

## 个体最优选择与社会合意行动\*

### ——兼论分配公平的逻辑内涵与实践特征

王国成 欧阳葵

(中国社会科学院数量经济与技术经济研究所, 北京, 100732; 西北大学经济管理学院, 西安, 710127)

**摘要:** 基于理性个体行为的策略交互性, 从博弈均衡的意义上提出社会合意行动的定义。通过引入一个公共参与人(其目标函数即为社会福利函数), 将博弈论与社会选择理论结合起来, 进一步提出社会最优契约的解概念。以此为工具, 分别比较了几种常见的公平准则的逻辑内涵及其实践特征, 并进一步探讨了群体行动计算的技术实现, 试图为解析个体与集体之间的复杂行为关系、解决现实中的公平分配等问题提供一条新的思路。

**关键词:** 个体行为; 公共参与人; 社会选择; 博弈分析框架; 分配公平

## Individual Optimal Choice and Social Consensus Activity: On the Logicity and Practice of Fair Distribution

**Abstract:** Based on the strategic interaction between rational individuals, this paper proposes the definition of social consensus activity in a game equilibrium sense. A (real or artificial) public player, whose objective function is just the social welfare function, is introduced to combine game theory and social choice theory together, and then the definition of social optimal agreement is proposed as a game solution concept. Logical connotation and practical characteristic of several kinds of common fair criteria with this solution concept are systematically compared, and the technical implementation of collective action calculation is further explored in order to provide a new approach for analyzing the complex behavioral relations between the individual and the collective and for solving realistic problems such as fair distribution.

**Keywords:** Individual behavior, Public player, Social choice, Analytical framework of game theory, Fair distribution

### 引言

如何在集体行动层面上合理地组织和有效地实现个体最优选择及其与集体目标的协调一致, 是不断考验和激发人类智慧的重大理论与实践课题。当今利益主体多元化与全球一体化、微观选择的个性化(行为方式多样化)与宏观经济的整体性之间的矛盾日益凸显和激烈, 尤其是在面对各类危机和冲突等日趋复杂的社会经济活动时, 传统理论在屡屡经受无情的冲击和检验时显露出滞后窘态, 这其实是在孕育着上佳的机遇和提出了迫切的现实需求, 将当代经济理论的发展推到了极其重要的

\*作者简介: 王国成, 中国社会科学院数量经济与技术经济研究所, E-mail: [wanggc@aliyun.com](mailto:wanggc@aliyun.com); 欧阳葵, 西北大学经济管理学院, E-mail: [wcsyok@163.com](mailto:wcsyok@163.com)。

项目资助: 本文为国家重大科研计划 973 项目(课题编号: 2012CB955802)的理论支持部分。

抉择关口。委托代理模型、激励机制设计具有明显的在非对称信息条件下以单方效用最大化为目标的理论倾向 (Rees,1985; Myerson,2008); 多边谈判和契约理论中着重研究对称行为假设下的联合效用最大化(Binmore, *et. al.*, 1986; Hart and Moore,1988); 而公共选择与对集体行动逻辑的探索, 更多地得到的是消极的负面结论, 虽然有森等人的改进, 但以阿罗的工作为代表、由严谨逻辑得出的一系列不可能定理, 为最大限度满足个体意愿的社会选择构筑的障碍似乎仍难以逾越(Olson,1965; Arrow,1951); 日渐炙热的用演化观点和方法研究合作行为的理论焦点(Nowak,2006), 也主要还是个体视角的; 福利经济学、发展经济学以及经济学家们在充填宏观微观人为沟壑所做的种种努力和沉淀, 或许能提醒我们从个体与集体的行为关系角度, 以博弈模型为分析框架, 更加集中地研究主体之间的策略行为, 有望揭秘个体选择与集体行动之间复杂的内在联动关系。

本文拓宽行为分析视野, 围绕个体与群体行为关系逐步展开探讨, 大致结构安排如下: 第一节在分析个体行为的策略性的基础上阐释社会合意行动; 第二节将博弈论与社会选择理论结合起来, 基于博弈均衡构建社会选择函数; 第三节侧重分析几种常见的公平准则的逻辑内涵及相互关系; 第四节进一步分析社会最优契约的实践特征与现实意义; 第五节概述集体行动计算的技术实现, 并简要介绍相关应用案例; 最后概括地给出结论性评述和展望。

## 1 个体自利行为的策略转向与社会合意行动

从自利角度考察个体理性行为及涉他性 (Other-regarding) 和从行为关系角度考察集体理性及社会效应, 能拓宽学术视野、丰富行为研究内容、发掘其本质特征。

### 1.1 个体行为的策略性

人类属于群居动物, 在与自然界进行物质交互过程中面临的可利用资源总是有限的, 其行动空间必然受不同主体相互影响的制约, 而且单一个体在相对独立决策时与多主体相互影响条件下决策时的行为模式还可能存在显著差异。因此, 将个体行为空间划分成两个区段或部位: 以个体效用最大化为目标的相对独立决策的区段 I 和以策略行为相互影响下考虑公共利益目标的区段 II (见图 1 中发生偏转的直线), 更有利于全面分析行为属性以寻求其本质特征。在相对独立的个体决策阶段, 可完全自主选择 (法律框架内想干啥就干啥), 追求效用最大化; 在相互影响 (策略行为) 的公共利益阶段, 需要遵循共同规则, 不允许干的就一定不能去干 (必须有法律明确赋权才能去行动)。两个区段的行动目标和方式也有显著不同: 个体行为追求最优, 群体行动避免最差; 相比较而言, 个体响应方式灵活敏感, 群体行动模式迟缓滞后; 不同区段的外界条件不同, 刺激和诱发出的行为表现也不同, 具体的决策行为方式也就不同, 区段 II 交互条件下的具体行为表现更多地受双 (多) 方策略选择的影响; 而且每一个体的转折点  $\lambda_0$  与其行为秉性、资源禀赋、外部环境 (包括与他人的关系) 密切相关, 由此表现出个体行为的异质性和策略性及非线性特征。

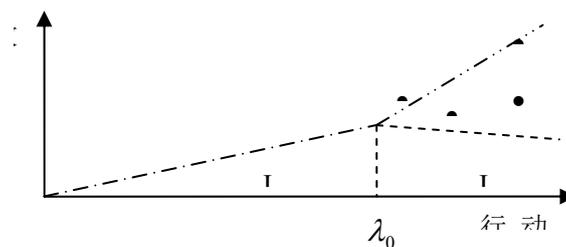


图1 个体非线性决策行为示意图

---

(图中, 单点虚线为边际效用相对稳定的个体决策区段;  $\lambda_0$  为行为分界点或阈值; 双点虚线的区段II是相互影响下的策略行为阶段, 还有可能呈波动震荡或折线状, 如实点所示; 虚线表示另一类可能的行为表现。)

在区段I可采用同质的代表性主体的行为假设及相关分析方法, 而在区段II则需要进行异质性主体的交互行为分析。个体行为的异质性具体表现为: 每一个体禀赋的多元行为属性的比例结构(权重分布)、与外界发生交互和演变过程中的条件依赖与临界点(图1中的 $\lambda_0$ )是不同的, 由此影响和可能导致群体行动的传导机理与产出结果的多重性。所以, 在考察个体决策行为时, 一定要注意在策略行为的交互影响下为实现公共利益(从而提高自身利益)有可能发生的转变。只要个体选择是非线性、差异化的, 必然会在集体行动中出现公共利益目标不一致、不协调的情况。换句话说, 宏观复杂的典型化事实等是纳什均衡多重性的表象, 根源上都是由异质性个体在交互行为中的多样化选择所致。通过分析个体自利行为的策略转向来拓宽视野和扩展疆域, 是理论发展的必然和现实需求的使然。

