

· 学界视点 ·

基于善意本性的公共品自愿供给与公平分配

——中美行为经济实验研究

柴 盈 王 翊 Marco A. Janssen*

摘 要：主流经济学的理性思维与现实中人的思维存在偏差。本文设计了一系列行为经济实验（信任、囚徒困境、公共品、独裁和最后通牒实验）并在中美两国大学课堂上实施。研究表明：人们天生具有善意本质和社会性思维，能够自愿供给公共品，偏好经济利益的公平分配；而在信任度、合作偏好、合作水平和公平偏好的量化方面，美方都表现出更高（强）的特征。本文建议政策制定者有必要充分了解人们的社会性思维及相应的微观行为，政策的重点有必要充分考虑到这些思维、能力和行动，以及为人们自愿解决集体行动问题提供信息平台。

关键词：善意 社会性思维 自愿供给 公平分配 行为经济实验

JEL 分类号：C92, H41, P43 **中图分类号：**F091.349

文献标识码：A **文章编号：**1000-6249 (2015) 10-099-16

一、引言

科学地制定经济政策的重要理论基础，是了解人们的真实性思维。自 20 世纪 80 年代以来，实验经济学家创建了一系列行为实验，用于探讨人们的真实思维，如 Berg 等（1995）设计的信任实验、Marwell 与 Ames（1979）的公共品实验、Guth 等（1982）的最后通牒实验以及 Kahneman 等（1986）的独裁实验，他们发现现实中人们行动的表现，并非只有顾及自己利益的理性思维，而是更多地表现出涉他偏好、互惠决策和条件性合作行为等特征，即人们在行动前的思维往往受到社会因素的影响。之后，这些经典实验在学术界得到广泛采用，带动了从微观层面探索公共事务领域人们的思维问题的热潮；同时，由于这批研究成果的卓越理论贡献，大

* 柴盈，广东财经大学国民经济研究中心研究员，硕士生导师，E-mail: chaiking19@163.com，通讯地址：广州市海珠区赤沙路 21 号，邮编：510320；王翊，广东财经大学经济贸易学院教授，E-mail: wangy969@163.com；Marco A. Janssen，亚利桑那州立大学可持续学院教授，E-mail: Marco.Janssen@asu.edu。感谢匿名审稿专家的意见，作者文责自负。

本研究受到国家自然科学基金项目（71303062），教育部人文社会科学研究规划基金项目（13YJA790119），广东省高等学校优秀青年教师培养计划资助（Yq2013081）和广东高校人文社科重点研究基地重大项目（2012JDXM_0024）的资助。

都发表在《Science》、《American Economic Review》、《Quarterly Journal of Economics》等国际顶级刊物上,被国内外广泛视为探究人们在社会经济活动中的思维和行动的典范。总体来说,这些研究初步证实了现实中的人们的偏好和思维与主流经济学的完全理性不同,而是介于完全理性与完全利他之间(Henrich等,2010a)。具体的代表性研究有,Andreoni(1995)以美国大学生为研究对象进行公共品实验,研究表明人们具有合作的偏好,愿意贡献初始禀赋的33%来生产公共品;而Karlan(2005)在秘鲁进行的公共品田野实验表明,当地人的思考方式具有涉他性,在决策时会考虑他人的行动,愿意贡献的份额高达禀赋的81%;Henrich等(2006)对5大洲的15个小型社会中人们进行了最后通牒实验研究,他们发现当存在同行惩罚机制的条件时人们偏好的分配意愿是自己禀赋的25%–51%。

在这些经典研究的基础上,有研究也发现受社会因素影响的人们的真实性思维能够引导人们走出集体行动困境,实现公共品自愿供给与公平分配。具体来讲,其功效主要体现于公共品自愿贡献机制(VCM, voluntary contribution mechanism),典型的成功经验如印度微金融行业,客户之间自愿维系的信用与互惠关系形成了良好的公共秩序、降低了违约风险,使得小额贷款服务实现了有效提供(Feigenberg等,2013)。而在公平分配方面,其主要体现在公共资源治理领域,成功地破解了“公地悲剧”的难题,验证了用户能够自我组织实现资源的长期持续利用(Ostrom, 1990; Janssen等, 2010);解释了市场经济发达地区的人们较为追求公平的目标、甚至愿意与陌生人进行交易以获取公平的收益(Henrich等, 2001),同时也解释了近年来我国国民对收入公平分配的强烈诉求的原因(陈叶烽等, 2011; 周业安等, 2013)。

显然,这些研究在明确人们的真实性思维上具有重要的贡献,也科学地精炼出人们具有信任、互惠、利他和条件性合作等社会性偏好(陈叶烽, 2009; 周业安, 2015),对理解人们的社会经济行为具有重大意义。但是这些研究尚存在如下两方面缺陷:缺乏对真实性思维一致的界定,以及忽略了真实性思维内在的逻辑关系。对于前者,已有研究主要描述了其关键表征,如社会偏好、互惠、利他、涉他等,没有清晰界定和深入分析这些偏好表征的本质属性与内涵,也就是说人们在从事集体行动时的根本思维性质尚不明确。对于后者,已有的一次性行为实验研究大都是独立进行,探讨真实性思维的某一方面,很少同时开展多个实验研究共同性,缺乏对它们之间逻辑关系的分析,割裂了真实性思维与公共品自愿供给和公平分配成果这两个行为结果之间的联系,即没有分析人们的真实性思维如何作用于经济生产和利益分配活动。这两个问题是解释人们思维、行动和经济运行机理的关键,因此,非常有必要对此进行深入分析。

针对这些不足之处,本文将采用来自中美双方的系列行为实验对其进行逐一探究。需要说明的是,早期在实验设计阶段,我们意在找出人们真实思维的一个共同特征,以统一术语,当时通过实验观察,一致认为人们在行动时的思考方式受到他人和社会环境的影响。幸运的是,在实验分析阶段,恰逢世界银行发布《2015年世界发展报告》,高屋建瓴地指出这是一种社会性思维(thinking socially),其研究对象与结论与我们的实验不谋而合。报告明确指出真实经济活动中人们的思维具有社会性,即人们偏好信任、互惠、利他与平等,

与主流经济学理性行为人的假设有些差异；同时指出，社会性思维能够引导人们有效地自愿供给公共物品，并且偏好公平分配这些物品。报告也阐述了社会性思维的根源是人们行动时会考虑他人的决策，会受到文化、社会规范和身份等因素的影响。基于此，报告认为，新时期的政策制定非常有必要洞悉这个理论基础，否则，不当的政策不仅会挤出人们自愿的投入、造成低效的公共品生产与分配，而且会产生贫困恶化、风险增加、资源浪费、发展退步等等重大问题（World Bank，2015）。

