

# 如何通过内生惩罚解决异质性群体的 集体行动困境？ ——博弈模型与案例分析

章平, 刘启超

(深圳大学 中国经济特区研究中心, 广东 深圳 518060)

**摘要:** 合作与协调是人类社会的基本问题。经典理论预测, 作为动机复杂且各不相同的行为主体, 人与人之间围绕共同的事或物而发生的集体行动往往会落入“个体理性导致集体非理性”的困境。文章基于对中国转型社会一个城中村长期观察发现, 异质性程度大的群体也可以实现自我合作治理。为了弥合现有理论与经验证据的鸿沟, 文章基于兰彻斯特法则拓展了现有的个体行为博弈模型和案例分析, 找到了解决异质性群体集体行动困境的关键变量——内生的监督和惩罚。相较于政府或市场干预, 自我治理组织的制度性建构及其运作, 有助于产生符合大部分个体预期的内生监督和惩罚, 更有效地维持公共事务自我治理的集体行动。

**关键词:** 合作治理; 集体行动; 异质性; 内生惩罚

**中图分类号:** F062.6    **文献标识码:** A    **文章编号:** 1001-9952(2020)05-0004-13

**DOI:** [10.16538/j.cnki.jfe.2020.05.001](https://doi.org/10.16538/j.cnki.jfe.2020.05.001)

## 一、问题的提出

标准理论认为, 理性自利的个体不会自愿为集体利益付诸努力, 个体理性将导致集体的非理性。因此, 有必要引入市场或政府(科层制)以解决集体行动困境, 实现个体的长远利益。然而, 大量经验证据表明, 市场会因为低排他性和产权无法界定而产生失灵(Samuelson, 1954), 政府也会因为成本、信息不对称和寻租行为而失灵(Buchanan 等, 1980)。大范围的田野调查也证实, 为实现自由贸易、分担风险和保护环境资源, 世界各地各行各业的人们会自愿组织起来, 并建立规则规范。这一理论模型和经验之间的鸿沟如何弥合? 原先假设的“理性人面对同样的情境时会采取同样的方式进行评价决策”显然与事实不符(Ostrom, 2010)。因此, 理论和实证研究者转而采用更加接近现实、更为复杂的理论框架和模型来理解集体行动的问题, 提出市场和国家之外的第三种模式——集体内部的自我治理。他们致力于回答: 人们如何围绕公共事物展开集体协商、避免搭便车的机会主义行为, 达成合作进而形成自我治理的集体行动? 而不同治理模式对不同情境、不同群体的适用边界又在哪里?

收稿日期: 2019-03-06

基金项目: 国家社会科学基金项目(14BZZ086); 广东省哲学社会科学规划项目(GD19CYJ24)

作者简介: 章平(1981—), 男, 浙江嘉善人, 深圳大学中国经济特区研究中心副教授;

刘启超(1991—), 男, 河南周口人, 深圳大学中国经济特区研究中心硕士生。

上述情况仍然是在一般意义上讨论经济学理论与现实差距的问题。在中国,公共品供给的集体行动嵌入在经济社会文化的多重转型过程中,理论解释的难度更大、更复杂。随着全球化、工业化和城市化的快速推进,城城、乡城之间出现了大量的人口迁移,费孝通意义上的“乡土中国”已经成为“城乡中国”(周其仁,2017),传统的熟人社会正在瓦解,安土重迁、低流动性的熟人社会日益远去。<sup>①</sup>特别是从乡村进入城市的流动人口,他们非亲非故,围绕大小、有形无形的公共品需求,在政府投入不足和市场失灵的情况下,群体内部能否走向自我合作、实现公共品供给?从理论和实践上来说,这是一个异质性群体下的集体行动问题——群体内部、异质性个体之间如何协调互动、如何相互制衡约束形成一种能够稳定各参与方行为预期的治理模式,又如何互相监督实施?

作为本文案例来源城市的深圳,2018年常住人口1302.66万人(其中户籍人口454.70万人),实际管理人口约2100万人。越来越多的非亲缘、非地缘和非血缘关系的人聚集在同一社区。调研发现,在这样的高度异质性个体面临的集体行动困境中,依然存在如Fehr和Gächter(2002)、Ostrom(2010)等人考察的自愿惩罚等监督机制来约束搭便车行为,用以维系公共品供给的现象。这一系列的事实与“规模越大、个体间异质性越大,集体行动问题越难解决(曼瑟尔·奥尔森,2014)”的标准理论相悖,两相比较,我们不禁要问:来自全国各地、千差万别的城市流迁人口,生活在同一社区空间,相互之间如何围绕共有、共享的公共品展开共治以解决集体行动问题?特别是在群体异质性程度大、个体间预期不稳定的情况下,相互监督和内生惩罚这一举措如何能够在维系集体行动过程中发挥作用?

为了回答上述问题,本文主要做了以下几项工作:(1)拓展兰彻斯特法则,刻画个体的不同行为动机,考察集体行动合作发生和维持的条件;(2)引入内生惩罚作为可置信威胁,研究其对非市场、非政府的公共品供给集体行动的影响,为解决传统的集体行动困境提供了一种可能性;(3)通过取自深圳这一典型移民城市的现实案例分析,实证检验了此类公共品供给中内生惩罚对于集体行动自发形成的作用,丰富了集体内部自我治理的理论和实践。

## 二、文献综述

公共品博弈作为研究人们在集体行动中个体行为的一般分析框架,具备人类现实社会生活的基本属性。在同质的理性人假设下,传统理论认为个体会在利益最大化驱动下选择搭便车的机会主义行为,合作会瓦解从而陷入“多人囚徒困境”。但是现实中的大量案例和实验表明,很多情况下人们实现了集体内部自我合作治理的集体行动(Ostrom,1990和2010)。

针对经典理论还不能很好解释、而现实中已经发生的成功集体行动,研究者陆续为我们提供了新的解释思路和角度:一是调整个体基本假定,将同质个体放松为更接近现实的、可用多个维度度量的异质性假定(Fischbacher等,2001;Burlando和Guala,2005;Fischbacher和Gächter,2010;周业安等,2013),通过群体内个体间差异的角度来解释集体行动达成合作的可能性;<sup>②</sup>二是

