



Working Paper No.202411

Jul.29, 2024

本文已发表于《世界经济与政治》2024年第7期

李冲: lichong@cass.org.cn

盛名之累：冲突预期与联盟声誉的再调节

【摘要】国家在联盟中的声誉包括直接声誉和间接声誉。直接声誉来源于国家与特定盟友的历史互动，间接声誉来源于国家与第三方盟友的历史互动。直接声誉反映可靠性，国家的直接声誉越好、其履行联盟承诺的概率越大，盟友挑起冲突的意愿越强。间接声誉预示资源分配偏好，国家的间接声誉越好、其资源分配越向第三方盟友倾斜，特定盟友挑起冲突的意愿越弱。直接声誉和间接声誉会通过正反馈机制影响国家“被牵连”和“被抛弃”的风险，促使其主动调控声誉水平，采取措施约束或安抚盟友。间接声誉较好的国家为降低“被抛弃”的风险，可能维持对特定盟友的支持力度；直接声誉较好的国家为降低“被牵连”的风险，可能限制特定盟友使用武力。由于强国调配资源的自由度相对较大，因此其间接声誉对盟友挑起冲突意愿的抑制作用较弱。间接声誉好的国家与特定盟友利益分化程度较大，因此同时拥有良好直接声誉和间接声誉的国家约束盟友的动力更强。基于1816-2001年联盟声誉与冲突数据的统计检验和1961-1976年美韩同盟的案例检验可发现，国家追求的是适度而非最大化的联盟声誉。这一发现为国家联盟声誉的波动与盟友挑起冲突概率的变化提供了解释。

【关键词】 声誉；联盟困境；国际冲突；美韩同盟；联盟管理

1975年4月底，美军开展了旨在撤离滞留越南南部美国公民和部分亲美人士的“常风行动(Operation Frequent Wind)”，不久之后越南人民军攻克西贡(今胡志明市)。美军最后一架直升机在使馆楼顶准备起飞的图片广为流传，被舆论描绘为美国的“西贡时刻”。无独有偶，2021年8月美军从阿富汗撤出上演了“喀布尔时刻”。从西贡到喀布尔，美国的声誉受到极大破坏。然而，在讨论美国声誉变化对其盟友的影响时各界存在不同看法。有观点认为，上述行为暴露了美国抛弃盟友的传统，折射出美国加速衰落的现实，对美国联盟体系构成了冲击，将导致地区盟友不再信任美国。也有观点指出，不同盟友在美国对外战略中的地位不同，美国从一地撤军并不影响对其他地区的安全承诺。相关争论为研究一国在联盟中的声誉(简称“联盟声誉”)如何影响其盟友的冲突行为提供了现实依托。既有研究主要关注直接声誉，即国家与特定盟友历史互动形成的声誉对该盟友发动冲突的影响，认为良好声誉是冲突爆发的“加速器”但既有理论与现实之间存在张力：一方面，“西贡时刻”等场景表明，除直接声誉外特定盟友挑起冲突的意愿还可能受国家与第三方盟友历史互动形成的声誉即间接声誉影响；另一方面，有研究发现良好的声誉在盟友冲突决策中有时不仅不是“加速器”，反而扮演了“制动阀”的角色。基于此，本文关注的核心问题是国家的联盟声誉对盟友挑起冲突的概率将产生何种影响？该问题包含两个层面：一是国家在联盟中的直接声誉和间接声誉的影响有何区别？二是良好声誉究竟是鼓励还是抑制了盟友挑起冲突？本文认为，直接声誉反映可靠性，国家的直接声誉越好、其履行联盟承诺的概率越大，特定盟友挑起冲突的意愿也越强。间接声誉预示资源分配偏好，国家的间接声誉越好、其资源分配越向第三方盟友倾斜，特定盟友挑起冲突的意愿越弱。直接声誉和间接声誉的正反馈会加剧联盟困境，提升国家“被牵连”和“被抛弃”的风险，促使国家主动调控声誉水平，采取措施约束或安抚盟友。

一、联盟声誉与冲突爆发

声誉是外界依据行为体过去的表现对其特征形成的信念(belief)。根据所反映特征的差别可以将声誉分为不同类型。例如，决心声誉(reputation for resolve)反映国家在危机中拒不让步的可能性，诚实声誉(reputation for honesty)体现国家言行一致的程度。联盟研究主要关注可靠声誉(reputation for reliability)，该声誉反映一国履行联盟安全承诺的概率。学界围绕联盟声誉的来源以及联盟声誉与冲突爆发的关系展开了相关争论

(一) 联盟声誉的来源

联盟声誉来源于历史行为，但在哪类历史行为可以构成联盟声誉上学界仍存在分歧。有观点认为，声誉作为一个关系型概念主要包含拥有者与观察者两类行为体。在国家 A 与国家 B 的互动中，如果国家 B 观察到国家 A 过去的行为并据此对国家 A 的某些特征做出判断，那么国家 A 就是声誉的拥有者、国家 B 是声誉的观察者。声誉本质上属于观察者的认知，不同的观察者可能对同一行为产生不同的判断。因此谈及联盟声誉时应明确观察者是哪个盟友，联盟声誉源自一国与特定盟友之间的定向历史互动。也有观点主张，对特定盟友而言，非定向历史即国家与第三方盟友的历史互动也能构成联盟声誉的来源。其一，盟友使用非定向历史具有必要性。但参与结盟谈判的各方尚未建立联盟关系，彼此不存在定向历史，只能根据对方与其他盟友的历史互动来判断其可靠性。其二，同类型联盟具有相似的安全承诺这也为盟友使用非定向历史判断一国的可靠性提供了依据。

对声誉来源的讨论反映了学界关于声誉是否具有通约性的争论:如果联盟声誉只来源于定向历史，说明声誉不可通约，直接声誉与间接声誉产生的影响不同:如果联盟声誉来源于非定向历史，说明声誉在不同盟友之间可以通约，直接声誉和间接声誉产生的影响相似。然而，既有研究存在两点改进空间:第一，主要关注间接声誉在多大程度上反映国家可靠性，忽视了间接声誉还能够预示国家在联盟中的资源分配态势。如果一国的间接声誉良好，说明该国可能将关注重点转向

第三方盟友第二，在直接声誉和间接声誉同时存在的情况下，如果两类声誉释放的信号相反盟友的冲突决策可能面临更复杂的考量，

（二）联盟声誉对冲突爆发的作用

从理性主义路径出发，部分学者认为声誉水平与声誉效果之间呈线性关系，即国家声誉越好，越有可能建立领导权威、收获盟友忠诚、巩固执政地位和威慑敌对国家。相关的冲突研究指出，一方面，良好的联盟声誉能够提升国家对盟友提供的延伸威慑的效果。如果国家在过往延伸威慑场景中表现坚定，就能减小敌人对盟友发动攻击的概率。另一方面，良好的联盟声誉也可能会给国家带来道德风险即国家声誉越好，盟友获得支持的预期越高，挑起冲突的可能性越大。从认知心理路径出发，有学者指出声誉的影响取决于观察者解读：如果盟友将国家的历史行为归因于外部环境而非其内在可靠性，联盟声誉的参考价值将受限，因此声誉的影响可能是非线性的。例如，林民旺和达里尔·普雷斯(Daryl G.Press)指出，国家在危机决策中更重视现实因素而较少援引历史。唐世平认为在无政府体系中国家会以“做最坏打算”的方式看待敌人和联盟，即假定敌人决心是坚定的，而联盟是不可靠的。

上述理论均存在将声誉作用极端化的问题：理性主义路径过于强调声誉中心地位，而认知心理路径怀疑国际体系中无法形成声誉。近年来一些学者调和了两类观点，认为声誉并不是完全独立地发挥影响，而是受国际格局、技术应用和地区特性等条件的限制。如相关研究指出，两极格局与核时代下的大国更加重视声誉，处在违约盛行地区的国家对负面联盟声誉的接受度更高。既有研究给本文带来了两点启示：第一，联盟声誉并非独立发挥作用，与盟友挑起冲突概率之间的关系可能呈现非线性。具体到本文的研究问题，直接声誉的影响可能受间接声誉的调节。第二，国家具有调控声誉的动机。有关延伸威慑和道德风险的研究表明联盟声誉具有“过犹不及”的特征。如果国家联盟声誉过差，敌人很有可能对其盟友发动攻击；如果国家联盟声誉过好，其盟友的机会主义行为则会增多，对外挑起冲突的可能性上升。因此，国家有必要根据冲突风险动态调整行为，使声