有鉴于此，本文将采用社会性思维这一表达方式来继续行文分析。首先，明确社会性思维的属性包括本质与表征两个方面，一方面，善意是社会性思维的本质，借此延伸了社会性思维这一理念的内涵，另一方面，信任、互惠和合作等社会性偏好是社会性思维的表征，厘清它们之间存在的递进性逻辑关系；其次，明晰基本要素、构建理论框架，梳理五个行为实验与社会性偏好之间的对应关系；再次，执行实验，并观察善意、社会性思维与公共品自愿供给与公平分配之间的因果关系；最后，提出相关的政策建议。

二、社会性思维的属性与相关假设

主流经济学的理性思维，以理性人假设作为分析人们在公共经济活动中行为的逻辑起点，但是这种理性的思维模式常常推演出与现实有些偏差的论断，日渐受到行为经济学的强烈挑战。理性假设将个人抽象化为同质化的、冷静且计算的自利行动者，免疫于他人和外部环境的影响，在面临公共事务时，搭便车问题无处不在，缺乏自愿合作解决的可能。但真实的情况是，人们的行为常常会表现出社会性思维，不完全符合理性思维和搭便车难题的预测，在决策时深受社会环境、社会规范和文化的影响。实际上，人们的初始理念具有善意或者恶意的差别，会产生积极或者消极两种不同的集体后果，例如公共品自愿供给与腐败。而本文意在考察人们如何通过社会性思维走出公共事务活动中的困境、实现集体利益，因此我们集中探讨社会性思维的积极性的一面，以善意作为假设的出发点。

为了说明社会性思维的内含属性，我们先定义如下概念：

（1）善意：经济活动中人们普遍怀有希望他人好的心理倾向和伦理认同，人们既有潜意识善举的内在动机，又有伦理认同的外在强化。

（2）信任：经济活动中各关联方对对方怀有善意的信心与忠诚；人们对他人的信任程度，受个体差异的影响，也受外在社会、经济、文化因素及对方行为的影响。

（3）互惠：经济活动中各关联方愿意采取相互有利的合作行为；互惠建立在相互信任基础之上，因而表现出强烈的正反馈效应：对我善者以善待之，对我恶者以恶报之。表现为“投桃报李”，或“冤冤相报”的行为模式。

（4）公平：经济活动中利益相关各方之间利益分配结果的平等程度。

沿着行为经济学的思路，本文认为，人们在公共事务决策中表现出的社会性思维包括如下两个方面的属性：一是作为社会性思维本质的善意。善意是个体与生具有的品性，是人性的重

要基础，并在社会过程中不断得到加强。与人为善是人与人之间行为关系的普遍准则。因此，在人们的经济行为中，必然反映出人性的这种影响。

二是作为社会性思维表征的社会性偏好，如信任、互惠、合作和公平等行为方式和特征，是人性内在的善意本性的外在表达。不同的社会偏好是善意本性在不同经济情景下的自然反映，善意与社会偏好两者之间具有内在的一致性。因此，如果个人是“性本善”的社会人，那么在公共品供给过程中，人们相互之间必然具有相互信任以达成合作、自愿供给的人性基础；对于生产结果的分配，他们追求的目标就不仅是效率，也内含公平。

基于上述认识，本文提出如下假设：

(1) 社会人假设：由于人性的善意属性，人们的经济行为在某种程度上受互惠的社会性思维支配，在不同类型或不同阶段的经济活动中，表现出倾向于信任、合作、公平等社会偏好的特征。这一基础假设在生产、分配等不同类型（阶段）的经济活动有不同的特点。根据这些特征，可以进一步派生出如下假设。

(2) 合作生产假设：人们的经济行为并非完全受理性的支配，也受非理性的善意的影响；心怀善意促使人们之间相互信任、互惠合作。信任程度越高，互惠程度越大，合作水平越高，人们得到的共同经济效益越高。

(3) 公平分配假设：善意使人们在意他人的成本与收益，从而对公共物品的分配具有平等化的偏好，因而注重公平结果。在一定条件下人们对公平的追求胜过效率，遇到不公平待遇时甚至愿意牺牲自己的福利而实施惩罚。

(4) 公平与合作互动假设：公平与信任具有内在的联系，它们是个体善意在经济活动不同阶段的表现。信任表现为合作策略的选择，是善意在生产过程中的反映；公平表现为利益均享，是善意在分配结果上的表达。善意大，则信任程度强，则公平程度高。

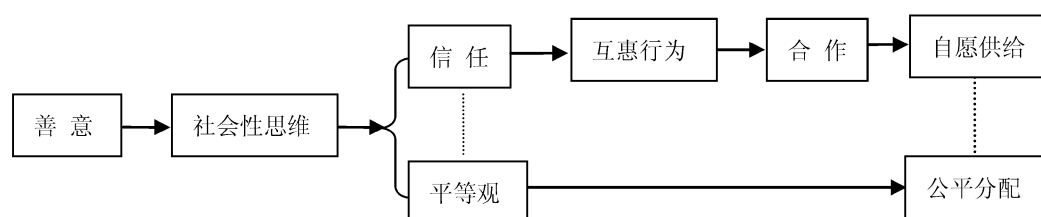


图 1. a 本文的分析框架

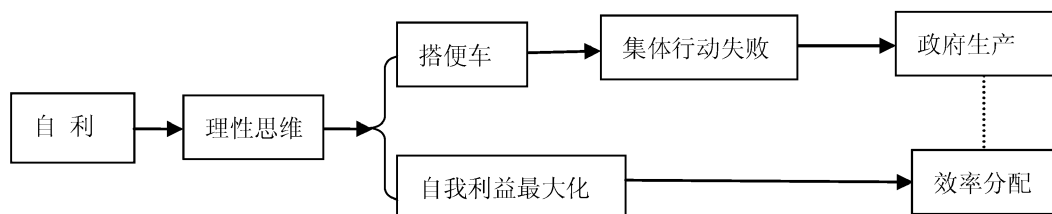


图 1. b 主流经济学的分析框架

根据上述假设,围绕从生产到分配的经济活动过程,可以得出如下的合理推论:心怀善意的人们在面对公共产品的供给和分配问题时,能够彼此互信,愿意采取互惠的合作策略,实现公共品的供给和共享。这个推论建立了上述不同类型(生产和分配)经济活动中不同行为假设之间的逻辑联系,同时也构建了利用不同的经济实验检验上述假设及其相互联系的分析框架(见图 1. a)。显然,这一框架与主流经济学的分析框架明显不同(见图 1. b)。

三、实验设计

为了检验上述假设,本文将采用经济学领域研究偏好问题的五个经典的行为实验来进行研究,分别是供给领域的信任实验、囚徒困境实验和公共品实验,分配领域的独裁实验和最后通牒实验。我们保持了这些实验的标准内涵,而根据本研究需要对禀赋、收益率和经济激励等参数进行了修改,主体设计工作由亚利桑那州立大学 Marco A. Janssen 教授完成,主要设计过程包括如下部分。