<sup>①</sup> 费孝通于1947年提出“乡土中国”这一论断时,中国的城市化率不足10%,到2018年常住人口的城镇化率达到59.58%。

<sup>②</sup> 在同质理性人(面对相同情境作出相同决策)假设下,“零供给(zero contribution)”在理论上是必然的;而现实中,人与人之间在不同维度的差异,在有搭便车行为时,“非零供给”现象非常普遍。因此,研究者尝试放松个体基本行为假定,用于弥合理论和现实鸿沟、解释社会合作如何发生——正如我们在生活中看到的:第一,合作是存在的,而非理论预测那样是“零供给”;第二,群体内个体间差异越大,合作可能性究竟是更小还是更大,这一对关系应该是非线性的(如因徒困境刻画的是参与人同质且理性的一般情况,在差异悬殊的极端情况下,转化为“智猪博弈”中大猪小猪间差异足够大的话,也可能产生非零供给的合作解)。这里感谢审稿人分享了自己的理解并提醒可能产生的歧义。

引入惩罚和奖励的激励机制来促进合作、监督机会主义行为(Balliet等,2011),这类文献主要研究有无惩罚、不同惩罚方式(Fehr和Gächter,2002;Xiao和Houser,2011)、惩罚实施主体的内生和外生(Gächter和Thöni,2005;Reuben和Riedl,2009)等变量对公共品供给集体行动的影响。

(一)从同质性假定到异质性假定的转变及其度量。第一类方向上的文献,重点在于个体异质性的表征和度量及其在公共品供给集体行动中的作用。这类研究试图回答“个体间的异质性是否有利于集体行动的发生和维持?是否可以通过拓展异质性来提升集体行动的可能性?”在异质性的表征和度量方面,Fehr和Gächter(2000)、Fischbacher和Gächter(2010)在其一系列的开创性研究中对“内在偏好差异对公共品贡献”的集体行动成员进行了分类。Ruttan(2008)在公共品集体行动中引入了个体的经济异质性和文化经济异质性,以此解释不同组织治理机制的选择。连洪泉和周业安(2015)从主观心理和客观存在的角度刻画了个体异质性,将其处理为内在偏好(利他、互惠等内在社会偏好)异质性和外在特征(性别、语言和社会角色)的异质性。<sup>①</sup>

在个体异质性影响公共品供给的集体行动方面,从实验室实验到实际案例都有大量文献,分别研究了不同维度的异质性如何影响了集体行动。已有研究关于异质性有利于还是有害于集体行动发生,尚未定论。Baldwin和Huber(2010)利用46个国家和地区的数据,分析得到了群体异质性和公共物品供给量的关系:群体间语言和不平等社会偏好的差异越大,公共品供给水平越低,而文化差异则没有显著影响。选取四川省卧龙自然保护区为实证对象,Yang等(2013)尝试了以森林管理人员数量来解释森林覆盖率的研究发现,第一,群体规模也并不显著影响公共品供给水平。第二,允许惩罚搭便车者和允许个体自主选择所属群体,可以提高集体行动成效。上述结论也在连洪泉等(2015)的实验研究中得到进一步论证,披露群体信息和群体选择,可有效提高合作水平;社会偏好程度高的个体及其组成的群体,能够显著提高公共品供给水平。<sup>②</sup>

(二)不同惩罚机制的设计引入及其效果比较。第二类研究方向上的文献,强调通过提供激励与约束的赏罚机制可以减少甚至杜绝搭便车行为。本文主要关注惩罚,一般认为,带有成本的惩罚可以维持社会合作,也是保障合作可持续的必要条件。群体内部实施的惩罚,可以提高搭便车策略的成本;而实施惩罚的成本,也具有规模经济,随着实施惩罚者数量的增加而降低(Boyd等,2003)。Becker(1968)认为惩罚通过改变目标行为的回报从而影响结果,惩罚产生的威慑作用——被惩罚者的预期损失等于实际成本与被惩罚概率的乘积,因此可以通过增加惩罚的概率或幅度来威慑违规行为,这会使违规行为的预期成本大于收益。但是Houser等(2008)研究了另一种可能性:当受到惩罚的威胁时,人们可能会变得不那么具有合作精神。

在惩罚的方式上,主要有物质惩罚和非物质惩罚。前者是对背叛者进行制裁、罚款等货币惩罚;后者是群体内部对背叛者实施“鄙夷”“社会排斥”和“谴责”等惩罚,主要通过作为社会资本重要组成部分的亲社会性和价值观来约束背叛行为(Masclot等,2003)。人类社会的合作是由物质和非物质的激励所驱动的。物质激励可以是积极的(赏),也可以是消极的(罚)。两者都可以由法律正式实施,也可以由同伴非正式实施。非物质激励通常与情感有关,如内疚、尊重或羞耻。

<sup>①</sup> 利他偏好、互惠偏好显然属于异质性的范畴:第一种情况,是0-1型的,也就是有没有社会偏好(利他、互惠)的差异,在这种情况下,如果所有人都有相同的利他偏好,那就没有异质性。第二种情况,是社会偏好的强度或程度不一样,比如分布在[0,1]区间上程度不等的社会偏好。本文的处理是:三类个体——背叛者(也就是传统的搭便车者,不具有社会偏好)、合作者(具有社会偏好并会合作)、惩罚者(具有社会偏好、自己选择合作,而且还会惩罚背叛者)。

<sup>②</sup> 一般来说,群体规模和异质性都是作为影响集体行动发生的两个重要因素。奥尔森在《集体行动的逻辑》中写道:“理性的自利个体不会为公共的或群体的利益付诸努力,除非群体规模很小,或者通过强力以及其他特殊机制。”但是很显然,规模和异质性之间也具有某种非线性相关关系,规模越大,可能的差异也越大,进而对公共品供给的行动影响也会分化(Zhang和Zhu,2011)。

这些非物质激励常常可以激励个人采取合作行动,它们有时与物质激励并不一致。虽然非物质激励可以是限制个人机会主义行为的内在机制,但是为了促进社会秩序,也往往需要设计物质激励进行外部干预(Xiao, 2018)。

在惩罚实施上,主要包括外生惩罚(第三方惩罚)和内生惩罚(第一、第二方惩罚),两种惩罚实施影响合作的机理和效果大不相同。Fehr 和 Gächter(2000)定义外生惩罚是指借助群体外部主体或正式法律制度,惩罚搭便车等机会主义行为;而内生惩罚是在群体内部自我组织、协商达成一致而实施的惩罚,或者是部分参与者实施的利他惩罚(Fehr 和 Gächter, 2002)。两种惩罚的有效性一直是大量文献研究的焦点问题。

