

极化现象与临界点

摘要

研究表明，党派分歧和极端主义立场的加剧在政治精英中比普通选民中更为显著。当前研究的焦点逐渐转向如何缓解这种极化现象。本文使用了一个通用的意见变化模型，探讨了影响力和同类性（homophily）所带来的自我强化动态是否会表现出使极化难以逆转的“临界点”。该模型适用于立法机构或其他小型、紧密连接的组织，并未假设特定国家的制度安排，以避免掩盖极化相变中的基本规律性。在模型中，个体初始随机分布于一个多维议题空间，该空间包括隶属于两个等规模政党之一的党派身份以及在 10 个议题上的立场。随后，个体根据其对分歧的容忍度及党派认同强度相对于其意识形态承诺的权衡，调整自己的议题立场，向与自己观点相近的邻居靠拢，远离观点相左者。研究通过计算实验操控了个体对分歧的容忍度和党派认同强度，并引入了外部冲击（如全球疫情）模拟共同威胁下的共同利益。政治极化的相图揭示了难以预测且可能不可逆的相变过程，这主要是由于不对称的滞回轨迹所致。研究结论指出，未来的实证研究需要更加关注极化临界点的识别及潜在对策的有效性。

关键词：极化；相变；临界点；迟滞动力学

1 引言

民主社会在分歧、辩论和多个利益集团之间的激烈竞争中茁壮成长 [1]。然而，人们越来越认识到，当多条争议线与党派身份一致时，政治分歧可能成为民主治理的一种负担 [2]。政治科学家将这种意见的结晶称为“约束” [3]，并指出两个主要原因：党派分歧和极端主义。首先，实质上不相关的问题（例如，死刑、生育权和枪支管制）的一致性减弱了党内分歧，消除了政治上的不和谐，并将意识形态的马赛克重新排列成两个截然相反的政治阵营 [4]。问题结盟和派系分歧使意见分歧相互加强，从而使温和的声音被压制，意见的分布变得越来越双峰化 [5]。唱名投票的分析表明，近几十年来，政治精英在核心政治问题上采取了更极端的立场 [6]。而选民之间的两极分化主要是情感上的，而不是意识形态上的 [7]。党派分歧和政治极端主义的结合可能会破坏多元多样性所依赖的“交叉分裂”，从而削弱有效民主治理所需的妥协能力。在本文中，我们指出了第三个危险，一个在越来越多的关于政治和文化两极分化的文献中很少受到关注的危险：存在一个临界点，超过这个临界点，即使面对共同的威胁，共同利益的激活也不能再使交战派系团结起来。我们对这个问题的兴趣源于一系列危机，这些危机可能会激发广泛的政治认同和统一的反应：大衰退、俄罗斯干预选举、迫在眉睫的气候灾难、全球流行病，以及最近的 1 月 6 日美国国会遇袭事件。毫不奇怪，这些事件引发了武装号召。令人惊讶的是手臂指向的方向。我们的目标不是解释两极分化，关于两极分化有大

量的文献和激烈的辩论。相反，我们将注意力集中在极化的相位分析上，这个问题在以前的研究中很大程度上没有引起人们的注意。我们使用舆论动态的一般模型来证明临界点的存在，在这个临界点上，即使是外部威胁也可能不足以扭转政治两极分化的自我强化动态。极化达到一个临界点时，增长率突然加快，当过程中表现出不对称的滞回环特征的相变。在自我强化的动态中，临界点的存在既不是不可避免的，也不是特别违反直觉的。然而，不能假设存在多个引爆点，一个在极化增加时，另一个在极化减少时。此外，如果下降轨迹的阈值低于上升轨迹的阈值，那么这种动态就很难逆转。这项研究的动机是需要引起对这种可能性的更多关注。

本文的主要贡献如下：

- 1) 提出了一个通用的政治极化模型，揭示极化的不可逆转性及其临界点：本文通过一个基于代理的模型，研究了政治极化的相变过程，发现了当极化程度超过某一临界值（称为“临界点”）后，极化可能变得难以逆转甚至不可逆。这种极化不可逆性源于非对称的迟滞回路（hysteresis loops），即极化增加和减少的临界点不同。当极化超过上升临界点后，即使控制参数回落到初始水平，极化也可能无法恢复。
- 2) 量化了党派认同、容忍度和外部冲击对极化的影响：通过操控三个关键参数（党派认同的强度、对分歧的容忍度和外部冲击的强度），本文系统地分析了这些因素如何驱动极化的发生和逆转。研究表明，党派认同和容忍度的增加会显著降低极化的可逆性，而外部冲击（如全球疫情）在极化程度较高时可能进一步加剧分裂，而非促进团结。
- 3) 揭示了外部冲击在不同极化背景下的双重效应：研究发现，在极化程度较低时，外部冲击可以作为统一力量，缓解党派分歧；但在极化程度较高时，冲击可能被现有分歧扩散并进一步恶化原有的极化状态。这一发现为理解现实中为何全球性危机未能缓解政治极化提供了理论支持。
- 4) 为未来研究提供了理论框架和实验基础：本文的模型不仅验证了极化过程中的相变现象，还提出了多个可供实证研究验证的假设，例如极化过程中党派重组的可能性和外部冲击的长期效果。通过这一理论框架，未来研究可以进一步探讨如何在临界点之前采取有效措施缓解释极化。
- 5) 强调了极化问题的紧迫性与政策干预的挑战：本文指出，极化的不可逆性可能对民主治理构成严重威胁，类似于气候变化中的不可逆转临界点。研究呼吁政策制定者和研究者尽早采取行动，避免极化趋势达到不可逆的阶段。