(一) 实验说明

五个实验相继在美国和中国开展。2013 年 10 月和 2014 年 10 月,我们先后在美国亚利桑那州立大学和广东财经大学招募实验对象,两校分别有 88 人和 130 人、共计 218 人参加了实验。实验对象分为中方与美方 2 组,每组全部实验对象同时集中在一间教室参加实验,这样做的目的是尽可能扩大实验对象之间随机分组的范畴,创造模拟陌生人之间互动的模式。每组实验对象依次参加信任、囚徒困境、公共品、最后通牒和独裁游戏,但是每个游戏中实验对象之间的分组都是不重复的随机组合。这些实验对象是来自一年级的不同专业的新生,将分别扮演不同身份连续进行 5 个(4 个双人一组和 1 个五人一组)实验,依次为每个实验的决策作答,整个实验过程以传统的纸笔方式进行,中方和美方的实验语言分别采用中文和英文。由于在中国进行的实验需要翻译为中文,为确保国内学生(实验对象)准确理解实验说明、随机组队的含义以及实验中收益支付的计算,以确保他们在实验中更好地决策,我们先选择一个 30 人的小班进行预实验,根据他们反馈的信息,将难以理解的术语加以详细解释,转化为易懂的中文语境,然后正式在国内实施实验。

我们要求实验对象在实验期间不准许交谈,只能独立作答,如果实验对象有任何不明白之处,只需要举手,实验员就会私下予以回答;初始禀赋与初始收益都是以筹码形式进行表示,最终收益将转化为平时成绩加分,1 个筹码等于增加 1 分的平时成绩^①;每个实验对象都需要参加全部实验,而且同时扮演两个角色;在每个决策做出之后,都需要附上说明,解释这样选择的原因;最终,对于每个实验对象来说,五项实验中只有一项可被抽选为计分,因此,这就要

^① 设计报酬机制的原则是,能够使实验对象的真实偏好与行为之间存在相容性,本研究的实验对象是大学生,平时成绩加分对他们来说是一种有效的收益表达方式,这也在周业安和宋紫峰(2008)的研究中得到验证,即采用的平时成绩的方式,主要为了剔除由于初始禀赋、收益分配不均引起的不适宜的激励效果,能够保证足够的、公平的激励。

求实验对象必须认真对待全部实验；每项实验的最高收益是 10 个筹码；实验设置中的分组抽选和计分抽选，都是事后由电脑随机匹配完成；5 项实验共需约 30 分钟。整个实验结束，每位实验对象因被抽中的身份不同以及实验收益不同，而最终收益在 0 - 10 分之间。

（二）实验设置

1. 供给实验

供给实验可分为两类，一类是信任与信任度实验，检验经济活动中人们对彼此的信任并衡量信任程度和合作可能性，包括信任实验和囚徒困境实验；另一类是公共品实验，检验相互信任的人们之间合作程度以及合作生产公共品的水平。

实验 1：信任实验（又称投资实验）。在该实验中，一个实验对象和另外一方随机地组成一组，分为实验对象 A 和 B；A 与 B 分别有一次决策机会，分为两个阶段进行。首先，在第一阶段，A 会从实验员处获得 3 个筹码，B 没有筹码，A 的任务是对这 3 个筹码在 AB 之间进行分配，可以选择 0 - 3 之间的任何一个筹码（x）分配给 B，实验规则是 B 得到的分数将变成 3 倍（3x）；然后，在第二阶段，B 可以将获得的筹码（3x）中的一部分比如 y 个筹码送回给 A，也可以选择不返还而全部占为己有。只要 B 选择 y 个筹码送回给 A，实验即结束，双方的收益分别为 A 的（3 - x + y）和 B 的（3x - y）个筹码。A 的分配（或投资）行为衡量的是对 B 的信任水平，给 B 的筹码越多说明信任度越高，不信任到完全信任依次为给 B 的（0，1，2，3）个筹码；B 的行为衡量的是互惠或者条件性合作水平，B 得到的筹码在增值三倍之后可能是（0，3，6，9），即被信任程度依次增加，如果 B 返还给 A 的筹码也依次增加的话，那么说明 B 是典型的条件合作者、而且合作水平随着被信任程度的增加而增加。如果依据主流经济学的逆向推断，A 会假设在第二阶段 B 为了收益最大而不会返还筹码，那么 A 认为 B 不值得信赖，于是在第一阶段 A 不会给 B 筹码，结果 A 的收益只有初始禀赋 3 个筹码，B 的收益为 0。也就是说，主流经济学逻辑下 A 与 B 之间不存在信任，B 不会回馈 A 筹码、A 自然也不会给 B 筹码。

实验 2：囚徒困境实验。同实验 1，两个实验对象随机地组成一组，分为 A 和 B；两个实验对象都没有禀赋，实验规则是有甲和乙两个方案可供 A 与 B 同时选择，一旦做出决策，游戏便结束；如果两个实验对象都选择甲方案，则都会获得 3 个筹码；如果一个实验对象选择甲，另一个选择乙，那么选择乙的实验对象将得到 5 个筹码作为收益，选择甲的实验对象将得到 0 个筹码；如果两个实验对象都选择乙，双方都会得到 1 个筹码。A 与 B 的行为衡量的是对对方的信任情况，以及信任产生的收益。如果依据主流经济学的博弈分析和纳什均衡原理，占优策略是两个实验对象都选择乙方案，双方之间根本不存在信任，也不会有合作，A 与 B 的集体收益只能最低（2 个筹码）。

实验 3：公共品实验。与上述两个实验不同，该实验扩大了成员规模，将五个实验对象随机组成一队；实验员分别给每个实验对象 3 个筹码作为初始禀赋，他们可以共同投资一个项目，可以全部保留而一个都不投，也可以投入 3 个筹码中的一部分比如 x，那么第 i 个实验对象投入的筹码则是 x_i ；投资的集体收益是 5 个实验对象的投资总额乘以 3，即 $(\sum_{i=1}^5 3x_i)$ ，无论其他实

验对象是否投资以及投资多少, 收益都是由他们平均共享 ($0.2 * \sum_{i=1}^5 3x_i$); 每个实验对象的收益是没有用于投资的筹码加上项目的平均共享收益 ($3 - x_i + 0.2 * \sum_{i=1}^5 3x_i$)。每个实验对象的投资行为衡量的是合作的意愿, 贡献的筹码越多 (从 0 到 3 个)、说明合作意愿越强烈, 而贡献 0 个筹码既是搭便车行为; 同时, 集体收益衡量的是合作水平, 收益越多说明合作水平越高。依据主流经济学的逻辑, 每个实验对象都会搭便车, 不会投资任何筹码, 结果集体收益为 0, 每个实验对象的收益只是当初的禀赋, 即 3 个筹码。

