

## 经济机制理论：信息效率与激励机制设计

田国强\*

**摘要** 本文主要对机制设计理论的一些基本结果及最新发展作一大致介绍。我们将介绍经济机制理论的两个领域：激励相容问题和信息效率问题。不同的机制会导致不同的信息成本、不同的激励和不同的配置结果。研究一个经济机制的信息有效性和激励相容性可评价其制度的优劣性。机制理论系统地研究经济制度的设计和这些制度是如何影响人们的互动行为和配置结果的。

**关键词** 经济机制，激励相容，信息效率

### 一、引言

经济机制理论主要研究，在自由选择、自愿交换、信息不完全及决策分散化的条件下，能否设计一套机制(游戏规则或制度)来达到既定目标，并且能够比较和判断一个机制的优劣性。世界上有许多(包括理论上)各式各样的经济制度，如市场经济机制、计划经济机制、公有制、私有制、集体合作制、混合所有制、边际成本定价机制等。那么，什么样的经济机制是好的呢？这是一个自上世纪20—30年代以来学术界一直在争论和想要回答的问题，并且在当今转型经济国家中还在争论的一个基本问题。比如中国大陆自改革开放以来，成为全世界经济发展最快的国家，年平均增长率达到9.7%左右，其中，集体经济是中国经济增长的主要贡献力量。于是不少人相信非私有经济可以像私有经济表现得一样好，或更好，而另外一些人却不同意此看法。

在讨论和判断一个经济制度的优劣时，人们需要首先给出评价一个经济制度优劣的标准。在经济学文献中，经济学家通常认为一个好的经济制度应满足三个要求：它导致了资源的有效配置、有效利用信息及激励相容。机制的有效资源配置要求资源得到有效利用，有效利用信息要求机制的运行具有尽可能低的信息成本，激励相容要求个人理性和集体理性一致。这些要求是评价一个经济机制优劣和选择经济机制的基本判断标准。如果一个资源配置不是有效的，则存在着资源的浪费和改进经济效率的余地；如果信息的利用不是有效的，机制运行的成本就比较大；如果一个机制不是激励相容的，个人在追求私利时可能会违背集体利益或影响社会目标的实施，从而个人利益与

\* 清华大学经济管理学院经济系和德州农工大学经济系。通信地址：Guoqiang Tian, Department of Economics, Texas A&M University, College Station, Texas 77843; 电话：(1-979) 845 7393; Email: gtian@tamu.edu。此文更完整的版本已收集在由田国强主编，即将由商务印书馆出版的《现代经济学、金融学前沿发展》论文集中。

集体或社会利益发生冲突。由于不同的经济机制会导致不同的信息成本、不同的激励反应、不同的配置结果,因而人们需要知道什么样的经济制度能满足以上三个要求。这样,仅仅把一个个机制分开考虑是不够的。当各种经济机制共存,可供选择时,一个国家则需要对经济制度做出选择。在所实施的经济制度出现问题时,人们也总想知道是否还存在着其他更好的经济制度。另外,在现实中,经济环境总是在不断地发生演变,特别在经济、社会制度转型时期更是如此,从而人们需要对制度进行选择或创新。这样我们需要有一个能研究和比较各种经济机制的更一般的理论来考虑制度的选择问题。在这个理论模型下,经济机制不必看成是给定的,而是未知、可设计的,并且在一定的标准下(如以上所提到的三个标准)可以研究和比较各种(已知和未知)经济机制的优劣。此外,人们所面临的是一个信息不完全的社会。由于任何人特别是上级部门没有,也不可能掌握其他人的所有私人信息,从而在直接指导社会经济活动时就会遇到很大的问题(如果所有人个人信息能够被全部掌握的话,直接控制或强制命令的集中化决策(比如像计划经济)方式就不会有问题。它可归结为一个简单的优化问题)。正是由于所有的个人信息不可能完全被一个人掌握,人们才希望分散化决策。用激励机制或规则这种间接控制的分散化决策方式来激发人们做设计者(制度或规章制定者)想做的事情,或实现设计者想达到的既定目标。这正是经济机制设计理论所要探讨的问题。

机制设计理论能系统地比较各种经济制度的优劣和研究不同的经济制度是如何影响人们的相互行为和资源配置结果的。经济机制理论把所有的经济机制放在一起进行研究。研究的对象大到对整个经济制度的一般均衡设计,小到对某个经济活动的激励机制设计。这个理论的基本框架是由美国经济学家利奥·赫维兹(Leo Hurwicz)最先严格给出的。它可用来研究和探讨各种经济问题,特别是在不完全信息情况下探讨和设计各种激励机制,以实施(implement)所要达到的社会或某个既定目标。概括地说,经济机制设计理论所讨论的问题是:对于任意的一个想要达到的既定目标,在自由选择、自愿交换的分散化决策条件下,能否并且怎样设计一个经济机制(即制定什么样的方式、法则、政策条令、资源配置等规则)使得经济活动参与者的个人利益和设计者既定的目标一致,即每个人主观上追求个人利益时,客观上也同时达到了机制设计者既定的目标。如可能的话,是否具有较小的信息运行成本。机制设计目标可以非常大,也可以非常小。大到可以是对整个经济社会的制度的设计,其目标是一个经济整体目标。也可以小到只是具有两个参与者的经济组织管理的委托人的目标,其目标只是他自己的最优利益。

本文对机制设计理论产生的背景、基本结果及最新发展作一大致介绍。需要提到的是本综述不是非常完整,对文献的引用具有很大的取舍性,许多没有被提到的结果并不表示它们不重要,只是个人的偏好及篇幅所致。

## 二、背景

我们前面提到经济机制的信息有效性和激励相容性是评价一个经济机制优劣的基本标准。对一个机制的信息有效性和激励相容性的研究是由上世纪30年代关于市场社会主义经济机制可行性的大论战引发的，争论的内容恰与当前中国经济改革所遇到的问题类似。上世纪20—30年代有一场非常著名的论战，称之为社会主义大论战。一批反对社会主义的经济学家试图证明社会主义在理论上是行不通的。他们的主要代表人物是米塞斯(Mises)和哈耶克(Hayek)。他们批评社会主义，不是针对社会主义理想是否合理，而是认为社会主义不可能获得维持经济有效运转的信息。他们把社会主义经济机制看作是一个高度集中的中央计划，每一个基层单位或企业向中央机构传送有关技术、成本、消费需求等方面的信息，再由中央计划机构制定出非常详细的计划并下达给企业。这样，中央计划机构需要知道消费者的偏好，企业的生产技术条件，并且要有解出数以百万计以上的供给和需求联立方程组的能力，即使在计算机非常发达的今天，这也是一件很困难的事。即使能知道这些信息并能解出这些方程式，由于收集信息和计算供求结果所需时间过长，人们的消费偏好和企业的技术条件也许早已发生了变化。所以他们认为经济社会不可能获得社会主义计划所需要的信息并合理的使用这些信息。论战另一方面的主要代表人物是兰格(Lange)和雷纳(Lerner)。他们认为即使在社会主义条件下人们仍然可以利用市场机制。他们的主张是：虽然生产资料收归国有，但资源的流动还应由供求关系确定(他们所说的资源不包括投资，仅仅只对消费领域而言)。对于企业而言，每个企业应该根据边际成本等于中央计划委员会所制定的产品价格来确定生产水平。在一定生产技术条件下，在数学上可以证明这种机制可导致资源的有效配置。兰格和雷纳所建议的其实是一种分散化的社会主义经济机制，或者说是市场社会主义经济机制。这种机制旨在解决信息要求过大的问题。以米塞斯和哈耶克为首的一批人认为社会主义式的计划经济机制不可能获得维持经济正常运转的信息。而以兰格和雷纳为代表的另一批人认为可以通过边际成本定价的方式来解决信息成本巨大的问题。兰格的这种分散化市场社会主义机制可能解决了信息量要求过大的问题，但它本身又产生了另外一个问题，那就是激励相容问题，也就是怎样激励基层单位完成上级计划部门下达的任务并且按照真实的边际成本订价来组织生产。由于边际成本是私人信息，上级部门不可能完全清楚。这样，企业为了更容易地完成上级下达的生产指标或利润，企业就会有激励高报生产成本，使得上级部门下达较低的生产指标，且能制定更高的产品价格。并且，当规模报酬递增的生产情况发生时，生产边际成本小于平均成本。如果按照边际成本定价，企业就会亏损，长久下去，企业就要破产。如果这种生产是必要的，即使在资本主义国家，也需实行补贴。但是对企业的补贴会引起许多其他的问题，其中之一就是财政问题，因为这些补贴要从其他企业上缴的利润(或税)中拿出来。另一个问题就是企业的激励(积极性)问题。如果

企业亏损了,政府会给他们补贴,那企业就缺乏提高效率的激励。这种情况说明:为了使整个社会提高效率而给予企业的补贴在客观上反而降低了企业内部的效率。分散化的社会主义经济或者是市场社会主义经济(在理论上或许能导致有效的资源配置)并没有解决激励问题,因此哈耶克他们认为兰格的设想仍然是不可行的。可以看出,他们争论的问题和今天中国改革的走向有很大的关系。

另外,资源配置机制理论的产生不仅与社会主义经济模式有关,也与资本主义经济模式有关。传统的经济分析把经济机制看作是给定的。比如新古典微观经济学主要把市场机制作为对象来进行研究,它讨论市场机制如何运转、有什么样的优越性及局限性。对计划经济机制的讨论也是如此。在早期的文献中,如 Bergson (1938)、Lange (1942)、Lerner (1944)、Arrow (1951)、Debreu (1959) 以及 Arrow 和 Hahn (1971),人们所讨论的中心问题是将某个给定的经济机制作为研究对象(例如,给定竞争市场机制),探讨在什么样的经济环境(即对于什么样的生产技术、消费者偏好、初始资源)下,它能导致帕累托最优(有效)资源配置。西方现代经济学的大多数研究都是从市场的角度研究最优资源配置的。然而,市场机制也有它的局限性。我们知道,在许多情况下(如不完全竞争市场、生产的外部性、公共商品、不完全信息市场、按规模报酬递增或不可分商品等),市场不能导致有效的资源配置。<sup>1</sup>因此,我们既要看到市场机制的优越性,又要看到它的局限性。在讨论其局限性时,仅仅只指出市场不能良好的运行是不够的,还需寻找其他方法或机制替代或改进市场的作用。于是人们想到也许有什么补救办法,即是否存在其他的经济机制,它能产生资源的有效配置。更一般地说,对于给定的经济环境类某个社会目标(这个社会目标可以是资源的有效配置、某种意义下的公平或公正配置,或某个其他配置),是否存在某个机制(配置规则),使得每个人即使追求个人目标,其客观效果正好能达到既定的社会目标。例如,我们知道在一般的情况下,完全竞争市场机制产生了资源的有效配置。那么是否还存在其他机制(如社会主义计划经济机制)同样地也产生资源的有效配置呢?如果回答是肯定的,这个机制是否能用比竞争机制更少的信息或成本来实现资源的有效配置呢?这就是经济机制设计理论所要讨论的主要问题。这些问题的提出对机制的信息理论和激励理论的产生有着直接影响。以上的第一个问题实际上与激励理论有关;第二个问题则与信息有效理论有关。

这样,机制设计需涉及到两个基本问题:一个是信息效率问题,即所制定的机制是否只需较少的信息传递成本,较少的关于消费者、生产者及其他

<sup>1</sup> 在微观经济学中有两个福利经济学定理给出了市场机制与所导致资源配置的有效性(最优性)之间的关系。第一福利经济学定理阐明完全竞争的市场机制导致了帕累托有效配置(正式定义在下面给出)。它假定不存在外部效应以及某种个人偏好的非充分满足(自利性)的特性。第二福利经济学定理阐明任何帕累托有效配置都可以通过合适的资产再分配后由完全竞争的市场机制来达到。它也假定不存在外部效应以及某种个人偏好的非充分满足的特性。但还要增加另外一些重要假定,如个人偏好的凸性及生产技术不存在规模报酬递增的现象等假设。帕累托有效(最优)配置指的是这样的一种配置:如果不存在能改善社会中某个成员的福利而又不损害其他人的福利的可供选择的可行的资源配置的话,那么这种资源配置就被说成是帕累托有效配置。

经济参与者的信息；另一个是机制的激励相容问题（也就是积极性问题），即在所制定的机制下，每个参与者即使追求个人目标，其客观效果是否也能正好达到设计者所要实现的目标。现在大多经济学家都已知道，当经济信息不完全并且直接控制的方式不可能或不恰当时，人们需要采用分散化决策的方式来进行资源的配置或做出其他的社会经济决策。这样，在制度或规则的制定者不可能了解所有个人信息的情况下，他所要掌握的一个基本原则就是所制定的机制能够提供给每个参与者某种激励使得参与者在追求个人的利益时也同时地实现了所制定的目标。这就是所谓的激励机制的设计。许多现实和理论问题都可归结为激励机制的设计。比如委托-代理 (principal-agent) 问题、最优合同设计、规章或法规制定、公共财政理论、拍卖机制的设计、最优税制设计、行政管理、政治社会制度设计，甚至处理家庭关系问题等。激励相容和信息成本是任何机制，特别是经济制度的设计所必须考虑的两个基本问题，人们把这两个因素作为判断一个经济机制优劣的标准。不同的机制会导致不同的信息成本、不同的激励和不同的配置结果。研究一个经济机制的信息有效性和激励相容性可评价其制度的优劣性。本文只对经济机制理论中具有多个代理人的信息效率问题和激励相容问题主要成果和最新进展作一大致介绍。由于只有一个代理人的委托-代理激励机制设计不涉及到代理人之间的对策问题，其激励机制的设计问题相对简单，可归结为有约束的最优化问题。对委托-代理理论和应用的系统讨论，请参阅 Laffont 和 Martimort(2002) 合著的新书：*The Theory of Incentives: The Principal-Agent Model*，其中文版已由中国人民大学翻译出版。

### 三、经济机制的信息效率问题

本节讨论一个经济机制实现某个目标所要求的最小信息量问题。在这里我们只注意经济机制的信息要求（即运行信息成本问题），而不考虑激励问题，即不要求个人自利行为（个人理性）与既定目标（集体理性）一致的问题。该问题放在下一节讨论。

#### （一）信息分散化经济机制模型

从信息的观点出发，一个经济机制可以看成是一个信息交换和调整的过程。像市场调整过程那样，当信息的交换处在平稳 (stationary) 位置上时，一个配置结果被决定。分散决策从本质上来说是信息不完全的一种特征——信息分散于各个生产和消费决策者的一种特征。人们通过对需求和供给等经济活动的信息交换和传递来做出生产和消费的决策。信息分散化与亚当·斯密-哈耶克-弗里德曼所论证的竞争市场机制的最优性的特征紧密相关。那么，信息是什么？信息分散化的严格定义又是什么？它应包括哪些内容？在什么意义下认为信息成本是大还是小？在讨论这些问题时，人们需要一个统一的模型来研究什么是经济机制。这个模型最好能包括信息分散过程、信息

集中过程、市场经济机制、计划经济机制以及它们的各种混合形式的机制,因为仅仅把一个个的机制(如市场机制和计划机制)分别加以考虑是不够的。我们下面介绍一个非常一般的信息调整过程模型,它能研究各种机制的信息成本问题。