2 相关工作

我们的模型基于 Lu 等人 [8] 提出的 Szymanski 模型。Szymanski 模型专门用于分析两党政治体系中立法者的两极分化现象。在该模型中，立法者通过投票来表达他们在不同议题上的立场，而 Rice 指数 [9] 则被用来量化这种党派两极分化的程度。模型的动态变化由一组微分方程描述，这些方程假设立法者会根据选民、政党及捐赠者的需求调整自己的投票行为，以最大化连任机会。这些外部因素如同“看不见的手” [10]，在立法者个人利益与政党策略之

间维持着微妙的平衡。Lu 等人将模型参数拟合到训练数据，并验证了其对过去 30 届美国国会投票记录具有较高的预测准确性。

根据 Lu 等人 [8] 的研究，我们在精英阶层而不是在普通人群中模拟党派极化。（通过赋予社交网络局部结构，我们的模型可以适用于互动密度较低的更大人群，但我们将这一调查留给未来的研究。）然而，与 Lu 等人 [8] 不同的是，本研究中的模型并不打算作为经验预测的工具。相反，我们使用理论模型来寻找系统级动态中的临界点，这些临界点来自于一系列关于人际互动的假设。虽然立法机构是典型的应用，我们的模型不是基于美国或任何其他国家的国会，也没有立法机构特定的制度假设来阻止该模型应用于任何小型、紧密联系的组织。通过设计，该模型省略了可能限制其适用于特定组织设置的独特特征，例如来自游说者或选民的压力，限制其他成员访问的内部组织结构，或者可以制裁不服从成员的组织领导者之间的权力差异。该模型避免了可能约束行为的特定制度实践、习俗或正式规则。这些复杂性被抽象出来，这样我们就可以研究在不同组织环境中可能出现的意见两极分化的引爆点的一般特性。因此，一个不对应于任何特定经验实例的模型也可以说捕捉到了可能广泛相关的规律。

本研究中的模型与 Lu 等人 [8] 的不同之处在于，它是基于 agent 而不是基于方程的。基于主体的模型可以被概念化为“模型的总体”，与基于方程的“总体模型”形成对比 [11]。在后者中，系统的特征相互作用（例如，人口的大小、密度、集群和极化）。在基于代理的模型中，代理相互作用，这样系统的属性就会从它们相互作用的动态中出现。与基于方程的模型一样，基于主体的模型可以根据经验校准预测的准确性。另外，这些模型也可以是高度抽象的，基于一组简单但貌似合理的行为假设，类似于用于探索合作复杂性的博弈论模型 [12] 或社区隔离的临界点 [13]。两种方法——经验预测和理论探索——对极化的研究都很有用，但这里的方法是理论性的。在博弈论应用中，目标是产生假设，将其检验留给后续的实证研究。

3 本文方法

3.1 属性集设置

每个 agent 都有一个属性集 M ，由 agent 的政党身份和 agent 在问题上的立场组成。虽然立法者偶尔会改变政党，但我们假设政党身份是一个固定的二元属性。为简单起见，论文假设只有两个政党，每个政党有 50 个成员（就像现在的美国参议院一样），但该模型可以推广到一个执政（多数）联盟和一个少数反对派的多党制。

与政党身份不同，对问题的立场是连续的，容易改变。座席在一个问题上的立场可以从 +1（极端支持）到 -1（极端反对）连续变化。我们将突出问题的数量限制在 10 个，这足以确保极化必须由交叉切割产生；也就是说，在某些问题上意见一致的两个代理可能在其他问题上意见不一致。

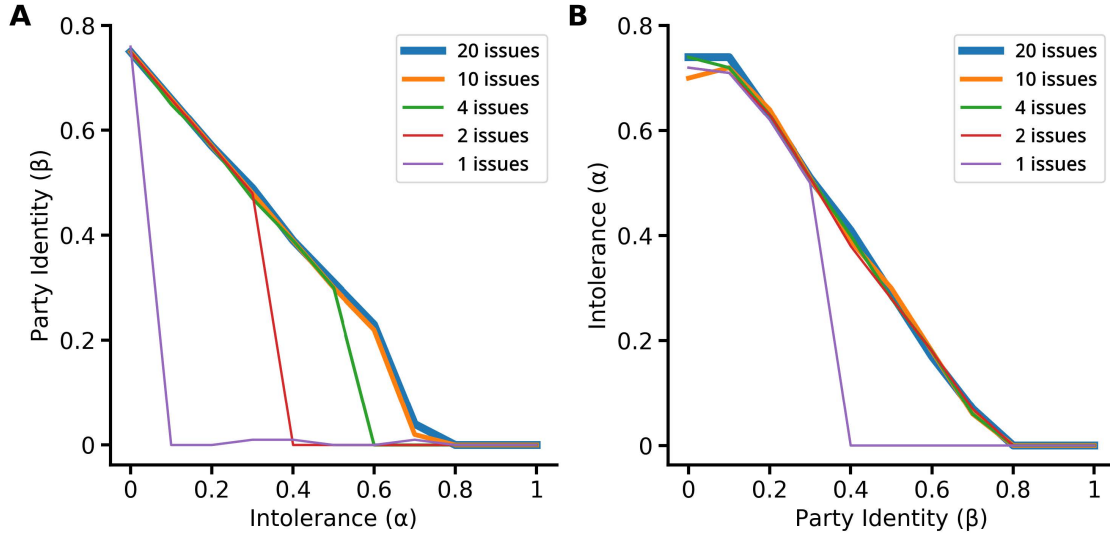


图 1. 党派认同和容忍度临界值随议题数量变化的稳健性

图中报告了极化相变中以极端化为衡量标准的临界点。测试的议题数量包括 1 个、2 个、4 个、10 个和 20 个。在 fig A 中，减少议题数量会降低发生相变的党派认同临界点。在面 fig B 中，对于 2 到 20 个议题的情况，临界点的位置表现出高度稳健性。然而，对于单一议题的情况，在中等到高水平的党派认同下，临界点位于较低的容忍度值。

