

机制设计理论及其突破性应用

——2007年诺贝尔经济学奖得主的重大贡献

何光辉 陈俊君 杨咸月

摘要: 机制设计理论及其广泛应用源自于2007年诺贝尔经济学奖三位得主的重大贡献。机制设计理论的关键内容包括激励相容机制、显示原理与执行理论。激励相容机制能对行为人是自利的且拥有私人信息经济体进行严谨分析;显示原理则极大简化了对机制设计问题的研究;占优策略均衡和贝叶斯-纳什均衡是机制设计解决方案的两个主要概念。而执行理论解决了多重均衡问题,能够设计使所有均衡结果对于给定目标函数都是最优的机制。机制设计理论已经在最优拍卖、规制和审计、社会选择等广泛的经济学领域获得了突破性应用,对经济政策和市场制度产生了重大影响。

关键词: 机制设计理论 激励相容机制 显示原理 执行理论

一、引言

经济交易通常发生于市场、公司和其他一系列的制度安排中。一些市场没有政府干预,而另一些则受到管制。在公司之间,一些交易以市场价格为导向,一些由谈判产生,还有一些受到管理控制。由赫维茨(Leonid Hurwicz)开创、马斯金(Eric S. Maskin)和梅尔森(Roger B. Myerson)进一步发展的机制设计理论着重于关注与激励和私人信息相关等问题,为这些不同制度或配置机制提供了一个分析框架。

市场或者类市场的制度,通常能有效地配置商品和服务。很久以前,经济学家就在商品生产和交易性质、参与者信息以及竞争程度等十分严格的假定之下从理论上证明了这种有效性。机制设计理论则允许研究者不太严格的假定下对各种制度进行系统性比较分析。利用博弈论,机制设计理论能够超越经典的方法,如能明确构建定价模型。在一些案例中,博弈论方法已经导致对市场机制进行新的评估。例如,理论表明:当交易者对其交易品价值拥有私人信息时,双重拍卖(买方和卖方报出买价和卖价)可能是有效的交易制度。随着交易者数量的增加,双重拍卖制度将越来越有效地加总私人信息,最终所有信息将反映到均衡价格中(Wilson, 1985)。这些结论支持了哈耶克(1945)“市场有效地加总私人信息”的观点。

机制设计理论表明何种机制对不同参与者(如买方或卖方)是最优的(Samuelson, 1984)。这已经被用于更好地理解我们经常看到的市场机制。例如,该理论被用于确认一些条件,在这些条件下通常观

测到的拍卖形式最大化了卖方的预期收入(Harris and Raviv, 1981; Myerson, 1981; Riley and Samuelson, 1981)。当这些条件不成立时,该理论也能包容最优拍卖形式的细节特征(Myerson, 1981; Maskin and Riley, 1984a)。同样地,机制设计理论能够使经济学家找到垄断价格问题的解决方案,例如,为使得卖方预期收入最大化,价格应当如何取决于数量和质量(Maskin and Riley, 1984b)。

在某些情况下,市场机制不能确保资源完全有效的配置。在此情形下,机制设计理论能用来确认其他更有效率的制度。经典的例子是诸如清新空气、国家安全之类的公共品。Samuelson (1954)推测资源配置机制不能确保完全有效的公共品水平。机制设计理论可以精确地分析萨缪尔逊的推测。更一般地,该理论能够分析可供选择的公共品供给制度的经济效率,包括从市场与通过多数规则实现协商一致的集体决策到独裁。重要之处在于协商的集体决策经常与经济效率不相容。因而该理论有助于证明政府通过征税为公共品融资的合理性。机制设计理论的应用已经在经济学的其他领域取得了突破,包括规制、公司财务和税收理论等领域。

机制设计理论由赫维茨(1960)开创。他把机制定义为一种信息系统,参与者相互或向信息中心传递信息。在这个框架中,市场或类市场的制度可与一系列备择制度相比较。起初,重点关注的是机制的信息和计算成本。Marshak和Radner的团队理论(1972)做出了重要贡献,引发了随后的大量研究(如Groves, 1973)。然而,在很多情况下为参与者提供激励是问题的重要组成部分。在赫维茨(1972)引入

“激励相容”的关键概念之后,机制设计理论才获得了广泛应用。激励相容允许将激励结合进自利参与者的分析中。特别地,它能对行为人是自利的且拥有私人信息经济体进行严谨分析。