## 1.2 社会合意行动的理论分析框架

群体中成员相互之间的策略行为可能导致合作创造剩余(增值), 也可能因不合作造成内耗降低个体效用。同质的理性主体在基数效用基础上可直接加总生成总量, 而对于不同维度、多元属性、多种影响因素、不同计量单位的异质性主体, 更多地应考虑其序数效用, 不具有可比性和可加性, 不宜简单加总而是某种序包含的关系。由此引出集体行动当且仅当是一种社会相对最优, 用合意来表示此时集体行动的基本属性和特征; 而合意行动基础上的社会选择(福利)函数的最优契约的存在与否, 取决于个体成员的选择行为属性(考虑策略效应的非线性决策行为)和对目标产出的评判标准(主流价值观)、组织形态结构演变和适应性群体认知(文化认同)以及对外部环境条件的响应模式等。

目标完全相同、步调绝对一致的集体行动(目标或价值函数的同构)在现实中通常是不可能存在的, 当个体选择延展到相互影响的策略行为区段, 众多差异化个体的交互作用会导致什么样的集体行动, 该如何选择尺度标准和描述刻画、量化测度的方法, 寻找和搭建什么样的理论分析框架? 由于个体是本能的自利与条件性的利他的行为复合体, 具有禀赋理性与学习适应性共生、独立行事与合作倾向并存等多元属性, 只是各部分(分量)所占比例程度、随外界条件变化的临界点等可能会因人而异、有所不同, 在考量策略交互性后更清晰地显现其内在的本质特征, 具体表现为分阶段选择的转折分岔轨迹(见图1); 由于集体行动的价值函数(目标)的非单一性、序关系的多途径关系链条和多向性、制度规则的作用以及政府行为的现实干预效应; 特定环境下差异性个体行为相互影响的非线性本征属性, 经过自身的信念与判断、相互之间的信任激荡、组织结构中的发酵与酿造, 潜藏了多种可能的总体产出, 在不同的结构网络和外界诱变条件下演化出多样态典型化事实, 使得个体选择与集体行动之间的关系错综复杂、扑朔迷离, 导致了理论上博弈均衡的多重性和实践中的宏观涌现, 这正是社会经济复杂性的根源和机理, 也在一定程度上表明了集体行动中有公共协调人参与(或某种力量的介入)的必要性和必然性。

博弈论中的囚徒困境<sup>①</sup>等经典案例, 在揭示个体优化选择与集体行动结果之间复杂内在联系(也称为社会选择悖论)的同时, 也构建了示范性的理论分析框架, 能够结构式、一体化地展现个体与群体的行为特征、传导机理和各种可能的结果, 将个体选择、总体产出和分布结构相互之间的有机

---

<sup>①</sup> 由 Merrill Flood, Melvin Dresher 和 Albert W. Tucker 共同莫立, 参见: Poundstone, W. (1992) *Prisoner's Dilemma*. NY: Doubleday.

联系整合后再现。在相互影响条件下，每一个体的效用或产出不仅仅由自身行为决定，还受他人行为和外界条件变化的影响（甚至在某些情况下要取决于他人如何选择）；同时，自身行为也在影响他人。“合”主要是指符合（程度），“意”即意向/意愿，可将大多数人认可接受、多数人有同感的意愿，理解成边界相对清楚的一个动态渐近过程，直至一致同意，而非一个点（值）的状态，适合在博弈框架内用现代集合论方法分析。进而，在博弈均衡意义上界定社会合意的集体行动和社会最优契约等，研究它们的实现条件和稳定（惟一）性等。如此定义的社会合意行动不可能保证所有参与人都能同时实现自身效用最大化，但却能避免集体最差结果：每一参与人不可能靠单方面的策略变换提高自身的收益，同时也不用担心其他人的策略行为变化会降低自己的效用，也没有人会主动偏离共同预期到的博弈均衡结果，因而又是相对稳定的。

## 2 基于博弈均衡构建社会选择函数

### 2.1 博弈均衡与社会选择

纳什均衡——非合作博弈论中的核心概念，是一个自我驱动型的契约(Brandenberger and Dekel,1987)，没有人愿意单方面偏离契约中的共同约定。然而，这样的契约却常常具有多重性，在社会选择时反映出现实中公平准则的不惟一性。如何从多个契约中选择合理可行的契约，而且是相对稳定的，这就需要由个体理性导致的纳什均衡结果施加某些集体理性的限制。这可通过考察如下表所示的具有三种备选事项的扩展情侣博弈来阐释。

表 1 扩展的情侣博弈

1 \ 2	芭蕾	电影	足球
芭蕾	1, 1	0, 0	0, 0
电影	0, 0	1, 5	0, 0
足球	0, 0	0, 0	4, 4

表中：（芭蕾，芭蕾）、（电影，电影）和（足球，足球）是该博弈的三个纯策略纳什均衡。然而，由于信息的不对称，个人的偏好与社会福利最大的集体产出并无必然的内在联系，那么，实际中到底会发生哪一种结果呢？虽然直观上很容易看出（芭蕾，芭蕾）的策略组合被其他两个纳什均衡所帕累托占优，但是，直觉缺乏严格的理论支撑，现实中并没有绝对把握保证其他两种结果更有可能发生。而若参与人 1 和 2 可以事先进行交流与协调，那么（足球，足球）便是一个更加可信的结果。这一结果对于双方来说不但是集体最优，而且也是稳定的。于是，参与人之间的交流与协调使得该类博弈本质上成了一个社会选择问题。具体来说，表 1 所示的情侣博弈便可转化为如下表所示的社会选择问题：

表 2 由扩展情侣博弈转化的社会选择问题

芭蕾，芭蕾	芭蕾，电影	芭蕾，足球
电影，芭蕾	电影，电影	电影，足球
足球，芭蕾	足球，电影	足球，足球

此时，若将参与人的行动空间看成是备择社会状态集，从而两个参与人在交流与协调过程中的契约选择过程就可转化为一个社会最优选择问题。如果基于参与人的支付定义或构建一个社会福利函数，那么社会最优选择问题就视同为一个社会福利最大化问题。在将社会福利视为个人效用之和

(简化为代数和)的情况下,表2中所示社会状态集所对应的社会福利水平如下:

表3 扩展情侣博弈的社会福利

2	0	0
0	6	0
0	0	8

由表3可见,(足球,足球)是社会福利最大化的行动组合。但是,一般而论,将此定义为社会最优契约选择仍存在两个问题:一是理论上并不能保证社会福利最大化的行动组合与纳什均衡等价(Bergson,1938),而且人们总是希望良好的集体行动结果具有稳健性,这就必须要求社会最优契约选择应包含于纳什均衡之中;二是把社会福利表示为个人效用的简单代数和可能存在极大的争议,因为个量如何生成总量至今仍是一个未解之谜。这种功利主义主张以及遵循伯格森—萨缪尔森传统(Samuelson,1947)的其他社会福利函数受到了强烈批判。

由阿罗不可能定理可知,如果备选社会状态至少有三个,那么在无限制的定義域下,当社会偏好满足完备性和传递性时,满足帕累托准则和不相关选择独立性的社会福利函数必然满足独裁性。这就是为达成集体理性施加限制性条件的必要理由。由此表明,在社会福利函数的选择上,缺乏精准度量的序数主义原则无法给出令人满意的答案。事实上,在很多情况下,社会选择问题必须依赖于能测度出偏好强度的基数性质。在策略行为影响下的人际比较中,社会选择所牵涉到的最具代表性的社会福利评价标准主要有三类:功利主义(Bentham,1789)、罗尔斯主义(Rawls,1971)和纳什谈判解(Nash,1950a)。在某种意义上,纳什谈判解可视为功利主义和罗尔斯主义之间的折中。而在这三种社会福利评价标准中,只有纳什谈判解才能满足基数不可比性。当然,这并不必然表明纳什谈判解比其他标准更具有优势。为保护公共利益,达成集体理性,本文采取的方法是在博弈中引入一个公共参与人,让其充当所有参与人事前交流磋商中的协调人,以便使构建更丰富含义的社会福利函数成为可行。

