报告等。此外,相比其他学术成果形式,经同行评议的期刊论文能够更好反映出领域的研究规范、学术特征和发展动态,也能够较为集中地显现出具有普遍性的倾向和问题。定量研究文章在形式上呈现出较为统一的标准,有助于在各重要维度上进行系统观察和分析。因此,期刊论文虽未完全覆盖定量研究领域,但以其为样本的分析结果能够大致而无偏地反映出领域特征和趋势。

本研究暂不包含中国学者(无论是在海外还是在国内工作和学习的学者)发表在国际期刊上的论文。虽然这些国际发表文章有一小部分通过社交媒体被介绍回中国学术界,但鉴于当前学界实际上用两套不同的标准审视中英文发表,以学者国籍而非期刊面向为选择标准可能模糊对我国国际关系定量研究的特点和问题所进行的观察和分析。

此外,“论”定量研究方法的文章也不在本研究分析中。关于定量研究的介绍、澄清、争论、历史梳理等工作可能对研究的发展产生重要意义,但其本身不属于定量研究。尤其值得强调的是,本研究不关注定量方法与定性方法在普遍意义上孰优孰劣的辨析,而是关心研究方法在具体研究中使用的恰当性和可改进性。需要特别说明的是,博弈论研究不属

于定量研究,但本研究对博弈论研究给予了额外的关注,是出于对博弈模型实证化方向和前景的关注。本研究不对国际关系定量研究的“学科建设”发展和现状(包含课程设置、教职分配、学术团体、专门发表渠道、代表性人物和流派等等)进行回顾,只在关于如何提高研究质量的探讨中会涉及这些内容。

在此基础上形成的研究报告,并非意在强化定量研究作为一个领域的存在,而是通过分析这一类型研究的发展轨迹和现状特征,以期促进定量方法在中国国际关系研究中得到更规范、更恰当、更有益的应用。本研究的主要目的不在于展示取得的成绩,而是将重点落脚在对现有问题的发现和对可能解决思路的探讨上,旨在促进我国国际关系定量研究的未来发展和对学科整体的价值发挥,并为学界就这一研究领域中的重要问题进行更为广泛、深入和持续的讨论提供基础。因此,本研究并不追求分析的全面性或权威性,而是希望既突破其他探讨形式的局限,也作为它们的补充。

在对定量研究进行回顾、反思和展望方面,我国学界已经做了很多积极的工作,多数以研讨会的方式进行,主要形式体现为学者们面对面口头交流其关切和意见。这种形式的探讨无疑有助于学术反观和共识形成,但也具有碎片化、过度依赖主观感知、难以累进等缺陷。更重要的是,研讨会的形式在参与讨论者和意见表达方面都可能存在着严重的选择偏差。“邀请—接受”的选择方式和面对面的交流氛围容易导致相似理解得到过度表达、观点相互印证,而实质性的异见往往难以得到足够重视和讨论。此外,研讨会上的交流往往难以深入,会议讨论议题可能涉及很多内容和问题,但由于时间限制和缺乏分析的共同资料基础,而往往流于吐槽或偏离以研究方法为主线的讨论。本研究并非要替代且无法替代这些更为生动、开放和动态的形式,而是为它们提供讨论的基础和聚焦。研究团队未来的跟踪报告也将更广泛地结合和吸收学界的观察、信息和思考,也希望更多的学者参与到后续研究过程中。

### (三) 本研究的主要发现

本研究基于近十年来中文学术期刊上发表的国际关系定量研究论文进行了以下三个方面的分析。第一是对这一领域研究成果在学术维度上的特

点及其变化趋势进行描述和提炼,包括定量研究发表成果的数量规模和年度变化趋势、研究议题在国际关系学次领域中的分布情况和比例、研究使用的数据情况、研究方法的选择和应用规范等。第二是对国际关系定量研究学者的职业发展阶段、组织形式、所属机构和地理分布进行报告。第三是对中国国际关系定量研究成果发布渠道(学术期刊)进行分析,观察国际关系学科对定量研究的接受程度以及定量研究对学科的影响范围。

从总体上看,近十年是中国国际关系定量研究开始作为一个具有高度辨识度和集体特征的领域得以持续和稳定发展的时期。这一时期的定量研究发表数量虽然年度波动较大,但整体呈明显的增长趋势,已经形成了一定的整体规模。定量研究方法在国际关系的主要议题领域都得到了有益的尝试,在国际政治经济学、安全研究和外交研究三大领域中产生了不少成果且分布较为均衡,并非只在某一个特定议题领域中生长。中国国关定量研究对现有成熟可靠的国际关系数据库的利用日益充分,对国际上数据库的创新和更新有较高的灵敏度,应用的数据呈现出形态的丰富性、来源的多样性和跟踪的及时性。但原创性数据较少是研究的一个明显短板。在方法应用方面,回归模型是最主要的方法类型,其中处理面板数据的模型应用最为常见。这与国际关系研究中最常见的数据是用于同时跨区间和跨时间进行比较的面板数据直接相关。具体模型设定根据面板数据的具体形态和问题而有较多的变换。不少研究对内生性问题和稳健性检验有着足够的重视,反映出我国国关定量研究学者对回归模型方法掌握的成熟度。与此同时,其他类型的方法也得到了探索性的运用。但从整体上看,因果推论和真正的大数据方法尚应用不足,尽管它们是当前方法研究领域中最重要的发展方向。

从研究人员的分布来看,国际关系定量研究论文的作者分布在学术职业发展的各个阶段。除硕士研究生外,作者的职业阶段分布较为平衡,并没有呈现出学者只在某个阶段从事这一类型研究的现象。但一个需要引起重视的现象是,近十年来青年学者对定量研究更为接受和积极,而罕有以资深学者身份进入这个领域的。统计中以教授职称发表定量研究文章的学者大都是在青年学者阶段就从事定量研究。参与者年轻化对于国关定量研究的发展而言是一个积极的信号,预示着未来可能会有更多的国关学者将定量

研究方法作为其日常研究工具。同时,这也显示出了中国国际关系定量研究在学术传承方面的缺乏,缺少多年累积性的学术直觉和理论思考的学术积淀注入定量研究中,无助于防止定量研究出现理论与实证脱节、研究与现实脱节的问题。从期刊平台上看,我国国际关系定量研究成果的发布高度集中于少数几个期刊,尤其是《世界经济与政治》期刊。尽管这些期刊具有较高的学科影响力,但高度集中现象表明定量研究在中国国际关系学界中的接受度仍然较为有限。此外,发表渠道缺乏多样性,实际上意味着我国定量研究发展中的脆弱性。

在这些分析的基础上,本研究希望学界对国际关系定量研究的问题、潜力和前景进行更为深入的分析 and 广泛的讨论。为此,本研究尝试寻找以上发现的原因和可能的改进之道,其中贯穿了撰写者作为亲历者的体验观察、个体思考和自我反省,并在最后结语部分提出了六点建议供学界进一步讨论。

## 二、本研究分析基础:中文期刊国际关系定量论文

### (一) 定量研究论文的选择范围

本研究界定国际关系定量研究论文为运用系统性数据分析回答国际关系领域研究问题的论文。根据这一界定,在选择近十年来的定量研究论文时,研究团队依据以下两个主要标准。第一是文章进行了系统性数据分析,即以计算机为工具对量化数据进行趋势和关系处理,并对报告结果可靠性加以实证检验。这意味着,仅引用第三方统计数据或现有研究中的图表、仅运用数学公式进行观点表达或理论推演的文章都不属于本研究所关注的定量研究。相反,进行了系统性数据分析的文章无论是否明确使用了数学模型表达都属于定量研究。第二个选择标准是文章回答国际关系问题。纯方法论文章虽然在测试和应用举例部分也进行系统性数据处理,但这些工作的目的不是回答国际关系问题,而是检测方法的可靠性和进步性,以及展示方法的有用性。这些研究不在本研究的分析范围中。

在考察时段上,研究团队选择 2013—2022 年十年中的期刊论文。2013

年之前已有中文期刊发表国际关系定量研究论文,但超出了本研究的分析时间范围,因此不纳入搜集分析之列。在对期刊的搜索范围方面,出于本阶段时间和人力资源考虑,研究团队本次采取定点选择的方式,集中在《中国社会科学》《世界经济与政治》《外交评论》《当代亚太》《国际政治科学》和《国际安全研究》等6种期刊。后5种期刊是近十年来国际关系定量研究论文的主要发表渠道,集中了绝大部分的此类文章。而《中国社会科学》作为社科综合类期刊,发表的国际关系研究数量和比例都较小,但鉴于该期刊具有跨学科的广泛关注度和影响力,研究团队也对其上发表的论文进行了采集。除了这6种期刊外的其他期刊也发表了少量国际关系定量研究,虽未纳入分析,但其数量和比例过小,不影响趋势性分析。

## (二) 定量研究的整体概况

### 1. 数量规模趋势

根据上述标准和搜集范围,研究团队共获得153篇定量研究文章、24篇采用博弈论进行研究的文章。这些文章既非随机抽样,也不构成全样本,这意味着,本研究是在启发意义上而非推论意义上对它们进行分析。

图1报告了2013—2022年6种期刊上发表的国际关系定量研究论文的

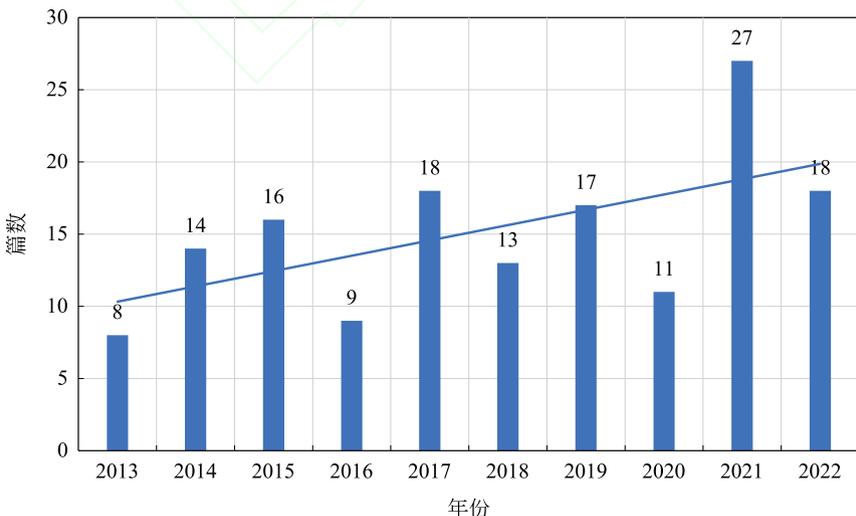


图1 2013—2022年国际关系定量研究论文发表数量趋势

年度数量和变化趋势。由图 1 可见,定量研究发表论文数量整体呈增长趋势,但年度波动较大。其增长趋势主要由 2013—2015 年 3 年的持续增长和 2021 年的大幅度跳涨形成。排除 2021 年后,2016—2022 年六年间实际上在一个平均水平线(平均 14 篇)上下波动,稳中略有上升。这和学界对定量研究发展的直观感受有明显差距,并没有如研讨会或日常所言的那种持续快速增长。这可能与定量研究在这 6 种期刊之外的刊物上呈扩散性发表有关,也有可能反映了对定量研究的兴趣、尝试和发表三者之间的实质性差距。

