第 16 卷第 1 期
2013 年 1 月

管 理 科 学 学 报
JOURNAL OF MANAGEMENT SCIENCES IN CHINA

Vol. 16 No. 1
Jan. 2013

从 FIPD 的合作现象看潜在参与人对 Nash 均衡的影响①

杨　城[1]，朱超威[2,3]，吕峻闽[1]
（1. 西南财经大学经济信息工程学院，成都 611130；2. 宜春学院经管学院，宜春 336000；
3. 中国人民大学商学院，北京 100872）

摘要： 针对大量现实博弈与 Nash 均衡预测相背离的现象，抛开“共同理性”假设，提出了 Nash 均衡理论的另一个隐性前提“博弈的封闭性假设”，进而引入“潜在参与人”的概念，扩展了博弈参与人的范围，从理论上指出开放环境下策略偏离者的支付增益可能导致博弈结果偏离原 Nash 均衡；然后举例说明，将随机扰动引入标准复制子动态，对一个两阶段重复囚徒博弈偏离预期均衡的原因进行剖析；最后基于 CAS 理论和多主体系统的建模思想，应用仿真实验进一步分析和验证了以上结论.

关键词： 纳什均衡；潜在参与人；进化博弈；有限次重复囚徒博弈；多主体系统

中图分类号： F224.32　　**文献标识码：** A　　**文章编号：** 1007－9807（2013）01－0087－08

0　引　言

纳什均衡理论的提出是 20 世纪最重大的科学进展之一，它对经济学和社会科学的影响尤为巨大，有着重要的实际意义，Nash 本人也于 1994 年获得诺贝尔经济学奖. Nash 均衡又称为非合作对策均衡，它是博弈论中最基本、最核心的概念. 作为理性推导的结果，Nash 均衡是建立在一个基本的假设前提之上的——博弈参与人是“共同理性”的：参与人都是完全理性的经济人，并且所有参与人对于其他参与人的理性和博弈结构具有共识. 然而，在现实生活和博弈实验中，经常出现参与人理性、能力的非对称性以及参与人行为与 Nash 均衡解的预测相背离的现象，尤其在一些类似“蜈蚣博弈”、“重复囚徒博弈”的悖论问题中，现实均衡与理论分析相去甚远. 有鉴于此，国内外已经有大量的学者对纳什均衡理论提出了批判和修正，他们强调“共同理性”假设过于理想和严格，与现实差异较大，应该区分博弈参与人的异质性和理性程度的有限性[1-5]. 本文不准备对“共同理性”假设进行深入分析，而是试图探讨 Nash 均衡的另一个潜在前提——“博弈的封闭性假设”，研究该隐性假设的不合理性和修正思路，并通过实例和仿真来分析验证.

1　开放环境下的 Nash 均衡

纳什均衡描述了若干追求自身利益最大化的博弈参与人的策略间的一种稳定状态，其规范定义如下，其中 S_i 和 u_i 分别表示参与人 i 的策略空间和支付函数[6].

在 n 人博弈 $G=\{S_1,\cdots,S_n;\ u_1,\cdots,u_n\}$ 中，策略组合 $s^*=(s_1^*,\cdots,s_n^*)$ 是一个纳什均衡，如果对每一个 i，s_i^* 是给定其他参与人选择 $s_{-i}^*=(s_1^*,\cdots,s_{i-1}^*,s_{i+1}^*,\cdots,s_n^*)$ 的情况下第 i 个参与人的最优策略，即

$$u_i(s_i^*,s_{-i}^*)\geqslant u_i(s_i,s_{-i}^*)\quad(\forall s_i\in S_i,\forall i)\tag{1}$$

① 收稿日期：2011－02－10；修订日期：2011－12－26.
基金项目：国家社会科学基金资助项目（11AZD077）；西南财经大学创新团队建设项目.
作者简介：杨　城（1977—），男，重庆人，博士，讲师. Email：mr. yangcheng@163.com

它表示一个稳定的策略组合，在其他人策略既定的情况下，任何一个参与人如果单独改变其策略都不可能使其收益增加，即没有人有足够的积极性去打破这种均衡状态. 因此，这样的结果是具有自我实施性和可预测性的. 在双人博弈中，Nash 均衡还可以通俗的表述为“给定你的策略，我的策略是我最好的策略; 给定我的策略，你的策略也是你最好的策略”.

