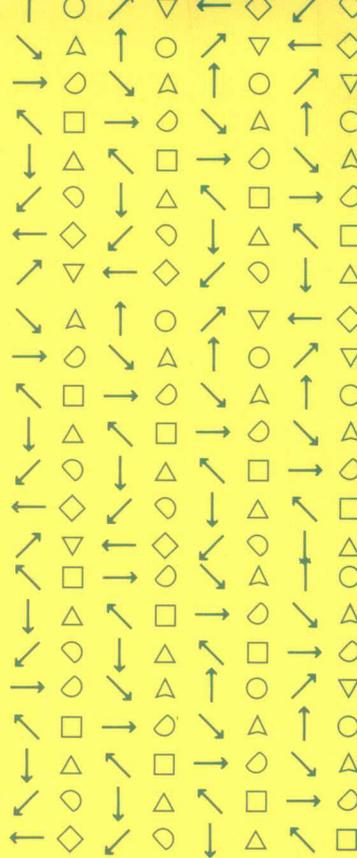


当代经济学系列丛书

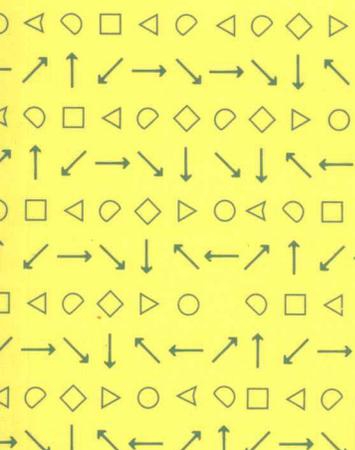
Contemporary Economics Series

主编 陈昕

国家“十三五”重点图书



# 机制设计理论



An Introduction to  
the Theory of  
Mechanism Design

当代经济学  
教学参考书系

[德] Tilman Börgers 提尔曼·伯格 著  
李娜 译



格致出版社  
上海三联书店  
上海人民出版社

当代经济学系列丛书

Contemporary Economics Series

主编 陈昕

# 机制设计理论

[德] 提尔曼·伯格 著  
李娜 译



当代经济学系  
教学参考书



格致出版

格致出版社  
上海三联书店  
上海人民出版社

## 图书在版编目(CIP)数据

机制设计理论/(德)提尔曼·伯格斯特著;李娜译.

—上海:格致出版社;上海人民出版社,2018.9

(当代经济学系列丛书;当代经济学教学参考书系)

ISBN 978-7-5432-2681-4

I. ①机… II. ①提… ②李… III. ①机制设计理论

IV. ①F062.5

中国版本图书馆 CIP 数据核字(2018)第 129710 号

责任编辑 唐彬源

封面设计 敬人设计工作室 吕敬人

## 机制设计理论

[德]提尔曼·伯格斯特 著

李娜 译

出版 格致出版社  
上海三联书店  
上海人民出版社  
(200001 上海福建中路 193 号)

发行 上海人民出版社发行中心

印刷 苏州望电印刷有限公司

开本 787×1092 1/16

印张 13.5

插页 3

字数 269,000

版次 2018 年 9 月第 1 版

印次 2018 年 9 月第 1 次印刷

ISBN 978-7-5432-2681-4/F·1118

定价 65.00 元

# 中文版序

## 主编的话

上世纪80年代,为了全面地、系统地反映当代经济学的全貌及其进程,总结与挖掘当代经济学已有的和潜在的成果,展示当代经济学新的发展方向,我们决定出版“当代经济学系列丛书”。

“当代经济学系列丛书”是大型的、高层次的、综合性的经济学术理论丛书。它包括三个子系列:(1)当代经济学文库;(2)当代经济学译库;(3)当代经济学教学参考书系。本丛书在学科领域方面,不仅着眼于各传统经济学科的新成果,更注重经济学前沿学科、边缘学科和综合学科的新成就;在选题的采择上,广泛联系海内外学者,努力开掘学术功力深厚、思想新颖独到、作品水平拔尖的著作。“文库”力求达到中国经济学界当前的最高水平;“译库”翻译当代经济学的名人名著;“教学参考书系”主要出版国内外著名高等院校最新的经济学通用教材。

