曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

译者序

2011 年，因教学需要我翻译了范里安（Varian）《中级微观经济学》第七版，后又翻译了第八版。这些版本经网络流出后，好评如潮（实事求是），这出乎我意料之外。我得到的反馈来自大陆 100 多所优秀大学（几乎包含了所有 211 大学）的师生。台湾、香港部分高校的师生以及海外留学生也有响应。

然而，必须指出，在以前的翻译版本中充斥着不少笔误，给读者造成了不必要的困扰，致歉。

此次翻译版本（**2014 版**）针对得正是我以前翻译的此书第八版，主要修正了符号和公式中的笔误。由于原来版本笔误过多，请读者果断放弃原来版本（第七版和第八版），改为使用 **2014 版**。

我相信 **2014 版**在质量上有质的提升，但仍不可能将笔误（尤其是文字上的笔误）降低为零。欢迎读者指出（这不是客套话），我的邮箱为 caoqianseu@163.com。

曹乾

东南大学 江苏南京

2013 年 9 月

译者原序

一切都应尽量简单，但不要太简单。

——爱因斯坦

对于大多数经管类专业的大学生来说，如果想深入学习微观经济学，那么范里安的《中级微观经济学：现代方法》（国内又译作《微观经济学：现代观点》）几乎是一本绕不过去的优秀教科书，你必须去读。这本书除了一些内容和数学工具没有涉及到外，在涉及到的内容中，基本已把中初级微观经济学写到了极致，市场中能超越这本书几乎没有。

2010 年，我在奥克兰大学经济系做访问学者期间，由于业余时间较多，仓促翻译了此书的第七版。翻译这本书的初衷是不忍心我的学生遭受市场上劣质教科书的折磨，当然也出于方便教学的目的。我的学生曾将零散教学资料上传到网络，得到了全国各地师生的好评，甚至在国外的留学生也来信索要全本。这其实在我的意料之中，原因就不多说了。

在第八版的翻译中，我改正了第七版中明显的笔误。我相信里面仍有笔误，但基本没有理解上的错误，所以称为“完美翻译版”。范里安教科书的所有内容的版权均属于范里安。但课后答案书的版权属于本人。仅供教学和学习参考。请勿用作商业用途。

曹乾　caoqianseu@163.com

东南大学

江苏南京

2012 年 2 月

第八版新增内容

第 10 章：跨期选择

- 新增案例：延长版权保护期

第 11 章：资产市场

- 新增一节：11.6 节资产泡沫。

第 13 章：风险资产

- 新增一节：13.3 节合同对方不履行合约责任的风险
- 新增案例：在险价值

第 17 章：拍卖

- 新增一节：17.4 节位置拍卖（根据第 7 版中的在线广告拍卖这个案例扩充而成）
- 新增案例：让墙报价
- 新增一节：17.7 稳定的婚姻问题
- 新增一节：17.8 机制设计

第 18 章：技术

- 新增案例：数据处理中心
- 新增案例：完全复制！

第 21 章：成本曲线

- 新增一节：21.4 节在线拍卖的成本曲线

第 23 章：行业供给

- 新增一节：23.11 烟尘排放税 Versus 排污上限与交易

第 24 章：垄断

- 新增案例：土豆供给管理

第 27 章：寡头垄断

- 新增案例：价格匹配（price matching）

第 35 章：信息技术

- 新增案例：苹果公司的 iPod 和 iTunes
- 新增案例：iPod 是谁生产的？
- 新增案例：关键词竞价广告（Adwords）和相关广告（Adsense）
- 新增一章：35.8 双边市场
- 新增案例：在线双边市场

第 36 章：公共物品

- 改写一节：将第 7 版中的"36.9 需求显示"改写为"36.9 维克里-克拉克-格罗夫机制"。
- 新增一节：36.10 维克里-克拉克-格罗夫机制的例子
- 改写一节：将第 7 版中的"36.10 克拉克税存在的问题"改写为"36.10 维克里-克拉克-格罗夫机制存在的问题"

注：此为翻译者总结。

目录

目录（续）

第八版前言

中级微观经济学前七版取得的成功让我很欣慰。这一事实验证了我的信念:市场欢迎用解析方法为大学生撰写的微观经济学教科书。

写该书第一版的目的在于向学生展示微观经济学的处理方法,以便让他们能自主运用这些工具，而不是被动地吸收教材中简化的案例（predigested cases)。我发现做此事的最好方法就是强调微观经济学的基本原理并且提供具体的案例,而不是试图提供百科全书式的术语和故事。

使用该方法的困难在于，很多大学缺乏为经济学课程准备的数学先修课。微积分知识和解题经验的缺乏，使得一些经济学分析方法难以展现。然而，这并非不可能做到。即使学生仅具有一些线性需求函数和供给函数以及基本代数知识，在经济学道路上也可以走得很远。不使用过多的数学又能以解析方法进行分析，这是完全可以做到的。

解析方法和其他方法的区别值得强调。经济学的解析方法是使用严格逻辑性的推理。这并不必然要求使用高等数学方法。借助数学语言可进行严谨分析，在需要时使用数学无疑是前进的最好方法，然而它并不适用于所有学生。

经济学专业的大学生应该掌握微积分，但很多学生没学——至少没学好。因此，本教材的主体没有使用微积分。然而在很多章节的附录部分，我还是提供了完整的微积分知识。这意味着你要是会用微积分就用它，你要是不会用也没有关系。

这种安排方式传达了这样的理念:微积分并非教材内容的脚注，而是研究经济学更强的工具，尽管这些内容你也可以仅使用语言和图形进行研究。借助一点数学工具，很多论证将更简单，所有经济学专业学生应掌握这些数学工具。在很多情形，我发现借助一些研究动机和一些漂亮的经济学案例，学生更有热情从解析的角度看待问题。

这本教材还有其它几个创新之处。首先，每一章通常比较短。我试图将大多数章节的长短控制在 “一次讲课的时间”，以便一下子读完。我也按照经济学标准的讨论顺序，首先分析消费者理论接着才是生产者理论，但是我在消费者理论上花费的时间稍微多些。这并不代表我认为消费者理论是微观经济学中最重要的部分；而是因为，我发现这部分内容是学生最难理解的，因此我想更详细地阐述。

其次，我置入了大量的例子来表明如何运用教材中的理论。在大多数经济学教科书中，学生观摩了很多曲线移动的图形，但是他们看不到多少代数，或者看不到多少该问题的计算。但在实践中要用代数来解决问题。图形可提供直觉见解，但经济分析真正的力量来自于对经济问题的数量计算。每个经济学专业的学生，都应该能够将经济故事转换为方程式或者数值的例子，但该能力的培养时常被忽视。因为此，我也提供了一本练习册，该习题册是本教材的有机补充。该习题册由我和我的同事西奥多.博格思众（Theodore Bergstrom）共同完成，

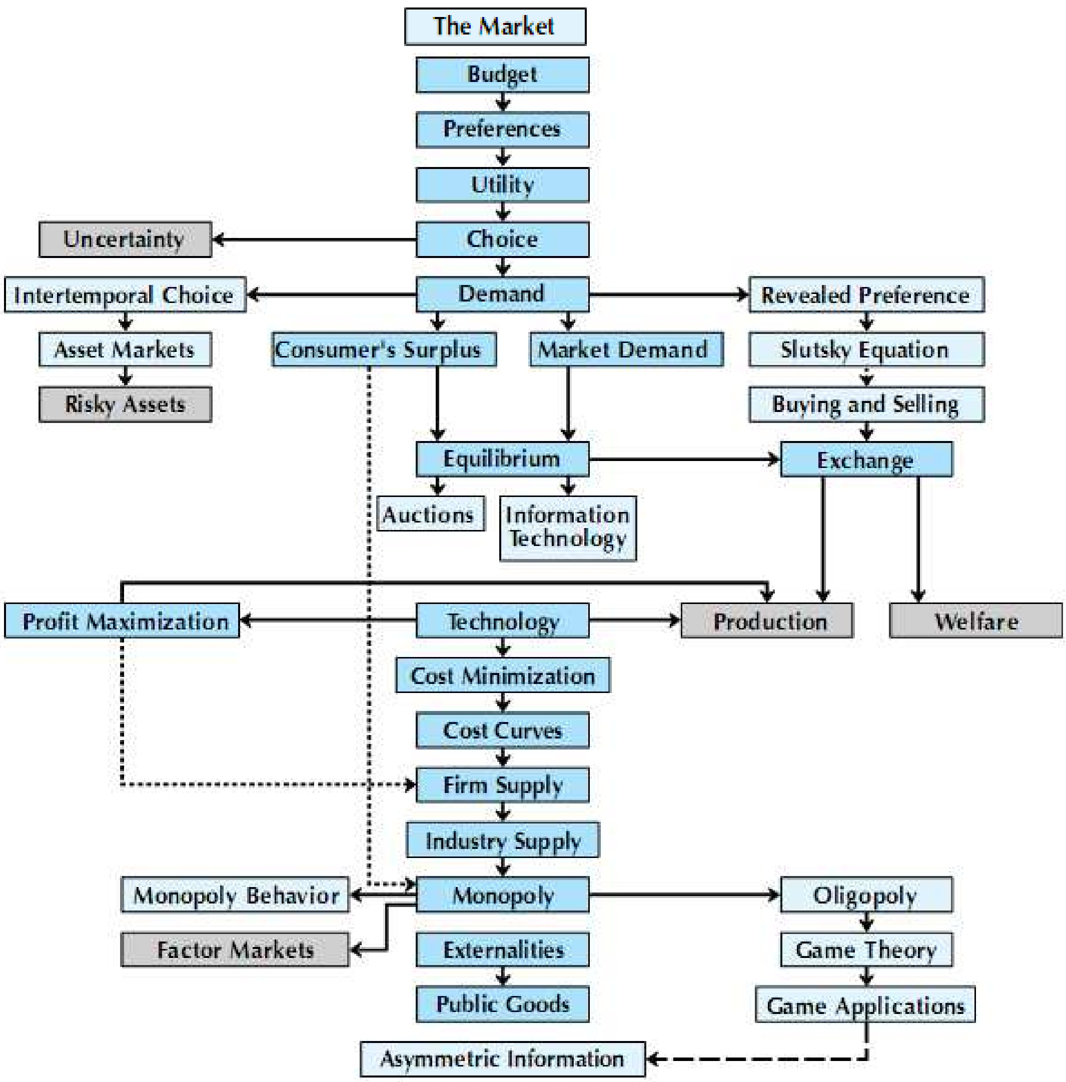
曹乾（东南大学 caoqianseu@163.com）

我们花了很大工夫来构造有趣而又有启发性的题目。我们认为该习题册为经济学专业的学生提供了重要的帮助。

第三，我相信我对本书主题的处理比一般中级微观经济学教科书更准确。需要承认，当分析一般情形很困难时，我有时会选择特殊的情形进行分析，这么做时我并没隐瞒。总的来说，在论证时我尽量详细列出每一步的步骤。我相信我提供的讨论不仅更完整更准确，而且注意细节的事实会让学生更理解论证过程，而一般教科书只是松散地讨论。

讲授本教材的路线

本书材料很多，很难在一学期轻松讲完，因此需要仔细地选择你希望研究的内容。如果你从第 1 页讲起，按照教材章节顺序向下讲，你肯定讲不完。本书的模块结构使得教师在选择讲授材料时有很大的自由度，我希望更多地人利用这个自由度。下列图形展示了各章之间的相关关系。



 曹乾（东南大学 caoqianseu@163.com）

深颜色的章节为 “核心”章节——每个中级微观经济学课程均应包括这几章。浅色章节为“可选”章节：每学期我都选择其中几个讲授。灰色章节通常不在我的讲授内容中，但在其他课程中，很容易将这些课程包括进来。从 A 章到 B 章的实线表明 A 章应该先学。若是虚线则意味着学习 B 章时应了解一些 A 章的内容，但在内容上并不显著依赖前一章。

我一般先讲“消费者理论”和“市场”然后直接过渡到“生产者理论”。另外一种流行的路线是在“消费者理论”之后，直接讲“交易”。 很多教师喜欢这条路线，我特意这么试讲后确信这条路是可行的。

一些教师喜欢把生产者理论放在消费者理论之前讲述。用这本书这么做也可行，但是教师需要补充些材料。例如，在介绍等产量曲线时，我假设学生已经了解无差异曲线。

有关公共产品，外部性，法律，和信息等章节的内容可在课程中提前介绍。我已把这些章节的材料安排好，你想放在哪介绍都很容易。

类似地，公共产品章节可用作埃奇沃思（Edgeworth）盒状图的分析材料。外部性可在学完成本曲线后立即介绍。信息这一章放在哪学习都行，只要学生熟悉了经济分析方法。

第八版的变化

在这一版中，我增加了一些和教材内容相关的例子，包括延长版权保护期、资产价格泡沫、合同对方不履行合约责任的风险、在险价值和烟尘排放税。我继续提供来自硅谷企业的例子，这些企业包括 Apple、eBay、Google、Yahoo 等。我讨论了下列一些主题：iPod 和 iTunes 之间的互补性，与 Facebook 之类的公司相关的正反馈，Google、微软和雅虎等使用的广告拍卖模型等。我相信这些现实经济中的例子新鲜而有趣。

我也进一步讨论了机制设计问题，包括双边配对市场和 Vickrey-Clarke-Groves 机制。人们一度认为这个领域主要是理论上的探讨，然而现在它具有十分重要的实践意义。

试题库与习题册

《中级微观经济学习题集》这本习题册是本课程的有机组成部分。它提供了很多填空练习题，旨在引导学生逐步实际运用教材中学到的工具。除了练习题之外，每章都有一个由多项选择题组成的小测验，这些选择题来源于习题册中的题目。小测验的答案也包含在习题册中。这些小测验帮助学生快速检验他们做题的成果。

不仅如此，如果教师授课时选用了该习题册，我们还提供与教材配套的习题库。试题库的题目来自于每章的小测验，但被改编为若干个版本。每个版本的内在逻辑相同，区别仅在于用于计算的数值不同。试题库可用于给学生布置额外的练习，或者用于课堂小测验。评分便捷且可靠，因为题目都是多项选择题，可用电子仪器读卡打分。

在我们的课堂中，我们让学生做完每章所有的测验题，可以独立完成也可以组建学习小

组协作完成。因此，讲课过程中每两周我们都有个课堂小测验，题目来源于试题库的不同版本。这些题目在本质上等同于习题册中的小测验，仅仅数值不同而已。因此，做家庭作业的学生会发现这些测验题目容易解答。

我们坚信不做题目很难学好经济学。习题册和试题库中提供的小测验让学习过程等容易。试题库的精装版本可从出版商买到，出版商还出售该教材的教师指导手册，该手册包括我的讲课建议、每一章的讲稿和习题册练习的答案。

本教材的其他一些有用辅助材料也可得到，包括完整的幻灯片以及诺顿经济通讯（Norton Economic News Service），该通讯可以提醒学生哪些经济新闻和教材中的特定材料有关。有关这些辅助材料的信息，请浏览本书的主页 http://www.wwnorton.com/varian.

本书的制作

这本教材是由我使用 TEX 进行排版的，该排版系统是由 Donald Knuth 开发的。我在 Linux 系统上操作，使用 GNU emacs 进行编辑，用 rcs 做版本控制以及用 TEXLive 进行加工我用 makeindex 制作索引，用特雷弗.达雷尔的 psfig 软件插入图形。

图书设计工作由 Nancy Dale Muldoon 完成，但由我和 Roy Tedoff 稍微做了部分修改。Anne Hellman 是原稿编辑，Jack Repchek 协助完成了编辑工作。

致谢

还有一些人对本书有贡献。首先，我必须感谢本书第一版的编辑助理，John Miller 和 Debra Holt. John 提供了很多建议。在使用本书草稿教学期间，他提供了很多练习。他还对本书章节的连贯性做了很大的贡献。在本书成书的最后阶段，Debra 仔细校对并准备了目录索引。

我在准备第一版的工作时，下列人员提供了有益的建议和评议：Ken Binmore (University of Michigan), Mark Bagnoli (Indiana University), Larry Chenault (Miami University), Jonathan Hoag (Bowling Green State University), Allen Jacobs (M.I.T.), John McMillan (University of California at San Diego), Hal White (University of California at San Diego), and Gary Yohe (Wesleyan University)。特别地，我要感谢 Reiner Buchegger 博士，他将本书翻译成德文，他仔细阅读了本书第一版，而且详细列出了一些错误之处。在第一版出书之前，我还要感谢 Theodore Bergstrom, Jan Gerson, Oliver Landmann, Alasdair Smith, Barry Smith 以及 David Winch。

第二版的编辑助理是 Sharon Parrott 和 Angela Bills，他们在书写和编辑上为我提供了有用的帮助。Robert M. Costrell (University of Massachusetts at Amherst), Ashley Lyman (University of Idaho), Daniel Schwallie (Case-Western Reserve), A. D. Slivinskie (Western

Ontario), and Charles Plourde (York University)等同行向我提供了详细的评价和建议，以便让第二版做得更好。

在准备第三版时，下列个人进行了有益评价：Doris Cheng (San Jose), Imre Csek´o (Budapest), Gregory Hildebrandt (UCLA), Jamie Brown Kruse (Colorado), Richard Manning (Brigham Young), Janet Mitchell (Cornell), Charles Plourde (York University), Yeung-Nan Shieh (San Jose), John Winder (Toronto). 我特别感谢以下诸位： Roger F. Miller (University of Wisconsin), David Wildasin (Indiana)，因为他们详细进行了评价、提出了有用的建议并纠正了一些错误之处。

第五版受益于下列个人：Kealoah Widdows (Wabash College), William Sims (Concordia University), Jennifer R. Reinganum (Vanderbilt University), and Paul D. Thistle (Western Michigan University).

在准备第六版时，我收到了以下个人的评价：James S. Jordon (Pennsylvania State University), Brad Kamp (University of South Florida), Sten Nyberg (Stockholm University), Matthew R. Roelofs (Western Washington University), Maarten-Pieter Schinkel (University of Maastricht), and Arthur Walker (University of Northumbria).

第七版得到了以下个人的评价：Irina Khindanova (Colorado School of Mines), Istvan Konya (Boston College), Shomu Banerjee (Georgia Tech) Andrew Helms (University of Georgia), Marc Melitz (Harvard University), Andrew Chatterjea (Cornell University), and Cheng-Zhong Qin (UC Santa Barbara).

最后，第八版得到了以下人士的有益评论：Kevin Balsam（Hunter College）, Clive Belfield (Queens College, CUNY), Jeffrey Miron (Harvard University), Babu Nahata (University of Louisville), and Scott J. Savage (University of Colorado).

伯克利，加利福尼亚

2009 年 10 月

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

1.市场（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

1 市场

微观经济学教科书的第一章的传统写法，是介绍经济学的“研究范围和方法论”。这样的安排比较有趣，但不太适合经济学**初学者**的学习。在学习一些实际经济分析的例子之后，你才能真正理解这样的讨论。

因此，我们以经济学**案例**分析作为本书的开始。在本章，我们将分析公寓市场模型。在分析过程中将引入一些经济学新概念和新工具。若我们介绍得太快你也不必担心，因为本章的目的仅仅在于让你了解这些概念是如何运用的，以后我们还将更加细致地讨论。

1.1 构建模型

经济学研究需要构建社会现象的**模型**（models）。模型是指现实的简化表示。此处需要强调“简化”二字。按 1:1 比例作出的地图毫无用处，试图全方位解释现实的经济模型也是如此。模型的力量来自去除无关紧要的细节，从而经济学家可以将主要精力集中于经济现实的本质特征，而这些现实是经济学家想要搞清楚的。

此处我们感兴趣的是公寓价格的决定因素，因此我们需要对公寓市场进行简化的描述。在构建模型时，如何进行简化是个艺术活。一般来说，我们希望采用最简单的模型来描述我们关注的经济形势。然后我们可以逐渐增加因素让模型变得更复杂，也希望这样的模型更符合现实。

我们选取的例子，是美国中西部中等规模大学城的公寓市场。城内的公寓有两类：离大学近的和离大学远的。通常认为学生想要离大学较近的公寓，因为上学方便。离大学较远的公寓，则需要搭公交车，或者骑自行车但要忍受寒冷和道路的漫长，因此大多数学生会选择离大学较近的公寓，当然前提是他能付得起租金。

我们按公寓离大学的距离将大学城分为两环：离大学近的位于内环，其余位于外环。我们仅分析位于内环的公寓。位于外环的公寓可以这样解释，在找不到内环的公寓时才选择外环。假设外环的公寓很多，价格固定在某已知水平。我们只关注内环公寓价格的决定因素以及谁能住在内环。

在公寓模型中，对于两类公寓的价格的差别，经济学家会这么说，外环公寓的价格是一个**外生变量**（exogenous variable），而内环公寓的价格是个**内生变量**（endogenous variable）。这是说，外环公寓的价格被认为是由未在该模型中讨论的因素所决定，而内环公寓的价格是由模型该模型本身的因素所决定。

在我们的模型中，我们做出的第一个简化是，假设所有公寓除了位置以外其余方面都相同。这样，公寓的“价格”这种说法才能讲得通，而无需担心公寓是一个卧室还是两个卧室，

曹乾（东南大学 caoqianseu@163.com）

或者其他方面的不同。

可是，这个价格的决定因素是什么？哪些因素决定了谁住在内环或外环？怎样评价不同经济机制在公寓配置中的合意性？我们可以使用什么概念来检验公寓不同配置方式的价值（merit）？我们希望我们的模型能解释上述这些问题。

1.2 最优化和均衡

在解释人们的行为时，我们需要一个能让我们的分析立足于其上的架构（framework）。在经济学的大部分内容中，我们使用基于下列两条基本原理的架构。

最优化原理（The optimization principle）：人们选择能支付得起的最优消费模式。

均衡原理（The equilibrium principle）：价格调整到人们对某商品的需求数量等于供给数量时为止。

思考一下这两个原理。第一个原理**几乎**是同义反复的。如果人们能自由选择其行为，自然可以假设：人们选择想要的而不选择不想要的。这个一般原理当然有例外，但例外情形一般不属于经济行为的范畴。

第二个原理有点疑问。你至少可以想到，在任何既定的时间，人们的需求和供给是不相等的，因此某些因素一定在变化中。这些变化可能要经历很长的时间才能平息，更糟的结果也有，它们可能引起其它因素的变化从而可能导致整个系统的“崩溃”。

上述情形可能发生，但通常不会。在公寓市场的情形中，我们通常发现租房价格在月与月之间保持稳定。我们对该**均衡**价格感兴趣，而不管市场是如何达到均衡的，也不管市场在较长的时期内是如何变化的。

需要注意，均衡的概念在不同模型中可能是不同的。在本章我们要研究的简单市场模型中，需求和供给均衡的概念已足够满足需要。但在更一般的模型中，我们将需要均衡的一般性定义。典型地，均衡要求经济人（economic agents）的行为彼此应该一致。

怎样使用上述这两个原理，来解答我们在前面提出的一系列问题？需要引入一些经济学概念。

1.3 需求曲线

假设我们考虑公寓的所有潜在的租赁者，逐一询问他所愿意支付的最高租房价。我们从高价开始。肯定有人愿意出最高价。也许这个人有很多钱，也许他很懒不愿意走那么远或其他原因。假设此人愿意出价 500 元/月租赁一套公寓。

若只有一人出此高价，若一套公寓的租金也为 500 元/月，则恰好只有一套公寓租出—

—租给了那个愿意出该价格的人。

假设第二最高出价为 490 元，若市场价格为 499 元，则仍然只有一套公寓租出：愿意出价 500 元者得到，而愿意出价 490 元者租不到。照这样下去，若市场价格为 498 元, 497 元, 496 元等等，市场仍然只一套公寓租出。直到价格为 490 元时，恰有两套公寓租出：一套租给出价 500 元者，一套租给出价 490 元者。

类似地，市场将一直只有两套公寓租出，直到市场价格等于第三最高出价时，才有三套公寓租出。如此进行下去。

经济学家将一个的愿意支付的最高价格称为该人的**保留价格**（reservation price）。保留价格是某个人愿意接受的最高价格但仍以该价格购买某商品。换句话说，某人的保留价格是指在该价格水平下，买或不买这两种结果对他来说恰好是无差异的。在我们的例子中，若某人保留价格为 p，这表示以下两种结果对他来说是无差异的：一是花费 p 元钱住在内环公寓；二是租住外环公寓。

因此，在既定价格 p^* 水平下租出的公寓数量，恰好等于保留价格大于或等于 p^* 的那些人的人数。因为如果市场价格为 p^*，则愿意至少出资 p^* 的那些人想住内环的公寓，而不愿意出此价钱的人则选择住在外环。

我们可以将这些保留价格在图形中画出来，如图 1.1 所示。我们以纵轴表示价格，以横轴表示愿意至少出这些价格的人数。

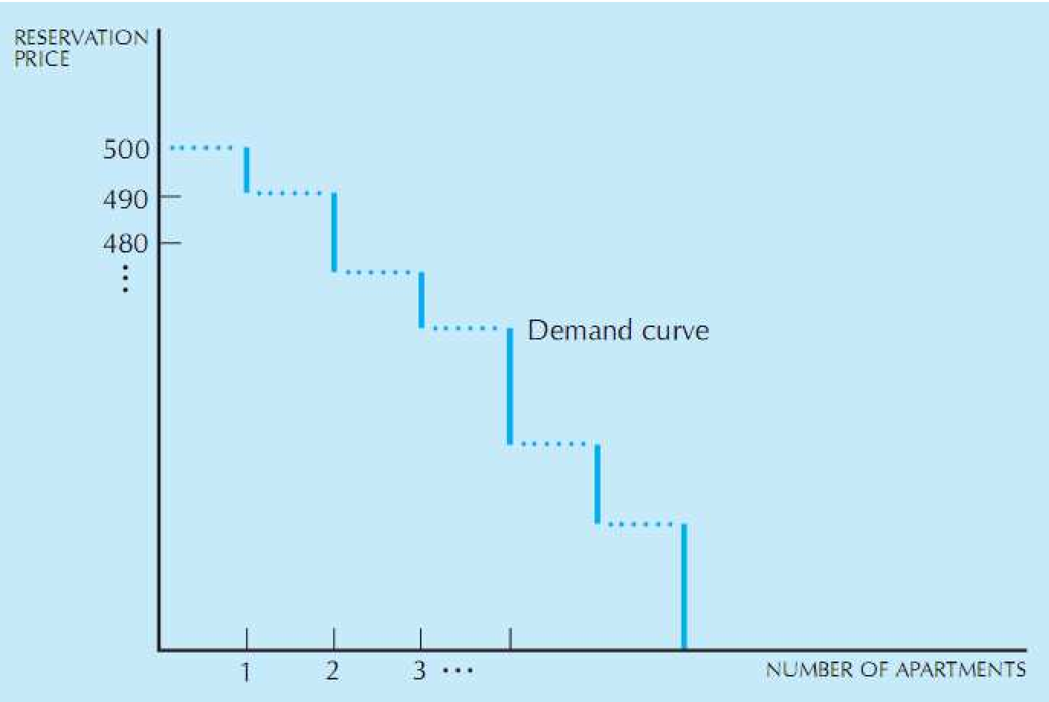


图 1.1：**公寓的需求曲线**。纵轴表示市场价格，横轴表示在每个价格水平下租出的公寓数量。

图 1.1 的另外一种看法是，将其作为一种衡量工具，用来衡量任一既定价格水平下愿意

租赁公寓的人数。图中所示的曲线称为**需求曲线**（demand curve）：将需求量和价格关联起来的曲线。当市场价格高于 500 元时，公寓的需求量为 0。当价格在 500 元和 490 元之间时，需求量为 1。当市场价格在 490 元和第三高的保留价格之间时，需求量为 2，以此类推。需求曲线描述了每个可能的价格水平下相应的需求量。

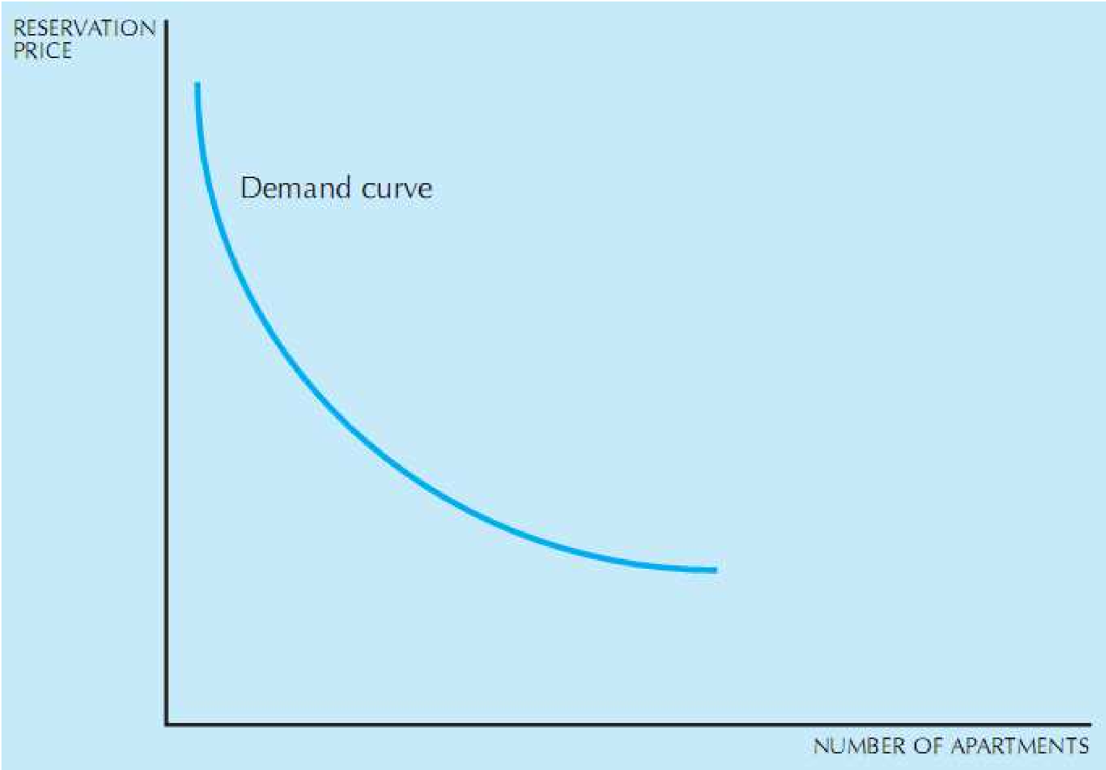


图 1.2：**很多需求者情形下的公寓需求曲线**。由于公寓需求者数量众多，不同价格之间的跳跃将变得很小，从而需求曲线通常是平滑的曲线。

需求曲线向下倾斜：随着公寓租金价格下降，更多的人愿意租住公寓。如果想租住公寓的人很多而且他们的保留价格相差很小，可以认为需求曲线是一条向下倾斜的平滑曲线，如图 1.2 所示。如果想租住公寓的人很多，图 1.1 中的曲线就趋近于图 1.2 中的曲线。也就是说这种情形下，图 1.1 曲线中的“跳跃”与市场的容量相比已变得非常小，因此在画市场需求曲线时，我们可以放心地忽略这些“跳跃”，从而将需求曲线画成向下倾斜的平滑曲线。

1.4 供给曲线

我们已有了描述需求行为的曲线即需求曲线，下面我们开始分析供给行为。我们首先必须说明我们所要研究的市场的特征。我们考虑的情形是这样的，即市场中有很多的房东，他们公开出租公寓，价格当然是越高越好。这种类型的市场称为**竞争性市场**（competitive market）。市场未必都是竞争性的，其他市场类型我们稍后分析。

目前，我们考虑的情形是市场中有很多房东，他们的行为彼此独立。显然，如果房东卖力招租，如果租房者完全知道房东索要的租金价格，那么内环所有公寓的均衡价格必然是相同的。说明这个结论并不困难。用反证法。假设租金价格不同，即存在着高价格 p_h 和低价格 p_l。租住高价公寓的人会跑到提供低价公寓的房东之处，以高于 p_l 但低于 p_h 的价格租房。这样价格下达成的交易对租房双方都有好处。只要租房双方不断追逐自己的利益，并且双方都知道市场中其他租房价格，那么均衡时就不会存在同物不同价的情形。

那么这个均衡价格为多大？我们使用已经在构建需求曲线过程中运用过的方法：任选一个价格水平，分析在该价格下市场供给公寓的数量。

市场供给公寓的数量在某种程度上取决于我们研究的时间段。如果这个时间段是几年，则可能有人建新公寓，公寓的数量当然对租金有反应。但如果这个时间段是"短期"，比如某个既定的年份，公寓的数量是固定的。如果我们仅考虑短期情形，则公寓的数量是既定的。

短期内，公寓市场的供给曲线是一条垂线，如图 1.3 所示。不管租金如何变化，公寓的数量都是既定的，也就是说当时所有的公寓都是现成的，拎包即可入住。

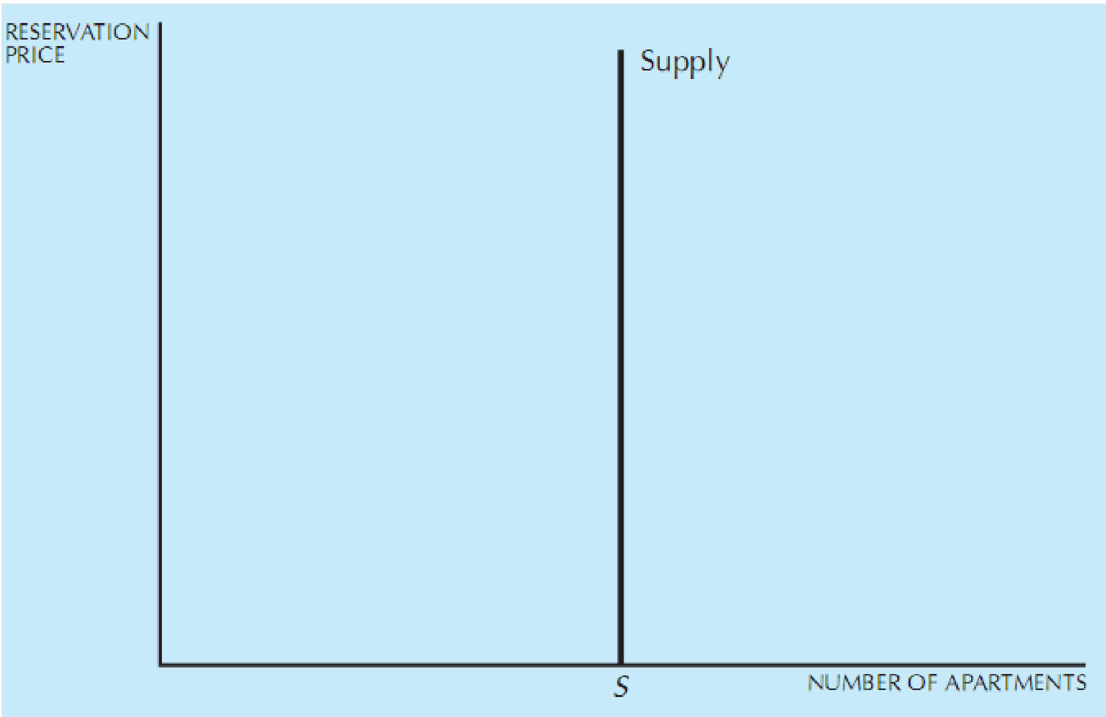


图 1.3：**短期供给曲线**。在短期，公寓的供给数量是固定不变的。

1.5 市场均衡

现在我们有了公寓市场的需求方和供给方的表示方法（即需求曲线和供给曲线）。我们将这两个工具放在一起去分析市场的均衡行为。如图 1.4 所示，我们在同一张图中同时画出了需求曲线和供给曲线。

在图 1.4 中，我们用 p^* 表示公寓的**均衡价格**（equilibrium price），即公寓需求量等于供给量时的价格。在这个均衡价格上，愿意至少支付 p^* 租金的每个消费者都能够找到公寓租住，每个房东都愿意按照现行的市场价格出租公寓。消费者和房东都没有任何理由改变自己的行为。这就是为什么我们把这种状态称为**均衡**的原因：无人改变自己的行为。

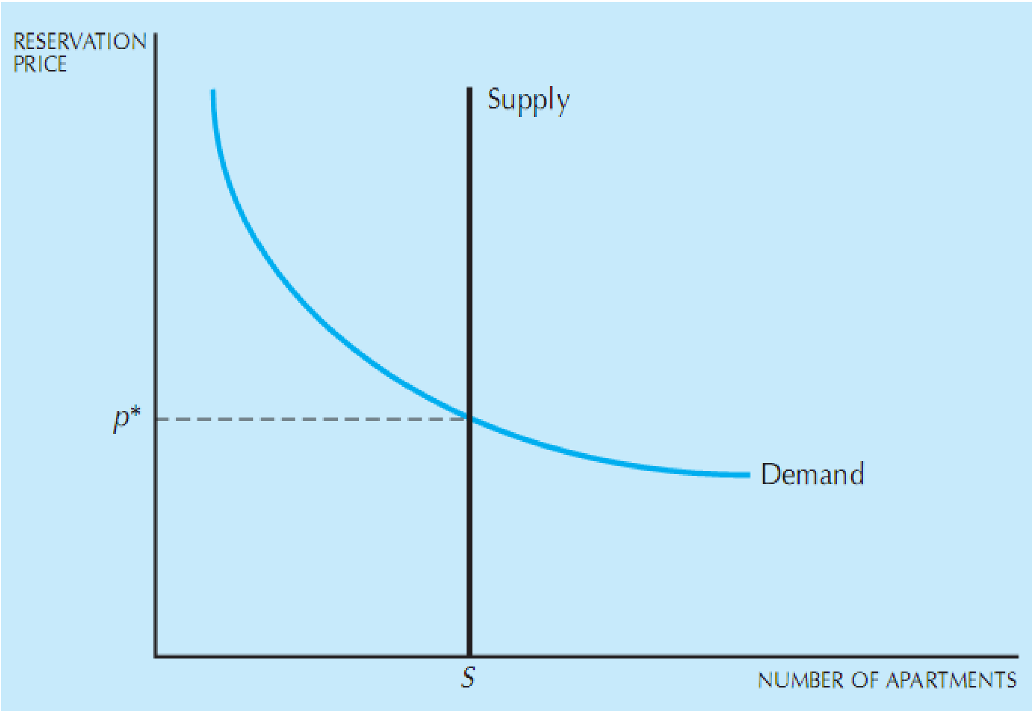


图 1.4：公寓市场的均衡价格。供给曲线和需求曲线相交时的价格为均衡价格 p^*。

为了更好理解这一点，我们考虑价格不是均衡价格 p^* 将会发生什么现象。例如，考虑价格 p 且 $p < p^*$，即该价格为需求大于供给时的价格。这样的价格能持续存在吗？在这样的价格水平上，必然出现某些房东无房可供的局面。在这样的价格水平上，愿意租房的人数必然大于公寓的供给数量。某些房东当然会提高价格，因为这对房东自己有利。

类似地，假设价格为 p 且 $p > p^*$，则有些公寓会空置：在此价格水平下，愿意租房的人数小于公寓的供给数量。某些房东可能一间公寓也租不出去。于是他们自然有降低价格吸引顾客的动机。

如果价格高于 p^* 则租房者太少；如果价格低于 p^*，则租房者太多。只有当价格等于 p^* 时，愿意租房者的人数才和公寓供给数量相等。只有在 p^* 时需求才等于供给。

价格为 p^* 时，房东的行为和租房者的行为是相容的（compatible），因为此时公寓的需求数量等于供给数量。 p^* 正是公寓市场的均衡价格。

我们已经确定了内环公寓市场的价格，现在问个问题：哪些人最终住进了内环公寓，哪些人被赶到较远的外环公寓？根据我们上述模型，这个问题容易回答。市场均衡时，愿意支付不低于 p^* 的价格的那些人住进了内环，而只愿意支付低于 p^* 的价格的那些人住进了外环。保留价格为 p^* 的人对于住在内环还是外环无所谓。其他住在内环的人，他们实际支付的价格小于他们的保留价格。因此，公寓这种资源的配置取决于租房者愿意支付多少钱。

1.6 比较静态分析

现在我们使用上面得到的公寓市场模型，分析均衡价格的行为。例如，你也许会问当市场中的各种因素变化时，公寓价格如何变化？这样的问题涉及到了**比较静态**（comparative statics）分析，因为需要比较两个“静态”的均衡结果，而不必关心市场如何从一个均衡移动到另一个均衡。

从一个均衡到另一个均衡的移动可能很耗时间，这种移动是怎样发生的问题无疑有趣而且重要。但在学会跑之前我们必须先学会走，因此我们暂时不研究这样的动态问题。比较静态分析只涉及均衡的比较，目前比较静态领域中的问题已足够你分析了。

我们先分析一个简单的情形。假设公寓的供给增加，如图 1.5 所示。

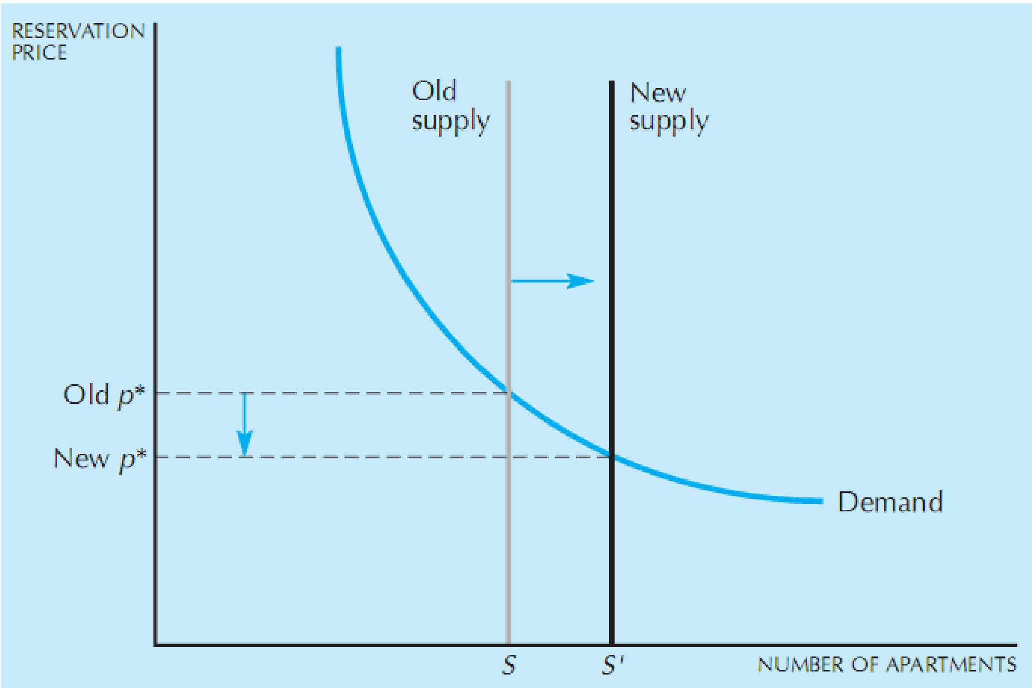


图 1.5：**公寓供给增加**。当公寓供给增加时，均衡价格下降。

由图易知，公寓供给增加，均衡价格下降。类似地，如果公寓供给减少，均衡价格上升。

我们再分析一个相对复杂但更有趣的例子。假设某开发商决定将某些公寓（apartments）

 曹乾（东南大学 caoqianseu@163.com）

改变为产权出售的公寓（condominiums）[一]。剩下公寓的价格如何变动？

你最初可能认为公寓的价格会上升，因为公寓的数量减少了。你的观点未必正确。用于出租的公寓供给确实下降了；但是别忘了**公寓的需求也减少**了，因为某些租住公寓的人可能决定购买那些产权出售的公寓（condominiums）。

自然地可假设购买公寓的人来自住在内环公寓的那些人，即愿意支付大于 p^* 的租金的那些人。例如，假设保留价格位于前 10 名的需求者决定购买公寓，而不是继续租住，则将原需求曲线任一价格水平下的需求量都减去 10，即可得到新需求曲线。由于需求者减少了 10 人而且公寓数量也减少了 10 个，新的均衡价格和原均衡价格相同，最终仍然是原来的那些人住在了内环。图 1.6 显示了这种情形。需求曲线和供给曲线都向左移动了 10 单位，因此均衡价格不变。

多数人会对该结论感到惊讶。他们想到了公寓的供给减少，但没想到公寓的需求也减少了。当然上述例子比较极端，因为我们假设**所有**购买公寓的人都是内环公寓的租住者。但是，如果你假设购买公寓的人都不是内环公寓的租住者，你的这种假设岂不是更极端？

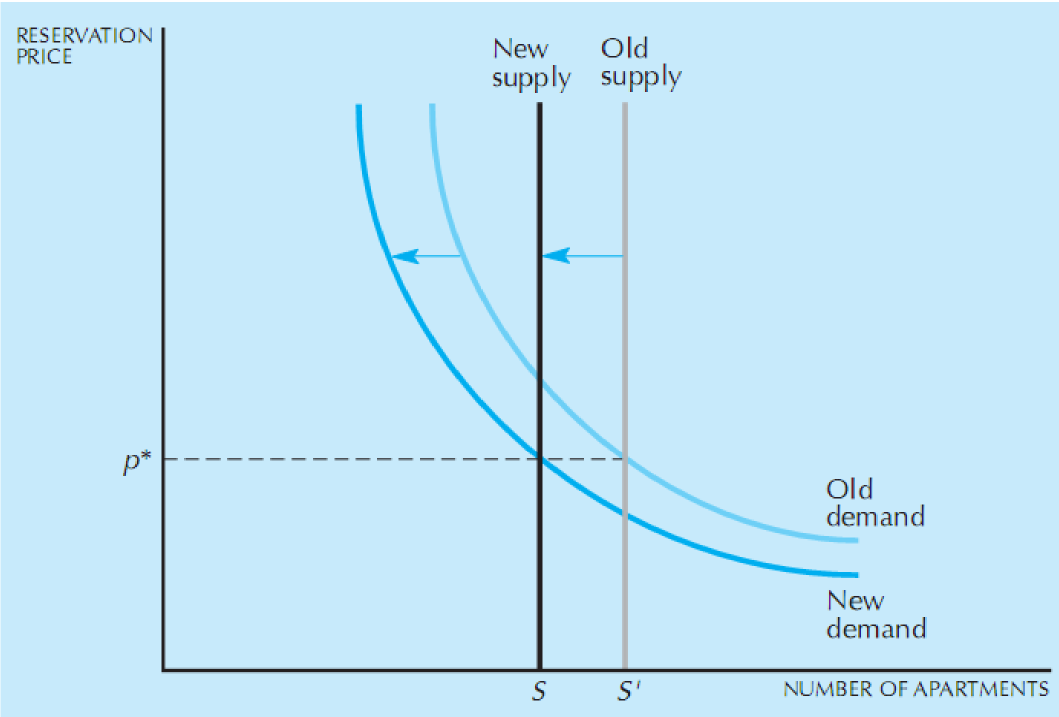


图 1.6：**改变公寓用途的效应**。如果需求和供给都向左方移动相同的数量，则均衡价格不变。

上述模型尽管简单，却使我们得到了一个重要结论。如果我们想确定改变公寓用途对公寓市场的影响，我们必须同时考虑它对公寓供给的影响以及对公寓需求的影响。

我们再看一个关于公寓税收效应的比较静态分析，这个例子的结论也会让你吃惊。假设市议会决定对每个公寓每年征收 50 元的税。因此每个房东按自己拥有的公寓数量缴税。

[一] apartment 和 condominium 都是“公寓”，二者在实体上差别不大，但前者用于租（产权不转让），后者用于卖（产权转让）。此处只要知道它们是两种商品即可。译者注。

这会对公寓的价格造成何种影响？

大多数人会认为至少部分税收会转嫁给租房者。但是，令人惊讶的是他们的观点是错的。事实上，公寓的均衡价格仍然维持不变！

为了验证这个结论，我们首先分析税收对需求曲线和供给曲线的影响。供给曲线不变，因为征税前后公寓的数量没变。需求曲线也不变，因为对于任一价格水平，公寓的需求量都没变。如果需求曲线和供给曲线都没有移动，征税后均衡价格不会变动。

你可以按照下面这种思路理解上述结论。在征税前，每个房东在保证自己公寓租出的前提下尽可能地索要高价。均衡价格 p^* 是与市场公寓供给数量相容的最高价格。征税后房东能否提高价格以补偿征税给他造成的损失？答案为否：如果他们能提高价格并且同时保证自己的公寓能顺利租出，那么他们在征税前就会提高价格。如果他们索要的最高价格已经是市场能承受的最高价格，房东们不可能进一步提高价格：转嫁给租房者的税收为零。因此房东必须独立承担全部税收。

需要指出，上述分析必须假设公寓供给保持不变。如果公寓数量随税收变动而变动，则租房者支付的价格通常会改变。等我们有了更强大的分析工具之后，我们再分析公寓供给随税收变动而变动的情形。

1.7 公寓配置的其他方法

在上一节，我们已经分析竞争市场中的公寓均衡问题。但是，配置资源的方法有多种，竞争市场只是其中一种。在本节我们将研究其他一些配置方法。有些方法听起来似乎比较奇怪，但每种方法都阐明了一种重要的经济思想。

歧视性的垄断者

我们首先分析的情形是，全部公寓都归某个房东所有。或者，一些房东联合起来共同行动，从而可以看成为一个大房东。某产品的市场若由一个卖方主导，则这样的情形称为**垄断**（monopoly）。

在出租公寓时，房东可以决定将公寓逐个拍卖，谁出的租金价格高就先租给谁。由于这意味着不同的租房者最终对公寓支付的租金价格不同，我们将这种情形称为**歧视性的垄断者**（discriminating monopolist）。为简单起见，假设歧视性的垄断者知道每个租房者对公寓的保留价格。（这个假设并不很符合实际，但它足以阐明一个重要的观点。）

上述假设意味着房东可将第一个公寓租给出价最高的人，在本例中这个最高价格为 500 元。沿着需求曲线依次向下，第二个公寓租给出价 490 元的人，依次类推。这样，每一个公寓都租给了出价相对最高的人。

曹乾（东南大学 caoqianseu@163.com）

歧视性垄断者的一个有趣特征是：**这种情形下租到公寓的人，恰好是竞争性市场中租到公寓的那些人**，即租到公寓的人都是愿意支付不低于 p^* 租金的那些人。歧视性垄断的情形下，租到公寓的最后一个人支付的租金为 p^*，而这恰好是竞争性市场的均衡价格。追求自身利润最大化的歧视性垄断者，它对公寓的配置竟然与竞争性市场的供给和需求机制是完全相同的。在这两种情形下，尽管每个人支付价格不同，但谁能得到公寓这一结果是相同的。这样的结果并非巧合，但是我们必须等到以后才解释其中的原因。

普通垄断者

在前面我们假设歧视性垄断者可以对每个公寓索要不同的价格。但是，如果规定他必须以同样的价格出租公寓，结果会如何？在这种情形下，该垄断者必须权衡下列两种选择：如果索要低价，租出的公寓数量也多；如果索要高价，可能最终因租出的公寓变少而少赚了钱。

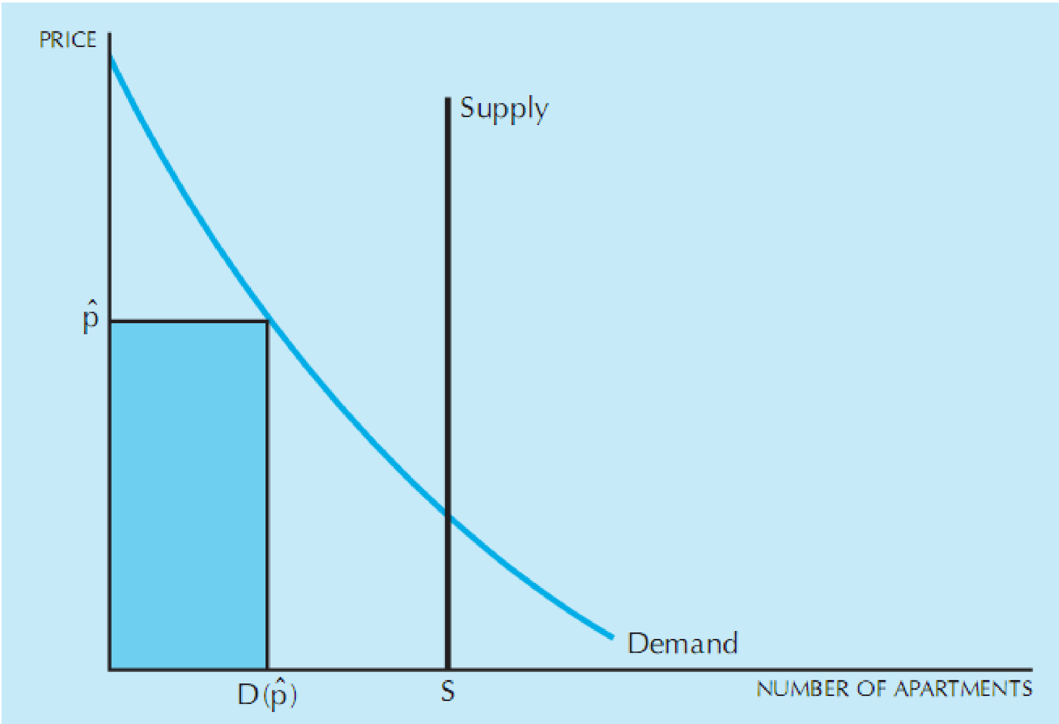


图 1.7：**用矩形的面积表示收入**。垄断者的收入等于价格乘以需求量，这可用图中的矩形面积表示。

令 $D(p)$ 表示公寓的需求曲线，即价格为 p 时公寓的需求量。如果垄断者定价为 p，则他租出的公寓数量为 $D(p)$，他的收入为 $pD(p)$。这个收入可以用矩形的面积衡量：矩形的高(即长)和宽分别为 p 和 $D(p)$。矩形的面积（长乘以宽）就是垄断者的收入。如图 1.7 所示。

如果垄断者出租公寓的成本为 0，他将选择使矩形面积最大的价格。如图 1.7 所示，矩

形面积最大时，价格为 $\hat{p}$ 。在这种情形下，垄断者发现只出租部分公寓对他有利。事实上，这正是垄断者的特征。垄断者通常限制供给量以使自己的利润最大。这意味着垄断者通常索要的价格高于竞争性市场的均衡价格 p^* 。与竞争性市场相比，在普通垄断者的情形下，垄断者供给的公寓数量变少，价格变高。

租金管制

我们要分析的第三种情形也是最后一种情形是租金管制。假设该城市决定对租金设定上限，比如定为 p_{max} 。假设 p_{max} 小于竞争性市场的均衡价格 p^*，此时会出现**超额需求**（excess demand）的现象：愿意以 p_{max} 租住的人数超过了公寓的供给数量。哪些人最终将住进公寓？

直到目前我们还不能用均衡理论分析上述问题。我们已经知道供求相等时结果会如何，但我们的模型还不能提供充分的细节让我们分析供求不相等时结果会如何。租金管制情形下谁最终住进了公寓，取决于谁最有时间到处找房，取决于谁对目前的租房者最了解，等等。这些细节都超出了我们模型的范围。租金管制情形下最终得到公寓的人，有可能正是竞争市场情形下得到公寓的那些人，但这样的可能性不大。最有可能的情形是某些原本住在外环公寓的人得到了内环公寓，从而挤走了在竞争市场情况下原本可以得到内环公寓的某些人。因此，租金管制时，在 p_{max} 价格下租出的公寓数量和竞争性价格下租出的公寓数量是相同的，区别在于租给了不同的人。

1.8 哪种配置方法最好？

目前我们已经分析了四种配置公寓的方法：

- 完全竞争市场。
- 歧视性的垄断者。
- 普通垄断者。
- 租金管制。

这是配置公寓的四种不同经济制度。每种制度都导致不同的结果：得到公寓的人不同，或者对公寓索要的价格不同。你可能会问哪种制度最好。回答这个问题之前我们要先界定“最好”这个词是什么意思。我们比较这些配置方法的标准是什么？

一种评价方法是分析相关人士的经济状况。显然，歧视性垄断情形下，房东的收入最高。类似地，租金管制情形下，房东的收入可能最少。

不同配置方法下租房者的经济状况如何？歧视性垄断情形下，租房者的状况平均来看，可能是最差的。因为他们中的大多数人支付的价格高于其他配置方式下支付的价格。租金管制情形下，租房者的状况变好了吗？有些人的状况变好了：**最终得到公寓的人**的状况比他们

在竞争性市场的情形下好。但是没得到公寓的人的**状况**却比竞争性市场情形下**差**。

因此，我们需要找到一种方法，这种方法能够评价所有相关人士（包括所有租房者和所有房东）的经济状况。站在所有人的角度上，我们如何判断哪种配置方法更受欢迎？站在所有人的角度上，如果你说一种配置方法比另外一种“更好”，你的评价标准是什么？

1.9 帕累托效率

在比较不同经济制度的结果优劣时，可以选择一种称为帕累托效率（Pareto efficiency）或经济效率的概念作为评价标准[一]。我们首先定义帕累托改进(Pareto improvement)这个概念：如果我们能找到一种方法，这种方法能使一些人的状况变好但同时又未使其他人的状况变差，则我们就得到了一个**帕累托改进**。如果某种配置存在帕累托改进，则将这种配置称为**帕累托无效率**（Pareto inefficient）;如果某种配置不存在这样的帕累托改进，则称为**帕累托有效率**（Pareto efficient）。

帕累托无效率的配置是不受欢迎的，因为存在着利己但不损人的改进方法。这种配置方式也许有其他的优点，但帕累托无效率对它给予了致命一击：既然可以设法让一些人的状况变得更好的同时又不损害别人的利益，为什么不这么做？！

帕累托思想是经济学中最重要的思想之一，在以后章节我们将详细研究。这个思想有很多微妙的细节，我们要慢慢分析，但是目前我们仍可以粗略地介绍一下它的思想。请看下面。

假设我们将内环公寓和外环公寓随机分配给租房者，但允许他们互相转租。有些想住在内环的人运气太差，最终被分配到外环。但是这些人可以从不想住在内环但又被分配在内环的人手里重新租得内环公寓。哪些人不想住在内环？认为内环公寓不值那么多租金的那些人。如果内外环公寓是随机分配的，必然有些人想转租公寓，只要他得到的补偿足够高。

例如，假设 A 分得了内环公寓但他认为它只值 200 元租金，B 被分配到了外环，但他愿意以 300 元的价格租入 A 的内环公寓。这种情形下存在着“交易收益（gain from trade）”，因为如果 B 向 A 支付某些租金（200 元到 300 元之间），从而二人交换公寓后，两人的状况都变好了。具体的交易额并不重要。重要的是支付意愿最高的人得到了内环公寓——否则，只要有两人对内环公寓的评价不同，就可能存在着转租的可能。

假设所有的自愿交易都已进行完毕，因此交易的所有收益都被取尽，这种情形下的配置必然是帕累托有效率的。如若不然，就会存在某种交易，这种交易使得两人的状况变得更好但又不损害其他人的利益，可是如果是这样，就意味着交易的收益没有被取尽，与我们的假设矛盾。如果某种配置使得所有自愿交易已进行完毕，则这种配置必然是帕累托有效率的。

[一] 帕累托效率的名字源于 Vilfredo Pareto(1848-1923)，他是 19 世纪的一位经济学家和社会学家。是他首先用这样的思想进行分析。

1.10 不同配置方法的比较

上面描述的交易过程如此粗略，对于它的结果，似乎我们已把能说的话全说完了。事实上，我们还可以多说一点。这一点我们以问题的形式给出：某种配置方法下，如果交易收益已被取尽，谁最终住进了内环公寓？

如果你注意到下列事实，你就能很快得到答案：对于内环公寓，内环的人的保留价格必然高于外环的人的保留价格；否则，他们就会交易从而使两方的状况都变好。因此，如果有 S 个内环公寓用于出租，则保留价格最高的前 S 名租房者最终得到了公寓。这种配置是帕累托有效率的，除此之外所有其他的配置都不是有效率的，因为其他任何配置方式都会存在交易，从而使至少两个人的状况变好，同时又没有损害别人的利益。

我们使用帕累托效率这个评价标准，来比较前面提到的四种资源配置方法的配置结果。假设有 S 个内环公寓用于出租。我们首先分析市场机制。不难知道，市场机制会将这些公寓分配给保留价格最高的前 S 名租房者。这 S 名租房者愿意支付的价格都不低于市场均衡价格 p^*。因此，完全竞争市场均衡时，交易的收益已取尽。完全竞争市场的配置结果是帕累托有效率的。

歧视性垄断者的配置结果如何？它是帕累托有效率的吗？如果你注意到歧视性垄断者将公寓分配给谁，答案就已经明了。前面我们已说过，这种配置方式下，得到公寓的人恰是在完全竞争机制下得到公寓的那些人。在这两种配置方式下，愿意支付不低于 p^* 价格的人都能得到公寓。因此，歧视性垄断者也能实现帕累托有效率的结果。

完全竞争和歧视性垄断者都能实现帕累托有效率的结果，因为在这两种方式下，当人们不再自愿交易时，交易的收益已被取尽。但是，需要注意，这两种配置方式下，收入分配是截然不同的。与完全竞争相比，歧视性垄断情形下：消费者的状况要差得多而房东的状况则要好得多。一般来说，帕累托效率几乎没有涉及到交易收益的分配问题。帕累托效率只与交易的**效率**相关：是否所有可能的交易已经全部进行。

普通垄断者和歧视性垄断者不同，它只能索要同一价格，这种配置方式的结果如何？它**不是**帕累托有效率的。验证这一结论并不难，只要你注意到下列事实：普通垄断者通常只租出部分而不是全部公寓，因此如果他将剩下的公寓以任何正的价格租出，他的利润都会增加。显然，可以找的这样的价格，它使得普通垄断者和租房者的状况都变好。只要该垄断者不改变已租出公寓的价格，这部分租房者的状况就不会改变。但是由于该垄断者将剩下的公寓租出，对他自己以及其他的租房者都有好处，因此我们找到了一个帕累托改进的方法。既然存在着效率改进余地，普通垄断者这种配置方式就不是帕累托有效率的。

最后一种情形即租金管制有效率吗？它也不是帕累托有效率的。我们可以论证如下。如果对租房者随机分配内外环公寓，通常会出现下列局面：某些人（比如 A）得到了内环公寓某些人（比如 B）得到了外环公寓，而且 A 对内环公寓的评价小于 B 的，比如假设 A 的保留价格为 300 元，B 的保留价格为 500 元。

为了证明这种租金管制无效率，我们必须找到一个帕累托改进，即找到一种使 A 和 B

的状况都变好但又不损害其他人利益的方法。但这容易做到：让 A 将内环公寓转租给 B 即可。内环公寓对于 B 来说值 500 元，但对 A 来说只值 300 元。因此 B 支付 400 元给 A 后二人交换公寓，则两人的状况都变好了：B 对内环公寓的评价为 500 元，但他只花 400 元就得到了内环公寓；A 得到了 400 元，但他对内环公寓的评价仅为 300 元。

1.11 长期均衡

我们已经分析了短期（short run）内的公寓均衡价格，短期意味着公寓的供给是固定不变的。但在长期（long run），公寓的供给可以改变。正如需求曲线描述不同价格水平下的公寓需求量一样，供给曲线衡量不同价格水平下的公寓供给量。市场价格的最终决定因素是供给曲线和需求曲线何时相交。

那么供给行为的决定因素是什么？一般来说，自由市场供给新的公寓的数量取决于这么做能赚多少钱。能赚多少钱又部分取决于房东索要的价格为多大。为了分析长期情形下的公寓市场的行为，我们必须同时分析供给者的行为和需求者的行为，这一任务我们留待后面章节解决。

当供给可变时，我们可以问的问题变多了，除了可以继续问谁最终得到了公寓之外，我们还可以问在不同类型的市场制度情形下，供给数量有何不同。与完全竞争相比，垄断情形下的供给数量是增加还是减少了？租金管制会增加还是减少公寓的均衡数量？用哪种市场制度供给的公寓是帕累托有效率的？为了回答这些问题以及与这些问题类似的其他问题，我们必须开发更系统更有力的工具，以进行经济分析。

总结

1.通常从构建社会现象的模型开始进行经济分析，这样的模型是现实的简化表示。

2.在经济分析的任务中，经济学家通常遵循最优化原则和均衡原则，前者是说人们通常尽力选择最好的东西，后者是说价格将一直调整直至需求和供给相等时为止。

3.需求曲线描述在每个价格水平上人们愿意需求的数量，供给曲线描述在每个价格水平上的供给数量。均衡价格是需求量等于供给量时的价格。

4.比较静态分析是研究当相关基本条件改变时，均衡价格和均衡数量如何改变。

5.某种经济状态是帕累托有效率的，如果不可能设法使一些人的状况变得更好但同时又不损害其他人的利益。帕累托效率这个概念可用于评估资源不同配置方式。

复习题

1.假设有 25 个人的保留价格都为 500 元，第 26 个人的的保留价格为 200 元。这种需求曲线的形状如何？

2.在上题中，如果只有 24 个公寓出租，均衡价格为多少？如果有 26 个公寓出租呢？如果有 25 个公寓出租呢？

3.如果人们的保留价格不同，为何需求曲线向下倾斜？

4.在教材中我们假设产权出售的公寓（condominium）的购买者来自内环，即已经住在内环公寓（apartment）的人。现在假设上述公寓（condominium）的购买者来自外环，即住在外环的人，内环公寓的租金价格如何变化？

5.假设现在产权出售的公寓（condominium）的购买者全部来自内环，但每个产权出售的公寓（condominium）是由两个公寓（apartments）改成。内环公寓的价格如何变化？

6.在长期中，税收将对公寓的数量产生何种影响？

7.假设需求曲线为 $D(p)=100-2p$.如果垄断者有 60 个公寓，他将如何定价？他将租出多少公寓？如果他有 40 个公寓，他将如何定价？他将租出多少公寓？

8.如果我们的租金管制模型允许不受限制地转租，谁最终获得内环的公寓？这个结果是帕累托有效率的吗？

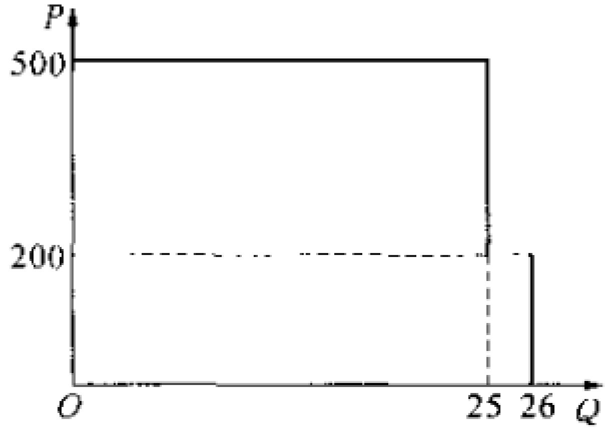
复习题答案

1.假设有 25 个人的保留价格都为 500 元，第 26 个人的的保留价格为 200 元。这种需求曲线的形状如何？

【复习内容】保留价格；需求曲线

【参考答案】

需求曲线衡量在不同价格上人们对某商品的购买数量（在此例中为租住的公寓数量）。模仿教材图 1.1，根据本题题意可画出如左图的需求曲线。



2.在上题中，如果只有 24 个公寓出租，均衡价格为多少？如果有 26 个公寓出租呢？如果有 25 个公寓出租呢？

【复习内容】均衡价格

【参考答案】

首先请读者分别根据题目已知条件，在上图中补充 S=24，S=25，S=26 的三条供给曲线，这三条供给曲线都是与横轴垂直的。

由于需求和供给相等（需求曲线和供给曲线相交）时的价格为均衡价格，因此，当 S=24 时，需求曲线和供给曲线相交时的价格为 500 元，因此 S=24 时，均衡价格为 500 元。当 S=25 时，供给曲线和需求曲线出现了部分重合，重合段对应的价格即 200 元（含）到 500 元（含）之间的价格都可能是均衡价格；当 S=26 时，供给曲线和需求曲线又出现了部分重合，重合段对应的价格即 0 元（含）到 200 元（含）的价格都可能是均衡价格。

3.如果人们的保留价格不同，为何需求曲线向下倾斜？

【复习内容】保留价格；需求曲线的斜率为负

【参考答案】

保留价格是对于某商品某人愿意支付的最高价格。保留价格不是实际支付价格，实际支付价格还要取决于卖方（供给方）索要多少。

假设卖方索要的价格为 P，不妨一开始把它想象为天价，也就是说所有消费者的保留价格都小于 P，此时需求量为 0；然后我们令卖方逐渐降低价格，当降低到比如 P_1 时，恰好有一个消费者的保留价格大于 P_1，则市场需求量就等于这个人的需求量；当商品价格持续降低时，此人可能会继续购买，另外，越来越多的人的保留价格开始大于该商品价格，因此越来越多的人开始购买该商品。从这个过程可以看出，随着商品价格下降，商品的需求量增加，这意味着需求曲线是向下倾斜的。

4.在教材中我们假设产权出售的公寓（condominium）的购买者来自内环，即已经住在内环公寓（apartment）的人。现在假设上述公寓（condominium）的购买者来自外环，即住在外环的人，内环公寓的租金价格如何变化？

【复习内容】供给变动（供给曲线移动）；均衡价格

【参考答案】

如果部分内环公寓被改变为产权出售的公寓，并且卖给了外环的人，这意味着内环公寓的供给减少（供给曲线向左移动），但内环公寓的需求不变（需求曲线位置固定），所以价格会上升。当然如果还有其他原因使内环公寓的需求也减少，则公寓的均衡价格难以判断，需要根据供给还是需求减少的幅度判断。请读者自行补充供给曲线和需求曲线。

5.假设现在产权出售的公寓（condominium）的购买者全部来自内环，但每个产权出售的公寓（condominium）是由两个公寓（apartments）改成。内环公寓的价格如何变化？

【复习内容】需求变动（需求曲线的移动）；供给变动（供给曲线的移动）；市场均衡

【参考答案】

内环公寓的价格会提高。因为如果部分内环公寓被改变为产权出售的公寓，并且卖给了内环的人，前半句话意味着供给减少（供给曲线向左移动），后半句话意味着需求也减少了（需求曲线向左移动）。但由于 2 个公寓才能改变为一个产权出售的公寓，这表明供给减少幅度大于需求减少幅度（供给曲线向左移动的距离大于需求曲线向左移动的距离），因此均衡价格会上升。请读者自行补充需求曲线和供给曲线。

6.在长期中，税收将对公寓的数量产生何种影响？

【复习内容】短期和长期的概念

【参考答案】

注意区分经济学中“短期”与“长期”的概念。

短期内，公寓的供给曲线为一条垂线。而在长期内，公寓的供给曲线是向右上方倾斜的，因为开发商可以建造新的公寓。

征税会提高公寓的成本，因此供给减少（即供给曲线向左移动），这种情形下均衡数量会下降。请自行补充供给曲线和需求曲线。

7.假设需求曲线为 $D(p)=100-2p$.如果垄断者有 60 个公寓，他将如何定价？他将租出多少公寓？如果他有 40 个公寓，他将如何定价？他将租出多少公寓？

【复习内容】垄断

【参考答案】

根据题意可知，垄断者的收入为 $pD(p)=p(100-2p)=100p-2p^2$；同样，沿袭教材中的假设，可假设垄断者出租公寓的成本为 0，因此上述表达式就是垄断者的利润表达式。（在以后章节我们会学习到，厂商的问题就是利润最大化问题）。

配方可知，p=25 时，他的利润最大。将 p=25 代入需求函数 $D(p)=100-2p$ 可知他供给 50 个公寓时利润最大。因此，如果他有 60 套公寓，他只会租出 50 套。

如果他有 40 套公寓，他会全部租出，将 D（P）=40 代入需求函数 $D(p)=100-2p$ 可知价格为 30.

