

# 偏好、信息与威慑博弈\*

向钢华 王永县

威慑理论是一种在冷战时期占主导地位的国际安全理论,曾被誉为社会科学中最有影响的学术创造之一。托马斯·谢林(Thomas Schelling)倡导的博弈论方法、罗伯特·杰维斯(Robert Jervis)倡导的心理分析方法和亚历山大·乔治(Alexander George)等倡导的个案分析方法,被认为是威慑理论研究的三种基本方法。在冷战后期特别是冷战结束以来,常规威慑日益受到重视,统计检验方法也愈来愈多地应用于威慑理论研究中。由于威慑具有行为者数量有限、行为者之间策略互动明显以及所涉及的利益重大等特点,因而有学者认为威慑等国际安全问题比较适合用博弈论研究方法。

---

\* 作者非常感谢匿名评阅专家对本文提出的宝贵意义。

Christopher Achen and Duncan Snidal, "Rational Deterrence Theory and Comparative Case Studies," *World Politics*, Vol 41, No 2, January 1989, p. 143.

克雷格·A. 斯奈德等编:《当代安全与战略》(徐纬地等译),长春,吉林人民出版社 2001年版,第 157、165页。

Andrew Kydd, "The Art of Shaker Modeling: Game Theory and Security Studies," in Sprinz and Wolinsky-Nahmias, eds., *Models, Numbers and Cases: Methods for Studying International Relations* (Ann Arbor, MI: University of Michigan Press, 2004), p. 348. 谢林作为在 20 世纪 50 年代最早采用博弈论方法研究威慑理论的学者,因运用博弈论方法在冲突与合作领域作出的杰出贡献而荣获了 2005 年诺贝尔经济学奖,这似乎为这种观点提供了佐证。

---

《国际政治科学》2006 年第 3 期(总第 7 期),第 1—33 页。

*Science of International Politics*

威慑理论和博弈论是在 20 世纪 40 年代中期几乎同时创建的两类学术理论,国际安全问题特别是核威慑问题是最早采用博弈论方法进行研究的领域之一。美国战略学界曾对博弈论抱有巨大希望,甚至有五角大楼分析家认为,“我们希望它起作用,正如我们在 1942 年希望原子弹起作用一样”。战略学家们很快意识到,美、苏之间的核威慑关系并不是早期博弈论所描述的那种完全对抗的零和博弈,而是有着极力避免核战争这一共同利益的非零和博弈,由此催生了后来非常有影响的“囚徒困境”和“胆小鬼游戏”等经典非零和博弈模型。目前建立在“纳什均衡”、“子博弈完美纳什均衡”等均衡概念基础上的主流博弈论方法,主要是在经济学领域发展起来的。1986 年巴里·纳莱巴夫(Barry Nalebuff)最早采用主流博弈论方法研究威慑问题,此后以罗伯特·鲍威尔(Robert Powell)、弗兰克·扎加雷(Frank Zagare)、詹姆斯·费伦(James Fearon)为代表的新一代威慑理论研究者普遍采用主流博弈论方法来建立威慑模型。在此之前,一些独具特色的非主流博弈分析方法在威慑与冲突研究中得到广泛应用,其中比较有影响的有元博弈(Metagames)方法、基于误博弈(Hypergames)的冲突分析方法以及行动理论(Theory of Moves)等。由于元博弈与重复博弈、误博弈与不完全信息博弈、行动理论与动态博弈分别类似,因此上述分析方法被认为是因错误的原因而反对主流博弈论。

---

冯·诺伊曼于 1944 年创建了博弈论,伯纳德·布罗迪在 1945 年最早提出了核威慑思想,参见 Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton: Princeton University Press, 1944); Bernard Brodie, ed., *The Absolute Weapon: Atomic Power and World Order* (New York: Harcourt, Brace and Co., 1946)。

J. McDonald, “A theory of strategy”, *Fortune*, No. 39, June 1949, p. 100.

参见 Barry Nalebuff, “Brinkmanship and Nuclear Deterrence: The Neutrality of Escalation”, *Conflict Management and Peace Science*, Vol. 9, No. 2, 1986, pp. 19—30; Robert Powell, “Nuclear Brinkmanship with Two-Sided Incomplete Information”, *The American Political Science Review*, Vol. 82, No. 1, 1988, pp. 155—178; Frank Zagare and Marc Kilgour, “Asymmetric Deterrence”, *International Studies Quarterly*, Vol. 37, No. 1, 1993, pp. 1—27; James Fearon, “Signaling versus the Balance of Power and Interests”, *Journal of Conflict Resolution*, Vol. 38, No. 2, 1994, pp. 236—269.

参见 Niall Fraser and Keith Hipel, *Conflict Analysis: Models and Resolutions* (New York: North-Holland, 1984), pp. 203—220, 235—259; Steven Brams, *Superpower Games: Applying Game Theory to Superpower Conflict* (New Haven: Yale University Press, 1985), pp. 59—62; Steven Brams, *Theory of Moves* (New York: Cambridge University Press, 1994), pp. 23—42.

Barry O'Neill, “Game theory models of peace and war”, in R. Aumann and S. Hart, eds., *Handbook of Game Theory* (Amsterdam: Elsevier, 1994), Vol. 2, pp. 997—1055.

本文尝试运用博弈论方法阐述威慑格局与可信威慑的形成条件,探讨威慑实力、决心与信息传递这三个要素起作用的机理,并结合四个实例进行具体分析。本文采用的假定与罗杰·迈尔森(Roger Myerson)对博弈论的两条基本假定一致:(1)博弈方是理性的,即国家是追求效用或期望效用最大化的理性行为体;(2)博弈方是具有理解力的,即国家之间的互动存在“共同知识”。本文分为四个部分。第一部分对理性威慑逻辑的三种分析框架进行了比较,提出了威慑格局概念与可信威慑的形成条件;第二部分探讨了不完全信息条件下威慑信号的作用机制;第三部分具体分析了海湾战争前美国与伊拉克之间的策略互动、中国在朝鲜战争中对美国跨过“三八线”的威慑、古巴导弹危机中美国对苏联的威逼以及马岛战争中英国对阿根廷的威逼四个实例;第四部分是对论文观点的总结。

### 一、威慑逻辑的一般分析

威慑作为一种使用武力威胁方式慑阻对手以防止出现不希望的结局的国际政治行为,在谋略、战略和理论三个不同层次上得到应用与研究。作为谋略,威慑与战争、贸易、联盟一样历史悠久,是一种非常重要的社会历史现象;作为战略,核威慑构成了冷战时期美国国家安全战略和军事战略的主题与基石;作为理论,它是任何威慑战略赖以存在的内在逻辑和原则。如果说,在核武器出现之前威慑是作为实际使用武力之前的一种谋略而存在,在冷战时期威慑战略是威慑研究的核心,那么,冷战结束后威慑理论尤其是威慑逻辑研究受到了研究者的普遍关注。

基于将国家作为理性行为体的假定,威慑逻辑的分析框架主要有三种,即成本—收益分析框架、理性选择分析框架和博弈论分析框架。

---

罗杰·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51页。

Claudio Cioffi-Revilla, "Origins and Age of Deterrence: Comparative Research on Old World and New World Systems," *Cross-Cultural Research*, Vol 33, No 3, August 1999, pp. 256—258

刘金质:《美国国家战略》,沈阳,辽宁人民出版社1997年版,第157页。

Patrick M. Mogan, *Deterrence Now* (New York: Cambridge University Press, 2003), p. 8

### (一) 威慑逻辑的成本—收益分析框架

成本—收益分析是一种源于经济学的“盈亏平衡”分析方法,它在国际政治研究中也得到广泛应用。其基本逻辑是:当收益高于成本时,经济或政治行为就会发生,否则该行为将不会出现。把这样的逻辑应用于威慑行为,其含义是:当战争的收益高于战争成本,即战争是有利可图时,潜在的进攻者将会发动战争,威慑者的威慑失败;否则,它将不会发动战争,威慑者的威慑成功。为此,威慑者将通过传递清晰明确的信息使潜在的进攻者认识到,如果它发起战争将会得不偿失、无利可图,以此慑阻对方可能发起的进攻。

在早期及后来的经典威慑理论中,对于威慑逻辑的探求充满了这样的成本—收益分析思维,认为“威慑概念是基于这样一种假设,即现代政府的政策制定者们在选择战争之前,通常会进行成本—收益分析”。乔治和理查德·斯摩克(Richard Smoke)在20世纪70年代中期曾给威慑下过一个被广泛引述的经典定义:“就其最一般的形式上,威慑就是让对手相信:他为某种行动而付出的代价或所冒风险大于收益。”在他们之前的早期研究者中,布罗迪(Bernard Brodie)就明确指出:“威慑的实现,是通过让潜在的进攻者确信,蓄意发动大规模核战争的收益,绝不会大于进行这种战争的成本。”亨利·基辛格(Henry Kissinger)认为:“威慑就是设法向对方提出危险性,这种危险将使对方认为与其要获得的利益不成比例,因而使对方不敢采取某种行动。”威慑理论研究中统计分析方法的倡导者保罗·胡思(Paul Huth)在20世纪末的一篇威慑研究综述文章中给威慑下了类似的定义:威慑是指一方通过发出威胁使另一方意识到可能招致的代价和损失,从而克制自己,避免采取某种行动。

---

詹姆斯·多尔蒂、小罗伯特·普法尔茨格拉夫:《争论中的国际关系理论》(阎学通、陈寒溪译),北京,世界知识出版社2003年版,第372页。

Alexander L. George and Richard Smoke, *Deterrence in American Foreign Policy: Theory and Practice* (New York: Columbia University Press, 1974), p. 13.

Bernard Brodie, *Strategy in the Missile Age* (Princeton, NJ: Princeton University Press, 1959), p. 271.

Henry A. Kissinger, *The Necessity for Choice* (Garden City, New York: Doubleday, 1962), p. 12.

Paul K. Huth, “Deterrence and International Conflict: Empirical Findings and Theoretical Debates,” *Annual Review of Political Science*, Vol 2, No. 1, 1999, p. 26.