誉‘恰到好处’。

二、资源分配、约束意愿与联盟声誉的再调节

在借鉴既有研究的基础上，本部分首先界定了直接声誉与间接声誉的概念，然后介绍两类声誉的来源，并分析两类声誉对盟友冲突行为的作用机制。直接声誉影响盟友对国家可靠性的认知，间接声誉影响盟友对国家资源分配态势的认知，两类联盟声誉共同影响盟友挑起冲突的意愿。由于联盟困境的存在，直接声誉或间接声誉过好会提升一国被特定盟友牵连或抛弃的风险。为调适联盟困境，国家会主动调节声誉，进而改变盟友挑起冲突的意愿。

（一）直接声誉与间接声誉

一国的联盟声誉可分为直接声誉和间接声誉。表 1 对两类声誉进行了比较。对特定盟友 B 而言，由于联盟承诺存在相互依赖，间接声誉虽然来源于国家 A 和第三方盟友 C 的历史互动，但也能够反映国家 A 的履约可靠性。尤其是第二次世界大战后联盟多于非战时环境中组建，盟友 B 可能在一段时间内不会卷入冲突，利用间接声誉判断国家 A 履约概率的必要性提升。需要指出的是，如果盟友 B 卷入冲突的次数较多，即直接声誉的来源较为丰富，其就会优先考虑利用直接声誉判断国家 A 的可靠性。

表 1 直接声誉与间接声誉的比较

| 声誉类型 | 直接声誉 | 间接声誉 |
|------|-----------------------|--------------------------|
| 拥有者 | 国家 A | 国家 A |
| 观察者 | 盟友 B | 盟友 B |
| 来源 | 国家 A 既往在盟友 B 卷入冲突时的表现 | 国家 A 既往在第三方盟友 C 卷入冲突时的表现 |
| 反映特征 | 国家 A 履约可靠性 | 国家 A 履约可靠性与资源分配偏好 |

资料来源：笔者自制。

除可靠性外，间接声誉还能反映国家 A 在联盟中的资源分配态势。国家 A 的间接声誉较好，说明第三方盟友 C 卷入的冲突较多、国家 A 对第三方盟友 C 的支持力度较大，且国家 A 未来将向第三方盟友 C 投入更多资源。由于国家 A 的资源总量有限且短时间内一般不会迅速增长，盟友 B 未来获得国家 A 的支持会减少。例如，冷战初期美国卷入朝鲜半岛和东南亚等地区的冲突，造成北约盟友

担心美国减少对欧洲的投入。在奥巴马政府实施“亚太再平衡”战略并计划将海军 60%的舰船交由亚太司令部指挥时，欧洲盟友也表达过类似担忧。

（二）联盟困境与声誉的正反馈

如果国家 A 的直接声誉较好，盟友 B 会认为国家 A 的履约可靠性较好、自己更有可能在潜在冲突中获得国家 A 的支持，因此挑起冲突的意愿更强；如果国家 A 的直接声誉较差，说明其履约可靠性较差，盟友 B 对获得国家 A 支持的预期较低、挑起冲突的意愿也较弱。如果国家 A 的间接声誉较好，盟友 B 认为国家 A 会继续将资源向第三方盟友 C 倾斜，在预期国家 A 减少支持的基础上，盟友 B 挑起冲突的意愿会下降；反之，国家 A 的间接声誉较差暗示其未来可能减少对第三方盟友 C 的投入而更加重视盟友 B，这会提升盟友 B 挑起冲突的意愿。

作为一类国际政治系统，联盟中不仅存在从输入端（声誉）到输出端（冲突）的影响，也存在从输出端到输入端的反馈。反馈使得行为体可以根据输出结果动态调整输入，使得系统能够持续和动态地适应与变化。反馈机制关注时间维度，输入对输出的影响与输出对输入的反馈发生在不同时间： $T-1$ 时刻的输入会影响 T 时刻的输出，而 T 时刻输出所反馈的是 $T+1$ 时刻的输入。图 1 和图 2 展示了联盟声誉的正反馈机制。图 1 展示了直接声誉的正反馈机制。如果国家 A 在 $T-1$ 时刻的直接声誉较好，盟友 B 在 T 时刻挑起冲突的概率更大。此时具有良好直接声誉的国家 A 会继续向盟友 B 提供支持，进一步提升自身在 $T+1$ 时刻的直接声誉。相反，如果国家 A 在 $T-1$ 时刻的直接声誉较差，盟友 B 在 T 时刻挑起冲突的概率较小，此时国家 A 会继续减少对盟友 B 的支持，进而导致自身在 $T+1$ 时刻的直接声誉进一步恶化。

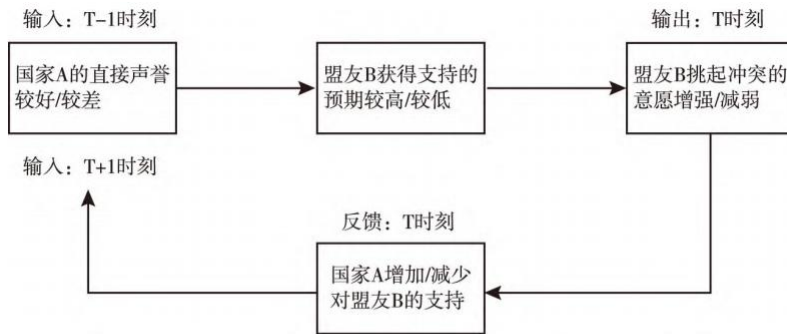


图1 直接声誉的正反馈

资料来源：笔者自制。

图2展示了间接声誉的正反馈机制。如果 T-1 时刻国家 A 的间接声誉较好，盟友 B 在 T 时刻挑起冲突的概率较小，这让国家 A 有机会将更多的资源投入第三方盟友 C，从而进一步提升自身在 T+1 时刻的间接声誉。如果国家 A 在 T-1 时刻的间接声誉较差，盟友 B 在 T 时刻挑起冲突的概率较大，由于没有被第三方盟友 C 牵扯更多精力，国家 A 可以继续向盟友 B 提供支持，这将导致国家 A 对第三方盟友 C 的投入进一步减少，其在 T+1 时刻的间接声誉进一步恶化。联盟声誉正反馈的结果就是格伦·斯奈德（Glenn H. Snyder）所描述的联盟困境，即国家履行联盟承诺的两难：如果坚定地履行联盟承诺并积极向盟友提供安全支持，国家可以收获盟友的忠诚，但同时也会因帮助盟友而卷入本不想参与的冲突——这就是联盟中的“被牵连”风险。如果国家履行承诺不坚定且不积极给盟友提供安全支持，该国可以降低“被牵连”的风险，但盟友会怀疑其承诺的可靠性并不再支持该国，这就是联盟中“被抛弃”的风险。

