

# 信息治理与国内冲突<sup>\*</sup>

——基于信号博弈模型的实证检验

桂晓书 吕蕙伊

**【内容提要】** 随着数字技术的不断进步，关于其如何影响国家安全治理的研究日益增多。然而，对于为何有些政府的信息治理措施能有效缓解国内抗争与冲突，而另一些却适得其反，仍是一个有待充分讨论的问题。有别于聚焦于单一案例或特定类型国家的既有研究，作者从宏观层面探究信息治理措施影响国内冲突风险的内在机制，基于信息不对称的框架构建了一个信号博弈模型，旨在阐明政府如何战略性地运用信息治理措施向潜在挑战者发出威慑信号。该模型发现，成本考量与政府实力是决定治理效果的关键：成本的增加会降低政府实施信息治理措施的可能性；而在相同的治理成本下，随着政府实力的削弱，信息治理作为威慑信号的功效会大打折扣，难以有效塑造政府强大的形象，反而可能加剧国内抗争与冲突爆发的风险。运用2016—2020年的跨国月度面板数据对理论假说进行实证检验后发现，统计结果显示，治理成本越高，政府采取信息治理措施的可能性越低。作为旨在抑制国内冲突的一种潜在手段，信息治理措施的有效性严格受限于政府实力的强大和政权的稳固性；而在政府实力较弱的国家，该措施非但不能遏制国内抗争与冲突，反而显著促进了此类事件的频发，这一负面效应在稳健性检验中依然保持显著。

**【关键词】** 信息政治；国家治理；国内冲突；国家安全；信号博弈

**【作者简介】** 桂晓书，中国人民大学澄海全球发展与安全高等研究院研究员（北京 邮编：100872）；吕蕙伊，清华大学社会科学学院国际关系学系博士研究生（北京 邮编：100084）。

**【中图分类号】** D815.5 **【文献标识码】** A **【文章编号】** 1006-9550 (2025) 03-0063-31

\* 本文系国家社会科学基金青年项目“应用大数据技术预测全球突发重大危机事件引发的社会骚乱研究”（项目批准号：20CGJ015）的阶段性成果。感谢《世界经济与政治》匿名审稿专家的意见与建议，文中疏漏由笔者负责。

## 一、引言

数字技术的迅猛发展正在深刻重构国家治理的权力架构。互联网的广泛普及既显著地增强了民众的政治表达与集体行动能力,<sup>①</sup>也提升了国家的信息能力与治理水平。<sup>②</sup>在数字时代,政府与民众围绕信息流动边界展开的博弈已成为各国政治生活的重要特征。在此背景下,众多政府将信息治理措施视为应对舆论风波、维护社会秩序和巩固自身权威的重要手段。例如,近年来英国和法国等国通过立法赋予了政府在网络空间打击虚假信息的合法权力。此举虽然取得了一定成效,但也引发了社会各界对于言论自由边界的深刻反思与广泛争论。<sup>③</sup>2024年7月,为平息国内愈演愈烈的大规模反政府抗议活动,孟加拉国政府采取了切断移动网络、暂停宽带服务及限制社交媒体访问等措施,深刻影响了全国范围内约1.7亿民众的生活,突显政府在维护社会秩序与回应民众诉求之间所面临的复杂权衡与挑战。<sup>④</sup>

既有信息治理研究主要集中于如何借助数字技术提升公共服务质量与政府响应效率,从而增强国家治理能力。<sup>⑤</sup>本文则聚焦于信息治理的另一重要维度,即通过对信息内容及其传播渠道的管理,以维护社会稳定与政治秩序。<sup>⑥</sup>随着数字技术的飞速发展,信息治理的效能已远超人工和传统通信手段所能及,不仅覆盖范围更加

① 黄贝、陈冲:《数字技术普及与国内暴力冲突》,载《国际政治科学》,2024年第3期,第36—71页。

② 孟天广、郑思尧:《国家治理的信息理论:信息政治学的理论视角》,载《政治学研究》,2023年第6期,第117—130页;吕俊延、刘燚飞:《国家的“视力”:技术革命与国家信息能力建构》,载《政治学研究》,2023年第5期,第97—113页。

③ Stephen A. Meserve and Daniel Pemstein, “Google Politics: The Political Determinants of Internet Censorship in Democracies,” *Political Science Research and Methods*, Vol. 6, No. 2, 2018, pp. 245–263; Ricki-Lee Gerbrandt, “Media Freedom and Journalist Safety in the UK Online Safety Act,” *Journal of Media Law*, Vol. 15, No. 2, 2023, pp. 179–212; Rachael Craufurd Smith, “Fake News, French Law and Democratic Legitimacy: Lessons for the United Kingdom?” *Journal of Media Law*, Vol. 11, No. 1, 2019, pp. 52–81.

④ “Bangladesh Shuts Offices, Imposes Curfew to Curb Deadly Job Quota Protests,” <https://www.reuters.com/world/asia-pacific/bangladesh-army-enforces-curfew-student-led-protests-spiral-2024-07-20/>.

⑤ 孟天广、郑思尧:《国家治理的信息理论:信息政治学的理论视角》,载《政治学研究》,2023年第6期,第117—130页;吕俊延、刘燚飞:《国家的“视力”:技术革命与国家信息能力建构》,载《政治学研究》,2023年第5期,第97—113页;黄其松:《数字时代的国家理论》,载《中国社会科学》,2022年第10期,第60—77页。

⑥ Jennifer Earl, Thomas V. Maher and Jennifer Pan, “The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review,” *Science Advances*, 2022, DOI: 10.1126/sciadv.abl8198.

广泛，还在识别并应对潜在社会稳定风险上展现出更高的精确度和效率。<sup>①</sup>

学界对信息治理的政治效应已展开广泛研究，<sup>②</sup>但对其能否有效遏制国内冲突的爆发尚未形成共识。一类观点认为，数字技术增强了政府精准识别反对派<sup>③</sup>并预测其潜在风险的能力。通过干预反对派的信息传播途径和组织动员渠道，政府的信息治理措施能显著提升反对派组织集体行动的难度，进而有助于降低国内冲突爆发的风险。例如，部分拉美国家的案例表明关停反对派网络平台能在一定程度上减弱其抗争活动频次和冲突的激烈程度。<sup>④</sup>但另一类观点指出，信息治理影响国内冲突的作用机制可能更为复杂，且存在意外触发“逆火效应（backfire effect）”的风险。<sup>⑤</sup>许多中东和非洲国家对媒体和互联网实施的管制措施不仅未能有效缓解国内社会动荡，反而加剧了矛盾，激发了更大规模的冲突浪潮。<sup>⑥</sup>这些案例彰显了信息治理效应的两面性——既可能作为抑制国内冲突的有效手段，也可能在不当使用下成为触发和加剧冲突的推手。

<sup>①</sup> Francesca Musiani, et al., eds., *The Turn to Infrastructure in Internet Governance*, Palgrave Macmillan, 2016, pp. 3–21; Anita R. Gohdes, *Repression in the Digital Age: Surveillance, Censorship, and the Dynamics of State Violence*, Oxford University Press, 2024, pp. 1–5; Dorota Glowacka, et al., “Digital Technologies as a Means of Repression and Social Control,” Policy Department for External Relations, Directorate General for External Policies of the Union, PE 653.636, April 2021, pp. 14–31.

<sup>②</sup> Jennifer Earl, Thomas V. Maher and Jennifer Pan, “The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review,” DOI: 10.1126/sciadv.abl8198; Margaret E. Roberts, “Resilience to Online Censorship,” *Annual Review of Political Science*, Vol. 23, 2020, pp. 401–419; Michael J. Ahn and Stuart Bretschneider, “Politics of E-Government: E-Government and the Political Control of Bureaucracy,” *Public Administration Review*, Vol. 71, No. 3, 2011, pp. 414–424.

<sup>③</sup> 本文中反对派泛指与当权者或主导力量在立场、观点或政策取向上存在显著分歧的个人或群体。反对派通常通过批评、监督或竞争等方式表达对现有政策、决策或制度的不满，并试图推动变革或替代现有秩序。

<sup>④</sup> Jason Gainous, Kevin Wagner and Tricia Gray, “Internet Freedom and Social Media Effects: Democracy and Citizen Attitudes in Latin America,” *Online Information Review*, Vol. 40, No. 5, 2016, pp. 712 – 738; Philip N. Howard and Muzammil M. Hussain, *Democracy’s Fourth Wave? Digital Media and the Arab Spring*, Oxford University Press, 2013, pp. 11–12.

<sup>⑤</sup> “逆火效应”主要是指民众因感到言论受限而被激起强烈的反抗情绪，导致他们更加坚决地表达不满，进而加剧了国内冲突。参见 Erica Chenoweth and Maria Stephan, *Why Civil Resistance Works: The Strategic Logic of Nonviolent Conflict*, Columbia University Press, 2011, pp. 50–51; Ruud Koopmans, “Dynamics of Repression and Mobilization: The German Extreme Right in the 1990s,” *Mobilization: An International Quarterly*, Vol. 2, No. 2, 1997, pp. 149–164; Jessica Beyer and Jennifer Earl, “Backfire Online: Studying Reactions to the Repression of Internet Activism,” in Lester Kurtz and Lee Smiley, eds., *The Paradox of Repression and Nonviolent Movements*, Syracuse University Press, 2018, pp. 102–142。

<sup>⑥</sup> Anita Breuer, Todd Landman and Dorothea Farquhar, “Social Media and Protest Mobilization: Evidence from the Tunisian Revolution,” *Democratization*, Vol. 22, No. 4, 2015, pp. 764–792; Navid Hassanpour, “Media Disruption and Revolutionary Unrest: Evidence from Mubarak’s Quasi-Experiment,” *Political Communication*, Vol. 31, No. 1, 2014, pp. 1–24.

同时，关于信息治理措施的最终作用方向及其复杂的影响路径尚待更加全面和深入的理论探讨与实证研究来加以澄清和验证。现实中，尽管政府采取了相似的信息治理措施，但反对派在面对政府信息治理措施时，其应对方式存在明显差异，有时保持克制，有时则发起抗争并引发激烈冲突。由此引出一个关键问题：反对派对政府信息治理措施的应对策略为何存在差异？信息治理措施如何影响国内冲突的爆发？

对此，既有解释框架主要聚焦于信息治理措施对国内冲突的直接效应，主要体现在两方面：一是信息治理措施能通过削弱反抗势力的集体行动能力有效降低国内冲突爆发的风险；二是信息治理措施可能会引发“逆火效应”，直接加剧国内冲突爆发的可能性。然而，这一框架忽视了信息治理措施作为一种政府传达的信号，能够间接地通过改变反对派对政府实力和决心的认知与评估来塑造其后续的抗争策略。深入剖析这一间接效应具有重要意义，因为在信息不对称的现实政治环境中，反对派往往难以直接洞察政府的真实实力，需要依据政府的某些行为进行揣测。理论上，作为政府调节信息流动的战略性行动，信息治理本身便是政府能力的重要载体。<sup>①</sup> 然而，关于反对派如何解读由信息治理措施传递出政府实力的信号，并据此调整自身应对决策的过程，学界对此议题并未进行充分的理论探讨。

为此，本文构建了一个以信息治理措施为核心的信号博弈模型，重点关注的是政府与反对派间复杂的互动关系，并进一步阐释这种动态交互如何影响国内冲突的爆发。有别于聚焦直接效应的传统视角，本文将信息治理措施视为一种传达政府实力与战略意图的信号，并通过博弈论探究这一机制如何影响反对派的认知与抗争决策。基于对参数条件与均衡结果的系统分析，本文揭示了完美信号、虚张声势、互相试探和信号失效四种可能的均衡状态，并剖析了每种均衡的形成条件及其背后的政治意涵。本文指出，成本考量与政府实力是决定治理效果的关键。成本的增加会降低政府实施信息治理措施的可能性，而随着政府实力的削弱，信息治理措施作为一种威慑信号，其功效会大打折扣，难以有效塑造政府强大的形象，反而可能加剧国内冲突爆发的风险。为检验上述理论，本文构建了一个跨国面板数据，在全球层面捕捉信息治理措施对国内冲突爆发的动态影响。统计结果显示，治理成本越高，政府采取信息治理措施的可能性越低。作为旨在抑制国内冲突的一种潜在手段，信

<sup>①</sup> Sandra Braman, *Change of State: Information, Policy, and Power*, The MIT Press, 2006, pp. 12–20; Joseph S. Nye and Jack Landman Goldsmith, “The Future of Power,” *Bulletin of the American Academy of Arts and Sciences*, Vol. 64, No. 3, 2011, pp. 45–52.