如果模型仅限于一个问题方面，就不可能出现这种跨领域的划分。相反，如果两个代理碰巧在一个问题上意见不一致，那么在第二个问题上的一致将大大增加代理之间的相似性，由此产生的吸引力的增加可能会导致两个代理中的一个在代理不同意的问题上改变立场。虽然一个问题是不够的，但超过 10 个问题也不会有什么收获。如果两个代理人碰巧在 10 个问题上意见不一致，那么在第 11 个问题上的一致意见对他们的观点概况的总体相似性的影响相对较小。总之，只允许一个问题将使极化变得微不足道，但随着更多维度被添加到模型中，每个额外维度上的意见分布对极化动态的影响越来越小。

3.2 核心假设

论文准备了两个核心假设，为之前的意见动态计算模型提供了信息：影响力和同质性 [14] [15] [16] [17]。影响是指代理在问题上的重新定位，以回应代理所吸引的邻居。同质性是指一个智能体对网络邻居的吸引力随着其属性集的相似性而增加。影响力和同质性可以创造一种自我强化的动态，在这种动态中，共识增强了影响力，而影响力又减弱了剩余的分歧。

3.2.1 影响力效应

已有研究意见动态模型几乎都假设影响是积极的，从而加强了共识。然而，Schelling [13] 著名地展示了稳定的分裂（例如，居住隔离）是如何在负面影响下出现的，从而加剧了差异，就像群体内的邻居受到影响而迁出，以回应群体外成员的迁入。在这种情况下，负面影响不应与采纳同伴负面价值特征的影响（例如，吸烟或破坏公物）相混淆。相反，负面影响是指群体内成员与群体外邻居区别开来的倾向。应用于舆论动态，负面影响（或“疏远”）也被证明会产生两极分化，形成对立的意识形态阵营 [18] [19] [20]。我们通过假设代理调整其在问题上

的立场，使其更接近吸引它的邻居，并与排斥它的邻居区分开来，从而对积极和消极影响进行建模 [4]。

3.2.2 同质性效应

Axelrod [14] 展示了同质性如何创造出由不可渗透的文化边界所划分的地方社区，这种文化边界排除了不同社区成员之间没有共同属性的影响。虽然 Axelrod 假设离散属性，但在“有界置信度”模型中，动力学被推广到连续属性 [15] [16]。即使具有离散属性，Axelrod 模型中的文化差异也被证明最终会因文化漂移而崩溃为单一文化 [21]。

重要的是，如果同质性与仇外心理相结合，就不会发生单一文化的崩溃 [4]。与有界置信模型 [15] [16] 一样，代理具有容忍分歧的阈值。阈值定义了一个参考距离，代理在决定是靠近还是远离邻居时，将该距离与到邻居的距离进行比较。当参考距离接近 1 时，代理变得高度不宽容，以至于在问题上的小分歧都是不可接受的。随着参考距离的减小，代理人变得更加开放，愿意倾听和妥协，即使是对那些在问题上意见截然不同的人。代理人对邻居的响应取决于他们在 m 维属性空间中与邻居的接近程度，其中包括政党身份和对问题的立场。一个代理在所有维度上离它的邻居越近，邻居的影响就越有可能是积极的而不是消极的。代理 i 到邻居 j 的距离 D_{ij} 是他们在 10 个问题上的欧氏距离和党派距离的加权平均值。来自同一党派的代理人的党派距离为零，而来自不同党派的代理人的党派距离为 1。代理在 10 个问题上也连续的欧几里得距离，范围从最小的 0（代理在所有问题上都相同）到 1（代理在所有问题上最大程度地不一致）。

$$D_{ij} = D_{ij}^{party} \beta + D_{ij}^{issues} (1 - \beta) \quad (1)$$

其中权重 β 是一个连续的外生参数，对应于政党认同。总之，模型借鉴了经典的 Schelling（负影响）和 Axelrod（同质）模型，以及有界置信（阈值距离）模型。我们的贡献不在于提出一种新的舆论动态模型或对两极分化的新解释。相反，我们狭隘地关注一个问题：是否存在一个临界点，超过这个临界点，两极分化就变得难以或不可能逆转。

3.2.3 外生冲击

根据 Condie 和 Condie [22]，我们将外生冲击建模为潜在的统一事件，如全球流行病，经济崩溃，外国对手或恐怖分子（无论是国外还是国内）的袭击，或即将发生的气候灾难。冲击是通过创建一个所有代理都强烈同意的附加问题来实现的。对冲击的反应是由它们出现时的两极分化程度所驱动的。冲击发生在作为控制参数操纵的极化水平上。当冲击发生在时间步骤 t_s 时，每个代理 i 添加一个新问题 $M+1$ ，而不考虑其他 10 个问题的党派身份或立场。在计算 D_{ij} 时，由另一参数控制冲击相对于其他 11 个属性的冲击的重要性。然后，通过其他代理的积极和消极影响，代理更新他们在冲击问题上的立场。

3.3 测量指标

模型在每个时间步记录两个极化测量指标：政党极化和极端主义。政党极化是指在随机选择的问题上，两个政治政党随机成员之间的预期差异。范围是从零（任何问题上都没有差异）到一（每个问题上都有最大的可能差异）。极端主义是指随机选择的问题上的预期标准差（即所有问题的标准差的平均值）。我们操纵四个假设会表现出临界动态的全局参数：