20多年过去了,本丛书先后出版了200多种著作,在很大程度上推动了中国经济学的现代化和国际标准化。这主要体现在两个方面:一是从研究范围、研究内容、研究方法、分析技术等方面完成了中国经济学从传统向现代的转轨;二是培养了整整一代青年经济学人,如今他们大都成长为中国第一线的经济学家,活跃在国内外的学术舞台上。

为了进一步推动中国经济学的发展,我们将继续引进翻译出版国际上经济学的最新研究成果,加强中国经济学家与世界各国经济学家之间的交流;同时,我们更鼓励中国经济学家创建自己的理论体系,在自主的理论框架内消化和吸收世界上最优秀的理论成果,并把它放到中国经济改革发展的实践中进行筛选和检验,进而寻找属于中国的又面向未来世界的经济制度和经济理论,使中国经济学真正立足于世界经济之林。

我们渴望经济学家支持我们的追求;我们和经济学家一起瞻望中国经济学的未来。

厉声

2014年1月

# 中文版序

很高兴得知本书将以中文形式与中国的读者见面。本书所阐述的机制设计理论几十年来一直是研究所关注的焦点。但是，我们还远未获知我们想知道的一切，实际上甚至连当前构建机制设计理论的框架是否就是要研究的最优框架也并不清楚。希望本中文版能鼓励许多有天赋的人加入机制设计理论这一领域的研究，提出新的原创性的想法。

机制设计理论基于经济学中一个由来已久的观念：只要我们把许多各自只拥有部分信息的人的信息而非所有的相关信息汇总起来，就可以有效地作出关键的经济决策（生产什么、如何生产、提供给谁等）。比如，绝大部分人只知道对于其自身或其亲友而言，不同行当的工作难不难以及有多难，或者有多想得到某种消费品，却不知道这些信息对于许多其他人又是怎样的。但为了决定应如何最佳地利用社会资源，我们需要以某种方式把所有这样的私人信息汇总起来。显然，对于社会顺利发展很重要的因素就在于能多有效地解决这一收集并整合分散信息的问题。

我们可以把现行的政治制度和经济制度视为借以形成如实披露信息的激励手段。最显而易见的是，市场机制可被看作一种收集并加工由价格汇总和传递的信息的制度。这是一种对于市场本质很通常的看法，哈耶克已经对之进行了广为人知的清晰阐述。对于政治制度而言，民主社会中的选举也可被视作一种收集并加工信息——有关选举人对于各种公共政策的偏好的信息——的制度。

机制设计理论所构建的数学模型允许我们把“哪

种制度是收集并整合信息的最佳制度”这一问题转变为定义良好的数学问题。通常而言,我们侧重的是与上述经济决策、政治制度和经济制度相比机制设计理论不那么宏大的应用,比如说拍卖。但在机制设计理论的具体应用和宏大应用中,数学模型的基本结构是一样的。我们应用机制设计理论所研究的这些问题很重要,可喜的是,对这些问题可以进行数学分析。

我非常希望读者看完本书后不会有机制设计理论领域已经“大功告成”的看法,也希望读者不会得出只有那些掌握了相当高等的数学技巧的人可以对机制设计理论的进一步发展有所贡献的定论。在本书各章节,我都尽量说明即使是机制设计理论中一些非常基本的概念性问题也是开放性的问题,其中有不少需要对如何构造模型提出新想法,而非对现有模型有更深的数学理解。我希望本书能让读者感受到鼓舞而在机制设计理论领域进行自主的研究。

提尔曼·伯格斯特(Tilman Börgers)