为了进一步观察对一种研究类型“冷”与“热”的直观感受或日常谈论与实际的研究成果之间的差距,研究团队对这 6 种期刊在同一时期发表的“冷门”研究博弈论文章的数量和趋势进行分析。图 2 显示,博弈论文章的数量确实很少,远远低于定量研究的数量,这与直观感受相一致。但与普遍认识不同的是,博弈论文章作为小众类型显示出了其生命力,在近年来其实有较大幅度的增长,并没有呈现出消逝的趋势。

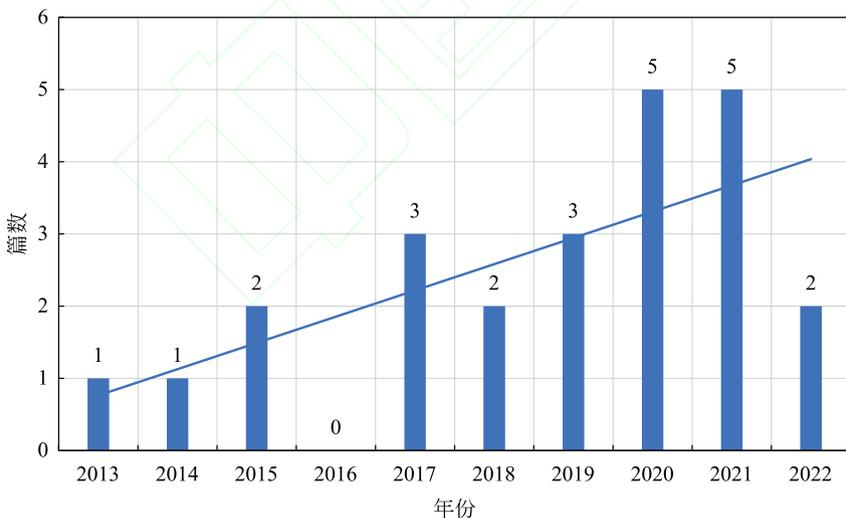


图 2 2013—2022 年国际关系博弈论研究论文发表数量趋势

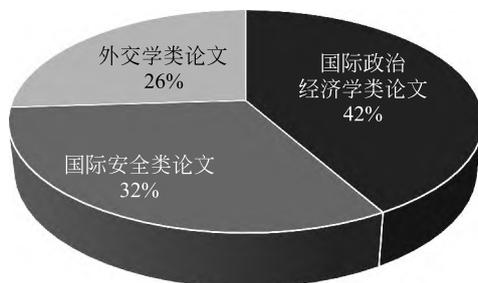
对定量研究和博弈论文章的变化趋势进行比较,揭示一个有趣的发现,那就是学术研究有其自身的规律,并非跟随通俗的流行和潮流而发生实时的变化。2013 年以来大数据席卷各个领域,包括学术界。国际关系领域中,“大数据”也成为研讨会和社交媒体文章中的热词,给人造成一个普遍的印

象,那就是国际关系定量研究成为热点。姑且不论大数据和定量研究之间能不能够画等号,对某一种新方向进行坐议立谈,或者在这个新方向上进行切实努力并取得研究成果,两者呈现的是迥然不同的态度和投入时间。议论的热情和学术的审慎,表面的拥抱和科研的积累,具有相当的差异。图1的趋势正是体现了国际关系定量研究并非乘潮而动,而是有其自身的发展逻辑和内在的成长动力。博弈论文章的增长趋势进一步印证了这种潮流与科研之间的区别。博弈论研究几乎没有任何热度,且由于其高度简化和抽象,以及通常被误认为是一种“不切实际”或“没有用”的方法,而遭到批评,但从图2可以看出仍然有学者不受外界影响,根据自身理解和需要进行博弈论应用研究并取得优秀成果。

## 2. 定量研究论文的议题领域分布

国际关系定量研究不是方法研究,而是应用量化数据和定量方法进行的国际关系研究。因此,国际关系议题领域的论文分布情况可以反映出近十年来中国国际关系的实质性知识积累如何与定量研究相关。

图3显示了样本论文在国际安全、国际政治经济学和外交学研究领域的总体占比情况。其中,国际安全领域包含国际和国内战争与冲突、军事危机与威胁、安全国际合作等议题,国际政治经济学领域主要涵盖经济全球化与国际国内政治之间的互动关系议题,而外交学领域研究主要是常态政治下国家以及次国家行为体在政治领域或带有明显政治目的的交往和互动(比较明显的标志为主题含有“民间外交”和“文化外交”等字样)。



■ 国际政治经济学类论文 ■ 国际安全类论文 ■ 外交学类论文

图3 定量类论文议题领域分布比例

图 3 显示,近十年来我国国关定量研究文章所占比例最高的是以国际政治经济学领域中议题为研究主题的文章,比重达到 42%。国际安全领域中定量研究论文的数量比例位居第二,为 32%。外交学领域定量研究文章的比例最小,为 26%。但这些比例差距并不悬殊,显示出定量研究在这三大领域中都得到较为均衡的发展。国际政治经济学领域的定量文章比例最高,与其和经济学之间的联系直接相关,不但体现在学科交叉的起源,还体现在样本文章中这是唯一一个拥有较多来自其他学科(经济学领域)作者的领域。国际安全领域定量研究的数量和比例超过外交学领域,其中一个重要的原因就在于国际安全领域数据库的可获得性更高和可使用性更为成熟。相对而言,外交学领域的研究在获取、采集和生成数据中需要付出更多的努力。

图 4 呈现了定量研究文章在 3 个领域中的逐年具体数量和变化趋势。由图 4 可见,国际政治经济学领域的定量研究文章数量的增幅在近十年中居于 3 个领域之首,外交学领域文章的增长比较平缓,国际安全领域文章居中。有趣的是,国际政治经济学和外交学领域的定量研究文章发表数量的变化在大部分年份中同步涨跌,而国际安全领域的文章数量则在一些年份上和其他两个领域有此消彼长的关系。这可能和各领域在现实国际关系中的热

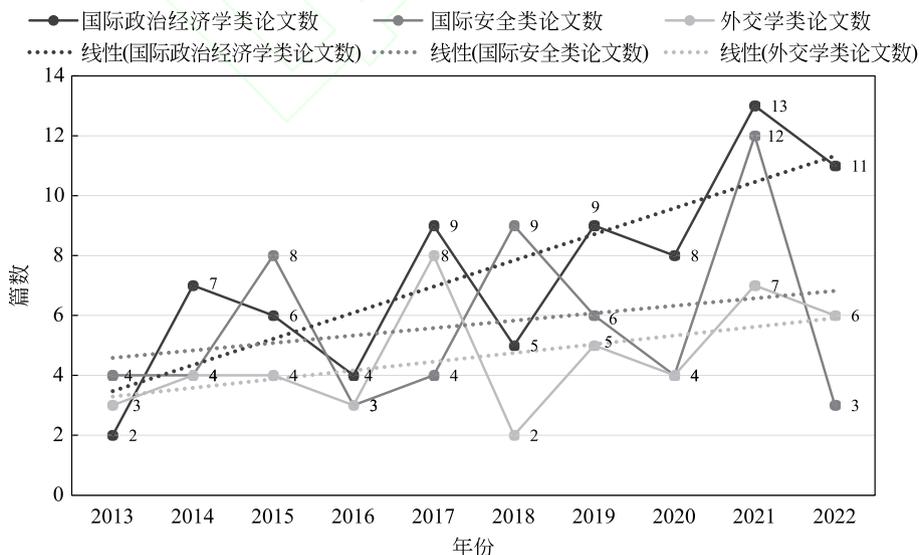


图 4 2013—2022 年定量研究的议题领域分布和变化

度有关——在安全问题受到更多关注的年份,非安全领域的议题热度下降。但由于发表文章有一个时间滞后问题,这种关系难以在时间上进行对应,仅为笔者猜测。

### 3. 定量研究采用的数据库及其原创性

量化数据是定量研究的分析材料,也是研究结论可信度的基础。数据的可得性通常限制了定量研究的问题范围,数据本身的质量对研究的成败和质量具有直接的影响,而数据创新的努力对推动议题的创新和测量的进步具有重要的作用。通过对153篇定量研究文章进行数据使用方面的分析,研究团队发现,数据来源主要以现有经过大量学术研究检验的数据库为主,如学术界通过“数据开发—研究应用”的长期迭代产生的战争相关数据库(Correlates of War,简称COW)、军事盟约ATOP数据库、国家政体数据项目Polity IV、军事冲突数据库(如Armed Conflict Dataset, ACD);UCDP Conflict Termination Dataset,乌普萨拉冲突数据项目)、清华大学国际关系研究院的中外关系衡量数据库等;政府或国际组织发布的官方统计数据,如中国海关数据、《中国贸易外经统计年鉴》、世界银行世界发展指标(WDI)、联合国大会投票数据;得到广泛使用并具有可靠声誉的专门机构搜集的数据,如皮尤研究中心(Pew Research Center)数据、GDELT项目事件数据(全球事件、语言和音调数据库)等等。这些数据库在全球或本国学术界得到了长期和广泛的应用,在可靠性上具有基本的保证。

