

公共物品实验中的合作行为与促进机制

徐菱¹

摘要

作为经济学领域研究搭便车行为，促进合作等问题的基本模型，公共物品实验在实验经济学中有重要地位。本文综述了公共物品实验与公共物品博弈的研究历史，合作出现的原因，多种促进合作的机制，并提出了加入持有损失的新机制。本文着重探讨了惩罚机制与交流机制对合作的影响。

关键词：公共物品博弈，公共物品实验

1 引言 Introduction

公共物品实验 (Public goods game, PGG) 源自一个历史悠久的问题——公共物品困境。早在 1776 年, Adam Smith 就讨论了政府在提供路灯中的作用, 引出了公共物品使个人受益却难以由个人提供的问题。随着上世纪七十年代博弈论的发展, 公共物品困境再次引起人们的注意。Olson (1965) 将此问题抽象并设计成公共物品实验, 并在论文中正式指出实验结果与理论预测有显著的偏差。在此之后, 许多模型和机制被提出, 以解释和解决这一困境。Ledyard 在 1995 年的论文中综合了大量实验结果, 并提出了 MPCR (Marginal Per Capita Return) 作为合作回报的重要衡量参数。此后该问题在零售市场策略 (Boyd, 1997), 团队绩效管理 (李建标 & 尉孟星, 2012) 等多个领域都得到了关注和研究。尽管已有多种解释理论和解决机制, 但公共物品困境的成因依旧没有明确结论, 其应用领域也始终期待新的有效机制设计。

本文的第二部分介绍公共物品实验和对实验现象的解释性理论; 第三部分介绍促进合作的机制, 包括惩罚机制 (punishments), 交流机制 (communication), 阈值机制 (threshold) 和其他解决机制; 第四部分提出了一种新的机制——持有损失; 第五部分为总结。

2 公共物品实验 Public Goods Game Experiment

2.1 公共物品实验概述 Introduction of PGG

公共物品具有三个重要特征: 合作生产 (jointly provided), 非排他性 (non-excludable) 和非竞争性 (non-rivalled)。具体而言, 公共物品是个人无法生产或不愿意生产的, 其使用权无法被个人独占, 不同人的使用不会彼此干扰 (Holt, 2007)。国防、道路等资源都是典型的公共物品, 但私人物品和公共物品并没有明确界限。公共池塘资源 (common-pool resource) 即指竞争性但非排他性的物品, 如地下水资源和渔业资源。本文讨论对象局限于理想的公共物品, 严格满足上述的三个重要特征。

由于公共物品具有非排他性, 导致购买公共物品的人不可能享受独占该物品所带来的全部效用和收益。当一个人提供一件公共物品时, 尽管对社会而言总体收益高于个人成本, 但对个人来说成本高于收益。根据 Ledyard (1995) 提出的 MPCR, 即单位资本收益与机会成本之比, 提供公共物品对个人具有严格小于 1 的 MPCR, 因此对个人而言提供公共物品一定不是帕累托最优。因此, 公共物品难以在市场交换机制下实现生产。

合作生产公共物品因此诞生, 即众人合资购买公共物品, 并同时获得其效用。比如某村居民合资修建马路, 从而共同享受马路带来的便利。其中最简单的合作模式为同等支出, 同等收益, 即参与者均为同质的。此时只要整体收益大于整体支出, 即可保证个人收益大于个人支出。但对于个人而言, 参与合作并不一定是帕累托最优, 成为免费搭车者 (free-rider) ——不参与支出但享受收益分成, 可能可以获得更高收益。在此情况下, 个人将趋于不合作, 而由于参与者的同质性, 所有人都将不合作, 从而无法购买公共物品, 也无法享受任何收益。公共物

¹ 2020110687, 金融实验班, 22 秋季学期, 实验经济学-0796 期末论文

品困境由此产生。

在公共物品实验中， N 名参与者各拥有初始资金 b_i ，在不清楚其他参与者选择的情况下，独立将初始资金中的 c_i 投入“项目”中，投入资金称“合作”。“项目”产生的公共物品将具有投入总资金 r 倍的价值，并均分给每一位参与者。因此参与者 j 的最终收益为

$$\frac{r \sum_{i=1}^N c_i}{N} + b_j - c_j = \frac{r \sum_{i \neq j} c_i}{N} + b_j - (1 - \frac{r}{N})c_j.$$

