

评估联合国大会表决数据： 应用、局限及建议^{*}

迟 永

【内容提要】 联大表决数据正被频繁用来衡量国家间外交立场相近程度等相关变量,分析大国外交关系等研究问题。由于能够处理国家异质性、从多议题角度提供全球不同国家的外交偏好与利益进而识别大国外交影响力,联大表决数据正成为研究大国关系的有效工具。基于联大表决数据生成的理想点指标更是迅速得到学术界认可,相关国内外研究成果也大量涌现。然而,既有联大表决数据及其相关数据指标并非完美无缺,仍存在以下局限:缺少协商一致的表决数据;对缺席现象的关注不足;难以克服议程变化的影响;表决数据的信息有限;单一维度的不充分问题。为避免数据估计结果与相关推论出现偏差,现阶段实际应用联大表决数据及其相关数据指标应注意以下事项:对理想点进行必要的技术处理;妥善处理弃权票记录;恰当筛选表决数据;探索新类型的联大数据集。

【关键词】 联合国大会 联大表决数据 外交立场相近程度 理想点

【作者简介】 迟永,中国政法大学全球化与全球问题研究所副教授、硕士研究生导师。

电子邮箱:chiyong1216@cupl.edu.cn

^{*} 本文系教育部基金项目“美国海外利益保护及对中国的启示研究”(项目编号:18YJCGJW003)、中国政法大学科研创新项目资助(项目编号:20ZFQ81002)和中央高校基本科研业务费专项资金资助的阶段性成果。感谢《国际政治科学》匿名审稿人和编辑部提出的宝贵修改建议。文中疏漏由笔者负责。

分析并比较国家间外交立场的相近程度一直是国际关系学科的重点研究问题,但以直观且准确的方式解答这一问题会面临诸多困难,不仅各个国家外交政策的关注点存在诸多差异,彼此间的利益关系也是异常复杂,比较并区别各种全球议题中国家间立场的相近程度或各自的外交影响力更是十分困难。当前较为可靠的实现路径是观察各种国际组织平台中的成员国行为,进而解读这些国家外交立场的相近程度。在众多国际组织中,联合国大会(以下简称联大)具备的表决机制与表决记录优势明显。尽管联大不具备安理会的强制力与执行力,但却是全球社会的缩影。联大几乎囊括全球所有主权国家,并会在相对固定的时间与地点,针对各种国际热点议题进行表决。联大的表决记录也逐渐被看作是全球各个主权国家对诸多国际议题政策与偏好的集中体现,并可以利用多种方式与方法进行分析比较。近年来,基于投票记录形成的联大表决数据已经成为研究或衡量国家间外交立场相近程度的主要方式,相关国内外研究成果也在大量涌现。

然而,既有联大表决数据及相关数据指标是否完美无缺?其能否有效反映出国家的外交政策立场?这些问题仍值得深入讨论。本文重点考察联大表决的原始数据及基于原始数据计算得出并得到广泛应用的理想点指标,发现既有联大表决数据及相关数据指标很容易造成“一叶障目”的问题,即衡量国家间外交立场相近程度等类似变量时效度不足,不能达到“洞若观火”的程度。联大表决数据存在继续开发完善的空间,要注意并处理一些技术细节方能更好地进行使用。发掘联大表决数据无疑会丰富学术界对联大的认识与理解,进一步充实联合国与国际组织领域的相关研究。

一、联大表决数据何以成为研究工具?

联大是依照《联合国宪章》(以下简称《宪章》)于1945年与联合国同步成立的主要分支机构,负责联合国的重要议事与决策事务,具备非常丰富的职能,包括表决联合国经费审查核准、新成员国的加入审议等重要问题。正如2000年的《联合国千年宣言》所强调,联大是当前联合国主要的议事、决策和代表机构,居于联合国核心地位。^①

^① 《联合国千年宣言》,2000年9月8日,联合国正式文件系统,文号:A/RES/55/2,第8页。

联大的表决机制沿袭国际联盟的国联大会制度,设立一国一票的表决制度,不区分国家大小与强弱,不设大国否决权,重点体现主权国家一律平等的理念,给予所有成员国相同的投票权利。联大表决机制及表决产生的投票数据使其迅速成为被研究的对象。诸多学者更是将其作为研究工具加以拓展,尝试衡量国家间外交立场相近程度等类似变量,并用来解释国际关系中的一系列重要理论与经验问题。联大表决数据之所以得到上述应用是因其具备以下理论与事实的基础条件。

其一是能解决国家外交关系的异质性问题。准确衡量国家间外交立场的相近程度必然要纳入尽可能多的主权国家,并观察这些国家就某些特定议题的看法与偏好,从而处理不同国家外交关系之间的异质性问题,实现测量国家间彼此外交关系远近的研究目标。在这方面,传统上利用国家间外交关系事件数据的方法存在一定局限,即无法实现不同国家之间的衡量比较。比如,分析特定时期内美国出现的外交事件确实可以观察到美国在这一时期同别国外交关系的发展变化,但这种方法仅限于分析美国外交,无法进行大国间外交影响力的比较。不同国家的外交关系之间存在极大的异质性,这对有效衡量国家间外交立场的相近程度造成较大阻碍,而联大表决机制则可以高效处理这类异质性问题。所有成员国在联大都面临较为一致的战略决策环境,并基于此环境表达出各自国家的外交立场偏好,使得全面比较各自国家间外交关系的亲疏远近成为可能。

其二是涵盖广泛的全球议题。除异质性问题外,衡量国家间外交立场相近程度还需避免研究视角与结论过于狭窄的问题,不能仅依赖少数议题。比如,虽然军事同盟常常被用来衡量国家间外交关系的亲密度,但该指标集中在高级政治领域。更为广泛的全球公共政策偏好高度相似的国家可能并没有结成军事同盟,仅通过军事同盟这一指标必然会忽视这类关键信息。而《宪章》规定联大会议可以讨论《宪章》涵盖范围内的所有国际事务,使得联大表决数据涵盖的议题非常广泛。

首先,联大下属的六大主要委员会鲜明体现出联大议题的广泛特征。六大主要委员更多在分担联大常规工作,涵盖裁军、全球发展、人道主义、去殖民化和国际法等诸多全球事务,所涉议题虽广泛但投票表决的争议性却不强。其次,联大还会基于重要的全球治理议题与突发的国际冲突事件组

织召开特别会议与紧急特别会议。这二者所涉议题的紧迫程度与表决投票阶段呈现的争议度明显更高,比如,自2022年2月以来反复召开的第十一届紧急特别会议围绕的议题是同时期爆发的俄乌冲突,该会议过程凸显了成员国之间立场的严重分歧。^①最后,联大在常会阶段也会直接审议诸多议题,包括全球反恐、国际合作、可持续发展、环境保护、资金筹措和经济制裁等广泛内容。这类议题的大部分内容关键但争议性却不强,比如,联大在审议《联合国全球反恐战略》草案时,通常会以全体一致或高比例赞成票的方式表决通过。然而,联大常会阶段还是会出现一定数量的争议性议题。比如,2022年12月30日,第七十七届联大讨论包括东耶路撒冷在内的巴勒斯坦被占领土问题,虽然相关决议最后被表决通过,但反对票与弃权票的总数接近赞成票数,表明关于这一问题成员国间存在较大分歧。^②

基于广泛全球议题形成的联大表决记录可以提供各个国家对几乎全方位的全球问题的立场与态度,从多个视角诠释哪些国家在何种议题上保持一致或存在分歧,多层次诠释出各个国家的外交偏好,从而极大提高衡量国家间外交立场相近程度等相关变量的效度。

其三是呈现国家外交偏好的发展趋势。恰当衡量国家间外交立场相近程度的研究工具还应具备时间维度上的优势,这可以帮助研究人员发现短期或长期国家外交政策偏好的发展变化、理解大国外交政策乃至全球政治的走势。以往的研究工具在这一方面都远逊于联大表决数据。截至2023年11月联合国官网的介绍内容,联合国已经拥有193个成员国。^③自设立至2023年,联大已召开过78届常会。每个成员国都应在联大召开之际,派出本国代表参加。经过近80年的运行发展,联大长时间累积的几乎囊括全球

① 联合国大会:《大会第十一届紧急特别会议》,来源:<https://www.un.org/zh/ga/sessions/emergency11th.shtml>,访问时间:2023年11月1日。

② United Nations Digital Library, “Israeli Practices Affecting the Human Rights of the Palestinian People in the Occupied Palestinian Territory, Including East Jerusalem: Resolution/Adopted by the General Assembly,” *United Nations Digital Library*, December 30 2022, <https://digitallibrary.un.org/record/3999158?ln=en>, 访问时间:2023年11月1日。

③ 联合国:《会员国的增长》,来源:<https://www.un.org/zh/about-us/growth-in-un-membership>,访问时间:2023年11月1日。

所有主权国家的详细表决记录已经成为分析研究国家间外交关系与大国外交影响力等问题的“宝藏”。同时,联大表决数据可以体现出各个国家在特定议题上的不同投票模式与走势,帮助我们了解各个国家对这些议题的立场与偏好选择。此外,当国家在联大某些议题的投票选择发生重大变化时,联大表决数据更可以提醒研究者这一国家的外交政策及其背后的国内政治正在或已经发生重大变革,导致其外交政策的立场与偏好也随之产生变化。

其四是评估大国外交影响力。联大表决数据还可以用于分析大国外交策略的有效性,进而评估大国外交的国际影响力。联大表决数据可以提供各种全球公共政策的发展状况与成员国之间的合作或分歧程度。一旦研究人员利用某种研究设计方法确定某个大国在这些合作或分歧上的作用,就可以确定并且量化这个大国在这些议题及其表决上的国际影响力。比如,大国常常可以通过幕后谈判、游说、其他外交努力以及场外的政治经济因素来促使别国与本国在联大选择一致投票,此时较高的投票一致度明显可以在一定程度上说明该大国的外交影响力。相反,一旦某个大国在联大某个议题的表决中处于少数状态,一定程度上表明该国在这一议题上出现被孤立的现象。虽然联大表决数据可以提供评估大国外交影响力的有效信息,但识别这类信息必须经过一定的数据技术处理,不能将联大投票一致现象简单解读或等同于大国的外交影响力。

总之,联大可以成为他者观察世界政治的“舞台”。研究人员可以在联大表决活动中观察对比几乎全球所有国家在诸多类型全球问题上的外交政策偏好。对联大表决众议题、多国家、长时段的技术性分析也有助于探索国家外交政策偏好的变化情况与国家间外交立场的相近程度。

诚然,衡量国家间外交政策相似程度的研究工具并非仅有联大表决数据,如 τ_b 系数、相似性分值(S-scores)等也曾经得到过学术界的广泛关注与使用。^① 特别是相似性分值更是得到经验验证,一度成为衡量国家间外交关

^① Bruce Bueno De Mesquita, “Measuring Systemic Polarity,” *Journal of Conflict Resolution*, Vol. 19, No. 2, 1975, pp. 187-216; Curtis S. Signorino and Jeffrey M. Ritter, “Tau-B or Not Tau-B: Measuring the Similarity of Foreign Policy Positions,” *International Studies Quarterly*, Vol. 43, No. 1, 1999, pp. 115-144.