2. 分配实验

分配实验有两个, 一个是最后通牒实验, 主要检验在存在同行惩罚的条件下, 人们的分配偏好。实际上, 上文的信任实验也可以用于考察分配行为, 它可以解释为存在同行奖励的条件下, 人们的分配偏好, 故此, 我们在本部分会对信任实验进行再次分析。另一个是独裁实验, 检验在没有任何惩罚的条件下人们的分配偏好。

实验 4: 最后通牒实验。同上文的两人实验类似, 两个实验对象被随机组成一队, 分为 A 和 B; 实验员仅仅给 A 分配 10 个筹码作为初始禀赋, 而没有给 B 任何筹码; 游戏规则分为两个阶段, 在第一阶段, A 要分配给 B 一些筹码, 可以是 10 个中的一部分比如 x 个, 也可以一个都不给; 在第二阶段, B 可以针对被分配的 x 筹码而选择“接受”或者“拒绝”, 如果选择了“接受”, 那么 A 的收益就是保留的筹码 ($10 - x$), B 的收益是被分配的筹码 (x), 但是, 若选择了“拒绝”, 那么 A 与 B 的收益都是 0。A 的行为考察的是公平分配意愿, 通常认为自己保留 5 个筹码、同时给对方 5 个筹码 5-5 方案是公平行为, 相应地自己保留 10 个筹码是绝对的自利行为, 而自己保留 0 个筹码是绝对利他行为; 而 B 的行为考察的是惩罚意愿, B 拒绝的筹码是采取惩罚而付出的代价。依据主流经济学的逆推逻辑, 既然 B 不会牺牲任何被赠与的机会, 因为拒绝就意味着失去利益, 因此会接受任何大于 0 的筹码, 而 A 则只需要提供最小的正数即可, 比如 1 个筹码, 结果, A 与 B 的收益分别为 9 与 1, 即存在惩罚会威慑 A 的完全自利分配方式。

另外, 对于信任实验, A 的行为可以看做是考察公平分配意愿, 相应地全部不分配和全部分配分别表示 A 的完全自利行为和完全利他行为; 而 B 的行为考察的是奖励意愿, B 返回给 A 的筹码是奖励的代价。

实验 5: 独裁实验。两个实验对象被随机组成一队, 分为 A 和 B; 实验员仅仅给 A 分配 10 个筹码作为初始禀赋, 而没有给 B 任何筹码; 游戏规则是, A 要分给 B 一些筹码, 可以是 10 个中的一部分比如 x 个, 也可以一个都不给, 不管 A 分配出去的筹码是多少, B 只能接受; 一旦 A 的分配决策完成, 实验便结束。A 的收益就是保留的筹码 ($10 - x$), B 的收益是被分配的筹码 (x)。与最后通牒实验类似, A 的行为是考察公平分配意愿, 5-5 方案是公平行为。可见, 独裁实验是最后通牒实验的上半部分, 也就是只有 A 有决策权时的分配选择, 检验的完全是 A 的偏好。依据主流经济学的逻辑, A 为了自己的利益最大化, 不会分配任何筹码给 B, A 偏好效率而不是公平。

四、实验结果

本部分将统计实验对象在五个实验中的决策行为，除了采用一般性统计方法评价中美双方实验对象的总体特征之外，我们还将采用 Wilcoxon 秩和检验方法分别考察中美双方实验结果的差异和每一方内部不同实验对象行为结果的差异，数据用 R 软件处理。

（一）信任实验中的信任与信任度

与主流经济学所预测的实验对象 A 不会信任 B、一个筹码都不会分给 B，100% 提供 0 个筹码的结果不同的是，中美双方实验结果都显示只有少数实验对象（中方 40%、美方不到 20%）才会表现得像理性经济人一样、完全不信任对方，大部分实验对象都提供了自身禀赋 33% - 100% 的筹码给对方，中方提供禀赋 33% 与 67% 的实验对象的比例分别为 24% 与 35%、美方则为 20% 与 62%，中方和美方都是只有 1 个实验对象提供了全部禀赋（即 3 个筹码）给对方（见图 2. a）。这些数据也提供信任程度在中美双方的比较，美方提供 67% 禀赋的实验对象比例远远高于中方，说明美方高信任度更为普遍一些，表 1 也说明了双方提供筹码方面的显著差异。我们问及这些赠送筹码给对方的实验对象 A 原因，多数人回答说，应该对他人友善一些，也相信对方会回馈一些筹码。实验结果说明，60% 以上的人们对他人无条件怀有善意，并对他人的善意怀有信心。

表 1 中方与美方实验结果的描述性统计与检验比较

序号	实验名称	中方与美方 实验对象的比较	均值（标准差）		Wilcoxon 秩和 检验结果（P 值）	差异显著与否 （P 值是否小于 0.05）
			中方	美方		
实验 1	信任实验	提供的筹码	0.95 (1)	1.47 (2)	0.0000	显著
		返回的筹码 (3x=3)	1.92 (2)	1.97 (2)	0.8227	不显著
		返回的筹码 (3x=6)	3.25 (3)	3.3 (3)	0.8321	不显著
		返回的筹码 (3x=9)	4.43 (4)	4.69 (5)	0.2124	不显著
实验 2	囚徒困境实验 [*]	选择的方案	0.55 (1)	0.3 (0)	0.0001	显著
实验 3	公共品实验	投资的筹码	2.01 (2)	2.37 (2)	0.0001	显著
实验 4	独裁实验	提供的筹码	5.76 (5)	5.12 (5)	0.1219	不显著
实验 5	最后通牒实验	提供的筹码	4.45 (5)	5.41 (5)	0.0000	显著
中方内部独裁与最后通牒实验提供的筹码					0.0000	显著
美方内部独裁与最后通牒实验提供的筹码					0.9190	不显著
中方内部独裁与信任实验提供的筹码					0.0000	显著
美方内部独裁与信任实验提供的筹码					0.0000	显著

说明：^{*}我们在量化处理囚徒困境实验的方案选择时，将数字 0 视为甲方案/合作，1 视为乙方案/不合作。

实验结果也表明, 实验对象 B 越被信任、赠予的筹码越多、返还给 A 的筹码越多。如图 2. b 和表 1 所示, 中方和美方回赠筹码方面没有差异, 双方都表明, 在 B 得到筹码由 3 增加至 9 时 (A 初始赠予的筹码为 1-3, 再乘以 3 倍), 返还给 A 的筹码约由 1 增加至 4.5。比如中方实验对象 B 在得到 1 筹码、而增值为 3 个筹码后, 会返还 1.01 个筹码、即收益的 33% 给 A, 相应地, 在 B 的筹码增值为 6 和 9 之后, 会分别返还 2.73 个筹码、即收益的 46%, 4.47 个筹码、即收益的 51% 给 A。由此, 验证了人们之间因信任而产生互惠的假设, 而且互惠程度随着信任度的增加而增加, 从而人们得到的回馈增加、总收益也越趋于最大化。