其中 $1 < r < N$ ，否则将不存在困境。收益函数为线性，且在线性效用函数下，该收益与效用等价。称此模型为线性公共物品博弈 (linear public goods games)。在这一博弈中，唯一的纳什均衡是 $c_i = 0$ ，即所有参与者都选择不合作，陷入囚徒困境 (prisoner's dilemma)。这一结果同时还是有限重复博弈的子博弈完美纳什均衡和演化博弈均衡。

平均投入率 (average contribution) 定义为参与者投入资金与拥有资金的比值的平均值，即 $\frac{1}{N} \sum_{i=1}^N \frac{c_i}{b_i}$ 。根据上述纳什均衡，平均投入率应为零。但实际的实验结果中，参与者的平均投入率往往显著不为零。根据 Ledyard 在 1995 年的实验结果，在单轮实验中平均投入率高达 40%~60%。Goeree et al. (2002) 和 Issac et al. (1988) 等的实验也均得到了显著为正的投入率。实验结果表明在一定条件下人们可能并不满足纳什均衡所需的一系列假设，尤其是理性人假设。下述的条件合作理论和经济利他主义是解释公共物品实验中参与者选择合作的重要理论。

2.2 合作为何出现 Why cooperate

条件合作理论 Conditional Cooperation

早在 1970 年，Kelly 和 Stahelski 就在研究囚徒困境的过程中提出了条件合作理论。这一理论认为参与者合作与否并不仅仅依赖于自身决策，而是关于其他参与者的行为的条件函数。条件可能包含其他参与者过去的行为，以及对其他参与者未来行为的期望等。大量实验中都观察到了条件合作的现象，将公共物品实验理解为条件问题也有助于得到一个合作的均衡。

经济利他主义 Economic Altruism

Anderson, Goeree 和 Holt (1998) 的实验结果表明参与者至多愿意付出 10 美分来给他人 1 美元。因此将利他主义加入公共物品实验参与者的效用函数 (utility function) 是合理的，且从结果上确实可以解释部分合作存在的原因。也有研究者认为经济利他主义可以归入社会价值取向 (social value orientation)，即将参与者分为利他取向、合作取向、个人取向和竞争取向，前两类为亲社会者，后两类为亲自我者 (Bogaert, Boone & Declerck, 2008)。公共物品实验中亲社会者的存在导致了显著的合作现象存在。

3 促进合作的机制 Mechanisms to Promote Cooperation

3.1 惩罚机制 Punishment

Fehr 和 Gächter 在 2000 年的研究中最早正式提出了参与者之间的惩罚机制，并证明有成本的惩罚机制下参与者的平均投入率会更高，即惩罚机制有利于促进合作。Gürrer et al. (2006) 进一步验证了惩罚机制对合作有显著的促进作用。但 Nikiforakis (2008) 和 Mulder et al. (2006) 等人的研究指出惩罚机制可能存在负作用。

Fehr 和 Gächter (2000) 的实验设计为 4 人一组，MPCR=0.4，初始资金为 20 元，共 20 轮实验。实验的前十轮为对照组，进行无惩罚实验，即标准的线性公共物品实验。后十轮为实验组，参与者可以看到其他人的投入，并选择是否惩罚及惩罚的力度。不同惩罚力度对应不同成本，进行越严厉的惩罚需要付出越高的成本 (见下表)。实验结果表明，无惩罚实验的平均投入率为 19%，而有惩罚实验的平均投入率为 58%，显著高于无惩罚实验。无惩罚实验的最后一轮 (Round 10) 平均投入率为 10%，有惩罚实验的最后一轮 (Round 20) 平均投入率为 62%，同样显著高于无惩罚实验。实验还发现，被惩罚者往往是那些投入率低于小组平均的人，且低于平均越多，受到惩罚越多。

惩罚力度 (%)	0	1	2	3	4	5	6	7	8	9	10
惩罚成本 (元)	0	1	2	4	6	9	12	16	20	25	30

表 1: 惩罚力度与惩罚成本的对应关系

来源: Fehr and Gächter (2000)

在 Gürer et al. (2006) 的实验中, 参与者可以在 30 轮中随时选择加入有惩罚的实验环境(sanctioning institution, SI) 或无惩罚的实验环境 (sanction free institution, SFI)。在有惩罚的实验环境下, 参与者可以选择奖励 (positive sanction, reward) 其他人或者惩罚 (negative sanction, punishment) 其他人。实验开始时, 63%的参与者选择了 SFI, SFI 的平均投入率为 37%, 显著低于 SI 的 64%。随着轮次进行, 选择 SI 的人逐渐增加直至接近 100%, SI 的平均投入率也逐渐增加到 90%。相对的, SFI 的投入率降低到了 0。整个实验过程中 SI 的平均投入率为 91%, 显著高于 SFI 的 14%。参与者的决策演化过程如下图所示。