现实中更常见的是不同的博弈参与人对备选项的偏爱程度、效用函数的取值是异质的,考虑个体选择中的自利与涉他不同区段选择的非线性特征,以及相应地由此引起的群体组织和支付结构的复杂多变,进而可推论,个体多元化的偏好分布、行为属性和选择导致了群体决策结果的不惟一性。正是因为异质性与交互性等个体多元行为属性(多样化行为表现)为分析集体行动结果和复杂宏观现象造成了诸多麻烦,因而也就必须以对异质与交互等行为属性的深化分析作为揭示复杂之谜的逻辑起点。基于互异主体(HIA: Heterogeneous Interacting Agent)研究复杂经济问题,不仅是对经济理论的重大贡献,也是认识论上重要的发展进步。在此情形下,可外生地將社会福利函数确定为公共参与人的效用函数,笔者认为,这一尝试不但在理论上是可行的,也是十分现实的。

## 2.2 公共参与人博弈

类似于完全竞争市场上一般均衡价格形成机制中的虚拟喊价人,以及一系列不可能性定理等相关研究结论的理论指向,启发我们构建引入公共参与人的博弈均衡分析框架。公共参与人的“存在”形式、<sup>①</sup>策略空间和职责是推动形成共同行为准则、增强互信和促进信息有效沟通交流,其效用(支付)函数就是代表公共利益的社会福利(选择)函数。为了便于对比说明,我们在此对相互关联的基础概念给出严格的形式化表述。设有限参与人集合为 $I=\{1, 2, \dots, N\}$ 。对于任意参与人 $i \in I$ , 设其

<sup>①</sup> 一个很容易想到的公共参与人的现实存在是政府或公共组织。

有限行动空间为  $A_i$ , 并记  $A = A_1 \times \cdots \times A_N$ . 参与人  $i$  的效用函数即为定义在  $A$  上的支付函数  $u_i: A \rightarrow \mathbb{R}$ . 对于任意  $i \in I$  和任意行动组合  $\mathbf{a} = (a_1, \cdots, a_N) = (a_i, \mathbf{a}_{-i}) \in A$ ,  $u_i(\mathbf{a}) \in U_i$  表示参与人  $i$  在行动组合  $\mathbf{a}$  下的支付 (或效用). 设  $u_i(A) = U_i$  为参与人  $i$  的支付 (或效用) 空间, 并记  $U = U_1 \times \cdots \times U_N \subseteq \mathbb{R}^N$ . 我们称  $G = (I, A, U)$  为参与人  $I$  之间的一个 (标准型) 博弈. 设定义在  $A_i$  上的任意概率分布的集合记为  $\Delta(A_i)$ , 称之为参与人  $i$  的混合策略空间, 其元素  $\alpha_i \in \Delta(A_i)$  称为参与人  $i$  的混合策略. 参与人的混合策略空间可记为  $\Delta(A_1) \times \cdots \times \Delta(A_N)$ .

将纳什均衡概念(Nash,1950b)简化而又严格地表示为: 如果存在某个混合策略组合  $\alpha^* = (\alpha_i^*, \alpha_{-i}^*) \in \Delta(A_1) \times \cdots \times \Delta(A_N)$ , 使得  $\forall i \in I, \forall a_i \in A_i$ ,

$$u_i(\alpha_i^*, \alpha_{-i}^*) \geq u_i(a_i, \alpha_{-i}^*),$$

则称  $\alpha^*$  为  $G$  的纳什均衡.

现在, 我们试图在博弈  $G$  中引入一个公共参与人. “公共参与人”可以是虚拟的某种形式, 也可以是真实存在的自然人或法人 (有时甚至可以存在于各参与人之间); 既可以是某种物质实体, 也可以是某种抽象的构造物……. 在博弈进行之前, 由公共参与人提出一个策略组合建议. 如果所有参与人都愿意遵循公共参与人所提出的建议, 则可称此建议为该博弈的一个契约. 如果公共参与人所提出的建议是一个纳什均衡, 那么将没有人愿意单方面违背该建议. 但是, 如果把纳什均衡视为该博弈的一个契约, 那么事实上就相当于假设所有参与人的行为选择都满足如下两条准则:

准则 1 (非合作原则): 每一参与人的最优策略选择基于其他人都遵循该建议;

准则 2 (纳什稳定原则): 在实施了准则 1 的情况下, 如果违背该建议对每一参与人没有任何好处, 那么其将遵循该建议.

公共参与人的基本特征: 需要考虑的主要因素包括整体形象、文化与主流价值观、战略选择、公信力等; 所执行的职能是实现必要的交流与沟通, 由此引导个体行为选择. 演化博弈中 (Weibull,1995)以对称协调博弈为平台, 研究合作行为的发生、持续和集体理性的实现条件, 相应地明确个体行为选择、个体约束条件和公共道德底线等. 而行为规则和个人信用, 是最具有可操作性的变量和控制因素.

准则 1 表明在纳什均衡中参与人策略选择的独立性, 从而使得纳什均衡在非合作原则下成为非常有说服力的解概念. 准则 2 体现了纳什均衡的稳定性, 即纳什均衡具有“自我驱动型”性质, 也可以直观地理解为: 因为违背该契约对参与人没有任何好处, 但却可能给其他人带来害处, 所以每一参与人将遵循该契约. 因此, 如果每一参与人都相信其他参与人不会采取“对自己无利而对他人有害”的行为, 那么准则 2 便是对非合作博弈的解概念所施加的一个非常合理的约束.

在这样的协调过程中, 公共参与人充当了一个协调人. 自然地, 公共参与人也可以与其他参与人一样参与博弈. 因此, 若将公共参与人用  $0$  来表示, 则参与人集合可设为  $I^\# = \{0, 1, \cdots, N\}$ . 对于公共参与人  $0 \in I^\#$ , 设其有限行动空间为  $A_0$ , 并记  $A^\# = A_0 \times A_1 \times \cdots \times A_N$ ,  $U^\# = U_0 \times U_1 \times \cdots \times U_N \subseteq \mathbb{R}^{N+1}$ . 公共参与人的效用函数即为社会福利函数, 且对于任意  $\mathbf{a} \in A^\#$ ,  $u_0(\mathbf{a}) = W(u_1(\mathbf{a}), \cdots, u_N(\mathbf{a}))$ . 特别地, 我们总是假设社会福利函数是个人效用函数的增函数, 即对于任意  $i \in I$ , 若  $u_i \geq u_i'$ , 则  $W(u_i, \mathbf{u}_{-i}) \geq W(u_i', \mathbf{u}_{-i})$ . 我们称  $G^\# = (I^\#, A^\#, U^\#)$  为一个公共参与人博弈. 如果公共参与人的行动空间是空集, 那么公共参与人博弈  $G^\#$  便退化为一般博弈  $G$ . 很明显, 公共参与人的行为与策略极易给我们留下充分的想象空间. 如果公共参与人拥有协调者和参与者双重角色, 将会使问题变得更加复杂然而也更加诱人.