息治理措施的有效性严格受限于政府实力是否强大；在政府实力较弱的国家，该措施非但不能遏制国内抗争与冲突，反而增加了冲突爆发的可能性。

## 二、既有研究回顾

数字技术通过赋权社会和赋能国家的双重路径，以前所未有的方式深刻重塑了国家治理的权力边界以及国家与社会的关系和互动模式。<sup>①</sup> 一方面，数字平台作为新兴媒介，极大拓宽了民众表达意愿和组织集体行动的渠道。互联网的普及和社交媒体的兴起不仅明显降低了民众政治参与的门槛与成本，还显著增强了反对派的组织效率与动员能力。这种技术赋权的现象在全球范围内日益增长的冲突事件中得到了鲜明体现。<sup>②</sup> 另一方面，技术进步为政府治理尤其是信息治理带来了革命性的变化。政府可以利用舆情分析等更丰富的形式开展信息领域的深度治理，以提高治理效能和稳定社会秩序。<sup>③</sup> 学界普遍认为，围绕信息流动展开的国家—社会互动不仅是衡量现代政府治理成效的重要标尺，还深刻影响着国家安全和社会稳定。<sup>④</sup> 然而，目前学界对信息治理能否有效降低国内冲突爆发以及哪些因素会影响其缓解冲突的效果等方面仍有争论。

首先，学界在信息治理措施能否有效缓和国内冲突爆发这一问题上存在明显分歧。一些研究指出，政府实施的信息治理措施能够抑制反对派的暴力抗争倾向，并明显降低国内冲突爆发的风险。这些措施增强了政府精准识别反对派领导层与核心成员的身份、位置和活动计划等关键信息的能力，同时也提高了情报获取效率，从

<sup>①</sup> 孟天广、郑思尧：《国家治理的信息理论：信息政治学的理论视角》，载《政治学研究》，2023年第6期，第117—130页；黄其松：《数字时代的国家理论》，载《中国社会科学》，2022年第10期，第60—77页。

<sup>②</sup> Marco Manacorda and Andrea Tesei, “Liberation Technology: Mobile Phones and Political Mobilization in Africa,” *Econometrica*, Vol. 88, No. 2, 2020, pp. 533–567; Ruben Enikolopov, et al., “Social Media and Protest Participation: Evidence from Russia,” *Econometrica*, Vol. 88, No. 4, 2020, pp. 1479–1514; Jan H. Pierskalla and Florian M. Hollenbach, “Technology and Collective Action: The Effect of Cell Phone Coverage on Political Violence in Africa,” *American Political Science Review*, Vol. 107, No. 2, 2013, pp. 207–224.

<sup>③</sup> Jennifer Earl, Thomas V. Maher and Jennifer Pan, “The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review,” DOI: 10.1126/sciadv.abl8198; Dorota Głowacka, et al., “Digital Technologies as a Means of Repression and Social Control,” pp. 14–31; Erica Frantz, Andrea Kendall-Taylor and Joseph Wright, “Digital Repression in Autocracies,” *Varieties of Democracy Institute Users Working Paper*, No. 27, 2020, pp. 1–22.

<sup>④</sup> 高奇琦、隋晓周：《数字政治学的未来图景与现实展开》，载《政治学评论》，2023年第1期，第65—87页；李舒：《互联网极端风险防范与大国博弈》，载《同济大学学报（社会科学版）》，2022年第4期，第48—57页；韩娜、邹初好：《国家安全视域下信息操纵的数字运行逻辑及风险应对》，载《公安学研究》，2023年第4期，第1—20页。

而能够在抗争活动尚未全面展开时实施精准的预防性打击，有效遏制抗争与冲突的爆发。<sup>①</sup> 威廉·马克恰克（William R. Marczak）等基于2010—2013年的网络数据分析发现，巴林与阿拉伯联合酋长国等中东国家采取一系列多元的信息治理措施，成功实现了对国内反叛组织的精准定位与高效打击，显著削弱了其活动能力和影响力，有效遏制了抗争活动在短期内的扩散，降低了冲突升级的风险。<sup>②</sup> 此外，从集体行动理论的资源动员角度，研究者发现信息治理措施可通过切断潜在参与者的信息获取渠道，削弱反对派对局势的评估以及对行动策略的共识构建，从根本上阻碍抗争活动的有效组织与动员。例如，对中东和非洲一些国家的研究表明，加强对互联网和社交媒体的监管能够有效削弱反对派的在线动员能力，迫使其转向传统的线下沟通方式，从而显著延缓国内冲突在空间上的扩散速度。<sup>③</sup>

还有一些研究则指出，信息治理措施可能会进一步激化政府与民众之间的矛盾并加剧国内冲突爆发的风险。如纳维德·哈桑普尔（Navid Hassanpour）认为，埃及政府切断互联网和手机信号的做法不仅没有阻止反对派的政治动员，反而激化了民众情绪，在全国范围内引发了大量的冲突事件，最终成为推翻政府的关键力量。<sup>④</sup> 除经验证据外，还有学者从规范角度对信息治理措施的正当性提出质疑，指出过度限制信息流通从长远角度也不利于维护社会秩序与政权稳定。<sup>⑤</sup> 这些研究发现信息治理措施对缓解冲突存在反效果的可能，相关措施虽然出于维护稳定的目的，但很有可能会不受控地成为国内冲突的“催化剂”和“导火索”。

其次，既有研究涉及政府信息治理措施能否有效缓解国内冲突爆发的争论主要聚焦于三个层面。第一，不同的治理措施及其力度对冲突产生的影响大相径庭。珍

<sup>①</sup> Jacob N. Shapiro and Nils B. Weidmann, “Is the Phone Mightier Than the Sword? Cellphones and Insurgent Violence in Iraq,” *International Organization*, Vol. 69, No. 2, 2015, pp. 247–274.

<sup>②</sup> William R. Marczak, et al., “When Governments Hack Opponents: A Look at Actors and Technology,” *Proceedings of the 23rd USENIX Security Symposium*, 2014, pp. 511–525.

<sup>③</sup> Lisa Garbe, “Pulling Through Elections by Pulling the Plug: Internet Disruptions and Electoral Violence in Uganda,” *Journal of Peace Research*, Vol. 61, No. 5, 2024, pp. 842–857; Babak Rahimi, “The Agonistic Social Media: Cyberspace in the Formation of Dissent and Consolidation of State Power in Postelection Iran,” *The Communication Review*, Vol. 14, No. 3, 2011, pp. 158–178.

<sup>④</sup> Navid Hassanpour, “Media Disruption and Revolutionary Unrest: Evidence from Mubarak’s Quasi-Experiment,” pp. 1–24.

<sup>⑤</sup> Tiberiu Dragu and Yonatan Lupu, “Digital Authoritarianism and the Future of Human Rights,” *International Organization*, Vol. 75, No. 4, 2021, pp. 991–1017; Adrienne LeBas and Lauren E. Young, “Repression and Dissent in Moments of Uncertainty: Panel Data Evidence from Zimbabwe,” *American Political Science Review*, Vol. 118, No. 2, 2024, pp. 584–601; Jessica Beyer and Jennifer Earl, “Backfire Online: Studying Reactions to the Repression of Internet Activism,” pp. 102–142.

妮弗·厄尔（Jennifer Earl）和杰茜卡·拜尔（Jessica Beyer）指出，过于高调且突如其来信息治理措施可能促使数字领域内活跃的个人与社群将线上的不满情绪转化为线下的抗议活动，进而加剧国内冲突爆发的风险。<sup>①</sup>当信息治理措施走向极端，严重扰乱民众正常生活秩序时（如全面中断互联网与移动通信服务），可能会引发更为激烈的冲突反应。有学者发现，2016—2018年，印度地方政府为预防冲突采取了互联网中断措施，反而导致暴力冲突事件的显著增加。<sup>②</sup>值得注意的是，限制访问特定社交媒体平台、在局部区域内暂时中断互联网或限制网速等较为温和的信息治理措施对冲突的影响在不同国家间表现出明显的差异。<sup>③</sup>因此，信息治理措施在可见性、持续性和强度上的差异深刻影响着反对派和民众的应对策略及冲突爆发的可能性。

第二，信息治理措施在抑制冲突效果上的异质性主要体现在其对不同社会群体的差异化影响上。相同的治理措施对不同的群体可能会产生迥异的影响，促使他们采取不同的冲突策略。例如，直接针对反对派的治理措施虽然可以暂时压制其抗争活动，但可能会激发旁观者的同情与支持。<sup>④</sup>因此，需要更加细致地区分信息治理措施对反对派和普通民众产生的不同影响，进而探究这些影响如何作用于冲突的爆发与后续发展。

第三，信息治理措施在抑制冲突方面的成效深受多重情境因素的影响。尼尔斯·韦德曼（Nils B. Weidmann）和埃斯彭·勒德（Espen G. Rød）的研究发现，尽管信息治理措施能够显著降低冲突爆发的风险，但若在冲突爆发后才采取此类措施，则可能适得其反，出现冲突激化的态势。<sup>⑤</sup>还有研究者指出，互联网中断后初期往往伴随着抗议活动和暴力冲突的急剧上升，但随着时间推移，数字通信和协调渠道

<sup>①</sup> Jennifer Earl and Jessica Beyer, "The Dynamics of Backlash Online: Anonymous and the Battle for WikiLeaks," *Intersectionality and Social Change*, Vol. 37, 2014, pp. 207–233.

<sup>②</sup> Jan Andrzej Rydzak, "A Total Eclipse of the Net: The Dynamics of Network Shutdowns and Collective Action Responses," Ph. D. Dissertation, The University of Arizona, 2018, pp. 123–170.

<sup>③</sup> Lisa Garbe, "Pulling Through Elections by Pulling the Plug: Internet Disruptions and Electoral Violence in Uganda," pp. 842–857; Philip N. Howard, et al., "Opening Closed Regimes: What Was the Role of Social Media During the Arab Spring?" The Project on Information Technology and Political Islam, Working Paper, January 2011; Philip N. Howard and Muzammil M. Hussain, *Democracy's Fourth Wave? Digital Media and the Arab Spring*, Oxford University Press, 2013, pp. 11–12.