- α 是容忍分歧的最小距离，用于调查随着 α 从零（完全容忍）到一（完全不容忍）增加时，对具有不同问题立场和政党身份的人的不容忍效果。
- β 是政党身份相对于问题立场在 D_{ij} 中的重要性，用于调查随着 β 从零（忽略政党）到一（忽略问题）增加时，政党认同的效果。
- γ 是冲击相对于所有其他维度在 D_{ij} 中的强度，用于找到逆转极化所需的冲击强度， γ 从零（忽略冲击）到一（忽略所有其他因素）增加。
- θ 是发生外部冲击时的极端主义水平，用于调查极化的可逆性， θ 从零（无极化）到一（完全极化）增加。

我们假设整个群体中的每个代理在这几个参数上都有相同的值，并观察当这些值在不同组织间变化时极化水平如何变化，公式2中的逻辑概率函数意味着即使所有其他条件相同，两个代理也可能对邻居产生相反的反应。

3.4 执行

在 $t = 0$ 时刻，每个代理 i 被分配一个初始向量 $Z_{i,0}$ ，该向量包含 $M = 11$ 个维度。第一个维度是二维静态的，对应于政党身份。其余 10 个维度是连续的，并随时间变化，对应于问题上的立场。代理被初始化为随机分配的政党身份（要么是 -1 或者 1），而每个问题上的位置则从均值为零、标准差为 0.25 的正态分布中随机抽取，必要时将范围截断到绝对值的单位区间内。因此，在 $t = 0$ 时刻，各政党之间没有预先存在的意识形态差异，无法根据代理的政党身份或任何其他维度上的位置预测其在某个问题上的立场。在基于代理的模型中重复使用随机起始位置进行初始化是一种可靠的寻找稳定点的方法。除了随机起点外，我们还通过从完全极化的配置开始建模滞后回路，以查看随着控制参数的下降，极化是否最终会崩溃。

在 $t = 0$ 的时间步完成初始化后，随机选择一个代理 i 在下一个时间步 $t + 1$ 更新其状态，受随机选择的邻居 j 的影响，而不考虑政党；然后，代理根据其邻居的状态 $Z_{j,t}$ 更新其状态 $Z_{i,t}$ 到新的状态 $Z_{i,t+1}$ ，具体如下。首先， i 根据一个对于所有代理都相同的连续不容忍参数 $\alpha \in [0, 1]$ 来评估其与 j 的相对接近程度，该参数代表容忍分歧的最小距离。如果 $\alpha = 0$ ，则每段距离 D_{ij} 都会被容忍，所有影响都是积极的；因此， i 将在每个问题上更接近 j 。如果 $\alpha = 1$ ，则任何距离都不会被容忍，所有影响都是消极的；因此， i 将在每个问题上远离 j 。

具体来说，正向影响的概率 $(P_+)_{ij}$ (D_{ij}) 的累积逻辑函数，陡度为 m ：

$$p_+ = \frac{1}{1 + e^{m(D_{ij} - (1-\alpha))}} \quad (2)$$

我们设置 $m = 10$ ，但是当函数接近 $m \gg 10$ 的确定性阈值时，结果是鲁棒的。然而，当 m 趋近于 1 时，函数变为线性，这引入了足够的噪声来衰减（但不能消除）极化。如果 j 的影响是正向的， $Z_{id,t+1}$ 是 $Z_{id,t}$ 和 $Z_{jd,t}$ 在每个动态维度 d 上的加权平均值：

$$Z_{id,t+1} = Z_{id,t} + (L - Z_{id,t})|D_{ij} - c| \text{ran} \quad (3)$$

其中 ran 是单位区间内的均匀随机实数， d 是更新的维度， $c = 1$ ， $L = Z_{jd,t}$ 。式 [3] 中的 $L - Z_{id,t}$ 定义了 i 在 d 上移动到 j 所需的距离。表达式 $|D_{ij} - c|$ 意味着 i 在给定问题上的移动比

例随 i 和 j 在所有 M 维度（包括 d ）上的总体距离的增加而减少。即对某一问题的分歧越小，解决起来就越容易，符合同质性偏好原则（即相似吸引）。

如果 j 的影响是负向的，那么 $c = 0$ ， i 在每个动态维度 d 上朝着使 i 和 j 之间的距离增大的方向移动。 L 的值取决于 $Z_{id,t}$ 和 $Z_{jd,t}$ 的值。如果 $Z_{id,t} > Z_{jd,t}$ 或 $Z_{id,t} = Z_{jd,t} < 0$ ，那么 $L = 1$ 。如果 $Z_{id,t} < Z_{jd,t}$ 或 $Z_{id,t} = Z_{jd,t} > 0$ ，那么 $L = -1$ 。如果 $Z_{id,t} = Z_{jd,t} = 0$ ，那么 L 随机选择为 1 的概率为 0.5，否则为 -1。随着 i 和 j 在每个动态维度 d 上的距离增加， i 移动到离 j 尽可能远的最大可能距离减小；即 i 需要移动的距离减小，而 i 远离 j 的距离比例随着整体距离 D_{ij} 的增加而增加。简而言之，负向影响是正向影响的镜像：邻居之间的分歧越大，解决起来就越困难，符合异质性厌恶原则（即对立排斥）。