同样地，可把公共参与人博弈中的混合策略空间记为 $\Delta(A_0) \times \Delta(A_1) \times \cdots \times \Delta(A_N)$ ，于是，就可以很自然地推广得到公共参与人博弈 $G^\#$ 的纳什均衡定义。

**定义 2.1** 如果存在某个混合策略组合 $\alpha^* = (\alpha_1^*, \alpha_{-1}^*) \in \Delta(A_0) \times \Delta(A_1) \times \cdots \times \Delta(A_N)$ ，使得 $\forall i \in I^\#, \forall a_i \in A_i$ ， $u_i(\alpha_i^*, \alpha_{-i}^*) \geq u_i(a_i, \alpha_{-i}^*)$ ，

则称 $\alpha^*$ 为公共参与人博弈 $G^\#$ 的纳什均衡。

公共参与人的引入，促进了在策略行为的区段 $\Pi$  参与人之间的信息交流和互信，确立公共行动规则(可共同接受的主流价值观)，协调集体行动，有利于达成社会最优契约。

### 3 社会选择中常见公平准则的逻辑内涵与比较

在社会选择时常用的三类公平准则中，功利主义是朴素的，罗尔斯主义附加了外生的公平观，纳什谈判解则显得有些变幻莫测。而在引入公共参与人的博弈分析框架中，通过比较能更清楚地看到各类公平准则的逻辑内涵与关键特征，以及它们之间的相互关联。

#### 3.1 集体合意行动与社会最优契约

在公共参与人博弈 $G^\#$ 中，由于一般假设参与人的行动组合空间是有限集，故参与人的混合策略空间必定是紧集。因此，至少存在一个混合策略组合 $\alpha^0 \in \Delta(A_0) \times \Delta(A_1) \times \cdots \times \Delta(A_N)$ ，使得 $\forall a \in A^\#$ ，有：

$$u_0(\alpha^0) \geq u_0(a)$$

也就是说， $\alpha^0$ 是使得社会福利最大化的混合策略组合，可称之为“社会最优策略组合”。遗憾的是，社会最优策略组合并不一定是纳什均衡，这使得社会最优策略组合通常难以实现。在纳什均衡中，当其他参与人都选择均衡策略时，公共参与人的均衡策略必定是针对其他参与人策略的最优反应(即给定其他参与人的均衡策略选择，公共参与人的均衡策略选择必定是社会最优的)，但是这个最优反应并不一定是 $\alpha^0$ 。由于纳什均衡通常具有多重性，在不同的纳什均衡中，公共参与人会选择不同的最优反应。因此，如果博弈的结果只能是纳什均衡，那么我们会很自然地期望得到一个“最好”的结果——即“最好”的纳什均衡，就称这个“最好”的纳什均衡为社会最优契约。即当其他人都选择均衡策略时，公共参与人应该选择使得社会福利最大化的行动。但是，由于博弈均衡的多重性，公共参与人博弈的纳什均衡却未必是惟一使得社会福利最大化的行为组合。因此，社会最优选择与纳什均衡常常是不一致的。

不失一般性，我们可以将公共参与人博弈中的混合策略空间 $\Delta(A_0) \times \Delta(A_1) \times \cdots \times \Delta(A_N)$ 视为一个可供选择的社会状态集。进一步，混合策略空间中的任意一个混合策略组合都可视为行动空间上的一个概率分布。设博弈 $G^\#$ 所有混合策略组合所对应的概率分布的集合为 $\mathcal{M}^\#$ ，而纳什均衡所对应的概率分布的集合则为 $\mathcal{E}^\# \subseteq \mathcal{M}^\#$ 。现在，如何将 $\mathcal{M}^\#$ 中的所有元素进行一个合理的社会排序便成为一个社会选择问题。据此，任何一个公共参与人博弈都相应地对应着一个社会选择问题。

然而，社会选择问题中的最优社会选择却与纳什均衡没有什么必然的联系。尽管公共参与人会倾向于社会最优的行为组合，但是其他参与人却未必同意这一点。而在纳什均衡中，虽然各参与人都不愿单方面偏离均衡策略，但均衡策略却未必使得社会福利最大化。为了在纳什均衡中兼顾策略选择的社会属性，本文提出社会最优契约的概念：即公共参与人在各种可行契约(纳什均衡)中提出社会最优的契约。不难发现，公共参与人的社会偏好只有在当纳什均衡具有多重性时才能体现出

来。这里的逻辑顺序是：

- (1) 公共参与者提出建议，其建议必须是纳什均衡；
- (2) 当存在多个纳什均衡时，公共参与者将选择一个最优的纳什均衡；
- (3) 各参与者（在准则 1 和准则 2 下）接受并履行社会最优契约。

然而，社会最优契约未必是最优社会选择，而只是多重纳什均衡中社会排序最高的纳什均衡。

我们现在将社会最优契约严格定义如下：

**定义 3.1** 在公共参与者博弈  $G^\#$  中，若存在  $\pi \in \mathcal{E}^\#$ ，使得

$$\pi \in \arg \max_{\pi' \in \mathcal{E}^\#} \sum_{\mathbf{a} \in A^\#} \pi'(\mathbf{a}) W(u_1(\mathbf{a}), \dots, u_N(\mathbf{a})),$$

则称  $\pi$  是一个社会最优契约。

根据此定义，由不同的社会福利函数假设可以得到不同的社会最优契约。如果社会福利函数  $W(\cdot)$  满足如下两个性质：

- (1) 若对于任意  $i \in I$ ， $u_i \geq u_i'$ ，则  $W(\mathbf{u}) \geq W(\mathbf{u}')$ ；
- (2) 若对于所有  $i \in I$ ， $u_i > u_i'$ ，则  $W(\mathbf{u}) > W(\mathbf{u}')$ ；

那么我们称社会福利函数  $W(\cdot)$  满足帕累托性质。如果社会福利函数  $W(\cdot)$  满足如下两个性质：

- (1) 若对于任意  $i \in I$ ， $u_i \geq u_i'$ ，则  $W(\mathbf{u}) \geq W(\mathbf{u}')$ ；

(2) 若至少存在一个  $i \in I$ ，使得  $u_i > u_i'$ ，则  $W(\mathbf{u}) > W(\mathbf{u}')$ ；那么我们称社会福利函数  $W(\cdot)$  满足严格帕累托性质。

如果社会福利函数满足严格帕累托性质，则称社会最优契约  $\pi^*$  是帕累托最优契约。在实践中，最为常用的满足严格帕累托性质的社会福利函数是功利主义。功利主义社会福利函数认为，社会福利等于个人效用之和(直接的代数和)。

**定义 3.2** 在公共参与者博弈  $G^\#$  中，若存在  $\pi \in \mathcal{E}^\#$ ，使得

$$\pi \in \arg \max_{\pi' \in \mathcal{E}^\#} \sum_{\mathbf{a} \in A^\#} \pi'(\mathbf{a}) \sum_{i \in I} u_i(\mathbf{a}),$$

则称  $\pi$  是一个功利主义最优契约。

设博弈  $G^\#$  的所有功利主义最优契约的集合为  $\mathcal{U}^\#$ 。功利主义已受到广泛的批评，与之对应的另外一个应用同样十分广泛的社会福利函数是罗尔斯主义社会福利函数，其认为社会福利水平取决于处境最差者的效用水平。

**定义 3.3** 在公共参与者博弈  $G^\#$  中，若存在  $\pi \in \mathcal{E}^\#$ ，使得

$$\pi \in \arg \max_{\pi' \in \mathcal{E}^\#} \sum_{\mathbf{a} \in A^\#} \pi'(\mathbf{a}) \min\{u_i(\mathbf{a})\}_{i \in I},$$

则称  $\pi$  是一个罗尔斯主义最优契约。

设博弈  $G^\#$  的所有罗尔斯主义最优契约的集合为  $\mathcal{R}^\#$ 。注意，罗尔斯主义社会福利函数虽然满足帕累托性质，却未必满足严格帕累托性质。因此，罗尔斯主义最优契约未必是帕累托最优契约。除了功利主义和罗尔斯主义之外，由纳什提出的所谓“纳什谈判解”也提供了一个非常重要的社会福利函数。这个社会福利函数认为，社会福利应当等于个人净增效用之积。特别地，在纳什社会福利函数中，假设存在着一个谈判的初始点。在公共参与者博弈中，我们可以将此初始点设为博弈中各参与者所能得到的相应的最差结果（保留效用）。

**定义 3.4** 在公共参与者博弈  $G^\#$  中，若存在  $\pi \in \mathcal{E}^\#$ ，使得

$$\pi \in \arg \max_{\pi' \in \mathcal{E}^\#} \sum_{\mathbf{a} \in A^\#} \pi'(\mathbf{a}) \prod_{i \in I} (u_i(\mathbf{a}) - \min\{u_i(\mathbf{a}')\}_{\mathbf{a}' \in A}),$$