上述学者的论述都清晰地反映出成本—收益分析思维。这一思维背后的逻辑是：威慑得以成立的条件是，“使对方相信：反击所造成的损失大于对方进攻所期待的得益；也就是使对方产生这样一个心理效应： $L$ （损失）-  $G$ （得益） $>0$ ”。

## （二）威慑逻辑的理性选择分析框架

威慑逻辑的成本—收益分析框架认为，潜在进攻者之所以屈从于威慑者的武力威胁，是因为其发起进攻的得益（或期望得益）小于损失（或期望损失），因此得不偿失。如果从决策视角考察，这种分析框架潜在地假定不采取行动的得失为 0，因此是否采取行动只需权衡行动本身的利弊得失（大于 0 即采取行动，否则不采取行动）。实际上，在威慑情景中普遍存在理性选择问题，即潜在进攻者需要对采取行动（如发动战争）与不采取行动（不发动战争）两种策略的效用进行比较，而不是简单地将不采取行动的得失默认为零。理性选择分析框架与成本—收益分析框架的差异在于：当发动战争得不偿失时，在成本—收益分析框架看来，挑战者将不会发动战争；而理性选择分析框架认为，如果不发动战争情况更糟，则发动战争虽然得不偿失但却是一种占优策略，挑战者将发动战争。

格伦·斯奈德（Glenn H. Snyder）在 20 世纪 60 年代初对潜在进攻者采取“进攻”和“不进攻”两种行动时的决策方案进行了比较分析，巴里·布赞（Barry Buzan）在 20 世纪 90 年代初也对潜在进攻者采取“挑战”和“不挑战”两种行动时的决策方案进行了讨论。在他们的分析中，都引入了对手策略的不确定性，即考虑了威慑者的可能行为对潜在进攻者的影响。理性选择框架的逻辑思路可以简要归纳如下：设潜在进攻者不采取行动的效用为  $U_s$ ，采取行动的效用为  $U_c$ 。由于它采取行动时，威慑者可能采取报复或不报复的回应方式，因此  $U_c$

---

参见张曙光：《威慑理论：美国国际战略学的一个重要领域》，《美国研究》1990年第 4 期，第 17 页。

参见 Glenn H. Snyder, *Deterrence and Defense* (Princeton: Princeton University Press, 1961), pp. 16—27; Barry Buzan, *An Introduction to Strategic Study: Military Technology and International Relations* (London: Macmillan Academic and Professional Ltd., 1991), p. 166。吴苑思对斯奈德和布赞的理性选择分析进行了详细论述，参见吴苑思：《威慑理论与导弹防御》，北京，长征出版社 2001 年版，第 18—26 页。

是一种期望效用。设当威慑者不报复时(设概率为  $p$ ),潜在进攻者效用为  $C$ ;当威慑者实施报复时(其概率为  $1 - p$ ),潜在进攻者效用为  $D$ 。则有:  $U_c = p \times C + (1 - p) \times D$ 。根据理性选择框架,当  $U_c > U_s$ 时,潜在进攻者采取行动;否则,潜在进攻者不采取行动。设  $p = p^*$  时有  $U_c = U_s$ ,可求得  $p^* = (U_s - D) / (C - D)$ 。 $p^*$ 被称为威慑“所需可信度”,即威慑者威慑成功的最低门槛概率。

由上述分析可知,在理性选择分析框架看来,潜在进攻者是否采取行动,既需要对采取行动的代价与收益的大小进行“代价评估”(即评估  $C$ 与  $D$ 的大小),也需要对威慑者兑现其威胁即实施报复的可能性进行“风险评估”(即评估  $p$ 值大小)。也就是说,“在被威慑方的评估系统中,至少有两组四个变数在起作用——代价与收益的大小和兑现与不兑现的可能性。如果被威慑方根据两种评估的总和,决定不采取行动,威慑就奏效;反之,如果被威慑方决定采取行动,威慑则失败”。

### (三) 威慑逻辑的博弈论分析框架

理性选择分析框架不仅考虑了被威慑者不同策略之间的比较与选择,而且考虑了威慑者在策略选择上的不确定性,但它并没有考虑威慑者与被威慑者之间的策略互动,而策略互动正是威慑行为的最根本特征之一。在这一点上,作为交互式决策理论的博弈论分析框架与前两种分析框架区分开来。与成本—收益分析和理性选择分析框架相比,博弈论分析框架除了“理性人”假定外还增加了博弈方之间存在“共同知识”的假定。

#### 1. 威慑格局的形成

从策略互动角度看,威慑是一种威慑者与被威慑者之间的动态博弈,博弈双方的行动有先后顺序之分。一方面,后行动的一方可以对先行动一方的行为进行观察,然后再选择自己的相应策略或行为,即存在相机选择问题;另一方面,先行动一方由于能够理性地预期到后行动的一方可能采取的策略,因此能够采取有针对性的行动。在从策略互动视角对威慑的早期研究中,谢林采用具

---

姚云竹:《战后美国威慑理论与政策》,北京,国防大学出版社 1998年版,第 17页。

有特定支付值的矩阵博弈模型进行了分析,丹尼尔·埃尔斯伯格(Daniel Ellsberg)在20世纪60年代初期构造了更一般的双矩阵博弈模型,并计算了美、苏相互威慑中美国威慑的“关键风险”和威慑的可靠性。由于当时动态博弈及子博弈完美纳什均衡的概念在博弈论中尚未建立,因此埃尔斯伯格在静态博弈模型中对博弈双方的“等待”、“打击”两种策略所进行的分析比较中,没有考虑博弈双方的策略互动,与上述理性选择分析框架并无实质差异。

在威慑逻辑的博弈分析中,被威慑者(C)又称进攻者或挑战者,威慑者(D)又称防卫者,本文在以下的分析中将威慑博弈双方分别称为挑战者与威慑者。威慑情景被看成是挑战者与威慑者之间的策略互动:挑战者试图打破现状以获取更大的利益,威慑者则通过发出武力威胁阻吓挑战者以保持现状。在如图-1所示的标准两阶段威慑模型中,挑战者有“挑战”和“维持”两种策略,威慑者有“抗击”和“容忍”两种策略。当挑战者采取“维持”策略时,博弈结局为维持现状(简称为SQ),此时挑战者、威慑者的效用分别为 $UC_s$ 和 $UD_s$ 。当挑战者采取“挑战”策略时,如果威慑者采取“抗击”策略,则博弈结局为冲突(简称为WAR),此时挑战者、威慑者的效用分别为 $UC_w$ 和 $UD_w$ ;如果防卫者采取“容忍”策略,则博弈结局为挑战者赢(简称为CW N),此时挑战者、威慑者的效用分别为 $UC_c$ 和 $UD_c$ 。按照威慑概念的定义,挑战者的偏好为 $CW N > SQ > WAR$ ,这意味着当挑战者发起挑战时,其结局或者优于SQ或者劣于SQ,因此

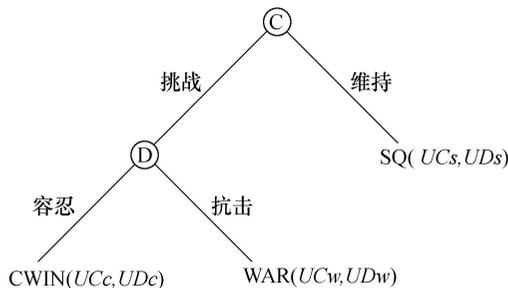


图-1 两阶段动态博弈模型

Thomas Schelling, "The Strategy of Conflict: Prospectus for a Reorientation of Game Theory," *Journal of Conflict Resolution*, Vol 2, No 3, September 1958, pp. 203—262

Daniel Ellsberg, "The Crude Analysis of Strategic Choices," *American Economic Review*, Vol 51, No. 2, 1961, pp. 472—478.

风险与收益并存;威慑者的偏好为  $CW_N < SQ$  且  $WAR < SQ$ , 即维持现状是威慑者最好的结局, 这也是威慑者慑阻挑战者发起挑战的根本动因, 否则威慑者将挑战现状, 转化为挑战者。在博弈双方的多种互动格局中, 威慑格局的形成是有条件的, 只有在上述挑战者与威慑者的偏好序同时成立的条件下, 才能形成与威慑定义相一致的威慑格局: 挑战者有动因发起挑战, 但要冒与威慑者产生冲突的风险; 威慑者力图慑阻挑战者的挑战, 因为维持现状是威慑者最好的结局。

### 2 可信威慑的充分必要条件

由上述对威慑格局的分析可知,  $CW_N$  是挑战者最偏爱的结局, 它也是挑战者之所以发起挑战的基本动因, 但挑战者的挑战也面临着风险: 如果威慑者“抗击”, 将导致结局  $WAR$ 。因此, 挑战者的策略选择是在“挑战”(将导致结局  $CW_N$  或结局  $WAR$ , 挑战者获得期望效用) 和“维持”(将导致结局  $SQ$ , 挑战者获得确定性效用) 两种策略之间进行权衡。根据成本—收益分析框架, 当挑战者“挑战”的收益大于损失时, 即  $UC_w > 0$ , 挑战者将发起挑战, 此时威慑者威慑失败; 否则挑战者将维持现状, 此时威慑者威慑成功。根据理性选择分析框架, 只有当挑战者发起挑战的效用大于维持现状的效用时, 挑战者才会“挑战”。设威慑者“容忍”的概率为  $p$ , 则其“抗击”的概率为  $1 - p$ , 挑战者发起挑战的期望效用  $UC$  为:  $UC = p \times UC_c + (1 - p) \times UC_w$ 。这样, 当  $UC > UC_s$  时, 挑战者将发起挑战, 此时威慑者威慑失败; 否则, 挑战者将维持现状, 此时威慑者威慑成功。令  $UC = UC_s$ , 可得  $p^* = (UC_s - UC_w) / (UC_c - UC_w)$ , 它为威慑者威慑成功的最低“可信度”。虽然理性选择分析框架考虑了威慑者的策略选择对挑战者得益计算的影响, 并以概率  $p$  来表征威慑者策略的不确定性, 但理性选择分析框架没有回答一个关键问题: 威慑者为什么采取或不采取“抗击”策略? 而这一点正是博弈论作为交互式决策理论得以发挥作用的地方。

根据博弈论分析框架, 挑战者与威慑者之间存在着策略互动, 后行为方即威慑者将根据先行为方即挑战者的策略选择做出最优反应, 同时威慑者的策略选择反过来将对挑战者的策略选择产生影响。也就是说, 先行为的挑战者由于能够理性地预期到后行为的威慑者可能采取的策略, 进而采取针对性的策略, 因此博弈论分析框架下的“理性选择”与理性选择分析框架下的“理性选择”将

有很大的不同。在完全信息条件下,即博弈方的效用已知的条件下,当  $UD_c > UD_w$ ,此时威慑者“容忍”的效用大于“抗击”的效用,威慑者将“容忍”,  $p = 1$ ;挑战者由于理性地预期到了这一点,威慑者策略运用上的不确定性消失,此时挑战者将发起挑战。当  $UD_c < UD_w$ ,此时威慑者“容忍”的效用小于“抗击”的效用,威慑者将“抗击”,  $p = 0$ ;挑战者由于理性地预期到了这一点,威慑者策略运用上的不确定性消失,此时挑战者将维持现状。由此可知,在完全信息条件下,挑战者的行为不仅取决于挑战者发起挑战的期望效用与维持现状时的确定效用之间的权衡比较,还取决于对方即威慑者将如何采取行动:如果威慑者“容忍”优于“抗击”,此时威慑者的威慑将是不可信的,挑战者将发起挑战;否则,威慑者的威慑将是可信的,挑战者将维持现状。因此,在效用确定的条件下,威慑逻辑的博弈论分析框架中并不存在理性选择分析框架中所谓威慑成功的最低“可信度” $p^*$ ,因为在前者中  $p$ 值是明确的,或者为 1,或者为 0。斯奈德和布赞在各自分析的实际举例计算中,既给出了具体的得益值(这意味着是完全信息条件),同时又计算出威慑者“抗击”的具体概率  $p$ ,这在博弈论分析框架看来是相互矛盾的。