8.如果我们的租金管制模型允许不受限制地转租，谁最终获得内环的公寓？这个结果是帕累托有效率的吗？

【复习内容】租金管制；帕累托效率

【参考答案】

我们在教材中指出租金管制是帕累托无效率的，那里的假定是不允许转租。但如果允许转租，还是无效率的吗？我们分析这个问题。

在租金管制的情形下，如果允许转租，那么最终保留价格排在前面的人得到了内环公寓。为什么？假设内外环的公寓随机分配，通常出现这样的局面，某些人（统称 A）想住内环却被分配到了外环，某些人（统称 B）被分配到了内环却认为内环的公寓不值那么多租金，通常 A 的保留价格会大于 B 的，因此 A 和 B 互相向对方转租公寓，当然 A 要补偿 B 一些钱。由此可见，保留价格较高的人最终得到了内环公寓。

最终的结果是帕累托有效率的，为什么？因为只要市场还存在互利的空间，转租将一直进行下去，直至交易的收益都被取尽，根据定义，这显然是帕累托有效率的。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

2.预算约束（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

2 预算约束

消费者的经济理论很简单：消费者们选择他们能够买得起的最佳的商品组合。为了说明该理论，需要更准确地阐述上句中 “最佳”以及 “能够买得起”的含义。本章将阐述如何描述消费者能够买得起：下一章则重点分析消费者如何确定什么样的商品组合为最佳。这样我们就可以详细研究消费者行为这一简单模型的意义。

2.1 预算约束

我们从**预算约束**(budget constraint)的概念入手分析。假设消费者可从某组商品中进行选择。在现实生活中有许多商品可供选择，但根据我们的目的，最好只考虑两种商品的情形，因为我们可以借助图形刻画消费者的选择。

我们将消费者的**消费束**(consumption bundle)用(x_1, x_2)表示。(x_1, x_2)是一个有序数对，它表明消费者选择消费商品 1 的数量（x_1）和商品 2 的数量（x_2）。有时用单个符号 X 来表示消费束更方便， X 为(x_1, x_2)的缩写。

假设两种商品的价格(p_1, p_2)和消费者拥有的钱数m都可知，则消费者的预算约束可以写为

$$p_1x_1 + p_2x_2 \le m \tag{2.1}$$

其中：p_1x_1是消费者花费在商品 1 上的钱数，p_2x_2是消费者花费在商品 2 上的钱数。消费者的预算约束要求花费在两种商品上的总钱数不超过消费者拥有的钱数。消费者可以**买得起**的消费束是指那些花费不超过m的消费束。我们把在价格(p_1, p_2)和收入m下可以买得起的消费束称为消费者的**预算集**(budget set)。

2.2 用两种商品进行分析通常已足够

用两种商品进行分析已具有足够的代表性，这可能超出你最初的想象。事实上，我们总可以将两种商品中的一种解释为，除了我们关注的某种特定商品之外的其余所有想要消费的商品。

例如，如果我们的兴趣在于研究消费者对牛奶的消费，令x_1表示他每月消费的牛奶数量，则可令x_2表示他想要消费的其他所有商品。

若接受上述解释，则可方便地将商品 2 想象为花费在其他商品上的货币数量，这种商品的价格自然为 1，因为一元的价格当然为一元。此情形下预算约束将变形为:

$$p_1x_1 + x_2 \le m \tag{2.2}$$

曹乾（东南大学 caoqianseu@163.com）

这个表达式只是说，花费在商品 1 上的钱数（p_1x_1）与花费在其他所有商品上的钱数（p_2x_2）之和，不能超过消费者的收入（m）。

我们说商品 2 表示一种**复合商品**（composite good），它代表除了商品 1 之外的所有其它消费品。我们总是用除了商品 1 之外其余所有商品的支出金额表示这种复合商品。从预算约束的代数表达式角度看，（2.2）式只是（2.1）式的一种特种情形，此时 $p_2=1$，因此（2.1）式的一切性质都适用于（2.2）式。

2.3 预算集的性质

预算线（budget line）是指将消费者的收入 m 恰好花完的那些消费束的集合：

$$p_1x_1+p_2x_2=m \tag{2.3}$$

预算集可用图 2.1 表示。粗实线代表预算线，即花费恰好等于 m 的消费束的集合。预算线下方的消费束，它们的花费都严格小于 m。

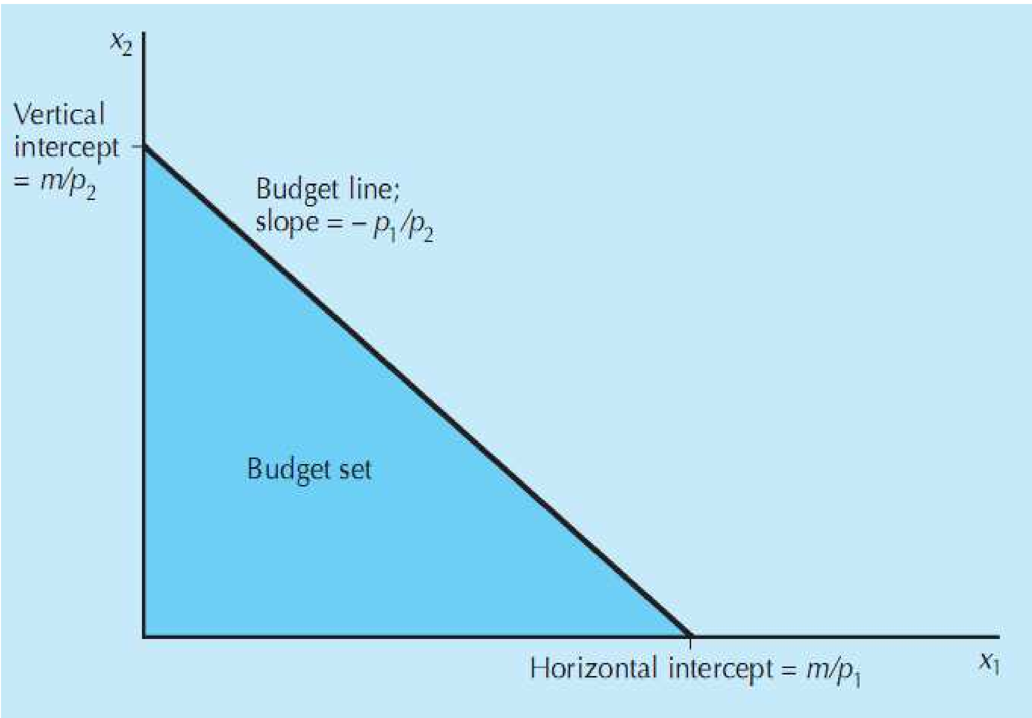


图 2.1：预算集。预算集包含了所有在既定价格和收入水平下能够买得起的消费束。

将预算线表达式(2.3)变形可得：

$$x_2=\frac{m}{p_2}-\frac{p_1}{p_2}x_1. \tag{2.4}$$

这是一条斜率等于$-p_1/p_2$且纵截距为m/p_2的直线。上式表示，如果消费者消费x_1单位的商品 1，他应该消费多少单位商品 2 才能使恰好满足预算线。

给定价格(p_1, p_2)和收入m，画出预算线并不难。首先，计算若将所有收入都用于购买商品 2 所能购买到的数量，答案是$\frac{m}{p_2}$。其次，计算若将所有收入都用于购买商品 1 所能购买到的数量，答案是$\frac{m}{p_1}$。

因此，横截距m/p_1和纵截距m/p_2，分别表示如果消费者将全部收入分别用于购买商品 1 和商品 2 所能购买到的数量。将这两个截距在坐标轴上标注出来，用一条直线连接它们就得到了预算线。

预算线的斜率有个美妙的经济学解释。它衡量市场中商品 1 替代商品 2 的比率。假设消费者打算增加商品 1 的消费量，增加量用Δx_1表示[一]。在这种情形下，商品 2 的消费量如何变化才能恰好满足预算约束？令Δx_2表示商品 2 消费量的变化量。

由于消费量变动前后，他的消费都满足预算约束，因此必然有：

$$p_1x_1 + p_2x_2 = m \quad 和$$

$$p_1(x_1+\Delta x_1) + p_2(x_2+\Delta x_2) = m.$$

第二个式子减去第一个式子可得

$$p_1\Delta x_1 + p_2\Delta x_2 = 0.$$

这就是说他的消费变动的价值之和等于 0。从上式解出$\Delta x_2/\Delta x_1$可得

$$\frac{\Delta x_2}{\Delta x_1} = -\frac{p_1}{p_2}$$

这个式子给出了在满足预算约束的前提下，用商品 1 替代商品 2 的比率。

这正是预算线的斜率。该式的符号为负，因为Δx_1和Δx_2变动的方向相反。如果你多消费（少消费）商品 1，你必须少消费（多消费）商品 2，才能继续满足预算约束。

经济学家有时会说预算线的斜率衡量了消费商品 1 的**机会成本**（opportunity cost）。为了多消费一些商品 1，你必须放弃商品 2 的一些消费。放弃消费商品 2 的机会是多消费商品 1 的真实经济成本，这样的经济成本以预算线的斜率衡量。

2.4 预算线的变动

[一] Δ为希腊字母，读作“德尔塔”。Δx_1表示商品 1 的变化量。本书数学附录部分列出了更多的变动符号和变动比率符号。

当商品价格和消费者收入变动时，消费者能够买得起的商品集也会变动。这样的变动对预算集有何影响？

我们首先分析收入的变化。从（2.4）式易知，收入增加会使纵截距变大，但不会影响预算线的斜率。因此，收入增加会使预算线**向外平行移动**，如图 2.2 所示。类似地，收入下降会使预算线向内平行移动。

商品价格变动会使预算线怎样变动？我们先分析商品 1 价格变动但商品 2 的价格和收入不变的情形。根据（2.4）式可知，p_1 上升不会改变纵截距，但会使预算线更陡峭，因为 p_1/p_2 变得更大。

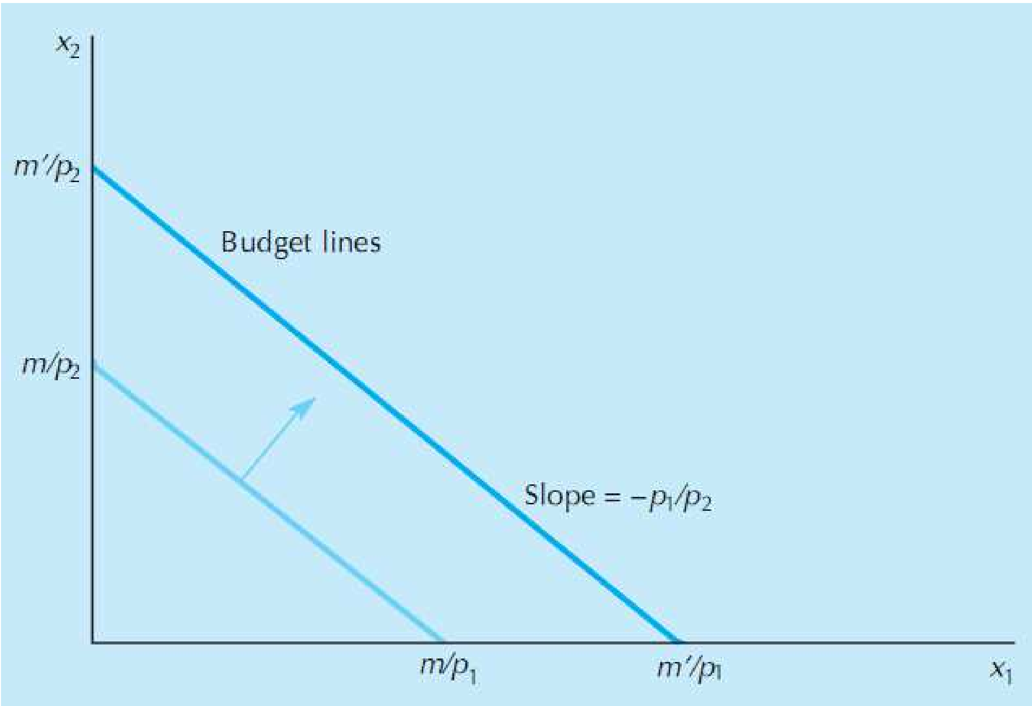


图 2.2：收入增加。收入增加会使预算线向外平行移动。

我们也可以使用前面介绍的预算线的画法技巧，分析预算线如何变动。如果你将所有的钱都花费在商品 2 身上，因此商品 1 价格上升，不会改变你能购买到的商品 2 的最大数量，于是预算线的纵截距不变。但是如果你将所有的钱都用于购买商品 1，由于商品 1 变得更昂贵，因此商品 1 的最大购买量必然下降，于是预算线的横截距必然减小，即预算线向内移动，如图 2.3 所示。

如果商品 1 和 2 的价格同时变动，预算线会如何变动？例如，假设商品 1 和 2 的价格都变为原来的 2 倍但收入不变，在这种情形下，横纵截距都变为原来的一半，因此预算线向内平行移动。收入不变但商品价格同乘以 2，等价于价格不变但收入除以 2。

上述结论也可用代数证明。假设初始预算线为

$$p_1x_1 + p_2x_2 = m.$$

现在假设价格变为原来的t倍，将两种商品的价格同时乘以t可得

$$tp_1x_1 + tp_2x_2 = m.$$

但是这个式子等价于下式

$$p_1x_1 + p_2x_2 = \frac{m}{t}.$$

因此，两种商品的价格同乘以一个常数t，等价于收入除以这个常数t。由此还可以知道，如果我们将两种商品的价格以及收入同乘以t，则预算线不变。

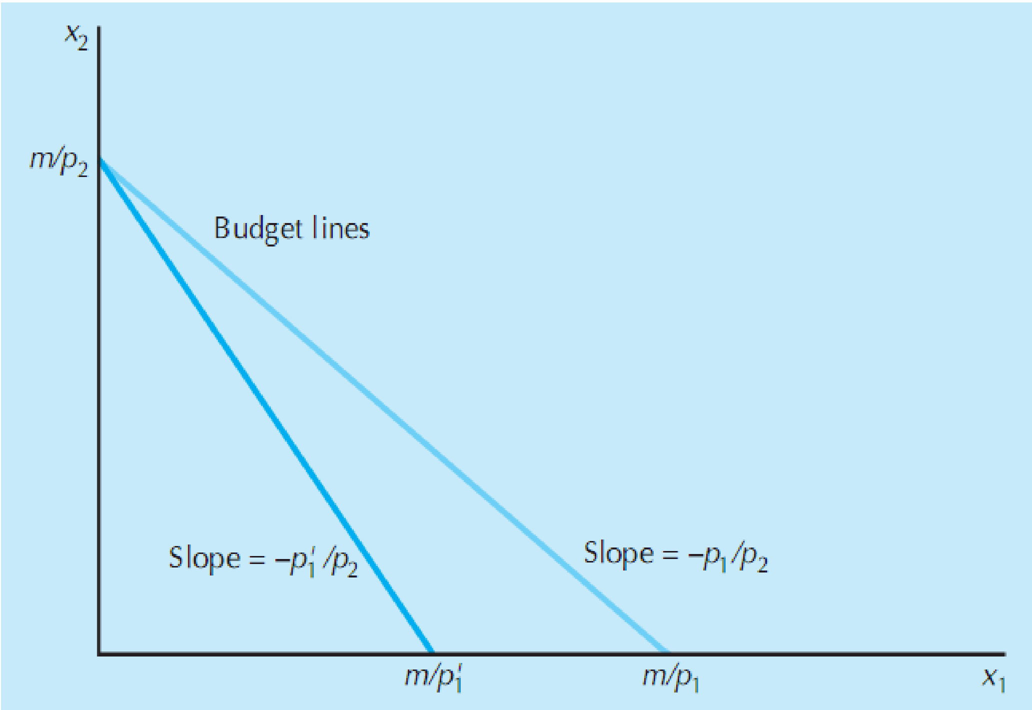


图 2.3：一种商品价格上升。如果商品 1 价格上升，预算线会变得更陡峭。

我们也可以考虑价格和收入同时变动的情形。当两种商品价格上升但收入下降时，预算线怎样变动？首先分析横纵截距如何变动。如果m下降但p_1和p_2上升，则横截距m/p_1和纵截距m/p_2都减小。因此，预算线会向内移动。预算线的斜率会如何变化？如果商品 2 的价格上升幅度比商品 1 的大，则斜率（$-p_1/p_2$）的绝对值会减小，即预算线会变得更平坦；如果商品 2 的价格下降幅度比商品 1 大，则预算线会更陡峭。

2.5 计价物

我们用两种商品价格和消费者的收入定义预算线，但这三个变量中有一个是多余的。我们可以将其中一种价格或者收入固定在既定水平上，调整其他两个变量也可以得到原来的预算集。于是，预算线

$$p_1x_1 + p_2x_2 = m \text{ 和}$$

$$\frac{p_1}{p_2}x_1 + x_2 = \frac{m}{p_2} \text{ 或}$$

$$\frac{p_1}{m}x_1 + \frac{p_2}{m}x_2 = 1$$

是同一条预算线的不同表达形式，因为第二式只是第一式左右两侧同除以 p_2，第三式是第一式左右两侧同除以 m。在第二个式子中，我们将 p_2 固定为 1；在第三个式子中，我们将 m 固定为 1。将其中一种商品价格或者收入固定为 1，相应调整其他商品的价格或收入不会改变预算集。

当我们将其中一种商品的价格设定为 1，正如上面我们所做的一样，我们通常将这种商品的价格称为**计价物价格或本位价格**（numeraire price）。计价物价格是我们选定某商品的价格作为基准价格，用来衡量其他商品的价格以及衡量收入。有时，将一种商品视为计价物将使分析变得简单，因为我们减少了一个价格变量。

2.6 税收、补贴和配额

在制定经济政策时，政府通常使用一些政策工具（比如税收）影响消费者的预算约束。例如，如果政府征收**从量税**（quantity tax），即消费者购买每单位某种商品时都要向政府缴纳一定数量的税收。在美国，消费者每消费 1 加仑汽油需要向联邦政府缴纳 15 分钱的汽油税。

从量税如何影响消费者的预算线？从消费者的角度来看，税收相当于商品的价格提高了。因此，对每单位商品 1 征收 t 元钱的从量税，会使商品 1 的价格从 p_1 变为 $p_1 + t$。我们已经知道，这意味着预算线将变得更陡峭。

另外一种税收叫做**从价税**（value tax；ad valorem tax）。顾名思义，这种税对销售商品的价格而不是购买的数量征税。从价税的税率通常以百分比表示。美国大多数州都征收销售税。如果销售税的税率为 6%，则价格为 1 元的商品的售价为 1.06 元。

如果商品 1 的价格为 p_1，但要缴纳税率为 τ 的销售税，则消费者面对的实际价格为 $(1+\tau)p_1$ [一]。消费者每购买一单位该商品，需要向卖方支付 p_1 元并且向政府缴纳 τp_1 元，因此购买一单位商品 1，消费者的总支出为 $(1+\tau)p_1$。

[一] τ 是希腊字母，读作“tau”。

补贴（subsidy）正好和税收相反。在**从量补贴**（quantity subsidy）的情形下，政府根据消费者购买某种商品的数量，**给与**他一定金额的补助。例如，如果喝牛奶有补贴，则政府需要向消费牛奶的每个消费者支付补贴，补贴金额取决于消费者消费牛奶的数量。如果每消费一单位商品 1 可以获得 s 元的补贴，则在消费者的眼里，商品 1 的价格变为 $p_1 - s$。这将会使预算线变得更平坦。

类似地，从价补贴是按照商品的价格计算补贴数量。如果你向慈善机构每捐献 2 元，政府返还你 1 元，则补贴率为 50%。更一般地，如果商品 1 的价格为 p_1，从价补贴率为 σ，则消费者面对的实际价格为 $(1-\sigma)p_1$[一]。

你已经看到税收和补贴对价格的影响方式是相同的，唯一不同的是代数符号：税收提高了商品价格而补贴则降低了商品价格。

另外一种类型的税收（或补贴）叫做**定额** (lump-sum)税（或补贴）。在定额税的情形下，政府对每个消费者征收一笔数额固定的税收，这种税收和消费者是否消费某种商品无关[二]。因此，征收定额税后，消费者的预算线向内平移，因为他的收入被拿走了一部分。类似地，定额补贴会使消费者的预算线向外平移。从量税和从价税会使预算线变得更陡峭或更平坦，这取决于对商品 1 还是商品 2 征税；但是定额税必然会使预算线向内平移。

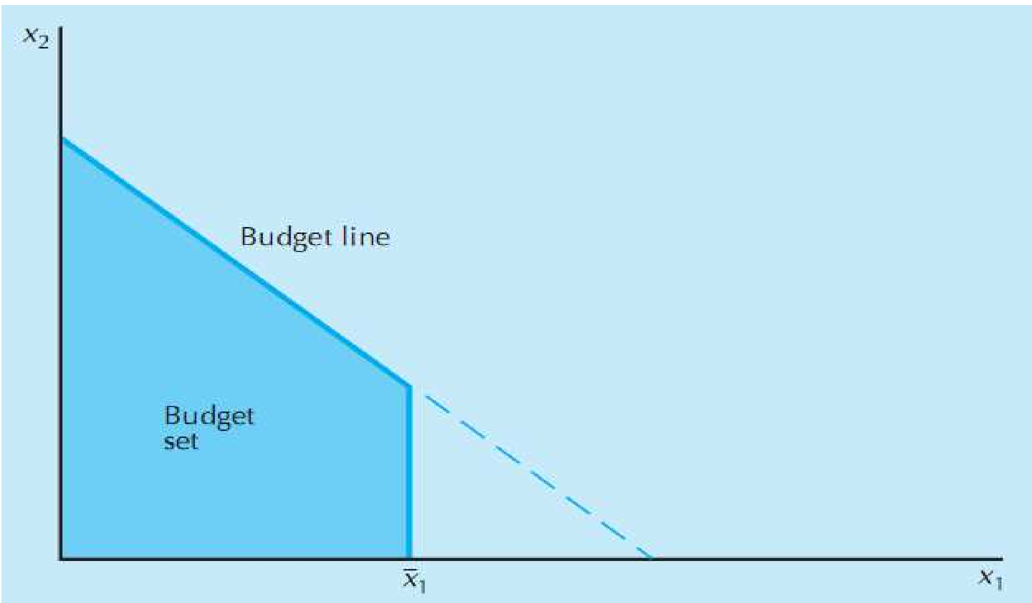


图 2.4：定量配给情形下的预算线。如果商品 1 是定量配给的，则超过配额的那部分预算集被砍掉了。

[一] σ 为希腊字母，读作“sig-ma”。
[二] 因此，定额税又称为人头税。译者注。

政府有时还会实施**配额**(rationing)约束或称为定量配给约束。这表示对某商品的消费水设立上限。例如，在第二次世界大战期间，美国政府对黄油和肉类等食品实施配额制度。

例如，假设商品 1 是定量配给的，每个消费者对商品 1 的消费量都不能超过 $\bar{x}_1$。此时消费者的预算集具有图 2.4 所示的形状：原来的预算集被砍掉了一块。也就是说所有满足 $x_1 > \bar{x}_1$ 的消费束都被砍掉了。

政府有时会将税收、补贴和配额联合使用。例如，消费者消费商品 1 时，他可以 p_1 的价格购买直到 $\bar{x}_1$ 的数量，如果他想继续购买，则每多购买一单位需要缴税 t 元。此时预算线的形状如图 2.5 所示。$\bar{x}_1$ 左边的预算线的斜率为 $-p_1/p_2$；$\bar{x}_1$ 右边的预算线的斜率为 $-(p_1+t)/p_2$。

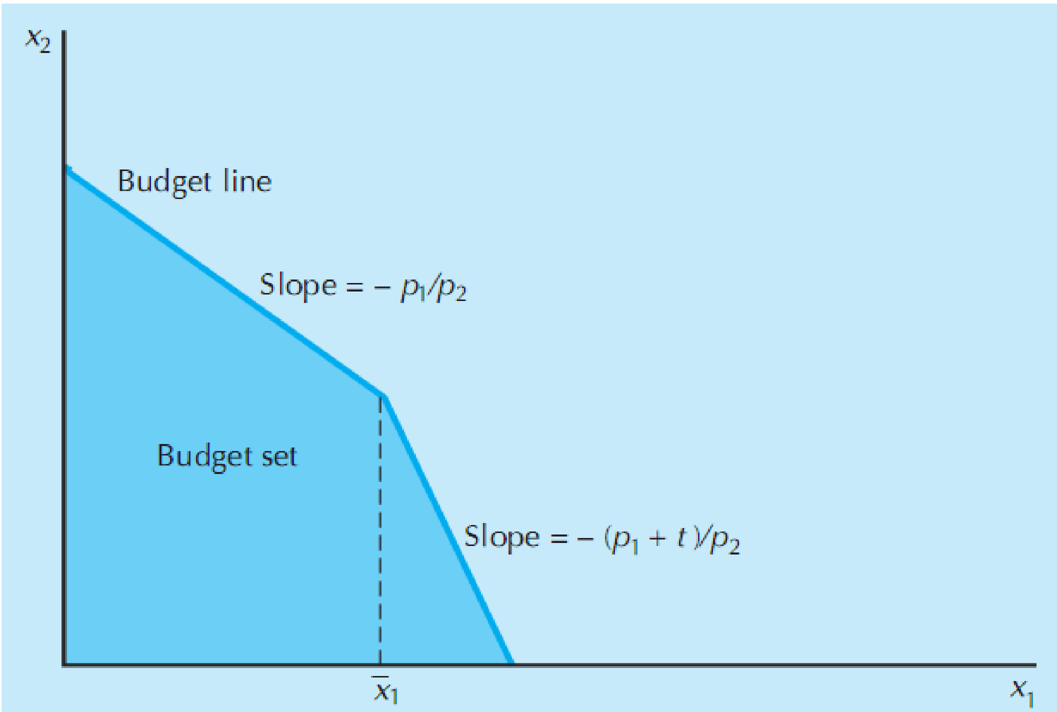


图 2.5：对超过配额 $\bar{x}_1$ 部分的消费征税。在这个预算集内，消费者需要对对超过配额 $\bar{x}_1$ 的那部分消费缴税，因此 $\bar{x}_1$ 右边的预算线更陡峭。

例子：食品票方案

1964 年美国联邦政府实施了食品票法案，对穷人消费食品给与补贴。这个法案曾调整过几次。我们分析其中一种调整的经济效应。

1979 年之前，穷人家庭可以购买食品票，然后用食品票在零售店购买食品。例如，在 1975 年 1 月，一个四口人的穷人家庭每个月最多可以得到价值 153 美元的食品优惠券。

优惠券的价格取决于家庭收入。一个四口之家，如果经过调整后的月收入为 300 美元，

则需要花 83 美元购买月度食品票，如果月收入为 100 美元，则只要花 25 美元购买[一]。

1979 年之前的食品票方案是一种从价补贴。食品的补贴率取决于家庭收入。花 83 美元购买整月食品票的四口之家，每支付 1 美元可以得到价值 1.84 美元的食品（1.84=153/83）。类似地，花 25 美元购买整月食品票的四口之家，每支付 1 美元可以得到价值 6.12 美元的食品（6.12=153/25）。

食品票方案对预算线的影响方式请见图 2.6A。此处我们用横轴表示消费者的食品支出，用纵轴表示他消费所有其他商品的支出。由于我们用支出的钱数衡量商品的数量，因此每种商品的“价格”自然为 1，因此预算线的斜率为－1。

如果政府允许某个家庭以 25 美元购买价值 153 美元的食品票，那么政府给与此家庭购买食品的补贴越为 84%（=1－25/153），因此预算线的斜率约为－0.16[=－（25/153）]，直到此家庭将这 153 美元的食品票用完。在该食品票用完之前，此家庭在食品上每花费 1 美元将减少 16 美分的其他商品的消费。在该食品票用完之后，预算线的斜率又重新变为－1。

这些效应使预算线出现了“折弯”（kink），如图 2.6 所示。家庭的收入越高，购买食品票的支出越多。因此，家庭收入增加时，预算线会变得更陡峭。

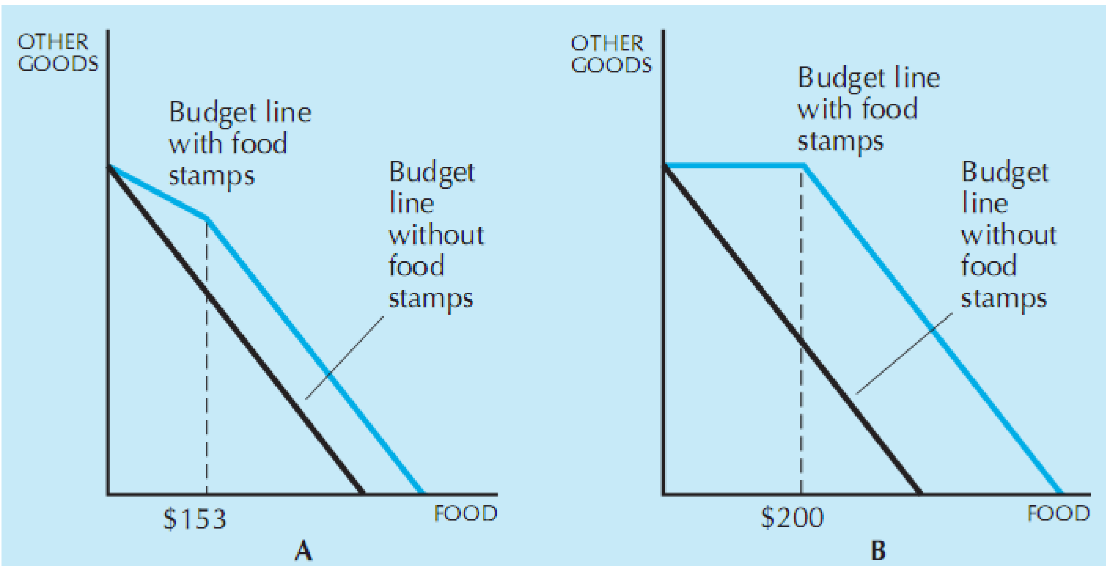


图 2.6：食品票。食品票方案如何影响预算线？A 图表示的是 1979 年之前的方案，B 图表示的是 1979 年之后的方案。

在 1979 年，政府修改了食品票方案。修改后的方案不再要求家庭购买食品票，而是对合格的家庭免费发放。图 2.6B 显示了修改后方案对预算集的影响。

假设某个家庭拿到了 200 美元的食品票。这意味着该家庭每个月可多消费 200 美元的食品，这个数量显然和该家庭消费在其他商品上的钱数无关。由此可知，预算线将向外移动，

[一] 资料来源： Kenneth Clarkson, Food Stamps and Nutrition, American Enterprise Institute, 1975.

移动距离为 200 美元。预算线的斜率不变，因为减少 1 美元的食品消费就可以增加 1 美元其他商品的消费。但是由于该家庭不能合法出售食品票，他消费其他商品的最大数量不变。食品票方案实质上是一种定额补贴，只不过食品票不能出售。

2.7 预算线变动

在下一章我们将分析消费者如何从他的预算集内选择最优的消费束。但在此之前，借助本章学习到的预算线移动知识，我们可以推导出一些结论。

首先，由于当所有的价格和收入同乘以一个正数不会改变预算集，最优消费束也不会变动。无需分析消费者的具体选择过程，我们就已得出了一个重要的结论：完全平衡的通货膨胀（即所有商品价格和收入都按相同比率上升），不会改变任何人的预算集，因此也不会改变任何人的最优选择。

其次，我们可以说明消费者在不同价格和收入水平下的福利情况。假设消费者的收入增加并且所有商品的价格不变。我们知道这表示预算线向外平移。因此，在较低收入水平下的任何消费束，在收入增加之后，自然仍然还能买得起。这意味着消费者在较高收入水平下的状况，不会比在较低收入水平下的状况差。原因在于，收入增加后，他的选择不仅包含原来的所有消费束，而且还能购买一些新消费束。类似地，如果一种商品价格下降但其余变量不变，消费者的状况也不会比原来的状况差。这个简单的结论作用可不小，我们在后面章节将会用到它。

总结

1.预算集包含了在给定价格和收入水平下消费者能买得起的所有消费束。我们通常假设只有两种商品，这种假设具有较高的代表性。

2.预算线的表达式为 $p_1x_1+p_2x_2=m$，它的斜率为 $-p_1/p_2$，横截距为 m/p_1，纵截距为 m/p_2。

3.收入增加会使预算线向外平行移动。商品 1 的价格上升会使预算线更陡峭，而商品 2 的价格上升会使预算线更平坦。

4.税收、补贴和配额改变了消费者支付的商品价格，从而改变了预算线的斜率和位置。

复习题

1.消费者的初始预算线为 $p_1x_1+p_2x_2=m$。现在商品 1 的价格变为原来的 2 倍，商品 2 的价格变为原来的 8 倍，收入变为原来的 4 倍。写出新预算的表达式，其中价格和收入要分别以原来的价格和收入表示。

2.如果商品 2 的价格上升，但商品 1 的价格和收入保持不变，那么预算线如何变动？

3.如果商品 1 的价格变为原来的 2 倍，商品 2 的价格变为原来的 3 倍，预算线变得更平坦还是更陡峭？

4.计价物的定义是什么？

5.假设政府最初对每加仑汽油征税 15 分钱，后来又决定对每加仑汽油给与 7 分钱的补贴。这种税收和补贴联合使用的方式等价于对每加仑汽油征多少税？

6.假设预算线方程为 $p_1x_1+p_2x_2=m$。现在政府决定：对消费者征收定额税，税额为 u；对每单位商品 1 征收 t 元从量税；对每单位商品 2 给与 s 元补贴。求新预算线的表达式。

7.如果消费者的收入增加，同时其中一种商品价格下降，消费者的状况必然至少与原来一样好吗？

复习题答案

1.消费者的初始预算线为 $p_1x_1+p_2x_2=m$。现在商品 1 的价格变为原来的 2 倍，商品 2 的价格变为原来的 8 倍，收入变为原来的 4 倍。写出新预算的表达式，其中价格和收入要分别以原来的价格和收入表示。

【复习内容】预算线；预算线的变动

【参考答案】

$$2p_1x_1+8p_2x_2=4m$$

2.如果商品 2 的价格上升，但商品 1 的价格和收入保持不变，那么预算线如何变动？

【复习内容】预算线；预算线的变动

【参考答案】

预算线绕着横截距点$(m/p_1,0)$向内转动（变得更平坦）。

设初始预算线方程为$p_1x_1+p_2x_2=m$，它的斜率为$-p_1/p_2$，横截距为m/p_1，纵截距为m/p_2。

当p_2上升而p_1和m不变时，则斜率绝对值变小，横截距不变，总截距变小，因此预算线绕着横截距点$(m/p_1,0)$向内转动（变得更平坦）。

3.如果商品 1 的价格变为原来的 2 倍，商品 2 的价格变为原来的 3 倍，预算线变得更平坦还是更陡峭？

【复习内容】预算线；预算线的变动

【参考答案】

设初始预算线方程为$p_1x_1+p_2x_2=m$，价格变动后得到的新预算方程为$2p_1x_1+3p_2x_2=m$，新预算线的斜率=$-2p_1/3p_2=2/3(-p_1/p_2)$，即新预算线的斜率为原预算线斜率的 2/3，斜率绝对值变小了，所以预算线变得更平坦。

4.计价物的定义是什么？

【复习内容】计价物

【参考答案】

计价物是用来衡量其他商品价格的某一商品，这样的商品可以任意选取，但要令它的价格为 1，从而可以计算出其他商品相对于这种商品的价格。例如商品 A 的价格为 2 元，商品 B 的价格为 4 元，如果选择商品 A 作为计价物，则商品 B 的相对价格为 2；类似地，如果选择商品 B 作为计价物，则商品 A 的相对价格为 1/2。

5.假设政府最初对每加仑汽油征税 15 分钱，后来又决定对每加仑汽油给与 7 分钱的补贴。这种税收和补贴联合使用的方式等价于对每加仑汽油征多少税？

【复习内容】征税或补贴对预算线的影响

【参考答案】

站在消费者的角度，征收消费税等价于价格提高，给他补贴等价于价格下降，因此政府征税 15 分钱后来又补贴 7 分钱，相当于政府的净税为 8 分钱。

6.假设预算线方程为 $p_1x_1+p_2x_2=m$ 。现在政府决定：对消费者征收定额税，税额为 u ；对每单位商品 1 征收 t 元从量税；对每单位商品 2 给与 s 元补贴。求新预算线的表达式。

【复习内容】征税或补贴对预算线的影响

【参考答案】

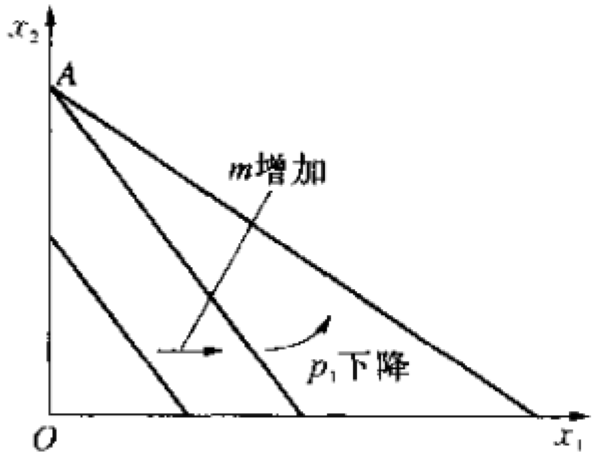
站在消费者的角度，征收定额税相当于收入减少，征收从量税相当于价格提高，给他补贴等价于价格下降，因此根据题目的条件可知，新预算线的表达式为 $(p_1+t)x_1+(p_2-s)x_2=m-u$ 。

7.如果消费者的收入增加，同时其中一种商品价格下降，消费者的状况必然至少与原来一样好吗？

【复习内容】预算集；预算集的变化；消费者的状况

【参考答案】

消费者收入增加后，预算线向外平行移动；不妨假设商品 1 价格下降，此时预算线绕着纵截距点向外转动。如下图所示。



预算线变动后，新的预算集包含原来的预算集，不仅如此，新的预算集还增加了一些新的消费选择，因此消费者的状况至少与原来一样好。事实上，我们后面章节会学习到，这种情形下，消费者的状况一定比原来好。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

3.偏好（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

3 偏好

在第 2 章我们已看到，消费者行为的经济模型很简单：人们在能购买得起的商品束中选择**最优的消费束**。上一章说明了“能够买得起”的含义，本章则旨在说明“最优商品束”的概念。

我们把消费者选择的东西称为**消费束**（consumption bundles）。消费束是我们研究选择问题中涉及到的**全部**商品（或服务）。“全部”二字值得强调：当分析消费者选择问题时，一定要将涉及到的商品全部包含在消费束的定义中。

如果我们是在最宽泛的水平上研究消费者选择问题，我们不仅需要知道消费者可能消费的所有商品，还需要知道消费的时间，地点以及在什么样情形下消费的。毕竟人们不仅关心今天的食物数量，还关心明天的食物数量。大西洋里的小船和撒哈拉沙漠里的同样的小船，意义是不同的，类似的还有晴天的雨伞与阴雨天的雨伞。通常将不同场所或环境中的“同种”商品视为不同的商品，因为消费者在这样的情形下对商品的评价会不同。

然而，当我们关注的是简单的消费选择问题时，相关的商品通常非常明显。我们经常采用前面介绍过的思想：只使用两种商品进行分析，而将其中一种商品视为“所有其他的商品”。这样做的好处是，我们可以重点关注一种商品和所有其他商品之间的权衡问题。这样，在涉及很多商品的消费选择问题时，我们仍可以使用二维图形进行分析。

因此，可以假设消费束只包含两种商品，令(x_1,x_2)表示消费束，其中x_1代表某种商品，x_2代表另外一种商品。有时可将这个消费束简写为X。

3.1 消费者的偏好

假设给定两个消费束(x_1,x_2)和(y_1,y_2)，消费者可按照他自己的意愿对这两个消费束排序。即消费者可以认为一个消费束严格好于另外一个消费束，或者认为这两个消费束无差异。

我们用符号$\succ$表示**严格偏好** (strict preference)，因此$(x_1,x_2)\succ(y_1,y_2)$表示消费者严格偏好(x_1,x_2)胜于(y_1,y_2)，意思是说他肯定想要消费束x而不是消费束y。这种偏好关系提供了判断消费者选择哪个消费束的依据。如果消费者偏好消费束x胜于消费束y，那么他会选择消费束x。因此，偏好的思想是基于消费者**行为**之上的。为了判断某消费者是否偏好某个消费束胜于另一个消费束，我们可以提供给他这两个消费束，观察他的选择行为。如果他原本可以选择(y_1,y_2)，但他总是选择(x_1,x_2)，那么可知他偏好(x_1,x_2)胜于(y_1,y_2)。

如果消费者认为两个消费束是**无差异的**（indifferent），我们记为$(x_1,x_2)\sim(y_1,y_2)$，其中符号~表示无差异。无差异的意思是说消费者根据自己的偏好，认为这两个消费束提供的满足程度是一样的。

我们用符号$\succeq$表示**弱偏好**（weak preference），$(x_1,x_2)\succeq(y_1,y_2)$意思是说消费者对前者的偏好胜于后者或者消费者对于这两个消费束无差异，简单地说就是消费者认为前者**至少**和后者一样好。

严格偏好、弱偏好和无差异不是三个独立的概念，它们互相相关！例如，若果$(x_1,x_2)\succeq(y_1,y_2)$并且$(y_1,y_2)\succeq(x_1,x_2)$，则$(x_1,x_2)\sim(y_1,y_2)$。即，如果消费者认为消费束$x$至少和消费束$y$一样好，**并且**消费束$y$至少和消费束$x$一样好，则消费者一定对这两个消费束是无差异的。

类似地，如果$(x_1,x_2)\succeq(y_1,y_2)$，但$(x_1,x_2)\sim(y_1,y_2)$不成立，则$(x_1,x_2)\succ(y_1,y_2)$。这只不过是说如果消费者认为消费束$x$至少和消费束$y$一样好，但这两个消费束对他来说不是无差异的，则他认为消费束x严格好于消费束y。

3.2 关于偏好的假设

经济学家通常对消费者偏好的“一致性(consistency)”或称“相容性”作出一些假设。例如，某情形下，既有$(x_1,x_2)\succ(y_1,y_2)$又有$(y_1,y_2)\succ(x_1,x_2)$，这样的偏好即使不说是矛盾的，也是不合理的。因为这意味着消费者认为消费束x比y好，但又反过来认为消费束y比x好。

因此需要对偏好关系作出某些假设。有些假设非常基本，通常被称为消费者理论的“公理”。此处我们给出消费者偏好的三条公理。

- **完备性**（complete）公理。假设任何两个消费束都可以比较，即给定消费束x和消费束y，必有$(x_1,x_2)\succeq(y_1,y_2)$或者$(y_1,y_2)\succeq(x_1,x_2)$，或者二者都成立，在最后一种情形中，消费者对于这两个消费束是无差异的。
- **反身性**（reflexive）公理。假设任何消费束都至少和它本身一样好，即$(x_1,x_2)\succeq(x_1,x_2)$。
- **传递性**（transitive）公理。如果$(x_1,x_2)\succeq(y_1,y_2)$并且$(y_1,y_2)\succeq(z_1,z_2)$，则$(x_1,x_2)\succeq(z_1,z_2)$。换句话说，如果消费者认为消费束$x$至少和$y$一样好，而且消费束$y$至少和$z$一样好，则他认为消费束$x$至少和$z$一样好。

第一个公理即完备性公理，几乎难以反驳，至少对于经济学家通常研究的选择问题来说是这样的。我们说任何两个消费束可以比较，只不过想表明消费者能在这二者中作出取舍。你也许会想到极端情形比如生或死的选择，这类问题中对选项的排序很困难甚至无法排序，幸好经济学基本不研究这样的选择问题。

曹乾（东南大学 caoqianseu@163.com）

第二个公理即反身性公理，是平凡的（trivial）[一]。任何消费束当然和同样的消费束至少一样好。父母偶尔会发现小孩的行为违背了反身性公理，但对于大多数成人的行为来说，这个假设似乎是合理的。

第三个公理即传递性公理，问题相对大些。人们并没搞清偏好的传递性是否为偏好的**必要**属性。从纯逻辑的角度看，偏好的传递性假设无法做到让人信服。事实上，传递性是对人们选择行为作出的假设，而不是纯逻辑的论断。当然，这个假设是否符合基本逻辑事实并不是问题的关键，关键在于这个假设能否合理地准确刻画消费者的选择行为。

如果某个消费者说他认为消费束 x 比 y 好，y 比 z 好，z 比 x 好。怎么评价他？我们通常认为他不正常。

更重要的是，如果给上述消费者三个消费束 x，y 和 z，他如何选出最喜欢的？他很难作出选择，因为无论他选哪个消费束，都还有比这个消费束更好的消费束。如果我们想研究消费者的“最佳”选择问题，必须假设偏好满足传递性公理。如果偏好不是传递的，就可能出现无法作出最优选择的情形。

3.3 无差异曲线

偏好的三个公理是消费者理论的基石，如果再加上其他一些技术性的假设，我们就构建起整个消费者理论的大厦。然而，借助图形分析更加直观，但我们先要引入刻画偏好的一个工具，这就是**无差异曲线**（indifference curve）。

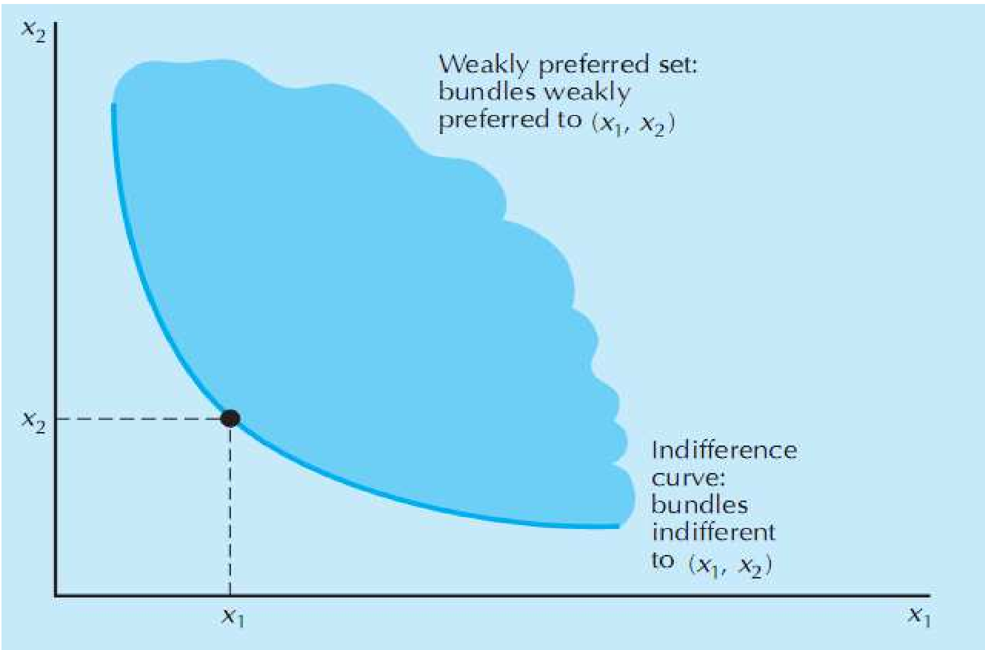


图 3.1：弱偏好集。阴影区域的所有消费束都至少和 (x_1, x_2) 一样好。

[一] trivial（平凡的），数学用语，是指某东西（例如向量空间）的结构非常简单，不值一提（但又不得不提）。有时我们也会遇到下列说法：某方程的某个解是平凡的。这是说这个解非常简单。

请看图 3.1，我们用横轴和纵轴分别表示商品 1 和商品 2 的消费数量。任选一个消费束 (x_1,x_2)，用阴影标记出所有比 (x_1,x_2) 更好或者一样好的消费束。这些消费束称为**弱偏好集**（weakly preferred set）。在弱偏好集边界上的消费束与 (x_1,x_2) 一样好，这些消费束组成了**无差异曲线**。

对于任何一个消费束都可以画出经过这点的无差异曲线。经过某个消费束的无差异曲线包含了所有与该消费束一样好的消费束。

使用无差异曲线刻画偏好时，在同一条无差异曲线上的消费束是无差异的。不同无差异曲线上的消费束怎样比较？在图形上如果只给你几条无差异曲线，你无法比较不同无差异曲线上的消费束。因此，有时候人们在无差异曲线上标记一些小箭头，指出哪个方向上的消费束更受偏好。我们一般不这么做，除非在容易混淆的情形时，我们才用箭头标记。

如果我们对偏好不作出进一步的假设，你可以画出任意形状的无差异曲线。即使如此，无差异曲线仍具有一个重要的性质：**不同无差异曲线（代表不同偏好水平）不能相交**。也就是说，不可能出现类似图 3.2 的情形。

为了证明上述结论，在图形上任意选取三个消费束：X，Y 和 Z，使得 X 只位于其中一条无差异曲线上，Y 只位于另一条无差异曲线上，Z 位于上述两条无差异曲线的交点处。由于不同无差异曲线代表不同的偏好水平，因此假设 X 严格好于 Y。由图可知，$X \sim Z$ 以及 $Z \sim Y$，由传递性公理可知 $X \sim Y$。但这与我们的假设 $X \succ Y$ 矛盾。这就说明无差异曲线不能相交。

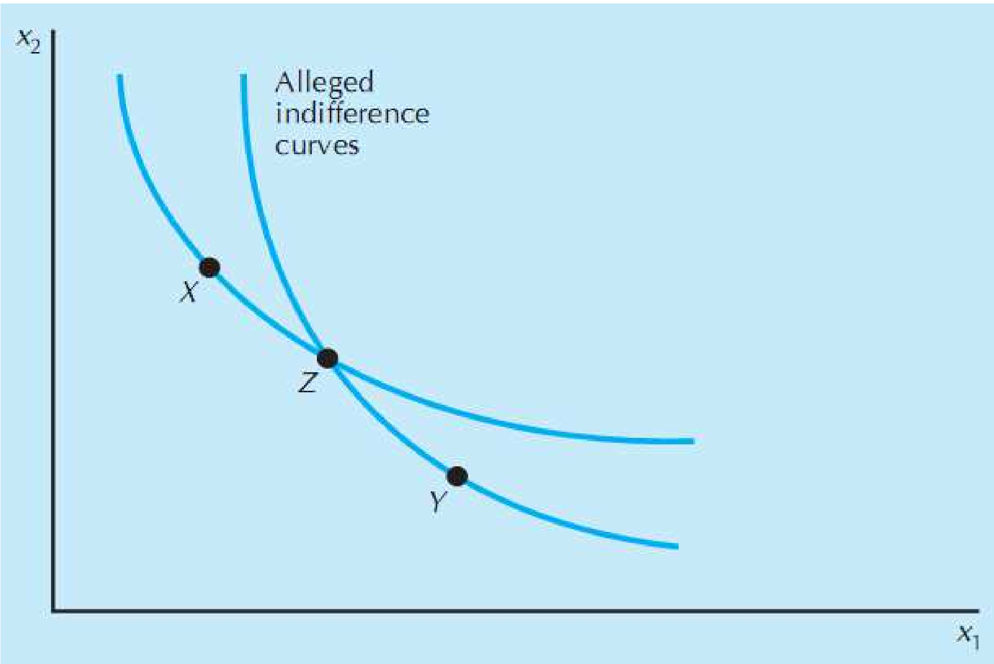


图 3.2：无差异曲线不能相交。如果相交，消费束 X，Y 和 Z 必然两两无差异，因此这三个消费束就不能位于不同的无差异曲线上。

无差异曲线还有其他什么样的性质吗？一般来说，答案为：并不多。无差异曲线是描述

偏好的一种方法。几乎任何“合理的”偏好都能用无差异曲线描述。你要掌握的技巧是给你一种偏好类型，你能找到相应的无差异曲线形状来描述它。下面我们就介绍这些知识。

3.4 偏好的例子

我们利用例子来考察偏好和无差异曲线之间的对应关系。我们先给出一些偏好类型，然后看看什么样形状的曲线能描述这样的偏好。

给定偏好类型，构建相应无差异曲线的通用程序如下。首先，画好横纵坐标，任意画出一个点，代表消费束 (x_1,x_2)。现在假设给消费者一些商品 1（Δx_1），则新消费束为 $(x_1+\Delta x_1,x_2)$。你来回答怎样**改变**商品 2 的数量才能恰好使消费者还在原来的无差异曲线上。假设商品 2 的变化量为 Δx_2。现在问题变为“给定商品 1 的消费量变化 Δx_1，Δx_2 为多大时，$(x_1+\Delta x_1,x_2+\Delta x_2)$ 和 (x_1,x_2) 这两个消费束恰好无差异？”一旦你知道如何移动 (x_1,x_2)，你就画出了一小段无差异曲线。再画出一个消费束，重复以上步骤…直到你画出比较全面的无差异曲线形状。

完全替代

两种商品是**完全替代**（perfect substitute）的，如果消费者愿意按**固定**比率将一种商品替换为另一种。完全替代最简单的情形是两商品为 1：1 替代。

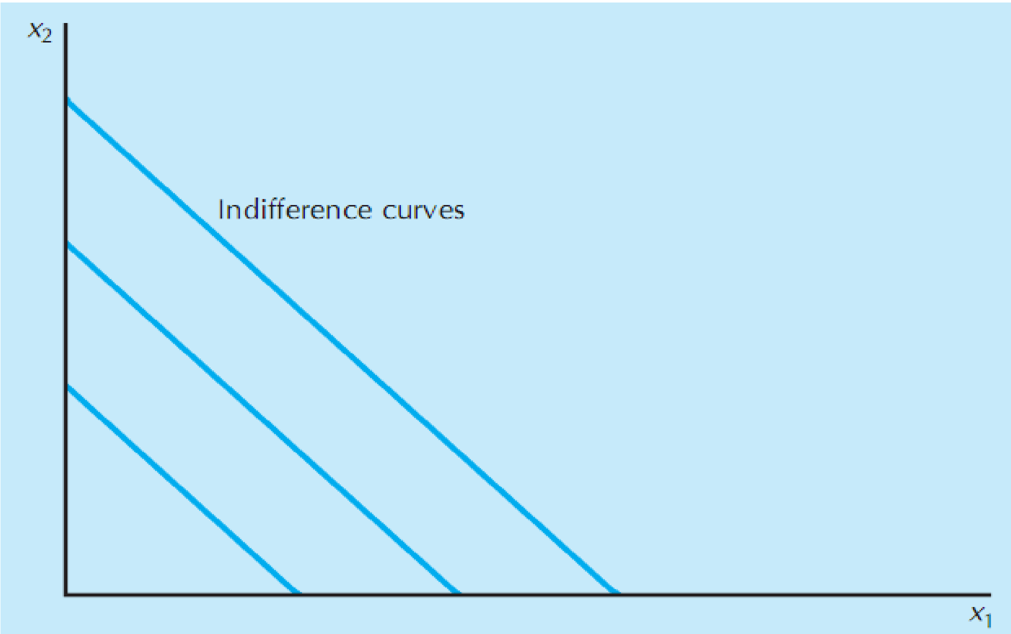


图 3.3：完全替代。消费者只关心铅笔的总数而不关心铅笔的颜色。因此无差异曲线为斜率等于-1的直线。

例如，假设两种商品分别为红色铅笔和蓝色铅笔，某消费者喜欢铅笔，但不在乎铅笔的颜色。任选一个消费束比如（10, 10），那么对于这个消费者来说，含有 20 支铅笔的任何消费束都和（10, 10）一样好。用数学语言表达，任何满足 $x_1 + x_2 = 20$ 的消费束都位于通过（10, 10）的那条无差异曲线上。因此该消费者的所有无差异曲线都是互相平行的直线，它们的斜率都等于 -1。如图 3.3 所示。由于铅笔数量越多的消费束越好，因此偏好增加的方向为右上方（含正上方和正右方）。如图 3.3 所示。

我们用前面介绍的通用程序进行分析。如果消费束为（10, 10），增加 1 单位红色铅笔则消费束变为（11, 10），蓝色铅笔的数量如何变动才能使我们回到原来的无差异曲线上？答案很明显：减少 1 单位蓝色铅笔。因此，通过（10, 10）的无差异曲线斜率为 -1。使用任何一个消费束，重复以上步骤，所得到的结果是一样的：无差异曲线的斜率等于 -1。

完全替代的一个重要特征是无差异曲线的斜率是**固定的**。例如，我们我们用纵轴表示蓝色铅笔，用横轴表示**成对**（每对 2 支）的红色铅笔，则无差异曲线的斜率为 -2，因为消费者为了得到**一对**红色铅笔，他愿意放弃 2 支蓝色铅笔。

为简单起见，教材里完全替代的例子基本是 1：1 替代的。我们把更一般的情形放在与本教材配套的习题书中。

完全互补

完全互补(perfect complements)是指几种商品按固定的搭配比例被一起消费。在某种意义上，商品之间是“互为补充的”。一个很好的例子是一双鞋中的右鞋和左鞋。消费者喜欢鞋子，但他总是穿一双而不是只穿左鞋或右鞋。只有一只鞋子对消费者来说几乎没什么用。

我们要画出完全替代的无差异曲线。假设消费束（右鞋的数量，左鞋数量）为（10，10）。现在增加一只右鞋得到（11，10）。因为这两种商品是 1：1 完全互补的，因此（11，10）和（10，10）对消费者来说无差异：额外多出的一只右鞋对消费者没什么好处。如果仅增加一只左鞋，则结果类似：（10，11）和（10，10）无差异。

因此完全互补的商品的无差异曲线都是 L 形，在 L 形曲线的顶点处，右鞋的数量等于左鞋的数量。如图 3.4 所示。

同时等量增加左鞋和右鞋的数量会使消费者移动到更好的位置，因此偏好增加的方向必然是右上方，如图 3.4 所示。

完全互补的一个重要特征是不同商品按固定比例搭配在一起消费，该搭配比例未必是 1：1。如果消费者喝一杯茶时总是放入两茶匙的糖，而且假设他买糖的目的就是为了喝茶，则无差异曲线是 L 形的。在这个例子中，若消费束为（糖的数量，茶的数量），则 L 形曲线的拐角出现在（2，1），(4，2)等等，而不再是鞋子例子中的（1，1），（2，2）等。

在教材中，我们主要研究 1：1 完全互补的情形。其他更一般的情形放在习题书中。

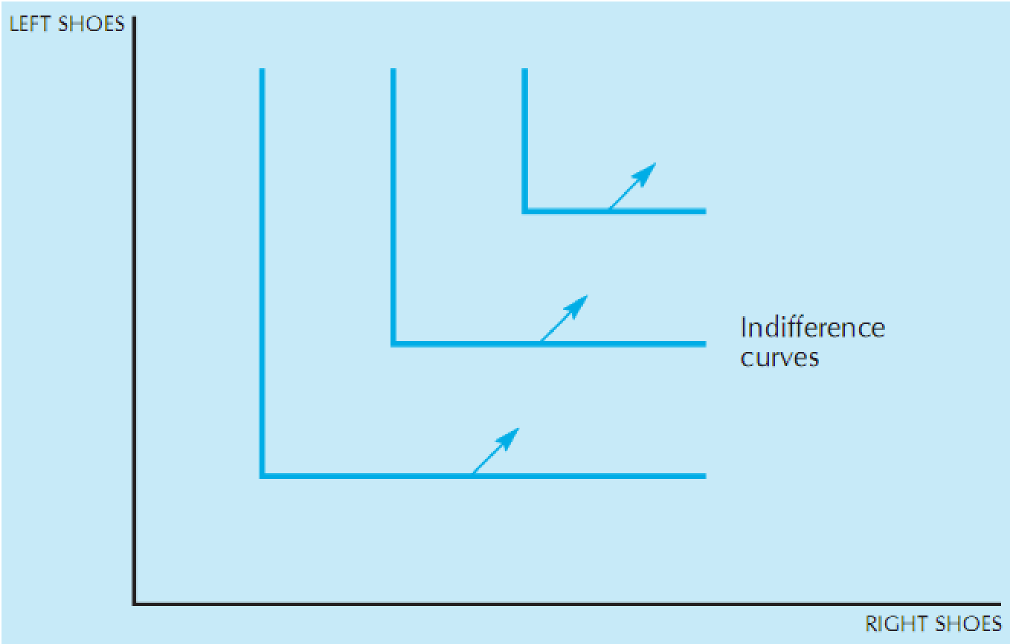


图 3.4：完全互补。消费者总是将不同商品按固定比例搭配在一起消费。因此无差异曲线都是 L 形的。

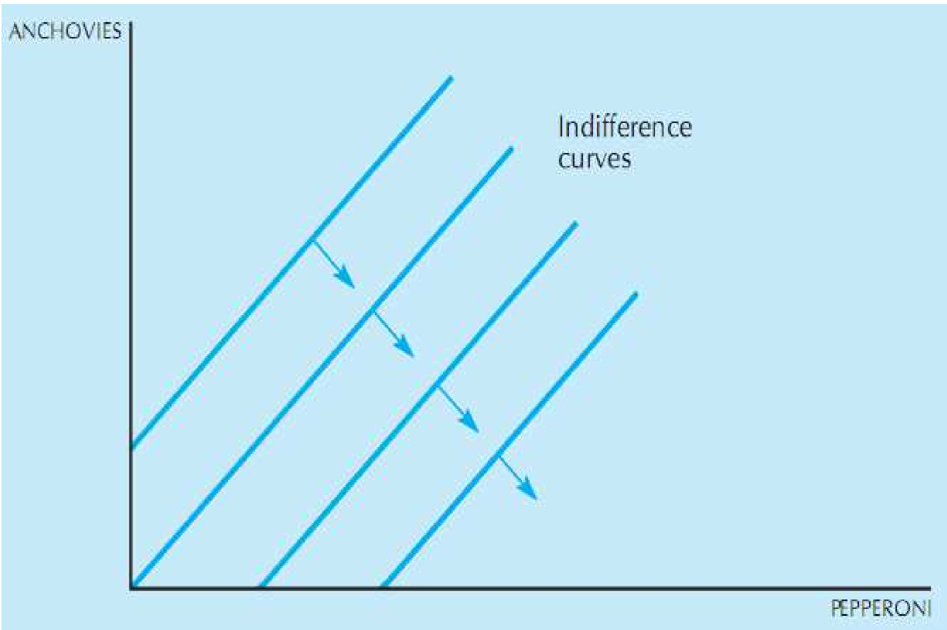


图 3.5：厌恶品。假设对某个消费者来说，凤尾鱼（anchovies）为“厌恶品（bad）”，辣香肠（pepperoni）为“好商品（good）”，因此他的无差异曲线斜率为正。

厌恶品

厌恶品(bad)是指消费者不喜欢的商品。例如，假设某消费者喜欢吃辣香肠(pepperoni)

但讨厌吃凤尾鱼（anchovies）。我们设计一种方案让该消费者在辣香肠和凤尾鱼之间进行选择，例如，如果他每吃 x_1 单位的凤尾鱼，我们会在比萨饼上添加 x_2 单位辣香肠作为补偿。这种情形下，该消费者的偏好如何用无差异曲线表示？

任选一个消费束 (x_1, x_2)，如果再给他一些凤尾鱼，那么怎样让消费者仍然位于经过 (x_1, x_2) 的无差异曲线上，即怎样变动辣香肠的数量？显然，我们必须多给他一些辣香肠，他才会多吃凤尾鱼。因此，他的无差异曲线必然向右上方倾斜，如图 3.5 所示。

在这个例子中，偏好增加的方向是右下方，也就是凤尾鱼的消费量减少而辣香肠的消费量增加的方向，如图 3.5 中的箭头所示。

中性商品

中性商品（neutral good）是指某种商品，消费者有它也行无它也可。如果凤尾鱼对于某个消费者来说是中性商品，他的无差异曲线是什么样子的？[一] 如果我们用纵轴表示凤尾鱼，以横轴表示辣香肠，则这个消费者的无差异曲线都是垂直于横轴的直线，如图 3.6 所示。

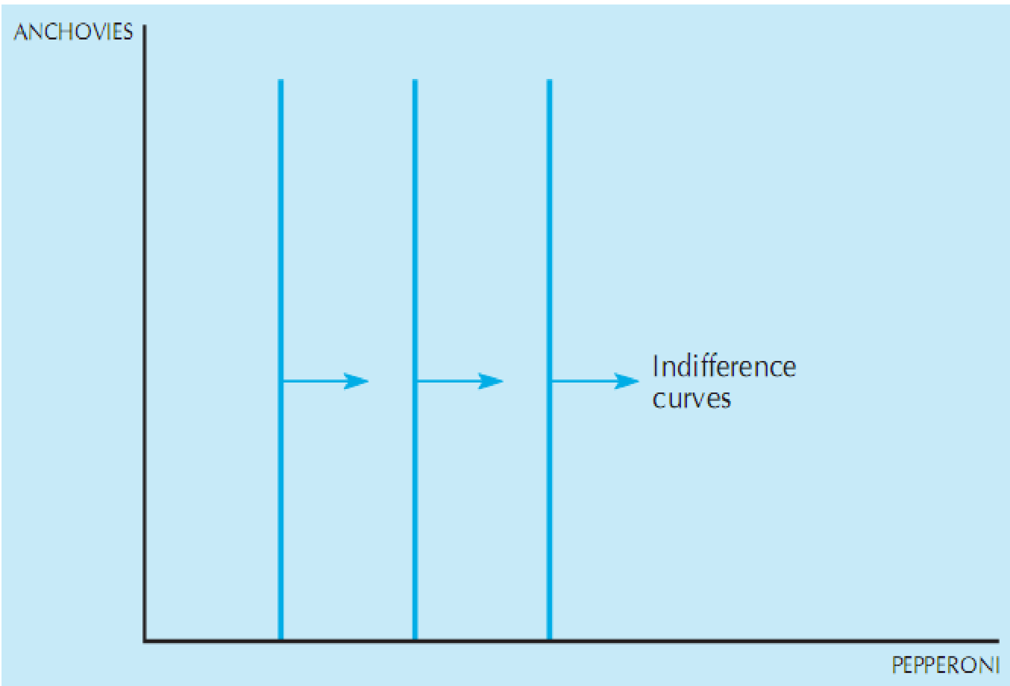


图 3.6：中性商品。消费者喜欢辣香肠（pepperoni），但对于凤尾鱼（anchovies）无所谓。凤尾鱼对他来说是中性商品，因此他的无差异曲线是垂直于横轴的直线。

这个消费者只关心辣香肠的数量，而不关心凤尾鱼的数量。辣香肠越多越好，但增加凤尾鱼的数量对他丝毫无影响。

[一] 凤尾鱼对于任何消费者来说都是中性商品吗？

饱和

有时候我们研究的消费选择含有饱和（satiation）点，饱和点的消费束对于消费者来说是最优的，消费束越接近饱和点，他越喜欢。例如，假设消费者最喜欢的消费束为 $(\bar{x}_1, \bar{x}_2)$，并且离这个消费束越远，他的状况越糟糕。在这种情形下，$(\bar{x}_1, \bar{x}_2)$ 就是一个**饱和点**（satiation point），或者称为**幸福点**（bliss point）。这种无差异曲线具有如图 3.7 所示的形状。最优点为 $(\bar{x}_1, \bar{x}_2)$，离这个饱和点越远的消费束，它所在的无差异曲线位置“越低”。

这种无差异曲线在某点处的斜率是正还是负？如果某点处两种商品都“过少”或者“过多”， 无差异曲线在该点的斜率为负，如果在某点处只有一种商品“过多”，那么无差异曲线在该点的斜率为正。如果这两种商品中一种过多，那么这种过多的商品就变成了厌恶品，因为减少它的数量，就会离饱和点更近些。如果两种商品都过多，那么它们都是厌恶品，分别减少各自的数量，离饱和点更近。例如，假设这两种商品为巧克力蛋糕和冰淇淋，很有可能你每周的最优消费量为 $(\bar{x}_1, \bar{x}_2)$。与饱和点相比，少吃不好，但过犹不及。

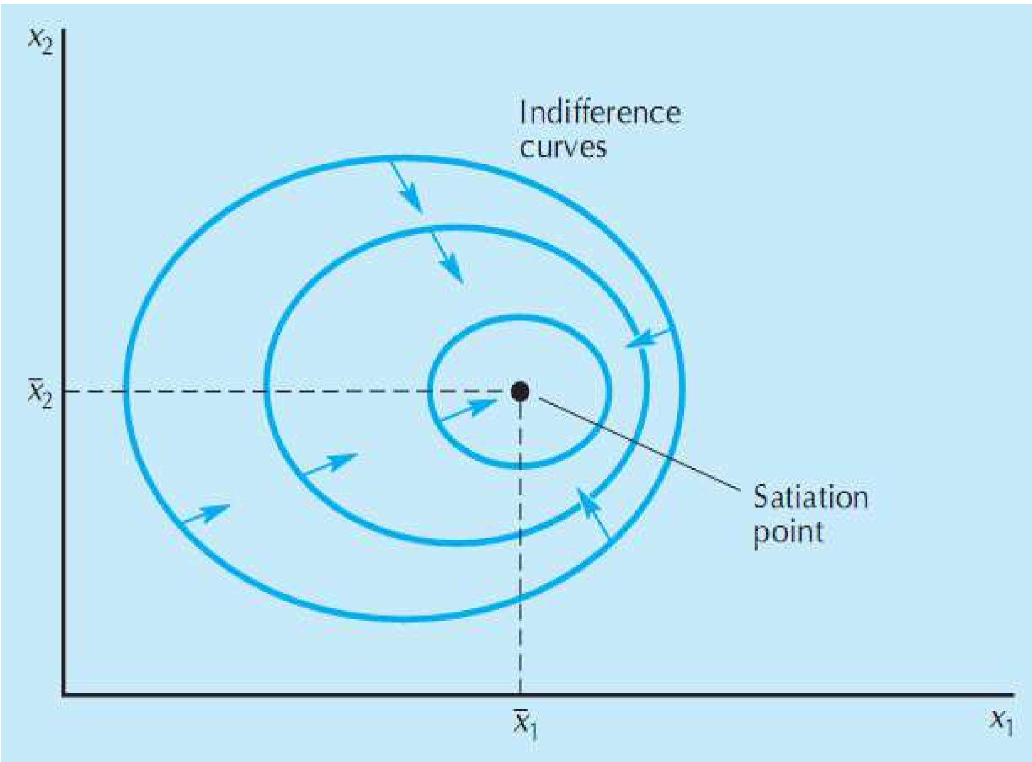


图 3.7 饱和型偏好。$(\bar{x}_1, \bar{x}_2)$ 是饱和点或称幸福点，所有的无差异曲线都围绕着饱和点。

如果你仔细想想，现实生活中绝大多数的商品都类似于这里的巧克力蛋糕和冰淇淋。但是人们一般不会自愿**选择**过多消费某种商品。如果你想要商品的数量为 $(\bar{x}_1, \bar{x}_2)$，你何必多要？因此，从消费者行为理论的角度来看，应该重点关注的区域是两种商品分别小于饱和数量 $(\bar{x}_1, \bar{x}_2)$ 的区域。人们实际关注的选择就是这样的情形，这也是我们所要研究的情形。