对成熟和通用数据库的应用无疑是“安全”的,而且获取便利,数据清理也比较容易,能够节约时间成本。但这种数据的运用不会产生由数据推动的研究创新,研究议题受限于现有数据的可得性,有时候不得不“削足适履”。此外,现有数据库对概念的通用性测量或者基于其特有研究目的的测量,往往会与研究者在特定理论和现象情景下对概念所界定的内涵和外延之间存在差距。研究者对数据进行重要修正、补充和原创,对于提高定量研究质量(数据质量以及理论与实证的匹配度)和促进研究创新具有重要的意义。本研究对“原创数据”采用宽泛定义来进行数据创新方面的分析,将对数据进行了较大处理(在原数据的测量上进行改变、更正、扩展)和对数据源进行采集编码形成的新数据都称为“原创数据”。根据这一较低的标准进行统计,本研究发现,具有数据原创性的论文仅占15%。一方面,鉴于统计口

径的宽松,这个百分比是一个值得警醒而非骄傲的数字。另一方面,15%的比例也意味着我国定量研究学者群在数据思维和数据生产方法等方面意识增强和探索积极。样本中的数据原创性主要有以下来源和方法。

第一类是根据研究问题的实证需求,设计民意调查问卷进行的第一手数据采集。此类数据生成方法在样本中的应用场景多为对民众针对国际关系行为体或国际事件的心理感知和政治态度进行询问和量表度量。关于感知、态度和情感的数据可以从一些通用民调项目(如皮尤中心民调数据库、“亚洲晴雨表”等等)获得。这些数据库中的民调问题较为宽泛,有时候可以用于具体研究问题中涉及的一般态度的衡量。但当涉及的问题或现象更为具体时,这些现有数据库中的态度衡量与研究所需存在较大的差距,研究者因此根据研究需要进行了量身定制。

第二类是根据公开政治文本进行编码生成的数据。样本中的政治文本有来自特定国家智库对某一国际重要现象的连续报告、对某类国际关系现象的新闻报道、国际组织出版的年度报告和文献库等等。文本数据是非结构化数据,无法直接进行统计分析,而需要进行结构化处理。文本数据也是高维度数据,其中包含着各类信息,需要从中进行有目的性和针对性的信息提取。因此,文本数据也是大数据的一种重要类型,需要进行提取、编码和结构化。样本文章中的此类工作绝大多数都采取人工编码、人工赋值的方式,罕有采用机器编码的大数据处理方式。此外,由于此类原创性数据在文章中披露信息不足,编码的可靠性和稳健性难以判断,在数据质量方面透明度较低。

第三类是对现有数据库中的指标进行改进或合成,生成新的测量指标。此类数据生成的样本文章中,有的研究根据来自不同数据库中的相关维度测量进行标尺调整和加权,进而生成研究所需要的新指标作为因变量或自变量,有的研究对现有指标进行分解,取其中与自己研究核心变量相符合的构成性部分进行重新加总后生成指标。这些调整性生成数据的原创性虽然低于前两类,但体现了研究者对于理论、测量和方法之间契合度的思考和要求,也是非常有益的工作。

#### 4. 方法特色和应用规范

定量研究中的方法选择取决于研究任务及其数据。任务和数据的多样

性会带来所运用研究方法的多样性,而前者的单一性也会让后者呈现出高度的同质性。当然,由方法驱动的研究在现实中也十分常见,进而导致发表文章在方法类型上的丰富性超过议题的活跃度。

#### 1) 方法应用特色

从本研究所基于的分析样本来看,最常见的数据类型是面板数据(比如  $N$  个国家在  $T$  年中的观察量形成的数据)。这由国际关系研究通常进行跨区间和跨时间同时比较的学科特点决定。相应地,近十年中国国际关系定量研究中使用最多的方法是各类面板数据分析方法,比如固定效应模型、随机效应模型、混合效应模型、多级模型(multi-level modeling),以及在它们基础上的一些变化(如根据因变量是连续变量或某类型的离散变量采取线性或非线性的模型设定)。除面板数据处理方法之外,还有一些研究旨在对时间持续长度进行分析和解释,运用了各种设定的生存模型。此外还有一些研究运用了因果推论的研究设计而非回归模型来进行因果假设的检验,如倾向值匹配法、双重差分法、合成控制法等等。回归模型除了用于因果假设检验外,也被用于考察时间规律,如时间序列分析中的向量自回归模型、多时间序列结构断点识别方法等。在因果假设检验之外,对规律和式样进行识别和描述的研究对社会网络分析方法进行了尝试,通常用于生成各种网络中心度等指标。此外,还有一些分类法、机器学习类的方法运用于预测研究,如分类器和贝叶斯随机变量模型选择等等。

从总体上看,近十年中国国关定量研究中所使用的方法较为丰富,并且集中在对因果关系的检验任务上。因为回归模型是迄今为止社会科学对因果假设进行检验的最常用方法,这就不难理解为什么在我国国关定量研究中,回归模型尤其是处理国际关系领域内最常见的面板数据的各种回归模型的运用最为常见。

#### 2) 内生性处理

本研究不探究学界长期存在的关于回归模型本身是否真的能够检验因果关系的问题,而是更关注我国国关定量研究文章运用回归模型进行的因果检验。笔者认为,研究人员对回归模型参数估计结果的因果解读应在一定程度上持谨慎态度。这就是定量方法应用的规范问题——回归模型的因果解读有其假定前提,某项研究在分析数据时是否思考和处理了这些假定

前提,是重要的研究规范问题。

对回归系数进行因果解读的最重要前提假定就是没有内生性问题。内生性问题有多种来源,如测量误差、循环因果、选择偏差等,而这些来源都可以被看作遗漏变量问题。只要对因果关系进行实证分析,无论是使用定量方法还是使用定性方法,无论是在定量方法中使用回归模型还是因果推论,内生性问题都必须得到细致考察和严格纠正。图 5 显示,样本文章中仅有 29% 的文章对内生性问题进行了讨论和相应处理,其他 71% 的文章中有些文章并非使用定量方法进行因果检验,比如定量趋势、博弈互动和关系描述等不需要考虑内生性问题的研究,但这些研究在比例上比较小。因此,尽管我国学界实际上对于内生性问题的意识和重视程度实际高于 29%,但对内生性问题的重视程度还有待大幅度提高。

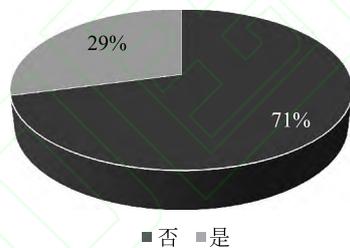


图 5 定量分析中是否对内生性问题进行了讨论和相应处理

在样本文章中,我国学界对内生性问题的处理方法主要有以下四类。第一类是根据原因发生在结果之前的原理,将原因变量进行滞后一期处理,这种做法主要是处理互为因果带来的内生性问题。但进行滞后一期处理的方法其实并不能解决互为因果的问题,因为在这种情况下,滞后一期的解释变量可能受到滞后一期或两期的因变量的影响,如此循环下去。换言之,滞后变量只是对一个时间点上的模型公式中的前因后果设定进行了明确,并不能解决内生性问题。第二类是寻找和使用工具变量,即找到一个变量与内生变量相关,但与误差项之间独立,且其对因变量只通过具有内生性的解释变量发生作用。这样的工具变量其实非常难找,尤其是要论证其严格的外生性很难令人信服。样本文章中的工具变量在选取和论证上大都存在这样的问题。第三类是考虑特殊类型的遗漏变量,即无法测量但不随时间改

变或观察对象改变的特征。这种特殊类型的遗漏变量可以通过面板数据的固定效应估计量得到控制。第四类是对选择偏差进行纠正的方法。在样本文章中的应用例子有回归模型中的赫克曼选择模型(Heckman Selection Model),也有运用因果推论范式进行研究设计(如双重差分、匹配和合成控制等)来进行处理机制纠偏。

### 3) 稳健性检验

定量分析结果可靠性的前提是稳健性。由于回归模型建立在一组特定假定的基础上,如因变量的分布假定、外生性假定、变量间线性或某种非线性关系假定等等。这些假定通常只能“辩护”不能“证明”,但分析结果的数值及其解读却依赖于这些假定。在合理范围内改变这些假定是否能够得到相似的、不改变结论的结果?数据的样本范围、变量的测量选择、控制变量的选择等等都是研究者的主观决定,这些决定又会在多大程度上改变分析结果?实证发现是否在部分样本中依然健在?这些都是稳健性检验要回答的问题。

在稳健性检验方面,样本文章显示出了比处理内生性问题更高的重视程度。文章中使用的检验包括:使用核心变量的不同测量、改变控制变量的选择、改变误差项的分布假定从而选择不同类型的模型、用因果推论替代回归模型来观察两者的结果差异、改变样本覆盖的地区和时间段、加入二次项以检验因变量和自变量之间是否存在关系的转折、加入交互项来观察自变量对因变量的影响程度是否受第三个变量的调节等等。这些稳健性检验反映了中国国际关系定量研究学者对分析结果可靠性的自我规范和同行要求。

## (三) 研究人员特征

中国国际关系定量研究作为一个领域,其活力不仅体现在研究成果的数量和质量上,研究人员及其学术互动作为产生这些成果的源头,也决定了领域的发展潜力和前景。因此,本研究对样本文章的作者基本信息进行了统计和分析。

### 1. 作者群的学术职业阶段分布

本研究以作者职称分类来观察定量研究人员在学术职业阶段的分布情

况,分为教授(含研究员)、副教授(含副研究员)、讲师(含助理教授和助理研究员)、博士研究生和硕士研究生。图6显示了样本文章的作者职称分布和年度变化。由图6可见,从整体上看并没有某一阶段的学者明显占据定量研究的多数,从事国际关系定量研究的学者贯穿了职业发展的各阶段。副教授群体相对而言占发文数量比重最高,这一结果与其处于相对高产的阶段有关。有趣的是,讲师群体在2019年前的定量研究发文数量要明显低于教授、副教授甚至博士生群体,并且一直保持在数量较低的状态,缺乏变化趋势。但在2021年突然增长至17篇,成为当年发文最多的群体,而在2022年又跌回之前水平。与之形成对比的是,除硕士生群体外的其他三个群体的增长趋势都比较稳定,这可能与2019年前后海外留学回国任教的青年教师中使用定量方法者明显增加有关。这个群体一方面致力于在英文期刊上发表其研究成果,另一方面也在中文期刊上发表定量研究成果,因此,尽管他们在持续产出国际关系定量研究,但中文期刊可能呈现出他们的发表量波动较大的情况。实际上,2021年的发表峰值本身就表明这个群体很有可能是当前定量研究最活跃的群体。博士群体在中文期刊上发表定量研究数量之高有些令人惊讶,因为他们在诸多年份超过讲师和教授群体,甚至在2019年成为发文量最大的一个群体。这展现了年轻学者对定量研究的学习能力和应用兴趣,同时也反映了博士群体的定量研究文章可能更多依赖中文期刊作为发表渠道。