则称  $\pi$  是一个纳什最优契约。

设博弈  $G^\#$  的所有纳什最优契约的集合为  $\mathcal{N}^\#$ 。需要指出的是，纳什最优契约也未必是严格帕累

托最优契约。这是因为，可能存在某个参与人，使得其相对效用水平为 0。如果所有参与人的相对效用水平都严格大于 0，则纳什最优契约一定是严格帕累托最优的。

### 3.2 社会最优契约排序

此时，根据上述三个社会福利函数<sup>①</sup>来计算表 1 所示博弈的最优社会契约。计算可分为三个步骤。首先，三种社会福利函数在行动空间上所对应的社会福利水平如下表所示：

表 4 扩展的情侣博弈的社会福利空间

2	0	0
0	6	0
0	0	8

(1)

1	0	0
0	1	0
0	0	4

(2)

1	0	0
0	5	0
0	0	16

(3)

其中，表 4 的 (1)、(2) 和 (3) 分别由功利主义、罗尔斯主义和纳什社会福利函数得到。然后，该博弈的七个纳什均衡所对应的行动分布(包括混合策略)分别如下表所示：

表 5 扩展的情侣博弈的纳什均衡（纯策略与混合策略）

1	0	0
0	0	0
0	0	0

(1)

0	0	0
0	1	0
0	0	0

(2)

0	0	0
0	0	0
0	0	1

(3)

5/12	5/12	0
1/12	1/12	0
0	0	0

(4)

16/25	0	4/25
0	0	0
4/25	0	1/25

(5)

0	0	0
0	16/45	4/45
0	20/45	5/45

(6)

80/261	80/261	20/261
16/261	16/261	4/261
20/261	20/261	5/261

(7)

然后，根据表 4 和表 5，可以得到三种社会福利函数所对应的各纳什均衡的社会福利水平如下表所示：

表 6 纳什均衡的社会福利评价

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
功利主义	2	6	8	4/3	8/5	136/45	296/261
罗尔斯主义	1	1	4	1/2	4/5	4/5	116/261
纳什社会福利	1	5	16	5/6	32/25	32/9	240/261

最后，由表 6 可知，功利主义对于七个纳什均衡的社会排序如下：

$$(3) > (2) > (6) > (1) > (5) > (4) > (7),$$

故功利主义最优契约为 (3)。再根据罗尔斯主义社会福利函数，可知

<sup>①</sup> 通俗点儿讲：功利主义社会福利就是个体成员效用的直接相加；罗尔斯主义社会福利则取决于社会成员中处境最差者的效用，反映出某种特定的公平倾向，实质上是外生的赋予权重的加权和；纳什最优（协商解）是乘积形式的联合效用最大化。与功利主义、罗尔斯主义和纳什协商解三种公平准则对应的总量生成方式分别为：代数/算术求和（和差关系）、最大最小值法和联合效用最大化的乘积关系。

$$(3) > (1) = (2) > (5) = (6) > (4) > (7),$$

故罗尔斯主义最优契约为 (3)。同样地，由纳什社会福利函数可知

$$(3) > (2) > (6) > (5) > (1) > (7) > (4),$$

故纳什最优契约亦为 (3)。由此说明，集体理性与社会最优契约的达成，可以在个体偏好一致或同序、<sup>①</sup>信息充分交流沟通的情况下实现；公共参与人则可通过提出共同行为准则建议，引导个体增强互信和交流，改善可能的均衡结果的权重分布，最终实现社会合意行动。

#### 4 社会最优契约的基本特征与现实意义

在上述案例的讨论中，三个不同的社会福利函数却得到了相同的社会最优契约。特别地，其中功利主义社会福利函数与纳什社会福利函数的社会排序结果高度接近。但是，这一结果并不具有普适性。因此，有必要分类讨论现实中可能存在的不同的社会福利函数及其相应的分配原则与政策选择有何种不同的含义。

##### 4.1 社会福利函数的线性

**定义 4.1** 在公共参与人博弈  $G^{\#}$  中，对于任意  $\pi \in \Delta(A_0) \times \Delta(A_1) \times \dots \times \Delta(A_N)$ ，有

$$\sum_{a \in A} \pi(a) W(u_1(a), \dots, u_N(a)) = W(\sum_{a \in A} \pi(a) u_1(a), \dots, \sum_{a \in A} \pi(a) u_N(a)),$$

则称该博弈的社会福利函数是线性的。

也就是说，如果社会福利函数是线性的，则对于任意  $\pi \in \Delta(A_0) \times \Delta(A_1) \times \dots \times \Delta(A_N)$ ，有

$$E[W(u_1(a), \dots, u_N(a))] = W(E[u_1(a)], \dots, E[u_N(a)]).$$

很明显，功利主义社会福利函数具有此线性性质，这是因为

$$\begin{aligned} \sum_{a \in A} \pi(a) W(u_1(a), \dots, u_N(a)) &= \sum_{a \in A} \pi(a) \sum_{i \in I} u_i(a) \\ &= \sum_{a \in A} \sum_{i \in I} \pi(a) u_i(a) \\ &= \sum_{i \in I} \sum_{a \in A} \pi(a) u_i(a) \\ &= W(\sum_{a \in A} \pi(a) u_1(a), \dots, \sum_{a \in A} \pi(a) u_N(a)). \end{aligned}$$

社会福利函数的线性特征可以理解为理论上的规模不变性与实践中的政策延续性，社会中每一成员的效用按相同比例增减时，社会福利函数的序位不变。

但是罗尔斯主义与纳什社会福利函数则不具有此性质。以表 4 所示的博弈为例，其各纳什均衡所对应的期望效用组合分别如下：

表 7 纳什均衡的期望效用组合

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$E[u_1]$	1	1	4	1/2	4/5	4/5	116/261
$E[u_2]$	1	5	4	5/6	4/5	20/9	180/261
$E[u_1] + E[u_2]$	2	6	4	4/3	8/5	136/45	296/261
$\text{Min}\{E[u_1], E[u_2]\}$	1	1	4	1/2	4/5	4/5	116/261
$E[u_1] \times E[u_2]$	1	5	4	5/12	16/25	80/45	160 \times 180 / (261^2)

不难发现，在上述七个纳什均衡效用组合中，只有 (2) 和 (3) 是帕累托最优契约。由此可知，功利主义满足线性，而纳什社会福利函数则不满足。特别地，罗尔斯主义在此例中也满足线性，这

<sup>①</sup> 当由  $u_i(a_i)$  扩展到  $u_i(a_i, a_{-i})$  为保序变换时，对任意  $i \in I$  均成立，两者在相同的点取得极值。

是因为在此例中，

$$\forall \mathbf{a} \in A, \text{Min}\{u_1(\mathbf{a}), u_2(\mathbf{a})\} = u_1(\mathbf{a}),$$

从而对于任意  $\pi \in \Delta(A_1) \times \Delta(A_2)$ ，可知

$$\begin{aligned} \sum_{\mathbf{a} \in A} \pi(\mathbf{a}) W(u_1(\mathbf{a}), u_2(\mathbf{a})) &= \sum_{\mathbf{a} \in A} \pi(\mathbf{a}) \text{Min}\{u_1(\mathbf{a}), u_2(\mathbf{a})\} \\ &= \sum_{\mathbf{a} \in A} \pi(\mathbf{a}) u_1(\mathbf{a}) \\ &= \text{Min}\{\sum_{\mathbf{a} \in A} \pi(\mathbf{a}) u_1(\mathbf{a}), \sum_{\mathbf{a} \in A} \pi(\mathbf{a}) u_2(\mathbf{a})\} \\ &= W(\sum_{\mathbf{a} \in A} \pi(\mathbf{a}) u_1(\mathbf{a}), \sum_{\mathbf{a} \in A} \pi(\mathbf{a}) u_2(\mathbf{a})). \end{aligned}$$