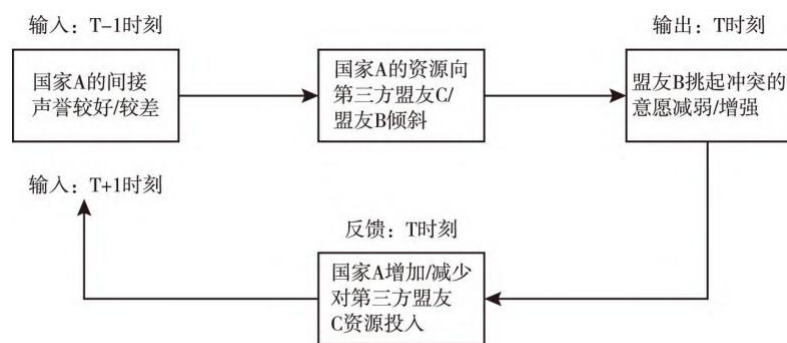


图2 间接声誉的正反馈

资料来源：笔者自制。

（三）联盟管理与声誉的负反馈

正反馈机制下输入持续增加或减少，最终会超过系统所能承受的极限或低于维持系统运行的必要水平，造成系统崩溃或转型。理论上，联盟声誉的正反馈机制会加剧联盟的不稳定性——直接声誉的螺旋上升会将国家 A 与盟友 B 深度捆绑并诱发全面战争，间接声誉的螺旋上升则导致国家 A 与盟友 B 分崩离析。然而现实中联盟的演化特点并非如此。1815 年至第一次世界大战爆发前，新建联盟平均存续时间为 8.1 年；1914—1944 年，新建联盟平均存续时间 9.1 年；冷战期间，新建联盟平均存续时间更是增加至 13.1 年。联盟稳定性之所以实现提升，是因为国家会通过联盟管理来调适联盟困境。联盟管理是联盟成员通过各类正式与非正式安排协调彼此内外政策的过程，其基本手段包括安抚与约束两类：安抚是指国家通过增加经济援助、部署军事人员和武器、宣示再保证与扩大联盟承诺范围等方式提升盟友向心力，约束是指国家以减少经济安全援助、限制联盟承诺范围与威胁不再提供安全保证等手段阻止盟友使用武力。联盟管理反映了联盟声誉的负反馈机制，国家会根据之前互动的结果调整后续行为选择。换言之，国家在联盟中追求的是适度而非最大化的声誉。

图 3 展示了联盟管理对间接声誉负反馈的影响。拥有良好的间接声誉使国家 A 面临被盟友 B 抛弃的风险。此时国家 A 倾向于安抚盟友 B，具体可能通过调整资源分配态势即适当提高对盟友 B 的援助来调适联盟困境。一旦国家 A 调整资源分配，盟友 B 获得国家 A 支持的预期会提升，其挑起冲突的意愿也随之增强；相应地，国家 A 未来对第三方盟友 C 的投入会下降，其在 T+1 时刻的间接声誉也会下滑。本文认为，国家对间接声誉的调节效果还受到自身实力调节的影响。如果国家 A 实力较强，在盟友 B 与第三方盟友 C 之间调配资源的压力就会较小，可以在安抚盟友 B 的同时维持对第三方盟友 C 的投入。例如，美国在 20 世纪 60 年代制定了“两个半战争”战略，准备同时在欧洲与亚洲方向同苏联和中国进行全面战争。当前拜登政府也对中俄两国实施了“双制衡”战略。基于此，本文提出两个假说。

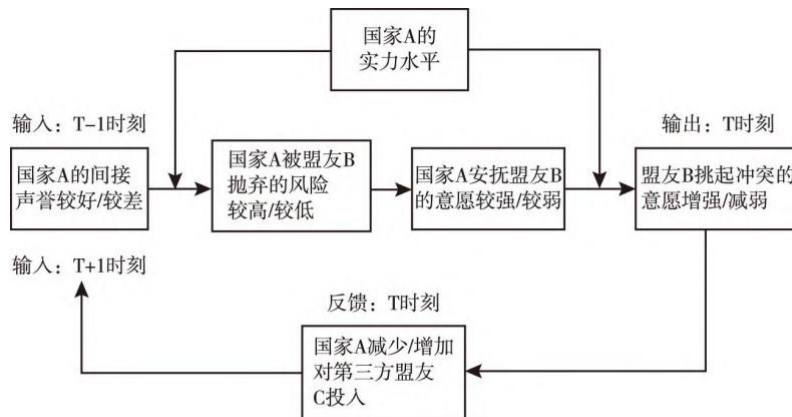


图3 间接声誉负反馈调节机制

资料来源：笔者自制。

假说 1-1：国家的间接声誉越好，盟友挑起冲突的概率越小，强国的间接声誉对盟友挑起冲突的抑制作用小于弱国。

假说 1-2：如果国家在 T-1 时刻的间接声誉过好，其在 T+1 时刻的间接声誉会下滑，强国间接声誉的负反馈小于弱国。

图 4 展示了联盟管理对直接声誉负反馈的影响。良好的直接声誉会提升国家 A 被盟友 B 牵连的风险，此时国家 A 倾向于约束盟友 B，可能减少对盟友 B 的投入并限制盟友 B 使用武力。一旦国家 A 实施约束，盟友 B 挑起冲突的意愿将弱于之前的预期，而约束行为会降低国家 A 在下一阶段的直接声誉。国家对直接声誉的调节效果也受到其间接声誉调节的影响。第一，间接声誉调节盟友 B 挑起冲突的意愿。如果国家 A 同时拥有良好的直接声誉和间接声誉，盟友 B 不仅获得了可靠性的“双保险”，还具备同第三方盟友 C 争夺资源的动机，挑起冲突的意愿较强。例如，在美国处理第一次柏林危机时，李承晚政权既从中看到了美国保卫盟友的可靠性，又担心美国过于关心欧洲事务而减少对自身的支持，因此在 1949 年下半年于“三八线”附近挑起多起武装摩擦。第二，间接声誉能够调节国家 A 实施约束的概率。是否实施约束很大程度上取决于国家 A 与盟友 B 之间的利益分化程度。间接声誉会影响国家 A 与盟友 B 的利益分化程度：间接声誉较好说明国家 A 更重视第三方盟友 C，此时国家 A 更可能约束盟友 B；间接声誉较差则表明国家 A 可能将战略重心收缩或转移至盟友 B，国家 A 与盟友

B 利益一致性较高,加之间接声誉下滑会对国家 A 的履约可靠性造成负面影响,国家 A 更有动力争取盟友 B 的支持。此时国家 A 不仅不会对盟友 B 实施约束,反而可能加大对其的支持。基于此,本文提出两个假说。

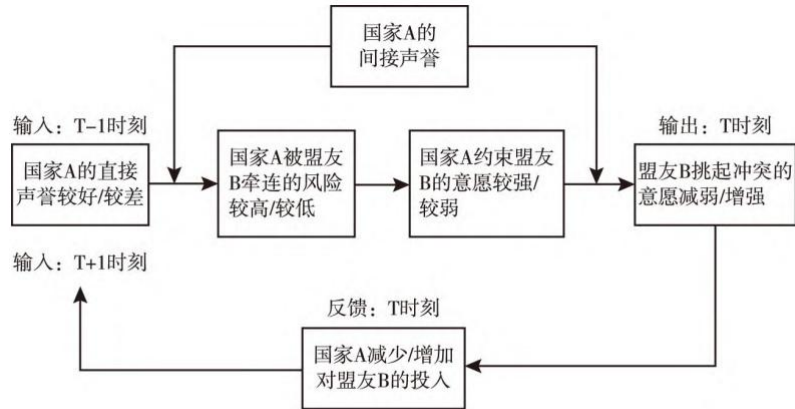


图4 直接声誉负反馈调节机制

资料来源:笔者自制。

假说 2-1: 当间接声誉较好时,国家的直接声誉越好,盟友挑起冲突的概率越小;当间接声誉较差时,国家的直接声誉越好,盟友挑起冲突的概率越大。

假说 2-2: 如果国家在 T-1 时刻的直接声誉过好,其在 T+1 时刻的直接声誉会下滑,间接声誉好的国家的直接声誉负反馈更强。

三、假说检验

在对直接声誉与间接声誉进行操作化的基础上,本文使用双向固定效应模型(two-way fixed effects model)检验相关假说。检验所使用的数据包括两类:一是 1816—2001 年联盟与国家间冲突的数据,二是 1816—2001 年国家的直接声誉与间接声誉数据。两类数据同属面板数据,均以定向二元/年(directed-dyad-year)基本分析单元。前者每条记录的基本形式为“国家 A—盟友 B—潜在冲突对象—年份—冲突结果”,后者每条记录的基本形式为“国家 A—盟友 B—年份—间接声誉/直接声誉”。