系远近的主要方式。^① 然而,无论是 τ_b 系数还是相似性分值都有难以克服的局限,且更注重国家间的军事联盟关系,忽视了国家间非军事联盟上的联系。^② 联大表决数据则在涵盖国家数量与国际议题种类上具备更明显的优势,既适用于高级政治领域也适用于低级政治领域的研究问题。联大表决数据逐步发展成熟后,迅速成为衡量国家间外交政策相似程度等相关变量的主要研究工具。相反, τ_b 系数与相似性分值等数据指标的使用频率大幅下降。只有遇到少数特殊的研究领域或问题,如分析联合国成立之前的国家外交政策相似程度时,二者才是更为合适的研究工具。

此外,联大表决数据的有效性也存在争议。最为主要的是联大表决行为背后是否蕴含足够的政治意义。联大本身并不是国际影响力与执行力较强的国际制度,决议一般也不具备强制力度,更多是号召与规范意义上的内容。^③ 联大的投票表决也很难具备足够的政治意义,不能明显体现国家实质性的外交偏好。然而,联大并非无关紧要的国际制度,其仍具有独特的全球影响力。正是由于代表的广泛性,联大通过的决议一般带有世界“舆论”的特征,国际道义上的影响力也非常显著。冷战终结后国际格局的重要特征就是中小国家的国际地位和影响力不断上升,这种全球政治格局必然要求一个更为广泛、更具兼容性的全球治理机制作为国际合作平台。联大的表决机制和国际合法性完全满足上述要求。即使诸多联大表决及产生的决议并不具备强制执行手段,但其仍与大多数国家重要的外交战略利益密切相关。每一届联大都会出现争议性极强的议题,诸多国家特别是大国不会对这类问题的投票表决漠不关心,反而是调动各种政治资源、积极引导投票表决更倾向于本国的偏好或利益。比如,2022年俄乌两国爆发冲突以来,联大成员国围绕这一议题展开激烈的讨论,相关决议的表决过程更是充满大国

① D. Scott Bennett and Matthew C. Rupert, “Comparing Measures of Political Similarity: An Empirical Comparison of S versus τ_b in the Study of International Conflict,” *Journal of Conflict Resolution*, Vol. 47, No. 3, 2003, pp. 367-393.

② Frank M. Häge, “Choice or Circumstance? Adjusting Measures of Foreign Policy Similarity for Chance Agreement,” *Political Analysis*, Vol. 19, No. 3, 2011, pp. 287-305.

③ M. J. Peterson, *The UN General Assembly*, Milton Park, Abingdon: Routledge, 2006, pp. 125-130.

间的争议与分歧。^① 基于此,当前学术界的主流观点更为认可联大表决数据的价值,争议点聚焦在如何更为有效地使用联大表决数据。

二、联大表决数据的实际应用

联大表决数据具备的特殊性质促成其不断被学界实际应用。当前使用范围最广、影响力也最大的联大表决数据集是艾利克·沃伊特恩(Erik Voeten)、安东·斯特列日涅夫(Anton Strezhnev)和迈克尔·贝里(Michael Bailey)三人共同完成的联合国大会表决数据(United Nations General Assembly Voting Data)。这一数据集于2009年公开发布第一版,截至2023年11月已经更新到第31版,包含联大1946到2022年间各届会议的唱名表决记录。^② 除联合国官方提供的表决记录外,联合国大会表决数据还囊括其他学者与研究机构编纂的联大表决记录,极大地提高了自身的丰富与准确程度。

然而,无论是直接使用联大表决数据还是基于其技术提炼出的各种数字指标都无法避免联大表决中的议题变化问题。由于可以讨论《宪章》涵盖范围内所有国际事务,联大会议议程经常会出现当时国际社会的热点问题,但热点问题往往具备时效特征,也往往会产生多次热点相关的会议辩论与表决。一旦热点问题得到解决或影响淡化,联大很可能减少对这一问题的关注与讨论,甚至不再将其纳入会议议程。联大这一特性对于忽视会议议程变化的表决数据及相关研究成果会造成致命影响,即国家间联大表决投票的相似程度会受联大议程变化的强烈影响,其数值变动情况可能非常显著,但实际上这期间国家间的外交立场相近程度可能没有发生相应幅度的变化。假如若干国家在联大关注某一热点问题期间在除这一热点问题之外议题的投票选择上没有发生变化,基于这一热点问题衍生出的诸多轮联大

^① Mohammad Reza Farzanegana and Hassan F. Gholipour, "Russia's Invasion of Ukraine and Votes in Favor of Russia in the UN General Assembly," *International Interactions*, Vol. 49, No. 3, 2023, pp. 454-470.

^② Erik Voeten, Anton Strezhnev and Michael A. Bailey, "United Nations General Assembly Voting Data," 2009, <https://doi.org/10.7910/DVN/LEJUQZ>, Harvard Dataverse, V31, 2023 version, 访问时间:2023年11月1日。

投票无疑会扭曲国家间外交立场相近程度的实际数值,进而使得后续数据估计结果出现严重偏差。

为了克服现有联大表决数据中的议程变化问题,迈克尔·贝里、安东·斯特列日涅夫和艾利克·沃伊特恩三位学者基于联大表决数据,结合空间模型(spatial models)分析与项目反应理论(Item Response Theory),设计开发出一种新的动态理想点(ideal point)数字指标,用来估计1946年至2012年参与联大表决的各个国家的外交政策偏好连续性与动态变化情况。^①理想点指标利用每一届联大都会投票表决的相同或相似决议作为跨时间的衡量尺度,并假定这些作为“桥梁”工具的决议草案及相关表决在较长的时间范围内固定不变。一旦有国家在这些表决中变更了原有的投票选择,研究人员就可以合理推断该国的外交政策偏好也在发生变化。以此为基础得出的理想点指标自然也会发生相应变化,且可以进行纵向的时空比较,这在一定程度上克服了联大议程变化对国家间外交立场相近程度的扭曲性影响。

理想点指标出现后迅速得到学术界的普遍认可,几乎成为最成功的联大表决数据。贝里等三名学者撰写发表的有关联大表决数据以及理想点指标论文也成为这一领域的权威成果。近年来几乎所有涉及联大表决数据的研究成果都会引用这篇论文,其引用数据一定程度上可以反映联大表决数据的实际应用情况。本文利用学术搜索引擎得出2016—2023年这篇论文的引用情况;虽然文章正式出版于2017年,但网络版本在2016年就得到学术界关注与引用,因而本文对其索引情况的统计起始于2016年;由于2024年的索引数据并不完整,故本文统计时间节点止于2023年。具体索引数值如图1所示,横轴为年份,纵轴为当年索引文章数量。总体趋势上,这篇论文的引用频次大体呈现出上升趋势,至2023年引用量更是高达97篇。值得说明的是,图1只是呈现科学引文索引数据库(Web of Science)的引用情况,索引这篇论文但没有得到该数据库收录的文章更是难以计数,远高于图1所示。

为了更为鲜明体现出当前联大表决数据的应用情况,本文将挑选若干最近时间发表的实证类论文,剖析并比较这些文章运用联大表决数据的方

^① Michael A. Bailey, Anton Strezhnev and Erik Voeten, “Estimating Dynamic State Preferences from United Nations Voting Data,” *Journal of Conflict Resolution*, Vol. 61, No. 2, 2017, pp. 430-456.

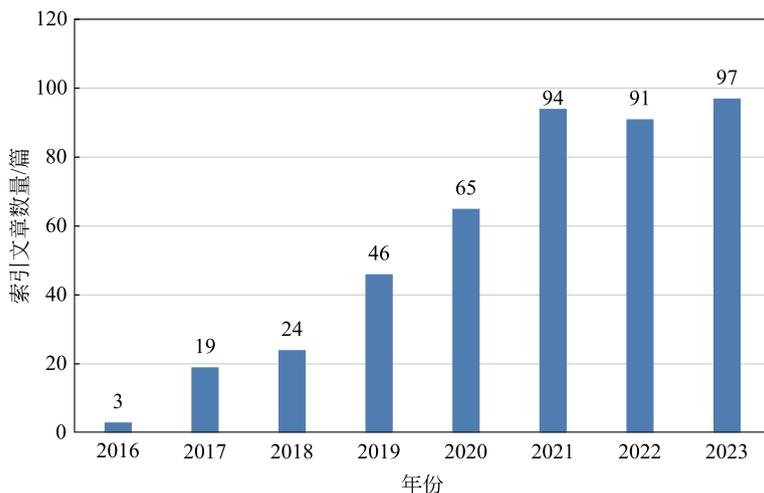


图1 “Estimating Dynamic State Preferences from United Nations Voting Data”
索引数据图(2016—2023年)

资料来源:科学引文索引数据库(Web of Science)。

式与相关的研究设计内容。依据既有研究的实际情况,本文将所示论文分为两类:一类是使用理想点指标的实证研究成果;另一类则是使用联大原始表决数据的实证研究成果。

相较于投票相似程度等指标,理想点指标无疑更科学也更具解释力度,同时也拥有直观、方便等优势,是当前众多研究人员用来衡量国家间外交立场相近程度等相关变量的主要方式。综合既有成果,理想点指标的实际运用方式可以归为以下三种。

其一,利用理想点指标衡量被解释变量。这类研究成果主要探讨何种因素的变化会影响国家在联大的投票表决行为。比如,戴维·莱克齐安(David Lektzian)与格伦·比柯莱泽(Glen Biglaiser)撰文分析美国制裁与援助的行为如何影响援助对象国在联大投票,并利用1990年至2014年美国与123个发展中国家的联大投票面板数据进行估计。^① 该文的被解释变量是

^① David Lektzian and Glen Biglaiser, “Sanctions, Aid, and Voting Patterns in the United Nations General Assembly,” *International Interactions*, Vol. 49, No. 1, 2023, pp. 59-85.