显然，在这个定义中，基于“共同理性”的推理对象仅限于与支付函数直接相关的参与人. 也就是说，在具体分析博弈均衡时，博弈模型是暂时与外界环境相隔离的(或者说环境因素已经体现在支付函数中) ，此时局中人只包含该博弈的若干直接参与人，策略分析、收益比较、胜负判断也仅仅局限于直接参与人之间. 因此，抛开“完全理性”和“公共知识”，Nash 均衡还蕴含着一个潜在的前提: 博弈是在一个封闭环境中进行的.

一个博弈最基本的要素包括参与人、策略和支付三方面. 其中，参与人指的是博弈中的决策主体，是以最大化自身收益(效用) 为目的，能够独立决策、独立承担结果的个人或组织. 所谓“直接参与人”就是直接参与当前博弈的决策主体，他们的策略组合直接作用于支付函数，共同决定着各自的最终收益.

通常，一个博弈的直接参与人不会太多，有效的策略空间也不会太大，故问题的总体复杂度一般都在人们的分析范围内(“象棋博弈”之类的极端情况除外) . 因此，在封闭环境中，Nash 均衡将博弈近似为“共同理性”的决策行为是可行的，此时“犯错误”的个体几乎不存在，博弈结果很难偏离预期均衡.

但是，现实博弈中的个体都是社会的人或组织，他们不是孤立存在的，而是生活在各自的小生境(Niche) 中的. 在同一个小生境内，长期进行着大量相同或相似的博弈，个体之间因此存在或多或少的、直接或间接的联系: 他们既可能因为直接交互成为某个博弈的直接参与人，也可能因为彼此收益的差距而产生间接的、潜在的影响. 因此，将博弈问题局限在一个直接交互的、封闭的环境中进行讨论是不恰当的. 本文把那些当前博弈之外的，隶属于同一个经济环境的，进行着相同或相似博弈的决策主体称为博弈的“潜在参与人”. 他们同样是博弈参与人的一部分，是决策分析的重要影响因素.

需说明的是，“潜在参与人”不同于部分博弈分析中引入的“虚拟参与人”(pseudo-player) ——“自然”[2]. “自然”是指决定外生随机变量的概率分布的机制，它没有自己的支付和目标函数(即所有博弈结果对它而言是无差异的) ，如寡头博弈中的市场需求; 而“潜在参与人”是一个在本质上与直接参与人相同的、具备自我意识、以收益最大化为目标的决策主体. 在开放环境中，一个博弈的潜在参与人往往有很多，数量远远超过直接参与人——后者仅仅是由参与人全集(包括直接和间接参与人) 随机抽取产生的、隶属于特定博弈单元的、一个参与人的子集. 此时，直接参与人的策略决定其所在单元的绝对收益，而所有潜在参与人的策略则决定彼此在整个博弈系统中的相对收益，他们共同影响着博弈的最终均衡状态.

在这种考虑大量潜在参与人的开放模型中，虽然个体“犯错误”的概率依旧很小，但由于总量很大，难免会有少量参与人的决策行为因为自身迁移、理性误差或者彼此缺乏信任等原因偏离一致预期，这一点类似于 Selten 的“颤抖的手”(trembling-hand) 在群体博弈中的体现. 大多数情况下，由于他们的整体比例太小，不会对原博弈的均衡产生影响. 但是在个别情况下，由于具体支付函数的特点，偏离者的收益有可能高于未偏离者的收益. 因为开放模型中的偏离者可能同时存在多个，他们可能单独分散在不同的博弈单元中(情况 1) ，也可能与别的偏离者共存于同一博弈单元中(情况 2) . 当博弈单元中只有一位偏离者时，如 Nash 均衡所描述的那样，偏离者不可能获得额外收益; 而当博弈单元中不止一位偏离者时，偏离者的收益就存在增加的可能性. 如果情况 2 中期望收益的增加值超过了情况 1 中期望收益的减少值，即当偏离者之间博弈的支付增益足以抵消他们在整体人数上的劣势时，偏离者的策略将更具有优势. 由于所有的参与人都是足够理性的，因此如果这种偏离优势存在，则它是能够被分析和预期的，而这样的利益驱动将吸引更多的参与人“故意”偏离先前的一致预期，进一步扩大偏

离者的支付增益,从而加剧博弈结果的偏离,最终出现与原 Nash 均衡不一致的均衡状态.

这类由于偏离者的支付增益导致博弈结果偏离理论均衡的现象,尤其集中在部分使用劣策略反复消去法的静态博弈问题和使用逆推归纳法的多阶段博弈问题中. 因为这类问题的推导分析过程是环环相扣的,任何一环出现微小的偏离都可能导致最终结果远离理想状态. 例如,在劣策略反复消去法中,依次消去的策略之间确实存在相对的优劣关系,但是这种优劣关系并不具备传递性. 如果原 Nash 均衡系统中新产生的偏离策略不全是最后一轮消去的劣策略,则原有的均衡策略未必就一定优于这些偏离策略,后者将可能成功入侵原系统.