<sup>④</sup> Jennifer Pan and Alexandra A. Siegel, "How Saudi Crackdowns Fail to Silence Online Dissent," *American Political Science Review*, Vol. 114, No. 1, 2020, pp. 109–125; Adrienne LeBas and Lauren E. Young, "Repression and Dissent in Moments of Uncertainty: Panel Data Evidence from Zimbabwe," pp. 584–601.

<sup>⑤</sup> Nils B. Weidmann and Espen G. Rød, *The Internet and Political Protest in Autocracies*, Oxford University Press, 2019.

的瘫痪能够遏制冲突蔓延。<sup>①</sup> 哈桑普尔则强调，在人口密集的城市区域，即便通信媒介被切断，人们仍能较为便利地通过面对面交流维持一定的沟通效率，这为反对派提供了便捷的动员途径。<sup>②</sup> 相反，在偏远地区，由于替代通信方式的匮乏与低效，破坏通信基础设施可能会有效地降低冲突爆发的可能性。<sup>③</sup> 由此可见，信息治理措施的实施时机、持续时间以及地理位置等因素均会影响冲突的效果。

综上，不同研究因在样本选择和分析层次上的差异而得出相悖的结论。既有研究大多聚焦于单一国家的个案分析，且将对象放在那些处于“民主转型”动荡期、冲突频发的国家作为研究蓝本的选择性偏见的问题。现实中，在一些国家能够有效抑制冲突爆发的信息治理措施在另一些国家可能会适得其反。为更全面深刻地理解信息治理措施对冲突爆发的影响，我们亟须从更宏观的理论框架出发并辅以扎实的实证支持。然而，受限于数据可得性以及可靠的信息治理测量方法，跨国比较研究仍显稀缺且普遍存在选择性偏见，未能全面覆盖所有国家与地区。这些研究往往忽视了西式民主国家的情况，从而导致其实证结果具有明显的局限性。<sup>④</sup>

此外，学界对于信息治理措施如何影响国内冲突爆发的既有分析视角存在静态和单向化的局限性。主流观点往往将信息治理措施视为政府单方面施加于社会的技术工具，强调其对反对派集体行动能力和抗争意愿的直接影响，却忽视了这一手段同样关键的间接效应：信息治理措施作为政府能力的一种信号，能够影响反对派对政府实力的认知与评估，进而间接塑造其抗争策略。信息治理措施之所以能够成为反映政府实力的信号，原因在于其顺利实施离不开技术、设备、人力资源和制度框架的全方位支撑，<sup>⑤</sup> 且政府在实施信息治理措施后，还需应对其可能带来的负面影响，包括对社会经济生活造成的扰乱、<sup>⑥</sup> 民众对政府信任度与满意度的下降<sup>⑦</sup>以及来

<sup>①</sup> Jan Andrzej Rydzak, “A Total Eclipse of the Net: The Dynamics of Network Shutdowns and Collective Action Responses,” pp. 85–120.

<sup>②</sup> Navid Hassanzpour, “Media Disruption and Revolutionary Unrest: Evidence from Mubarak’s Quasi-Experiment,” pp. 1–24.

<sup>③</sup> Margaret Coker and Charles Levinson, “Rebels Hijack Gadhafi’s Phone Network,” *The Wall Street Journal*, April 13, 2011.

<sup>④</sup> Erica Frantz, Andrea Kendall-Taylor and Joseph Wright, “Digital Repression in Autocracies,” pp. 1–22; Nils B. Weidmann and Espen G. Rød, *The Internet and Political Protest in Autocracies*, 2019.

<sup>⑤</sup> Jennifer Earl, Thomas V. Maher and Jennifer Pan, “The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review,” DOI: 10.1126/sciadv.abl8198.

<sup>⑥</sup> Anirudh Tagat, Amreesh Phokeer and Hanna M. Kreitem, “Net Loss: An Econometric Method to Measure the Impact of Internet Shutdowns,” *ACM Journal on Computing and Sustainable Societies*, 2024, DOI: 10.1145/3659466.

<sup>⑦</sup> Jennifer Pan and Alexandra A. Siegel, “How Saudi Crackdowns Fail to Silence Online Dissent,” pp. 109–125.

自国际社会和人权组织的谴责等。<sup>①</sup>正是鉴于这些不可忽视的成本与代价，政府能够选择并成功执行此类信息治理措施，实际上是对其治理能力与决心的有力证明。<sup>②</sup>信息治理措施不仅彰显了政府对信息领域进行监管的强烈意愿，也更深层次地反映了其在该领域内施加干预和影响的实质性能力。然而，既有研究对反对派如何解读政府释放的信号并据此调整其行动策略尚未展开充分的理论讨论，仅仅简单地将信息治理措施视为政府单向影响社会的工具而忽视了政府与反对派间复杂的博弈全貌。事实上，反对派并不总会消极接受信息治理措施，而是能够积极主动地根据政府释放的信号调整对局势的理解与应对策略，展现了极高的策略灵活性与环境适应性。但既有文献鲜有探讨信息治理措施如何改变反对派对政府实力的认知及其对后续抗争决策和冲突行为的影响。考察信息治理效应的既有实证研究也常常忽视政府的实际实力与其通过信息治理措施刻意塑造和展示出来的实力之间存在一定差距。<sup>③</sup>

### 三、信息治理与国内冲突的信号博弈模型

为深入探究信息治理措施如何塑造反对派对政府实力的认知并进一步评估这一认知对国内冲突风险的影响，本文基于不完全信息博弈的理论框架，结合既有研究对信号机制的见解，构建了信息治理—国内冲突的博弈模型。<sup>④</sup>该模型围绕政府与反对派之间的动态关系展开，尤其关注政府的信息治理措施与反对派的抗争策略之间的相互作用，旨在清晰展现两者间复杂的互动模式及其可能产生的后果。为更准确地聚焦于信息治理措施中信号传递的复杂过程，模型对复杂的现实世界进行了必要的简化和抽象处理。

<sup>①</sup> Tiberiu Dragu and Yonatan Lupu, "Digital Authoritarianism and the Future of Human Rights," pp. 991–1017.

<sup>②</sup> Kristin Eichhorn and Eric Linhart, "Election – Related Internet – Shutdowns in Autocracies and Hybrid Regimes," *Journal of Elections, Public Opinion and Parties*, Vol. 33, No. 4, 2023, pp. 705–725; Tiberiu Dragu and Yonatan Lupu, "Digital Authoritarianism and the Future of Human Rights," pp. 991–1017.

<sup>③</sup> Espen G. Rød and Nils B. Weidmann, "Empowering Activists or Autocrats? The Internet in Authoritarian Regimes," *Journal of Peace Research*, Vol. 52, No. 3, 2015, pp. 338–351.

<sup>④</sup> 尹继武、贾宇娴：《双向蔑视、非优策略选择与危机僵局》，载《世界经济与政治》，2024年第6期，第69—105页；常晓燕：《施压信号、沟通渠道与国际危机管控失败》，载《世界经济与政治》，2022年第9期，第4—34页；祁昊天：《不确定性视角下的美国战略威慑与信号博弈》，载《世界经济与政治》，2021年第7期，第48—77页；漆海霞：《崛起信号、战略信誉与遏制战争》，载《国际政治科学》，2020年第4期，第1—37页；杨原、曹玮：《核边缘、信号博弈与小国的“自我孤立”悖论》，载《当代亚太》，2018年第6期，第69—105页；Tiberiu Dragu and Yonatan Lupu, "Digital Authoritarianism and the Future of Human Rights," pp. 991–1017。

## (一) 模型设定

在本文博弈模型中存在政府  $G$  与反对派  $O$  两个参与方。政府拥有一项实力指标  $\theta$ , 可据此分为强政府  $S$  与弱政府  $W$  两类 (即  $\theta \in \{S, W\}$ ) , 用以评估其在面临挑战时实施有效社会控制及维持政治秩序的能力。具体而言, 当遭遇反对派时, 强政府被挑战成功的概率为  $\pi_s$ , 相应地, 其维持统治的概率为  $1-\pi_s$ ; 而弱政府被挑战成功的概率为  $\pi_w$ , 维持统治的概率为  $1-\pi_w$ 。其中  $\pi_s < \pi_w$ , 这意味着挑战强政府成功的概率低于挑战弱政府。换言之, 强政府在面临挑战时更可能维持其统治, 我们简单假定反对派挑战成功的收益为 1, 挑战失败收益为 0。

需要强调的是, 政府的能力类型并非公开信息,  $N$  决定政府的类型, 这一信息仅政府知晓, 反对派认为政府类型是强政府  $S$  的先验概率为  $p$ , 是弱政府  $W$  的先验概率为  $1-p$ 。政府首先选择实施信息治理措施  $IG$  或不实施信息治理措施  $\sim IG$ , 为体现出政府和反对派对信息治理措施的偏好差别, 我们假设政府实施相关措施的政策收益为  $U_G^{Policy}(IG) = X$ , 不实施相关措施则享受不到额外的信息优势, 即  $U_G^{Policy}(\sim IG) = 0$ 。由于信息治理措施会加剧反对派相较于政府的信息劣势, 因此反对派的政策收益为  $U_O^{Policy}(IG) = -X$ ; 在没有任何信息治理措施的情况下, 反对派没有额外的信息损失, 即  $U_O^{Policy}(\sim IG) = 0$ 。当政策收益  $X$  增大时, 政府和反对派关于信息治理措施的偏好差距会越来越大。此外, 政府的能力能够影响信息治理成本。本文假设两种类型的政府实施信息治理措施的成本  $c$  分别为  $c_s$  和  $c_w$ , 由于相较于弱政府, 强政府有更好的基础条件 (如软硬件设备、技术和人员等) 和制度保障来支持信息治理措施, 因而实施相关措施的成本更低, 即  $c_s < c_w$ 。<sup>①</sup> 进一步考虑信息治理措施的成本与其给政府带来的信息收益之间的关系, 为使信息治理措施成为政府社会控制能力强的可信信号, 我们假设  $c_s < X < c_w$ 。

反对派在观察到政府信息治理措施后选择发起抗争  $R$  或不发起抗争  $\sim R$ 。当其选择不反抗  $\sim R$  时, 博弈自动结束, 现状得以维持; 当其选择发起抗争  $R$  时, 则爆发冲突。一方面, 信息治理措施本身具有政策收益  $X$ , 即给政府带来更多信息优势; 另一方面, 信息治理措施能在抗争爆发时通过直接干涉信息传播的内容和渠道来增加组织集体行动的难度, 从而减小抗争行动成功的概率。由此, 本文进一步假设, 信息治理措施  $IG$  会在反对派选择反抗  $R$  时令其获胜的概率减小  $\alpha$ 。因此, 当存在信息治理措施  $IG$  时, 反对派在面对强政府与弱政府时反抗成功的概率分别为  $\pi_s - \alpha$  和

<sup>①</sup> 为了简化, 本文只聚焦于政府选择是否实施信息治理措施, 并不关注相关措施的严格程度, 实际上可能存在多种类型的信息治理措施。

$\pi_w - \alpha$ 。如果反对派发起抗争的成本为  $k$ , 为使其抗争决策符合理性选择, 我们假设  $\pi_s < k < \pi_w$ ,  $\pi_s < \alpha < \pi_w$  以及  $\pi_s < \alpha + k < \pi_w$ 。<sup>①</sup>