2018年8月

# 序 言

本教材的首要目标,是给出机制设计理论中经典结论的严密而易懂的解释。第二个目标是将相关领域的读者带到研究前沿。与对经典结论的描述相比,对最新研究领域的介绍难免多了一些主观性。读者可以去阅读机制设计理论的近期文献,以了解那些不同于我们的研究视角。

本教材的读者对象是有一定博弈论基础的经济学专业高年级本科生和研究生。Fudenberg 和 Tirole (1993)的内容对于本书而言绰绰有余。本书也需要读者具备实分析的基本知识。这方面可以参阅 Rudin (1976)。

本书源自 2006 年冬季学期我在密歇根大学所讲授的一门课程的讲义。当时就我所知,还没有一本关于机制设计的书存在。当我正费力地着手将讲义写成一本书时,有几本内容类似的优质图书相继出现了。每当发现这样的书,为了保持写作动力,我就需要说服自己:其他作者的书与本书之间存在重大差异。下面我按时间顺序把这些好书列出来,同时列出本书的存在理由。尽管这些书在某种意义上是本书的竞争者,我还是将其列出来。这是为了服务于我们这些同类书的初衷:帮助读者获得对于机制设计领域的完整理解。希望下面这些书和本书一起会对有志于此的读者们大有裨益。

• Leonid Hurwicz and Stanley Reiter, *Designing Economic Mechanisms*, Cambridge: Cambridge University Press, 2006.

机制设计理论是由 Leonid Hurwicz 与 Eric Maskin

和 Roger Myerson 一起创建的,他们因此而获得 2007 年的诺贝尔经济学奖。上面这本书重点讨论了机制中的信息效率与隐私保护,但和我这本书相比,较少涉及激励问题。

• Steven R. Williams, *Communication in Mechanism Design: A Different Approach*, Cambridge: Cambridge University Press, 2008.

该书涵盖的内容与 Hurwicz 和 Reiter 的书类似。这两本书都把重点放在机制的信息空间(message space)的大小上,从而不同于机制设计的现代处理方式。但是,正如本书中特别是第 10 章中所讨论到的,机制的复杂性或简单性——其中一个方面就是信息空间的大小——仍然很重要,而且可能是未来研究的核心问题。

• Dmitrios Diamantaras, Emina I. Cardamone, Karen A. Campbell, Scott Deacle, and Lisa A. Delgado, *A Toolbox for Economic Design*, New York: Palgrave MacMillan, 2009.

在这里列出的所有书中,该书最接近本书,只是涵盖的内容更多一些,包括纳什实施理论、匹配市场理论以及有关机制的实证证据等。有时候我真希望该书是我写的。本书关注的领域则要更窄一些,也许更加深入,并且更加强调博弈论基础与机制设计之间的关系。

• Rakesh Vohra, *Mechanism Design, A Linear Programming Approach*, Cambridge: Cambridge University Press, 2011.

这是一本极好的书,展示了机制设计理论的大部分内容是如何运用线性规划理论而发展出来的,但和本书相比,较少强调机制设计的博弈论基础。

我对上面各书的评论已经表明了本教材的一些局限,下面还有两点需要强调。首先,我没有介绍纳什实施理论。在本段接下来的内容中,我会解释这样做有什么特别含义。但需要指出的是,读者也许只有在看完本书后才能理解本段的内容。根据我的理解,在有关机制设计的文献中,允许机制设计者从其提出的机制的各个均衡中选择其一。而在有关机制的文献中,则要求机制设计者考虑其提出的机制的所有均衡。有关纳什实施理论的文献只是有关机制的文献的一部分。纳什实施理论这一术语没有被普遍运用,但在本书中会很有用。请注意,只要提到显示原理(revelation principle),就是在讨论机制设计理论而非纳什实施理论。如前所述, Diamantaras 等人的著作中包含了有关纳什实施理论的章节。特别有用的文献还包括:

• Matthew O. Jackson, “A Crash Course in Implementation Theory”, *Social Choice and Welfare*, 18, 655—708, 2001.