国外大量的理论和实证研究认为,群体内部成员平等协商形成的内生制度(典型如惩罚),对集体行动中的合作水平具有十分关键的影响。运行制度或规则的产生方式本身也会促进合作产生溢价,为数不少的研究表明(Hyde, 2011),同样一项制度,相较于外生强加的方式,由群体内部成员协商选择形成的内生制度下发生的合作水平更高,即“内生溢价”。早期曼瑟尔·奥尔森(2014)的经典研究认为,民主治理可以使公共品供给达到集体理性状态下对应的量。Decker 等(2003)进行的允许群体内成员投票选择惩罚机制来惩罚他人的实验发现,一致同意是效率最高的策略。Sutter 等(2010)实验显示,被试拥有自主参与权可以提升合作,亦即内生赏罚制度显著优于外生。Mathew 和 Boyd(2011)研究发现,在特定制度条件下,有代价的内生惩罚不仅能够解释小型群体合作的演化,还能够维持较大群体发生的集体行动用以对抗外来入侵。

与上述结论不同的是,周晔馨等(2014)研究发现了惩罚过程中存在“内生溢价”现象,但是其适用边界不一而足,不同群体中可能产生的溢价方向不同、机理各异。连洪泉等(2013)的公共品实验发现,由于存在报复和反惩罚等行为带来的风险,其结果是外部惩罚比内生惩罚更能够提升合作水平。闫佳和章平(2016)也通过实验室实验验证了来自外部主体的集中式惩罚可以带来合作水平的显著提高,而作为对照组的内生惩罚效果则并不明显。

最后,如何实施惩罚机制也会直接影响效果。Xiao 和 Houser(2011)的公共品博弈实验表明,与没有惩罚的基线处理相比,私人实施的惩罚减少了合作。然而,当同样的激励措施公开实施时,合作率会持续保持且显著高于基线或私人惩罚时的水平。可见,使个体的私人信息转化为公共信息,可以增加他人参与、实施惩罚的意愿,关键是如何让正确的惩罚信号向群体中其他成员扩散。人类社会中的群体可以采取会议协商一致和独立司法程序的办法,为私有信息转变为公共信息提供了可能(Ertan 等, 2009)。

(三)文献评述总结及本文可能的边际贡献。回顾上述文献的发展脉络,可以归纳如下两点:

第一,对异质性个体假设的拓展是目前文献的主攻方向,且一般认为个体内在偏好的异质性会增加达成集体行动的难度。异质性如何界定、度量?绝大部分研究都是从某一特定视角切入,比如外在特征、内在偏好。第二,惩罚(激励)机制能够提升合作,但是产生的效果可能因人因地因情境而异。谁来实施惩罚监督(第一方、第二方还是第三方),如何产生惩罚制度(外生还是内生),货币还是非货币的惩罚形式以及惩罚的净效率(实施惩罚的代价和提升的合作收益之差)都会影响其效果。实验研究中还会存在被试群体间的国家地域文化差异的作用。

在国内为数不多的使用中国被试群体的实验研究中,“内生溢价”的结论尚不稳健,它会因为具体研究对象、规模、社会和文化环境的变化而发生变化,即因人而异、因地制宜。现有文献虽然在异质性个体的假设上进行了很好的拓展,但是未曾阐述正确的惩罚信号如何向群体中其他成员扩散,来促使惩罚和监督机制的实施。在中国转型过程中,大量丰富生动、具有典型意义的集体内部自我治理案例并没有得到重视。

本文的研究与上述两类文献一脉相承、也有所推进,在已有文献基础上展开了理论模型分析和基于案例的实证研究,尝试解释:在群体与群体之间、个体与个体之间异质性程度较大的情形下,从出现发起群体到集体行动发生,进而当出现搭便车等卸责行为时,如何选择运用内部监督和惩罚策略以维持已经发生的集体行动。

与已有文献集中运用实验室控制实验数据的分析解释相比,本文的可能贡献主要体现在:借鉴实验室实验(Fischbacher和Gächter,2010)中对个体异质性的普遍分类度量方式,即以在行为的方式来识别内在偏好,根据动机的不同将群体划分为合作者、背叛者和惩罚者,并将其纳入演化生物学模型(Lanchester's Law),刻画群体内个体间如何协商一致对背叛行为采取惩罚措施——考虑在长期重复博弈中内生惩罚制度对集体行动的影响,进而为我们从微观动机的层面理解集体行为中的合作涌现提供一种新的视角,并给出一个实际发生的案例研究作为经验证据。

### 三、基于兰彻斯特法则的行为博弈模型与研究假设

针对“异质性群体互动博弈是否会收敛到合作均衡?其条件又是什么?”我们尝试在Hwang(2009)扩展的兰彻斯特法则行为博弈模型的基础上,做进一步扩展,将强互惠理论引入该模型,假设没有激励条件下,个体也会成为利他惩罚者。这样就能用于刻画现实中涌现的“不惜成本、不计得失”地惩罚搭便车行为,这样即使会降低自身收益(适应性),但是从长远看,可以提高群体的平均适应性。

Hwang(2009)扩展了兰彻斯特法则,证明在集体行动中,当群体规模和惩罚者增多时会降低惩罚成本和提高集体行动成效。本文进一步对上述模型进行了扩展,放松其模型假设。在Hwang(2009)的模型中,个体在一定激励条件下如奖励,才会站出来对搭便车者进行惩罚,本文把塞缪尔·鲍尔斯和赫伯特·金迪斯(2015)的强互惠理论引入该模型,认为不需要激励,为了公平个体也会成为利他惩罚者。这样就可以解释现实中有许多个体或群体不惜在自己付出成本下去惩罚搭便车者,即降低自身的适应性,在接下来的交互中来提高群体的平均适应性。

对应到下一部分的深圳故事中,在整个业主群体中,凡是“按时缴纳物业费”的业主记为合作者,否则记为背叛者;而惩罚者是指,面临“有业主连续多期未缴物业费”一事进行协调会商、决定将卸责行为诉诸司法途径处理的业主,当然,这部分业主首先是被作为合作者定义的。拓展后的模型具体如下:

一个由若干亚群体组成的群体,亚群体成员之间非亲非故、随机匹配形成,规模为 $n$ ,内部进行重复博弈。可选策略集: {背叛( $D$ ),合作( $C$ ),惩罚( $P$ )}。背叛( $D$ )界定为不合作、公共品零供给;合作( $C$ )是指公共品非零供给,但是个体选择合作策略后不再会对背叛者实施惩罚;惩罚( $P$ )是不光会供给公共品,还会对背叛者实施惩罚。

就集体行动全过程而言,首先是发起阶段,但本文模型重点关注集体行动发生之后的维持阶段,这一阶段个体的理性选择决定了是否会出现内生的监督惩罚,三类效用函数如果用兰彻斯特法则表征如式(1)–式(3)所示:

$$\pi(P) = \frac{b}{n} \sum_j E_j - c - dH(I_P) \quad (1)$$

$$\pi(D) = \frac{b}{n} \sum_{j \neq i} E_j - sH(I_D) \quad (2)$$

$$\pi(C) = \frac{b}{n} \sum_j E_j - c \quad (3)$$

式(1)–式(3)分别表示惩罚者、背叛者和合作者类型的收益。其中,  $b$  表示从公共品中获得的收益,  $c$  是贡献公共品所付出的成本,  $b$  和  $c$  满足  $c < b < nc$ 。  $E_i$  表示个体  $i$  捐赠的公共品量(亦即合作水平)。为简便起见, 我们将此类公共品集体行动界定为 0–1 型, 有成功供给为 1, 反之为 0。式(1)中的  $H(I_p)$  和式(2)中的  $H(I_D)$  分别为因实施惩罚和选择背叛而付出成本的概率,  $d$  和  $s$  分别为实施惩罚和选择背叛而产生的代价。可以看出, 当惩罚发生时, 即  $s$  与  $d$  均不为 0 时, 对强互惠者或惩罚者( $P$ )而言, 其收益  $\pi(P)$  会具有额外的损失  $dH(I_p)$ ; 对惩罚者而言, 其收益  $\pi(D)$  会具有额外的损失  $sH(I_D)$ 。不同的是, 强互惠者是自愿付出额外的成本或以损失自己的适应性来促进合作的维持, 而背叛者是因惩罚的发生而损失额外的成本。在这种情况下, 背叛者会权衡合作的收益  $\sum_{j \neq i} E_j$  与惩罚产生的  $sH(I_D)$  来进一步决定自己的策略。因此, 正是因为强互惠者不惜付出成本牺牲自己的适应性来惩罚背叛者, 这才使合作有得以维持的可能。此外, 若  $s=d=0$ , 则表示惩罚未发生或未进入惩罚阶段, 可以看出此时采取背叛策略或搭便车策略是占有策略, 进而个体会倾向于选择搭便车策略, 这会使合作产生瓦解的可能性。

兰彻斯特法则最初是一个军事模型, 用来刻画基于军队规模和单兵作战能力的战斗效率(包括平方律和线性律)。Adams 和 Mesterton-Gibbons(2003)等人用该法则解释了社会动物的集体行为规律。Johnson 和 MacKay(2015)较早地尝试了运用平方律来预测人的集体行为。

本文旨在分析集体行动维持阶段背叛者与惩罚者之间的较量, 用  $r$  表示一个惩罚者所面临的背叛者, 具体用背叛者人数的倒数计算: 当  $r=1$  时, 意为只有一个人背叛; 当  $r \rightarrow 0$  时, 背叛者数量越来越多,  $r$  越来越小, 人们选择背叛策略就越来越普遍。在平方律作用下, 式(4)为惩罚策略(惩罚者)战胜背叛策略(背叛者)的条件为:

$$mx^2 \geq a[ry](1/r)^2 + a(y - [ry]1/r)^2 \quad (4)$$

其中:  $[x]$  代表取  $x$  的整数部分。我们令:

$$\varphi_{(x,y)} = mx^2 - a[ry](1/r)^2 - a(y - [ry]1/r)^2 \quad (5)$$

显然, 失败的策略(亦即持该策略类型的个体)更容易付出代价而被淘汰, 可以假设惩罚者、背叛者付出代价的概率  $H(I_p)$  和  $H(I_D)$  仅取决于式(5), 可以得  $H(I_p)$  和  $H(I_D)$  具体表达式为:

$$H(I_D) = F(\varphi_{(x,y)}) \quad (6)$$

$$H(I_p) = 1 - F(\varphi_{(x,y)}) \quad (7)$$

由上述可知  $\lim_{t \rightarrow -\infty} F(t) = 0$ ,  $\lim_{t \rightarrow +\infty} F(t) = 1$ 。即当  $t \geq 0$  时,  $F(t)=1$ ; 当  $t < 0$  时,  $F(t)=0$ 。

合作策略、惩罚策略占优于背叛策略, 并成为纳什均衡策略的条件如式(8)和式(9)。

$$\pi(P) > \pi(D) \quad (8)$$

$$\pi(C) > \pi(D) \quad (9)$$

结合式(1)、式(2)和式(3)可知:

$$\lim_{n \rightarrow +\infty} [\pi(C) - \pi(D)] = \lim_{n \rightarrow +\infty} [\pi(P) - \pi(D)] = s - c > 0 \quad (10)$$

由此可得, 合作策略、惩罚策略占优于背叛策略需满足  $s > c$ , 其含义为: 个体在维持集体行动的惩罚阶段付出的代价, 高于集体行动发生时为公共品供给而付出的代价。背叛带来的收益同时低于合作策略和惩罚策略时, 没有人会选择搭便车。

我们将公共品集体行动分为两个阶段: 第一是集体行动的发起; 第二是本文解释的集体行动的维持(监督惩罚)阶段, 然后给出以下两个假说。

假说 1: 在集体行动维持阶段, 对非亲缘、非地缘关系的异质性群体而言, 若不存在有效监督和可置信惩罚(或尚未进入惩罚阶段, 即  $s=d=0$ )时, 由于个体的理性选择是搭便车策略, 结果是合

作的集体行动将难以维持。

假说 2: 在集体行动过程中引入惩罚选项(或进入维持集体行动的惩罚阶段,即  $s \neq 0$  且  $d \neq 0$ )时,存在对背叛行为的内生监督惩罚,可以形成可置信的威慑,促使背叛者形成会受到惩罚进而付出代价的行为预期,这可以减少甚至杜绝搭便车等机会主义行为的发生,从而使成功发起的集体行动得以维持。