2　一个重复囚徒博弈的例子

在这一节中,将借用一个两阶段重复囚徒博弈(2-IPD) 的例子,来具体剖析开放环境中的潜在参与人是如何影响 Nash 均衡的.

例如,表 1 是一个标准的囚徒博弈的支付矩阵,C 代表合作,D 代表背叛. 在这个博弈中,无论外界条件如何,即无论在封闭环境还是在开放环境,无论其他的潜在参与者选择策略 C/D 的比例怎样,策略 D 都是上策. 因此,策略组合(D,D) 是囚徒博弈的一个上策均衡(dominant-strategy equilibrium) ,也是唯一的 Nash 均衡,并且它是一个稳定的均衡解.

表 1　一个标准囚徒博弈的支付矩阵 $\boldsymbol{M}$

Table 1 Payoff matrix $\boldsymbol{M}$ of the classic prisoner's dilemma

	2nd Player: D	2nd Player: C
1st Player: D	10, 10	50, 9
1st Player: C	9, 50	49, 49

如果把上面这个博弈重复两次,就变成一个两阶段的重复囚徒博弈. 关于有限次重复囚徒博弈问题,已经有大量的文献进行研究,其中参与人的行为策略通常基于一种"类触发策略"(quasi trigger strategies,QTS) : 参与人在博弈刚开始的前 i 个阶段主动选择 C,一旦对方选择 D 或者重复阶段数大于 i 时,则在余下的阶段博弈中永远选择 D[7-9]. 其中,i 称为 QTS 策略的合作截止点,表示参与人的最大可能合作次数.②具体在本例的 2-IPD 中,每一位参与人有三种策略 $s[0]$、$s[1]$、$s[2]$,分别代表"永远背叛"(all defect,AllD) 、仅首次合作和针锋相对(tit-for-tat,TFT) ,其支付矩阵的扩展式如表 2 所示.

表 2　一个两阶段的重复囚徒博弈的扩展支付矩阵 $\boldsymbol{M}(2)$,其原博弈矩阵为 $\boldsymbol{M}$

Table 2 Extended payoff matrix $\boldsymbol{M}(2)$ of 2-IPD based on the original payoff matrix $\boldsymbol{M}$

$1^{st}\backslash 2^{nd}$	$s[0]$	$s[1]$	$s[2]$
$s[0]$	20, 20	60, 19	60, 19
$s[1]$	19, 60	59, 59	99, 58
$s[2]$	19, 60	58, 99	98, 98

经典博弈理论可知,当原博弈 G 有唯一的纯策略 Nash 均衡时,有限次重复博弈 G(T) 有唯一的子博弈完美 Nash 均衡,即各博弈方每个阶段都采用 G 的 Nash 均衡策略. 也就是说,2-IPD 只是标准囚徒博弈的简单重复,依然不可能实现合作,也不会改变原博弈的低效率均衡. 因此,策略组合($s[0]$, $s[0]$) 应该是唯一的均衡解,其分析过程一般基于"最后阶段博弈效应"(end-game effect,即随着最后阶段的临近,博弈双方进一步合作的潜在利益愈来愈小,停止合作的可能性会越来越大,故将合作持续到最后一刻是不理性的) ,运用"逆推归纳法"进行推导,本例也可以对矩阵 $\boldsymbol{M}(2)$ 运用"劣策略

② QTS 策略实际上是多种常见策略的一个集合. 基于 QTS 的策略思想,G(T) 中每个参与人的策略空间包含 $T+1$ 个纯策略,分别表示为 $s[i]$,$i=0,1,\cdots,T$. 其中,$s[0]$ 表示自始至终选择背叛,等价于 AllD 策略; $s[T]$ 表示始终以合作对待合作,以永久性的背叛反击任何背叛,等价于 TS 触发策略. 相对于重复囚徒博弈中的其它几种常见策略,如 TFT、GTFT(与 TFT 类似,但同时以小概率的合作回应背叛) 、Pavlov(赢则保持,输则变换) 等,QTS 通常更为有效. 例如,TFT 等策略和 QTS 虽然前期都能形成高收益的合作局面,但前者只能被动的反击背叛行为,而 QTS 却能够基于"最后阶段博弈效应",策略性的主动选择背叛,从而取得优势. TFT 等策略希望通过反击促使对手重新选择合作,甚至在对方背叛的情况下仍然以小概率尝试合作,以避免意外背叛所导致的不合作(如 GTFT) ; 但 QTS 通常只在最后某个阶段主动背叛,此时合作的潜在利益已经很小,后续的阶段博弈已经很难形成有效的反击空间. 尤其当 T 值较小时,仅有的一两次占优的阶段博弈,将使得 QTS 的策略优势显得尤为明显.

