

实力与决心的较量

——三方不对称军事威慑博弈分析

曹金绪◎

【内容提要】 “狐假虎威”现象在国际政治中较为普遍，其主要特征是弱国借助强大盟国的庇护或支持，挑战非盟友国家并获取利益。为解释这一现象，本文建构了一个三方不对称军事威慑博弈模型，简要分析了三个不同类型的国家行为体之间的策略互动过程。基于对完全信息和不完全信息条件下三方不对称军事威慑博弈的均衡分析，本文揭示了隐藏在“狐假虎威”现象背后的内在规律和因果逻辑。本文发现，威慑者、挑战者和庇护者三类国家在实力、决心上的不同较量，影响并决定了“狐假虎威”现象的发生与演变结局。

【关键词】 军事威慑 实力 决心 不对称

一、研究问题与方法

在众多的国际关系研究议题中，军事威慑和同盟政治是两个非常经典并且成果极为丰硕的研究领域。目前有关同盟与威慑相互关系的研究主要是在延伸威慑（extended

《国际政治科学》2013年第2期（总第34期），第1—34页。

Quarterly Journal of International Politics

deterrence) 概念框架下进行。延伸威慑研究关注的核心问题是不对称同盟中处于盟主地位的大国如何依靠自身的军事实力慑阻他国攻击自己的盟友。延伸威慑, 就是扩展威慑效应, 不仅维护自己国家的安全, 而且还将盟友, 或者准盟友的安全也纳入保护的领地。在冷战时期, 西方延伸威慑研究的核心问题是如何发挥美国核保护伞的作用以保护其盟友的安全, 这一现实需求推动众多学者深入研究延伸威慑现象。^① 冷战后, 美国成为唯一的超级大国, 拥有庞大的同盟网络, 其他国家则不具备这样的优势和地位, 为维护美国的霸权地位以及同盟关系, 美国学者在延伸威慑研究上继续投入大量精力, 力求更全面地解释延伸威慑现象。^② 不过在这些研究中, 研究者总是从同盟内部成员的角度思考问题, 只关心如何慑阻同盟外部国家挑战, 以维护同盟的安全利益。对于非西方盟友而言, 关注点则有所不同, 他们更关心的是如何应对强大的西方军事同盟造成的安全威胁。

① Erich Weede, "Extended Deterrence by Superpower Alliance," *Journal of Conflict Resolution*, Vol. 27, No. 2, 1983, pp. 231 - 254; John J. Mearsheimer, "Nuclear Weapons and Deterrence in Europe," *International Security*, Vol. 9, No. 3, 1984/1985, pp. 19 - 46; Janice Gross Stein, "Extended Deterrence in the Middle East: American Strategy Reconsidered," *World Politics*, Vol. 39, No. 3, 1987, pp. 326 - 352; Paul K. Huth, "Extended Deterrence and the Outbreak of War," *American Political Science Review*, Vol. 82, No. 2, 1988, pp. 423 - 443; Bruce Martin Russett, "Extended Deterrence with Nuclear Weapons: How Necessary, How Acceptable?" *The Review of Politics*, Vol. 50, No. 2, 1988, pp. 282 - 302; Erich Weede, "Extended Deterrence, Superpower Control, and Militarized Interstate Disputes, 1962 - 1976," *Journal of Peace Research*, Vol. 26, No. 1, 1989, pp. 7 - 17; Fred Chernoff, "Stability and Alliance Cohesion: The Effects of Strategic Arms Reductions on Targeting and Extended Deterrence," *Journal of Conflict Resolution*, Vol. 34, No. 1, 1990, pp. 92 - 101; Samuel S. G. Wu, "To Attack or Not to Attack: A Theory and Empirical Assessment of Extended Immediate Deterrence," *Journal of Conflict Resolution*, Vol. 34, No. 3, 1990, pp. 531 - 552.

② D. Marc Kilgour and Frank C. Zagare, "Uncertainty and the Role of the Pawn in Extended Deterrence," *Synthese*, Vol. 100, No. 3, 1994, pp. 379 - 412; Alastair Smith, "Extended Deterrence and Alliance Formation," *International Interactions*, Vol. 24, No. 4, 1998, pp. 315 - 343; Vesna Danilovic, "The Sources of Threat Credibility in Extended Deterrence," *Journal of Conflict Resolution*, Vol. 45, No. 3, 2001, pp. 341 - 369; Frank C. Zagare and D. Marc Kilgour, "Alignment Patterns, Crisis Bargaining, and Extended Deterrence: A Game - Theoretic Analysis," *International Studies Quarterly*, Vol. 47, No. 4, 2003, pp. 587 - 615; Curtis S. Signorino and Ahmer Tarar, "A Unified Theory and Test of Extended Immediate Deterrence," *American Journal of Political Science*, Vol. 50, No. 3, (转下页注)

由于威慑理论研究的重镇是美国和欧洲，他们并不关心这些“非西方视角”的威慑现象和问题。在大量的延伸威慑理论研究文献中，有一类现象长期被学者们忽视。这类现象俗称“狐假虎威”，即实力较弱的国家依靠强大盟友的庇护和威慑力，主动挑战非盟友国家，寻求难以单独获取的利益。一般情况下，离开了盟友的庇护，实力较弱的国家不会主动挑战强大对手，但是由于强大盟友的存在，弱国的挑衅性将会增加，这给非盟友国家造成极大的困扰和压力。同盟外部国家在决策时不得不考虑强大庇护者干预的问题，如果策略失当将会面对与整个军事同盟对抗的危险。扎加雷和基尔戈（Zagare and Kilgour）注意到这一现象，不过因其“西方视角”，他们仅研究盟主如何约束盟友的问题，探讨了威慑与约束困境（deterrence – versus – restraint dilemma），即盟主国家一方面要尽力发挥延伸威慑的作用保护盟友，另一方面不能因为保护过度而导致盟友肆无忌惮，趁机借助盟主国家的力量攫取最大利益，却伤害了盟主的利益。^① 扎加雷和基尔戈将同盟外部国家视为“修正主义国家”（Revisionist State），将同盟内部成员视为“维持现状国家”（Status Quo State）。但是，在“狐假虎威”现象中，同盟内部成员却扮演修正主义国家角色。同盟内弱国利用同盟外国家与盟主国家的实力差距和矛盾，主动发起盟主国家认可或支持的对同盟外国家的挑战，试图获取国家利益，其中同盟外国家是偏好现状的国家。对同盟外部国家而言，核心关切是双重威慑问题，不仅要慑阻同盟内部某一成员的挑衅，同时还要慑阻同盟内部其他国家的干

（接上页注②）2006，pp. 586 – 605；Stephen L. Quackenbush, “Not Only Whether but Whom: Three – Party Extended Deterrence,” *Journal of Conflict Resolution*, Vol. 50, No. 4, 2006, pp. 562 – 583；Frank C. Zagare and D. Marc Kilgour, “The Deterrence – Versus – Restraint Dilemma in Extended Deterrence: Explaining British Policy in 1914,” *International Studies Review*, Vol. 8, No. 4, 2006, pp. 623 – 641；Andrew O’Neil, “Extended nuclear deterrence in East Asia: redundant or resurgent?” *International Affairs*, Vol. 87, No. 6, 2011, pp. 1439 – 1457.

① Frank C. Zagare and D. Marc Kilgour, “The Deterrence – Versus – Restraint Dilemma in Extended Deterrence: Explaining British Policy in 1914,” pp. 623 – 641.

预。由于同盟并非铁板一块，不同国家的利益很少完全一致，因此有效威慑军事同盟不能简化为双边威慑或者延伸威慑问题，需要深入研究同盟因素给双边威慑带来的影响。已有的中西方威慑文献多是在延伸威慑的框架下分析同盟与威慑关系，并没有对“狐假虎威”现象给予清晰的解释。

如果把“狐假虎威”现象进行抽象和一般化，显然是一个涉及三个国家行为体的相互威慑现象，也可以称之为三方不对称军事威慑博弈问题。三类国家行为体分别为盟内弱国、盟内强国、盟外国家，它们扮演的角色是不同的，盟内弱国为挑战者，盟内强国为庇护者，盟外国家为威慑者。对盟外国家来说，问题是如何实施双重威慑；对盟内强国来说，问题则是如何解决威慑与约束困境；对盟内弱国来说，则是如何获得盟内强国保护、利用盟内强国获取利益，同时避免盟内强国对自身利益的损害。研究三方不对称威慑博弈现象，核心是描述、解释和预测三类国家行为体（一个国家对一个同盟）之间的冲突或战争现象。本文主要对盟外国家如何有效威慑盟内弱国挑战的问题进行研究，力求深刻揭示这一现象背后的因果逻辑。重点分析在盟外国家为现状满意国家、盟内弱国为挑战者、盟外国家与盟内强国实力不对称的情景下，盟内弱国主动发起挑战，最后引发冲突或战争的现象。本文研究的核心问题可以归纳为：为什么盟外国家威慑盟内弱国有时成功，有时失败？或者说，在什么样的条件下，哪些关键因素决定了三方不对称军事威慑现象的不同结局？更加具体的问题包括：第一，在何种条件下，挑战者没有实施挑战，而是维持现状，即威慑者一般威慑（general deterrence）的成功；第二，在何种条件下，挑战者发起挑战，即一般威慑失败；第三，在何种条件下，挑战者发起挑战，威慑者反抗，然后挑战者让步，即威慑者即时威慑（immediate deterrence）成功；第四，在何种条件下，挑战者发起挑战，威慑者反抗，挑战者实施升级行动，导致双边冲突或战争，即威慑者即时威慑失败；第五，在何种条件下，庇护者会干预挑战者和威慑者之间的军

事冲突，导致三边冲突或战争。^①

研究军事威慑问题所采用的科学方法很多，最常用的有两种，一种是结构主义分析法，另一种是博弈均衡分析法。结构主义分析法是一种整体研究方法，它一般不能在微观层次上提供对事件的描述和预测，但它能够在更宏观、更一般的层次上提供有说服力的解释。当然，结构主义方法也有很多不足，对结构的强调意味着很多现象被忽略，得出的一般性结论往往显而易见，或者不言自明；同时对结构理论的检验也较为困难。在威慑理论研究的早期，许多经典威慑理论家使用结构主义方法分析和解释威慑现象。^② 不过，更多的学者使用博弈论方法构建威慑理论模型，这些威慑理论被称为决策理论的威慑理论。早期学者使用“斗鸡模型”研究核威慑现象，^③ 使用“元博弈”解决核威慑的理性悖论；^④ 近期学者则更加娴

① 有关一般威慑和即时威慑的区别可参见：Patrick M. Morgan, *Deterrence: A Conceptual Analysis* (Beverly Hills, CA: Sage, 1983), p. 30; Vesna Danilovic, "Conceptual and Selection Bias Issues in Deterrence," *Journal of Conflict Resolution*, Vol. 45, No. 1, 2001, pp. 100 – 102; Vesna Danilovic, *When the Stakes Are High: Deterrence and Conflict among Major Powers* (Ann Arbor: The University of Michigan Press, 2002), pp. 54 – 55; Stephen L. Quackenbush, "Deterrence Theory: Where Do We Stand?" *Review of International Studies*, Vol. 37, No. 2, 2011, pp. 741 – 762.