#### 四、来自深圳西头新村的合作治理集体行动案例

深圳西头新村是高速高度城市化、城市化滞后于工业化进程的产物,建成初期的“脏乱差”现象与一般城中村的形象并无二致,之后进入基层治理政策的实践者、研究者视野的主要原因是:第一,业主聚集在一起并非因为地缘和亲缘关系,而是短时间购地建房使然;成员之间差异大,依然能够自发形成公共品供给的集体行动,并且前期没有政府、市场等外部组织治理形式参与。第二,从理论层面看,异质性程度大的群体的自治组织,不仅成功发起了集体行动,还进入到了第二阶段——利用自治组织形成的规范进行监督与惩罚,体现了自治组织可持续的制度性作用。因此,西头新村的实践可以作为一个准自然实验,供我们观察,异质性群体如何解决自组织治理中集体行动的可持续问题。与一般政府主导介入的城中村治理案例不同,西头新村公共品供给的集体行动,从产生到运行,始终没有外部主体介入,直到后期主管部门将其作为自治成功的典型予以推广复制、并给予一定的经费支持。因此可以确认,该案例是一次自发的、公共品合作治理的集体行动,且个体间一开始就处于非亲非故、异质性较大的陌生人社会状态。

(一)集体行动发起阶段的案例发展描述。西头新村地处深圳市龙华区民治街道办上芬社区。龙华区原隶属于宝安区,2011年成立龙华新区(功能区),2016年国务院批准设立龙华行政区,辖区面积102平方公里。全区常住人口约167万,实际管理人口超过300万,人口密度约3万人/平方公里,最大的社区达12万人/平方公里,基层治理的人力、物力和财力与人口规模不相匹配,存在“小马拉大车”的现象。

西头新村地块,原属村集体经济——西头社区股份合作公司所有,2002年打包给某地产商进行开发。地产商又将地块切割零售(并非是经过国有化、再市场化的土地,而是集体所有制的小产权土地),来自五湖四海的业主购得土地之后,各自在土地上盖房出租。形成了由138栋楼宇、约7160间出租屋、居住约2万外来人口的城中村,原名“宇丰城”,后改现名。建成初期,基本属于“三不管”的“插花地”,村内环境差、治安乱,业主租户上访不断。尽管出租屋租金相对周边地区更低,但是出租率也提不上去,极大地影响了业主们的主要收入来源。

地块原所有者——西头社区股份合作公司,村改居、集体经济股份制改革之后基本只负责原村民的公共事物,西头新村作为外来人口并非在基于地缘、亲缘关系的集体经济覆盖范围之内,因此一直没有设立物业管理机构。此间,有业主曾向负责开发地块的地产商和基层政府呼吁,但也没有回复。

2004年9月,在第一次的碰头会上,十多名业主不约而同地提出业主自治的想法。在之后召开的会议中,业主们慢慢协商、达成一致,决定申请组建业主自治委员会,代表全体建房业主自我管理大家的公共事物。

民治办事处根据当时西头新村的实际情况,对上述行动采取了默认和试试看的态度。2004年12月31日,全体建房业主大会召开。村内138栋楼宇的业主,“一栋一票”,严格按照选举章程,公开透明、投票选举12名业委会成员,正式成立西头新村业主自治委员会,制定完善了业委会章程和自治公约。规定,业委会成员不拿工资,因公因事产生的费用可享受合理补贴。

自 2005 年始，业委会按照 0.28 元/平米·月的标准收取物业费，近年来开始收取停车费。每月的收支及时张榜公布：每月收入 4 万多元，用于聘请 10 余名保安和保洁工人，安装监控和广播系统。短期内成效显著，警情逐年下降，特别是原先屡见不鲜的入室盗窃，直线下降（见表 1）。

表 1 西头新村 2008—2013 年总警情和入室盗窃案件数

年份	2008	2009	2010	2011	2012	2013
总警情	292	320	300	276	255	80
刑事警情	26	—	—	—	—	6
入室盗窃案	—	28	22	20	4	2

资料来源：作者调研统计和相关报道，其中“—”表示数据缺失。

值得注意的是，在村内的一个显著位置，围绕一棵大树撑起了一个简易建筑——“恳谈亭”，一开始参与发起的业主便是自发在这棵树下商议公共事物，延续至今。后来，业委会还提供报刊和茶水，欢迎其他业主和租住人员表达“心声”，提出自己对家园建设管理的意见、建议。此外，还牵头注册深圳市龙华区民治西头新村社区共建促进会（共促会），实行单位会员制和个人会员制，社区内的各类组织、单位和个人都可以参与进来。

2014 年新设立的西头新村党支部和社区共促会一起，将大树下的“恳谈亭”改建升级为居民议事会场所，为社区居民参与公共事务提供了开放包容的平台。

经过逐年一系列的自治理举措，原来“脏乱差”的问题村——西头新村，逐渐往“净宁畅”的先进村演变，作为“政府治理和社会自我调节、居民自治良性互动”的样板，也成为广东省内外同类社区治理学习的模范村。

（二）集体行动维持阶段的监督惩罚实施。在上述自治理过程中的 2014 年，出现了拖欠物业费的问题。当事人钟某，是 138 名业主之一，名下的一栋楼建筑面积约为 1 100 平米，按照公约商定的当年约 0.3 元/平米·月的标准，钟某应缴纳 340 元/月的物业费。事实是，自该年 1 月至 6 月，钟某共拖欠费用 2 040 元。但此间钟某依然享受了业委会组织提供的公共物业服务。

对应到上一部分的模型设定，钟某可以被记为选择搭便车策略的背叛者，因为这种行为危害到了集体行动的可持续性，如果其他业主纷纷效仿，那么集体行动导致的公共服务供给将出现枯竭的可能。业委会注册成立的企业实体——新丰物业管理有限公司<sup>①</sup>多次联系钟某催缴物业费未果。业委会在充分讨论之后决定，以新丰物业管理有限公司的名义将钟某告上法庭，通过法律途径解决其拖欠物业费问题。法院判决钟某补交拖欠的物业费及产生的其他相关费用共计 4 575.1 元。