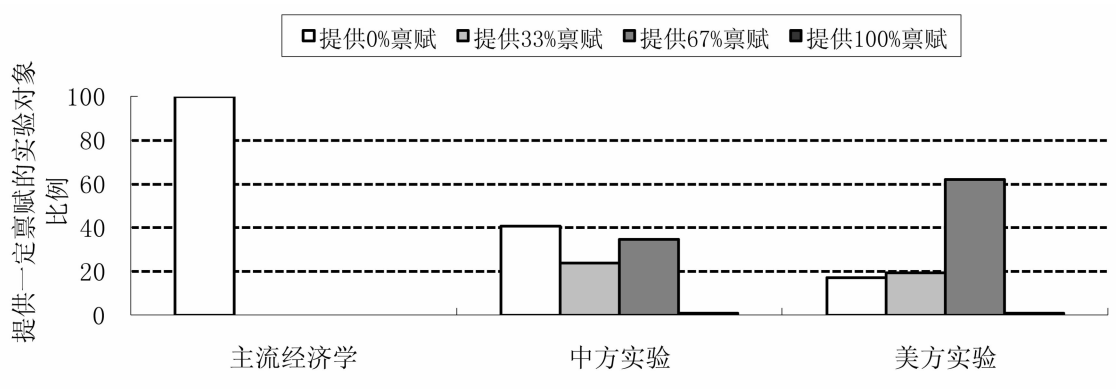


图 2. a 信任与信任度

结果 1: 信任实验表明, 人们对他人不仅持有信任, 还表现出互惠特征, 对对方信任程度越高、对方回馈的越多。善意本性引发人们互信和互惠偏好。信任程度在一定程度上受不同文化差异的影响, 美方的高信任度现象比中方更为普遍, 而但双方在互惠表现上则不存在差异。

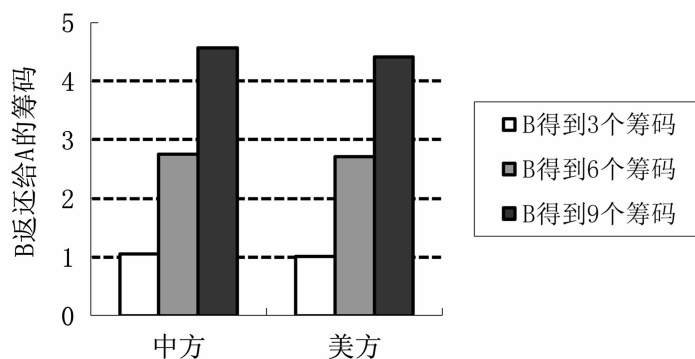


图 2. b 信任与互惠

(二) 囚徒困境实验中的信任与合作

同理, 与主流经济学预测的所有实验对象都会选择方案乙 (即原始游戏中的坦白)、或者说不合作的结果不同的是, 而中美双方实验对象并没有全部选择乙方案, 而是有一部分选择了合作的甲方案 (偏离纳什均衡) (见图 3)。也就是说, 在双方都没有任何禀赋的前提下, 人们的经济收益完全取决于对方的选择时, 人们还是会选择信任对方、共同合作以获取共同的成果。选择甲方案的实验对象告诉我们, 他们相信对方会为了双方共同的利益而进行合作, 信任对方的思维超越了理性计算的思维, 让他们决定听从前者, 即使会有一定的风险。表 1 说明, 在不

同文化背景下，实验对象的合作意愿存在显著差异，中美双方分别有 45% 和 70% 的实验对象表现出合作意愿，显然美方存在更为普遍信任与合作偏好，比中方多 25%。

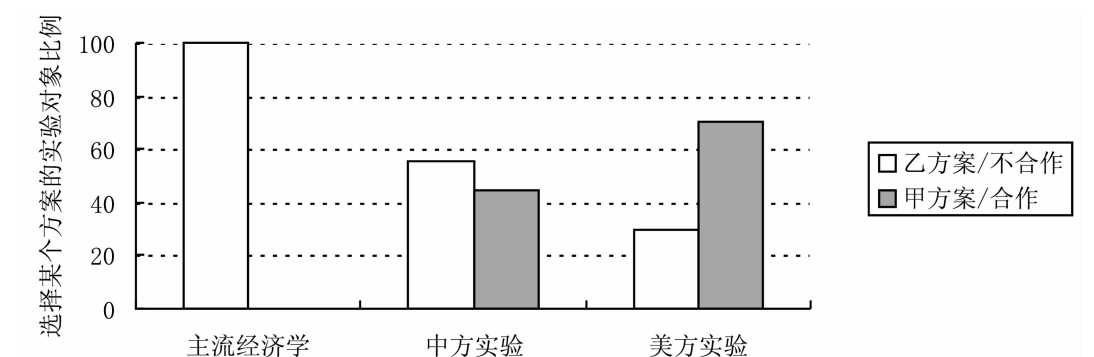


图3 信任与合作

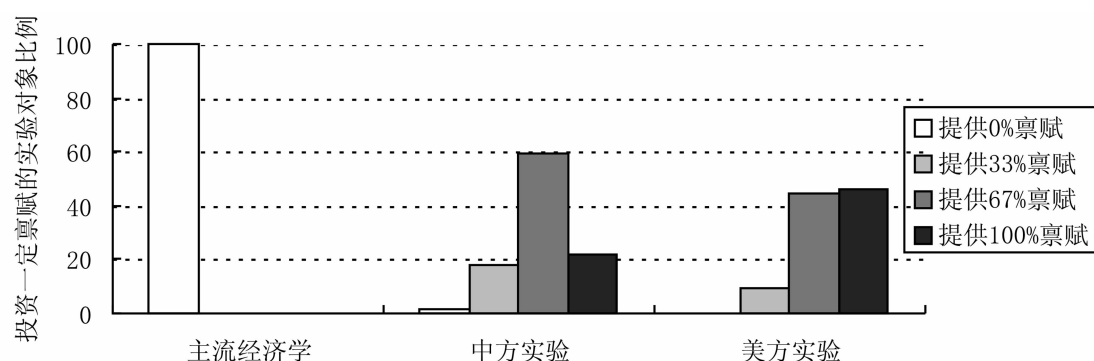


图4 信任度、合作水平与自愿生产

结果 2：囚徒困境实验表明，人们对他人不仅持有善意的信任，还表现出合作取得共同收益的偏好。合作偏好在不同文化之间表现出显著差异，美方的合作偏好现象比中方更为普遍。