美国和他国在联大表决投票的差异程度,并以理想点指标操作被解释变量,具体操作化为美国与他国平均理想点估值之间的绝对差值。但和贝里等学者的理想点指标不同,该文选择美国政府认定的联大重要投票作为计算理想点的材料。1985年以来,美国国务院每年向国会提交的《美国参与联合国》报告会附加名为《联合国表决情况》(Voting Practices in the United Nations)的文件,详细记录美国在联合国大会和安理会的表决情况,并着重强调美国政府认为的事关美国国家利益的联大决议及其表决情况。美国政府往往视其为重要投票,并给予极大的关注,甚至会利用游说、威胁和援助等政治手段尽可能促使别国与美国投票一致或至少持弃权的中立态度。该文应用理想点的创新点是着重强调这些重要投票的作用,并采用和贝里等学者同样的计算方法,得出只基于重要投票的理想点数值,并利用新得出的理想点数值进行后续的计量估计。

再比如,漆海霞在其分析中国同其伙伴关系国家,特别是欧洲伙伴国家之间贸易关系对双方政治关系影响的文章中也使用理想点操作被解释变量。^①由于这些国家在联大参与表决的过程中大概率会受到中美关系及其他因素的复合影响,这篇文章则创新性地设计出两个被解释变量:一个是中国同20个拥有伙伴关系升级情况的国家在联大的投票差异情况;另一个是这些国家同美国在联大的投票差异情况。这种设计无疑会更加突出20个国家在中美关系影响下外交立场的选择情况。此外,文章应用范围内的联大表决数据很可能存在内生性与自相关问题,同时也会产生反向因果关系,即被解释变量也可能引发解释变量的变化。为了解决上述问题,该文选择广义矩估计法(generalized method of moments, GMM)处理1993至2014年间的联大表决面板数据,尽可能提高统计估计的有效程度。最后,该文还基于英国、波兰、法国和德国四个欧洲国家在联大的表决行为作出稳健性检验,进一步充实文章的结论内容。

其二,理想点也可以作为解释变量与控制变量或作为衡量二者的主要手段。这类成果重在发现并解释国家在联大的投票表决行为同其研究对象

^① Qi Haixia, "China's Partners or US Allies: the Dual Status of Major European States and their Voting Behaviour in the UNGA," *Asia Europe Journal*, Vol. 21, No. 2, 2023, pp. 225-250.

之间的联系与规律。比如,易正容与孙楚仁2022年探讨中国与潜在国家以外的国家的政治关系空间溢出效应对中国从潜在国家进口大宗商品的影响。^① 该文重要的解释变量——国家间政治关系——使用国家间理想点的差值作为主要衡量方式。该文对理想点的技术处理是利用1996—2014年196个国家之间的理想点的差值计算出彼此的年度双边关系矩阵,再将其与中国同他国双边政治关系的数值相乘,得出文章所需的中国与他国双边政治关系空间溢出效应数值。实证结果表明中国与他国双边政治关系的空间溢出效应确实会对中国的大宗商品进口产生影响。此外,理想点指标更多被当作研究中的控制变量,帮助研究人员控制国家外交政策相似程度等类似变量对计量估计模型的影响。绝大多数涉及联大表决数据的研究成果都可以归到这一类别,但由于重要程度相对较低,这类研究成果很少对这类控制变量进行更为复杂精妙的设计,往往对理想点进行简单操作后直接应用到文章数据中。囿于篇幅,本文不再赘述介绍此类研究成果。

其三,基于理想点研究国家间的凝聚力问题。分析国家间或政府间国际组织成员国之间在对外政策上的凝聚力问题一直是联大表决数据主要的应用目标,即观察相关国家在联大某个特定时期或特定议题上的表决行为,推导这些国家在外交政策偏好方面的相近程度。美国学者玛格丽特·鲍尔(M. Margaret Ball)早在1951年就撰文分析联大集团式投票(bloc voting)与集团凝聚力的问题,指出联大成员国之间会因地区、意识形态和共同利益等原因组成国家集团并在联大会议中采取较为一致的表决行为。^② 在理想点指标出现后,这类研究也正在探索使用理想点分析这类问题的方式与方法。同上述两类研究不同的是,分析国家间凝聚力问题并不强调解释变量与被解释变量的关系,更重视研究对象所涉国家对外政策的差异以及如何用理想点指标有效诠释这种差异等相关问题。

这类研究成果的代表性研究是分析欧盟或欧洲国家的凝聚力,相关领

^① Zhengrong Yi and Churen Sun, "Bilateral Political Relations, Space Spillovers and the Import Expansion of China's Bulk Commodity," *Journal of Asian Economics*, Vol. 81, 2022, pp. 1-19.

^② M. Margaret Ball, "Bloc Voting in the General Assembly," *International Organization*, Vol. 5, No. 1, 1951, pp. 3-31.

域学者正在积极尝试利用理想点深化丰富这一研究主题。比如,阿里·居内什(Ali Gunes)与穆拉特·奥兹卡莱利(F. Murat Ozkaleli)2023年的发文。该文在详细梳理既有研究的基础上,认为使用静态数据或短时间段的数据分析欧盟凝聚力问题往往会出现偏差,但利用跨时期统计分析法估计欧盟国家联大表决的理想点可以得出更为准确的实证结果。^① 该文继而利用理想点指标构建三个变量:“联盟凝聚力”(union cohesion),指欧盟所有成员国在联大表决中投票一致的数值;“成员国偏离程度”(member deviation),指欧盟成员国在联大表决中出现分歧程度的数值;“非成员国接近程度”(non-member adhesion),指非欧盟成员的域外国家同欧盟在联大表决中投票一致程度的数值。该文认为上述指标可以有效捕捉欧盟发展过程中的重要节点事件中主要成员国的外交立场变化,并发现欧盟主要的域外国家同欧盟外交政策相似的不同程度。

除了分析欧洲国家或欧盟,一系列诠释中国外交影响力的研究也可以归纳为凝聚力问题研究。虽然这类研究关注中国同别国在联大表决中的相近程度,但其本质仍是在处理国家间的凝聚力问题,并得到国内外学者的广泛讨论。比如,理查德·图萨尼(Richard Q. Turcsanyi)等2023年利用理想点指标探究中国对16个中东欧地区国家的外交影响力问题。^② 为了尽可能全面比较16个中东欧国家的外交偏好变化情况,这篇文章在研究设计上整理出两组数据分别估计这些国家外交政策偏好的动态变化:第一组是16个中东欧国家各自的理想点数值变化情况;第二组则是这些国家同中国、美国和德国理想点数值的差异变化。文章还特别强调2008年后两组数字的变化情况。这种处理方式无疑更为鲜明直观,也可以更为准确测量16个中东欧国家在联大表决问题上的立场。此外,这篇文章还在使用理想点指标的基础上,增加针对联合国一般性辩论语料库的文本分析,使用复杂向量解析并

^① Ali Gunes and F. Murat Ozkaleli, "An Intertemporal Statistical Analysis to Ideal Point Estimates: EU Cohesion at the UNGA Revisited," *Comparative European Politics*, Vol. 21, 2023, pp. 333-355.

^② Richard Q. Turcsanyi, Kamil Liškutin and Michal Mochtak, "Diffusion of Influence? Detecting China's Footprint in Foreign Policies of Other Countries," *Chinese Political Science Review*, Vol. 8, No. 3, 2023, pp. 461-486.

比较 16 个中东欧国家代表团在历年联合国大会一般性辩论发表的演讲文本。多类型联合国数据的使用不仅大幅提高了该文的严谨程度,也为探索联大其他类型的研究数据作出积极尝试。

除了理想点指标外,当前学术界也在积极创新联大原始投票数据的使用方式。比如,《和平研究杂志》(*Journal of Peace Research*)于 2024 年第 2 期正式发表的一篇论文,专门采用 1994—2017 年间 1178 项有关军控问题的联大决议及各国的原始投票数据,分析各国对军控问题的立场变化情况。^①为了凸显各个国家针对军控问题的共识与分歧,这篇论文采用人工编码赋值的方式,直接处理上述军控问题决议及其相关的表决投票数据。如图 2 所示,作者先是将研究范围内的联大决议根据所涉内容分为四类,每一类涉及不同的军控议题。各个国家针对这些决议的投票也可以体现出对军控问题的不同立场,并删除无法鲜明体现出国家具体偏好的决议。

作者继而根据决议所属类别与各国的具体投票选择为所有参与投票的国家赋值,从而计算出各个国家对军控问题的立场得分。这种新颖的筛选方式不仅可以有效识别出上述联大决议涉及军控内容的幅度与类别,更可以利用这些决议的表决情况凸显国家间围绕军控问题的共识与分歧。这篇论文在技术处理联大表决原始数据方面具备极强的创新性,不仅有效避免使用联大原始表决数据容易出现的各种问题,更是将观察对象严格控制在研究范围之内,尽可能降低议程变化等争议问题。出于严谨,作者更是在附录中详细介绍了数据的处理方法、过程、有效程度等细节内容。^②

总体上看,近年来的研究成果几乎都以理想点与原始表决数据为主要研究工具,特别是前者的应用范围更为广泛,后者研究数量相对较少。同时,现有的联大表决数据信息丰富,涵盖时间、国家、议题、具体的投票选择等关键信息,方便研究人员从不同维度选择合适的量化模型进行数据估计。事实上,大部分现有成果会基于理想点设计出包括连续变量、二元变量、序数变量等多种类型的被解释变量,或基于联大原始表决数据将国家-年或单

① Tobias Risse, "External Threats and State Support for Arms Control," *Journal of Peace Research*, Vol. 61, No. 2, 2024, pp. 214-227.