我们认为，钟某付出的代价不止  $s_1=4\ 575.1$  元，除了这部分显性的拖欠物业费、诉讼等成本，还应包括非货币成本  $s_2$  部分，也就是因为受到群体内部实施的惩罚而付出的无形代价，如法院判决结果及相关费用清单会在恳谈亭公开给所有居民，无形中对钟某的行为形成压力，也即来自其他居民道义上谴责的压力。<sup>②</sup>

① 在集体所有制土地上建造的物业属于小产权房，其房主不能称为受法律保护的业主，因而成立多年的业主自治委员会并不具备法律地位，只能以企业实体注册的方式代为行使业委会的职责。

② “诉诸法院”本身是一项集体行动，该项行动的行为主体是“新丰物业管理有限公司（即私建住宅业主自治委员会不具备法律地位，只能注册企业实体）”。“诉诸法院”是业主委员会集体协商达成的对该拖欠物业费的业主进行惩罚的一个选项，这等于集体达成一致选用某种具体的惩罚方式（罚款、报警报案或聘请律师诉诸法院等司法途径）进行惩罚，因此，本文将其视为“内生惩罚”，当然有如审稿人所说“不纯粹”、不理想的一面，但现实中这已经是一个能够观测到的较为理想的样本案例了。



综合理论和现实可见,集体行动过程中出现搭便车、退出合作的机会主义行为并不意外,但是,是否存在内部的监督和惩罚机制,这对集体行动能否持续有效运作具有关键作用。能够实时有效地实施内部监督和惩罚,并由此产生激励,这会使背叛者在理性选择的前提下放弃搭便车、重新选择合作策略。同时,这一过程也会对其他可能想要选择搭便车策略的个体构成威慑和警示,进而使集体行动得以维持。

结论 1: 在集体行动维持阶段,如果只有制度安排,而没有有效监督和可置信的惩罚,即出现模型中  $s=d=0$  的情况时,非亲缘、非地缘的陌生人群体维持集体行动的可能性极低。<sup>①</sup>

Nowak(2011)关于合作发生的五种机制(亲缘选择、直接互惠、间接互惠、空间博弈和群体选择)认为,合作首先发生于亲缘或具有较强身份认同的群体内。同时,对非亲缘、非地缘关系的个体之间,因为利他等社会偏好的普遍存在,也可以产生合作。那么,在合作维持的过程当中,个体的理性选择,不应局限于物质层面的收益成本分析,而应扩展到非物质、非货币层面,如声誉、群体排斥等。

如何让合作成为个体的理性选择和集体的理性选择是制度安排的关键。首先,将以社会偏好为代表的广义效用引入集体行动加以考察,更容易寻找达成自治理组织集体行动所需满足的条件。其次,变一次性博弈为重复博弈,也为扭转合作式微的趋势提供了可能——改变后续互动发生的可能性与成本收益比之间的大小关系,从而使在非亲缘的异质群体内获得合作解。

表 2 统计数据显示,138 名业主来源地不一、非亲非故,文化背景差异大。较为接近理论模型按照“不同动机划分三类个体”的区分——满足了异质性程度大的这一假设。再结合社会偏好理论和个体间互动频率高,村内村外边界界定清晰、小范围的业主自治公约为集体选择提供了组织制度保障、存在有效监督和惩罚、冲突解决机制和个体对自治组织业委会的认可(Ostrom, 1990)。

实验研究(连洪泉等,2013)也认为,当在个体社会偏好异质性程度大的情形下,惩罚机制可以解决搭便车难题。此时,当个体选择搭便车、而不愿为提供公共品付出成本  $c$  时,即  $s=d=0$  条件不成立,此时对个体实施惩罚才能维持集体行动中的合作水平。

结论 2: 在集体行动维持过程中,引入有效监督和可置信的惩罚(即  $s \neq 0$  且  $d \neq 0$ )时,群体内部形成监督、惩罚,促使个体形成“一旦背叛便会受到来自内部的惩罚并付出代价”的行为预期,可以减少甚至杜绝搭便车等机会主义行为的发生,使集体行动得以持续。

一般而言,可以将集体行动过程分为两个阶段:发起阶段和维持阶段。

表 2 西头新村(宇丰城)业主来源地构成

来源地	人数	来源地	人数
广东省内	108	广东省外	30
茂名	48	福建	12
梅州	24	广西	3
湛江	15	香港	2
潮汕	8	东北	2
清远	6	甘肃	1
肇庆	3	湖南	1
阳江	2	四川	1
增城	2	江西	1
深圳	2	海南	1
-	-	其他	6
合计	138		

资料来源:作者调研统计。注:由于西头新村房屋属于私建民宅,无房产证,因此不符合《物权法》等法律法规中业主的概念。而在正式文件和日常用语中,一直以“业主”称呼。

<sup>①</sup> 结论 1 对应模型部分的“假说 1”,因为是演化博弈模型,其结果是经过相对经典的博弈和更为漫长的演化过程收敛呈现的。在均衡收敛过程中,会反复,即合作者、背叛者和惩罚者之间此消彼长、势均力敌甚至急转直下都是可能的,但是最后都会收敛到理论预测的这个解,即无法达成集体行动。本文观测案例也只是一个阶段(2002-2017 年),严格来说只是模型演绎的整个尘埃落定过程中的一部分。

在发起阶段,首先需要有先行动的发起者。他们可以是群体内的社会、经济或政治精英(能人),依赖公共事物获得的效用相对一般个体更高。围绕发起者逐渐形成小群体,也就是制度变迁意义上的第一行动团体或关键群体(*key mass*)。西头新村,一开始在恳谈亭聚集商议自治事宜的12名业主就是关键群体(第一次12名业主参会、第二次30名、第三次60人)。此时,形成公开的信息平台,线上线下的信息发布和交流互动,以此扩散,完成制度安排的供给(业委会章程和业主自治公约)。

在实现制度供给之后,进入本文关注的维持阶段,即群体内部依照自治公约形成稳定的行为预期,此时的信息平台依然发挥重要作用。依模型设定,个体可行策略集包括合作与背叛,合作需要付出成本 $c$ ,则所有业主因此获得的人均收益 $b/n$ ,若满足假设的 $b > c > b/n$ ,则业主的占优策略为搭便车,走向传统理论预测的 $n$ 人囚徒困境。