由上述分析可知,在完全信息条件下,可信威慑的充分必要条件是:(1)挑战者的偏好为  $CW_N > SQ > WAR$ ; (2)威慑者的偏好为  $SQ > WAR > CW_N$ 。条件(1)意味着威慑者是有实力的,它能使挑战者发起挑战的可能结局  $WAR$ 劣于  $SQ$ ,此时挑战者的挑战要冒风险。条件(2)意味着威慑者是有决心的,“抗击”是威慑者的占优策略,此时如果挑战者挑战必然导致冲突结局。只有当上述两个条件同时成立,即威慑者既有实力又有决心时,威慑者的威慑才是可信的。成本—收益分析框架与理性选择分析框架都强调了条件(1)而没有考虑条件(2),但条件(1)只是可信威慑的必要条件而不是充分必要条件。博弈论分析框架从威慑格局成立的第二个条件( $CW_N < SQ$ 且  $WAR < SQ$ )中,将威慑者没有决心的情形(此时  $SQ > CW_N > WAR$ ,“容忍”是挑战者的最优选择),即不可置信的威慑,从策略集中剔除出来,使“抗击”成为唯一的可信选择(此时  $SQ > WAR > CW_N$ )。

### 3. 对可信威慑的进一步讨论

在威慑文献中,基辛格最早将威慑能力、威慑决心和威慑信息传递列为可

信威慑必须具备的三个要素：“威慑需要实力、使用实力的意志，以及潜在进攻者对这两方面因素的评估等三方面的结合。而且，威慑是所有这些因素之乘积，而不是它们的和。如果其中任何一个因素是零，威慑就会失效。”

在上述要素中，威慑实力是可信威慑的必要条件但不是充分条件，只有同时具有实力与决心，才是可信威慑的充分必要条件。“决心”一词，长期以来一直被作为一种内隐、主观或心理概念使用，在本文中基于理性人假定把决心看成是一种外显的、客观或物理的概念。这样，威慑者作为一种理性行为体，其决心的有无取决于它对不同策略下利益的权衡。如果“容忍”比“抗击”效用更大，威慑者无论内心是否情愿都将选择“容忍”策略，此时威慑者是没有决心的（如果威慑者宣称它是有决心的，则是一种不可信的虚张声势）；反之，当“抗击”比“容忍”效用更大时，威慑者将选择“抗击”，此时威慑者是有决心的。因此，威慑者是否有决心，不是建立在主观愿望而是建立在客观的利益权衡基础上的，这也使挑战者有可能通过对威慑者行为方式的观察和对其利益的计算，判断其决心的有无。

根据可信威慑的充分必要条件，威慑者在完全信息条件下的威慑或者是可信的，或者是不可信的，但在这两种情形下信息传递所起的作用是不同的。当  $UD_c < UD_w$  时，威慑者的威慑是可信的。由于挑战者与威慑者都是理性行为体，挑战者能够理性地预期到威慑者的策略选择，而威慑者知道挑战者能够预期这一点，即存在共同知识，此时无须威慑信息传递起作用。当  $UD_c > UD_w$  时，威慑者的威慑是不可信的，此时为了将不可信威慑转化为可信威慑，威慑者需要采取承诺行动（commitments）。承诺行动意味着，威慑者为了使威慑可信，需要付出额外的沉淀成本  $SC$ （sunk costs），以使  $UD_w > UD_c - SC$ 。通过实施承诺行动，威慑者的“抗击”策略由劣策略转化为占优策略，其威慑由不可信威慑转化为可信威慑。承诺行动实质上是一种威慑者内心意愿的外显表达，由于需要付出沉淀成本且满足  $SC > UD_c - UD_w$ ，因此它可以看做是一种昂贵信号。也就是说，只有当威慑信号是昂贵时，信息传递才会在承诺行动中起作用。

---

Henry A. Kissinger, *Nuclear Weapons and Foreign Policy* (New York: Harper&Row, 1957), p. 12

## 二、不完全信息下威慑信号传递

上面的分析表明,在完全信息条件下即当威慑博弈双方的支付都明确时,威慑者的可信威慑需要满足既有实力又有决心这一充分必要条件,并且只有昂贵信号才会在承诺行动中起作用。然而,完全信息是相对少见的情形,现实威慑情景中更普遍存在的是不完全信息的情形。在后一种情况下,由于挑战者不具有关于威慑者支付的真实信息,威慑者的可信威慑就不仅取决于威慑者是否具有威慑实力与决心,同时也取决于挑战者是否相信威慑者具有这样的实力和决心。也就是说,在不完全信息条件下,由于挑战者不知道威慑者支付的真实信息,因此威慑博弈的结局与挑战者对威慑者威慑实力与决心的信念密切相关:如果挑战者认识到威慑者具有威慑实力与决心,它将不发起挑战,威慑者威慑成功;如果挑战者不相信威慑者具有威慑实力与决心,它将发起挑战,威慑者威慑失败。下文将对不完全信息条件下威慑者的类型、威慑信号的类型以及信息传递方式对博弈结局的影响进行探讨。

### (一) 不完全信息条件下威慑博弈基本模型

自 20 世纪 70 年代以来,不完全信息博弈研究取得了突破性进展,这为研究不完全信息条件下的威慑博弈提供了强有力的分析手段。一般而言,博弈信息结构分为不完全信息和 imperfect information 两个方面,前者是指博弈方在支付信息方面的不对称,后者是指博弈方在博弈进程信息方面的不对称。“海萨尼转换”(Harsanyi Transformation)是一种处理不完全信息博弈的基本方法,它引入一个虚拟的参与人——“自然”,并把收益信息不完全的博弈方划分为不同的类型,从而将不完全信息博弈转换为“完全但不完美信息博弈”。这里,“类型”的概念在不完全信息博弈中具有重要作用:虽然某些博弈方在不完全信息条件下不能确定其他博弈方在一定策略组合下的效用,但至少能知道其他博弈方的效

---

在 1967—1968 年的 3 篇系列论文中,海萨尼开辟了不完全信息博弈研究的新领域,发展了“贝叶斯纳什均衡”概念,参见 J. C. Harsanyi, “Games with Incomplete Information Played by ‘Bayesian’ Players, Parts I, II and III,” *Management Science*, Vol 14, 1967—1968, pp. 159—182, 320—334, 486—502.

用有哪几种可能类型。这样,支付信息不完全的博弈方的策略选择将是类型依赖的,即不同类型的博弈方具有不同的策略偏好和不同的策略选择。根据“海萨尼转换”,具有不完全信息的博弈方究竟属于哪种类型,这是该博弈方的私有信息,但其属于每种类型的概率分布是共同知识,“自然”的作用就是在博弈开始前为实际博弈方抽取他们各自的类型。由此,“海萨尼转换”通过类型划分及其分布概率的共同知识假定,将不完全信息博弈转化为完全但不完美信息博弈。

鉴于类型划分在不完全信息条件下威慑博弈中的作用,以“威慑实力”与“威慑决心”作为两个维度,可以将威慑者划分为4种类型:(1)无实力无决心类型;(2)无实力有决心类型;(3)有实力无决心类型;(4)有实力有决心类型。由前面对可信威慑条件的分析已经知道,只有第4种类型的威慑者是可信威慑者,而前三种类型威慑者的威胁是不可信的。由此,在不完全信息条件下可将威慑者划分为“强硬”和“软弱”两种类型:“强硬”类型威慑者是可信的威慑者,“软弱”类型威慑者是不可信的威慑者。威慑者的策略选择是类型依赖的:在“抗击”和“容忍”两种策略中,“强硬”类型威慑者总是“抗击”,而“软弱”类型威慑者总是“容忍”。在威慑博弈的三种结局SQ、WAR和CW N中,不同类型的威慑者具有不同的偏好:“强硬”类型威慑者的偏好为 $SQ > WAR > CW N$ ,而“软弱”类型威慑者的偏好为 $SQ > CW N > WAR$ 。与威慑者具有两种类型不同,挑战者只有一种类型,其偏好为 $CW N > SQ > WAR$ 。挑战者不确切知道威慑者属于哪一种“类型”,但知道其为某一类型的概率,这一概率表征了挑战者对威慑者属于该类型的信念。

不完全信息条件下威慑博弈的基本模型,与完全信息条件下的两阶段威慑模型不同,由于增加了一个虚拟博弈方即“自然”,因此它是一种三阶段博弈模型。首先,由“自然”抽取威慑者类型的分布概率。其次,由挑战者决定是否发起挑战。挑战者不知道威慑者的真实类型,但知道它的类型分布概率。如果挑战者选择“维持”,博弈结局为SQ;如果挑战者“挑战”,博弈结局不确定。此时,当挑战者遭遇“软弱”类型威慑者时,结局为CW N;当挑战者遭遇“强硬”类

---

这是威慑格局得以形成的必要条件。

型威慑者时,结局为 WAR。当挑战者发起“挑战”后,博弈进入到第三阶段,将由威慑者进行策略选择:“强硬”类型威慑者将“抗击”,而“软弱”类型威慑者将“容忍”。设挑战者对威慑者为“强硬”类型的信念为 $p$ ,则对其为“软弱”类型的信念为 $1-p$ (当威慑者类型的信息完全时,可看做 $p$ 取1或0时的特殊情况)。令 $p^* \times UC(WAR) + (1-p^*) \times UC(CWN) = UC(SQ)$ ,可得:

$$p^* = \frac{UC(CWN) - UC(SQ)}{UC(CWN) - UC(WAR)}$$

$p^*$ 被称为成功威慑的转折概率。

威慑博弈的贝叶斯均衡为:(1)当 $p > p^*$ 时,威慑者“抗击”,挑战者不挑战,威慑者威慑成功;(2)当 $p < p^*$ 时,威慑者“容忍”,挑战者挑战,威慑者威慑失败。

## (二) 威慑信号传递机制分析

如上所述,在不完全信息条件下,挑战者不完全清楚威慑者对每种结局的真实效用,因而将威慑者划分为“强硬”和“软弱”两种类型,不同类型的威慑者具有不同的结局偏好和策略选择。挑战者对威慑者属于某一类型的信念本质上是一种主观概率。当威慑者的信号传递能够改变挑战者的这种信念时,它对威慑的可信性具有重要影响。这种信念改变在博弈论中被假设为一种贝叶斯过程,改变前的信念称为先验信念,改变后的信念称为后验信念,后验信念的值根据先验信念的值与威慑信号对信念的影响程度,按照贝叶斯公式计算。当挑战者对威慑者属于“强硬”类型的信念高于转折概率时,挑战者将不发起挑战,威慑者威慑成功,此时无须威慑信息传递起作用;当挑战者对威慑者属于“强硬”类型的信念低于转折概率时,挑战者将发起挑战,威慑者威慑失败,此时为了阻挑战者发起挑战,威慑者需要主动选择并发出清晰的威慑信号,准确表达自身的真实类型,以使对方对自身类型的后验信念大于转折概率,从而使威慑能获得成功。这种威慑者主动发出威慑信号以表达自身真实类型的威慑博弈,实质上是一种信号博弈(Signaling Game)。