20世纪70年代,“显示原理”和“执行理论”的发展使得机制设计理论获得重大进展。显示原理极大简化了对机制设计问题的分析。应用这个原理,研究者在为配置问题寻找最优的可能解决机制时,可以把注意力限制在机制的一个小子集里,即“直接机制”。直接机制并非对现实制度的描述,其数学结构使分析相对简单。对给定配置问题下所有直接机制的集合进行最优化在数学上能够很好地加以定义,一旦一个最优直接机制被找到,研究者就能够把这个机制转化成更符合现实的机制。靠这种貌似迂回的方式,研究者能够解决用其他方式很难处理的制度设计问题。Gibbard(1973)第一个系统阐述了显示原理。一些学者独立地将其扩展成贝叶斯-纳什均衡的一般思想(Dasgupta, Hammond and Maskin, 1979; Harris and Townsend, 1981; Holmstrom, 1977; Myerson, 1979; Rosenthal, 1978)。Myerson(1979, 1982, 1986)在最一般化方面发展了这个原理,并开创性地将之应用于规制(Baron and Myerson, 1982)和拍卖理论(Myerson, 1981)等的重要领域。

显示原理极其有用,但不能解决多重均衡问题:尽管最优结果可能在一个均衡点达到,但其他次优均衡也可能存在。这样,参与者有可能在次优均衡上结束他们的游戏。能否设计这样一个机制:所有均衡都是最优的?对此,马斯金(1977)第一个提出了一般性解决方案。其理论即众所周知的执行理论,构成了现代机制设计的关键部分。

二、机制设计理论的关键概念及其洞察力

这里首先阐述激励相容和显示原理,然后讨论从占优策略均衡和贝叶斯-纳什均衡获得的一些结论,并特别考察了公共品最优供给这一经典配置问题,最后对执行问题进行了论述。

(一) 激励相容和显示原理

赫维茨(1960, 1972)的原创性工作标志着机制设计理论的诞生。在赫维茨的框架中,机制是一个信息系统,参与者相互交换信息,这些信息共同决定了结果。信息可以包含私人信息,如为公共品付费的个人意愿(真实的或假装的)。这个机制就像机器那样编译和处理收到的信息,由此汇集了由许多行为人提供的(真的或假的)私人信息。每个行为人力图最大化其预期报酬(效用或利润),可以决定保留不利信息或者发送虚假信息(如希望为公共品少付费)。这产生了信息博弈均衡“执行”结果的概念:机制定义了信息博弈的“规则”。

对于给定目标函数(如卖方利润或社会福利),为确定最优机制,研究者必须首先描绘可行机制集,

然后明确说明用以预言参与者行为的均衡标准。假设我们关注“直接机制”集,其中行为人报告其私人信息。假定行为人讲的不一定是真话,他们仅在对自已有利的情形下是诚实的。以个人报告为基础,直接机制指派一个结果(如公共品的供给数量和融资费用)。假定采用占优策略均衡概念作为行为准则,则赫维茨(1972)的激励相容概念的表述如下:如果每个参与者真实报告其私人信息是占优策略,那么机制是激励相容的。此外,还要施加一个参与约束:没有人因参与这个机制而使其境况变坏。在一些弱假设下,赫维茨(1972)证明了如下相反的结论:在一个标准的交换经济中,满足参与约束条件的激励相容机制不能产生帕累托最优结果。换言之,私人信息无法实现完全有效性。

源自赫维茨(1972)经典研究的一个自然疑问是:如果考虑到更多类机制或要求不像占优策略均衡那样严格的均衡概念(如纳什均衡或者贝叶斯-纳什均衡),帕累托均衡能否实现?如果不能,我们想知道社会福利的损失有多大、效率的适当标准是什么。更一般地,我们想知道什么样的机制可以最大化一个给定的目标函数,如利润或社会福利(不论这个结果是否完全有效)。在随后的研究中,这些问题得到了回答,而其中大部分成就均归功于显示原理的发现。

显示原理表明:独裁机制的任何均衡结果都可以由某个激励相容的直接机制所复制。在梅尔森(1979, 1982, 1986)所发展的最一般化版本中,显示原理不仅在行为人拥有私人信息、在采取观测不到的行动时(道德风险)是有效的;在机制有多重阶段时也是有效的。尽管所有可能的机制集是庞大的,但显示原理意味着:在由直接机制组成、构造良好的子集中,总能找到最优机制。相应地,大量文献已经关注通过数学推理寻找能最大化目标函数的直接机制,这个目标函数有激励相容约束。