更新完所有 10 个问题后，再随机抽取一对新的 i, j ，时间步长增加 1。此过程迭代进行，直到系统要么稳定在一个稳态，要么达到任意的时间限制。

4 复现细节

4.1 与已有开源代码对比

论文假设只有两个政党，每个政党有 50 个成员（就像现在的美国参议院一样），但该模型可以推广到一个执政（多数）联盟和一个少数反对派的多党制。

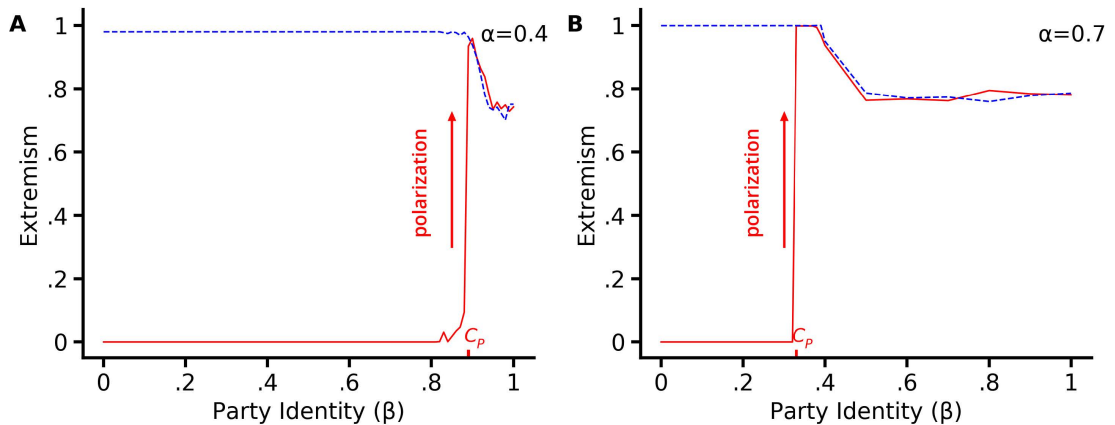


图 2. 党派认同临界值 β 的稳健性

当在政党数量从 2 增加到 5 个时，党派认同临界值 β 的稳健性，与图 5 相似，但由于政党数量增加导致动力学的变化，参数调整为 $\alpha = [0.4, 0.7]$ 和 $m = 100$ 。

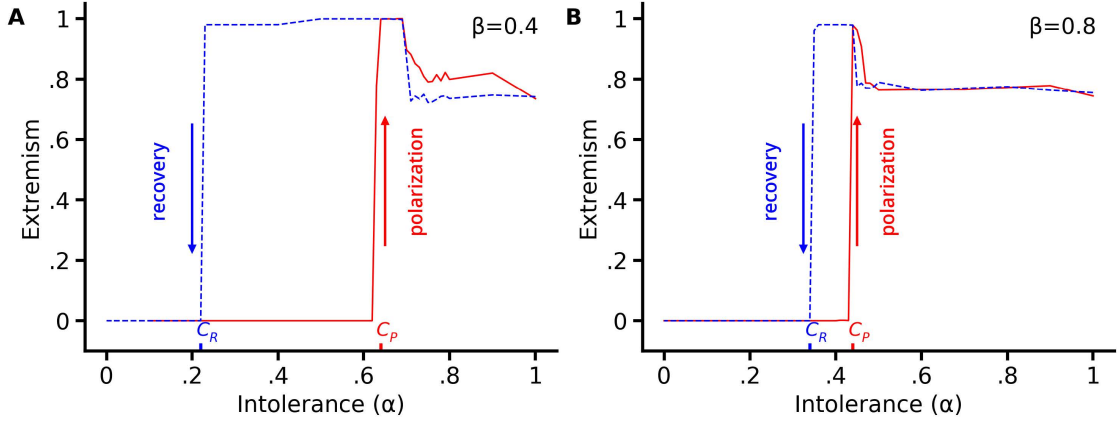


图 3. 容忍度临界值 α 的稳健性

容忍度临界值 α 在政党数量从两个增加到五个时的稳健性。结果以极端化为衡量标准进行报告，其质量上与正文中的 fig3 相似。但由于政党数量增加导致动力学变化，参数调整为 $\beta = [0.4, 0.8]$ 和 $m = 100$ 。

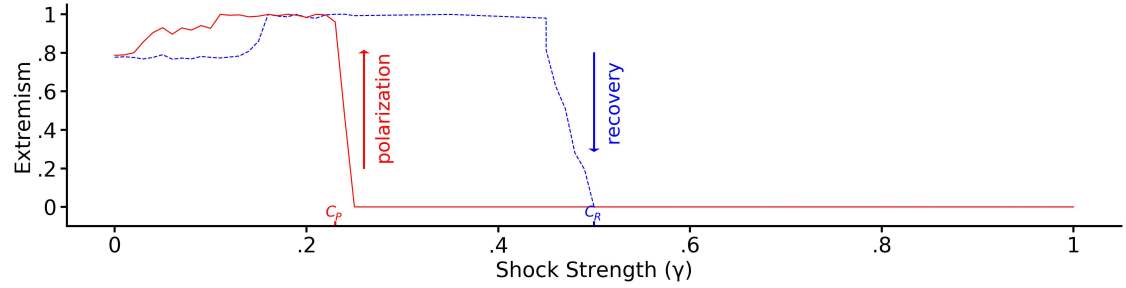


图 4. 外部冲击临界值 γ 的稳健性

外部冲击强度 γ 临界值在政党数量从两个增加到五个时的稳健性。结果以极端化为衡量标准进行报告，其质量上与正文中的 fig4 相似。但由于政党数量增加导致动力学变化，参数调整为 $\alpha = 0.6$ 、 $\beta = 0.6$ 和 $m = 100$ 。

4.2 实验环境搭建

```
python 3.6.9
numpy 1.23.5
matplotlib 3.6.2
networkx 3.0
scipy 1.10.1
```