### 1. 信号博弈及信号类型区分

信号博弈是一种应用十分广泛的具有信息传递机制的不完全信息动态博

弈,由美国学者斯宾塞 (Michael Spence)在 20 世纪 70 年代首次提出并运用。信号博弈的博弈方由信号发送者与信号接收者构成,其中拥有私有信息的博弈方是信号发送者,而信息完全的博弈方是信号接收者。在威慑情景中,威慑方是信号发送者,它通过精心选择并发送信号,传递自身类型的信息;挑战者是信号接收者,它通过对威慑者所发送信息的性质判断,决定是否改变对威慑者类型的先验信念和是否发起挑战。正如将威慑者区分为不同类型一样,威慑信号也需要区分不同的类型,因为威慑信号并不总是有效或必然起作用。根据后验信念与转折概率的比较,可以把威慑者发出的信号分为“强信号”与“弱信号”:“强信号”是一种威慑者发送后,能够使挑战者的后验信念从小于转折概率转换到大于转折概率的信号,而“弱信号”则不能实现这样的转换。根据对威慑者发出信号的类型判断,挑战者确定是“维持”还是“挑战”。

设“强硬”和“软弱”威慑者类型分别以  $T$  和  $W$  表示,“强”和“弱”威慑信号类型分别以  $R_S$  和  $R_W$  表示,则在威慑博弈中威慑者与威慑信号之间有四种组合:(1)“自然选定  $T$  时发送  $R_S$ ,”自然选定  $W$  时仍发送  $R_S$ ;(2)“自然选定  $T$  时发送  $R_S$ ,”自然选定  $W$  时则发送  $R_W$ ;(3)“自然选定  $T$  时发送  $R_W$ ,”自然选定  $W$  时则发送  $R_S$ ;(4)“自然选定  $T$  时发送  $R_W$ ,”自然选定  $W$  时仍发送  $R_W$ 。其中,在组合 (1) 和组合 (4) 中,“自然选定的不同类型威慑者都发送相同的威慑信号,称为混同 (pooling) 信号;在组合 (2) 和组合 (3) 中,“自然选定的不同类型威慑者分别发出不同的威慑信号,称为分离 (separating) 信号。显然,威慑者对不同类型威慑信号的发送,将影响挑战者对威慑者类型的判断,进而影响威慑可信性。为了实施有效威慑,“软弱”类型威慑者力求表现得“强硬”,会选择混同信号;而“强硬”类型威慑者则通过发送“分离”信号,充分显示威慑实力和决心,以揭示自身真实类型。

与信号博弈相对应的均衡概念,是混同均衡和分离均衡,它们本质上是不完全信息动态博弈中的“完美贝叶斯均衡”。完美贝叶斯均衡是贝叶斯均衡、子博弈完美均衡和贝叶斯推断的结合,它不再是策略组合,而是策略加信念的

---

信号博弈模型由斯宾塞 1973 年在劳动力市场模型中首次运用,并在 1974 年的论文中使其论证更为一般化,参见 Michael Spence, “Job Market Signaling,” *Quarterly Journal of Economics*, Vol 87, No 3, August 1973, pp. 355—374; Michael Spence, “Competitive and Optimal Responses to Signaling: An Analysis of Efficiency and Distribution,” *Journal of Economic Theory*, Vol 7, No 3, March 1974, pp. 296—332。

组合。

2 基于信号博弈的威慑模型

威慑博弈的信号模型如图-2所示，图中圆圈内的  $N$ 、 $D$ 、 $C$  分别表示“自然”、威慑者和挑战者的决策节点。“自然”选择威慑者的类型，威慑者为“强硬”类型  $T$  和“软弱”类型  $W$  的先验信念分别为  $\pi$  和  $1 - \pi$ 。无论哪种类型威慑者，都有“强”信号  $R_s$  和“弱”信号  $R_w$  可以发送，设  $p$  和  $q$  分别是威慑者发出  $R_s$  和  $R_w$  后挑战者对威慑者为  $T$  类型的后验信念，则威慑者发出  $R_s$  和  $R_w$  后挑战者对威慑者为  $W$  类型的后验信念分别为  $1 - p$  和  $1 - q$ 。挑战者不知道威慑者的真实类型，只能根据威慑者发出的信号类型来修正其对威慑者先验信念的判断，它有“挑战”和“维持”两种策略可以选择。当遭遇  $T$  类型威慑者时，如果挑

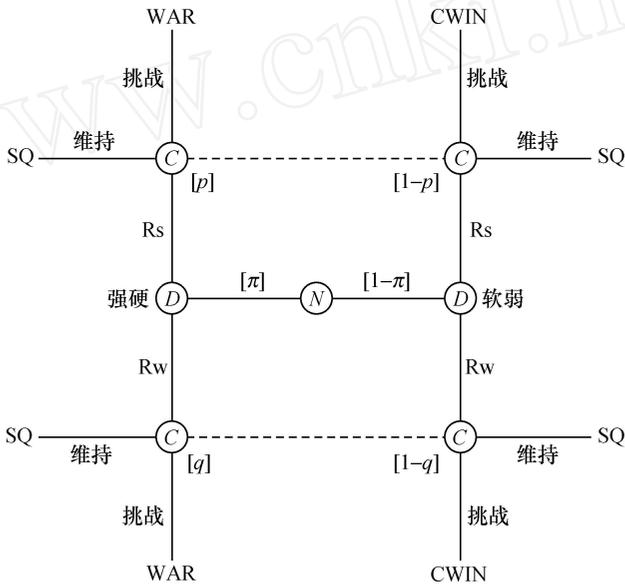


图-2 威慑的信号博弈模型

不完全信息动态博弈中的“完美贝叶斯均衡”概念，其正式定义是由弗登博格和梯若尔在 1991 年给出的，参见 D. Fudenberg and J. Tirole, “Perfect Bayesian Equilibrium and Sequential Equilibrium,” *Journal of Economic Theory*, Vol 53, No 2, April 1991, pp. 236—260.

图-2 表现形式源自曹 (I K Cho) 和克莱普斯 (D. Krep)。本文结合问题情景对之进行了变形，参见 I K Cho and D. Krep, “Signaling Games and Stable Equilibria,” *Quarterly Journal of Economics*, Vol 102, No 2, May 1987, pp. 179—221。

挑战者“挑战”，结局为 WAR；如果挑战者“维持”，结局为 SQ。当遭遇 W 类型威慑者时，如果挑战者“挑战”，结局为 CW N；如果挑战者“维持”，结局为 SQ。

设争夺目标对挑战者和威慑者的效用分别为  $u_c$  和  $u_d$ ，威慑者发出  $R_s$  和  $R_w$  成本的效用分别为  $m_s$  和  $m_w$ ，挑战者在威慑者发出  $R_s$  和  $R_w$  时敢于挑战的声誉的效用分别为  $\xi$  和  $\eta$ 。当结局为 SQ 时，威慑者在发出  $R_s$  和  $R_w$  后的效用分别是  $u_d - m_s$  和  $u_d - m_w$ ，挑战者相应的效用分别为  $-\xi$  和  $-\eta$ ；当结局为 CW N 时，威慑者在发出  $R_s$  和  $R_w$  后的效用分别是  $-u_d - m_s$  和  $-u_d - m_w$ ，挑战者相应的效用分别为  $u_c + \xi$  和  $u_c + \eta$ 。对于结局 WAR，设威慑者和挑战者战争代价的效用分别为  $c_d$  和  $c_c$ ，赢得战争的概率分别为  $p$  和  $1 - p$ ，则在发出  $R_s$  时威慑者的期望效用是：

$(u_d - c_d - m_s) + (1 - p)(-u_d - c_d - m_s) = (2 - 1)p u_d - c_d - m_s$ ，挑战者相应的期望效用是： $(1 - p)(u_c - c_c + \xi) + p(-u_c - c_c + \xi) = (1 - 2p)u_c - c_c + \xi$ ；同理，在发出  $R_w$  时威慑者和挑战者的期望效用分别是：

$$(2 - 1)p u_d - c_d - m_w \quad \text{和} \quad (1 - 2p)u_c - c_c + \eta$$

当挑战者观察到“强”信号  $R_s$  后，如果“挑战”和“维持”两种策略无差异时，满足：

$$p^* [(1 - 2p)u_c - c_c + \xi] + (1 - p^*)(u_c + \xi) = p^*(-\xi) + (1 - p^*)(- \xi) \quad (1)$$

$$p^* = \frac{u_c + 2\xi}{2u_c + c_c} \quad (2)$$

当挑战者观察到“弱”信号  $R_w$  后，如果“挑战”和“维持”两种策略无差异时，有：

$$q^* [(1 - 2p)u_c - c_c + \eta] + (1 - q^*)(u_c + \eta) = q^*(-\eta) + (1 - q^*)(- \eta) \quad (3)$$

$$q^* = \frac{u_c + 2\eta}{2u_c + c_c} \quad (4)$$

后验信念  $p = p(T/R_s)$  和  $q = q(T/R_w)$  根据贝叶斯公式计算为：

$$p = \frac{p(T)p(R_s/T)}{p(T)p(R_s/T) + p(W)p(R_s/W)} \quad (5)$$

$$p = \frac{p(R_S/T)}{p(R_S/T) + (1 - )p(R_S/W)} \quad (6)$$

同理可得：

$$q = \frac{p(R_W/T)}{p(R_W/T) + (1 - )p(R_W/W)} \quad (7)$$

### 3 威慑博弈均衡

在以上分析中,  $p^*$  和  $q^*$  分别是威慑者发出“强”信号和“弱”信号后, 挑战者认为威慑者属于“强硬”类型的转折概率, 而  $p$  和  $q$  分别表征威慑者发出“强”信号和“弱”信号后挑战者认为威慑者属于“强硬”类型的后验信念。显然有: 当  $p > p^*$  时, 挑战者的最优反应是“维持”, 反之则“挑战”; 当  $q > q^*$  时, 挑战者的最优反应是“维持”, 反之则“挑战”。

威慑者信号传递组合有  $(R_S, R_S)$ 、 $(R_W, R_W)$ 、 $(R_S, R_W)$  和  $(R_W, R_S)$  四种, 括号内第一、第二个符号分别对应威慑者为“强硬”和“软弱”类型时发出的信号。挑战者的回应策略组合也有  $(Ch, Ch)$ 、 $(Su, Su)$ 、 $(Ch, Su)$  和  $(Su, Ch)$  四种, 括号内第一、第二个符号分别对应挑战者对“强”信号和“弱”信号的回应策略, 其中以  $Ch$  表示“挑战”,  $Su$  表示“维持”。

#### (1) 挑战均衡 ( $p < p^*$ 且 $q < q^*$ )

当  $p < p^*$  且  $q < q^*$  时, 挑战者的最优反应是  $(Ch, Ch)$ , 即无论威慑者发出什么样的信号, 挑战者都“挑战”。由于  $0 < m_W < m_S$ , 即威慑者发出“强”信号的成本效用要高于发出“弱”信号的成本效用, 此时“强硬”和“软弱”类型威慑者都将发出“弱”信号, 即  $q = \frac{m_S}{m_S + (1 - )m_W}$ ,  $q = \frac{m_S}{m_S + (1 - )m_W}$ 。

故  $[(R_W, R_W) (Ch, Ch); p, q = \frac{m_S}{m_S + (1 - )m_W}]$  是此种情况下的混同均衡。根据本文前面对“强”信号和“弱”信号的定义, 这一混同均衡意味着: 无论“强硬”还是“软弱”类型威慑者, 都没有有效的威慑信号可以采用, 以使挑战者的后验信念大于转折概率, 因此威慑必然失败。该混同均衡也可以称为“挑战均衡”, 因为此时挑战者必然挑战。