假定在一个经济社会中,有  $n$  个参与者,每个参与者可以既是生产者也是消费者,也可以只是生产者或只是消费者,或是一个家庭、政府的某一个部门或机构,所有参与者的集合记为  $N$ 。作为一个生产者,企业有一个生产可能性集合(生产技术条件约束),记为  $Y_i$ 。作为一个消费者,他有一个消费空间,记为  $X_i$ ,有一个消费偏好关系或效用函数(如存在的话),记为  $R_i$ (或  $u_i$ ),即对任何两组商品组合,他能比较哪一组商品对他更为有利。每个单位  $i$  还有一个初始资源,记为  $w_i$ 。这样,消费空间、初始资源、消费偏好关系、生产技术这四项合起来就构成了这个参与者的经济特征,记为  $e_i = (X_i, w_i, u_i, Y_i)$ 。抽象地说,一个经济社会就是由所有参与者的特征组成的,它也被称之为经济或经济环境,记为  $e = (e_1, e_2, \dots, e_n)$ 。所有可能的经济环境形成了一个集合,记为  $E$ 。所有资源配置的集合称为资源配置空间,记为  $Z$ 。从信息传播的角度讲,所谓经济机制就是把信息从一个经济单位传递到另一经济单位。从信息物质形态讲,信息的传播形式可以是一封信、一个电话、一个图像等;从信息量化的角度讲,传递的内容可以是一组数、一个向量或一个矩阵。机制设计所需要考虑的一个重要问题就是尽量简化传递过程中的复杂性,或使一个机制合理运行而使用较少的信息,因为较少的信息意味着较少的机制运行(交易)成本。由第  $i$  个人传递出的信息我们记为  $m_i$ ,也称为语言(messages)。所有这些信息的集合称为第  $i$  个人的语言空间,记为  $M_i$ 。 $n$  个人在时间  $t$  的一组语言记为  $m(t) = (m_1(t), \dots, m_n(t))$ 。所有这些语言的集合称为语言空间,记为  $M$ 。由于人们根据所接收到其他人的信息不断调整和反馈自己所发出的信息,在一个简单的一阶差分模型中,第  $i$  个参与者在时间  $t+1$  对时间  $t$  时的信息响应由差分方程

$$m_i(t+1) = f_i(m(t), e), \quad i \in N, \quad (1)$$

给出。这里  $f_i: E \rightarrow M$  被称为响应函数。一旦这种调整过程达到平稳点,人们不再改变信息,即  $m$  是响应函数的不动点  $m_i = f_i(m, e), i \in N$ , 或达到规定的终点时刻  $T$  时,通过某个资源配置规则(称为结果函数)  $h(\cdot): M \rightarrow Z$  来决定资源配置结果,即资源的配置由  $z = h(m)$  来决定。这样的一个信息调整,资源配置过程决定了一个经济机制。它是由语言空间、响应函数及结果函数组成的,记为  $\langle M, f, h \rangle$ 。信息空间规定了每个人依据自己的特征送出什么样的信息;响应函数表示了下一时刻输出的信息,它反映了如何在接到前一时刻的信息以后以怎样的形式反应出来,当然这种响应与经济环境  $e$  有关,反映函数决定了平稳信息状态;配置规则  $h$  是依据各单位送来的信息做出资源配置的。任何一个机制都是在一定的约束下运行的,约束的规则却是由政

府，或由立法机关制定，或由经济系统中每个参与人共同制定的。每一个人在这种约束下选择认为对他有利的信息。信息集合的元素可以是他自己对某种商品的需求或供给量；或是自己对商品的偏好关系；或是对产品成本的描述等等。配置规则决定了资源的配置。这个规则把信息的传递过程转化为物质资源的配置过程。这样，它建立了从信息空间到资源配置空间的一个关系（映射）。它根据个人、企业、或其他经济单位从信息集合中所选的信息来决定社会的生产及个人消费。

注意，式 (1) 中响应函数平稳点的集合定义了一个从经济环境空间  $E$  到信息空间  $M$  一个对应，记为  $\mu_i : E \rightarrow M$ ，即  $\mu_i(e) = \{m \in M : m = m_T \text{ 或 } m_i = f_i(m, e), i \in N\}$ 。令

$$\mu(e) = \bigcap_{i=1}^n \mu_i(e). \quad (2)$$

我们得到了一个从  $E$  到  $M$  的多值对应： $\mu : E \rightarrow M$ ，并且  $m \in \mu$  当且仅当  $m$  是 (1) 的平稳点。这样，一个信息调整机制可等价地定义为  $\langle M, \mu, h \rangle$ ，这里， $\mu = \bigcap_{i=1}^n \mu_i$  称为平稳信息对应。

方程 (1) 所反映的调整过程包括了信息集中化调整过程，这是由于参与者  $i$  在下一时刻输出的信息，可能不仅与自己的经济特征  $e_i$  有关，也可能与其他人的特征有关。如果在一个经济机制中，每个经济单位只需要知道自己的经济特征，而不知道其他单位的经济特征来决定下一时刻所传递的信息，这样的机制将称为信息分散化机制。它是方程 (1) 的一个特殊情况，即当参与者  $i$  在下一时刻输出的信息只依赖于自己的经济特征  $e_i$ ，而与其他人的特征无关时，方程 (1) 成为

$$m_i(t+1) = f_i(m(t), e_i). \quad (3)$$

方程 (3) 定义了一个信息分散决策过程或称为隐私保障 (privacy-preserving) 调整过程，相应的均衡信息对应成为  $\mu(e) = \bigcap_{i=1}^n \mu_i(e_i)$ ，这里  $\mu_i(e_i) = \{m \in M : f_i(m, e_i)\}$ 。

在实际中，交流的信息内容通常是向量。这样，一旦调整过程和信息分散化被定义后，从一个机制的信息空间的维数的大小可以评价这个机制的好坏。当考虑实际机制时，我们也许会发现，有些经济机制需要传递非常多的指标，而有些经济机制只需传递很少的指标。从信息的观点来看，对于想要实现的某个社会目标，人们总想找到一个既能实现这个社会目标又有尽可能小的运行成本（一种交易成本）的机制。

当信息空间是无限维时，人们不能通过信息空间维数来比较它的大小，从而需要更一般的比较信息空间大小的方法。一种方法是通过比较信息的拓朴空间的大小来决定所用信息量的大小。于是我们有下面比较信息空间大小的定义，它是由 Walker(1977) 最先给出的。

令  $S$  和  $T$  是两个信息拓扑空间。空间  $S$  被说成至少用到像空间  $T$  一样的信息, 记为  $S \geq_F T$ , 如果  $T$  能够被同胚地坎入在  $S$  中, 即存在着  $S$  的一个子空间  $S'$  使得它与  $T$  同胚, 也就是存在着一个从  $T$  到  $S'$  一一对应, 其逆也是连续连续函数。

一个信息分散且导致了资源有效配置的机制  $\langle M, \mu, h \rangle$  被说成是信息有效的, 如果它的信息空间  $M$  在所有导致了有效配置的信息分散化机制中是最小的。

一个从经济环境空间  $E$  到结果空间  $Z$  的信息机制  $\langle M, \mu, h \rangle$  定义了一个对应, 称作为机制的表现对应 (performance correspondence), 记为  $G$ :

$$G(e) = \{z \in Z : z = h(m), m \in \mu(e) \text{ 对某个 } m \in M\}.$$

给定一个社会选择对应  $F : E \rightarrow Z$  和信息机制  $\langle M, \mu, h \rangle$ , 如果对所有的经济环境  $e \in E$ ,  $G(e) \neq \emptyset$  并且  $G(e) \subset F(e)$ , 我们称信息机制  $\langle M, \mu, h \rangle$  实现了社会选择目标  $F$ 。如果对所有的经济环境  $e \in E$ ,  $G(e) \neq \emptyset$  并且  $G(e) = F(e)$ , 我们称信息机制  $\langle M, \mu, h \rangle$  完全地实现了社会选择目标  $F$ 。<sup>2</sup>

资源的有效配置(即帕累托最优配置)是大多数人所能接受的一个社会标准(目标)。我们知道竞争的市场机制导致了资源的有效配置。那么人们也许会问: 对新古典经济环境类(即商品是完全可分的, 消费者偏好是连续的、单调的及凸的, 生产集是闭的, 没有规模报酬递增)是否还存着其他信息分散机制(如市场社会主义经济机制)在信息方面比竞争市场机制更有效(即比竞争市场机制利用了更少的信息(交易)成本)而实现了最优配置? 赫维兹等人在 70 年代证明: 纯交换的新古典经济环境类, 没有什么其他经济机制既能导致资源有效配置而又比竞争市场机制使更少的信息。美国数理经济学家乔丹(Jordan)在 1982 年更进一步证明了对纯交换的新古典经济环境类, 竞争市场机制是唯一的利用最少信息并且产生了有效配置的机制。由于纯交换经济并没有考虑到包括生产的经济环境类, 这显然离现实太远, 从而这些理论结果从政治经济学的角度考虑就大大打了折扣: 因为人们如果要想论证市场机制的信息有效性, 就必须证明即使在包括生产的经济环境类情况下, 人们也会得到类似的结果。也许是由于在包括生产的经济环境类情况下证明类似的结果涉及到更多的复杂技术和高深的数学, 市场机制在一般的生产经济环境下是否也是信息最有效在过去 20 年成为一个一直没有解决的问题。笔者最近终于取得了进展。在 Tian (2000e) 中, 通过应用代数拓扑学等数学工具, 证明了即使具有生产的私有竞争市场机制是唯一的利用最少信息并且产生了有效配置的机制。由于有生产的经济环境包括没有生产的经济环境作为一个特殊情况, 我们只需对一般的生产经济环境进行讨论。

<sup>2</sup> 在经济学文献中, 人们一般用“实现 (realize)”和“实施 (implement)”来分别表示一个经济机制在达到社会目标时的信息和激励因素。

## (二) 市场竞争机制的信息有效性及唯一性

为了证明市场机制的信息有效性和唯一性, 我们考虑一个特别的生产经济环境类, 记为  $E^{cq} = \prod_{i=1}^N E_i^{cq}$ , 这里消费者的偏好关系是由 Cobb-Douglas 函数给出, 有效生产技术是由二次型生产函数给出。对  $i = 1, \dots, I$ , 消费者  $i$  的经济特征的集合  $E_i^{cq}$  是由所有的  $e_i = (X_i, w_i, R_i)$  给出, 使得  $X_i = R_i^L, w_i > 0$ ,  $u(x_i + w_i, a_i) = \prod_{l=1}^L (x_i^l + w_i^l)^{a_i^l}$ , 这里  $a_i \in \Delta^{L-1}$ 。

对  $i = I+1, \dots, n$ , 生产者  $i$  的经济特征的集合是由所有的  $e_i = \mathcal{Y}_i = \mathcal{Y}(b_i)$  给出:

$$\mathcal{Y}(b_i) = \left\{ y_i \in R^L : b_i^1 y_i^1 + \sum_{l=2}^L \left( y_i^l + \frac{b_i^l}{2} (y_i^l)^2 \right) \leq 0 \right. \\ \left. -\frac{1}{b_i^l} \leq y_i^l \leq 0 \text{ 对所有的 } l \neq 1 \right\}, \quad (4)$$

这里,  $b_i = (b_i^1, \dots, b_i^L)$ ,  $b_i^l > \frac{J}{w_i^l}$ 。

给定初始资源  $\bar{w} \in R_{++}^L$ , 定义子集合  $\bar{E}^{cq} \subset E^{cq}$  如下:

$$\bar{E}^{cq} = \{ e \in E^{cd} : w_i = \bar{w}_i \forall i = 1, \dots, I \}, \quad (5)$$

即初始资源在  $\bar{E}^{cq}$  中保持不变。令  $E^c$  为保证瓦尔拉斯均衡存在的所有生产经济环境的集合。

于是我们有下列定理。

**定理 1(信息有效性定理)** 假定  $\langle M, \mu, h \rangle$  是定义在生产经济环境上  $E^c$  的资源配置机制使得:

- (i) 它是信息分散化的;
- (ii) 它导致了帕累托有效配置;
- (iii)  $M$  是 Hausdorff 拓扑空间;
- (iv)  $\mu$  在某点  $e \in \bar{E}^{cq}$  有一个局部连续选择。

则这个机制的信息空间  $M$  至少像竞争的市场机制的信息空间一样大, 即我们有  $M \geq_F M_c =_F R^{(L-1)I+LJ}$ 。

这个定理说明了对包括生产的一般新古典经济环境类, 没有任何其他经济机制既能导致资源有效配置而又比私有制下的竞争市场机制用到了更少的信息。下面的定理 2 更进一步证明了定义在  $E^{cq}$  上的市场机制是唯一信息有效的机制。

**定理 2(唯一性定理)** 假定  $\langle M, \mu, h \rangle$  是一个定义在经济环境集合  $E^{cq}$  的机制使得:

- (i) 它是信息分散化的;
- (ii) 它导致了帕累托有效配置;
- (iii) 它相对于的给定保证分享结构  $\gamma_i(e; \theta)$  是个人理性的;
- (iv)  $M$  是一个  $(L-1)I + LJ$  维数的流型;
- (v)  $\mu$  在  $E^{cq}$  上是一个连续函数。

则存在一个从  $\mu(E^{cq})$  到  $M_c$  上的同胚映射  $\phi$  使得

- (a)  $\mu_c = \phi \cdot \mu$ ;
- (b)  $h_c \cdot \phi = h$ .