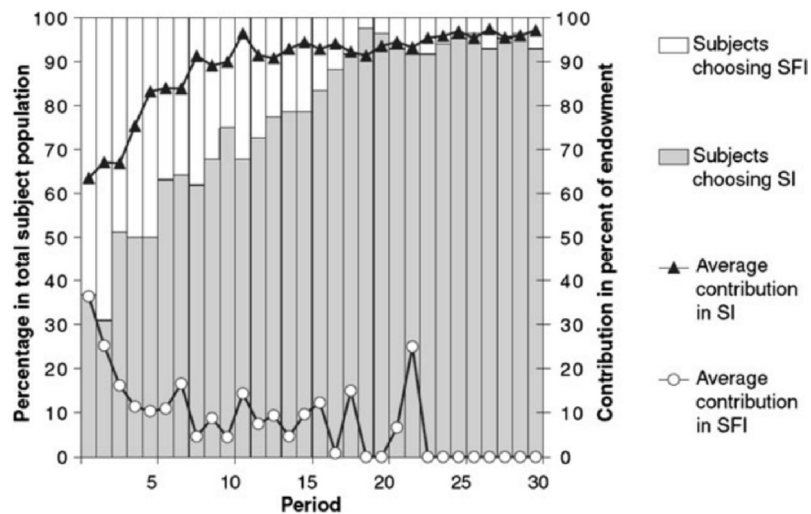


图 1: SFI 与 SI 组的人数比例与平均投入率

来源: Gürer et al. (2006)

尽管在惩罚机制下, 参与者的合作行为得到了很大促进, Nikiforakis (2008) 的实验却指出, 惩罚也可能是“反社会的” (anti-social)。Nikiforakis 发现 Fehr 和 Gächter (2000) 实验中除了高投入者对低投入者的惩罚, 还出现了搭便车者对高投入者的惩罚。他认为这种行为一方面出于策略目的, 另一方面则是为了转嫁自己受到的惩罚。在与 Fehr 和 Gächter (2000) 相同的惩罚机制下, Nikiforakis 将实验组设计为每轮包含三个阶段。第一阶段为标准的线性公共物品实验, 第二阶段为惩罚, 第三阶段为对抗惩罚 (counter-punishment)。在对抗惩罚阶段, 被惩罚且有剩余财富的人可以得知惩罚的施加者, 并选择是否付出一定成本来对抗惩罚这些施加者。对照组仅包含前两个阶段。参与者可以自由选择参与含对抗惩罚的实验或不含对抗惩罚的实验。实验结果表明, 含对抗惩罚实验的平均投入率显著低于不含对抗惩罚的实验, 且参与者的最终收益也要更低。而 Mulder et al. (2006) 的实验结果则发现, 当惩罚撤除时参与者的合作水平会下降, 甚至低于惩罚前的水平, 即惩罚可能会损害后续的合作倾向。王沛和陈莉 (2011) 以中国大学生为实验对象, 也发现惩罚机制会削弱参与者的人际信任水平, 对亲社会者的合作意愿有显著的减弱作用。

3.2 交流机制 Communication

在实验前的交流中参与者会往往做出关于投入值的承诺, 以促成较高的合作水平。在 Isaac 和 Walker (1998) 的实验中, 交流后的平均投入率可以达到 80%以上。而以中国高校学生为对象的实验中, 彼此熟悉的同学之间交流后的平均投入率甚至可以高达 96% (龚欣等, 2010)。Bochet et al. (2006) 对比了三种交流方式: 面对面 (face-to-face), 聊天室 (chat room) 和数字廉价谈判 (numerical cheap talk), 并将其与惩罚机制结合, 证明只有前两者能够显著提高平均投入率, 但惩罚机制对三种交流的效果都有明显促进作用。

在 Bochet et al. (2006) 的实验中, 共有八组不同的实验设置, 如下表所示。B 为标准的公共物品实验, FF 组

在所有实验开始前有 5 分钟的面对面交流时间，CR 组在第 1、4、7 轮实验开始前有一段线上聊天室的交流时间，NCT 组在每轮实验开始前都能向所有人发送若干次自己的“计划投入值”。R 组与对应的 wR 组中加入了惩罚机制。