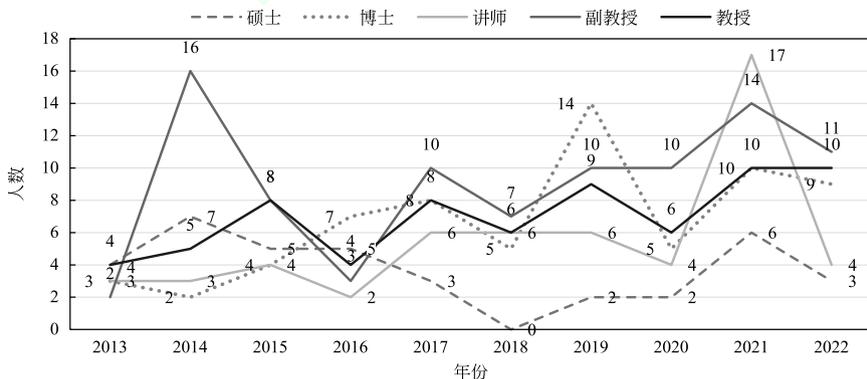


图6 定量研究文章作者群学术职业阶段分布

此外,十年时间中学者成长也会影响到定量文章作者职称类型的分布和变化。一个学者在十年期间可以完成图6中不同折线间的多次跳跃。但是图6并没有呈现定量研究发文数量在不同群体之间此消彼长的现象,这说明有源源不断的不同阶段的研究者加入到定量研究中来,而不是大致同一群研究者在十年之中发表定量文章。图6表明,定量研究人数规模增长比较快,在2013年作者的总体数量是16人,而最高点的2021年为57人,2022年是37人,十年之间人数增长数倍。需要注意的是,这仅是发表文章的作者数量,实际从事国际关系定量研究的中国学者群体规模要更大。

这一分析也显现出我国国关界定量研究作者群具有年轻化特征,这表明量化方法在近十年来对学界年轻学者的吸引力较大,他们在定量文章发表方面相对其他类型研究更有优势。定量研究方法的学习有一定难度,但在学习定量方法和编程技术上年轻人相对于更资深学者具有一定的优势。与此同时,相对于理论工作和思辨方法而言,定量研究更能够进行程式化学习,也能够实现更快捷的实际研究应用。此外,定量研究的问题切口一般比较小,研究任务能够被限制在一定范围中,不像理论和思辨工作具有更模糊的边界或者更容易牵扯出越来越多的研究任务,所以定量文章在一定时间中的完成度更高,这些都对年轻学者更具吸引力。而更资深的学者倾向于对大问题进行思考,对小切口问题兴趣更低,对文章的完成时效也更不看重。这些都可能解释了中国国际关系研究在人员分布上年轻化的现象。定量研究者年轻化意味着在可见的将来定量研究会具有持续的发展前景。

## 2. 作者所属机构分布

本研究对样本文章作者所属高校或研究所的情况也进行了分析,意在观察定量研究发展的机构基础。这一分析揭示出中国国关定量研究发展中有趣但又令人担忧的一些情况。

表1统计了作者所属机构的近十年发文数总量,图7显示了机构的占比。在表1中,笔者只统计了发文3篇以上的机构。<sup>①</sup> 一个非常引人注目的

---

<sup>①</sup> 中南财经政法大学、中共江苏省委党校、浙江师范大学、上海外国语大学、香港中文大学和台湾政治大学等都有学者发表了定量研究论文,但由于总数不足3篇,不展现在表1中。

表 1 发表定量研究文章作者所属机构及发文数量

| 高校/研究所    | 篇数 |
|-----------|----|
| 清华大学      | 37 |
| 中国社会科学院   | 18 |
| 中国人民大学    | 16 |
| 南开大学      | 10 |
| 中山大学      | 9  |
| 辽宁大学      | 8  |
| 国际关系学院    | 7  |
| 北京大学      | 6  |
| 上海国际问题研究院 | 4  |
| 南京大学      | 4  |
| 华东政法大学    | 4  |
| 复旦大学      | 3  |
| 其他各高校     | 27 |

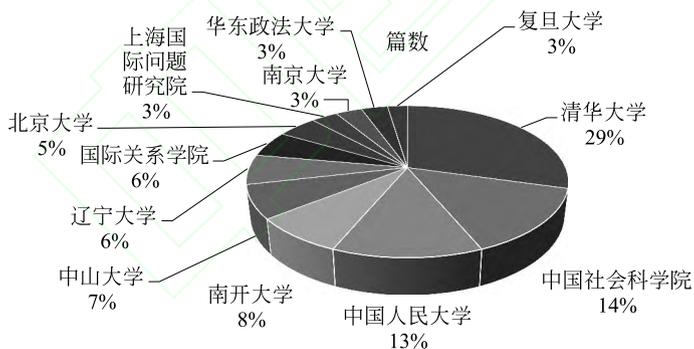


图 7 发表 3 篇以上定量研究文章的作者单位分布

数字就是清华大学的总发文量为 37 篇,占比 29%,远高于其他机构,比位居第二和第三的中国社会科学院和中国人民大学加在一起的总发文量还要多 3 篇,占比高 2%。这印证了中国定量研究发展和清华大学国际关系学系之间的密切关系,这也是近十年来中国国际关系学界的普遍感受。前 3 家机构一共占到总发文量的 56%,超过第四位及其之后的其他机构发文量的加总。换言之,一大半的定量研究文章来自前 3 家机构的国关研究者。值得注意的是,这种差距与机构的整体学科规模和研究实力之间并无关联,表明差

距并不是由人员数量或研究实力造成的,而是定量研究发文量稀少的机构对这种类型研究主动回避。这表明我国国关学界对定量研究的接受度和热情度在分布上十分不均,只有少数机构具有持续的积极态度。

在机构的地理位置分布上,本研究也发现了一个明显的特征,那就是对国际关系的定量研究探索主要集中在北方高校和研究所,南方高校发表的定量研究文章相对较少。背后的原因值得进一步探索。其中一个可能的原因是学术流动性特征,北方高校的博士毕业生更多在北方机构寻求职位,南方亦然。

从总体上看,中国国际关系定量研究在地理上和机构上都呈现出“分化—集中”的特征,表明定量研究并没有成为这一学科的常备研究工具,而且在对其作为工具的认识和态度上学界共识也较低,部分地区和机构采取非常积极的态度和行动,但其他地区 and 机构则相对怀疑和冷淡。

### 3. 独作或合作比例

相对于思想性、理论性和定性研究,定量研究更容易将研究任务分解和组合,从而更为适合合作性研究。与此同时,由于国际关系领域无论在国际还是国内都没有分化出“理论国际关系”和“实证国际关系”的专门领域,一项国际关系研究需要涵盖从问题的提出、理论的构建和假设提出以及数据搜集清理、方法应用等等全过程,但其实每一个环节的专业性都较强且在思维上要求各有侧重。这意味着,如果合作者之间有合理的专长配置和紧密的协作,合作进行的国际关系定量研究的质量会比独立研究更高。当然,并非所有的联合署名研究都是配置合理的合作研究,甚至有一些联合署名研究在实际中的合作度相当低。但可以肯定的是,促进良性而有效的研究合作是推动定量研究发展的重要途径。本研究因此对样本中的合作和独作比例进行了统计和分析。

由图8可见,对于定量和博弈论文章,研究人员之间的合作呈上升趋势,而独作数量维持稳定。尤其考虑到近年来的学术评价体系的变化趋势是朝向不鼓励合作研究的方向发展,合作论文数量在近几年中的持续上升表明了定量研究合作的必要性和学者遵循研究规律而非完全按照评价指挥棒进行工作。需要注意的一点是,其中较多的合作是教师和学生之间的合作,这在一定程度上规避了学术评价带来的不利因素,同时也是科研育人的方式



图 8 定量和博弈类文章的合作和独作情况

和体现。但教师之间的合作相对少很多,这不利于提升定量研究质量、探索前沿议题、进行长期投入。而且,独作文章的数量一直保持较为平稳的趋势,这也反映了学术评价体系的影响。

#### (四) 研究成果正式发表平台特征

学术期刊是学术成果发布的重要平台。学术期刊实行同行评议制度,对学术成果的可靠性和学术贡献进行的程序性和实质性把关的严格程度最高,学术期刊对其发表的文章在质量上和信誉上都进行背书,不但对学术成果起到了传播作用,而且发表本身即成为对学术价值的肯定。因此,学术期刊发表什么类型、什么主题、什么风格的论文,对于一个领域的发展十分重要,对于学科学术观念、研究任务、学术品味和论文规范等都产生深刻而广泛的影响。相对于大部分英文期刊而言,中国的学术期刊在研究议题引导和学科发展塑造方面扮演着更加积极和主动的角色,这就让本研究统计中国定量研究论文的发表期刊分布具有了更重要的意义。

图 9 报告了样本文章的 6 种来源期刊各自在近十年发表国际关系定量研究和博弈论文章的数量。图 9 中最为显著的一点是,《世界经济与政治》期刊发表的定量和博弈论研究文章数量在十年中一直居于首位,而且远高出



图9 2013—2022年6种期刊发表定量和博弈论文章数量及变化趋势

其他期刊。除2014年和2020年外,《世界经济与政治》上发表的定量博弈论文章数量甚至多于或等于其他5种期刊的总和,无疑是这类研究的最主要发表渠道。这与该期刊由中国社会科学院世界经济与政治研究所主办有关。该机构的研究领域包含经济学,对定量分析采取更为常态化的认知态度,并且在论文评审过程中能够在国际关系学科外引入评审资源,降低了为定量和博弈论文章寻找合适评审人的难度。除《世界经济与政治》之外,《当代亚太》和《国际政治科学》相对而言在刊发定量和博弈论文章方面较为积极,尤其是《当代亚太》从2017年以来发表此类文章的数量有较为明显的增长趋势,但这一趋势在2022年中断。《中国社会科学》刊发的此类文章最少,这与该刊是人文社科综合性期刊有关,该期刊上发表的国际关系类论文总量就非常少。