这样, 这个定理证明了任何定义在  $E^{cq}$  上, 有相同维数且导致了帕累托有效配置的机制事实上是和竞争机制等价的。由于  $E^{cq}$  是新古典经济环境类的一个子集合, 于是我们有

**定理 3** 假定  $\langle M, \mu, h \rangle$  是定义在经济环境  $E^{cq}$  上的非竞争市场机制使得:

- (i) 它是信息分散化的;
- (ii) 它导致了帕累托有效配置;
- (iii) 它相对于的给定保证分享结构  $\gamma_i(e; \theta)$  是个人理性的;
- (iv)  $\mu$  在  $E^{cq}$  上是一个连续函数。

则它的信息空间  $M$  一定会比竞争市场机制的信息空间大, 即,  $M >_F M_c =_F R^{(L-1)I+LJ}$ 。

对以上三个定理的证明以及对如何将市场机制定义成信息分散化决策机制, 请参见 Tian (2000d)。

这样, 私有产权的竞争市场机制是唯一的利用最少信息并且产生了有效配置和个人理性配置的经济机制。对公共商品的情况, 笔者也得到了类似的结果。在 Tian (2000f) 一文中, 对具有公共商品的经济环境类, 笔者证明了没有任何其他经济机制既能导致资源有效配置而又比林道机制利用了更少的信息, 并且林道机制是唯一的利用最少信息并且产生了有效配置和个人理性配置的机制。<sup>3</sup> 对具有副产品的公共商品的经济环境类, Tian (1994a) 讨论了信息有效性机制设计的问题。由于篇幅有限, 我们就不做详细讨论。

从以上的这些结果, 我们可以得到一个重要的结论: 无论是指令性计划经济机制、国有经济、集体经济、还是股份合作制, 以及任何其他非市场

<sup>3</sup> 林达尔均衡配置是指: 在具有公共商品的经济环境中, 如果存在着的一组私人商品价格及个人化公共商品价格(即公共商品的价格对不同的人也许是不同的)向量使得所导致的总需求等于总供给, 这样的配置被称为林达尔均衡配置。我们在下面讨论公共商品经济环境下的激励机制设计时将会给出林道均衡的严格定义。

的经济制度，它为了实现资源有效配置所需要的信息一定要比竞争市场机制所需要的多，从而这些机制不是信息有效率的，即需要花更多的运行成本（或代价）来实现资源的最优配置。这个结果告诉人们，在竞争市场机制能够解决资源的最优配置的情况下，应让市场来解决。只有在竞争市场无能为力的情况下，才采用其他一些机制来补充市场机制的失灵。这个结果可能对中国为什么要搞市场化的经济改革，国有经济民营化提供了一个理论基础。它也部分地回答了早期社会主义大论战所争论的信息效率问题，这个推论实际上是比较直观的。从以上对竞争市场机制定义为一个信息分散化经济机制可看出，市场机制的信息空间是由两个向量组成：一个是价格向量，另一个是资源配置向量（商品供给和需求所组成的向量）。而当人们运用指令性计划经济机制时，下面的企业必须向上面汇报、传递各种信息，其中包括生产函数（它反映了企业的技术条件和生产能力）。比如，即使假定生产函数是用多项式函数给出，它可能有任意高的次数。这样当一个企业向中央计划部门传递关于多项式生产函数的信息时，信息空间的维数可变得任意大。中央计划部门同时还可能需要得到消费者需求方面的情况。由于不同的人有不同的消费偏好，也就是具有不同的效用函数，这样计划机制的信息空间的维数可能会变得非常大，从而使得政府计划部门要做出生产和消费的决策所需要的信息就会变得非常大，使得机制的运行成本非常大。

当然，对一个具有较少维数的经济机制，它的配置规则也许可能会变得非常复杂。运转这个机制总的代价也许比运转某个具有较大维数的机制的总代价还要大。不管怎样，对机制的最小信息空间的研究能够使人们知道运转一个机制至少需要多大的信息量或运行成本。当然对探索机制的其他方面（如机制的复杂性）也是重要的。

另外一个是，对更一般的包括非新古典的经济环境类（比如，不可分的商品、非凸的偏好关系或生产可能性集），是否存在着导致了最优资源配置的信息分散决策机制？如果存在的话，它和所需信息（交易成本）的大小之间的关系是什么？赫维兹等人对非常一般的经济环境证明了这种机制的存在。但是，这样的机制是以非常高的信息成本作为代价的。Calsamiglia(1977)和Hurwicz(1999)分别证明了对一类非古典的经济环境类，特别是对非凸和具有外部性的经济环境类，需要一个无限维的语言（信息）空间使得一个机制导致帕累托最优的资源配置。

#### 四、经济机制的激励相容问题

上节主要关心的是一个经济机制实现资源有效配置所需最小信息空间——信息成本问题，而忽视了机制的激励相容问题，即参与者的行为是由响应函数或信息平稳对应描述而不是由参与者的偏好和对其他人的策略如何反应来决定的。对激励相容问题的探讨是当前信息经济学研究的主要课题之一。我们知道人的利己行为、经济自由选择、分散化决策、引进各种激励机制是一个

经济制度运行良好的先决条件。然而,经济学文献中的大多数研究只是讨论在市场制度下的各种激励问题。一个人所做的每一件事都涉及到利益与代价(收益与成本)。这种利益和代价可以是有形的,也可以是无形的。只要利益和代价不相等,就存在着激励问题。既然个人、社会和经济组织的利益不可能完全一致,激励问题在每一个社会经济单位中都会出现。由于每个人从所要做的事中获得利益并付出代价,在自利的驱动下,他将做出合理的选择:利益大于代价,就做这件事,或把它做好;否则就不做,或不想把它做好。一个经济制度要解决的一个根本性问题就是如何调动人们积极性的问题,即如何通过某种制度或游戏规则的安排来诱导人们努力工作,使得努力工作的收益大于所付出的代价。这样的激励机制能够把人们的自利和社会利益有机地结合起来。这样,检验一个机制或规则是否运行良好的一个基本标准就是,看它能否提供内在激励(动力)使人们努力工作、做出高质量的工作,激励决策者作出有利于他主管的经济组织的决策,激励企业尽可能有效地生产,从而使整个经济能健康发展。一个经济制度如果不能激发其成员的积极性,反而却压抑了其成员的创造力,制造出一批又一批的懒人、闲人,这个制度就不可能长期存在下去。

我们在前面提到早期对激励问题的探讨是由对社会主义经济机制的可行性的争论所引起的,它导致了机制设计理论的产生。其实,激励机制设计的例子在历史上很早就存在。例如,圣经《旧约全书》就有一个关于以智慧著称的古以色列国王所罗门(Solomon,大卫之子)如何设计激励机制来解决两个妇女争夺婴儿所属权的著名故事。它讲的是两个妇女来到国王所罗门面前要求解决婴儿归属问题。两人都宣称自己是婴儿的真正母亲。尽管这两个妇女自己知道谁是小孩真正的母亲,但国王不知道谁在说谎。国王的目标是想要把小孩判给他真正的母亲。国王所采用的办法(所设计的机制)是对两个妇女采用威胁的方式:把这个活着的小孩砍成两半,一人分一半。听到国王的命令,孩子的亲妈赶忙说:“把孩子给她吧,千万不要杀了他。”但另一个妇女说:“这孩子不应属于我们中的任何一个,把他分成两半吧。”于是国王做出判决:“这孩子归第一个女人,她是孩子的亲妈”。在得不到儿子的情况下,小孩的母亲不愿意看到自己的亲生儿子被杀死,而只好放弃。国王通过这种间接的方式达到了目标,确定了谁是孩子真正的母亲。读者将在后面看到所罗门国王的激励机制是有问题的,如果第二个女人聪明一点也要求不要把孩子杀了,把孩子给对方,则所罗门国王的激励机制不能解决问题。另外一个例子是几人分一个饼。如何把它分得公平使大家都没有意见?那么如何设计一种激励机制使得无论是谁分饼都可能把饼分得尽可能均等?一个简单的激励机制是:动手分的人最后拿,不动手分的人先拿。采取这种方式,即使分饼的人非常想为自己多分一点,她一定有很大的激励想把饼分得均匀。

那么,什么是激励相容问题呢?假定机制设计者(或称委托人)有一个经济目标,称为社会目标,这个目标可以是资源的帕累托最优配置、在某种意义下的资源公平配置、个人理性配置、某个经济部门或企业主所追求的目标、或

在其他准则下的配置等。设计者认为这个目标是好的，想要达到的。那么，是否能够激励每个参与者（消费者、企业、家庭、基层机构等）按照这个目标去做呢？换句话说，应制定什么样的规则才能使经济活动中每个成员的利己行为的实际结果与给定的社会或集体目标一致呢？或者说，应制定什么样的规则使得每个人在追求个人利益的同时也使既定的社会目标实现呢？激励机制设计理论可以回答这个问题。应当注意，这里所指的设计者是一个抽象的设计者，并不一定指某个人。根据不同的问题，设计者可以是一个人，一组人，立法机构，政府部门，政策制定者，经理厂长，部门主管，提出各种经济模式的经济学家等，甚至是约定都要遵守既定游戏规则的所有参与者，或其他制定规则或法则的某种机构。设计者知道哪些社会目标是好的，值得达到的。例如，他们认为有效地配置资源，公平分配，减少企业亏损等这些目标是好的。经济学家或改革者们的任务则是制定具体计划来实现这个目标。实际上，往往一些很具体的经济政策问题需要以一些很抽象的数学模型来严格描述。当我们认为某种方案不能实施时，我们应该追问究竟是什么阻碍了它的实施。当然一个明显的限制或障碍就是物资和技术条件。除此之外，还有一个因素：激励相容问题。如果一个经济机制不是激励相容的，则会导致个人行为与社会目标的不一致，往往导致所谓的“上有政策，下有对策”或者“和尚把经念歪了”的现象，使得制定的政策或制度不能发挥既定的作用。离开人的积极性和主动性，社会目标自然无从实现（至少是与理想的状态相差太远）。为什么经常个人或企业的行为结果与政策、法规制定者所想达到的目标不一致呢？就是因为这些规章制度不是激励相容的。因为在所制定的规则下，个人或企业不按照设计者所制定的社会目标那样去做可以得到更大的好处。那么我们应该采取什么样的机制（或规则）使得每个人的行为（不管利己与否）与社会目标一致呢？这就是本节要介绍的激励机制设计理论所要回答的核心问题之一。

### （一）基本模型

在考虑激励相容问题时，经济学家引进了一个基本理论模型来研究激励机制的设计。这个理论模型包括五个组成部分：（1）经济环境；（2）配置空间与社会目标；（3）经济机制；（4）个人自利行为策略均衡解假设；（5）社会目标的实施。

#### 1. 经济环境

在所考虑的经济社会中，有  $n$  个参与者。像上节信息调整机制一样，参与者  $i$  的经济特征记为  $e_i = (X_i, w_i, R_i, Y_i)$ ，这里  $X_i$  是  $i$  的消费集， $w_i$  是  $i$  的初始资源， $R_i$  是  $i$  的偏好关系（如果效用函数存在，则用  $u_i$  表示  $i$  的偏好关系），及  $Y_i$  是  $i$  的生产集。所有允许的经济特征的集合记为  $E_i$ 。所有参与者经济特征的一个组合  $e = (e_1, e_2, \dots, e_n)$  被称为一个经济或经济环境。假定  $e$  属于某个先验允许 (priori admissible) 集合  $E = \prod_{i \in N} E_i$ 。这里的  $E$  是经济环境的一般表达形式，根据所面临的不同问题，经济环境也许只是由参与者特

征的某些分量组成。例如,经济环境只是由参与者的效用函数组成。假定机制设计者知道经济环境  $E$  的区域,但不知道各个参与者  $i$  的真实的经济特征  $e_i$ 。每个参与者知道自己的经济特征,但根据所考虑的具体问题,可假定他知道或不知道其他参与者的经济特征。如果知道的话,我们称这种情况为完全信息 (complete information) 情况。如果不知道的话,我们称这种情况为不完全信息 (incomplete information) 情况。在这两种情况下,机制设计者都假定不知道参与者的经济特征。

## 2. 配置空间与社会目标

在给定的经济环境下,每个人都参与经济活动,做出决策,并从经济活动中得到配置结果。令  $Z$  表示所有配置结果的集合,称为配置结果空间。配置结果空间的点从社会的角度来看,也许并不都是可行的或最优的。令  $A \subset Z$  表示所有可行的配置结果集合。在某种社会最优的标准下,可行集的某个子集构成了一个社会目标(或称为社会选择对应),记为  $F$ 。于是它是从经济环境空间到可行集的一个对应  $F: E \rightarrow A$ 。比如,对给定的经济环境  $e, F(e)$  是帕累托有效配置的集合,个人理性的集合,市场一般均衡的集合,或是其他所期望达到的配置结果集合。当社会选择对应成为一个单值映射时,我们称它为社会选择函数,记为  $f$ 。设计者的任务是对所有的经济环境  $e \in E$ ,找出某种配置规则(经济机制)使得所导致的配置结果符合社会目标。

## 3. 经济机制

由于机制设计者缺乏关于个人经济特征方面的信息,设计者需要制定恰当的激励机制(游戏规则)以此诱导每个人能真实显示他们的信息。这样,为了实施某个社会目标,设计者可先告诉参与者他所收集到的信息将如何被用来决定配置结果(即先告诉游戏规则)。然后,根据游戏规则和参与者所报告或传递的信息,决定配置结果。于是一个机制是由信息空间  $M$  和结果函数(配置规则) $h$ 两部分所组成,记为  $\Gamma = \langle M, h \rangle$ 。令  $M_i$  表示参与人  $i$  的信息空间,它是参与人  $i$  所有可能交换和传送信息  $m$  的集合。令  $M = \prod_{i \in N} M_i$ 。信息空间  $M$  规定了各参与者送出什么样的信息的范围,配置规则  $h$  则依据各参与者所报的信息  $m$  给出配置结果,于是结果函数  $h: M \rightarrow Z$  是从信息空间  $M$  到结果空间  $Z$  的一个映射。任何一个机制  $\langle M, h \rangle$  都是在一定的规则下运行的。需要指出的是,激励机制的设计与上节所考虑信息调整机制有所不同。在激励机制设计中,参与者的行为不再由响应函数或信息对应来描述而是由参与者根据他们的偏好和采用策略的方式所决定。不过,将激励机制看作为信息调整机制的一种方式是将激励机制的信息空间的所有点看作为信息调整机制的平稳点。

在经济学文献中,机制通常也称作为博弈形式 (game form)。注意它和博弈论中博弈不一样。在机制中,参与者策略的相互作用导致的是配置结果,而不是效用或支付量。当然,一旦个人的偏好被确定后,一个经济机制(博弈形式)导致了一个博弈,它的策略空间为  $M$ ,效用或支付函数为  $u_i(h(m))$ 。另外一个不同点是,博弈论中参与者的偏好是给定的,而在机制设计中,机制

设计者不知道参与者的真正的经济特征，只知道它属于某个集合范围，从而不是给定和不变的。

传统的经济学的研究方法是将机制作为已知，研究它能导致什么样的资源配置。例如，把市场机制作为给定，而把市场机制的运行结果作为未知，研究在什么样的经济环境下，市场机制导致了资源的最优配置。然而，对经济机制的设计者来说，他们提出来的问题往往是相反的。他们把社会目标作为已知（即设计者知道哪个社会目标是好的，想要达到的），而想找到一个经济机制能实现既定的社会目标，也就是找到一个机制使得人们自利行为的结果和想要达到的社会目标一致，当然并不是所有的社会目标都是可达到的。经济机制设计理论的一个目标就是研究什么样的社会目标能实现，什么样的社会目标不能实现。通过对这一问题的研究，可以帮助解决经济理论中一些具有争论性的问题。

#### 4. 个人自利行为策略均衡解假设

在机制理论中，一个基本的假设是每个人在主观上都追求个人利益，根据个人私利行事。除非得到好处，他们一般不会真实地显示有关他们经济特征方面的信息。不同的经济环境和机制（游戏规则）将导致参与者个人自利行为的不同反应。每一个人在规则下选择认为对自己最有利的信息。每个人行事的策略（即所送出的信息）取决于他的自利行为（行为方式）。个人的自利行为不仅取决于他的经济特征，也取决于经济制度或游戏规则，不同的规则显示出不同的利己行为。令  $b(e, \Gamma) \in M$  表示在经济环境为  $e$ ，及机制  $\Gamma$  给定下的均衡自利行为策略解的集合。这样，给定经济环境  $E$ 、信息空间  $M$ 、配置规则  $h$  及自利行为准则  $b$ ，所导致的所有均衡配置结果是由配置规则和均衡自利行为策略复合而成的，即  $h(b(e, \Gamma))$ 。

#### 5. 社会目标的实施与激励相容

激励机制设计的目的是要实施某个给定的社会目标  $F$ 。首先注意社会选择对应  $F$  依赖于经济环境。其次，给定一个经济机制  $(M, h)$  和均衡自利行为决策集  $b(e, \Gamma)$ ，社会目标的实施问题涉及到  $F(e)$  和  $h(b(e, \Gamma))$  这两个集合相交的状态关系问题。我们有下列关于一个社会目标可实施的定理：

完全实施：如果对所有的经济环境  $e \in E$ ，有 (1)  $b(e, \Gamma) \neq \emptyset$ ，(2)  $h[b(e, \Gamma)] = F(e)$ ，我们就说机制  $\Gamma = (M, h)$  按策略行为  $b$  完全地实施了社会目标  $F$ 。