② Tobias Risse, "External Threats and State Support for Arms Control," online appendix.

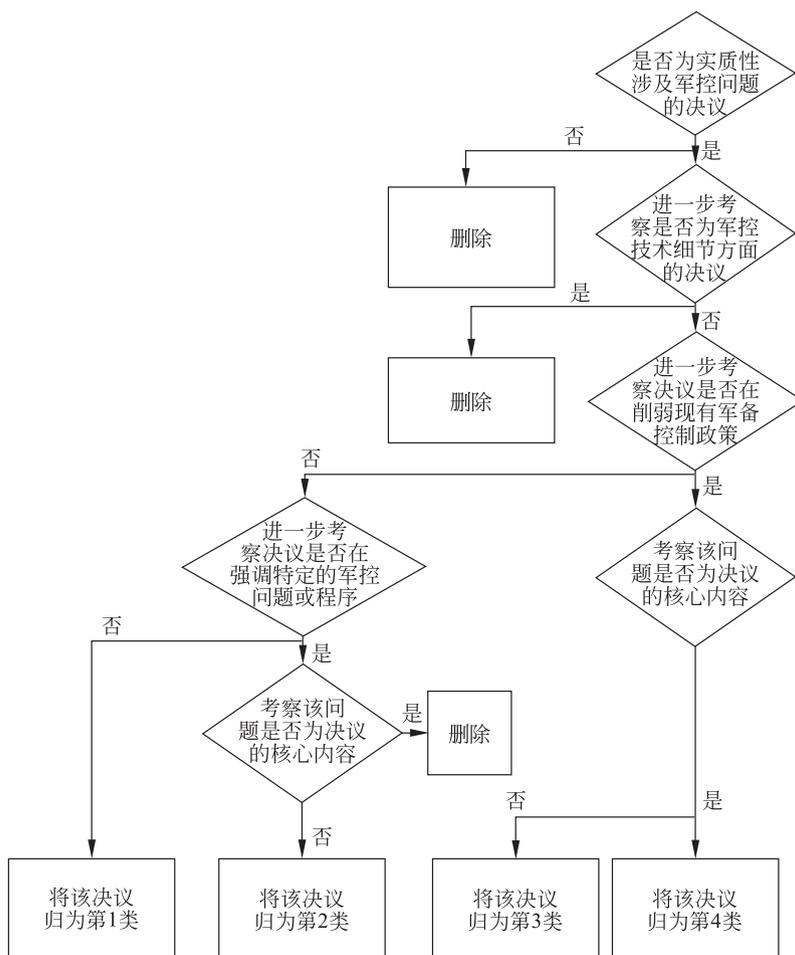


图2 《外部威胁和国家支持军控》一文处理联大军控决议的步骤图

资料来源：“External Threats and State Support for Arms Control”。^①

次投票作为分析单位构建截面数据、时间序列数据、面板数据等。无论采取何种类型的研究数据,既有研究成果普遍会对其进行一定程度的技术再处理,特别会基于实际研究需要精心筛选数据,尽可能提高研究设计与变量操作的可信程度。

^① Tobias Risse, “External Threats and State Support for Arms Control,” p. 218.

三、既有联大表决数据的局限

当前绝大部分研究成果更多将联大表决数据作为研究工具,但既有联大数据及相关数值指标远非完美无缺。即使已经得到部分学者的大幅改进,但联大表决数据仍存在一定局限,具备进一步完善的空间。必须强调的是,下述局限之处的前四点是现有联大原始表决数据与理想点指标二者共同有待解决的问题。第五点则是单独针对理想点指标,应用原始表决数据时几乎不会出现所涉及的单一维度问题。

第一,缺少协商一致的表决数据。

在实际表决过程中,联大主要采取两种形式,即协商一致(consensus)与唱名表决(roll-call vote)。联大的协商一致近似于全票通过,或以成员国代表团全体鼓掌的形式通过。设置协商一致表决方式的初衷就是促进各成员国达成普遍共识,至少是不反对或弃权,共同推动相关决议的实施。虽然协商一致只是联大在会议实践中逐步形成的表决形式,并没有得到《宪章》和《联大议事规则》的正式规定,但自冷战结束后,联大在实际中大力鼓励成员国采用协商一致的表决方式,并且以协商一致方式通过的决议数量也在不断增加。近年来的联大绝大多数决议都经由协商一致的表决形式通过,使用唱名表决通过的决议仅占一小部分。比如,2021—2022年召开的第76届联大会议公布了307项决议,其中229项是以未经表决的方式通过,而使用唱名表决通过的决议数量不足三分之一。根据彼得森(M. J. Peterson)的计算,在20世纪50年代后期,联大20%至30%的决议是以协商一致的方式通过的,这一比例在20世纪80年代中期达到50%至60%,到20世纪90年代初约为75%,近年来已经接近80%。^①以协商一致的方式通过决议无疑表明成员国已经对决议草案形成广泛共识,更能凸显联大塑造国际共识与决议合法性的能力。^②虽然协商一致的表决方式并非联大首创,但其在联大深

^① M. J. Peterson, "The UN General Assembly," in Thomas G. Weiss and Rorden Wilkinson, eds., *International Organization and Global Governance*, Milton Park, Abingdon: Routledge, 2018, p. 239.

^② M. J. Peterson, "The UN General Assembly," pp. 74-79.

化自身职能与功能上具备重要意义。

虽然协商一致是联大主要的表决方式,每年常会都得到大量采用,但绝大多数有关联大表决数据的研究仍是基于唱名表决,还是会忽略协商一致。这种情况相当于大量赞成票没有纳入统计,造成成员国表决投票的相似数值远比实际情况低。即使目前使用范围最广的理想点指标也存在这一问题。如果纳入协商一致的表决记录,各国的理想点数值必然出现较大变动,以理想点指标为基础形成的研究推论很可能出现严重偏差。比如,艾尔波托·艾莱斯纳(Alberto Alesina)与杜大伟(David Dollar)讨论发达国家基于何种因素选择援助国以及援助规模的问题。文章也关注并利用联大表决数据,分析发达国家的援助是否会受到联大投票相似程度的影响,结果表明美国、日本、法国、德国、英国的对外援助确实会受到发展中国家在联大投票相似程度的影响,特别是日本的对外援助同这一因素高度相关,即与日本在联大投票一致程度越高的国家会获得越多的日本对外援助。^①然而,弗兰克·海格(Frank M. Häge)与西蒙·雨歌(Simon Hug)在复现这篇论文数据的基础上,纳入协商一致的数据,重新进行模型估计后,发现新数据的估计结果出现偏差,甚至部分估计结果显著不支持原文结论。^②虽然在既有唱名表决数据的基础上增加协商一致数据确实要面临较高的技术难题,即使海格与雨歌上篇论文也只是作出初步开拓,但为了减少数据估计结果的偏差,拓展相关研究工作与数据建设还是非常必要的,特别要在理想点指标中纳入协商一致的数据,从而更为高效准确地利用联大表决数据。

第二,没有重视缺席现象。

除了“赞成”“弃权”和“反对”三种投票选择外,联大在事实上还在频繁出现一种特殊的投票选择,即缺席。缺席并不同于“弃权”,后者仍然表明当事国实际参与过联大表决环节,也是联大议事制度正式认可的投票选择,并会被记录在案。但缺席则表明当事国根本没有参与表决,极少甚至不会被

^① Alberto Alesina and David Dollar, “Who Gives Foreign Aid to Whom and Why?” *Journal of Economic Growth*, Vol. 5, No. 1, 2000, pp. 33-63.

^② Frank M. Häge and Simon Hug, “Consensus Decisions and Similarity Measures in International Organizations,” *International Interactions*, Vol. 42, No. 3, 2016, pp. 503-529.

记录在案。比如,2014年3月27日,联大就当时的克里米亚事件以及俄乌关系问题展开激烈的会议辩论,并就西方众多国家支持的决议草案进行表决。虽然该决议草案以100票赞成、11票反对、58票弃权的结果获得通过,但当时包括伊朗、塞尔维亚、以色列在内的20多个国家代表并没有参加这次会议,也没有参与投票。^①如图3所示,纵观联大的发展史,成员国缺席表决并非罕见现象。联大公开的全体会议正式记录中经常可以发现成员国的缺席,且表决记录中参与投票的国家数量往往也会少于联合国成员国总数,甚至会出现某些国家在当天会议中选择性参与表决的现象,即在同一天参与部分会议表决,也会缺席其他的会议表决。

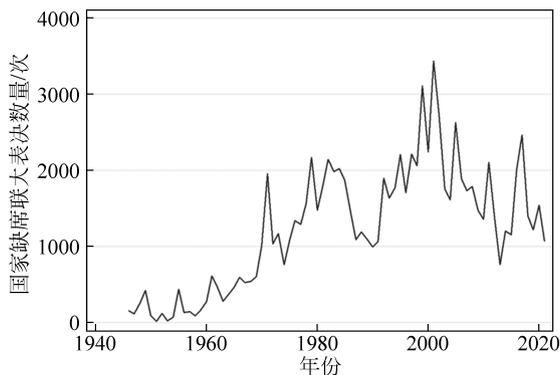


图3 联大表决缺席数量图(1946—2022年)