#### 四、现状、进展与问题分析总结

本研究对中文期刊上发表的国际关系定量研究成果进行的回顾和总结,显现出中国国际关系定量研究近十年来的持续发展,也的确形成了一个初具规模的领域。虽然总体数量占比并不大,但国际关系定量研究在文章结构形态的外观、分析手段和学术理念上都具有明显的类型化,这让数量不

多的这一类文章形成了某种可以被清晰感知到的合力,超出了文章数量本身所显现出来的发展势头和影响范围。与此同时,国关定量研究数据使用和相应技术的多样性,让定量研究文章给人以“异彩纷呈”和“推陈出新”的观感。由于定量研究发表的主要渠道《世界经济与政治》期刊在中国国际关系学界拥有很高的声誉和广泛的影响,也让国关定量研究在近十年的发展形成了令人瞩目的学科级现象。但需要清醒意识到的是,尽管中国国际关系定量研究在近十年来的确有长足的发展,但实际的推进速度、范围、质量和数量都并没有直观印象中的那么显著。其中隐藏着一些还未充分发掘的机会,潜伏着一些令人担忧的障碍,也出现了一些有可能陷入误区的苗头。对此,本研究结合近十年我国国关定量研究发展的一些重要特征进行分析和探讨。

### (一) 实证研究任务的选择特征

#### 1. 定量研究类型以因果检验为主

就科学研究任务而非国际关系议题而言,定量研究可以分为描述、测度、因果(解释)和预测等不同的类型。近十年来中国国际关系定量研究以解释性研究为主,实证任务高度集中于对因果假设进行检验,旨在提供一类现象是否以及在何种程度上影响到另一类现象的实证证据。相对而言,对其他类型的研究任务关注和投入过少,取得的研究成果的比例过小。诚然,因果关切可以说是科学研究中最核心、最终极的任务,但其他类型的研究是因果解释的前提、基础和应用,在知识体系构筑和推进中的重要性并不逊于最终的因果解释。

我国国关定量研究对解释性研究尤其专注,其中一个可能的原因是,定量方法最初在中国国关学界进行推广时,为了简明扼要,通常将“科学”研究等同于回答“为什么”问题(why question)。这不但与观念有关,也是一种在推广初期能够产生很好效果的策略。“为什么”(解释性)的定量研究可以被简化为“因变量—自变量”的程式化思考,容易上手,便于普及。此类文章写作也具有比较成熟的框架,即“导论—文献回顾—论点和假设—定量检验—结论”。对于初学者而言,根据框架进行模仿学习比较有迹可循。

探索性、描述性、测量和预测研究所占的比例过少,不但是因为对它们

在知识创造中的价值认识不足,也与它们在构思、执行和报告等方面要求更高的创造力、更难程式化、文章也不符合刻板印象中的“科学”或“定量”形态有关。目前,我国国关学界尤其需要关注测量型研究过少、质量也参差不齐的问题。在样本文章中,测量通常是在因果研究中作为一个变量的数据生成任务来进行的,没有严格遵照测量研究的程序和标准。这样生成的原创性数据能够被重复性、累积性应用的可能性很小,往往成为“一次性”数据。这不仅导致中国国际关系定量数据创新和积累方面进步缓慢,也不利于建立和加强数据思维。其中一个需要纠偏的趋势就是指标研究在测度型研究中占据了主导地位。现有指标研究常常显现出一些重要的不足,如缺乏对所关注的理论概念进行深入思考、没有在国际关系知识体系背景中进行概念化及操作化、采取“赋值一加总”方式过于主观等等。相反,建立在测度模型和信息提取算法基础上的测度型研究还非常薄弱,需得到研究者更多的关注和投入。

## 2. 数据便利性: 选题的重要考量因素

与测量研究相对稀缺密切相关的是,数据便利性在很大程度上划定了中国国际关系定量研究的选题范围。数据“便利性”不同于数据“可得性”。数据可得性是研究任务所需数据是否能够通过可行的手段加以获得。可得的数据是在现有的信息、技术、时间和伦理条件下具有获得的可能性和可行性。它们可以是已经存在于结构化数据库中能够加以提取的数据,或者是需要研究者花费大量时间进行数据挖掘、访谈、民调或实验等得以生成、搜集、整理和甄别的数据。而便利性的数据则是指数据可以用低成本的方式获取,通过简单合并和清理即可进行分析。这些数据通常是结构化数据,而且在他人研究中已经得到了重复使用,比如复刻数据和专题或通用数据库。数据便利性极大地降低了研究的时间、人力和智力成本,也可以不必就数据可靠性向论文评议人进行说服。在发表绩效考核压力大、将学者研究水平和贡献与发表数量高度挂钩的科研评价体系下,根据数据便利性来选择研究课题不失为一种个体层面上理性又无奈的策略。但在研究领域和学科的层面上看,数据便利性驱动研究选题带来的后果是消极的,导致国际关系定量研究在议题方向上趋同性程度较高,显露出以分析技术翻新来追求新意的现象以及数据创新和积累程度低的问题等。

数据便利性驱动研究产生的另一个后果就是,定量研究集中在国际关

系传统议题领域和宏观层面,因为在这里数据经过长时间的积累具有更高的便利性。这与国际学界同领域的定量研究议题动态发展方向形成鲜明对比。国际上的研究议题近年来明显转向气候变化、全球治理/国际组织、政治心理学、对国际关系学基本概念的反思和测量(权力、权威、结构、利益/偏好、规范等等)等。这些更具开拓性的议题要求研究者对所用数据进行原创,而数据的原创性也就成为评价定量研究价值和贡献越来越重要的考量。实验方法在国际上的大量使用更是加大了对数据投入的要求,且这些数据无法被重复使用,下一个研究即使做同样的问题也必须进行自己的实验,生成自己的数据。<sup>①</sup>中国的国际关系定量研究在数据工作上投入不够导致议题的更新和活跃程度远低于国际水平,且尚未发挥定量研究对微观层次的国际关系现象和理论的推动作用。

## (二) 研究的方法特征

与研究议题方向相比,中国国关定量研究近十年来在技术应用方面推陈出新的速度较快,显示了对方法新颖性的热情和技术复杂性的追求。这反映出对定量方法的兴趣是本学科定量研究发展的主要驱动力,也印证了将定量研究在很大程度上视同“技术型研究”的普遍观念。此外,一些期刊对有技术亮点的文章较为偏好强化了这种特征。

### 1. 以回归模型为主导性方法的原因和问题

中国国际关系定量研究以因果检验为主,因此对多样性方法的尝试相应集中在回归模型领域。回归系数具有“其他因素不变”的因果解读,是经济学和其他社会科学中长期用于因果检验的主要方法。近十年来,中国国关研究中用到的回归模型以面板模型和多级模型为主,在模型设定的复杂程度和参数估计技术难度上都有着明显上升。在使用回归模型时,对内生性问题的意识和处理明显加强,使用工具变量、两阶段回归、样本选择模型等技术来有意识处理内生性问题的文章开始增多。但是,对于谨慎选择控

---

<sup>①</sup> 后来研究可以重复实验设计来复刻实验结论,但实验本身生成的数据是一次性使用的。这是另一类型的“一次性”数据,但这种数据生成的目的就是一次性的,可积累的是数据生成的方法,即实验设计。

制变量来避免或减轻内生性问题的基本路径却不够重视。定量文章中对控制变量选择的讨论(或不加讨论)体现出对控制变量在回归模型中扮演角色和发挥功能的了解不足或误解。不少研究只讨论控制变量和因变量之间的关系,或者以现有研究中的“惯例”来合理化控制变量的选择。控制变量选择的随意性或“从众”风格体现出对回归模型的“统计控制”原理解理解和掌握不够,也反映出对回归模型用于因果研究和预测研究的不同设定标准和要求未能进行清楚区分。

对混杂因子(confounders)进行细致分析,在此基础上尽量寻找到它们的测量作为控制变量来进行纠正,这应该是第一步,也是最重要的一步,尽管这一步在方法和技术上并没有什么可观之处。实际上,对无法测量的变量带来的内生性问题,无论是用回归方法还是用因果推论的研究设计,能够纠正的偏差都非常有限,而且只能在一些非常特殊的场景下才能够起作用。选择控制变量就是思考“数据生成过程”或者“选择机制”,其重要性在因果检验中怎么强调也不为过。我国国关定量研究需要切实加强这方面的意识和工作。

## 2. 大数据方法的应用和不足

### 1) 社会网络分析

在回归模型之外,社会网络分析在国际关系定量研究中的使用频率在近年来呈明显上升趋势。社会网络分析在一定程度上可以被视为大数据方法,并非因为它处理的数据量一定有多大,核心是它不仅打破传统方法所依赖的观察量独立假定,而且以分析观察量之间相互关系及其形成的结构特征为主要任务,适用于“行动者—结构”的双向分析,与国际关系学科关注问题的性质十分契合,因此社会网络分析方法一经引入国际关系学科后就深受国关学者青睐。与此同时,二元数据(dyadic data)库在近年来的大量出现提供了构建社会网络的基本数据单元(行为体及其关系)。

社会网络分析并非单一分析方法,而是包含回归模型(如 ERGM)、系统描述(如各种网络指标)、模拟预测等等。这也表明社会网络分析并非服务于某一特定研究任务,而是以社会网络视角来理解国际关系现象。但现有中国国关定量研究呈现的一个特征就是更为注重作为技术的社会网络分析,而对其作为一种重要的观察和分析视角的理解、重视和探索尚待加强。

对社会网络理论和视角的重视不够,在样本文章也体现出来,导致方法的一些误用,如将在逻辑和直觉上缺乏传递性和高阶性质的二元观察量进行技术连接而构建起牵强的“网络”来展现分析技术等。

### 2) 数据可视化

数据可视化的进步是大数据对中国国际关系定量研究的一个清晰可见的影响。尽管可视化技术与大数据之间没有必然的联系,但在我国国际关系学界,正是对于大数据的热情激发了对数据可视化的兴趣。定量研究的文章逐渐以图形替代数字表格来进行更为高效的表达,信息传递更为直观,也节约了期刊版面。与此同时,可视化的图形和色彩具有强大的视觉冲击力,比枯燥的数字和沉闷的表格更能引起阅读的兴趣。可视化的代码能够进行轻松复制,初学者通过跑代码也能够计算机中生成令人惊叹的图形,带来较快的成就感和强烈的吸引力。这些都增加了定量研究对初学者的吸引力,有助于激发他们学习数据处理和编程技能的热情,是推广定量研究的有效途径。