② William W. Kaufmann, *The Requirements of Deterrence* (Princeton: Princeton University Press, 1954); Bernard Brodie, "The Anatomy of Deterrence," *World Politics*, Vol. 11, No. 2, 1959, pp. 173 – 179; Glenn H. Snyder, *Deterrence and Defense: Toward a Theory of National Security* (Princeton: Princeton University Press, 1961); Michael D. Intriligator and Dagobert L. Brito, "Can Arms Races Lead to the Outbreak of War?" *Journal of Conflict Resolution*, Vol. 28, No. 1, 1984, pp. 63 – 84; Michael D. Intriligator and Dagobert L. Brito, *The Stability of Mutual Deterrence*, in Jacek Kugler and Frank C. Zagare, eds., *Exploring the Stability of Deterrence* (Boulder: Lynne Rienner, 1987), pp. 13 – 39; John J. Mearsheimer, "Back to the Future: Instability in Europe after the Cold War," *International Security*, Vol. 15, No. 1, 1990, pp. 5 – 56. 表达结构威慑理论主要观点的一个最好模型是 Intriligator 和 Brito 发展的导弹战正式模型，揭示了体系结构、战争成本和武器系统特征之间错综复杂的关系，有关这一模型的中文解释也可参见阎学通、阎梁《国际关系分析》，北京大学出版社 2008 年版，第 171 页。

③ Daniel Ellsberg, "The Crude Analysis of Strategic Choice," *American Economic Review*, Vol. 51, No. 2, 1961, pp. 472 – 478; Herman Kahn, *On Thermonuclear War* (Princeton: Princeton University Press, 1960); Thomas Schelling, *Arms and Influence* (New Haven: Yale University Press, 1966).

④ Nigel Howard, *Paradoxes of Rationality: Theory of Metagames and Political Behavior* (Cambridge, MA: MIT Press, 1971).

熟地使用博弈论方法,鲍威尔(Powell)的核威慑理论、^①扎加雷(Zagare)的完美威慑理论、^②费伦(Fearon)的观众成本理论^③等威慑理论研究成果已经普遍采用博弈论中的子博弈完美均衡、完美贝叶斯均衡等概念来构建威慑理论模型。博弈论方法是一种研究理性行为体之间策略互动的数学和逻辑分析方法,它依靠形式化逻辑和演绎推理,生动地刻画威慑现象的微观基础和内在机理,能够提供超越直觉的洞察和预见。罗杰·迈尔森(Myerson)指出,博弈论是“智能的理性决策者之间冲突与合作的数学模型的研究”。^④基于威慑现象本质上是一种策略互动现象,使用博弈论方法构建理论模型具有明显的优势和价值。本文将构建一个三方不对称军事威慑博弈模型,分析并求解博弈均衡和行为体最优策略,力求揭示“狐假虎威”现象中的各种因果关系。

二、博弈模型与均衡分析

隐藏在“狐假虎威”现象背后的是三类国家行为体之间的策略互动,其中两类国家行为体签订了军事同盟条约,盟内弱国为挑战者,盟内强国为庇护者,作为第三方的盟外国家是威慑者。下文将展开博弈分析。

(一) 基本博弈模型

三方不对称军事威慑博弈基本模型可以表示为图 1 所示。三个行为体

① Robert Powell, *Nuclear Deterrence Theory: The Search for Credibility* (New York: Cambridge University Press, 1990).

② Frank C. Zagare and D. Marc Kilgour, *Perfect Deterrence* (Cambridge: Cambridge University Press, 2000).

③ James D. Fearon, "Signaling Versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model," *Journal of Conflict Resolution*, Vol. 38, No. 2, 1994, pp. 236-269; "Domestic Political Audiences and the Escalation of International Disputes," *American Political Science Review*, Vol. 88, No. 3, 1994, pp. 577-592.

④ 罗杰·B. 迈尔森 《博弈论: 矛盾冲突分析》(于寅、费剑平译),北京,中国经济出版社 2001 年版,第 1 页。

之间进行一个分阶段的动态博弈过程。第一阶段（在节点1），挑战者选择是否挑战，如果不挑战，结局为现状（SQ）；如果挑战者选择挑战，这时挑战者通常被视为修正主义国家，它对现状不满。挑战的方式主要包括三类：①对不曾拥有或不曾有效控制的领土提出权利主张；②试图推翻或改变另一个国家的政府；③宣布不再遵守另一个国家的政策。第二阶段（在节点2），针对挑战者的挑战行为，威慑者选择让步还是反抗，如果选择让步，挑战者的要求得到满足，结局是挑战者胜利（CS）；如果威慑者选择反抗，则表示威慑者实施即时威慑行动。即时威慑的目的有两个，一是让挑战者让步，二是让庇护者不干预。第三阶段（在节点3），庇护者选择干预还是不干预，如果选择干预，则会导致三边冲突或战争（TW），威慑者的即时威慑行动失败；如果庇护者选择不干预，则表示威慑者对庇

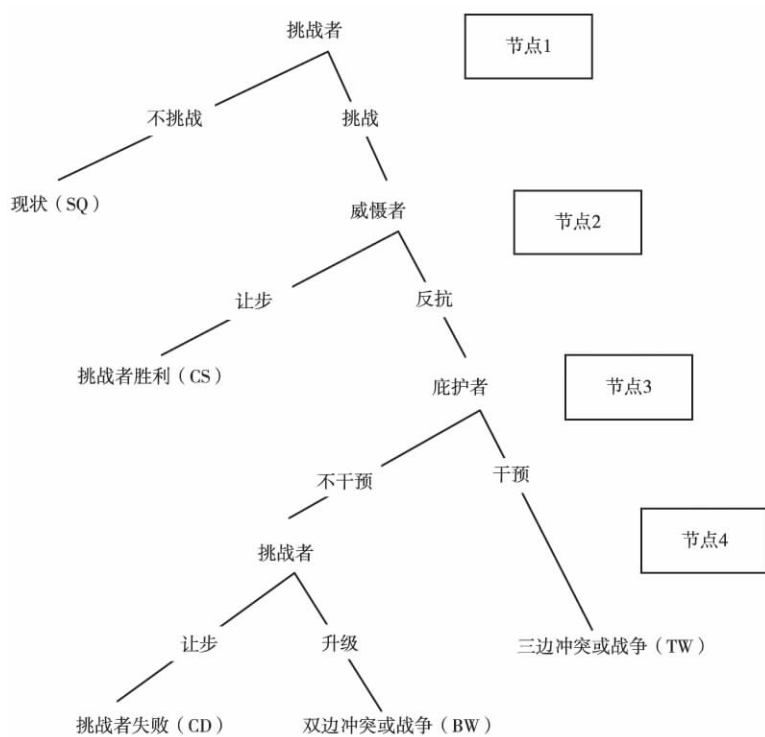


图1 三方不对称军事威慑博弈基本模型

护者的即时威慑成功。第四阶段（在节点 4），挑战者将选择究竟是让步还是升级行动，当挑战者选择升级行动时，结局为威慑者和挑战者之间的双边冲突或战争（BW），威慑者对挑战者的即时威慑失败；如果挑战者选择让步，结局为挑战者失败（CD），威慑者对挑战者的即时威慑成功。三方博弈的结局共分五种情况，按照威慑者的军事威慑成功或失败的类型来划分，现状（SQ）结局是威慑者一般威慑的成功，其他结局（CS，CD，BW，TW）是威慑者一般威慑的失败；挑战者失败（CD）是威慑者即时威慑的成功；双边冲突或战争（BW）、三边冲突或战争（TW）是威慑者即时威慑的失败。

（二）国家偏好结构

在三方不对称军事威慑博弈中，不同国家行为体对五种博弈结局具有相异的偏好，这些不同偏好关系构成的整体可称为国家偏好结构，当某一国家的偏好关系出现变化时，预示着国家偏好结构产生转变，对博弈结局产生决定性影响。军事威慑理论揭示，国家偏好关系能够反映国家的实力和决心。在威慑理论中，实力主要是指威慑性威胁的实力，这和通常所理解的国家实力的概念是不同的。谢林（Schelling）指出，有实力的威胁是能够产生伤害的威胁。^① 扎加雷（Zagare）提出，一个行为体的威胁是有实力的，仅当受威胁方更喜欢现状，而不是威胁实施后的结局；一个威胁是缺乏实力的，当被威胁方在威慑性威胁被执行后仍然愿意行动。^② 简而言之，威慑者有威慑实力，说明挑战者更愿意维持现状，而不是和威慑者发生冲突或战争。也就是说，威慑者相对挑战者而言是有实力的，则表明挑战者对现状（SQ）和双边冲突或战争（BW）两个结局之间的偏好关系为 $SQ >_{ch} BW$ 。^③ 对威

① Thomas C. Schelling, *Arms and Influence* (New Haven, CT: Yale University Press, 1966), p. 7.

② Frank C. Zagare, "Reconciling Rationality with Deterrence: A Re - Examination of the Logical Foundations of Deterrence Theory," *Journal of Theoretical Politics*, Vol. 16, No. 2, 2004, p. 124.

③ 本文使用 " $>_{ch}$ ， $>_{Def}$ ， $>_{Pro}$ " 分别表示挑战者、威慑者和庇护者对两个不同结果之间的严格偏好关系。

慑者而言，决心是其实施威胁的意志，是一个主观变量；对挑战者而言，决心反映了威慑者实施威胁的可信性（credibility），是一个客观变量。如果一个威胁在挑战者看来是不可信的，那么该威胁通常被认为是一种虚张声势，威慑将很难成功。斯莫克（Smoke）认为，可信威胁一般等同于应该被相信的威胁。^① 贝茨（Betts）指出，只有当威胁是理性实施的，这些威胁才能够被相信。^② 勒博（Lebow）也认为，只有理性威胁是可信的。^③ 扎加雷和基尔戈（Zagare and Kilgour）将以上观点进行了总结^④：威胁是可信的当且仅当其能够被相信，并且威胁是可理性实施的，可信威胁是满足塞尔腾（Selten）的完美标准^⑤的威胁，即威胁方更愿意执行它。如果威慑者是理性的，那么它有无决心不是由主观愿望决定，而是取决于它对不同结局的效用权衡，如果威慑者认为让步的效用高于冲突的效用，则威慑者的决心是不可信的，相反，如果威慑者认为冲突的效用高于让步的效用，则威慑者的决心是可信的。因此，威慑者是否有决心，不是建立在主观愿望而是客观利益权衡基础上的。威慑者是有决心的，则其在双边冲突或战争与让步之间的偏好关系为： $BW >_{Def} CD$ 。在威慑博弈中，庇护者的偏好关系中最重要的是它对于三边冲突或战争（TW）和双边冲突或战争（BW）两者的偏好。我们可以假定具有军事同盟的国家之间，如果盟友是可靠的，即当发生冲突或战争时，盟友站在自己一边参战，则该盟友的偏好关系为 $TW >_{pro} BW$ ；如果盟友是不可靠的，则说明该盟友的偏好关系为 $BW >_{pro} TW$ 。

笔者根据实力、决心和可靠性等指标将参与三方不对称军事威慑

① Richard Smoke, *National Security and the Nuclear Dilemma* (Reading, MA: Addison - Wesley, 1987), p. 93.

② Richard K. Betts, *Nuclear Blackmail and Nuclear Balance* (Washington, D. C.: Brookings, 1987), p. 12.