	Without reduction option	With reduction option
Baseline	B 3 sessions 4 groups of 4 each session	R 3 sessions 4 groups of 4 each session
Face to face	FF 2 sessions 4 groups of 4 each session	FFwR 2 sessions 4 groups of 4 each session
Chat room	CR 3 sessions 4 groups of 4 each session	CRwR 3 sessions 4 groups of 4 each session
Numerical cheap talk	NCT 3 sessions 4 groups of 4, 2 sessions 3 groups of 4, 1 session	NCTwR 3 sessions 4 groups of 4, 2 sessions 3 groups of 4, 1 session

表 2：八个实验组的设置

来源：Bochet et al. (2006)

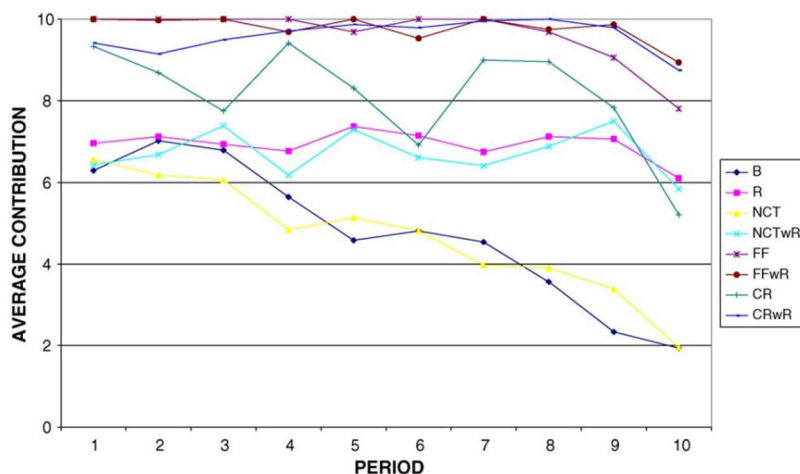


图 2：不同组的平均投入率

来源：Bochet et al. (2006)

各组的平均投入率如上图所示。面对面交流后的平均投入率跃升到了 96%，加入惩罚机制的影响并不明显，只将平均投入率提高到了 97%。聊天室交流的效果不如面对面对交流，但依旧有 81% 的平均投入率，这在加入惩罚机制后提高到了 96%。而数字廉价谈判组与普通对照组的平均投入率都显著低于这些组，且呈现递减趋势。尽管在加入惩罚机制后略有提高，但仍显著低于其他实验组。这一结果表明，一些交流机制确实可以提高合作的程度，但交流中也可能存在无效承诺甚至欺骗。对于较高合作程度的交流机制，加入惩罚机制的作用相对较小；但对较低平均投入率的实验，惩罚机制能显著提高合作表现。

3.3 其他机制 Other Mechanisms

其他解决机制还包括阈值机制 (Croson & Marks, 1998)，领导机制 (马博, 2013) 和非对称机制 (Hauser et al., 2019) 等。阈值机制要求参与者的总投入值高于某一阈值方能得到回报，否则投入将不带来任何收益。这一机制能够促进参与者投入更多以避免达不到阈值而浪费投入的资金，但当集体的平均投入率很低时也会抑制个体的投入意愿。领导机制下每轮实验将被分为两个阶段，由一名参与者即领导者首先投入一定金额并公开，再由其他参与者进行投入。实验结果表明领导者的投入率对平均投入率有较大影响，高投入的领导者能促进合作，而低投入的领导者不利于合作。非对称机制包括初始资金差异和 MPCR 差异等。其中 MPCR 差异包括生产力差异和份额差异，前者指不同参与者投入资金被乘以的倍数不同，后者指不同参与者从公共物品中获得收益的份额不同。非

对称机制的设计和效果可以非常复杂。

4 持有损失 Holding loss

在金融学中，货币具有时间价值，收益也具有时间偏好，即相同收益在当期比在未来更具价值。用折现率 ρ 来描述，即参与者认为当期的1单位收益与下一期的 $1 + \rho$ 单位收益等效。这可以理解为持有货币的机会成本。出于这种原因，持有现金而不投入到公共物品的生产中也存在一定损失。参与者选择持有的 $b_j - c_j$ 现金在下一期的价值将降低为 $\frac{b_j - c_j}{1 + \rho}$ ，因此收益转化为

$$\frac{r \sum_{i=1}^N c_i}{N} + \frac{b_j - c_j}{1 + \rho} = \frac{r \sum_{i \neq j} c_i}{N} + \frac{b_j}{1 + \rho} - \left(\frac{1}{1 + \rho} - \frac{r}{N} \right) c_j$$