4.3 界面分析与使用说明

直接运行代码就可以

4.4 创新点

当政党数量不在是两党时，党派认同 β 临界值、容忍度 α 临界值和外部冲击强度 γ 临界值与两党制时相似。

5 实验结果分析

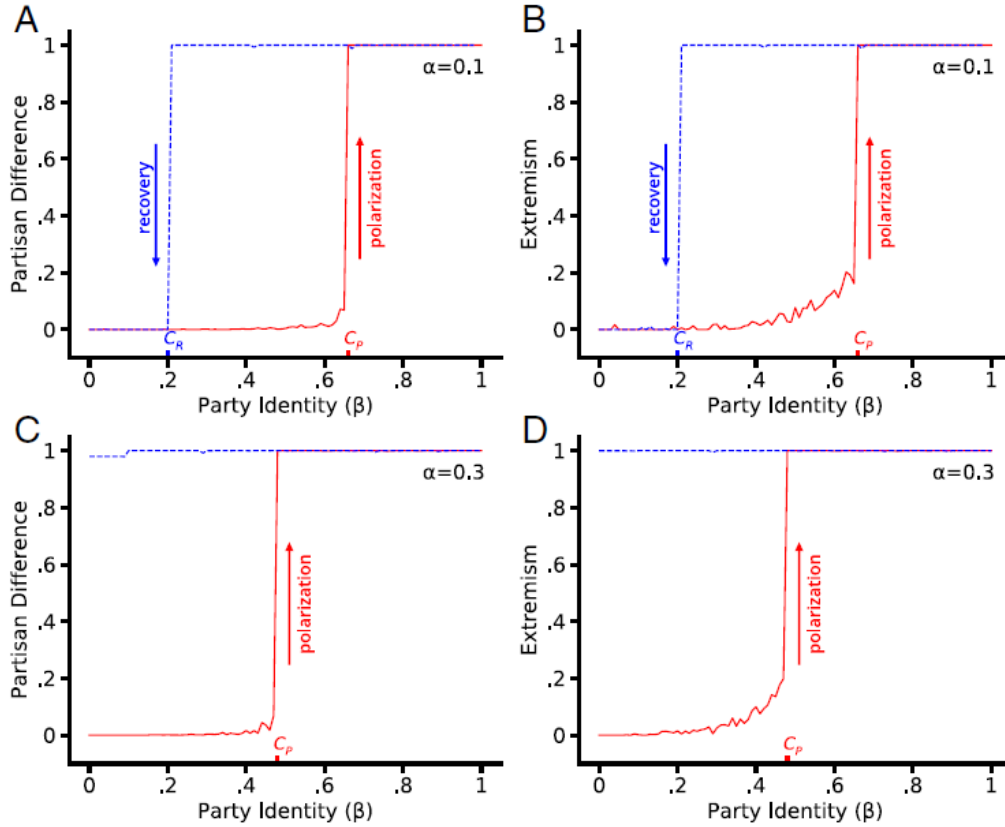


图 5. 政党认同相变图

图 5 展示了使用两个不同的不容忍值的政党认同相变。极化被测量为政党差异（图 5A 和 C）和极端主义（图 5B 和 D）。当 $\alpha = 0.1$ 时（图 5A），恢复轨迹的临界值（CR）远小于前向轨迹。这意味着只有当政党认同强度大幅下降至 CP（极化增加的临界值）以下，降低政党认同强度才会大大减少极化水平。