资料来源:联合国大会投票数据。^②

国家缺席联大表决的现象何以出现?戴安娜·庞克(Diana Panke)认为,国家的外交能力与联大议题领域的不同会影响缺席出现的概率。部分小国往往会因参会代表团人数少、缺少相关专业知识、工作效率较低等原因缺席表决。同时,联大不同议题领域出现缺席的概率也大为不同,如涉及联

^① 《第六十八届会议第八十次全体会议正式记录》,2014年3月27日,联合国正式文件系统,文号:A/68/PV.80,第1—25页。

^② Erik Voeten et al., "United Nations General Assembly Voting Data," 2009, <https://doi.org/10.7910/DVN/LEJUQZ>, Harvard Dataverse, V31, 2023 version, 访问时间:2023年11月1日。

大行政、预算、法律等议题的表决时,成员国缺席现象更为频繁。一旦涉及经济、裁军、发展、人道主义等更具实质性与关键性意义的议题领域,缺席现象出现的概率骤降。^① 布里奇特·考金斯(Bridget Coggins)和朱莉娅·莫尔斯(Julia C. Morse)则将联大表决缺席出现的原因分为两类,即制度性剥夺和战略性缺席。制度性剥夺指联大可以依据《联合国宪章》的规定,剥夺没能及时足额缴纳会费的国家参与联大表决投票的权利。自20世纪90年代以来,联合国会费总额不断增长,越来越多的国家已经无法承担相应份额,随之被剥夺表决投票权利。战略性缺席指部分国家出于政治考虑,不愿陷入某些争议性较高的联大会议议题,避免同强国发生矛盾。^② 相较于弃权,缺席更凸显国家立场与偏好的模糊性,更适合中小国家在联大表决中选择“明哲保身”,回避一定的国际政治压力。

现有对缺席现象的研究更加凸显其重要意义,有时缺席甚至比直接参与表决更能显示出国家的外交政策偏好,特别是考金斯与莫尔斯提出的“战略性缺席”问题更在明确指出缺席所蕴含的复杂政治内涵及其外交战略意义。虽然沃伊特恩等学者搜集的联大原始表决数据集确实具备成员国的缺席记录,但没有进一步发掘缺席数据记录背后的意义。理想点指标也没有系统性处理国家的缺席信息,只是强调缺席并非针对联大表决本身,更可能是当事国因内战或政变等原因没能形成有效政府,因而可能出现持续性的表决缺席现象。^③ 现今绝大部分有关联大表决数据及相关的实证研究成果并没有重视缺席问题,部分原因是缺席蕴含的政治意义过于复杂:既可能是当事国因没能及时足额缴纳联合国会费从而被临时性剥夺投票权利,也可能是因国内政治动乱、避免卷入大国纷争等原因主动拒绝参与投票。过多

① Diana Panke, “Absenteeism in the General Assembly of the United Nations: Why some Member States rarely Vote,” *International Politics*, Vol. 51, No. 6, 2014, pp. 729-749.

② Bridget Coggins and Julia Morse, “Your Silence Speaks Volumes: The Politics of Absence in the UN General Assembly,” the 14th Annual Conference, The Political Economy of International Organization, Oxford, United Kingdom, July 7-9, 2022, pp. 1-57.

③ Michael A. Bailey et al., “Estimating Dynamic State Preferences from United Nations Voting Data,” p. 432.

的政治动机导致相关数据搜集分析的客观困难较大。

联大制度设计的初衷是为了体现主权国家不论大小强弱一律平等,都拥有投票参与表决的权利。国家主动回避,甚至放弃这一权利无疑表明联大制度设计并非完美,也表明联大的投票表决并非只具备象征意义,相反也会产生政治压力,迫使一些国家陷入“左右为难”的境地。基于此,增加对缺席现象的经验研究已经是当前联大表决数据有待突破的难点问题。后续研究可以在既有研究基础上探索建设更为全面优质的缺席数据集,进一步分析其所蕴含的政治意义及完善相关的数字指标。

第三,难以克服议程变化的影响。

虽然诸多研究成果都在利用联大表决数据在各自研究范围内进行经验验证,但大多没有仔细区分联大表决投票所针对的具体议程,研究设计与所得结论都存在一定“隐患”。即使理想点指标已经作出重要突破,但这种处理方法也并非完美无缺。不足之处恰恰也是理想点指标的创新之处,即国家围绕联大每年重复出现决议草案的表决投票行为能否有效反映出其外交政策偏好?塞缪尔·布瑞兹(Samuel Brazys)与戴安娜·庞克曾在考察国家对于联大重复出现议题的投票选择发生变化的现象时,强调国家往往具有强烈动机规避孤立状态,特别是在联大高级政治议题的表决上,国家有时为了不当表决结果的“少数派”,会主动改变立场并同大多数国家的投票选择保持一致。^①国家主动规避孤立的行为无疑在动摇联大重复出现决议草案的“桥梁”性作用,进而影响理想点指标的解释效力。同时,联大重复出现的决议草案更多涉及的是国际规范的内容,国家围绕其产生的表决投票也往往只代表国家对这些国际规范的基本立场与态度,同国家实质性与关键性的利益相距甚远。

当前解决联大表决数据的议程变化问题与提高联大表决数据的解释力度逐渐陷入“鱼和熊掌”的困境,二者顾此失彼,很难同时得到根本性改善。

^① Samuel Brazys and Diana Panke, “Push and Pull Forces in the UNGA: Analyzing Foreign Policy Change in the Context of International Norms,” *International Politics*, Vol. 54, No. 6, 2017, pp. 760-774.

为了克服议程变化的影响,研究人员不得不将研究重心置于每年联大会议议程的重复性内容,但基于这些议程的投票表决大多只具有象征性和规范性意义,不会对成员国外交政策造成实质性影响。围绕这些议题产生的国家间投票差异往往并非基于国家彼此间外交立场的矛盾,大多数也并不适合解释为国家间外交政策偏好的差异。然而,利用解释力度更强的联大表决记录则很难排除议程变化的隐患,无法有效揭示出长时段表决数据背后的国家间外交立场差异程度。基于此,现有的联大表决数据仍要继续探索解决议程变化问题的新方式,尽可能补齐这一短板,提升相关研究成果的说服力。

第四,表决数据的信息有限。

即使生成囊括协商一致与唱名表决的数据集,联大表决数据也必然还拥有一个严重缺陷,即联大表决数据所包含的国家外交立场与偏好的信息量很可能达不到预期。现有的联大表决数据大多过分关注表决这一环节,导致自身忽略掉诸多联大其他工作环节的内容及其背后所蕴含的有用信息。早在1985年,布莱恩·汤姆林(Brian W. Tomlin)就曾撰文讨论联大表决数据的限度问题,强调当时大多数研究人员并没有对联大表决数据的效度问题进行深入研究,导致基于联大表决数据产生的数据指标可信程度并不高。^①彼得森也曾总结,联大发起投票表决一般基于如下情况:一是多个国家决定放弃继续谈判,期望利用联大表决程序通过其支持的决议草案;二是国家计划利用投票表决,使自身或其他国家对决议草案的观点记录在案;三是大多数国家会利用投票表决程序,展示持反对态度的国家正处于少数孤立状态。^②表决只是联大的程序之一,也只蕴含部分国家偏好。如图4所示,联大成员国间的会议辩论、决议草案的生成、委员会的研究报告、法律文本的编纂等大量工作极少付诸表决,或者表决只是诸多工作完成后执行的会议程序。

联大投票表决程序的背后往往隐藏着复杂的成员国间互动与偏好展

^① Brian W. Tomlin, "Measurement Validation: Lessons From the Use and Misuse of UN General Assembly Roll-Call Votes," *International Organization*, Vol. 39, No. 1, 1985, pp. 189-206.

^② M. J. Peterson, "The UN General Assembly," p. 240.

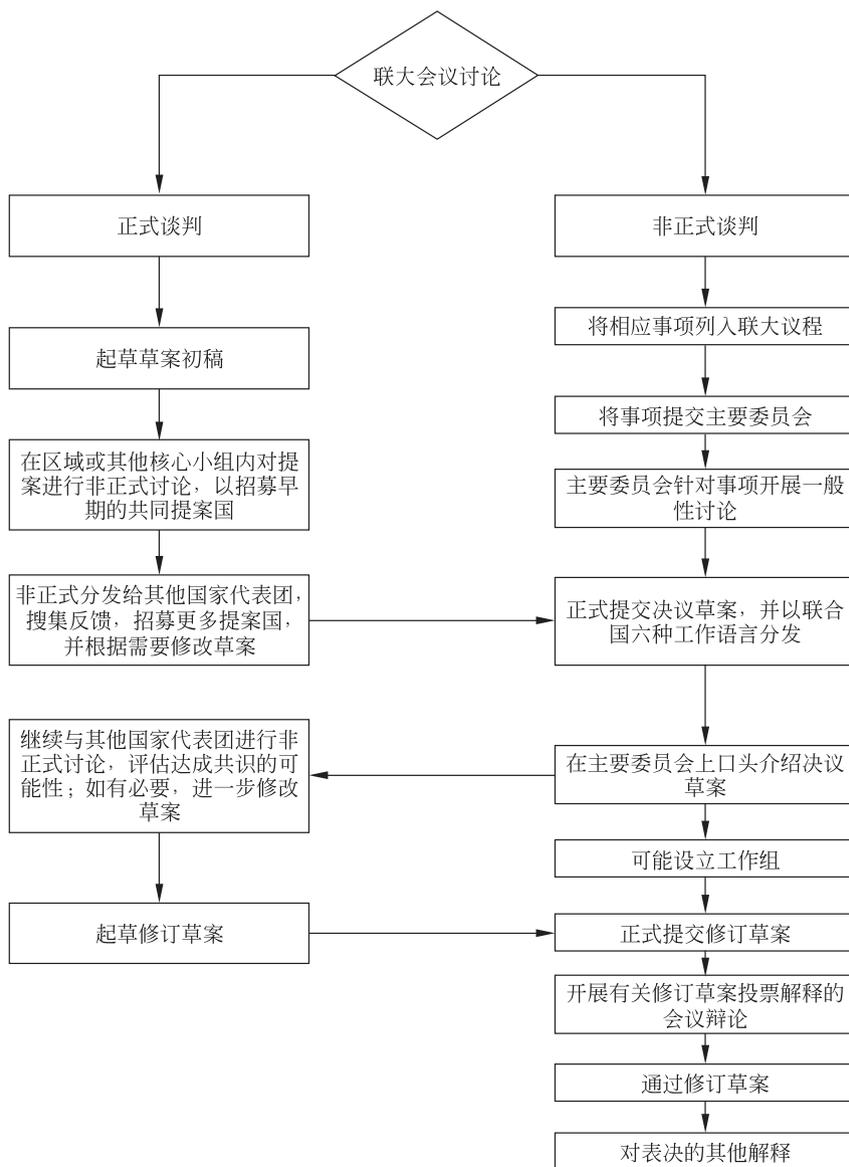


图4 联大的会议辩论流程图

资料来源：“The UN General Assembly”。^①

^① M. J. Peterson, “The UN General Assembly,” p. 240.