离散商品

通常商品消费数量可以含有小数，例如尽管你每次买一品脱（等于 8 加仑）牛奶，但你每个月可能平均只喝 12.43 加仑。但是有时商品消费数量只能是整数，这样的商品称为**离散商品**（discrete good）。

例如消费者对汽车的需求。我们可以将汽车需求定义为开车时间，显然开车时间是个连续变量，但很多时候我们关心的是人们需求多少辆汽车。汽车的辆数是一个离散变量，它只能是整数。使用偏好刻画消费者对离散商品的选择行为，并不困难。假设 x_1 是离散商品，x_2 为花费在所有其他商品上的资金。在图 3.8 中，我们说明了这种无差异“曲线”的形状以及一个弱偏好集。在这种情形下，与给定消费束无差异的消费束构成了一个离散点集；与给定消费束至少一样好的消费束构成了一个线段集。

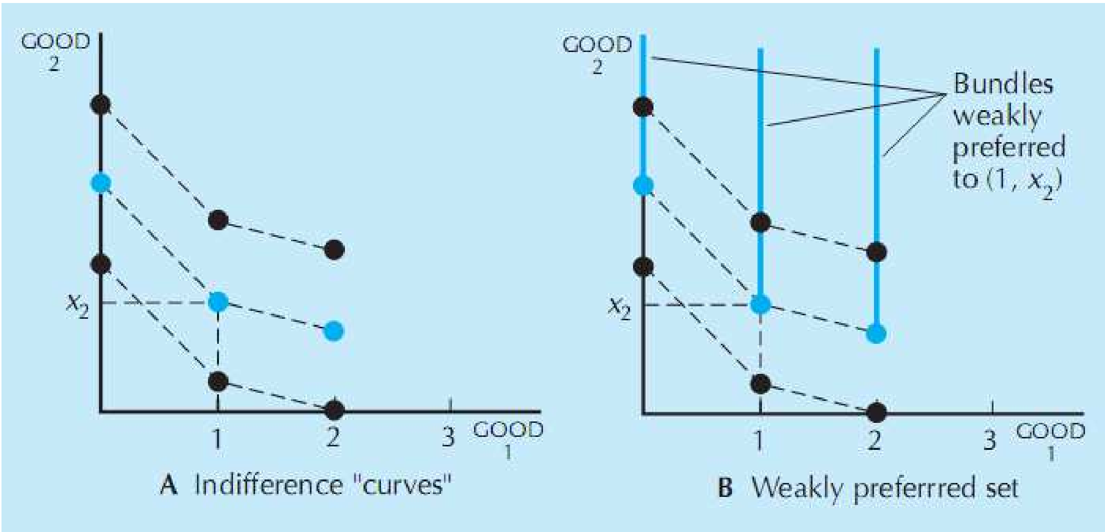


图 3.8：离散商品。商品 1 为离散商品，只能以整数计量。A 图中，同一条虚线上的消费束是无差异的；B 图中，垂线表示和给定的消费束至少一样好的消费束集合。

通常根据实际情况确定是否强调商品的离散性质。在我们分析消费者行为时，如果某消费者只消费一两个单位的某商品，可能有必要强调这种商品的离散性质。如果他消费三四十单位，将这种商品视为连续商品会更方便。

3.5 良好性状的偏好

我们已经分析了几种无差异曲线。我们已经知道，很多类型的偏好，无论合理与否，都可用无差异曲线描述。但如果我们想在一般意义上刻画偏好，最好重点研究一些一般形状的无差异曲线。

在前面我们曾给出了偏好的三个公理性假设，在本节我们进一步给出偏好的几个一般性

假设，并阐明这些假设对无差异曲线形状的影响。当然，对偏好所做的假设不限于我们这里列出的；在一些情形下，根据你具体的研究目的，你也可以作出其他的假设。我们这里所作的假设目的在于定义**良好性状的无差异曲线**（well-behaved indifference curves）。

第一个假设是多多益善。也就是说我们研究的是**好的商品**(goods)，而不是**厌恶品**（bads）。更准确地说，给定两个商品束 (x_1, x_2) 和 (y_1, y_2)，如果 $y_1 \geq x_1$ 且 $y_2 > x_2$ 或者 $y_1 > x_1$ 且 $y_2 \geq x_2$，则必有 $(y_1, y_2) \succ (x_1, x_2)$。这个假设有时叫作偏好的**单调性**（monotonicity）。需要指出，如果存在饱和点，则在达到饱和点之前，多多益善的假设仍然成立。因此，单调性的假设是说我们研究的情形仅限于消费束在达到**饱和之前**的情形，既然没饱和，多多益善**仍然成立**。如果每个人都达到了饱和点，经济学这门学科存在的意义就不大。

单调性假设意味着无差异曲线的形状是什么样的？它意味着无差异曲线的斜率为**负的**。请看图 3.9。我们从任意一个消费束 (x_1, x_2) 开始分析，如果我们从这个消费束向右上方（含正上方和正右方）移动，我们必定达到一个更好的位置。如果我们向左下方（含正下方和正左方）移动，我们达到的位置必定更差。因此如果我们移动到一个和 (x_1, x_2) **无差异**的位置，那么我们移动的方向必然是左上方或者右下方，即这个位置的消费束与 (x_1, x_2) 相比，一种商品增加的同时另一种商品减少。换句话说，无差异曲线的斜率必然为负。

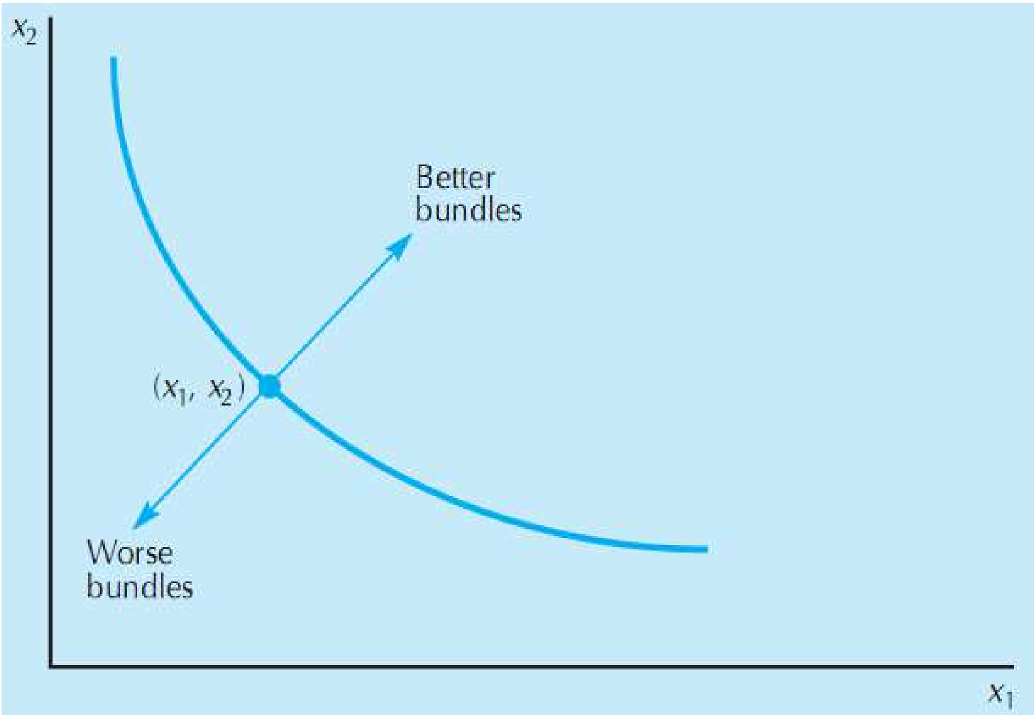


图 3.9：单调类偏好。(x_1, x_2) 右上方的消费束比 (x_1, x_2) 好；(x_1, x_2) 左下方的消费束比 (x_1, x_2) 差。

第二个假设是平均束好于端点束（averages are preferred to extremes）。也就是说，如果我们在同一条无差异曲线上取两个消费束 (x_1, x_2) 和 (y_1, y_2)，那么由它们加权平均得到的消

费束比如$(\frac{1}{2}x_1+\frac{1}{2}y_1,\frac{1}{2}x_2+\frac{1}{2}y_2)$弱偏好于或者严格偏好于上述两个端点消费束。该加权平均消费束中商品 1 和 2 的数量，分别等于两个端点消费束商品 1 的平均数和商品 2 的平均数。因此它位于连接消费束 x 和 y 的线段的中点处。

事实上，这个假设中的权重可取任意数t（其中$0\le t\le 1$），上例中权重取$1/2$只是特例。因此第二个假设更一般的表述是，如果$(x_1,x_2)\sim(y_1,y_2)$，则对于任意t（其中$0\le t\le 1$）下式都成立

$$(tx_1+(1-t)y_1,tx_2+(1-t)y_2)\succeq(x_1,x_2)$$

上式中的左端是加权平均消费束，它由消费束 x 和 y 加权平均得到，其中 x 商品束的权重为t， y 商品束的权重为$(1-t)$。因此，消费束 x 与该加权平均消费束之间的直线距离，等于消费束 x 与消费束 y 之间的直线距离乘以$(1-t)$。

上述偏好的第二个假设的有什么几何意义？它表明弱偏好于消费束(x_1,x_2)的那个消费束集是**凸集**（convex set）。因为(x_1,x_2)和(y_1,y_2)是不同的消费束，则平均消费束弱偏好于端点消费束，即(x_1,x_2)和(y_1,y_2)的所有加权平均消费束弱偏好于这两个消费束本身。凸集的一个性质是如果你在某集合中**任取**两点，那么连接这两点的直线段也全部位于该集合内。

图 3.10A 画出了凸偏好的一个例子，而图 3.10B 和 3.10C 中的例子都是非凸偏好。图 3.10C 中的偏好为严重非凸的，这种情形我们称为“凹（concave）偏好。”

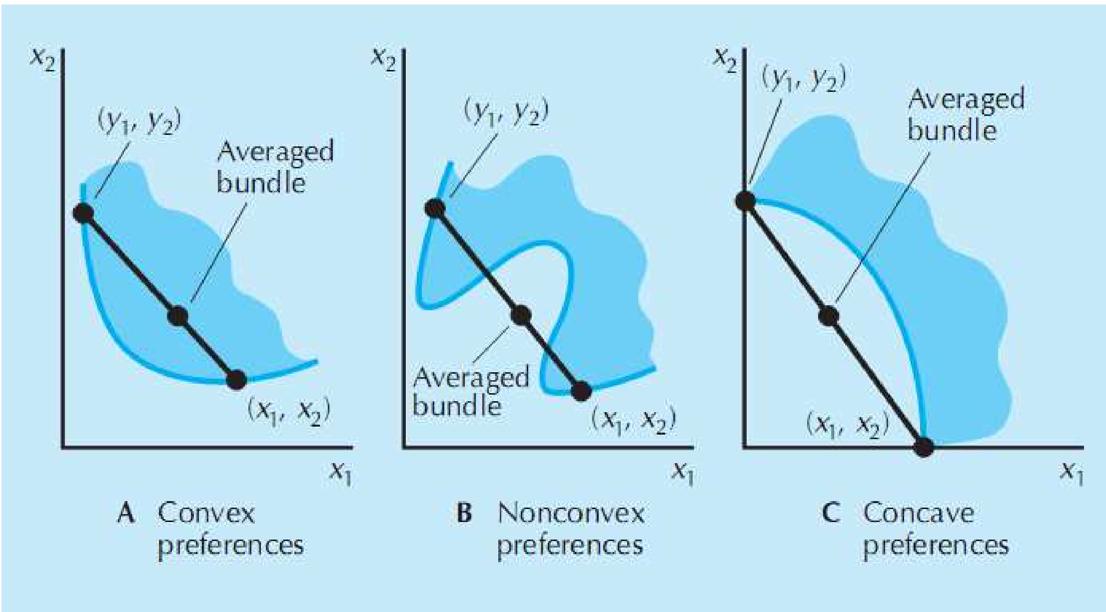


图 3.10：凸偏好和非凸偏好。图 A 为凸偏好；图 B 和图 C 为非凸偏好，其中图 C 为凹偏好。

你能想出非凸偏好的例子吗？我举个例子，比如我对冰淇淋和橄榄的偏好。我喜欢吃冰淇淋也喜欢吃橄榄，但我不喜欢把它们放在一起吃！我打算一个小时后吃点东西，（8 单位冰淇淋，2 单位的橄榄），或者（2 单位的冰淇淋，8 单位的橄榄），这两种选择对我来说是无差异的。但这两个消费束的任意一个都比消费束（5 单位冰淇淋，5 单位橄榄）好。我的

这种偏好类型可用图 3.10C 表示。

为什么我们假设良好性状的偏好是凸的？因为，不同商品通常是一起消费的。图 3.10B 和 3.10C 意味着消费者存在着某种程度的“挑食”现象，他只喜欢消费其中一种商品。然而，通常的情形是，消费者愿意用一种商品换得另外一种商品，最终每种商品都消费些，而不是只消费其中的一种。

事实上，如果观察我的冰淇淋和橄榄**月度**消费数据，而不是很短时期内的消费数据，那么你就会知道我的偏好更像图 3.10A 而不是图 3.10C。在每个月内，某时我会吃些冰淇淋，另一个时刻我会吃些橄榄，而不是在整个月份内只吃其中的一种。

最后指出一点，对凸偏好的假设可以进一步强化，这就是**严格凸**（strictly convex）偏好。这表示两个端点消费束的加权平均得到的消费束，**严格好于**这两个端点消费束本身。凸偏好的无差异曲线可能包含直线段，而**严格**凸偏好的无差异曲线则必须是“弧形”的（即不能含有直线段）。两种商品为完全替代时，这种偏好是凸的，但不是严格凸的。

3.6 边际替代率

在消费者行为理论研究中会频繁使用**无差异曲线在某点处的切线的斜率**，这样的斜率如此有用以至于它有个专门的名字：**边际替代率**（Marginal Rate of Substitution, MRS）。如此命名的原因是，MRS 衡量某消费者恰好愿意以一种商品替代另外一种商品的比率。

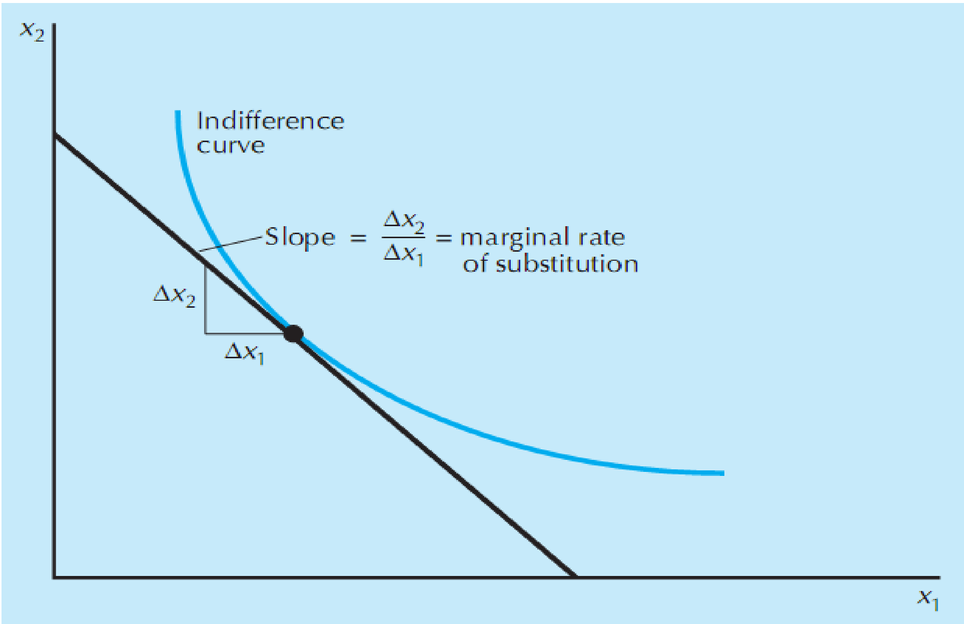


图 3.11：边际替代率（MRS）。边际替代率等于无差异曲线的斜率。

你可能对 MRS 通常为**负值**感到少许困惑。事实上，我们已经知道单调偏好的假设意味

着无差异曲线的斜率必然为负。既然 MRS 等于无差异曲线的斜率，因此 MRS 自然为负值。

假设我们从消费者手里拿走一些商品 1（Δx_1），再给他一些商品 2（Δx_2），这些商品 2 足够让他恰好呆在原先的无差异曲线上，因此，在用一些商品 2 替代一些商品 1 之后，该消费者的状况不变。$\Delta x_2 / \Delta x_1$ 这一比率就是该消费者愿意用商品 2 替代商品 1 的比率[一]。

现在假设 Δx_1 很小（称为边际变动），则 $\Delta x_2 / \Delta x_1$ 衡量商品 2 替代商品 1 的边际比率。随着 Δx_1 趋向于 0，$\Delta x_2 / \Delta x_1$ 趋近于无差异曲线的斜率，如图 3.11 所示。

在比率 $\Delta x_2 / \Delta x_1$ 中，我们总是认为分子和分母的数值很小，以描述原消费束的**边际**变化。因此，MRS 等于无差异曲线的斜率：消费者恰好愿意按照该比率，用少量商品 2 替代商品 1。

MRS 衡量了消费者行为的一个有趣性质。假设消费者的偏好的是良好性状的，即偏好单调而且凸。再假设他目前的消费束为 (x_1, x_2)。现在让他做如下交易：他可以用商品 1 交换商品 2，或用商品 2 交换商品 1，规定"交换率"为 E（即一单位商品 1 可以交换 E 单位商品 2）。

因此，如果消费者放弃 Δx_1 单位商品 1，他可以得到 $E\Delta x_1$ 单位商品 2；或者相反地，如果他放弃 Δx_2 单位商品 2，他可以得到 $\Delta x_1 / E$ 单位商品 1。从几何图形上看，这表明该消费者可以沿着一条直线任意移动，这条直线的斜率为 $-E$ 且经过 (x_1, x_2) 点，如图 3.12 所示。从点 (x_1, x_2) 向左上方移动意味着用商品 1 交换商品 2；从点 (x_1, x_2) 向右下方移动意味着用商品 2 交换商品 1。在这两种情形下，交换率都为 E。因为交换意味着放弃一种商品以获得另外一种商品，**交换率** E 对应着**斜率** $-E$。

现在问个问题，如果消费者宁可选择呆在点 (x_1, x_2)，即他不愿意交易，那么交换率应为多大？如果你注意到下面的事实，这个问题就不难回答：每当这条交易线穿过无差异曲线时，这条交易线上就存在比点 (x_1, x_2) 更好的某些点，这些更好的点位于被穿过的那条无差异曲线的上方。

所以，如果消费者宁可呆在 (x_1, x_2) 点而不愿意移动，交易线必然与预算线相切。也就

[一]计算商品 1 和 2 之间的边际替代率时，涉及到这两种商品变化量的比率，那么这个比率到底是 $\Delta x_2 / \Delta x_1$ 还是 $\Delta x_1 / \Delta x_2$？为了去除这种模糊性，通常定义 $MRS_{12} = \dfrac{\Delta x_2}{\Delta x_1}$，$MRS_{21} = \dfrac{\Delta x_1}{\Delta x_2}$。在计算时用这两个式子中的任何一个都可以，而且都可以读成"商品 2 对商品 1 的边际替代率"（或"商品 1 对商品 2 的边际替代率"），也就是说在读法上我们通常不明确区分这两个式子的区别。然而，你一定要搞清楚它的意思。例如 $MRS_{12} = \dfrac{\Delta x_2}{\Delta x_1} = -\dfrac{3}{5}$，它的意思可能有两种：一是增加 3 单位的 x_1 就必须减少 5 单位 x_2 才能使他的满足程度不变（位于同一条无差异曲线上）；二是减少 3 单位 x_1 就必须增加 5 单位 x_2 才能使他的满足程度不变。具体是哪种，需要根据更具体的信息分析。请在图 3.11 中比划一下这两个意思的区别。另外，我们通常用横轴表示商品 1 的数量，用纵轴表示商品 2 的数量，在这种情形下，用 $MRS_{12} = \dfrac{\Delta x_2}{\Delta x_1}$ 更直观些。译者注。

是说，这条交易线的斜率（$-E$）必然等于无差异曲线在(x_1, x_2)点的斜率。如果交换率不为 E，则交易线会穿过无差异曲线，这样消费者就会移动到其他更好的点。

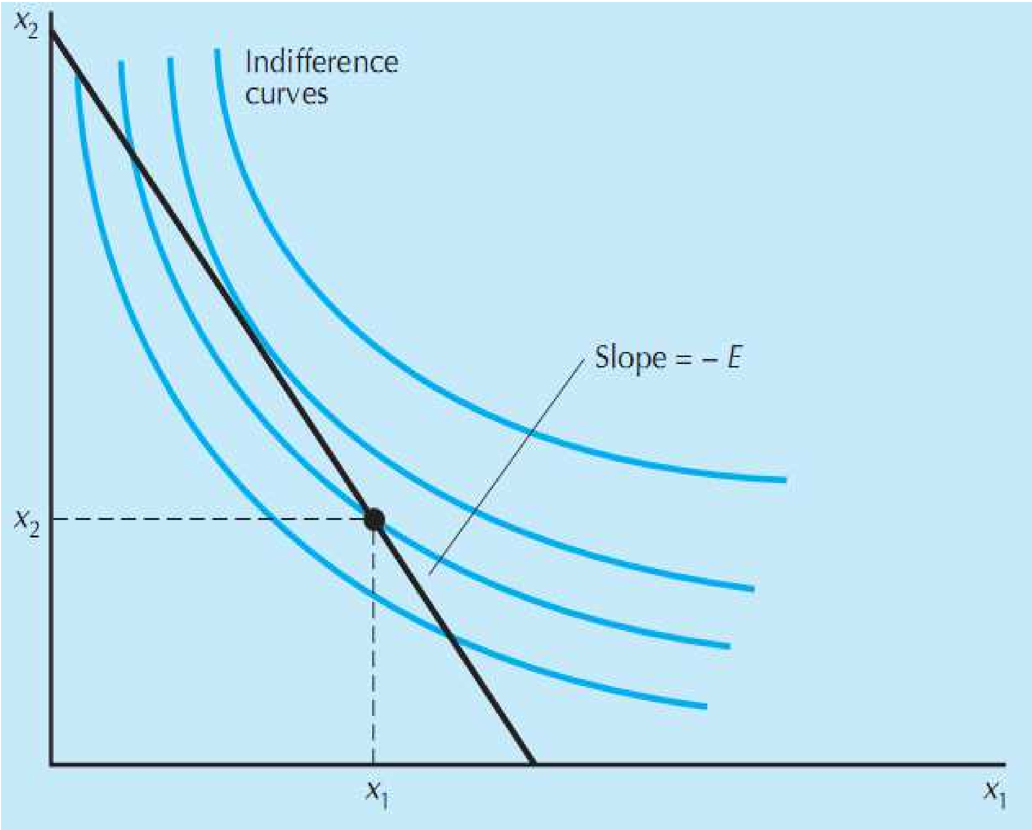


图 3.12：按某交换率进行的交换。此处我们让消费者按照交换率 E 进行交换，这意味着他可以沿着斜率为$-E$的一条直线任意移动。

因此，无差异曲线的斜率即边际替代率，意味着在该比率下消费者恰好处于交易与不交易的边际点上。如果交易率不等于 MRS，消费者就会用一种商品交换另一种商品；而如果交易率等于 MRS，则消费者就不会进行交易。

3.7 MRS 的其他解释

我们说过，MRS 是一种比率，在该比率下消费者恰好处于用商品 1 替代商品 2 的边际上。换句话说，消费者恰好愿意“支付”一些商品 1 来购买一些商品 2。因此，有时人们说无差异曲线的斜率衡量消费者的**边际支付意愿**（marginal willingness to pay）。

如果商品 2 代表除了商品 1 之外的“所有其他商品”的消费，它用货币单位（元）来衡量，那么你就很容易理解边际支付意愿的这种解释方法。用商品 2 替代商品 1 时的 MRS，它表示为了得到少许商品 1，你愿意支付多少元钱。这正是边际支付意愿的定义。

如果你用边际支付意愿来解释 MRS，则你应该强调两个词：“边际”和“意愿”。MRS

衡量消费者**愿意**支付一定数量的商品 2 来购买**边际**数量（即少许）的商品 1。你实际支付的钱数未必等于你愿意支付的钱数：实际支付的钱数取决于你想购买的商品的价格；愿意支付的钱数取决于你的偏好，而不取决于价格。

类似地，额外少量（即边际）和额外大量购买这两种情形下，你愿意支付的钱数也不同。额外大量购买的情形下，你实际支付的钱数取决于你对商品的偏好和商品的价格。额外少量购买的情形下，你愿意支付的钱数仅取决于你的偏好。

3.8 MRS 的行为

有时候，我们可以通过描述边际替代率的行为来刻画无差异曲线的形状。例如，1：1 完全替代情形下的无差异曲线，它的特征是 MRS 恒为-1。如果两种商品中，若商品 1 是中性商品（以横轴表示它的数量），则无差异曲线为**平行**于**横轴**的直线（请自行画出此图并比较它与图 3.6 的差异），这样的无差异曲线的特征是 MRS_{12} 处处为 0（MRS_{21} 处处为无穷大）。完全互补类型的无差异曲线的特征是：除了顶点外，MRS 要么为 0 要么无穷大（在顶点处 MRS 无定义）。

我们已经指出过，偏好的单调性假设意味着无差异曲线的斜率必然为负，因此对于单调性的偏好来说，MRS 必然为负，即为了多消费一种商品必须少消费另外一种商品。

偏好的凸性假设使 MRS 展现出另外一种行为。对于严格凸的无差异曲线来说，随着 x_1 逐渐增加，无差异曲线的斜率的绝对值即$|MRS_{12}|$会逐渐减少。因此，无差异曲线展现出**边际替代率递减**（diminishing marginal rate of substitution）的性质。这表明随着商品 1 数量的增多，每增加一单位商品 1，消费者愿意放弃商品 2 的数量是递减的。或者反过来说，即随着商品 1 数量的增多，为了多得到一单位商品 2，消费者愿意放弃商品 1 的数量是递增的。由此可见，无差异曲线的凸性假设是很自然的：它表明你拥有某种的数量越多，你越愿意放弃一些这种商品，用来交换其他商品。（但要记住前面介绍过的冰淇淋和橄榄的例子：对某些商品组合，凸性假设不成立！）

总结

1.经济学家假设消费者可对各种消费束进行排序。消费者对消费束的排序方法刻画了他的偏好。

2.可用无差异曲线描述不同类型的偏好。

3.良好性状的偏好，既是单调的（即多多益善）又是凸的（即平均束好于端点束）。

4.边际替代率（MRS）等于无差异曲线的斜率。它的意思是说，为了多得到一些商品 1

消费者愿意放弃商品 2 的数量，即 $MRS_{12}=\Delta x_2/\Delta x_1$。

复习题

1.如果我们有一次看到，某消费者在(x_1,x_2)和(y_1,y_2)都可选的情况下选择了(x_1,x_2)，我们能否断定$(x_1,x_2)\succ(y_1,y_2)$？

2.假设有三个人 A,B 和 C，身高关系为“至少和…一样高”，比如“A 至少和 B 一样高”。这样的关系是传递的吗？是完备的吗？

3.假设有三个人 A,B 和 C，身高关系为“严格高于”。这样的关系是传递的吗？是反身的吗？是完备的吗？

4.某大学橄榄球教练说，任意给定两个前锋比如 A 和 B，他永远偏好身材更高大和速度更快的那个。他的这种偏好关系是传递的吗？是完备的吗？

5.某条无差异曲线能否与自身相交？例如，图 3.2 能否是一条无差异曲线而不是两条？

6.如果偏好是单调的，能否把图 3.2 看成一条无差异曲线而不是两条？

7.如果辣香肠和凤尾鱼都是厌恶品，那么无差异曲线的斜率为正还是负？

8.解释为什么凸偏好意味着“平均束好于端点束”。

9.面值 1 元的钞票与面值 5 元的钞票，计算它们之间的边际替代率。

10.如果商品 1 是中性商品，计算用商品 1 替代商品 2 的边际替代率。

11.举例说明你的偏好在什么样的情形下为凹的。

复习题答案

1.如果我们有一次看到，某消费者在(x_1,x_2)和(y_1,y_2)都可选的情况下选择了(x_1,x_2)，我们能否断定$(x_1,x_2)\succ(y_1,y_2)$？

【复习内容】弱偏好；严格偏好

【参考答案】

注意弱偏好和严格偏好的区别。

不能断定，因为也可能是消费者恰好在这两个消费束之间无差异。也就是说，根据题目的已知条件我们只能断定$(x_1,x_2)\succsim(y_1,y_2)$，这是弱偏好。

那么对本题加上什么样的假设前提，题目中的断定就是正确的？如果加上消费者的偏好是严格凸的这一限制条件，断定$(x_1,x_2)\succ(y_1,y_2)$就是正确的。因为严格凸性条件下，最优解若存在则只有一个。

2.假设有三个人 A,B 和 C，身高关系为“至少和…一样高”，比如“A 至少和 B 一样高”。这样的关系是传递的吗？是完备的吗？

【复习内容】二元关系

【参考答案】

这种关系是传递的也是完备的。

首先它是完备的。因为这三个人中任意给定两个人，比如 A 和 C，则必然有$A \geq C$或者$C \geq A$或者二者都成立。

其次它是传递的，因为例如如果$A \geq B$并且$B \geq C$，则必然有$A \geq C$。

3.假设有三个人 A,B 和 C，身高关系为“严格高于”。这样的关系是传递的吗？是反身的吗？是完备的吗？

【复习内容】二元关系

【参考答案】

首先它**不是**完备的，因为任给两个人，比如 A 和 C，我们不能肯定必然有$A > C$或$C > A$，因为这两人可能一样高。

其次它不是反身的，因为任意给定一个人比如 A，不可能有 A>A 成立。

再次它是传递的，因为比如如果 A>B 而且 B>C，则必然有 A>C。

4.某大学橄榄球教练说，任意给定两个前锋比如 A 和 B，他永远偏好身材更高大和速度更快的那个。他的这种偏好关系是传递的吗？是完备的吗？

【复习内容】二元关系

【参考答案】

首先它不是完备的。反证一下。假设它是完备的，则必然有（$A_1 \geq B_1$并且$A_2 \geq B_2$）或者（$B_1 \geq A_1$并且$B_2 \geq A_2$）。但是我们立刻可以举出反例，例如$A_1 > B_1$但$A_2 < B_2$，即 A 的身材更高但速度更慢，而 B 的身材更矮但速度更快，这种情形下选择谁？（注：此处我们用下标 1 表示身高；用下标 2 表示速度）

其次它是传递的，因为如果（$A_1 \geq B_1$并且$A_2 \geq B_2$），以及（$B_1 \geq C_1$并且$B_2 \geq C_2$），

则必然有（$A_1 \geq C_1$ 并且 $A_2 \geq C_2$）。

5.某条无差异曲线能否与自身相交？例如，图 3.2 能否是一条无差异曲线而不是两条？

【复习内容】无差异曲线的性质

【参考答案】

无差异曲线可以与自身相交（但两条无差异曲线不能相交）。

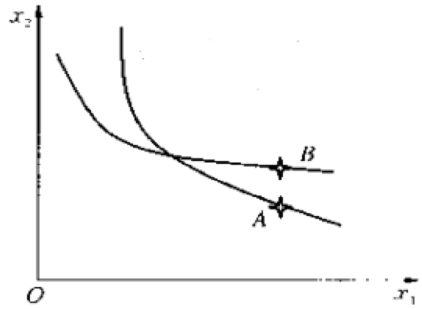
其实这个题目无任何意义，因为如果对消费者的偏好不加以假设限制，完全可以画出一条任何形状的无差异曲线，包括“这里的山路十八弯”类型。

因此，当然可以把教材图 3.2 看成一条无差异曲线而不是两条。

6.如果偏好是单调的，能否把图 3.2 看成一条无差异曲线而不是两条？

【复习内容】无差异曲线的性质

【参考答案】



如果偏好是单调的，则教材图 3.2 不能看成为一条无差异曲线。反证。

如果是一条无差异曲线（上图），则 A 和 B 无差异。由图可知，B 与 A 相比，商品 1 的数量相同，但商品 2 的数量较多，因为偏好是单调的，消费者更偏好 A。矛盾。所以不能把左图看成一条无差异曲线。

7.如果辣香肠和凤尾鱼都是厌恶品，那么无差异曲线的斜率为正还是负？

【复习内容】厌恶品；无差异曲线的斜率

【参考答案】

无差异曲线的斜率恒为负。

两种商品都是厌恶品的情形下，如果只给消费者增加一些辣香肠，则他的效用会减低，为使他仍在同一条无差异曲线上，就必须减少凤尾鱼的数量。类似地，如果只给消费者减少一些辣香肠，则他的效用会增加，为使他仍在同一条无差异曲线上，就必须增加凤尾鱼的数量。由此可见，一种商品增加的同时另外一种商品减少，因此无差异曲线的斜率恒为负。

曹乾（东南大学 caoqianseu@163.com）

8.解释为什么凸偏好意味着“平均束好于端点束”。

【复习内容】凸偏好

【参考答案】

因为凸偏好就是如此定义的。

例如，假设$(x_1, x_2) \sim (y_1, y_2)$，即 X 消费束和 Y 消费束位于同一条无差异曲线上，根据凸偏好的定义可知，对于任意t, $0 \le t \le 1$恒有

$$(tx_1 + (1-t)y_1, tx_2 + (1-t)y_2) \succeq (x_1, x_2)$$

$$(tx_1 + (1-t)y_1, tx_2 + (1-t)y_2) \succeq (y_1, y_2)$$

上两式中的左端是加权平均消费束，它由消费束 X 和 Y 加权平均得到，其中 X 商品束的权重为 t，Y 商品束的权重为 1-t；右端是端点消费束。

显然上述两个式子是说，加权平均消费束**不会差于**端点消费束。注意只有在偏好为**严格凸**时，我们才能断言加权平均消费束**好于**端点消费束。

但是，这样的假设是基于对人们消费行为观察基础上的，或者说这个假设符合我们的直觉，即人们在消费时不喜欢极端，而喜欢多样化。

9.面值 1 元的钞票与面值 5 元的钞票，计算它们之间的边际替代率。

【复习内容】边际替代率

【参考答案】

如果把面值 1 元的钞票和面值 5 元的钞票分别看成商品 1（画在横轴）和商品 2（画在纵轴），则由边际替代率的计算公式可知，

$$MRS_{12} = \frac{\Delta x_2}{\Delta x_1} = \frac{1}{-5} = -\frac{1}{5} \qquad (1)$$

$$= \frac{-1}{5} = -\frac{1}{5} \qquad (2)$$

；

时刻记住，边际替代率为无差异曲线的斜率，第一个式子是说减少 5 张 1 元的钞票，要增加 1 张 5 元的钞票才能使消费者还在原来的无差异曲线上；第二个式子是说增加 5 张 1 元的钞票，要减少 1 张 5 元的钞票才能使消费者还在原来的误差与曲线上。

读者可以思考这样的问题，在本题中，如果把 1 元的钞票看作商品 2，把 5 元的钞票看作商品 1，则边际替代率为多大？(答案-5)。

10.如果商品 1 是中性商品，计算用商品 1 替代商品 2 的边际替代率。

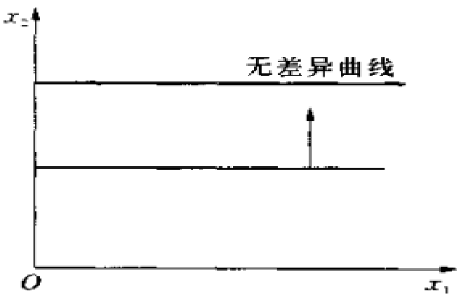
【复习内容】中性商品；边际替代率

【参考答案】

如果对某个消费者来说，商品 1 是中性商品，商品 2 是他喜欢的。则他的无差异曲线形状如左图所示。

之所以是水平线，是因为由中性商品的定义可知，消费者根本不关心它的数量，因此在本例中消费者只关注商品 2 的数量。

根据边际替代率的定义可知 $MRS_{12}=\dfrac{\Delta x_2}{\Delta x_1}=0$ 。



11.举例说明你的偏好在什么样的情形下为凹的。

【复习内容】凹偏好

【参考答案】

考察凹偏好的定义。从几何图形上看，凹偏好意味着加权平均消费束不会好于端点消费束。举出这样的例子并不困难，例如日常生活中我们不能或者不喜欢混搭着吃的食品就是最好的例子。

比方说，你日常服用维生素 C 片剂，而且喜欢吃虾，但有谣言说如果混搭着吃，就会在你体内合成砒霜。你还会混搭着吃吗？

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

4.效用（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

4 效用

在维多利亚时代，哲学家和经济学家轻率地把“效用”作为衡量一个人总体福利（well-being）的指标。他们用效用数值衡量一个人的幸福程度。在这种思想下，自然可认为消费者做出选择的目的是使他们的效用最大（使他们尽可能地幸福）。

问题在于这些古典经济学家从未真正地阐述过怎样衡量效用。我们应该如何量化不同消费选择下的效用“数额”？某人的效用与另外一人的效用相同吗？额外一颗糖块的效用是额外一根胡萝卜效用的两倍，这句话到底是什么意思？效用是人们希望最大化的“东西”，这里的“东西”是指什么？

由于这些概念难以准确界定，现代经济学家已经抛弃了上述老套的观点，即他们不再把效用看成幸福的衡量指标。取而代之的是，他们用**消费者偏好**（consumer preferences）重新改写了消费者行为理论，效用仅仅被当作为**一种描述偏好的方法**。

经济学家逐渐认识到，对于选择行为而言，效用最要紧的事情是一个商品束的效用是否比另外一个商品束的高，至于高多少并不重要。

起初，偏好是用效用定义的：说一个商品束(x_1, x_2)比另外一个商品束(y_1, y_2)更受偏好，表示商品束x比商品束y效用更高。但现在我们的观点正好反过来。消费者的**偏好**是研究选择行为的最基本工具，效用只是描述偏好的一种方法。

效用函数（utility function）是对每个可能的消费束都赋予数值的一种方法，这种方法要做到对于更受偏好的消费束，赋值更大。也就是说，消费束(x_1, x_2)比(y_1, y_2)更受偏好当且仅当前者的效用值要比后者大：用符号表示，$(x_1, x_2) \succ (y_1, y_2)$当且仅当$u(x_1, x_2) > u(y_1, y_2)$.

效用赋值的唯一重要性能是它是如何对商品束**排序**的。效用函数的重要性就在于它将不同商品束排了顺序。两个消费束效用值的差额有多大并不重要。由于强调排序，这种效用称为**序数效用**（ordinal utility）。

Bundle	U_1	U_2	U_3
A	3	17	−1
B	2	10	−2
C	1	.002	−3

表 4.1: 赋予效用值的不同方法

例如，在表 4.1 中，我们对三个商品束用三种方法赋予了效用值，在这些方法中，三个商品束的排序是相同的。在本例中，消费者偏好 A 胜过 B，偏好 B 胜过 C。这三种方法（三种效用函数）都描述了相同的偏好，都是可行的，因为它们对 A 的赋值大于 B 的，对 B 的

赋值大于 C 的。

由于重要的事情只是商品束的排列顺序，对商品束赋予效用值不可能只有一种方法。如果我们能找到赋予效用值的一种方法，我们就能找到无数多方法。如果$u(x_1,x_2)$代表对商品束(x_1,x_2)赋值的一种方法，那么将$u(x_1,x_2)$乘以 2（或者乘以任何正数）就是另外一种赋值方法。

将$u(x_1,x_2)$乘以 2 是**单调变换**（monotonic transformation）的一个例子。单调变换是将一组数字转换为另外一组数字的方法，这种方法要保留转换前后数字的顺序不变。

通常用函数 $f(u)$ 表示单调变换，它将每一个u的数值转换为其他的数值 $f(u)$，而且要保留数字的顺序，即$u_1>u_2$意味着 $f(u_1)>f(u_2)$。单调变换和单调函数是一回事。

单调变换的例子有：乘以一个正数（例如 $f(u)=3u$；加上任一常数（例如，$f(u)=u+17$）；变为自身的奇数次幂（例如，$f(u)=u^3$）；等等[一]。

$f(u)$随u变化的变化率可以这样衡量：先求 $f(u_2)$与 $f(u_1)$的差，然后除以u_2与u_1的差，即：

$$\frac{\Delta f}{\Delta u}=\frac{f(u_2)-f(u_1)}{u_2-u_1}.$$

对于单调变换来说，上式右端的分子与分母同号。因此，单调函数总有正的变化率。这表明，单调函数图形的斜率总是正的，如图 4.1A 所示。

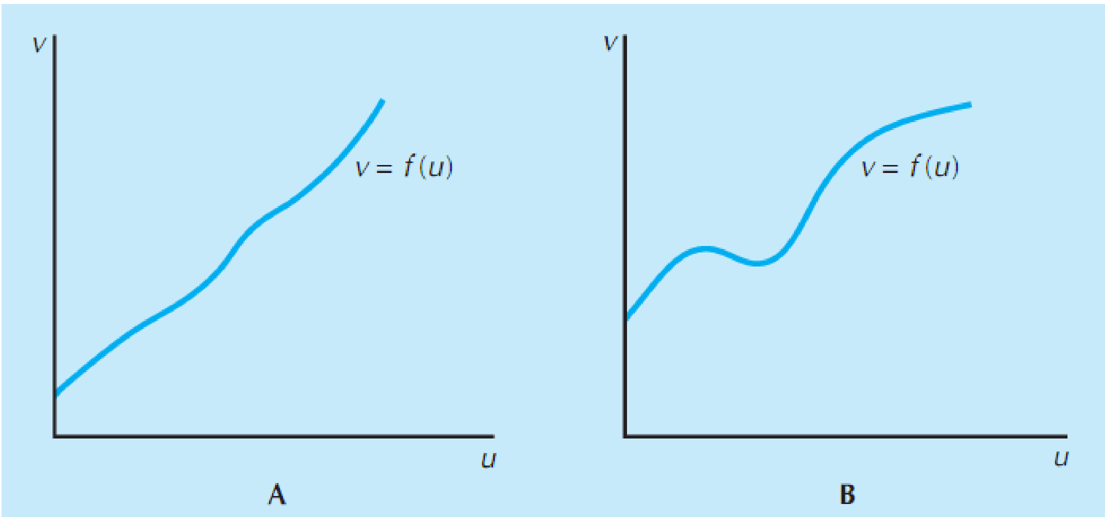


图 4.1:一个正单调变换。A 图为单调函数，它是递增的函数。B 图的函数不是单调的，因为它有时增有时减。

[一] 此处的单调变换，严格地说，应称为“正单调变换”（positive monotonic transformation），以便和“负单调变换”（negative monotonic transformation）区分开。负单调变换使得变换后数字的顺序与原顺序**相反**。单调变换有时称为“乏味的变换”（monotonous transformations），这似乎不公平，因为它们实际上很有趣。

如果 $f(u)$ 是任何一个效用函数（代表某种特定的偏好）的单调变换，那么 $f(u(x_1,x_2))$ 这个效用函数代表的偏好与上述偏好是相同的。

为什么？论证过程可以用下面三句话表示：

1. $u(x_1,x_2)$ 代表某个特定的偏好，这句话的意思是：$u(x_1,x_2)>u(y_1,y_2)$ 当且仅当 $(x_1,x_2)\succ(y_1,y_2)$。

2. 但是如果 $f(u)$ 是单调变换，那么：$u(x_1,x_2)>u(y_1,y_2)$ 当且仅当 $f(u(x_1,x_2))>f(u(y_1,y_2))$。

3.因此，$f(u(x_1,x_2))>f(u(y_1,y_2))$ 当且仅当 $(x_1,x_2)\succ(y_1,y_2)$，因此 $f(u)$ 代表的偏好与原效用函数 $u(x_1,x_2)$ 代表的偏好是一样的。

我们将以上的讨论总结为下列原理：**效用函数的单调变换是一个新的效用函数，它代表的偏好与原效用函数相同**。

在几何图形上，效用函数是标记无差异曲线的一种方法。因为位于同一条无差异曲线的所有商品束必然有相同的效用，所以可用效用函数为无差异曲线赋值，位置更高的无差异曲线，赋值应越大。从这个视角看，单调变换只是重新标记无差异曲线的一种方法。只要将含有更受偏好商品束的无差异曲线标记更大的数值，那么任何一种标记方法代表的偏好都是相同的。

4.1 基数效用

有些效用理论认为效用大小很重要，这样的理论称为**基数效用**（cardinal utility）理论。基数效用理论认为，两个商品束效用的差值大小多少有些意义。

我们已知道，如何判断某人是否偏好某个商品束胜过另一个：我们只要让他在这两个商品束之间做出选择即可。这样我们就知道如何对两个商品束赋予序数效用值：只要被选中的商品束的赋值大于放弃的那个商品束即可。任何这样的赋值方法都是一个效用函数。于是我们就得到了一个便于操作的判断标准，用于判断对某人来说一个商品束的效用是否高于另一个。

但是，我们如何判断某人喜欢一个商品束的程度是否是另一个的两倍？以你为例，你怎么判断自己喜欢一个商品束的程度是否是另一个的两倍？

你可能提出几种方法对上述情形进行赋值：对于更喜欢的那个商品束——我愿意支付两倍的价格；我愿意跑两倍的距离；我愿意等待两倍的时间；我愿意以两倍的输赢机率为它下注。

上述方法都没错，每种方法赋予的效用值都有一定程度的可操作性。但是这些方法也都不对。尽管每种方法都能说明喜欢一个商品束的程度是另一个的两倍这层意思，但是每种方

曹乾（东南大学 caoqianseu@163.com）

法的说服力都不够强大。

即使我们找到了一种能对上述情形赋值的方法，而且该方法很有说服力，它对我们描述选择行为有什么好处？为了判断是选择这个还是那个商品束，我们只要知道哪个商品束更受偏好（具有较高的效用值）即可。知道这个商品束的效用值比那个商品束大了多少，对于我们的分析选择行为毫无用处。既然我们分析选择行为时不需要基数效用，而且也没有好的方法对商品束赋予基数效用值，我们坚持完全使用序数效用理论。

4.2 构造效用函数

但是，我们真能找到赋予序数效用值的方法吗？给定一个偏好排序关系，我们总能找到某个效用函数使得它对商品束的排序和这个偏好排序相同吗？是否存在能刻画任何合理偏好排序的效用函数？

不是所有的偏好都能用效用函数刻画。例如，假设某人的偏好是非传递的，比如 $A \succ B \succ C \succ A$。如果某个效用函数能刻画这个偏好关系，那么它必然含有数值 $u(A)$,$u(B)$ 和 $u(C)$ 使得 $u(A) > u(B) > u(C) > u(A)$，但这是不可能的。

然而，只要我们去除掉非传递性偏好这样的反例，我们通常可以找到代表偏好的效用函数。此处我们构造一个效用函数，在第 14 章我们将构造另外一个。

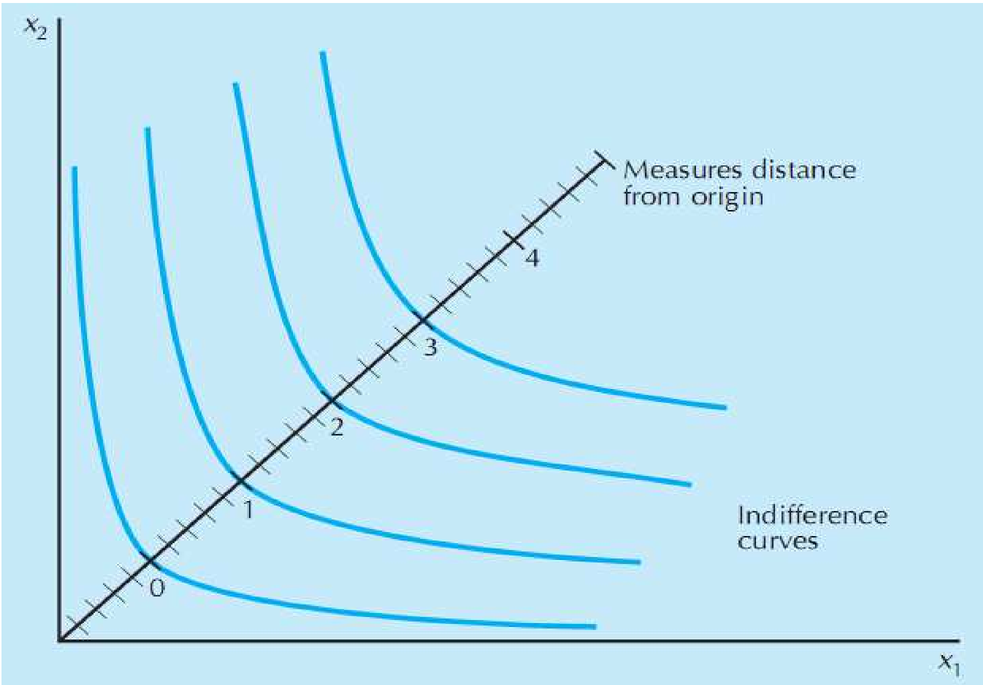


图 4.2：**从无差异曲线构造一个效用函数**。画出一条对角线，测量对角线与每条无差异曲线交点到原点的距离，将这些距离标记在相应的无差异曲线上。

假设我们有图 4.2 所示的无差异曲线。我们知道，效用函数是标记无差异曲线的一种方法，使得更高位置的无差异曲线赋值更大。我们如何进行标记？

一种简便的方法是画出一条如图所示的对角线，测量对角线与每条无差异曲线的交点到原点的距离，然后把这些距离标记在相应的无差异曲线上。

我们怎样知道这是一个效用函数？不难知道，如果偏好是单调的，那么对角线与每条无差异曲线的交点只有一个。因此，每个商品束都得到了标记，位于更高无差异曲线上的商品束标记的数值越大，这样就得到了一个效用函数。

这样我们就有了一种标记无差异曲线的方法，至少只要偏好是单调时该法可行。在具体情形中，这种方法未必是最自然的方法，但至少它表明了序数效用函数的通用性：几乎任何“合理的”偏好都能用效用进行刻画。

4.3 效用函数的一些例子

在第 3 章，我们已分析了一些偏好的例子，知道用什么样的无差异曲线表示这些偏好。我们也可以使用效用函数来表示这些偏好。如果给你一个效用函数 $u(x_1,x_2)$，画出一条无差异曲线就比较容易：你只要画出所有的 (x_1,x_2) 使得 $u(x_1,x_2)$ 等于一个常数即可。在数学上，满足 $u(x_1,x_2)$ 等于一个常数的所有 (x_1,x_2) 称为**水平集**（level set）。对于每个不同的常数值，你得到一条不同的无差异曲线。

例子：从效用到无差异曲线

假设效用函数为 $u(x_1,x_2)=x_1x_2$。那么它的无差异曲线的形状是什么样的？

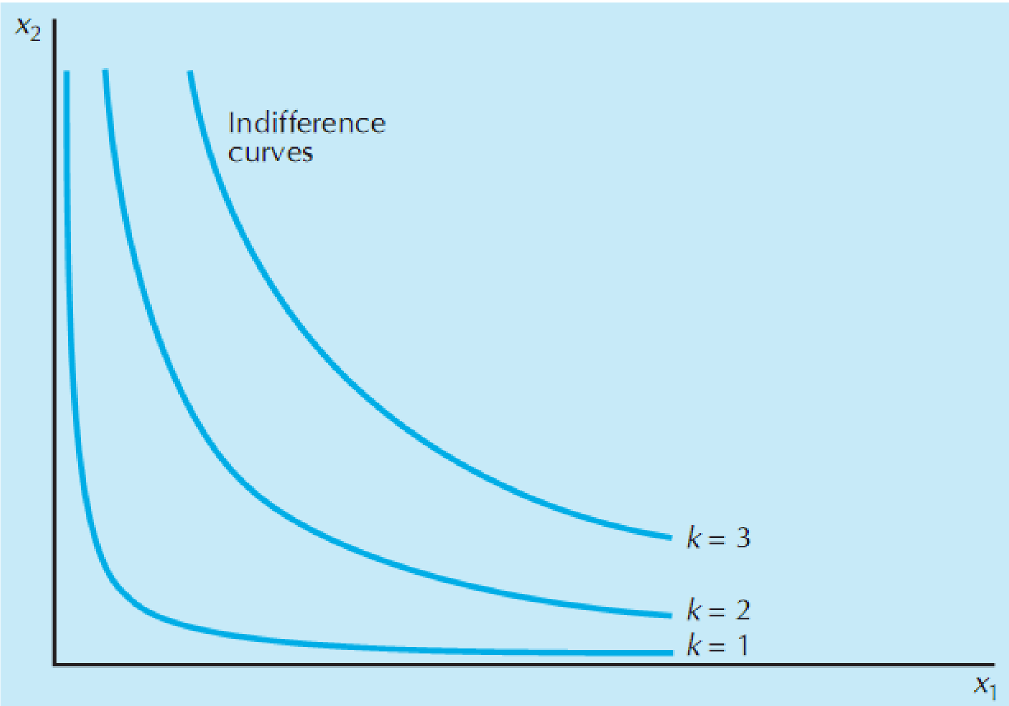


图 4.3：无差异曲线。不同 k 值时的无差异曲线 $k=x_1x_2$。

我们知道，无差异曲线是满足 $k=x_1x_2$（其中 k 为常数）的所有 x_1 和 x_2 的集合。将 x_2 视

为 x_1 的函数解之可得

$$x_2 = \frac{k}{x_1}$$

上面的式子就是典型无差异曲线的表达式。$k = 1,2,3$ 时的相应无差异曲线请见图 4.3.

我们分析另一个例子。假设效用函数为 $v(x_1, x_2) = x_1^2 x_2^2$。无差异曲线形状如何？由代数的标准法则可知

$$v(x_1, x_2) = x_1^2 x_2^2 = (x_1 x_2)^2 = [u(x_1, x_2)]^2$$

因此，效用函数 $v(x_1, x_2)$ 只是效用函数 $u(x_1, x_2)$ 的平方。因为 $u(x_1, x_2)$ 不能为负，所以 $v(x_1, x_2)$ 是效用函数 $u(x_1, x_2)$ 的一个单调变换。这表示效用函数 $v(x_1, x_2) = x_1^2 x_2^2$ 无差异曲线的形状，与图 4.3 中 $u(x_1, x_2)$ 的无差异曲线形状相同。这两个函数的无差异曲线的标记是不同的，原先的标记为 1，2，3…，现在则为 1，4，9…,但满足 $v(x_1, x_2) = 9$ 的商品束恰好就是满足 $u(x_1, x_2) = 3$ 的商品束。这样，$v(x_1, x_2)$ 代表的偏好和 $u(x_1, x_2)$ 代表的偏好是相同的，因为它们对所有商品束的**排序**是相同的。

将上面的事情反过来做，即根据一些无差异曲线找到效用函数，比较困难。有两种方法可做此事。一是数学方法。给定无差异曲线，我们试图找到一个函数，这个函数沿着每条无差异曲线的函数值必须为常数，而且更高的无差异曲线上的函数值必须更大。

第二种方法更直观一些。给定偏好的描述，我们考虑消费者试图最大化的是什么——什么样的商品束刻画了消费者的选择行为。这种方法暂时有些模糊，在我们分析几个例子之后，它就会明朗化。

完全替代

还记得红铅笔和蓝铅笔的例子吗？对消费者来说，只有铅笔总数才是最重要的。于是，自然可用铅笔总数衡量效用。因此我们暂时选择一个效用函数 $u(x_1, x_2) = x_1 + x_2$。这个函数可行吗？就问下面两个问题即可：该函数沿着无差异曲线的函数值是常数吗？更受偏好的商品束赋予了更大的数值了吗？这两个问题的答案都是肯定的，因此我们就得到了一个效用函数。

当然，这不是唯一可行的效用函数。我们也可以使用铅笔总数的平方。于是效用函数 $v(x_1, x_2) = (x_1 + x_2)^2 = x_1^2 + 2x_1 x_2 + x_2^2$ 同样可以表示完全替代的偏好，当然 $u(x_1, x_2)$ 的其他任何单调变换都可以表示这一偏好。

如果两种商品不是 1：1 替代的，结果会怎样？例如，假设消费者要求得到**两**单位的商品 2，他才愿意放弃一单位的商品 1。这表示对消费者来说，商品 1 的价值是商品 2 的**两倍**。因此效用函数的形式为 $u(x_1, x_2) = 2x_1 + x_2$。注意，由该效用函数产生的无差异曲线的斜率

为-2。

一般来说，完全替代类型的偏好可用下列形式的效用函数表示

$$u(x_1, x_2) = ax_1 + bx_2 .$$

此处，a 和 b 均为正数，分别表示商品 1 和 2 的“价值”。注意，它的无差异曲线的斜率为$-a/b$。

完全互补

这是左鞋和右鞋的情形。在这种偏好中，消费者只关心他有多少双鞋子，因此自然可以选择鞋子的双数作为效用函数。鞋子的双数取决于右鞋数量 x_1 和左鞋数量 x_2 中**最小**（minimum）的那个。于是，完全互补的效用函数采取的形式为$u(x_1, x_2) = \min\{x_1, x_2\}$。

为证实该函数实际可行，选择一个商品束比如（10, 10）。若商品 1 增加一单位则得到（11, 10），它应该和（10, 10）位于同一条无差异曲线上。这是因为$\min\{10,10\} = \min\{11,10\} = 10$。

因此，用效用函数$u(x_1, x_2) = \min\{x_1, x_2\}$刻画 1:1 完全互补的偏好是可行的。通常，该效用函数的任何单调变换同样可以刻画上述偏好。

如果完全替代商品不是 1:1 替代的，比如消费者每喝一杯茶时总是放入 2 茶匙糖，结果会如何？如果用 x_1 表示茶的杯数，x_2 表示糖的茶匙数，那么糖茶的杯数为$\min\{x_1, \frac{1}{2}x_2\}$。

这样的问题有点麻烦，所以我们停下来思考一下。如果茶的杯数大于糖的茶匙数的一半即 $x_1 > (1/2)x_2$，则不可能做到每杯茶里都放入 2 茶匙糖。在这种情形，糖茶的数量等于$(1/2)x_2$。（你可以用某些具体数字代替 x_1 和 x_2 验证一下。）

当然，该效用函数的任何单调变换仍然能刻画这个偏好。例如，我们可以乘以 2 消灭掉分数，这样得到的效用函数为$\min\{2x_1, x_2\}$。

一般地，刻画完全互补类型偏好的效用函数具有下列形式

$$u(x_1, x_2) = \min\{ax_1, bx_2\} .$$

其中a和b是正数，表示两商品的搭配比例。

拟线性偏好

我们以前未见过拟线性偏好（quasilinear utility）的无差异曲线。假设消费者的无差异曲线族是由无差异曲线互相垂直移动得到，如图 4.4 所示。这表明将一条无差异曲线垂直移动就可以得到所有的无差异曲线。由此可推知，无差异曲线的表达式为 $x_2 = k - v(x_1)$，其中k

为常数，不同的无差异曲线 k 值不同。这个式子是说，每条无差异曲线的高度等于 x_1 的函数 $-v(x_1)$ 加上某个常数 k 。k 值越大，无差异曲线位置越高。（$-v(x_1)$ 中的负号只是一种惯例做法；下面我们将知道为什么加了负号比较方便。）

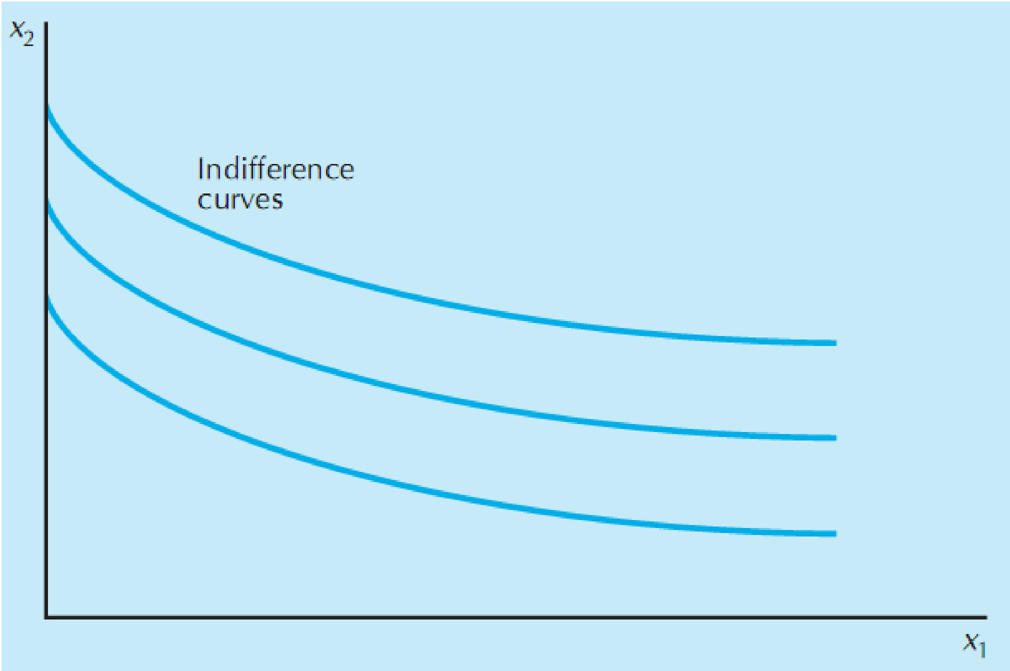


图 4.4:拟线性偏好。所有的无差异曲线都可由一条无差异曲线垂直移动得到。

标记这种无差异曲线自然的方法是使用 k 进行标记。大致来说，k 值是无差异曲线沿着纵轴的高度。解出 k 并令其等于效用，可得

$$u(x_1, x_2) = k = v(x_1) + x_2 .$$

上述效用函数对于商品 2 来说是线性的，但对于商品 1 来说（可能）是非线性的；因此，**拟线性效用**（quasilinear utility）的意思是“部分为线性”的效用。拟线性效用具体的例子有 $u(x_1, x_2) = \sqrt{x_1} + x_2$ 或者 $u(x_1, x_2) = \ln x_1 + x_2$ 等等。拟线性效用函数可能不太符合实际，但便于分析，我们在以后章节的例子中就会知道它比较简便。

柯布-道格拉斯偏好

另外一种常用的效用函数类型是**柯布-道格拉斯**（Cobb-Douglas）效用函数

$$u(x_1, x_2) = x_1^c x_2^d .$$

其中 c 和 d 为正数，表示消费者对两商品的偏好[一]。

[一] 保罗.道格拉斯（Paul Douglas）是 20 世纪的经济学家，在芝加哥大学任教，后来成为美国的参议员；查尔斯.柯布（Charles Cobb）为阿默斯特学院（Amherst College）的数学家。柯布-道格拉斯函数最初用于分析生产行为。

柯布-道格拉斯效用函数在一些情形下比较有用。柯布-道格拉斯效用函数描述的偏好通常具有图 4.5 所示的形状。在图 4.5A 中，我们画出了当 c=1/2，d=1/2 时的无差异曲线；在图 4.5B 中，画出了 c=1/5，d=4/5 时的无差异曲线。注意观测当参数 c 和 d 变化时无差异曲线的形状怎样变化。

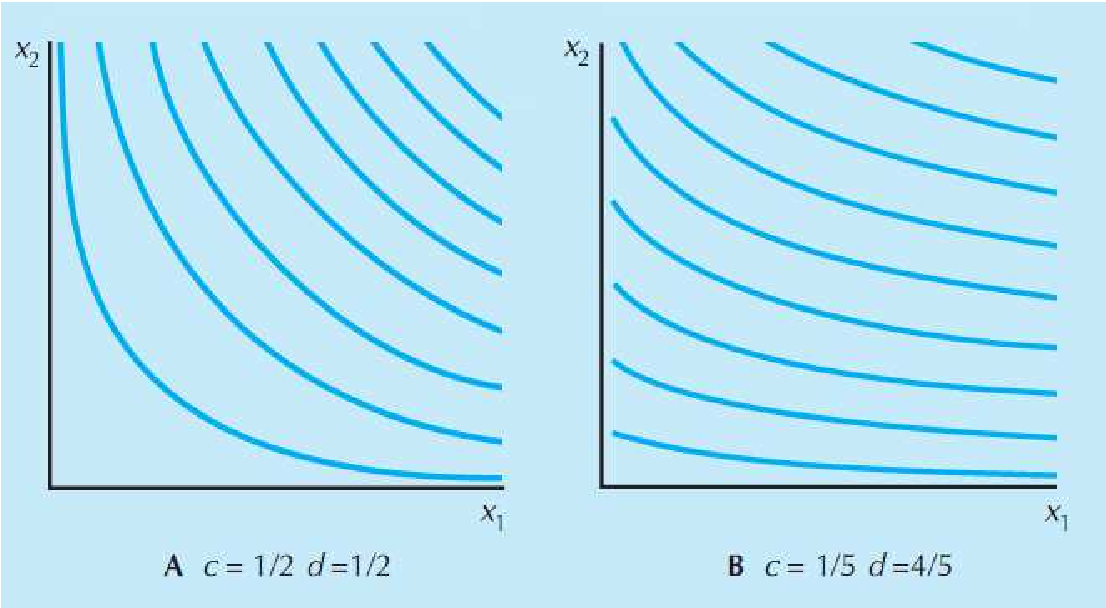


图 4.5: **柯布-道格拉斯无差异曲线**。当 c=1/2,d=1/2 时的无差异曲线(图 A); 当 c=1/5,d=4/5 时的无差异曲线（图 B)。

柯布-道格拉斯无差异曲线的形状类似于凸且单调的无差异曲线，我们在第 3 章把后者称为“具有良好性状的无差异曲线”。柯布-道格拉斯偏好是性状良好无差异曲线的标准例子，事实上，它的表达式大概是能产生良好性状偏好的最简单的代数表达式。在以后章节你会发现，使用柯布-道格拉斯函数表达经济思想很是方便。

当然，柯布-道格拉斯效用函数的任何单调变换也能表示同样的偏好，此处分析几个变换的例子比较有好处。

首先，如果取效用函数的自然对数，项的乘积将变成项的加和，于是我们有

$$v(x_1, x_2) = \ln(x_1^c x_2^d) = c \ln x_1 + d \ln x_2 .$$

该效用函数的无差异曲线的形状和 $u(x_1, x_2) = x_1^c x_2^d$ 的无差异曲线的形状类似，因为取对数是单调变换。(对自然对数知识的简要回顾，请见本书末的数学附录。)

第二个例子，假设柯布-道格拉斯效用函数为

$$u(x_1, x_2) = x_1^c x_2^d .$$

将效用变为原来的 $1/(c+d)$ 次幂，即

$$x_1^{\frac{c}{c+d}} x_2^{\frac{d}{c+d}}.$$

现在令 $a=\dfrac{c}{c+d}$，我们得到下列新的效用函数

$$v(x_1,x_2)=x_1^a x_2^{1-a}.$$

这意味着我们总可以将柯布-道格拉斯效用函数进行单调变换，使得指数的总和等于 1. 以后我们将看到这样的处理有个有用的解释。

柯布-道格拉斯效用函数的表达形式很多，你应该学会识别它们，因为这类偏好很有用。

4.4 边际效用

假设某消费者消费的商品束为 (x_1,x_2)，如果我们多给他一点商品 1，他的效用将怎样变化？这个变化比率称为商品 1 的**边际效用**（marginal utility）。我们将其记为 MU_1，它是一个比率

$$MU_1=\frac{\Delta U}{\Delta x_1}=\frac{u(x_1+\Delta x_1,x_2)-u(x_1,x_2)}{\Delta x_1}.$$

它衡量效用变化 Δu 与商品 1 数量变化 Δx_1 的比率[一]。

这个定义表明，为了计算商品 1 消费数量的微小变化引起的效用变化，只要将消费的变化量乘以该商品的边际效用即可：

$$\Delta U=MU_1\Delta x_1.$$

类似地，可以定义商品 2 的边际效用：

$$MU_2=\frac{\Delta U}{\Delta x_2}=\frac{u(x_1,x_2+\Delta x_2)-u(x_1,x_2)}{\Delta x_2}.$$

注意，当计算商品 2 的边际效用时，我们保持商品 1 的消费量不变。我们可以用下式计算商品 2 消费量变化引起的效用变化

$$\Delta U=MU_2\Delta x_2.$$

很重要的一点，是要知道边际效用的大小取决于效用的大小。因此，边际效用取决于我们选取的效用衡量方法。如果我们将效用（函数）乘以 2，则边际效用也乘以 2。这样处理我们仍可得到一个完全有效的效用函数，因为它表示同样的偏好，区别仅在于标记的数值不同。

这表明边际效用本身不含有选择行为的内容。我们如何根据消费者的选择行为计算边际

[一] 使用微积分处理边际效用的内容请见本章后的附录。

效用？无法计算。选择行为仅仅揭示了消费者如何将不同商品束排序的信息。边际效用取决于我们使用什么样的效用函数来反映偏好排序，边际效用本身没有特别意义。然而，我们可以使用边际效用，计算含有选择行为内同的其他事情，在下一节我们将看到这一点。

4.5 边际效用和边际替代率

可以使用效用函数$u(x_1, x_2)$计算第 3 章定义的边际替代率（MRS）。我们已知道 MRS 衡量无差异曲线在给定商品束那一点的斜率；可以认为 MRS 是一个比率，消费者恰好愿意按此比率用一定数量的商品 2 替代商品 1。

这样我们就有了计算 MRS 的一种简便方法。假设两种商品消费量的变化（$\Delta x_1, \Delta x_2$）恰好使效用不变，即消费量沿着一条无差异曲线变动。于是必然有

$$MU_1 \Delta x_1 + MU_2 \Delta x_2 = \Delta U = 0 .$$

解出无差异曲线的斜率，可得

$$MRS_{12} = \frac{\Delta x_2}{\Delta x_1} = -\frac{MU_1}{MU_2} .$$

(注意上式左端 2 在 1 的上边，而在上式右端 1 在 2 的上边。别搞混了！)

MRS 的代数符号为负：如果你增加商品 1 的消费那么你必须减少商品 2 的消费，才能使效用不变。然而，时刻牢记这个负号让人厌烦，因此经济学家通常将 MRS 的绝对值（即一个正数）当作 MRS。在不至于混淆的前提下，我们遵循这种惯例。

关于 MRS 的计算有个有趣的事情：可以通过观察某人的实际行为来计算 MRS——我们找到恰好使他不愿买也不愿卖的那个交换率，我们已在第 3 章探讨过这个问题。

效用函数，从而边际效用函数，不是唯一的。某个效用函数任何单调变换得到的效用函数同样可行。例如，如果我们将效用乘以 2，则边际效用也乘以 2。因此，边际效用函数的大小取决于我们主观选取的效用函数。边际效用不依赖于选择行为，而是依赖于我们用于刻画选择行为而选取的效用函数。

然而，边际效用比率提供了一个可观测的尺度，即边际替代率。对效用函数进行单调变换不会影响边际效用比率大小。举例验证，例如如果你将效用函数乘以 2，此时MRS_{12}变为

$$MRS_{12} = -\frac{2MU_1}{2MU_2} .$$

消去分子分母中的 2，可知边际替代率的大小没变。

这个结论对任何单调变换都成立。进行单调变换只是对无差异曲线重新标记，MRS 的计算仍然沿着同一条无差异曲线进行。尽管单调变换改变了边际效用数值，但边际效用的比率与你标记偏好的方法无关。

4.6 交通效用

效用函数在本质上是描述选择行为的方法：如果可买得起商品束 Y，但却选择了商品束 X，则 X 的效用一定比 Y 大。对消费者的选择行为进行分析，就可以估计出描述他们行为的效用函数。

上述想法广泛应用于交通经济学，用来研究消费者的交通行为。在多数大城市，人们可以选择乘公交车或者开车上班。每一种选择代表了一个消费束，这个消费束由不同的特征变量组成：行程时间，等待时间，花费的现金费用，舒适度，便利性等等。令 x_1 表示行程时间， x_2 表示等待时间，等等。

如果令 $(x_1, x_2, ..., x_n)$ 表示开车的 n 个不同特征变量，$(y_1, y_2, ..., y_n)$ 表示乘公交车的相应特征变量，我们可以用一个模型分析人们选择开车还是乘公交，他们做出选择的依据为是否更偏好其中的一个消费束。

更具体地说，假设典型消费者对上述特征变量的偏好可用下列效用函数表示

$$U(x_1, x_2, ..., x_n) = \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n .$$

其中，系数 $\beta_1, \beta_2, ... \beta_n$ 为 n 个未知参数。当然，上述效用函数的任何单调变换可以同样地描述人们的选择行为，但是线性形式的函数更容易用统计方法处理。

假设我们获得了若干消费者基于上述 n 个变量的交通方式选择的数据。可使用统计方法找到 β_i（i=1,…n）的数值，使得这些数值最好地拟合了消费者们的上述选择。这些统计方法可以估计出不同交通模式的效用函数。

研究发现效用函数具有下列形式[一]

$$U(TW, TT, C) = -0.147TW - 0.0411TT - 2.24C \qquad (4.2)$$

其中，TW=步行往返公交车或自驾车的总时间

TT=行程总时间（分钟）

C=行程总成本（元）

Domenich-McFadden 的书中估计出的效用函数，正确地刻画了家庭样本中 93%家庭的交通方式选择行为。

方程（4.2）中变量的系数刻画了典型家庭对交通行程不同特征变量赋予的权重；即每种特征变量的边际效用。两个系数之间的比率衡量两个特征变量之间的边际替代率。例如，步行往返时间的边际效用与行程总时间的边际效用的比值，表明对于典型消费者来说，步行

[一] 请见 Thomas Domenich and Daniel McFadden, Urban Travel Demand (North-Holland Publishing Company, 1975)。该书中的统计估计除了包括我们上述列举的变量外，还包括家庭的若干人口学特征。Daniel McFadden 因为研发了估计这类模型的方法，于 2000 年获得了诺贝尔经济学奖。

时间代表的劳累程度大约是行程时间劳累程度的 3 倍。换句话说，消费者为了少步行 1 分钟，愿意增加 3 分钟的行程。

类似地，行程总成本与行程总时间的比率，表明典型消费者对于这两个变量的权衡选择。在该研究中，消费者认为每分钟交通时间的价值为 0.0411/2.24=0.0183 元，即每小时的价值为 1.10 元。而同期（1967 年）的每小时工资大约为 2.85 元。

这样估计出的效用函数对于下列决策非常重要：对公交系统做出某些改变是否值得。例如，在上面的效用函数中，影响交通方式选择的一个显著变量是行程时间。城市交通管理部门可以花钱增加公交车的数量，以减少乘客行程时间。但是，因此增加的乘客能弥补公交系统额外增加的成本吗？

给定效用函数和消费者样本数据，我们可以预测出哪些消费者将自己开车，哪些消费者将搭乘公交车。这样我们就可以大体计算出收入是否能够弥补额外的成本。

而且，我们可以使用边际替代率估计行程时间减少对每个消费者的**价值**。在上面的交通行为研究中，我们已知道，1967 年人们认为交通时间的价值是每小时 1.10 元。因此，如果将行程减少 20 分钟，人们愿意支付 0.37 元。这个数字衡量了提供更准时公交服务的金钱收益。这个收益必须与提供更准时服务的成本相比较，以便确定这样的服务是否值得做。对收益进行量化，无疑有助于管理部门作出交通政策的合理决策。

附录

首先，我们解释一下“边际效用”的意思。在经济学中，“边际”表示导数。因此，商品 1 的边际效用就是

$$MU_1=\lim_{\Delta x_1\to 0}\frac{u(x_1+\Delta x_1,x_2)-u(x_1,x_2)}{\Delta x_1}=\frac{\partial u(x_1,x_2)}{\partial x_1}.$$

注意，此处我们使用的是偏导数，因为计算商品 1 的边际效用时，我们保持商品 2 的数量不变。

现在我们使用微积分重新推导课文中的 MRS。我们使用两种方法：第一种方法为微分法；第二种方法是使用隐函数。

对于微分法，假设消费量变动 (dx_1,dx_2) 使得效用不变，因此

$$du=\frac{\partial u(x_1,x_2)}{\partial x_1}dx_1+\frac{\partial u(x_1,x_2)}{\partial x_2}dx_2=0.$$

第一项表示由商品 1 的消费量微小变动（dx_1）引起的效用变动量；第二项的解释类似。我们选取的 dx_1 和 dx_2 恰好使效用变化量之和（du）等于 0。解出 dx_1/dx_2 可得

$$\frac{dx_1}{dx_2}=-\frac{\partial u(x_1,x_2)/\partial x_1}{\partial u(x_1,x_2)/\partial x_2},$$

这个式子正好等价于课文中的（4.1）式。

对于隐函数方法，假设无差异曲线可由函数 $x_2 = x_2(x_1)$ 刻画。也就是说，对于 x_1 的每个值，$x_2(x_1)$ 相应给出了 x_2 的值，恰好使所有的 (x_1, x_2) 位于给定的无差异曲线上。因此 $x_2(x_1)$ 必须满足下列恒等式

$$u(x_1, x_2(x_1)) \equiv k,$$

其中，k 表示那条给定的无差异曲线标记的效用值。

上面的恒等式两端分别对 x_1 微分，可得

$$\frac{\partial u(x_1,x_2)}{\partial x_1} + \frac{\partial u(x_1,x_2)}{\partial x_2}\frac{\partial x_2(x_1)}{\partial x_1} = 0.$$

注意，在恒等式 $u(x_1, x_2(x_1)) \equiv k$ 中，x_1 出现在两个地方，因此，x_1 的变动将对函数有双重影响，我们必须在 x_1 出现的地方都求导数。

在上面的微分等式中求出 $\partial x_2(x_1)/\partial x_1$，可得

$$\frac{\partial x_2(x_1)}{\partial x_1} = -\frac{\partial u(x_1,x_2)/\partial x_1}{\partial u(x_1,x_2)/\partial x_2},$$

这个式子和第一种方法（微分法）得到的式子是一样的。

隐函数方法稍微严格些，但是微分法比较直接，在计算时小心些即可。

假设我们对某个效用函数进行单调变换，比如 $v(x_1, x_2) = f(u(x_1, x_2))$。我们计算一下该效用函数的 MRS。使用链式法则

$$\begin{aligned} MRS_{12} = -\frac{\partial v/\partial x_1}{\partial v/\partial x_2} &= -\frac{\partial f/\partial u}{\partial f/\partial u}\frac{\partial u/\partial x_1}{\partial u/\partial x_2} \\ &= -\frac{\partial u/\partial x_1}{\partial u/\partial x_2} \end{aligned}$$

因为分子和分母中的 $\partial f/\partial u$ 可以约去，这表明 MRS 不受单调变换的影响。

上面这一结论，可以帮助我们识别不同效用函数代表的偏好是否相同：给定两个效用函数，只要分别计算边际替代率，看看它们是否相等。如果相等，那么这两个效用函数具有同样的无差异曲线。如果这两个效用函数的偏好增加的方向相同，那么潜在的偏好必然是一样的。

例子：柯布-道格拉斯偏好

使用上面推导出的式子，很容易就能计算出柯布-道格拉斯偏好的 MRS。

如果我们使用对数形式的表达式

$$u(x_1, x_2) = c\ln x_1 + d\ln x_2\ ,$$

则有

$$MRS_{12} = -\frac{\partial u(x_1, x_2)/\partial x_1}{\partial u(x_1, x_2)/\partial x_2}$$
$$= -\frac{c/x_1}{d/x_2} = -\frac{c}{d}\frac{x_2}{x_1}$$

注意，MRS 仅取决于两个参数(c 和 d)的比率以及两种商品的数量。

如果我们选取的是下列指数表达式，MRS 为多大？

$$u(x_1, x_2) = x_1^{\ c} x_2^{\ d}$$

计算 MRS 如下：

$$MRS_{12} = -\frac{\partial u(x_1, x_2)/\partial x_1}{\partial u(x_1, x_2)/\partial x_2}$$
$$= -\frac{cx_1^{c-1}x_2^d}{dx_1^c x_2^{d-1}} = -\frac{cx_2}{dx_1}$$

这个结果和前面的计算结果是一样的。当然，你自始至终都知道，单调转换不会改变边际替代率！

总结

1.效用函数只是偏好排序的一种表示方法。效用水平的数值大小没有实质含义。

2.因此，给定任何一个效用函数，它的任何单调变换仍然表示相同的偏好。

3. 边际替代率 MRS，可由效用函数计算出，计算公式为 $MRS_{12} = \Delta x_2/\Delta x_1 = -MU_1/MU_2$.