（三）公共品实验中的信任度、合作水平与自愿供给

与主流经济学预测的所有实验对象都搭便车、一个筹码都不投资的结论不同，实验结果表明，中方只有 1.5% 的实验对象、而美方一个实验对象都没有符合这种情况。相反的是，大部分实验对象都提供了筹码，中方实验对象提供 1、2、3 个筹码的比例分别为 17.8%、59.2% 和 21.5%，美方则为 9.2%、44.8% 和 46%。中方实验对象大部分愿意提供 2 个筹码（33% 禀赋），而美方实验对象则愿意提供 2 个筹码（33% 禀赋）与 3 个筹码（100% 禀赋）（见图 4）。从这一点可以看出，美方实验对象之间的信任度更高、合作水平也更高，表 1 中的对比检验也说明了中美双方存在的显著差异。此外，中方实验对象的反馈信息也显示，他们并不完全相信其他 4 位队友，偏好至少保留一部分筹码、以保证初始禀赋不会功亏一篑，所以提供 2 个筹码是他们的普遍决策，结果大部分都持此理念，总体投资水平也并非最高。据此，我们可以说，信任度

与合作水平之间存在相关关系，信任度越高、合作水平越高；同时也表明，公共品是可以由用户自愿（合作）供给的，供给水平则取决于用户之间的信任度。

结果 3：公共品实验表明，人们对他人的信任产生互惠与合作，信任度与互惠程度和合作水平之间存在正比关系，合作水平越高、自愿供给的公共品数量越多。合作水平在不同文化之间表现出显著差异，美方的平均合作水平远远高于中方。

（四）独裁实验中的平等化与公平分配

当人们有绝对权力进行分配时，他们并没有像主流经济学所预测的那样，为了最大化自己的利益而自己保留 10 个筹码、一个筹码都不提供给其他人，实际上，中美双方的实验对象几乎没有这样的自私行为，相反，60% 以上的实验对象都选择给他人 5 个筹码（50% 禀赋）、自己保留 5 个筹码的这种平等化分配方式，说明他们天生偏好公平分配结果（见图 5）；此外，中方和美方学生平均提供分别为禀赋的 57.61% 和 53.4%，由表 1 可知，双方并无显著性差异。但是，少数实验对象也有强利他偏好，自己保留的筹码少于提供给对方的筹码，选择了提供禀赋 60%—100% 的行为。实验对象偏好平等分配的原因在于，他们相信对方也做出同样的决定。我们在他们的选项说明中发现，他们认为“由于一人扮演提供者和接受者两个角色，就意味着自己有 50% 的概率被选为接受者，而此时自己希望被赠与一半的筹码，他们也相信其他人也会有同样的预期，反过来，换位思考，如果自己是提供者时，那么就应该提供 50% 禀赋。”由此可见，是信任对方也有平等观，才会使提供者做出平均分配的决定。

但是，表 1 也说明，当比较中方内部实验对象在独裁实验和信任实验中提供行为时，却存在显著差异，因为在信任实验中，实验对象完全不提供的比例高达 40%，远远高于独裁实验中的 0%；而提供 67% 禀赋的人数比较高、达到 35%，而在独裁实验中提供相对程度禀赋的人数不足 20%。这说明，在存在同行奖励的情况下，人们的公平分配结果减弱，被奖励这种经济激励“挤出”。同理，美方内部两种实验结果的比较也存在相似情况。

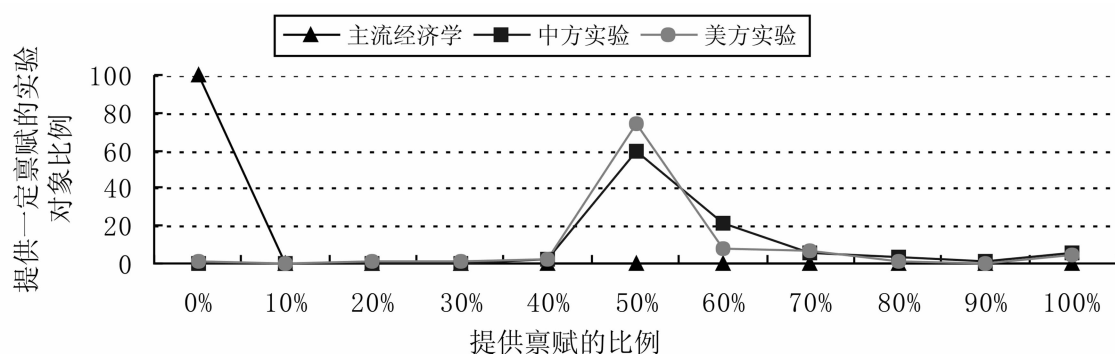


图 5 信任、平等观与公平分配

结果 4：独裁实验表明，人们天生具有平等观，相信他人也有共同的偏好，愿意提供 50% 的禀赋给他人，做出公平分配的决定；分配结果在不同文化之间并未存在明显差异。

结果 5：对比独裁实验与信任实验的结果，同行奖励机制减弱了公平分配结果，存在挤出效应。

（五）最后通牒实验中的惩罚与公平分配

主流经济学预测当理性经济人有相对权力进行分配时，即存在同行惩罚时，仅仅会给对方提供 1 个筹码（10% 禀赋），虽然不多，但 1 个筹码也大于一无所有，对方应该会接受这种分配，而且提供给对方的筹码越多、对方越会接受，也就是说提供 1 - 10 个筹码，对方都会接受，实验结果表明这种推断几乎不成立。首先，与独裁实验的结果类似，中美双方的实验对象几乎没有提供禀赋 10% 的筹码，仍然有 60% 以上的实验对象选择了 5 - 5 公平的分配方案（见图 6. a）；此外，中方和美方的实验对象平均提供分别为禀赋的 44.27% 和 45.9%，说明美方分配更趋向于公平，而是双方差异显著（见表 1）。其次，50% 以上的实验对象都认为被分配 1 个筹码是非常不公平的，果断拒绝这种方案，结果双方都没有得分。再次，实验对象 100% 接受的分配方案是禀赋的 50% 或者 60%，即他们偏好公平分配（见图 6. b）。最后，当实验对象受到不同平分配时，他们会拒绝，即使被提供的筹码高于 6 个，他们宁愿牺牲这些收益，例如，在被提供 10 个筹码（100% 禀赋）时，他们也认为这是非常不公平的分配，社会性思维已经远远胜过理性思维，拒绝 10 个筹码就意味着双方一分都没有、这是对对方的惩罚（见图 6. b）。可以说，人们为了获得公平分配的结果，不惜付出高昂的代价。