无道德风险情况下的显示原理可大略证明如下。首先,确定任一给定机制的一个均衡。行为人的私人信息被认为具有他自己的“类型”。假设具有 t 类型的行为人在这个均衡中发送信息 $m(t)$ 。考虑与之相关的直接机制,行为人简单地报告了一个类型 t , t 可能是他自己的真实类型或者他类型。报告的类型 t 是直接机制中的信息,其结果定义为与行为人在起始机制的均衡中发送 $m(t)$ 信息的相同。根据假设, t 类型的行为人更喜欢在起始机制中发送信息 $m(t)$ (行为人不可能单边偏向另外信息)。特别地,对任意 $t \neq t'$,行为人更乐意发送信息 $m(t)$ 而不是 $m(t')$ 。因此,在直接机制中他也更乐意报告自己的真实类型而不是虚假地报告任何其他类型 t' 。所以直接机制是激励相容的:行为人没有误报自己类型的动机。这样,直接机制与起始机制产生了相同结果。因而任何(独断的)均衡水平都能够被

激励相容的直接机制所复制。

如下所述,可用显示原理将赫维茨(1972)的不可能性结果推广到贝叶斯-纳什均衡的情形。因此,在参与者拥有私人信息时,经典意义上的帕累托最优一般是得不到的。需要一种将激励考虑在内的新的有效性标准。如果一种直接机制在激励相容约束下最大化了行为人经加权的预期支付,则称它是激励有效的。以这一定义为基础,研究者能够回答源自赫维茨(1972)研究中的许多问题。一个关键性问题是市场机制是否是激励有效的。在局部均衡下,Myerson和Satterthwaite(1983)、Wilson(1985)证明了双重拍卖是激励有效的。Prescott和Townsend(1984)对竞争性的一般均衡是激励有效的信息结构特征进行了刻画。下文我们以公共品为例,讨论与占优策略均衡和贝叶斯-纳什均衡相关的一些结论。

(二) 公共品供给的占优策略机制

如上所述,公共品的最优供给是一个经典问题。当个人对公共品支付的意愿拥有私人信息时,他们有可能试图假装不怎么感兴趣以便减少对供给成本的分摊。这个问题随处可见:如农民应当如何为公共灌溉或排水系统分摊成本;子女应当如何负担对年长父母的照顾?

在1970年之前,经济学家一般认为公共品供给不能达到有效水平,是因为人们在支付上不愿显示真实意愿,所以人们十分惊讶于Clarke(1971)、Groves(1973)提出的如下观点:如果公共品需求没有收入效应(技术上讲,如果效用函数是拟线性的),那么存在一类机制,(a)某人支付意愿的真实显示是一种占优策略;(b)公共物品的均衡水平最大化了社会盈余。在二元决策中(如是否建桥梁),Clarke-Groves机制的最简化表述是这样的:要求每个人报告其支付项目的意愿;当且仅当报告的支付意愿总和超过项目成本,项目才被承建。然后,每个人按照项目成本与其余所有人报告的支付意愿差额支付税款或费用。在这种税制下,每个人“内在化”了整个社会盈余,讲真话是一种占优策略。这种机制的主要缺陷在于总税收收入不等于项目成本:机制一般不满足预算平衡(Green and Laffont, 1979)。集资过多和过少都成问题。例如,参与者共享盈余资金会破坏其讲真话的动机,而浪费盈余资金又导致无效率。

在Clarke和Groves研究的拟线性经济环境之外,占优策略机制并不总能实现。Gibbard(1973)、Satterthwaite(1975)给出了这一效应的结论。他们认为在完全一般的环境中,唯一的占优策略机制是独裁,预先挑选的行为人即独裁者总是得到其中意的可选方案。由于这个及其他一些否定结果,研究焦点由占优策略转向贝叶斯机制设计。