复习题

1.课文中说，将某数字变为它的奇次幂是一种单调变换。那么，将其变为它的偶次幂是单调变换吗？（提示：要考虑类似 $f(u) = u^2$ 的情形。）

曹乾（东南大学 caoqianseu@163.com）

2.下列哪些是单调变换？（1）$u=2v-13$；（2）$u=-1/v^2$；（3）$u=1/v^2$；（4）$u=\ln v$；（5）$u=-e^{-v}$；（6）$u=v^2$；（7）$u=v^2$（其中$v>0$）；（8）$u=v^2$（其中$v<0$）。

3.课文中有个结论，即如果偏好是单调的，那么经过原点的对角线与每条无差异曲线只能有一个交点。你能严格证明这个结论吗？（提示：如果它与某条无差异曲线有两个交点，结果会如何？）

4. 效用函数 $u(x_1,x_2)=\sqrt{x_1+x_2}$ 表示什么类型的偏好？效用函数 $v(x_1,x_2)=13x_1+13x_2$ 呢？

5. 效用函数 $u(x_1,x_2)=x_1+\sqrt{x_2}$ 表示什么类型的偏好？效用函数 $v(x_1,x_2)=x_1^2+2x_1\sqrt{x_2}+x_2$ 是 $u(x_1,x_2)$ 的单调变换吗？

6.假设效用函数为 $u(x_1,x_2)=\sqrt{x_1x_2}$。它表示什么类型的偏好？函数 $v(x_1,x_2)=x_1^2x_2$ 是 $u(x_1,x_2)$ 的单调变换吗？$w(x_1,x_2)=x_1^2x_1^2$ 是 $u(x_1,x_2)$ 的单调变换吗？

7.你能说明为什么对一个效用函数进行单调变换不会改变边际替代率？

复习题答案

1.课文中说，将某数字变为它的奇次幂是一种单调变换。那么，将其变为它的偶次幂是单调变换吗？（提示：要考虑类似 $f(u)=u^2$ 的情形。）

【复习内容】单调变换。

【解题思路】

作者范里安在教材中指出：“单调变换是将一组数字转换为另外一组数字的方法，这种方法要保留转换前后数字的顺序不变。…单调变换和单调函数是一回事。”换一种表述方法是：如果某既定偏好可用函数 $u=u(x_1,x_2)$ 进行刻画，则可以对函数 u 进行复合（构造复合函数）比如 $f(u)=f(u(x_1,x_2))$，但要保证这个复合函数是正单调函数（即单调递增的函数），即对 $u=u(x_1,x_2)$ 进行正单调变换。所以说正单调变换和正单调函数是一回事。

因此，我们对原效用函数并不要求它是单调递增的，但对这个函数进行正单调变换后得到的新效用函数一定是单调递增的。原因在于正单调变换和正单调函数是一回事。

也正因为此，题目中的“将某数字变为它的奇次幂是一种单调变换”这种说法并不严格，因为当幂指数等于 1 时，这正是原效用函数本身，但它不是正单调变换，否则这意味着要求

原效用函数必须为单调的。而我们未必一定要求原效用函数是单调的，尽管我们通常这么要求。所以，幂指数应将 1 除外。

通常情况下，$u = u(x_1, x_2) \geq 0$，因此 $f(u) = u^2$ 是（正的）单调变换，也就是说 $f(u) = u^2$ 是单调递增的函数。

但是，也有可能存在 $u = u(x_1, x_2) \leq 0$ 的情形。举个例子，给某消费者两种商品，但这两种商品都是他非常讨厌的，由于这种情形下，他的效用不可能为正，即 $u = u(x_1, x_2) \leq 0$。所以该情形下 $f(u) = u^2$ 就不是（正的）单调变换，但它是（负的）单调变换，但根据我们的目的，我们不考虑负单调变换的情形。

【参考答案】见上述解题思路中的最后两段文字。

2.下列哪些是单调变换？（1）$u = 2v - 13$；（2）$u = -1/v^2$；（3）$u = 1/v^2$；（4）$u = \ln v$；（5）$u = -e^{-v}$；（6）$u = v^2$；（7）$u = v^2$（其中 $v > 0$）；（8）$u = v^2$（其中 $v < 0$）。

【复习内容】单调变换。

【参考答案】

（1）是（正的）单调变换。因为复合函数 u 是函数 v 的单调递增函数。判断依据：$du/dv = 2 > 0$。以下题目的原因请类推。

（2）$v > 0$ 时是（正的）单调变换，$v < 0$ 时是（负的）单调变换（即变换后得到的函数 u 是递减函数），由于我们不考虑负单调变换的情形，因此自此以后凡是说到单调变换就是指正单调变换。以下各题不再一一说明。

（3）$v < 0$ 时是单调变换，$v > 0$ 时不是。

（4）是单调变换（顺便指出，此题暗含着 $v > 0$ 的假设，否则 $u = \ln v$ 无定义）。

（5）是单调变换。

（6）$v \geq 0$ 时是单调变换，$v \leq 0$ 时不是。

（7）和（8）请见（6）。

3.课文中有个结论，即如果偏好是单调的，那么经过原点的对角线与每条无差异曲线只能有一个交点。你能严格证明这个结论吗？（提示：如果它与某条无差异曲线有两个交点，结果会如何？）

【复习内容】单调偏好的定义；无差异曲线的特征

【参考答案】

反证法。假设这条对角线与一条无差异曲线交于两个不同的点：(x_1, x_2)；(y_1, y_2)。由于这两个点在同一条无差异曲线上，因此 $(x_1, x_2) \sim (y_1, y_2)$。另外，这两个点都在该对角线上，因此必有 $x_1 > y_1, x_2 > y_2$ 或者 $x_1 < y_1, x_2 < y_2$。不妨设 $x_1 > y_1, x_2 > y_2$，由于偏好是单

调的，这意味着$(x_1,x_2)\succ(y_1,y_2)$。显然$(x_1,x_2)\succ(y_1,y_2)$和刚才得到的$(x_1,x_2)\sim(y_1,y_2)$矛盾。所以如果偏好是单调的，那么经过原点的对角线与每条无差异曲线只能有一个交点。

4.效用函数$u(x_1,x_2)=\sqrt{x_1+x_2}$表示什么类型的偏好？效用函数$v(x_1,x_2)=13x_1+13x_2$呢？

【复习内容】单调变换；完全替代类型的偏好

【参考答案】

对效用函数$u(x_1,x_2)$作单调变换$f(u)=u^2$，可得新的效用函数$f(u)=u^2=x_1+x_2$；不妨令$w(x_1,x_2)=x_1+x_2$。

对效用函数$v(x_1,x_2)$作单调变换$g(v)=\dfrac{1}{13}v$，可得新的效用函数$g(v)=\dfrac{1}{13}v=x_1+x_2=w(x_1,x_2)$。

由于单调变换不改变原有的偏好关系，因此可知$u(x_1,x_2)=\sqrt{x_1+x_2}$，$v(x_1,x_2)=13x_1+13x_2$和$w(x_1,x_2)=x_1+x_2$这三个效用函数代表的偏好关系是相同的。而我们已知道$w(x_1,x_2)=x_1+x_2$表示 1:1 完全替代的偏好类型。所以$u(x_1,x_2)=\sqrt{x_1+x_2}$和$v(x_1,x_2)=13x_1+13x_2$都表示 1:1 完全替代的偏好类型。

5.效用函数$u(x_1,x_2)=x_1+\sqrt{x_2}$表示什么类型的偏好？效用函数$v(x_1,x_2)=x_1^2+2x_1\sqrt{x_2}+x_2$是$u(x_1,x_2)$的单调变换吗？

【复习内容】拟线性偏好；单调变换。

【参考答案】

$u(x_1,x_2)=x_1+\sqrt{x_2}$表示拟线性的偏好类型。

因为$x_1\geq 0,x_2\geq 0$（消费数量不能为负），所以$u(x_1,x_2)=x_1+\sqrt{x_2}\geq 0$，对效用函数$u(x_1,x_2)$做单调变换$f(u)=u^2$（请参考第 2 题第（6）小题的答案），可得$f(u)=x_1^2+2x_1\sqrt{x_2}+x_2$，而这正是效用函数$v(x_1,x_2)$，因此$v(x_1,x_2)=x_1^2+2x_1\sqrt{x_2}+x_2$是$u(x_1,x_2)$的单调变换。

6.假设效用函数为$u(x_1,x_2)=\sqrt{x_1x_2}$。它表示什么类型的偏好？函数$v(x_1,x_2)=x_1^2x_2$是$u(x_1,x_2)$的单调变换吗？$w(x_1,x_2)=x_1^2x_1^2$是$u(x_1,x_2)$的单调变换吗？

【复习内容】柯布—道格拉斯型偏好；单调变换

【参考答案】

（1）$u(x_1,x_2)=\sqrt{x_1x_2}$表示柯布—道格拉斯型偏好。

（2）判断一个函数是否为另一个函数的单调变换，有时使用一个小技巧反而更方便。这个小技巧就是第 7 题的结论，即对一个效用函数进行单调变换不会改变边际替代率。$u(x_1,x_2)=\sqrt{x_1x_2}$ 的边际替代率 $MRS_{12}=-\dfrac{MU_1}{MU_2}=-\dfrac{x_2}{x_1}$；$v(x_1,x_2)=x_1^2x_2$ 的边际替代率 $MRS_{12}=-\dfrac{MU_1}{MU_2}=-\dfrac{2x_2}{x_1}$，二者不等，因此 $v(x_1,x_2)=x_1^2x_2$ 不是 $u(x_1,x_2)$ 的单调变换。

（3）因为 $u(x_1,x_2)=\sqrt{x_1x_2}\geq 0$（请参考第 2 题第（6）小题的答案），所以对该函数作单调变换 $f(u)=u^4$ 可得 $f(u)=u^4=x_1^2x_2^2=w(x_1,x_2)$，所以答案为:是。

7.你能说明为什么对一个效用函数进行单调变换不会改变边际替代率？

【复习内容】单调变换；边际替代率的序数性质

【参考答案】

假设我们对某个效用函数进行单调变换，比如 $v(x_1,x_2)=f(u(x_1,x_2))$。我们计算一下该效用函数的 MRS_{12}。使用链式法则

$$\begin{aligned}MRS_{12}&=-\frac{\partial v/\partial x_1}{\partial v/\partial x_2}=-\frac{\partial f/\partial u}{\partial f/\partial u}\frac{\partial u/\partial x_1}{\partial u/\partial x_2}\\&=-\frac{\partial u/\partial x_1}{\partial u/\partial x_2}\end{aligned}$$

因为分子和分母中的 $\partial f/\partial u$ 可以约去，这表明效用函数进行单调变换不会改变边际替代率。这也就是说边际替代率具有序数性质。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

5.选择（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

5 选择

在本章，我们将把预算集和偏好理论放在一起，以便研究消费者的最优选择。我们之前说过，刻画消费者选择行为的经济模型，是指人们在能买得起的商品束中选择最好的。现在我们把这句话改成听起来更专业的说法，即“消费者从他们的预算集中选择最受偏好的商品束。”

5.1 最优选择

图 5.1 是一种典型情形。此处我们在同一张图中，画出了某个消费者的预算集和几条无差异曲线。我们想找到位于最高无差异曲线上的商品束。因为偏好是良好性状的，商品束中商品数量越多越受偏好，所以我们将注意力集中在预算线上（on）的商品束，而不必考虑在预算线下方（beneath）的商品束。

现在我们从预算线的右端开始逐渐向左端移动。在沿着预算线移动的过程中，我们注意到：我们接近的无差异曲线位置越来越高。在到达位置最高的无差异曲线（恰好触及预算线）时，我们停下来。将这两条曲线触及点的商品束标记为 (x_1^*, x_2^*)，见图 5.1。

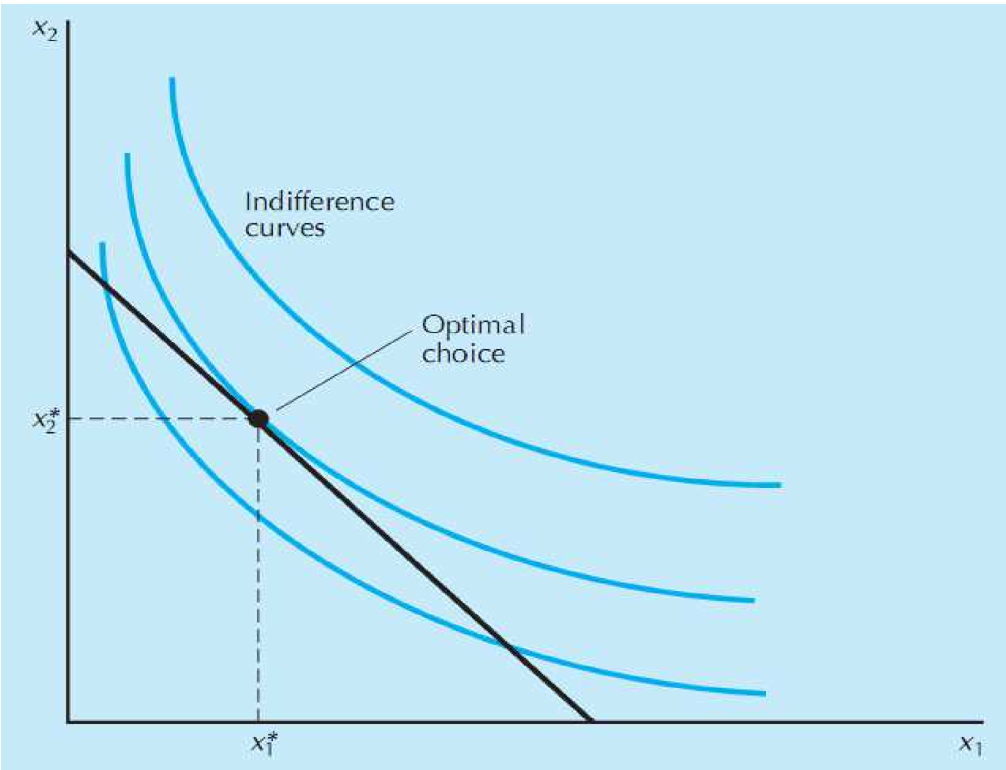


图 5.1：最优选择。最优消费选择是无差异曲线和预算线的切点处的商品束。

(x_1^*, x_2^*) 是消费者的**最优选择**（optimal choice）。比 (x_1^*, x_2^*) 好的商品束集合，位于恰好

 曹乾（东南大学 caoqianseu@163.com）

触及预算线的那条无差异曲线的上方（这些消费束他买不起）；比(x_1^*, x_2^*)差的商品束集合，位于预算线的下方（这些商品束他能买得起）。上述两个商品束集合不相交，因此(x_1^*, x_2^*)是他能买得起的商品束之中最好的。

注意最优消费束的重要特征：在最优选择处，无差异曲线和预算线相切。稍加思索你就会明白为何需要相切：如果无差异曲线和预算线不相切，那它就会预算线相交[一]，沿着预算线在交点附近的某些点就会位于这条无差异曲线的上方，这表明交点不是最优消费束。

在最优选择处，相切条件必须满足吗？事实上不是所有情形都要求必须相切，但对于良好性状的偏好，要求相切。无论如何，无差异曲线和预算线的交点都不可能是最优选择。那么，“不相交”何时意味着相切？我们先研究一种例外的情形。

首先，某些形状的无差异曲线可能没有切线，如图 5.2 所示。无差异曲线在最优选择处出现了拐折（kink），此时不存在切线，因为切线的定义要求每一点只有唯一的切线。这种情形的经济学意义不大，它是一种最讨厌的情形。

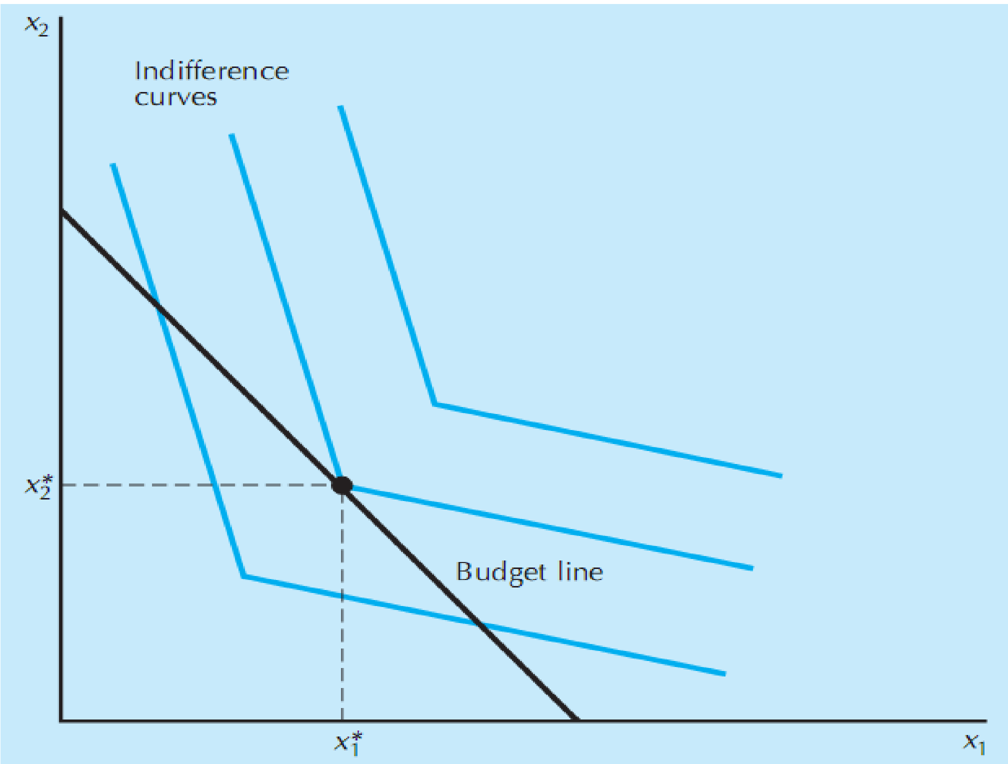


图 5.2 拐折的偏好。在最优选择处，无差异曲线无切线。

第二种例外比较有趣。假设在最优点，某种商品的消费量为零，如图 5.3 所示。因此，无差异曲线的斜率与预算线斜率不相等，无差异曲线与预算线不相交。我们说图 5.3 表示的

[一] 如果不相切也不相交，即无差异曲线位于预算线的上方，这条无差异曲线代表的消费束，消费者都买不起。

曹乾（东南大学 caoqianseu@163.com）

是**边界最优**（boundary optimum），而图 5.1 表示的是**内部最优**（interior optimum）。

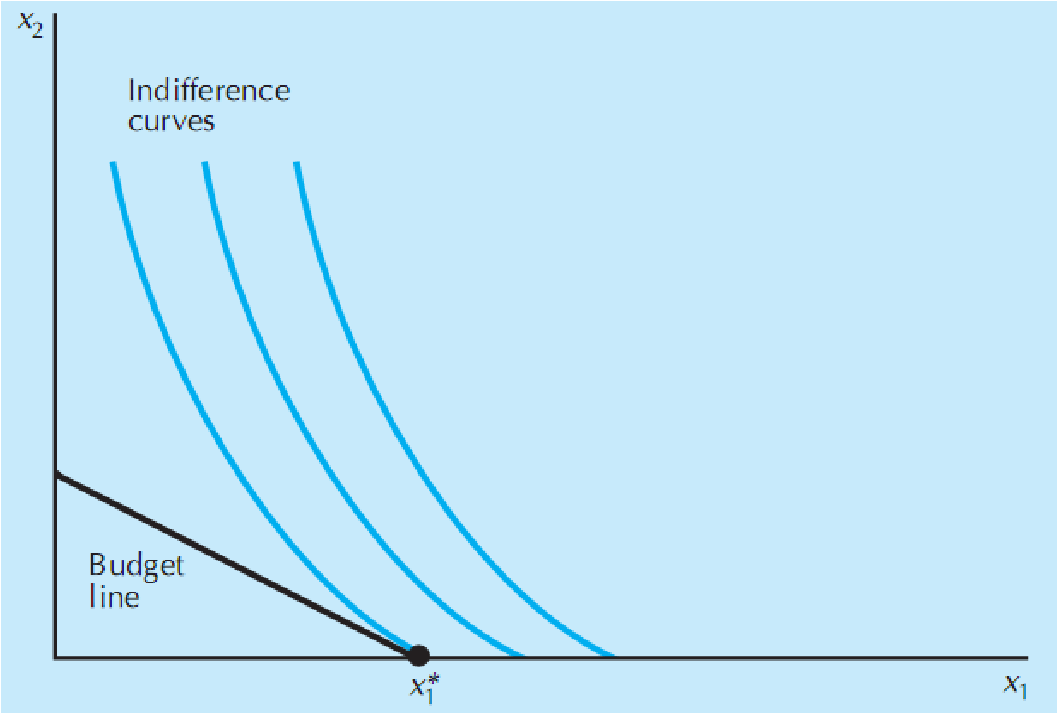


图 5.3：边界最优。最优选择处，商品 2 的消费量为 0。无差异曲线与预算线不相切。

如果我们愿意除掉“拐折的偏好”，我们可以不用再考虑图 5.2 所示的情形[一]。更进一步，如果我们仅考虑**内部最优**，我们可排除边界最优等其他情形。如果平滑的无差异曲线有内部最优解，则无差异曲线的斜率和预算线的斜率必须相等…因为如果它们不相等，无差异曲线将与预算线相交，交点不是最优点。

我们已经发现最优选择必须满足的必要条件。如果最优商品束中两种商品的数量都大于 0，即它是内部最优而不是边界最优，那么无差异曲线必然与预算线相切。但是，相切条件是某商品束为最优商品束的**充分条件**吗？即，如果无差异曲线与预算线相切，那么切点一定是最优选择吗？

请看图 5.4。图中有三个商品束满足相切条件，它们都位于预算线内部，但三个切点中只有两个切点是最优选择。因此，一般来说，相切的条件只是最优化的必要条件而不是充分条件。

然而，在凸偏好的情形下，相切的条件是最优化的充分条件。凸偏好是一种比较重要的情形。如果偏好为凸，任何满足相切条件的点必定是最优点。这个结论在几何图形上是直观的：由于凸无差异曲线必然弯曲地远离预算线，它不可能再拐回头与无差异曲线相切。

[一] 否则，本书有可能被列入 R 级（限制级）。以下为译者注： kinky taste（拐折的偏好），另外的一个意思是“性变态的爱好”，所以才有限制级这种戏说。

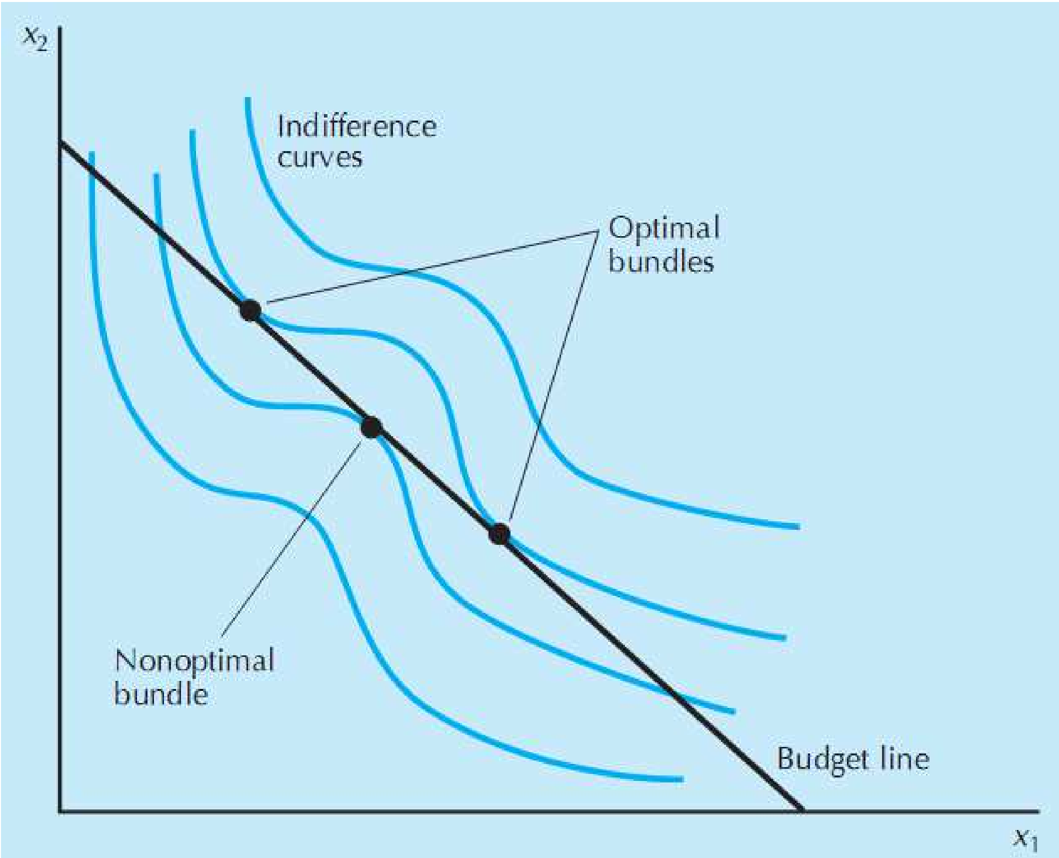


图 5.4：**多个切点的情形**。图中有三个切点，但其中只有两个为最优点，因此相切的条件是必要条件而不是充分条件。

由图 5.4 可知，通常有多个最优商品束满足相切条件。然而，凸偏好限制了上述情形的发生。如果无差异曲线为**严格**凸，即无差异曲线上不含有直线段，那么每条预算线上仅有唯一的最优点。尽管可用数学证明这个结论，但借助图形也可说明这个结论是很合理的。

在内部最优点，MRS 必须与预算线相等，这个条件在图形上表现比较明显，但是它的经济学意义是什么？我们知道 MRS 的其中一种解释是，消费者恰好不愿买也不愿意卖时两种商品的交换比率。市场提供给消费者的交换比率是$-p_1/p_2$，即：如果你放弃一单位商品 1，你可以购买 p_1/p_2 单位商品 2. 如果对于某个消费束，消费者不愿意买也不愿意卖，则必然有 MRS 等于商品的交换率：

$$MRS_{12}=-\frac{p_1}{p_2}$$

另外一种分析方法是，假设 MRS 不等于两商品的价格之比，看看将会发生什么情形？例如假设 $MRS_{12}=\Delta x_2/\Delta x_1=-1/2$，两商品的价格比率为 1/1。前者表明消费者为得到一单位商品 2，恰好愿意放弃两单位的商品 1；而后者表示市场愿意按 1：1 的比例交换两种商

品。因此，消费者当然愿意放弃一些商品 1，多购买一些商品 2。只要 MRS 不等于两商品的价格之比，消费者不可能已实现最优选择。

5.2 消费者需求

在某组价格和消费者的收入水平下，商品 1 和 2 的最优选择称为消费者的**需求束**（demanded bundle）。通常当价格和收入改变时，消费者的最优选择也会改变。**需求函数**（demand function）是将最优选择（即商品的需求数量）与不同价格和收入水平关联起来的函数。

我们将需求函数写为依赖于两种商品价格和消费者收入的函数：$x_1(p_1, p_2, m)$ 和 $x_2(p_1, p_2, m)$。对每组不同的价格和收入，消费者最优选择的商品组合也不同。不同的偏好对应不同的需求函数；很快我们就会分析一些例子。在接下来的几章，我们的主要任务是研究这些需求函数下的消费者行为，即最优选择如何随价格和收入变动而变动的。

5.3 一些例子

我们将使用第 3 章介绍的那些偏好类型来研究消费者的选择。每个例子分析的基本程序是相同的：画出预算线和若干无差异曲线，找到位置最高的那条无差异曲线和预算线的触及点。

完全替代

图 5.5 显示了 1：1 完全替代的情形。最优选择有三种可能性。如果 $p_2 > p_1$，那么预算线的将比无差异曲线平坦。在这种情形下，最优选择是消费者将所有的钱都购买商品 1。如果 $p_1 > p_2$，则消费者只会购买商品 2。最后一种情形，如果 $p_1 = p_2$，那么将有一些列的最优选择，即满足预算线的所有商品 1 和 2 的组合都是最优的。因此，商品 1 的需求函数为

$$x_1 = \begin{cases} m/p_1 & p_1 < p_2; \\ x_1 \in [0 \quad m/p_1] & p_1 = p_2; \\ 0 & p_1 > p_2. \end{cases}$$

这些结果和符合常识吗？这些结果是说如果两种商品完全替代，则消费者只购买比较便宜的那一种商品。如果两种商品价格相等，则消费者并不关心他购买了哪种商品。

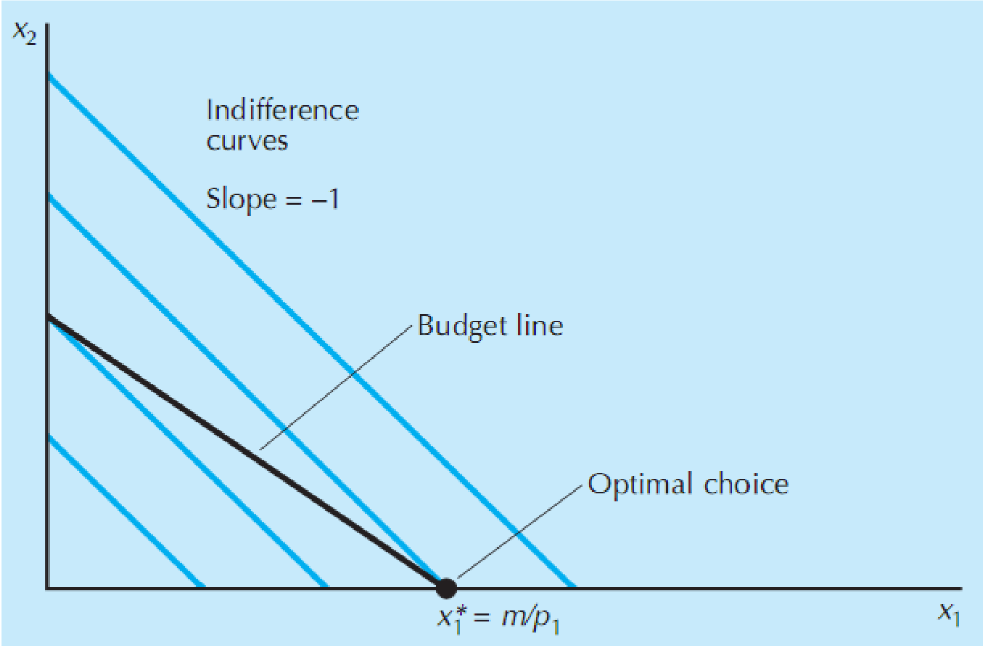


图 5.5：**完全替代情形的最优选择**。如果两种商品是完全替代的，最优选择通常位于边界上。

完全互补

两商品 1：1 完全互补的情形见图 5.6. 由图可知，最优选择必然位于对角线上，即无论价格如何，消费者购买的两种商品的数量是相等的。对这个例子来说，这表明有两只脚的人购买的鞋子都是成对的 [一]。

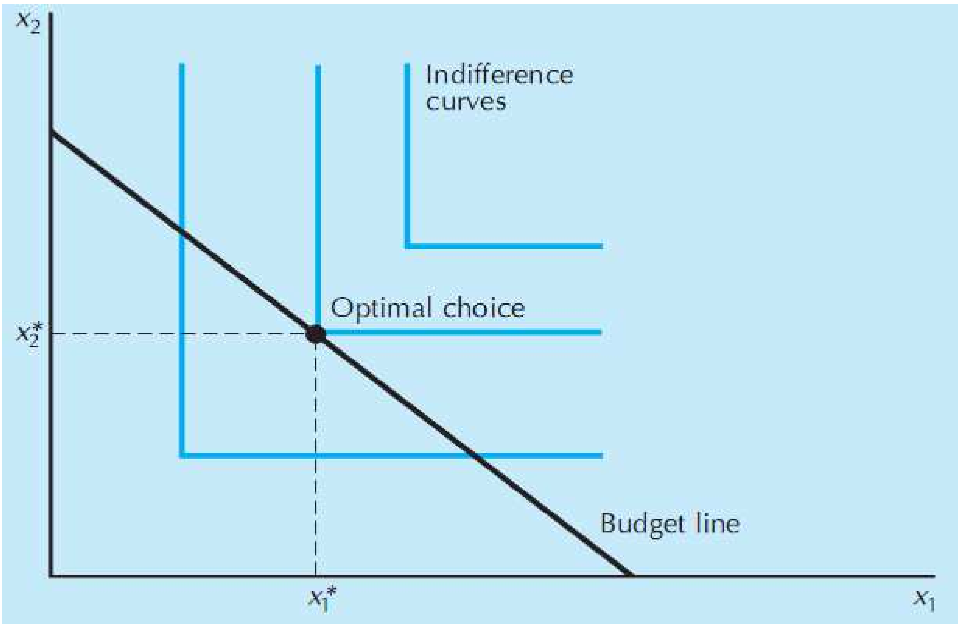


图 5.6：完全互补情形下的最优选择。如果两种商品是 1：1 完全互补的，则需求束总位于对角线上，这是因为最优选择时 $x_1 = x_2$

[一] 别担心，以后我们还会给出更有趣的例子。

下面我们用代数方法分析最优选择问题。我们知道消费者不管价格如何都会购买相同数量的商品 1 和 2。令 x 表示该数量，则我们必须满足下列预算约束

$$p_1x_1 + p_2x_2 = m .$$

解出 x 就得到商品 1 和 2 的最优选择

$$x_1 = x_2 = x = \frac{m}{p_1 + p_2} .$$

此处最优选择的需求函数非常直观。由于两种商品总是被一起消费，可以把它们视为一种整体商品，消费者将全部资金用于购买这种商品，它的价格为 $p_1 + p_2$。

中性商品和厌恶商品

如果商品束中含有中性商品（neutral good），消费者将会把全部资金用于购买他喜欢的商品，购买中性商品数量为零。对于厌恶品（bad），上述结论仍然成立。因此，如果商品 1 是消费者喜欢的商品(good)，商品 2 为厌恶品，则它们的需求函数分别为：

$$x_1 = \frac{m}{p_1}$$
$$x_2 = 0.$$

离散商品

假设商品 1 是离散商品（discrete good），离散商品只能以整数计数，而商品 2 是花费在所有其他商品上的资金。如果消费者购买 1，2，3，…单位的商品 1，这意味着他选择的消费束为 $(1, m - p_1), (2, m - 2p_1), (3, m - 3p_1)$，等等。只要比较一下上述消费束的效用，看看哪个商品束的效用最高即可。

或者，我们可以使用无差异曲线进行分析，如图 5.7 所示。通常，最优消费束是那个位于最高无差异“曲线”上的消费束。如果商品 1 的价格很高，则消费者的购买数量为零；随着价格下降，消费者会发现购买一单位商品 1 是最优的。典型地，如果价格进一步下降，消费者会购买更多单位的商品 1.

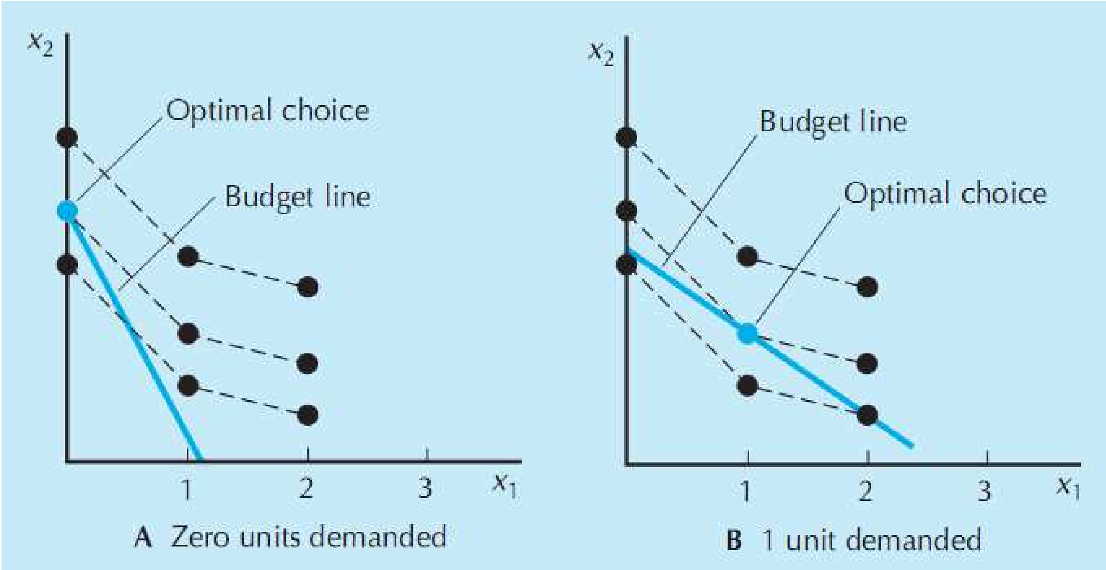


图 5.7：离散商品。在 A 图，消费者对商品 1 的需求量为 0；在 B 图，其需求量为 1.

凹偏好

思考图 5.8 所示的情形。X 是最优选择吗？不是！这类偏好下的最优选择总是边界选择，比如消费束 Z。思考一下非凸偏好的意思。如果你想购买冰淇淋和橄榄，但你不喜欢同时消费这两种商品，你会将所有的资金全部用于购买其中一种商品。

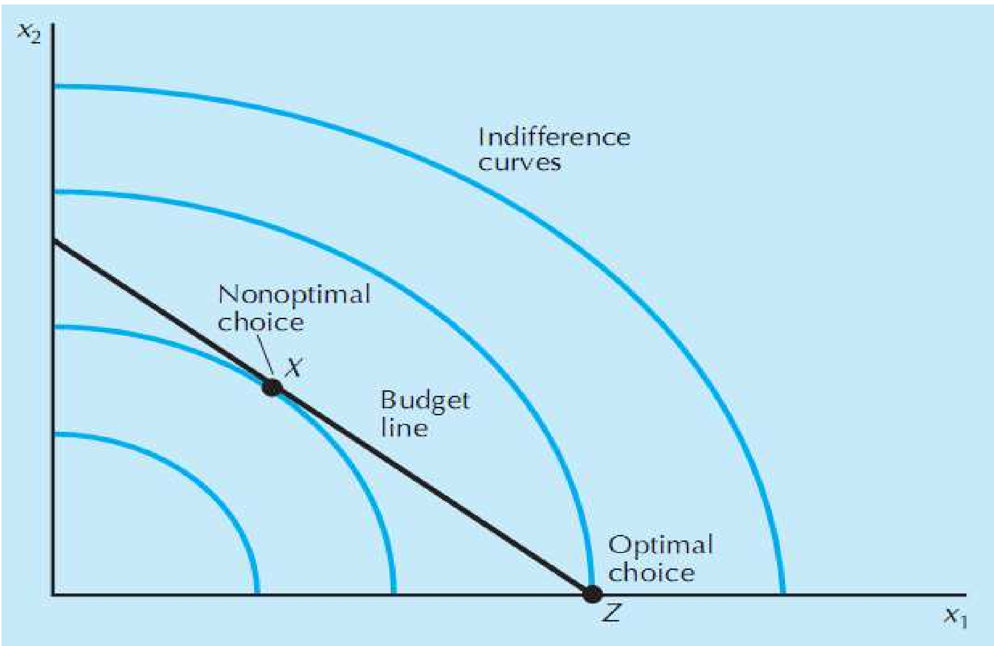


图 5.8：凹偏好（concave preferences）**情形下的最优选择**。最优选择是边界点的商品束 Z，而不是内部切点的商品束 X，因为 Z 位于更高的无差异曲线上。

柯布−道格拉斯偏好

假设效用函数为柯布-道格拉斯形式，$u(x_1, x_2) = x_1^c x_2^d$。在本章的附录部分，我们用微积分推导了这个效用函数的最优选择。结果为

$$x_1 = \frac{c}{c+d}\frac{m}{p_1}$$

$$x_2 = \frac{d}{c+d}\frac{m}{p_2}.$$

这类需求函数在代数计算中比较有用，因此你也许应该记住它们。

柯布-道格拉斯偏好有一个较好的性质。这种偏好类型下，消费者花费在商品 1 上的资金占收入的比例为多大？如果他购买 x_1 单位的商品 1，花费为 $p_1 x_1$，因此占收入的比例为 $p_1 x_1 / m$。将 x_1 的需求函数代入可得

$$\frac{p_1 x_1}{m} = \frac{p_1}{m}\frac{c}{c+d}\frac{m}{p_1} = \frac{c}{c+d}.$$

类似地，消费者花费在商品 2 上的资金占收入的比例为 $d/(c+d)$。

于是，消费者花费在每种商品上的资金，占收入的比例是固定的。这一比例的大小取决于柯布-道格拉斯函数中变量的指数。

因此，使用变量的指数之和为 1 的柯布-道格拉斯函数比较方便[一]。如果 $u(x_1, x_2) = x_1^a x_2^{1-a}$，则我们立即可将 a 解释为花费在商品 1 上的收入比例。正因为此，我们通常将柯布-道格拉斯偏好写成这种形式。

5.4 估计效用函数

我们已看到几种不同类型的偏好和效用函数，并相应分析了它们的需求行为。但在现实生活中，我们通常必须反过来分析：我们可观测需求行为，由需求行为推知相应的偏好。

例如，我们观察消费者在不同价格和收入水平下的选择行为。表 5.1 就是这样一个例子。表中的数据是两种商品在不同年份的需求数据。我们还计算出了每年每种商品的支出占收入的比例，计算公式为 $s_1 = p_1 x_1 / m$ 和 $s_2 = p_2 x_2 / m$。

对于这些数据来说，支出比例的数据相对稳定。支出比例在不同年份间存在较小的变动，而不是大的让人担心的变动。商品 1 的平均支出比例大约为 1/4，商品 2 的平均支出比例大约为 3/4。效用函数 $u(x_1, x_2) = x_1^{\frac{1}{4}} x_2^{\frac{3}{4}}$ 很好地拟合了这些数据。也即是说，该效用函数产生的

[一] 事实上 $u(x_1, x_2) = x_1^a x_2^{1-a}$ 是 $v(x_1, x_2) = x_1^c x_2^d$ 的单调变换，这两个函数代表的偏好是相同的。译者注。

 曹乾（东南大学 caoqianseu@163.com）

选择行为和表中的选择行为非常接近。为了便于说明，我们将表中相应的观测值代入这个柯布-道格拉斯效用函数，从而计算出了效用值。

Year	p_1	p_2	m	x_1	x_2	s_1	s_2	Utility
1	1	1	100	25	75	.25	.75	57.0
2	1	2	100	24	38	.24	.76	33.9
3	2	1	100	13	74	.26	.74	47.9
4	1	2	200	48	76	.24	.76	67.8
5	2	1	200	25	150	.25	.75	95.8
6	1	4	400	100	75	.25	.75	80.6
7	4	1	400	24	304	.24	.76	161.1

表 5.1：描述消费行为的一些数据。

我们从表中的观测行为能得出的结论是，似乎消费者最大化的效用函数为 $u(x_1,x_2)=x_1^{\frac{1}{4}}x_2^{\frac{3}{4}}$。当然，如果进一步观测这个消费者的行为，我们也有可能推翻这个结论。但我们仅有表 5.1 的数据，根据这些数据估计出的上述效用函数很好地拟合了最大化模型。

这个结论具有重要意义，因为现在我们可以使用这个“拟合的”效用函数来估计政策变动的影响。例如，假设政府考虑征收税收，并且假设该税收使得消费者面对的价格为（2，3），收入为 200。根据我们的估计，在这样的价格下，消费束为

$$x_1=\frac{1}{4}\frac{200}{2}=25$$

$$x_2=\frac{3}{4}\frac{200}{3}=50.$$

这个消费束的效用估计值为

$$u(x_1,x_2)=25^{\frac{1}{4}}50^{\frac{3}{4}}\approx 42\,.$$

这表明征税政策使消费者的状况，比表 5.1 中第二年他的状况好，但比第三年他的状况差。因此，我们可以使用观测到的选择行为去评估税收政策对这个消费者的影响。

由于这是经济学中很重要的一个思想，我们再次回顾它的逻辑。给定选择行为的某些观测值，我们试图确定最大化的是什么东西。一旦我们模拟估计出上述用于最大化的东西，我们既可用它预测新形势下的选择行为，又可用它评估经济环境变化导致的影响。

当然，我们前面给出的例子非常简单。在现实中，我们通常难以得到个人消费选择的详

细数据。但我们通常可获得人口组别的消费数据，比如青少年，中产阶级家庭，老年人，等等。这些不同的人口组别对不同的商品可能有不同的偏好，这些偏好可从他们的消费支出模式中看出。我们可以估测出刻画他们消费模式的效用函数，再使用这个估测出的效用函数预测需求以及评估政策变化的影响。

在前面的例子中，由于支出比例相对稳定，使用柯布-道格拉斯函数进行拟合，效果很好。在其他情形下，可能需要使用更为复杂的效用函数进行拟合。计算过程可能因此变得繁琐，我们可以使用计算机进行估计，但这些估计程序的基本思想是相同的。

5.5 MRS 条件的意义

在上一节，我们学习了一个重要的思想：观察需求行为可知道相应的潜在偏好的一些重要信息。如果我们能获得足够多的消费者选择行为的观察资料，通常就能够估计出产生这些行为的效用函数。

即使仅知道某个消费者在一组价格下的一个选择，我们也可以推测出一些有用的信息，比如他的效用将如何随消费束变化而变化。我们看看怎样做出这样的推测。

在功能良好的市场中，每个人面对的商品价格通常大致相等。以黄油和牛奶这两种商品为例。如果每个人面对同样的价格，每个人都追求效用最大化，每个人的最优解都是内部解…那么黄油和牛奶之间边际替代率对于每个人来说都是相同的。

由上面的分析可知，市场使每个人的黄油和牛奶的交换率都相等，每个人都在调整他的消费束，一直调整到他对这两种商品的“内部”评价（valuation）等于市场对这两种商品的“外部”评价时为止。

上述结论有趣的地方在于，它不受收入和偏好的影响。人们对于这两种商品总消费的评价，可能差异很大。有些人会消费很多的黄油和很少的牛奶，有些人则相反。有些富人可能消费很多的牛奶和很多的黄油，而有些穷人可能消费很少。但是只要消费这两种商品，每个人必然拥有相同的边际替代率。

人们必须对商品的交换率达成一致：为了多得到一单位的一种商品，他们愿意放弃多少单位另外一种商品。

商品的价格比率等于它们的边际替代率这一事实非常重要，因为它是评价消费束变化的一种方法。例如，假设牛奶和黄油的价格分别为 1 元/品脱、2 元/磅，则边际替代率为 2：若让消费者放弃 1 磅黄油，必须补偿他们 2 品脱牛奶。或者反过来说，若让他们放弃 2 品脱牛奶，必须补偿他们 1 磅黄油。因此每个人对消费的边际变动的评价是相同的。

假设某人发明了将牛奶变成黄油的新机器：用 3 品脱牛奶可生产 1 磅黄油，而且生产不出其他的有用副产品。问：这样的机器有市场吗？答：绝对没有。因为人们能够以 2 品脱牛奶换得 1 磅黄油，他们怎么可能傻到用 3 品脱牛奶换 1 磅黄油，所以这样的机器一钱不值。

如果他发明的机器能用 1 磅黄油生产 3 品脱牛奶，这样的机器有市场吗？答：有！因为两种商品的价格表明，人们恰好愿意用 1 磅黄油换 2 品脱牛奶。一换三显然比市场中的一换二合算。风险投资者会投资这样的机器（以及黄油）！

市场价格表明第一种机器无法赚钱：生产出 2 元黄油需要投入 3 元牛奶。也就是说，人们对投入的评价比对产出的评价高。第二种机器产出 3 元牛奶仅需投入 2 元黄油，它能赚钱，因为人们对产出的评价高于对投入的评价。

上述论述的要点在于，价格衡量商品之间的交换率，因此可使用价格来评估相关政策（导致消费变化的政策）的影响。价格并不是任意的数字，而是反映人们在边际上如何评价商品，这一事实是经济学中最基本和最重要的思想之一。

给定一组价格和消费数据，我们就可以计算出该观测点上的 MRS。如果价格变动，观察消费者的另外一个选择，就可得到另外一个 MRS。观测点越多，就能越准确地估计出产生这种选择行为的潜在偏好的形状。

5.6 税种选择

到目前为止，我们学习了一点消费者理论，但即使这一丁点理论也能让我们推导出有趣而重要的结论。下面税种选择的分析就是一个很好的例子。**消费税**（quantity tax）是对消费者消费某种数量所征税收，例如每加仑汽油征税 15 分。**所得税**（income tax）是对收入征税。如果政府希望借助税收增加财政收入，用上述两种税中的哪种更好一些？我们用消费者理论进行分析。

我们首先分析征收消费税的情形。假设征税前预算线为

$$p_1x_1 + p_2x_2 = m\,.$$

如果对商品 1 征收消费税，税率为 t，则预算线怎样变动？答案很简单。对于消费者来说，征收消费税就好比商品 1 的价格上升了 t 元，因此征税后的预算线为

$$(p_1+t)\,x_1 + p_2x_2 = m\,. \tag{5.1}$$

征收消费税所获得的财政收入为 $R^* = tx_1^*$。

因此消费者认为，对某种商品征收消费税相当于提高了该商品的价格。价格变动后需求怎样变动？请看图 5.9。目前，我们无法确切知道征税后商品 1 的消费量是增加还是下降，尽管通常假设消费量会下降。但是，无论消费量是增还是减，有一点是肯定的，即最优消费束 (x_1^*, x_2^*) 必然满足下列预算线

$$(p_1+t)\,x_1^* + p_2x_2^* = m\,. \tag{5.2}$$

接下来我们分析所得税，假设征收所得税所获得的财政收入仍为 R^*。征缴所得税后的预算线变为

$$p_1x_1 + p_2x_2 = m - R^*.$$

或者将 $R^* = tx_1^*$ 代入上式得

$$p_1x_1 + p_2x_2 = m - tx_1^*.$$

在图 5.9 中，这条预算线是什么样的？

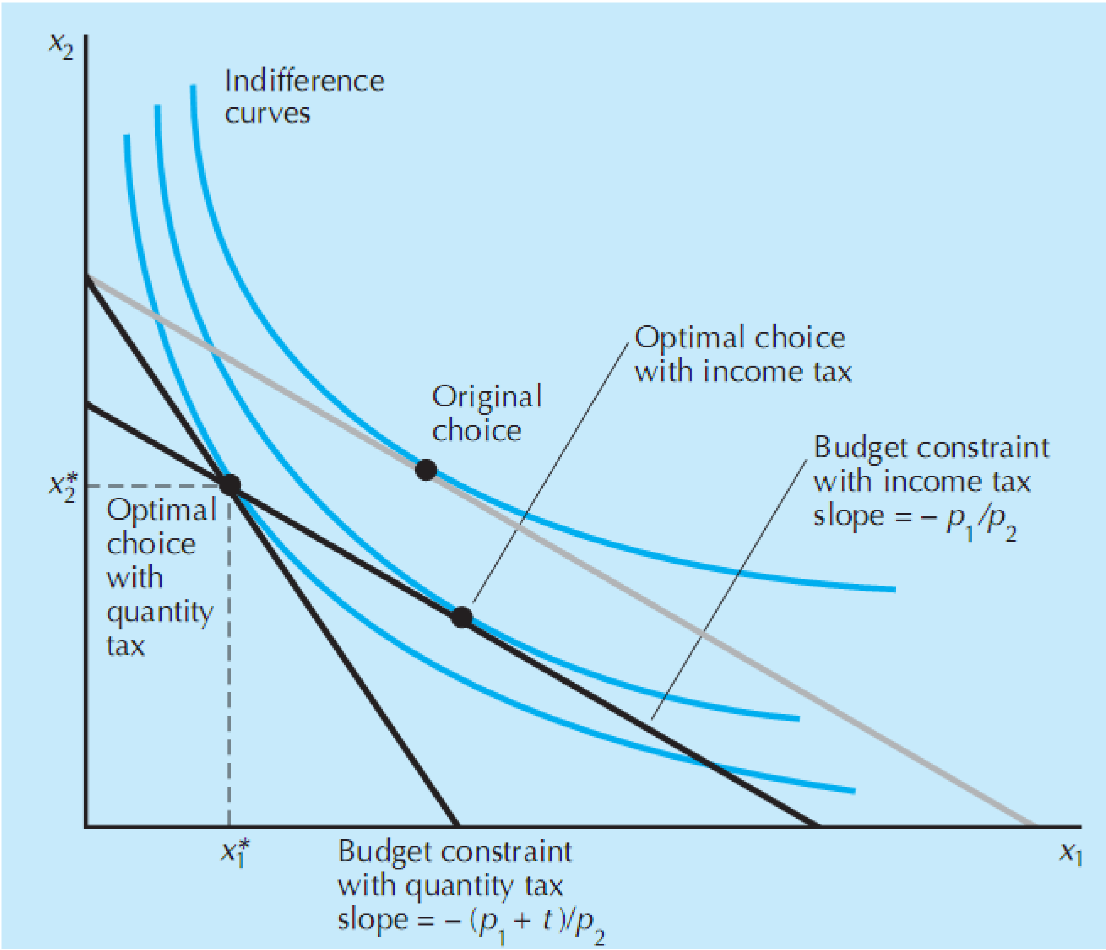


图 5.9：**所得税还是消费税(quantity tax)？**。此处我们分析无论哪种税征收的税收总额都是 R*的情形下，政府应该选择所得税还是消费税。由于与征收消费税相比，征收所得税能使消费者位于更高的无差异曲线上，即消费者的状况更好，因此政府应征收所得税。

容易看出上述预算线与征税前的预算线斜率相同，都等于$-p_1/p_2$，问题是要确定这条预算线的位置。容易证明这条预算线必然经过(x_1^*, x_2^*)：将(x_1^*, x_2^*)代入该预算线可知预算线等式仍成立，因为

$$p_1x_1^* + p_2x_2^* = m - tx_1^*$$

只不过是（5.2）式的变形，而我们知道（5.2）式是成立的。

这就证明了(x_1^*, x_2^*)位于征缴所得税后的预算线上：消费者能买的起这个消费束。但这个消费束是最优选择吗？容易看出答案为否。在点(x_1^*, x_2^*)，MRS 等于$-(p_1+t)/p_2$。但是征所得税后我们可以$-p_1/p_2$的交换率进行交易。于是预算线和无差异曲线交于(x_1^*, x_2^*)点，这表明预算线上存在比(x_1^*, x_2^*)更好的商品束。

因此，征收所得税肯定比征收消费税好，其中的道理在于，无论征哪种税，所获得的财政收入都相同，但征所得税后消费者的状况好于征收消费税后的消费者状况。

这个结论很漂亮，值得记住，但同时你要知道这个结论的局限性。首先，它只使用于一个消费者的情形。从前面的论证可以看出，对于既定的某个消费者，无论缴纳哪种税，他的税收负担都相同，但他宁可选择缴纳所得税。但由于每个人缴纳的所得税税额通常存在差异，因此对所有消费者都征收所得税，未必一定好于对所有消费者都征收消费税。（例如，某消费者不消费商品 1，此人肯定相对喜欢消费税而不喜欢所得税。）

其次，我们假设征收所得税后，消费者的收入不变，即假设税收为一种定额税（a lump sum tax）——这种税会改变消费者的商品支出总额但不会改变其最优选择。这种假设不合理。如果这里的所得为消费者的工作收入，可以预期征收所得税会打击消费者工作的积极性，因此征税后消费者的收入损失可能远大于他所缴纳的所得税税额。

最后，我们没有考虑税收对商品供给的影响。我们已说明税收如何影响需求，但税收同样会影响供给。完整的分析应将这两方面因素都考虑进来。

附录

你应该学会求解效用最大化问题，并且能得到实际需求函数。我们在本章正文中已做过这样的事情，不过针对的是完全替代和完全互补这类简单的情形。在此处，我们将针对更一般的情形。

首先，我们通常希望以效用函数$u(x_1, x_2)$表示消费者的偏好。在第 4 章我们已知道这不是一个约束性很强的假设。绝大部分良好形状的偏好都能用效用函数刻画。

事实上我们已知道怎样求解最优选择问题。我们要做的只是把我们前 3 章学过的知识汇集起来。在本章我们知道最优选择(x_1, x_2)必须满足下列条件

$$MRS_{12} = -\frac{p_1}{p_2}, \tag{5.3}$$

在第 4 章的附录我们知道，MRS 的绝对值等于效用函数偏导数的比率。由此可知

$$\frac{\partial u(x_1, x_2)/\partial x_1}{\partial u(x_1, x_2)/\partial x_2} = \frac{p_1}{p_2}. \tag{5.4}$$

从第 2 章可知，最优选择必须也满足预算约束

$$p_1x_1 + p_2x_2 = m\,. \tag{5.5}$$

由以上可得两个方程——MRS 条件和预算约束，两个未知数 x_1 和 x_2。我们所要做的就是用这两个等式解出最优选择 x_1 和 x_2，它们都是价格和收入的函数。两个方程两个未知数的求解方法有多种。第一种方法是用预算约束解出其中一个变量，然后代入 MRS 条件，这种方法通常可行，但未必最简单。

将预算约束变形可得

$$x_2 = \frac{m}{p_2} - \frac{p_1}{p_2}x_1 \tag{5.6}$$

将其带入（5.4）式可得

$$\frac{\partial u(x_1, m/p_2 - (p_1/p_2)x_1)/\partial x_1}{\partial u(x_1, m/p_2 - (p_1/p_2)x_1)/\partial x_2} = \frac{p_1}{p_2}\,.$$

这个式子看上去有些可怕，但它只有一个变量 x_1，解出 x_1，它通常的形式是 $x_1 = x_1(p_1, p_2, m)$。将其代入预算约束可解出 x_2，x_2 也是价格和收入的函数。

我们也可以使用最大值的微积分条件，求出效用最大化问题的解。这种方法的规则如下。首先，提出效用最大化问题，它是一个约束最大化问题：

$$\max_{x_1,x_2} u(x_1, x_2)$$

$$\text{使得 } p_1x_1 + p_2x_2 = m\,.$$

这个最大化问题让我们选择 x_1 和 x_2 的合适数值使得：一是它们必须满足约束条件；二是使 $u(x_1, x_2)$ 最大。

这类问题通常有两种求解方法。第一种方法是从约束条件解出其中一个变量，它是另外一个变量的函数；然后将其带入目标函数 $u(x_1, x_2)$。

例如，给定 x_1 的任一数值，则满足预算约束的 x_2 由下式给出

$$x_2(x_1) = \frac{m}{p_2} - \frac{p_1}{p_2}x_1\,. \tag{5.7}$$

现在用 $x_2(x_1)$ 代替效用函数中的 x_2，就得到了下列**无约束的**（unconstrained）最大化问题

$$\max_{x_1} u(x_1, \frac{m}{p_2} - \frac{p_1}{p_2}x_1)\,.$$

这个无约束最优化问题中的选择变量只有一个，即 x_1，这是因为我们使用了函数 $x_2(x_1)$，这个函数可以保证 x_2 必然满足预算约束，无论 x_1 的数值为多少。

对上面的目标函数求关于 x_1 的微分并令它等于 0，可解得一阶条件:

$$\frac{\partial u(x_1, x_2(x_1))}{\partial x_1} + \frac{\partial u(x_1, x_2(x_1))}{\partial x_2}\frac{dx_2}{dx_1} = 0. \tag{5.8}$$

上式左端第一项是 x_1 增加对效用增加的直接效应。第二项由两部分相乘：第一部分是 x_2 增加时效用的增加率 $\partial u / \partial x_2$；第二部分即 dx_2 / dx_1，表示当 x_1 增加时为使预算方程仍然成立 x_2 应增加的比率。dx_2 / dx_1 可从（5.7）式解得，对（5.7）式微分可得

$$\frac{dx_2}{dx_1} = -\frac{p_1}{p_2}.$$

将其带入（5.8）式可得

$$\frac{\partial u(x_1^*, x_2^*)/\partial x_1}{\partial u(x_1^*, x_2^*)/\partial x_2} = \frac{p_1}{p_2},$$

这个式子是说，商品 1 和 2 在最优选择点 (x_1^*, x_2^*) 边际替代率的 MRS_{12} 必须等于商品 1 与商品 2 的价格之比。这个条件正是我们前面推导出来的条件：无差异曲线的斜率必须等于预算线的斜率。当然最优选择点必须还要满足预算约束 $p_1x_1^* + p_2x_2^* = m$，这样我们又得到了两个二元方程。

这类问题的第二种解法是使用**拉格朗日乘数**（Lagrange multipliers）。这种方法需要先定义一个函数，这个函数称为拉格朗日函数：

$$L = u(x_1, x_2) - \lambda(p_1x_1 + p_2x_2 - m).$$

上式中有个新变量 λ[1]，这个变量称为拉格朗日乘数，因为它与预算约束相乘。拉格朗日定理是说，最优选择 (x_1^*, x_2^*) 必须满足下列三个一阶条件：

$$\begin{aligned}
\frac{\partial L}{\partial x_1} &= \frac{\partial u(x_1^*, x_2^*)}{\partial x_1} - \lambda p_1 = 0 \\
\frac{\partial L}{\partial x_2} &= \frac{\partial u(x_1^*, x_2^*)}{\partial x_2} - \lambda p_2 = 0 \\
\frac{\partial L}{\partial \lambda} &= p_1x_1^* + p_2x_2^* - m = 0.
\end{aligned}$$