③ Richard Ned Lebow, *Between Peace and War: The Nature of International Crisis* (Baltimore, MD: The Johns Hopkins University Press, 1981), p. 15.

④ Frank C. Zagare and D. Marc Kilgour, *Perfect Deterrence*, pp. 66 - 69.

⑤ Reinhard Selten, "A Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, Vol. 4, 1975, pp. 25 - 55.

博弈的国家偏好结构归纳为表 1。在表 1 中，笔者没有列出挑战者更偏好于现状而不试图挑战威慑者以求获胜的情况，因为在这种情况下博弈的结局非常容易预测。如果挑战者的偏好为 $SQ >_{Ch} CS$ ，即现状优于挑战者胜利，则挑战不会发生，结局为现状（SQ）。当威慑者属于有决心类型时，还可以根据它对双边冲突或战争（BW）与三边冲突或战争（TW）之间的不同偏好关系细分为有完全决心（U）和有部分决心（C）两种子类型，当威慑者的偏好关系为 $BW >_{Def} TW >_{Def} CS$ 时，则表示有完全决心，即无论是双边还是三边的冲突或战争，都优于让步；当威慑者的偏好关系为 $BW >_{Def} CS >_{Def} TW$ 时，则表示有部分决心，即威慑者面对双边冲突或战争时有决心，但是面对三边冲突或战争时表现为缺乏决心。

表 1 三方不对称军事威慑博弈国家偏好结构

国家	类型		偏好关系
挑战者	有决心(H)	威慑者有威慑实力(C)	$CS >_{Ch} SQ >_{Ch} TW >_{Ch} BW >_{Ch} CD$
		威慑者无威慑实力(N)	$CS >_{Ch} TW >_{Ch} BW >_{Ch} SQ >_{Ch} CD$
	无决心(S)	威慑者有威慑实力(C)	$CS >_{Ch} SQ >_{Ch} [CD, TW] >_{Ch} BW$
威慑者	有决心(H)	有完全决心(UH)	$CD >_{Def} BW >_{Def} TW >_{Def} CS$
		有部分决心(CH)	$CD >_{Def} BW >_{Def} CS >_{Def} TW$
	无决心(S)		$CD >_{Def} CS >_{Def} BW >_{Def} TW$
庇护者	可靠(R)		$TW >_{Pro} [BW, CD]$
	不可靠(U)		$[BW, CD] >_{Pro} TW$

说明：在三方不对称军事威慑博弈中，威慑者策略选择导致的结局只有四种情况（CD，BW，TW，CS），因此在威慑者偏好关系的规定中不再涉及威慑者对于现状（SQ）的偏好。同样，对庇护者而言，其策略选择导致的结局只有三种情况（BW，CD，TW），对于庇护者偏好关系的规定不再考虑现状（SQ）和挑战者胜利（CS）两种结局。

在三方不对称军事威慑博弈中，如果威慑者不具有威慑实力，不能使挑战者认识到维持现状优于冲突或战争，而是使挑战者的偏好关系为 $BW >_{Ch} SQ$ 时，则挑战者必然会发起挑战，威慑者的军事威慑必然会失败，因此，威慑实力是威慑成功的必要条件。对于威慑者不具备威慑实

力的情景，逻辑推理非常简单，博弈结局一目了然。基于以上原因，在军事威慑理论研究中，一般假定威慑者具有威慑实力，重点研究威慑者的决心大小对于威慑成功或失败的影响。“狐假虎威”现象通常发生在威慑者具备威慑实力的情况下，此时挑战者更喜欢现状而非冲突或战争，同时要求威慑者的决心类型属于有部分决心（CH）类型或者无决心（S）类型，即在庇护者不干预的情况下有决心而当庇护者干预时威慑者表现为无决心，或者无论庇护者干预与否威慑者都无决心。论文在讨论三方不对称军事威慑各博弈方策略选择时，将讨论重点放在威慑者具备威慑实力的情形，对于无威慑实力的情形，由于预测相对简单，笔者不再做详细分析。

（三）完全信息条件下博弈均衡分析

在完全信息条件下，参与博弈的各方都知道自己和其他方的偏好关系。在不完全信息条件下，我们假定参与者仅知道自己的偏好关系，但是对于其他两方的实际偏好关系具有不完全信息，只是对其他各方所属类型的概率具有完全信息。博弈各种解的概念是建立在两个重要假定基础之上的，一是理性（rationality）假定，二是共同知识（common knowledge）假定。理性假定是指博弈者能够正确计算出各种不同行为组合会带来什么样的结局，并总是采用效用最大化的行为。共同知识假定是指如果每个博弈者都知道该事实，每个博弈者都知道每个博弈者都知道该事实，这样无穷推演下去。也就是说，（每个博弈者都知道）^k 每个博弈者都知道该事实， $k=0, 1, 2, \dots, \infty$ 。^① 三方不对称军事威慑博弈属于多阶段动态博弈，在完全信息条件下对应的均衡概念为“子博弈完美纳什均衡”。^② 求解子博弈完美纳什均衡的方法是逆向归纳法（backward induction），最后解出的均衡结果如表2所示。

① 涂志勇 《博弈论》，北京大学出版社2009版，第51页。

② 谢识予编著 《经济博弈论》（第三版），上海，复旦大学出版社2002年版，第119页。

表 2 完全信息条件下三方不对称军事威慑博弈均衡结果

参与者类型			最优策略组合			均衡结果
挑战者	威慑者	庇护者	挑战者	威慑者	庇护者	
有决心	有完全决心	可靠	不挑战/升级	反抗	干预	SQ
有决心	有完全决心	不可靠	不挑战/升级	反抗	不干预	SQ
有决心	有部分决心	可靠	挑战/升级	让步	干预	CS
有决心	有部分决心	不可靠	不挑战/升级	反抗	不干预	SQ
有决心	无决心	可靠	挑战/升级	让步	干预	CS
有决心	无决心	不可靠	挑战/升级	让步	不干预	CS
无决心	有完全决心	可靠	不挑战/让步	反抗	干预	SQ
无决心	有完全决心	不可靠	不挑战/让步	反抗	不干预	SQ
无决心	有部分决心	可靠	挑战/让步	让步	干预	CS
无决心	有部分决心	不可靠	不挑战/让步	反抗	不干预	SQ
无决心	无决心	可靠	挑战/让步	让步	干预	CS
无决心	无决心	不可靠	不挑战/让步	反抗	不干预	SQ

根据表 2，我们可以发现，若威慑者想要确保威慑成功，需要满足以下条件：

首先，威慑者必须具备威慑实力，这是确保威慑成功的必要条件。即挑战者在现状和冲突或战争之间的偏好关系为 $SQ >_{Ch} TW >_{Ch} BW$ 。

其次，在庇护者不可靠的情况下，威慑者的决心必须高于或者等于挑战者的决心。这里包括两种情况。一是威慑者有决心（UH 或者 CH 类型），即威慑者在冲突和让步之间的偏好关系为 $BW >_{Def} CS$ ；二是当威慑者无决心（S 类型）时，挑战者必须无决心（S 类型）。

最后，在庇护者可靠的情况下，威慑者必须是有完全决心类型（UH 类型），即威慑者在冲突和让步之间的偏好关系为 $BW >_{Def} TW >_{Def} CS$ 。

以上条件说明，在完全信息条件下，三方不对称军事威慑博弈结局主要由挑战者、威慑者两类国家行为体在实力、决心方面的差异以及庇护者是否可靠等因素所共同决定。只有当威慑者具有威慑实力和完全威慑决心

(属于 UH 类型) 时, 威慑才能够确保成功; 如果威慑者不具备完全威慑决心, 则要求庇护者属于不可靠类型 (U 类型), 同时挑战者不能属于有决心类型 (H 类型), 而属于无决心类型 (S 类型), 此时才会威慑成功, 否则威慑会失败。

(四) 不完全信息条件下博弈均衡分析

如果博弈各方对彼此决心具有不完全信息, 则需要进行“海萨尼转换”^①, 将不完全信息转化为不完美信息处理。假定在不完全信息条件下三方不对称军事威慑博弈中, 挑战者为 H 类型和 S 类型的先验概率分别为 p_{Ch}^H 和 p_{Ch}^S , 满足 $p_{Ch}^H + p_{Ch}^S = 1$, $0 < p_{Ch}^H < 1$, $0 < p_{Ch}^S < 1$; 威慑者为 H 类型和 S 类型的先验概率为和, 满足 $p_{Def}^H + p_{Def}^S = 1$, $0 < p_{Def}^H < 1$, $0 < p_{Def}^S < 1$; 庇护者为 R 类型和 U 类型的先验概率为 p_{Pro}^R 和 p_{Pro}^U , 满足 $p_{Pro}^R + p_{Pro}^U = 1$, $0 < p_{Pro}^R < 1$, $0 < p_{Pro}^U < 1$ 。以上先验概率为共同知识。使用一个九元组 $[x_H, x_S, y_H, y_S, q, z_H, z_S, w_H, w_S]$ 表示三方不对称军事威慑博弈的完美贝叶斯均衡。其中 q 为后验概率, 表示威慑者关于挑战者属于 H 类型的后验概率。 x_H, x_S 分别表示 H 类型和 S 类型的挑战者在博弈模型节点 1 选择挑战的概率, y_H, y_S 分别表示 H 类型和 S 类型的威慑者在博弈模型节点 2 选择反抗的概率, z_H, z_S 分别表示 R 类型和 U 类型的庇护者在博弈模型节点 3 选择干预的概率, w_H, w_S 分别表示 H 类型和 S 类型的挑战者在博弈模型节点 4 选择升级行动的概率。假定在威慑者有威慑实力的情况下, 使用 c_X, d_X, i_X 分别表示挑战者、威慑者和庇护者对于 X 结局的效用。对于三方不对称军事威慑博弈而言, 分别使用 H 表示博弈者有决心, S 表示博弈者无决心, 使用 R 表示庇护者是可靠的, 使用 U 表示庇护者是不可靠的。使用 $c_{TW}^H, c_{BW}^H, c_{TW}^S, c_{BW}^S, d_{TW}^H, d_{BW}^H, d_{TW}^S$ 和 d_{BW}^S 分别表示挑战者和威慑者属于 H 和 S 类型时关于结局 TW 和 BW 的效用; 使用 i_{TW}^R, i_{TW}^U 表示庇护者属于 R

^① John C. Harsanyi, “Games with Incomplete Information Played by Bayesian Players,” *Management Science*, Vol. 14, 1967/1968, pp. 159–182, 320–334, 486–502.