实施：如果对所有的经济环境  $e \in E$ ，有 (1)  $b(e, \Gamma) \neq \emptyset$ ，(2)  $h[b(e, \Gamma)] \subset F(e)$ ，我们就说机制  $\Gamma = (M, h)$  按策略行为  $b$  实施了社会目标  $F$ 。

弱实施：如果对所有的经济环境  $e \in E$ ，有 (1)  $b(e, \Gamma) \neq \emptyset$ ，(2)  $F(e) \cap h[b(e, \Gamma)] \neq \emptyset$ ，我们就说机制  $\Gamma = (M, h)$  按策略行为  $b$  弱实施了社会目标  $F$ 。

对于某个给定的社会目标对应  $F$ ，如果存在着某个经济机制，它按给定的均衡概念（完全或弱）实施了这个社会目标对应，我们就称这个社会目标按给定均衡概念是可（完全或弱）实施的。

如果一个社会目标  $F$  按给定均衡概念是可（完全或弱）实施的，我们称机制  $\Gamma$  与社会目标  $F$  按给定均衡概念是激励相容的。

注意,在定义以上社会目标的可实施性和激励相容性时,我们没有给出具体的均衡概念。读者将在后面看到一个社会目标是否可实施,在很大的程度上依赖于不同的自利行为均衡解假设。在文献中,当信息是完全时,对人的利己行为的均衡解假设通常包括优势策略(dominant strategy)均衡,纳什策略(Nash strategy)均衡,强纳什均衡(strong Nash equilibrium),子博弈完备纳什均衡(subgame perfect Nash equilibrium)解,非优势纳什均衡(undominated Nash equilibrium)等解的概念。我们将分别对这些均衡概念下的社会目标的实施问题进行讨论。在此之前,我们先给出一些例子。

### 6. 例子

具有公共商品的经济环境通常会导致激励不相容的问题。世界上的商品大致可分为两类:私人商品和公共商品。私人商品的特征是他们在使用上具有相互排斥性:一个人使用了它,另一个人就不能再使用它了。例如有一个苹果,我吃了,你就再吃不到那个苹果了。公共商品的特征是一个人对于同一个单位的商品的使用不降低另外一个人对同一单位的商品使用的可能性。这个特征在使用上是一种非对抗性的关系,大家不需要通过互相竞争而使用这种商品。国防、路灯、公路、公共设施、基础研究、电视台和广播电台等都是公共商品或局部公共商品的例子。

例1(公共项目) 假定政府需要决定是否应该修建某个成本为 $C$ 的公共设施,比如修建一个体育馆、公共图书馆或公园。假定公共项目的成本由所有居民平摊。配置结果空间为 $Y = \{0, 1\}$ 是公共项目的集合。这里,0表示不修建这个项目,1表示修建。居民 $i$ 从这个公共项目所获得的效益为 $r_i$ 。在这种情况下,项目不修建时居民 $i$ 的净效益为0,而修建的净效益为 $v_i \equiv r_i - \frac{C}{n}$ 。于是, $i$ 对公共项目的价值(valuation)函数可写为

$$v_i(y, v_i) = yr_i - y\frac{C}{n} = yv_i.$$

例2(可分公共商品) 在例1中,公共商品只能取两个值,0和1,因而并没有规模大小的问题。但在许多情况下,公共商品的大小取决于集资的多少,因而公共商品 $y \in R_+$ 可以是任意一个非负值。令 $C(y)$ 表示修建规模为 $y$ 公共商品的成本。这样,配置结果空间为 $Z = R_+ \times R^n$ ,可行集 $A = \{(y, z_1(y), \dots, z_n(y)) \in R_+ \times R^n : \sum_{i \in N} z_i = C(y)\}$ ,这里 $z_i(y)$ 是为修建规模为 $y$ 的公共商品居民 $i$ 所分担的费用。居民 $i$ 从修建这个规模为 $y$ 的公共商品所获得的效益为 $r_i(y)$ 。假定 $r_i(0) = 0$ 。这样,项目不修建时居民 $i$ 的净效益为0,而修建的净效益为 $r_i(y) - z_i(y)$ 。于是, $i$ 的价值函数可写为

$$v_i(y) = r_i(y) - z_i(y).$$

例3(分配某个物品) 将某种不可分的物品分给经济社会中的某个人,比如将一个国有企业承包或拍卖给某个人,或将单位的一间房分给某个职工。在这种情况下,配置结果空间为 $Z = \{y \in \{0, 1\}^n : \sum_{i=1}^n y_i = 1\}$ ,这里 $y_i = 1$

表示第  $i$  人得到这个物品,  $y_i = 0$  表示这个人没有得到物品。如果第  $i$  人得到这个物品, 则他从得到这个物品所获得的净收益为  $v_i$ 。如果没得到物品, 则他的净收益为 0。于是,  $i$  的价值函数可表示为

$$v_i(y) = v_i y_i.$$

注意, 在这里我们可以把  $y$  看做为一个  $n$  维的公共商品向量。

从以上这些例子, 一个社会最优的决策显然依赖于个人的真实效用  $v_i(\cdot)$ 。比如, 例 1 中公共项目应该被修建当且仅当所有人收益的总和超出它的修建成本, 即如果  $\sum_{i \in N} r_i > C$ ,  $y = 1$ , 而当  $\sum_{i \in N} r_i < C$  时,  $y = 0$ 。令  $V_i$  是所有价值函数  $v_i$  集合, 令  $V = \prod_{i \in N} V_i$ , 令  $h: V \rightarrow Z$  是一个决策规则。  $h$  是有效的, 当且仅当下列不等式成立:

$$\sum_{i \in N} v_i(h(v_i)) \geq \sum_{i \in N} v_i(h(v'_i)) \quad \forall v' \in V.$$

一般来说, 如果政府只是简单地根据个人报出对公共商品的受益程度来决定摊派成本或自报对成本分摊的份额, 这样的规则会导致每个人在做贡献时有激励低报, 从而导致无效配置, 因而个人利益和社会目标是不相容的。在文献中, 解决激励不相容的一个方法是通过采用转移支付 (即对个人进行税收或补贴) 的方式来诱导人们真实地显示自己的偏好。为了让读者了解转移支付的这个作用, 考虑以上例 1 中的公共项目的实施问题。不难发现, 任何收益小于平摊成本 ( $r_i < C/n$ ) 的个人都不希望修建此公共项目, 而任何收益大于平摊成本的个人却希望修建此公共项目。想象政府只是简单地根据人们所报出的总收益  $\sum_{i \in N} r_i$  是否大于成本来决定是否修建这个公共项目。如果每个人报出他的真实收益, 导致的收益将是有效的。但是, 这种简单的方法不会给人们激励真实地报出他们的收益。这是由于对那些收益小于平摊成本的人来说, 他们会有激励尽可能低报他们的收益, 而对那些收益大于平摊成本的人来说, 他们会有激励尽可能高报他们的收益, 这将会导致一个错误的决策。这样, 为了真实地显示个人的偏好  $v_i$ , 像恰当的转移支付这样的调整是有必要的。后面要讨论的格罗夫斯 (Groves) 机制就给出了这样的激励相容机制。

## (二) 优势均衡实施与真实显示机制

对利己行为最强的均衡解假设是所谓的优势策略 (dominant strategy) 均衡解。在优势均衡解假设下, 每个人所作出的决策总是最优的而不管其他人的决策如何。在博弈论中的一个基本公理是, 只要优势均衡解存在, 博弈的参加者就会采用它。

对给定的  $e \in E$  和机制  $\Gamma = \langle M, h \rangle, m^*$  被说成是一个优势均衡, 当且仅当对所有的  $i \in N$ , 我们有

$$h(m^*, m_{-i}) R_i h(m) \quad m \in M,$$

这里,  $(m_i^*, m_{-i}) = (m_1, \dots, m_i^*, \dots, m_n)$ 。令  $D(e, \Gamma)$  表示机制  $\Gamma$  的所有优势均衡策略的集合。

在优势均衡假设下, 由于每个人的最优策略的选择都不依赖其他人的策略选择并且不需要知道其他人的经济特征, 每个人做决策时所要求的信息最少。这样, 如果优势均衡策略存在, 它将是最理想的一种情况。并且, 尽管一个机制的信息空间可以任意的决定, 但下面的显示原理 (revelation principle) 告诉人们在优势行为假设下没有必要寻找更复杂的机制, 而只需要考虑所谓的直接显示机制 (direct revelation mechanism) 就足够了, 这将大大地减少了构造机制的复杂性。

一个机制  $\langle E, h \rangle$  被称为直接显示机制, 如果信息空间完全是由参与者的经济特征集合组成, 即  $M_i = E_i$ 。

**定理 4 (显示原理)** 如果一个机制  $\Gamma = \langle M, h \rangle$  按优势均衡弱实施了社会选择函数  $f$ , 则直接显示机制  $\langle E, f \rangle$  同样地按优势策略弱实施了  $f$ 。

对一个直接显示机制  $\langle E, h \rangle$ , 如果对每个  $e \in E$ ,  $e$  是一个优势均衡, 且  $h(e) \subset F(e)$ , 则它被认为是和社会选择对应  $F$  按优势策略是激励相容的。当社会选择对应成为单值函数  $f$  时,  $\langle E, f \rangle$  能被看作为一个直接显示机制。如果对所有的  $e \in E$ ,  $e$  是一个优势均衡, 则直接显示机制  $\langle E, f \rangle$  被认为在  $E$  上是激励相容的。

由显示原理, 我们知道, 对于任意给定的社会选择函数  $f$ , 如果一个机制  $\langle M, h \rangle$  按优势均衡弱实施了  $f$ , 则直接显示机制  $\langle E, f \rangle$  是按优势策略激励相容的。真实显示优势均衡在文献中也称为强激励相容 (strongly incentive compatible)。

需要强调的是, 显示原理只是对弱实施有效, 而对实施或完全实施的结论却可能不成立。即当机制  $\Gamma = \langle M, h \rangle$  有多个优势均衡时, 可能对某些均衡策略  $m$ , 并没有  $h(m) = f(e)$ 。并且, 显示原理只是将原机制一个优势均衡与直接显示机制  $\langle E, f \rangle$  的真实显示均衡对应起来, 这个直接显示机制也可能存在着其他非真实显示均衡, 它并不对应原机制的任何均衡。这样, 即使一个机制  $\Gamma = \langle M, h \rangle$  按优势均衡实施了社会选择函数  $f$ , 直接显示机制  $\langle E, f \rangle$  也可能只是按优势策略弱实施了, 而不是实施或完全实施了  $f$ 。当然, 原机制和直接显示机制都只有唯一均衡策略时, 这个问题就不存在。在许多应用问题中, 所考虑的机制就只有一个均衡, 或所有均衡导致了等价的结果。

#### 1. Gibbard-Satterthwaite 不可能性定理

显示原理表明了没有必要要求个人显示与自己无关的任何信息。显示原理在如何找到一个优势机制方面是非常有用的。如果人们希望社会选择目标  $f$  能按优势均衡被实施, 人们只需要证明直接显示机制  $\langle E, f \rangle$  是激励相容的。但不幸的是, 下面的 Gibbard-Satterthwaite 不可能性定理表明, 当经济环境的区域不加任何限制时, 按优势策略可实施的社会目标几乎不存在。我们先引进下列概念。

一个规则是独裁的，如果存在某个  $i$  使得  $f(e) \in \{z \in Z : u_i(z) \geq u_i(z') \quad z' \in Z\}$ ，也就是说个人  $i$  的最优选择完全地决定社会的最优选择。

**定理 5(Gibbard-Satterthwaite 不可能性定理)** 如果结果空间至少包括三个元素，经济环境空间没有限制，且社会选择函数  $f$  按优势均衡是可弱实施的，则  $f$  是一个独裁规则。

Gibbard (1973) 和 Satterthwaite(1975) 的不可能性定理是一个非常负面的结果，这个结果在某种意义上是非常接近阿罗 (Arrow) 的不可能性定理。不过，如果对经济环境的区域做出一定的限制，则对一些社会选择可能会得出按优势策略是激励相容的正面结果(像后面要介绍的定义在拟线性效用集合上的格罗夫斯机制)。不幸的是，对导致了资源有效配置的社会选择对应，即使对于新古典的经济环境类，其结果也是否定的。这就是下面要介绍的著名的赫维兹不可能性定理。

## 2. 赫维兹不可能性定理 - 真实显示偏好与最优资源配置的不相容性

在赫维兹不可能性定理给出之前，人们以为只是对具有公共商品的经济环境类，真实显示与导致了个人理性和帕累托最优配置的社会目标是激励不相容的，而对私人商品的经济环境类不存在激励不相容的问题。人们知道完全的竞争市场机制可以很好地处理私人商品，但不能很好地处理公共商品。其原因是，每个人都想“搭便车”，都想从别人对公共商品的贡献中得到好处。例如，从前面处理公共设施开支的那个例子可以看出，如根据每一个人所报的自己享受这个公共商品的程度(边际替代效用)来决定这个人所应付的税，那么就有些人可能为了少付而低报自己的真正偏好，但这些人仍然可以同样地从公共设施中得到同样的好处，长此以往就没有人对公共设施支出付钱感兴趣了。这与私人商品大不一样，你花钱为你自己买日常用品不会使别人得利。

在机制设计理论产生以前，一谈到公共商品和私人商品的其他差别时，大多数经济学家以为对于只有私人商品的经济社会，资源最优配置与个人的利己行为是一致的。认为在竞争市场中，价格是作为参数给定的(即每个人的购买量不会影响价格的高低)，每个消费者没有必要隐瞒自己的真正偏好，即没有必要讲假话；而对于具有公共商品的经济社会，资源最优配置与个人的自利行为不一致，因为每个人都有激励想“搭便车”，想从别人对公共商品的贡献中得到好处，因而不愿报告自己对公共商品的真正偏好，即都宣称公共商品对他不重要以减少他自己对此应承担的贡献。这种不真实显示自己真正偏好的策略现象最初由 Samuelson(1954, 1955) 针对配置公共商品的林达尔均衡(Lindahl equilibrium) 解的批评而提出的。他进一步猜想对具有公共商品的经济环境，不存在任何分散化经济机制，它能导致帕累托最优配置并且使每个人有激励去真实地报告他自己的偏好。按以上直接显示机制和激励相容的术语来说，Samuelson 的论断意味着每个参与者为了追求个人利益，不可能真实地显示自己的经济特征，即真实显示偏好策略不是优势均衡。令人吃

惊的是这一论断不只对公共商品的环境类成立而且也对有限个人的私有商品的环境类成立。

赫维兹在1972年给出了著名的“真实显示偏好”不可能性定理。他证明了：即使对纯私人商品的经济社会，只要这个经济社会中的成员的个数是有限的，在参与(participation)约束条件下(即导致的配置应是个人理性的)，就不可能存在任何分散化经济机制(包括竞争市场机制)，它在新古典经济环境类下能导致帕累托最优配置并且使每个人有激励去真实地报告自己的经济特征。正式的表述如下：