本书没有介绍纳什实施理论,并非因为它不重要,而是因为其他学者已经很好地对它进行了解释,此外,它需要的许多方法和论证与本书大不相同。尽管如此,我依然为略去了这部分内容而感到遗憾。<sup>①</sup>

<sup>①</sup> 请注意本书中使用了“实施”这个动词。但当用到这个词时,所指的并不是纳什实施理论意义上的实施。但愿“实施”这个动词的用意在本书内容的展开中会变得清楚起来。

本书也没有介绍计算机科学家对机制设计理论的贡献。对这些贡献的综述参见下面这本书的第9—16章：

• Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani, *Algorithmic Game Theory*, Cambridge: Cambridge University Press, 2007.

本书没有提及这部分内容，因为我对该领域不是足够了解。目前我正在以平常的蜗牛速度阅读上述的综述章节。但与此同时，计算机科学家所写的有关机制设计的文章的数量似乎在以惊人的速度增长。寻求如何跟上这些发展以及如何将我所了解的经济学家的研究与计算机科学家的研究融合在一起是一项庞大的任务。因此，要在本书中完成这一任务是不现实的。

读者可能觉得令人着迷的一个很大的主题是实践中的机制设计问题。但这一主题需要单独一本书或者几本书来介绍。读者可以参阅下面这些有用的综述：

• Alvin E. Roth, “What Have We Learned from Market Design?” *Economic Journal*, 118, 285—310, 2008.

• Peter Cramton, “Spectrum Auction Design”, *Review of Industrial Organization*, 42, 161—190, 2013.

与这些研究相反，本书的重点放在机制设计理论的方法论上。至此我也许已经失去了一些更加注重实践的读者，但现在也正是那些对本书感兴趣的读者继续看下去的时候。<sup>①</sup>

提尔曼·伯格斯(Tilman Börgers)  
2014年8月于安娜堡(Ann Arbor)

① 如果你在本书中发现了任何一种错误，请发邮件至 [tborgers@umich.edu](mailto:tborgers@umich.edu)。本书的勘误表会在 <http://www-personal.umich.edu/~tborgers/>上列出来。

# 致 谢

本教材阐述了我所理解的机制设计理论的内容。我的理解大部分来自其他学者的书和文章，同时也来自认真听我课并指出我的错误的学生以及跟我一起做研究的合作者们。

我特别缺乏全面了解的一个领域是动态机制设计理论。Daniel Krähmer 和 Roland Strausz 慷慨地同意就这一主题替我写一章，即第 11 章。

Christoph Kuzmics 把他用本书内容上课时给他的学生们出的一些题目给我了。我把受这些题目启发而生成的一部分习题也纳入了本书。

我还要感谢牛津大学出版社的诸位匿名审稿人以及 Stefan Behringer, George Chen, Shaowei Ke, Christoph Kuzmics, Xian Li, Vitor Farinha Luz, Stephen Morris, Colin von Negenborn, Martin Pollrich, Arunava Sen, Xianwen Shi, Jan-Henrik Steg, Roland Strausz 和 Tobias Widmer 向我提出的非常有帮助的校正、评论和建议。

Trevor Burnham, Yan-Min Choo 和 Nikoleta Scekcic 校对了本书的各个部分。他们发现了许多令人尴尬的小错误，对此我感激不尽。\*

我从 Vijay Krishna (2002) 与 Paul Milgrom (2004) 关于拍卖理论的书获益良多。本教材第 2 章和第 3 章中的阐述就很大部分归功于这两本书。在本

\* 翻译期间成书后的修订内容见 <http://www-personal.umich.edu/~tborgers/Corrections.pdf>。译成中文版时已将英文原书中的错误内容直接替换为相应的修订内容；对于未包含在这些修订内容中的内容上的错误，翻译时也相应作了修改并用注释加以注明。——译者注