和 U 类型时关于结局 TW 的效用。三方不对称军事威慑博弈中不同国家对博弈结局的效用如下所示。

挑战者: $c_{CS} > c_{SQ} > c_{TW}^H > c_{BW}^H > [c_{TW}^S, c_{CD}] > c_{BW}^S$;
威慑者: $d_{CD} > d_{BW}^H > [d_{CS}, d_{TW}^H] > d_{BW}^S > d_{TW}^S$;
庇护者: $i_{TW}^R > [i_{BW}, i_{CD}] > i_{TW}^U$ 。

我们求解不完全信息条件下三方不对称军事威慑博弈均衡，可以得出表 3 所示的结果。

表 3 三方不对称军事威慑博弈完美贝叶斯均衡*

均衡	策略选择								存在条件	
	挑战者				威慑者		庇护者			
	x_H	x_S	w_H	w_S	y_H	y_S	z_R	z_U		
威慑均衡	0	0	1	0	1	1	1	0	$q \leq q_t^S$	
	0	0	1	0	1	0	1	0	$p_{Def}^H \geq c_t^1$ 且 $q_t^S \leq q \leq q_t^H$	
	0	0	1	0	v_1	0	1	0	$y_H \geq f_1$ 且 $q = q_t^H$	
	0	0	1	0	1	v_2	1	0	$y_S \geq f_2$ 且 $q = q_t^S$	
	0	0	1	0	v_1	v_2	1	0	$(y_H p_{Def}^H + y_S p_{Def}^S) \geq c_t^3$ 且 $q = q_t^H = q_t^S$	
分离均衡	1	0	1	0	1	0	1	0	$p_{Pro}^R \leq d_t$ 且 $c_t^2 \leq p_{Def}^H \leq c_t^1$	
	1	0	1	0	v_1	0	1	0	$p_{Pro}^R = d_t$ 且 $f_3 \leq y_H \leq f_1$	
攻击均衡	1	1	1	0	1	0	1	0	$p_{Def}^H \leq c_t^2$ 且 $q_t^S \leq p_{Ch}^H \leq q_t^H$	
	1	1	1	0	0	0	1	0	$p_{Ch}^H \geq q_t^H$	
	1	1	1	0	v_1	0	1	0	$p_{Ch}^H = q_t^H$ 且 $y_H \leq f_3$	
	1	1	1	0	1	v_2	1	0	$p_{Ch}^H = q_t^S$ 且 $y_S \leq f_4$	
	1	1	1	0	v_1	v_2	1	0	$p_{Ch}^H = q_t^H = q_t^S$ 且 $(y_H p_{Def}^H + y_S p_{Def}^S) \leq c_t^2$	
虚张声势均衡	1	u_2	1	0	1	0	1	0	$p_{Def}^H = c_t^2$ 且 $d_1 \leq x_S \leq d_2$	
	1	u_2	1	0	1	v_2	1	0	$x_S = d_2$ 且 $y_S = f_4$	
	1	u_2	1	0	v_1	0	1	0	$x_S = d_1$ 且 $y_H = f_3$	
	1	u_2	1	0	v_1	v_2	1	0	$x_S = d_1 = d_2$ 且 $(y_H p_{Def}^H + y_S p_{Def}^S) \leq c_t^2$	
	u_1	0	1	0	1	0	1	0	$p_{Def}^H = c_t^1$ 且 $p_{Pro}^R \leq d_t$	
	u_1	0	1	0	v_1	0	1	0	$p_{Pro}^R = d_t$ 且 $y^H = f_1$	

* 后验概率 q 参数的值：在威慑均衡中无限制，在分离均衡中 $q = 1$ ，在攻击均衡中 $q = p_{Ch}^H$ ，
在虚张声势均衡中 $q = \frac{p_{Ch}^H}{(p_{Ch}^H + p_{Ch}^S x_S)}$ ；表中 $0 \leq [u_1, \mu_2, \nu_1, \nu_2] \leq 1$ 。

在表 3 中, 有关博弈均衡存在条件的重要参数值主要有四个,^① 分别为:

$$\begin{aligned} q_t^H &= \frac{(d_{CD} - d_{CS}) + p_{Pro}^R (d_{TW}^H - d_{CD})}{(1 - p_{Pro}^R) (d_{CD} - d_{BW}^H)}, \\ q_t^S &= \frac{(d_{CD} - d_{CS}) + p_{Pro}^R (d_{TW}^S - d_{CD})}{(1 - p_{Pro}^R) (d_{CD} - d_{BW}^S)}, \\ c_t^1 &= \frac{(c_{CS} - c_{SQ})}{p_{Pro}^R (c_{BW}^H - c_{TW}^H) + (c_{CS} - c_{BW}^H)}, \\ c_t^2 &= \frac{(c_{CS} - c_{SQ})}{p_{Pro}^R (c_{CD} - c_{TW}^S) + (c_{CS} - c_{CD})}. \end{aligned}$$

根据表 3, 三方不对称军事威慑博弈共存在四类均衡形态: 威慑均衡、分离均衡、攻击均衡和虚张声势均衡, 根据 p_{Ch}^H , p_{Def}^H , p_{Pro}^R , q 的不同, 以上四类博弈均衡所在位置可以用图 2 表示。

1. 威慑均衡

威慑均衡包含两类, 一个是确定威慑均衡, 另一个是不确定威慑均衡。威慑者威慑决心的先验概率较高时 ($p_{Def}^H \geq c_t^1$), 挑战者将不实施挑战, 博弈结局为现状 (SQ)。从 c_t^1 函数可以看出, 在其他条件不变的情况下, 随着庇护者可靠性 (p_{Pro}^R) 的增加, c_t^1 的值增大, 威慑者威慑成功的门槛值上升, 要求威慑者展示更高的决心 (p_{Def}^H), 才能确保威慑的成功。确定威慑均衡不受挑战者决心大小的影响, 即 c_t^1 这一函数和挑战者决心的先验概率 (p_{Ch}^H) 无关。在其他条件不变的情况下, 当现状的效用 (c_{SQ}) 对于挑战者而言变大时, c_t^1 的值变小, 威慑者威慑成功的门槛值下降, 即在威慑者决心保持不变的情况下, 威慑者更可能威慑成功。在其他

^① 除了 c_t^1 , c_t^2 , q_t^H , q_t^S 四个重要的参数外, 表 3 中其他几个参数的表达式如下:

$$\begin{aligned} c_t^3 &= \frac{(c_{CS} - c_{SQ})}{p_{Pro}^R (c_{BW}^S - c_{TW}^S) + (c_{CS} - c_{BW}^S)}; \quad d_1 = \frac{p_{Ch}^H (1 - q_t^H)}{q_t^H (1 - p_{Ch}^H)}; \quad d_2 = \frac{p_{Ch}^H (1 - q_t^S)}{q_t^S (1 - p_{Ch}^H)}, \\ d_t &= \frac{d_{BW}^H - d_{CS}^H}{d_{BW}^H - d_{TW}^H}; \quad f_1 = \frac{c_t^1}{p_{Def}^H}; \quad f_2 = \frac{c_t^1 - p_{Def}^H}{1 - p_{Def}^H}; \quad f_3 = \frac{c_t^3}{p_{Def}^H}; \quad f_4 = \frac{c_t^2 - p_{Def}^H}{1 - p_{Def}^H}. \end{aligned}$$

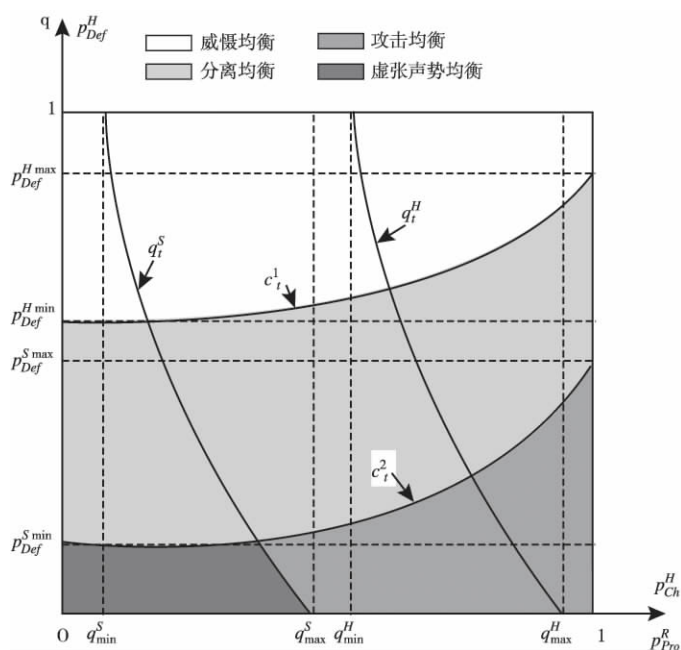


图 2 三方不对称军事威慑博弈均衡位置示意图

条件不变的情况下，当三边冲突或战争的效用 (c_{TW}^H) 对于挑战者而言更高时， c_t^1 的值增大，威慑者威慑成功的门槛值升高。在其他条件不变的情况下，当挑战者关于双边冲突或战争的效用 (c_{BW}^H) 变大时， c_t^1 的值增大，威慑者威慑成功的门槛值升高。

当威慑者决心的先验概率较低时 ($p_{Def}^H < c_t^1$)，仍然可能威慑成功，因为存在一个不确定威慑均衡。在不确定威慑均衡中，要求威慑者必须以足够高的概率选择反抗 ($\gamma_H \geq f_1, \gamma_S \geq f_2$) 不过这一威慑均衡是非常不确定的，因为当威慑者决心的先验概率较小时，还可能存在其他均衡。威慑者选择反抗还是让步，取决于挑战者属于 H 类型的后验概率 q 。当威慑者不考虑挑战者决心的大小，而承诺一旦挑战者发起挑战，将以较大的概率实施反抗，这时也会促使挑战者放弃挑战，获得现状的结局。

2. 分离均衡

分离均衡在两种情况下存在。当威慑者决心的先验概率低于确定威慑

均衡的门槛值，高于某一门槛值，同时庇护者的先验概率低于某一门槛值（即 $p_{Pro}^R \leq d_t$ 且 $c_t^2 \leq p_{Def}^H \leq c_t^1$ ）时，或者当庇护者的先验概率等于某一门槛值，同时威慑者决心的先验概率高于某一门槛值（ $p_{Pro}^R = d_t$ 且 $p_{Def}^H \geq c_t^3$ ）时，H 类型挑战者将实施挑战，S 类型的挑战者将不实施挑战，而是维持现状。在分离均衡中，挑战者的偏好关系可以通过它的策略选择揭示出来，H 类型和 S 类型的挑战者在策略选择上是不同的。从最低门槛值（ c_t^2 ）的特征可以发现，在其他条件不变的情况下，当现状的效用（ c_{SQ} ）增大时， c_t^2 的值变小，即威慑者威慑住 S 类型的挑战者的可能性增大；当挑战者让步的效用（ c_{CD} ）变小时， c_t^2 的值变小，威慑者威慑住 S 类型的挑战者的可能性增大。

3. 攻击均衡

当威慑者决心的先验概率低于分离均衡的门槛值（ $p_{Def}^H \leq c_t^2$ ），同时挑战者的先验概率高于某一门槛值时（ $p_{Ch}^H \geq q_t^S$ ），则无论是 H 类型的挑战者还是 S 类型的挑战者都会实施挑战。根据 q_t^S 函数形式可以发现，随着庇护者的可靠性（ p_{Pro}^R ）上升，挑战者实施挑战的门槛值 q_t^S 降低，所以挑战者实施挑战的可能性上升。当威慑者让步的效用（ d_{CS} ）增大时，挑战者实施挑战的门槛值 q_t^S 降低，挑战者实施挑战的可能性上升。在其他条件不变的情况下，当双边冲突或战争的效用（ d_{BW}^S ）对于 S 类型的威慑者而言增大时，挑战者实施挑战的门槛值 q_t^S 升高，挑战者实施挑战的可能性下降。当三边冲突或战争的效用（ d_{TW}^S ）对于 S 类型的威慑者而言增大时，挑战者实施挑战的门槛值 q_t^S 升高，挑战者实施挑战的可能性下降。