上面三个式子有几个有趣的地方。首先，注意这三个式子是拉格朗日函数分别对 x_1, x_2 和 λ 分别求导而得到。最后一个式子就是预算约束。其次，现在我们有了三个含有三个未知数（即 x_1, x_2 和 λ）的方程。我们有希望解出以 p_1, p_2 和 m 表示的 x_1, x_2。

任何高等微积分教材都有拉格朗日定理的证明。这种证明广泛地应用于高等经济学课程，但对于我们的目的来说，知道这个定理是怎么回事以及如何使用就足够了。

在上面的例子中，如果我们将第一个条件与第二个条件相除，可得到

[1] λ 为希腊字母，读作“lamb-da”.

$$\frac{\partial u(x_1^*,x_2^*)/\partial x_1}{\partial u(x_1^*,x_2^*)/\partial x_2}=\frac{p_1}{p_2},$$

这个式子是说 MRS 必须等于价格比率，这一点我们前面已经得出。预算约束给出了另外一个方程，因此我们就得到了两个方程，这两个方程都含有两个未知数。

例子：柯布—道格拉斯函数

在第 4 章我们引入了柯布—道格拉斯效用函数

$$u(x_1,x_2)=x_1^c x_2^d .$$

由于任何效用函数的单调变换都保持原来的偏好关系，因此出于计算方便的考虑，可将上式取对数变成对数形式

$$\ln u(x_1,x_2)=c\ln x_1+d\ln x_2 .$$

下面我们求解该效用函数中商品 1 和商品 2 的需求函数。我们希望求解的最大化问题是

$$\max_{x_1,x_2}\ c\ln x_1+d\ln x_2$$
$$\text{s.t}\ \ p_1x_1+p_2x_2=m$$

上式中的“s.t”是英文单词“such that”的缩写，意思是“使得”，即表示约束条件。

解这个最优化问题至少有三种方法。一种方法是分别写出 MRS 条件和预算约束条件，联立求解。我们在第 4 章已得出效用条件的表达式，因此将上述两个条件联立可得

$$\frac{cx_2}{dx_1}=\frac{p_1}{p_2}$$
$$p_1x_1+p_2x_2=m.$$

由上面的两元方程组可以解出 x_1 和 x_2 的最优数量。求解方法也有好几种，第一种方法是将第二式代入第一式，可得

$$\frac{c(m/p_2-x_1p_1/p_2)}{dx_1}=\frac{p_1}{p_2}.$$

交叉相乘可得

$$c(m-x_1p_1)=dp_1x_1.$$

重新整理这个式子可得

$$cm=(c+d)\,p_1x_1.$$

由此可得

$$x_1 = \frac{c}{c+d}\frac{m}{p_1}.$$

这个式子就是商品 1 的需求函数。将上式代入预算约束方程可得商品 2 的需求函数

$$x_2 = \frac{m}{p_2} - \frac{p_1}{p_2}\frac{c}{c+d}\frac{m}{p_1}$$
$$= \frac{d}{c+d}\frac{m}{p_2}.$$

第二种方法是一开始就将预算约束代入目标函数。如果这么做，最大化问题变为

$$\max_{x_1} c\ln x_1 + d\ln(m/p_2 - x_1 p_1/p_2).$$

这个最大化问题的一阶条件为

$$\frac{c}{x_1} - d\frac{p_2}{m - p_1 x_1}\frac{p_1}{p_2} = 0.$$

由此可解出 x_1

$$x_1 = \frac{c}{c+d}\frac{m}{p_1}.$$

将其代入预算约束 $x_2 = m/p_2 - x_1 p_1/p_2$ 可得

$$x_2 = \frac{d}{c+d}\frac{m}{p_2}.$$

于是我们就得到了商品 1 和 2 的需求函数。欣慰地是，它们和第一种方法得到的结果是一样的。

下面我们使用拉格朗日方法求解需求函数。建立拉格朗日函数

$$L = c\ln x_1 + d\ln x_2 - \lambda(p_1 x_1 + p_2 x_2 - m)$$

分别对 x_1, x_2 和 λ 求导并令等于 0，可得下列三个一阶条件

$$\frac{\partial L}{\partial x_1} = \frac{c}{x_1} - \lambda p_1 = 0$$

$$\frac{\partial L}{\partial x_2} = \frac{d}{x_2} - \lambda p_2 = 0$$

$$\frac{\partial L}{\partial \lambda} = p_1 x_1 + p_2 x_2 - m = 0.$$

剩下的问题就是怎么求解上述方程了。最好的方法是先求出 λ，然后再求出 x_1, x_2。所以我们重新整理前两个式子可得

$$c = \lambda p_1 x_1$$
$$d = \lambda p_2 x_2.$$

将这两个式子相加可得

$$c + d = \lambda(p_1 x_1 + p_2 x_2) = \lambda m,$$

由此可得

$$\lambda = \frac{c+d}{m}.$$

将其代入前两个一阶条件可解得 x_1, x_2

$$x_1 = \frac{c}{c+d}\frac{m}{p_1}$$

$$x_2 = \frac{d}{c+d}\frac{m}{p_2}.$$

这个结果和我们第一种方法解得的结果是一样的。

总结

1.消费者的最优选择是这样的商品束，它位于消费者预算集内的最高无差异曲线上。

2.在最优商品束之处，通常有下列条件：无差异曲线的斜率（MRS）等于预算线的斜率。

3.如果能获得某消费者的一些消费选择数据，则可能估计出描述他选择行为的效用函数。这样的效用函数可用来预测他将来的消费选择，或者用于估计新经济政策对消费者效用的影响。

4.如果每个人面对两种商品的价格是相同的，那么每个人都有相同的边际替代率，因此都愿意按该比率进行商品交易。

复习题

1.如果两种商品是完全替代的，求商品 2 的需求函数。

曹乾（东南大学 caoqianseu@163.com）

2.假设某消费者的无差异曲线为斜率等于$-b$的直线。给定任意的价格和货币收入即(p_1, p_2, m)，他的最优选择是什么样的？

3.假设某消费者每喝一单位咖啡总是放两单位的糖，如果糖和咖啡的价格分别为p_1, p_2，收入为 m，分别计算出他购买的咖啡和糖的数量。

4.假设你对冰淇淋和橄榄有高度的非凸偏好（即凹偏好），正如教材中描述的那样，这两种商品的价格分别为p_1, p_2，收入为m。求最优消费束。

5.如果某消费者的效用函数为$u(x_1, x_2) = x_1 x_2^4$，计算他花费在商品 2 上的资金支出占他的收入的比例。

6.如果对某消费者来说，无论征收所得税还是消费税，他的状况是一样的。那么他有什么样的偏好类型？

复习题答案

1.如果两种商品是完全替代的，求商品 2 的需求函数。

【复习内容】完全替代；需求函数

【解题思路】如果计算量较少，最好画图分析。即按照教材中所说的：画出预算线和若干无差异曲线，找到位置最高的那条无差异曲线和预算线的触及点。请读者自行画图。

【参考答案】

不妨假设两种商品的价格为p_1, p_2；收入为m。

由于商品 1 和 2 是完全替代的，但题目未告知它们之间的替代比率，不妨假设商品 1 和 2 的的替代比率为$1:b$，即 1 单位商品 1 可以替代b单位商品 2。

如果$bp_2 > p_1$，则消费者只会购买商品 1，因此此时商品 2 的需求量为 0；如果$bp_2 < p_1$，则消费者只会购买商品 2，此时商品 2 的需求量为m/p_2。如果$bp_2 = p_1$，则预算线上的任何一点都是最优选择点，此时商品 2 的需求量可以为闭区间$[0 \quad \frac{m}{p_2}]$中的任一数。

因此，商品 2 的需求函数为

$$x_2 = \begin{cases} 0 & bp_2 > p_1 \\ x_2 \in [0 \quad \frac{m}{p_2}] & bp_2 < p_1 \\ \frac{m}{p_2} & bp_2 = p_1. \end{cases}$$

 曹乾（东南大学 caoqianseu@163.com）

2.假设某消费者的无差异曲线为斜率等于 $-b$ 的直线。给定任意的价格和货币收入即 (p_1, p_2, m)，他的最优选择是什么样的？

【复习内容】完全替代；最优选择

【解题思路与参考答案】

其实本题和第一题是一回事。这是完全替代的偏好，效用函数为 $u(x_1, x_2) = bx_1 + x_2$。

同样可以画图分析。由于本题和第 1 题是相同的。因此此处只给出简略答案：如果 $bp_2 > p_1$，则消费者只会购买商品 1，因此此时商品 2 的需求量为 0；如果 $bp_2 < p_1$，则消费者只会购买商品 2，此时商品 2 的需求量为 m/p_2。如果 $bp_2 = p_1$，则预算线上的任何一点都是最优选择点。

3.假设某消费者每喝一单位咖啡总是放两单位的糖，如果糖和咖啡的价格分别为 p_1, p_2，收入为 m，分别计算出他购买的咖啡和糖的数量。

【复习内容】完全互补；最优选择。

【解题思路和参考答案】

令 x_1, x_2 分别表示糖和咖啡的购买量。

由题意可知

$$x_1 = 2x_2 \qquad (1)$$

预算线为

$$p_1x_1 + p_2x_2 = m \qquad (2)$$

联立（1）和（2）可解得：$x_1 = \dfrac{2m}{2p_1 + p_2}; x_2 = \dfrac{m}{2p_1 + p_2}.$

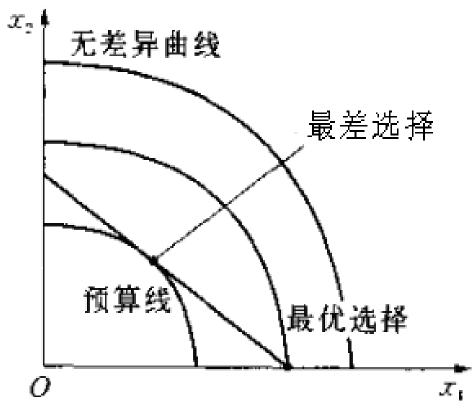
4.假设你对冰淇淋和橄榄有高度的非凸偏好（即凹偏好），正如教材中描述的那样，这两种商品的价格分别为 p_1, p_2，收入为 m。求最优消费束。

【复习内容】凹偏好；最优选择。

【解题思路与参考答案】

凹偏好情形下，端点消费束好于平均消费束，注意，此时由无差异曲线和预算线相切条件解出的切点解，不仅不是最优的，反而是最差的。

因此，最优选择是角点解（corner solutions），如下图所示，最优选择为 $(m/p_1, 0)$ 当然，$(0, m/p_2)$ 也为最优选择。此时的图形读者自行画出。



5.如果某消费者的效用函数为$u(x_1,x_2)=x_1x_2^4$，计算他花费在商品 2 上的资金支出占他的收入的比例。

【复习内容】柯布—道格拉斯效用函数；最优选择。

【解题思路和参考答案】请读者务必学习本章的附录部分。附录部分提供了详细的分析。我们只给出其中一种解题方法。

由题意可知，消费者面临的最大化问题为

$$\max_{x_1,x_2} \ x_1x_2^4$$
$$s.t. \ \ p_1x_1+p_2x_2=m.$$

构造拉格朗日函数$L(x_1,x_2)=x_1x_2^4-\lambda(p_1x_1+p_2x_2-m)$，其一阶条件为：

$$\frac{\partial L}{\partial x_1}=x_2^4-\lambda p_1=0$$

$$\frac{\partial L}{\partial x_2}=4x_1x_2^3-\lambda p_2=0$$

$$\frac{\partial L}{\partial \lambda}=p_1x_1+p_2x_2-m=0.$$

由前两个式子可得$\frac{x_2}{4x_1}=\frac{p_1}{p_2}$，即$p_2x_2=4p_1x_1$，将其代入预算方程$p_1x_1+p_2x_2=m$可知$p_1x_1=\frac{1}{5}m$，$p_2x_2=\frac{4}{5}m$。这就是说消费者将 1/5 的收入花费在商品 1 上，将 4/5 的收入花费在商品 2 上。

6.如果对某消费者来说，无论征收所得税还是消费税，他的状况是一样的。那么他有什么样的偏好类型？

 曹乾（东南大学 caoqianseu@163.com）

【复习内容】税收类型对消费者福利的影响

【分析思路】

本题的解题思路就是猜。通过教材中的学习可知，对于一般形状的无差异曲线(即偏好为良好性状时)，某消费者相对更喜欢所得税一些。因为与征收消费税相比，征收所得税能使消费者位于更高的无差异曲线上，即他的状况更好一些。请参考教材图 5.9。

因此，题目中的消费者，他的偏好必然是特殊类型的。根据教材中介绍的几种特殊类型的偏好，多猜几次并简单画图分析一下，就可知道他的偏好是完全互补类型的。事实上，也只有完全互补类型的偏好，才能出现题目中所说的现象。

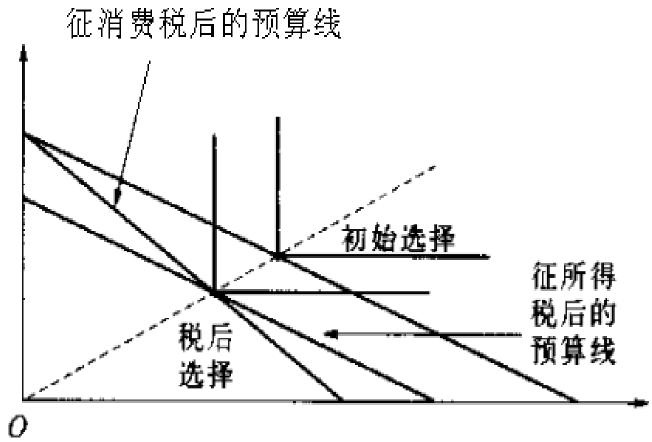
猜中了答案后，要分析一下原因。首先注意以下两个事实：一是如果你对某种税收的规避能力越强，你的状况越好；二是从宽泛意义上来说，任何两种商品之间总存在着替代性和互补性。现在考虑这样的问题：如果对你消费的两种商品中的其中一种征收消费税，那么这两种商品的替代性越强（即互补性越弱）还是越弱时，你对税收的规避能力越强？答案显然为替代性越强时。如果两商品价格相同，且为 1：1 完全替代，则政府对商品 1 征收消费税，对你的状况毫无影响，因为此时你可以选择消费商品 2。

因此题目可以转化为下列问题：政府征收所得税，你无法规避；什么样情形下，政府对某种商品征收消费税你也无法规避？答案为这两种商品为完全互补时。因为两商品若为完全互补的，可以把它们视为一种商品。

【参考答案】

当消费者的偏好为完全互补类型时，政府无论征收所得税或消费税，他的状况是一样的。

如下图所示。如果暂时理解不了这个图也不要紧，等学完了第 8 章斯勒茨基方程后再来看这个图就会豁然开朗。



曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

6.需求（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

6 需求

在上一章，我们提出了消费者选择的基本模型：预算约束下的效用最大化如何给出了最优选择。我们已知道，消费者的最优选择取决于他的收入和商品的价格，我们还分析了一些例子，即在某些简单类型偏好情形下，相应的最优选择是什么样的。

消费者**需求函数**（demand function）的定义为：每种商品的最优数量是商品价格和消费者收入的函数。可将需求函数写为

$$x_1 = x_1(p_1, p_2, m)$$
$$x_2 = x_2(p_1, p_2, m).$$

每个等式的左端表示商品的需求数量。每个等式右端是将价格和收入与需求数量关联起来的函数。

在本章，我们将分析某种商品的需求如何随价格和收入的变化而变化。研究选择如何随经济环境变化而变化，称为**比较静态**分析（comparative statics），这个术语我们已在第一章介绍过。“比较”表示我们将比较两种情形：经济环境变化之前的选择与之后的选择。“静态”的意思是说，我们不关注从一种选择到另外一种选择变动的调整过程，我们只关注已达到均衡的选择。

在上述消费者的例子中，模型中只有两类因素影响最优选择：价格和收入。因此，消费者理论中的比较静态问题，就是研究当价格和收入变化时需求如何变化。

6.1 正常商品和劣等商品

我们首先分析消费者对某种商品的需求，是如何随他的收入变动而变动的。我们想知道，在不同收入水平下，如何比较相应的最优选择。在分析过程中，假设价格保持不变，我们只关注由收入变动引起的需求变动。

我们已知道，价格不变而货币收入增加会对预算线有何影响——预算线会向外平行移动。这一移动对需求有何影响？

我们通常认为当消费者的收入增加时，他对每种商品的需求都会增加，如图 6.1 所示。经济学家由于特别缺乏想象力，把这样的商品称为**正常**商品（normal goods）。如果商品 1 是正常商品，则收入增加时需求增加，收入减少时需求减少。对于正常商品来说，需求量的变动方向总是与收入变动的方向相同：

$$\frac{\Delta x_1}{\Delta m} > 0 .$$

有叫做正常商品的，那么**可能**也有叫不正常商品的。还真存在不正常商品。图 6.2 展现了良好性状的无差异曲线，但此处随着收入增加，其中一种商品的需求量却**减少**。这样的商

品称为**劣等**商品（inferior good）。这有些“不正常”，但仔细想想你就知道，劣等商品并不罕见。很多商品的需求随收入的增加而下降，可能的例子包括稀饭，简陋的房子，粗制腊肠，或者其他低质量商品。

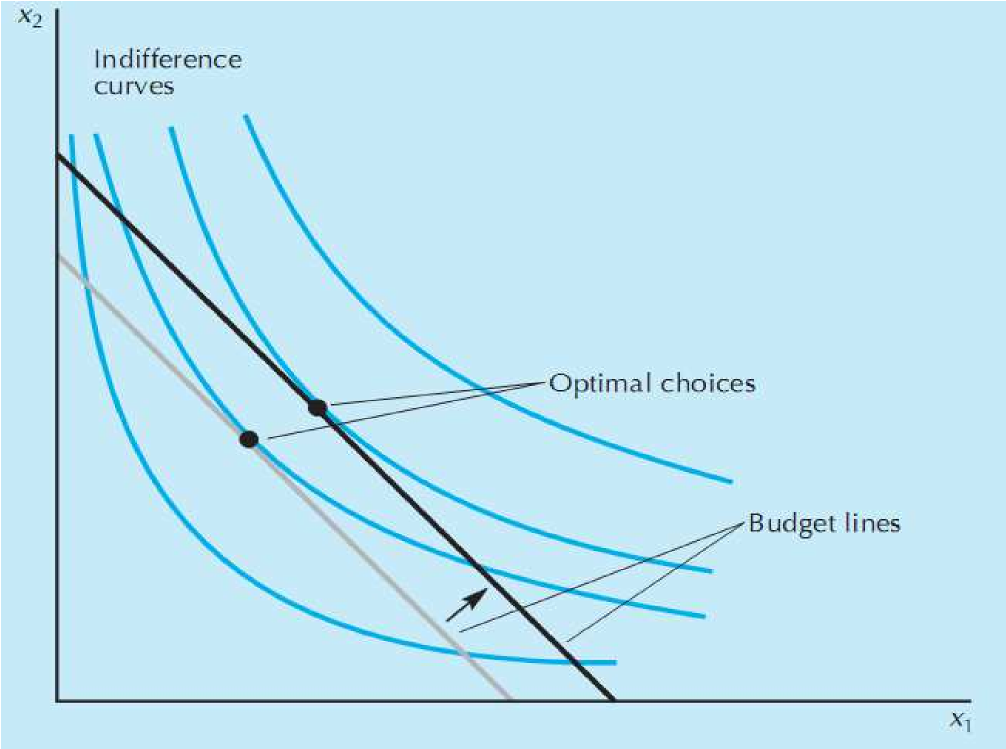


图 6.1：正常商品。收入增加时两种商品的需求都增加，因此都为正常商品。

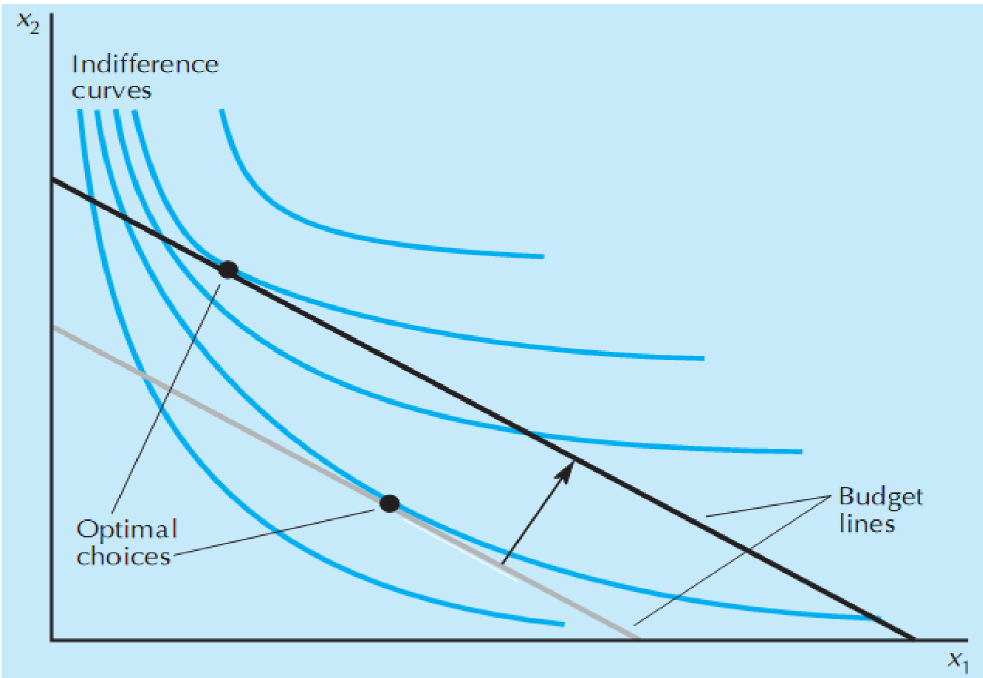


图 6.2：劣等商品。当收入增加时商品 1 的需求反而下降，因此它为劣等商品。

某商品是否为劣等商品取决于我们研究的收入水平。很有可能出现这样的事情，即非常穷的人们在收入增加时，腊肠的需求数量增加。但增加到一定点，需求量可能会随收入增加而下降。在现实生活中，当收入增加时，某种商品的需求可能增加也可能减少，幸好经济理论允许这两种情况并存。

6.2 收入提供曲线和恩格尔曲线

我们已经看到收入增加时预算线向外平移。预算线平移时会产生一系列需求束，将这些需求束连接起来就得到了**收入提供曲线**（income offer curve）。这条曲线说明了不同收入水平下的相应的需求束，如图 6.3 所示。收入提供曲线又叫作**收入扩展路径**（income expansion path）。如果两种商品都是正常商品，那么收入扩展路径的斜率将为正，如图 6.3A 所示。

在每个收入水平 m 上，任何一种商品都有最优需求量。我们重点考虑商品 1，分析在每组价格和收入水平 $x_1(p_1, p_2, m)$ 上的最优选择。这就是商品 1 的需求函数。如果我们保持商品 1 和商品 2 的价格不变，看看需求怎样随收入变动而变动，这样我们就得到了**恩格尔曲线**（Engel curve）。恩格尔曲线是在所有商品价格不变情形下，将某种商品的需求视作收入的函数而得到的曲线，如图 6.3B 所示。

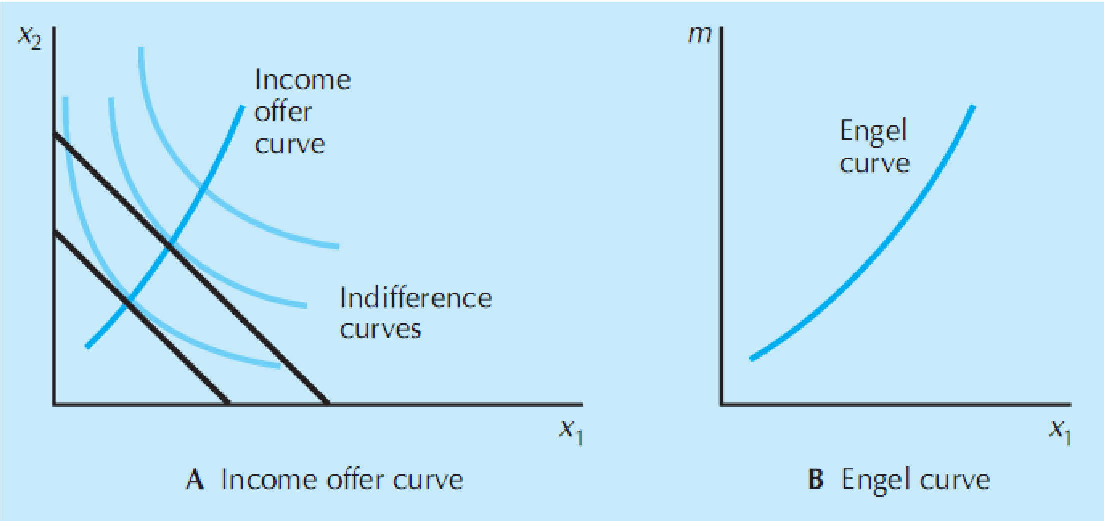


图 6.3：需求如何随收入变动而变动。图 A 展示了收入提供曲线（或收入扩展路径），它描述了价格不变时不同收入水平下的最优选择。如果我们画出不同收入水平下商品 1 的最优选择，就得到了恩格尔曲线，见图 B。

6.3 一些例子

我们再次分析第 5 章研究过的几种偏好，看看在这些不同的偏好下，收入提供曲线和恩格尔曲线是什么样子的。

完全替代

图 6.4 描述了完全替代的情形。如果 $p_1 < p_2$，则消费者只消费商品 1，随着收入增加，他会增加商品 1 的消费。由于我们以横轴表示商品 1 的需求量，收入提供曲线就是横轴，如图 6.4A 所示。因为此种情形下商品 1 的需求为 $x_1 = m / p_1$，恩格尔曲线是一条斜率为 p_1 的直线，如图 6.4B 所示。（因为纵轴表示收入 m，横轴表示商品 1 的需求量 x_1，由 $x_1 = m / p_1$ 得 $m = p_1 x_1$，显然该恩格尔曲线的斜率为 p_1）。

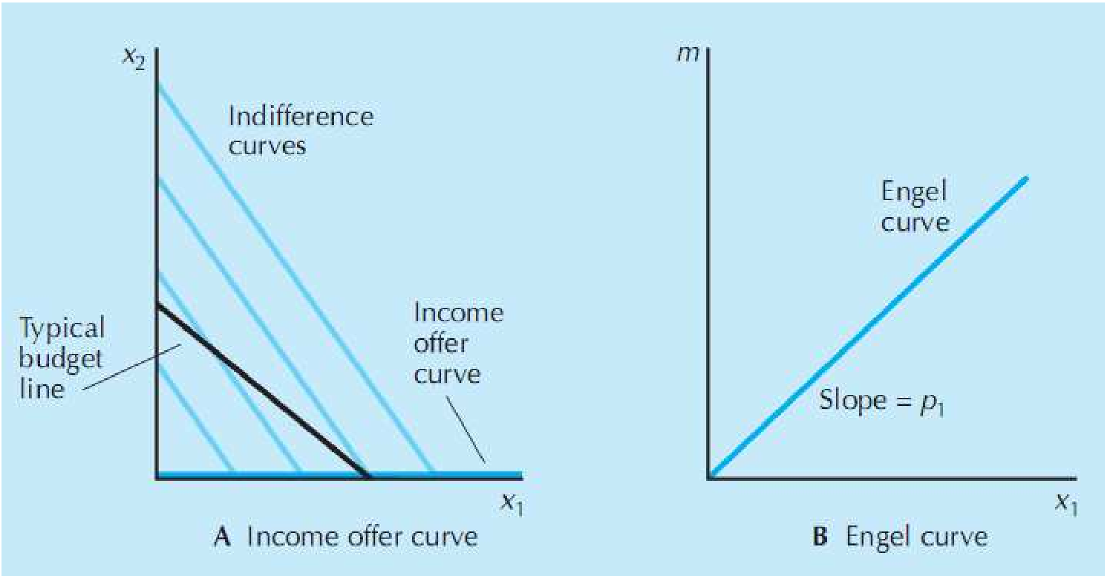


图 6.4：完全替代。用横轴表示商品 1 的需求，假设 $p_1 < p_2$，则完全替代情形下的收入提供曲线（图 A）和恩格尔曲线（图 B）。

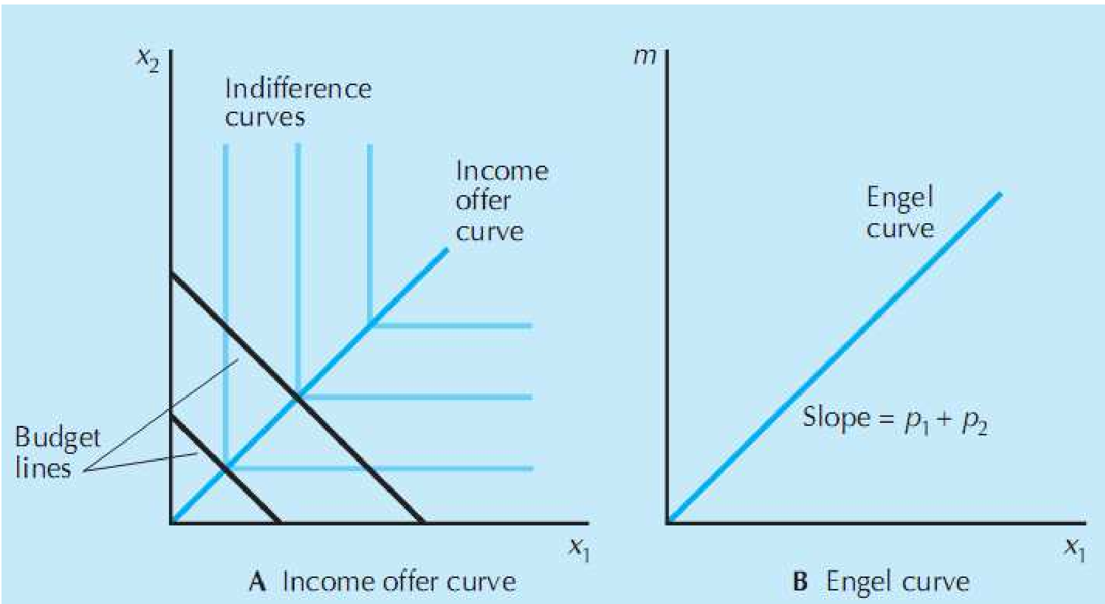


图 6.5：完全互补。1：1 完全互补情形下的收入提供曲线和恩格尔曲线。

曹乾（东南大学 caoqianseu@163.com）

完全互补

图 6.5 描述了完全互补情形下商品的需求。假设两种商品 1：1 完全互补。因为该情形下，每种商品的消费量总是相同，收入提供曲线是一条经过原点的对角线，如图 6.5A 所示。我们已知道商品 1 的需求为 $x_1 = m/(p_1 + p_2)$，因此商品 1 的恩格尔曲线是一条斜率为 $(p_1 + p_2)$ 的直线，如图 6.5 所示。

柯布-道格拉斯偏好

对于柯布-道格拉斯类型的偏好，比较容易的做法是分析代数形式的需求函数，来看看函数的图形是怎么样的。如果 $u(x_1, x_2) = x_1^a x_2^{1-a}$，则商品 1 的需求为 $x_1 = am / p_1$。如果价格 p_1 不变，则商品 1 的上述需求函数是收入 m 的**线性**函数，因此收入翻一倍需求也翻一倍，收入翻二倍需求也翻二倍，以此类推。事实上，收入 m 乘以任何正数 t，则需求也乘以 t。

商品 2 的需求为 $x_2 = (1-a)m / p_2$，若价格 p_2 不变，这显然也是线性函数。由于假设价格不变时，两种商品的需求函数都是线性的，收入扩展路径将是通过原点的直线，如图 6.6A 所示。商品 1 的恩格尔曲线是一条斜率为 p_1 / a 的直线，如图 6.6B 所示。

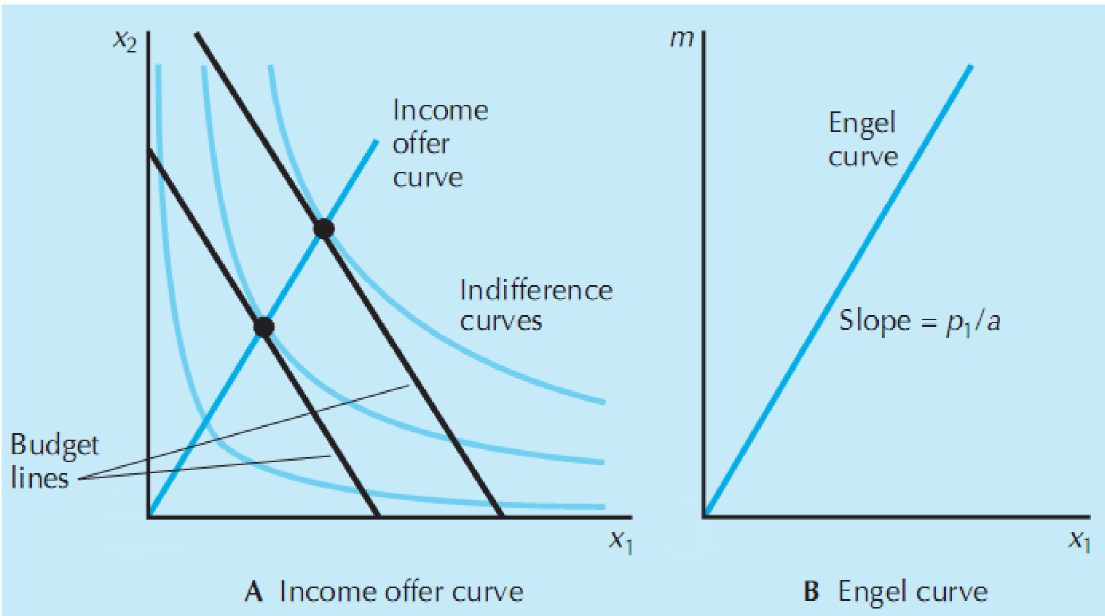


图 6.6：**柯布-道格拉斯偏好**。A 图为该偏好类型的收入提供曲线；B 图为该偏好类型的恩格尔曲线。

位似偏好

直到目前我们看到的收入提供曲线和恩格尔曲线都很简单,事实上，它们都是直线！其中原因在于我们的例子都很简单。真正的恩格尔曲线未必为直线。一般来说，当收入增加，某种商品需求增加的速度可能比收入增加的速度快，也可能比收入增长慢。前者情形的商品称为**奢侈品**（luxury good），后者情形的商品称为**必需品**（necessary good）。

上述两种情形的分界线是需求增长速度和收入增长速度相同。前面分析过的三类例子都属于这种情形。什么样的偏好能导致这种情形发生？

假设消费者的偏好仅取决于商品 1 和商品 2 的**比率**,这意味着如果消费者偏好(x_1, x_2)胜过(y_1, y_2)，那么他将自动偏好$(2x_1, 2x_2)$胜过$(2y_1, 2y_2)$，偏好$(3x_1, 3x_2)$胜过$(3y_1, 3y_2)$，以此类推，原因在于这些商品束中商品 1 和商品 2 的比率是相同的。事实上，对于任意$t > 0$消费者都将偏好(tx_1, tx_2)胜过(ty_1, ty_2)。具有这种性质的偏好称为**（同）位（类）似偏好**（homothetic preferences）。不难证明我们研究过的三类偏好，即完全替代，完全互补和柯布—道格拉斯偏好，都是位似偏好.

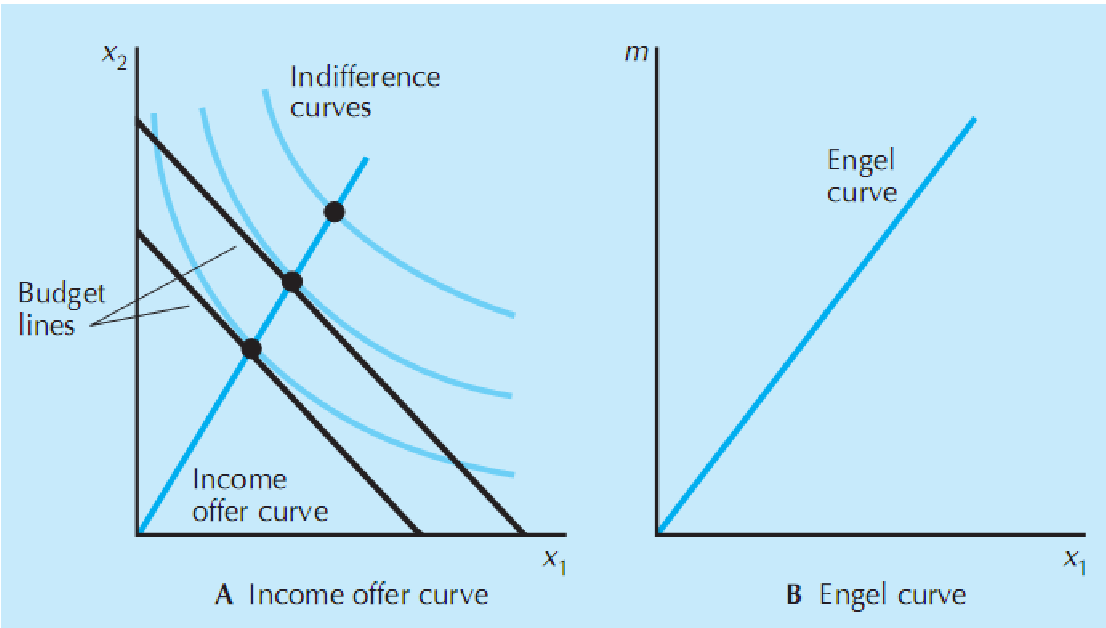


图 6.7: **位似偏好**。 位似偏好情形下的收入提供曲线（A 图）和恩格尔曲线（B 图）。

如果消费者的偏好为位似偏好，那么收入提供曲线都为通过原点的直线，如图 6.7 所示。更具体地说，如果偏好为位似的，收入乘以t，那么需求也乘以t，其中$t > 0$。这个结论可以严格证明，但图形表示也很清晰。如果某条无差异曲线和预算线相切于点(x_1^*, x_2^*)，那么经过(tx_1^*, tx_2^*)的另一条无差异曲线，必与平移后的预算线（收入变为原来t倍但价格不变）相切于点(tx_1^*, tx_2^*)。这意味着恩格尔曲线是直线。如果你将收入翻一番，每种商品的需求也将翻一番。

使用位似偏好很方便，因为收入效应很简单。不幸的是，因为同样原因位似偏好不太符合现实。但我们还是要经常使用位似偏好的例子，还是因为它简单。

拟线性偏好

拟线性偏好的收入提供曲线和恩格尔曲线比较特别。回想一下第 4 章中拟线性偏好的定义，在该情形下，所有无差异曲线是由一条无差异曲线“平移”得到。如图 6.8 所示。等价地，拟线性偏好的效用函数形式为 $u(x_1,x_2)=v(x_1)+x_2$，注意图 6.8 中我们以横轴表示 x_1。如果预算线向外平移，将会发生什么事？在拟线性偏好的情形下，如果一条无差异曲线和预算线相切于点 (x_1^*,x_2^*)，那么对于不等于于零的任意常数 k，在点 (x_1^*,x_2^*+k) 必有另一条无差异曲线和平移到此位置的预算线相切。收入增加根本不影响商品 1 的需求，所有多余的收入都用于消费商品 2。如果偏好是拟线性的，我们有时说商品 1 的 “收入效应为零”。因此，商品 1 的恩格尔曲线是一条垂线，因为你改变收入时，商品 1 的需求保持不变。

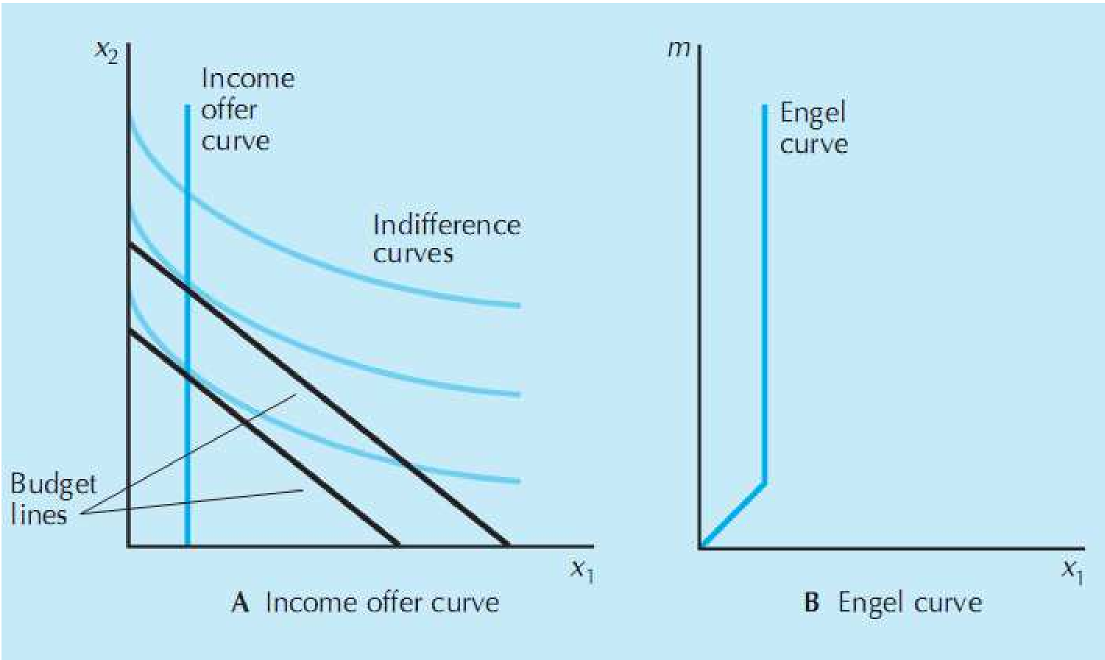


图 6.8：**拟线性偏好**。拟线性偏好情形下的收入提供曲线（A 图）和恩格尔曲线（B 图）。

我们举个现实生活中的例子说明拟线性偏好情形。假设商品 1 为铅笔，商品 2 为花费在其他商品上的金钱。最初，我可能只买铅笔，但收入大幅增加时，我不会再买铅笔。也就是说，多余的收入都被我花费在其他商品上。拟线性偏好的其他例子可能还有盐或牙膏。当我们研究两种商品之间的选择时，若其中一种商品占消费者预算的比例不大，另外一种“商品”

是其余所有商品，则拟线性的假设是合理的，至少在消费者收入足够大时是合理的。

6.4 普通商品和吉芬商品

现在分析价格变化。假设商品 1 的价格下降，维持商品 2 的价格和货币收入不变，那么商品 1 的需求量将怎样变化？直觉告诉我们商品 1 的需求量会增加。的确，这是普通情形，如图 6.9 所示。

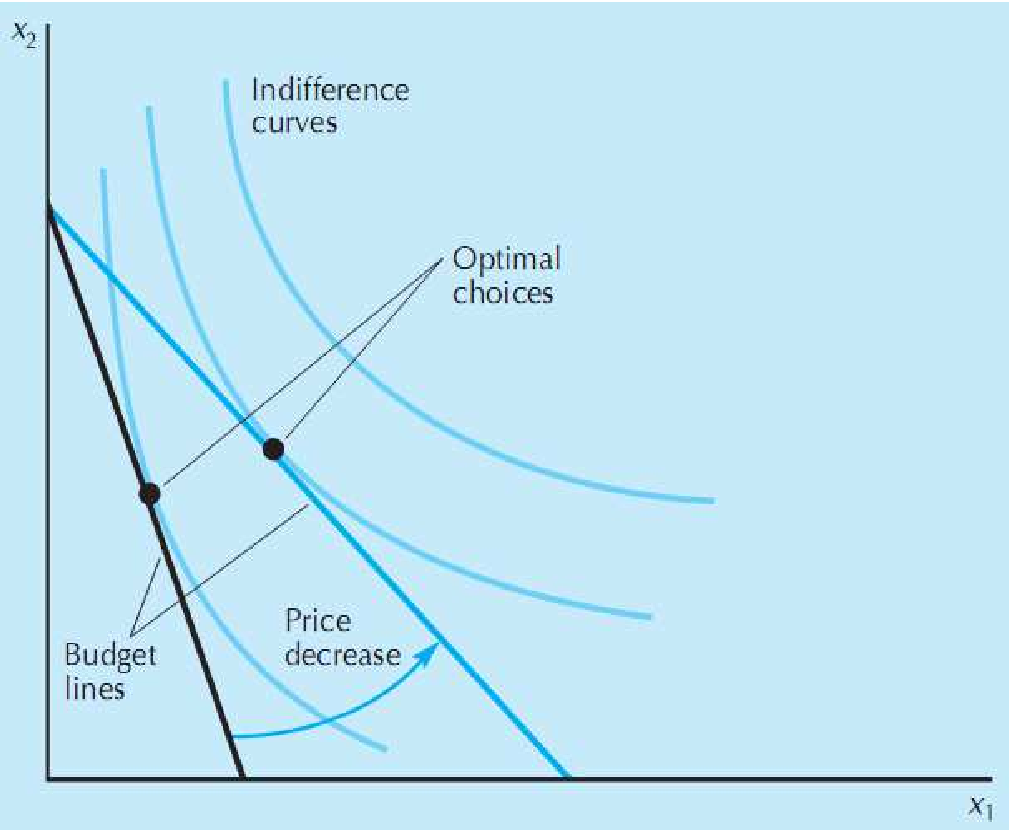


图 6.9：普通商品。普通商品的价格下降，需求量上升。

商品 1 的价格下降时，预算线变得更平坦，即纵截距不变但横截距向右移动，在图 6.9 中，商品 1 的最优选择也随之向右移动：商品 1 的需求量增加了。但我们怀疑情形是否总是如此，即不管消费者具有任何类型的偏好，价格下降时，某种商品的需求量都会上升？

我们将看到，答案为否。我们可以构造出，在良好性状的偏好下，商品 1 的价格下降反而导致其需求量下降的情形。这样的商品称为**吉芬商品**（Giffen good）。这种商品以 19 世纪经济学家吉芬的名字命名，吉芬首先注意到了存在这样商品的可能性。图 6.10 就是吉芬商品的一个例子。

怎么用经济学的语言解释这种现象？什么类型的偏好能产生图 6.10 这种特殊的行为？假设你消费的两种商品为稀饭和牛奶，目前你每周消费 7 碗稀饭和 7 杯牛奶。现在稀饭的价格下降。如果你还消费原来数量的稀饭和牛奶，你会余下一些钱，用这些钱可以购买更多的

牛奶。事实上，你可能决定减少稀饭的消费量以进一步增加牛奶的消费量。稀饭价格下降节省了部分资金，这部分资金可用于购买其他的商品，但你有可能减少稀饭的消费量。因此，价格的变化在某种程度上**类似**收入的变化。即使**货币**收入不变，某商品价格的变化也会改变购买力，因此改变了需求。

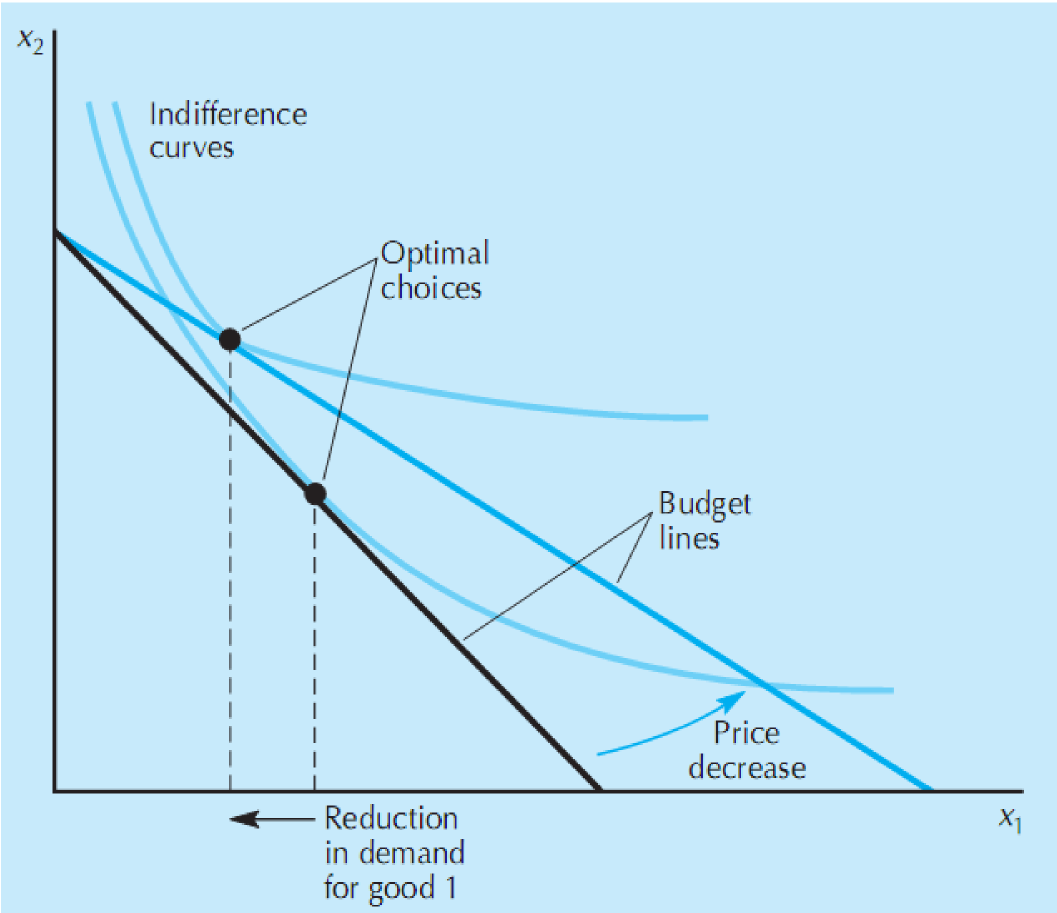


图 6.10：吉芬商品。商品 1 为吉芬商品，因为价格上升时，其需求量上升。

因此，尽管现实生活中很难遇到吉芬商品的情形，但在理论上它是完全可能存在的。绝大多数商品都是普通商品，即价格上升时，它们的需求下降。稍后我们将看到为什么称这种情形为普通情形。

顺便说一句，我们使用稀饭作为劣等商品的例子，又用它作为吉芬商品的例子，这么做不是偶然的。在后面章节我们将知道，这两种商品之间存在着密切的关系。

目前我们对消费者理论的分析，也许会让你感觉几乎一切都有可能发生：如果收入上升，某种商品的需求可能上升也可能下降；如果价格上升，需求可能下降也可能上升。难道消费者理论能解释**任何**类型的消费行为？或者消费者理论将某些类型的消费行为排除在外？事实上，效用最大化模型的确**需要**对消费行为作出某些假设限制。我们在下一章将看到这些假设是什么。

6.5 价格提供曲线和需求曲线

假设商品 1 价格变动，保持商品 2 的价格和货币收入不变。在几何图形上，这涉及预算线的转动。每一条无差异曲线与相应的预算线相切，把最优点（即切点）连接起来就得到了**价格提供曲线**(price offer curve)，如图 6.11A 所示。这条曲线代表了在不同价格下的需求束。

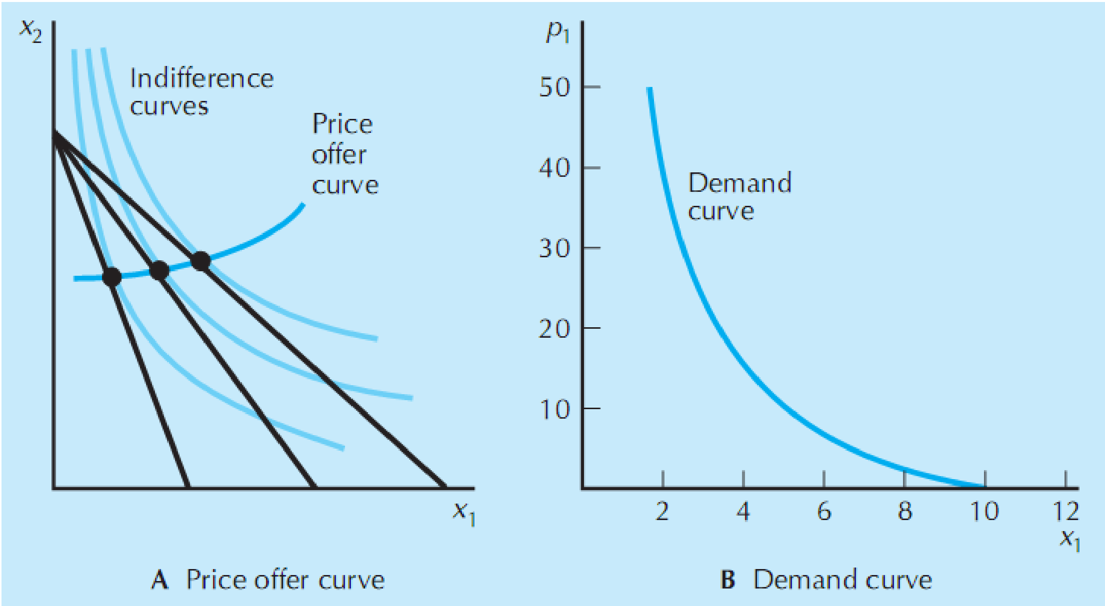


图 6.11：价格提供曲线和需求曲线。A 图展示了价格提供曲线，该曲线描述了价格变动时，商品 1 的最优选择的变动。B 图展示了需求曲线，该曲线表明商品 1 的最优选择是其价格的函数。

我们可以用另外一种方法描述价格变动和最优选择变动之间的关系。同样假设商品 2 的价格和货币收入不变，对每个不同的 p_1 值相应画出商品 1 的最优消费水平，就得到了**需求曲线**（demand curve），如图 6.11B。需求曲线是在事先给定 p_2 和 m 的数值并保持不变时，需求函数 $x_1(p_1, p_2, m)$ 的图形表示。

通常，当某商品的价格上升时，该商品的需求会下降。因此，价格变动方向和商品的需求量变动方向是**相反的**，这就是说需求曲线的斜率通常为负。上句话的意思用变化率表示为：

$$\frac{\Delta x_1}{\Delta p_1} < 0.$$

然而，我们已经知道吉芬商品情形下，价格下降，需求量也会下降。因此，需求曲线的斜率也可能为正，尽管这种可能性较小。

6.6 一些例子

使用第 3 章讨论过的若干类型的偏好，我们来看看需求曲线的一些例子。

完全替代

我们仍以红铅笔和蓝铅笔的例子，说明完全替代类型偏好的价格提供曲线和需求曲线，如图 6.12 所示。在第 5 章我们已知道，两种商品若 1：1 完全替代，则：当 $p_1 > p_2$ 时，商品 1 的需求量为 0；$p_1 = p_2$ 时，商品 1 的需求量为预算线上的任何数量；$p_1 < p_2$ 时，商品 1 的需求量为 m / p_1。价格提供曲线描述了所有上述可能性。

为了得到商品 1 的需求曲线，我们将商品 2 的价格固定在某一价格水平 p_2^*，画出商品 1 的不同价格水平及其相应的需求量，就得到了商品 1 的需求曲线，如图 6.12B 所示。

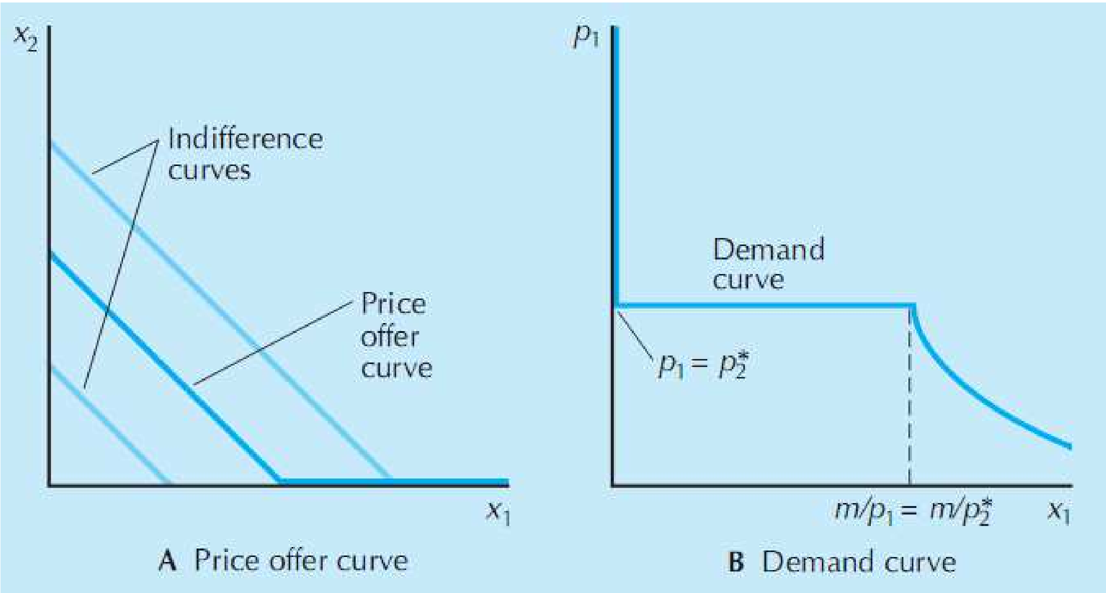


图 6.12：完全替代。完全替代情形下的价格提供曲线（A 图）和需求曲线（B 图）

完全互补

我们仍以左鞋和右鞋的例子说明完全互补的情形，如图 6.13 所示。由于这两种商品 1：1 互补，因此无论两商品的价格如何，消费者必须购买同样数量的商品 1 和商品 2。这样，他的价格提供曲线将是一条通过原点的 45^o 直线。

在第 5 章我们已知道商品 1 的需求为

$$x_1 = \frac{m}{p_1 + p_2}.$$

如果保持m和p_2不变，画出x_1和p_1之间的关系，就得到了需求曲线，如图 6.13B 所示。

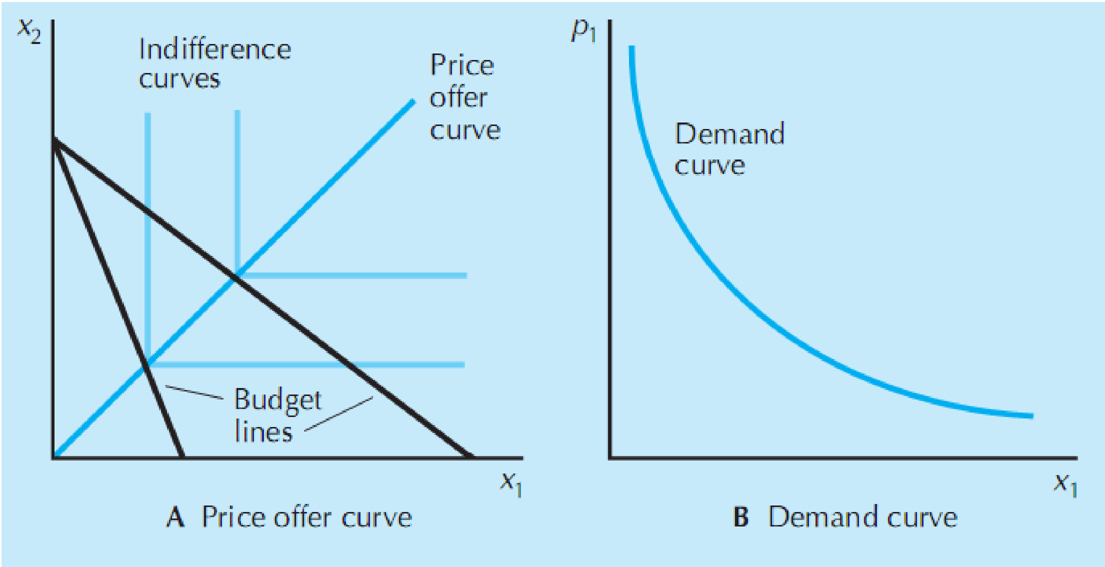


图 6.13：完全互补。两商品 1：1 完全互补情形下的价格提供曲线（A 图）和需求曲线（B 图）

离散商品

假设商品 1 是离散商品。如果p_1很高，则消费者将严格偏好于消费 0 单位；如果p_1很低，消费者将严格偏好消费 1 单位。在某个价格水平r_1下，消费者会感觉到消费商品 1 和不消费商品 1 这两种选择没有差异（indifferent）。这样的价格称为**保留价格**（reservation price）[一]。图 6.14 画出了无差异曲线和需求曲线。

从图 6.14 可知，需求行为可以用一系列保留价格刻画，在这些保留价格水平下，消费者愿意购买另外一单位商品。在保留价格r_1，消费者愿意购买 1 单位商品；如果价格降到保留价格r_2的水平，消费者愿意购买另外一单位，依次类推。

这些保留价格可用效用函数表述。例如，r_1是消费者在两种选择（消费 0 单位商品 1 和 1 单位商品 1）之间恰好无差异的价格水平，因此r_1必须满足下式：

$$u(0,m)=u(1,m-r_1) \tag{6.1}$$

类似地，r_2必须满足

[一] 术语保留价格源自拍卖市场。在拍卖市场，如果某人想卖某东西，他通常报出愿意出售此商品的最低价格。如果买方所出的最高价格仍低于他报出的价格，他就保留自己买回的权利。这个价格称为卖方的保留价格。保留价格最终演化为描述某人恰恰好愿意购买或出售某商品的价格。

$$u(1,m-r_2)=u(2,m-2r_2) \tag{6.2}$$

(6.2)式子的左端是在价格 r_2 时消费一单位商品的效用，式子的右端是消费两单位商品的效用，每个单位的价格都是 r_2。

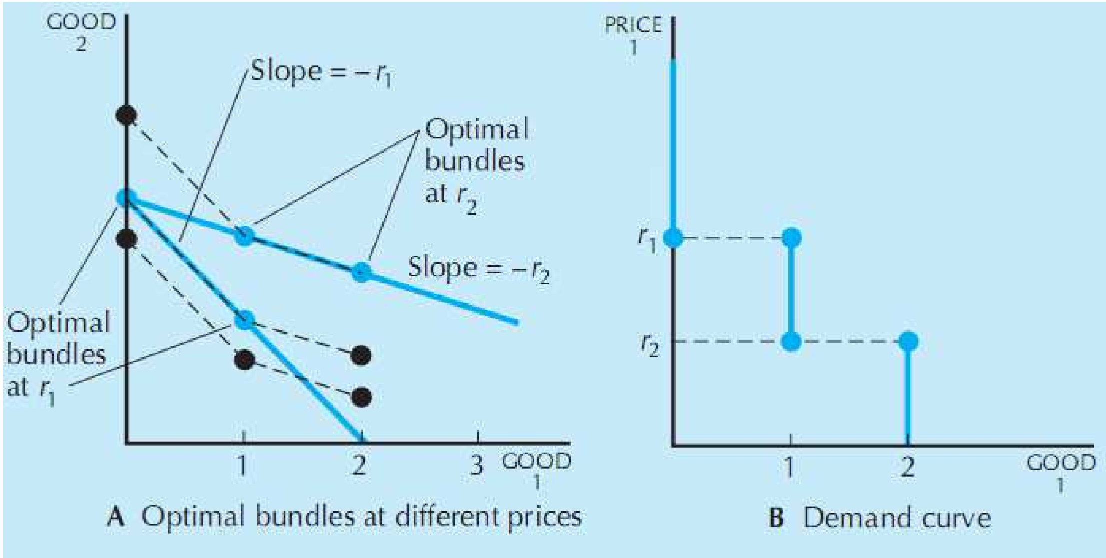


图 6.14：离散商品。当商品 1（离散商品）的价格从高价位下降时，就有可能达到消费者的保留价格水平。在该价格下，他在买和不买商品 1 之间无差异。当价格继续下降时，他就会需求更多的商品 1。

如果效用函数是拟线性的，则保留价格的效用函数表达式就变得比较简单。若 $u(x_1,x_2)=v(x_1)+x_2$ 而且 $v(0)=0$，则可将（6.1）式写为：

$$v(0)+m=m=v(1)+m-r_1.$$

因为 $v(0)=0$，可解得

$$r_1=v(1). \tag{6.3}$$

类似地，可将（6.2）式写为

$$v(1)+m-r_2=v(2)+m-2r_2.$$

移项整理可得

$$r_2=v(2)-v(1).$$

照此处理，可得第三单位商品的保留价格

$$r_3=v(3)-v(2).$$

 曹乾（东南大学 caoqianseu@163.com）

依次类推，等等。

在每个保留价格的表达式中，保留价格衡量必要的效用增量以诱使选择增加一单位商品的消费。大致来说，保留价格衡量商品 1 不同消费水平的边际效用。凸偏好的假设意味着保留价格必须是递减的：$r_1 > r_2 > r_3 \dots$

由于拟线效用函数的特别结构，商品 1 的保留价格不依赖于商品 2。这当然是一种特殊的情形，但它便于描述需求行为。给定任何价格 p，我们要做的只是看看它位于保留价格序列中的哪个位置。例如，假设 p 位于 r_6 和 r_7 之间。$r_6 > p$ 表示消费者愿意放弃 p 元钱来得到 6 单位商品 1；$p > r_7$ 表示消费者不愿意放弃 p 元钱来得到 7 单位的商品 1.