教材正式提到机制设计之前就在第 2 章中介绍甄别问题的想法,则来自与 Mark Armstrong 的一次谈话。

我感谢 Daniel Krähmer 和 Roland Strausz 同意将所有作者得自本书的收入都捐给国际特赦组织(Amnesty International)。即使你看完本书后感到浪费了你的时间,没看完本书也会确保你没有彻底浪费掉你的钱,因为其中的一部分助力了一项善事。

提尔曼·伯格斯特(Tilman Börgers)  
2014 年 8 月于安娜堡(Ann Arbor)



076	5	激励相容性	
076	5.1	导言	
076	5.2	设定	
077	5.3	弱单调性	
079	5.4	周期单调性	
082	5.5	结果为彩票时的周期单调性	
083	5.6	一维类型空间	
087	5.7	丰富的类型空间	
087	5.8	收入等价	
088	5.9	个人理性	
089	5.10	文献评论	
089		习题	
091	6	贝叶斯机制设计	
091	6.1	导言	
091	6.2	设定	
093	6.3	类型独立	
095	6.4	类型相关	
103	6.5	文献评论	
104		习题	
105	7	占优策略机制设计	
105	7.1	导言	
105	7.2	占优策略激励相容性	
107	7.3	实施帕累托有效的决策规则	
108	7.4	差异正关联	
110	7.5	个人理性和预算平衡	
113	7.6	文献评论	
113		习题	
114	8	不可转移效用	
114	8.1	导言	
115	8.2	Gibbard-Satterthwaite 定理	
121	8.3	机制在受限制定义域上的占优策略激励相容性	
123	8.4	贝叶斯激励相容性	
124	8.5	文献评论	
125		习题	

126	9 信息互依
126	9.1 引言
127	9.2 一个例子
128	9.3 实施福利最大化决策规则的不可能性
131	9.4 刻画所有的激励相容机制
133	9.5 文献评论
133	习题
134	10 稳健机制设计
134	10.1 引言
135	10.2 一个例子
139	10.3 为不完全信息建模
144	10.4 机制设计者的不确定性
146	10.5 机制
146	10.6 贝叶斯纳什均衡与显示原理
149	10.7 可以实施什么机制
155	10.8 具有共同先验信念的稳健机制设计
156	10.9 没有共同先验信念的稳健机制设计
158	10.10 拟线性偏好情形中的概念问题
160	10.11 再论投票
165	10.12 文献评论
167	习题
168	11 动态机制设计
168	11.1 引言
168	11.2 动态私人信息
186	11.3 动态配置
189	11.4 文献评论
190	习题
192	参考文献

# ▶ 1 导 论

假如你想把你的房子卖了,并且你的房地产经纪人已经确定了几个愿意支付你的要价的潜在买者,那么你也许想在这些买者之间进行一场拍卖以获得更高的卖价。有许多不同的拍卖形式可以供你采用。比如,可以要求每个买者递交一个有约束力的最后出价。或者,众买者可以分几轮报价,每一轮他们都会被告知上一轮的最高出价,然后被要求调整其这一轮的出价。你也可以采取这些拍卖形式的某种组合。你应该如何在这些不同拍卖形式之间作选择呢?这正是机制设计理论旨在回答的问题之一。

现在想象你和你的同事们正在考虑是否要买一台新的冰箱放在办公室。你们可以把从家里带来的食物存放在该冰箱里面。尽管每个人都赞成,但并不清楚该冰箱对于不同人的价值具体是多少。你怎么能知道每个人最多愿意出的价格之和是否会超出这台冰箱的成本呢?你可以要求每个人同时递交认购款,然后看看认购款的总额是否超出了成本。此外,你还可以告诉每个同事其他人目前的认购款是多少。或者,你可以将这台冰箱的成本除以认购的人数,并承诺只要每个人愿意付出其份额就去买冰箱。这些做法中哪一种最好呢?这又是机制设计理论要阐述的一个问题。