（一）联盟声誉的测量

测量联盟声誉是假说检验的前提。布雷特·阿什利·利兹（Brett Ashley Leeds）等倾向于使用违约历史测量联盟声誉。该方法存在两个问题：一是过于强调条约文本。联盟条约往往比较模糊，或者给使用武力设置了较为复杂的约束条件。当盟友卷入冲突时，国家不提供安全支持可能在法律上不违约，但由于“联盟光环（alliance halo）”的存在，盟友依然会认为该国不可靠。过度拘泥于条约文本会将大量冲突场景排除在声誉测量之外，减少国家的“表现”机会。二是标准过高。利兹等主要关注足以导致联盟解体的严重违约行为，在国家间大规模战争与联盟解体的发生概率减小的背景下，此类事件相对罕见。因此，本文同时将国家与盟友围绕低烈度和高烈度冲突展开的历史互动作为声誉来源。相关互动分为履约行为和违约行为两类：前者包括国家同盟友一道挑起针对第三方的冲突以及国家在盟友遭遇攻击时与盟友共同回应两种情况，后者包括国家未同盟友一道挑起冲突以及盟友遭到第三方攻击时国家未向盟友提供支持两种情况。

如前所述，当讨论国家 A 的声誉对盟友 B 的冲突行的影响时，国家 A 的直接声誉来源于国家 A 在盟友 B 卷入冲突时的表现，国家 A 的间接声誉来源于国家 A 在第三方盟友 C 卷入冲突时的表现。本文使用过去十年中国家 A 与盟友 B 或第三方盟友 C 的互动来刻画其声誉水平。式 1 展示了直接声誉的计算方法。分子是国家 A 对盟友 B 的履约记录与违约记录的差值，差值越大表明国家 A 的直接声誉越好。鉴于过去行为的影响力会随着时间的增长而降低，在式 1 中，过去的时间每增加一年，相关行为的重要性减少 0.1。分母是过去十年国家 A 对盟友 B 的履约记录与违约记录之和，代表国家 A 履行承诺的机会。^⑧如果盟友 B 同时与多个国家结盟，本文会选取实力最强的国家作为测量声誉的对象。国家实力根据国家综合实力指标（CINC）确定。^⑨这样的处理主要基于两点考虑：其一，实力是一国声誉发挥作用的基础。其二，在一些多边联盟中，成员之间距离较远、联系较少，只是因在某个（或某些）强国的推动下才结成联盟，并且都依赖强国提供安全保障。

$$\text{直接声誉} = \frac{\sum_{y \in [t-10, t-1]} [(\text{履约}_{AB} - \text{违约}_{AB}) \times (1 - (t - y)/10)]}{\sum_{y \in [t-10, t-1]} [(\text{履约}_{AB} + \text{违约}_{AB}) \times (1 - (t - y)/10)] + 1} \quad \text{式 1}$$

式 2 展示了间接声誉的计算方法。其基本原理与式 1 一致，只是在计算时使用国家 A 与第三方盟友 C 的历史互动。如果国家 A 同一时期拥有多个盟友，本文将除盟友 B 外的所有盟友视作第三方盟友，利用国家 A 与相关盟友的履约和违约记录的差值计算国家 A 对盟友 B 的间接声誉。此外，在稳健性检验中本文还分别使用国家 A 与同盟友 B 处在相同地区（大洲）的盟友以及国家 A 与同盟友 B 处在不同地区（大洲）的盟友的履约违约记录计算国家 A 的间接声誉。

$$\text{间接声誉} = \frac{\sum_{y \in [t-10, t-1]} [(\text{履约}_{AC} - \text{违约}_{AC}) \times (1 - (t - y)/10)]}{\sum_{y \in [t-10, t-1]} [(\text{履约}_{AC} + \text{违约}_{AC}) \times (1 - (t - y)/10)] + 1} \quad \text{式 2}$$

（二）调节变量与控制变量

除间接声誉外，调节变量还包括国家实力。本文使用国家综合实力指标测量国家实力，以国家 A 实力占国家 A 与盟友 B 实力之和的比重刻画该变量。本文在模型中加入了部分控制变量。第一类控制变量是国家 A 的特征，包括大国身份、政体类型和政权更迭。大国身份根据战争相关因素（COW）数据库中的“国家体系资格（state system membership）”数据集确定。政体类型取值来源于“政体 5（Polity5）”数据库。政权更迭根据“领导支持来源变化（CHISOLS）”数据库确定。第二类控制变量是国家 A 与盟友 B 以及盟友 B 与潜在冲突对象的关系，包括毗邻性、实力对比、政体相似性、联盟关系、领土争端与盟友 B 是否美国盟友。毗邻性反映地理接近程度。实力对比是指盟友 B 与潜在冲突对象之间的实力差距，根据国家综合实力指标计算。政体相似性表示盟友 B 与潜在冲突对象是否都为西方式民主国家，政体性质根据“政体 5”确定。联盟关系反映国家 A、盟友 B 与潜在冲突对象是否结盟，根据“联盟条约义务和条款（ATOP）”数据集确定。领土争端反映的是盟友 B 与潜在冲突对象之间是否存在领土纠纷，根据“主题战争相关因素（ICOW）”数据库确定。第三类控制变量属于时间序列变量。冲突可能具有时间依赖性。如果盟友 B 在过去一段时间中频繁挑起冲

突，表明其安全环境较差、外部威胁较严重，其在当下挑起冲突的概率可能较大；如果长时间处在和平中，盟友 B 在当下挑起冲突的概率可能较小。借鉴既有研究，本文将在模型中加入盟友 B 在上一年卷入冲突的情况、过去十年间的冲突历史以及迄今为止经历的和平时间（同时包括和平时间的平方项与立方项）。冲突历史与和平时间根据“国家间军事冲突数据集（MID）”确定，将和平时间做对数处理。

（三）统计结果与分析

假说 1-1 与假说 2-1 中因变量“冲突爆发”的取值根据国家间军事冲突数据集确定，如果盟友 B 在特定年份对潜在冲突对象威胁、展示或使用军事力量，则因变量的取值 1，反之取值为 0。假说 1-2 与假说 2-2 中的因变量为国家 A 在 T+1 时刻的间接声誉和直接声誉。

表 2 是假说 1-1 的检验结果，由于因变量是取值为 0 或 1 的虚拟变量，因此使用 logit 模型。对比模型 1 与模型 2，“间接声誉”的系数符号由正变负，交互项“间接声誉*国家实力”在 99%的置信水平上显著。这说明国家 A 的间接声誉对盟友 B 挑起冲突概率的“净效应”为负，间接声誉的作用与国家实力相关。借助边际效应图（图 5）可以更好理解间接声誉与冲突爆发的关系。在国家 A 的实力较弱时，其间接声誉与盟友 B 挑起冲突的概率呈负相关；在国家 A 的实力较强时，其间接声誉与盟友 B 挑起冲突概率的负相关程度逐渐下降；当“国家实力”的取值趋近 1 时，国家 A 的间接声誉与盟友 B 挑起冲突的概率之间甚至转为正相关。

控制变量结果基本符合预期：“冲突历史”和“领土争端”系数符号始终显著为正，说明盟友 B 面临外部威胁越严重，其挑起冲突的可能性越高。“政权更迭”的系数符号显著负，表示国家 A 领导人的变化会降低盟友 B 获得支持的预期，进而减少其挑起冲突的概率。“实力对比”始终显著为正，代表盟友 B 具有实力优势时更容易挑起冲突。

模型 3 与模型 4 分别使用国家 A 与同盟友 B 处于相同地区（大洲）的第三

方盟友以及同盟友 B 处于不同地区（大洲）的第三方盟友的互动来衡量国家 A 的间接声誉（下文区分为“地区间接声誉”与“地区外间接声誉”）。两个模型中“间接声誉”和“间接声誉*国家实力”的系数符号同模型 2 一致。“盟友 B 是否美国盟友”变量的系数在四个模型中均不显著，这表明相关理论并非只适用于美国联盟体系。考虑到资源分配可能受到冲突类型的影响，模型 5 将因变量限制在爆发高烈度冲突。高烈度冲突的判断标准参考“国家间军事冲突”数据集关于国家在冲突中最强敌意分类。该数据集将敌意程度分为“无冲突行为”“威胁使用武力”“展示武力”“使用武力”“战争”五类。本文以是否使用武力为标准对冲突烈度分类。如果盟友 B 在冲突中的最强敌意行属于“使用武力”或“战争”，则视为挑起高烈度冲突。模型 5 表明统计结果依然稳健。