上述论证比较直观，但为使论证更清晰我们可以使用数学。假设消费者需求 6 单位商品 1，我们需要证明 p 必须满足 $r_6 \geq p \geq r_7$。

如果消费者是追求效用最大化，对于所有可能的选择 x_1，我们必然有

$$v(6) + m - 6p \geq v(x_1) + m - px_1$$

特别地，我们有

$$v(6) + m - 6p \geq v(5) + m - 5p$$

整理可得

$$r_6 = v(6) - v(5) \geq p,$$

到此处，我们的证明完成了一半。

根据同样的逻辑可得，

$$v(6) + m - 6p \geq v(7) + m - 7p$$

整理可得

$$p \geq v(7) - v(6) = r_7.$$

至此证明完毕。

6.7 替代和互补

我们已经使用过替代和互补这两个概念，现在是时候给出它们的正式定义了。既然我们已看到过几次完全替代和完全互补的说法，有必要分析一下不完全替代和不完全互补的情形。

我们先分析替代的情形。我们说过红铅笔盒蓝铅笔可以被认为是完全替代的，至少对不在乎颜色的人是这样的。但是铅笔和钢笔呢？这是“不完全”替代的情形。也就是说，钢笔

和在铅笔在某种程度上是互相替代的，尽管替代程度不如红铅笔和蓝铅笔互相替代那般完全。

类似地，我们说一双鞋中的右鞋和左鞋是完全互补的。但一双鞋和一双袜子呢？右鞋和左鞋几乎总是被一起消费的，鞋和袜子**通常**是被一起消费的。互补的商品是指类似鞋和袜子这样的商品，它们通常被一起消费，虽然不是总是被一起消费。

既然我们已经探讨了替代和互补的基本思想，我们可以给出精确的经济学定义。回忆一下商品 1 需求函数的概念，它是商品 1 的价格和商品 2 的函数，因此写为 $x_1 = x_1(p_1, p_2, m)$。现在的问题是：商品 2 的价格改变时，商品 1 的需求如何变动，是上升还是下降？

如果商品 2 的价格上升时，商品 1 的需求上升，那么我们说商品 1 是商品 2 的**替代品**（substitute）。用替代率来表达，商品 1 是商品 2 的替代品，如果下式成立

$$\frac{\Delta x_1}{\Delta p_2} > 0.$$

其中的思想是，当商品 2 变得更昂贵时，消费者转向消费商品 1：消费者用相对便宜的商品**替代**比较昂贵的商品。

另一方面，商品 2 的价格上升，若商品 1 的需求下降，则商品 1 是商品 2 的**补充品**（complement）。这意味着

$$\frac{\Delta x_1}{\Delta p_2} < 0.$$

互补品是一起被消费的商品，例如咖啡和糖，因此，当其中一种商品价格上升，两种商品的需求都会下降。

完全替代和完全互补的情形可以漂亮地说明上述观点。注意，完全替代时 $\Delta x_1 / \Delta p_2 \geq 0$；完全互补时 $\Delta x_1 / \Delta p_2 < 0$。

关于替代和互补的概念，需要提醒以下几点。首先，用两种商品来分析替代和互补的问题，比较特殊。因为收入固定，在商品 1 上多花了钱，必然在商品 2 上少花钱。这是一个缺陷，因为它限制了更多的可能行为的展现。如果我们用两种以上的商品来分析替代和互补问题，上述限制就不再是个问题。

其次，尽管用消费者的需求行为来定义替代和互补是合理的，但在更一般的环境下，这种定义方式存在着一些障碍。例如，当涉及两种以上商品时，如果我们按照上述定义方式进行分析，很可能出现下面的情形：商品 1 是商品 3 的替代品，但商品 3 却是商品 1 的补充品。因为这种特别的性质，更高级的处理通常使用稍微不同的定义方式。我们上面定义的替代和互补在比较高级处理方法中，被称为**总替代**（gross substitutes）和**总互补**（gross complements）;但对我们这本书而言，我们的定义方式已经足够用了。

6.8 反需求函数

如果保持 p_2 和 m 不变，画出 x_1 和 p_1 的对应关系 $x_1 = x_1(p_1)$，就得到了**需求曲线**（demand curve）。如上一节所述，我们通常认为需求曲线向下倾斜，因此高价格导致低需求，尽管吉芬商品表明存在相反的情形。

只要某需求曲线向下倾斜，这也是通常的情形，那么谈及**反需求函数**（inverse demand function）就有意义。反需求函数将价格视为需求量的函数。也就是说，对于商品 1 的任一需求水平，反需求函数衡量商品 1 的价格为多大时才能使消费者选择该消费水平。因此，反需求函数和相应的原需求函数描述了相同的关系，只不过视角不同。图 6.5 画出了反需求函数的曲线，当然如果你将需求量看作价格的函数，这条曲线也是原需求函数的曲线。

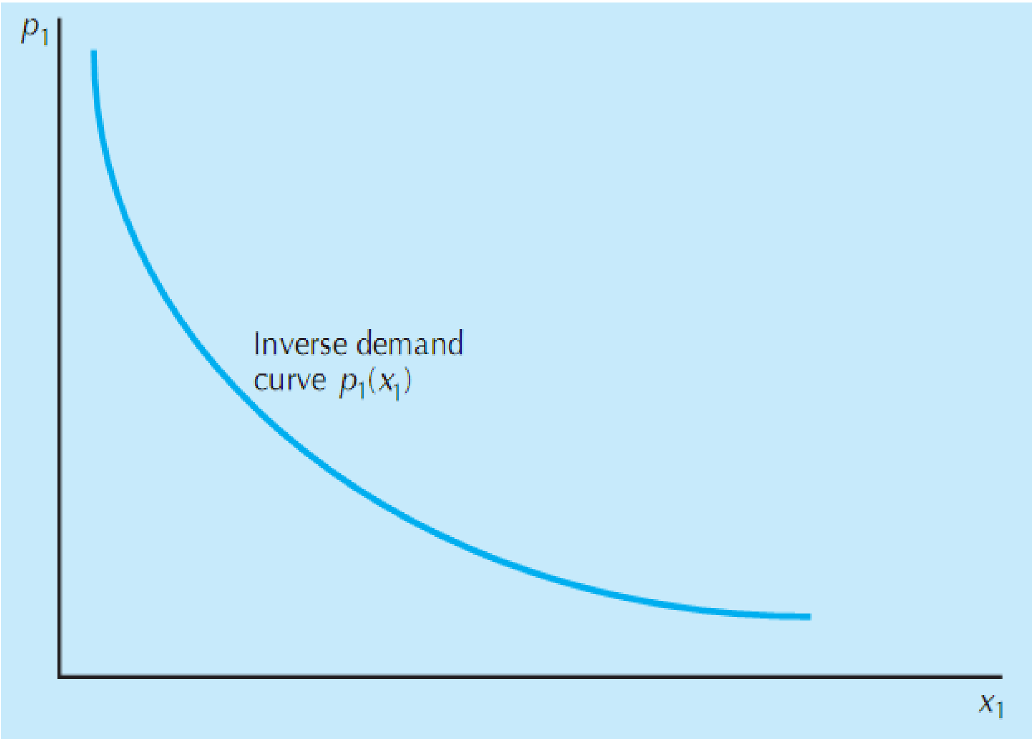


图 6.15：**反需求曲线**。如果你将需求曲线视为价格是需求量的函数，你就得到了反需求曲线。

例如，我们知道商品 1 的柯布-道格拉斯需求函数为 $x_1 = am/p_1$。我们可以将这个价格和需求量之间的关系写为 $p_1 = am/x_1$。第一个式子为需求函数，第二个式子为反需求函数。

反需求函数有一个有用的经济学解释。我们已知道只要两种商品的消费量都为正数，则最优选择必须满足下列条件，即边际替代率的绝对值等于价格之比：

$$|MRS_{12}| = \frac{p_1}{p_2}.$$

曹乾（东南大学 caoqianseu@163.com）

这意味着在商品 1 的最优消费数量处，必有

$$p_1 = p_2\left|MRS_{12}\right|.$$

因此，商品 1 的消费数量为最优时，商品 1 的价格与两商品边际替代率的绝对值成正比。

为简单起见，假设商品 2 的价格为 1。则(6.4)式表明，在商品 1 的最优需求处，商品 1 的价格衡量消费者为了多消费一点商品 1，他愿意放弃商品 2 的数量。在本例中，反需求函数衡量了边际替代率的绝对值。对商品 1 的任一最优需求水平，反需求函数衡量为了让消费者愿意少消费一点商品 1，应该补偿给他的商品 2 的数量。或者反过来说，反需求函数衡量消费者为了多消费一点商品 1，他愿意放弃商品 2 的数量。

如果我们将商品 2 视为花费在所有其他商品上的货币，则边际替代率表示消费者为了多得到一点商品 1，他所愿意放弃的货币数量。在前面我们已说过，在本例中，由于假设商品 2 的价格为 1，可将边际替代率的绝对值视作消费者的边际支付意愿。因为在本例中，商品 1 的价格等于边际替代率的绝对值，这表明商品 1 的价格本身衡量了边际支付意愿。

对于每个数量 x_1，反需求函数衡量消费者为了多得到一些商品 1 所愿意花费的钱数；用另外一种方式表达，消费者为了购买最后一单位商品 1 所愿意花费的钱数。如果商品 1 的数量足够小，这两种说法是一回事。

从上述角度看，向下倾斜的需求曲线有了新的含义。当 x_1 很小时，消费者愿意放弃很多的钱（即很多其他商品），以多获得一点商品 1。当 x_1 很大时，在边际上消费者愿意放弃很少的钱来多获得一点商品 1。因此，边际支付意愿，即消费者放弃商品 2 以获得商品 1 的边际意愿，随着商品 1 消费量的增加，是递减的。

附录

如果偏好是某种特别的形式，这意味着从这些偏好推导出的需求函数也采取特别的形式。在第 4 章我们描述了拟线性偏好是什么样子的。这类偏好的无差异曲线是互相平行的，因此可用下列效用函数表示拟线性偏好

$$u(x_1, x_2) = v(x_1) + x_2 .$$

它的最大化问题为

$$\max_{x_1, x_2} \; v(x_1) + x_2$$
$$\text{s.t.} \; p_1x_1 + p_2x_2 = m.$$

从预算约束解出 x_2，它是 x_1 的函数。将 x_2 代入目标函数可得

$$\max_{x_1} \; v(x_1) + m / p_2 - p_1x_1 / p_2.$$

上式对 x_1 求导并令其等于 0，可得一阶条件

$$v'(x_1^*)=\frac{p_1}{p_2}.$$

这个需求函数有个有趣的性质：商品 1 的需求和收入无关。这个结论我们在教材正文中已用无差异曲线的图形推出过。反需求函数为

$$p_1(x_1)=v'(x_1)p_2.$$

也就是说，商品 1 的反需求函数是效用函数对 x_1 的导数乘以 p_2。一旦我们计算出了商品 1 的需求函数，将其代入预算约束方程即可得到商品 2 的需求函数。

例如，请根据下列效用函数求出需求曲线

$$u(x_1,x_2)=\ln x_1+x_2.$$

应用一阶条件可得到

$$\frac{1}{x_1}=\frac{p_1}{p_2},$$

因此商品1的需求函数为

$$x_1=\frac{p_2}{p_1}.$$

它的反需求函数为

$$p_1(x_1)=\frac{p_2}{x_1}.$$

将商品1的需求函数 $x_1=p_2/p_1$ 代入预算约束方程可得到商品 2 的需求函数：

$$x_2=\frac{m}{p_2}-1.$$

对拟线性偏好的需求函数需要提醒读者注意以下的问题。也许你已注意到了本例中商品1的需求和收入无关。这是拟线性函数的一般特征：当收入改变时，商品1的需求保持不变。然而，这个结论只对某些收入值成立。我们不能说无论收入为多大时，商品1的需求都和收入无关。毕竟，当收入为0时，所有的需求都未0。这个结论只有当每种商品的消费量都为正数时才成立。

在本例中，当 $m<p_2$ 时，商品2的消费量为0。当收入增加时消费商品1的边际效用递减。当 $m=p_2$ 时，消费商品1的边际效用恰好等于消费商品2的边际效用。超过了这一点，消费者会将额外的收入都花费在商品2上。

因此，最好将商品2的需求函数写为：

$$x_2 = \begin{cases} 0 & \text{当} m \le p_2 \text{时} \\ m/p_2 - 1 & \text{当} m > p_2 \text{时} \end{cases}$$

如果想了解拟线性函数更多的性质，请参见 Hal R. Varian, Microeconomic Analysis, 3rd ed.（New York : Norton, 1992）。

总结

1.消费者对某商品的需求函数一般取决于所有商品的价格以及他的收入。

2.正常商品的需求随着收入的增加而上升；劣等商品的需求随着收入增加而下降。

3.普通商品的需求随着价格上升而下降；吉芬商品的需求随着价格上升而上升。

4.若商品 2 的价格上升时商品 1 的需求增加，则商品 1 是商品 2 的替代品；如果该情形下，商品 1 的需求不是增加而是下降的，则商品 1 是商品 2 的补充品。

5.反需求函数衡量任一既定需求量对应的价格。（反）需求函数在既定消费水平上的高度，衡量在该消费水平上若增加额外一单位商品的消费，消费者的边际支付意愿是多少。

复习题

1.如果某消费者只消费两种商品，而且他总是将钱全部花完，那么这两种商品能都是劣等商品吗？

2.说明完全替代类型的偏好是位似偏好（homothetic preferences）。

3.说明柯布—道格拉斯类型的偏好是位似偏好。

4.收入提供曲线对于恩格尔曲线来说，正如价格提供曲线对于什么曲线？

5.如果消费者对某两种商品的偏好是凹的，他会同时消费这两种商品吗？

6.汉堡包（hamburgers）和葡萄干夹心小面包(buns)是互补的还是替代的？

7.完全互补情形下，商品 1 的反需求函数是什么样子的？

8.判断对错。如果需求函数为 $x_1 = -p_1$，则反需求函数为 $x_1 = -1/p_1$.

曹乾（东南大学 caoqianseu@163.com）

复习题答案

1.如果某消费者只消费两种商品，而且他总是将钱全部花完，那么这两种商品能都是劣等商品吗？

【考察内容】劣等商品的概念。

【参考答案】

第一种方法。

不妨假设该消费者的收入增加，增加了 Δm 。如果他仍然将 Δm 花完，则至少有一种商品的购买量增加。根据劣等商品的定义可知，这种商品显然不是劣等商品。由此可见，两种商品不可能都是劣等商品。

第二种方法（反证法）。

假设这两种商品都为劣等商品，由劣等商品的定义可知 $\frac{dx_i(p_1,p_2,m)}{dm}<0$ ，其中 i=1,2 表示商品 1 和 2； $x_i(p_1,p_2,m)$ 表示第 i 种商品的需求。因此必有 $p_1\frac{dx_1}{dm}+p_2\frac{dx_2}{dm}<0$ 。

由题意知， $p_1x_1+p_2x_2=m$ 。该式两边同时对 m 求导可得： $p_1\frac{dx_1}{dm}+p_2\frac{dx_2}{dm}=1$ 。矛盾。因此不可能都是劣等商品。

2.说明完全替代类型的偏好是位似偏好（homothetic preferences）。

【考察内容】完全替代；位似偏好。

【参考答案】

完全替代的效用函数具有以下形式： $u(x_1,x_2)=ax_1+bx_2$ 。因此，如果 $u(x_1,x_2)>u(y_1,y_2)$ ，这等价于 $ax_1+bx_2>ay_1+ay_2$ 。由此可知，对于 $t>0$ ， $tax_1+tbx_2>tay_1+tay_2$ ，因此 $u(tx_1,tx_2)>u(ty_1,ty_2)$ 。根据位似偏好的定义可知， $u(x_1,x_2)=ax_1+bx_2$ 为位似偏好。证毕。

3.说明柯布—道格拉斯类型的偏好是位似偏好。

【考察内容】柯布—道格拉斯类型的偏好；位似偏好。

【参考答案】

假设柯布—道格拉斯效用函数为 $u(x_1,x_2)=x_1^a x_2^{1-a}$ 。

则 $u(tx_1,tx_2)=(tx_1)^a(tx_2)^{1-a}=t^at^{1-a}x_1^a x_2^{1-a}=tx_1^a x_2^{1-a}=tu(x_1,x_2)$ 。

因此，若$u(x_1,x_2)>u(y_1,y_2)$，则$u(tx_1,tx_2)=tu(x_1,x_2)>tu(y_1,y_2)=u(ty_1,ty_2)$，根据位似偏好的定义，可知柯布—道格拉斯类型的偏好是位似偏好。证毕。

4.收入提供曲线对于恩格尔曲线来说，正如价格提供曲线对于什么曲线？

【考察内容】收入提供曲线；恩格尔曲线；价格提供曲线；需求曲线。

【参考答案】需求曲线。

收入提供曲线和恩格尔曲线都是描述收入变动对商品需求影响的工具。

收入提供曲线是收入增加时，预算线向外平移，预算线平移时会产生一系列需求束，将这些需求束连接起来就得到了收入提供曲线。简单地说，收入提供曲线说明了不同收入水平下的相应的需求束。

由收入提供曲线可得到恩格尔曲线。如果我们在上述商品束中，关注其中一种商品（比如商品 1）的需求与收入变动的关系，这就是恩格尔曲线。当然，由两种商品的恩格尔曲线也可以得到收入提供曲线。

价格提供曲线和需求曲线都是描述价格变动对商品需求影响的工具。

价格提供曲线：假设只有一种商品（比如商品 1）价格变动（商品 2 的价格和消费者的收入都不变），则预算线会随商品的价格变动而转动，每一条无差异曲线与相应的预算线相切，把最优点（即切点）连接起来就得动了价格提供曲线。简单地说，价格提供曲线说明了不同价格水平下相应的需求束。

由价格提供曲线可得到需求曲线。在上述需求束中，如果我们仅商品 1 的需求与它自身价格变动的关系，这就是商品 1 的需求曲线。当然，由两种商品的需求曲线也可以得到价格提供曲线。

5.如果消费者对某两种商品的偏好是凹的，他会同时消费这两种商品吗？

【考察内容】凹偏好

【参考答案】不会。凹偏好情形下的最优解为角点解（corner solutions）,因此必然有一种商品的消费量为 0。请参考第 4 章复习题的第 4 题的解答。这两个题目完全一样。

6.汉堡包（hamburgers）和葡萄干夹心小面包(buns)是互补的还是替代的？

【考察内容】互补；替代

【参考答案】此题涉及西方人的饮食习惯。在美国，很多人将汉堡包和葡萄干小面包搭配着吃，因此它们是互补的。但对于一些人来说，它们也可能是互相替代的。

7.完全互补情形下，商品 1 的反需求函数是什么样子的？

【考察内容】完全互补；反需求函数

【参考答案】假设商品 1 和商品 2 的按 a:b 的比例互补（$\frac{x_1}{x_2}=\frac{a}{b}$），即 a 单位商品要和 b 单

位商品搭配。由此可知$\frac{b}{a}x_1 = x_2$将其代入预算方程$p_1x_1 + p_2x_2 = m$可得商品 1 的反需求函数为$p_1(x_1) = \frac{m}{x_1} - \frac{b}{a}p_2$.

8.判断对错。如果需求函数为$x_1 = -p_1$，则反需求函数为$x_1 = -1/p_1$.

【考察内容】需求函数；反需求函数。

【参考答案】

错误。由$x_1 = -p_1$可知其反需求函数为$p_1(x_1) = -x_1$。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

7.显示偏好（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

7 显示偏好

在第 6 章，我们已知道怎样使用消费者的偏好和预算约束来确定他的需求。在本章我们颠倒这一过程，看看怎样利用消费者的需求信息去发现他的偏好情况。直到目前，我们的思考方式都是利用偏好推测消费者的行为。但在现实生活中，偏好不可直接观察到，也就是说我们必须通过人们的行为来推测他们的偏好。为做此事，本章将研发一些工具。

在谈及通过观察行为来确定偏好时，必须假设：在观察期消费者的偏好保持不变。如果时间间隔很长，这个假设不太合理。但经济学家研究的情形通常是月度或季度时间段。在这样的时间段，消费者的爱好（taste）似乎不可能发生大的变动。因此，我们接受上述假设，即在观察消费者的选择行为期间，他的消费偏好保持不变。

7.1 显示偏好的思想

在分析之前，我们按照惯例作出约定：不管人们的潜在偏好如何，都必须为严格凸的。这一假设贯穿本章。如此假设的目的是保证每个预算集都有唯一的需求束。研究显示偏好理论并不一定需要作出上述假设，但这样做能使阐述过程简单明了。

请看图 7.1，图中画出了消费者的需求束(x_1,x_2)以及另外一个任意商品束(y_1,y_2)，该商品束位于消费者的预算线之下。假设此消费者的目的是效用最大化，那么它对这两个商品束的偏好如何？

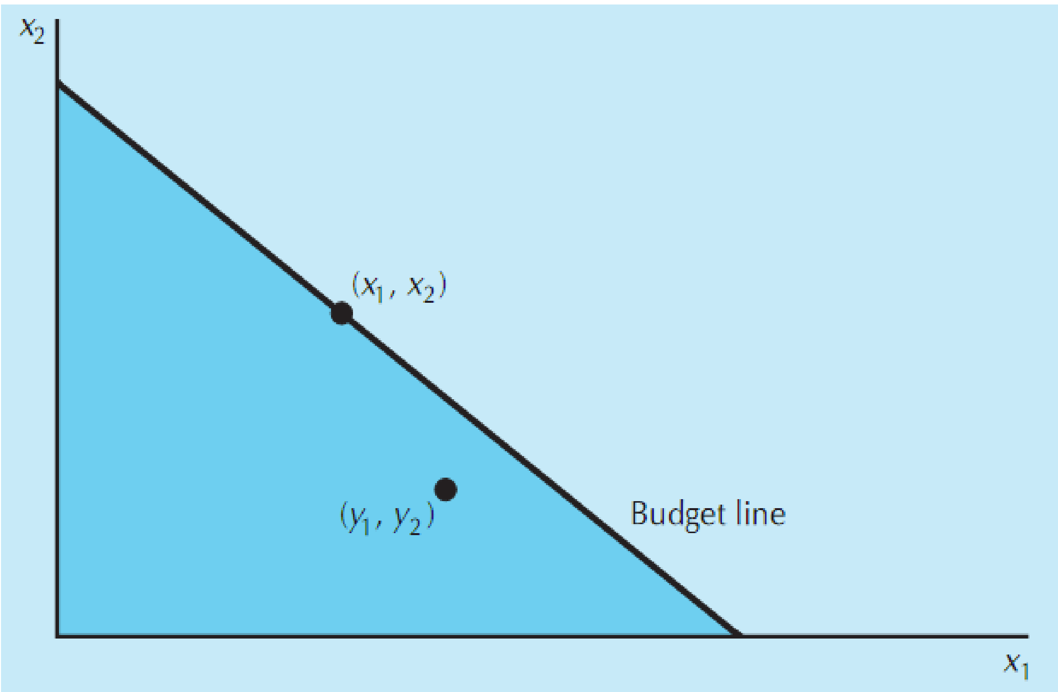


图 7.1：显示偏好。消费者选择了商品束(x_1,x_2)，这表明(x_1,x_2)被显示偏好于(y_1,y_2)，因

为他原本可以选择(y_1, y_2)但没选。

如图所示，在给定的预算情形下，(y_1, y_2)当然能买得起，也就是说，如果消费者想买该商品束，他就能买而且还有余钱。因为(x_1, x_2)是**最优的**，它肯定比消费者能买得起的其他任何商品束都要好。特别地，它肯定比(y_1, y_2)好。

类似的结论对于预算线上（on）以及预算线下面的其他消费束都成立。既然这些消费束都**可以买**但并未买，因此**实际购买的**商品束一定更好。此处我们用到了上面提及的假设，即每个预算都有**唯一的**需求束。如果偏好不是严格凸，比如无差异曲线有直线线段，那么**预算线上**（on）就可能存在和需求束一样好的其他点。处理这种稍微复杂点的情形并不难，但更容易的做法就是使用上面的假设排除这种情况。

在图 7.1 中，预算线下面阴影区域中的商品束，都被显示比需求束(x_1, x_2)差。这是因为它们都可以买得到但并未买，实际购买的是(x_1, x_2)。现在我们将用几何图形表示的显示偏好理论翻译成代数语言。

假设当消费者收入为m，两商品的价格为(p_1, p_2)时，需求束为(x_1, x_2)。在同样的收入和价格水平下，(y_1, y_2)能买得起的意思是什么？它的意思是说(y_1, y_2)满足预算约束

$$p_1 y_1 + p_2 y_2 \le m.$$

既然在这一给定的预算下，实际购买的是(x_1, x_2)，它必须满足预算等式约束

$$p_1 x_1 + p_2 x_2 = m.$$

把这两个式子放在一起比较。预算为(p_1, p_2, m)时，能买得起(y_1, y_2)意味着

$$p_1 x_1 + p_2 x_2 \ge p_1 y_1 + p_2 y_2.$$

如果上述不等式成立，且两个商品束不同，我们说(x_1, x_2)**被直接显示偏好**（directly revealed preferred）于(y_1, y_2)。

注意，不等式的左端是，在价格(p_1, p_2)下**实际购买的**商品束的支出。因此，显示偏好表示的是在某预算情形下，下列二者之间的关系：实际购买的商品束与**本来可买但未买**的商品束。

术语“显示偏好”有些误导。它在本质上和偏好一点关系也没有，尽管我们已知道如果消费者做出最优选择，偏好和显示偏好密切相关。因此，与其说“X 被显示偏好于 Y”，不如说“X 优先于 Y 被选择。”当我们说 X 被显示偏好于 Y，我们全部的意思是说，在 X 和 Y 都能被选择的情况下，实际选择的是 X，即$p_1 x_1 + p_2 x_2 \ge p_1 y_1 + p_2 y_2$。

7.2 从显示偏好到偏好

我们将上一节的内容简单总结一下。根据消费者行为理论—人们选择他们能买得起的商

品束中最好的，即他们实的选择的商品束比他们能选择但未选择的商品束更受偏好。用上一节的术语来说，如果(x_1, x_2)被直接显示偏好于(y_1, y_2)，则(x_1, x_2)比(y_1, y_2)更受偏好。用更正式地语言表达：

显示偏好原理（The Principle of Revealed Preference）。**令(x_1, x_2)是价格水平为(p_1, p_2)时消费者选择的商品束，(y_1, y_2)是满足条件$p_1x_1 + p_2x_2 \geq p_1y_1 + p_2y_2$的其他商品束，则如果消费者在他能买得起的商品束中选择最好的，必有$(x_1, x_2) \succ (y_1, y_2)$。**

如果你第一次遇到这个原理，你会感觉它似乎是循环论证。如果 X 被显示偏好于 Y，这难道不是自动意味着 X 比 Y 更受偏好？答案是否定的。“显示偏好”表示当 Y 可选时实际选择的是 X；“偏好”则表明消费者将 X 排在 Y 的前面。如果消费者选择能买得起的消费束中最好的，则“显示偏好”意味着“偏好”，但这个结论是根据消费者行为模型推知，而不是由显示偏好的定义推知。

这就是我们建议使用某商品束比其他商品束“被优先选择”这种表达方法的原因。这样，我们可将显示偏好原理表达如下：“如果 X 被先于 Y 选择，则 X 一定比 Y 更受偏好。”这种表述明确告诉我们，如何根据消费者行为模型来推测他潜在偏好的信息。

不管你使用哪种术语，关键要点是明晰的：如果我们看到消费者在 Y 商品束可选的情况下，选择了 X，则我们就得到了关于这两种商品束的偏好信息，即 X 比 Y 更受偏好。

假设我们碰巧知道，价格为(q_1, q_2)时需求束为(y_1, y_2)。(y_1, y_2)被显示偏好于另一消费束比如(z_1, z_2)，即：

$$q_1y_1 + q_2y_2 \geq q_1z_1 + q_2z_2.$$

于是我们可得$(y_1, y_2) \succ (z_1, z_2)$，在前面我们已得出$(x_1, x_2) \succ (y_1, y_2)$，由传递性假设可知$(x_1, x_2) \succ (z_1, z_2)$。

这个结论可用图 7.2 说明。显示偏好和传递性告诉我们，对于作出上述选择的消费者，(x_1, x_2)一定比(z_1, z_2)好。

这种情形下，自然可说(x_1, x_2)被间接显示偏好（indirectly revealed preferred）于(z_1, z_2)。当然，观察的选择“链条”可长于三个：若 A 被直接显示偏好于 B，B 被直接显示偏好于 C，C 被直接显示偏好于 D，…比如，直到 L 被直接显示偏好于 M，那么 A 仍然被间接显示偏好于 M。用于直接比较的链条可以任意长。

如果一个商品束直接或间接被显示偏好于另一个商品束，我们可说第一个商品束被显示偏好（revealed preferred）于第二个。显示偏好的思想比较简单，但它的威力非常强大。仅仅观察消费者的选择，就可以得到那么多的关于他的潜在偏好的信息。例如，考察图 7.2，

该图描述了若干不同预算下的需求束。这些信息表明(x_1, x_2)被直接或间接显示偏好于阴影区域中的所有商品束。对于作出这样选择的消费者而言，他最偏好(x_1, x_2)。换一种说法，这意味着通过(x_1, x_2)的无差异曲线，无论形状如何，必然位于阴影区域的上方。

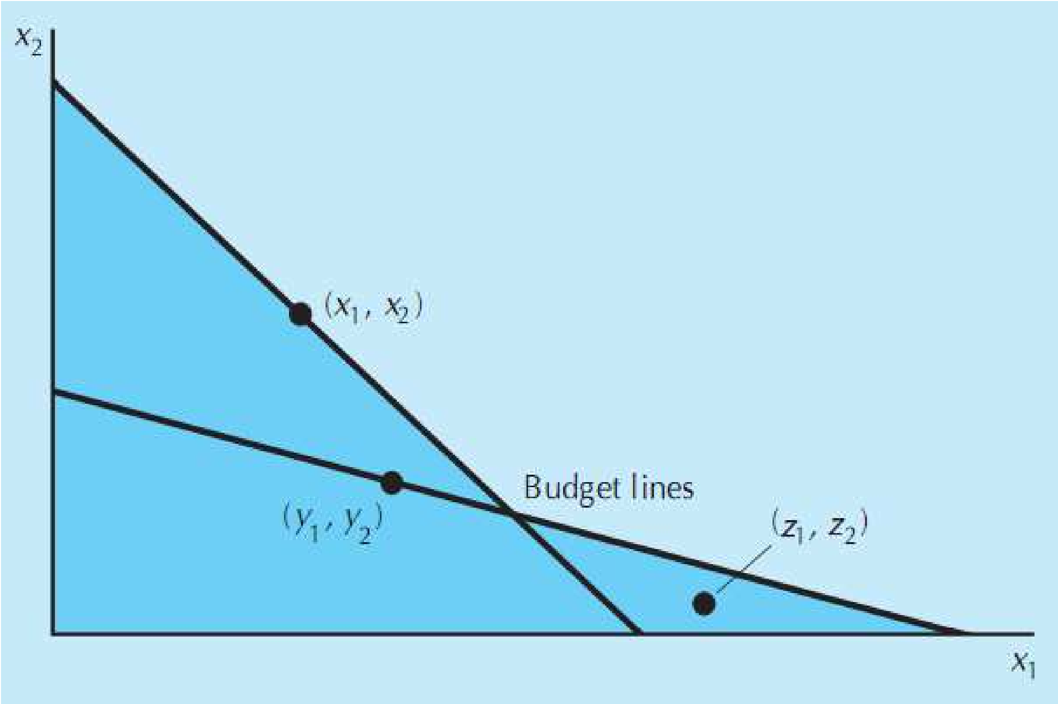


图 7.2：间接显示偏好。(x_1, x_2)被直接显示偏好于(y_1, y_2)，(y_1, y_2)被直接显示偏好于(z_1, z_2)，则称(x_1, x_2)被间接显示偏好于(z_1, z_2)。

7.3 复原偏好

观察消费者的选择，可知他的偏好信息。他作出的选择越多，对他的偏好形状估计越准确。

这些偏好信息对于政策制定非常重要。大多数经济政策涉及不同商品之间的权衡选择（trade-off），比如我们对鞋子征税但对衣服给与补贴，最终结果可能是鞋子变少衣服变多。为了评价这样的政策是否受大众欢迎，需要知道消费者对于鞋子和衣服的偏好信息。通过观察消费者的选择，我们可以使用显示偏好及其相关技术提取这类偏好信息。

如果我们对消费者的偏好增加更多的限制，我们就能更准确地估计出无差异曲线的形状。例如，假设我们观测到，Y 和 Z 这两个商品束被显示偏好于 X，如图 7.3 所示。如果增加偏好为凸的假设，那么可知 Y 和 Z 的所有加权平均值也比 X 好。如果再增加偏好为单调的假设，那么比 X，Y 和 Z 含有更多两种商品的所有商品束，都比 X 好；X，Y 和 Z 的任一加权平均数也比 X 要好。

图 7.3 中标记为“更差的商品束”表示，X 被直接显示偏好于该区域的所有商品束。也

就是说，该区域的由这样的商品束组成：花费比 X 小的所有商品束（这些商品束记为 F）；花费比 F 商品束还小的所有商品束，如此类推。

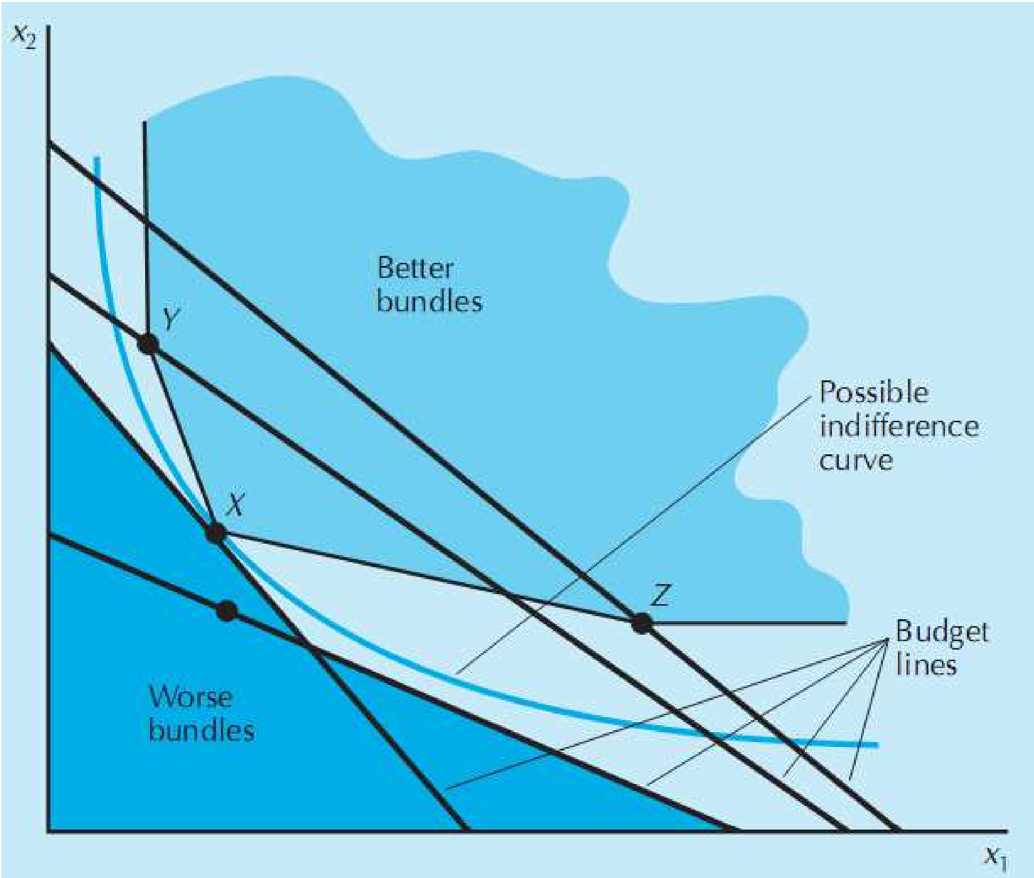


图 7.3:复原无差异曲线。X 上方的阴影区域包含的商品束都比 X 好，X 下方的阴影区域包含的商品束都比 X 差，因此包含 X 的无差异曲线必然位于这两个阴影区域之间。

根据作出上述选择的消费者的偏好信息，可得结论：在图 7.3 中，位于上方阴影区域中的所有商品束都比 X 好，位于下方阴影区域中的所有商品束都比 X 差。穿过 X 的无差异曲线必然位于两个阴影区域之间的某个位置。你看，你仅仅巧妙使用了显示偏好的思想以及关于偏好的一些简单假设，就已经紧紧地捕捉到无差异曲线。

7.4 显示偏好弱公理

以上内容依赖于下列假设：消费者有偏好而且他总是从买得起的商品束中选择最好的那个。如果消费者的行为并非如此，我们在上一节对无差异曲线的“估计”就没意义。由此自然产生的问题是：我们如何判断消费者的行为符合效用最大化模型？或者反过来问，即根据什么样的观测信息可知消费者的行为**不符合**效用最大化模型？

考虑图 7.4 阐述的情形。这两个选择都是由某个追求效用最大化的消费者作出的吗？根据显示偏好的逻辑，由图 7.4 可知两件事：（1）(x_1,x_2) 比 (y_1,y_2) 更受偏好；（2）(y_1,y_2) 比 (x_1,x_2) 更受偏好。这显然是荒谬的。在图 7.4 中，显然消费者在可以选择 (y_1,y_2) 的情况下，选择了 (x_1,x_2)，表明 (x_1,x_2) 比 (y_1,y_2) 好；但是，他在可以选择 (x_1,x_2) 的情形下，又选择了 (y_1,y_2)，这又表明 (y_1,y_2) 比 (x_1,x_2) 好！

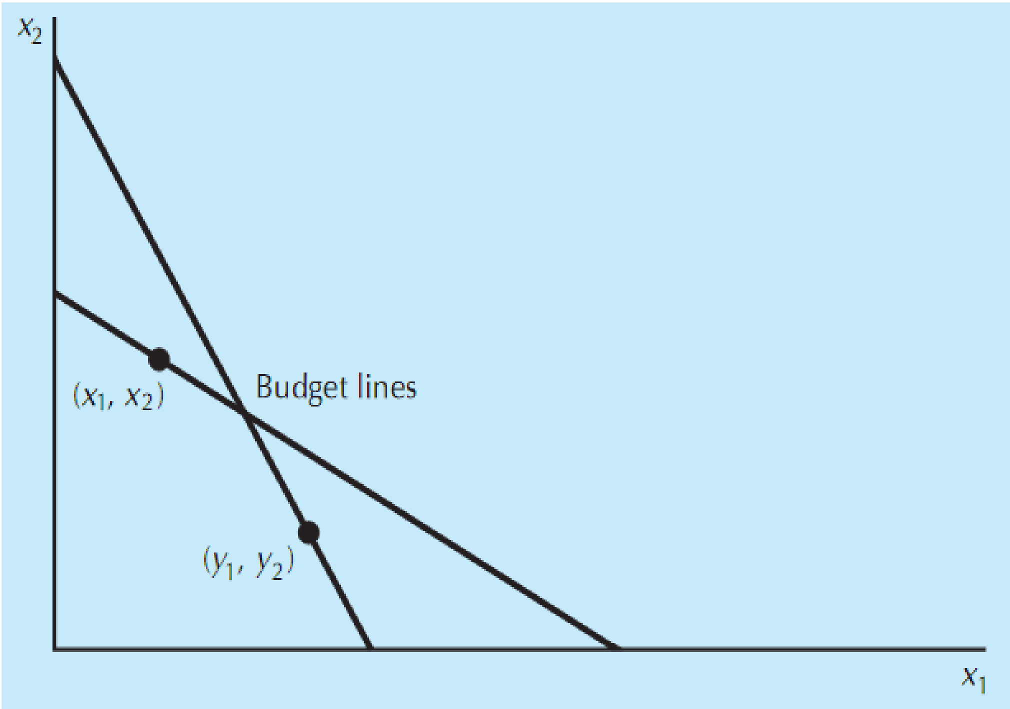


图 7.4：违背显示偏好弱公理的情形。某个消费者如果选择了 (x_1,x_2) 和 (y_1,y_2)，他的行为就违背了显示偏好弱公理。

显然，该消费者不是追求效用最大化的消费者。要么该消费者选择的不是他能买的起消费束中最好的，要么是他的选择决策的其他方面发生了变化，但这个变化我们还未观测到。任何类似上述的情形都不符合不变环境下的消费者选择模型。

消费者选择理论意味着图 7.4 中的情形不会发生。如果消费者选择他们能买得起的商品束中最好的那个，那么能买得起但并未买的商品束，一定比实际购买的商品束差。经济学家把这个简单的观点归纳成为消费者理论中的一个基本公理，即：

显示偏好弱公理（Weak Axiom of Revealed Preference，WARP）。**如果 (x_1,x_2) 被直接显示偏好于 (y_1,y_2)，而且这两个商品束不是同一个商品束，则 (y_1,y_2) 被直接显示偏好于 (x_1,x_2) 的情形不可能发生。**

换句话说，如果价格为(p_1, p_2)时购买(x_1, x_2)，而价格为(q_1, q_2)时购买另一个商品束(y_1, y_2)，则如果有：

$$p_1x_1 + p_2x_2 \geq p_1y_1 + p_2y_2 .$$

那么下式就**不成立**：

$$q_1y_1 + q_2y_2 \geq q_1x_1 + q_2x_2 .$$

用语言表述：如果能买得起 Y 商品束但却购买了 X 商品束，那么如果购买的是 Y 商品束，则说明此时买不起 X 商品束。

图 7.4 中的消费者**违背**了显示偏好弱公理。因此，我们知道这个消费者的行为不可能是效用最大化行为[一]。

在图 7.4 中，我们无法画出使 x 和 y 这两个商品束都为效用最大的无差异曲线族。但图 7.5 中的消费者行为满足显示偏好弱公理，在此图中我们可以找到使其行为为最优行为的无差异曲线族。满足这样行为的无差异曲线族可能有很多，我们在图中画出了其中一种。

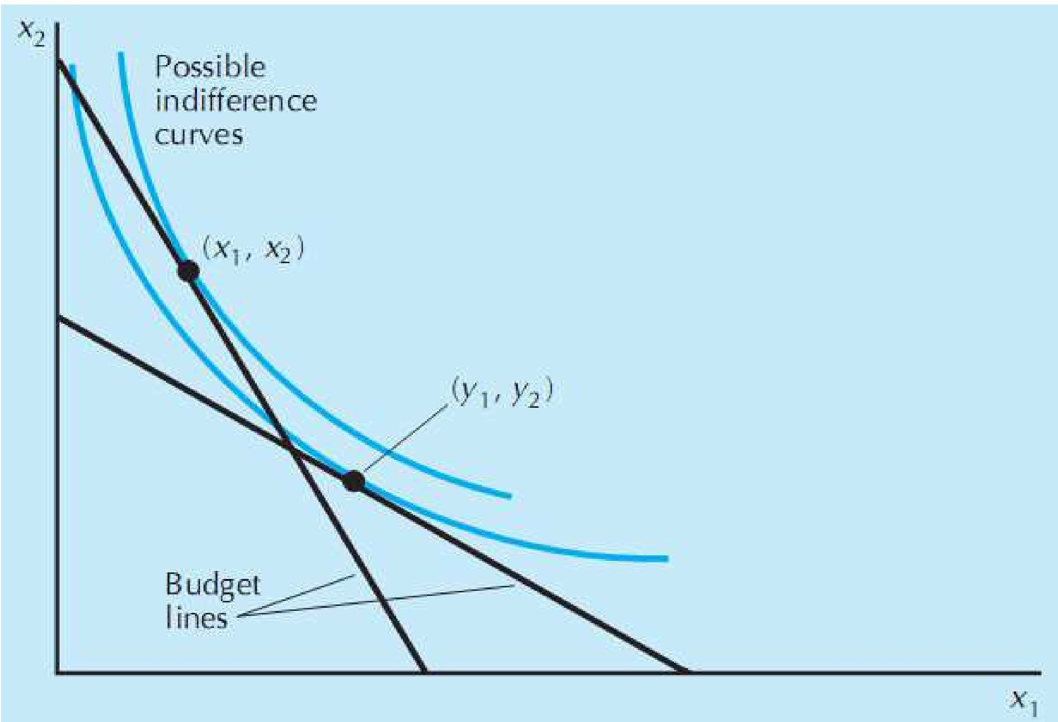


图 7.5：满足显示偏好弱公理。消费者的选择满足显示偏好弱公理以及可能的无差异曲线。

7.5 检验显示偏好弱公理

需要强调，显示偏好弱公理是一种条件，如果消费者总是在他能买得起的商品束中选择最好的，他必须满足这个条件。显示偏好弱公理是消费者这种行为模型的逻辑应用，因此，

[一] 我们能说他的行为满足显示偏好弱公理吗？嗯，可以这么说，但正式场合不要这么说。

我们可以使用该公理来检验某个消费者的行为，是否和我们的经济模型一致；如果将某个经济实体视作消费者，也可以使用该公理来检验它的行为是否和经济模型一致。

在实践中，我们怎样有条理地检验显示偏好弱公理？假设我们观察到在不同价格水平下，某个消费者选择了不同的商品束。假设我们第 t 次观察时价格为 (p_1^t, p_2^t)，选择的消费束为 (x_1^t, x_2^t)。我们使用具体的例子说明，相应数据请看表 7.1.

Observation	p_1	p_2	x_1	x_2
1	1	2	1	2
2	2	1	2	1
3	1	1	2	2

表 7.1：某些观测数据

给定表格中的数据，我们就可以计算出不同价格水平下购买不同商品束的支出。计算的结果列在表 7.2. 例如，第三行第一列的数字，表示在第三组价格水平下购买第一个消费束的费用。

		Bundles		
		1	2	3
Prices	1	5	4*	6
	2	4*	5	6
	3	3*	3*	4

表 7.2：不同价格水平下每个商品束的支出.

表 7.2 对角线中的数字表示选择束在当时价格下的支出。每一行中的其他数字表示在该价格水平下，如果购买其他商品束，应该花费多少钱。那么，如何判断第 3 个商品束是否被显示偏好于第 1 个商品束？我们检验第三行第一列的数字（在第三组价格水平下，若购买第一个商品束的支出）是否小于第三行第三列的数字（在第三组价格水平下购买第三个商品束的支出）。在这个具体的例子中，消费者在能买得起第 1 个商品束的情形下，却选择了第 3 个商品束，这表明第 3 个商品束被显示偏好于第 1 个商品束。因此，我们在第三行第一列的数字旁标记一个星号。

从数学的观点来看，我们只是在第 s 行第 t 列的数字旁标记一个星号，如果该数字小于第 s 行第 t 列的数字。

我们可以使用表 7.2 检验违背显示偏好弱公理的情形。在这种分析框架下，违背显示偏

好弱公理的情形包含 t 和 s 这两个观测时间点，使得第 t 行第 s 列含有一个星号，并且第 s 行第 t 列也含有一个星号。因为这表示 s 时刻购买的商品束被显示偏好于 t 时刻购买的商品束，但反过来说也成立。这就违背了显示偏好弱公理。

我们可以使用计算机（或者研究助理人员）进行检验，看看消费者的商品束中是否存在类似上述的情形，如果有，则他的选择和消费者理论不一致。不一致的原因可能是该理论对于这个特定的消费者来说是错的，也可能是因为他的消费环境发生了变化而我们没有控制这种变化。因此，显示偏好弱公理可以让我们轻松验证，消费者的某些选择是否符合消费者理论。

在表 7.2 中，我们看到第一行第二列含有一个星号，第二行第一列也含有一个星号。前半句话表明在第 2 观测点消费束可以买得起的情形下，消费者选择了第 1 观测点的消费束；后半句话的意思正好相反。这就违背了显示偏好弱公理。我们由此可以得出下列结论：表 7.1 和 7.2 中的数据不可能是理性消费者的数据，因为理性消费者的偏好是不变的，而且总是在他能买得起的商品束中选择最好的那个。

7.6 显示偏好强公理

上一节介绍的显示偏好弱公理，给我们提供了一个可观测到的条件，追求效用最大化的消费者必须满足这个条件。但是有时我们也会用到一个更强条件。

我们已经知道若 X 商品束被显示偏好于 Y，Y 商品束被显示偏好于 Z，则 X 一定被显示偏好于 Z。如果消费者的偏好不变，那么我们永远不可能看到下列情形，即从一系列选择中推出 Z 被偏好于 X。

显示偏好弱公理要求，若 X 被直接显示偏好于 Y，则我们绝不可能看到 Y 被显示偏好于 X。显示偏好强公理要求类似的条件对间接显示偏好也适用。更正式地，我们有：

显示偏好强公理（Strong Axiom of Revealed Preference，SARP）。**如果 (x_1,x_2) 被直接或间接显示偏好于 (y_1,y_2)，而且 (x_1,x_2) 和 (y_1,y_2) 是不同的商品束，那么 (y_1,y_2) 不可能被直接或间接显示偏好于 (x_1,x_2)。**

显然，如果我们观测到某消费者的行为是效用最大化行为，那么它必须满足显示偏好强公理。因为若消费者追求效用最大化，而且 (x_1,x_2) 被直接或间接显示偏好于 (y_1,y_2)，必然有 $(x_1,x_2)\succ(y_1,y_2)$. 然而若我们同时有 (x_1,x_2) 被显示偏好于 (y_1,y_2) 以及 (y_1,y_2) 被显示偏好于 (x_1,x_2)，即 $(x_1,x_2)\succ(y_1,y_2)$ 并且 $(y_1,y_2)\succ(x_1,x_2)$，这显然矛盾。由此我们可以得知，要么消费者不是追求效用最大化，要么消费环境（比如口味，其他商品价格等）发生了变化。

粗略地说，既然消费者的潜在偏好是传递的，则**显示**的偏好也是传递的。因此，显示偏好强公理是效用最大化行为的**必要**条件：若消费者总是选择他能买得起的消费束中最好的那个，那么观测到的行为符合显示偏好强公理。更奇妙的是，任何满足强公理的行为，都可以被认为是追求效用最大化的消费者表现出的行为。其中的原因在于，如果观测到的行为符合强公理，我们总能找到**能够**产生该行为的精确的、性状良好的偏好。在这种意义上，强公理是效用最大化行为的**充分**条件：若观测到的选择行为符合强公理，那么总能找到相应的偏好，使得观测到的行为是效用最大化行为。不幸的是，上述充分和必要条件的证明超出了本书的范围，但理解这两个条件的重要性并不难。

这个重要性在于，显示偏好强公理提供了判断某消费行为是否是效用最大化行为的**全部**条件。因为如果观测到的行为符合强公理，我们可以“构建”产生该行为的偏好。因此，显示偏好强公理是观测到的行为符合消费者理论的必要和充分条件。

这是否意味着，我们构建的偏好就是由观测到的行为实际产生的？当然不是。就象其他科学论断一样，我们只能说观测到的行为和构建的偏好模型不是不一致的。我们无法证明该经济模型是正确的；我们只能确定该模型的结论，看看观测到的行为是否和模型的结论一致。

7.7 检验显示偏好强公理

假设我们有表 7.2，如果观测点 t 的商品束被直接显示偏好于 s，则 t 行 s 列处标记了星号。我们怎样使用该表检验强公理？

最简单的方法是首先将该表转化。表 7.3 就是这样的一个例子。该表类似表 7.2，区别在于数字不同。此处，星号表示直接显示偏好，括号内星号的意思将在后文介绍。

		Bundles		
		1	2	3
Prices	1	20	10^*	$22^{(*)}$
	2	21	20	15^*
	3	12	15	10

表 7.3：怎样检验显示偏好强公理？

现在我们系统地审视表中的数字，看看是否存在这样的观测点链条，使得某个商品束被间接显示偏好于另一个。例如，商品束 1 被直接显示偏好于商品束 2，因为第 1 行第 2 列处有个星号。商品束 2 被直接显示偏好于商品束 3，因为第 2 行第 3 列处也有个星号。因此，商品束 1 被**间接**显示于偏好商品束 3，我们因此在第 1 行第 3 列处的**括号内**标记个星号。

一般来说，如果观测点很多，我们必须检查任意长的比较链条，看看是否存在间接显示偏好。用手工做此事比较繁琐，而编写个简单的计算机程序就可以解决问题。将表中的直

接显示偏好的数据输入程序，就可以计算出间接显示偏好关系，如果观测点 s 被间接显示偏好于观测点 t，计算机就在表格中的 st 处标记个星号。

一旦计算完毕，我们就可以容易地检验强公理。我们只要看看第 t 行第 s 列和第 s 行第 t 列是否都有星号。如果是这样，观测点 t 就被显示偏好于 s（直接或间接都可以），同时，观测点 s 也被显示偏好于 t。这就违背了显示偏好强公理。

与此相反，如果我们发现这样的情形，则可知这些观测值符合消费者理论。这些观测值可能就是理性消费者作出的，理性消费者追求效用最大化并且具有良好形状的偏好。因此我们就有了检验某消费者的行为是否符合消费者理论的检验方法。

这一点比较重要，因为我们可以把某些经济机构的行为，建立类似消费者行为的模型。例如，一个由若干人组成的家庭的行为。家庭的消费选择是否使"家庭效用"最大化？如果我们获知家庭消费的数据，我们可以使用显示偏好强公理进行判断。

另外一种可以视为消费者的机构是诸如医院或大学的非盈利组织。大学在做出经济决策时是否达到了效用最大化？如果我们得到大学在不同价格水平下的一系列选择的数据，我们可以回答这样的问题。

7.8 指数

假设我们分析某消费者在两个不同时期的消费束，我们希望比较消费的变动。令 b 表示基期（base period），t 表示其他时期。第 t 年的"平均"消费如何与基期的消费相比较？

假设在 t 期价格为 (p_1^t, p_2^t)，消费者的选择为 (x_1^t, x_2^t)。在基期 b，价格为 (p_1^b, p_2^b)，消费者的选择为 (x_1^b, x_2^b)。我们想问的问题是"平均"消费束是如何变动的。

令 w_1 和 w_2 表示计算加权平均值时的"权重"，则我们可以看看下面这样的数量指数（quantity index）：

$$I_q = \frac{w_1 x_1^t + w_2 x_2^t}{w_1 x_1^b + w_2 x_2^b}.$$

如果 $I_q > 1$，则表明 t 期与 b 期相比，"平均"消费上升了；相反如果 $I_q < 1$，则表明"平均"消费下降了。

问题在于我们使用什么东西作为权重？一个自然的选择是使用商品的价格，因为它们在某种程度上代表两种商品的相对地位。但由于这里有两组价格即 t 期的价格与 b 期的价格，我们应该用哪一组价格？

如果我们用基期的价格作为权重，由此计算出的指数称为拉斯贝尔指数（Laspeyres index）简称**拉氏指数**，如果我们用 t 期的价格作为权重，由此计算出的指数称为**帕氏指数**

（Paasche index）[一]。这两个指数都可用来衡量“平均”消费的变化，区别在于它们所用的权重不同。

在前面的数量指数表达式中若用 t 期的价格作为权重，可得到**帕氏数量指数**

$$P_q = \frac{p_1^t x_1^t + p_2^t x_2^t}{p_1^t x_1^b + p_2^t x_2^b}.$$

如果上式中用 b 期的价格作为权重，即可得到**拉氏数量指数**

$$L_q = \frac{p_1^b x_1^t + p_2^b x_2^t}{p_1^b x_1^b + p_2^b x_2^b}.$$

拉氏指数和帕氏指数的大小可用来衡量消费者福利的变化。假设某情形下帕氏数量指数大于 1：

$$P_q = \frac{p_1^t x_1^t + p_2^t x_2^t}{p_1^t x_1^b + p_2^t x_2^b} > 1$$

与 b 期相比，消费者在 t 期的状况有何变化？

用显示偏好理论可得出答案。将上述不等式交叉相乘可得

$$p_1^t x_1^t + p_2^t x_2^t > p_1^t x_1^b + p_2^t x_2^b,$$

这个式子立即表明消费者在 t 期的状况必然比在 b 期好，因为在 t 期，他本来可以购买 b 期的消费束但没买。

如果帕氏数量指数**小于** 1，结果又会如何？如果它小于 1，我们可得到

$$p_1^t x_1^t + p_2^t x_2^t < p_1^t x_1^b + p_2^t x_2^b,$$

这个式子是说当消费者选择 t 期的消费束 (x_1^t, x_2^t) 时，他买不起 b 期的消费束 (x_1^b, x_2^b)。需要注意的是，这个式子没有涉及对消费束进行排序。你**不能**因为你买不起某消费束，就认为这个消费束比你正在消费的消费束好。

下面来看看拉氏指数。拉氏指数的原理和帕氏指数的原理类似。假设拉氏指数**小于** 1：

$$L_q = \frac{p_1^b x_1^t + p_2^b x_2^t}{p_1^b x_1^b + p_2^b x_2^b} < 1.$$

将上式交叉相乘可得

$$p_1^b x_1^b + p_2^b x_2^b > p_1^b x_1^t + p_2^b x_2^t,$$

这个式子表明 (x_1^b, x_2^b) 被显示偏好于 (x_1^t, x_2^t)。因此消费者在在 b 期的状况必然比在 t 期好。

[一] 拉斯贝尔（1834-1913），是一位法国经济学家。帕氏（1851-1925）为德国统计学家。译者注。

7.9 价格指数

价格指数和上述数量指数的原理类似。一般来说，价格指数是价格的加权平均：

$$I_p = \frac{p_1^t w_1 + p_2^t w_2}{p_1^b w_1 + p_2^b w_2}.$$

在这种情形下，自然可以选择商品的数量作为权重。和数量指数一样，价格指数也有两种。如果我们选取t期的数量作为权重，就得到了**帕氏价格指数**：

$$P_p = \frac{p_1^t x_1^t + p_2^t x_2^t}{p_1^b x_1^t + p_2^b x_2^t},$$

如果我们选择基期的数量作为权重，就得到了**拉氏价格指数**：

$$L_p = \frac{p_1^t x_1^b + p_2^t x_2^b}{p_1^b x_1^b + p_2^b x_2^b}.$$

如果帕氏指数小于 1，根据显示偏好理论，你能比较一下消费者在t期和在b期的福利状况吗？无法比较。因为上述指数的分子和分母用不同的价格作为权重，因此显示偏好理论用不起来。

下面我们引入衡量总支出变动的指数，定义如下

$$M = \frac{p_1^t x_1^t + p_2^t x_2^t}{p_1^b x_1^b + p_2^b x_2^b}.$$

它是t期总支出与b期总支出的比值。

现在假设你已知道帕氏价格指数大于 M，即

$$P_p = \frac{p_1^t x_1^t + p_2^t x_2^t}{p_1^b x_1^t + p_2^b x_2^t} > \frac{p_1^t x_1^t + p_2^t x_2^t}{p_1^b x_1^b + p_2^b x_2^b}.$$

消去不等式左右两端的分子，交叉相乘，可得

$$p_1^b x_1^b + p_2^b x_2^b > p_1^b x_1^t + p_2^b x_2^t.$$

这个式子表明，消费者在b期选择的消费束被显示偏好于t期的消费束。这意味着如果帕氏价格指数大于支出指数，则消费者在b期的状况比在t期好。

这个结论符合我们的直觉。因为从b期到t期，如果价格上升的幅度大于收入上升幅度，则消费者的状况会恶化。我们已在上面使用显示偏好理论证明了这种直觉的正确行。

对于拉氏价格指数也可作出类似的结论。如果拉氏价格指数小于 M，则消费者在t期的状况比在b期好。这再一次证明了我们的直觉：如果价格上升幅度小于收入上升幅度，则消费者的状况会改善。使用价格指数进行分析时，要紧的事情不是该指数是否大于 1，而是它

是否大于支出指数。

例子：社会保障金的指数化

对于很多老年人来说，社会保障金是他们收入的唯一来源。正因为此，政府希望调整社会保障金，以保证当商品价格变化时他们的社会保证金的购买力不变。由于这样的做法是将社会保障金的数额与价格指数或生活成本指数挂钩，因此这样的方案称为**指数化**（indexing）。

有些经济学家提出了这样的一种指数化方案。在某个基期年度 b 测量老年人的平均消费束。在随后的年度，社会保障金系统根据消费物价相应调整保障金得数额，使得一般（average）老年人的"购买力"保持不变，即让领取社会保障金的一般老年人恰好还能买得起 b 期的消费束[一]。如图 7.6 所示。

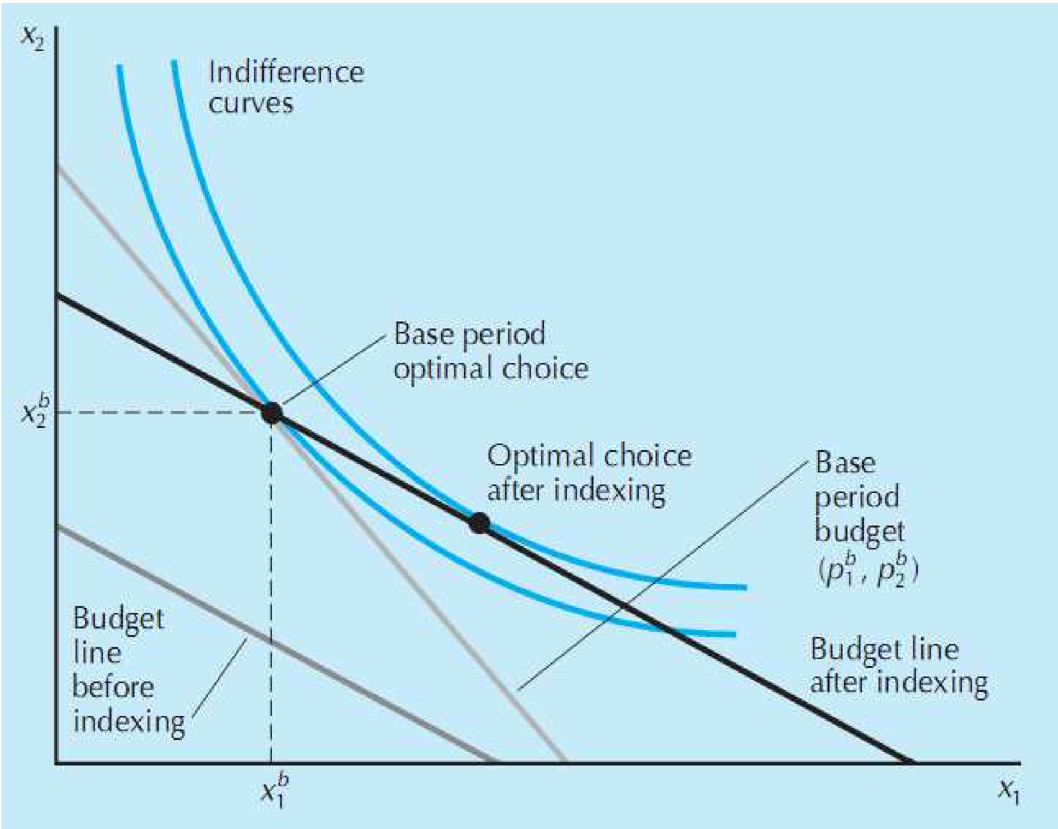


图 7.6：社会保障金。价格变动通常会使消费者的状况变好(与基期的状况相比)。

上述指数化方案有个奇特的结果：一般老年人的状况基本总是比他在 b 期的状况好。假设选择年度 b 作为价格指数的基期，则价格为 (p_1^b, p_2^b) 时的最优消费束为 (x_1^b, x_2^b)。这表明

[一] 这句话中有两个"一般的（average）", average 的意思是"一般的，通常的，典型的"，也就是"具有代表性的"意思；因此此处，"一般老年人"是指具有代表性的老年人。译者注。

价格为(p_1^b, p_2^b)时的预算线必然与通过(x_1^b, x_2^b)的无差异曲线相切。

现在假设价格变动，比如价格上升，因此如果没有社会保障，预算线将向内移动但不与原预算线平行。向内移动是因为两种商品的价格都上升；与原预算线不平行是因为两种商品价格上升幅度不同从而相对价格不同。由于商品价格上升，指数化方案必然是增加社会保障金以使得消费者在新价格下仍然能买得起原来的消费束(x_1^b, x_2^b)。但这意味着指数化后的预算线将穿过无差异曲线，因此，指数化后的预算线上必然存在着比(x_1^b, x_2^b)更好的消费束。所以，这种情形下，消费者选择的消费束通常好于他在基期选择的消费束。

总结

1.如果消费者在原本可以选择消费束 Y 的情况下却选择了消费束 X，则消费束 X 被显示偏好于消费束 Y，或简单地说 X 比 Y 好。

2.如果消费者总是在他能买得起的消费束中选择他最偏好的，则被选中的消费束一定好于其他能买得起但又没买的消费束。

3.可以通过观察消费者的选择行为，来“还原”或者估计隐藏在他的选择行为下的偏好。我们获得的该消费者的选择数据越多，对产生这些选择的潜在偏好的估计就越准确。

4.如果消费者的行为符合最优化选择的经济模型，则他的选择行为必然要满足显示偏好弱公理（WARP）和显示偏好强公理（SARP）。

复习题

1.当价格为$(p_1, p_2) = (1, 2)$时，某消费者的需求为$(x_1, x_2) = (1, 2)$；当价格为$(q_1, q_2) = (2, 1)$时，他的需求为$(y_1, y_2) = (2, 1)$。那么他的行为符合最大化行为模型吗？

2.当价格为$(p_1, p_2) = (2, 1)$时，某消费者的需求为$(x_1, x_2) = (1, 2)$；当价格为$(q_1, q_2) = (1, 2)$时，他的需求为$(y_1, y_2) = (2, 1)$。那么他的行为符合最大化行为模型吗？

3.在上题中，消费者更偏好哪个消费束，是 x 消费束还是 y 消费束？

4.在教材中我们已经知道社会保障金根据价格变动而相应调整，会使保障金领取人的状况至少和基期一样好。价格怎样变动才能做到下列这一点：无论他的偏好是什么类型，价格变动后他的状况恰好和基期一样好。

5.本题与上题的基本信息相同，但问题变为：对于什么样的偏好才能做到下列这一点：无论价格怎样变动，他的状况仍恰好和基期一样好。

复习题答案

1. 当价格为 $(p_1, p_2) = (1, 2)$ 时，某消费者的需求为 $(x_1, x_2) = (1, 2)$；当价格为 $(q_1, q_2) = (2, 1)$ 时，他的需求为 $(y_1, y_2) = (2, 1)$。那么他的行为符合最大化行为模型吗？

【复习内容】显示偏好弱公理

显示偏好弱公理的基本思想如下：

如果价格为 (p_1, p_2) 时购买 (x_1, x_2)，而价格为 (q_1, q_2) 时购买另一个商品束 (y_1, y_2)，则如果有：

$$p_1x_1 + p_2x_2 \geq p_1y_1 + p_2y_2 .$$

那么下式就不成立：

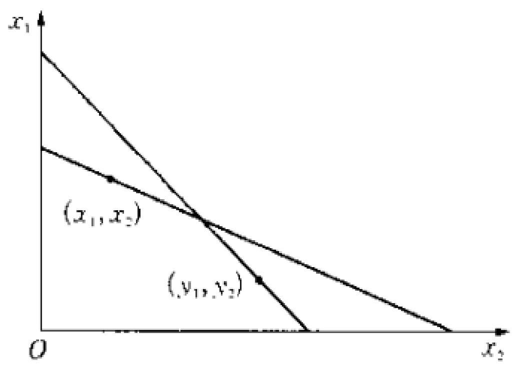
$$q_1y_1 + q_2y_2 \geq q_1x_1 + q_2x_2 .$$

用语言表述：如果能买得起 y 商品束但却购买了 x 商品束，那么如果购买的是 y 商品束，则说明此时买不起 x 商品束。

【参考答案】

方法一：画图分析。

由 $(p_1, p_2) = (1, 2)$ 时 $(x_1, x_2) = (1, 2)$ 可知，此时预算线斜率 $k = -\dfrac{p_1}{p_2} = -\dfrac{1}{2}$，并且通过点（1,2），容易画出这样的预算线。同理可以画出价格 $(q_1, q_2) = (2, 1)$ 时 $(y_1, y_2) = (2, 1)$ 的预算线。见下图。



由图可知，该消费者在原本可以选择 y 消费束的情形下却选择了 x，因此如果他的行为

符合显示偏好弱公理，则意味着他选择 y 时，x 肯定买不起。

但此图却又表明，他选择 y 时，x 是可以买得起的。

因此，他的行为不符合最大化模型。

方法二：代数计算。

根据题目已知条件可得 $p_1x_1+p_2x_2=5;\ p_1y_1+p_2y_2=4$，因此 $p_1x_1+p_2x_2\geq p_1y_1+p_2y_2$。根据显示偏好弱公理可知，此时不可能再有 $q_1y_1+q_2y_2\geq q_1x_1+q_2x_2$。但根据题目已知条件可以计算出 $q_1y_1+q_2y_2=5$；$q_1x_1+q_2x_2=4$，该式成立。这显然违背了显示偏好弱公理，也就意味着该消费者的行为不符合最大化模型。

2. 当价格为 $(p_1,p_2)=(2,1)$ 时，某消费者的需求为 $(x_1,x_2)=(1,2)$；当价格为 $(q_1,q_2)=(1,2)$ 时，他的需求为 $(y_1,y_2)=(2,1)$。那么他的行为符合最大化行为模型吗？

【复习内容】显示偏好弱公理

【参考答案】

本题的解题思路同上题相同。也可以使用几何或者代数方法说明。此处只给出代数方法。

根据题目已知条件可得

$$p_1x_1+p_2x_2=4;\ p_1y_1+p_2y_2=5。$$

因此 $p_1x_1+p_2x_2\leq p_1y_1+p_2y_2$。这就是说消费者选择 x 消费束时，他买不起 y 消费束。类似地，$q_1y_1+q_2y_2=4$；$q_1x_1+q_2x_2=5$，因此 $q_1y_1+q_2y_2\leq q_1x_1+q_2x_2$，这就是说消费者选择 y 消费束时，他买不起 x 消费束。因此该消费者的行为没有违背显示偏好弱公理，意味着他的行为符合最大化行为模型。

这个题目初看起来有些“别扭”，所以多说几句。显示偏好弱公理要求：如果消费者从若干能买得起的消费束中选择了某个消费束比如 x 消费束，则只要这些消费束能够买得起，他就必须一直选择 x 消费束。换句话说，如果他选择了其他的消费束（比如 y 消费束）必然意味着他已经买不起 x 消费束。从这个角度来看，本题中的消费者显然没违背显示偏好弱公理。

3.在上题中，消费者更偏好哪个消费束，是 x 消费束还是 y 消费束？

【复习内容】显示偏好弱公理；偏好关系

【参考答案】

如果能比较哪个消费束更好，比如 x 比 y 好，意味着什么？根据显示偏好弱公理，这意味着只要 x 和 y 都能买得起，则消费者应该一直选择 x。

然而，根据第 2 题的分析可知，消费者在选择 x 消费束时他买不起 y 消费束，在选择 y 消费束时他买不起 x 消费束，也就是说根据第 2 题已知的条件，我们无法比较 x 和 y 哪个消

费束更好，换句话说我们无法判断该消费者更偏好哪个消费束。

4.在教材中我们已经知道社会保障金根据价格变动而相应调整，会使保障金领取人的状况至少和基期一样好。价格怎样变动才能做到下列这一点：无论他的偏好是什么类型，价格变动后他的状况恰好和基期一样好。

【复习内容】指数化；价格变化对消费者状况（即福利）的影响。

【参考答案】

注意本题有一个很强的限制条件，即无论消费者的偏好是什么类型，这种价格变动都对他的状况无影响。那么这种价格变动到底是怎样的？

如果你还记得本书第 2 章第 2.7 节第 2 段："由于当所有的价格和收入同乘以一个正数不会改变预算集，最优消费束也不会变动。无需分析消费者的具体选择过程，我们就已得出了一个重要的结论：完全平衡的通货膨胀（即所有商品价格和收入都按相同比率上升），不会改变任何人的预算集，因此也不会改变任何人的最优选择。"那么，你已经知道了答案，即两种商品的价格上升幅度相同时，就会出现题目中要求的现象。

具体地说，假设基期的最优选择为 (x_1^b, x_2^b)，则基期的预算线为 $p_1x_1 + p_2x_2 = p_1x_1^b + p_2x_2^b$。现在假设两种商品的价格上升幅度相同，都为 c 个百分点，则新预算线为 $p_1(1+c)x_1 + p_2(1+c)x_2 = p_1(1+c)x_1^b + p_2(1+c)x_2^b$，将此式中的（1+c）消去可知，它仍是基期的预算线。因此，预算集未变动，从而最优选择也未变动，仍为基期的最优选择 (x_1^b, x_2^b)。

5.本题与上题的基本信息相同，但问题变为：对于什么样的偏好才能做到下列这一点：无论价格怎样变动，他的状况仍恰好和基期一样好。

【复习内容】指数化；价格变化对消费者状况（即福利）的影响。

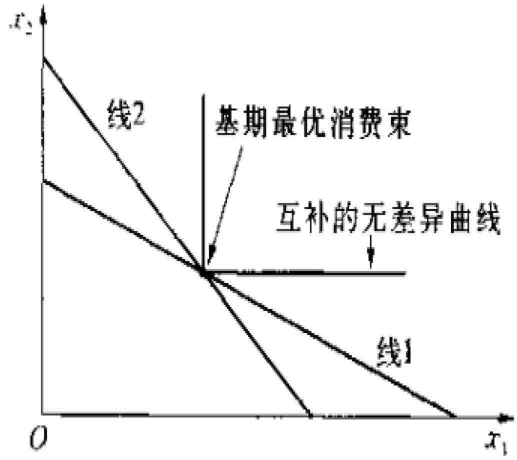
【参考答案】

本题同样有一个很强的限制条件，即无论价格怎样变动。如果读者对第 5 章复习题第 6 题的答案还有印象，则容易猜测出答案可能为他的偏好是完全互补类型的。事实上，这是正确答案。

解题思路和第 5 章复习题第 6 题的思路类似。由于在教材中我们已经知道，对于一般偏好（即良好性状的偏好），根据本章上题可知，如果两种商品价格变动幅度不同，则指数化后消费者的状况会比基期好。因此该消费的偏好应是特殊类型的。

由于在宽泛意义上，两种商品之间既有替代性和互补性。如果两商品价格的替代性越强，则在价格变动时，消费者的选择余地越大。例如，两种商品如为 1:1 完全替代，且在基期价格相同，现在如果商品 1 的价格上升 10%，商品 2 的价格上升 20%。由于指数化要求价格上升时相应增加社会保障金使得该消费者仍能买得起原来的消费束。但此时消费者明显不会选择原来的消费束，他只会购买商品 1。相反，如果两种商品是完全互补的，比如 1:1 互补，

此时无论两商品价格怎样变化，消费者都没有任何选择余地，他必须 1:1 消费这两种商品。



如上图所示。线 1 表示基期的预算线。完全互补类型的无差异曲线为 L 型。L 型曲线的顶点为基期的最优选择点。

现在价格变化时，新的预算线必然会绕着这个基期的最优选择点转动，换句话说这个点必然位于任何一条新的预算线上，为什么？因为指数化！指数化要求保障金根据价格调整而调整，使得消费者恰好还能买得起原来的最优消费束。

由于偏好是完全互补类型的，因此价格变化后最优选择点并未变动。

建议读者在学完第 8 章和第 9 章之后，再回头看看这个题目，那时相信你就能完全理解答案为什么是完全互补类型的了。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

8.斯勒茨基方程（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

8 斯勒茨基方程

经济学家通常关心随经济环境变化，消费者的行为是如何变化的。本章我们研究消费者的选择怎样随着价格变化而变化。自然地，我们认为当某商品的价格上升时，需求会下降。然而，正如在第 6 章所见，我们可以构造一个反例，即当某商品价格下降时，最优需求也会下降。具有这种属性的商品称为**吉芬商品**（Giffen good）。

吉芬商品比较特殊，通常是个理论产物。但在某些情形下，价格的变化会有“反常的”效应，若再考虑一下，就会发现它并非那么不合理。例如，我们通常认为，若人们工资增加，则他们会干得更多。可是如果你的工资从 10 元/小时增加到 1000 元/小时，结果会如何？难道你不想减少工作时间，并将你挣的钱用来做其他事情？如果你的工资是 1,000,000 元/小时，结果会如何？你难道不想干得更少吗？

另外一个例子是，想想当苹果价格上升时，你对苹果的需求将怎样变化？你很有可能减少苹果的消费。但是种植并销售苹果的家庭会减少苹果的消费吗？如果苹果的价格上升，该家庭的收入可能增加，因此他们可能认为他们有能力消费自己的苹果。对于该家庭的消费者来说，苹果价格上升很可能导致他们对苹果的消费增加。