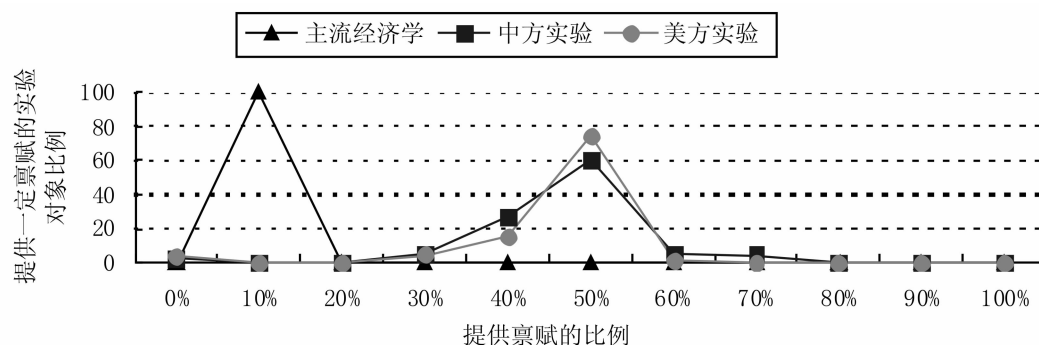


图 6. a 同行惩罚与公平分配

但是，同行惩罚对公平分配行为的影响存在文化差异。从表 1 可以看出，对比中方的独裁实验和最后通牒实验结果，实验对象提供的禀赋存在显著差异，在前者中，实验对象分配的方案倾向于大于等于禀赋的 50%，除了平等化的 5 - 5 分配比例都在 60% 左右以外，独裁试验中提供禀赋 60% 的强利他行为的比例占 20%，远远高于最后通牒实验中的比例、几乎为零；而提供禀赋 40% 的弱利他行为在最后通牒实验中却显著增加。由此可以说明，在中方实验中，同行惩罚弱化了利他行为。但是，同行惩罚对美方实验对象的分配行为却没有显著影响。

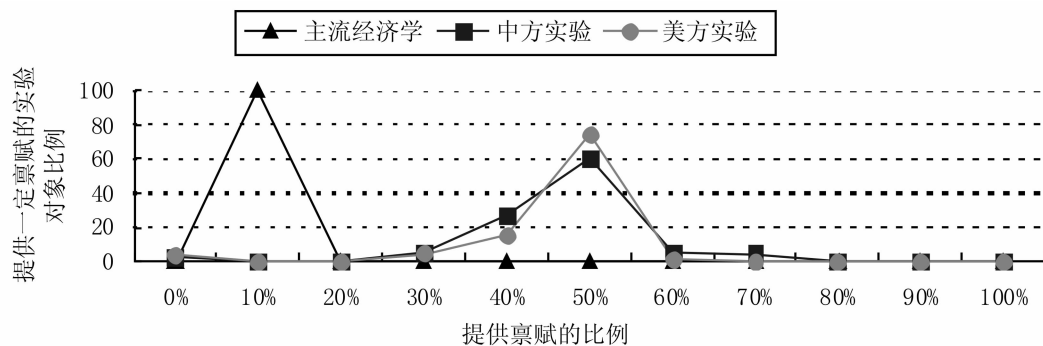


图 6. b 不公平分配与同行惩罚

结果 6：最后通牒实验表明，在存在同行惩罚的情况下，人们也会做出公平分配的选择，愿意提供 50% 的禀赋给对方，但是分配偏好在不同文化之间存在明显差异。

结果 7：对比最后通牒实验与独裁实验的结果，同行惩罚机制在不同文化条件下对公平行为具有不同影响，在中方则减弱了强利他行为，存在挤出效应；而在美方则没有影响。

五、结论

我们的实验结果不仅验证了人们天生具有善意的本质，以及人们的社会性思维与自愿供给和公平分配之间的关系，而且精炼出社会性思维的属性，即善意是根本、社会性偏好是表征。具体来讲，信任实验、囚徒困境实验和公共品实验，以层层深入的方式验证了公共品自愿供给假设，而独裁实验和最后通牒实验，则以程度相继加深的方式验证了公平分配假设，这些实验共同验证了主流经济学自利偏好和理性思维的局限性。由此，可以说，人们天生具有的善意本性主导着人们的社会性思维，产生信任、互惠和合作等社会偏好，能够自愿供给公共品、追求公平分配结果。本文的研究结论支持了周业安和宋紫峰（2008）的观点，即使我国的公共部门基本上由政府统一控制，社会成员自愿供给公共品的空间非常小，但是公共品自愿供给行为在某种程度上会超越文化因素的影响而显著存在，是社会性思维所导致的一种固有的行为模式。本文也支持了叶航（2012）的研究，善意是人们的道德情感，而其产生的终极原因是，人类天性中的正义感是人类在长期演化过程中适应生存环境的结果，社会规范内部化的产物，是维护人类合作秩序的社会正义，通过自然选择被固化在我们身体和心智中的禀赋。同时，本文的研究结果部分与 Bowels（2008）的观点一致，即人们的道德情操会被政策所削弱，尤其是政府制定的具有惩罚性质的外部规章，虽意在监管人们的违规行为，但往往会挤出原本道德本身的约束效应，结果事倍功半，这与我们中方最后通牒实验挤出部分强利他行为类似。

同时,本项研究的另一个重要贡献是,将5个实验结果联系起来,验证了“善意本性——信任、公平偏好——公共品自愿供给和公平分配”之间的逻辑关系,进而说明了理论预测在实践应用中的科学性。其中,信任是生产与分配的关键变量。信任是人们的一种本性,是人们行为策略选择的基础;信任在个体之间存在程度的差异、同时也相互影响个人之见的策略选择,信任度越高、生产水平越高、集体收益和个人收益均增加,信任度达到最大化时、个人收益和集体收益都会实现最大化;同时,信任度越高、越选择考虑对方利益的合作策略,受惠的对方也越将维护另一方的利益,结果导致双方的利益差别越小,分配结果越公平,但总的收益越大。可见,在公共品的提供方面,最大化信任度能导致类似生产帕累托最优的结果。信任的基础是个体心存善意,以及对对方的善意保持信息,因此,信任程度越高、相互信任的个体之间必然是公平的分配。一方面,信任导致合作策略的选择,本身包含考虑对方利益的同时也信任对方会同样考虑自己的利益,任何一方不会过度占用合作产生的利益。不会过度占用合作生产的利益,意味着双方的利益差别不会很大,或者说差不多,结果就表现为公平分配的实际结果。另一方面,信任导致的互惠行为的另一种表现是,当碰到明显偏离信任——合作应有的结果时,采取的不合作或对抗行为。如果碰到这种不符合“信任——合作——互惠——效率”的情况而仍然接受,实际上就否定了前面所谓的互惠行为的存在。相反,在出现明显不公平分配的情况时,可以看作是对“互信——合作”行为的否定,那作为具有“互信——合作”本性的个体,选择不合作甚至对抗以制止这种行为,就是一种逻辑的必然。继而,通过惩罚来制止不公平的分配,可以看作是前面互惠行为的另外一面。我们认为,相信人们的平等化观念,制定公平的分配方案,才能以最低的成本实现有序的分配共享利益。