当 $\alpha = 0.3$ 时，在图 5C 和 D 中，政党认同不需要很强就能触发高极化的相变，并且，极化出现不可逆性，即使政党认同强度降为零。

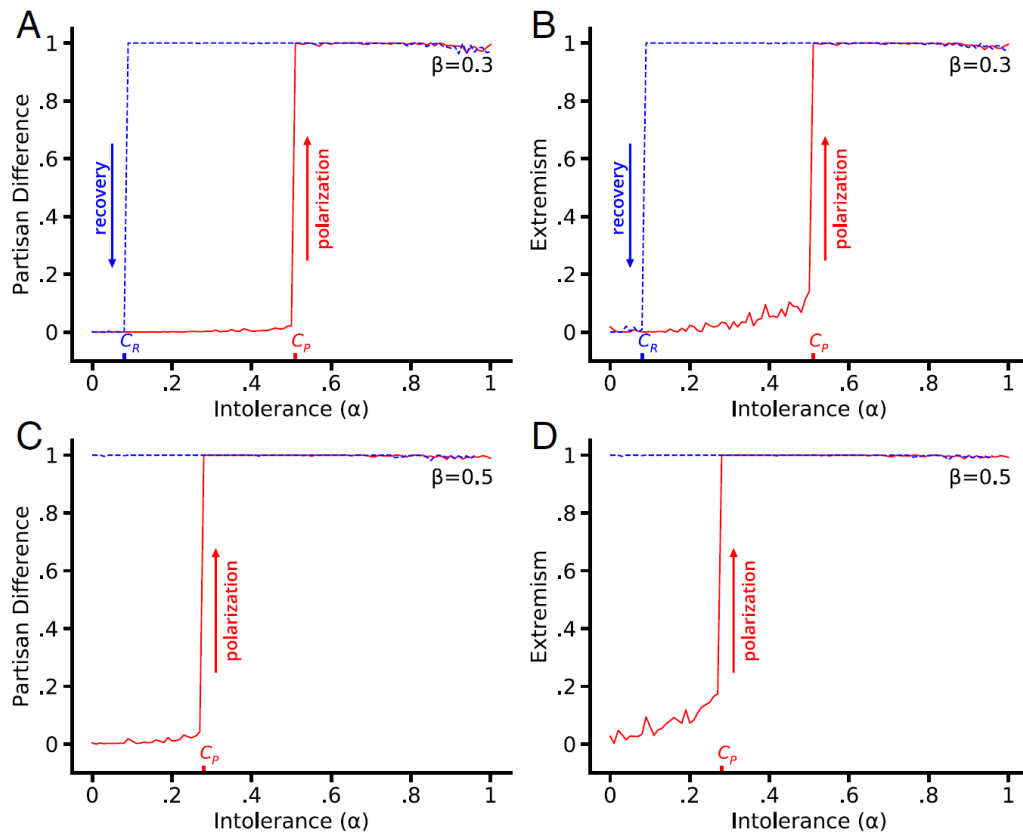


图 6. 容忍度相变图

图 6A 和 B 展示了在两个政党认同值 $\beta = 0.3$ 和 $\beta = 0.5$ 下的容忍度 α 的相变。图 6A 和 C 测量极化为政党差异，图 6B 和 D 测量为极端主义。在极化临界点以上，随后的政治宽容度增加并不会将政党重新团结在一起（图 6A）或减少极端主义（图 6B），除非不容忍水平大幅降至极化增加的水平以下。图 6C 和 D 显示，政党认同的增加缩小了滞后回路的宽度，政党认同的增加进一步降低了 C_R ，极化无法逆转。

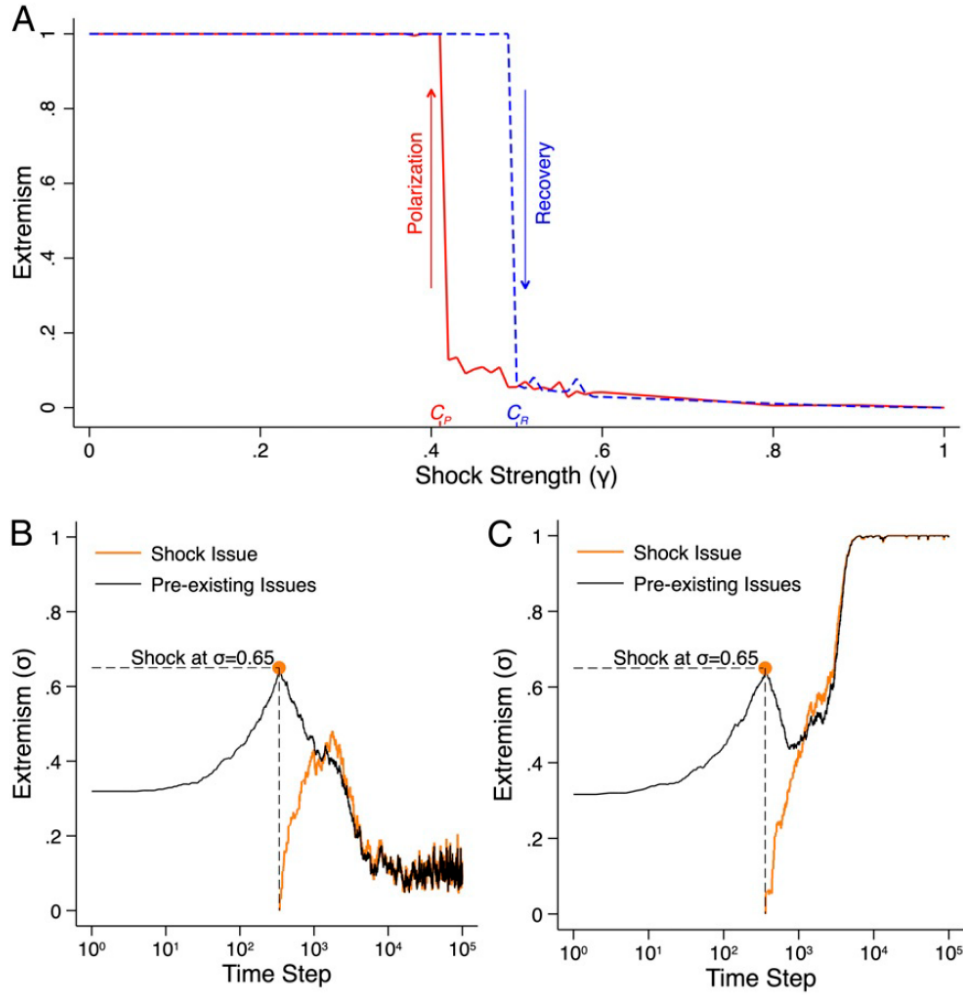


图 7. 外部冲击的相变图

图 7 报告了关于外部冲击的相变分析。在所有三个面板中， $\alpha = \beta = 0.5$ 。图 A 显示了随着冲击强度增加的顺时针滞后回路。存在一个临界水平的 γ (C_R)，超过该水平后，冲击几乎总是能导致恢复，即使冲击发生在或接近完全极化的情况下。存在第二个临界水平的 γ ($C_P < C_R$)，低于该水平时，冲击几乎永远不会导致恢复，即使冲击发生在或接近零极化的情况下。在这两个临界点之间，冲击可能导致或不导致恢复，这取决于冲击发生时的极化水平。

图 7B 和 C 描绘了发生在图 7A 所示关键区域内的冲击 ($\gamma = 0.45$) 的极化轨迹。黑线显示了 10 个现有问题（不包括冲击）的轨迹，橙色显示了关于冲击的分歧轨迹。冲击发生在极化水平 ($\theta = 0.65$) 处，此时极化的概率达到 0.5。图 7B 显示了冲击如何逆转极化轨迹，而图 7C 显示了即使所有参数与图 7B 完全相同，恢复也可能失败。