#### (2) 威慑均衡 ( $p > p^*$ 且 $q > q^*$ )

当  $p > p^*$  且  $q > q^*$  时, 挑战者的最优反应是  $(Su, Su)$ , 即无论威慑者发出什么样的信号, 挑战者都“维持”。由于威慑者发出“强”信号的成本效用要高于发出“弱”信号的成本效用, 此时“强硬”和“软弱”类型威慑者都发出“弱”信

号。故  $[ (Rw, Rw) (Su, Su); p, q = ]$  是此种情况下的混同均衡。

这一混同均衡意味着,即使是发出“弱”信号,威慑者也能威慑成功,因此它可以称为“威慑均衡”。挑战者对威慑类型的先验信念大于转折概率时的博弈均衡,就是这种情形。

(3) 理论均衡 ( $p < p^*$  且  $q > q^*$ )

当  $p < p^*$  且  $q > q^*$  时,挑战者的最优反应是  $(Ch, Su)$ ,即挑战者在威慑者发出“强”信号时“挑战”,在威慑者发出“弱”信号时“维持”。此时无论什么类型威慑者都选择“弱”信号,挑战者在观察到“弱”信号后“维持”,故  $[ (Rw, Rw) (Ch, Su); p, q = ]$  是此种情况下的混同均衡。

这一混同均衡意味着,威慑者发出“强”信号威慑失败而发出“弱”信号则威慑成功,因此它只是一种在形式化分析中的“理论均衡”,而不是理性行为假设下的真实均衡。

(4) 条件均衡 ( $p > p^*$  且  $q < q^*$ )

当  $p > p^*$  且  $q < q^*$  时,挑战者的最优反应是  $(Su, Ch)$ ,即挑战者在威慑者发出“强”信号时“维持”,在威慑者发出“弱”信号时“挑战”。这一情形意味着,“强硬”与“软弱”类型威慑者在发出信号的能力上存在区别,即“强硬”类型威慑者能够发出“弱”信号,“软弱”类型威慑者难以发出或无法发出的“强”信号。设“软弱”类型威慑者的信号集为  $WR$ ,“强硬”类型威慑者的信号集为  $TR$ ,则  $WR$  是  $TR$  的子集。令  $DR = TR - WR$ , $DR$  是“强硬”类型威慑者能够发出但“软弱”类型威慑者没有能力发出的信号集。再设  $tr \in TR, wr \in WR, dr \in DR$ 。此时存在分离均衡  $[ (dr, wr) (Ch, Su); p = 1, q = 0 ]$ 。在这样的分离均衡下,挑战者能够通过对威慑信号类型的判断,将不同类型的威慑者区分开来:“强硬”类型威慑者发出“强”信号,挑战者“维持”,结局为  $SQ$ ,威慑者威慑成功;“软弱”类型威慑者因无法或难以发出“强”信号,只能发出“弱”信号,挑战者“挑战”,结局为  $CWN$ ,威慑者威慑失败。

分离均衡又可以称为“条件均衡”,即只有在威慑者满足发出“强”信号条件时,威慑才会成功。

### (三) 威逼信号传递模型

威慑概念假设威慑者能够在挑战者发起挑战之前,通过武力威胁方式慑阻

挑战者的挑战,从而避免不想要的结局出现。如果威慑者事先未能感知挑战者可能的挑战,或者说挑战者在没有受到威慑的情况下发起了挑战,则出现了威慑者不想要的结局,此时威慑者仍将发出武力威胁,以迫使挑战者撤回挑战。这种在挑战者实施了实质性的挑战之后的武力威胁,称为威逼(compellence)。谢林最早对威慑与威逼这两个概念做了区分,并认为威逼比威慑更难成功。威逼之所以比威慑更难以成功,是因为威逼涉及挑战者已经付出的实质性代价和撤回挑战承受的声誉损失,因此在威逼中威慑者通常更加注重“胡萝卜加大棒”策略的运用,以在一定程度上“补偿”挑战者的某些利益和声誉损失。

威逼是一个时常被混淆的概念,研究者一方面容易将威逼与威慑两种概念不加区分,另一方面又常常忽视威逼概念的被动特征,而后一点尚没有被明确指出。例如,罗伯特·阿特(Robert Art)认为,威逼与威慑的区别“就是主动与被动使用武力的区别”。理查德·哈斯(Richard N. Haass)也认为,威逼的“一个传统例子或者模式,就是炮舰外交”。克劳迪奥·乔菲-雷维利亚(Claudio Cioffi-Revilla)在区分威慑与威逼概念时认为:尽管两者均以“实力”为手段,但威慑的目标是维持现状,消除对现状挑战的潜在威胁,而威逼旨在改变现状,以取得新的平衡。笔者认为,威逼概念从根本上说是一个具有被动特征的概念,虽然它采用了逼迫或强制挑战者撤回挑战这样的主动形式,但其根本目的是为了回避不想要结局的出现,此时威慑者与挑战者之间对不同结局的偏好或威慑格局并没有改变。如果将威逼概念表征为“主动使用武力”或“炮舰外交”以追求所想要的结局,那么此时威慑者对结局的偏好已经发生了根本改变,威慑格局形成的条件也不再存在,威慑者实质上已经转化为挑战者。因此,威逼虽然采用了主动形式,但它仍然是威慑者的被动行为,旨在避免不想要的结局出现,而不是旨在主动改变现状的挑战者的主动行为。

根据上述分析,威逼博弈基本模型是一个与两阶段威慑博弈模型不同的三阶段模型。在第一阶段,挑战者发起挑战。在第二阶段,威慑者有“容忍”和

---

Thomas Schelling, *Arms and Influence* (New Haven: Yale University Press, 1966), pp. 69—91.  
转引自理查德·N.哈斯著《新干涉主义》(殷雄等译),北京,新华出版社2000年版,第52页。  
同上书,第54页。

Claudio Cioffi-Revilla, “Origins and Age of Deterrence: Comparative Research on Old World and New World Systems,” *Cross-Cultural Research*, Vol 33, No 3, August 1999, p. 241.

“逼迫”两种策略。当威慑者“容忍”时,结局为 CW N;当威慑者“逼迫”时,博弈进入第三阶段。在第三阶段,挑战者有“后退”与“坚持”两种策略。当挑战者“后退”时,结局为“恢复现状”RSQ;当挑战者“坚持”时,结局为 WAR。对于不完全信息条件下的威逼模型,可以看做是一种带有信号子博弈的“试探”模型。挑战者的“挑战”可以看做是一种“试探”行为,由此引发威慑者做出反应,博弈进入信号子博弈。与上述威慑信号博弈模型类似,威慑者可以发出“强”信号和“弱”信号两种信号。挑战者根据威慑者发出信号的性质,决定“后退”还是“坚持”。如果挑战者“后退”,结局为 RSQ;如果挑战者“坚持”,当遇到“强硬”类型威慑者时结局为 WAR,遇到“软弱”威慑者时结局为 CW N。信号子博弈仍有三种混同均衡和一种分离均衡,威逼博弈的完美贝叶斯均衡为:

- (1) 当  $p > p^*$  且  $q < q^*$  时,“强硬”类型威慑者发出“强”信号,挑战者“退却”,结局为 RSQ;“软弱”类型威慑者发出“弱”信号,挑战者“坚持”,结局是 CW N。
- (2) 在其他情况下,挑战者无法从混同信号中区分出不同类型的威慑者,它将按照对威慑者类型的先验信念采取行动。具体的形式化分析这里不再展开。

### 三、对威慑博弈的案例分析

从前面两个部分的博弈分析可以看出,威慑格局与可信威慑的形成都是有条件的,威慑信息传递也并不总是起作用。以下结合四个个案,分别进行简要分析。在第一个个案中,本文认为伊拉克入侵科威特前美伊之间的策略互动并未形成一种威慑格局,因此伊拉克入侵科威特并不像学术界所认为的那样是美国延伸威慑的一次失败,以此解释美国当时何以会对伊拉克发出含糊不清的信息。在第二个个案中,本文从“挑战均衡”的角度,认为在中国出兵朝鲜前中美之间的策略互动中,即使信息沟通渠道畅通、信息传递者立场中性、威慑信息表达明确,中国对美国的威慑也难以获得成功。在第三、四两个威逼个案中,本文从分离均衡的角度,对比分析了美国在古巴导弹危机中对苏联威逼成功、英国在马岛危机中对阿根廷威逼失败的原因。

#### (一) 伊拉克入侵科威特前美伊之间的策略互动

1990年7月17日,伊拉克总统萨达姆·侯赛因发表广播电视讲话,不点

名地指责海湾一些国家不遵守石油输出国组织所规定的配额,大量倾销原油,造成国际市场石油价格下跌,致使伊拉克经济利益受到损害。翌日,伊拉克官方媒体点名抨击科威特和阿联酋等国大量超额出口石油,重点指责科威特侵占伊拉克边境领土,掠夺伊石油资源达 10 年之久,称此为“对伊拉克的侵略与战争”。针对伊拉克咄咄逼人的行动,经埃及、沙特等国居中斡旋,伊、科两国于 7 月 31 日赴沙特举行直接谈判。由于伊拉克提出的要求太高,科方无法接受如此苛刻的条件,阿拉伯方面调解受阻。8 月 1 日,伊、科谈判破裂,第二天伊拉克大举入侵科威特。

伊拉克入侵科威特的举动,被学术界看做是美国延伸威慑的一次失败。对于美国何以未能威慑伊拉克,哈斯认为,是海湾“当地的政治环境不允许美国军队在该国驻扎”。由于美军“必须驻扎在该国公众看不见的地方”,也就“不得不驻扎在该国的外部敌人所看不到的地方”,这样就无法形成一种“通过军队进行威慑的压力”。因此,当地缺少美国军事能力的存在是未能劝阻萨达姆入侵科威特的主要原因之一。此外,由于担心激怒伊拉克,海湾各国(阿联酋除外)拒绝参加美国提议的于 1990 年 7 月底举行的军事演习。不过哈斯也认为,萨达姆不理解美国对科威特所承诺义务的深刻意义是未能劝阻萨达姆入侵科威特的另一个主要原因。“延伸威慑是一种微妙的手段。对于美国来说,这种手段并不总是能够起到明确的威慑效果,而且如果对这种手段怀有三心二意,那就更不会有什效果了。”贾尼丝·格罗塞·斯坦(Janice Gross Stein)在对海湾战争中进行威慑的教训分析后认为:虽然美国犯了没有明确表达自己利益的错误,但还是有充分的理由认为威慑是一项不可能完成的任务,这是由萨达姆的心态所决定的。从信号传递角度分析,20 世纪 90 年代初,随着苏联的衰弱,美国不仅具有强大的军事实力(即属于“强硬”类型的先验信念很高),而且如果要表达自己的威慑意愿,其可供选择的威慑信号也有很多,但美国最

---

赵国忠主编:《海湾战争后的中东格局》,北京,中国社会科学出版社 1995 年版,第 10 页。

理查德·N. 哈斯:《新干涉主义》,第 79—80 页。

同上书,第 80 页。

参见《海湾战争:深刻的军事教训》,华盛顿战略与国际研究中心,1991 年 7 月,第 42 页。转引自理查德·N. 哈斯:《新干涉主义》,第 79 页。

Janice Gross Stein, “Deterrence and Compellence in the Gulf, 1990—1991: A Failed or Impossible Task?” *International Security*, Vol 17, No 2, Autumn 1992, pp. 147—179.