但是在更一般的情况下，功利主义社会福利函数仍然具有线性性质，而罗尔斯主义社会福利函数与纳什社会福利函数一样，则不具有线性性质。

## 4.2 社会最优契约的不变性

**定理 4.1** 假设在博弈  $G = (I, A, U)$  中，混合策略组合  $\alpha^* \in \Delta(A_1) \times \dots \times \Delta(A_N)$  是博弈  $G$  的纳什均衡。如果在博弈  $G' = (I, A, U')$  中，对于任意  $i \in I$ ， $u'_i = f_i(u_i) \in U'_i$ ，其中  $f_i$  为任意正线性（或正仿射）变换，那么  $\alpha^*$  是博弈  $G'$  的纳什均衡当且仅当  $\alpha^*$  是博弈  $G$  的纳什均衡（证明略）。

由定理 4.1 可知，在博弈  $G = (I, A, U)$  中，纯策略纳什均衡具有序数不可比性，混合策略纳什均衡具有基数不可比性。但是，阿罗不可能定理使得这一性质在公共参与人博弈  $G^\#$  中对于社会最优契约不再成立。也就是说，社会最优契约一般不满足序数不可比性。此外，即使允许某种程度上的人际比较，我们依然面临一个本质上的难题，即人际比较需要极为丰富的效用信息。为了解决这一难题，我们不再对个人效用函数的任意正单调变换，而是集中于考虑线性变换对于社会福利函数的影响。研究结果已经表明，不同的社会福利函数对于不同类型的线性变换具有不同的不变性特征 (Sen, 1970, 1974; d'Aspremont and Gevers, 1977; Roberts, 1980)。

**定理 4.2** 假设在公共参与人博弈  $G^\# = (I^\#, A^\#, U^\#)$  中， $\pi \in \mathcal{N}$  是一个纳什最优契约。如果在公共参与人博弈  $G^{\#'} = (I^\#, A^\#, U^{\#'})$  中，对于任意  $i \in I^\#$ ， $u'_i = a_i u_i + b_i \in U^{\#'}$ ，其中  $a_i$ 、 $b_i$  为任意实数且  $a_i > 0$ ，那么  $\pi$  是博弈  $G^{\#'}$  的纳什最优契约当且仅当  $\pi$  是博弈  $G^\#$  的纳什最优契约。

证明略。由定理 4.2 可知，纳什最优契约满足基数不可比性。

**定理 4.3** 假设在公共参与人博弈  $G^\# = (I^\#, A^\#, U^\#)$  中， $\pi \in \mathcal{U}$  是一个功利主义最优契约。如果在公共参与人博弈  $G^{\#'} = (I^\#, A^\#, U^{\#'})$  中，对于任意  $i \in I^\#$ ， $u'_i = a u_i + b \in U^{\#'}$ ，其中  $a$ 、 $b_i$  为任意实数且  $a > 0$ ，那么  $\pi$  是博弈  $G^{\#'}$  的功利主义最优契约当且仅当  $\pi$  是博弈  $G^\#$  的功利主义最优契约。

证明略。由定理 4.3 可知，功利主义最优契约满足基数单位可比性。

**定理 4.4** 假设在公共参与人博弈  $G^\# = (I^\#, A^\#, U^\#)$  中， $\pi \in \mathcal{R}$  是一个罗尔斯主义最优契约。如果在公共参与人博弈  $G^{\#'} = (I^\#, A^\#, U^{\#'})$  中，对于任意  $i \in I^\#$ ， $u'_i = a u_i + b \in U^{\#'}$ ，其中  $a$ 、 $b$  为任意实数且  $a > 0$ ，那么  $\pi$  是博弈  $G^{\#'}$  的罗尔斯主义最优契约当且仅当  $\pi$  是博弈  $G^\#$  的罗尔斯主义最优契约（证明略）。<sup>①</sup>

由定理 4.4 可知，罗尔斯主义最优契约满足基数完全可比性。

社会最优契约不变性的政策蕴意是社会总福利不受财富形式和分配方式的影响，即独立于所选择的价值尺度，与度量工具无关。如果我们更加关注社会福利函数的线性性质，功利主义社会福利函数无疑是最佳选择。但如果我们希望强调社会福利函数的不变性，那么纳什社会福利函数则是最佳选择。因此，社会福利函数的线性与不变性特征不可兼得，这无疑给公共参与人博弈的求解过程

<sup>①</sup> 对本节中定理证明感兴趣者，可参见学术专著：《公共参与人博弈与社会最优契约》，中国社会科学出版社，2013 年。

带来了无法避免的困难，这也是实践中公平分配原则难以确立和实施的理论原由。

### 4.3 社会最优契约的协调性

由于纳什均衡的概念可以推广到相关均衡，那么社会最优契约的概念也可自然地推广到相关均衡。事实上，相关均衡的概念为参与人之间的相互协调提供了更多的可能性。我们借助交通博弈和相关均衡的概念来说明社会最优契约应具备的协调性。

关于相关均衡的一个经典例子是“交通博弈”：在一个十字路口，参与人 1 和参与人 2 驾车相遇；双方可选择让路或者不让路，若双方都选择不让路则将发生交通事故。该类博弈的支付结构如下表所示。

表 8 交通博弈

		2	
		让路	不让路
1	让路	0, 0	0, 1
	不让路	1, 0	-10, -10

该博弈有三个纳什均衡，即（让路，不让路），（不让路，让路），以及双方都以 10/11 的概率让路、以 1/11 的概率不让路，其各自所对应的行动分布分别如下表所示：

表 9 交通博弈的纳什均衡

0	1	0	0	100/121	10/121
0	0	1	0	10/121	1/121
(1)	(2)	(3)			

如果双方可以进行事先协调，那么协调的结果必定是这三个纳什均衡中的其中一个，而使得没有人会单方面主动偏离协调的结果。但是，在此例中，我们还可更进一步考虑在十字路口设置一个红绿灯装置的情况，使得双方总是以一定的概率看到其他方向道路上行使车辆的信号，并且共同遵守“红灯停、绿灯行”的交通规则。此时，“红灯停、绿灯行”便构成了一个所谓的相关均衡，从而避免了（让路，让路）与（不让路，不让路）等具有纳什均衡意义的行动组合。一般情况下，设（让路，不让路）和（不让路，让路）发生的概率均各为 1/2，则行动分布如下表所示：

表 10 “红绿灯”相关均衡

0	1/2
1/2	0

在这个相关均衡中，没有人敢单方面主动偏离“红灯停、绿灯行”的约定，而这与纳什均衡的思想极为一致。<sup>①</sup>只不过，相关均衡的概念事实上将参与人策略的随机化过程一般化了。因此，任何一个混合策略都赋予行动空间一个概率分布，即一个能在相应的群体间进行协调的行动分布，可称行动空间上与任意一个概率分布相应的协调行动分布为相关策略组合，对应的博弈均衡即为相关均衡，反映出经过协调使群体收益提高，从而有可能增加每一参与人的收益或效用。

由此可见，混合策略即为相关策略要求各参与人的策略分布是相互独立的特殊情形。根据各自的期望支付，相关策略均衡（亦即相关均衡）使得各参与人都没有单方面偏离均衡策略的动机。而交通灯的设置实质上是行使了公共参与人的职能，协调所有参与人的行动达成相关均衡。经过简

<sup>①</sup> 该示意性案例中的支付结构和概率假设是对真实交通状况的一种近似替代，当然，实际中红绿灯的时长控制是要根据交错通过路口的车流量比例而定。

---

单的计算比较，可得知交通博弈中相关均衡的社会福利评价，并由此可以看出，与相关均衡对应的社会福利评价要优于某些纳什均衡对应的社会福利，如在罗尔斯主义和纳什社会福利函数中，“红绿灯”相关均衡都要优于所有的纳什均衡。<sup>①</sup>

在相关均衡意义上表述社会最优契约的整体协调性，是指集体行动中有与相关均衡相应的实现帕累托改进意义上的社会福利最大化的可能性，如前面表 5 中的 (7)，公共参与人的引入，有助于社会福利最大化的实现。本节中分析阐述的社会最优契约所具有的性质，表明改革实践中的收入再分配必须要有明确的价值取向和相应的政策措施，为此寻求有力的理论解释和可行的解决途径，也对基于公共利益调节收入分配差距和促进和谐社会建设等有积极的启发意义。