反复消去法”得到同样的结论. 但是，这样的结论往往与人们的直觉经验和事实不符，尤其是当重复次数较多的时候.

由于经典博弈理论无法解释这一悖论，考虑从进化博弈的观点来重新分析. 进化博弈理论从有限理性的个体出发，以群体为研究对象，认为现实中的个体并非天生的行为最优者，个体的决策是一个向优势策略渐变的过程，是通过相互之间的模仿、学习和突变等动态过程来实现的. 显然，这种理论是建立在一个开放的博弈环境之上的，博弈群体就是一个包含大量潜在参与人在内的参与人全集[10-13].

复制子动态(replicator dynamics) 是进化博弈的一种主要分析机制. 公式(2) 是 2-IPD 博弈的复制子动态方程组，其中 $x[i]$ 表示使用策略 $s[i]$ 的人数占博弈总人数的比例，$u[i]$ 表示 $s[i]$ 的期望收益，$\bar{u}$ 表示所有策略的平均期望收益.

$$\frac{\mathrm{d}x[i]}{\mathrm{d}t} = x[i] \times (u[i] - \bar{u}), \quad i = 0,1,2 \qquad (2)$$

图 1 展示了基于矩阵 $M(2)$ 的标准复制子动态的微分方程组的向量场(运用 Maple 软件) . 其中，三角形为策略单纯形在($x[0]$，$x[1]$) 平面上的投影，它的每个点代表一种策略组合，例如顶点 $s[0]$ 表示所有参与人都使用 AllD 策略，$s[0]$ 与 $s[2]$ 的中点表示使用策略 AllD 和 TFT 的参与人各占一半; 红色的小箭头表示该点处对应的策略组合的演化方向，蓝色的实线表示部分解轨道(即部分初始策略组合的演化路径) . 如图所示，无论初始状态如何，经过足够长时间的演化，系统都将收敛到单一同态的纯策略 $s[0]$; 并且，其动态演化过程与劣策略反复消去法的推导过程是相似的: 群体中的策略 $s[2]$ 逐渐过渡到 $s[1]$，并最终全部收敛于 $s[0]$. 但另一方面，图 1 的向量场也显示 $s[0]$ 的吸收域非常狭小，系统在该点处并不稳定，很小的扰动就会导致系统远离稳态. 也就是说，只要有少数几个参与人偏离均衡策略，系统就将按照先前的过程重新演化一遍. 由于现实环境中，各种随机扰动总会存在，系统常常受到来自突变和其它偶然事件的冲击. 因此，本例的 2-IPD 博弈不可能稳定在 $s[0]$ 点处，策略 $s[0]$ 不是演化稳定的.

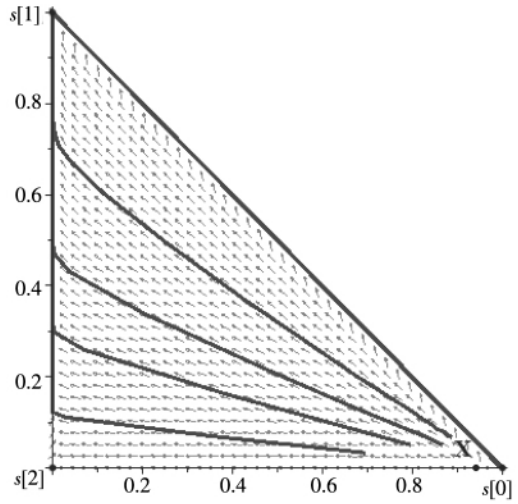

图 1　基于标准复制子动态的微分方程组的向量场，其中支付矩阵为 $M(2)$，3 个顶点分别代表 3 种纯策略，$s[0]$ 是吸引点

Fig. 1 Vector fields of standard replicator equations based on the payoff matrix $\boldsymbol{M}(2)$, where the vertices represent three pure strategies, and $s[0]$ the attractive point, respectively