根据上述模型的描述, 政府和反对派各自都有四种纯策略 (pure strategy) 可供选择。<sup>②</sup> 政府可基于其自身实力和具体形势决定是否实施信息治理措施  $IG$ 。具体来说, 政府的策略集由四种策略  $S$  组成:  $S_G = \{(IG | S; IG | W), (IG | S; \sim IG | W), (\sim IG | S; IG | W), (\sim IG | S; \sim IG | W)\}$ 。在这里,  $(A | S; B | W)$  表示强政府  $S$  会采取  $A$  行动, 弱政府  $W$  则会采取  $B$  行动。同样地, 反对派也有四种纯策略来回应政府的行动:  $S_O = \{(R | IG; R | \sim IG), (R | IG; \sim R | \sim IG), (\sim R | IG; R | \sim IG), (\sim R | IG; \sim R | \sim IG)\}$ 。这个策略集展示了反对派在面对政府不同的信息治理措施时的复杂决策过程, 也揭示了他们在追求目标时所面临的权衡和考量。<sup>③</sup> 值得注意的是, 除纯策略外, 双方还可以采用混合策略 (mix strategy), 即每个局中人都能以概率  $q$  采取任一种行动, 从而使情况复杂化, 后文会进一步讨论这种情况。在解这个模型时, 本文寻求每个参与者在给定其他参与者信念和策略的基础上各自寻求最优化的策略这一均衡状态。本文假设反对派认为, 实施信息治理措施  $IG$  的政府为强政府  $S$  这一信念用条件概率  $Pr(S | IG)$  表示。这个模型有四种均衡情况, 分别代表四种政府与反对派之间以信息治理措施为信号的互动模式。

## (二) 均衡情况讨论

命题 1: 如果  $c_w \geq \pi_w + X$ , 以下策略和信念组成一个完美贝叶斯均衡, 即  $S^* = \{(IG | S; \sim IG | W), (\sim R | IG; R | \sim IG)\}$ ,  $Pr(S | IG) = 1$ ,  $Pr(S | \sim IG) = 0$ 。在这个均衡中, 反对派可通过观察是否存在信息治理措施完全获悉政府的实力。实施信息治理措施的就是强政府, 反之就是弱政府; 反对派只在没有信息治

① 详见图 1, 有两点问题值得说明: (1) 在不存在信息治理措施时, 若  $k > \pi_w$ , 反对派反抗的成本大于挑战弱政府的预期收益, 那么他们永远不会选择反抗; 若  $k < \pi_s$ , 即反对派反抗的成本小于反抗强政府带来的预期收益, 那么他们总是会选择反抗。(2) 当存在信息治理措施时, 同理可知, 若  $\alpha + k > \pi_w$ , 反对派总是会选择不反抗; 若  $\alpha + k < \pi_s$ , 反对派则总是会选择反抗。为使反对派的抗争决策符合理性选择的预设, 我们假定  $\pi_s < k < \pi_w$  和  $\pi_s < \alpha + k < \pi_w$ , 并由此得出  $\pi_s < \alpha < \pi_w$ 。囿于篇幅, 如需要进一步的附录信息, 请联系作者获取。

② 纯策略是指在某些条件下一定会采取某一行动。当  $(R | IG)$  是一个纯策略时, 是指反对派在面临政府信息治理措施时一定会选择反抗; 而当其是一个混合策略时, 则代表反对派在政府实施信息治理措施时  $q$  反抗存在的概率。

③ 需要注意的是, 这些策略的选择可能还受反对派内部的意见分歧、外部支持者的态度以及政府对外部环境的反应等其他因素影响。在实际情况中, 反对派的行动可能会更加复杂多变, 本文的博弈模型则注重探讨政府的信息治理措施如何影响反对派是否发起冲突抗争的决策。

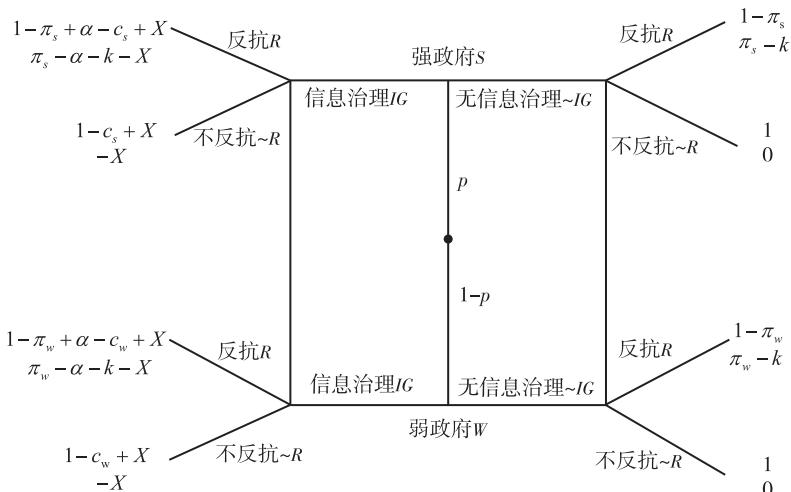


图 1 信息治理措施的信号博弈

资料来源：笔者自制。

理措施时选择反抗，而在观察到存在信息治理措施时选择不反抗（即只向弱政府发起抗争）。

我们称这是一个“完美信号”的均衡，其核心特质在于政府的行为完全透明地反映了其实力状况，没有故作声势或刻意掩藏。在此均衡中，政府的信息治理措施成为一个完美的信号传递机制：当实力强时，政府倾向于强化信息治理措施，当实力较弱时，政府放任信息自由流通。此均衡仅当信息治理措施对弱政府而言成本过高（即满足  $c_w > \pi_w + X$ ）时才成立，因此其通常不会采取此策略。这种高昂的成本门槛使得反对派能够合理推断，有能力实施信息治理措施的政府必然实力雄厚，因而在面对挑战时能够进行强有力的回应。相反，若政府允许信息自由流通，则被视为实力不足，可能激发反对派的抗争意愿。值得注意的是，在此均衡下的政府行为保持了高度的透明性和真实性，构成了一个清晰无误的信号。尽管这种直接明了的做法在复杂的博弈环境中较为罕见，但它正是信息治理措施高成本背景下的独特产物。然而，一旦信息治理成本降低至弱政府能够承受的水平，它们可能会尝试实施信息治理措施，这将削弱政府实力信号的清晰度，增加反对派评估政府实际实力的难度，进而左右其抗争策略的制定，最终导致这一均衡状态的瓦解。

**命题 2：**当  $c_w < \pi_w + X$ ,  $c_w \geq X + \alpha$  和  $p < \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$  时，以下策略和信念组成了一个完

美贝叶斯均衡，即强政府总是实施信息治理措施，弱政府实施信息治理措施的概率

为  $q = \frac{p}{(1-p)} \frac{(k+\alpha-\pi_s)}{(\pi_w-k-\alpha)}$ ; 反对派在观察到没有信息治理措施时总是选择反抗，在观察到存在信息治理措施时选择反抗的概率为  $r = \frac{\pi_w - c_w + X}{\pi_w - \alpha}$ 。在这个均衡中，反对派在观察到存在信息治理措施时相信此时政府实力强的概率为  $Pr(S \mid IG) = \frac{p}{p+q(1-p)}$ 。

我们称这是一个“相互试探”的均衡，其维持的先决条件是弱政府信息的治理成本较低 ( $c_w < \pi_w + X$ ) 但仍需跨越一个阈值 ( $c_w \geq X + \alpha$ )，且强政府出现的概率较小 ( $p < \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$ )。在此均衡状态下，弱政府有时会以实施信息治理措施的方式营造强势假象，试图误导反对派相信政府实力雄厚，进而遏制其潜在的反抗意愿。对于反对派而言，在观察到政府实施信息治理措施后，他们会有选择地进行反抗。换言之，当弱政府能以较低的成本实施信息治理措施并可能被反对派识破其虚张声势时，双方均倾向于采取一种灵活多变的混合策略进行试探。具体而言，弱政府会根据时局权衡是否值得冒险展现出强硬姿态、伪装成实力强的政府以维持威慑；而反对派也会权衡利弊，评估是否值得冒险通过反抗来揭露政府表面的强硬之下可能隐藏的虚弱本质。

**命题 3：**如果  $c_w < \pi_w + X$ ,  $p \geq \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$ , 以下策略和信念组成了一个完美贝叶斯均衡： $S^* = \{ (IG \mid S; IG \mid W), (\sim R \mid IG; R \mid \sim IG) \}$ ,  $Pr(S \mid IG) = p$ , 以及  $Pr(S \mid \sim IG) \leq \frac{\pi_w - k}{\pi_w - \pi_s}$ 。在这个均衡中，当政府实施信息治理措施时，反对派认为此时政府实力强的概率为  $Pr(S \mid IG) = p$ 。政府无论实力强弱，总是会实施信息治理措施；反对派在观察到政府实施相关措施时选择不反抗，在观察到没有相关措施时则选择反抗。

我们称这是一个“虚张声势”的均衡，其存在的前提是弱政府能够以足够低的成本 ( $c_w < \pi_w + X$ ) 实施信息治理措施，并且此时政府实力强的概率较大 ( $p \geq \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$ )。鉴于当前情境下强政府出现的可能性较大，弱政府若采取强硬姿态模仿强政府的行为，往往能有效震慑反对派，使其不敢轻举妄动。基于反对派的这种认知与行为模式，政府坚信反对派在观察到存在信息治理措施时，无论其真实实力如

何，反对派都会选择不抗争。因此，政府倾向于持续采取强硬的姿态实施信息治理措施，以维持其强势形象。然而，一旦政府有所放松不实施信息治理措施，反对派便会敏锐捕捉机会发起反击。为规避这一潜在风险，对于弱政府而言，即便需要付出一定代价来维持表面的强势，这依然一个更为稳妥且优选的策略。

**命题 4：**如果  $c_w < \alpha + X$ ,  $p < \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$ , 以下策略和信念组成了一个完美贝叶斯均衡,

即  $S^* = (IG \mid S; IG \mid W), (R \mid IG; R \mid \sim IG)$ ,  $Pr(S \mid IG) = p$ ,  $Pr(S \mid \sim IG) \leq \frac{\pi_w - k}{\pi_w - \pi_s}$ 。在这个均衡中，当政府实施信息治理措施时，反对派相信此时政府实力强的概率  $Pr(S \mid IG) = p$ 。政府无论实力强弱，总是实施信息治理措施；反对派无论是否观察到政府实施相关措施，均会选择反抗。

我们把这个均衡称为“信号失效”，意指信息治理措施不再能够作为有效传达政府实力强弱的可靠信号。在此情境下，弱政府能以极低的成本 ( $c_w < \alpha + X$ ) 实施信息治理措施，它便更加倾向于频繁诉诸此类手段。这种低成本方式逐渐演化为政府一种日常的治理手段，从而不再构成一个能有效体现政府实力强的可靠信号。因此，反对派不再单纯依据信息治理措施的实施情况来评估政府的实际能力；信息治理措施的威慑力也大打折扣，不再能直接遏制其反抗意图，反对派因而可能倾向于寻求其他途径来评估政府的真正实力。鉴于此时强政府出现的概率  $p < \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$  本身就比较低，因此无论政府是否采取信息治理措施，反对派均倾向于选择反抗，以应对其感知到的政府脆弱性。