但可视化的价值并非仅是为了表达高效,更不是为了吸引眼球,其实质是进行创造性和探索性工作。研究人员通过尝试合适的可视化技术来发现数据中不易观察的式样、趋势或问题。在数据思维薄弱的人群中,可视化的吸引力却集中在了外观的炫酷和技术的夸耀,并鼓励了这些倾向。从样本文章来看,数据可视化在问题激发方面的功能使用不足,文中对因变量(待解释现象)和自变量(关注的原因或处理状态分布)的统计描述可视化不够,更集中在对分析结果的可视化上。应用可视化技术对国际关系现象进行真正意义上的大数据描述,从复杂和海量数据中发现国际关系互动和结构模式,产生新颖的研究问题,探索对问题的创造性切入角度等,是未来需要鼓励和加强的方向。

### 3) 文本数据

从文本数据中提取信息生成结构化数据进行描述、测量、因果和预测研究是大数据革命带给社会科学研究的最大红利之一。但中国国际关系定量研究的“大数据热”却极大地冷落了文本数据这种重要的大数据。样本论文中有使用文本数据进行人工编码生成所需变量的努力,但这尚处于“前大数据”阶段。文本分析方法(对文本进行机器编码和机器学习等)的应用严重

低于国际水平和频度,导致在使用文本的效率、数量、范围和任务等方面都十分受限。中国国际关系定量研究使用的第三方结构化数据很多已经是国际学术界通过文本挖掘和分析并加以结构化的数据。在这个意义上,文本数据对中国国际关系定量研究产生了重大的影响,但这种影响到目前为止主要是单向的。中国国际关系定量研究尚缺乏对文本数据和文本方法应用产生的学术贡献。我国国际关系定量学者未来应在这方面作出进一步努力,尤其是展开跨学科的合作,充分利用当前基于文本数据的深度学习的巨大进展给国际关系学研究开辟的新机遇。

### 3. 因果推论方法应用的滞后

近十年来中国国际关系定量研究以对因果关系的实证检验为主,但中国学者对因果推论作为一种方法范式的接受程度却相对较低,且热情不高。因果推论的核心不是数据处理的计算技术,而是关于如何设计因果效应识别策略的逻辑思考。因果推论大幅度提高了对研究信度的评判标准,增加了研究者自我质疑和相互质疑的维度和深度。因果推论将大量的功夫下在反复推敲而不是最后能够在文章中呈现出来的工作量,并明确承认技术无法对思考进行替代或补偿。因果推论工作的艰辛和精彩具有内在性,不以数据量、可视化、复杂模型和吸睛技术的外在方式体现。这些都让因果推论这种方法显得“吃力不讨好”,其受到冷遇也不难理解。

在近些年来的中国国际关系定量研究中,也零星出现了一些应用因果推论的文章,比如利用倾向性匹配、合成控制法、差分法、工具变量等来识别因果效应。这些尝试有助于将识别意识和信度标准引入国际关系学科的发展中。然而,这些因果推论的定量研究中也显现出一些“设计意识”上的重要不足,体现在将倾向性匹配和差分法视为技术而不是识别策略、对“自然实验”概念和设计的误用、将差分法等同于一种具有处理内生性神奇功效的回归模型、选择的工具变量缺乏对严格外生性的充分合理化论证等。对这些问题的克服有赖于因果推论系统性和思想性训练的加强,尤其要避免将因果推论当成一种技术和技巧来加以学习、应用和推广。

此外,中国国际关系因果推论定量研究基本为观察性研究,实验研究尚未展开。这与中国国际关系研究以宏观层面研究为主有直接关系,对于“人”(决策者或公民)在国际关系中的角色不够重视。以国家作为研究单元

很难(或不可能)开展实验研究。而缺乏通过实验研究对因果推断的直观体验,也阻碍了中国国际关系研究者对因果推论的理解和接受。因果推论的思想和原则以随机实验为基准,着眼于数据生成过程而非数据分析技术来进行因果识别。缺乏对实验方法的直接感知,会让因果推论中的原则和思想显得过于抽象甚至不近人情。更为重要的是,实验方法无用武之地反映的问题并非“有一种重要方法我们还没有用到”,而是我国对国际关系的微观基础缺乏关注。与之相对比,国际学界更重视对微观层面上产生的态度、情绪、认知、互动等对国际关系的影响进行理解,积极开展实验室实验和田野实验,产生了很多优秀研究成果,值得我们学习。与中国政治学的横向比较也表明,政治学对微观基础和个人角色的关注让这个领域在近十年内对实验方法的接受和开展速度很快,同时也促进了因果推论的思想体系和设计方法在政治学领域更为学者和学生所理解和接受。

#### 4. 数据思维开始形成但尚待加强

研究议题及其涉及的数据决定选取何种方法更为可能和恰当。近十年来中国国关定量研究显示出了有关的研究规范和研究方式的形成和加强,呈现出方法上的多样性和丰富性。数据思维,即对数据如何联系理论和方法进行思考,在国际关系学界逐渐形成。对数据作为信息载体的理解提升,推动中国国关学者重视数据在研究中的作用,对于所关心的议题领域中国际国内新发布的数据库具有敏感性,以及将数据搜集和分析工作作为研究工作的重要组成部分而投入大量精力和智力。

但本研究分析也发现,我国国关定量研究中的数据思维还比较薄弱,原因之一是上文提到的数据便利性对议题和研究的推动,这不利于研究者对数据进行深入思考、探索和考察。研究本应按照“议题—理论—数据”的顺序来展开,数据应根据特定理论中概念的内涵和适用范围来进行搜集或搜寻,这让数据工作变得艰难,需要对数据、现象、变量、概念之间的关系进行深入思考。但由便利性数据驱动的研究实际上沿着“数据—议题—理论”的路径进行,研究者根据手中已有的数据来确定因变量、自变量和控制变量等,甚至将变量的名称直接等同于理论概念,由变量来构建理论假设,省去对理论概念和实证数据之间跨度和差异的深思熟虑,也错过了训练和强化数据思维的机会。

数据思维薄弱不是指数据分析技能不足,而是如何看待数据,包括数据的可靠性、数据不同的测量误差和偏差、数据与所关注现象/问题/概念的匹配程度、数据(不)能够用于识别什么、拥有的数据和必要的数据之间的缺口等等。样本文章对数据的检视大都还停留在报告数据的来源和描述性统计值,表现出将数据视为给定的倾向。实际上,“定量”中最核心和最基础的部分是数据,数据的特点和问题决定了方法的需求和选择。数据是实证研究的枢纽,数据思维薄弱会产生向前的影响(实证与理论的不相契合)和向后的影响(方法和数据的不相契合)。

大数据热是近几年来中国国际关系定量研究中的一个显著现象,但数据思维薄弱在其中也表现出来,那就是倾向于以“数据大”和“分析技术复杂”来开展大数据研究,而对大数据在推动学术创新上更为重要和核心的潜力的理解和挖掘不够。相对于数量规模,大数据更重要的是极大地拓展了数据形态(如文本、图像、声音、视频)和数据来源,将研究问题和思路从数据限制中解放出来,而不是像过去那样只能根据有什么数据来决定做什么研究。中国国际关系定量研究对大数据展现出了很高的热情,但过度强调量大和技术却阻碍了真正投入大数据革命。

综上,数据思维是当前和未来中国国际关系定量研究需要特别加强的方面。

### (三) 中国国际关系定量研究的领域特征

将中国国际关系定量研究作为一个领域(尽管是“过渡性”的)来探讨,本研究在对人员、专业标准和发表渠道等分析的基础上作一些扩展性探讨。

首先,本研究对样本文章作者的职业阶段进行的分析表明,国际关系定量研究对青年学者的吸引力较大,且这种吸引力具有一定的持久性。这对于该领域的未来发展无疑是积极的。但另一方面,缺乏成熟学者的中途加入(当前处于教授职称阶段的定量研究者大都是在十年前青年学者阶段就从事定量研究的学者)也带来潜在问题。这表明中国国际关系定量研究较少传承到通过长时间观察和思考而获得理论深度和对现象把握的能力,缺乏资深学者对定量研究进行的理论引导和现实指引,不利于避免重技术而轻实质、重数据而轻理论、重流程而轻思想等可能倾向。

国际发展经验也表明了资深学者的参与和引领对于国际关系定量研究发展的重要性。例如,在美国国际关系学中,从事定量研究的学者形成更为合理的梯队,其中不乏以理论成就著称的资深学者。这些资深学者尽管并不掌握新颖的定量技术,在其青年学者时期也可能没有受过定量方法训练,但他们之后在与后辈学者的合作中逐渐掌握和深化了对定量方法原理和技术理解和掌握。这些学者在国际关系定量研究中并不是简单的参与者,而是真正的领航者,把握选题的大方向和审视研究的国际关系理论价值,并让青年学者在技术层面上发挥优势,产生了为数不少的将理论、现实和技术很好结合的佳作,也有助于后辈学者加强理论修养和现实理解能力,促进了定量方法为国际关系研究议程服务(而非相反)的共识。但当前我国的情况却是,除了极少数特例外,资深学者不大涉足定量研究领域。中国国际关系定量研究的未来发展不仅要培养青年一代正确掌握定量工具以及具备不断更新的能力,还应该进行特别的努力让资深学者能够更多参与到定量研究中来,关心定量研究的发展。青年学者也应该卸下自己的“技术盔甲”,创造和抓住向资深学者虚心学习、耳濡目染和研究合作的机会,扩展理论视野和培养思想深度。

从发表渠道来看,定量研究文章的发表高度集中在少数几个期刊,而在为数众多的其他国际关系期刊却毫无踪迹或非常零星。尤为重要的是,近十年来定量研究文章的发表高度依赖《世界经济与政治》这一种期刊。尽管这一期刊具有很高的国内行业声望,但其对定量研究的接受程度高却并没有对其他期刊的发表偏好产生明显影响。定量研究发表渠道的高集中度其实也是其高度的脆弱性,对于未来发展的稳定性而言令人担忧。