示,比如联大在正式会议谈判环节之外往往会产生大量的非正式谈判。由于成员国数量与会议议程不断增多,联大常会的正式谈判环节不断挤压,国家需要在极短的时间内高度简洁地表达观点和意见。在这种情况下,无论联大行政机构还是国家都需要进行更多非正式谈判,充分进行磋商。非正式谈判往往没有固定形式,甚至大多发生在休息室、走廊等处。外界对于会议表决前出现过何种磋商与谈判的经历几乎一无所知,但这些却是深入了解国家偏好的重要资料。尼古拉斯·布尔梅斯特(Nicolas Burmester)与迈克尔·扬科夫斯基(Michael Jankowski)在对欧盟国家联大投票凝聚力的研究中,发现欧盟表面上会呈现出较低的投票集团凝聚力,但在控制联大分歧程度后,就会发现欧盟国家往往会在争议性较高的决议草案表决投票中保持高水平的集团凝聚力,在特定时期甚至居于全球主要区域组织之首。^①如果对投票表决的背景了解不多,无法确定投票表决的决议草案对不同国家政策与利益的影响程度,研究人员仅凭简单的投票选择数据很难有效推断出国家间外交立场的具体差异。

第五,单一维度的不充分问题。

这一问题主要针对理想点指标及其相关的研究成果。理想点指标的计算过程时刻都在遵循一个前提,即美国在联大的投票选择会严重影响其他联合国成员国的选择,特别是纳入理想点指标计算模型的“桥梁”性投票表决记录更要信守这个前提。这种假定只强调联大当中的美国维度,同美国理想点数值差距越大表明该国同美国外交政策偏好越远。美国与在联大极端反美的国家处于两个极端,其余成员国居于其中。理想点指标的主要任务就是利用长时期的联大唱名表决数据反映出联合国其他成员国对美国领导的霸权秩序与外交意识形态的接受程度。

然而,国家间的外交立场相近程度不可能仅靠美国单个维度得到“通解”。一方面,美国在联大的影响力很可能被学术界过高估计,诸多联大投票表决记录都显示出美国是极少数投反对票的国家。比如,拉斐尔·梅斯基塔(Rafael Mesquita)等在对联大提案数据估计的基础上,发现美国在联大

^① Nicolas Burmester and Michael Jankowski, “Reassessing the European Union in the United Nations General Assembly,” *Journal of European Public Policy*, Vol. 21, No. 10, 2014, pp. 1491-1508.

并没有达到预期中的积极活跃水平,其在联大的活动轨迹和部分小岛国相仿,先验地将美国作为分析联大数据的单一核心维度,很可能导致后续数据估计推论出现偏差。^①另一方面,理想点指标很难处理“特立独行”的成员国。一旦某些国家没有考虑美国立场而独立作出投票选择,这些国家的行为就没有满足计算理想点指标的前提,得出的数值自然也同实际不符。事实上,联大表决记录中也确实存在这类成员国。比如,本杰明·劳德黛尔(Benjamin E. Lauderdale)在讨论理想点估计方法时,发现印度、俄罗斯、以色列、土耳其四国在联大的表决行为非常类似美国国会投票中的“特立独行者”(mavericks),四国很少参与其他国家组成的投票集团,其投票偏好与实际选择同样很难得到准确预测,导致四国的理想点数值很容易出现失效等问题。^②

事实上,贝里与沃伊特恩也注意到这一问题,并尝试建立联大表决数据中的第二维度。两位学者2018年增加“南北方矛盾”这一维度,拓展出双维度的估计模型,即在估计联大表决数据时也考虑发达国家与发展中国家在联大表决中凸显的偏好分歧。但两位研究者指出,第二维度并不稳定且也不容易得到充分解释,只在特定时期才会强烈影响国家在联大的投票选择,并建议后续应用理想点指标的研究最好还是只采用第一维度进行数据估计。^③

总的来说,理想点指标出现单一维度的问题根源于研究人员为方便利用联大数据而不得已作出的选择。由于联大涉及太多全球性议题,国家在这些议题上的外交政策偏好也各有不同。如果没有恰当维度作为线索将国家的外交政策偏好统一分类比较,联大表决数据就几乎无法得到有效应用。然而,联大海量的议题与表决投票记录不可能仅仅遵循“反美”与“亲美”的单一维度展开。不同于冷战时期激烈的超级大国对抗,近年来的国际关系形势更为复杂,国家间既有合作也有分歧。在参与联大投票表决的过程中,

① Rafael Mesquita, Rodrigo Martins, Pedro Seabra, “Estimating Ideal Points from UN General Assembly Sponsorship Data,” *International Interactions*, Vol. 48, No. 6, 2022, pp. 1233-1252.

② Benjamin E. Lauderdale, “Unpredictable Voters in Ideal Point Estimation,” *Political Analysis*, Vol. 18, No. 2, 2010, pp. 166-167.

③ Michael A. Bailey and Erik Voeten, “A Two-Dimensional Analysis of Seventy Years of United Nations Voting,” *Public Choice*, Vol. 176, 2018, pp. 33-55.

国家需要考虑的因素也明显增多。在这种复杂情况下,仅依靠单一维度计算得出的理想点指标难以准确概括出国家对各个全球议题的外交偏好,基于此得出的研究推论也很容易出现偏差。

四、合理应用联大表决数据的建议

鉴于一系列客观存在的难点问题,既有联大表决数据及相关数据指标在短期内无法得到根本性完善。研究人员应在了解联大表决数据特征与局限的基础上,顾及若干实际应用的注意事项,更为合理高效地完成研究设计工作,或进一步发掘联大表决数据的有效信息,使其成为更为完善的研究工具。

(一) 对理想点进行必要的技术处理

作为当前联大表决数据中最具影响力的研究成果,理想点指标已经得到国内外学术界的普遍关注与应用。部分研究成果往往采用对子国家同一年份理想点的差值衡量其外交立场的相近程度。如作为控制变量等非关键研究对象的操作方法,这种简单快捷的研究设计方式本无可非议,但一旦涉及解释变量或被解释变量等实际研究中的关键对象,这种操作方法则略显不足,须进行更为精细的研究设计。这是由于理想点指标主要针对联合国其他成员国对美国国际秩序偏好的认可程度,进而被当作比较他国与美国外交立场相近程度的“尺度”。本质上,除美国外的两国理想点差值更适合说明这两国与美国外交立场相近程度的差异,但这种差异很难充分说明两国本身之间的外交立场差异。

为了提高衡量效度,研究人员最好避免简单使用理想点指标,尽可能根据自身研究范畴对理想点指标采取进一步的技术处理。参考既有研究,本文推荐的简明操作方法是设计双变量,即在生成原有研究对象对子国家同一年份理想点差值变量的基础上,另外再生成与美国同一年份理想点差值的变量,具体选择何种国家生成后者变量则需要根据研究实际情况进行单独设计。这种方法可以提高相关研究设计与数据估计推论的可信度与稳健性。

更为理想的做法是基于理想点的理论基础与计算方法,在技术上开发新维度甚至多维度的新理想点指标。虽然贝里与沃伊特恩没有充分认可联大表决数据中的其他维度,但国家在联大表决中表现出的外交政策偏好必然还存在其他维度的解释。特别是南北国家对立背后的全球财富和权力分配不均衡问题已经成为近年来联大诸多成员国的主要分歧,围绕这一问题展开的会议辩论与表决投票也非常具备冲突性,适合作为新的分析维度,^①忽略这一维度的表决记录无疑会浪费大量国家外交政策偏好的信息。如果研究问题与对象契合,研究人员完全可以根据南北冲突这一维度计算得出解释力度更强的理想点指标,大幅提升理想点的技术处理水平。

(二) 妥善处理弃权票记录

在正式表决环节中,联大使用“赞成”“弃权”和“反对”三者之一记录各个参与表决国家的投票选择。“赞成”和“反对”两种投票选择非常清晰,相关数据的技术处理也非常容易,但“弃权”的情况较为复杂。虽然现有联大表决数据已经大范围涵盖成员国“弃权”的投票记录,但目前学术界对弃权票的理解与数据化处理的方式仍存在明显差异。

联大自身强调弃权票的中间意义。《联大议事规则》明确写明,只有投赞成票或反对票的成员国才纳入票数统计,统计结果才能决定表决是否有效与通过,“弃权的成员国应被认为没有参加表决”。^② S值也将弃权票看作中间情况,表达赞成与反对之间的中性意义。^③ 沃伊特恩则认为弃权更接近于反对,因为联大表决结果并不具有强制力,即使国家投赞成票也不会付出巨大政治成本,国家选择弃权是在间接表达不赞成态度。^④ 然而,在计算理想点

① M. J. Peterson, “The UN General Assembly,” p. 245.

② 《大会议事规则》,2022年,联合国正式文件系统,文号:A/520/Rev. 20,第21页,来源:<https://www.un.org/zh/ga/about/ropga/>,访问时间:2023年11月1日。

③ Curtis S. Signorino and Jeffrey M. Ritter, “Tau-B or Not Tau-B: Measuring the Similarity of Foreign Policy Positions,” pp. 128-129.