以上均衡结果揭示出信息治理措施影响国内抗争和冲突风险的作用路径，这一过程受信息治理成本和政府实力两个要素的制约（见表 1）。其中，治理成本的高低直接决定了弱政府伪装成强政府的难度，强政府出现的可能性则影响着反对派对政府实际实力的评估。在完美信号均衡中，高昂的治理成本成为塑造可靠威慑信号的关键，使反对派将信息治理措施视为强大政府的象征而放弃反抗意图。在虚张声势均衡中，较低的治理成本使弱政府倾向于高调实施相关措施以震慑反对派。而信号失效反映出，若信息治理措施被当作低成本治理工具频繁使用，其作为威慑信号的效力会被削弱，从而激起更多反抗。互相试探均衡则表明，当治理成本处于“中间地带”时，政府实力的不确定性会使反对派和政府的策略更为复杂，体现出反复试探与动态调试的特点，增强了双方行为的难以预测性。

表1 四种均衡状态的特征描述

均衡类型	信息治理成本	政府类型为强政府的可能性	政府是否实施信息治理措施	反对派回应
完美信号	最高	—	否	选择反抗
互相试探	较低	低	有可能实施	有可能选择反抗
虚张声势	较低	高	是	选择不反抗
信号失效	最低	低	是	选择反抗

资料来源：笔者自制。

注：表中展示了弱政府的信息治理措施以及反对派基于相关措施的回应。四种均衡状态下强政府的信息治理措施及反对派回应情况囿于篇幅未能呈现，如需相关信息请联系笔者获取。

### (三) 治理成本与政府实力对信息治理措施与国内冲突的影响

鉴于强政府实施信息治理措施的成本始终低于其带来的效益，强政府往往倾向于实施相关措施。在强政府行动策略比较稳定的情况下，本文重点关注弱政府策略的动态演变。弱政府是否实施信息治理措施主要受限于治理成本，这一因素在图2中的横轴上得以体现；图2纵轴则表示政府类型为强政府的概率。这一二维空间中的成本—实力差异精准捕捉到了影响政府和反对派之间潜在冲突的两个因素：一是对政府实力的主观评估，二是弱政府实施信息治理措施的成本。具体而言，随着成本的攀升，弱政府实施信息治理措施的意愿减弱；相应地，当反对派观察到政府没有实施信息治理措施的迹象时，其反抗意愿会增强。反之，如果政府展现出实力强的迹象，其实施信息治理措施的动力就会更为强劲，从而有效抑制反对派的反抗意愿。

图2还揭示了四种均衡背后的含义。在图2的右侧区域，信息治理成本很高，甚至超越了政府不干预信息流动、任由反对派抗争的预期效用，导致完美信号的均衡成为主导。然而，当信息治理成本降至 $\pi_w + X$ 这一关键阈值之下时，情况就会发生变化：在图2左上角，仅有虚张声势这一均衡；而在图2左下方，则并存信号失效与互相试探两种均衡。在互相试探均衡中，弱政府有时会实施信息治理措施，其决策存在一定不确定性。与之相比，在另外两种均衡中，弱政府则始终选择实施信息治理措施，尽管其背后的动机各不相同。在虚张声势均衡中，由于弱政府被误认为强政府的概率较大，它便通过实施信息治理措施模仿强政府的行为，从而使反对派因惧怕“强政府”的高压措施而不敢轻易反抗。相反，在信号失效均衡中，弱政府总是选择实施信息治理措施，是因为此时治理成本足够低，甚至低于其放任信息流动时的预期效用。换言之，当信息治理成本高于 $\pi_w + X$ 这一阈值时，对弱政府而

言代价太大，此时其宁愿面对反对派的抗争也不愿意实施信息治理措施。相反，如果成本低于该值，弱政府虽需付出一定代价，但尚在可接受的范围内；假如政府是强政府的概率较大 ( $p > \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$ )，弱政府始终会模仿强政府实施信息治理措施以震慑反对派；假如政府是强政府的概率较小 ( $p < \frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$ )，但治理成本极低 ( $c_w < \alpha + X$ ) 时，即便存在被识破的风险，弱政府仍会实施信息治理措施以最大化其效用；当信息治理成本较低但仍有一定门槛时 ( $\alpha + X \leq c_w < \pi_w + X$ )，是否实施信息治理措施对弱政府而言预期效用相当，因此它们时而实施治理措施时而采取放任策略。

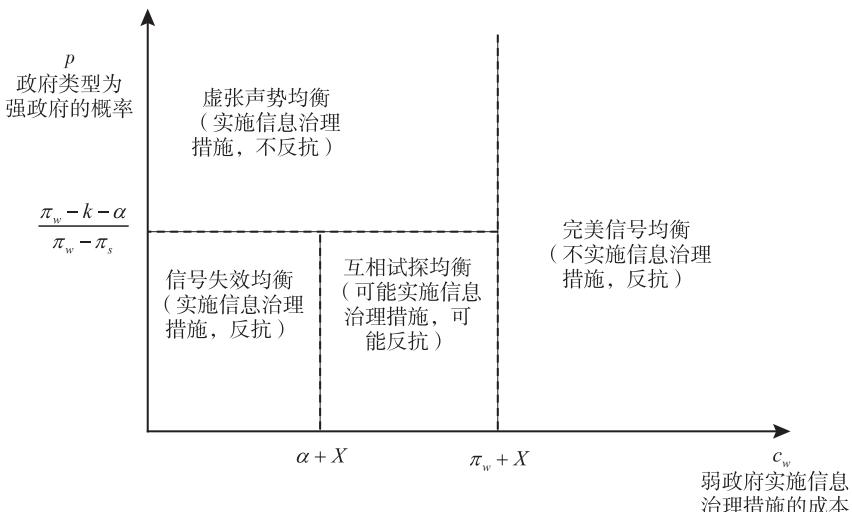


图 2 基于信息治理成本与政府实力的均衡预测

资料来源：笔者自制。

为验证以上理论发现的现实影响力，本文提出了可用实证证据检验的研究假说。两个核心假说分别对应政府实施信息治理措施的决策逻辑以及反对派据此做出抗争决策的行为机理。

我们首先关注成本对政府信息治理措施的影响。通过分析由高至低四种不同成本水平的均衡，可以清晰地发现这一事实：随着信息治理成本的下降，政府更倾向于实施信息治理措施；相反，随着治理成本的上升，政府实施信息治理措施的可能性逐渐降低。本文据此提出假说 1。

假说 1：信息治理成本的增加会降低政府实施信息治理措施的可能性。

需要指出，该假说非常直观且符合常识预期，并得到了大量研究结果的支持，<sup>①</sup>因此我们将主要聚焦于第二个发现，即政府的信息治理措施如何通过影响反对派对政府实力的认知来影响其抗争决策。通过观察图2的y轴（代表政府类型为强政府的概率 $p$ ），可以发现在左侧的三个均衡状态下，随着 $p$ 值的增加，即强政府的可能性变大时（特别是当 $p$ 值超过阈值 $\frac{\pi_w - k - \alpha}{\pi_w - \pi_s}$ 时），政府实施的信息治理措施遭到反抗的可能性会降低。而在完美信号均衡中，反对派通过观察政府是否采取信息治理措施，就能准确获知政府的实力类型并据此制定抗争策略。具体而言，若政府实施信息治理措施，则被视为强政府（ $p=1$ ），反对派选择不反抗；若未实施，则被视为弱政府（ $p=0$ ），反对派选择反抗。因此，在完美信号均衡中，政府实力对反对派抗争决策的影响与其他三个均衡类似，即政府实力增强会降低其信息治理措施遭遇反抗的可能性，只不过在完美信号均衡中， $p$ 值表现为二元而非连续。本文据此提出假说2。

**假说2：**当政府的实力较弱时，政府实施的信息治理措施更有可能加剧反对派的抗争，增加国内冲突的数量；当政府实力强时，政府实施的信息治理措施更有可能削弱反对派的抗争，降低国内冲突的数量。

#### 四、实证统计分析

前文的博弈模型揭示了信息治理措施在抑制国内冲突爆发方面的作用，主要受信息治理成本以及政府实力的影响。治理成本决定了政府实施信息管控措施的可能性，政府实力则会调节其对国内冲突爆发的影响。下文将通过实证研究检验该理论框架的解释力。

##### （一）研究设计

###### 1. 数据和因变量

本文基于全球148个国家和地区2016年1月至2020年12月的月度面板数据，对信息治理措施如何影响国内冲突进行了实证考察。选择进行全球范围的跨国研究主要基于两点考量：其一，鉴于近年来全球范围内冲突事件频发且呈上升态势，探究提升信息治理在预防国内冲突方面的作用显得尤为迫切和重要。区别于冲突爆

<sup>①</sup> Tiberiu Dragu and Yonatan Lupu, “Digital Authoritarianism and the Future of Human Rights,” pp. 991–1017; Anita R. Gohdes, “Repression Technology: Internet Accessibility and State Violence,” *American Journal of Political Science*, Vol. 64, No. 3, 2020, pp. 488–503.

后政府应对（responsive）措施效果评估的既有研究，本文着重关注信息治理措施作为预防性（preventive）策略的实际效果，<sup>①</sup> 旨在通过宏观层面的跨国分析描绘出信息治理措施在预防冲突中的基本作用轮廓。其二，纳入全球范围内的国家作为研究样本，能够更全面地揭示国家间的差异性以及特定情境因素对信息治理效果的影响。这种跨国视角不仅能丰富了我们对信息治理措施作用机制的理解，还能促进了对信息治理在维护社会稳定和预防冲突方面潜在效能的深入思考。

本文借助“互联网中断跟踪优化项目（STOP）”<sup>②</sup> 和“民主多样性数据集（V-Dem）”<sup>③</sup> 两个数据库对假说1的因变量信息治理措施进行测量。其中STOP数据集提供了2016—2023年全球互联网关停事件的数据，包含每次事件的日期、地点、持续时间和影响范围等。V-Dem数据集则覆盖了1993—2023年各国政府对传统媒体和互联网的监管情况。通过整合两个数据集，本文得到2016—2023年各国政府信息治理措施的面板数据。

假说2主要聚焦于探讨信息治理措施对国内冲突爆发频率的影响，即分析其是否促使国内冲突抗争事件的增多或减少。作为这一分析的关键指标，本文将国内冲突抗争事件的发生次数作为因变量。为确保数据的权威性和全面性，我们依托“武装冲突地点与事件数据项目（ACLED）”<sup>④</sup> 和乌普萨拉冲突数据项目（UCDP）两大数据库，它们广泛收集了各类社会与政治冲突事件，并进行了系统的时间与地点编码。为更细致地剖析信息治理措施对不同类型冲突的具体作用并增强研究发现的稳健性，本文细化了因变量的测量方式，将冲突抗争事件按照暴力程度由低至高划分为抗议（protest）、暴动骚乱（riot）和武装冲突（armed conflict）等三种类型，<sup>⑤</sup> 并在后续的分析模型中分别就这些不同类型的冲突抗争事件进行讨论，以期更为全面地理解信息治理措施在不同冲突情境下的实际效果。

## 2. 自变量与控制变量

假说1中的自变量是政府实施信息治理措施的成本，由于目前缺乏直接反映该

<sup>①</sup> Tiberiu Dragu and Adam Przeworski, “Preventive Repression: Two Types of Moral Hazard,” *American Political Science Review*, Vol. 113, No. 1, 2019, pp. 77–87; Nathan Danneman and Emily Hencken Ritter, “Contagious Rebellion and Preemptive Repression,” *Journal of Conflict Resolution*, Vol. 58, No. 2, 2014, pp. 254–279.