表 2 Logit 回归：间接声誉、国家实力与冲突爆发

| 自变量/控制变量 | 因变量：冲突爆发 | | | | |
|----------|------------------|-------------------|---------------------|-------------------|------------------|
| | 模型 1 | 模型 2 | 模型 3 | 模型 4 | 模型 5 |
| 间接声誉 | 0.227 (0.149) | -0.280 (0.224) | -0.552** (0.227) | -0.069 (0.210) | 0.064 (0.280) |

| 自变量/控制变量 | 因变量：冲突爆发 | | | | |
|----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | 模型 1 | 模型 2 | 模型 3 | 模型 4 | 模型 5 |
| 国家实力 | -0.397 (0.344) | 0.130 (0.383) | -0.011 (0.369) | -0.122 (0.359) | 0.199 (0.473) |
| 间接声誉 * 国家实力 | | 1.141 *** (0.363) | 1.169 *** (0.343) | 0.705 ** (0.334) | 1.009 ** (0.435) |
| 冲突历史 (过去 1 年) | 0.096 *** (0.024) | 0.097 *** (0.024) | 0.100 *** (0.024) | 0.096 *** (0.024) | 0.096 *** (0.029) |
| 联盟关系 (国家 A 与潜在冲突对象) | -0.263 * (0.158) | -0.262 * (0.158) | -0.290 * (0.160) | -0.274 * (0.158) | -0.346 * (0.196) |
| 政权更迭 | -0.431 *** (0.095) | -0.421 *** (0.094) | -0.429 *** (0.094) | -0.420 *** (0.094) | -0.544 *** (0.124) |
| 国家 A 政体类型 | 0.038 (0.185) | 0.129 (0.186) | 0.046 (0.188) | 0.103 (0.189) | 0.237 (0.229) |
| 毗邻性 (国家 A 与潜在冲突对象) | 0.021 (0.036) | 0.020 (0.036) | 0.020 (0.036) | 0.019 (0.036) | 0.018 (0.046) |
| 实力对比 | 1.086 * (0.589) | 1.107 * (0.588) | 1.070 * (0.589) | 1.110 * (0.588) | 2.012 *** (0.750) |
| 毗邻性 (盟友 B 与潜在冲突对象) | 0.010 (0.070) | -0.005 (0.070) | -0.017 (0.070) | 0.004 (0.070) | -0.005 (0.096) |
| 政体相似性 (盟友 B 与潜在冲突对象) | -0.198 (0.145) | -0.220 (0.145) | -0.191 (0.145) | -0.218 (0.144) | 0.084 (0.187) |
| 联盟关系 (盟友 B 与潜在冲突对象) | -0.233 (0.151) | -0.237 (0.151) | -0.230 (0.152) | -0.233 (0.152) | -0.452 ** (0.183) |
| 领土争端 | 0.790 *** (0.227) | 0.803 *** (0.227) | 0.789 *** (0.227) | 0.783 *** (0.226) | 0.643 ** (0.282) |
| 盟友 B 是否为美国盟友 | 0.080 (0.293) | 0.097 (0.291) | 0.193 (0.296) | 0.179 (0.293) | 0.316 (0.345) |
| 和平时间 | -0.192 (1.128) | -0.170 (1.129) | -0.233 (1.131) | -0.041 (1.130) | -0.963 (1.449) |
| 和平时间 (平方项) | -1.663 (1.230) | -1.659 (1.230) | -1.583 (1.233) | -1.829 (1.231) | -0.987 (1.595) |
| 和平时间 (立方项) | 0.898 ** (0.392) | 0.892 ** (0.392) | 0.865 ** (0.393) | 0.956 ** (0.393) | 0.741 (0.511) |
| 观测值 | 16733 | 16733 | 16733 | 16733 | 16733 |
| 个体固定效应 | 是 | 是 | 是 | 是 | 是 |
| 时间固定效应 | 是 | 是 | 是 | 是 | 是 |

资料来源：笔者自制。

注：括号中数字是稳健标准误，***、**、* 分别代表在 1%、5% 和 10% 水平上的统计显著性。

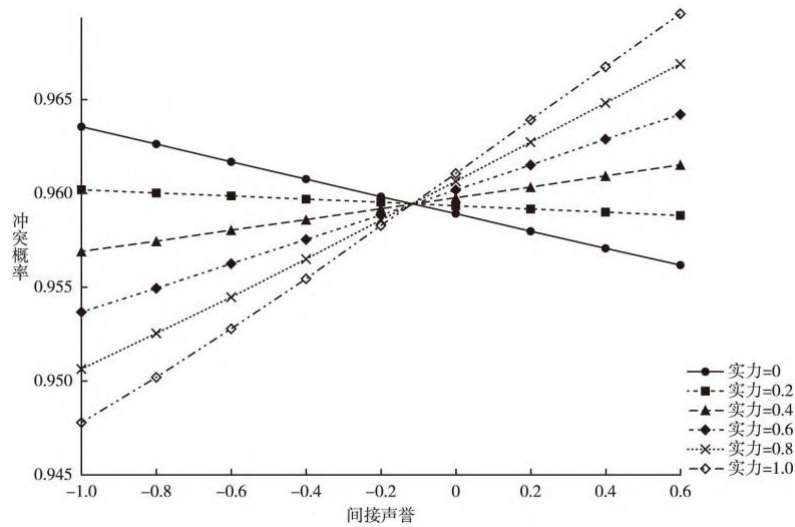


图5 国家实力对间接声誉影响的调节

资料来源：笔者自制。

表3是对假说1-2的检验。因变量间接声誉属于连续变量，因而使用普通最小二乘法（OLS）回归模型。本文认为，当间接声誉在一定范围内时，国家A被盟友B抛弃的风险不明显、安抚盟友B的动力不强，其在T-1时刻的间接声誉与其在T+1时刻的间接声誉仍呈正相关。如果间接声誉过好，国家A会安抚盟友B，从而导致其在T+1时刻的间接声誉下滑，因此在模型中加入了间接声誉的平方项。在模型6中，“间接声誉”的系数符号显著为正，“间接声誉”的平方项显著为负。这印证了本文的假说，即一国T-1时刻的间接声誉与T+1时刻的间接声誉呈倒U形关系，较好的间接声誉会促进下一阶段的间接声誉，但过好的间接声誉反而会抑制下一阶段的间接声誉。交互项“间接声誉*间接声誉*国家实力”的系数符号为正，说明与弱国相比，强国的间接声誉负反馈较弱。模型7和模型8中的自变量分别为国家A在T-1时刻的地区间接声誉和地区外间接声誉，交互项回归系数符号与模型6一致。

表 3 OLS 回归：国家实力调节下的间接声誉负反馈

| 自变量/控制变量 | 因变量：间接声誉 (T+1) | | |
|--------------------|----------------------|----------------------|----------------------|
| | 模型 6 | 模型 7 | 模型 8 |
| 间接声誉 | 0.355*** (0.023) | 0.341*** (0.020) | 0.344*** (0.036) |
| 国家实力 | 0.210*** (0.036) | 0.071** (0.029) | 0.159*** (0.019) |
| 间接声誉 * 间接声誉 | -0.161*** (0.056) | -0.218*** (0.055) | -0.103 (0.067) |
| 间接声誉 * 间接声誉 * 国家实力 | 0.186** (0.081) | 0.390*** (0.085) | 0.313*** (0.093) |
| 毗邻性 | 0.025*** (0.005) | 0.007* (0.004) | 0.021*** (0.003) |
| 大国身份 | 0.113*** (0.029) | 0.104*** (0.036) | 0.080*** (0.028) |
| 政体相似性 | 0.169*** (0.021) | 0.167*** (0.019) | 0.108*** (0.015) |
| 冲突历史 | 0.008* (0.005) | 0.009** (0.004) | 0.006* (0.003) |
| 常数项 | -0.192*** (0.023) | -0.120*** (0.020) | -0.100*** (0.012) |
| 观测值 | 19614 | 19614 | 19614 |
| R ² | 0.294 | 0.243 | 0.424 |
| 个体固定效应 | 是 | 是 | 是 |
| 时间固定效应 | 是 | 是 | 是 |