(三) 公共品供给的贝叶斯机制

在贝叶斯模型中,行为人是预期效用最大化者。

典型的概念是贝叶斯-纳什均衡。Dasgupta、Hammond和Maskin(1979)、Myerson(1979)、Harris和Townsend(1981)对一般贝叶斯机制设计问题进行了详细阐述。显示原理被发现之后,贝叶斯机制设计理论的发展主要体现在Myerson(Myerson, 1979, 1981, 1983; Baron and Myerson, 1982; Myerson et al., 1983)的一系列论文中。在这些论文中,可能的配置集是一维的,并且行为人拥有满足单交(Single-crossing)性质的拟线性偏好,这也常见于Mirrlees和Spence的研究之中。梅尔森获取了激励约束的雅致描述方法,相同的机理在随后的应用问题中被广泛使用。

如上所述,公共品供给的Clarke-Groves占优策略机制违背了预算平衡。Claude d'Aspremont和Louis-André Gérard-Varet(1979)认为这一问题可用贝叶斯模型来解决。在占优策略机制中,激励相容约束要求每个行为人通过报告真实情况最大化其效用,而不管其他行为人如何报告。在贝叶斯模型中,行为人是预期效用最大化者,激励相容约束仅需在期望中成立。相应地,激励相容约束在贝叶斯模型中更易得到满足,d'Aspremont和Gérard-Varet比在占优策略下更有可能得到有利结果。实际上d'Aspremont和Gérard-Varet(1979)机制可以看作Clarke-Groves机制在贝叶斯背景下的扩展。

d'Aspremont和Gérard-Varet(1979)机制产生了完全符合帕累托有效的结果,但违背了参与约束。已经观测到其类型但还不曾采取行动的一些个体更希望不参与,所以这个机制仅在强制参与时才可行。如果参与是自愿的,并且项目上马的决定一定要在一致同意时才能做出,那么搭便车的问题变得很严重。利用Myerson(1981)、Mailath和Postlewaite(1990)发展的技术可以证明:为公共项目融资的可能性随着行为人数量的增加将趋向于零。他们给出这样一个例子:尽管每个人都知道如果项目获得融资,其共同处境会改善,但是为公共品融资的渐近概率是零。

这些结论为Samuelson(1954)关于公共品的不利推测提供了严格基础。他们为公共品供给失败提供了似是而非的解释。例如,在诸如公地圈占和沼泽排水之类的公共品决策方面,英国村庄比法国要早得多,这可归因于这样一个事实:法国村庄在这类问题上需要一致同意,而英国的不需要。这至少可以部分地解释1600-1800年间为什么英国农业生产力量会超过法国(Grantham, 1980; Rosenthal, 1992)。

在大量同类模型中,即便没有公共品,经典帕累托有效与自愿参与也是不相容的。在这些模型中,帕累托有效的经典概念通常被相关性更强的激励有效所替代。这个效应的两个基础性的“不可能结果”自愿参与与帕累托有效不相容被Laffont和Maskin(1976)以及Myerson和Satterthwaite(1983)所证明。

(四) 执行理论

激励相容保证了讲真话是一种均衡,但并不能保证它是唯一均衡。许多机制都产生了不同结果的多重均衡。例如,Leininger、Linhart 和 Radner (1989)发现双重拍卖有无限多种(实际上是数不清)非线性均衡,这些均衡的福利从激励有效到 0。显然,多重均衡降低了双重拍卖的吸引力。

Wilson (1979) 分析了可分品的一价拍卖均衡,其中出价者在讨价还价中以非常低的价格拆分商品。在这些“串谋”的均衡中,每一出价者以任何低于其预期均衡份额的价格报价,这种做法阻止了其他出价者获取高于他们(暗中协商的)份额的尝试。这种串谋严重损害了卖方利益。根据 Klemperer (2004) 的观点,这种暗中串谋确实给包括英国电力市场在内的许多现实中的拍卖造成了困扰。

社会选择理论也难摆脱多重均衡问题。那些将从许多候选人中选出一个来的投票者,实际上面临着合作问题。将票投给几乎没机会赢的人意味着“浪费选票”。相应地,如果选举中大家都认为某位候选人没有机会获胜,那这可能成为自我实现的预期。这种现象容易产生多重均衡,其中一些导致了次优结果。

鉴于这些困难,需要设计使所有的均衡结果对于给定目标函数都是最优的机制,这就是众所周知的“执行问题”。Groves 和 Ledyard (1977)、赫维茨和 Schmeidler (1978) 认为,在某些情况下,构建使所有纳什均衡都是帕累托最优的机制是可能的。马斯金 (1977) 则给出了纳什可执行社会选择函数的一般性描述。他认为纳什执行需要一个现在被称为“Maskin 单调”的条件。马斯金 (1977) 也说明:如果 Maskin 单调和被称作“无否决力量”的条件得到满足,若至少有三个行为人,那么纳什均衡的实现是可能的。