此时,群体内部存不存在有效监督、有没有业主组织(机制)实施惩罚,是决定集体行动走向的关键。考虑惩罚的经济性:惩罚是有成本的,同时,惩罚实施也具有规模效应——惩罚者越多(或实施惩罚越多),越容易遏制背叛行为。

在本案例的集体行动维持阶段,出现了以钟某为典型代表的拒交物业费、退出合作个案,根据上述案例发生经过,钟某因此受到惩罚付出的货币成本为 $s_1=4\ 575.1$ 元。此前若钟某选择合作供给而非背叛策略,则其应付公共品供给的成本为 $c=2\ 040$ 元,显见 $s_1 > c$ 。

与外生惩罚不同的是,决议对簿公堂前后,集体行动的组织——业委会判决前后经过了充分的交流沟通,是对其组织动员及应变能力的一次考验,这一事件全程通过小区信息平台公开透明——私人信息成为共同知识,自治公约这一规范在整个社区的居民群体内获得了认可和监督执行。

综上,钟某承担的成本由两部分组成:货币成本 $s_1$ 和非货币损失 $s_2$ ,总成本 $s=s_1+s_2$ 远大于公共品合作供给成本 $c$ ——选择背叛并不是其理性的占优策略。

所以,西头新村集体行动在实施监督惩罚后能得以持续,有赖于有效监督和可置信的内生惩罚,也就是模型设定作为博弈一方的合作者实施惩罚。在监督惩罚过程中,业委会、共促会和居民议事会等组织形式得到了居民认可,也在解决冲突、健全规范等治理过程中积累了经验,关键群体秉持“自治从信任出发、信任从公心而来”(业委会成员老周总结)的治理理念,使上述组织常态化、稳定化,进而对集体行动的可持续起到了关键作用。

## 五、结论与可能的政策含义

本文针对一类非市场、非政府的自治理集体行动,通过模型推演和深圳西头新村外来人口治理公共事物的案例分析,给出了解决集体行动的困境、维持合作治理的一种可能方案。

理论研究上,基于兰彻斯特法则拓展的演化博弈模型,可以用策略表征个体类型,找到不同行为动机的个体间集体行动的成功条件,即个体理性与集体理性相符的均衡策略。发起集体行动、完成制度供给之后,在集体行动维持阶段有以下两种情况:第一,没有有效监督和可置信的惩罚的条件( $s=d=0$ )下,非亲缘、非地缘关系的外来人口群体很难维持集体行动,集体物品枯竭。第二,内部形成有效监督和可置信的惩罚的条件( $s \neq 0$ 且 $d \neq 0$ )下,实施惩罚的实际行动,会使被惩罚者付出代价,群体内部个体就形成“一旦背叛便会受到来自内部的惩罚并付出代价”的行为预期,合作供给公共品的集体行动得以持续。

综之,本文初步回答了“有无惩罚选项会对集体行动的成败产生什么影响?有了惩罚选项之后能否发生协商一致的集体惩罚?”如果把惩罚作为一种新的公共品处理的话,是否实施惩罚就如传统理论中“合作还是背叛的社会困境(二阶困境)”。

政策实践中,通过取自深圳这一典型移民城市的案例,检验了此类公共品供给中内生监督和惩罚机制对集体行动维持的影响作用,丰富了集体内部自我治理的理论和实践,对提升城市治理能力的现代化有所裨益。

在快速城市化、城市化又滞后于工业化的中国情境下,如何解决异质性群体的集体行动困境问题,是对治理体系和治理能力现代化的挑战或考验。根据本文的案例,作为政府决策部门,应该掌握利用基层非正式制度来实现科层组织所不具备的灵活性,激励居民主动参与管理家门口的公共事物治理;同时,充分发挥党建引领、注重党组织在城市治理中的作用,因地制宜,总结出更多的基层治理工作的经验,让社区治理更接地气、更有效率。既减轻政府负担,又整合、融合了民间资源,形成“政府治理和社会自我调节、居民自治良性互动”的生动局面。

利用基层民间资源,重要一点就是发挥好方方面面的精英人物(能人)的关键作用,党建引领,支持精英、搭建居民议事渠道,实现多中心治理。特别是在集体行动已经发生之后的维持阶段,因“内生溢价”的存在,群体内部行之有效的赏罚规范必须是协商一致、符合各方稳定预期的可行方法。这种方法并非是政府部门或其他主体在缺乏对群体状况了解和掌握时所直接给定的,而是在社区群体内部共享信念体系意义上的一种制度。

#### 主要参考文献:

- [1][美]曼瑟尔·奥尔森.集体行动的逻辑[M].陈郁,郭宇峰,李崇新,译.上海:格致出版社,2014.
- [2]连洪泉,周业安.异质性和公共合作:调查和实验证据[J].经济学动态,2015,(9):131-144.
- [3]连洪泉,周业安,陈叶烽,等.信息公开、群体选择和公共品自愿供给[J].世界经济,2015,(12):159-188.
- [4]连洪泉,周业安,左聪颖,等.惩罚机制真能解决搭便车难题吗?——基于动态公共品实验的证据[J].管理世界,2013,(4):69-81.
- [5][美]塞缪尔·鲍尔斯,赫伯特·金迪斯.合作的物种——人类的互惠性及其演化[M].张弘,译.杭州:浙江大学出版社,2015.
- [6]闫佳,章平.集中式惩罚与公共品自愿供给:一项实验研究[J].经济学动态,2016,(6):85-99.
- [7]周其仁.城乡中国[M].北京:中信出版社,2017.
- [8]周业安,连洪泉,陈叶烽,等.社会角色、个体异质性和公共品自愿供给[J].经济研究,2013,(1):123-136.
- [9]周晔馨,涂勤,胡必亮.惩罚、社会资本与条件合作——基于传统实验和人为田野实验的对比研究[J].经济研究,2014,(10):125-138.
- [10]Adams E S, Mesterton-Gibbons M. Lanchester's attrition models and fights among social animals[J]. *Behavioral Ecology*, 2003, 14(5): 719-723.
- [11]Baldwin K, Huber J D. Economic versus cultural differences: Forms of ethnic diversity and public goods provision[J]. *American Political Science Review*, 2010, 104(4): 644-662.
- [12]Balliet D, Mulder L B, Van Lange P A M. Reward, punishment, and cooperation: A meta-Analysis[J]. *Psychological Bulletin*, 2011, 137(4): 594-615.
- [13]Burlando R M, Guala F. Heterogeneous agents in public goods experiments[J]. *Experimental Economics*, 2005, 8(1): 35-54.