## 5 计算社会合意行动的技术实现及应用

理性个体独立决策的优化行为和均衡求解，是数学规划中典型的最优解问题，<sup>②</sup>而集体行动可通过数值模拟或计算实验等技术手段实现理论上相对最优的渐近解，并与现实问题平行地交互验证。集体行动的建模与模拟技术的发展，有力地促进了经济理论的发展，展现特定情境和时期下的经济活动方式、人类认知能力和科技发展水平的交互作用，得出相应的结论；同时也催生了社会计算这类新的分支学科和技术。

### 5.1 集体行为计算

现代高性能计算和大数据分析处理等技术，使得对从个体（优化）行为到群体（合意）行动的聚变和传导机理，以及结构演变过程的数值计算由不可能变为可能。从异质性个体及相互之间的行为关系入手，着重探讨人文复杂性，利用（人文）复杂性建模和数值模拟技术方法，拓展传统的计量实证，将经验数据与用实验方法获取的数据相结合，扩宽数据来源、丰富数据内容、提升数据质量，将对经典理论的验证和新规律的发现提炼结合、共同促进；以及方法（论）上的演绎与归纳、还原论与整体论、解析最优解与模拟近似解技术、描述性与构造性求解的融合；由个体基本简单的属性入手，经过对组织结构和传导机理的过程分析，发现整体规律性特征及其如何影响个体决策行为；虽然总体呈非单一的多形态表现/涌现，其具体形态会随个体行为属性赋值的不同发生相应的变化，与个体行为属性参数或许存在着内在对应关系（函数）；因而，借助分布式计算技术来实现，是从个体层面（单一行为属性假设）上升到集体层面（多种形态的涌现或典型化事实）的固有其本征和必然需求。

由于计算技术的发展和数据的积累（尤其是大数据概念的提出及相应行动），为在微观层面上多方位、系统深入地分析主体行为，认知复杂多变的宏观形态奠定了良好基础。综合系统（动力学）仿真、微观模拟、数值逼近和基于（多）主体建模（ABMs）的优势，并且引入能更好地满足人文社会科学研究需要和更具有针对性的计算实验方法，开辟了新的更广阔的学术视阈。人类集体行为的建模和模拟，对于认知个体行为是如何导致群体层面上的组织形式变换和结果涌现，无疑是大大提高了经验性学科的定量化和科学解释力。近年来，侧重描述众多个体之间、与环境之间交互的ABMs技术方法，提供了过程导向、非程式化的、可条件性选择的并且是直观描述性与逻辑构造性相结合

---

<sup>①</sup> 本节中严格的数理分析证明和示意性算例可参见学术著作：《公共参与人博弈与社会最优契约》，中国社会科学出版社，2013年；有兴趣的读者也可直接与作者联系。

<sup>②</sup> 相关均衡的求解通常是求解不等式组，在公共参与人博弈中就成了典型的数学规划问题，因此相关均衡通常比纳什均衡具有更好的可计算性质。但我认为，纳什均衡一般来说不具有良好的可计算性质。

---

的数学模型，其在解释群体行为方面的成功发展及应用非常令人瞩目(Goldstone and Janssen, 2005)，尤其是在规模、结构和过程以及群体响应模式、不同主体/子群体之间的传染扩散和传导机理、合作行为发生的诱因与持续条件等方面取得了突破性进展，不仅能解释、预测，而且还能实时地定位操控；不仅在很大程度上缓解传统理论中基本行为假设的局限性，而且也将个体作为认知的基本单元扩展到群体水平上，无论是自觉或不自觉、可观察或不可观察，群体组织中的个体相互之间、与组织及其外界条件之间，都在影响和被影响着；如此的计算建模和分布式数值模拟，为探讨个体行为选择与社会行动意愿之间的复杂关系提供更有说服力的注解，开辟了符合逻辑、切实可行、相对有效、直观可信的科学检验方式和实现途径。

当现实形态与传统理论中单一行为属性的经典假设和黑箱式处理方法相去甚远、出现了难以用技术性手段弥补或消除的显著性偏离时，必要与必须的可行做法是：相比而言应更加强调和着重对真实行为的属性和表现的分析，以及与宏观环境条件的相互作用关系；作为（多情景下多重）博弈均衡的集体行动的计算、组织结构相应的权重分布、动态演变过程、内生化的个体行为参数，一方面是需要在这些基础环节上展开集体行动计算，另一方面在各个环节上所表现出的问题也可内化于主体行为中在集体行动计算中一并解决。集体行动计算可重点验证几个方面：（1）个体行为属性及演变，尤其是合作行为；（2）组织形态和响应行为模式；（3）扩散机理和传导路径；（4）个体行为与宏观涌现的联动关系。类比委托代理理论中具有更广适应性的泛函分析机制和模型方法，只有当委托代理关系中授权方向和方式决定的各类主体的行为单一稳定时，模型中待定参数相对集中突出，分布函数的参数化方法(Mirrlees, 1976; Holmstrom, 1979)才更为有效。具体到前述的功利主义、罗尔斯主义和纳什谈判解之类的社会选择问题，集体行动计算以其独特优势为之提供可行有效的技术实现途径。

一般的模型方法相应地是服务于一定的理论及所研究的现实问题，基于单一行为属性（无论是完全理性还是有限理性）的优化模型和均衡分析方法，对应于完全竞争市场；而分布式模拟或仿真、计算建模等更适合日趋复杂的现实经济，能在很大程度上克服传统计量实证方法在样本选择性偏差、共线性或互为因果、波动积聚及转折分岔等方面的局限性。由于人文社会科学研究中人类行为主导的本质特征和主观性特色等，在ABMs等技术工具的支持下，在深化行为分析基础上进行自下而上与自上而下结合的微观宏观一体化的建模与计算实验，适宜的模型方法无疑有助于剖析和揭示个体与群体之间复杂的行为关系(王国成，2013)。

基于单一行为属性假设考察因果关系的传统计量实证方法，容易在处理利用经验数据/数字掩盖或忽略许多真实而有价值的信息；而群体行动的计算模拟/实验方法，相对地能更准确地找到问题的要害与关键点，深入到机理分析层面和结构演变、全方位全过程的可视化再现，大大减轻了对一般性结论的偏倚和对特殊现象的舍弃；经验数据对典型化事实的支持乏力，但运用实验方法中的样本选择性、大规模模型的协调性和计算量巨大及由此引起的可信性与有效性等问题，仍将是集体行动计算乃至ACE (Agent-based Computational Economics) 发展在未来相当长一段时期内致力攻克的主要难关。

## 5.2 初步应用体会

（1）行为-制度-增长的内在关系研究。差异化的个体行为是经济活动的源动力和分析基元，制度是个体与集体行为关系的枢纽，经济增长（总产出）是集体行动的结果；制度（广义的也包括契约和文化等）实质上是在起着引导和协调个体行为的公共参与人的职能，能够促进和调控经济增长。

---

研究中的核心观点是：主体行为多样化是制度存在的前提，制度的主要功能是协调个体与群体的行为关系；通过放松基本行为假设，多视角、多层面和多手段地对微观层面上个体行为的异质性、交互性、主观性和可变性等多属性的深化分析，构建相应的博弈模型，并辅助性地运用包括中国改革历程的经验、公共品供给案例、制度与绩效关系的计量和基于行为一体化建模的模拟等组合实证方法，探究差异性个体行为、不同的社会组织结构和形式、制度类型与增长（集体行动的结果）的共生演化关系，既能更深入、更广泛地研究制度与增长的内在关系，又表明如此的方法选择是必要的、适宜的和有效的。进而比较不同的制度设计和实施措施的现实效果，在不同程度上实现了制度这类半（弱）结构化变量相对规范的分析及数值模拟的方法与流程，试图为制度的科学定量分析开辟可行有效的新途径（实践中最需要而传统方法又难以提供的）。<sup>①</sup>