当公共物品的生产期较短时， ρ 较小，因此 $\frac{1}{1 + \rho} - \frac{r}{N} > 0$ ，不会对博弈的均衡造成影响。但当公共物品生产期很长时， ρ 可能很大，导致 $\frac{1}{1 + \rho} - \frac{r}{N} < 0$ ，从而使博弈的均衡变为100%投入。然而在 ρ 较小的情况下，可以预计到参与者并不会完全理性地选择零投入，而依旧会选择投入相当一部分，甚至比不加入持有损失时的投入率更高。这一猜想可以由两个角度解释。一方面，参与者从心理上不愿意承担持有现金的损失，因为这一损失类似于无谓损失 (dead loss)。另一方面，斜率 $\frac{1}{1 + \rho} - \frac{r}{N}$ 变小使得演化过程的收敛趋势更弱，参与者选择减小投入带来的收益增益会更小，从而其他促进投入的因素效果会更明显。

5 总结 Conclusion

本文研究了公共物品实验中的合作行为与促进机制，对公共物品实验的研究进行了综述，着重探讨了惩罚机制和交流机制，并指出两者对合作的双重影响。本文提出了加入持有损失对公共物品实验可能的影响，但由于条件限制未能进行实验验证。公共物品实验相关理论极其丰富，本文仅能覆盖其中较小的一部分，希望未来有机会能够有更加深入的研究。

参考文献 Reference

- [1] Ananish, C. (2011). Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature. *Experimental Economics*, 14, 47-84.
- [2] Anderson, S., Goeree, J., & Holt, C. (1998). A theoretical analysis of an altruism and decision error in public goods games. *Journal of Public Economics*, 70, 297-323.
- [3] Bochet, O., Page, T., & Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization*, 60(1), 11-26.
- [4] Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, 47, 453-480.
- [5] Boyd, D. W. (1997). From "Mom and Pop" to Wal-Mart: the impact of the consumer goods pricing act of 1975 on the retail sector in the United States. *Journal of economic issues*, 31(1), 223-232.
- [6] Charles, A, Holt, C. (2007). *Markets, Games, & Strategic Behavior*. Princeton: Princeton University Press (Chap. 26).
- [7] Croson, R., & Marks, M. (1998). Identifiability of individual contributions in a threshold public goods experiment. *Journal of Mathematical Psychology*, 42(2-3), 167-190.
- [8] Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, 90(4), 980-994.
- [9] Goeree, J. K., C. A. Holt, & T. R. Palfrey. (2002). Quantal response equilibrium and overbidding in private-value auctions. *Journal of Economic Theory*, 104(1), 247-272.
- [10] Gürer, Ö., Irlenbusch, B., & Rockenbach, B. (2006). The competitive advantage of sanctioning institutions. *Science*, 312, 108-111.
- [11] Isaac, R., & Walker, J. (1988). Communication and free riding behavior: the voluntary contributions mechanism. *Economic Inquiry*, 26(4), 686-608.
- [12] Kelly, H., & Stahelski, A. (1970). Social interaction basis of cooperators and competitors' beliefs about others. *Journal of Personality and Social Psychology*, 16, 66-91.
- [13] Ledyard, O. (1995). Public Goods: A Survey of Experimental Research. In A. Roth & J. Kagel (Eds.), *A Handbook of Experimental Economics*. Princeton: Princeton University Press (Chap. 2).
- [14] Mulder, L. B., Van Dijk, E., & Wilke, H. A. M. (2006). Undermining trust and cooperation: The paradox of sanctioning systems in social dilemmas. *Journal of Experimental Social Psychology*, 42, 147-162.
- [15] Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: can we really govern ourselves? *Journal of Public Economics*, 92, 91-112.
- [16] Olson, M. (1965). *The logic of collective action: public goods and the theory of groups*. Cambridge: Harvard University Press.
- [17] Smith, A. (1976, originally 1776). *The Wealth of Nations*, E. Cannan, ed., Chicago: University of Chicago Press.
- [18] 李建标 & 尉孟星. (2012). 适当流动性能提高团队绩效——来自公共物品实验的证据. *经济与管理研究*, 10, 113-120.
- [19] 马博. (2013). 公共品博弈实验中的合作行为研究进展. *经济论坛*, 11, 163-169.
- [20] 王沛 & 陈莉. (2011). 惩罚和社会价值取向对公共物品两难中人际信任与合作行为的影响. *心理学报*, 43(1), 52-64.
- [21] 龚欣, 刘文忻 & 张元鹏. (2010). 公共品私人自愿提供决策的实验研究. *中南财经政法大学学报*, 181(4), 26-32.