4. 虚张声势均衡

当威慑者决心的先验概率等于威慑均衡的门槛值（ $p_{Def}^H = c_t^1$ ），同时庇护者的先验概率低于某一门槛值（ $p_{Pro}^R \leq d_t$ ）；或者当威慑者决心的先验概率低于威慑均衡的门槛值（ $p_{Def}^H \leq c_t^1$ ），同时庇护者的先验概率等于某一门槛值（ $p_{Pro}^R = d_t$ ）时，此时 H 类型的挑战者将实施混合策略，S 类

型的挑战者选择不挑战。当威慑者决心的先验概率低于分离均衡的门槛值 ($p_{Def}^H \leq c_i^2$)，并且挑战者的先验概率低于某一门槛值时 ($p_{Ch}^H \leq q_i^S$)，则 H 类型的挑战者将实施挑战，而 S 类型的挑战者将实施混合策略，即以一定的概率实施挑战。从 d_1 d_2 两个公式可以看出，在其他条件不变的情况下，随着庇护者可靠性 (p_{Pro}^R) 的上升， d_1 d_2 值增大，S 类型的挑战者以更高的概率实施挑战。在 S 类型的挑战者实施混合策略时，H 类型的威慑者和 S 类型的威慑者也会实施混合策略，即以一定的概率 ($y_H = f_3$ ， $y_S = f_4$) 选择反抗。

三、理论假设与研究设计

(一) 理论假设

不完全信息条件下三方不对称军事威慑博弈均衡分析结果表明，“狐假虎威”现象中的三类国家行为体在进行动态博弈的过程中，会由于所拥有的私有信息和共同知识的差异而选择不同的优势策略，从而导致最终的博弈结局出现多种结果。按照不同信息条件以及博弈各方的偏好关系，可以将三方不对称军事威慑博弈的结局进行推测和归纳，结果如表 4 所示。根据表 4 的内容我们可以提出许多不同的理论假设，由于论文研究的核心问题是“狐假虎威”现象发生的原因与演变结局，即明确三方不对称军事威慑成功或失败的充分必要条件，所以本论文仅聚焦于以上问题提出几个重要的研究假设。在三方不对称军事威慑博弈中，威慑者必须具备威慑实力，这是威慑成功的必要条件，此时挑战者更喜欢维持现状而不是冲突或战争，即挑战者的偏好关系满足 $c_{BW} < c_{TW} < c_{SQ}$ 。如果不能满足以上条件，威慑将失败。因此，论文在提出核心理论假设前，首先假定威慑者具备足够的威慑实力，这样本文所提出的理论假设将只关注挑战者、威慑者和庇护者决心大小对威慑结局的影响。

表 4 不完全信息条件下三方不对称军事威慑博弈结局预测

信息条件	挑战者偏好		威慑者偏好	庇护者偏好	博弈结局	
无限制	$c_{SQ} > c_{CS}$		无限制	无限制	SQ	
无限制	$c_{TW} > c_{BW} > c_{SQ}$		H 类型	R 类型	TW	
				U 类型	BW	
			S 类型	无限制	CS	
$p_{Def}^H \geq c_i^1$		无限制	无限制	无限制	SQ	
$c_i^2 \leq p_{Def}^H \leq c_i^1$, $p_{Pro}^R \leq d_i$		H 类型	H 类型	R 类型	SQ ,TW	
			U 类型	U 类型	SQ ,BW	
			S 类型	无限制	SQ ,CS	
$p_{Def}^H < c_i^2$, $p_{Ch}^H > q_i^H$		S 类型	无限制	无限制	SQ	
		无限制	无限制	无限制	SQ ,CS	
$p_{Def}^H < c_i^2$, $q_i^S \leq p_{Ch}^H \leq q_i^H$		H 类型	H 类型	R 类型	SQ ,TW	
			U 类型	U 类型	SQ ,BW	
			S 类型	无限制	SQ ,CS	
		S 类型	H 类型	R 类型	SQ ,TW	
			U 类型	U 类型	SQ ,CD	
			S 类型	无限制	SQ ,CS	
$p_{Def}^H < c_i^2$, $p_{Ch}^H \leq q_i^S$		H 类型	H 类型	R 类型	SQ ,TW	
			U 类型	U 类型	SQ ,BW	
		S 类型	R 类型	R 类型	SQ ,TW ,CS	
			U 类型	U 类型	SQ ,BW ,CS	
		S 类型	H 类型	R 类型	SQ ,TW	
			U 类型	U 类型	SQ ,CD	
		S 类型	R 类型	R 类型	SQ ,TW ,CS	
			U 类型	U 类型	SQ ,CD ,CS	

研究假设 H1：如果挑战者对威慑者决心的判断高于阈值 c_i^1 ，那么威慑者的一般威慑能够成功。

研究假设 H2：如果挑战者对威慑者决心的判断低于阈值 c_i^1 ，高于阈值 c_i^2 ，同时庇护者的可靠性低于阈值 d_i ，那么威慑者针对有决心的挑战者的一般威慑会失败，针对无决心的挑战者的一般威慑能够成功。

研究假设 H3：如果挑战者对威慑者决心的判断低于阈值 c_i^2 ，同时威

慑者对挑战者决心的判断高于门槛值 q_t^S , 那么威慑者的一般威慑会失败。

研究假设 H4: 如果挑战者对威慑者决心的判断低于门槛值 c_t^2 , 同时威慑者对挑战者决心的判断低于门槛值 q_t^S , 那么威慑者针对有决心的挑战者的一般威慑会失败, 而无决心的挑战者会采取混合策略, 即以一定的概率实施挑战。

根据以上假设中的几个门槛值的函数表达式, 我们还可以发现, 如果挑战者对威慑者决心的判断高于 p_{Def}^{Hmax} , 则无论庇护者是否可靠, 威慑者的一般威慑都会成功; 当对威慑者决心的判断处于区间 $[p_{Def}^{Hmin}, p_{Def}^{Hmax}]$ 和 $[p_{Def}^{Smin}, p_{Def}^{Smax}]$ 之中时, 威慑者的一般威慑成功与否取决于庇护者的可靠性, 随着庇护者可靠性 (p_{Pro}^R) 的增加, 当威慑者具有更高的决心时, 一般威慑才会成功。^① “狐假虎威” 现象主要发生在挑战者认为威慑者决心处于以上两个区间之内时。

(二) 概念操作化与测量

本文中需要进行操作化和测量的概念主要有两类: 一类是因变量, 是需要解释的变量; 另一类是自变量, 是用来解释因变量发生变异的变量。

1. 因变量

根据研究假设, 需要解释的因变量是一般威慑结局, 该变量是个二分类变量, 具有一般威慑成功和一般威慑失败两个取值。区分一般威慑成功和失败的关键指标是现状是否被打破。所有的军事威慑行动都是围绕维持现状和改变现状而展开的。如果作为挑战者和威慑者的两个国家之间未发生军事争端事件, 则表明三方不对称威慑博弈的结局为现状。以军事争端事件的发生与否作为区分现状和非现状的指标, 并非意味着没有军事行动的争端就不重要。国家没有采取军事行动, 只能说明国家间的冲突议题不是非常重要, 而当国家采取军事行动时, 则意味着国家间的争端非常严重

① $p_{Def}^{Hmin} = (c_{CS} - c_{SQ}) / (c_{CS} - c_{BW}^H)$, $p_{Def}^{Hmax} = (c_{CS} - c_{SQ}) / (c_{CS} - c_{TW}^H)$, $p_{Def}^{Smin} = (c_{CS} - c_{SQ}) / (c_{CS} - c_{CD})$, $p_{Def}^{Smax} = (c_{CS} - c_{SQ}) / (c_{CS} - c_{TW}^S)$.

了。由于本文在实证检验部分使用 MID 数据集,^① 所以笔者使用该数据集中于对军事争端概念的定义和测量。所谓军事争端 (Militarized Interstate Dispute) 是指某一国家明确地威胁、展示或者使用军事力量针对另一个国家, 前一个国家被称为挑战者, 也可以称之为修正主义国家, 它不满足于现状, 提出挑战。挑战者针对的对象国是威慑者。当挑战者发起对威慑者的挑战预示着威慑者的一般威慑失败, 如果挑战者未发起挑战, 亦即军事争端事件未发生, 则表明威慑结局为现状, 即威慑者的一般威慑成功。

军事争端事件在等级、演变和结局上是不同的, 在等级上可以区分为五级: 无军事行动、武力威胁、武力展示、武力运用、战争。^② 武力威胁是敌对意图的口头指标, 依据外交语言来界定, 威胁行动通常以最后通牒 (ultimatum) 的方式来表达。武力展示涉及军事行动但是没有接触战斗, 是非暴力的。武力运用则代表了更严重的等级, 包括封锁、抵触、占领领土等。当战斗人员死亡人数超过 1000 人时, 则为战争。军事争端的后果主要有七种情况: 胜利 (victory) 、屈服 (yield) 、僵局 (stalemate) 、妥协 (compromise) 、获释 (release) 、不清楚 (unclear) 以及战争 (joins ongoing war) 。^③ 胜利是指某一国家使用军事行动获得满意的现状改变结果, 同时给对方造成失败, 包括获得有形的领土、使对方外交政策发生显著改变或者成功地使对方政府垮台。屈服是指服从于另一个国家的要求, 通过让步换取不再被对方威胁或者攻击的局面。僵局是指缺乏决定性的现状改变, 双方对后果都不满意。妥协是指一方同意放弃某些要求或者采取有关现状的让步。获释是指被捕获的人员或者物资得到释放或归还。不清楚是指历史资料提供了矛盾的或者模棱两可的信息。

① Faten Ghosn, "Glenn Palmer and Stuart A. Bremer, The MID3 Data Set, 1993 - 2001: Procedures, Coding Rules, and Description," *Conflict Management and Peace Science*, Vol. 21, No. 2, 2004, pp. 133 - 154.

② Daniel M. Jones, et al., "Militarized Interstate Disputes, 1816 - 1992: Rationale, Coding Rules, and Empirical Patterns," *Conflict Management and Peace Science*, Vol. 15, No. 2, 1996, pp. 170 - 174.

③ Daniel M. Jones, et al., "Militarized Interstate Disputes, 1816 - 1992: Rationale, Coding Rules, and Empirical Patterns," p. 180.