（2）收入差距演变的模拟实证。近年来，中国收入分布中的差距过大和拉大速度过快及相关问题备受关注，影响因素及相互关系错综复杂、谜团笼罩，不仅要探讨资源配置与财富分配，更有必要考察人的行为属性、主观感知等人本层面的因素；不仅是经济现象研究，还大量地触及社会、政治、文化和心理等非经济领域深层问题……。如何破解此谜局是理论和实践的重大课题，即使在动态结构框架下，基于单一的同质行为假设运用传统计量方法，也难以刻画和揭示收入差距的内在动因与演变轨迹。因而，不仅关注总体的不平等程度，更重要的是探究收入分布基本形态演变的关键点、动力学机理、方向、方式与路径，以及可能产生的社会经济影响和应对措施。从与人们利益关系最密切的、最容易受关注的人和事出发，以货币化意义上刻画和度量收入分配为基础，还要考虑主观情感和认知等人文因素，不仅是一般的理论与实践、定性与定量以及国际比较，而是探索微观与宏观、客观与主观、计量实证与模拟实验、行为分析与要素分析、自然的与人文的、总体指标与关键特征的一体化建模方法，着重分析揭示收入差距的成因交织和复杂机理；方法论上是将个体视角的主观满意度、公平感与整体视角的引入公共参与人、基于社会福利函数研究收入差距演变相结合，运用经典计量和计算实验模拟的组合实证，探讨微观个体和社会整体对收入差距及演变的心理阈值与结构承受限度，由此得出更具有针对性和更富有启发意义的结论及建议。<sup>②</sup>

另外，我们运用本文中所述的思想方法，还探讨了真实投资行为与股市波动之间的联动关系，基于异质性农户行为评估了惠农政策效应，开展了全球气候保护以及公共品提供的实验研究等。<sup>③</sup>如此能跳出原有分析框架的窠臼，与传统理论方法相辅相成、相互促进，厘清经济分析中归纳与演绎推理的逻辑关联，摆脱实验样本代表性和规模效应等方面的局限，更有针对性地深入地解决现实复杂问题，具有诱人的发展前景：理论上是可进可退、进退有余，既可退守基于经典行为假设、运用传统理论和计量实证方法研究常态规律性，又可进取基于微观主体多元行为属性及相互关系、探索宏观复杂（非常态）的典型化事实；在实践中能有效协调个体与集体的行为关系、不同社会阶层和利益群体之间的摩擦矛盾甚至是冲突等，在政策方向和着力点上都会有新的思路和做法。

## 结 语

在充分尊重个体选择的基础上，萃取和树立能代表集体意愿的主流价值取向、合理的公平衡量标准与评价准则、需要共同遵守的行为规则是必须的，现实生活中虽然不可能达成绝对一致的基于

---

<sup>①</sup> 详见：中国社会科学院重大科研项目（YZDA2009）的最终结项成果《行为-制度-增长：基于博弈模型的分析框架及应用》，中国社会科学出版社，2012。

<sup>②</sup> 详见：国家自然科学基金项目（70971138）的最终结项成果《收入差距演变的动因与机理 -- 行为分析与微观模拟视角》，经济科学出版社，2013。

<sup>③</sup> 国家重大科学研究计划 973 项目（课题编号：2012CB955802）的阶段成果。

---

社会福利最大化的集体行动，但又有可能和必须在群体认知和意愿的基础上凝聚形成合意行动和驱动力量，而且是需要培育、扶持和引导的，这是经济学乃至人文社会科学研究继续深化的主攻方向。本文只是基于引入公共参与人的博弈，通过比较不同意义上按不同的评判标准的社会选择与社会福利函数，剖析了社会最优契约的实质、逻辑内涵和实践特征，在为深入分析个体与群体行为关系构建基本框架、选择更为适宜的模型方法和技术实现路径等方面，做了初步的然而积极的探索。

公共参与人博弈的社会最优契约的本质特征和内在要求是：公共参与人相对最优的建议应当是稳健的。公共参与人的引入使得交流与协调更加通畅，非合作博弈的均衡求解事实上成为参与人之间事前沟通协商、相互合作的结果，这是不以人的意志为转移、无论是能观察到或不易观察到的社会存在。当人类面临竞争与冲突时总是会试图寻找最优的合作方式，公共参与人博弈的社会最优契约无疑为解决这一问题提供了一条新的思路。发现和尊重个体的异质性，明确各自的行为方式、边界和相互关系，充分调动各类主体的积极性及协调其与群体目标相一致，凝聚和形成促进社会和谐进步、经济健康发展的主流价值观，如此的共同行动必然是最根本和最强大的推动力。

在人类欲望的膨胀远远快于行为能力（财富提供）的提升这一广泛存在的约束条件下，个体决策的有限认知和效用最优的条件依赖性，任何公平都是相对于一定的价值取向和评价标准而言的，相对应的是只可能存在合意的而非绝对一致的集体行动。公共参与人博弈的一个难题是如何确定恰当的社会福利函数。理论表明，满足人们最基本的良好愿望、完美无缺的社会福利函数的存在性是空洞的；然而一旦实践中达成社会合意行动、相应地构建适当的社会福利函数，借助数学方法和计算求解工具，相对意义上的社会最优契约是可认知和可实现的。

## 参考文献

- 王国成. 2013. 深化微观行为分析探索复杂经济问题——HS与CA结合的实现技术. *江苏社会科学*, (3):20~28.
- Arrow, K. J. 1951. *Social Choice and Individual Values*, 1st edition, New York: Wiley.
- Bentham, J. 1789. *An Introduction to the Principles of Morals and Legislation*, London: Payne.
- Bergson, A. 1938. A reformulation of certain aspects of welfare economics. *Quarterly Journal of Economics*, 52: 310~334.
- Binmore, K.; Rubinstein, A.; Wolinsky, A. 1986. The Nash Bargaining Solution in Economic Modelling. *RAND Journal of Economics* 17: 176~188.
- Brandenberger, A. and E. Dekel. 1987. Rationalizability and correlated equilibrium. *Econometrica*, 55(6): 1391~1402.
- d'Aspremont, C. and L. Gevers. 1977. Equity and informational basis of collective choice. *Review of Economic Studies*, 44:199~210.
- Goldstone R. and M. A. Janssen. 2005. Computational models of collective behavior. *TRENDS in Cognitive Sciences*, 9(9): 424~430.
- Hart, Oliver and Moore, John. 1988. Incomplete Contracts and Renegotiation. *Econometrica*, 56(4):755~785.
- Holmstrom B. 1979. Moral Hazard and Observability. *The Bell Journal of Economics*, 10(1):74~91.
- Mirrlees J. 1976. The Optimal Structure of Incentives and Authority within an Organization. *Bell Journal of Economics and Management Science*, 7(1) :105~131.
- Myerson R. B. 2008. Mechanism design. *The New Palgrave Dictionary of Economics*.
- Nash, J. F. 1950(a). The bargaining problem. *Econometrica*, 18:155~162.
- Nash, J. 1950(b). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences of the United States of America*, 36:306~310.

---

States of America, 36(1): 48~49.

Nowak MA. 2006. Five rules for the evolution of cooperation. *Science*. 314: 1560~1563.

Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

Rawls, J. 1971. *A Theory of Justice*. Cambridge: Harvard University Press, and Oxford: Clarendon Press.

Rees, R., 1985. The Theory of Principal and Agent—Part I, Part II. *Bulletin of Economic Research*, 37(1), 3~26.

Roberts, K. W. S. 1980. Interpersonal comparability and social choice theory. *Review of Economic Studies*, 47:421~439.

Samuelson, P. A. 1947. *Foundations of Economic Analysis*. Cambridge: Harvard University Press.

Sen, A. K. 1970. *Collective Choice and Social Welfare*. North-Holland, Amsterdam.

Sen, A. K. 1974. Informational bases of alternative welfare approaches: Aggregation and income distribution. *Journal of Public Economics*, 3:387~403.

Weibull J. W. 1995. *Evolutionary Game Theory*. MA: MIT Press.