最后补充一点样本文章分析之外的观察和思考,那就是关于训练和认证专业人士的公认标准问题。中国国际关系定量研究已有不少定期和不定期的培训班和训练营,一些机构也向完成培训的学员发放认证。这些努力促进了定量研究领域诸多共识的形成,如研究规范性、问题意识、研究任务的明确、数据的报告、研究透明度和可复刻、对研究合作的积极态度等。但中国国际关系学界定量方法培训班的大受欢迎也折射出系统性方法课程的缺失。培训班长期作为课程系统缺失的补充而不是在课程体系之上的专门化提升,在一定程度上甚至成了课程体系的替代。高校和研究所里开设

国际关系定量方法系列课程的风毛麟角,而这些课程也很少形成一个前后衔接、循序渐进的系列。系统性的课程训练和严格的课程考核对于形成学科标准不可或缺,课程建设的滞后局面需要尽快得到改观。

#### (四) 对学科“科学化”的推动

定量研究对中国国际关系学科发展的影响,是一个长期备受关注也富有争议的话题。定量研究推动国际关系学科学化,这是由定量方法本身作为科学方法的重要组成部分而决定的。在国际关系研究中应用定量方法,这本身就表明研究采取了科学而非人文的范式。科学和人文范式只是两种不同的范式,并无高下对错之分。选择了定量研究也就是选择了科学范式,并不代表进步或倒退,只是具体研究任务下的具体选择。中国国际关系学是否应该提高其科学化程度,这也是仁者见仁、智者见智的问题,超出本研究讨论的范畴。本研究仅就定量研究在近十年来推动学科科学化的影响加以分析,而不进行任何价值判断。

定量研究推动中国国际关系学科科学化,在浅层上表现为采取程式化的科学研究程序(假设提出—实证检验)、科学分析工具的使用(定量方法)以及科学论文的写作格式和规范等。在更深一层的意义上,学科的科学化表现在对研究效度和信度的重视程度提高和对实证工作的投入加大。国际关系定量研究以明确的方式承认,在论点的逻辑阐述之外,研究者还负有举证的责任,对实证证据的获取、处理、呈现和鉴别成为研究的主要工作之一,对实证证据的系统性、无偏性、一致性以及与论点之间的匹配度等进行严格要求,这些是研究效度和信度的重要保障。我国国际关系定量研究近十年来在实证操作上有着长足的进步,例如,对内生性问题处理意识和技术的加强,就体现了对研究信度的自我要求和同行要求的提高;对理论与实证脱节现象的自我反省和公开讨论,也正是提升研究效度行业标准的自觉努力。

在深层次上,科学化是科学精神的内化,仅仅是技术化和实证化无法达到这个层次上的科学化,且可能损害科学精神的内化。科学精神的核心是“求真”,而科学中的“真”只能通过不断而坚决的“去伪”来接近和显现。这里的“伪”既包括形式上的“伪装”,也指内容上的错误或不可靠。中国国际关系定量研究在近十年的发展中遭受最大也最普遍的质疑正是关于其浅层

科学化(形式和技术)是否对研究内容的肤浅、分析的偏差和证据的错误造成掩盖,不是“去伪存真”而是以技术门槛来庇护了“伪”。这样的质疑往往引发“范式之争”,但国关定量研究者与其面向质疑者进行自我辩护,不如在质疑中进行自我检视和提高,尤其是在研究中首先做到严格的自我质疑和相互质疑,这也正是科学精神的集中体现。

在这一方面,我国定量研究未来还需做得更为自觉和严格。目前来看,定量研究文章大都采取全方位“自我捍卫”的写作风格,而缺少对研究缺陷、局限性、前提假定、数据问题、方法受限、结论不确定性等问题的公开承认和讨论。这可能是出于对研究贡献的误解。但其实,无懈可击和研究贡献之间并没有必然的关系。所谓前沿探索的价值,不仅在于看到从前未知的风景,还在于触碰到新的困难和局限。报告这些困难和局限本身构成了研究的价值。此外,相互质疑在我国定量研究中主要集中于匿名评审阶段,在研究发表前和发表后,定量研究学者之间进行公开的相互质疑还非常罕见,相互质疑的建设性学术文化有待形成。这种文化不仅指质疑时的“对事不对人”,更重要的是,质疑不是彰显自己的犀利或学养,而是尽量帮助作者改进研究工作和找到弥补方案。对研究进行总体而模糊性的负面判断不是科学质疑,在不考虑时间和资源限制的条件下对研究进行“菜单式要求”也不是科学质疑。加强自我质疑和相互质疑,形成建设性的学术质疑文化,是定量研究未来在自我发展和推动学科发展的道路上应该切实加强的地方。

## 六、结语：进路探讨

从总体上看,近十年中国国际关系定量研究展现了良好的发展势头和持续的生命力,体现了中国国际关系学科积极探索创新空间、锐意进取和拥抱变化的精神,是学科自身在发展阶段中产生的内在强大驱动力,也是学科对国际国内学术整体变化发展的响应和回应。定量方法在自然科学和社会科学中的应用和发展有着悠久的历史,已被证明既不可能(也不应该)是“一统天下”的科研工具,也绝不是一时的潮流,而是有着持久且与其他方法互补的功能和用途。定量方法在中国国际关系学中曾经被视为新奇之物,但经过短短一二十年的发展如今已成了很多学者工具箱中的常备和常规研究

工具。近年来的大数据潮流激发了对中国国际关系定量研究的热情,但建立在科学原理和累积性知识基础上的定量方法本身并不是潮流的产物,在学术研究中的命运和前景也不会受大数据潮起潮落的影响。定量方法对中国国际关系学者的吸引力归根结底在于这种研究范式和方法本身的学术之美,这也是定量研究具有持久生命力的根源。

本研究分析发现,从总体上看,我国国关学界在最近十年对定量方法的应用取得了长足的进步,对所采用回归模型的应用前提和合理检验都有规范性操作。随着技术的进步,我国国关学界除了对变量间关系进行回归分析,还加强了对相关作用机制的探讨,并对可能存在的内生性问题进行了稳健性检验。我国国关学界不仅能采用各种合适的模型进行分析,而且有一定数量的原创数据,也进行了合适的检验,研究程序的规范性和结论的可信性得到加强。当然,我国国关学界在定量研究上也存在不足,定量研究对理论突破的推动作用有待加强。囿于数据所限,目前我国国关学界有的文章还存在数据驱动的现象,进而导致对深层作用机理的挖掘探析尚显不足。有的文章虽然发现了变量间的统计关系,但对背后的机理和作用机制考虑不深,存在重视数据而忽视理论的现象。

本研究的开展并非出于扩大这个领域或提升其学科地位的关切。进行这样的分析、总结和反思,目的在于挖掘和发挥定量方法对国际关系学科知识创造和积累、对学科标准和专业规范与时俱进的贡献潜力。基于此,本研究提出以下几条进路供学界进一步探讨和探寻。

**第一,进一步提高研究透明度以加强科学规范性、学术声誉、研究质量和知识可积累性。**学科公认和内化的学术标准不由任一领域来制定,而是通过不同领域的实践和交流而逐渐形成。每种类型的研究通过提高自身的规范性和质量,来为整个学科理念规范的形成作出贡献。透明度是定量研究的核心要求、特色和优势,国际关系定量研究应在这一方面继续努力,为学科作出更大的贡献。以下是一些可行的建议。

首先是建立数据和代码公开机制,将其作为期刊发表定量研究文章的要求。当前国内有一些学者在个人网站或国外公开数据平台上公布其国关论文的复现资料,但这些还属于个人自愿行为,资料的分散性也造成了国内检索查询的不便,尚未形成规范、合力和整合的资源库。中文期刊可以考虑

要求定量研究论文一经采用后将复现资料公布在国内公开学术数据平台上。国内已有这样的公开平台可以使用,如北京大学开放研究数据平台是一个可行的选项。<sup>①</sup>此平台是哈佛大学 Dataverse 的中国版,兼具国际平台的标准化和国内平台的数据安全性,创建账号、数据空间和下载数据等对校内外人员均免费。期刊在上面建立自己的数据空间,供在本期刊发表论文的复现资料上传,既可提高期刊的影响力,也方便使用者检索。统一平台对上传资料的规范格式要求能够避免不同期刊各自设定标准带来的问题。

其次是要加大在教学和培训中对中文期刊发表的定量研究论文的复刻和批判训练。一些学校已经将研究复刻形式引入定量研究教学中,但由于目前可得国内研究复刻资料较少,能够给教师和学生提供的选择范围太小,大多数复刻研究选择来自英文期刊,这不利于对我国国际关系定量研究在透明度和质量上进行监督和促进。在第一步实行之后,中国国关定量研究的复刻资料将会很快的积累,在这样的情况下就能够增加对中文期刊论文的复刻,更好地平衡对英文期刊论文的复刻,对于推进中国国际关系定量研究的行业规范和学术进步具有重要而长远的影响。

最后是要充分发挥在线附录的作用。期刊评审和发表定量研究时,可以通过在线附录的形式克服由于版面限制而造成的透明度不足问题。由于期刊版面有限,一般会对投稿稿件做字数规定,正式发表中更是有严格的篇幅限制。规定的篇幅一般远不能保证定量研究在数据来源、分析方法、实证结果报告、稳健性检验等方面的报告内容达到足够透明度。一种在其他学科和地区的长期实践已经证明行之有效的解决方法是,大幅度精简文章正文中公式图表和稳健性检验等内容,把技术性和信息性的报告尽可能放到附录。将附录与正文一起送审,供评审人对论文材料、分析和结论可靠性进行判断。在论文发表时,附录与复刻资料一起上载到公开数据平台而不占期刊版面,这样不但提高了研究透明度、节约了期刊版面资源,而且有利于提高定量研究文章的可读性和突出研究的实质性发现及贡献。

**第二,对方法和技术祛魅,以实质性贡献来评判研究价值。值得一再强**

---

<sup>①</sup> 北京大学开放研究数据平台官方网址为 <https://opendata.pku.edu.cn/>,可供期刊免费创立自己的数据空间,详情见网站。

调的是,“定量研究”不同于“定量方法研究”,其差异性不在于定量方法的复杂程度或先进程度。定量研究选择和使用定量方法的目的是回答国际关系问题,而定量方法研究则旨在解决方法领域中的理论和技术问题。混淆这两种不同类型的研究会使得定量研究的评判标准产生偏差和混乱。若研究人员试图通过方法的推陈出新来替代或补偿实质性贡献的不足,则会造成对方法和技术作用的不实夸大甚至盲目崇拜。我国国际定量研究应进一步摆正定量方法在研究中的工具性位置,根据特定国际关系研究问题来评判什么是好的(即恰当)方法选择和应用,根据方法和数据对国际关系重要问题的回答来判断研究价值。换言之,中国国际关系定量研究要进行方法祛魅,既不要过度热衷于方法本身的求新求异,也要避免在文风上就方法对研究的作用和贡献夸大其词。