这是怎么一回事？价格的变化怎么可能对需求产生不同的影响？在本章和下一章我们将研究这些效应。

8.1 替代效应

某种商品价格改变时，将产生两种效应：你用一种商品交换另外一种商品的比率发生改变；你收入的总购买力也发生了改变。例如，假设商品 1 变得更便宜，这表明为购买商品 1 你需要放弃商品 2 的数量变少了。商品 1 价格变化，改变了你用商品 2“替代”商品 1 的比率。消费者对于这两种商品的权衡（trade-off）改变了。

与此同时，如果商品 1 变得更便宜，这意味着你的货币收入可以购买更多的商品 1。你货币的购买力已上升；尽管你拥有的钱数是相同的，但它们能够买到的商品数量增加了。

第一部分效应，即需求变化是由两种商品之间的交换比率变化所导致，称为**替代效应**（substitution effect）；第二部分效应，即需求变化是由购买力变化所导致，称为**收入效应**（income effect）。这只是两种效应的粗略定义。准确的定义需要对这两种效应进行更详细地研究。

我们将价格的变动分解为两步：首先，令**相对**价格改变，相应调整货币收入，以保持购买力不变；其次，保持相对价格不变，调整购买力。

用图 8.1 解释这两种效应。此处假设商品 1 价格下降，这意味着预算线绕着纵截距 m/p_2

转动，并且变得更平坦。我们将预算线的这一移动分割为两步：第一步，预算线以原需求束为轴心转动（pivot）；第二步，将转动后的预算线平移(shift)到新需求束之处。

这种“转动-平移”操作让我们方便地将需求变化分解为两部分。第一步为转动，转动后的预算线斜率改变但购买力不变；第二步为平移，平移后的预算线斜率不变但购买力改变。两步分解法只是一个理论上的解释，真实情形没那么复杂。真实情形为消费者看到价格变化后，自然的反应是选择一个新的消费束。但是为了分析消费者选择的变化，假设预算线的变动分为两步，即先转动然后平移，这种方法比较有用。

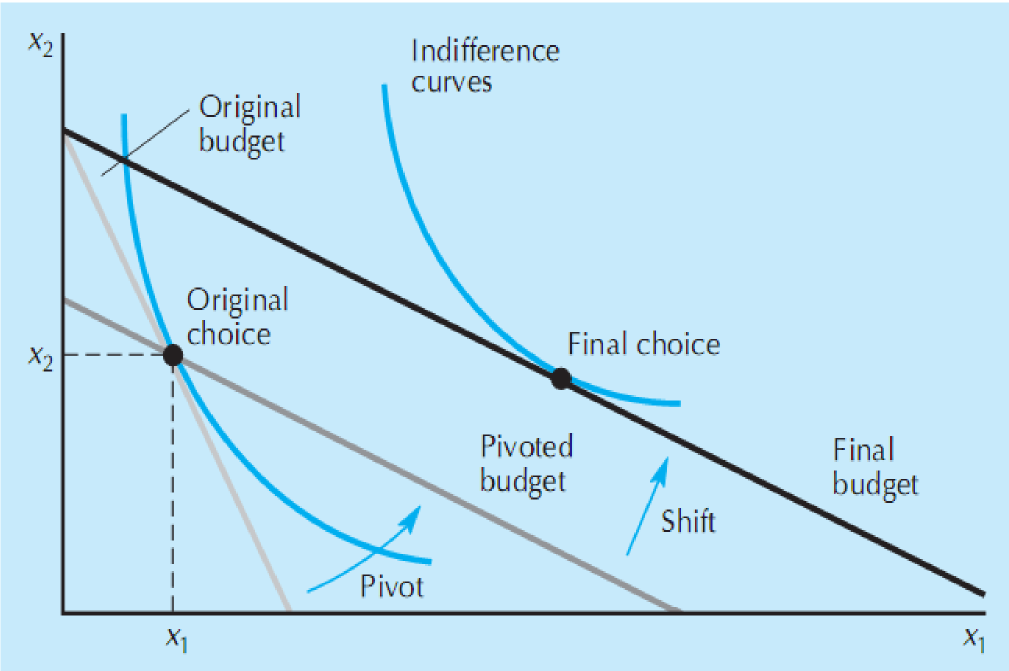


图 8. 1：转动和平移。当商品 1 的**价格**改变而收入不变时，预算线绕纵轴转动。我们将这一转动视为两步变化的总和：首先是转动，即预算线绕原需求束转动；其次是平移，即预算线向外平移到新需求束。

预算线的转动和平移有什么经济学意义？首先考虑转动后的预算线。转动后预算线的斜率和最终预算线的斜率相等，因此相对价格也相等。然而这两条预算线所代表的货币收入数额不同，因为纵截距不同。由于原消费束 (x_1, x_2) 位于转动后的预算线上，而且消费者的购买力保持不变，这意味着该消费束恰好能被买得起。

我们计算一下，为恰好买得起原消费束，货币收入调整数额为多大？令 m 表示在该收入水平下我们恰好能买得起原消费束的货币收入数额；这个收入水平也是转动后的预算线所代表的收入水平。由于在 (p_1, p_2, m) 和 (p_1', p_2, m') 这两种价格收入水平下，都能恰好买得起 (x_1, x_2)，我们有：

$$m' = p_1' x_1 + p_2 x_2$$

 曹乾（东南大学 caoqianseu@163.com）

$$m = p_1x_1 + p_2x_2\text{。}$$

第一式减去第二式得：

$$m' - m = x_1[p_1' - p_1].$$

这个式子是说，为了在新价格下还能恰好买得起原消费束，货币收入改变量应恰好等于商品 1 的原来消费数量乘以价格的改变量。

在上式中，令 $\Delta p_1 = p_1' - p_1$ 以及 $\Delta m = m' - m$，可得：

$$\Delta m = x_1\Delta p_1. \tag{8.1}$$

注意，收入改变和价格改变的运动方向是相同的：如果价格上升，我们就必须增加收入以便还能买得起原来的消费束。

让我们用一些实际的数值。假设消费者原来每周消费 20 块糖，每块糖的价格为 0.50 元。如果糖块的价格上升了 0.10 元，因此 $\Delta p_1 = 0.60 - 0.50 = 0.10$。收入应改变多少才能买得起原来的消费束？

我们可以使用前面的公式。如果消费者的收入增加了 2 元钱，他刚好还能消费得起原来数量的糖（20 块）。用公式来表达：

$$\Delta m = x_1\Delta p_1 = 20 \times 0.10 = 2\text{ 元}.$$

现在我们有了转动后的预算线的表达式：该预算线就是在新价格下，收入改变了 Δm 的预算线。注意，如果商品 1 的价格下降，那么收入的改变量就是负的。当一种商品价格下降，消费者的购买力增加，因此我们必须减少消费者的收入，以使得购买力不变。类似地，当一种商品价格上升，购买力下降，因此为使购买力不变，必须让收入改变量为正。

在转动后的预算线上，尽管 (x_1, x_2) 还能买得起，但通常已不是最优消费束。在图 8.2 中，我们把转动预算线上的最优消费束用 Y 表示。当我们改变价格并且相应调整收入以便还能恰好买得起原消费束时，Y 这个消费束就是最优的。消费束从 X 向 Y 的移动，称为**替代效应**。它表明，当商品价格变化但购买力不变时，消费者如何用一种商品“替代”另外一种商品。

更准确地说，替代效应 Δx_1^s，是指当商品 1 的价格变为 p_1'，同时收入变为 m' 时，商品 1 需求的变动量：

$$\Delta x_1^s = x_1(p_1', m') - x_1(p_1, m).$$

为了计算替代效应，我们必须使用消费者的需求函数，计算在 (p_1', m') 和 (p_1, m) 水平下，相应的最优选择。商品 1 需求的该变量可大可小，这取决于消费者无差异曲线的形状。但在给定需求函数的情形下，只要将相应数值代入即可计算出替代效应。（当然，商品 1 的需求可能还受商品 2 价格的影响，但由于此处假设商品 2 的价格不变，因此在

需求函数中我们省略了商品 2 的价格 p_2，这样做是为了避免函数中变量的符号太多。）

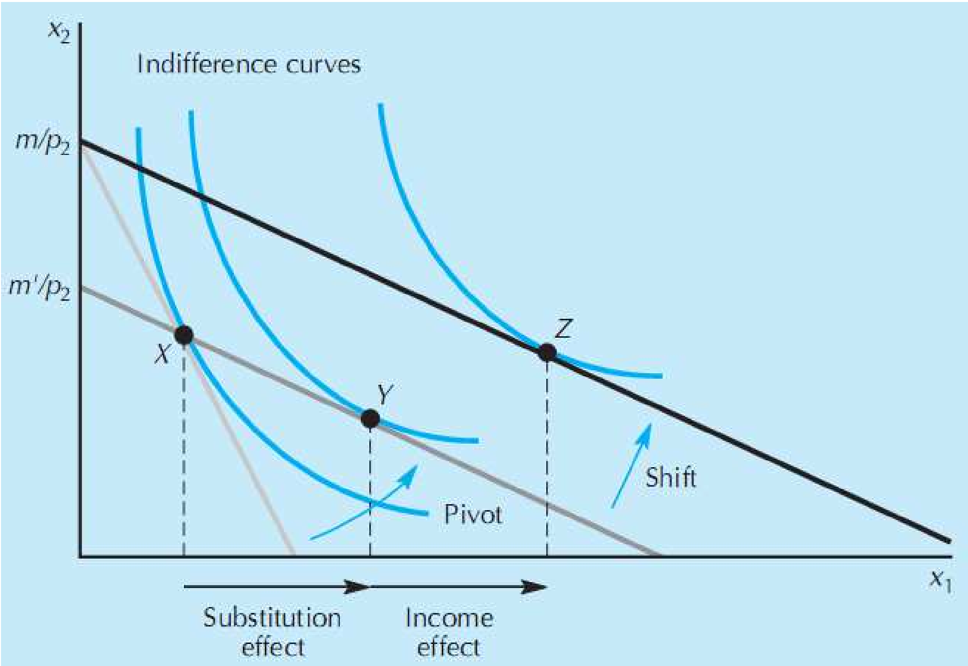


图 8.2：替代效应和收入效应。预算线转动给出了替代效应；预算线平移给出了收入效应。

替代效应有时称为**补偿需求**（compensated demand）的变动量。理由是，商品价格上升时，补偿给消费者足够的资金，以便他能买得起原来的消费束。当然，如果商品价格下降，“补偿”方式是拿走他的部分资金。我们坚持使用“替代”这个术语，以便保持一致性。不过你要知道，经济学文献也广泛使用“补偿”这个术语。

例子：计算替代效应

假设消费者对牛奶的需求函数为

$$x_1 = 10 + \frac{m}{10p_1}.$$

起初他的收入为每周 120 元，牛奶的价格为每单位 3 元。于是每周他对牛奶的需求为 $10 + 120/(10 \times 3) = 14$ 单位。

现在假设牛奶的价格降低为每单位 2 元。在该新价格下他每周的需求为 $10 + 120/(10 \times 2) = 16$ 单位。需求的总变动为 $+2$ 单位每周。

为了计算替代效应，我们必须首先计算，当牛奶的价格为 2 元每单位时，为了让他

曹乾（东南大学 caoqianseu@163.com）

还能恰好买得起原来牛奶的消费量，他的收入应该变动多少。使用公式（8.1）：

$$\Delta m = x_1 \Delta p_1 = 14 \times (2-3) = -14\text{元}.$$

因此能让购买力不变的收入水平为 $m' = m + \Delta m$ =120－14=106 元。在该收入水平下并且牛奶的新价格为 2 元时，消费者对牛奶的需求量为多少？将相应数值代入需求函数可得

$$x_1(p_1', m') = x_1(2, 106) = 10 + \frac{106}{10 \times 2} = 15.3\text{单位。}$$

因此替代效应为

$$\Delta x_1^s = x_1(2, 106) - x_1(3, 120) = 15.3 - 14 = 1.3\text{单位。}$$

8.2 收入效应

我们现在转向价格调整的第二步，即预算线的平移。它的经济学意义也容易解释。我们知道，预算线的平行移动发生在收入改变而相对价格不变时。因此，价格调整的第二步称为**收入效应**。第二步，只要将消费者的收入从 m' 改变为 m，但维持价格 (p_1', p_2) 不变即可。在图 8.2 中，这一变动将使我们从点 (y_1, y_2) 移动到 (z_1, z_2)。自然地可将这一移动称为收入效应，因为我们只是改变了收入但维持价格 (p_1', p_2) 不变。

更准确地说，收入效应 Δx_1^n，是指当收入从从 m' 改变为 m，但维持商品 1 的价格 p' 不变时，消费者对商品 1 的需求量的变动：

$$\Delta x_1^n = x_1(p_1', m) - x_1(p_1', m')\text{。}$$

在 6.1 节我们已研究过收入效应。那时我们已知道收入效应的方向有两个：它可能增加商品 1 的消费，也可能减少商品 1 的消费，这取决于商品 1 是正常品还是劣等品。

如果一种商品价格下降，我们需要减少消费者的收入以便维持购买力不变。如果该商品为正常品，那么收入的下降会导致需求的下降。如果该商品是劣等品，收入的下降会导致需求增加[一]。

例子：计算收入效应

在本章前面的例子中，我们已知

$$x_1(p_1', m) = x_1(2,\ 120) = 16$$

[一] 这一段当属作者犯的低级错误。因为按照这一段的表达，这不是收入效应而是替代效应了。正确的表达如下：如果一种商品价格下降，相当于消费者收入增加了，因此如果该商品为正常品，那么收入的增加会导致需求的增加。如果该商品是劣等品，收入的增加会导致需求减少。译者注。

曹乾（东南大学 caoqianseu@163.com）

$$x_1(p_1',m')=x_1(2,\ 106)=15.3\text{。}$$

因此，该问题的收入效应为

$$\Delta x_1^n=x_1(2,120)-x_1(2,106)=16-15.3=0.7$$

因为牛奶对于消费者来说是正常商品，收入增加时牛奶的需求也增加了。

8.3 替代效应的符号

在上面我们已知道收入效应可正可负，取决于商品是正常品还是劣等品。替代效应的符号是怎样的？如果一种商品的价格下降，如图 8.2 所示，那么由替代效应引起的需求变动**必然**非负。也就是说，如果 $p_1>p_1'$，就**必然**有 $x_1(p_1',m')\geq x_1(p_1,m)$，因此 $\Delta x_1^s\geq 0$。

上述结论的证明如下。考虑图 8.2 中位于转动后的预算线上的这样的点，在这些点上商品 1 的消费数量小于消费束 X。在原价格水平 (p_1,p_2) 时，这些消费束都能买得起但并未被购买。相反，实际购买的是消费束 X。如果消费者总是从他能买得起的消费束中选择最优的，那么与转动后预算线上并且位于原来预算集内部的所有其他的消费束相比，X 必然是最受偏好的。

这表明位于转动后的预算线上的最优选择，必然不是那些位于原预算线之下的消费束，这个最优选择要么是 X，要么是位于 X 右侧的某个点。但这意味在新的最优选择中，商品 1 的数量至少和原消费束一样多，这正是我们想要证明的。在图 8.2 表示的情形中，转动后预算线上的最优选择是消费束 Y， Y 中商品 1 的数量当然比原消费束 X 中的商品 1 的数量多。

替代效应的运动方向总是与价格运动方向相反。我们说**替代效应为负**，因为由替代效应引起的需求变动和价格变动方向相反：如果价格上升，商品的需求因替代效应而下降。

8.4 需求的总变动

需求的总变动 Δx_1，是指由于价格变动但收入不变引起的需求变动：

$$\Delta x_1=x_1(p_1',m)-x_1(p_1,m)\text{。}$$

在前文我们已知道如何将总需求的变动分解为两个部分：替代效应和收入效应。用前文使用的符号表示：

$$\Delta x_1=\Delta x_1^s+\Delta x_1^n$$

曹乾（东南大学 caoqianseu@163.com）

$$\Delta x_1 = x_1(p_1', m) - x_1(p_1, m) = [x_1(p_1', m') - x_1(p_1, m)] \\ + [x_1(p_1', m) - x_1(p_1', m')]$$。

这个式子是说需求的总变动等于替代效应加上收入效应。这个式子叫做**斯勒茨基恒等式**（Slutsky identity）[一]。注意，这是个恒等式：它对 p_1, p_1', m, m' 的所有取值都成立。恒等式右端第一项和第四项抵消，因此右端**恒等于**左端。

斯勒茨基恒等式的内涵是，它不仅仅是代数恒等式，后者在数学上是平凡的。该恒等式的内涵来自它对等式右端两项的解释能力：替代效应和收入效应。特别地，我们可以运用收入效应和替代效应的符号来确定总效应的符号。

尽管替代效应总为负，即与价格变动方向相反，但收入效应可正可负，因此总效应可正可负。然而，对于正常品来说，替代效应和收入效应的运行方向相同。如果价格上升，替代效应使需求减少；价格上升相当于收入下降，对于正常品来说，收入效应使需求下降。两种效应互相加强。正常品的价格上涨而导致的需求变化，用我们前面的符号可表示如下：

$$\begin{array}{ccccc} \Delta x_1 & = & \Delta x_1^s & + & \Delta x_1^n \\ (-) & & (-) & & (-) \end{array}$$

（式子中每一项下面的负号表示该项的符号为负）。

注意收入效应的符号。因为我们此处考虑的情形是价格上涨，这意味着购买力下降——对于正常品来说，这将导致需求下降。

另一方面，对于劣等品来说，收入效应可能超过替代效应，因此价格上升引起的需求总变动可能为正。这个情形用符号表示：

$$\begin{array}{ccccc} \Delta x_1 & = & \Delta x_1^s & + & \Delta x_1^n \\ (?) & & (-) & & (+) \end{array}$$。

如果等式右端第二项即收入效应很大，需求的总变动将为正。这表示价格上升导致需求上升。前文描述的吉芬商品反例就是这样的：价格上升大幅减少了消费者的购买力，他不得不增加劣等品的消费。

斯勒茨基恒等式表明，这种反常的效应只可能发生在劣等品的身上：如果商品为正常品，那么收入效应和替代效应就会互相加强，因此需求的总变动总是呈现“正常的”方向。

因此，吉芬商品必为劣等品，但劣等品未必是吉芬商品。因为对于吉芬商品来说，收入效应不仅有“错误的”方向符号，而且数值还要足够大，以便超过替代效应“正确的”方向符号。这就是为什么吉芬商品在现实生活中很难看到的原因：它们不仅必为劣等品，而且必须**非常**劣等。

[一] 以 Eugen Slutsky (1880-1948)的名字命名，他是位研究需求理论的前俄国经济学家。

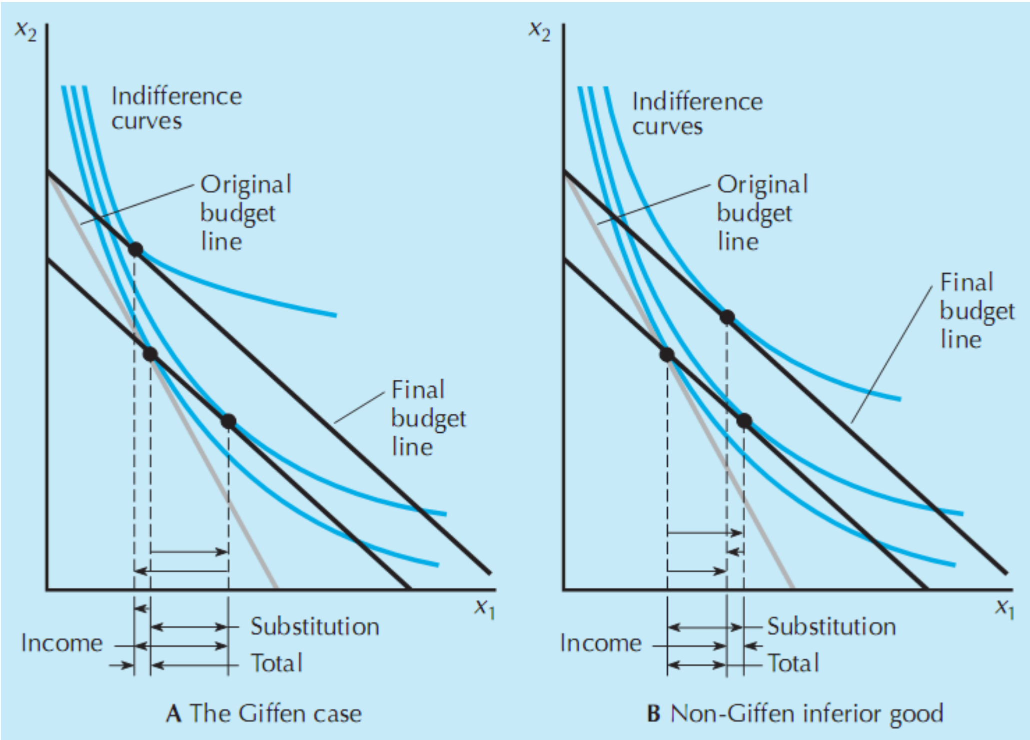


图 8.3：劣等品。图 A 表明，商品 1 非常劣等，因此为吉芬品；图 B 表明，商品 1 是劣等品，但其效应不够强，因此不是吉芬品。

劣等品和吉芬品的情形见图 8.3。此处我们运用通常的转动—平移方法，来确定替代效应和收入效应。在两种情形下，商品 1 都是劣等品，收入效应因此为负。在图 8.3A 中，收入效应足够大，超过了替代效应，因此为吉芬品。在图 8.3B，收入效应较小，因此商品 1 对于本身的价格变化，有通常的反应。

8.5 变化比率

我们已看到，收入和替代效应可使用转动—平移方法进行图形描述，也可以用斯勒茨基恒等式进行代数描述：

$$\Delta x_1 = \Delta x_1^s + \Delta x_1^n \text{。}$$

这个恒等式是说需求的总变动等于替代效应加上收入效应。此处，斯勒茨基恒等式是用绝对变化表述的，但更常见的表述是用变化率。

当用变化率表述斯勒茨基恒等式时，比较方便的做法是将 Δx_1^m 定义为收入效用的**相反数**（negative）：

$$\Delta x_1^m = x_1(p_1', m') - x_1(p_1', m) = -\Delta x_1^n \text{。}$$

使用这个定义，斯勒茨基恒等式变为

$$\Delta x_1 = \Delta x_1^s - \Delta x_1^m.$$

等式两端同除以 Δp_1，可得：

$$\frac{\Delta x_1}{\Delta p_1} = \frac{\Delta x_1^s}{\Delta p_1} - \frac{\Delta x_1^m}{\Delta p_1}. \tag{8.2}$$

等式右端第一项是，当价格变化并且相应调整收入以使得恰能买得起原消费束时，需求的变动比率，这就是替代效应。我们主要分析第二项。因为该项分子包收入变化因素，如果分母能改为收入变化会更好。

我们知道收入变化 Δm 和价格变化 Δp_1 之间存在关系式：

$$\Delta m = x_1 \Delta p_1.$$

由上式解出 Δp_1：

$$\Delta p_1 = \frac{\Delta m}{x_1}$$

将其带入（8.2）得到我们的最终表达式：

$$\frac{\Delta x_1}{\Delta p_1} = \frac{\Delta x_1^s}{\Delta p_1} - \frac{\Delta x_1^m}{\Delta m_1} x_1.$$

这就是以变化率表示的斯勒茨基恒等式。我们可将恒等式中的每一项解释如下：

$$\text{左端}\ \frac{\Delta x_1}{\Delta p_1} = \frac{x_1(p_1', m) - x_1(p_1, m)}{\Delta p_1}$$

是价格变化而收入不变时，需求的变化率。

$$\text{右端第一项}\ \frac{\Delta x_1^s}{\Delta p_1} = \frac{x_1(p_1', m') - x_1(p_1, m)}{\Delta p_1}$$

是价格变化并且相应调整收入以使得恰好能买得起原来消费束时，需求的变化率。这就是替代效应。

$$\text{右端第二项}\ \frac{\Delta x_1^m}{\Delta m_1} x_1 = \frac{x_1(p_1', m') - x_1(p_1', m)}{m' - m} x_1 \tag{8.3}$$

是维持价格不变（注意此处的价格是新价格），但调整收入后的需求变动率。这就是收入效应。

收入效应本身是两部分的乘积：收入变动引起的需求变动；乘以原来的需求水平。当价格变动 Δp_1 时，由收入效应引起的需求变动为

$$\Delta x_1^m = \frac{x_1(p_1', m') - x_1(p_1', m)}{\Delta m} x_1 \Delta p_1.$$

但上式右端最后一项即 $x_1\Delta p$ ，是为了恰好能买得起原消费束，收入的调整数额。也就是说， $x_1\Delta p = \Delta m$ 。因此，由收入效应引起的需求变动可以简化为

$$\Delta x_1^m = \frac{x_1(p_1', m') - x_1(p_1', m)}{\Delta m}\Delta m$$

这个式子正是我们前面已得到的。

8.6 需求法则

在第 5 章我们曾提及过对消费者理论下列事实的关注，即该理似乎没什么特别之处，因为价格上升时，需求可上升也可下降；收入增加时，需求也可上升或下降。如果一个理论不能以**某种**方式控制所观测的行为，它就不足以称为真正的理论。一个能解释一切行为的模型无实际意义。

然而，我们知道消费者理论的确有意义，因为我们已看到理性消费者的选择必须满足显示偏好强公理。而且，我们已看到价格变动可以分解为两个部分的变化：替代效应，它的符号必然为负，即与价格变动方向相反；收入效应，它的符号取决于商品是正常品还是劣等品。

尽管消费者理论没有限制当价格变动时需求怎样变动，也没有限制当收入变动时需求怎样变动，但它的确限定了这两类变化之间的关系。特别地，我们有下列法则。

需求法则。**如果收入增加时某种商品的需求上升，那么价格上升时，该商品的需求必然下降**。

这个法则可由斯勒茨基等式直接推导出：如果收入增加时，某种商品的需求上升，那么它必为正常品。对于正常品，替代效应和收入效应互相加强，因此价格上升必然减少需求。

8.7 收入效应和替代效应的例子

现在我们考虑某些特别类型偏好下价格变动的例子，并将需求变动分解为收入效应和替代效应。

完全互补的偏好

我们首先分析完全互补的情形。斯勒茨基分解由图 8.4 说明。当预算线绕原消费束旋转时，在新预算线上的最优选择就是原消费束，这意味着替代效应为零。需求的变动完全由收入效应引起。

曹乾（东南大学 caoqianseu@163.com）

完全替代的偏好

完全替代情形的结果如何？图 8.5 说明了该情形。预算线绕原消费束转动时，消费束从纵轴跳跃到横轴。这里没有涉及到预算线的平移！需求变动完全由替代效应引起。

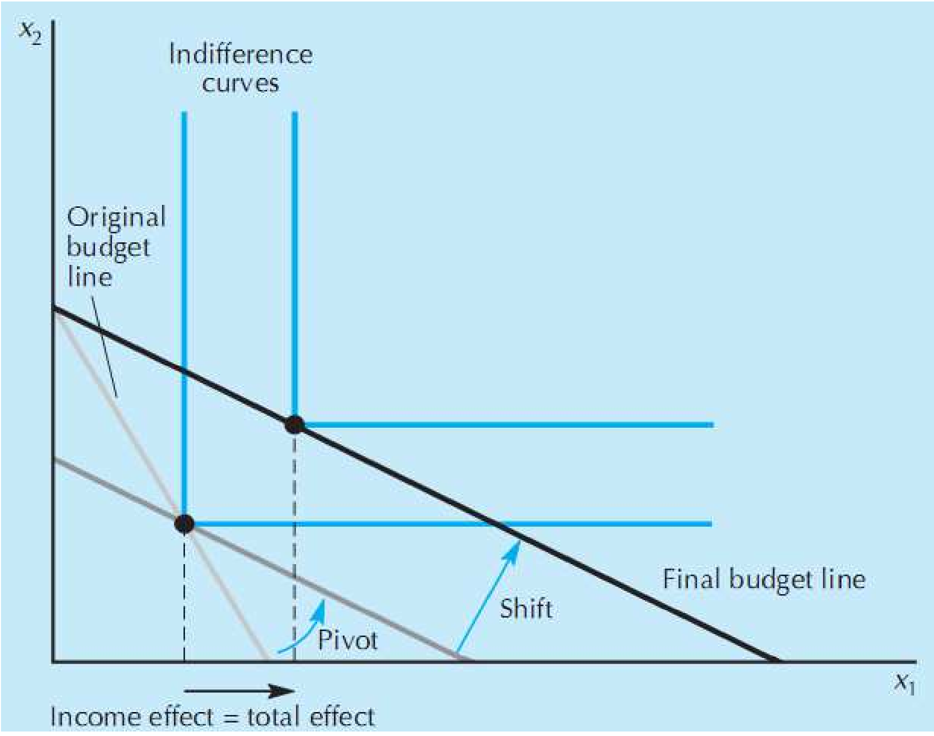


图 8.4：完全互补。完全互补时的斯勒茨基分解。

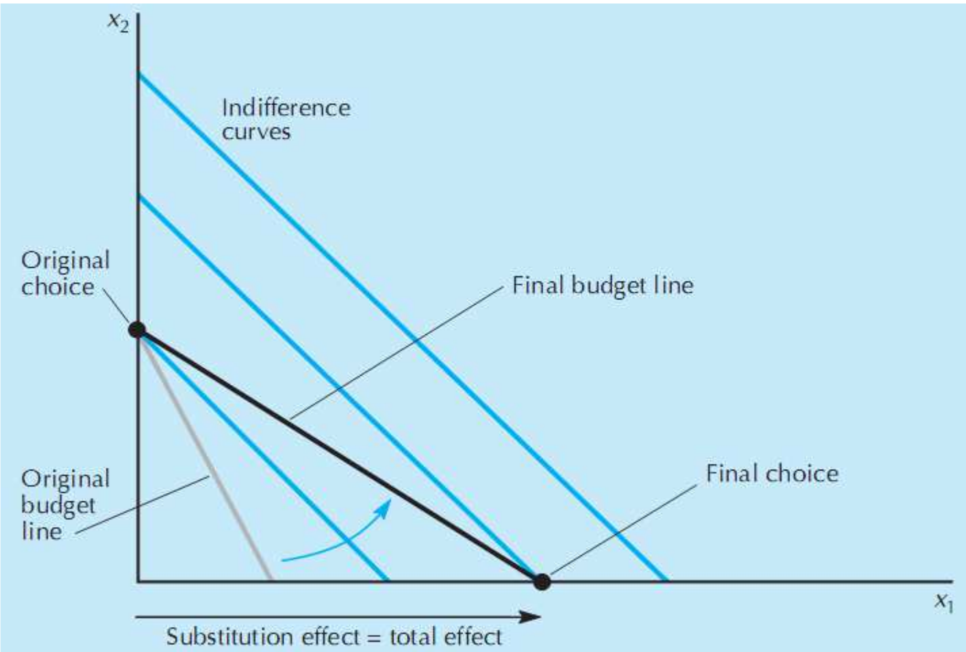


图 8.5：完全替代。完全替代情形时的斯勒茨基分解。

拟线性偏好

第三个例子是拟线性偏好。请看图 8.6。拟线性偏好的情形有些特殊。我们已知道，当偏好为拟线性时，收入变化对商品 1 的需求没有任何影响。这表明商品 1 的全部需求变动都是由替代效应引起，收入效应为零。

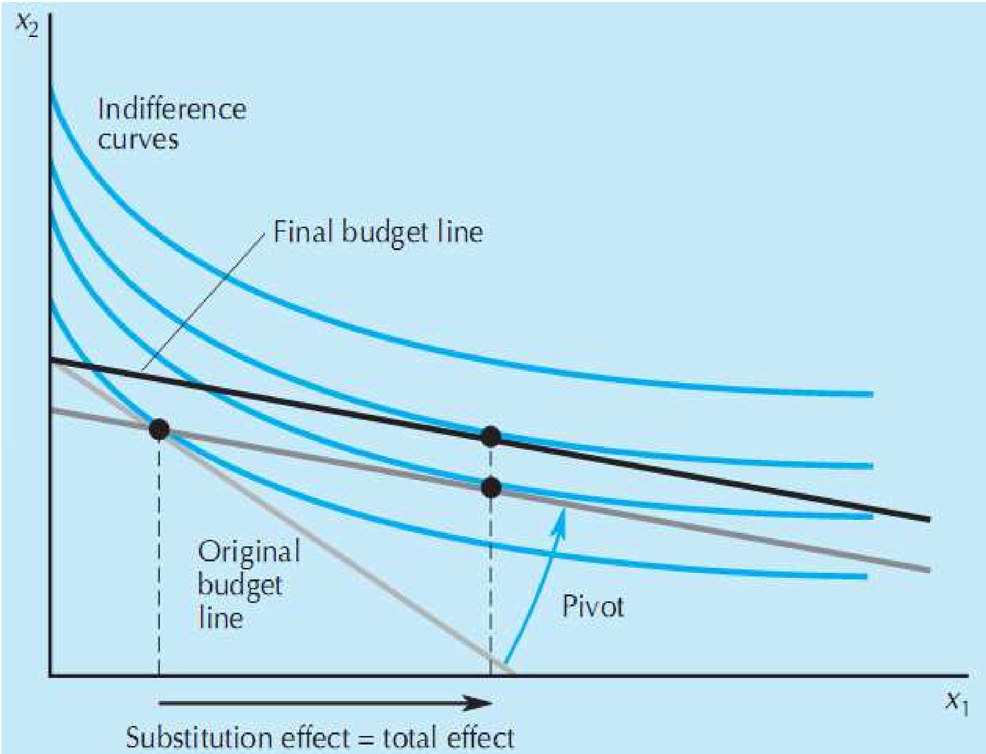


图 8.6：拟线性偏好。拟线性偏好情形下，全部需求的变动都是由替代效应引起。

例子：退税

1974 年，石油输出国组织（OPEC）对美国实施石油禁运。OPEC 有能力做到连续几个星期不向美国港口运输石油。美国深受石油禁运之害，这让议会和总统非常苦恼。当时人们提出了很多方案，旨在减少美国对国外石油的依赖。

其中一种方案是提高汽油税。消费者使用汽油的成本上升，会减少他们对汽油的消费，这又会减少对国外石油的需求。但直接提高汽油税会伤害消费者，伤害了消费者等同于打击了税收来源，政府税收可能因此减少。所以该方案在政治上不可行。有人建议政府应将征缴的汽油税返还给消费者，返还方式可以直接返还，也可通过减少消费者的其他税收形式间接返还。

批评家认为税收返还将使干预消费者需求的愿望落空，因为消费者可用返还的资金购买更多的汽油。经济学家会怎样评价退税方案？

为简单起见，假设汽油税完全转嫁到汽油消费者身上，因此汽油价格涨幅恰好等于税额。（一般来说，只能转嫁税收的一部分，但此处我们不考虑这种复杂的情形）。假设税收使汽油价格从 p 上升到 $p' = p + t$，单个典型消费者作出的反应是需求从 x 减少至 x'。征税后，典型消费者每单位汽油多支付了 t 元，他消费的汽油量为 x'，因此该人贡献的税收为

$$R = tx' = (p' - p)x'.$$

注意，征税收入取决于消费者最终消费的汽油量，而不是最初消费量 x。

如果令 y 表示用于其他商品的支出，令其价格等于 1，则原来的预算约束为

$$px + y = m, \tag{8.4}$$

退税后的预算约束为[一]

$$(p+t)x' + y' = m + \hat{m} \quad \text{或}$$

$$px' + tx' + y' = m + \hat{m}$$

在上式中，典型消费者选择等式左端的变量，即选择每种商品的消费数量 x' 和 y。但等式右端即他的收入 m 和政府对他的退税数额 $\hat{m}$，都是既定不变的。政府对他的退税数额 $\hat{m}$ 不变，因为政府根据**所有**消费者对汽油的消费情况测算对每个消费者的退数额，而不是根据个别消费者的汽油消费情况。这也就是说，在上式中 tx' 通常不等于 $\hat{m}$。

但是，如果某消费者对汽油的消费量处于平均水平，tx' 可能等于 $\hat{m}$。在这种情况下，上式变为

$$(p+t)x' + y' = m + tx'. \tag{8.5}$$

如果我们消去（8.5）式左右两端的 tx'，即可得到：

$$px' + y' = m.$$

因此，(x', y') 这个消费束在原预算约束下可以买得起，但该消费者选择的是 (x, y)。这表明，(x, y) 比 (x', y') 更受偏好，因此退税方案使消费者的状况变坏了。这也许正是这个方案没有付诸实施的原因！

图 8.7 说明了退税方案的消费者均衡问题。税收使商品 1 变得昂贵，退税增加了货币收入。原来的消费束已购买不起，该消费者的状况必然变差。退税方案中消费者的选择为：减少汽油消费，增加“所有其他商品”的消费。

[一] 紧跟着此脚注的两个预算线表达式都是译者所加，目的是更清楚地说明作者的思想。原书中作者在此处直接给出了（8.5）式。

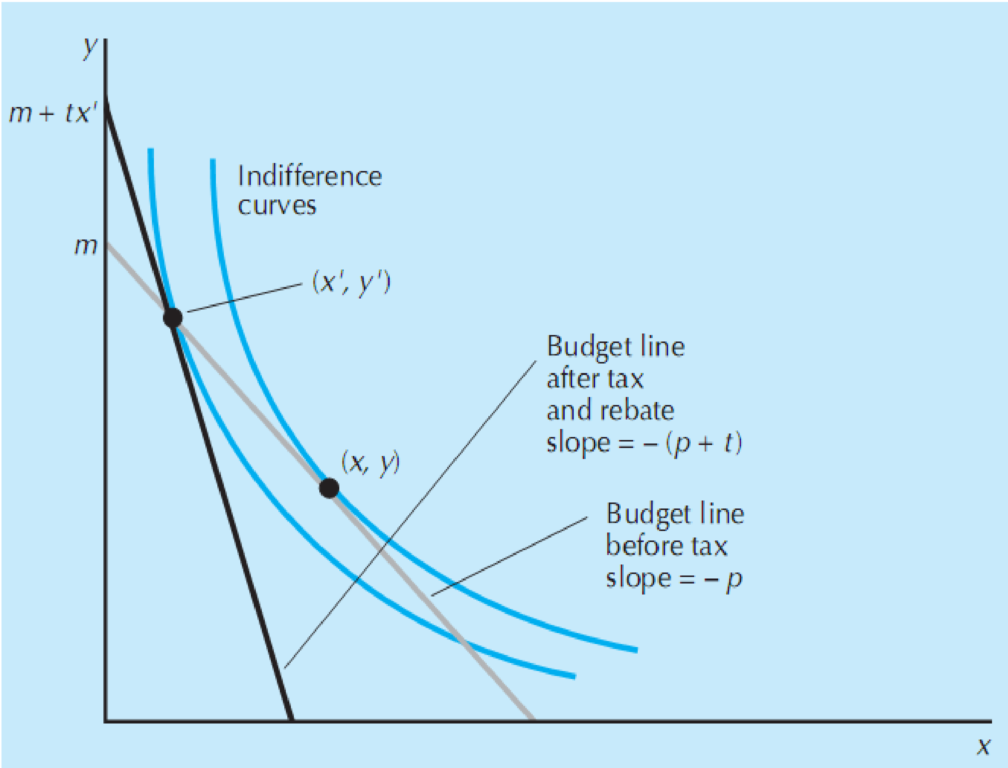


图 8.7：**退税**。向某消费者征税然后再退税，这使得消费者的状况变差。

我们对汽油消费量有何看法？典型消费者能买得起原来数量的汽油，但由于税收使汽油变得昂贵。一般来说，消费者会减少汽油的消费。

例子：自愿实时定价

电力生产会遭遇极限容量难题：在达到极限容量之前，电力生产相对便宜。根据定义，在极限容量点，不可能生产得更多。容量建设及其昂贵，因此，从经济学的观点看，设法减少高峰需求期的用电量更好些。

气候温暖的州例如乔治亚州，在高峰需求期，大约 30%的用电量是由空调引起。而且，提前一天预报气温比也不难做到，因此消费者有时间调整电力需求，比如调高空调温度，穿比较薄的衣服等等。问题在于要设计出一套定价系统，使得那些能降低用电量的用户有减少用电的动机。

一种方法是使用实时定价（Real Time Pricing，RTP）方案。在该方案中，大型企业

用户配备了特别的仪表，电价根据发电厂发送的信号实时变化。电力需求接近极限容量时，发电厂提高电价以鼓励用户降低用电量。电价由电力的总需求函数确定。

乔治亚电力公司宣称，它当时经营的实时定价系统为世界最大的定价系统。1999 年，该公司成功将高电价时段的电力需求量减少了 750 兆瓦，因为实时定价系统使某些大客户的用电量减少了 60%。

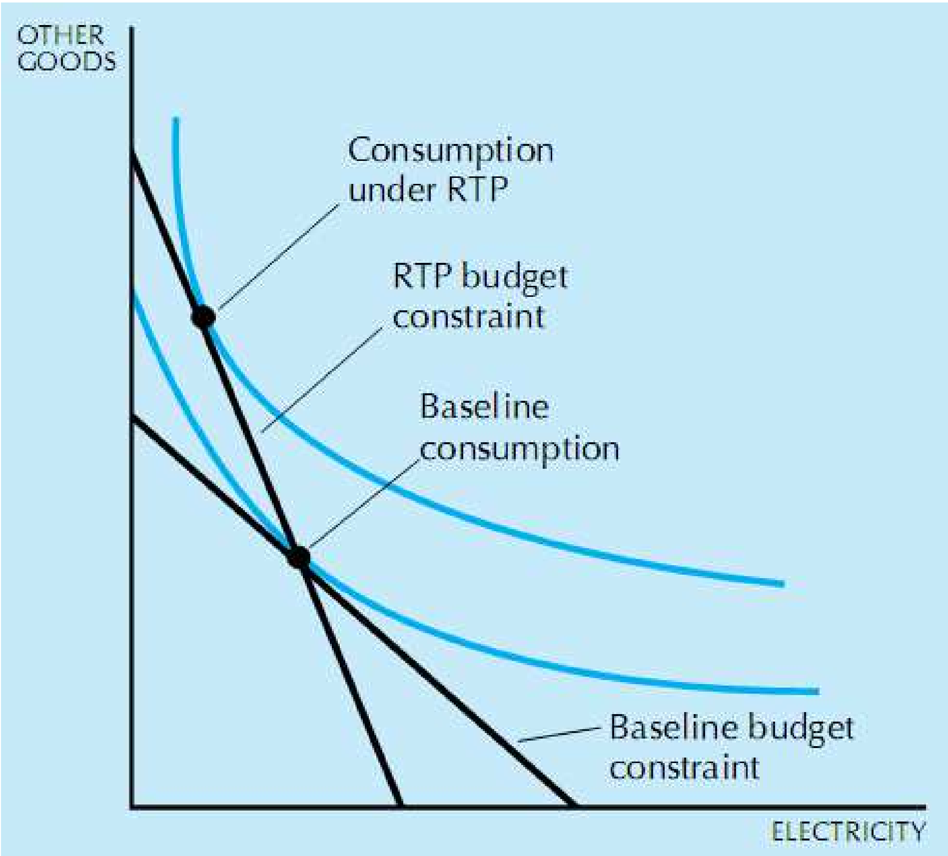


图 8.8：自愿实时定价。当实时电价上升时，如果客户的用电量超过基准数量，对超过部分要支付高价格，但如果客户能将用电量降低，则能享受价格折扣。这种定价方法使预算线绕基准消费束转动，使客户的状况更好。

在基本实时定价模型的基础上，乔治亚电力公司又研发了几种实时定价方法。其中一种定价方法是，该公司给每个客户制定一个基准数量，即正常用电量。在电力供应短缺时，对超过基准数量部分的用电量索要高价。当然，如果客户能将用电量减至基准数量以下，也可享受价格**折扣**。

可用图 8.8 分析这种定价方法对客户预算线的影响。纵轴表示“花费在电力以外的其他商品上的资金”，横轴表示“用电量”。在正常时段，预算约束由基准电力价格决定，客户在这种预算约束下选择用电量，以使自身的效用最大化。此情形下的用电量就是他们的基准消费量（baseline consumption）。

气温上升时，实时价格上升使电价更贵。但电价上升对部分客户来说是件好事，因为这部分客户可以减少用电量，电力公司按照客户用电量减少的千瓦数给与他们价格折

扣。如果用电量等于基准数量，电价维持在实时价格水平上，没有折扣。

不难看出这种定价方法使得预算线绕着基准消费束旋转。因此我们可以断言：用电量会下降，并且客户在实时定价情形下的状况至少和在基准价格情形下的状况一样好。说真的，这种定价方法很受欢迎， 1600 多个大客户自愿使用这种方法。

8.8 另外一种替代效应

当价格改变但保持消费者购买力不变（仍能买得起原消费束）的情形下，需求会改变。经济学家将上述现象称为替代效应。但这只是替代效应的其中一种定义。还有一种定义也很有用。

我们在前面已学习过的替代效应的这种定义，称为**斯勒茨基替代效应**（Slutsky substitution effect）。本节将要介绍的定义称为**希克斯替代效应**（Hicks substitution effect）[一]。

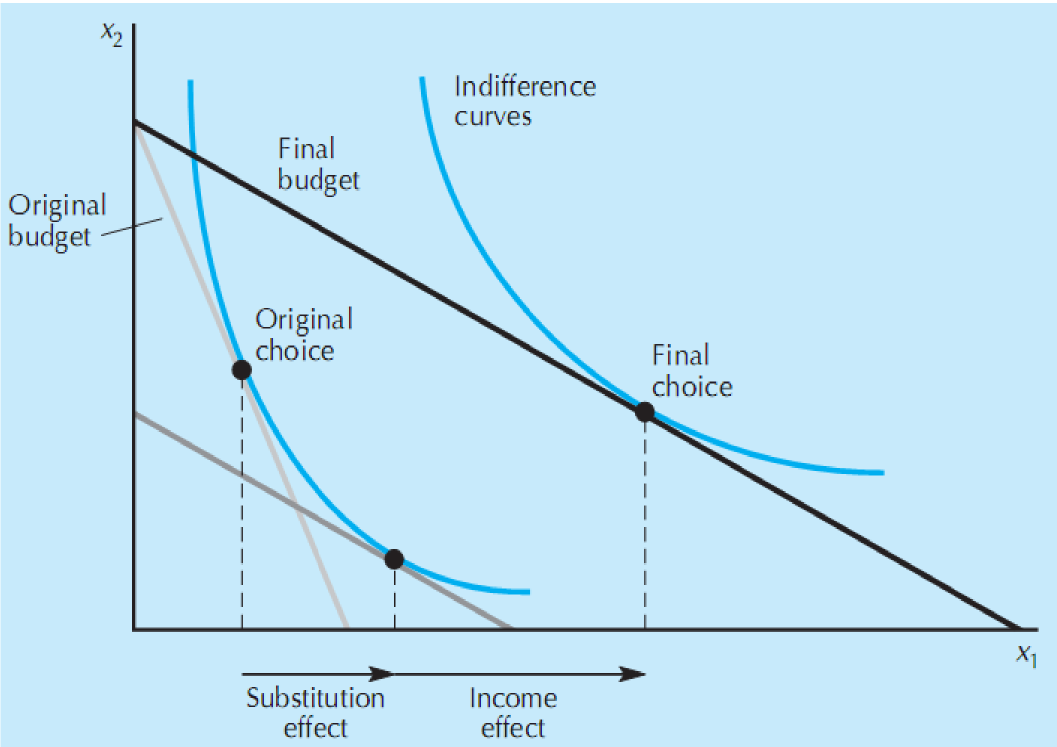


图 8.9：希克斯替代效应。此处我们将预算线绕通过原消费束的那条无差异曲线转动，而不是绕原消费束转动。

假设我们不再让预算线绕着原消费束转动，现在我们将预算线绕着通过原消费束的那条无差异曲线旋转，也就是说使旋转后的预算线不仅要与上述无差异曲线相切，还要

[一] 该名字源自约翰.希克斯男爵，英国经济学家，诺贝尔经济学奖获得者。

与最终的预算线平行。如图 8.9 所示。这样，我们就得到了一条新预算线，这条预算线与最终预算线有相同的相对价格（两条预算线平行），但与最终预算线代表不同的收入水平。消费者在这条预算线下的购买力不足以购买原消费束，但足以买到和原消费束无差异的其他消费束（因为这条预算线与通过原消费束的无差异曲线相切）。

因此，**希克斯替代效应保持效用不变，而斯勒茨基替代效应保持购买力不变**。斯勒茨基替代效应给消费者一定的资金，使他恰好能买得起原消费束，而希克斯替代效应给消费者一定的资金，使他恰好能回到通过原消费束的那条无差异曲线上。尽管两种替代效应的定义不同，但希克斯替代效应也必然为负，即替代效应与价格变动方向相反，这一点和斯勒茨基替代效应是相同的。

证明方法也是使用显示偏好理论。令 (x_1, x_2) 表示价格为 (p_1, p_2) 时的需求束，令 (y_1, y_2) 表示价格为 (q_1, q_2) 时的需求束。假设消费者的收入水平恰好使得 (x_1, x_2) 和 (y_1, y_2) 无差异。由于消费者对 (x_1, x_2) 和 (y_1, y_2) 无差异，因此它们互相不能显示偏好。

根据显示偏好的定义，这意味着下述两个不等式都不成立：

$$p_1x_1 + p_2x_2 > p_1y_1 + p_2y_2$$

$$q_1y_1 + q_2y_2 > q_1x_1 + q_2x_2\,.$$

由此可知下面两个不等式是成立的：

$$p_1x_1 + p_2x_2 \le p_1y_1 + p_2y_2$$

$$q_1y_1 + q_2y_2 \le q_1x_1 + q_2x_2\,.$$

将这两个不等式相加并整理可得

$$(q_1 - p_1)(y_1 - x_1) + (q_2 - p_2)(y_2 - x_2) \le 0\,.$$

这个式子表明，当价格改变后，如果相应调整收入使消费者还呆在原来的无差异曲线上，需求如何变动。由于我们关注只有一种商品价格变动的情形，比如商品 1 的价格变动。因此 $q_2 = p_2$，上式因此变为

$$(q_1 - p_1)(y_1 - x_1) \le 0\,.$$

这个式子表明需求量的变动和价格变动的符号相反，这正是我们想证明的。

需求的总变动仍然等于替代效应与收入效应之和，但此处显然应该使用希克斯替代效应。由于希克斯替代效应也为负，斯勒茨基方程的形式和前面的一样，解释也完全相同。这两种替代效应都可行，具体用哪一种取决于你研究什么样的问题。可以证明，如果价格变动很小，这两种替代效应完全相同。

曹乾（东南大学 caoqianseu@163.com）

8.9 补偿需求曲线

我们已经分析了下列三种情形下需求量是如何随价格的变动而变动的：保持价格不变（标准情形）;保持购买力不变（斯勒茨基替代效应）；保持效用不变（希克斯替代效应）。对于这三种情形中的任何一种，我们都可以画出价格和需求量之间的关系。由此可以得出三条不同的需求曲线：标准需求曲线；斯勒茨基需求曲线和希克斯需求曲线。

本章的分析表明斯勒茨基需求曲线和希克斯需求曲线总是向下倾斜的。而且，普通需求曲线（即前面提及的标准需求曲线）也是向下倾斜的，但它有个前提即这种商品是正常商品，因为我们已知道劣等商品（包括吉芬商品）的需求曲线在理论上可能向上倾斜。

希克斯需求曲线（即保持效用不变得到的需求曲线），有时称为**补偿需求曲线**（compensated demand curve）。这么叫其实挺自然，因为你可以这样想：当价格上升时，你补偿他一些收入，使他的效用不变；当价格下降时，你从他手里拿走一部分收入（相当于补偿为负），使他的效用不变。由于价格变动后消费者得到了“补偿”，因此希克斯需求曲线上效用处处相等。普通需求曲线不是效用处处相等，在普通需求情形下，价格较高时消费者的状况比价格较低时差，这是由于他的收入既定不变。

补偿需求曲线在高级经济学课程中的用处很大，尤其是当我们研究收益—成本分析时它很有用。在这类分析中，由于政策变动会对消费者造成影响，你自然会问应该对消费者补偿多少这样的问题。我们可用补偿数额的大小衡量政策变动的成本。然而，补偿需求曲线的计算要求更多的数学工具，这超出了本书的范围。

附录

我们用微积分推导斯勒茨基方程。考虑斯勒茨基替代效应，在该效应中，我们调整消费者的收入使其恰好能买得起原消费束$(\bar{x}_1,\bar{x}_2)$。若价格为(p_1,p_2)，则收入进行上述调整后消费者的实际选择取决于(p_1,p_2)和$(\bar{x}_1,\bar{x}_2)$。将这种关系称为商品 1 的**斯勒茨基需求函数**，记为$x_1^s(p_1,p_2,\bar{x}_1,\bar{x}_2)$。

假设原需求束$(\bar{x}_1,\bar{x}_2)$是在价格(p_1,p_2)和收入$\overline{m}$下的消费束。斯勒茨基需求函数告诉我们消费者面对不同的价格(p_1,p_2)和收入$p_1\bar{x}_1+p_2\bar{x}_2$时，他的需求变为多少。因此，在$(p_1,p_2,\bar{x}_1,\bar{x}_2)$条件下的斯勒茨基需求函数就是价格为$(p_1,p_2)$且收入为$p_1\bar{x}_1+p_2\bar{x}_2$时的普通需求函数。即，

$$x_1^s(p_1,p_2,\bar{x}_1,\bar{x}_2)\equiv x_1(p_1,p_2,p_1\bar{x}_1+p_2\bar{x}_2).$$

这个式子是说价格为(p_1,p_2)时，消费者对商品 1 的斯勒茨基需求数量，恒等于将其收入调整使其恰好能买得起原消费束$(\bar{x}_1,\bar{x}_2)$时，他对商品 1 的（普通）需求数量。这正

是斯勒茨基需求函数的定义。

将上述恒等式对 p_1 求导，可得

$$\frac{\partial x_1^s(p_1,p_2,\overline{x}_1,\overline{x}_2)}{\partial p_1}=\frac{\partial x_1(p_1,p_2,\overline{m})}{\partial p_1}+\frac{\partial x_1(p_1,p_2,\overline{m})}{\partial m}\overline{x}_1.$$

整理可得

$$\frac{\partial x_1(p_1,p_2,\overline{m})}{\partial p_1}=\frac{\partial x_1^s(p_1,p_2,\overline{x}_1,\overline{x}_2)}{\partial p_1}-\frac{\partial x_1(p_1,p_2,\overline{m})}{\partial m}\overline{x}_1.$$

注意在求导时我们使用了链式法则。

这是导数形式的斯勒茨基方程。它表明价格变动的总效应由替代效应（调整收入使 $(\overline{x}_1,\overline{x}_2)$ 恰好能买得起）和收入效应构成。从教材中我们知道替代效应恒为负，但收入效应的符号取决于商品是正常商品还是劣等商品。你已经看到，这个方程正是我们在教材中已经推导出的斯勒茨基方程，唯一不同的是此处我们用导数符号代替了那里的差商符号 Δ/Δ。

希克斯替代效应是什么样的？也可以定义一个斯勒茨基方程。令 $x_1^h(p_1,p_2,\overline{u})$ 表示希克斯需求函数，它表示当价格为 (p_1,p_2) 且调整消费者的收入使其效用恰好等于原效用水平 $\overline{u}$ 时，他对商品 1 的需求数量为多少。可以证明在该情形下，斯勒茨基方程的表达式为

$$\frac{\partial x_1(p_1,p_2,m)}{\partial p_1}=\frac{\partial x_1^h(p_1,p_2,\overline{u})}{\partial p_1}-\frac{\partial x_1(p_1,p_2,m)}{\partial m}\overline{x}_1$$

该式的证明取决于下式

$$\frac{\partial x_1^h(p_1,p_2,\overline{u})}{\partial p_1}=\frac{\partial x_1^s(p_1,p_2,\overline{x}_1,\overline{x}_2)}{\partial p_1}$$

当价格变动无穷小（$\Delta p_1\to 0$）时，这个式子恒成立。也就是说当价格变动无穷小时，斯勒茨基替代效应和希克斯替代效应是相同的。这个结论的证明不是非常困难，但它涉及一些新的概念，这些概念超出了本书的范围。一个相对简单的证明请见范里安，微观经济分析，第 3 版（纽约：诺顿，1992）。

例子：小额征税且退税

我们可使用微分形式的斯勒茨基方程，分析退税情形下如果税收稍微变动对消费者需求的影响。

和以前一样，假设税收使得商品价格上升为商品本身价格加上全部税额。令 x 表示汽油的消费量，p 表示汽油的初始价格，t 表示单位税额。征税且退税后消费量的变动

可用下式表示

$$dx = \frac{\partial x}{\partial p} t + \frac{\partial x}{\partial m} tx .$$

上式右端第一项衡量价格变动引起的需求变动$(\partial x/\partial p)$与价格变动量（t）的乘积，这就是税收的价格效应；上式右端第二项表示由收入变动引起的需求变动（$\partial x/\partial m)$与收入变动量（tx）的乘积，此处消费者收入增加额等于政府对他的退税额。

现在使用斯勒茨基方程扩展上式右端第一项(扩展后上式第一项变为下式括号内的项)，由此可得商品价格自身变动引起的替代效应和收入效应：

$$dx = (\frac{\partial x^s}{\partial p} t - \frac{\partial x}{\partial m} tx) + \frac{\partial x}{\partial m} tx = \frac{\partial x^s}{\partial p} t .$$

可以看出，在上式中，收入效应互相抵消掉了，剩下的只有纯替代效应。政府征收少量税收并且退税，相当于使价格升高然后补偿消费者让他还能买得起原来的消费束。只要税收足够小，这种导数近似就是可行的。

总结

1.商品 1 的价格下降将会对消费产生两种效应。首先，商品 1 价格下降后，两商品的相对价格因此变动，这会使消费者希望多消费相对便宜的商品 1；其次由于商品 1 价格下降，消费者的购买力因此增加，这将导致消费增加（正常商品情形）或者消费减少（劣等商品情形）。

2.由相对价格变动引起的需求变动称为替代效应；由于购买力变化引起的需求变动称为收入效应。

3.替代效应是指当价格变动但保持购买力不变（仍能买得起原消费束）时，需求如何变动。为保持实际购买力不变，必须调整货币收入。货币收入调整数额$\Delta m = x_1 \Delta p_1$。

4.斯勒茨基方程是说需求的总变动等于替代效应与收入效应之和。

5.需求定理是说正常商品的需求曲线必然向下倾斜。

复习题

1.假设消费者对两种商品的偏好是完全替代类型。在此情形下，你能变动价格使得

需求总变动全部由收入效应引起吗？

2.偏好为凹的情形下，替代效应还为负吗？

3.在教材汽油税的例子中，如果退税的依据是他们汽油初始消费量 x 的多少，而不是最终消费量 x'，结果会如何？

4.在上题中，政府退税总额大于还是小于它的税收收入？

5.如果根据汽油初始消费量退税这一提案生效，消费者的状况是改善了还是恶化了？

复习题答案

1.假设消费者对两种商品的偏好是完全替代类型。在此情形下，你能变动价格使得需求总变动全部由收入效应引起吗？

【复习内容】斯勒茨基替代效应和收入效应；完全替代

【解题思路与参考答案】

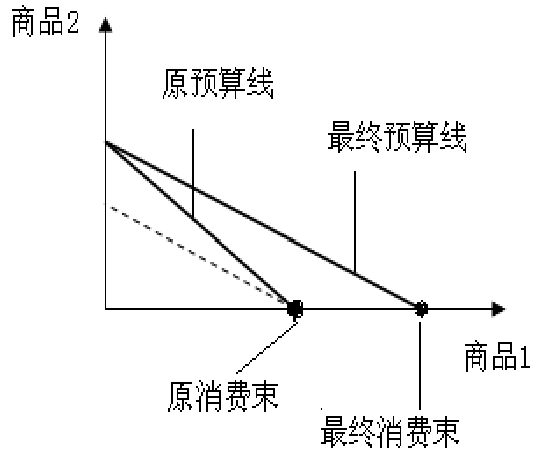
教材中的那个完全替代的例子表明，需求的变动完全由替代效应引起。本题的旨意是打破你的思维惯性。

我们知道，替代效应，粗略地说，是指当一种商品比如商品 1 价格变动后，如果商品 1 相对（商品 2）更贵，则消费者会减少商品 1 的消费（即用商品 2 替代商品 1）；如果商品 1 相对更便宜，则消费者会增加商品 1 的消费，即用商品 1 替代商品 2，也就是说减少商品 2 的消费。本题要求替代效应为 0，即什么情形下不能再减少商品 2 的消费？答案显然为商品 2 的消费量已经为 0，因为我们要求消费量不能为负。

这就是说如果两种商品是 1：1 完全替代的，而且商品 1 的价格低于商品 2 的价格，则消费者只会消费商品 1。现在令商品 1 的价格下降，消费者仍然只会消费商品 1，但商品 1 消费量的增加完全由收入效应引起，因为商品 2 的消费量原本已为 0，不可能再用商品 1 替代商品 2。由下图表示。

下图说明了下列道理：在 1：1 完全替代情形下，因为商品 1 的价格低于商品 2，消费者只会消费商品 1，现在商品 1 的价格下降，消费者仍然只会消费商品 1。

商品 1 消费量的增加全部由于预算线平移（由虚线表示的预算线平移到最终预算线）引起，这正是收入效应。所以该情况下，商品 1 消费量的总变动完全由收入效应引起。



2.偏好为凹的情形下，替代效应还为负吗？

【复习内容】凹偏好；替代效应

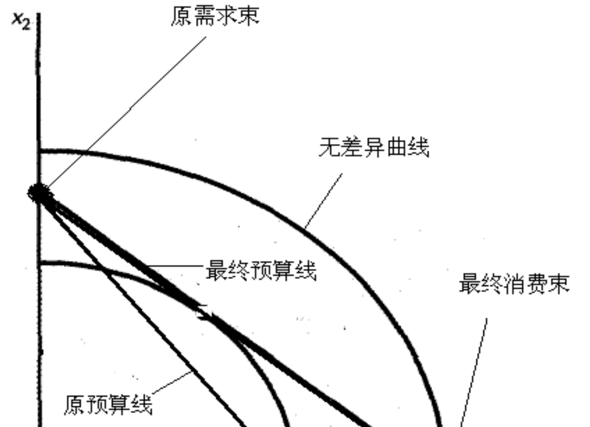
【解题思路与参考答案】

注意对替代效应的理解。最好这样理解“替代效应恒为负”：如果存在替代效应（即替代效用不为 0），则替代效应使商品需求量的变动方向与价格变动方向相反。

如果你注意到了完全替代类型的偏好（效用函数为线性函数）既是凹偏好又是凸偏好（请根据定义自行证明），那么你至少已经举出一个凹偏好情形下替代效应为负的例子。请参考教材本章中的那个完全替代的例子。

其实，即使在凹偏好的情形下，替代效应也恒为负。

可以画图分析，见下图。该图形与教材图 8.5 几乎完全相同，区别在于此处我们用凹偏好（无差异曲线向下凹）代替了图 8.5 的线性无差异曲线，解释几乎也相同。



由于凹偏好意味着端点束好于平均束，因此凹偏好情形下必有一种商品的消费量为

曹乾（东南大学 caoqianseu@163.com）

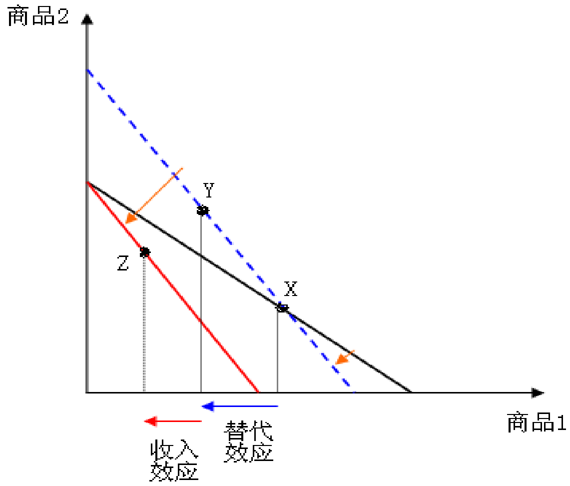
0（请参考第 6 章复习题第 5 题）。在我们画出的上图中，起初商品 1 的消费量为 0，消费者只消费商品 2。现在假设商品 1 价格下降后，商品 1 相对商品 2 更便宜，则消费者只会消费商品 1，商品 2 的消费量为 0。在图形中可以可出，商品 1 需求量的总变动完全由预算线转动引起，也就是说需求量总变动完全由替代效应引起。商品 1 价格下降，替代效应使商品 1 需求量增加，因此替代效应符号为负。

3.在教材汽油税的例子中，如果退税的依据是他们汽油初始消费量 x 的多少，而不是最终消费量 x'，结果会如何？

【复习内容】征税与退税对消费的影响

【解题思路与参考答案】

如果政府按初始汽油消费量退税，则对于汽油消费量为 x 的某个既定的消费者来说，这意味着他得到的退税额 tx，其中 t 为单位税额。



由于征税相当于汽油价格提高，但现在政府给他“补贴”（退税）tx 元，因此该消费者恰好还能买得起原来的需求束。从教材中可知将预算线变动分解为转动和平移，分别对应着替代效应和收入效应。政府的上述做法恰好相当于抵消了收入效应，因此总效应等于替代效应，而替代效应恒为负，又因征税相当于汽油价格提高，所以他的汽油需求量会减少。

最容易的理解思路是首先分析政府只征税**不退税**的情形。图形分析如下，如果政府只征税不退税，这就等价于汽油(商品 1)价格升高。假设原需求束为 X，价格升高后需求束变为 Z。我们要将需求量的变动分解为两部分。

商品 1 价格升高，为了得到纯替代效应，必须**假定**补偿他一些资金使得他恰好还能买得起原需求束，补偿金额为 tx。这样原预算线（黑线）**转动**到补偿预算线（蓝色虚线），此时需求束为 Y。由于商品 1 价格升高，他的收入相当于**减少**，因此将补偿预算线向左**平行移动**到最终预算线，此时需求束为 Z。由图可以看出，替代效应和收入效应都使汽油的需求量下降。

然而在本题中，政府的确给与消费者补贴（退税额为tx），因此如果继续上面的变化过程，可以看到最终预算线又返回到补偿预算线，因此此时消费束为 Y。这就表明商品 1 价格升高的收入效应被政府的退税政策抵消了，剩下的是纯粹的替代效应。

4.在上题中，政府退税总额大于还是小于它的税收收入？

【复习内容】征税与退税对消费者的影响

【参考答案】由上题可知道，对汽油征税后又退税，消费者对汽油的消费量会下降，即$x' < x$，所以$tx' < tx$，这表明政府征税收入tx'小于政府支付的退税额tx，因为政府是根据消费者初始汽油消费量（x）退税的，但征税却要根据消费者的实际消费量（x'）。

5.如果根据汽油初始消费量退税这一提案生效，消费者的状况是改善了还是恶化了？

【复习内容】征税与退税对消费者的影响

【参考答案】

请看第 3 题答案中的图形，根据显示偏好弱公理可知道，此情形下消费者的状况改善了。原因在于，由第 3 题的分析或由第 3 题的图都可知道，此情形下消费者仍然能够买得起原来的需求束 X。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

9.买与卖（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

9 购买与销售的决策

在前面几章的消费者行为模型中，消费者的收入是给定的。在现实生活中，人们为了增加自己的收入，可以出售东西。所出售的东西包括：生产出的商品；积累的资产；或者更一般地，他们的劳动。在本章，我们研究：为了描述这些行为，必须如何修改前几章的模型。

9.1 净需求和总需求

和前几章一样，我们仅限于研究两种商品的情形。我们从消费者的禀赋（endowment）开始分析，**禀赋**是指消费者在进入市场之前就拥有的财富。假设某消费者的禀赋为两种商品，记为(ω_1,ω_2) [一]。例如，某个农民的禀赋为ω_1单位的胡萝卜和ω_2单位的马铃薯。在进入市场后，他会看看市场价格如何，以便决定在这两种商品中买进哪一种卖出哪一种。

为便于说明，有必要区分两个概念：总需求（gross demands）和净需求(net demands)。我们以某消费者对商品 1 的需求为例说明这两个概念的区别。某消费者对商品 1 的**总需求**是指他最终实际消费的商品 1 的数量，即最终他从市场带多少商品 1 回家。消费者对商品 1 的**净需求**是指他对商品 1 的总需求减去他的商品 1 的初始禀赋数量。消费者对商品 1 的净需求就是他买进或卖出商品 1 的数量。商品 2 的总需求和净需求的区别可以类推。

令(x_1,x_2)表示总需求，则$(x_1-\omega_1,x_2-\omega_2)$为净需求。注意，总需求通常为正值，但净需求可能为正值也可能为负值。如果商品 1 的净需求为负，表明消费者商品 1 的消费量小于其禀赋量，也就是说他是商品 1 的**供给者**。负的净需求量就是供给量。

对于经济分析的目的来说，总需求的概念比较重要，因为消费者最终关心的就是总需求。但是，净需求是市场实际显示的需求（或供给），因此净需求的概念接近于初学者眼中的需求或供给。

9.2 预算约束

我们首先分析预算约束的形式。消费者最终消费的约束条件是什么？这个约束条件必然是：他带回家商品束的价值等于他带到市场中的商品束的价值。用代数表达式表达：

$$p_1x_1+p_2x_2=p_1\omega_1+p_2\omega_2$$

也可以用净需求表示上述预算线，将上式变形即可得到：

$$p_1(x_1-\omega_1)+p_2(x_2-\omega_2)=0\ .$$

[一] ω是希腊字母，读作欧米茄（o–may–gah）。

如果 $(x_1-\omega_1)$ 为正，则称消费者为商品 1 的**净购买者**（net buyer）或**净需求者**(net demander)；如果 $(x_1-\omega_1)$ 为负，则称消费者为商品 1 的**净出售者**（net seller）或**净供给者**(net supplier)。因此用净需求表达的上述预算线是说，消费者购买商品的价值必定等于他出售商品的价值。

在消费者拥有禀赋的情形下，我们也可以将上述预算线，按照我们在前几章的预算线形式表达，但要分为两个等式：

$$p_1x_1+p_2x_2=m$$
$$m=p_1\omega_1+p_2\omega_2\,.$$

如果商品价格固定不变，那么禀赋的价值从而消费者的货币收入，都是固定不变的。

这样的预算线对应什么样的图形？如果商品价格不变，则消费者的货币收入也不变，因此这样的预算线和我们前几章学习的预算线是一样的。所以，这条预算线的斜率一定也等于 $-p_1/p_2$。剩下的问题就是确定它的位置。

如果你知道下列事实：<u>禀赋束一定在预算上</u>，这条预算线的位置就容易确定。禀赋束必定在预算线上，这是因为你将 $x_1=\omega_1$ 和 $x_2=\omega_2$ 代入预算线，等式显然是成立的。这表明，消费者总能买得起他的禀赋，原因在于他花费的钱数恰好等于禀赋的价值。