资料来源：笔者自制。

注：括号中数字是稳健标准误，***、**、* 分别代表在 1%、5% 和 10% 水平上的统计显著性。

表 4 是对假说 2-1 所做的统计检验。模型 9 中交互项“间接声誉*直接声誉”的系数符号显著为负。图 6 是根据模型 9 结果绘制的边际效应图：当国家 A 的间接声誉较差时，其直接声誉越好，盟友 B 挑起冲突的概率越大；当国家 A 的间接声誉较好时，其直接声誉越好，盟友 B 挑起冲突的概率越小。模型 10 与模型 11 中的调节变量分别为国家 A 的地区间接声誉和地区外间接声誉，两个模型中交互项“间接声誉*直接声誉”的系数符号与模型 9 一致。模型 12 以高烈度冲突为因变量，交互项系数符号依然稳健。

表 4 Logit 回归：间接声誉、直接声誉与冲突爆发

| 自变量/控制变量 | 因变量：冲突爆发 | | | |
|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 模型 9 | 模型 10 | 模型 11 | 模型 12 |
| 间接声誉 | -1.028** (0.446) | -1.072*** (0.358) | -0.235 (0.413) | -0.748 (0.546) |
| 直接声誉 | -0.115 (0.468) | 0.026 (0.435) | 0.162 (0.345) | -0.433 (0.592) |
| 间接声誉 * 直接声誉 | -1.066** (0.439) | -0.839*** (0.321) | -0.300 (0.380) | -1.199** (0.540) |
| 国家实力 | -0.430 (0.572) | -0.674 (0.586) | -0.621 (0.469) | -0.026 (0.734) |
| 间接声誉 * 国家实力 | 1.531*** (0.475) | 1.569*** (0.411) | 0.503 (0.402) | 1.516*** (0.572) |
| 直接声誉 * 国家实力 | -0.917* (0.530) | -0.875* (0.515) | -0.592 (0.492) | -0.557 (0.665) |
| 冲突历史 (过去 1 年) | 0.079*** (0.025) | 0.074*** (0.025) | 0.132*** (0.019) | 0.082*** (0.031) |
| 联盟关系 (国家 A 与潜在冲突对象) | -0.403** (0.175) | -0.393** (0.176) | -0.313* (0.166) | -0.486** (0.218) |
| 政权更迭 | -0.440*** (0.098) | -0.456*** (0.098) | -0.314*** (0.087) | -0.558*** (0.131) |
| 国家 A 政体类型 | 0.252 (0.202) | 0.265 (0.207) | 0.040 (0.181) | 0.414 (0.252) |
| 毗邻性 (国家 A 与潜在冲突对象) | -0.003 (0.041) | 0.000 (0.041) | -0.006 (0.040) | -0.043 (0.053) |
| 实力对比 | 1.006 (0.630) | 1.061* (0.628) | 0.913 (0.593) | 2.124*** (0.800) |
| 毗邻性 (盟友 B 与潜在冲突对象) | -0.009 (0.070) | -0.011 (0.070) | 0.002 (0.069) | -0.013 (0.097) |
| 政体相似性 (盟友 B 与潜在冲突对象) | -0.233 (0.147) | -0.190 (0.146) | -0.265* (0.142) | 0.079 (0.192) |
| 联盟关系 (盟友 B 与潜在冲突对象) | -0.282* (0.159) | -0.282* (0.159) | -0.119 (0.149) | -0.523*** (0.193) |
| 领土争端 | 0.754*** (0.245) | 0.749*** (0.244) | 0.657*** (0.240) | 0.582* (0.313) |

| 自变量/控制变量 | 因变量：冲突爆发 | | | |
|--------------|-------------------|-------------------|--------------------|-------------------|
| | 模型 9 | 模型 10 | 模型 11 | 模型 12 |
| 盟友 B 是否为美国盟友 | -0.236 (0.311) | -0.084 (0.313) | -0.234 (0.239) | 0.111 (0.372) |
| 和平时间 | -0.476 (1.198) | -0.513 (1.200) | -0.579 (1.169) | -0.930 (1.559) |
| 和平时间 (平方项) | -1.297 (1.299) | -1.244 (1.302) | -1.266 (1.274) | -0.997 (1.713) |
| 和平时间 (立方项) | 0.792* (0.413) | 0.769* (0.414) | 0.811** (0.407) | 0.764 (0.548) |
| 观测值 | 14941 | 14941 | 14941 | 14941 |
| 个体固定效应 | 是 | 是 | 是 | 是 |
| 时间固定效应 | 是 | 是 | 是 | 是 |

资料来源：笔者自制。

注：括号中数字是稳健标准误，***、**、* 分别代表在 1%、5% 和 10% 水平上的统计显著性。

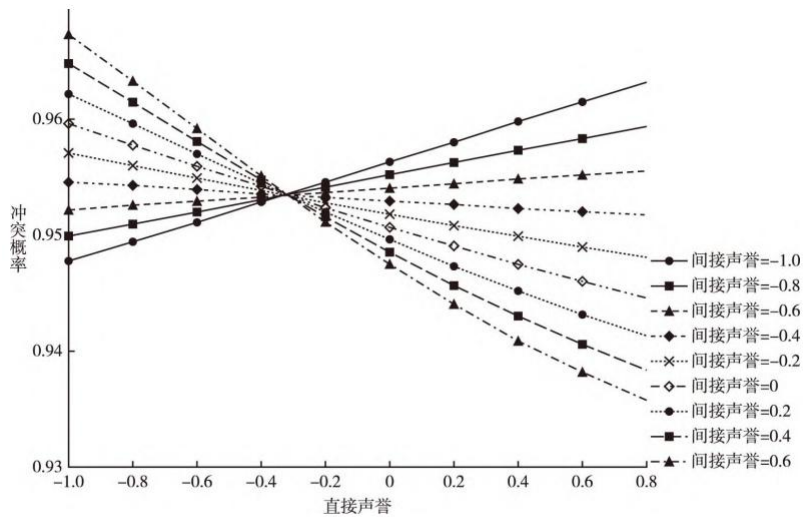


图6 间接声誉对直接声誉影响的调节

资料来源：笔者自制。

表5是针对假说2-2的检验。模型13中“直接声誉”的系数符号为正，平方项“直接声誉*直接声誉”的系数符号为负，表明当国家A在T-1时刻的直接声誉过高时，其在T+1时刻的直接声誉可能下滑。交互项“直接声誉*直接声誉*间接声誉”的系数符号显著为负，且数值均大于平方项。这表示与间接声誉较差的国家相比，间接声誉较好的国家的直接声誉负反馈效应更强。模型14与模型15中的调节变量分别为地区间接声誉和地区外间接声誉，其中交互项“直接声誉*直接声誉*间接声誉”的系数符号与模型13一致，这表明统计结果具有稳健性。