<sup>②</sup> 互联网中断跟踪优化项目由全球数字权利保护组织Access Now与#KeepitOn Coalition共同提供，具体链接参见 <https://www.accessnow.org/issue/internet-shutdowns/>。

<sup>③</sup> 参见“The V-Dem Dataset,” <https://www.v-dem.net/data/>。

<sup>④</sup> Clionadh Raleigh, et al., “Introducing ACLED: An Armed Conflict Location and Event Dataset,” *Journal of Peace Research*, Vol. 47, No. 5, 2010, pp. 651–660.

<sup>⑤</sup> 谢艳、周羿：《社会冲突的跨国扩散效应——基于1995—2012年冲突事件数据的分析》，载《世界经济与政治》，2022年第9期，第74—98页。

成本的系统可靠的跨国数据，本文借鉴克里斯廷·艾希霍恩（Kristin Eichhorn）和埃里克·林哈特（Eric Linhart）的做法，采用服务业占国内生产总值（GDP）的比重作为衡量治理成本的一个间接指标。<sup>①</sup> 理论上，由于服务业更加依赖于信息的高效流通与交换，服务业占GDP比重越高的国家，信息治理措施对经济造成的潜在负面影响可能越大。<sup>②</sup>

假说2的自变量为政府实施的信息治理措施。本文选择互联网关停、媒体内容治理和互联网内容治理三个指标代表政府的信息治理措施。其中互联网关停的数据来自STOP数据集，我们主要构建了两个变量：一是互联网关停状态的哑变量，0代表某国在某月未出现互联网切断现象，1代表发生了互联网关停；二是互联网关停次数的连续变量，即统计某国在某个月内发生的互联网关停事件总数。此外，本文也考虑了政府对媒体与互联网信息内容的治理强度，相关数据来自V-Dem数据集，我们将原始的连续变量转换成了一个哑变量：若政府对传统媒体与互联网信息内容进行常规化监管，表明其治理力度更严格，则该变量被编码为1；反之，若政府的监管行为仅限于在应对特别敏感事件时偶尔使用，则此变量被赋值为0，以示其较小的治理力度。

假说2主要关注政府实力的差异如何调节信息治理措施在国内冲突中的作用，并预测当政府实力较弱时其信息治理措施将增加国内冲突的数量。为验证这一假说，本文采用脆弱国家指数（FSI）来衡量政府实力。<sup>③</sup> 该指数评估了国家凝聚力、经济繁荣、政治稳定和社会发展状况，分数越高意味着国家在这些方面表现越糟糕且面临更高的国家失败（state failure）风险，因而是评估政府实力的理想工具。将政权脆弱程度与信息治理措施的交互作用纳入模型，我们能直接考察政权脆弱性如何调节政府信息治理措施对国内冲突的作用。

本文在既有文献基础上选取并控制一系列关键变量。通过军事人员数量与军事支出的对数来把握国家军事能力；同时，利用人口总数、人均GDP的对数以及互联网普及率分别控制了人口规模、经济和信息技术发展水平。上述变量均来自联合国世界发展指标（WDI）数据库。<sup>④</sup> 此外，政治制度可能会同时影响信息治理措施和

<sup>①</sup> Kristin Eichhorn and Eric Linhart, "Election – Related Internet – Shutdowns in Autocracies and Hybrid Regimes," pp. 705–725.

<sup>②</sup> Kristin Eichhorn and Eric Linhart, "Election – Related Internet – Shutdowns in Autocracies and Hybrid Regimes," pp. 705–725.

<sup>③</sup> 参见“Fragile State Index,” <https://fragilestatesindex.org/>。

<sup>④</sup> 参见“World Development Indicators,” <https://databank.worldbank.org/source/world-development-indicators>。

冲突的爆发，本文也根据 V-Dem 数据库设置了相关变量。鉴于本文的主要核心因变量（互联网关停和不同类型的冲突）是以日期为单位的数据，而部分自变量（如媒体和网络内容治理的强度、脆弱国家指数）以及所有控制变量均为年度数据，为融合这两类数据并尽可能地捕捉因变量的变化，本文采用国家一月为分析单位。同时，考虑到前一时间段的信息治理措施和冲突水平对当前时段可能存在的滞后影响，我们将因变量的一阶滞后项纳入控制变量，即控制  $t-1$  月的信息治理措施和冲突事件数量。

### 3. 模型选择

假说 1 选取互联网是否关停、互联网关停次数、媒体内容治理力度和互联网内容治理力度作为因变量，除互联网关停次数是计数变量外，其余三个指标均为哑变量。因此，我们对哑变量使用逻辑回归模型（logistic regression），对计数变量使用负二项回归模型（negative binomial regression）。假说 2 选取抗议、骚乱和武装冲突的次数作为因变量，这些均为计数变量。由表 2 可知，它们的均值均小于方差，表明数据分布存在过离散（overdispersion）现象，不满足线性回归模型对正态分布的基本假定，因此本文对假说 2 部分负二项回归模型分析。作为稳健性检验，本文也使用了其他模型对其进行估计。

表 2 变量的描述性统计

变量	观测值	均值	标准差	最小值	最大值
抗议次数	5697	59.749	171.441	0	3441
骚乱次数	5697	9.682	29.764	0	561
武装冲突次数	5697	0.016	0.251	0	8
关停互联网次数	5697	0.083	0.788	0	17
关停互联网	5697	0.028	0.164	0	1
媒体内容治理力度	5697	0.335	0.472	0	1
互联网内容治理力度	5697	0.371	0.483	0	1
脆弱国家指数	5697	76.887	20.807	14.627	114
互联网普及率	5502	48.266	28.492	1.880	99.702
国家人口总数（对数）	5697	16.670	1.473	11.497	21.068
国家人均 GDP（对数）	5589	24.735	1.799	20.917	30.611
国家政治制度	5697	0.476	0.223	0.014	0.915

续表 2

变量	观测值	均值	标准差	最小值	最大值
军事支出（对数）	5697	20.213	3.192	0	27.381
军队人员数量（对数）	5697	10.827	1.805	0.000	14.937
服务业占 GDP 比重	5529	52.189	9.655	30.493	87.421

资料来源：笔者自制。

需要明确指出的是，本文的数据尤其是控制变量存在一定缺失现象。鉴于当前广泛使用的多重差补法和最大似然法等处理缺失数据的策略均基于数据随机缺失的假定，而这一假定在本研究中并不成立。具体而言，一些国家和地区的数据缺失并非随机产生，而是源于它们在数据收集与统计能力上的局限性。因此，为避免采用基于随机缺假设的方法可能引入的额外估计偏差，本文采取表列删除（listwise deletion）的方式来处理缺失数据，在不满足随机缺失假定时，这一方法在处理缺失数据方面处理缺失数据最为有效和稳妥的方式。<sup>①</sup> 最终，本文的原始数据覆盖 2016—2020 年的 148 个国家和地区，有效观察数为 5502。

## （二）实证分析结果

根据本文的理论预期，治理成本越高，政府实施信息治理措施的可能性越低；当政府实力较弱即脆弱国家指数较高时，实施信息治理措施会促使国内冲突事件数量的增加。

### 1. 主要模型结果

#### （1）成本对政府信息治理措施的影响（假说 1）

表 3 展示了针对假说 1 的实证检验结果。模型 1、模型 3 和模型 4 采用了逻辑回归模型，其因变量分别对应互联网是否关停、媒体内容治理力度以及互联网内容治理力度；模型 2 则运用了负二项回归模型，聚焦于互联网关停次数。模型 1 的结果支持了假说 1，在控制了其他变量的情况下，服务行业占 GDP 的比重越高，该国关停互联网的可能性就越低，这一结果在 99% 的置信区间显著。模型 2 进一步巩固了假说 1 的可靠性，发现当服务行业占 GDP 的比重上升时，该国互联网关停的次数也明显更少。模型 3 和模型 4 考察了成本对媒体和互联网内容治理力度

<sup>①</sup> Paul D. Allison, *Missing Data*, Sage Publications, 2001, p. 6; Thomas B. Pepinsky, "A Note on Listwise Deletion Versus Multiple Imputation," *Political Analysis*, Vol. 26, No. 4, 2018, pp. 480–488; 苏若林：《国内危机烈度与冲突对象选择》，载《世界经济与政治》，2021 年第 8 期，第 84—105 页。

的影响。结果显示，服务行业占 GDP 的比重对媒体内容治理的力度并无显著影响。然而，当服务行业占 GDP 的比重上升时，该国倾向于采取更为宽松的互联网内容治理措施的可能性显著增加。这可能是由于服务业企业越来越依赖于互联网进行市场推广、业务运营和客户服务，而传统媒体对服务业的影响程度相对较低。<sup>①</sup> 综上，实证结果支持了假说 1，即信息治理的成本与政府采取信息治理措施的可能性呈负相关。换言之，信息治理成本越高，政府采取此类措施的可能性越低。

表 3 成本对信息治理措施的影响

解释变量	模型 1	模型 2	模型 3	模型 4
	互联网关停	互联网关停次数	媒体内容治理力度	互联网内容治理力度
服务行业占 GDP 比重	-0.069 (0.017) ***	-0.071 (0.017) ***	-0.007 (0.006)	-0.019 (0.005) ***
军队人员数量（对数）	0.349 (0.122) ***	0.334 (0.123) ***	-0.165 (0.076) **	-0.522 (0.079) ***
军事支出占 GDP 比重	0.940 (0.172) ***	1.057 (0.180) ***	0.324 (0.089) ***	1.235 (0.094) ***
国家人口总数（对数）	1.021 (0.275) ***	1.705 (0.304) ***	0.968 (0.134) ***	1.020 (0.169) ***
国家人均 GDP（对数）	-1.159 (0.280) ***	-1.759 (0.294) ***	-0.599 (0.132) ***	-1.217 (0.170) ***
国家政体类型	0.826 (0.578)	0.840 (0.555)	-19.140 (0.760) ***	-13.152 (0.491) ***
互联网普及率	-0.015 (0.008)*	0.001 (0.007)	0.011 (0.005) **	0.003 (0.005)
脆弱国家指数	0.002 (0.013)	-0.005 (0.013)	-0.047 (0.006) ***	-0.007 (0.005)

<sup>①</sup> James Manyika and Charles Roxburgh, "The Great Transformer: The Impact of the Internet on Economic Growth and Prosperity," McKinsey Global Institute Report, 2011, pp. 1–2.