马斯金在完全信息博弈中考察纳什均衡,但他的结论已被推广到不完全信息博弈中的贝叶斯-纳什均衡 (Postlewaite and Schmeidler, 1986; Palfrey and Srivastava, 1989; Mookherjee and Reichelstein, 1990; Jackson, 1991)。例如, Palfrey and Srivastava (1991) 表明,对双重拍卖进行修正可使所有均衡是激励有效的。

马斯金的结论还被拓展到其他许多方面,如虚拟(或近似)执行 (Matsushima, 1988; Abreu and Sen, 1991)、“防止重新谈判”均衡中的执行 (Maskin, Moore, 1999)、序列机制路径 (Moore and Repullo, 1988)。执行理论已经并继续在经济学的一些领域发挥重要作用,如社会选择理论 (Moulin, 1994)、不完备契约理论 (Maskin and Tirole, 1999)。

三、机制设计理论的应用

在许多情况下,机制设计使现有研究路线更加现代化和统一化。例如, Vickrey (1961) 已经认识到

广为人知的拍卖“收入等价”原理,机制设计方法则给出了更一般化的收入等价理论。而只有机制设计技术才能确定最普通拍卖的最优形式(属于所有可能的出售机制)。机制设计理论家们还就其他一些情形开发出了一整套新的研究方法。下文仅简单讨论一下机制设计的一些重要应用。

(一) 最优出售和采购机制

拍卖和类拍卖机制是现代经济生活的重要组成部分。梅尔森 (1981) 对最优拍卖所做的开创性研究及随后出现的大量研究 (Krishna, 2002) 有助于经济学家理解现实世界的这些重要制度。在一个典型场景中,经济行为人有一个物品要出售,但是他不知道潜在的买主(出价者)愿意出什么价钱。要想最大化卖方的期望收入,哪一个机制将是最优的? 梅尔森 (1981) 分析了这个问题。应用显示原理,梅尔森研究了出价者报告其支付意愿的激励相容直接机制。作为这些报告的函数,机制明确阐述了谁将以什么价格得到物品。激励相容保证了讲真话是贝叶斯-纳什均衡。因为参与是自愿的,均衡一定也满足参与约束:参与拍卖的每个出价者,至少一定要与他放弃购买时一样好。在这种情景中,梅尔森证明了一般化的收入等价理论。这一理论有一些条件,如风险中性和类型不相关。在该条件下,对(均衡时)物品售予出价最高者的任何拍卖,卖方从中获得相同的预期收入。特别地,四种广为人知的拍卖形式(英式拍卖、荷式拍卖、第一次和第二次报价密封投标拍卖)产生相同的预期收入。梅尔森 (1981) 认为,如果出价者是“对称的”(从同一类型库中抽取)、卖方设定一个合适的“保留价”(最低价),那么所有四种拍卖形式实际上都是最优的。如果出价者的类型是从 0 至 100 区间上的均匀分布中独立抽取,则最优保留价是 50,与出价者数量无关。这个保留价诱使那些超过 50 的出价者报出更高的价格,从而增加了预期收入。另一方面,如果买方认为物品不值 50,那么即便物品对一些买方有正的价值而对卖方一文不值,物品仍不会售出。这个结果从经典意义上说显然不是帕累托有效的。然而,从上述定义上讲,上文提及的拍卖形式是激励有效的。

梅尔森 (1981) 假定卖方的目标是最大化预期收入。但当政府要将一种资产(如无线电波段、公共生产设施)私有化的时候,收入最大化也许不是唯一(或甚至最重要的)激励。更多关注的也许是社会福利最大化:资产应当到最重视其价值的个人或企业手中。

马斯金 (1992) 发现,在某些条件下,即便每个出价者的价值取决于其他出价者的私人信息,英式拍卖也最大化了社会福利。有人可能以下述理由来贬损为最大化社会福利而进行的政府拍卖。假设有两个潜在的买主 A 和 B, B 认可的资产价值比 A 高,那么,即便政府把资产错配给了 A, B 是否简单地从 A