表 5 OLS 回归：间接声誉调节下的直接声誉负反馈

| 自变量/控制变量 | 因变量：直接声誉 (T+1) | | |
|--------------------|----------------------|----------------------|----------------------|
| | 模型 13 | 模型 14 | 模型 15 |
| 间接声誉 | 0.316*** (0.032) | 0.290*** (0.030) | 0.300*** (0.028) |
| 直接声誉 | 0.024* (0.014) | 0.042*** (0.014) | 0.002 (0.019) |
| 国家实力 | 0.011 (0.016) | 0.010 (0.016) | 0.017 (0.016) |
| 直接声誉 * 直接声誉 | -0.020 (0.040) | -0.040 (0.038) | -0.022 (0.041) |
| 直接声誉 * 直接声誉 * 间接声誉 | -0.136** (0.061) | -0.077 (0.065) | -0.077 (0.071) |
| 毗邻性 | 0.014*** (0.003) | 0.014*** (0.003) | 0.014*** (0.003) |
| 大国身份 | 0.088** (0.043) | 0.086** (0.043) | 0.088** (0.042) |
| 政体相似性 | 0.097*** (0.013) | 0.095*** (0.013) | 0.100*** (0.013) |
| 冲突历史 | 0.007** (0.003) | 0.007** (0.003) | 0.007** (0.003) |
| 常数项 | -0.068*** (0.014) | -0.066*** (0.014) | -0.073*** (0.014) |
| 观测值 | 22363 | 22363 | 22363 |
| R ² | 0.182 | 0.183 | 0.181 |
| 个体固定效应 | 是 | 是 | 是 |
| 时间固定效应 | 是 | 是 | 是 |

资料来源：笔者自制。

注：括号中数字是稳健标准误，***、**、* 分别代表在 1%、5% 和 10% 水平上的统计显著性。

四、美韩同盟中的联盟声誉与冲突爆发

本部分通过分析 1961—1976 年美韩同盟案例展示联盟声誉与盟友挑起冲突之间的关系。选择美韩同盟作为分析案例主要基于三点考量：第一，美韩同盟是典型的防御性联盟。美韩同盟是朝鲜战争的产物，在美苏冷战的大背景下，朝鲜半岛成为美苏两个超级大国在亚洲对峙的前沿。在 1953 年《朝鲜停战协定》即将达成之际，韩国向美国提出结盟要求，希望在战争结束后仍能获得美国的保护。出于遏制“共产主义扩张”的目的，美国同韩国签署了《美韩共同防御条约》。第二，美国具有较强的实力优势，有能力对韩国政策发挥影响。第三，在所关注的时段内，美国的联盟声誉与韩国挑起冲突的行为均发生了显著变化。

（一）“青瓦台事件”与“普韦布洛”号危机

对韩国而言，美国在 20 世纪 50 年代形成了较好的间接声誉并延续至 60 年代。在《朝鲜停战协定》尚未签署之际，美国开始将目光投向东南亚地区。1953 年 4 月越南人民军与老挝巴特寮武装发动联合行动，艾森豪威尔政府表示，“如果老挝丢失，可能失去东南亚其他地区”。1954 年 1 月，美国出台国家安全委员会第 5405 号文件，设想美军在法国撤出后一定程度上参与中南半岛的战争。1954 年 9 月，美国与英国、法国、菲律宾和泰国等国基于《东南亚集体防务条约》组建东南亚条约组织，旨在对抗所谓共产党力量的“武装进攻”与“颠覆”。进入 20 世纪 60 年代，美国对东南亚的干涉力度加大，主要包括两方面：一是强化联盟建设。由于东南亚条约组织相对松散且采取联合行动须所有成员一致同意，力提高行动效率美国尝试同地区国家建设双边联盟。1962 年 3 月，泰国外交部部长他纳·科曼（Thanat Khoman）访美并与美国国务卿迪安·腊斯克（Dean Rusk）发表联合公报称，美国在帮助泰国抵御“共产党侵略和颠覆”上立场坚定，会全面执行东南亚条约组织规定的义务，且相关行动不需要“条约其他成员国事先同意”。二是军事干涉力度升级。1963 年年底约翰逊政府上台后，美国逐渐扩大在越南战场的投入，轰炸越南北方，与泰国、菲律宾和澳大利亚等盟友开展联合干涉，并直接派遣军队赴越南作战。

20 世纪 60 年代的美国在维持良好间接声誉的同时并未对韩国置之不理。作为超级大国，美国的对外政策目标具有全球性，在重视东南亚的同时依然对东北亚予以关注。1961 年 12 月，美国国务院提交了一份关于军事援助的报告，认为韩国应“保持足够的军事力量”，只有确定“常规军事力量为改变驻韩美军态势创造额外基础”时，才考虑将对韩国的军事援助转向发展援助。在驻韩美军问题上，1963 年 4 月，美国参谋长联席会议提出缩减韩国军队与驻韩美军的规模，但被国务院否定。美国之所以维持和加强对韩国的军事援助与安全承诺，除军政部门之间的竞争、传统政策惯性和应对朝鲜等原因外，声誉因素也发挥了作用。由美国国防部、参谋长联席会议与国际开发署等部门代表组成的军事援助问题指导小组在 1962 年 1 月提交报告，建议美国应促使受援国自身承担更多的国防开支，把节省下来的军事援助资金投入国际开发援助。但美国国家安全顾问麦乔治

•邦迪（MeGeorge Bundy）认为，在欧洲和东南亚局势因第二次柏林危机和老挝内战再度紧张背景下，削减美国的军事援助会提升盟友面临的风险；即使在相对平静的时期，盟友也会认为军事援助缩减意味着美国的安全保障弱化。美国的外交部门也持相似的观点，美国副国务卿帮办乌拉尔·亚历克西斯·约翰逊（Ural Alexis Johnson）在给参谋长联席会议的信件中称，维持美国驻韩军队的规模是必要的，减少美军数量会让韩国对美国的承诺产生怀疑。

美国良好的直接声誉与间接声誉影响了韩国在朝韩非军事区的冲突行为。1953年《朝鲜停战协定》在朝鲜半岛设置了以“三八线”为基础、以实际控制地为起点的军事分界线，各方军队从分界线后退2千米，由此形成一个宽度4千米的非军事区。到20世纪60年代中期，朝韩双方在非军事区的武装行动逐渐增多。据不完全统计，1964年非军事区发生的摩擦有32起，到1967年则超过400起。在美国拥有良好直接声誉的背景下，虽然包括副总统和国务卿等在内的美国高级官员均希望韩国限制在非军事区的行动，但韩国总统朴正熙（Park Chung-Hee）始终不为所动。韩国国防部组织了一支2400余人的反渗透部队专门在非军事区执行任务。自1966年年底至1968年2月，韩国平均每个月在非军事区发起2起突袭。面对这样的情势，美国担心遭到韩国牵连，其良好的间接声誉进一步加剧了美国的担忧程度。

1967年6月，美国中央情报局分析认为，朝鲜在非军事区的行动“主要是情报收集和分析，而不是武力破坏”，针对美韩军人的袭击数量增多是对美韩两国在越南战场作战的回应。围绕非军事区的武装行动在1968年达到顶峰。1968年1月21日，朝鲜发动了“青瓦台事件”，31名朝鲜特工越过非军事区计划突袭韩国总统府，该行动虽未成功但却震惊韩国朝野。两天后，美国电子侦察船“普韦布洛”号在朝鲜东海岸进行情报侦察时遭到朝鲜扣押，23名船员成为人质。“青瓦台事件”与“普韦布洛”号危机让朝鲜半岛再次处在战争边缘。韩国将上述事件视作朝鲜企图发动战争的证据，主张予以强硬回击。而美国分析认为，朝鲜扣押“普韦布洛”号是为在越南问题上向美国施压。基于这样的认知，美国在制定

应对措施时的出发点是避免在越南以外再开辟新战场。因此美国选择对韩国实施约束，防止朴正熙政府挑起大规模冲突。1968年2月，美国国防部副部长塞勒斯·万斯（Cyrus Vance）作为总统特使出访韩国。万斯声称，如果出现因韩国单独行动导致战争爆发的局面，美军将从韩国撤出。在美国的约束下，韩国挑起冲突的意愿显著下降。朴正熙政府向美国保证，韩国不会因“青瓦台事件”与“普韦布洛”号危机对朝鲜实施报复，韩国政府将安抚民众情绪，在采取相关行动前会同美国协商。

美国的表现破坏了其对韩国的直接声誉。“普韦布洛”号危机发生后，韩国对美国可靠性的怀疑程度上升，尝试走自主国防道路。1968年2月，朴正熙在讲话中表示，“如果没有首先依靠自身力量来防御的决心的话，就不能拥有这样的国防”。