- [14]Decker T, Stiehler A, Strobel M. A comparison of punishment rules in repeated public good games: An experimental study[J]. *Journal of Conflict Resolution*, 2003, 47(6): 751–772.
- [15]Fehr E, Gächter S. Cooperation and punishment in public goods experiments[J]. *American Economic Review*, 2000, 90(4): 980–994.
- [16]Fischbacher U, Gächter S. Social preferences, beliefs, and the dynamics of free riding in public goods experiments[J]. *American Economic Review*, 2010, 100(1): 541–556.
- [17]Fischbacher U, Gächter S, Fehr E. Are people conditionally cooperative? Evidence from a public goods experiment[J]. *Economics Letters*, 2001, 71(3): 397–404.
- [18]Gächter S, Thöni C. Social learning and voluntary cooperation among like-minded people[J]. *Journal of the European Economic Association*, 2005, 3(2–3): 303–314.
- [19]Houser D, Xiao E T, McCabe K, et al. When punishment fails: Research on sanctions, intentions and non-cooperation[J]. *Games and Economic Behavior*, 2008, 62(2): 509–532.
- [20]Hwang S H. Larger groups may alleviate collective action problems[R]. Working Papers 2009-05, 2009.
- [21]Hyde S D. Catch us if you can: Election monitoring and international norm diffusion[J]. *American Journal of Political Science*, 2011, 55(2): 356–369.
- [22]Johnson D D P, MacKay N J. Fight the power: Lanchester’s laws of combat in human evolution[J]. *Evolution and Human Behavior*, 2015, 36(2): 152–163.
- [23]Masclot D, Noussair C, Tucker S, et al. Monetary and nonmonetary punishment in the voluntary contributions mechanism[J]. *American Economic Review*, 2003, 93(1): 366–380.
- [24]Mathew S, Boyd R. Punishment sustains large-scale cooperation in prestate warfare[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(28): 11375–11380.
- [25]Ostrom E. *Governing the commons: The evolution of institutions for collective action*[M]. Cambridge: Cambridge University Press, 1990.
- [26]Ostrom E. Beyond markets and states: Polycentric governance of complex economic systems[J]. *American Economic Review*, 2010, 100(3): 641–672.
- [27]Reuben E, Riedl A. Public goods provision and sanctioning in privileged groups[J]. *Journal of Conflict Resolution*, 2009, 53(1): 72–93.
- [28]Ruttan L M. Economic heterogeneity and the commons: Effects on collective action and collective goods provisioning[J]. *World Development*, 2008, 36(5): 969–985.
- [29]Sutter M, Haigner S, Kocher M G. Choosing the carrot or the stick? Endogenous institutional choice in social dilemma situations[J]. *Review of Economic Studies*, 2010, 77(4): 1540–1566.
- [30]Xiao E T. Punishment, social norms, and cooperation[A]. Teitelbaum J C, Zeiler K. *Research handbook on behavioral law and economics*[M]. Northampton, Mass: Edward Elgar Publishing, 2018.
- [31]Xiao E T, Houser D. Punish in public[J]. *Journal of Public Economics*, 2011, 95(7-8): 1006–1017.
- [32]Yang W, Liu W, Viña A, et al. Nonlinear effects of group size on collective action and resource outcomes[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(27): 10916–10921.
- [33]Zhang X Q, Zhu F. Group size and incentives to contribute: A natural experiment at Chinese Wikipedia[J]. *American Economic Review*, 2011, 101(4): 1601–1615.

# How to Solve the Dilemma of Collective Actions among Heterogeneous Individuals by Endogenous Punishment? A Theoretical Model and Case Study

Zhang Ping, Liu Qichao

(China Center for Special Economic Zone Research, Shenzhen University, Shenzhen 518060, China)

**Summary:** Cooperation and coordination are the basic problems of human society. The classical theory predicts that, as individuals with complex and different motivations, collective actions on the provision of the commons often fall into the social dilemma, in which individual rationality leads to collective irrationality. Evidence from the long-term observation of an urbanizing village in transitional China suggests that heterogeneous groups can also achieve efficient cooperation and self-governance.

We focus upon the problems that, in the case of different groups with heterogeneous individuals, from the formation of groups to collective actions taking place, how can internal monitoring and punishing strategies be selected to sustain collective actions by preventing free riding and other opportunistic behaviors?

The main work of this paper is as follows: (1) It expands Lanchester's Law to describe different behavioral motivations of individuals and investigate the conditions for the occurrence and maintenance of collective cooperation; (2) It introduces endogenous punishment as a credible threat and investigates its impact on the provision of public goods (non-market and non-government), which provides a possible solution to the traditional dilemma of collective actions; (3) Based on the analysis of a real case from Shenzhen, it empirically examines the effect of endogenous punishment on the spontaneous formation of collective actions in the provision of public goods, which enriches the theory and practice of self-governance within the community.

This paper makes the following contributions to the literature: First, we employ the general classification and measurement of individual heterogeneity from a laboratory experiment (Fischbacher and Gächter, 2010), which reflects implicit preferences through explicit behavior. According to different motivations, groups are divided into cooperators, free-riders and punishers, and individual types can be revealed by their strategies. Our evolutionary model based on Lanchester's Law, depicts how individuals in a group take punitive measures against free riding. Then we investigate the influence of endogenous punishment system on collective actions in a long-term repeated game. This paper also provides a new perspective to understand the emergence of cooperation in collective actions and gives a case study as empirical evidence.

The results show the conditions for the success of collective actions among individuals with different behavioral motivations, that is, the equilibrium satisfies both of the individual and collective rationality. After initiating collective actions and completing the supply of the system, the maintenance stage of collective actions is equally important. First, in absence of effective supervision and credible punishment, it is difficult for the population who are non-kinship or non-geographical relationship to promote collective actions, and then public goods will be out of provision. Second, in the presence of effective supervision and credible punishment, individuals within the group form the behavioral expectation that "once betrayed, they will be punished and pay the price", and the collective actions of cooperation on the provision of public goods can be sustained.

**Key words:** cooperative governance; collective actions; heterogeneity; endogenous punishment

(责任编辑 许 柏)