终对萨达姆发出了含糊不清的信息。也就是说,美国能够发出分离信号,却发出了混同信号。那么,伊拉克入侵科威特,确实是美国延伸威慑的一次失败吗?

1990年2月12日,美国负责中近东事务的助理国务卿约翰·凯利(John Kelly)正式访问巴格达,他在与萨达姆的谈话中称:“你们是本地区的一支温和力量,美国希望发展与伊拉克的关系。”他还表示美国对伊拉克同科威特的边界冲突不感兴趣。7月23日,美国间谍卫星侦察到伊拉克将3个师约3.5万人的兵力部署在离科威边境30公里的地区。7月24日,美国为了表示关切,与阿拉伯联合酋长国举行了海军联合演习。但美国国务院女发言人玛格丽特·塔特怀勒在回答记者提问时强调:美国同科威特没有防务条约,对它也没有什么特殊的防务承诺。7月24日深夜,在萨达姆召见美国驻伊拉克大使阿普里尔·格拉斯佩的会晤中,美国大使对于伊科之间的边界纠纷表示:“我们不作评论。60年代末我曾在科威特任职,当时我们接到的指示是,在这个同美国无关的问题上不表态”,“我们的全部希望,就是你们迅速提出一项解决办法”。会晤结束后,这位大使打电话给华盛顿,敦促美国在伊拉克与科威特自行解决它们的争端以前,“少作批评”。在尔后的一份致萨达姆的电报中,布什(George Bush)称他的政府“继续愿与伊拉克保持较好的关系”。

7月27日,美国中央情报局向白宫传送了一些卫星拍摄的照片,照片显示出,集结的人员和装备越来越多。7月28日,美国中央情报局起草的报告更为具体,萨达姆为其部署在边界的部队建立起重要的供应线,从照片上还可以看出大批提供后勤支援的卡车。中央情报局局长威廉·韦伯斯特(William Webster)确信,如果仅仅是一次恫吓行动,这么大规模的后勤是不必要的。7月30日,美国中央情报局已经能够提出伊拉克集结在科威特附近的军力情况:包括

---

皮埃尔·塞林格、埃里·洛朗:《海湾战争:秘密档案》(廖先旺等译),北京,世界知识出版社1991年版,第3—4页。

哈立德·本·苏尔坦、帕特里克·西尔:《沙漠勇士:联合国部队司令对海湾战争的己见》(郑道根等译),北京,军事谊文出版社1996年版,第214页。

同上书,第215页。

军事科学院军事历史研究部:《海湾战争全史》,北京,解放军出版社2000年版,第62—63页。

科林·鲍威尔:《我的美国之路》(王振西主译),北京,昆仑出版社1996年版,第513页。

解力夫编著:《海湾战争实录》,北京,世界知识出版社1994年版,第61页。

精锐部队共和国卫队在内的 10 万大军, 300 辆坦克和 300 门重型火炮。华盛顿依然保持沉默。7 月 31 日, 美国助理国务卿凯利到美国众议院“中东小组委员会”作证, 针对众议员李·汉密尔顿的提问, 凯利回答说: “我们同海湾各国没有防务协定, 这是显而易见的。但是我们支持那个地区友好国家的独立和安全。”当李·汉密尔顿又问: “说我们没有一个条约、没有一项义务迫使我们派出美国部队, 这对吧?” 凯利对此作了肯定的回答。在这个战争与和平的紧要关头, 英国 BBC 广播电台立刻播出了美国助理国务卿凯利的上述讲话, 向伊拉克总统转达了美国不卷入海湾冲突的保证。有研究者指出, 美国“这些错误的外交信号一定毫无疑问地鼓励了萨达姆采取行动”。

笔者认为, 美国在海湾危机中并不存在对伊拉克入侵的延伸威慑及所谓延伸威慑的失败, 因为美伊之间的互动关系并不构成一种威慑格局, 美国对伊拉克的策略实质上是“诱导”而不是威慑。从对结局的偏好进行分析, 在“维持现状” $SQ$ 、“干涉伊科冲突” $WAR$ 和“默认伊拉克侵占科威特” $CW N$ 这三种结局中, 如果美国进行干预或发出明确威慑信号, 则伊拉克对结局的偏好是  $CW N > SQ > WAR$ , 此时由于伊拉克对美国属于“强硬”类型的信念很高, 它将不发起挑战, 即美国威慑成功; 如果美国不准备进行干预或发出含糊不清甚至带鼓励性的信息, 由于伊拉克军事实力远高于科威特, 则伊拉克对结局的偏好变为  $CW N > WAR > SQ$ , 此时  $SQ$  是最差结局, 发起挑战是伊拉克的占优策略, 伊拉克必然入侵伊拉克。长期以来, 美国一直想驻兵海湾而由于海湾国家的戒备而未能如愿 (即  $SQ$  并不是美国的最好结局), 一旦伊拉克发起挑战, 将给美国出兵海湾以契机。美国至少可以达成以下目标: (1) 打出反侵略的旗号, 大规模出兵海湾, 以遏制伊拉克的扩张势头; (2) 利用伊拉克入侵科威特事件对沙特等国造成的恐惧心理和现实威胁, 促使它们同意美军大规模进驻, 实现美国多年来企图在海湾地区重建军事基地的愿望; (3) 利用联合国带头对伊拉克进行制裁和封锁, 显示美国在处理国际事务中的“先锋”作用; (4) 利用美国在处

解力夫编著:《海湾战争实录》,第 62 页。

皮埃尔·塞林格、埃里·洛朗:《海湾战争》,第 56—57 页。

军事科学院军事历史研究部:《海湾战争全史》,第 64 页。

哈立德·本·苏尔坦、帕特里克·西尔:《沙漠勇士》,第 214 页。

张相元等主编:《海湾战争纵览》,北京,海潮出版社 1992 年版,第 11 页。

理海湾危机中的重要影响和赢得的广泛支持,提高布什在国内的威望,同时也借海湾危机缓解国内大幅度削减军费的压力,继续维持一支庞大的军事力量。

由上述分析可知,维持现状虽然是科威特的最好结局,但并不是美国的最好结局,因此通过武力威胁的方式慑阻伊拉克入侵科威特以维持现状并不符合美国的最大利益。换句话说,不是美国不能发出清晰明确的“强”信号,而是美国不愿意传递这样的信息。美国既希望伊拉克发起挑战,又不愿意被海湾国家认为未尽责任,因此美国既与阿联酋举行了象征性的威慑演习,同时又多次强调与科威特没有防务协定。其含糊不清的信息传递,实际上起到了“诱导”伊拉克入侵的行动。此后的事实证明,美国是海湾危机的最大受益者。

### (二) 中国出兵朝鲜前中美之间的策略互动

1950年6月24日,朝鲜战争爆发。6月25日傍晚,美国总统杜鲁门命令美国海空军投入战斗。在9月15日美军于仁川登陆成功并逼向“三八”线后,中国多次明确地对美军发出不要越过“三八”线的警告。9月27日,中央军委代总参谋长聂荣臻授命向力图调停的印度政府转达了中国方面的强硬姿态,指出:“如果帝国主义果真要发动战争,那么,我们也只有起而抵抗了。”9月30日,周恩来总理出席政协全国委员会时庄严声明:“中国人民决不能容忍外国的侵略,也决不能听任帝国主义者对自己的邻邦肆行侵略而置之不理。”10月3日凌晨1时,周恩来紧急约见印度驻华大使潘尼迦,再次郑重表明中国立场:如果美军越过“三八”线,中国不能坐视不顾,中国一定要管。这是中国避免出现被迫出兵的情况而通过外交途径向美国发出的最明确的警告。然而,中国发出的威慑信号未能阻止美国的进一步军事行动。10月7日,美国开始越过“三八”线。10月19日,中国人民志愿军入朝参战,与美国进行了面对面的军事较量。

---

D. Acheson, *Present at the Creation* (New York: W. W. Norton & Company Inc., 1969), p. 408.

柴成文、赵勇田:《板门店谈判》,北京,解放军出版社1989年版,第74页。

转引自谢益显主编:《中国当代外交史》,北京,中国青年出版社1997年版,第45页。

参见《周恩来外交文选》,北京,中央文献出版社1990年版,第25页。

对于中国的警告何以未起作用,一些学者从危机管理角度进行了分析。国内有学者对朝鲜战争和越南战争期间中美信息沟通进行了比较研究,也得出了一些思考。国外有学者从声誉的角度分析了中国出兵朝鲜前对美国的“诚实的威胁”(honest threat),认为外交等“廉价信号”在威慑互动中也会发挥有价值的作用,因为威慑者为了获得声誉以在未来的争端中发出的威慑有效,也会发出“诚实的威胁”信号;而没有决心的威慑者可能因为虚张声势的声誉,影响到其未来发出威胁信号的真实性。朝鲜战争中中国对美国不要越过“三八”线发出的外交威慑,就是这样一种“诚实的威胁”,而美国却误读了这种信号。

本文从信号博弈角度认为,中美之间的威慑博弈结局是一种典型的“挑战均衡”。具体地说,由于美国对中国属于“强硬”类型的先验信念很低(即  $v$  值很小),中国要想威慑成功的转折概率很高(即  $p^*$  值很大)以及中国由于当时条件限制无“强”信号可选,中国只能发出混同信号,无法从根本上改变美国对自身类型的认知,因此威慑很难成功。

鉴于中国是一个不发达的国家,刚刚建国,百废待兴,综合实力和武器装备差,而美国是世界上综合实力和军事实力最强大的国家,并且是刚结束的第二次世界大战的最大赢家,因此美国对中国属于“强硬”类型的先验信念很低,认为中国不会出兵是一种很自然的想法。1950年10月5日,杜鲁门总统在威克岛会见麦克阿瑟和其他美军高级将领时,麦克阿瑟向杜鲁门汇报朝鲜战局说,虽然中国在中国东北集中了40万军队,“但根据我的专业知识判断,在联合国军拥有空中优势的情况下,中国共产党领导人不会蠢到派遣大规模常规部队到朝鲜前线。因为中国军队的主要力量在到达平壤前就会遭遇联合国军的重创,并被消灭”。当杜鲁门提出“中国或苏联干预(朝鲜战争)的可能性时,麦克

---

参见 Phil Williams, *Crisis Management: Confrontation and Diplomacy in the Nuclear Age* (London: Martin Robertson & Co Ltd, 1976), pp. 185—187; 夏立平:《美国关于危机管理的理论与实践:以中美关系为例》,《美国研究》2003年第2期,第73—86页;王缉思、徐辉:《中美危机行为比较分析》,《美国研究》2005年第2期,第22—46页。

Anne Sartori, "The Might of the Pen: A Reputational Theory of Communication in International Disputes," *International Organization*, Vol 56, No 1, 2002, pp. 123—151.