**定理 6(赫维兹不可能性定理)** 对于私人商品的纯交换经济环境类  $E$ ，假定下列条件成立：

- (1)  $n$  是有限的，(2) 效用函数是单调、凹及连续的。

则任何导致个人理性和帕累托最优配置的机制在  $E$  上不是激励相容的。

证明见 Hurwicz(1972)。

以上赫维兹的不可能性定理告诉人们：即使对于只有私人商品的一般经济环境类，当参与人的个数是有限的，也不可能存在任何信息分散化经济机制(无论是市场机制，还是计划经济机制)，使得当人们的行为按优势策略决策时，它实施了资源最优配置。其原因在于有限人数经济环境与完全竞争假设不相容。不过，当经济社会中的成员数目与实数轴上的点一样多时(无穷不可数多个点)，“真实显示偏好”是可能的，但这与现实相差太远。当我们要设计某种经济机制时，首先必须牢记这个定理。如果想要某个机制能产生帕累托最优配置，我们必须放弃优势均衡假设，即放弃每个人都说真话办真事的假定。对于具有公共商品的经济社会，无论这个经济社会的人员的数目是多少，我们也能得到类似的不可能定理，即“激励相容”不可能性结果。从这一点说，这两种经济环境(即具有公共商品的经济环境与不具有公共商品的经济环境)没有什么大的差别。

### 3. 格罗夫斯-克拉克-威克瑞机制与真实需求显示

以上的结果说明了真实显示偏好与资源的帕累托最优配置一般来说是不可能同时达到的。那么，如果人们放弃帕累托最优配置标准，比如只考虑解决某个公共商品的有效生产问题，是否有可能设计出激励机制使得每个参与人有激励真实地显示自己的偏好并能有效地生产出所需要的公共商品呢？答案是肯定的。对拟线性类效用函数(quasi-linear utility function)，所谓的格罗夫斯-克拉克-威克瑞(Groves - Clark - Vickrey)需求显示机制能够解决公共商品的有效生产问题。

为了说明问题，我们先考虑不可分公共商品的供应问题(见例1)。假定某个社区有  $n$  个人，修建某个公共设施所需费用是  $C$ 。这  $n$  个人为修建这个公共设施所愿做出的捐献记为  $g_1, g_2, \dots, g_n$ 。当且仅当  $\sum_{i \in N} g_i > C$ ，这个公共设施被修建。公共设施给这  $n$  个人带来的效用(好处)记为  $r_1, r_2, \dots, r_n$ 。

不难看出, 当且仅当  $\sum_{i \in N} r_i > \sum_{i \in N} g_i$  时, 这个公共设施应该被修建。既然修建公共设施带给第  $i$  个人的净效用为  $v_i = r_i - g_i$ , 于是这个公共设施应被修建的充分必要条件成为:  $\sum_{i \in N} v_i > 0$ 。由于每个人的净效用  $v_i$  只有他自己知道, 他可能有激励谎报他的净效用。那么, 如何设计税制机制使得每个人都有激励报出他真正的净效用  $v_i$  呢? 我们在前面谈到引入转移支付也许能解决说假话的问题。事实确实如此。假定参与者  $i$  的效用函数  $u_i(x_i, y)$  为:

$$u_i(x_i, y) = x_i - w_i + r_i y, \quad (6)$$

这里,  $x_i - w_i$  可解释为净消费增量。他的预算方程应满足

$$x_i + g_i y = w_i + t_i. \quad (7)$$

代预算式 (7) 到 (6), 并且注意  $v_i = r_i - g_i$ , 参与者  $i$  的  $u_i(\cdot)$  效用函数于是成为:

$$u_i(t_i, y) = t_i + v_i y.$$

格罗夫斯给出了以下真实显示机制, 它很好地解决了说假话的问题。

首先, 格罗夫斯机制要求每个人报出他的净效用。记每个人所报的净效用为  $b_i$ 。这样, 信息空间为  $M_i = R$ 。由于每个人有可能真报或假报,  $b_i$  不一定就等于  $v_i$ 。然后, 根据所有个人总净效用的和是否大于总成本来决定这个公共设施是否被修建, 即格罗夫斯机制规定: 公共设施是否被生产由下式决定:

$$y(b) = \begin{cases} 1 & \text{如果 } \sum_{i \in N} b_i > 0 \\ 0 & \text{如果 } \sum_{i \in N} b_i \leq 0. \end{cases}$$

每个人的转移支付 (transfer payment) 记为  $t_i$  (如果  $t_i < 0$ , 它被解释为附加税; 如果  $t_i > 0$ , 它被解释为补偿)。  $t_i$  由下式决定:

$$t_i(b) = \begin{cases} \sum_{j \neq i} b_j + d_i(b_{-i}) & \text{如果 } \sum_{i \in N} b_i > 0 \\ d_i(b_{-i}) & \text{如果 } \sum_{i \in N} b_i \leq 0. \end{cases}$$

这里  $d_i(b_{-i})$  是一个独立于  $b_i$  可任意给定的函数。当参与者  $i$  的效用函数为  $u_i(t_i, y) = t_i + v_i y$  时, 将以上  $y(b)$  和  $t_i(b)$  代入到效用函数  $u_i(t_i, y) = t_i + y v_i$ , 我们得到了  $i$  的支付函数:

$$\phi_i(b) = \begin{cases} v_i + \sum_{j \neq i} b_j + d_i(b_{-i}) & \text{如果 } \sum_{i \in N} b_i > 0 \\ d_i(b_{-i}), & \text{如果 } \sum_{i \in N} b_i \leq 0. \end{cases}$$

我们现在来证明每个人真实显示他的净效用  $v_i$  是优势均衡策略, 即每个人都有激励说真话。有两种情况需要考虑。情况 1.  $v_i + \sum_{j \neq i} b_j > 0$ 。这就意味着参与者  $i$  希望公共设施被修建, 因为修建公共设施为第  $i$  个人所带来的效用要大于不修建时的效用 (即  $v_i + \sum_{j \neq i} b_j + d_i(b_{-i}) > d_i(b_{-i})$ )。当  $i$  真

实地报出  $b_i = v_i$  时, 他能保证公共设施被修建。因为, 如果  $b_i = v_i$ , 则个人的福利和修建公共设施这一社会目标一致, 即它们都保证了  $\sum_j b_j > 0$ 。  
 情况 2.  $v_i + \sum_{j \neq i} b_j \leq 0$ 。这就意味着参与者  $i$  不希望公共设施被修建, 因为修建公共设施给第  $i$  个人所带来的效用要小于或等于不修建时的效用 (即  $v_i + \sum_{j \neq i} b_j + d_i(b_{-i}) \leq d_i(b_{-i})$ )。当  $i$  真实地报出  $b_i = v_i$  时, 他能保证公共设施不被修建。因为, 如果  $b_i = v_i$ , 则个人福利和不修建公共设施这一社会目标一致, 即它们都保证了  $\sum_j b_j \leq 0$ 。

这样, 我们证明了真实显示是一个优势均衡解。由于  $d_i(b_{-i})$  是一个独立于  $b_i$  的任意给定的函数, 如果我们特意地令

$$d_i(b_{-i}) = \begin{cases} -\sum_{j \neq i} b_j & \text{如果 } \sum_{i \in N} b_i > 0 \\ 0 & \text{如果 } \sum_{i \in N} b_i \leq 0. \end{cases}$$

则支付转移结果函数成为

$$t_i(b) = \begin{cases} -|\sum_{j \neq i} b_j| & \text{如果 } (\sum_{i \in N} b_i)(\sum_{j \neq i} b_j) < 0, \\ 0 & \text{如果 } (\sum_{i \in N} b_i)(\sum_{j \neq i} b_j) > 0. \end{cases}$$

这意味着, 只对那些个人决策改变了总体决策的人增收附加税, 即只对那些使得  $(\sum_{i \in N} b_i)(\sum_{i \neq j} b_j) < 0$  的人收附加税。这种特殊情况的格罗夫斯机制成为所谓的克拉克机制或关键者 (pivotal) 机制。

我们现在考虑具有不可分公共商品的经济环境。在这个经济环境中, 有  $n$  个参与者, 一种私人商品  $x_i, K$  种公共商品  $y \in R_+^K$ 。生产公共商品  $y \in R_+^K$  的总成本为  $C(y)$ 。每个参与者  $i$  为生产  $y$  所需分担的费用为  $z_i(y)$ , 从而有  $\sum_{i \in N} z_i = C(y)$ 。于是他的预算方程应满足

$$x_i + g_i(y) = w_i + t_i. \quad (8)$$

参与者  $i$  的效用函数  $u_i(x_i, y)$  假定是拟线性 (quasi-linear) 的:

$$u_i(x_i, y) = x_i - w_i + r_i(y), \quad (9)$$

这里,  $x_i - w_i$  可解释为净消费增量,  $r_i(y)$  为参与者  $i$  消费公共商品  $y$  所获的效益。代预算式 (8) 到 (9), 参与者  $i$  的  $u_i(t_i, y)$  效用函数成为:

$$u_i(t_i, y) = t_i + v_i(y),$$

这里  $v_i(y) = r_i(y) - g_i(y)$ , 称为  $i$  对公共商品的价值函数 (valuation function)。许多经济模型都具有拟线性的效用函数。例如, 在信息经济学中的逆向选择模型中, 有一个委托人和  $n$  个代理人。委托人想要实施某个项目  $y$ , 代理人  $i$  从这个项目的实施会得到的收益为  $v_i(y)$  (或要负担的成本为  $-v_i(y)$ ) 和支付 (或收到) 转移税为  $t_i$  (或转移补偿支付  $-t_i$ )。这样代理人的支付函数将会为拟线性函数。

从  $\sum_{i \in N} z_i = C(y)$ , 我们得到可行性条件为:

$$\sum_{i=1}^n t_i = 0. \quad (10)$$

我们知道对于拟线性效用函数, 帕累托最优配置完全由最大化  $\sum_{i \in N} [t_i + v_i(y, \theta_i)]$  的配置  $(y, t_1, \dots, t_n)$  决定。这样, 对以上所考虑的公共商品经济类, 帕累托最优配置由总和最大化条件

$$\max_{y \in Y} \sum_{i=1}^n v_i(y) \quad (11)$$

和可行性条件 (10) 所特征化。

我们假定机制设计者不知道真实的价值函数  $v_i(y)$ 。设计激励机制的目的是要选择最优的公共商品水平 (即, 它是 (11) 的解), 并且每个参与者有激励真实地显示他的价值函数。用直接显示机制  $\langle V, h \rangle$  的语言来说, 机制要求每个参与者报出他的价值函数  $b_i$ , 所报的价值函数可以是真实或非真实的, 然后设计者给出配置的决策规则 (结果函数),  $h = (g, t_1, \dots, t_n) : V \rightarrow R_+ \times R^n$  使得  $y = g(b)$  和  $t_i = t_i(b)$ 。

一个直接显示机制  $\langle V, g, t_1, \dots, t_n \rangle$  是一个 格罗夫斯机制, 当且仅当满足下列两个条件:

(i)  $g(b)$  是 (11) 的解, 即,

$$\sum_{i=1}^n b_i(g(b)) \geq \sum_{i=1}^n b_i(y) \quad \text{对所有的 } y \in R_+^K \text{ 和 } b \in V. \quad (12)$$

(ii) 转移结果函数为

$$t_i(b) = \sum_{j \neq i} b_j(d(b) + d_i(b_{-i})), \quad (13)$$

这里,  $d_i(\cdot)$  是任意一个从  $V_{-i}$  到  $R$  的函数。

和不可分的公共商品情况一样, 我们可类似地证明每个人真实显示他的净效用  $v_i$  是优势均衡策略。同样, 如果令

$$d_i(v_{-i}) = - \max_{y \in R_+^K} \sum_{j \neq i} v_j(y),$$

则支付转移结果函数成为

$$t_i(b) = \sum_{j \neq i} b_j(g(b)) - \max_{y \in R_+^K} \sum_{j \neq i} v_j(y),$$

我们就得到了对可分公共商品情况下的克拉克机制 (关键者) 机制。有趣的是, 克拉克机制包括了著名的威克瑞拍卖机制 (也称为第二价格拍卖机制) 作

为一种特殊情况。在威克瑞拍卖机制下, 报价最高者获得拍卖品, 但成交价等于第二高报价。为了看出威克瑞拍卖机制是克拉克机制的特殊情况, 让我们回到前面例3关于拍卖某种物品的问题。在这种情况下, 配置结果空间为  $Z = \{y \in \{0, 1\}^n : \sum_{i=1}^n y_i = 1\}$ 。  $y_i = 1$  表示第  $i$  人得到这个物品,  $y_i = 0$  表示这个人没有得到物品。第  $i$  人的价值函数可表示为

$$v_i(y) = v_i y_i.$$

由于我们可以将  $y$  看做是一个  $n$  维的公共商品向量, 从以上克拉克机制, 我们知道

$$g(b) = \{y \in Z : \max_{i=1}^n v_i y_i\} = \{y \in Z : \max_{i \in N} v_i\}.$$

这样, 如果  $g_i(b) = 1$ , 则  $t_i(v) = -\max_{j \neq i} v_j$ 。如果  $g_i(b) = 0$ , 则  $t_i(v) = 0$ 。这就意味着报价最高者获得拍卖品, 但成交价等于第二高报价, 这正好是威克瑞拍卖机制的结果。

需要提到的是, 格罗夫斯机制一般不是平衡的。一个格罗夫斯机制是平衡的, 如果对所有的  $b \in V$ , 有

$$\sum_{i=1}^n t_i(b) = 0. \quad (14)$$

这样, 尽管格罗夫斯机制可导致公共商品的有效配置, 但以上等式一般来说不成立, 它并不会导致帕累托有效配置。这是由于一个帕累托最优配置必须满足平衡条件(10)和最大化条件(11)。如果, 平衡条件不满足, 它就不是帕累托有效的。将  $t_i$  代入(14), 可看出一个格罗夫斯机制导致了帕累托最优配置当且仅当它是平衡的, 即

$$(n-1) \sum_{i=1}^n b_i(d(b)) + \sum_{i=1}^n g_i(b_{-i}) \equiv 0. \quad (15)$$

然而, 当经济环境类特别小时(在效用函数空间上是无处稠密的), 有可能存在着平衡的格罗夫斯机制。Green和Laffont(1979)及Laffont和Maskin(1980)证明了存在一个平衡格罗夫斯机制的充分必要条件是价值函数满足下列微分方程:

$$\sum_{i=1}^n \frac{\partial^{n-1}}{\partial \theta_{-i}} \left[ \frac{\partial V_i}{\partial y} \frac{\partial y^*}{\partial \theta_i} \right] \equiv 0.$$

这里  $\frac{\partial^{n-1}}{\partial \theta_{-i}}$  表示  $\frac{\partial^{n-1}}{\partial \theta_1 \dots \partial \theta_{i-1} \partial \theta_{i+1} \dots \partial \theta_n}$ 。然而, 直到Tian(1993), 人们只是对特殊的二次效用函数形式

$$V_i(y, \theta_i) = \theta_i y - y^2/2$$

及参与人的个数大于 2 证明了这种可能性 (参见 Groves and Loeb, 1975; Green and Laffont, 1979; Laffont and Maskin, 1980)。笔者 (Tian, 1996a; Liu and Tian, 1999) 对更宽类的价值函数

$$\{V_i(y, \theta_i) = \psi_i(\theta_i)\phi(y) - (b\phi(y) + c)^\alpha\}_{i=1}^n$$

和

$$V_i(y, \theta_i) = \psi_i(\theta_i)\phi(y) - G_i(y) + \varphi_i(\theta_i)$$

研究了平衡的格罗夫斯机制的可能性。

### (三) 纳什均衡与纳什实施

以上所有的结果都是在优势自利行为策略解假设下, 讨论一个社会目标的激励相容性问题。赫维兹的不可能性定理说明了真实显示偏好 (优势均衡) 与资源的帕累托最优配置一般来说不可能同时达到。我们知道优势策略均衡是一个非常强的解的概念。如果我们采用较弱的纳什均衡解的概念来描述人的自利行为, 是否能设计某个激励机制, 它能导致帕累托最优配置呢? 本小节来回答这个问题。