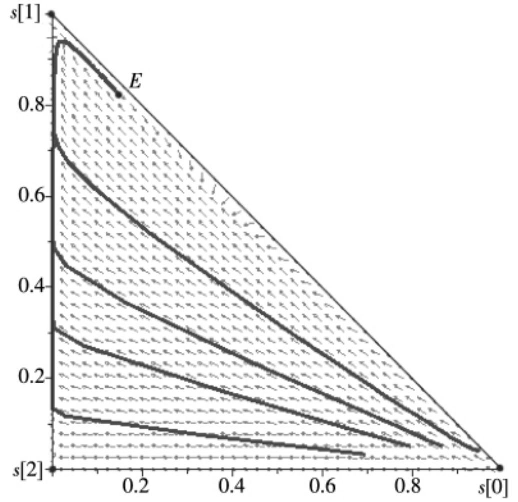

图 2　引入随机扰动后的复制子动态微分方程组的向量场，其中支付矩阵为 $M(2)$，3 个顶点分别代表 3 种纯策略，E 为吸引点，扰动率 $p = 5\%$

Fig. 2 Vector fields of replicator equations after introduction of stochastic disturbance, where the payoff matrix is $\boldsymbol{M}(2)$, the vertices represent three pure strategies, respectively. E is the attractive point, and the disturbance rate $p = 5\%$

以上是针对博弈群体的分析，对于博弈个体而言，假设原系统受某种随机扰动的影响发生轻微偏离至点 X，此时 3 种策略的比例($x[0]$，$x[1]$，$x[2]$) 分别为(95%，0%，5%) ，对应的收

益($u[0]$，$u[1]$，$u[2]$) 为(22，23，22.95)，$s[1]$ 和 $s[2]$ 的收益相对较高. 理性的参与人由于预期到博弈群体中存在类似的偏离以及对应的收益差，因此他们有积极性选择策略 $s[1]$ 或 $s[2]$，从而进一步加速系统偏离原均衡点 $s[0]$.

由于标准复制子动态强调的是选择机制，只能够在原有策略集中进行优胜劣汰式的竞争，但不包含任何形式的变异机制，因此它无法体现图 1 中吸引点处的随机扰动. 由此，考虑在标准复制子动态中引入非确定性的随机扰动，用更为现实的随机动态模型(stochastic dynamical model)[14-16] 来研究 2-IPD 中的策略演化过程.

假定每一轮博弈结束后，都会有很小一部分的参与人随机地变换其策略，每种策略变异个体的数量与使用该策略的参与人比例成正比，即系统存在一个常数化的随机扰动率 p. 当原博弈体系引入随机扰动后，其直接效应导致当期的策略构成比中 $s[i]$ 的成分有 $p \cdot x[i]$ 移出，同时有 $p \cdot (1 - x[i])/2$ 的成分由其它两种策略移入. 于是，完整的融入随机扰动因素后的复制子动态微分方程如下

$$\frac{\mathrm{d}x[i]}{\mathrm{d}t} = x[i] \cdot (u[i] - \bar{u}) + p \cdot \left(\frac{1 - x[i]}{2} - x[i]\right)$$

$$i = 0, 1, 2 \qquad (3)$$

图 2 显示了矩阵 $M(2)$ 对应的随机动态模型的演化过程，其中扰动率 $p = 5\%$. 对比两张图的向量场，它们的演化轨迹大体上是类似的，但前者收敛于策略单纯形的一个顶点 $s[0]$，后者收敛于单纯形的一个内点 E. 前者的均衡是脆弱的，而后者的均衡是演化稳定的. 图 2 中吸引点 E 的位置表明，该 2-IPD 博弈均衡于混合策略，其中主体策略为 $s[1]$，即绝大多数参与人第一次选择 C 而第二次选择 D. 值得一提的是，E 点处的均衡并非是静止不动的，系统仍然会在 E 点附件的区域波动，它是一种动态的、基于时间平均的均衡. 此时，3 种策略的收益值虽然并不相等，但是不同期望收益产生的复制效应与随机扰动产生的变异效应在 E 点处正好相互抵消，使得各成分策略的比例趋于平稳. 由此可见，有限次重复囚徒博弈并非一定均衡于 AllD 策略，当开放环境中的博弈个体考虑潜在参与人的影响时，系统也可能动态均衡于一个非 AllD 的状态，从而实现部分合作.

最后，关于本例还有两点说明: (1) 在原博弈矩阵 M 中，不同策略间的收益差较大，这主要是为了加速合作可能，便于在二维空间作图分析. 对于任何其它形式的囚徒博弈，只要重复博弈的次数足够大，同样会促使参与人的策略偏离 AllD，涌现合作. 与之类似的例子还有“蜈蚣悖论”，“连锁店悖论”等. (2) 在扩展支付矩阵 $M(2)$ 中，策略 $s[2]$ 相对于 $s[1]$ 是弱劣策略，但即使修改为严格劣策略(比如把表 2 中($s[1]$，$s[0]$)的支付改为(19.5，60))，其效果也是类似的，不影响均衡的最终状态.