6 总结与展望

我们使用一般的意见动态模型展示了存在临界点，在这些临界点上，即使是全球大流行病等外部威胁，也可能不足以逆转党派极化的自增强动态。极化在临界点到达时突然加速增加，并且过程表现出由不对称滞后环特征的相变。该模型适用于小型、密集连接的组织，如立法机构。

我们的发现的意义不仅在于证明极化可以削弱对外部冲击的统一响应。未来的研究还可

以考察冲击对政党重新排列的影响。如果政党在应对冲击时内部存在分歧，我们的模型预测了一种动态，在这种动态中，反对党内的派系可能会互相吸引，并与其所在政党的成员疏远。简而言之，我们提出了关于冲击时机、长度和影响的假设，为未来的实证研究提供了一个兼具理论意义又具实践重要性的议程。

参考文献

- [1] D. B. Truman, *The Governmental Process* (Alfred A. Knopf, New York, 1951).
- [2] D. Baldassarri, A. Gelman, Partisans without constraint: Political polarization and trends in American public opinion. *AJS* 114, 408–446 (2008).
- [3] P. E. Converse, “The nature of belief systems in mass publics” in *Ideology and Discontent*, D. E. Apter, Ed. (Free Press, New York, 1964), pp. 206–261.
- [4] D. DellaPosta, Y. Shi, M. Macy, Why do liberals drink lattes? *AJS* 120, 1473–1511 (2015).
- [5] P. Dreyer, J. Bauer, Does voter polarisation induce party extremism? The moderating role of abstention. *West Eur. Polit.* 42, 824–847 (2019).
- [6] N. McCarty, K. T. Poole, H. Rosenthal, *Polarized America: The Dance of Ideology and Unequal Riches* (MIT Press, Cambridge, MA, 2006).
- [7] S. Iyengar, Y. Lelkes, M. Levendusky, N. Malhotra, S. J. Westwood, The origins and consequences of affective polarization in the united states. *Annu. Rev. Polit. Sci.* 22, 129–146 (2019).
- [8] X. Lu, J. Gao, B. K. Szymanski, The evolution of polarization in the legislative branch of government. *J. R. Soc. Interface* 16, 20190010 (2019).
- [9] S. A. Rice, *Quantitative Methods in Politics* (Alfred A. Knopf, New York, 1928).
- [10] A. Smith, *The Theory of Moral Sentiments* (Penguin, New York, 2010).
- [11] M. Macy, R. Willer, From factors to actors: Computational sociology and agent-based modeling. *Annu. Rev. Sociol.* 28, 143–166 (2002).
- [12] R. Axelrod, *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton University Press, Princeton, NJ, 1997).
- [13] T. Schelling, Dynamic models of segregation. *J. Math. Sociol.* 1, 143–186 (1971).
- [14] R. Axelrod, The dissemination of culture: A model with local convergence and global polarization. *J. Conflict Resolut.* 41, 203–226 (1997).
- [15] G. Deffuant, F. Amblard, G. Weisbuch, T. Faure, How can extremism prevail? A study based on the relative agreement interaction model. *J. Artif. Soc. Soc. Simul.* 5, 1 (2002).

- [16] R. Hegselmann, U. Krause, Opinion dynamics and bounded confidence models, analysis and simulation. *J. Artif. Soc. Soc. Simul.* 5, 2 (2002).
- [17] A. Flache, M. Macy, Local convergence and global diversity: From interpersonal to social influence. *J. Conflict Resolut.* 55, 970–995 (2007).
- [18] M. W. Macy, J. A. Kitts, A. Flache, S. Benard, “Polarization in dynamic networks: A hopfield model of emergent structure” in *Dynamic Social Network Modelling and Analysis*, R. Brieger, K. Carley, P. Pattison, Eds. (The National Academies Press, Washington, DC, 2003), pp. 162–173.
- [19] N. P. Mark, Culture and competition: Homophily and distancing explanations for cultural niches. *Am. Sociol. Rev.* 68, 319–345 (2003).
- [20] D. Baldassarri, P. Bearman, Dynamics of political polarization. *Am. Sociol. Rev.* 72, 784–811 (2007).
- [21] K. Klemm, V. M. Eguíluz, R. Toral, M. S. Miguel, Global culture: A noise-induced transition in finite systems. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 67, 045101 (2003).
- [22] S. A. Condie, C. M. Condie, Stochastic events can explain sustained clustering and polarisation of opinions in social networks. *Sci. Rep.* 11, 1355 (2021).
- [23] M. J. van Strien, S. H. Huber, J. M. Anderies, A. G. Regamey, Resilience in socioecological systems: Identifying stable and unstable equilibria with agent-based models. *Ecol. Soc.* 24, 8 (2019).
- [24] J. M. Brändle, G. Langendijk, S. Peter, S. H. Brunner, R. Huber, Sensitivity analysis of a land-use change model with and without agents to assess land abandonment and long-term re-forestation in a swiss mountain region. *Land (Basel)* 4, 475–512 (2015).
- [25] R. Dahl, *Pluralist Democracy in the United States* (Rand McNally, Chicago, IL, 1967).
- [26] E. C. Ladd, C. D. Hadley, *Transformations of the American Party System* (Norton, New York, 1978).
- [27] K. Arceneaux, M. Johnson, “More a symptom than a cause: Polarization and partisan news media in America” in *American Gridlock: The Sources, Character, and Impact of Political Polarization* (Cambridge University Press New York, NY, 2015), pp. 309–336.
- [28] J. C. Rocha, G. Peterson, O. Bodin, S. Levin, Cascading regime shifts within and across scales. *Science* 362, 1379–1383 (2018).