④ Erik Voeten, “Clashes in the Assembly,” *International Organization*, Vol. 54, No. 2, 2000, p. 193.

指标时,沃伊特恩等学者还是将弃权当作区别于赞成与反对的中性投票选择。^①

之所以出现上述不同认识很大程度上缘于弃权本身具备的模糊性特征,并且弃权也会因所处时代与具体政治氛围产生出不同的内涵与意义。比如,研究一旦涉及联大成立早期的表决数据,可以将当时的弃权记录操作化为“反对”。但在分析联大1960年后的表决数据时,研究人员应将“弃权”单列,区别于赞成票与反对票。这是因为早期的联大确实可以将弃权票等同于“反对”,但在1960年后联大出现的“弃权”更多表示对“赞成”“反对”任何一方都没有强烈意见,或是成员国为避免成为联大“孤独的反对方”。^②在当前联大的政治氛围中,反对票又在部分表决环节上发展成为“非常反对”,表达成员国对表决所涉内容的强烈不满,而弃权票的内涵又几乎等同于“反对”。如表1所示,即使频繁投出反对票的美国也会在部分自身不支持的表决中选择“弃权”。这种情况则允许研究人员将近年来的弃权票数据视为“反对”。此外,在涉及与美国投票是否一致的问题时,研究人员应将“弃权”视为同美国投票一致。这是由于美国通常会在自身关注的联大表决议题中游说别国选择与美投票一致,最低限度也要实现别国态度不相左,即在表决中选择“弃权”。^③

弃权票的模糊性特征极易影响相关的经验研究,不同的处理方式往往会得出差异较大的研究结果。为了充分说明这一问题,本文将复现古斯塔沃·弗洛雷斯-马西亚斯(Gustavo A. Fbres-Macias)和萨拉·克雷普斯(Samh E. Kreps)2013年的研究,在调整其对弃权票操作方式的基础上重新估计原文的数据集。这篇论文主要利用1992年至2006年间联大针对具体国家人权决议草案的表决数据,分析中国同非洲和拉丁美洲国家的贸易变

① Michael A. Bailey et al., “Estimating Dynamic State Preferences from United Nations Voting Data,” pp. 435-436.

② M. J. Peterson, “The UN General Assembly,” pp. 240-241.

③ David B. Carter and Randall W. Stone, “Democracy and Multilateralism: The Case of Vote Buying in the UN General Assembly,” *International Organization*, Vol. 69, No. 1, 2015, p. 14.

表1 美国联大投票选择情况表(1946—2022)

投票选择	选择频次	所占百分比/%
赞成票	1879	29.08
弃权票	1272	19.68
反对票	3263	50.5
缺席	48	0.74

资料来源:联合国大会投票数据。^①

化是否会影响这些国家同中国外交立场相近程度的问题,并发现这些国家与中国的对外贸易量越大越可能在联大特定国家人权相关决议草案表决上与中国投票相近,即在一定程度上与中国外交政策趋同。^② 文章的被解释变量是两大区域的国家同中国外交立场相近程度,具体操作化为这些国家同中国在上述联大表决中年度投票相似数平均值的自然对数,其中一国投票与中国一致的情况赋值为1;如果一国投票与中国不一致且不为弃权票则赋值为0,弃权票没有被视为反对票或不同意,并将赋值定为0.5。^③

然而,这种“中间立场”式的操作方式并不合理。一方面,原文针对弃权票的操作方式是参考斯特罗姆·塞克尔(Strom C. Thacker)1999年的研究成果,但塞克尔的论文关注美国政府认为的联大重要投票及相关国家的投票记录。^④ 弗洛雷斯-马西亚斯和克雷普斯的主要研究对象是联大针对具体国家人权决议草案的表决数据以及中国同非洲和拉丁美洲国家的投票选择。二者的关注对象并不完全一致,完全照搬操作方式并不严谨。另一方

^① Erik Voeten et al., “United Nations General Assembly Voting Data,” 2009, <https://doi.org/10.7910/DVN/LEJUQZ>, Harvard Dataverse, V31, 2023 version, 访问时间:2023年11月1日。

^② Gustavo A. Flores-Macias and Sarah E. Kreps, “The Foreign Policy Consequences of Trade: China’s Commercial Relations with Africa and Latin America, 1992—2006,” *The Journal of Politics*, Vol. 75, No. 2, 2013, pp. 357-371.

^③ Gustavo A. Flores-Macias and Sarah E. Kreps, “The Foreign Policy Consequences of Trade: China’s Commercial Relations with Africa and Latin America, 1992—2006,” pp. 361-362.

^④ Strom C. Thacker, “The High Politics of IMF Lending,” *World Politics*, Vol. 52, No. 1, 1999, pp. 46-49.

面,原文筛选的表决记录都在针对具体国家,决议草案往往极具争议性。中国在联大此类人权议题上的政策立场也一贯鲜明直接,即坚决主张捍卫国家主权和领土完整的原则,并公开强烈反对西方国家利用人权问题肆意干涉别国内政,投票选择也以反对或弃权为主,二者内涵相近,都是中国对相关决议草案表达否定态度。非洲和拉丁美洲各个发展中国家在联大同类议题上的投票选择必然受到中国立场及其他外部因素的影响,投出的弃权票也应视为对相关决议草案的反对态度。此外,原文关注的联大表决记录全部处于冷战后时代,各个国家也往往倾向利用联大弃权票表达反对态度。

基于上述原因,本文将提升弃权票的内涵,将其等同为反对票,并更改原文的操作方式,重新操作被解释变量。具体操作指标如下:当中国投赞成票时,非洲和拉丁美洲国家也投赞成票则赋值为 1,投反对票或弃权票则赋值为 0;当中国投反对票或弃权票时,非洲和拉丁美洲国家也投反对票或弃权票则赋值为 1,投赞成票则赋值为 0。此外,成员国投票缺席与非联合国成员国的特殊情况则为缺失值。参照原文的操作方式,全部赋值会以年和国家的形式相加并得出年度的平均数,进而完成被解释变量的操作化。

修改完被解释变量后,本文采用两位作者公开的论文数据,纳入其他解释变量与控制变量,生成新的数据集,重新估计原文展示的前六个计量模型。数据与模型均采用 STATA12.0 软件运行,分析结果如表 2 所示。

表 2 中国对外贸易对拉非国家联大投票相似度的复现数据分析结果表

	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
贸易额度(中)	0.035* (0.011)					
贸易额度(中) ($t-1$)		0.034* (0.011)	0.035* (0.011)			
贸易显著度(中)				0.004 (0.003)		
贸易显著度(中) ($t-1$)					0.005* (0.002)	0.007* (0.003)
国家实力	0.062 (0.123)	0.078 (0.121)		0.092 (0.124)	0.087 (0.125)	

续表

	模型 1	模型 2	模型 3	模型 4	模型 5	模型 6
国家实力 ($t-1$)			0.067 (0.116)			0.097 (0.122)
美国援助水平	-0.008 (0.008)	-0.007 (0.008)		-0.011 (0.008)	-0.005 (0.009)	
美国援助水平 ($t-1$)			-0.003 (0.007)			-0.010 (0.007)
贸易额度(美)	0.021 (0.022)					
贸易额度(美) ($t-1$)		-0.007 (0.017)	-0.001 (0.023)			
贸易显著度(美)				0.002 (0.003)		
贸易显著度(美) ($t-1$)					0.001 (0.003)	0.001 (0.002)
政体相似度	0.123 (0.072)	0.122 (0.071)		0.115 (0.073)	0.112 (0.071)	
政体相似度 ($t-1$)			0.110 (0.065)			0.085 (0.063)
人权水平	0.018 (0.024)	0.015 (0.024)		0.015 (0.025)	0.016 (0.025)	
人权水平 ($t-1$)			0.031 (0.022)			0.032 (0.023)
2003年后	0.127* (0.035)	0.153* (0.032)	0.151* (0.032)	0.194* (0.032)	0.196* (0.030)	0.192* (0.030)
固定效应	是	是	是	是	是	是
p 值	0.000	0.000	0.000	0.000	0.000	0.000
R^2	0.156	0.151	0.157	0.134	0.134	0.141
样本数	843	843	840	843	843	840

资料来源:作者自制。

注:复现数据结构为面板数据,估计方法为固定效应的混合最小二乘法模型, $t-1$ 指相关变量的滞后一年值;所有数值均取至小数点后第三位,括号内为稳健标准误;* $p < 0.05$ 。

对比原文的主要发现,复现数据的分析结果呈现出巨大差异。首先,复现数据分析结果中解释变量的系数普遍高于原文的分析结果,表明调整后的被解释变量同解释变量间存在更强联系,即中国对外贸易在对拉非地区国家联大投票选择上的影响力比原文预期的要更高。其次,复现数据的分析结果并不支持原文中有关贸易额度滞后效应的发现。原文的分析结果发现中国贸易额度滞后一年的解释力要高于同期,但复现数据结果则表明二者解释力度的差别极小,甚至同期的解释力还略高。然而,复现数据的分析结果并没有在中国贸易显著度同期与滞后一年上与原文产生根本不同,大体支持原文相关部分的发现。再次,虽然复现数据分析结果也表明原文中有关 2003 年后全球反美主义浪潮兴起对于被解释变量拥有较强解释力度,但复现数据的解释方向与原文完全相反,前者的系数为正,表明 2003 年后兴起的全球反美浪潮对于被解释变量可以产生正向影响,但后者系数为负,呈现的是反向解释。虽然原文认为 2003 年伊拉克战争之后高涨的全球反美浪潮会促使他国与美国的外交分歧越来越大,特别会在联大投票上更倾向中国立场,但其数据估计结果并未支持这一点,反而是本文复现数据的分析结果验证了这一假设。最后,复现数据分析结果并不支持其他控制变量的解释力,特别是原文有关国家实力的相关变量,复现数据的分析结果并不显著。

基于上述分析,本文认为,联大弃权票的复杂特性要求研究人员必须谨慎对待。妥善处理联大的弃权票记录很可能会进一步完善既有的联大表决数据分析结果,提升数据研究发现与结论的稳健程度。简而言之,本文建议使用近 20 年以来联大表决原始数据的相关研究应将弃权票等同为“反对”,并进行相关的操作化。这种处理方式既基本符合当前联大的发展现实,又可以减轻研究设计的复杂程度。诚然,以上所列情况并不能涵盖与解决操作弃权票时面临的所有问题。研究人员还是要根据自身的与研究对象妥善处理弃权票数据,最好做到“具体问题具体分析”,判断其所具备的意义,避免“一刀切”式的简单化操作,尽可能提高衡量工作的效度。^①

^① Guillermo Rosas, Yael Shomer, Stephen R. Haptonstahl, “No News Is News: Nonignorable Nonresponse in Roll-Call Data Analysis,” *American Journal of Political Science*, Vol. 59, No. 2, 2015, pp. 511-528.