3　仿真实验分析

为了作进一步分析验证，运用一个类似 SWARM 的软件平台，按照复杂适应系统(complex adaptive system，CAS)的思想建立多主体系统的仿真模型，将大量的具有有限理性的参与人设计为适应性主体(adaptive agent)，通过随机配对式的反复博弈来考察系统内策略构成比的演化趋势和稳定性，以及合作涌现的情况. 这里所谓的适应性主体，是指参与人采用类似生物进化的方式，通过不断地相互学习模仿和动态调整来适应环境的变化，以获取更高的收益. 仿真模型中所有 Agent 之间的交互过程实际上就是彼此博弈的过程: 每个“自利的”Agent 根据各自的规则集预测其它 Agent 的行为，并采用恰当的行为最大化自己的利益(适应度)，最终所有 Agent 的决策将形成一个整体稳定的策略格局，即博弈的 Nash 均衡. 显然，这些 Agent 的决策不仅依赖于直接对手，也同样依赖于其它博弈单元的选手状况，因为模型演化的依据是所有 Agent 的相对适应度.

仿真模型包含若干采用不同 QTS 策略的 Agent，初始态 3 种策略按照等概率随机分布，并且同样以 $M(2)$ 作为支付矩阵. 在策略的进化方式上，模型没有采用传统的基于 3 种遗传算子(选择、交叉和变异)的演化算法，而是直接将每一个

演化代内(包含 1 000 轮独立博弈，在每一轮博弈中所有 Agent 以随机配对的方式展开 2-IPD 博弈)总收益最低的、比例值为 p 的那些 Agent 的策略进行随机变异，生成新策略. 这样，既体现了自然选择的思想，以保证优势策略的动态增长；又体现了随机扰动的思想，以保证策略的非确定性变异.

共设计了两组仿真实验，分别用于研究模型在不同总体规模下的 Nash 均衡和在开放环境下各种成分策略的均衡分布.

在第一组实验中，模型的 Agent 数量从 2 至 2 000 不等，以模拟不同程度的封闭环境和开放环境. 由于部分实验的 Agent 数量太少，为了避免偶然性，每次实验包含 20 组独立运行的仿真博弈，实验数据采用它们的平均状态值. 博弈模型的整体均衡状态主要通过各种策略所占比例的加权平均数 C_{avg} 来反映，其大小体现了全体 Agent 的平均合作意愿.

$$C_{avg} = (0 \cdot x[0] + 1 \cdot x[1] + 2x[2])/2 \tag{4}$$

如图 3 所示，当人数较少时，AllD 策略的比例几乎达到 100%，模型几乎没有合作现象；而当人数增至一定规模时，AllD 策略的比例急剧下降，同时 C_{avg} 曲线迅速上升，模型涌现出明显的、比较稳定的合作现象. 这一结果与前文关于相对收益和绝对收益的分析是一致的：演化模型中个体优劣的评判标准是他们的相对收益，而非绝对收益. 单一个体收益的绝对值再高，当其他个体的收益值更高时，该个体仍然会遭到排挤. 由此决定了两种博弈环境下，参与人决策的不同指导方针——封闭时趋向于相对保守的策略，宁可自己收益少也不能便宜对手；开放时趋向于互惠的策略，即使直接对手的获利会更多也要尽可能扩大自身收益，因为此刻的参照对象要比前一种情况广泛得多[17]. 此外，图 3 还说明封闭与开放是一组相对的概念，"封闭"并不仅限于博弈单元内的直接参与人，过少的博弈人群仍然应该视为"封闭"(因为总人数太少时，单元博弈中很难同时出现多位偏离者)，二者的界限与具体博弈的支付矩阵、随机扰动率以及参与人的理性程度密切相关. 例如在本例中，人数 40 至 60 这个区间就是"封闭"与"开放"的质变区间.