基于此,本文建议政策制定者有必要充分了解人们的社会性思维和微观行为。既然人们天生有合作的意愿,能够本能地监督他人的行为,并毫不吝惜代价地惩罚违规者,那么政策的重点有必要考虑这些思维、能力和行动,为人们自愿解决集体行动问题提供信息平台。如果政策仅仅以主流经济学的理性假设为前提,强调外部的物质激励措施,比如价格机制和行政管理等,而忽略了人们内在的社会性思维,可能将付出高成本、低效率的干预代价。

同时,我们也注意到美方在信任度、合作偏好、合作水平和公平偏好方面都表现出较高(强)的特征,借助已有研究我们初步总结出中美双方在行为效果差异方面的原因,包括两个方面。一方面是基于市场整合因素的解释,即市场整合程度越高的地区,人们的合作水平越高(Henrich等,2010a);美国显然具有比中国更为发达的市场经济水平,在信任和合作等行为结果方面表现出更高水平的绩效也就不言而喻了。另一方面是基于基本认知和动机过程因素的解释,即西方国家的大学生,尤其是美国人,与亚洲国家的相比,更多地采取分析式的推理策略,不受环境影响而是更为依赖规则来解释和预期行为(Henrich等,2010b)。

需要说明的是,尽管中美双方的文化和制度等背景不同,双方的实验对象在信任、合作等程度上表现出一定差异,但是他们的总体行为趋势表现出极大地相似性,即没有一方出现主流经济学所预测的后果占据主导地位的情况,这也恰恰说明具有善意的人们行为

受不同环境的影响而表现出社会性思维。此外,本文也存在一些不足之处,例如只设计了静态单次博弈、没有具体比较个人的理性思维和社会性思维、没有考虑个人的异质性特征等,仅仅基于学生作为实验对象得出的结论或许与社会一般现状之间有些偏差,对中美双方行为结果量化差异的原因探讨还有待深入,我们后续会进一步研究这些因素并将实验扩展到田野进行实施。

参考文献

- Andreoni, J., 1995, "Warm - Glow versus Cold - Pickle: The Effects of Positive and Negative Framing on Cooperation in Experiments", *Quarterly Journal of Economics*, 110 (1), pp. 1 - 21.
- Berg, J., Dickhaut, J., and McCabe, K. A., 1995, "Trust, Reciprocity, and Social History", *Games and Economic Behavior*, (10), pp. 290 - 307.
- Bowles, S., 2008, "Policies Designed for Self - interested Citizens May Undermine 'the Moral Sentiments': Evidence from Economic Experiments", *Science*, 320, pp. 1605 - 1609.
- Feigenberg, B., Field, E., and Pande, R., 2013, "The Economic Returns to Social Interaction: Experimental Evidence from Microfinance", *Review of Economic Studies*, 80 (4), pp. 1459 - 1483.
- Guth, W., Schmittberger, R., and Schwarze, B., 1982, "An Experimental Analysis of Ultimatum Bargaining", *Journal of Economic Behavior and Organization*, (3), pp. 367 - 388.
- Henrich, J., Boyd, R., Bowles, S., et al., 2001, "In Search of Homo Economics: Behavioral Experiments in 15 Small - scale Societies", *American Economic Review*, 91 (2), pp. 73 - 78.
- Henrich, J., McElreath, R., Barr, A., et al., 2006, "Costly Punishment across Human Societies", *Science*, 312 (23), pp. 1767 - 1770.
- Henrich, J., Ensminger, J., McElreath, R., et al., 2010a, "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment", *Science*, (327), pp. 1480 - 1484.
- Henrich, J., Heine, S. J. and Norenzayan, A. 2010b. Most People Are Not WEIRD. *Nature*, 466, pp. 29.
- Janssen, M. A., Holahan, R., Lee, A., and Ostrom, E. 2010, "Lab Experiments for the Study of Social - ecological Systems", *Science*, 328, pp. 613 - 617.
- Kahneman, D., Knetsch, J., and Thaler, R., 1986, "Fairness and Assumptions of Economics", *Journal of Business*, (59), pp. s285 - s300.
- Karlan, D., 2005, "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions", *American Economic Review*, 95 (5), pp. 1688 - 1699.
- Marwell, G., and Ames, R., 1979, "Experiments on the Provision of Public Goods I: Resources, Interest, Group Size and the Free - rider Problem", *American Journal of Sociology*, 84 (6), pp. 1335 - 1360.
- Ostrom, E., 1990, *Governing the Commons*. Cambridge: Cambridge University Press.
- World Bank, 2015, *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: World Bank. doi: 10.1596/978-1-4648-0342-0. License: Creative Commons Attribution CC BY 3.0 IGO.
- 陈叶烽、周业安和宋紫峰, 2011, 《人们关注的是分配动机还是分配结果?》, 《经济研究》第 6 期 31 - 44 页。
- 陈叶烽, 2009, 《亲社会性行为及其社会偏好的分解》, 《经济研究》第 12 期 131 - 144 页。

叶航, 2012, 《公共合作中的社会困境与社会正义》, 《经济研究》第 8 期 132 – 143 页。

周业安、连洪泉、陈叶烽等, 2013, 《社会角色、个体异质性和公共品自愿供给》, 《经济研究》第 1 期 123 – 136 页。

周业安和宋紫峰, 2008, 《公共品的自愿供给机制：一项实验研究》, 《经济研究》第 7 期 90 – 104 页。

周业安, 2015, 《论偏好的微观结构》, 《南方经济》第 4 期 106 – 115 页。

Voluntary Provision and Fair Distribution of Public Goods based on Goodness: Behavioral Economic Experiments between China and USA

Chai Ying Wang Yi Marco A. Janssen

Abstract: The rational thinking assumption by main stream economics has a far distance from real people. This paper designs a series of behavioral economic experiments (trust, prisoners' dilemma, public good, dictator and ultimatum games) and conducts them in two university classrooms of China and USA. It shows that people intrinsically have goodness and social thinking, and can provide public goods voluntarily and have fair preference of economic interest distribution. However, the American students show higher trust degree and cooperation level, and stronger preferences of cooperation and fairness than the Chinese ones. Therefore, this paper suggests that policy makers are necessary to fully understand people's social thinking and micro behavior and they are necessary to emphasize on publicizing this kind of thinking, ability and action, and should provide information platform for people to solve collective dilemmas voluntarily.

Keywords: Goodness; Thinking Socially; Voluntary Provision; Fair Distribution; Behavioral Economic Experiments.

(责任编辑: 金保)