2. 自变量

解释三方不对称军事威慑现象的自变量是国家偏好结构和信息分布。其中国家偏好结构是指挑战者、威慑者和庇护者三类国家行为体对于威慑结局的不同偏好关系,国家偏好关系的度量一般使用效用函数表示。在研究假设中,挑战者的偏好使用效用函数 c_X 表示,威慑者的偏好使用效用函数 d_X 表示,庇护者的偏好使用效用函数 i_X 表示。信息分布则是指三方不对称军事威慑博弈的参与者关于其他博弈方所属类型的主观估计,使用先验概率进行度量。在研究假设中,挑战者属于有决心类型(H 类型)的先验概率用 p_{Ch}^H 表示,威慑者属于有决心类型(H 类型)的先验概率用 p_{Def}^H 表示,庇护者属于可靠类型(R 类型)的先验概率用 p_{Pro}^R 表示。对自变量国家偏好结构和信息分布的测量就是测量以上三类效用函数和三个先验概率。

(1) 国家偏好测量。对于国家偏好的测量问题,最初学者普遍认为难以进行有效测量。布恩诺·德·梅斯奎塔(Bueno de Mesquita)首次提出了一种测量国家互动博弈期望效用的方法。^①这一方法是使用国家同盟组合的相似性(similarity in national alliance portfolios)来测量国家期望效用,其原理是:两个国家的对外政策选择的方式越类似,两个国家之间的利益越一致,其中任何一个国家挑战另一方的利益越小,双方互动博弈中任何一方胜利,赢得的效用越小。如果使用 Δ_i 表示国家 i 索要的利益的大小,那么使用 $U^i(\Delta_i)$ 表示挑战胜利后得到利益产生的效用,使用 $U^i(\Delta_j)$ 表示挑战失败后失去利益产生的效用。国家间同盟组合越类似,则双方互动博弈时向对方提出 Δ_i 越小,产生的效用也越小,即 $U^i(\Delta_i)$ 和 $U^i(\Delta_j)$ 的差异越小。测量国家间同盟组合的相似性,可以使用同盟组合之间的肯德尔等级相关系数(Kendall's τ_b rank correlation coefficient)^②来度量,使用

① Bruce Bueno de Mesquita, *The War Trap* (New Haven: Yale University Press, 1981).

② 肯德尔相关系数是以 Maurice Kendall 命名的,并经常用希腊字母 τ (tau) 表示其值。肯德尔相关系数是一个用来测量两个随机变量相关性的统计值。肯德尔相关系数的取值范围在 -1 到 1 之间,当 τ 为 1 时,表示两个随机变量拥有一致的等级相关性;当 τ 为 -1 时,表示两个随机变量拥有完全相反的等级相关性;当 τ 为 0 时,表示两个随机变量是相互独立的。

符号 K_j^i 表示。如果两个国家盟友组合相同, 则 $K_j^i = 1.0$ 。布恩诺·德·梅斯奎塔还提出了测量国家冒险倾向 (risk-taking propensity) 的方法。^① 以上期望效用测量方法不断得到修正和改进, 不久之后布恩诺·德·梅斯奎塔和拉曼 (Bueno de Mesquita and Lalman) 提出了更加完善的期望效用测量方法。^② 在梅斯奎塔和拉曼的测量方法中, 使用三个效用作为基础测量互动博弈其他结局的期望效用。三个基础效用为: $U^i(\Delta_i)$ 表示得到利益的效用, $U^i(\Delta_j)$ 表示失去利益的效用, $U^i(SQ)$ 表示维持现状的效用。^③ 三个基础效用的测量, 依赖于两个更加基础性的概念, 即同盟组合的相似性 (K_j^i) 以及国家的冒险倾向 (使用 r^i 表示)。

在布恩诺·德·梅斯奎塔和拉曼提出期望效用测量方法之后, 许多学者对其展开研究并提出改善方法。希格诺里诺和里特 (Signorino and Ritter) 指出, 最初研究者们普遍使用同盟组合相似性的 τ_b 作为两个国家是否具有共同利益或冲突利益的大致测度, 但存在一种特殊情况, 即某些国家之间具有非常相似的同盟组合, 不过它们是否具有共同的安全利益却是不清楚的, 或者具有很强的共同安全利益的国家之间的同盟组合却有非常大的差异。^④ 应该说, 使用同盟组合相似性作为测量国家间利益共同性的指标有其合理性, 但是这一方法也有问题, 如有些国家没有任何盟友, 是否表示这些国家之间无共同利益呢? 因此同盟组合相似性的 τ_b 分数并不能很好地代表国家间利益共同性。为此希格诺里诺和里特提出了另一种测量方法, 使用 S 分数表示国家间外交政策相似性 (foreign policy similarity)。贝内特和鲁珀特 (Bennett and Rupert) 系统开展了有关 S 分

① Bruce Bueno de Mesquita, "The War Trap Revisited: A Revised Expected Utility Model," *The American Political Science Review*, Vol. 79, No. 1, 1985, pp. 156-177.

② Bruce Bueno de Mesquita and David Lalman, *War and Reason: Domestic and International Imperatives* (New Haven: Yale University Press, 1992).

③ Bruce Bueno de Mesquita and David Lalman, *War and Reason: Domestic and International Imperatives*, p. 47.

④ 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. 123.

数与 τ_b 分数的优劣的实证研究。^① 使用 S 分数相比 τ_b 分数在测量国家间利益差异方面更可靠一些。因此, 论文选择使用 S 分数测量三个国家对不同博弈结局的期望效用。

挑战者效用的测量公式为:

$$\begin{aligned} c_{SQ} &= U_c(SQ); c_{CS} = U_c[\Delta_c - (1 - \beta_c) \varphi_c P_c^b - \beta_c \varphi_c P^i]; \\ c_{CD} &= U_c\{\Delta_d - \gamma_c\{1 - [\beta_c P^i + (1 - \beta_c) P_c^b]\}\}; \\ c_{BW} &= P_c^b\{U_c[\Delta_c - \varphi_c P_c^b - \alpha_c(1 - P_c^b)]\} + (1 - P_c^b)\{U_c[\Delta_d - \varphi_c P_c^b - \alpha_c(1 - P_c^b)]\}; \\ c_{TW} &= P^i\{U_c[\Delta_c - \varphi_c P^i - \alpha_c(1 - P^i)]\} + (1 - P_c^i)\{U_c[\Delta_d - \varphi_c P^i - \alpha_c(1 - P^i)]\}. \end{aligned}$$

其中 Δ_c 表示挑战者对威慑者提出的利益要求的大小, Δ_d 表示威慑者对挑战者提出的利益要求的大小, P_c^b 表示挑战者和威慑者之间的双边冲突挑战者胜利的概率, P^i 表示挑战者、庇护者和威慑者之间的三边冲突挑战者胜利的概率, φ_c 表示挑战者使用武力的国内观众成本, α_c 表示冲突本身给挑战者带来的冲突成本, γ_c 表示挑战者挑战失败的国内观众成本, β_c 表示挑战者关于庇护者干预的概率估计。

其中, 战争胜利概率 (P_c^b, P^i) 的测量根据国家间的相对实力计算。测量公式:

$$\begin{aligned} P_c^b &= ACINC_{Ch} / (ACINC_{Ch} + ACINC_{Def}); \\ P^i &= (ACINC_{Ch} + ACINC_{Pro}) / (ACINC_{Ch} + ACINC_{Def} + ACINC_{Pro}). \end{aligned}$$

$ACINC_{Ch}, ACINC_{Def}, ACINC_{Pro}$ 分别是挑战者、威慑者和庇护者的调整后的国家综合实力分数。为了对国家实力进行测量, 辛格、布雷默和斯塔基 (Singer, Bremer and Stuckey) 提出了一种较为合理的测量指标和方法。^②

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 J. David Singer, Stuart Bremer and John Stuckey, "Capability Distribution, Uncertainty, and Major Power War, 1820-1965," in Bruce Russett, ed., *Peace, War, and Numbers* (Beverly Hills: Sage, 1972), pp. 19-48; J. David Singer, "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985," *International Interactions*, Vol. 14, No. 2, 1987, pp. 115-132.

布恩诺·德·梅斯奎塔认为国家相对实力需要根据国家之间的距离进行修正。^① 本论文所使用的调整后的国家综合实力分数是指对原始综合实力分数 CINC (COW 数据) 进行距离修正后的国家综合实力分数。

φ_i 表示挑战者使用武力的国内观众成本。布恩诺·德·梅斯奎塔和拉曼 (Lalman) 认为国内观众成本 φ_i 与国家 i 对现状的评价正相关, 而其测量为 $U^i(SQ)$, 如果国内民众更渴望现状, 那么使用武力改变具有吸引力的现状更可能激发反对, 而当现状不再具有吸引力时, 使用武力改变现状的反对声音更小。^② 夸肯布什 (Quackenbush) 研究发现, 以上测量存在内在矛盾, 一方面布恩诺·德·梅斯奎塔和拉曼假定 $\varphi_i > 0$, 但是根据 $\varphi_i = U^i(SQ)$, 这一值在 -0.83 和 1 之间变化, 所以既然是成本, 那么应该限制国内观众成本为非负值: $\varphi_i = \max[U^i(SQ), 0]$ 。^③ 布恩诺·德·梅斯奎塔和拉曼在测量国家 i 屈服的成本和冲突的成本时, 认为没有很好的方法来测量两者, 所以假定 $\alpha_i = \gamma_i = 1$, 这导致让步的成本等同于冲突的成本。以上假定是存在问题的, 使得冲突在任何情况下都好于让步, 从而任何威胁都是有实力的。目前并没有成熟的有关让步和冲突相对成本的理论, 不过在平均水平上, 让步的成本可能小于冲突的成本, 夸肯布什认为, 不妨假定 $\alpha_i = 4\gamma_i = 1$, 即让步的成本等于冲突成本的 $1/4$, 这些修正获得了实证研究的支持。^④ 本论文将根据夸肯布什的参数修正方法进行挑战者效用的计算, 并且对所有测量后的效用值进行标准化处理, 将其变换为在 $[0, 1]$ 区间内。

如果 $c_{CD} > c_{BW}$, 则以上 c_{BW} 和 c_{TW} 测量公式测量的是 S 类型的挑战者

① Bruce Bueno de Mesquita, *The War Trap*, p. 105.

② Bruce Bueno de Mesquita and David Lalman, *War and Reason: Domestic and International Imperatives*, p. 297.

③ Stephen L. Quackenbush, "General Deterrence and International Conflict: Testing Perfect Deterrence Theory," *International Interactions*, Vol. 36, No. 1, 2010, p. 85.

④ Stephen L. Quackenbush, "General Deterrence and International Conflict: Testing Perfect Deterrence Theory," p. 84; 相关附录的网址为 <http://web.missouri.edu/~quackenbush/>。

的效用 c_{BW}^S 和 c_{TW}^S , 此时 H 类型挑战者的效用未知; 如果 $c_{BW} > c_{CD}$, 则以上 c_{BW} 和 c_{TW} 测量公式测量的是 H 类型的挑战者的效用 c_{BW}^H 和 c_{TW}^H , 此时 S 类型挑战者的效用未知。为了测量未知的效用, 可以假定 H 类型和 S 类型的挑战者在冲突或战争选项上的效用与让步选项上的效用的距离相同, 三边冲突或战争选项上的效用与双边冲突或战争选项上的效用的距离相同, 同时不能违反基本偏好关系。

当 S 类型的挑战者的效用已知时, 则 H 类型的挑战者的效用可以用以下公式测量:

$$\begin{aligned} c_{BW}^H &= \min [c_{CD} + (c_{CD} - c_{BW}^S) \rho_{SQ} - 0.02]; \\ c_{TW}^H &= \min [c_{BW}^H + (c_{TW}^S - c_{BW}^S) \rho_{SQ} - 0.01]。 \end{aligned}$$

当 H 类型的挑战者的效用已知时, 则 S 类型的挑战者的效用可以用以下公式测量:

$$\begin{aligned} c_{BW}^S &= \max [c_{CD} - (c_{BW}^H - c_{CD}) \rho_{SQ} - 0.01]; \\ c_{TW}^S &= \min [c_{BW}^S + (c_{TW}^H - c_{BW}^H) \rho_{SQ} - 0.01]。 \end{aligned}$$