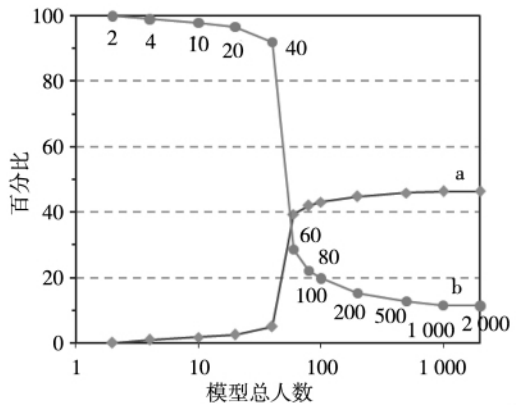

图 3 2-IPD 在不同人数规模下的演化状态图(X 轴为对数坐标)，其中支付矩阵为 $M(2)$，$p = 5\%$，a 表示 C_{avg} 曲线，b 表示 AllD 策略 $s[0]$ 的百分比曲线

Fig. 3 Evolution of 2-IPD with different agent numbers (X axis is logarithmic coordinates), where the payoff matrix is $\boldsymbol{M}(2)$, $p = 5\%$, a represents C_{avg} curve, and b is the proportion curve of AllD strategy $s[0]$

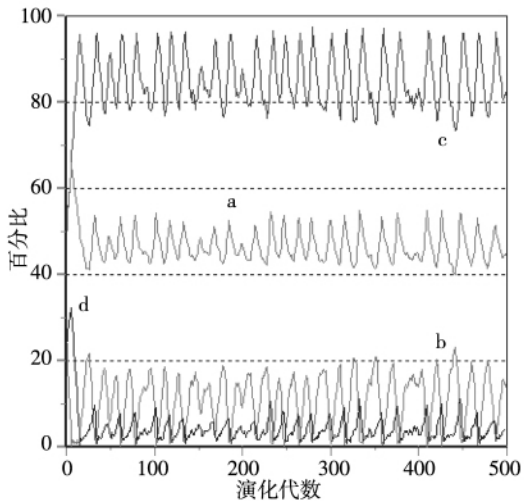

图 4 2-IPD 在开放环境下的演化状态图，其中支付矩阵为 $M(2)$，$p = 5\%$，a 表示 Cavg 曲线，b/c/d 分别表示策略 s[0]、s[1] 和 s[2] 的百分比曲线

Fig. 4 Evolution of 2-IPD in an open environment, where the payoff matrix is $\boldsymbol{M}(2)$, $p = 5\%$, a represents the evolution curve for C_{avg}, and b/c/d represents the proportion curves of three strategies $s[0]$, $s[1]$, and $s[2]$, respectively

在第二组实验中，总共包含 1 000 个 Agent，仿真一个典型的开放环境. 如图 4 所示，C_{avg} 曲线远离 X 轴，在 45% 上下波动，说明系统有明显的合作现象. 在三种可选策略中，$s[1]$ 的比例最高，

约占 85%; s [J] 其次,约占 11%; s [I] 最少,约占 4%. 仿真结果与图 2 中向量场的分析结论趋于一致. 同时,图 4 中系统的平均合作意愿曲线(a) 和策略构成比曲线(b/c/d) 都表现得相当不稳定,演化曲线呈周期性波动. 该现象进一步表明,2-IPD 模型的演化稳定态并非表现为策略构成比固定不变,而是一种循环的、不确定的、基于时间平均的均衡,这一点同样印证了前文的分析.

4　结束语

Nash 均衡定义了一种博弈的稳定状态,其中的每一个参与人没有单独偏离预期均衡的利益驱动,这在参与人总体规模不大的封闭环境下是可行的; 但在开放环境下,除了考虑自身博弈单元的直接参与人外,还需要考虑大量潜在参与人的偏离可能. 因此,在实际生活中分析博弈均衡时,不能完全忽略 Nash 均衡解之外的可选策略,尤其是当这些非均衡策略之间蕴含着较原 Nash 均衡收益大得多的支付增益时,它就可能形成新的利益驱动,使得非均衡策略成为一种优势策略,进而改变原博弈的均衡状态.

不仅如此,本文关于"潜在参与人"的概念还可以作进一步扩展. 由于经济联系的普遍性,个体优劣的比较往往是其综合收益的整体比较,它涉及到广泛的、各种形式的经济博弈. 因此,开放环境下的"潜在参与人"不仅应该包括同类博弈的潜在博弈群体,还应该包括小生境中所有与该博弈单元的任意一个直接参与人有直接或间接联系的经济个体. 显然,这是一个更加庞大、更加复杂的博弈系统,但也是一个更加真实的博弈系统,它将进一步修正纳什均衡理论与现实博弈之间的差距.

参 考 文 献:

[1] Patrick D, Andrzej S, Witold L, et al. A modal characterization of Nash equilibrium [J]. Fundamenta Informaticae, 2003, 57(2-4): 281-321.

[2] Ben-Porath E. Rationality, Nash equilibrium and backwards induction in perfect-information games [J]. The Review of Economic Studies, 1997, 64(1): 23-46.