(三) 恰当筛选表决数据

合理应用联大表决数据应尽可能在自身研究问题或研究对象所需的基础上,进行恰当筛选,聚焦更具解释力度的表决记录,特别是在没有利用如理想点等现有指标,而是采用原始联大表决数据进行研究时,这一步骤更为关键。纵观联大的发展历程,大多数表决都以大比例赞成票通过,围绕决议草案展开激烈辩论或投票表决分歧严重的现象相对较少。部分表决更是针对联大会议的程序问题,极少产生争议点。这些表决并不会对联大成员国的国家利益产生实质性影响,成员国也很少投入注意力关注表决结果或影响别国的投票选择。国家在这部分联大表决中无论是投票趋同还是趋异,都很难被用来衡量彼此间外交立场的远近。近年来具备一定学术影响力的研究成果也很少基于全部的联大表决数据,普遍进行过不同程度的数据精简与筛选工作。综合既有研究,本文认为筛选联大表决数据可以遵循如下两条标准。

一是挑选国家更为重视的表决议题。在使用联大原始表决数据进行研究时,研究人员往往要利用联大议题差异的特征突出更重要与更具解释力度的表决记录。一旦投票表决所涉具体议题没有得到仔细区分,以此为基础形成的有关国家外交政策偏好的推断难免会受到影响。研究人员更应关注更具争议性的联大表决记录,特别是研究所涉国家极为重视的表决议题,或是会对该国国家利益产生实质性影响的表决议题,进而观察这些议题与具体的表决投票情况,更为高效获取国家间外交立场相近程度的信息。比如,前文提及的弗洛雷斯-马西亚斯和克雷普斯认为中国更为关注联大的人权议题表决;塞克尔则强调美国政府认定的联大重要投票对于观察美国联大行为的重要意义。研究当前俄罗斯联大表决行为问题无疑更适合利用联大涉及俄乌冲突议题的会议与表决。^①

二是限定研究数据的时段。一方面,研究人员不必执着于长时间的研

^① Mohammad Reza Farzanegana and Hassan F. Gholipour, "Russia's Invasion of Ukraine and Votes in Favor of Russia in the UN General Assembly," *International Interactions*, Vol. 49, No. 3, 2023, pp. 454-470.

究范围,应根据实际需要尽可能限定数据记录时段,达到类似时间类控制变量的作用。比如,一旦研究涉及中国在联大的表决投票行为,研究人员应尽可能将1978年作为研究起点。虽然新中国于1971年10月25日成功恢复联合国合法席位,但在之后的几年里,中国对联合国涉及的各种全球性问题大多谨慎参与,特别是在联大表决时,中国代表常常选择出席会议但不参与表决。出现这一现象部分是源于中国政府刚接触联大各项日程与议题,缺少相关情况的调研导致中国代表在联大的行为较为谨慎。这种状态在1978年后才得到大幅改变,中国开始大力参与联大各项议程,并积极参与会议表决。基于中国的特殊情况,研究人员应将1978年作为观察中国联大表决记录的起点,避免过多缺失值造成数据估计出现偏差。

另一方面,限定数据的研究时段还要注意联大表决时间上的滞后效应。国家间联大表决趋同与否大多是其彼此联系互动后的结果。在实际研究中,我们应该注意诸多影响国家在联大投票选择的因素极少局限于同期,更多会出现一定期限的滞后。比如,某国的对外援助行为不仅可能造成该国与受援国当年联大表决趋同的比例上升,二者大概率仍会在援助之后的一到两年间保持较高的表决趋同水平。这种情况使得研究人员要纳入相关的一期甚至二期滞后变量,多观察一定时段的联大表决情况,满足相关变量滞后效应的要求。

上述建议内容只是常出现的研究情况。无论研究人员采用何种方式筛选联大表决数据,都应基于自身研究需要与联大实际发展情况,灵活设计数据筛选方式。

(四) 探索新类型的联大数据集

鉴于现阶段联大表决数据仍存在多种难以回避的难题,当代一部分学者开始尝试发掘联大尚未被开发的学术研究资源,探索新的数据内容。目前,这类探索研究工作也取得初步成效,形成一些颇具特色的联大数据集,比如有关联大一般性辩论的数据。一般性辩论通常设置在每年9月联大召开新一届会议之初,届时联大会组织一次大型的高级别会议,邀请各国的国家元首或其他高级别代表齐聚联合国总部,并为每个成员国都提供发言机会供其阐述本国的核心外交理念和针对任何国际问题的观点。一般性辩论

早已经成为联大正式且固定的会议环节,即使面对新冠疫情的冲击也只是转为线上举办,没有改变每年举行的既定规律。相较于投票表决,国家很少在一般性辩论环节面临外部政治束缚与压力,也会更多表达本国的外交立场与优先关注的议题,这些特征使得一般性辩论非常适合用来系统分析国家的外交政策偏好。基于联大一般性辩论的特征,亚历山大·巴图罗(Alexander Baturo)等学者利用搜集到的7300多份国家发言文本,组建出1970—2014年间联大一般性辩论的文本语料集,并进行初步应用。^①

除一般性辩论之外,也有学者注意到联大的提案可以搜集整理成为新的数据集。由于联大决议大多会在起草阶段组织国家讨论修改,关注相关提案得到国家何种支持程度或提出过何种修改意见也可以反映出国家的外交政策偏好。拉斐尔·梅斯基塔等学者就将注意力放到联大的提案数据,并利用项目反应理论估计2009年至2019年联大提案数据中国家的理想点,并得到表决数据计量估计没能取得的新发现:如美国在联大的重要性与影响力并没有预期高;七十七国集团和欧盟两个大型多边性质的国家集团正在联大呈现出“两极化”的现象。^②

鉴于联大制度本身内容丰富,研究人员可以在其中发掘更新且具备高度研究价值的联大数据,比如联大决议草案的段落表决数据仍有待开发。由于联大部分决议草案内容庞杂且篇幅较大,联大有时会针对这类决议草案实施段落表决,即每一段的内容都会产生一轮表决。段落表决结束后,联大再发起针对决议草案整体内容的表决。多轮的段落表决结果通常不会和整体内容的表决相同,且更易凸显出成员国的立场偏好,但由于数据分析存在一定技术困难,目前关注联大段落表决数据的研究成果尚未出现。

① Alexander Baturo, Niheer Dasandi and Slava J. Mikhaylov, “Understanding State Preferences with Text as Data: Introducing the UN General Debate Corpus,” *Research & Politics*, Vol. 4, No. 2, 2017, pp. 1-9.

② Rafael Mesquita, Rodrigo Martinsa and Pedro Seabra, “Estimating Ideal Points from UN General Assembly Sponsorship Data,” *International Interactions*, Vol. 48, No. 6, 2022, pp. 1233-1252; Pedro Seabra and Rafael Mesquita, “Beyond Roll-Call Voting: Sponsorship Dynamics at the UN General Assembly,” *International Studies Quarterly*, Vol. 66, No. 2, 2022, pp. 1-12.

五、结论

联合国自 1945 年成立以来就设立全体成员国参与且具有平等表决权的联合国大会。联大的制度设计使其可以几乎涵盖所有的国际重要议题。不同于安理会更为重视五大常任理事国的作用,联大不区分国家大小与强弱,可以使其公正且平等地参与联大表决,但代价是牺牲联大的制度约束力与决议执行力,造成当今联大的影响力远不及安理会。尽管联大审议的议题与相关表决大多只具备象征意义,很少对主要国家的外交战略与国家利益产生实质性影响,但国家的投票选择仍能凸显其外交政策的基本立场与偏好,特别在联大出现一些争议性较强的会议议题时,引发的会议辩论与表决使得国家间外交政策的共识与分歧更为显著。由于关键数据记录获取难度低,且几乎可以囊括全球所有主权国家,联大表决数据已经成为探究国家间外交立场相近程度的有效研究工具。

然而,目前开发出的联大表决数据集及相关的重要指标仍存在进一步完善的空间。该数据集一旦被不合理使用,很容易使得数据估计结果及相关推论出现严重偏差。在讨论联大表决数据实际应用与现有局限的基础上,本文提出研究人员可以基于以下建议更为合理地应用联大表决数据:在涉及关键研究对象时,应尽可能对理想点进行必要的技术处理;要根据自己的研究问题与研究对象妥善处理弃权票数据,可以将近 20 年以来联大出现的弃权票等同为“反对”;采用联大原始表决数据进行研究时应恰当筛选出国家更为重视的表决议题,并关注研究数据的时段,特别要注意联大表决普遍存在的时间滞后效应;探索采用新类型的联大数据集,进一步开发联大表决数据的研究资源。

鉴于联大机制的复杂性以及影响国家投票选择因素的多样性等原因,既有联大表决数据及相关数值指标远非完美无缺。相关领域的研究人员须将解决联大表决数据的既有局限作为这一领域未来研究的根本方向。本文提及的任一局限能得到妥善解决,联大表决数据的质量就会得到大幅提升,衡量国家外交立场相关变量的效度更会获得实质性改善,由此推动其逐步达到“洞若观火”的水平。