学过博弈论的读者都知道所谓纳什策略是每个人将其他人的策略视为给定, 选择对自己最有利的策略。一个策略是纳什均衡 (Nash equilibrium) 当且仅当每个人的均衡策略是对其他人的均衡策略的最佳反应。对给定的  $e \in E$ , 一个机制  $\Gamma = \langle M, h \rangle$  被说成有纳什均衡  $m^*$ , 当且仅当对所有的  $i \in N$

$$u_i[h(m^*)] \geq u_i[h(m_i, m_{-i}^*)] \quad m_i \in M_i.$$

$N(e, \Gamma)$  表示机制  $\Gamma$  的所有纳什均衡策略的集合。

下面的定理证明了, 如果真实显示策略是一个直接显示机制的纳什均衡, 则它也是一个优势均衡。这样, 如果只是将注意力集中在直接显示机制, 采用纳什均衡解并不会得到任何新的结果。

**定理 7** 如果一个社会选择目标  $f$  可由一个直接显示机制按纳什均衡弱实施, 则这个社会选择目标  $f$  可由一个直接显示机制按优势均衡弱实施。

证明见 Dasgupta, Hammond 和 Maskin (1979)。

由显示原理, 真实显示纳什均衡与优势策略均衡配置是等价的。这样, 只要人们坚持采用直接显示机制, 就不会得到任何新结果。为了得到比优势均衡更满意的结果, 人们就不能采用直接显示机制, 而应采用具有更一般信息空间的机制, 即策略空间不完全是由参与者的经济特征所组成。当采用一般的信息空间后, 即使我们仍然假定人们的自利行为是按纳什均衡原则行事, 激励相容与最优资源配置同时达到也并不是不可能的。即使每个人都从个人的利益出发, 只要我们用一定的规则去引导, 也能够导致资源的最优配置或其他社会目标。不过, 如采用纳什均衡策略来描述人的自利行为, 弱实施不是

一个很有用的概念。为了说明这点,考虑任何社会选择目标  $f$  和下列机制: 每个人信息空间完全由经济环境空间组成, 即  $M_i = E$ , 定义结果函数  $h: E \rightarrow Z$  如下: 如果所有人报出的经济环境相同, 即,  $m_i = e \quad \forall i \in N$ , 则  $h(m) = f(e)$ , 否则每个人都被惩罚得到一个最差的结果  $z_0 \in Z$ , 即  $h(m) = z_0$ 。很容易看出, 真实地报出经济特征组合  $e$  是这个机制的纳什均衡, 所以这个机制弱实施了社会目标  $f$ 。但是, 这个机制还存在许多其他的纳什均衡。比如, 大家合伙一齐谎报任何一个假的经济特征组合  $e'$  也是一个纳什均衡。显然, 它是一个非常弱的实施形式,  $D(e, \Gamma)$  不仅包括了真实的  $e$ , 也包括了整个经济环境类  $E$ , 从而它有无穷不可数这么多个解。因此, 在纳什实施时, 人们要求一个社会目标被实施或完全实施。

### 1. 纳什可实施社会目标的特征化

在本节, 我们讨论什么样的社会选择目标是可以通过纳什激励机制达到。Maskin 在 1977 年(直到 1999 年才正式发表)对一般的社会目标对应给出了它是纳什激励相容(可完全实施)的充分必要条件。Maskin 的研究不仅能帮助我们理解什么样的社会目标是纳什可实施的, 而且也提供了在其他均衡假设下研究一个社会目标能否被实施的基本技巧和方法。Maskin 给出了一个纳什实施的直观必要条件, 称之为 Maskin 单调性条件。这个条件可以用两种不同方式来表述。不难证明它们是等价的。在严格陈述 Maskin 单调性条件之前, 我们先讨论一个纳什可完全实施的社会对应需要满足什么样的条件。

假定一个社会选择对应是纳什可完全实施的。这就意味着存在一个可实施的机制  $\Gamma = \langle M, h \rangle$  使得  $h[N(e, \Gamma)] = F(e)$  对所有的  $e \in E$  成立。考虑一个经济环境  $e$  和一个配置结果  $x \in F(e)$ 。既然  $F$  是纳什可完全实施的, 于是存在着一个纳什均衡  $m \in N(e, \Gamma)$  使得  $h(m) = x$ 。如果存在另外一个经济环境  $\bar{e}$  使得  $x \notin F(\bar{e})$ , 则  $\Gamma = \langle M, h \rangle$  纳什完全实施了  $F$  这一事实就意味着  $m$  不可能是在经济环境  $\bar{e}$  下的一个纳什均衡。这样, 就必定存在一个  $i$  和  $\bar{m}_i$  使得  $\bar{u}_i[h(\bar{m}_i, m_{-i})] > \bar{u}_i[h(m)]$ 。既然  $m$  是在经济环境  $e$  下的一个纳什均衡, 于是有  $u_i[h(m)] \geq u_i[h(\bar{m}_i, m_{-i})]$ 。令  $y = h(\bar{m}_i, m_{-i})$ , 我们有下列条件:

Maskin 单调性条件: 如果对任何两个经济环境  $e$  和  $\bar{e}$  及  $x \in F(e)$  使得  $x \notin F(\bar{e})$ , 存在某个  $i$  及另外  $y$  使得  $u_i(x) \geq u_i(y)$  并且  $\bar{u}_i(y) > \bar{u}_i(x)$ , 社会选择对应  $F$  被认为是 Maskin 单调的。

于是有以下定理:

**定理 8 (Maskin 定理 (1998))** 如果一个社会选择对应  $F$  是纳什完全可实施的, 则它是 Maskin 单调的。

为了较容易地验证 Maskin 单调性条件, 我们介绍它的一个等价条件。考虑一个经济环境  $e$  和一个社会最优结果  $x \in F(e)$ 。假定  $\bar{e}$  是另外一个经济环境使得  $u_i(x) \geq u_i(y)$  对所有的  $y$  成立, 就意味着  $\bar{u}_i(x) \geq \bar{u}_i(y)$  对所有的  $y$  成立。既然  $F$  是纳什可完全实施的, 于是存在着一个纳什均衡  $m \in N(e, \Gamma)$  使得  $h(m) = x$ , 并且  $u_i[h(m)] \geq u_i[h(\bar{m}_i, m_{-i})]$  对所有的  $i$  和  $\bar{m}_i$  成立。这样, 从

$u_i(x) \geq u_i(y)$  对所有的  $y$  成立就得到  $\bar{u}_i(x) \geq \bar{u}_i(y)$  对所有的  $y$  成立, 我们必定有  $\bar{u}_i[h(m)] \geq \bar{u}_i[h(\bar{m}_i, m_{-i})]$  对所有的  $i$  和  $\bar{m}_i$  成立。因此,  $m$  也是在  $\bar{e}$  下的一个纳什均衡。于是我们有下面等价的 Maskin 单调性条件:

等价的 Maskin 单调性条件: 如果对任何两个经济环境  $e$  和  $\bar{e}$  及  $x \in F(e)$  使得对所有的  $i \in N$  和所有的  $y$ ,  $u_i(x) \geq u_i(y)$  意味着  $\bar{u}_i(x) \geq \bar{u}_i(y)$  对所有的  $y$  成立, 则  $x \in F(\bar{e})$ , 我们就说社会选择对应  $F$  是 Maskin 单调的。这个等价的 Maskin 单调性条件通常被用来检查一个社会选择对应是否满足 Maskin 单调性条件。

Maskin 单调性条件本身并不能保证一个社会选择对应是纳什可完全实施的。但是, 如果再加上下面所谓的“个人无否决权”条件, Maskin 单调性条件成为纳什可完全实施的充分条件。

“个人无否决权”条件: 一个社会选择对应  $F$  满足“个人无否决权”(no veto power) 条件当且仅当: 对于任意的  $i, e, x \in Z$ , 使得  $u_j(x) \geq u_j(y)$  对所有的  $y \in Z$  和  $j \neq i$  成立, 则  $x \in F(e)$ 。“个人无否决权”条件意味着如果  $n-1$  个人认为一个资源配置  $x$  对他们来说是最优的, 则  $x$  是属于社会目标集合  $F(e)$ 。这是一个相当弱的条件。例如, 对具有至少三个参与者的私人商品经济, 如果每个人的效用函数是单调的, 则不存在任何资源配置, 它对一个以上的参与者同时是最好的, 从而“个人无否决权”条件显然满足。于是有下面定理:

**定理 9 (Maskin 定理)** 如果  $F$  是 Maskin 单调的, 满足“个人无否决权”条件, 及  $n \geq 3$ (至少有三个参与者), 则它是纳什完全可实施的。

证明见 Maskin(1999)。

前面提到的以色列国王所罗门解决两个妇女争夺婴儿所属权的例子可看成是一个纳什实施问题。这是由于这两个妇女自己知道谁是小孩真正的母亲, 但国王不知道谁在说谎。然而, 国王所罗门的机制却不满足 Maskin 单调条件, 因而它不是纳什可实施的。如果第二个女人聪明一点也要求不要把孩子杀了, 把孩子给对方, 则所罗门国王的激励机制不能解决问题。现在我们将国王所罗门解决婴儿归属问题的方法用机制设计的语言来描述。两个妇女分别叫 Anne 和 Bath。有两种状态(两种经济环境):  $\alpha$  表示 Anne 是婴儿的真正母亲和  $\beta$  表示 Beth 是婴儿的真正母亲。国王有三种选择可能: (a) 将婴儿给 Anne; (b) 将婴儿给 Bath; (c) 将婴儿分割(cut)成两半。他的目的是想要将婴儿判给真正的母亲这样一个目标:  $f(\alpha) = a$  及  $f(\beta) = b$ 。Anne 的偏好关系是: 在状态  $\alpha$  下有  $a P_A^\alpha b P_A^\alpha c$ , 而在状态  $\beta$  下有  $a P_A^\beta c P_A^\beta b$ 。Beth 的偏好关系是: 在状态  $\alpha$  下有  $b P_B^\alpha c P_B^\alpha a$ , 而在状态  $\beta$  下有  $b P_B^\beta a P_B^\beta c$ 。即, 每个妇女都希望得到婴儿。如果她是婴儿的真正母亲, 在得不到的情况下, 她情愿放弃小孩, 而不愿看到小孩被杀死。如果她不是婴儿的真正母亲, 她宁愿小孩被杀死, 也不愿意给与竞争对手。注意,  $f(\alpha) = a$ , 并且对 Anne 来说, 无论是在状态  $\alpha$  还是在状态  $\beta$  下,  $a$  相对于其他两种选择都是最佳选择。

于是根据 Maskin 单调条件, 我们应该有  $f(\beta) = a$ , 但显然这并不成立(因为  $f(\beta) = b$ )。这样, 国王的社会目标函数不是 Maskin 单调的, 从而由 Maskin 定理, 它不是纳什完全可实施的。所罗门国王所面临判决问题比《旧约全书》所给出的方法要复杂的多。

## 2. 性质优良机制设计与资源有效配置的纳什实施

在证明纳什可实施社会目标的特征化的定理中所用到的机制一般来说非常复杂。比如, 它用到了一个无穷维的信息空间, 结果函数是不连续的。性质优良机制是指一个机制具有某些好的性质, 包括信息空间的维数有限, 甚至达到最小, 机制是连续的、可行的等优良性质。如果一个机制具有以上提到的某些优良性质, 它也许就更具有可操作性, 可能被应用到实际当中去。

### (1) 格罗夫斯-利加德机制

格罗夫斯(Groves)和利加德(Ledyard)在1977年第一个给出了性质优良的非显示经济激励机制。在纳什均衡的原则下, 对具有公共商品的经济社会, 他们的机制产生了资源最优配置。

为了说明格罗夫斯-利加德机制的基本结构, 考虑一个简单化的格罗夫斯-利加德机制。考虑一类公共商品经济环境, 它只有一个私人商品  $x_i$ , 一个公共商品  $y$  和三个参与者 ( $n=3$ )。生产公共商品的生产函数为:  $y = x$ 。

格罗夫斯-利加德机制定义如下:

信息空间为:  $M_i = R$ , 它的一个点可解释为参与者  $i$  为生产此公共商品所愿意作出的捐献。结果函数:  $t_i(m) = m_i^2 + 2m_j m_k$   $i \neq j \neq k$  是由机制决定的参与者  $i$  的实际转移支付;  $y(m) = (m_1 = m_2 + m_3)^2$  是由机制决定的公共商品水平, 及  $x_i(m) = w_i - t_i(m)$  是  $i$  的私人商品的消费量。由于  $\sum_i [x_i(m) + t_i(m)] = \sum_i x_i(m) + y(m) = \sum_i w_i$  对所有的  $i$  和  $m \in M$  都成立, 这个机制是平衡的。参与者  $i$  的支付函数为:

$$v_i(m) = u_i(x_i(m), y(m)) = u_i(w_i(m) - t(m), y(m)).$$

从纳什均衡内点所满足的一阶条件, 我们有  $\frac{\partial u_i / \partial y}{\partial u_i / \partial x_i} = \frac{m_i}{\sum_i m_i}$ 。从而, 总和起来有  $\sum_i \left[ \frac{\partial u_i / \partial y}{\partial u_i / \partial x_i} \right] = \sum_i \left[ \frac{m_i}{\sum_i m_i} \right] = 1$ 。于是帕累托最优配置的一阶条件成立。

当效用函数是拟凹时, 一阶必要条件也是充分的。这样, 格罗夫斯-利加德机制纳什实施了帕累托有效配置。

于是他们认为他们自己解决了“搭便车”的问题。但经济学往往是复杂的: 有的人认为他们的确解决了“搭便车”的问题, 而另外一些人则不以为然。理由有两点: 一是这个机制不能保证导致的配置是个人理性配置, 即通过机制分配的结果对某些人来说比他们以前持有的初始资源的效用还要低, 从而就有人可能不愿意参与这个机制来进行资源再配置, 因为参加后反而损害了自己的利益; 二是对于某些非均衡策略, 导致了  $x_i(m) < 0$ , 从而它是个人不可行配置, 即通过机制配置的资源不在个人的消费约束集之内。于是人们也许会问: 能否设计这样的机制——它能产生资源的最优配置, 而这个配

置又是个人理性的配置呢？答案是肯定的。我们知道林达尔配置及瓦尔拉斯配置导致了帕累托最优配置。赫维兹在 1979 年分别对公共商品的经济环境类及私人商品的环境类给出了这样的机制，它们分别实施了林达尔配置和瓦尔拉斯配置，从而它们导致了资源最优和个人理性配置。沃克 (Walker) 在 1981 年也给出了类似的机制。我们先给出林达尔均衡的定义。

对只有一个私人商品  $x_i$ ，一个公共商品  $y$  和  $n$  个参与者，生产函数为  $y = x$  的经济环境类，资源配置  $z = (x, y) = (x_1, x_2, \dots, x_n, y) \in R_+^n \times R_+$  被称之为一个林达尔均衡配置，当且仅当它是可行的，并且存在着一组个人化价格向量  $(q_1, \dots, q_n) \in R_+^n$  使得