[3] George J M. Do people play Nash equilibrium? Lessons from evolutionary game theory [J]. Journal of Economic Literature, 1998, 36(3): 1347-1374.

[4] 张　峰. 论博弈论的分析方法——纳什均衡分析法 [J]. 北京理工大学学报(社会科学版), 2008, 10(2): 95-99.
Zhang Feng. On the analysis method of Nash equilibrium in game logic [J]. Journal of Beijing Institute of Technology (Social Science Edition), 2008, 10(2): 95-99. (in Chinese)

[5] 邵　鹏. 纳什均衡的行为经济学精炼 [J]. 上海经济研究, 2009, 9: 77-83.
Shao Peng. Behavioral refinement of Nash equilibrium: Cognitive hierarchy model based on the group perception and experimental test [J]. Shanghai Economic Review, 2009, 9: 77-83. (in Chinese)

[6] 谢识予. 经济博弈论 [M]. 第 3 版. 上海: 复旦大学出版社, 2010.
Xie Shiyu. Economic Game Theory [M]. 3 Edition. Shanghai: Fudan University Press, 2010. (in Chinese)

[7] Normann H T, Wallace B. The impact of the termination rule on cooperation in a prisoner's dilemma experiment [EB/OL]. http://ssrn.com/abstract=952953. 2006, 12.

[8] Jiawai L, Graham K. Finite iterated prisoner's dilemma revisited-belief change and end-game effect [C]. BQGT 10 Proceedings of the Behavioral and Quantitative Game Theory: Conference on Future Directions, 2010.

[9] Kreps D, Milgrom P, Roberts J, et al. Rational cooperation in the finite repeated prisoner's dilemma [J]. Journal of Economic Theory, 1982, 27: 245-252.

[10] Weibull J. 演化博弈论 [M]. 上海: 上海人民出版社, 2007.
Weibull J. Evolutionary Game Theory [M]. Cambridge: MIT Press, 1995.

[11] Chong S Y, Humble J, Kendall J, et al. Iterated prisoner's dilemma and evolutionary game theory [M] // Chapter 2 of The Iterated Prisoner's Dilemma: 20 Years On. World Scientific, 2007.

[12] 敬 嵩，雷良海. 利益相关者参与公司管理的进化博弈分析[J]. 管理科学学报，2006，9(6)：82－86.
Jing Song, Lei Lianghai. Analysis of stakeholders management by evolutionary games theory[J]. Journal of Management Sciences in China, 2006, 9(6): 82－86. (in Chinese)

[13] 张洪潮，何 任. 非对称企业合作创新的进化博弈模型分析[J]. 中国管理科学，2010，18(6)：163－170.
Zhang Hongchao, He Ren. An analysis of evolutionary game model on cooperative innovation between asymmetric corporations[J]. Chinese Journal of Management Science, 2010, 18(6): 163－170. (in Chinese)

[14] Kandori M, Mailath G J, Rob R. Learning, mutation, and long-run equilibrium in games[J]. Econometrica, 1993, 61: 29－56.

[15] Ellison G. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution[J]. Review of Economic Studies, 2000, 67: 17－45.

[16] Young H P, Foster D. Cooperation in the short and in the long run[J]. Games and Economic Behavior, 1991, 3: 145－156.

[17] 杨 城，孙世新. 基于双层演化的多人囚徒博弈研究[J]. 计算机应用，2008，28(1)：108－111.
Yang Cheng, Sun Shixin. A 2-layer evolutionary model of n-player iterated prisoner's dilemma[J]. Journal of Computer Applications, 2008, 28(1): 108－111. (in Chinese)

The impact of potential players on Nash equilibrium: A perspective of cooperation in FIPD

YANG Cheng[1], *ZHU Chao-wei*[2,3], *LV Jun-min*[1]

1. School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China;
2. School of Economics and Management, Yichun University, Yichun 336000, China;
3. School of Business, Renmin University of China, Beijing 100872, China

Abstract: Considering that large number of real-life games inconsistent with the prediction of Nash equilibrium theory, this article argues that the closed game environment is another potential assumption of Nash equilibrium, in addition to the assumption of common rationality. Accordingly, this article expands the concept of potential players and assumes that the extra gains of deviating-players in an open environment might lead to a new equilibrium different from Nash equilibrium. To illustrate, an example of two-stage iterated prisoner's dilemma is introduced and the reason behind the new non-excepted-equilibrium is analyzed. In addition, according to the idea of CAS theory and multi-agent system, a simulation model is adopted to test the hypotheses proposed.

Key words: Nash equilibrium; potential players; evolutionary game; finite iterated prisoner's dilemma (FIPD); multi-agent system