在上述两个例子中,你可能考虑的每一种做法都在参与者之间产生了非合作博弈论意义上的一个策略博弈。这些做法的参与者知道,结果将不仅取决于他们自己的选择,而且还取决于其他人的选择,因此,他们自己的最优策略(optimal strategy)也许取决于其他人的策略。换言之,这些做法的参与者明白他们在进行非合作博弈。机制设计理论因此建立在博弈论的基础上(Fudenberg and Tirole, 1993)。博弈论视博弈规则为给定的,并对策略博弈者的行为作出预测。机制设计理论则是关于博弈规则的最优选择(optimal choice)。

实际上,博弈规则的设计无处不在。股东的投票该如何进行?公司内部的晋升机制应如何设计?最优婚前协议应如何拟定?所有这些问题都与最优博弈规则有关。机制设计理论试图研究所有这些应用背后的一般结构,但也详细考察了许多特别重要的应用。

研究机制设计至少在两个层面有意义。首先,机制设计理论在实践中有助于

真实世界里的机制的设计者。例如,在有关政府拍卖和产业拍卖的设计的讨论中往往会用到最优拍卖理论。可以把机制设计理论的第一个层面称为机制设计的“规范”(normative)层面。其次,可以把真实世界里的制度理解成其设计者的理性选择,以此来解释这些制度为什么是这样的。例如,我们也许会试图借助机制设计理论来解释拍卖在一些房屋销售中的作用以及标价在另一些房屋销售中的作用,这表明,标价在有些情况下是最优的,而拍卖在另一些情况下是最优的。可以把机制设计理论的第二个层面称为机制设计的“实证”(positive)层面。

对博弈规则的选择所产生的激励是机制设计理论的关键所在。激励也是合约理论(Bolton and Dewatripont, 2005)的关键所在。乍看之下,机制设计理论与合约理论之间的区别很简单:在合约理论中,研究的是单个行为人的最优激励设计;在机制设计中,研究的是一群行为人的最优激励设计,比如前述第一个例子中的众买房者和第二个例子中的诸位同事。因此,合约理论与机制设计理论不同之处在于,前者不必处理策略互动问题。

然而,合约理论与机制设计理论之间的关系更加微妙。事实上,机制设计理论的一部分是对合约理论的洞见的直接扩展。这有些出人意料,因为人们原本料想机制设计理论中有而合约理论中无的策略互动元素会产生大量的新问题。理解为什么事实不是这样不仅很有趣,而且在概念上很重要。下面将详加阐释。合约理论与机制设计理论之间的这种密切对应关系只存在于机制设计理论的某些部分。机制设计理论的其他部分则与合约理论无关。

传统上,合约理论被分为两部分:隐藏信息理论(也被称为“逆向选择”理论)和隐藏行动理论(也被称为“道德风险”理论)。以医疗保险合约为例,很容易解释这两部分之间的区别。你知道你过去是否经历过严重的胸痛,但你打算购买的医疗保险的所属公司不知道这事。这就是“隐藏信息”。一旦你为心脏手术购买了全覆盖的保险,定期锻炼抑或有所松懈是你所作的选择。保险公司对此并不知情,除非安排了监视你。这就是“隐藏行动”。隐藏的无论是信息还是行动,对于合约设计都很重要。比如,保险公司通过给你提供保险合约的清单并观察你在这一清单中的选择,也许就能推断你可能想向该保险公司隐瞒你的健康风险的相关信息。通过引入免赔额,保险公司可以试图维持你关注自己健康的激励,从而减轻道德风险问题。