（二）“板门店事件”

1969年1月尼克松政府上台，其任内美国对韩国的直接声誉和间接声誉均出现下滑。美国的直接声誉受损主要受两个因素影响：一是削减驻韩美军。1969年7月，尼克松发表亚洲政策演讲，要求亚洲盟友承担更多的防卫责任。1970年3月，尼克松签署第48号国家安全行动备忘录（NSDM），计划在1971年结束前削减2万名驻韩美军。随后，美国陆军第七师从韩国撤出。二是不介入朝韩海上争端。20世纪70年代，朝韩两国的海上争端逐渐增多。《朝鲜停战协定》并未明确划分海上停火线，只是将韩国西北部海域“联合国军”所控制的白翎岛、大青岛、小青岛、延坪岛及隅岛诸岛连线作实际控制线。1965年1月，美国控制的“联合国军”海军司令部单方面宣布将实控线作为停火线在海上的延伸，即“北方界线（NLL）”。朝鲜方面并不承认“北方界线”，并针对性地提出“南方警戒线”。两条界线之间的海域成为朝韩海上冲突的焦点区。1974年6月，朝鲜炮艇与韩国海警船交火并最终将该船击沉。1975年2月，韩国驱逐舰在“北方界线”以南击沉一艘朝鲜渔船。事后，朝鲜向相关海域派遣包含扫雷艇、炮舰与导弹舰的舰队，并先后出动85架次战斗机飞越相关岛屿。在这一系列冲突中，

美国不仅未向韩国提供军事支持，反而一定程度上默许了朝鲜的活动。对其间接声誉造成更大破坏的事件则属美国坐视越南南方政权败亡。越南南北双方在《巴黎协定》签署后仍未停止作战。1975年4月30日，越南北方军队占领西贡，越南南方政权宣告灭亡。美国坐视“西贡陷落”引发东南亚盟友对其的深刻怀疑。1975年7月，泰国和菲律宾发表公报宣布准备解散东南亚条约组织，并同意中南半岛国家加入东南亚国家联盟。

越南战争结束后，美国在东南亚的影响力大大下降，但也为其恢复在东北亚的影响提供契机。福特政府上台后提出“新太平洋主义”，试图恢复美国对韩国的直接声誉，强调美国利用实力“维持太平洋均势”，再次发挥驻韩美军的作用。美国国防部部长詹姆斯·施莱辛格（James R. Schlesinger）于1975年8月访问韩国，向朴正熙承诺不会继续削减驻韩美军。1976年6月28日至7月7日，美韩两国举行了海军与海军陆战队的联合训练。

除宣示安全承诺与展示武力外，福特政府还倾向于在危机或冲突中展现强硬姿态来提高声誉、彰显安全保证的可靠性。恰在此时，“板门店事件”为美国提供了表现机会，这是一起发生在共同警备区（JCA）的流血冲突。共同警备区位于非军事区内，以板门店为中心，直径约800米。朝鲜和美韩方面各自在共同警备区内设置哨所、执行警戒任务。1976年8月18日，11名美韩军人与5名韩国工人在共同警备区内对一棵杨树进行砍伐作业。作业过程中，朝鲜士兵出现并阻止，双方爆发肢体冲突，导致两名美军军官死亡、多名美军士兵受伤。事件发生后，美国决定“必须予以强硬的回应”，将驻韩美军的防卫等级从四级提升至三级，属于爆发战争前的最高等级，也是1953年以来驻韩美军首次将防卫等级提升至三级。此外，美国还进行了大规模的武力部署和展示，如从日本调遣一支F-4战斗机中队、从美国本土派出18架F-111战斗机、举行由可搭载核武器的B-52轰炸机参加的演习、将“中途岛”号航母部署至朝鲜西南部黄海海域。同年8月21日，美韩军队继续执行砍树任务，其间未受到朝鲜方面的干扰。行动结束后1小时，朝美两国的停战委员会代表在板门店举行了私人会见。朝鲜方面

转交了金日成（Kim Il-Sung）给理查德·史迪威（Richard G. Stilwell）的信件。金日成在信中对“板门店事件”表达了遗憾，并表示朝鲜不会率先挑衅。1976年9月7日，驻韩美军将防卫等级降回4级，并减少了B-52轰炸机的演习次数。

在应对“板门店事件”时，美国的做法比韩国更加激进。在1976年8月18日的会议后，美国向朴正熙做了通报。朴正熙同意解决砍树问题，但提醒美国必须对朝鲜可能的回应做准备。朴正熙希望“不使用武器”，以免进一步刺激朝鲜，并表示可以挑选精通跆拳道的韩国士兵参与行动。美国设想了朝鲜两种可能的军事反应：一是朝鲜军队与美韩军队在共同警备区内展开混战，二是朝鲜发动了旨在占领共同警备区的军事行动。作为对第一种情况的回应，美国将对朝鲜设施进行大规模炮击，为美韩军队撤离提供掩护；如果第二种情况发生，美韩支援部队将进入共同警备区。

五、结论

一国的联盟声誉对盟友挑起冲突的概率将发挥何种影响？为回答这一问题，本文区分了国家在联盟中的直接声誉和间接声誉，聚焦两类声誉对盟友挑起冲突概率的不同影响机制。本文认为，对特定盟友而言，国家的直接声誉和间接声誉一定程度上不可通约。直接声誉主要反映国家可靠性，国家的直接声誉越好、其履行联盟承诺的概率越大，盟友挑起冲突的意愿也越强。间接声誉主要反映国家资源分配偏好，国家的间接声誉越好、其资源分配越向第三方盟友倾斜，特定盟友挑起冲突的意愿随之越弱。同时，本文特别指出国家会根据自身面临的“被牵连”和“被抛弃”风险程度主动调节联盟声誉水平。因此，联盟声誉并非单纯只是盟友挑起冲突的“加速器”或“制动阀”，二者的关系呈现非线性特点。

本文的理论价值主要包含三方面：一是厘清了直接声誉与间接声誉的联系与区别。与直接声誉相似，间接声誉的来源也是国家与盟友的既往履约和违约记录，在一国与盟友定向历史互动不足的情况下，间接声誉也会影响盟友对该国可靠性的判断。与直接声誉不同，间接声誉主要体现国家对不同盟友的资源分配态势。

二是指出联盟系统中存在声誉的正负反馈机制。声誉具有时间依赖性，国家在 T-1 时刻的声誉对 T 时刻盟友的冲突行为发挥作用，同时也影响自身在 T+1 阶段的声誉水平。正负反馈机制揭示了国家联盟声誉水平变化的动因和盟友冲突行为的反复。三是对国家的联盟声誉进行了操作化测量。本文依据国家在联盟中的履约和违约历史构建了直接声誉和间接声誉数据，并使用统计模型进行假说检验。

历史上英国、德国、美国和苏联等大国都曾在地区或全球范围建立联盟体系，与不同地区的多个盟友产生互动。本文理论为解释此类国家何时安抚与约束特定盟友提供了解释。此外，本文还有一定的现实政策意义。当前，联盟已成为美国对华政策的重要工具，中国周边安全环境受到美国亚太联盟体系的深刻影响。^①根据本文的理论预测，一方面，拜登政府推行对中俄“双遏制”政策，对美国亚太盟友同时维持较高水平的直接声誉和间接声誉，这会提升美国亚太盟友挑起冲突的意愿。当亚太盟友的行为逐渐提升美国“被牵连”的风险时，美国以限制安全承诺触发条件等方式约束其亚太盟友的概率会增大。另一方面，在国际格局“东升西降”持续演化的背景下，美国同时对亚太盟友维持较高水平的直接声誉（加大对亚太盟友的投入）和间接声誉（加大对欧洲盟友的投入）的难度会越来越大，美国未来可能会减少对欧洲盟友的投入。

全球治理研究团队

任琳 熊爱宗 吴国鼎 陈兆源 黄宇韬
韩永辉 宋锦 沈陈 田旭 彭博

研究助理团队

兰馨彤 苏山岳 孟思宇 杨嘉豪