参见韩国国防部战争史编辑委员会:《朝鲜战争》第1卷,牡丹江,黑龙江朝鲜民族出版社1988年版,第4页。

阿瑟回答说：“非常小。”10月底，当中国军队首次出战，狠狠打击了孤军深入的美军和南朝鲜军时，麦克阿瑟还认为这无关紧要，只不过会暂时迟滞他向鸭绿江的大举进攻。不仅美国对中国进行了这样的类型定位，即使在中国党内和军内对是否出兵朝鲜意见也不一致。“朝鲜战局的恶化使得中共中央领导内部对出兵条件的估计变得谨慎多了，也顾虑多了。除了毛泽东以外，与会者几乎都对迅速出兵感到没有太多把握。”“在10月2日下午举行的政治局扩大会议上，几乎所有领导人都对现在出兵朝鲜持怀疑和反对态度。而最重要的是军队领导人几乎一致对同美军作战表示没有把握。”事实上，中国出兵朝鲜的决策经历了一次分析框架的转换：从单纯考虑出兵的风险，到对出兵与不出兵两种方案进行权衡。决策伊始，充分考虑了“要管”的风险：“首先，美军在釜山的有效抵抗、实施仁川登陆大举围歼朝鲜人民军主力以及对北朝鲜的狂轰乱炸，不仅显示出美国军队在装备及火力上拥有巨大优势，而且表明它在指挥和战术上也不可小视。在装备及火力对比如此悬殊的条件下贸然投入对美作战，是否有利呢？”其次，在经过了多年战争之后，人民刚刚得到和平和稳定，国家刚刚开始进入经济的全面恢复和建设时期。出兵朝鲜，对美作战，不能不考虑美国向中国宣战的严重性。”但在反复讨论后，政治局领导人一致认为，虽然出兵风险很大，但与不出兵相比，出兵是最好的选择：“让敌人压至鸭绿江边，国内国际反动气焰增高，则对各方都不利，首先是对东北更不利”；“总之，我们认为应当参战，必须参战，参战利益极大，不参战损害极大。”

除了上述美国对中国的类型定位因素外，中国要获得威慑成功的转折概率  $p^*$  也非常高。在  $p^*$  值的计算式中，美国对自身军事实力看得过高（由此低估战争代价  $c_c$ ），对作为强国在中国威慑下被慑阻的声誉损失看得很重（即公式中  $k_3$  值大），因此同样的  $u_c$  值下  $p^*$  值增大。由于先验信念与转折概率  $p^*$  之

---

参见韩国国防部战争史编辑委员会：《朝鲜战争》第1卷，牡丹江，黑龙江朝鲜民族出版社1988年版，第4页。

阿伦·米利特、彼得·马斯洛斯基：《美国军事史》（军事科学院外国军事研究部译），北京，军事科学出版社1989年版，第491页。

杨松奎：《毛泽东与莫斯科的恩恩怨怨》，南昌，江西人民出版社2005年版，第378页。

同上书，第379页。

同上书，第377页。

参见《建国以来毛泽东文稿》第1册，北京，中央文献出版社1987年版，第556页。

间的差距扩大,必然增大选择“强”信号的难度,而中国当时几乎没有能够揭示自身真实类型的“强”信号可供选择,因此只能发出混同信号。美国无法从中国发出的混同信号中进行信念更新,仍按原有信念越过“三八”线,最终中国被迫出兵朝鲜。因此,即使信息传递渠道畅通、信息传递者立场中性、信息表述十分明确,美国对中国类型的认知也不会因中国发出的混同信号而改变,中国的威慑很难成功。再从声誉角度看,虽然声誉在长期互动关系中起着十分重要的作用,中国为了不破坏声誉有动机通过外交声明等“弱”信号表达“诚实的威胁”,但声誉毕竟是建立在对以往行为的认知基础上的,它具有统计特性,需要在“重复”中建立,其本身就是一种后验信念。而朝鲜战争是中美之间第一次真正的“交手”,此前缺乏深入了解或声誉基础,即使中国有建立自身声誉的动机,也不足以通过发出“弱”信号使美国相信来自“陌生”的新中国的声音。

### (三) 古巴导弹危机中的美苏策略互动

1962年10月14日,美国一架U-2高空侦察机第一次拍摄到了苏联在古巴设置地对地中程导弹的照片。根据美国中央情报局的估计,在古巴的导弹基地尚需10天才可投入使用。美国总统肯尼迪很快成立了专门小组研究并讨论对策,10月19日他下定决心要用海军封锁古巴。10月22日晚7时整,肯尼迪向全国发表电视演说,宣布美国将采取包括对一切正前往古巴的进攻性军事装备实行海上“隔离”的七项措施。根据肯尼迪的命令,40艘军舰和2万名海军战斗人员将用于对古巴的封锁,美国在全球的军队立即进入紧急戒备状态,一场震惊世界的古巴导弹危机拉开了帷幕。在美国强大的军事压力下,苏联最终以美国承诺不入侵古巴并答应撤除在土耳其针对苏联的导弹作为交换条件,从古巴撤出弹道导弹。

美国在古巴导弹危机中的行动,是一次典型的威逼行为,其结局是一种典

---

杨存堂:《美苏冷战的一次极限:加勒比海导弹危机》,桂林,广西师范大学出版社2002年版,第22页。

参见罗伯特·肯尼迪:《十三天:古巴导弹危机回忆录》(复旦大学历史系拉丁美洲研究室译),上海人民出版社1977年版,第83—95页。

崔金久等:《20世纪重大军事危机述评》,北京,军事科学出版社1995年版,第121页。

戈登·克雷格、亚历山大·乔治:《武力与治国方略:我们时代的外交问题》(时殷弘、周桂银等译),北京,商务印书馆2004年版,第162页。

型的分离均衡。如前所述,威逼互动中挑战者的挑战,可以看做是一种“试探”行为:如果威慑者发出“强”信号,则“后退”,否则就“坚持”。由于威逼涉及挑战者的声誉,它比威慑行为更难成功;但美国在古巴导弹危机中通过“核战争边缘”方式发出“强”威慑信号,并灵活采取“胡萝卜加大棒”策略,最终得以成功。如果不是采取“强”信号而是“弱”信号,如外交谈判、单纯的利益交换,甚至是单纯的海上封锁而不是把局势推向核战争边缘,美国的威逼行动都将难以奏效。

苏联在古巴部署导弹的行动是在未察觉的情况下秘密开始的,赫鲁晓夫的设想是:“如果我们秘密地装置导弹,如果导弹已经装好可供发射之后才为美国发现,美国人在试图用军事手段摧毁我们的设备之前就得仔细考虑考虑。”他要求下属:“我们必须采取一切预防措施,悄悄地运送和部署导弹,以便给美国造成一个既成事实。”在发现并确认了苏联的行动后,美国选择了利用海军力量实施封锁的方案。肯尼迪认为这是除军事袭击(以后如有必要,可以采取)以外的最不妥协的行动,它既能为对方所理解又给他们留有回旋余地,而对美国来说可以在自己控制下逐步升级,也避免了对古巴直接实施空中打击或入侵古巴可能导致的同苏联发生一场核大战的风险。从信号博弈角度看,在转折概率  $p^*$  的计算式中,比较得益  $u_c$ 、声誉损失  $s$  与核战争代价  $c_c$ ,由于后者数值巨大,因此转折概率  $p^*$  值较低。 $p^*$  值较低,为美国运用“核战争边缘”作为“强”威慑信号而不致局势失控提供了可能:只要美国最大限度地表现出威慑决心,迫使苏联从古巴撤出导弹是可能的。到 10 月 20 日,美国已命令在全世界的武装力量都处于戒备状态,四个空军中队随时处于空中打击状态,部分空军开进佛罗里达州的东南部,导弹人员处于高度的戒备之中,B-52 轰炸机队

---

赫鲁晓夫:《赫鲁晓夫回忆录(下)》(赵绍棣等译),北京,中国广播电视出版社 1988 年版,第 697—699 页。

阿纳托利·多勃雷宁:《信赖:多勃雷宁回忆录》(肖敏等译),北京,世界知识出版社 1997 年版,第 81 页。

杨存堂:《美苏冷战的一次极限》,第 37 页。

与美国相比,苏联危机时在核力量上处于劣势,美国拥有 5 000 个核弹头,而苏联仅有 300 个,运载工具也未达到相当的程度。同上书,第 13 页。

肯尼迪当时估计爆发核战争的危险在三分之一至二分之一之间,这可以看做是对转折概率值的一种粗略估算。参见 Avinash Dixit and Susan Skeath, *Games of Strategy* (New York: W. W. Norton & Company Ltd, Second Edition, 2004), p. 450.

奉命携带原子武器进入空中,一架着陆,立即由另一架飞上去顶替。所有这些行动,都构造了核战争一触即发的威逼态势,传递了最强烈的威慑信号。“核战争边缘 通过制造“带风险的威胁”(threats of chance)或“受控的失控”,使赫鲁晓夫产生了“一旦交火必然引起一系列连锁反应,甚至可能引发一场全世界的核大战”的担忧,最终决定从古巴撤出导弹。

为了降低苏联采取“后退”策略的声誉损失,美国在不断表现出咄咄逼人“刚性”一面的同时,也灵活采取了“柔性”的策略,承诺不入侵古巴,并答应撤除在土耳其针对苏联的导弹。为了“使对手让步,但不遭受羞辱”,美国国务院的法律顾问建议将对古巴的封锁称为“防御性隔离”,这个新词可以表明美国采取的行动在政治上是防御性的。他们还建议,由美洲国家组织援引 1947年《里约热内卢相互协助条约》授权进行对古巴的“隔离”行动,而不仅是援引《联合国宪章》第 51 款规定的自卫权。因为要获得像美洲国家组织这样一个地区机构在《联合国宪章》第八条下的授权是非常困难的,其他国家今后难以援引这件事作为先例。美国由于选择了正确的信息传递方式,灵活采取了“胡萝卜加大棒”策略,最终迫使苏联从古巴撤出了导弹,威逼得以成功。

#### (四) 马岛危机中的英阿策略互动

1982年 4月 2日爆发的英国与阿根廷之间的马岛危机,最终演化为自第二次世界大战结束以来南大西洋首次爆发的一场规模较大的、历时 74天的海上冲突。英阿马岛冲突的焦点,在于争夺马尔维纳斯群岛、南乔治亚群岛、南桑德韦奇群岛及其附近海域的主权。危机伊始,阿根廷集中兵力于 4月 2日在马岛首府斯坦利港登陆,3日在南乔治亚岛首府格里特维肯港登陆。针对突发事态,英国做出了迅速反应并成立了战时内阁,翌日决定动员并组建、派遣特混舰队。4月 5日首批舰船出航,接着空军 1 个大队向南大西洋阿森松岛转场,12

---

参见杨存堂:《美苏冷战的一次极限》,第 33—34页。

赫鲁晓夫在与肯尼迪的来往信件中这样告诫下属和相互告诫,同上书,第 109页。

夏立平:《美国关于危机管理的理论与实践:以中美关系为例》,第 79页。

同上。

日核动力潜艇进入战区开始海上封锁,15日飞机在战区开始空中巡逻。此后,英军逐步扩大海空封锁,阿根廷进行反封锁,双方封锁和反封锁较量十分激烈,最终英军于6月14日重新夺回马岛。