和传统上所理解的一样,机制设计与多个行为人的隐藏信息而非隐藏行动有关。在前述第一个例子中,卖房者试图找出的隐藏信息是买房者的真实支付意愿。在前述第二个例子中,我们试图找出的隐藏信息是同事们对办公室冰箱的真实支付意愿。在投票中,我们要找出的隐藏信息是个人对不同选项或不同候选人的真实排序。当然,涉及众多行为人的隐藏行动也是一个非常有趣的领域,而且处理隐藏行动的理论就同机制设计理论一样,关注的是博弈规则的最优选择。例如,一家公司内部的各种晋升机制为一群雇员确定了工作激励。对这些晋升机制的最优选择就是经济理论的一个重要研究领域。不过,这并非传统上所理解的机制设计的

研究领域。

当选择多个行为人之间的策略互动规则时,我们也许会将自己限定于所有可以想到的规则的一个小的子集中,也或许会尽力将网撒大一些,考虑各种规则的一个尽可能大的集合。例如,你在考虑如何拍卖掉你的房子时,也许会只考虑对必需的最低出价的选择,并理所当然地认为拍卖会以全部潜在买房者同时递交出价的方式进行,那么,你可以在可行拍卖方式的一个大得多的集合中只考虑仅有一个参数的拍卖方式。抑或是,除了拍卖以外,你也许还会考虑与全部买房者按照某种预定形式同时谈判等所有可以想到的处理方式。机制设计理论的成果之一,就是发展出一个人可用于在所有可以想到的规则之中找到最优博弈规则的框架。事实上,机制设计传统上被理解成在所有可以想到的规则之中进行全局优化(grand optimization)的领域。本书中将坚持这一理解。

假定你考虑了用于处理你的房屋销售问题的所有可能规则,并且你得出如下结论:仅有一轮报价的拍卖是最优的。在最高出价已被揭示之后,报价失败的买房者中有一位接洽你,并提出一个高于拍卖中获胜出价的新出价。你会接受吗?这是一个明显的诱惑,但如果你接受了后来的新出价,你还愿意进行仅有一轮报价的拍卖吗?本书假设机制设计者具有充分承诺能力。规则一旦被宣布就一锤定音了,即机制设计者不会背离最初的承诺。在我们的例子中,拍卖结果已被揭示之后,机制设计者绝对会拒绝再谈判。这显然是一个强假设。在合约理论中,大量注意力被放在不能总是获得充分承诺时的最优合约设计上,而且在这条研究路径上的研究成果颇丰。在机制设计中,这也可能是一条有趣的研究路径。但我在本书中不会考虑这条研究路径,因为我想继续集中讨论传统机制设计理论的核心观点,而这些观点假设机制设计者作出了充分承诺。

我为本书选择了一个有些不同寻常的开头,即在第2章中介绍了基本的甄别理论。甄别理论有时不被看作机制设计理论的一部分,因为它仅为单个行为人而非多个互动的行为人构建激励机制。不过,第2章介绍的甄别理论与后面章节特别是第3章所阐述的最优机制设计理论紧密相关。我希望通过将甄别理论与机制设计理论并列对照,能有助于读者理解最优机制设计的哪些特征来自策略互动,以及最优机制设计的哪些特征等同于最优甄别的相应特征。

接下来,第3章借助一些著名例子回顾了经典的“贝叶斯”机制设计理论。第5章运用了非常大的一类模型,对第2章的单个行为人理论进行了推广。第6章接着以类似的方式将“贝叶斯”机制设计理论发展为一个普遍适用的框架。

第3章和第6章所采用的经典博弈论模型建立在几个限制性假设之上,我对之进行了详细讨论。其中一个假设是,对于某一机制的各给定规则,在对各行为人关于彼此的私人信息的信念的特定描述下,各行为人采用的是该机制的贝叶斯纳什均衡策略。与现实相比,这一假设也许会为机制设计者提供更多的信息。因此,研究文献已经试图发展出不需要机制设计者有多了解各行为人的信念的机制。其经典方法就是寻找占优策略机制。第4章用几个例子给出了该方法,第7章则给