那里把资产买过来(假定资产可交易)?如果是这样,那么最看重这份资产的B最终总能得到资产,所以政府不必为获得初始配置权过分担心。然而,这种论点是不正确的,因为没有考虑到信息约束。Laffont - Maskin 与 Myerson - Satterthwaite 不可能结果意味着,B即便最看重这份资产依然可能不从A处购买。因此,正确获取所有权的初始配置可能极其重要。确实,并没有假定要求初始配置应该给单个人,因为共同所有也许能产生更高的社会福利(Cramton, Gibbons and Klemperer, 1987)。

如果卖方能以固定成本生产额外物品,买方之间的直接竞争并不重要。Mussa 和 Rosen (1978)、Maskin 和 Riley (1984b) 得出了不知道其客户类型(即品味参数)的垄断者的最优销售机制。最优机制包含数量折扣(而非单位固定价格)。Stole (1995) 将理论拓展到寡头情形。其他一些拓展主要涉及了多产品垄断和多维类型(Armstrong, 1996)。从收入转向社会福利最大化,Maskin (1992) 的效率结论没有一般化至多维类型的情形。在这种情形下,Jehiel 和 Moldovanu (2001) 发现一般并不存在激励相容机制,总是能够把物品配置给认为最有价值的人。

(二) 规制和审计

对垄断和寡头的规制是经济学中古老而重要课题。正如 Laffont (1994) 所讨论,早期研究对规制过程给出相当随意的假设。监管者被假定面临某些约束,如垄断者的回报一定要高于市场水平。该水平不是由基本的最优化过程导出,而是简单地特别设定。在这种框架中,对规制过程很难做出合理标准的判断。而 Baron 和 Myerson (1982)、Sappington (1982, 1983) 的原创性贡献以及 Weitzman (1978)、Loeb 和 Magat (1979) 的建设性研究急剧改变了这种状况。在这些论文中,规制过程根据不完全信息博弈来建模。监管者没有直接获得垄断者真实生产成本的信息。使用显示原理,Baron 和 Myerson (1982)、Sappington (1982, 1983) 在没有特别假定的情况下导出了最优规制方案。在最优机制中,监管者(通常是政府机构)在从垄断者那里攫取租金(政府收入)与鼓励有效产出水平这两个目标之间进行权衡。此外,一定要给予垄断者充分的参与激励(即留在市场中)。

规制经济的文献因 Baron 和 Myerson、Sappington 的贡献而大量涌现。他们为诸如价格上限、成本利润分担方案之类的备择规制机制的评估提供了坚实的理论基础。经济学家应用 Baron - Myerson 模型来对公司行为规制效应进行经验估计(Wolak, 1994)。原始的静态模型从各个方向进行了扩展。Freixas、Guesnerie 和 Tirole (1985) 以及 Laffont 和 Tirole (1988) 等分析了信息在随时间推移逐步揭示时的最优时间一致性机制问题,尤其是“棘轮效应”。Laffont 和 Tirole (1987)、McAfee 和 McMillan (1986)、Riordan 和

Sappington (1987) 提出了最优拍卖和最优规制综合的理论。Baron 和 Besanko (1984)、Laffont 和 Tirole (1986) 引入了公司成本事后审计的可能性(概率)。其他很多主题如规制企业、审计人员甚至监管机构之间的串谋已在最优规制文献中进行了广泛分析。

(三) 社会选择理论

在 Arrow (1951) 开创的社会选择理论中,存在可行备择集 X 和对此有偏好的 n 个人。社会选择法则是这样一种规则:对于给定偏好,在个人偏好基础上从 X 中选一个或数个备择方案。Arrow 的开创性研究主要关注社会选择规则如何表达“人民的普遍愿望”这一规范性问题。20 世纪 70 年代,在可选的选举程序下选民策略性行为的问题受到关注。能否设计这样一种机制即一种投票程序,能够诱使选民显示真实偏好? Gibbard (1973)、Satterthwaite (1975) 的不可能结论给出了相反答案。他们认为,如果 X 包含至少三个可选方案,则存在在显示真实偏好是占优策略的机制中能被执行的独裁社会选择。Gibbard - Satterthwaite 定理的证明可直接转变成阿罗著名的规范性社会选择不可能定理的证明(Muller and Satterthwaite, 1985)。这证实了阿罗的“不相关可选方案的独立性”公理与占优策略机制密切相关的推测。因此, Gibbard - Satterthwaite 定理在规范性与策略性分析之间架起了桥梁。其后的研究是放松占优策略要求。相关文献受到了 Maskin (1977) 在纳什执行方面研究的重大影响。关于社会选择理论在策略方面的综述,参见 Moulin (1994)。