英国在马岛危机中的行动,与美国在古巴导弹危机中的行动一样,也是一种威逼行为,并且也选择了出动特混舰队并实施海空封锁这样的威慑形式,其目的就是以武力威胁方式逼迫阿军撤出马岛。当时英国战时内阁中几乎所有人都认为,凭借这支部队的威慑力量足以令阿根廷放弃武力抗衡的想法并坐到谈判桌前来,“他们从未想过它会真的参战”。然而,透过英国的海空封锁,阿根廷并未读出英国的威慑决心,而是“把英特混舰队缓慢的航行看成是单纯的示威行动”。从占领马岛伊始,阿方就认为英国武力夺回马岛的可能性很低,当时的阿总统加尔铁里事后曾说:“尽管认为英国作出某种反应是一种可能性,我们不相信大不列颠为马尔维纳斯群岛进行动员。”“在阿根廷看来,战后,英国经济萧条,国内反对生产核武器和要求裁军的呼声很高。而且,英国废弃了在这种远距离作战中必不可少的攻击型航空母舰,就连刚刚竣工的轻型航空母舰也要出售给外国。因而未想到英国会断然决定,从北纬50度南下6000海里,远涉重洋到南纬50度实施登陆作战。”且在此之前,英阿之间曾探讨过以“主权换治权”的可能性。此外,阿根廷对英军特混舰队的实施威逼行动的军事实力也是怀疑的。由战舰、补给舰等100余艘舰船组成的特混舰队,看起来确是一支庞大的舰队,但英海军力量主要是用于反潜作战的力量,而现在不得不把它转用于登陆作战。“北约赋予英国海军的主要任务是反潜,其武器装备是为执行这一任务而配备的。皇家海军因而不配备攻击型航空母舰,其他

---

吴艾军:《英阿马岛战争概况》,载于军事科学院外国军事研究部编:《外国对英阿马岛战争经验教训的评论》,北京,中国对外翻译公司1983年版,第2页。

雨果·杨格:《铁女人撒切尔夫人传》(汤玉明等译),西安,西北大学出版社1990年版,第259页。

玛格丽特·撒切尔:《撒切尔夫人回忆录:唐宁街岁月》,呼和浩特,远方出版社1997年版,第112页。

亨特:《战争的经验教训》,载于军事科学院外国军事研究部编:《外国对英阿马岛战争经验教训的评论》,第41页。

参见《加尔铁里谈马岛战争:答记者法拉奇问》,载于军事科学院外国军事研究部编:《外国对英阿马岛战争经验教训的评论》,第64页。

筑土龙男:《英军登陆成功的主要原因》,同上书,第228页。

舰艇也不再装备适合于轰击海岸目标的大口径火炮及火箭发射装置。这就使再次占领马岛的任务,至少可以算是一种冒险行动。”相比之下,阿军以逸待劳,离本土近,最近处仅 500 公里,有岸上基地作依托,人员和飞机在数量上也占有较大的优势。

从战争代价  $c_c$  与声誉损失  $k_s$  比较看,  $c_c$  比  $k_s$  小得多,因此同样的得益  $u_c$  下转折概率  $p^*$  值增大。对阿根廷来说,马岛危机即使转化为马岛战争,也只是局限在马岛周围的局部战争,“如果英国要入侵阿根廷本土,那是一种不切实际的幻想,因为拥有 130 000 人的阿根廷陆军能够在局部地区形成绝对的优势”。而在占领马岛前,马岛本来就在英国的治理之下,阿根廷并未有多大实际损失。相比而言,阿根廷的声誉要重要得多。阿军攻占马岛的消息传到阿国内,首都布宜诺斯艾利斯群情激昂,数十万名群众聚集在总统府所在的“五月广场”上庆祝胜利;全国十几个政党顿释前嫌,一致表示支持以总统兼陆军总司令加尔铁里为首的军政府的行动。外长门德斯神采飞扬地说:“今天是我一生中最幸福的一天。”总统本人更是激动,表示阿根廷决不屈服于武力威胁,“民族的骄傲和尊严,必须不惜任何代价地去恢复”。由于英阿双方都高度重视声誉,因此两国都很难在国际调解中做出实质的让步。最终,英国为了在军事上占据主动,在海上禁区之外击沉了阿根廷的“贝尔格拉诺将军”号巡洋舰,战争正式开始,威逼行动失败。

比较马岛危机与古巴导弹危机这两次威逼行为,虽然都运用了海军力量实施海上封锁的形式,但在许多方面还是有很大不同,最终美国获得了成功,而英国的威逼行为失败。首先,两次危机中挑战者对威慑者的先验信念不一样。在阿根廷看来,英国缺乏相应的威慑实力,其海上力量不适于登岛作战;而在苏联看来,美国在核力量上占有优势,在古巴秘密部署弹道导弹本身是苏联试图取得军事平衡努力的一部分。其次,可信威慑的转折概率也不相同。在古巴导弹危机中,核战争的巨大毁伤后果使苏联将声誉损失考虑退居次要地位,成功威慑的转折概率降低,只要表现出足够的威慑决心,威逼行为就能够取得成功;而

胡安·卡洛斯·穆尔吉苏尔博士:《论马岛之战》,同上书,第 69 页。

艾德里安·英格利希:《阿根廷的军事潜力》,同上书,第 54 页。

沈根林、徐红兵:《从马岛之战到海湾战争》,上海人民出版社 1994 年版,第 6 页。

在马岛危机的得失权衡中,局部战争只局限于几个岛屿及附近的海域,而且英国取胜的概率并不高,因此声誉比常规冲突的实际代价更重要,成功威慑的转折概率提高。由于英国属于“强硬”类型的先验概率不高,而成功威慑转折概率较高,使得英国选择能使后验信念高于转折概率的“强”信号的难度大大增加。同样是军事封锁形式,在古巴导弹危机中能够成为可信威慑的“强”信号,而在马岛危机中仍然是“弱”信号,使得美国在古巴导弹危机中威逼成功,英国在马岛危机中威逼失败。

### 四、结 论

本文通过对“成本—收益”分析、理性选择分析和博弈论分析三种威慑逻辑分析框架的比较,认为前两种分析框架由于忽视了威慑中博弈方之间的策略互动,在研究可信威慑时不可避免地存在局限性。根据博弈论分析框架,本文得出以下结论:第一,挑战者与威慑者之间的策略互动存在多种互动格局,威慑格局只是这些互动格局中的一种。威慑格局成立的条件是,挑战者的偏好为  $CW_N > SQ > WAR$ ,威慑者的偏好为  $CW_N < SQ$  且  $WAR < SQ$ 。第二,在完全信息条件下可信威慑的充分必要条件是:(1)挑战者的偏好为  $CW_N > SQ > WAR$ ; (2)威慑者的偏好为  $SQ > WAR > CW_N$ 。威慑实力只是可信威慑的必要条件,只有当威慑者既有实力又有决心时,威慑才是可信的。第三,在完全信息条件下威慑者的承诺行动中,只有昂贵信号才起作用。

本文以威慑实力与决心作为两个维度,将威慑者划分为4种类型,其中只有既有实力又有决心的威慑者才是可信威慑者。在此基础上,根据“海萨尼转换”的研究思路,将不完全信息条件下的威慑者划分为“强硬”和“软弱”两种类型,并计算了威慑成功的转折概率:只有当挑战者对威慑者为“强硬”类型的先验信念大于转折概率时,威慑者的威慑才是可信的,否则是不可信的。针对不完全信息条件下的动态威慑博弈,本文构造了基于信号博弈的威慑模型,求解了模型的混同均衡和分离均衡,分析了这些均衡的内在含义,指出在不完全信息威慑情景下的信息传递并不总是有效,它只在分离信号等某些情形下起作用。

通过案例分析,本文得出以下观点:(1)伊拉克入侵科威特并不是美国的一次“延伸威慑”失败,因为美伊互动关系并不构成威慑格局,美国的模糊信号传递实质上不是威慑而是一种“诱导”。(2)即使信息传递渠道畅通、信息传递者立场中性、信息表述十分明确,中国对美国越过“三八线”的威慑也会因为无法选择“强信号”以有效改变美国对中国“类型”的先验信念而使得威慑很难成功。(3)由于涉及声誉损失,威逼行为更难以成功。在古巴导弹危机中,声誉损失与核战争的灾难后果不成比例,加之美国灵活采取了“胡萝卜加大棒”策略,使威逼得以成功;在马岛危机中,对阿根廷而言国家声誉的价值远高于局部冲突的代价,而英国未能给阿根廷以适当的“补偿”,致使威逼失败。

## 作者简介

**向钢华** 清华大学经济管理学院博士候选人。1985年获浙江大学工学学士学位,1990年获北京理工大学工学硕士学位。主要研究方向为军事威慑、军事决策与博弈分析。

电子邮箱: xianggh\_03@em.tsinghua.edu.cn

**王永县** 清华大学经济管理学院教授。1966年在清华大学电机工程系获工业企业电气化学士学位。研究方向为决策理论与管理决策优化。著有:《国外的国家经济安全研究与战略》(2000年)、《运筹学》(1993年)。

电子邮箱: wangyx@em.tsinghua.edu.cn

**王莉欢** 1986年毕业于新加坡国立大学,获社会科学学士学位;2006年毕业于清华大学,获国际关系学硕士学位。现在新加坡政府任职。

电子邮件: laifoonwong@pacific.net.sg

**阎梁** 清华大学国际问题研究所2005级博士研究生。1995和2001年分别在西北大学和兰州大学获经济学学士和硕士学位。2002—2005年在中国现代国际关系研究院世界经济所任助理研究员。著有:《社会危机事件处理的理论与实践》(2003年)。

电子信箱: yanliang05@mails.tsinghua.edu.cn

**潘成鑫** 澳大利亚迪肯大学国际关系与政治学院讲师。在北京大学国际政治系获法学学士和硕士学位,2004年在澳大利亚国立大学获哲学博士学位。1996—1997年为墨尔本大学政治学系访问学者。研究兴趣为中国外交、中美关系、东亚安全。论文曾发表在 *Alternatives, Political Science* 等学术刊物上。专著 *Understand China's International Relations: Western Representations of an Emerging Superpower*, 将由英国 Edward Elgar Publishers 出版。

电子信箱: chengxin.pan@deakin.edu.au

**李巍** 复旦大学国际关系与公共事务学院2006级博士研究生。2003年在华中科技大学新闻学院获文学学士学位,2006年在北京大学国际关系学院获法学硕士学位。研究兴趣为国际关系理论、国际政治经济学、美国政治与对外政策。

电子信箱: kindyleeway@gmail.com

**王勇** 北京大学国际关系学院副教授、国际政治经济研究中心主任。1987、1990和1996年先后在北京大学国际政治系获法学学士、硕士和博士学位。研究兴趣为国际政治经济学、国际贸易政治、中美经贸关系。著有:《最惠国待遇的回合:1989年至1997年美国对华贸易政策》(1998年)、《知识经济对策:运作与案例》(1998年)。

电子信箱: yowang@pku.edu.cn