续表 3

解释变量	模型 1	模型 2	模型 3	模型 4
	互联网关停	互联网关停次数	媒体内容治理力度	互联网内容治理力度
常数项	-12.938 (2.897) ***	-12.092 (2.854) ***	4.149 (1.384) ***	-0.524 (1.017)
lnalpha		-0.189 (0.297)		
样本量	5145	5145	5145	5145

资料来源：笔者自制。

注：\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ 。括号内为稳健标准误。

## (2) 政府实力如何调节信息治理措施对国内冲突爆发的影响（假说 2）

表 4 展示了针对假说 2 进行验证的结果，其中因变量为抗议次数。所有模型均控制了前一个月该国的冲突水平。模型 5 的核心自变量为互联网关停状况，其回归系数为负，且在 99% 的置信区间显著，说明在控制其他因素后，互联网关停在短期内降低了抗议的发生频率。然而，我们更关注的是政府实力如何调节信息治理对冲突抗争的影响，即具体的信息治理措施与脆弱国家指数的交互作用。我们发现这一交互项的回归系数为正，且在 99% 的置信区间显著。这意味着在同等条件下，当国家能力弱时关停互联网不但不能有效抑制抗议活动，反而显著增加了抗议活动的数量。这一结果表明，信息治理措施对国内冲突的作用受政府实力差异的影响，弱政府的信息治理措施适得其反，可能会激发更多抗争和冲突，从而验证了假说 2。

为确保上述发现并非由自变量的测量误差导致，我们替换了模型 5 中的核心解释变量（互联网关停的哑变量）。模型 6、模型 7 和模型 8 分别采用互联网关停次数、对传统媒体及互联网的内容治理力度作为自变量的替换测量。结果显示，这三个新变量与脆弱国家指数的交互项均呈正相关，且除了模型 7 的交互项系数在 95% 的置信区间显著外，其他的均在 99% 置信区间显著。这一发现巩固了模型 5 结论的稳健性，表明在政府实力薄弱的背景下，无论选择何种形式的信息治理措施，均会加剧冲突抗争态势，并导致抗议次数的显著增加。

基于模型 5 和模型 7 的结果，本文分别预测了互联网关停以及政府对媒体的内容治理如何影响国内冲突水平，并聚焦政府实力差异对此影响的调节效果，具体结

果如图 3 所示。在图 3a 中，基于保持其他条件不变的前提，模型预测当政府实力强、脆弱国家指数低于 90 时，关停互联网较之于保持互联网畅通，能显著减少抗议活动的次数。反之，当政府实力较弱、脆弱国家指数攀升至 90 及以上时，关停互联网的举措不但无法抑制抗议，反而显著增加了抗议活动的数量，并且这种加剧效应随着政府实力的削弱（即脆弱国家指数的上升）而进一步增强。这意味着作为一种潜在的抑制国内冲突的手段，切断互联网的有效性严格受限于政府实力的强大和国家政权的稳固。然而，在国家能力薄弱且面临国家崩溃高风险的环境下，切断互联网非但不能平息冲突，反而会显著加剧国内的抗争活动和冲突态势。图 3b 进一步揭示了媒体内容治理力度与抗议活动之间的类似趋势：当政府实力较弱、脆弱国家指数大约超过 80 时，高强度的媒体内容治理不仅无法抑制抗议，反而会促使抗议活动数量显著增加，且这种增加的趋势随着政府实力的削弱而加剧。综合这两点可以发现，信息治理措施（包括切断互联网和高强度的媒体内容治理）在抑制国内冲突方面的有效性很大程度依赖于政府实力的强大与政权的稳固。一旦政府实力薄弱，面临国家失败的风险，这些信息治理措施不但起不到预防冲突的作用，反而会显著加剧国内的抗争活动和冲突水平。这一结论为假说 2 提供了有力支持。

表 4 负二项回归模型结果：信息治理措施对国内冲突的影响（因变量：抗议次数）

解释变量	模型 5	模型 6	模型 7	模型 8
抗议次数（上个月）	0.006 (0.000) ***	0.006 (0.000) ***	0.006 (0.000) ***	0.006 (0.000) ***
脆弱国家指数	0.020 (0.002) ***	0.019 (0.002) ***	0.019 (0.002) ***	0.019 (0.002) ***
互联网关停	-2.637 (0.684) **			
互联网关停 * 脆弱国家指数	0.030 (0.008) ***			
互联网关停次数		-1.765 (0.319) ***		
互联网关停次数 * 脆弱国家指数		0.020 (0.004) ***		

续表 4

解释变量	模型 5	模型 6	模型 7	模型 8
媒体内容治理力度			-0.648 (0.245)***	
媒体内容治理力度 * 脆弱国家指数			0.006 (0.003)**	
互联网内容治理力度				-0.922 (0.247)***
互联网内容治理力度 * 脆弱国家指数				0.010 (0.003)***
军队人员数量 (对数)	0.282 (0.019)***	0.276 (0.019)***	0.285 (0.019)***	0.281 (0.019)***
军事支出占 GDP 比重	-0.043 (0.004)***	-0.043 (0.004)***	-0.044 (0.004)***	-0.043 (0.004)***
国家人口总数 (对数)	0.126 (0.044)***	0.144 (0.044)***	0.118 (0.044)***	0.115 (0.044)***
国家人均 GDP (对数)	0.034 (0.040)	0.019 (0.040)	0.046 (0.041)	0.043 (0.041)
国家政体类型	0.857 (0.123)***	0.841 (0.121)***	0.468 (0.162)***	0.685 (0.143)***
互联网普及率	0.032 (0.002)***	0.031 (0.002)***	0.032 (0.002)***	0.033 (0.002)***
常数项	-5.909 (0.489)***	-5.684 (0.491)***	-5.781 (0.514)***	-5.868 (0.503)***
样本量	5169	5169	5169	5169
Wald chi2 (df)	4459.73 (10)	4491.29 (10)	4389.65 (10)	4313.97 (10)
Prob>chi2	0.000	0.000	0.000	0.000
Pseudo R <sup>2</sup>	0.1257	0.1272	0.1255	0.1256

资料来源：笔者自制。

注： \* p&lt;0.1； \*\* p&lt;0.05； \*\*\* p&lt;0.01。括号内为稳健标准误。

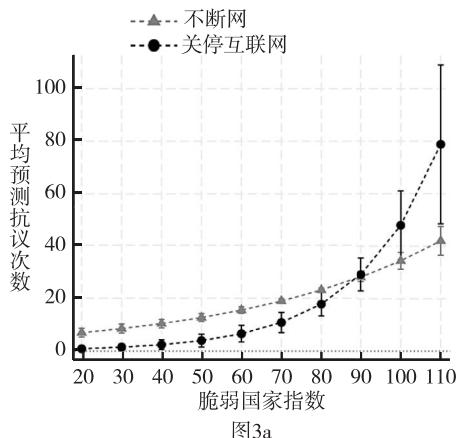


图3a

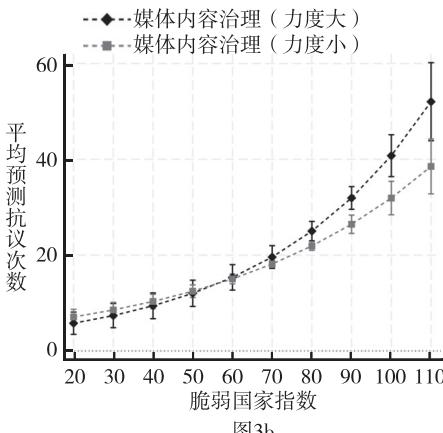


图3b

图3 变量的预测边际效应

资料来源：笔者自制。

注：图中圆形或三角形代表估测系数，线段为95%的置信区间。图3a根据模型5展示了政府实力在互联网关停对抗议次数影响方面的调节效用，图3b根据模型7展示了政府实力在媒体内容治理力度对抗议次数影响方面的调节作用。

本文使用骚乱与武装冲突次数作为因变量的替换测量，同时沿用与表4完全相同的自变量和控制变量，重新估计了表4中的所有模型。<sup>①</sup>为了直观展示关键发现，图4描绘了这些模型核心交互项的回归系数，其中每个点代表回归系数，线段为95%置信区间，虚线为零线。若置信区间与零线相交，则表示估测结果在0.05层次统计上不显著；若置信区间完全位于零线一侧，则表明在0.05层次统计上显著。估测结果显示，除图4b中的最后一个模型外，所有核心自变量与脆弱国家指数交互项的回归系数均呈显著正相关（95%置信区间）。这一发现揭示了在弱政权的环境下，政府实施的信息治理措施会显著增加该国的骚乱与武装冲突次数。此结论与表4的结果一致，因而可以认为核心交互项的效应并不是虚假相关的影响。

## 2. 稳健性检验与内生性讨论

为确保本文核心发现的稳健性，我们还进行了一系列稳健性检验。

第一，在探讨信息治理措施和冲突爆发之间的复杂关系时，我们需要警惕逆向因果和选择性偏差的可能。具体而言，政府实施的信息治理措施可能并非反对派抗争的原因，而是其在冲突爆发后的应对策略。因此存在一种竞争性解释：抗争和冲

<sup>①</sup> 如需要具体结果，请联系作者获取。

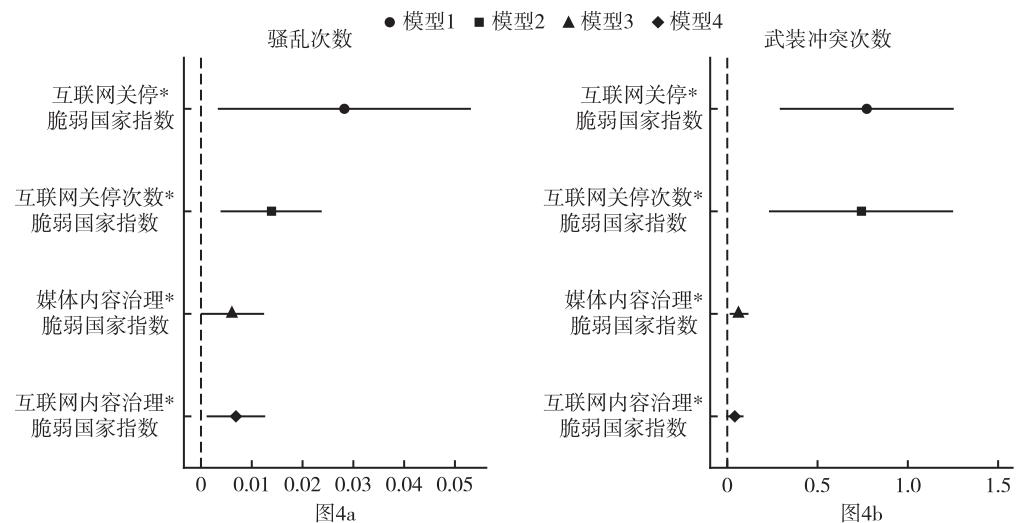


图4 政府实力在不同信息治理措施情况下对国内冲突水平影响的调节效用分析

资料来源：笔者自制。

注：图中点代表估测系数，线段为95%的置信区间，图4a体现了政府实力在不同信息治理措施（互联网关停、媒体和互联网内容治理力度）对骚乱次数影响方面的调节效用；图4b则反映了政府实力在不同信息治理措施对武装冲突次数影响方面的调节作用。

突事件促使政府采取更为严格的信息治理措施以平息矛盾、缓解紧张局势。特别是冲突频发的弱政府更倾向于加强信息治理措施以抑制冲突升级。为此，我们在所有模型中均控制了前一个月的冲突水平。此外，为明确冲突与信息治理措施的时间先后顺序，我们在所有模型中均采用了前一个月的自变量数据，以确保冲突发生在信息治理措施之后。结果显示，前一个月的信息治理措施与政府实力交互项的回归系数为正，且在99%的置信区间显著，这进一步排除了虚假相关的疑虑，增强了交互项解释力的可信度。<sup>①</sup>