该领域的早期研究的洞见在于如果要求社会选择法则是“单一值”(Singleton-valued),即如果一定要选中一个可选方案,那么 Gibbard - Satterthwaite 不可能结论对纳什执行也成立。为理解这个否定结论,试想一下社会选择的纳什执行的一个必要条件是 Maskin 单调条件。该条件表明:如果初始可选方案 $a \in X$ 被社会选择规则选中,且 a 不在任何选民偏好次序等级中,那么 a 一定仍然被选中。

为阐明该条件的力度,考虑一个特定的社会选择法则即多数法则。如果 X 中的一种可选方案被最大多数选民置于最高等级,则称该方案为“多数可选方案”。多数法则简单地规定多数可选方案应该总被选中。现假设有七位选民, X 包含三种可选方案 a, b, c 。假定选民的偏好次序是这样的:选民 1、2、3 认为 a 是最好的可选方案, b 是次优的, c 最差, $a > b > c$; 选民 4、5 认为 b 是最好方案, a 次优, c 最差, $b > a > c$; 最后,选民 6 和 7 认为 c 是最好的, b 次优, a 最差, $c > b > a$ 。显然 a 是多数可选方案,因为三位选民将其置于最高等级中,而 b 和 c 分别只有两位选民将其置于最高等级中。现假设选民 6 和 7 改变了主意,认为无论如何可选方案 c 比 a 和 b 要差。在其新的偏好次序中, b 上升到第一位, a 上升到第二位: $b > a > c$ 。如果其余选民的等级如前不

变,则方案 a 在任何人的偏好次序等级中均未下降,但不再是多数可选方案,因为现在有四位选民把 b 置于最高等级,但仍然只有三位选民把 a 置于最高等级。因此,多数法则不是 Maskin 单调的。

根据马斯金(1977)定理,不存在“纳什执行”多数法则的决策机制。精确地说,只要有至少三个可选方案,多数法则就不是 Maskin 单调的;相反,很容易证明如果只有两种可选方案,多数法则是 Maskin 单调的。类似地,文献中提出的其他社会选择法则,如著名的 Borda 法则(J. C. de Borda 于 1781 年提出)也不满足 Maskin 单调性。更一般地, Muller 和 Satterthwaite(1977)认为无“单一值”的社会选择法则可能是 Maskin 单调的。这意味着选民的策略性行为将导致任何可想象的投票机制将产生纳什均衡,根据给定的社会选择法则都是次优的。

摆脱两难的一个途径是去掉社会选择法则是单一值的要求。许多多值社会选择法则(如总是选择所有“帕累托有效”方案)是 Maskin 单调的,能够被纳什执行。但缺点是我们被迫接受一个基础不确定的状态:在一些偏好下,应该被社会接受的方案多于一个。因此,不止一个纳什均衡存在着。这种不确定性可能是非独裁制度中不可避免的方面。最终结果可能取决于选民之间的谈判和讨价还价。在 Thomas Schelling 的术语中,选民可能在“焦点”均衡下进行调整,这种均衡在给定文化背景、历史或其他社会和心理因素的情况下自然出现。获得更有利结果的另一种途径是假定纳什均衡的精致设计能够把握选民的行为,如颤手完美纳什均衡(Selten,1975)。

参考文献:

1. Abreu, D. and Sen, A., 1991. "Virtual Implementation in Nash Equilibria." *Econometrica* 59, pp. 997 - 1 022.
2. Armstrong, M. 1996. "Multiproduct Nonlinear Pricing." *Econometrica* 64, pp. 51 - 75.
3. Arrow, K., 1951. *Social Choice and Individual Values*. Wiley, New York.
4. Arrow, K. J., 1979. "The Property Rights Doctrine and Demand Revelation under Incomplete Information." in M. Boskin (ed.), *Economics and Human Welfare*. Academic Press, New York.
5. Baron, D. and Besanko, D., 1984. "Regulation, Asymmetric Information, and Auditing." *Rand Journal of Economics* 15, pp. 447 - 470.
6. Baron, D. and Myerson, R., 1982. "Regulating A Monopolist With Unknown Costs." *Econometrica* 50, pp. 911 - 930.
7. Dasgupta, P.; Hammond, P. and Maskin, E., 1979. "The Implementation Of Social Choice Rules: Some General Results on Incentive Compatibility." *Review of Economic Studies* 46, pp. 181 - 216.
8. d'Aspremont, C. and Gérard - Varet, L. A., 1979. "Incentives and Incomplete Information." *Journal of Public Economics* 11, pp. 25 - 45.
9. Freixas, X.; Guesnerie, R. and Tirole, J., 1985. "Planning under Incomplete Information and the Ratchet Effect." *Review of Economic Studies* 52, pp. 173 - 192.

10. Gibbard, A., 1973. "Manipulation of Voting Schemes: A General Result." *Econometrica* 41, pp. 587 - 602.
11. Green, J. and Laffont, J. J., 1979. *Incentives in Public Decision Making*. North - Holland, Amsterdam.
12. Groves, T., 1973. "Incentives in Teams." *Econometrica* 41, pp. 617 - 663.
13. Grantham, G., 1980. "The Persistence of Open - Field Farming in Nineteenth Century France." *Journal of Economic History* 40, pp. 515 - 531.
14. Groves, T. and Ledyard, J., 1977. "Optimal Allocation of Public Goods: A Solution to the 'Free Rider' Dilemma." *Econometrica* 45, pp. 783 - 811.
15. Harris, M. and Raviv, A., 1981. "Allocation Mechanisms and the Design of Auctions." *Econometrica* 49, pp. 1 477 - 1 499.
16. Harris, M. and Townsend, R., 1981. "Resource Allocation under Asymmetric Information." *Econometrica* 49, pp. 33 - 64.
17. Hayek, F., 1945. "The Use of Knowledge in Society." *American Economic Review* 35, pp. 519 - 530.
18. Holmstrom, B., 1977. "On incentives and Control in Organizations." Ph D dissertation, Stanford University.
19. Hurwicz, L., 1960. "Optimality and Informational Efficiency in Resource Allocation Processes." in Arrow, Karlin and Suppes (eds.), *Mathematical Methods in the Social Sciences*. Stanford University Press.
20. Hurwicz, L., 1972. "On Informationally Decentralized Systems." in Radner and McGuire, *Decision and Organization*. North - Holland, Amsterdam.
21. Hurwicz, L., 1973. "The Design Of Mechanisms For Resource Allocation." *American Economic Review* 63, *Papers and Proceedings*, pp. 1 - 30.
22. Hurwicz, L. and Schmeidler, D., 1978. "Construction of Outcome Functions Guaranteeing Existence and Pareto - Optimality of Nash Equilibria." *Econometrica* 46, pp. 1 447 - 1 474.
23. Jackson, M., 1991. "Bayesian implementation." *Econometrica* 59, pp. 461 - 477.
24. Jehiel, P. and Moldovanu, B., 2001. "Efficient Design with Interdependent Valuations." *Econometrica* 69, pp. 1 237 - 1 259.
25. Klemperer, P., 2004. *Auctions: Theory and Practice*. Princeton University Press, Princeton.
26. Krishna, V., 2002. *Auction Theory*. Academic Press, New York.
27. Laffont, J. - J. and Maskin, E. 1979. "A Differentiable Approach to Expected Utility - Maximizing Mechanisms." in Laffont (ed.), *Aggregation and Revelation of Preferences*. North - Holland, Amsterdam.
28. Laffont, J. - J. and Tirole, J., 1986. "Using Cost Observation to Regulate Firms." *Journal of Political Economy* 94, pp. 614 - 641.
29. Laffont, J. - J. and Tirole, J. 1988. "The Dynamics of Incentive Contracts." *Econometrica* 56, pp. 1 153 - 1 175.
30. Laffont, J. - J. and Tirole, J., 1987. "Auctioning Incentive Contracts." *Journal of Political Economy* 95, pp. 921 - 937.
31. Laffont, J. - J. and Tirole, J., 1993. *A Theory of Incentives in Procurement and Regulation*. MIT Press, Cambridge.

(以下文献略)

(作者单位:复旦大学经济学院 上海 200433
上海社科院部门经济研究所 上海 200020)
(责任编辑:刘成奎)