- ①  $x_i + q_i y = w_i, \quad i = 1, \dots, n;$
- ② 对所有的  $i = 1, \dots, n, (x'_i, y')$   $P_i(x_i, y)$  意味着  $x'_i + q_i y' > w_i;$
- ③  $\sum_{i=1}^n q_i = 1.$

在单调性假设下，每个林达尔配置显然是个人理性的，同时也是帕累托最优的。我们现在简单介绍沃克机制。

### (2) 沃克机制

信息空间为： $M_i = R$ 。结果函数： $y(m) = \sum_{i \in N} m_i$  是由机制所决定的公共商品的实际水平； $i$  的个人化的林达尔价格为  $q_i(m) = 1/n + m_{i+1} - m_{i+2}$ ，显然有  $\sum_i q_i(m) = 1$ ， $t_i(m) = q_i(m)y(m)$  是由机制决定的实际转移支付，及  $x_i(m) = w_i - t_i(m)$  是  $i$  的私人商品的消费量。由于  $\sum_i [x_i(m) + t_i(m)] = \sum_i x_i(m) + y(m) = \sum_i w_i$  对所有的  $i$  和  $m \in M$  都成立，这个机制是平衡的。参与者  $i$  的支付函数为：

$$v_i(m) = u_i(x_i(m), y(m)) = u_i(w_i(m) - t(m), y(m)).$$

由纳什均衡内点所应满足的一阶条件，我们有  $\frac{\partial u_i / \partial y}{\partial u_i / \partial x_i} = q_i(m)$ 。这也是林达尔解的一阶条件。当效用函数是拟凹时，一阶必要条件也是充分的。这样我们知道沃克机制纳什完全地实施了林达尔配置。然而人们对赫维兹和沃克的机制仍不太满意，因为他们的机制不能保证个人可行性条件，并且赫维兹的机制利用了一个较大维数的信息空间。赫维兹在他的另外一篇文章里给出了一个保证个人可行性条件的配置机制，这个机制产生了资源有效配置。然而，它不是连续的，即使微小的信息传递误差也会导致较大差异的资源配置结果，这在实际应用中就会出现精确性问题；另外它也不是预算平衡的，即通过机制配置的资源超过了社会的总资源。那么人们是否能够设计一个既是个人可行又是预算平衡的机制呢？赫维兹等人 (Hurwicz, Maskin, and Postlewaite, 1984) 证明一个机制如果产生了个人可行同时又是预算平衡的机制，则信息空间必依赖于初始资源。

### (3) 田氏机制

以上所提到的机制总有这样或那样的一些令人不太满意的缺点。于是人们也许会问:是否存在着一个机制,它可以产生有效、个人理性的资源配置,并且是连续的、个人可行的、平衡的?笔者在 Tian (1989) 一文中回答了这个问题。对具有公共商品的经济环境类,在纳什行为下,笔者证明并给出了这样的一个机制——它具有以上提到的所有性质。在 Tian (1990) 中给出了一个具有同样性质,但具有最小维数信息空间的激励机制。这个机制和沃克机制的其他地方都相同,除了对公共商品的结果函数由下式给出:

$$y(m) = \begin{cases} 0 & \text{如果 } \sum_i m_i \leq 0 \\ \sum_i m_i & \text{如果 } 0 \leq \sum_i m_i \leq a(m) \\ a(m) & \text{如果 } \sum_i m_i \geq 0 \end{cases}$$

这里  $a(m) = \min_{i \in N(m)} \frac{w_i}{q_i(m)}$  及  $N(m) = \{i \in N : q_i(m) > 0\}$ 。可以证明,这个机制完全实施了林达尔配置。

对非完全非传递的偏好关系、具有生产报酬递减或当个人初始资源是不完全信息时,笔者在一系列已发表的文章中给出了类似的机制 (Tian, 1991, 1992; Tian and Li, 1991; Li, Nakamura, and Tian, 1995)。对于私人商品的经济环境类,笔者 (Tian, 1992) 也给出了类似的机制。对以上这些结果的详细讨论见后面的参考文献书中那些论文,它们都已发表在欧美经济学主要期刊上。另外,有趣的是赫维兹在 1979 年的另外一篇文章中证明了对于新古典的经济环境类,通过任何经济激励机制所产生的资源有效和个人理性配置都可以通过实施瓦尔拉斯配置来达到。这个结果对于修补市场的局限性有很大的帮助。例如,对于只有几个买者和卖者的市场,我们有充分的理由相信市场不是完全竞争的。这样,市场所导致的配置一般不是有效的。然而人们可以利用其他经济机制,使得它的纳什均衡配置与假定下的完全竞争市场机制所导致的配置一样,从而它导致的配置是帕累托有效的。

以上所设计的激励经济机制都是以私有制经济环境为前提的。前面我们已提到兰格针对公有制提出的边际成本定价机制不是激励相容的。人们也许会问,是否能为公有制设计出一些有效和个人理性的激励机制呢?答案也是肯定的。至少对公共商品的情况是如此。笔者一直从事着这方面的工作 (Tian, 1994b, 2000b, 2000c, 2000d; Tian and Li, 1994, 1995b)。笔者通过设计具体的机制证明了即使国营企业在不追求利润最大化的条件下,只要把生产公共商品的成本让消费者根据他们自己的偏好来分担,在适当的成本分担机制下,所设计的机制能导致资源的有效和个人理性的配置。这样,这些机制解决了市场机制不能很好解决公共商品的问题。如果把这些机制和市场机制结合起来,即使在具有公共商品的情况下,也能解决资源的有效配置问题。当然所给出的机制离实际应用还有一段距离,并且由前面的信息有效性结果,这些机制不可能是信息有效的。不过,这些模型说明,在具有公共商品的情况下,这种由国营企业生产,让消费者分担成本的方式,尽管不是信息有效的,至少在理论上能解决资源有效配置的问题。详细的讨论见相关文章。

需要提到的是，在经济机制设计理论中，大多都是把激励相容问题和信息效率问题分开来考虑的。激励相容理论只考虑在给定的自利行为准则下，一个既定目标可实施的条件，而不考虑机制的信息要求量问题。信息效率理论只考虑实现一个社会目标所需要的信息量（即信息空间的维数）的问题，而忽略了机制的激励问题。Reichelstein 和 Reiter(1988) 同时考虑了这两个问题。他们证明了，在纳什激励相容的条件下，实施一个社会目标所需要的信息量不会少于与不考虑激励问题而实现同一既定目标所需要的信息量。事实上，对公共商品的环境类，实施林达尔机制的最小维数和实现林达尔配置的最小维数一样，Walker(1981) 及笔者(Tian, 1990, 1991) 给出了具体的这样的激励机制。但是，对私人商品的环境类，Reichelstein 和 Reiter(1988) 证明实施瓦尔拉斯机制的最小维数要比实现瓦尔拉斯配置的最小维数要大。

#### (四) 精练纳什实施与近似纳什实施

尽管上面给出了一些纳什可实施的社会目标的例子及纳什可实施社会目标的充分必要条件，但有许多社会目标不是单调的（比如国王所罗门解决婴儿归属问题），因而在纳什均衡解意义上它是不可实施的。经济学文献中给出了可大大地扩大可实施社会目标范围的两种方法。一种是通过采用精练纳什均衡解的方法。利用纳什策略，可能导致多个纳什均衡解，精练纳什均衡解的概念给出了剔除那些缺乏说服力的纳什均衡点的方法。另外一种是采用近似地纳什实施一个社会目标的方法，它只要求配置结果任意地接近社会选择对应。

##### 1. 精练纳什实施

什么样的社会目标在精练纳什均衡解的假定下可实施呢？尽管纳什均衡结果的集合  $N(e)$  可能不是社会目标集合  $F(e)$  的子集时（于是这个社会目标可能不是纳什可实施的），但由于精练纳什均衡结果的集合比纳什均衡结果的集合要小得多，精练纳什均衡结果的集合却可能是社会目标集合  $F(e)$  的子集，因而在精练纳什均衡解的情况下也许是可实施的。本小节介绍几种精练纳什可实施的概念。

一个精练纳什均衡的概念是强纳什均衡 (strong Nash equilibrium) 假设。它意味着：当一个策略处于强均衡状态时，对任何一组人所形成的小集团，当其他人策略给定时，这个集团中的每个人都不可能从合作中得到更大的好处。显然，强纳什均衡策略是比纳什均衡策略要求更强的均衡假设，每一个强纳什均衡显然是纳什均衡，反之却不成立。这样，强纳什均衡的集合是纳什均衡的一个子集。由于这个均衡集合比纳什均衡集合可能要小，通过强纳什均衡实施的社会选择目标集合可能会更大。Maskin (1999) 证明：对适当的偏好关系的集合类，任何导致了帕累托最优配置和理性配置的社会选择目标都是强纳什可实施的。

人的利己行为策略假设还可以是塞尔顿 (Selton) 所引进的子博弈纳什均衡 (subgame perfect Nash equilibrium) 策略假设或其他精练纳什均衡 (refinement of Nash equilibrium) 解。另外还有许多种策略均衡解来表达人的个人利己行为, 如非优势的纳什均衡 (non-dominanted Nash equilibrium) 或其他精练纳什均衡解给出。Moore 和 Repullo (1988), Abreu 和 Sen (1990) 等人证明了在子对策纳什均衡假设下, 几乎所有的社会目标都是可实施的。Palfrey 和 Srivastava (1990) 在非优势的纳什均衡的假设下也证明了同样的结果。

## 2. 近似纳什实施

人们也可通过近似地纳什实施一个社会目标的方法来扩大可实施社会目标的范围。尽管纳什均衡结果的集合  $N(e)$  不能完全包含在社会目标  $F(e)$  集合中, 但只要每个纳什均衡配置可以任意地接近  $F(e)$  中某个配置 (即这两个配置的差 (距离) 可以任意小), 我们就说这个社会目标是可近似实施的。Matsushima (1988), Abreu 和 Sen (1991) 等人证明了几乎所有的社会目标都是可近似实施的。

## (五) 双实施: 纳什实施和强纳什实施

纳什均衡策略是一种非合作的策略均衡概念, 它完全排除任何合作的可能性。尽管纳什均衡相对容易达到, 但它也许是不稳定的: 参与者往往会通过某种形式的合作来对付机制设计者, 因为通过合作可能会得到更大的好处。这样, 人的利己行为由纳什策略均衡来描述的假设也许是不现实的, 从而采用强纳什策略均衡的假设也许会更合理。由于强纳什均衡的集合显然是纳什均衡的子集合, 通过强纳什均衡实施一个社会选择目标的可能性会更大。

尽管强纳什均衡假设更为合理, 但它可能不存在或较难被解出。为了能够相对容易达到, 同时也是稳定的, 人们自然要求一个社会选择目标能同时被纳什和强纳什均衡实施。这就是所谓的双实施 (double implementation)。Suh (1997) 给出了一个社会选择目标能被纳什和强纳什同时双实施的充分必要条件。由于 Suh 的特征化结果只是考虑了一个社会选择目标被双实施的充分必要条件, 而没有考虑机制的复杂性, Peleg (1996a, 1996b) 及笔者 (Tian, 1999a, 2000a, 2000b, 2000c, 2000d) 给出了一系列激励机制, 它们双实施了瓦尔拉斯配置、林达尔配置和其他导致了帕累托最优配置和理性配置的社会选择目标。

## (六) 不完全信息与贝叶斯实施

上面的纳什实施、精练纳什实施及近似纳什实施对信息的要求用到了一个非常强的假定。尽管机制设计者不需要知道所有参与者的经济特征, 但纳什均衡及近似纳什均衡解的一个缺点是它假定各个参与者知道其他参与者的经济特征。在现实中这个假设难以满足。这个假设是否可去掉呢? 答案是肯定的。人们可用海萨尼 (Harsanyi) 引进的贝叶斯 - 纳什均衡 (Bayesian-Nash equilibrium) 解来假定参与者的利己行为。尽管各个参与者不知道其他参与者

的经济特征，贝叶斯解假定每个人只要知道其他人的经济特征的概率分布情况。在这种情况下，人们仍然可以设计出某类激励相容的机制。

为了简单起见，假定效用函数的变动只依赖于参数  $\theta_i$  的变动。它可表示为  $u_i(x, \theta_i)$ 。假定所有的参与者和机制设计者都知道效用类型向量  $\theta = (\theta_1, \dots, \theta_n)$  在先验集  $\Theta$  上按概率密度函数  $q(\theta)$  分布。每个参与者知道他自己的类型  $\theta_i$ ，并能够计算其他人属于什么类型的条件分布：

$$q(\theta_{-i}|\theta_i) = \frac{q(\theta_i, \theta_{-i})}{\int_{\Theta_{-i}} q(\theta_i, \theta_{-i}) d\theta_{-i}}$$

像前面一样，一个机制是由  $\Gamma = \langle M, h \rangle$  给定。给定机制  $\langle M, h \rangle$ ，每个参与者  $i$  信息  $m_i$  的选择是  $\theta_i$  的函数： $\sigma_i : \Theta_i \rightarrow M_i$ 。令  $\Sigma_i$  是  $i$  的所有策略的集合。给定  $\sigma = (\sigma_1, \dots, \sigma_n)$ ， $i$  在  $t_i$  类型下的期望效用为

$$\Pi_{\langle M, h \rangle}^i(\sigma; \theta_i) = \int_{\Theta_{-i}} u_i(h(\sigma(\theta)), \theta_i) q(\theta_{-i}|\theta_i) d\theta_{-i}. \quad (16)$$

一个策略  $\sigma$  是机制  $\langle M, g \rangle$  的贝叶斯 - 纳什均衡，当且仅当对所有的  $\theta_i \in \Theta_i$ ，

$$\Pi^i(\sigma; \theta_i) \geq \Pi^i(\hat{\sigma}_i, \sigma_{-i}; \theta_i) \quad \forall \hat{\sigma}_i \in \Sigma_i$$

成立。给定一个机制  $\langle M, h \rangle$ ，它的贝叶斯均衡集合依赖于经济环境，记为  $B(e)$ 。象纳什实施一样，贝叶斯激励相容也涉及到  $F(e)$  和  $B(e)$  这两个集合的关系问题。当对  $E$  中所有的经济环境  $e$ ， $B(e)$  是  $F(e)$  的一个非空子集时，我们称机制  $\langle M, h \rangle$  贝叶斯实施了社会目标  $F$ 。如果对于某个给定的社会目标对应  $F$ ，存在着某个经济机制贝叶斯实施了这个社会目标对应，我们就称这个社会目标是贝叶斯可实施的。

Pastlewaite 和 Schmeidler (1986)，Palfrey 和 Srivastava (1989)，Mookherjee 和 Reichelstein (1990)，Jackson (1991)，Dutta 和 Sen (1994) 等人，以及笔者 (Tian, 1996a, 1999b) 对一般的社会目标及各种经济环境类给出了它是贝叶斯可实施的充分必要条件。在各种技术性的条件下，他们证明了一个社会目标对应是贝叶斯可实施的充分必要条件是： $F$  是贝叶斯单调的并且是贝叶斯激励相容的。

同样的理由，人们可考虑精练贝叶斯实施及近似贝叶斯实施。Palfrey 和 Srivastava (1989) 在非优势的贝叶斯均衡解的假设下证明了，对至少有三个参与者的经济环境类，一个社会目标对应是非优势贝叶斯可实施的充分必要条件是： $F$  是贝叶斯激励相容的。Abreu 和 Matsushima (1990)，Matsushima (1993)，Duggan (1993) 以及笔者 (Tian, 1997) 等人在各种技术性条件下证明了一个社会目标对应是近似贝叶斯可实施的充分必要条件是：这个社会目标是贝叶斯激励相容的。