对于威慑者效用的测量采用与挑战者效用测量相似的方法, 在此不单独列出。

庇护者效用的测量公式可表达为如下:

$$\begin{aligned} i_{BW} &= P_c^b U_i(\Delta_c) + (1 - P_c^b) U_i(\Delta_d); \\ i_{TW} &= P^t \{ U_i[\Delta_c - \varphi_i P^t - \alpha_i(1 - P^t)] \} + (1 - P^t) \{ U_i[\Delta_d - \varphi_i P^t - \alpha_i(1 - P^t)] \}。 \end{aligned}$$

其中 Δ_c , Δ_d , P_c^b , P^t 四个值的含义和测量公式与前述一致, φ_i 表示庇护者使用武力的国内观众成本, α_i 表示三边冲突给庇护者带来的冲突成本。

(2) 国家类型测量。国家类型测量是指对参与博弈的挑战者、威慑者属于有决心类型 (H 类型) 以及庇护者属于可靠型 (R 类型) 的主观先验概率的测量, 即对 p_{Def}^H , p_{Ch}^H 和 p_{Pro}^R 的测量。对挑战者和威慑者的决心的主观先验概率的测量主要依据该国家在冲突和让步两者之间的偏好强度来决定国家类型, 偏好强度是 $\Delta U = U(\text{冲突}) - U(\text{让步})$ 。最初, 金

和布恩诺·德·梅斯奎塔 (Kim and Bueno de Mesquita) 发展了一个测量国家类型的公式, 其中国家 i 是 H 类型的概率是这个国家对冲突和让步效用差异的线性函数。^① 但是, 夸肯布什认为进行非线性变化更能反映对博弈参与者类型的主观估计, 他提出了新的测量参与者类型的测量公式。^② 夸肯布什的测量方法更加合理一些, 但是夸肯布什仅考虑了双边威慑博弈的情景。本论文将参照夸肯布什的测量方法, 对三方威慑博弈情景中挑战者和威慑者的决心类型进行测量, 并将其扩展到对庇护者可靠性的测量。

挑战者、威慑者和庇护者分别属于 H 类型和 R 类型的主观先验概率测量公式为:

$$\begin{aligned} p_{Ch}^H &= \left[\frac{e^{3(c_{BW}-c_{CD})}}{1 + e^{3(c_{BW}-c_{CD})}} \right]^{(1+R_{Def})}; \\ p_{Def}^H &= \left[\frac{e^{3(d_{BW}-d_{CS})}}{1 + e^{3(d_{BW}-d_{CS})}} \right]^{(1+R_{Ch})}; \\ p_{Pro}^R &= \left[\frac{e^{3(i_{TW}-i_{BW})}}{1 + e^{3(i_{TW}-i_{BW})}} \right]^{(1+R_{Def})}。 \end{aligned}$$

四、案例分析

案例研究、大样本统计分析和实验方法是社会科学研究理论假设检验的三种主要方法。长期以来, 许多学者认为案例研究是理论假设检验方法中最弱的, 案例研究难以控制其他变量的干扰性影响。但是斯蒂芬·范埃弗拉 (Stephen Van Evera) 指出, 对案例研究的这种批评是不公正的, 案例研究在控制被忽略变量的影响方面能够提供颇为有力的方法。^③ 案例研

① Woosang Kim and Bruce Bueno de Mesquita, "How Perceptions Influence the Risk of War," *International Studies Quarterly*, Vol. 39, No. 1, 1995, pp. 51-65.

② Stephen L. Quackenbush, "General Deterrence and International Conflict: Testing Perfect Deterrence Theory," p. 85.

③ 斯蒂芬·范埃弗拉 《政治学研究方法指南》(陈琪译), 北京大学出版社 2006 年版, 第 48—50 页。

究能否成功,关键在于选择的案例是否能够进行变量控制。变量控制的目的是控制可能干扰自变量与因变量关系的因素,创造一个研究条件相同或相似的环境,以便于观察和检验研究假设中的变量关系。^①为了更好地控制变量,本论文选择 1978—1990 年间中越苏三方威慑博弈关系进行分析,试图检验论文的某些理论假设。

(一) 1978—1990 年中越苏三国关系简况

1978 年至 1990 年共计 13 年的时间内,中越苏三边关系呈现以中国与越苏同盟对抗为主轴的关系。1978 年 11 月 3 日,越南和苏联在莫斯科签订了《越苏友好合作条约》,第六条规定“缔约双方将对涉及两国利益的一切重要国际问题进行协商。一旦双方中之一成为进攻的目标或受到威胁的目标时,缔约双方立即进行协商,以便消除这种威胁和采取相应的有效措施来保障两国和平与安全。”^②以上条款说明《越苏友好合作条约》本质上是军事同盟性质的条约。中苏关系自 1967 年 1 月进入敌对状态,并于 1979 年 4 月达到对抗水平,从此长期维持在这一水平,1982 年开始缓和,1989 年 5 月才实现双边关系的正常化。^③中越关系在 1975 年前是友好关系,但是从 1976 年开始由友好转向不和,并且逐渐恶化,1979 年 2 月 17 日爆发了双边战争,从此开启了长达十余年的对抗之路,直到 1991 年双边关系才实现了正常化。

(二) 1978—1990 年中越的军事争端事件

1978—1990 年,中越之间冲突不断,多次发生军事争端事件,有些军事争端事件演变为战争。13 年间中越的军事争端事件如表 5 所示。

① 阎学通、孙学峰 《国际关系研究实用方法》(第二版),北京,人民出版社 2007 年版,第 106 页。

② 郭明主编 《中越关系演变四十年》,南宁,广西人民出版社 1992 年版,第 112—113 页。

③ 阎学通等著 《中外关系鉴览 1950—2005: 中国与大国关系定量衡量》,北京,高等教育出版社 2010 年版,第 389 页。

表 5 1976—1990 年间中越军事争端事件

时间	军事争端 事件名称	冲突议题	中国最高 军事对抗 水平	越南最高 军事对抗 水平	冲突后果	伤亡 人数(人)
1978 年 4 月 18 日— 1979 年 3 月 15 日	第一次中越战争 (对越自卫反击战)	政策	战争	战争	僵局	超过 1000
1981 年 5 月 5 日— 1981 年 8 月 17 日	法卡山和扣林山 自卫反击战	政策	武力运用	武力运用	僵局	251—500
1983 年 4 月 10 日— 1983 年 4 月 23 日	无	政策	武力运用	武力运用	僵局	26—100
1984 年 1 月 14 日— 1984 年 7 月 12 日	老山和者阴山 自卫反击战	政策	武力运用	武力运用	僵局	缺失
1985 年 1 月 9 日— 1986 年 2 月 6 日	无	政策	武力运用	武力运用	僵局	缺失
1986 年 10 月 1 日— 1987 年 3 月 15 日	第二次中越战争	政策	战争	战争	僵局	超过 1000
1987 年 4 月 15 日— 1987 年 6 月 6 日	南沙群岛争端 (柏礁岛)	领土	武力展示	无军事 行动	僵局	无伤亡
1987 年 10 月 5 日	越南米格战斗机 入侵事件	政策	武力运用	武力展示	僵局	1—25
1988 年 2 月 20 日— 1988 年 3 月 16 日	南沙群岛争端 (赤瓜礁)	领土	武力运用	武力运用	中国胜利	缺失

说明：该数据库对大多数军事争端事件并未进行命名，笔者主要根据中国官方的说法对一些冲突进行了命名。

资料来源：数据来源于战争相关指数（COW）项目国际军事争端数据（MID）。参见 Faten Ghosn, Glenn Palmer and Stuart A. Bremer, “The MID3 Data Set, 1993 – 2001: Procedures, Coding Rules, and Description,” *Conflict Management and Peace Science*, Vol. 21, No. 2, 2004, pp. 133 – 154。

从表 5 可以看出，中越两国在 13 年期间除 1980 年、1982 年和 1989 年这三年外，其余年份均发生军事冲突，其中有些冲突延续时间长达一年之久。1978 年至 1990 年间，中越两国共发生 9 次军事冲突，其中 7 次冲突起因于严重的外交政策分歧，主要原因是越南企图建立“印度支那联邦”。在苏联的支持下，越南派遣大规模军队侵占了柬埔寨，而中国坚定

支持民主柬埔寨政权,反对越南的侵略,并强烈要求越南撤出柬埔寨;其余两次冲突则是因为南沙群岛的岛屿主权归属问题。在 9 次中越军事冲突中,除南沙群岛(赤瓜礁)军事冲突外,其余冲突中越南均为挑战者。将越南视为挑战者(修正主义国家)的根本原因在于,7 次冲突的主要矛盾是外交政策分歧,越南入侵柬埔寨并持续扩大侵略成果,而中国试图阻止越南侵略柬埔寨行动的升级;同时,以上冲突的表现特征具有明显的相似性,每一次冲突都是越南率先在中越边境地区实施武装挑衅和入侵活动,然后中国采取外交部照会、外交声明或发言等多种形式对越南的挑衅行为进行强烈谴责或抗议。最后中国在外交抗议无效的情况下,直接使用军事力量惩罚和打击了越南。^① 1987 年 4 月发生的中越岛屿争端事件是由于越南公然派遣军队非法占领中国南沙群岛的柏礁岛所引发的,在这次冲突中,越南也是挑战者。1988 年中越“赤瓜礁海战”冲突的起因是中国于 1988 年 2 月开始实施永暑礁改建工程以及在周围岩礁上进行“高脚屋”工程建设。在对峙一个多月后,1988 年 3 月 14 日,双方在南沙群岛九章群礁海域赤瓜礁海域发生冲突。根据上述 9 次中越冲突历史事实的分析,可以发现,其中 1980 年、1982 年、1988 年、1989 年和 1990 年五年时间内中国对越南的一般威慑是成功的,其余年份中国对越南的一般威慑是失败的。

(三) 1978—1990 年越中苏三方不对称军事威慑博弈均衡分析

对于越中苏三个国家的综合实力分数(CINC)、对外政策相似性分数(S)以及风险倾向分数(R)等基础数据,笔者使用 MID 数据集和 EUGene 软件生成,^② 结果如表 6 所示。

① 在 7 次因外交政策分歧而引发军事冲突中,有关中国外交抗议行为可以参见每次冲突发生后一段时间内的《人民日报》。其中 1978 年 4 月至 1979 年 3 月的中越冲突期间,中国针对越南在两国边境地区实施的挑衅行为而进行的外交抗议高达 20 多次;其余几次冲突分别参见《人民日报》(1981 年 5 月 6 日,1983 年 4 月 11 日,1985 年 1 月 11 日,1987 年 10 月 6 日等)。

② D. Scott Bennett and Allan Stam, "EUGene: A Conceptual Manual," *International Interactions*, Vol. 26, No. 2, 2000, pp. 179-204, 网址为 <http://eugenesoftware.org>。