当前中国国际关系定量研究中出现了过度追求方法和技术的新颖性、前沿性、先进性的倾向,并以此显示研究水平和论文价值。且不论这一文风带来了不良观感,这些关于方法和技术的声称往往不符合事实。研究方法是一个日新月异的领域,且包含处理各种类型问题的众多方法,并没有所谓“最前沿”方法一说。此外,在方法研究领域新发展出来的方法用于实质性研究会有一段合理的时滞,具体的国际研究也不可能采用“最前沿”和“最新”的方法。第一个原因是,研究人员应根据应用领域和研究问题选择方法,最新的方法和最恰当的方法之间没有必然联系,不应该去盲目追逐新鲜出炉的方法。第二个原因是,新发展出来的方法往往还没有经过足够的应用检验,应用领域的研究者对新方法应持谨慎态度而不是一拥而上。定量研究文章应该严谨求实地论证方法选择的恰当性,而不是过分强调新颖性。

**第三,定量研究者在专业训练和学术修养方面不能只注重方法的迭代更新,更要着重加强理论思维和定性分析能力。**这一倡议似乎违背了“术业有专攻”的原则。然而,定量研究并非一种“专攻”研究,而是在国际关系研究中使用了定量方法在目前会被人们视为一种研究类型。之所以在一项研究中使用某种定量方法,是因为根据研究者的判断,这种方法对这个研究而言是最恰当的,最能够帮助研究产生贡献。也就是说,使用定量方法是视情况而定的选择,而不是一以贯之的“专业忠诚”。一种做法可称为“选择”,其前提是研究者充分知晓回答这个国际关系问题需要什么、不需要什么,且既

掌握选中的方法、也掌握没有选中的方法,否则就不能称之为“选择”。定量研究者要提高对定量方法进行选择的水平,就不能局限于对定量方法本身的学习,功夫还要下在“诗外”。

诚然,学术界有不少优秀的定量研究者会频繁地使用某类方法,看上去是对这类方法有所“专攻”所致。但事实上因果往往是倒过来的——研究者聚焦于某类政治学或国际关系问题,研究这类问题最可行的是某类方法,于是其经常性地使用这种方法。比如,政治心理学研究者通常会依赖实验方法,不是因为他们“专攻”了实验方法,而是因为实验方法是这类研究可靠且可行的方法。当然,长期的积累也可能让政治心理学者逐渐成为实验方法方面的专家,在方法上也能够作出足够的创新性贡献而成为有所建树的方法学者。但如若这些学者想要在政治心理研究方面作出持久的贡献,成为实验方法的专家是远远不够的,还需要了解更多的方法,比如根据实验室产生的文本、图像和声音等分析心理和情感,以便在面对特定的研究问题时选择是否采用实验方法或其他方法。与此同时,这些学者也需要对政治心理的概念和理论进行深入思考和持续更新,才能够设计出新的实验来回答重要问题。

学术期刊在定量研究者学养的推动上需要发挥更为积极的作用。建议期刊在对定量研究稿件送审时,邀请至少一位文章实质性问题领域中的专家在无需读懂技术细节的情况下给出关于研究贡献的评审意见。这种做法不仅有利于排除技术要素对实质性贡献判断的干扰,也有利于消除人为“技术门槛”在国际关系学者之间造成的不必要和不健康的分野。实际上,对一项国际关系定量研究的意义、价值和贡献等方面进行专业判断并不需要看懂技术细节。比如,一项定量研究的任务是测量某个国际关系学中的概念,但该研究对国际关系学界对此概念的理解、理论应用场景、现实显现场景、背景性和操作性概念、学界关于概念的共识部分和争议部分、与临近概念之间的关系等都缺乏足够的了解,这个研究无论使用了多大的数据量、进行了多么复杂的建模和计算,也是一项毫无价值的研究,而其他国际关系学者不需要读懂那些技术细节就能在上述方面进行专业判断。换言之,加强定量研究的同行质疑和同行评价中的同行作用,适用于国际关系学者而非仅限于定量研究学者。

第四,大力加强测度研究的比重,积累可重复使用、可积累的原创性研究数据。数据是实证分析的前提、基础和要义。数据的可得性、质量、创新性从源头上决定了实证发现的可靠性和启发性。生成数据并不是研究中的额外或附带工作,而是一项严肃而重要的研究工作,即测度类型研究。测度是联系理论和实证的纽带,在社会科学中,测度的对象不仅是现象,还包括理论概念。对重要现象(如战争、国际援助、国际条约等等)的测量对于推进国际关系议题研究十分关键,对重要概念(如政治极化、国家声誉、国际权威等)的测量直接影响理论的发展和检验。测量本身就是一类集显示观察、问题思考、概念探索为一体的重要研究。但近十年来中国国际关系定量研究中测量类型研究过少,这不仅影响了数据创新和积累,也不利于议题开拓和理论创新。

近十年来我国国际关系定量研究的发展与对于数据的重视相辅相成。尤其在近几年大数据潮流的影响下,我国国际关系学界对数据库建设表现出高涨的热情并进行快速增长的投入,产出了一些数据库成果。但这个潮流也显现出了测度研究不足带来的缺陷和误区,以工程建设的形式来开展对数据的搜索和编码,缺乏对现象、概念、理论的细致观察和深入思考。由于并非建立在累积性的测度研究基础上,数据库建设不是由一个又一个具体研究议程进行内在推动,这就导致了建成的数据库实用性不足,难以运用到严格的学术研究中。我们提倡和鼓励以“分散建立—逐渐整合—不断积累—动态修正”的路径来形成中国国际关系定量数据资源,即研究者根据研究实际需要来创建对某一类现象和某个概念的数据集,这些数据集在学术界流通(如复刻数据的公开),并根据相似的研究议程进行汇聚和整合,在研究议程推进中得到时空维度和精细程度上的扩展,在不断使用中实现对其概念化、编码和其他偏差与错误的动态公开纠正。

此外,提高数据质量也需要依赖测度型研究的成熟和发展。测度作为一类研究,有其特定的执行程序 and 检测标准,并不是一个简单的量化过程。只有以学术研究的标准来要求数据生成,才能够给予数据质量起码的保证,达到学术研究所要求的数据使用标准。而从源头上保证数据质量,也有利于进行方法祛魅。实证方法(包含定量和定性)是对真实世界的信息(数据)进行提取和分析。在研究方法领域中有一条众所周知的规律,那就是数据

越干净,需要用到的方法越简单。方法用以保证数据生成的干净(如精心设计、谨慎执行的实验/调查问卷就是为了避免数据被有意图或无意图地污染)或者处理有问题但还有救的数据(如特定的选择偏差,某些情况下的缺失数据,或某类测量误差等)。但每一种方法都只能够处理一种或有限几种数据问题,有些问题是无论什么方法都无法挽救的,也不存在同时对所有重要的数据问题进行一揽子处理的方法。提升定量研究质量和创新的关键在于数据,而不是寄希望于方法。

**第五,加强定量方法课程体系建设,提供数据思维和方法原理的体系性训练。**学术研究是专业性活动和工作,要求严格的学术训练和长期的学术积累,并不存在“短平快”路径。将定量方法等同于编程技术的认识如今并不少见,无论是提供培训一方或是接受培训一方,都不乏有人真心认为定量方法可以速成。诚然,国际关系定量研究以议题决定方法,在方法选择上并不需要遵循体系性逻辑。但对方法恰当性进行判断和正确操作,却要求研究者具有体系性的方法训练。

定量方法本身是一套严格的知识体系,最简单的方法或模型正是最基本和最重要的科学原理得到的明显体现。这些方法或模型可能在技术上不值一提,毫无基础的人都可以在几分钟之内学会如何把数据放入软件程序,把结果做成表格或图形,将参数估计的解读背诵下来。但这些基础方法背后的科学原理和逻辑却通常需要一个学期课程的系统专业讲解,掌握这些原理和逻辑对于进阶学习各种方法的复杂变换十分必要,否则方法的世界将会呈现出万花筒似的碎片化的纷乱,让人除了追逐潮流外无所适从。体系化的定量方法教学不仅对培养定量方法人才有必要,而且对把定量方法妥善地放入未来研究者的工具箱中至关重要。

**第六,增强研究成果的国际交流和传播,提高中国国际关系定量研究的国际知名度和影响力。**中国国际关系定量研究在近十年来产生了不少优秀的成果,但这些成果在国际学术界中的传播力和影响力目前还非常有限,国际学术界对我国国际关系定量研究的学术现状和发展动态知之甚少。即使是具有国际知名度和影响力的学者,他们在中文期刊上发表的国际关系定量研究论文也鲜为国际同行所知,其国际学术可见度和贡献力基本来自英文期刊上发表的文章以及国际学术活动。

定量方法是一种跨越语言和文化背景的方法,使用这种方法的研究本应该在国际传播中具有一定的优势。但这种优势目前尚未发挥出来,限制了我国国际关系学的国际影响力发挥,也不利于引入国际同行的学术质疑、交流和合作以加速中国国际关系定量研究的发展。此外,由于国际传播渠道不畅,一些中国学者为获得更为广泛的学术影响而选择将自己更为看重的研究成果发表在国际期刊上,也让中国学术期刊错过不少优秀的研究成果。

造成这种现象的最主要原因在于国际传播不足、渠道不畅、力度不够。值得一提的是,一些国内期刊组织专门的翻译人员将文章译为英文进行传播,这种努力值得赞赏。然而,翻译人员常常不掌握英文专业论文的写作风格和表达习惯,英文译本在语法上可能无懈可击,但在专业用词和表达习惯上却让整个论文读起来有外行之感。学术的国际传播说到底还是学者之间思想的传播和交流,所以最好的译者是作者本人,最好的做法不是翻译而是用英文重写。但这样的工作对于研究者本人来说并不具有现实吸引力,不但因为这些工作不计入绩效考核,而且占用研究者稀缺的时间资源会使新成果延迟出现。创办英文期刊也无法有效解决这个问题。本研究将国际传播力和国际影响力作为一个开放性的问题提出,希望学界集思广益,找到有效的途径。