## 五、结束语

本文大致地介绍了机制设计理论一些基本成果,讨论了自利行为、自由选择、不完全信息、分散决策、激励与经济机制设计的关系问题。激励机制设计理论所讨论的问题是:对于任何给定的社会目标,在自由选择、自愿交换、信息分散化决策条件下,能否设计、怎样设计一个经济机制来达到既定的社会目标或其他经济目标。由于篇幅有限,还有许多重要的成果没有或得不到详细介绍,有兴趣更进一步了解机制理论的读者,可阅读本文所列出的有关参考文献。笔者希望读者通过本文能够对所介绍的机制设计理论有一大致的了解,更希望所介绍的理论对读者在思考、分析和解决中国经济改革、制度转型和企业管理问题时有所帮助。

## 参考文献

- [1] Abreu, R. and A. Sen, "Subgame Perfect Implementation: A Necessary and Almost Sufficient Condition." *Journal of Economic Theory*, 1990, 50, 285-299.
- [2] Abreu, R. and A. Sen, "Virtual Implementation in Nash Equilibrium." *Econometrica*, 1991, 59, 997-1021.
- [3] Arrow, K. J., "An Extension of the Basic Theorems of Welfare Economics." *Proceedings of the 2nd Berkeley Symposium*, University of California Press, 1951, 507-532.
- [4] Arrow, K. J. and Hahn, F. H., *General Competitive Analysis*, San Francisco: Holden Day, 1971.
- [5] Bergson, A., "A Reformulation of Certain Aspects of Welfare Economics." *Quarterly Journal of Economics*, 1938, 52, 310-334.
- [6] Calsamiglia, X., "Decentralized Resource Allocation and Increasing Returns." *Journal of Economic Theory*, 1977, 14, 263-283.
- [7] Dasgupta, P., P. Hammond and E. Maskin, "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility." *Review of Economic Studies*, 1979, 46, 185-216.
- [8] Debreu, G., *Theory of Value*. New York: Wiley, 1959.
- [9] Duggan, J., "Virtual Implementation in Bayesian Equilibrium with Infinite Types." Parts I and II, Mimeo, 1993.
- [10] Groves, T. and Ledyard, J., "Optimal Allocation of Public Goods: A Solution to the Free Rider Problem." *Econometrica*, 1977, 45, 783-811.
- [11] Groves, T. and Ledyard, J., "Incentive Compatibility Since 1972." Chapter 2: *Information, Incentive, and Economic Mechanisms*, T. Groves, R. Radner, and S. Reiter, eds., University of Minnesota Press, 1987.
- [12] Hayek, F. A. von, "The Present State of the Debate." in *Collectivist Economic Planning*, F. A. von Hayek, ed., London, 1935 201-243.
- [13] Hayek, F. A. von, "The Use of Knowledge in Society." *American Economic Review*, 1945, 35, 519-530.
- [14] Hurwicz, L., "On Informational Decentralized Systems." in *Decision and Organization*, Radner, R. and C. B. McGuire, eds., in Honor of J. Marschak, North-Holland, 1972, 297-336.

- [15] Hurwicz, L., "Outcome Function Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Point." *Review of Economic Studies*, 1979a, XLVL, 397–419.
- [16] Hurwicz, L., "On Allocations Attainable Through Nash Equilibria." *Journal of Economic Theory*, 1979b, 21, 140–165.
- [17] Hurwicz, L., "Balanced Outcome Functions Yielding Walrasian and Lindahl Allocations at Nash Equilibrium Points for Two or More Agents." in *General Equilibrium, Growth, and Trade*, Jerry R. Green and Jose A. Scheinkman, eds., New York: Academic Press, 1979c.
- [18] Hurwicz, L., "Socialism and Incentives: Developing a Framework." *Journal of Comparative Economics*, 1979d, 3, 207–216.
- [19] Hurwicz, L., "Incentive Aspects of Decentralization." in *Handbook of Mathematical Economics*, K. J. Arrow and M. D. Intriligator, eds., Vol. III, NorthHolland, 1986a.
- [20] Hurwicz, L., "On Informational Decentralization and Efficiency in Resource Allocation Mechanism." in *Studies in Mathematical Economics*, S. Reiter, ed., Mathematical Association of America, 1986b.
- [21] Hurwicz, L., "On the Implementation of Social Choice Rules in Irrational Societies." in *Social Choice and Public Decision Making Essays in Honor of Kenneth J. Arrow*, ed., Vol. I, Cambridge: Cambridge University Press, 1986c.
- [22] Hurwicz, L., "Revisiting Externalities." *Journal of Public Economic Theory*, 1999, 1, 225–246.
- [23] Hurwicz, L., E. Maskin, and A. Postlewaite, "Feasible Implementation of Social Choice Correspondences by Nash Equilibria." Mimeo, 1984.
- [24] Hurwicz, L. and M. Walker, "On the Generic Nonoptimality of Dominant-Strategy Allocation Mechanism: A General Theorem that Includes Pure Exchange Economies." *Econometrica*, 1990, 58, 683–704.
- [25] Jackson, M.O., "Bayesian Implementation." *Econometrica*, 1991, 59, 461–477.
- [26] Jordan, J. S., "The Competitive Allocation Process in Informationally Efficient Uniquely." *Journal of Economic Theory*, 1982, 28, 1–18.
- [27] Kreps, D., "Signalling." in *Handbook of Game Theory*, R. J. Aumann and S. Hart, eds., Vol. II, North-Holland, 1994.
- [28] Laffont, J. J. and E. Maskin, "A Differential Approach to Expected Maximizing Mechanisms." in *Aggregation and Revelation of Preferences*, J. J. Laffont, ed., North Holland, 1979.
- [29] Laffont, J. J. and D. Martimort, *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press, 2002.
- [30] Lange, O., "On the Economic Theory of Socialism: part one." *Review of Economic Studies*, 1936, 4, 53–71.
- [31] Lange, O., "On the Economic Theory of Socialism: part two." *Review of Economic Studies* 1937, 4, 123–142.
- [32] Lange, O., *On the Economic Theory of Socialism*. Philadelphia: Lippincott, 1938.
- [33] Lange, O., "The Foundations of Welfare Economics." *Econometrica*, 1942, 10, 215–228.
- [34] Lange, O. and F. M. Taylor, *On the Economic Theory of Socialism*. B.E. Lippincott, ed., New York, 1938.
- [35] Lerner, A. P., *Economics of Control*, New York, 1944.

- [36] Li, Q., S. Nakamura and G. Tian, "Nash Implementation of the Lindahl Correspondence with Decreasing Returns to Scale Technology." *International Economic Review*, 1995, 36, 34-50.
- [37] Liu, L., and G. Tian, "A Characterization of Optimal Dominant Strategy Mechanisms." *Review of Economic Design*, 1999, 4, 205-218.
- [38] E. Maskin, "Nash Equilibrium and Welfare Optimality." *Review of Economic Studies*, 1999, 66, 23-38.
- [39] Matsushima, H., "A New Approach to the Implementation Problem." *Journal of Economic Theory*, 1988, 45, 128-144.
- [40] Matsushima, H., "Bayesian Monotonicity with Side Payments." *Journal of Economic Theory*, 1993, 59, 107-121.
- [41] Mookherjee, D. and S. Reichelstein, "Implementation via Augmented Revelation Mechanisms." *Review of Economic Studies*, 1990, 57, 453-475.
- [42] Moore, J. and R. Repullo, "Subgame Perfect Implementation." *Econometrica*, 1988, 56, 1191-1220.
- [43] Moore, J. and R. Repullo, "Nash Implementation: A Full Characterization." *Econometrica*, 1990, 56, 1083-1099.
- [44] Mount, K., and S. Reiter, "Informational Size of Message Spaces." *Journal of Economic Theory*, 1974, 8, 161-191.
- [45] Palfrey, T. and S. Srivastava, "On Bayesian Implementable Allocations." *Review of Economic Studies*, 1987, 54, 193-208.
- [46] Palfrey, T. and S. Srivastava, "Mechanism Design with Incomplete Information: A Solution to the Implementation Problem." *Journal of Political Economy*, 1989, 97, 668-691.
- [47] Palfrey, T. and S. Srivastava, "Implementation with Incomplete Information in Exchange Economies." *Econometrica*, 1989, 57, 115-134.
- [48] Palfrey, T. and S. Srivastava, "Nash Implementation Using Undominated Strategy." *Econometrica*, 1991, 59, 479-502.
- [49] Postlewaite, A., "Implementation in Nash Equilibria in Economic Environments." in *Social Goal and Social Organization Essays in Memory of Elisha Pazner, L. Hurwicz, D. Schmeidler, and H. Sonnenschein*, eds., Cambridge: Cambridge University Press, 1985.
- [50] Postlewaite, A. and D. Schmeidler, "Implementation in Differential Information Economies." *Journal of Economic Theory*, 1986, 39, 14-33.
- [51] Postlewaite, A. and D. Wettstein, "Continuous and Feasible Implementation." *Review of Economic Studies*, 1989, 56, 603-611.
- [52] Reichelstein, S. and S. Reiter, "Games Forms with Minimal Message Spaces." *Econometrica*, 1988, 56, 661-692.
- [53] Repullo, R., "The Revelation Principle under Complete and Incomplete Information." in *Economic Organization as Games*, Binmore, K. and P. Dasgupta, eds., Oxford: Basil Blackwell, 1986.
- [54] Samuelson, P., "A Pure Theory of Public Expenditure." *Review of Economics and Statistics*, 1954, 36, 387-389.
- [55] Samuelson, P., "Diagrammatic Exposition of a Theory of Public Expenditure." *Review of Economics and Statistics*, 1955, 37, 360-366.

- [56] Schmeidler, D., "Walrasian Analysis via Strategic Outcome Functions." *Econometrica*, 1980, 48, 1585–1593.
- [57] Tian, G., "Implementation of the Lindahl Correspondence by a Single-Valued, Feasible, and Continuous Mechanism." *Review of Economic Studies*, 1989, 56, 613–621.
- [58] Tian, G., "Completely Feasible and Continuous Nash-Implementation of the Lindahl Correspondence with a Message Space of Minimal Dimension." *Journal of Economic Theory*, 1990, 51, 443–452.
- [59] Tian, G., "Implementation of Lindahl Allocations with Nontotal-Nontransitive Preferences." *Journal of Public Economics*, 1991, 46, 247–259.
- [60] Tian, G., "Implementation of the Walrasian Correspondence without Continuous, Convex, and Ordered Preferences." *Social Choice and Welfare*, 1992, 9, 117–130.
- [61] Tian, G., "Implementing Lindahl Allocations by a Withholding Mechanism." *Journal of Mathematical Economics*, 1993, 22, 169–179.
- [62] 田国强, "中国国营企业改革与经济体制平稳转轨的方式与步骤", 《经济研究》, 1994 年第 11 期第 3—9 页。
- [63] Tian G., "On Informational Efficiency and Incentive Aspects of Generalized Ratio Equilibria." *Journal of Mathematical Economics*, 1994a, 23, 323–337.
- [64] Tian, G., "Implementation of Linear Cost Share Equilibrium Allocations." *Journal of Economic Theory*, 1994b, 64, 568–584.
- [65] Tian, G., "On the Existence of Optimal Truth-Dominant Mechanisms." *Economics Letters*, 1996a, 53, 17–24.
- [66] Tian, G., "Continuous and Feasible Implementation of Rational Expectation Lindahl Allocations." *Games and Economic Behavior*, 1996b, 16, 135–151.
- [67] Tian, G., "Virtual Implementation in Incomplete Information Environments with General Sets of Alternatives and Types." *Journal of Mathematical Economics*, 1997, 28, 313–339.
- [68] Tian, G., "Double Implementation in Economies with Production Technologies Unknown to the Designer." *Economic Theory*, 1999a, 13, 689–707.
- [69] Tian, G., "Bayesian Implementation in Exchange Economies with State Dependent Preferences and Feasible Sets." *Social Choice and Welfare*, 1999b, 16, 99–119.
- [70] Tian, G., "Double Implementation of Lindahl Allocations by a Continuous and Feasible Mechanism." *Social Choice and Welfare*, 2000a, 17, 125–141.
- [71] Tian, G., "Incentive Mechanism Design for Production Economies with Both Private and Public Ownership." *Games and Economic Behavior*, 2000b, 33, 294–320.
- [72] Tian, G., "Implementation of Balanced Linear Cost Share Equilibrium Solution in Nash and Strong Nash Equilibria." *Journal of Public Economics*, 2000c, 76, 239–261.
- [73] Tian, G., "Double Implementation of Linear Cost Share Equilibrium Allocations." *Mathematical Social Sciences*, 2000d, 40, 175–189.
- [74] Tian, G., "On Uniqueness of Informational Efficiency of the Competitive Mechanism in Production Economies." Mimeo, 2000e.
- [75] Tian, G., "A Unique Informationally Efficient Allocation Mechanism in Economies with Public Goods." Mimeo, 2000f.
- [76] Tian, G. and Q. Li, "Completely Feasible and Continuous Implementation of the Lindahl Correspondence with Any Number of Goods." *Mathematical Social Sciences*, 1991, 21, 67–79.

- [77] Tian, G. and Q. Li, "An Implementable and Informational Efficient State-Ownership System with General Variable Returns." *Journal of Economic Theory*, 1994, 64, 268-297.
- [78] Tian, G. and Q. Li, "On Nash-Implementation in the Presence of Withholding." *Games and Economic Behavior*, 1995a, 9, 222-233.
- [79] Tian, G. and Q. Li, "Ratio-Lindahl Equilibria and an Informationally Efficient and Implementable Mixed-Ownership System." *Journal of Economic Behavior and Organization*, 1995b, 26, 391-411.
- [80] 田国强、张帆, 《大众市场经济学》(市场经济学普及丛书, 田国强主编), 上海: 上海人民出版社, 1993年。
- [81] Thomson, W., "Maximum Strategies and Elicitation of Preferences." in *Aggregation and Revelation of Preferences*, J. J. Laffont, ed., North-Holland, 1979.
- [82] Varian, H.R., *Microeconomic Analysis* (Third Edition), W.W. Norton and Company, 1992.
- [83] Walker, M., "A Note on the Characterization of Mechanisms for the Revelation of Preferences." *Econometrica*, 1978, 46, 147-152.
- [84] Walker, M., "On the Nonexistence of a Dominant Strategy Mechanism for Making Optimal Public Decisions." *Econometrica*, 1980, 48, 1521-1540.
- [85] Walker, M., "A Simple Incentive Compatible Scheme for Attaining Lindahl Allocations." *Econometrica*, 1981, 49, 65-71.

## Economic Mechanism Theory: Informational Efficiency and Incentive Mechanism Design

GUOQIANG TIAN

( *Texas A&M University* )

**Abstract** This paper introduces some of the basic results, insights, and recent developments in the economic mechanism theory. The theory of mechanism design takes a systematic look at the design of institutions and how these affect the outcomes of interactions. We discuss both subjects of the mechanism theory: realization theory and implementation theory that mainly study the informational requirements and incentives of an economic system, respectively. The two most fundamental contributions of the mechanism theory are (1) the competitive mechanism is the unique economic institution that is both informational and allocative efficient over the class of classical economic environment, and (2) the incentive problem may or may not occur, depending on the kind of game (mechanism) that agents play and the game theoretical solution concept.

**JEL Classification** D71, D83, P51