表 6 越中苏三国效用测量基础数据

年份	越南 CINC	中国 CINC	苏联 CINC	中越 S	苏越 S	中苏 S	越南 R	中国 R	苏联 R
1977	0.009	0.117	0.175	0.630	0.375	0.100	0.599	0.688	0.677
1978	0.009	0.120	0.172	0.463	0.545	0.089	0.638	0.645	0.672
1979	0.009	0.119	0.169	0.468	0.550	0.100	0.825	0.672	0.662
1980	0.011	0.118	0.170	0.463	0.547	0.095	0.564	0.712	0.699
1981	0.011	0.119	0.171	0.459	0.543	0.092	0.621	0.714	0.721
1982	0.011	0.117	0.173	0.456	0.530	0.078	0.844	0.705	0.729
1983	0.012	0.117	0.173	0.453	0.527	0.074	0.788	0.709	0.703
1984	0.011	0.111	0.165	0.461	0.528	0.084	0.813	0.697	0.736
1985	0.010	0.110	0.170	0.470	0.520	0.079	0.739	0.722	0.737
1986	0.011	0.109	0.169	0.465	0.517	0.078	0.587	0.725	0.735
1987	0.013	0.106	0.168	0.466	0.513	0.084	0.666	0.718	0.720
1988	0.011	0.107	0.165	0.471	0.520	0.089	0.424	0.751	0.735
1989	0.011	0.109	0.131	0.489	0.572	0.168	0.319	0.718	0.728
1990	0.010	0.106	0.124	0.514	0.166	-0.213	0.046	0.712	0.721

根据表 6 我们可以看出，从 1978 年开始，中越两国外交政策相似性发生重大变化，S 分数由 1977 年的 0.630 变为 0.463 左右，这表明两国的外交政策分歧迅速扩大，其主要原因是苏越结盟以及越南入侵柬埔寨。从 1989 年开始，中越两国 S 分数开始上升，说明两国紧张关系开始缓和。将表 6 的基础数据分别代入三方不对称军事威慑博弈效用测量公式，可以求得越中苏三方对博弈不同结果的效用值，如表 7 所示。根据效用计算结果可以发现 $d_{BW} > d_{CS}$ ， $c_{CD} > c_{BW}$ ， $i_{BW} > i_{TW}$ ，因此 1978—1990 年间，中国属于有决心类型（H 类型），而越南实际为无决心类型（S 类型），苏联为不可靠类型（U 类型）。从中我们还可以发现 $d_{BW} > d_{SQ}$ ，即对中国而言，与越南发生双边冲突的效用优于维持现状，这说明越南对中国没有威慑实力。

表 7 越中苏三方不对称军事威慑博弈结果效用

年份	c_{SQ}	c_{CS}	c_{CD}	c_{BW}	c_{TW}	d_{SQ}	d_{CS}	d_{CD}	d_{BW}	d_{TW}	i_{BW}	i_{TW}
1978	0.490	0.583	0.365	0.231	0.433	0.489	0.394	0.582	0.556	0.354	0.408	0.293
1979	0.461	0.548	0.343	0.209	0.402	0.484	0.392	0.576	0.550	0.351	0.410	0.295
1980	0.502	0.598	0.374	0.246	0.446	0.478	0.386	0.570	0.538	0.344	0.404	0.291
1981	0.493	0.588	0.366	0.24	0.437	0.478	0.385	0.570	0.537	0.343	0.401	0.289
1982	0.458	0.546	0.340	0.212	0.403	0.479	0.385	0.572	0.539	0.342	0.398	0.286
1983	0.466	0.557	0.346	0.219	0.413	0.479	0.384	0.572	0.538	0.341	0.400	0.289
1984	0.462	0.551	0.344	0.216	0.409	0.481	0.387	0.573	0.539	0.343	0.397	0.286
1985	0.474	0.563	0.354	0.224	0.424	0.477	0.385	0.567	0.536	0.335	0.396	0.287
1986	0.498	0.593	0.372	0.247	0.449	0.476	0.384	0.567	0.532	0.335	0.395	0.287
1987	0.485	0.578	0.363	0.242	0.438	0.477	0.384	0.568	0.528	0.334	0.396	0.289
1988	0.525	0.623	0.394	0.270	0.475	0.472	0.382	0.561	0.527	0.332	0.396	0.288
1989	0.543	0.640	0.411	0.286	0.471	0.477	0.389	0.564	0.530	0.357	0.405	0.290
1990	0.591	0.688	0.455	0.328	0.514	0.478	0.394	0.560	0.531	0.361	0.348	0.235

说明: c_{BW} d_{BW} 两个效用的初始计算结果表明, $d_{BW} > d_{CS}$ $c_{CD} > c_{BW}$, 因此在表 7 中测量的是 c_{BW}^S d_{BW}^H 两个值。

若我们将计算出来的效用值分别代入三方不对称军事威慑博弈不同均衡存在条件的门槛值计算公式, 可以求出各门槛值。同时对于表 7 所示的不同年份的博弈结果效用值和表 6 所示的国家风险倾向分数, 我们可以将之代入假定为共同知识的国家类型计算公式, 计算出博弈各方对挑战者、威慑者属于 H 类型以及庇护者属于 R 类型的主观先验概率。根据理论假设比较主观先验概率和门槛值, 可以推理出理论预测值, 即每一年三方博弈均衡类型与威慑结局, 结果如表 8 所示。其中, 1978 年越中苏三方博弈预测均衡结果为攻击均衡, 1980 年、1988 年、1989 年和 1990 年这四年越中苏三方博弈预测均衡结果为分离均衡, 其余年份预测均衡结果为虚张声势均衡。相应地, 根据均衡类型可以推理出参与博弈的各方的最优策略以及威慑结果。从表 6 中我们已知越南是无决心 (S) 类型, 中国为有决心 (H) 类型, 苏联是不可靠 (U) 类型, 所以当预测均衡为攻击均衡时, 越南将会挑战中国, 中国对越南的一般威慑是失败的; 当预测均衡为分离

均衡时，越南将会维持现状，中国对越南的一般威慑将会成功；当预测均衡为虚张声势均衡时，越南将会实施混合策略，即以一定的概率挑战中国。

表 8 越中苏三方不对称军事威慑博弈门槛值与均衡类型

年份	p_{Ch}	p_{Def}	p_{Pro}	c_t^1	c_t^2	q_t^H	q_t^S	d_t	均衡类型	预测结果
1978	0.223	0.456	0.235	0.841	0.461	6.639	0.219	0.802	攻击均衡	一般威慑失败
1979	0.217	0.414	0.229	0.831	0.455	6.591	0.228	0.797	虚张声势均衡	一定概率挑战
1980	0.213	0.465	0.223	0.844	0.461	5.486	0.251	0.783	分离均衡	一般威慑成功
1981	0.213	0.451	0.223	0.842	0.462	5.265	0.254	0.783	虚张声势均衡	一定概率挑战
1982	0.214	0.406	0.225	0.833	0.460	5.216	0.250	0.781	虚张声势均衡	一定概率挑战
1983	0.214	0.417	0.224	0.837	0.463	5.133	0.253	0.781	虚张声势均衡	一定概率挑战
1984	0.216	0.411	0.228	0.833	0.461	5.157	0.246	0.777	虚张声势均衡	一定概率挑战
1985	0.210	0.424	0.224	0.834	0.461	5.427	0.241	0.75	虚张声势均衡	一定概率挑战
1986	0.212	0.456	0.224	0.842	0.466	4.858	0.252	0.751	虚张声势均衡	一定概率挑战
1987	0.217	0.434	0.226	0.839	0.467	4.196	0.260	0.739	虚张声势均衡	一定概率挑战
1988	0.208	0.491	0.219	0.846	0.463	4.827	0.259	0.744	分离均衡	一般威慑成功
1989	0.214	0.515	0.221	0.845	0.449	4.947	0.271	0.817	分离均衡	一般威慑成功
1990	0.214	0.587	0.223	0.845	0.440	5.257	0.260	0.808	分离均衡	一般威慑成功

（四）理论预测与历史事实的比较

在表 5 中，本文已经详细归纳了 1978—1990 年中越两国间的实际军事冲突情况，表 8 中总结了相同时期内中国威慑越南成败的理论预测值。通过比较表 7 和表 8，我们可以清楚地发现，理论预测基本符合历史事实。在理论预测为一般威慑成功的年份，中越两国间未发生军事冲突，即相互维持现状；在理论预测为越南会以一定概率实施挑战的年份，除 1982 年外，中越两国间都发生了军事冲突，而且越南都是挑战者。根据虚张声势均衡的定义，越南会实施混合策略，即越南维持现状和实施挑战两个选项的效用等同，这说明 1982 年的历史数据也符合理论预测。因此，以上有关越中苏三方不对称威慑博弈的案例分析报告支持了论文所提出的主要理论假设。

五、研究结论

本文依据简化的博弈模型深入分析了“狐假虎威”现象，清楚地揭示了参与博弈的三类国家行为体之间复杂的互动关系，成功辨识出同盟外部国家有效威慑同盟内部弱国挑战的充分必要条件。研究发现，同盟外部国家要实现成功威慑同盟内部弱国挑战的目标，需要满足两项基本条件。第一，同盟外部国家必须具备威慑实力，即所拥有的威慑性威胁在实施后能够给同盟内部弱国的利益造成重大损失，确保同盟内部国家认识到维持现状要优于和同盟外部国家发生冲突。第二，同盟外部国家的威慑决心必须高于某一门槛值，当同盟内部弱国为有决心类型（H 类型）时，门槛值用 c_i^1 表示；当同盟内部弱国为无决心类型（S 类型）时，门槛值用 c_i^2 表示。由于三方不对称威慑博弈还存在一个不确定威慑均衡，当同盟外部国家的威慑决心较低，也有可能威慑成功，不过这种成功是不稳定的，因为不确定威慑均衡存在的条件与攻击均衡和虚张声势均衡的存在条件存在交集。从 c_i^1 和 c_i^2 两个门槛值的函数特征可以发现，随着盟内强国庇护或支持力度的增大，盟外国家成功威慑盟内弱国挑战的决心门槛值上升，也就是说盟内弱国的挑衅性上升，这清楚地反映了“狐假虎威”效应。观察两个有关威慑决心的门槛值，还可以发现几个有价值的推论：在其他条件不变的情况下，当现状对盟内弱国而言具有更大的价值时，即挑战者的现状满意度越高，其挑衅性会下降；当冲突或战争成本对盟内弱国来说变大时，盟外国家更容易威慑成功；当挑战获胜得到的利益对盟内弱国而言变大时，盟外国家的威慑更容易失败；当让步对盟内弱国来说具有更高的成本时，盟外国家的威慑更容易成功。

作者简介

曹金绪 第二炮兵指挥学院博士候选人。

电子信箱: jinxucao@163.com

吴宇 南开大学历史学院博士候选人。

电子信箱: fairstein@gmail.com

曹玮 清华大学当代国际关系研究院博士后。2012年在复旦大学国际关系与公共事务学院获法学博士学位。主要研究方向为公共外交。

电子信箱: cwangel@163.com

赵可金 清华大学当代国际关系研究院副教授。1997年毕业于山东曲阜师范大学政治系,2002年和2005年在复旦大学国际关系与公共事务学院分别获得法学硕士和法学博士学位。专著有《治理美国:国会的角色与模式》《当代中国外交制度的转型与定位》《外交学原理》等。

电子信箱: kejinzhao@tsinghua.edu.cn

张伟玉 清华大学国际关系学系2011级博士生。

电子信箱: zhangwy08@139.com

陈哲 天津外国语大学讲师。

电子信箱: chenzhecom@gmail.com

表娜俐(韩) 清华大学国际关系学系2011级博士生。

电子信箱: djmousai@hotmail.com

沃德·沃姆丹姆(Ward Warmerdam) 荷兰伊拉斯姆斯大学社会研究所博士研究员。

电子信箱: warmerdam@iss.nl