我们进一步使用双变量probit模型来探讨可能存在的选择性偏差问题。由于该模型的两阶段方程均要求因变量为哑变量，我们对部分变量进行了重新编码处理。表5展示了该模型估计的结果。在模型1至模型3中，第一阶段的因变量均为是否发生了互联网关停事件，第二阶段的因变量则分别为是否发生了抗议、骚乱和武装

<sup>①</sup> 我们还通过格兰杰因果检验来考察信息治理措施和国内冲突的时间关系，结果表明并没有证据支持信息治理措施是国内冲突的格兰杰原因。需要注意的是，格兰杰因果关系并非两个变量之间真正的因果关系，只能提供一定的参考价值。

冲突事件。模型 1 与模型 2 的结果进一步支持了本文的两个理论假说，但模型 3 的结果并不显著。值得注意的是，尽管这三个模型中两个方程的相关性均为正，但在 90% 的置信区间不显著，这意味着两个方程之间的相关性并不显著，因此可以分别运行两个独立的模型。这实际上与本文主模型的处理方式相吻合，即分别对假说 1 和假说 2 进行回归分析，进一步加强了我们对主模型结果可靠性的信心。

表 5 双变量 Probit 模型结果

解释变量	模型 1		模型 2		模型 3	
	第一阶段	第二阶段	第一阶段	第二阶段	第一阶段	第二阶段
	因变量是否发生断网	因变量是否发生抗议	因变量是否发生断网	因变量是否发生骚乱	因变量是否发生断网	因变量是否发生暴力冲突
第三产业占 GDP 比重	-0.080 (0.031) ***		-0.082 (0.031) ***		-0.080 (0.032) **	
冲突水平（上个月）		1.045 (0.077) ***		1.110 (0.045) ***		0.654 (0.354) *
互联网关停		1.326 (1.187)		-47.546 (13.118) ***		-5.293 (1.114) ***
脆弱国家指数	0.018 (0.020)	0.024 (0.004) ***	0.017 (0.020)	0.020 (0.002) ***	0.020 (0.020)	0.007 (0.007)
互联网关停 * 脆弱国家指数		0.029 (0.014) **		0.636 (0.171) ***		0.014 (0.012)
军队人员数量 (对数)	0.378 (0.244)	0.598 (0.052) ***	0.405 (0.247)	-0.040 (0.026)	0.359 (0.255)	0.020 (0.088)
军事支出占 GDP 比重	0.769 (0.317) **	-0.297 (0.078) ***	0.768 (0.320) **	-0.029 (0.009) ***	0.789 (0.322) **	0.223 (0.098) **
国家人口总数 (对数)	1.283 (0.308) ***	-0.332 (0.095) ***	1.261 (0.306) ***	0.343 (0.054) ***	1.297 (0.309) ***	0.088 (0.168)
国家人均 GDP (对数)	-1.232 (0.282) ***	0.286 (0.118) **	-1.236 (0.278) ***	-0.060 (0.046)	-1.238 (0.283) ***	-0.157 (0.170)
政治制度	1.823 (0.886) **	2.683 (0.254) ***	1.799 (0.885) **	1.298 (0.129) ***	1.873 (0.913) **	-0.437 (0.405)
互联网普及率	0.013 (0.008)	0.013 (0.003) ***	0.012 (0.008)	0.011 (0.002) ***	0.013 (0.008) *	-0.008 (0.006)
常数	-14.653 (4.699) ***	-4.756 (0.893) ***	-14.308 (4.524) ***	-5.986 (0.536) ***	-15.171 (4.729) ***	-5.075 (1.110) ***

续表 5

解释变量	模型 1		模型 2		模型 3	
	第一阶段	第二阶段	第一阶段	第二阶段	第一阶段	第二阶段
因变量是否发生断网	因变量是否发生抗议	因变量是否发生断网	因变量是否发生骚乱	因变量是否发生断网	因变量是否发生暴力冲突	
athrho		0.089 (0.170)		0.155 (0.222)		0.274 (0.220)
样本量	5145	5145	5145			

注： \* p<0.1； \*\* p<0.05； \*\*\* p<0.01。括号内为稳健标准误。

第二，考虑到负二项回归模型的局限性，我们采取了多元化的模型策略以进行更为精准的估计。假说 2 因变量中零值占据相当大的比重，<sup>①</sup> 表现出显著的“零膨胀”特征。针对这一特性，我们选用了零膨胀负二项回归（zero-inflated negative binomial regression）模型，将数据中零值的产生和冲突抗争事件的出现视为两个独立的过程，分开进行建模分析。其结果与负二项回归模型的结果非常接近，因此进一步支持了假说 2。<sup>②</sup>

第三，我们还对假说 2 的因变量进行了对数变换处理，改变其分布形态，使其能够利用广义线性模型（GLM）框架下的泊松拟最大似然估计值（QMLE）。“长期以来，人们已经认识到泊松分布假设对因变量的条件矩（conditional moments）施加了限制，而这些限制尤其是条件方差与均值相等这一限制在实际应用中常常被违反。”<sup>③</sup> 杰弗里·伍德里奇（Jeffrey M. Wooldridge）指出，泊松回归能在任何方差均值比（variance-mean ratio）条件下产生一致的均值参数估计。这一对方差均值之间关系的宽松假设，被称为“泊松广义线性模型方差假设”，并广泛存在于广义线性模型的应用中。<sup>④</sup> 泊松广义线性模型的结果与本文主模型结果接近，其回归系数的估计值比负二项回归模型的结果更小，标准误也更小，进一步支持了假说 2。<sup>⑤</sup>

综上，既有实证研究通常指出信息治理措施对国内冲突抗争具有双重作用——

① 暴力冲突、骚乱和抗议次数中零值分别占比 99%、29% 和 9%。

② 如需要具体结果，请联系作者获取。

③ Jeffrey M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, 2010, pp. 502–517.

④ Jeffrey M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, pp. 502–517.

⑤ 如需要具体结果，请联系作者获取。

既可能加剧冲突，也可能抑制冲突发生。<sup>①</sup> 本文则基于信号博弈模型分析了信息治理措施对国内冲突爆发影响的方向和大小如何受到政府实力强弱的调节。具体而言，当政府实力强的可能性大（脆弱国家指数低）时，信息治理措施往往会被反对派视为政府强势的信号和象征，从而抑制了它们发起抗争的意愿；相反，当政府实力强的可能性较小（脆弱国家指数高）时，反对派在观察到存在信息治理措施后，评估政府实力强且能有效镇压其抗争的可能性较低，因此更倾向于发起反抗。这一发现揭示了信息治理措施与国内冲突之间关系的复杂性，并强调了政府实力在其中的关键调节作用。

## 五、结论

本文基于 2016—2020 年覆盖 148 个国家和地区的月度面板数据分析了治理成本与政府实力如何影响信息治理措施的效果，并特别强调了政府实力在这一过程中的调节作用。研究发现，随着政府实力的削弱，信息治理措施在遏制冲突方面的效果不断下降；尤为重要的是，在政权脆弱的情境下，此类措施不但不能缓和冲突态势，反而会在短期内触发更多的抗争活动，从而加剧国内冲突的风险。需要注意的是，受限于数据可得性和本文的研究设计，我们只估计了这种短期影响，因此上述结果与相关研究发现的信息治理措施的长期影响并不矛盾。<sup>②</sup> 此外，由于本文采用的是单阶信号博弈模型，该模型的一个固有局限是没有系统地将历史博弈行动信息纳入考量。因此，本文的实证结果主要揭示了近期治理措施的影响。若要探究更长期的历史互动如何塑造政府的信息治理措施以及反对派的回应，有必要在后续研究中构建一个多阶博弈模型。

本研究具有一定的理论贡献和现实意义。在理论层面，既有关于信息治理措施如何影响国内冲突的研究往往局限于个别案例或特定类型国家尤其是非西式民主国家，过于强调其对言论自由的负面影响，却忽视了信息治理作为一种政府治理手段的普遍性以及政府实力在这一过程中的核心调节作用。<sup>③</sup> 本文通过构建信号

<sup>①</sup> Jennifer Earl, Thomas V. Maher and Jennifer Pan, “The Digital Repression of Social Movements, Protest, and Activism: A Synthetic Review,” DOI: 10.1126/sciadv.abl8198.

<sup>②</sup> Nils B. Weidmann and Espen G. Rød, *The Internet and Political Protest in Autocracies*, 2019; Jan Andrzej Rydzak, “A Total Eclipse of the Net: The Dynamics of Network Shutdowns and Collective Action Responses,” pp. 123–170.

<sup>③</sup> Erica Frantz, Andrea Kendall-Taylor and Joseph Wright, “Digital Repression in Autocracies,” pp. 1–22.

博弈模型，揭示了政府实力是决定国内反对派是否发起抗争的关键因素，并在实证研究中将样本扩展至全球国家和地区。研究发现，信息治理措施对国内冲突爆发的影响并非单一的抑制或激化，而是取决于政府实力的强弱：在政府实力强的背景下，信息治理措施被视为政府实力雄厚且有能力应对反对派抗争活动的信号，进而抑制反对派的抗争活动；相反，当政府实力较弱时，此类治理措施可能引发“逆火效应”，加剧冲突。未来的研究还可以进一步细化，从微观或中观层面探索影响信息治理效果的具体条件。此外，既有研究多聚焦于信息治理措施直接干预信息传播以阻碍抗争活动的机制（即信号模型中的参数 $\alpha$ ），而本研究在此基础上着重探讨了信息治理措施作为政府实力和态度信号的间接作用。这意味着信息治理措施不仅直接加剧了信息不对称并提高了组织集体行动的成本，还能作为一种威慑手段，间接影响反对派的抗争决策，从而丰富了对信息治理复杂影响机制的认识。

这一研究发现的现实意义主要体现在两方面：首先，本文通过探讨政府实力在信息治理措施对国内冲突影响中的调节作用，明确指出决定信息治理措施能否有效遏制冲突的关键因素在于政府实力与政权稳固性，而非政治制度的简单划分。这一发现挑战了部分西方学者关于信息治理措施在非西式民主国家中必然产生负面影响的片面观点，<sup>①</sup> 揭示了即便是在西式民主国家中，若政权根基不稳、国家能力脆弱，信息治理措施也可能成为触发更为激烈社会抗争和冲突的催化剂。因此，与政治制度价值层面的评判相比，更为迫切和务实的议题应是增强国家的综合实力和政权的稳定性。其次，本研究为审视信息治理措施在国家治理中的作用提供了新视角。与以往静态分析信息治理效果的研究不同，本文构建了一个涵盖政府、反对派和信息治理措施的信号博弈模型。这一模型揭示了信息治理措施的效果是高度动态且复杂的，会随着双方行为偏好的变化以及对彼此策略的认知深化而不断演变。可见，一味地增加或减小信息治理措施的力度并不能实现预防冲突、维护社会稳定的最优效果。政府需要综合考虑多种情境因素，深入理解反对派的偏好与行动策略并据此灵活调整信息治理措施。

（截稿：2024年12月 责任编辑：郭泉）

<sup>①</sup> Tiberiu Dragu and Yonatan Lupu, “Digital Authoritarianism and the Future of Human Rights,” pp. 991–1017; Andrea Kendall-Taylor, Erica Frantz and Joseph Wright, “The Digital Dictators: How Technology Strengthens Autocracy,” *Foreign Affairs*, Vol. 99, No. 2, 2020, p. 103.