综上可知，这条预算线的斜率为 $-p_1/p_2$ 且经过禀赋点。如图 9.1 所示。

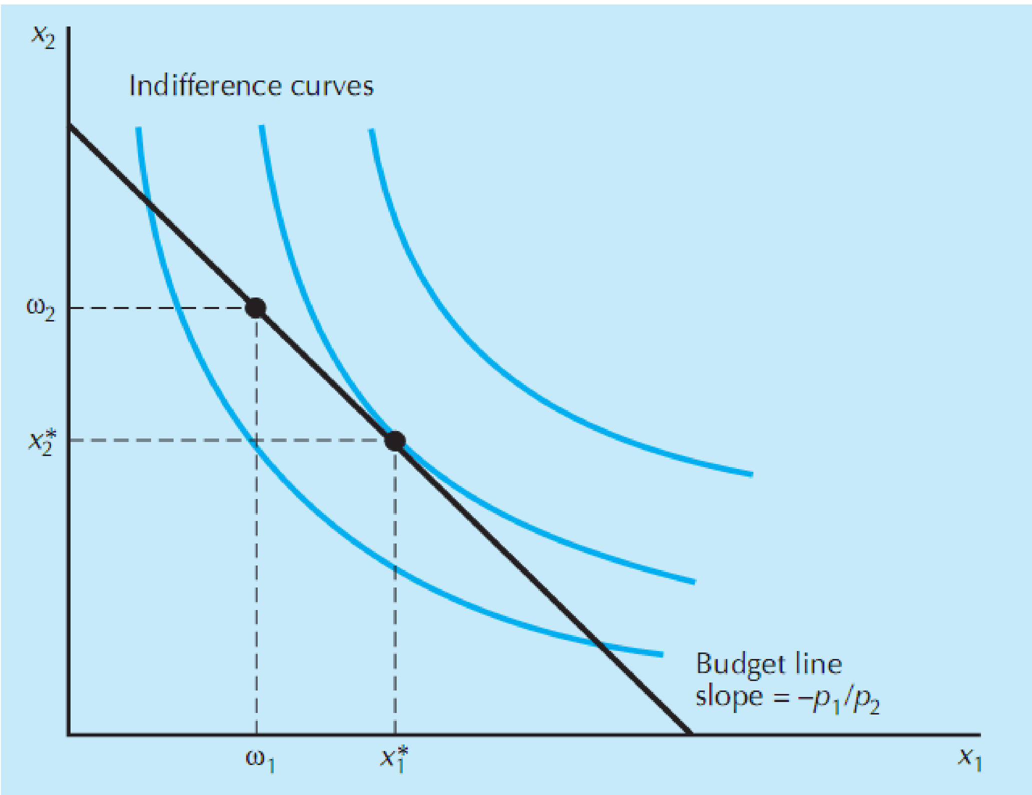


图 9.1 预算线：预算线经过禀赋点，且斜率为 $-p_1/p_2$。

给定预算约束，消费者就可以像前几章一样选择最优的消费束。在图 9.1 中，最优消费束为 (x_1^*, x_2^*)。像前几章一样，这个消费束满足最优条件：边际替代率的绝对值等于价格之比。

在图 9.1 中，容易看出 $x_1^* > \omega_1$ 和 $x_2^* < \omega_2$，因此消费者是商品 1 的净购买者和商品 2 的净出售者。净需求就是消费者买进或卖出两种商品的数量。一般来说，消费者根据这两种商品的相对价格，决定是买进还是卖出商品。

9.3 禀赋变动

在前面几章，我们已分析过：当价格不变而消费者的货币收入变动时，他的最优选择如何变动。类似地，我们可以研究当价格不变而**禀赋数量**变动时，他的最优选择如何变动的。

假设禀赋数量从 (ω_1, ω_2) 变为 (ω_1', ω_2') 使得

$$p_1\omega_1 + p_2\omega_2 > p_1\omega_1' + p_2\omega_2'.$$

这个不等式表明，新禀赋束 (ω_1', ω_2') 的价值小于原禀赋束：新禀赋束销售所得货币收入变小了。

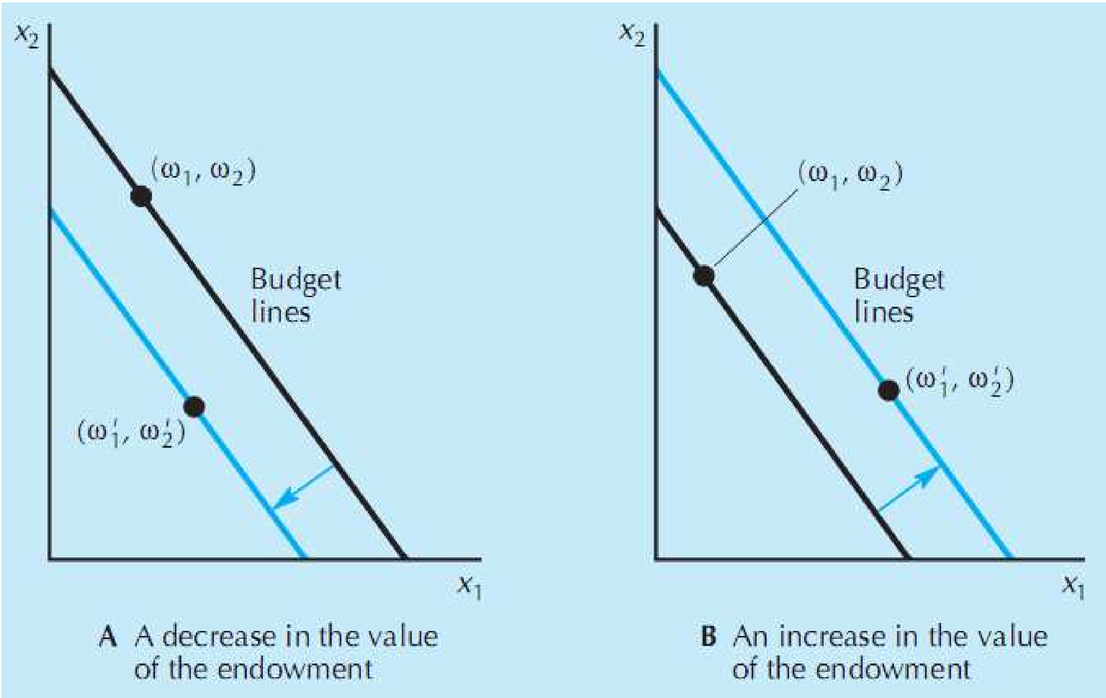


图 9.2：禀赋价值的变动。A 图表示禀赋价值减少，B 图表示禀赋价值增加。

这一变化可用图 9.2A 说明：预算线向内平移。由于这种变化等价于货币收入减少，我们可以得到以下两个结论（和前面章节货币收入减少的结论相同）。首先，与原禀赋束相比，新禀赋束代表的消费者状况更差，因为他的消费可能性减小了。其次，他对每种商品的需求

将发生变化，变化方向取决于商品是正常商品还是劣等商品。

例如，如果商品 1 是正常商品，上述禀赋变化减少了商品的价值，我们可以推知消费者对商品 1 的需求将下降。

图 9.2B 描述的情形是禀赋的价值增加了。预算线向外平移，消费者的状况必定变好。用代数语言表达，若禀赋从 (ω_1,ω_2) 变为 (ω_1',ω_2') 使得 $p_1\omega_1 + p_2\omega_2 < p_1\omega_1' + p_2\omega_2'$，则消费者的新预算集必定包含他的原预算集。这意味着消费者在新预算集下的最优选择，一定好于原禀赋情形下的最优选择。

上述结论值得思考。在第 7 章我们说过，商品束 1 的花费比商品束 2 高，并不意味着商品束 1 比 2 更好。但第 7 章的这个结论仅在商品束必须**被消费**的情形下才成立。如果消费者能在竞争市场以既定价格销售商品束，那么他对价值高的商品束的偏好总是胜于价值低的商品束。原因在于价值高的商品束销售收入更多，从而消费可能性越多。因此，价值高的**禀赋**总是比价值低的禀赋更受偏好。稍后我们将看到这个简单的结论有重要的应用价值。

还有一种情形需要考虑：如果 $p_1\omega_1 + p_2\omega_2 = p_1\omega_1' + p_2\omega_2'$，结果会如何？这种情形下，预算集未变：在 (ω_1,ω_2) 和 (ω_1',ω_2') 两种情况下，消费者的状况是一样的，他的最优选择也是一样的。禀赋只是沿着原来的预算线移动。

9.4 价格变动

以前，我们分析价格变动对需求变动的影响时，我们假设消费者的货币收入保持不变。现在，货币收入取决于禀赋的价值，货币收入不变的假设就不合理了：若你销售的商品价格改变，你的货币收入当然也改变。因此，在消费者拥有禀赋的情形下，价格变动自动意味着收入变动。

我们先用图形分析。若商品 1 的价格下降，预算线会变得更平坦。因为消费者总能买得起禀赋束，所以商品 1 价格下降意味着预算线必然会绕着禀赋束转动，如图 9.3 所示。

在这种情形下，消费者最初为商品 1 的出售者，在价格**下降**后他仍然是出售者。消费者的福利如何变化？由图 9.3 可知，与价格下降前相比，价格下降后， 消费者在位置更低的无差异曲线上，但这个结论具有一般性吗？我们用显示偏好原理来回答这个问题。

如果消费者仍然是商品 1 的供给者，即 $x_1 < \omega_1$，这意味着他的新消费束必然位于新预算线在禀赋点以上的那一段。但这段新预算线在原来预算集的内部，这表明这段新预算线上的消费束，在商品 1 价格下降之前，消费者就能买的起。因此，根据显示偏好理论可知，这些消费束比原来的消费束差。由此可得下面的结论：如果消费者销售的某种商品价格下降，并且他继续充当销售者，则他的福利一定是下降的。

如果消费者销售的某种商品价格下降，并且他决定转变为这种商品的购买者，消费

者的福利如何变化？这种情形下，我们无法判断消费者的福利是增加还是下降。

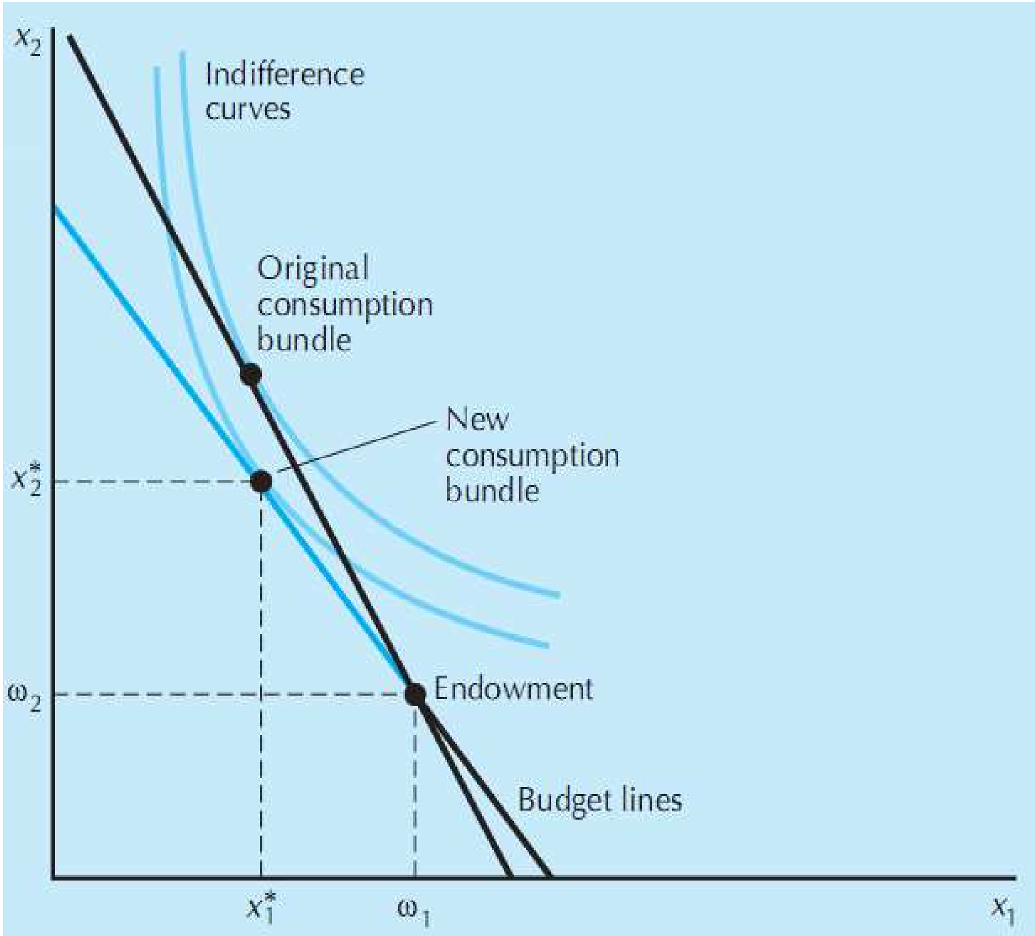


图 9.3：商品 1 价格下降。商品 1 价格下降，预算线会绕着禀赋束转动。若消费者仍是供给者，他的状况会变差。

现在我们开始研究消费者是某种商品的净购买者的情形。这种情形与消费者是净销售者的情形正好相反：如果他是某种商品的净购买者，当该商品价格上升时，如果消费者继续选择作为销售者，那么他的状况必定变差。但是，如果价格上升后，他决定变为销售者，则他的状况可能变好也可能变差。和前面一样，这些结论是从显示偏好理论推得，但建议你自己画图分析一下，以确保你理解了这些推导过程。

我们还可以使用显示偏好理论分析当价格改变时，消费者是继续充当购买者还是应该变为销售者。假设，如图 9.4 所示，消费者是商品 1 的净购买者，如果商品 1 的价格下降，结果会如何？请看图 9.4，该情形下预算线变得更平坦。

和以前一样，我们不知道该消费者会多买还是少买商品 1，因为这取决于他的偏好。然而，有一点是肯定的：**消费者仍然选择作为商品 1 的净购买者，他不会选择变为销售者**。

为什么？用反证法。如果消费者变为销售者，结果会怎样？如果变为商品 1 的销售者，即 $x_1 < \omega_1$，则他的消费束必然位于新预算线禀赋点以上的那一段，见图 9.4。但是这些消费

束他原来就能买得起，因为它们位于原预算线的内部；能买得起但他却未买，因为他实际选择的是(x_1^*, x_2^*)。因此，(x_1^*, x_2^*)必然比这些消费束都好。在新预算线的情形下，他仍然可以买得起(x_1^*, x_2^*)。所以，当在新预算线的情形下，消费者选择的消费束必然比(x_1^*, x_2^*)好，从而也比新预算线禀赋点以上的那一段的消费束都好。这意味着商品 1 的消费量必然位于新预算线禀赋点以下的一段，也就是说，他仍然选择作为商品 1 的净购买者。

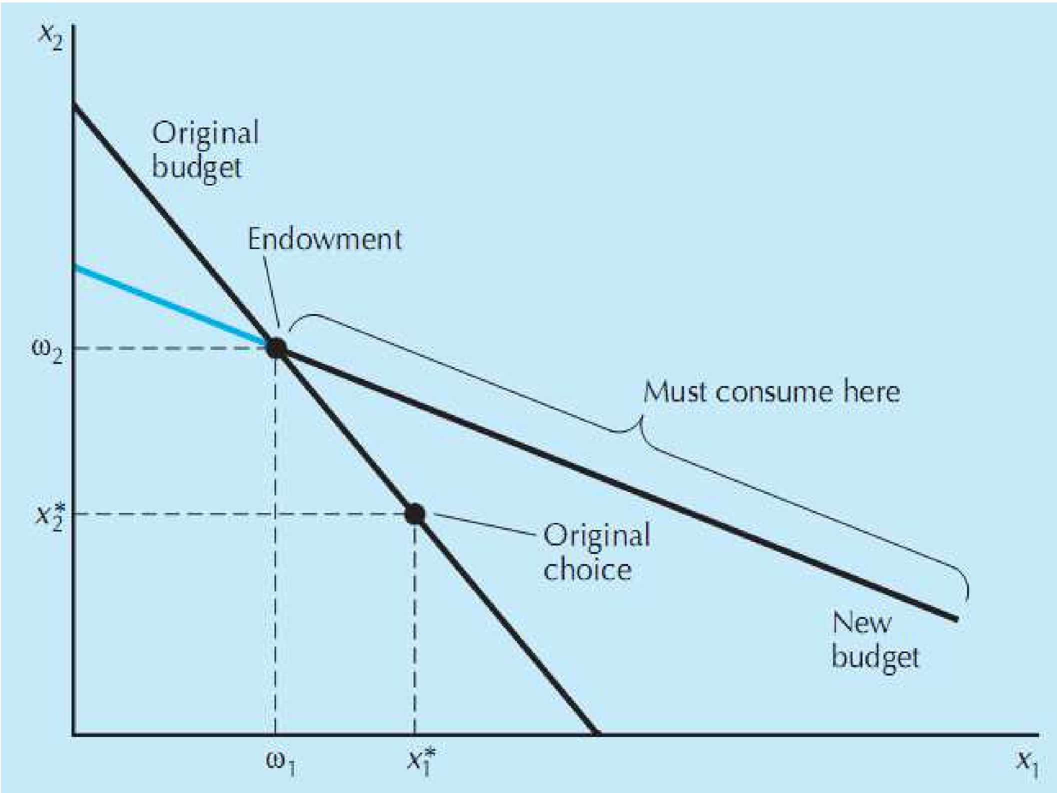


图 9.4：商品 1 价格下降的情形。如果消费者是商品 1 的购买者，在商品 1 价格下降时，他会选择继续作为商品 1 的购买者。

类似地，我们可推知如果消费者是商品 1 的净销售者，那么在该商品价格上升时，他仍然选择作为商品 1 的净销售者，而不会选择变为净购买者。我们无法确定这个消费者是否会多消费还是少消费商品 1，但是我们可以肯定，当商品 1 价格上升时，他会继续销售商品 1.

9.5 价格提供曲线和需求曲线

我们在第 6 章已知道，价格提供曲线描述消费者对商品束的需求，而需求曲线描述的是某商品的需求量和该商品价格之间的关系。在消费者拥有禀赋的情形下，我们也可以构造价格提供曲线和需求曲线。

例如，图 9.5 给出了某消费者的价格提供曲线和需求曲线。价格提供曲线必然经过禀赋点，因为在某价格下消费者会选择禀赋束，也就是说在某价格下消费者的最优选择是不买也不卖。

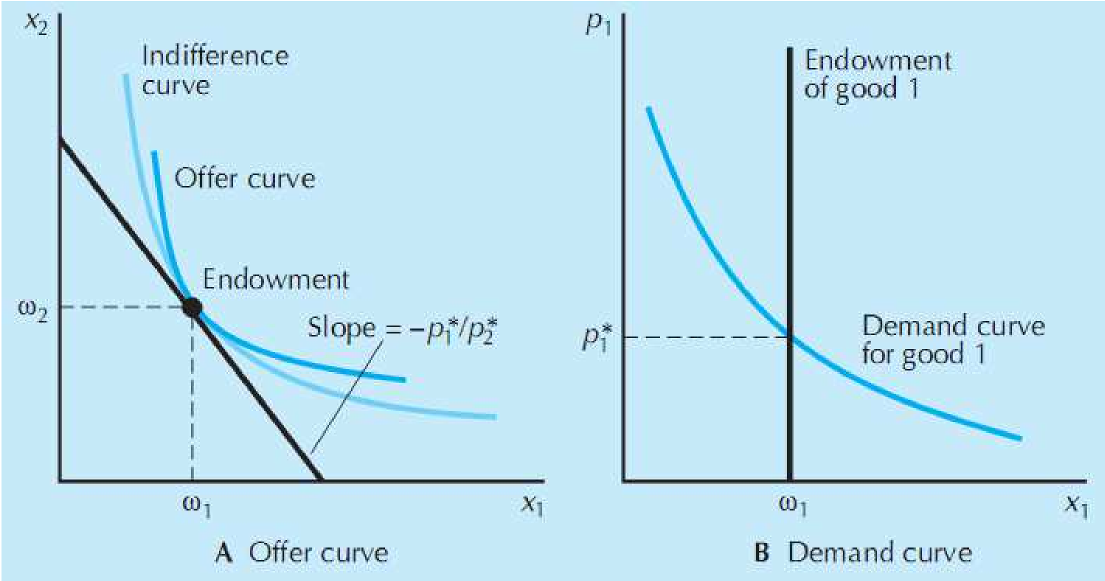


图 9.5：价格提供曲线和需求曲线。在消费者拥有禀赋的情形下，可用价格提供曲线或需求曲线描述需求束和价格的关系。

我们已经知道，在某些价格下消费者会卖出商品 1（即 $x_1 < \omega_1$），而在另外的价格下他会买进商品 1（即 $x_1 > \omega_1$）。因此随着商品 1 价格下降，商品 1 的价格提供曲线通常从禀赋点的左侧穿过禀赋点然后向右侧延展。

图 9.5B 中的需求曲线是总需求曲线，他衡量不同价格下消费者消费商品 1 的数量。与该总需求曲线相应的净需求曲线请见图 9.6。

注意，在某些价格下，商品 1 的净需求为负。这时商品 1 的价格应该很高，因此消费者出售而不是继续买进商品 1。也就是说如果商品 1 的价格很高，消费者就会选择从净需求者转变为净销售者。

尽管可以将供给视为负的净需求，但人们习惯将供给曲线画在正的像限中。我们遵从这一惯例，将供给曲线画成向右上方倾斜的曲线，如图 9.6 所示。

如果我们将某消费者对商品 1 的净需求 $d_1(p_1,p_2)$ 和净供给 $s_1(p_1,p_2)$ 都规定为**非负数**，则可用以下代数表达式分别定义商品 1 的净需求和净供给：

$$d_1(p_1,p_2)=\begin{cases} x_1(p_1,p_2)-\omega_1 & \text{若} x_1(p_1,p_2)>\omega_1; \\ 0 & \text{其他情形.} \end{cases}$$

$$s_1(p_1,p_2)=\begin{cases}\omega_1-x_1(p_1,p_2) & \text{若}\ \omega_1-x_1(p_1,p_2)>0;\\ 0 & \text{其他情形.}\end{cases}$$

也就是说，当商品 1 的总需求 $x_1(p_1,p_2)$ 大于商品 1 的禀赋 ω_1 时，净需求为二者之差，其他情形下净需求为 0；当商品 1 的禀赋 ω_1 大于商品 1 的总需求 $x_1(p_1,p_2)$ 时，净供给为二者之差，其他情形下净供给为 0.

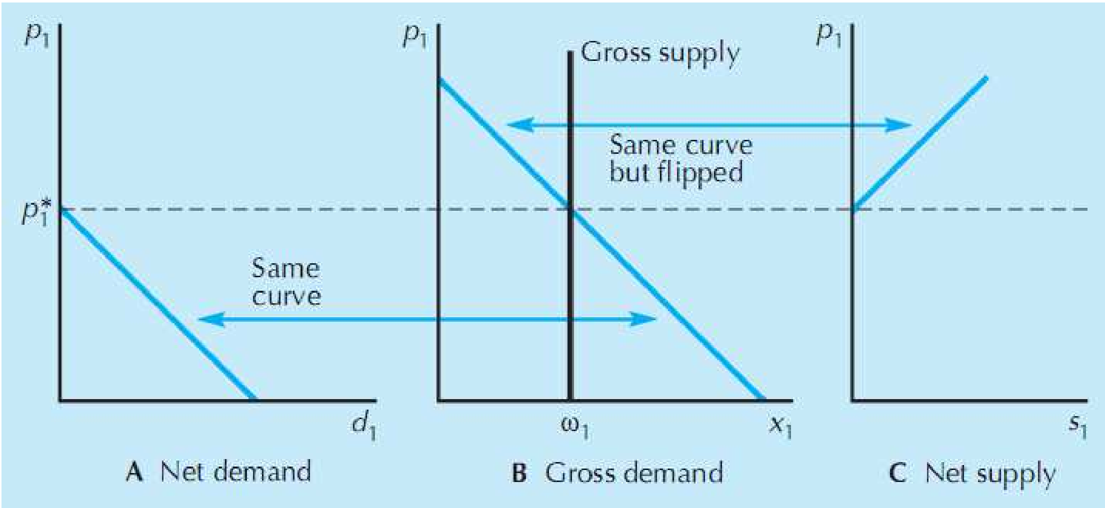


图 9.6：总需求、净需求和净供给。使用总需求和净需求来刻画需求行为和供给行为。

需求行为的一切性质，可以直接应用于供给行为，因为供给就是负的需求。如果**总需求**曲线总是向下倾斜，则净需求曲线也是向下倾斜的，而供给曲线是向上倾斜的（原因请参见上面的代数表达式）。如果价格上升使净需求曲线更陡峭，那么供给曲线也会更陡峭。

9.6 斯勒茨基方程再思考

在本章前面几节我们看到，显示偏好理论比较好用，但它们无法回答我们的主要问题：价格变动如何影响商品的需求。在第 8 章我们已知道，如果货币收入不变，并且商品为正常商品，那么价格下降必然导致需求增加。

上述结论有个假设前提，即货币收入不变。我们此处研究的情形是货币收入会发生变动，因为，在消费者拥有禀赋的情形下，价格变动当然会使禀赋价值变动。

在第 8 章，我们使用斯勒茨基方程把由价格变动引起的需求变动，分解为替代效应和收入效用。收入效应是由于价格变动导致购买力变动。但现在，当价格变动时，购买力改变的原因有两个。第一个原因涉及斯勒茨基方程的定义：例如，当某种商品价格下降，你还能购买原来的消费量并且还有余钱。我们把这种效应称为**普通收入效应**（ordinary income effect)。第二种收入效应我们以前没介绍过。当某种商品价格改变时，你的禀赋价值也改变，从而你的货币收入也改变。例如，如果你是某种商品的净供给者，该商品价格下降会直接减少你的收入，因为禀赋价格下降了，你销售既定数量禀赋所得收入必然减少。这种由于商品价格改变导致禀赋束价值改变的收入效应，我们称为**禀赋收入效应**（endowment income effect）。由此可见，在消费者拥有禀赋的情形下，总收入效应等于普通收入效应（即第 8 章介绍的收入效应）加上禀赋收入效应。

在第 8 章的斯勒茨基方程中，货币收入是固定不变的。但现在我们必须考虑禀赋价值变动引起的货币收入变动。因此，当计算价格变动对需求变动的效应时，斯勒茨基方程将是下面的形式：

需求的总变动=替代效应引起的需求变动+普通收入效应引起的需求变动+禀赋收入效应引起的需求变动。

上式右侧中前两个效应我们并不陌生。和以前一样，令 Δx_1 表示需求的总变动，Δx_1^s 表示由替代效应引起的需求变动，Δx_1^m 表示由普通收入效应引起的需求变动。将这些符号代入上述式子可得到用变化率表示的斯勒茨基方程：

$$\frac{\Delta x_1}{\Delta p_1} = \frac{\Delta x_1^s}{\Delta p_1} - x_1 \frac{\Delta x_1^m}{\Delta m} + \text{禀赋收入效应} \qquad (9.1)$$

（9.1）式右侧最后一项即禀赋收入效应的表达式是怎样的？在推导它的表达式之前，我们思考一下它的结构。禀赋价格改变导致货币收入改变，货币收入改变又会导致需求改变。因此，禀赋收入效应包含两项：

禀赋收入效应=收入改变导致的需求改变×禀赋价格改变导致的收入改变　（9.2）

我们先分析（9.2）式右侧的第二个效应。由于货币收入定义为

$$m = p_1\omega_1 + p_2\omega_2.$$

由此可得

$$\frac{\Delta m}{\Delta p_1} = \omega_1.$$

这个式子表示由商品 1 价格改变引起的货币收入改变量为多大：如果你有 10 单位商品 1 出售，当它的价格上升了 1 元，你的货币收入增加了 10 元。

（9.2）式右侧的第一项是，当收入改变时需求的改变量为多大。在第 8 章我们已经知道它的表达式为 $\Delta x_1^m / \Delta m$，即需求改变量除以收入改变量。因此，禀赋收入效应的表达式为：

$$\text{禀赋收入效应} = \frac{\Delta x_1^m}{\Delta m}\frac{\Delta m}{\Delta p_1} = \frac{\Delta x_1^m}{\Delta m}\omega_1 \qquad (9.3)$$

将（9.3）代入（9.1）即可得到斯勒茨基方程的最终表达式：

$$\frac{\Delta x_1}{\Delta p_1}=\frac{\Delta x_1^s}{\Delta p_1}+(\omega_1-x_1)\frac{\Delta x_1^m}{\Delta m}.$$

我们可使用该方程回答上面提出的商品价格变动引起的需求变动问题。我们已知道，替代效应的符号总为负，即与价格变动方向相反。假设商品为正常商品，因此，$\Delta x_1^m/\Delta m>0$。于是，总收入效应的符号取决于消费者是商品 1 的净购买者还是净销售者：当他为商品 1 的净购买者（即 $w_1-x_1<0$）时，总收入效应为负；当他为商品 1 的净销售者（即 $w_1-x_1>0$）时，总收入效应为正。接下来我们看总效应的符号。如果消费者是某正常商品的净需求者，当该商品价格上升时，由于替代效应和总收入效应均为负，则总效应必然为负，也就是说消费者会减少购买量。如果消费者是某正常商品的净供给者，则总效应的符号不明朗：因为此时替代效应为负，而总收入效应为正。也就是说，这种情形下消费者可能会增加商品 1 的购买量(当替代效应小于总收入效应时)，也可能减少商品 1 的购买量（当替代效应大于总收入效应时）。

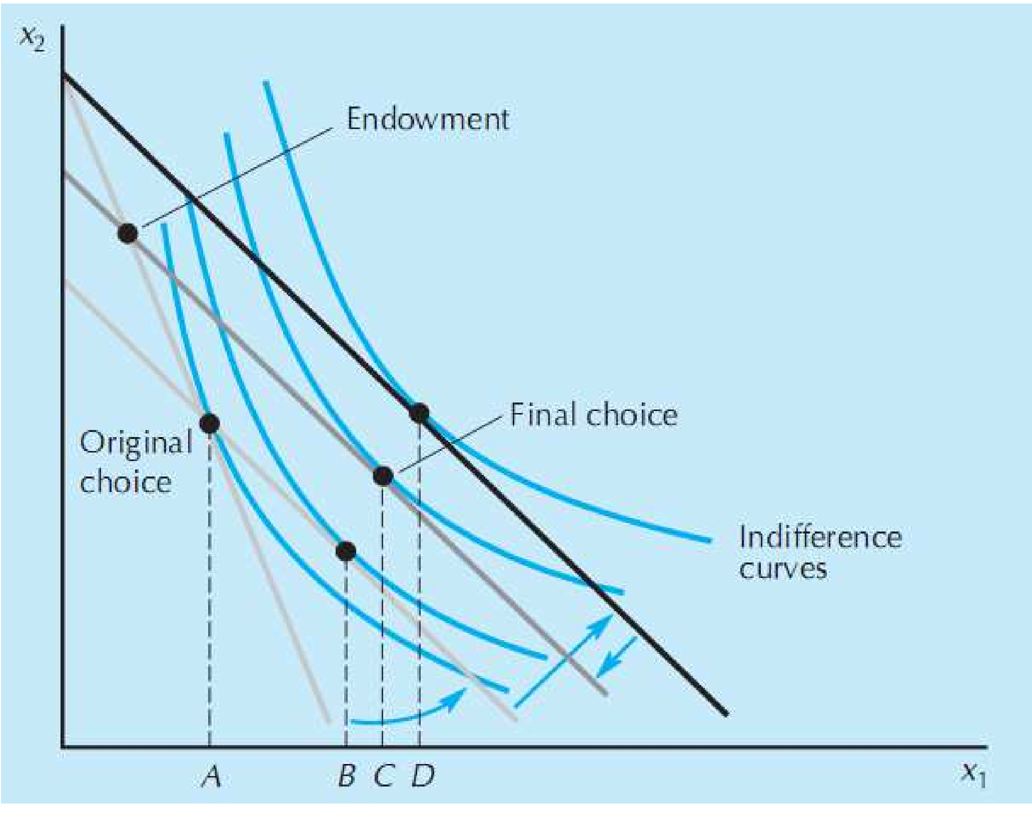


图 9.7：斯勒茨基方程的再思考。价格变动的总效应分解为三个部分：替代效应（AB）；普通收入效应（BD）；禀赋收入效应（DC）。

和以前一样，上述每种变动都可以用图形表示，只不过此时图形看上去有些乱。请看图 9.7。这个图刻画了价格变动的斯勒茨基分解。商品 1 的需求的总变动由 A 点到 C 点的位移表示。这个位移可分解为三部分：替代效应（A 点到 B 点的位移）和两种收入效应。普

通收入效应（B 点到 D 点的位移），是**维持货币收入不变时**需求的变动量，也就是说普通收入效应就是第 8 章介绍的收入效应。但是由于商品价格改变引起禀赋价值改变，此时有个额外的收入效应，即禀赋价值改变引起消费者的货币收入改变。货币收入的这种变动使得预算线向内平移并且通过禀赋束。由禀赋收入效应引起的需求变动可用 D 点到 C 点的位移表示。

9.7 斯勒茨基方程的应用

假设某个消费者在他家后院种植了一些柑橘和苹果，收获后拿到市场中销售。在第 8 章我们已经分析过类似的例子，那时我们得出的结论是，如果苹果价格上升，这个消费者会多消费苹果。使用本章推导出的斯勒茨基方程可很快明白其中的原因。令 x_a, p_a 分别表示该消费者对苹果的需求和苹果的价格，则有

$$\frac{\Delta x_a}{\Delta p_a} = \underset{(-)}{\frac{\Delta x_a^s}{\Delta p_a}} + \underset{(+)}{(\omega_a - x_a)}\underset{(+)}{\frac{\Delta x_a^m}{\Delta m}} .$$

这就是说，由苹果价格变动引起的苹果需求变动等于替代效应与收入效应之和。替代效应为负，苹果价格上升时，它的需求量下降；但由于对该消费者来说，苹果是正常商品，而且他是苹果的净销售者，苹果价格上升使他的货币收入增加，因此他会多消费苹果，也就是说收入效应的符号为正。如果收入效应超过了替代效应，总效应必然为正，即该消费者最终会多消费苹果。由此可见，某正常商品价格上升需求量也上升，这种似乎“反常”的现像已得到了合理解释。

例子：计算禀赋收入效应

举一个数值计算的例子。假设某奶农每周可生产 40 单位牛奶(x_1)。牛奶的初始价格 p_1 为 3 元，他对牛奶的需求即自己消费的数量为

$$x_1 = 10 + \frac{m}{10p_1} .$$

根据已知条件，可计算出他的货币收入为 120 元每周，他对牛奶的需求 x_1=14。现在假设牛奶的价格降为 2 元，他的货币收入变为 $m' = 2 \times 40 = 80$ 元，他的需求量为 $x_1' = 10 + 80/20 = 14$。

如果他的货币收入固定在 m=120 元，牛奶价格为 2 元时，他将消费 $x_1 = 10 + 120/20 = 16$ 单位的牛奶。因此，禀赋收入效应，即由于禀赋价值变动引起的需求变动为-2。替代效应和普通收入效应的问题我们在第 8 章已解决，此处不再赘述。

9.8 劳动供给

我们使用禀赋思想来分析消费者劳动供给决策。消费者可以多干活多挣钱，从而多消费；也可以少干活少挣钱，从而少消费。劳动的供给数量和商品的消费数量取决于消费者偏好和预算约束的共同作用。

预算约束

假设消费者的初始收入为M，这个收入和他是否劳动无关，比如，M为以前投资所得或者亲戚的馈赠。这部分收入称为**非劳动收入**（nonlabor income）。（非劳动收入当然可以为 0，但我们允许它为正数。）

令C表示消费量，p表示消费品的价格。再令w表示工资率，L表示劳动供给量，则预算线为：

$$pC = M + wL\,.$$

这就是说消费者消费的商品价值应等于非劳动收入与劳动收入之和。

将上式与以前的预算约束相比，可以发现主要的区别在于，此式右端存在消费者的选择变量：劳动供给。将这一项移到左边可得：

$$pC - wL = M\,.$$

这个式子更像我们以前的预算线。但是此式中有个减号，而以前的预算线表达式中通常为加号。怎么办？我们可以假设劳动供给数量存在上限，这个上限可以规定为一天 24 小时，也可以一周 7 天或者其他合理的时间限制。令$\overline{L}$表示该劳动时间的上限，在上式两边同时加上$w\overline{L}$并整理可得

$$pC + w(\overline{L} - L) = M + w\overline{L}\,.$$

再令$\overline{C} = M / p$表示消费者初始消费量（他不工作时的消费量），或者将$\overline{C}$称为他的消费禀赋（endowment of consumption），则有：

$$pC + w(\overline{L} - L) = p\overline{C} + w\overline{L}\,.$$

现在这个式子更像我们以前的预算线表达式。在此式左侧我们有两个选择变量（即 C 和 L），在此式右侧我们有两个禀赋变量（即$\overline{C}$和$\overline{L}$）。等式左侧中的$\overline{L} - L$可以解释为“闲暇”的数量，即这些时间内消费者在享受闲暇而不是在工作。令R表示闲暇，则$R = \overline{L} - L$。显然，闲暇的时间上限$\overline{R} = \overline{L}$，于是预算约束变为

$$pC + wR = p\overline{C} + w\overline{R}\,.$$

上式在本质上等同于本章开头给出的第一个预算约束。然而上式的解释更加耐人寻味。它的意思是说，某人消费的消费品价值与闲暇的价值之和，必然等于他的消费禀赋价值与时

间禀赋价值之和，其中时间禀赋的价格按劳动工资率计算。工资率不仅是劳动的价格，而且还是**闲暇**的价格。

为什么工资率也是闲暇的价格？因为如果工资率是每小时 10 元，而你决定消费闲暇一小时，这一小时的成本为多大？答案为它等于你放弃的 10 元劳动收入，因此每小时闲暇的价格也为 10 元。经济学家有时将工资率称为闲暇的**机会成本**（opportunity cost）。

上式的左端有时称为消费者的**全部收入**（full income）或**隐性收入**（implicit income）。它衡量消费者拥有的价值，即消费禀赋价值（如果有的话）和时间禀赋价值。这样命名的目的是和消费者的**已评定收入**（measured income）这个概念区分开，已评定收入是指消费者出售部分时间以获得收入。

上述预算线美妙的地方在于，它和我们以前见过的预算线很相似。它经过禀赋点$(\overline{L},\overline{C})$且斜率为$-w/p$。禀赋是消费者参加市场交易之前原本就拥有的东西，预算线的斜率则表明两种商品在市场中的交换率。

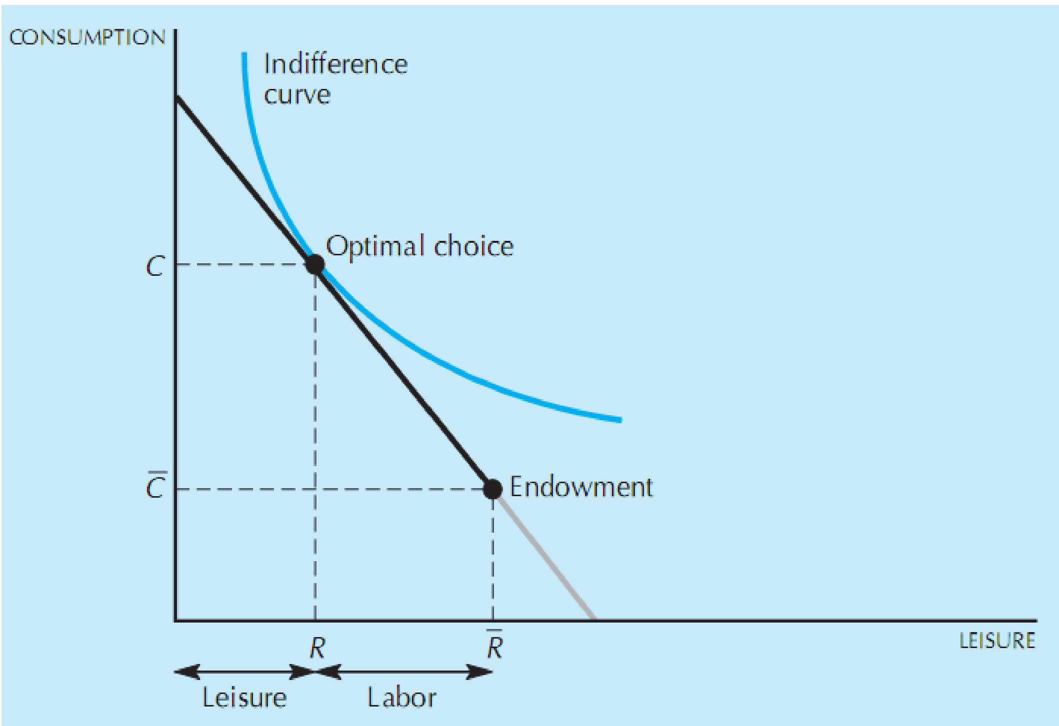


图 9.8：劳动供给。横轴从原点向右的方向表示闲暇，横轴从$\overline{R}$点向左的方向表示劳动。从该图最优选择点可以看出，消费者对闲暇的需求量为原点到 R 的距离，对劳动的需求量为$\overline{R}$到 R 的距离。

如图 9.8 所示，最优选择点位于边际替代率（即商品消费和闲暇的权衡选择）等于w/p之处。w/p称为**实际工资率**（real wage rate）。上述最优条件的意思是说，消费者对于下列二者的评价必须相等:多干一点活增加的消费；因多干那点活而损失的闲暇价值。实际工

资率是如果消费者放弃单位闲暇时间（从而工作挣钱）所能购买到的消费量。

9.9 劳动供给的比较静态分析

我们首先分析当价格和工资率不变时，某个消费者的劳动供给是如何随货币收入的改变而改变的。如果你中了彩票，你的非劳动收入大幅增加，你的劳动供给会如何变化？你对闲暇的需求会如何变化？

对于大多数人来说，如果他们的收入增加，他们会减少劳动供给。换句话说，闲暇对于大多数人来说可能是正常商品：收入增加，对闲暇的需求也增加。这个结论有不少证据，因此我们假设闲暇是一种正常商品。

这意味着工资率的变动将如何影响消费者的劳动供给？当工资率上升时，有两种效应产生：劳动报酬增加；消费闲暇的成本增加。我们可以使用收入效应和替代效应的思想，并且使用斯勒茨基方程将这些效应隔离开，进而逐一分析。

当工资率上升时时，闲暇变得更加昂贵，这本身会使人们减少闲暇的消费（替代效应）。由于闲暇是正常商品，根据第 8 章斯勒茨基方程可知，工资率上升必然会导致闲暇需求的减少，也就是说劳动供给会增加。正常商品需求曲线的斜率必然为负。如果闲暇是正常商品，则劳动供给曲线必然有正的斜率。

但是，这个分析有些问题。首先，在直觉上，工资率上升**总是**导致劳动供给增加似乎并不合理。如果我的工资已经很高，也许我会“花费”额外的收入去消费闲暇。这种行为显然是合理的。但是，我们怎样才能使得这种行为与上述经济理论（工资率上升导致劳动供给增加）兼容？

如果从某理论推导出错误答案，原因可能在于我们并没有正确使用该理论。前面的推理就犯了这样的错误。前面使用斯勒茨基方程给出的需求变动有个假设前提：**维持货币收入不变**。但是，如果工资率变动，则工作所得到的货币收入也会变动。由货币收入改变引起的需求变动是额外的收入效应，即禀赋收入效应。它是普通收入效应以外的另外一种收入效应。

因此应该使用**包含两种收入效应**的那个斯勒茨基方程进行分析，如下所示：

$$\frac{\Delta R}{\Delta w} = \underset{(-)}{\text{替代效应}} + \underset{(+)}{(\overline{R} - R)} \underset{(+)}{\frac{\Delta R}{\Delta m}}. \tag{9.4}$$

在（9.4）中：替代效应为负（因为它总是与价格变动方向相反）；$\Delta R/\Delta m$ 为正，因为我们假设闲暇是正常商品；$(\overline{R}-R)$ 也为正（因闲暇时间 R 不会超过上限 $\overline{R}$ ）。因此，总效应的符号不明朗。即使闲暇是正常商品，闲暇需求的符号也不明确，在这一点上，闲暇和其他商品的需求不同。当工资率增加时，劳动供给量可能增加也可能减少。

为什么总效应的符号不明确？当工资率上升时，替代效应说，闲暇变得更昂贵，你应该

多干活多挣钱，以消费更多其它商品（以其它商品替代闲暇）。但是，当工资率上升时，你的时间禀赋的价值也增加了。这就类似于你得到了额外的收入，你很可能将这部分收入用于消费更多的闲暇。到底哪种效应更大些？这需要实证研究来回答，理论本身不能回答这样的问题。我们必须观察人们实际的劳动决策以确定哪个效应占主导地位。

工资率上升导致劳动供给量下降的情形，可用**向后弯曲的劳动供给曲线**（backward-bending labor supply curve）表示。由上述斯勒茨基方程可知，（$\overline{R}-R$）越大（即劳动时间已够多）时，越有可能出现劳动供给曲线向后完全的现像。当$\overline{R}=R$时，即消费者全部时间用于闲暇时，由斯勒茨基方程可知，此时收入效应为 0 从而总效应等于替代效应，因此工资率上涨必然导致闲暇消费减少，即劳动供给量增加。但是，随着劳动供给量增加，工资率的每次上升(Δw)，他都多赚了一部分钱，这个额外收入等于他的劳动时间乘以Δw，因此在达到某种临界点时，他可能金额定使用这部分额外收入去“购买”更多的闲暇，也就是说**减少**他的劳动供给量。

图 9.9 展示了一个向后弯曲的劳动供给曲线。当工资率很低时，替代效应大于收入效应，工资率的增长会导致闲暇需求减少，因此劳动供给增加。但工资率很高时，收入效应超过了替代效应，工资率增长会导致闲暇需求增加，因此劳动供给**减少**。

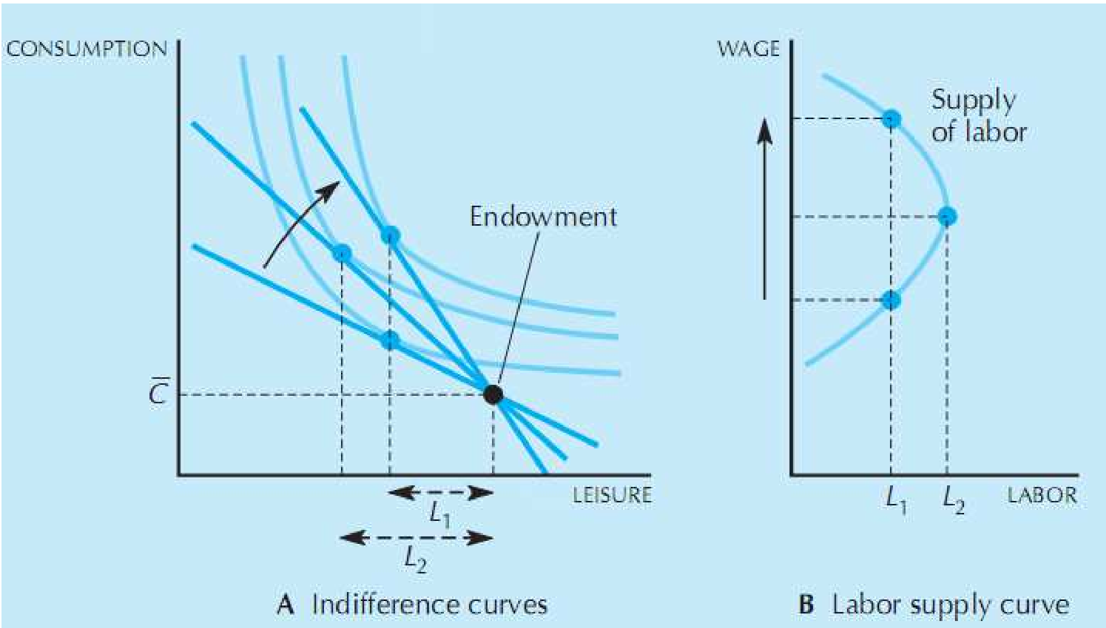


图 9.9：向后弯曲的劳动供给曲线。随着工资率上升，劳动供给从L_1增加到L_2。但是工资率继续上升时，劳动供给可能又从L_2减少到L_1。

例子：加班费与劳动供给

某工人在工资率为 w 时，他的劳动供给量为 $L^* = \overline{R} - R^*$，如图 9.10 所示。假设如果他愿意加班，则**加班工资**（overtime wage）按 w' 计算（$w' > w$）。

在图形上，这种情形下的预算线是一条折线，请看图 9.10，对应于 $L > L^*$ 那段预算线比对应于 $L < L^*$ 的那一段更为陡峭。此时消费者的最优劳动供给量必然大于 L^*，这个结论由显示偏好理论推知：在无加班的情形下，消费者本来可以选择小于 L^* 的劳动供给量，但是他并没这么做。

注意：加班工资会刺激劳动供给量增加；但是如果对于该工人全部工作时间（不仅是加班时间）都按较高的工资率 w' 计算报酬，则劳动供给量是否增加并不明确，这一点我们在前面已分析过。二者为何存在着区别？原因在于工人对加班工资的反应是纯粹的替代效应：由预算线绕 (R^*, C^*) **转动**而导致的劳动供给量变化。加班费是加班时间的报酬，而工资率的直接增长是对工人全部工作时间都按较高的工资率计算报酬。因此，直接增加工资率包含收入效应和替代效应，而加班工资只涉及替代效应。图 9.10 展示的就是这样的情形。在图 9.10 中，直接增加工资率导致劳动供给下降，而加班工资率的增加则会导致劳动供给增加。

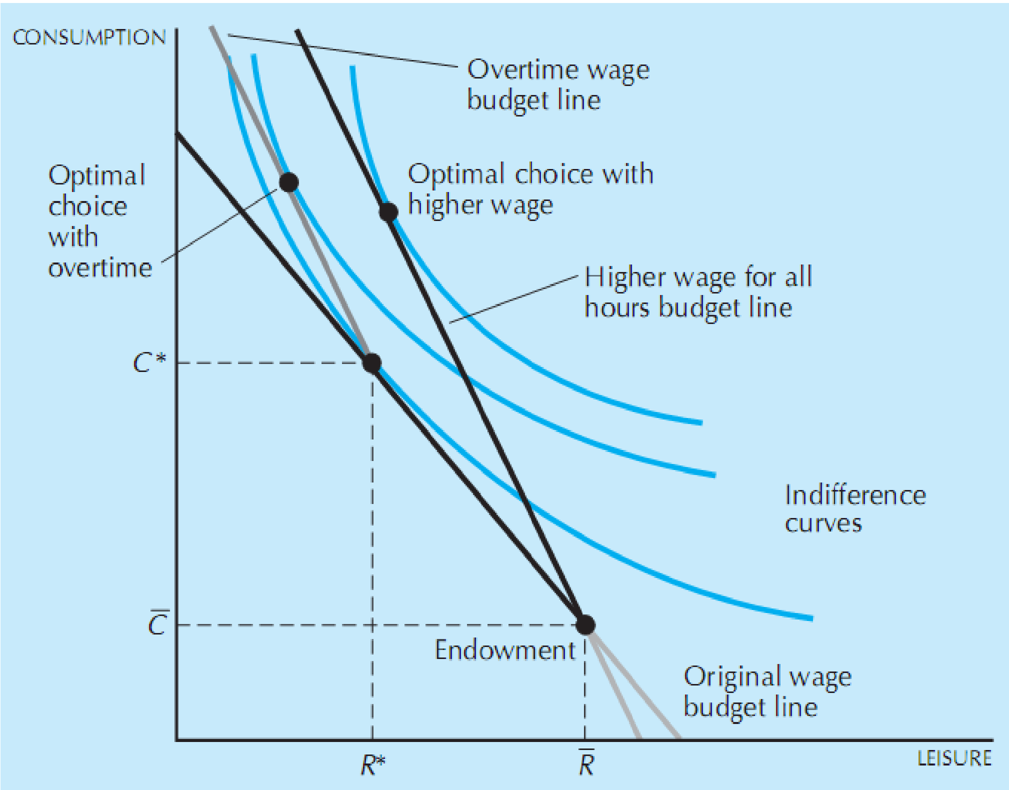


图 9.10：加班费和普通工资增长。加班工资率增加必然导致劳动供给量增加，而直接增加工资率则有可能导致劳动给量减少。

附录

正文中斯勒茨基方程的推导稍微有些武断，即在分析禀赋货币价值变动对需求的影响时，我们武断地认为这种影响等于$\Delta x_1^m / \Delta m$。在老版本的斯勒茨基方程中，它是对的。它表示为了在收入变动后为了还能买的起原来的消费束，需求应变动多大（以比率表示）。但在禀赋价值变动的情形下，$\Delta x_1^m / \Delta m$未必等于需求的变动率。下面我们更细致地进行分析：

令商品 1 的价格从 p_1 变为 p_1'，令 m'' 表示价格为 p_1' 时的货币收入。假设商品 2 的价格固定不变，因此在需求函数中可以省略 p_2 这个变量。

根据 m'' 的定义可知

$$m'' - m = \Delta p_1 \omega_1 .$$

注意，下列式子是个恒等式

$$\begin{aligned}\frac{x_1(p_1', m'') - x_1(p_1, m)}{\Delta p_1} = \\ + \frac{x_1(p_1', m') - x_1(p_1, m)}{\Delta p_1} & \quad \text{(替代效应)} \\ - \frac{x_1(p_1', m') - x_1(p_1', m)}{\Delta p_1} & \quad \text{（普通收入效应）} \\ + \frac{x_1(p_1', m'') - x_1(p_1', m)}{\Delta p_1} & \quad \text{（禀赋收入效应）}\end{aligned}$$

（合并同类项即可知道这是个恒等式.）

根据普通收入效应的定义可知，

$$\Delta p_1 = \frac{m' - m}{x_1}$$

由禀赋收入的定义可知，

$$\Delta p_1 = \frac{m'' - m}{\omega_1} .$$

将这两个式子代入前面的恒等式可得，

$$\begin{aligned}\frac{x_1(p_1', m'') - x_1(p_1, m)}{\Delta p_1} = \\ + \frac{x_1(p_1', m') - x_1(p_1, m)}{\Delta p_1} & \quad \text{(替代效应)} \\ - \frac{x_1(p_1', m') - x_1(p_1', m)}{m' - m} x_1 & \quad \text{（普通收入效应）} \\ + \frac{x_1(p_1', m'') - x_1(p_1', m)}{m'' - m} \omega_1 & \quad \text{（禀赋收入效应）}\end{aligned}$$

曹乾（东南大学 caoqianseu@163.com）

用Δ符号可把上式写得更紧凑些

$$\frac{\Delta x_1}{\Delta p_1}=\frac{\Delta x_1^s}{\Delta p_1}-\frac{\Delta x_1^m}{\Delta m}x_1+\frac{\Delta x_1^m}{\Delta m}\omega_1 .$$

与老版本的斯勒茨基方程相比，上式只有最后一项（$\frac{\Delta x_1^m}{\Delta m}\omega$）是新增的。它是收入变动引起的消费变动，与收入变动的比值然后再乘以商品 1 的**禀赋**。这正是禀赋收入效应。

注意，在上式左侧第二项和第三项的分数，我们都写成了$\Delta x_1^m/\Delta m$。为什么？因为假设价格变动量很小，因此收入变动也很小。这意味着收入从m变为m'以及从m'变为m''时，商品 1 需求的变动率是相同的，于是上式中两种收入效应中的分数是相同的。在这种情形下，合并同类项可将两种收入效应合并成（总）收入效应：

$$\frac{\Delta x_1^m}{\Delta m}(\omega_1-x_1) .$$

于是我们得到了和以前一样的新版本斯勒茨基方程

$$\frac{\Delta x_1}{\Delta p_1}=\frac{\Delta x_1^s}{\Delta p_1}+\frac{\Delta x_1^m}{\Delta m}(\omega_1-x_1) .$$

如果想用微积分的形式表达斯勒茨基方程，只要直接取偏微分便可得到正确的表达式。下面我们来做此事。令$x_1=x_1(p_1,m(p_1))$表示商品 1 的需求函数，其中我们维持商品 2 的价格不变，因此在商品 1 的需求函数中可省略p_2这个变量。另外还要意识到货币收入m取决于商品 1 的价格，即$m(p_1)=p_1\omega_1+p_2\omega_2$。商品 1 需求函数$x_1=x_1(p_1,m(p_1))$的两边同时对变量$x_1$求导可得：

$$\frac{dx_1(p_1,m(p_1))}{dp_1}=\frac{\partial x_1(p_1,m)}{\partial p_1}+\frac{\partial x_1(p_1,m)}{\partial m}\frac{dm(p_1)}{dp_1} \tag{9.5}$$

根据$m(p_1)$的定义即$m(p_1)=p_1\omega_1+p_2\omega_2$，可知当价格变动时收入如何变动：

$$\frac{\partial x_1(p_1,m)}{\partial m}=\omega_1 , \tag{9.6}$$

以前我们曾推导出当商品 1 价格改变但货币收入不变时的斯勒茨基方程：

$$\frac{\partial x_1(p_1,m)}{\partial p_1}=\frac{\partial x_1^s(p_1)}{\partial p_1}-\frac{\partial x_1(p_1,m)}{\partial m}x_1 . \tag{9.7}$$

将（9.6）式和（9.7）式分别代入（9.5）式可得：

$$\frac{dx_1(p_1,m(p_1))}{dp_1}=\frac{\partial x_1^s(p_1)}{\partial p_1}-\frac{\partial x_1(p_1,m)}{\partial m}(\omega_1-x_1) .$$

这就得到了我们想要的新版本的斯勒茨基方程。

总结

1.消费者可以出售禀赋获得收入。

2.某消费者对某商品的总需求是指他对这种商品的最终消费量。净需求则是指他实际购买的该商品的数量。因此，净需求等于总需求减去禀赋。

3.消费者拥有禀赋束情形下，他的预算线的斜率为$-p_1/p_2$，并且经过禀赋束。

4.当一种商品价格改变时，消费者出售商品的价值也改变，因此在斯勒茨基方程中多出了一项收入效应即禀赋收入效应。

5.收入效应与替代效应相互作用的一个有趣例子是消费者的劳动供给决策。由于这两种效应共同作用而且理论上无法说明哪个效应更大，因此工资率上升，劳动供给可能增加也可能减少。

复习题

1.如果某消费者的净需求为（5，-3），他的禀赋束为（4，4），则他的总需求为多少？

2.如果价格$(p_1,p_2)=(2,3)$，消费者当前消费的商品束为$(x_1,x_2)=(4,4)$。假设这两种商品的市场都是完美的，不存在买卖成本。那么，这个消费者是否一定更喜欢消费商品束$(y_1,y_2)=(3,5)$？他是否更喜欢拥有商品束(y_1,y_2)？

3.如果价格$(p_1,p_2)=(2,3)$，消费者当前消费的商品束为$(x_1,x_2)=(4,4)$。现在价格变为$(q_1,q_2)=(2,4)$，消费者的状况是否变得更好些？

4.美国当前石油消费量中约有一半依赖进口，另一半依靠美国国内生产。如果石油价格上升，美国的状况是否变得更好些？

5.假设一天 24 小时神奇地延长为一天 30 小时（这种奇迹恰巧发生在考试周之前），那么预算线有何变化？

6.如果闲暇是劣等商品，那么劳动供给曲线的斜率为正还是负？

复习题答案

1.如果某消费者的净需求为（5，-3），他的禀赋束为（4，4），则他的总需求为多少？

【复习内容】总需求；净需求；禀赋

【参考答案】

消费者对某商品的净需求是指他对该商品的总需求减去他此种商品的初始禀赋数量。因此总需求等于净需求加上禀赋。因此他的总需求为（9,1）。

2.如果价格 $(p_1, p_2) = (2,3)$，消费者当前消费的商品束为 $(x_1, x_2) = (4,4)$。假设这两种商品的市场都是完美的，不存在买卖成本。那么，这个消费者是否一定更喜欢消费商品束 $(y_1, y_2) = (3,5)$？他是否更喜欢拥有商品束 (y_1, y_2)？

【复习内容】显示偏好弱公理；总需求与净需求；总供给与净供给；价格变动对消费者状况的影响

【参考答案】

消费束 $(x_1, x_2) = (4,4)$ 在当前价格 $(p_1, p_2) = (2,3)$ 下的支出为 $20 = 2\times4+3\times4$。商品束 $(y_1, y_2) = (3,5)$ 在当前价格 $(p_1, p_2) = (2,3)$ 下的支出为 21。这种情况下无法应用显示偏好弱公理，即我们无法根据现有信息判断该消费者是否更喜欢消费 (y_1, y_2)，换句话说，你不能认为你买不起某商品束你就更喜欢消费此商品束。

但是该消费者必定更喜欢拥有 (y_1, y_2)，原因在于如果他将该商品束全部售出，他得到 21 元的收入，可以用来购买和消费他更喜欢的商品束。

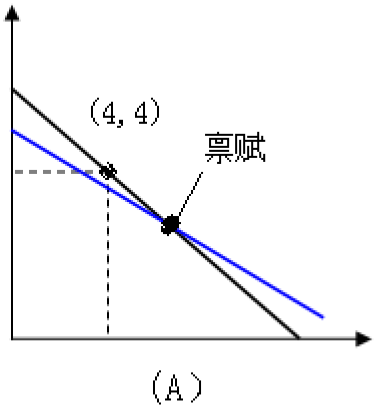
3.如果价格 $(p_1, p_2) = (2,3)$，消费者当前消费的商品束为 $(x_1, x_2) = (4,4)$。现在价格变为 $(q_1, q_2) = (2,4)$，消费者的状况是否变得更好些？

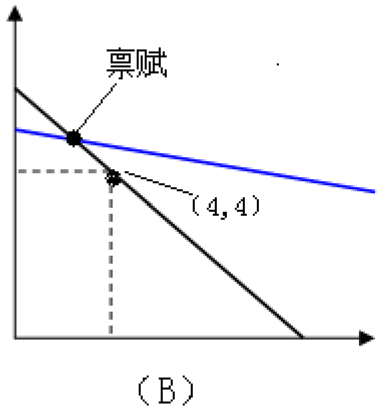
【复习内容】显示偏好弱公理；总需求与净需求；总供给与净供给；价格变动对消费者状况的影响

【参考答案】

消费者的状况是否变得更好取决于他是商品 2 的净购买者还是净销售者。用纵轴表示商品 2，横轴表示商品 1。如下图所示。

(A)图表示该消费者为商品 2 的净购买者，商品 2 价格上升后，预算线绕禀赋旋转到蓝色预算线的位置，由显示偏好弱公理可知，消费束（4,4）至少和禀赋一样好，但商品 2 价格上升后，消费者已经消费不起（4,4）这个消费束，因此该情形下，消费者的状况与原来相比，一般会变差（但也不排除与原来一样好，因为此情形下，消费者仍然可以买得起他的消费束）。【当然如果你假设他的偏好是凸的，则消费束（4,4）一定比禀赋好，由于他是商品 2 的净购买者，当商品 2 价格上升后，他已消费不起（4,4），因此他的状况一定变差。但这样一来，我们就把偏好理论和显示偏好理论扯在一起，我们并不赞成这样做。】





（B）图表示该消费者为商品 2 的净销售者，商品 2 价格上升后，他的状况与原来相比，一般会变好（但不排除与原来一样好）。分析思路与（A）图类似，读者自行补充。

4.美国当前石油消费量中约有一半依赖进口，另一半依靠美国国内生产。如果石油价格上升，美国的状况是否变得更好些？

【复习内容】显示偏好弱公理；总需求与净需求；总供给与净供给；价格变动对消费者状况的影响

【参考答案】

不一定。如果石油价格上升后，美国转变为石油的净出口国（净销售者），则其状况一般会变好，但若价格上升后，美国反而变为石油的净进口国（净购买者），则其状况一般会变差。

此题与上一题实质为同一题。因为你可以将本题中的（美国石油进口量，美国石油国内产量）视为上题的消费束（4，4），把本题的价格（1,1）视为上题中的原来价格（2,3）。因此具体细节请读者自行补充。

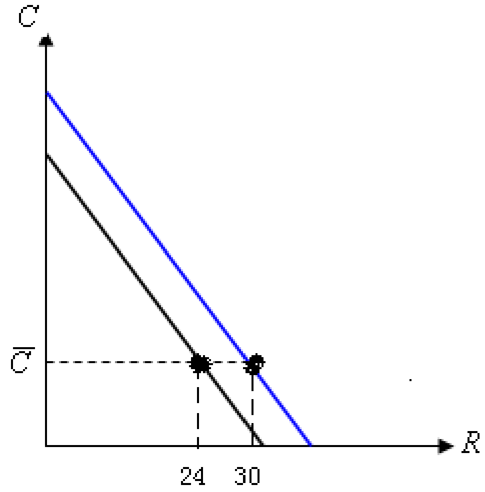
5.假设一天 24 小时神奇地延长为一天 30 小时（这种奇迹恰巧发生在考试周之前），那么预算线有何变化？

【复习内容】禀赋数量变化对预算线的影响

【参考答案】

可以模仿劳动供给决策的预算线。此处你可以粗略地将学习等价于工作，学习的报酬是获得奖学金（类似于工资），这样一路推导下来就可以得到类似教材中的预算线，即：

$$pC + wR = p\overline{C} + w\overline{R}\ .$$



现在$\overline{R}$由 24 增加到 30 小时，预算线会向外移动，由于两商品相对价格不变，因此预算线的这一移动是向外平行移动。这种移动纯粹因为其中一种禀赋数量（时间）变动而引起的，消费品 C 的禀赋未变。请看上图。

6.如果闲暇是劣等商品，那么劳动供给曲线的斜率为正还是负？

【复习内容】劣等商品；劳动供给决策；劳动供给的比较静态分析

【参考答案】

如果闲暇是劣等商品，则劳动供给曲线的斜率为正。分析如下。

由于劣等商品在收入上升时需求量会下降。现在假设工资率上升。我们使用斯勒茨基方程进行分析：

$$\frac{\Delta R}{\Delta w} = \underset{(-)}{\text{替代效应}} + \underset{(+)}{(\overline{R}-R)}\underset{(-)}{\frac{\Delta R}{\Delta m}}. \tag{9.4}$$

在（9.4）中：替代效应为负（因为它总是与价格变动方向相反）；$(\overline{R}-R)$为正（因闲暇时间 R 不会超过上限$\overline{R}$），但$\Delta R/\Delta m$为负，因为我们假设闲暇是劣等商品。这样，收入效应也为负。因此，总效应的符号为负。也就是说，工资率（闲暇的机会成本）上升时，闲暇的需求会下降，从而劳动供给增加，所以劳动供给曲线斜率为正。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

10.跨期选择（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

10 跨期选择

本章我们继续研究消费者理论。以前各章研究的消费选择都是在同一时期内发生的，本章研究的情形是多个时期内的消费选择问题，这样的选择称为**跨期选择**（intertemporal choices），跨期消费选择不仅涉及消费还涉及到储蓄问题。

10.1 预算约束

假设有两个时期，某个消费者在每个时期内都要作出消费商品数量的选择。我们通常假设他消费的商品为复合商品(composite good)，我们在第 2 章已介绍过这个概念。当然，你也可以假设这种商品为任何具体的商品（比如苹果）。令消费者在这两个时期消费的商品数量为(c_1, c_2)，假设商品价格恒为 1。他在两个时期内的收入为(m_1, m_2)。

消费者可以将时期 1 的资金转移到时期 2，假设最初消费者转移资金的方法是储蓄但不计利息，而且假设他无法借钱，因此在时期 1 他可以消费的资金最多为m_1。这样的情形下，他的预算线的形状将如图 10.1 所示。

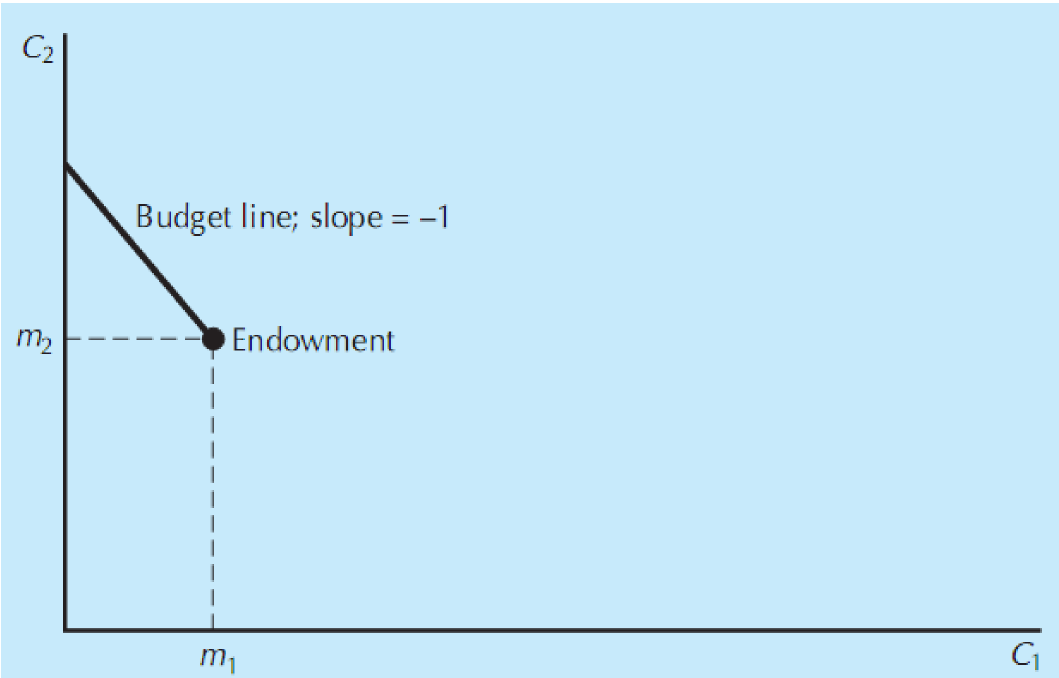


图 10.1：预算约束。利息率为 0 且不允许借钱的情形下，消费者的预算线如图所示。他在时期 1 的消费越少，在时期 2 他能消费的数量越多。

这种情形下，消费者有两种消费选择：一是他可以选择消费(m_1, m_2)，即每期消费支出等于每期的钱数；另外一种选择是，时期 1 他的消费支出小于m_1，这样他储蓄了部分资金用于时期 2 的消费。

现在，允许消费者借入或贷出资金，利率为r。为方便起见，仍然假设商品价格在两个时期内恒为 1，现在我们推导出预算约束。假设消费者在时期 1 决定做储蓄者，因此$c_1 < m_1$。在这种情形下，他储蓄的本金为$m_1 - c_1$，因利率为r，下一期他的消费量为：

$$\begin{aligned} c_2 &= m_2 + (m_1 - c_1) + r(m_1 - c_1) \\ &= m_2 + (1+r)(m_1 - c_1) \end{aligned} \tag{10.1}$$

这个式子表明，消费者在时期 2 的消费量，等于他在时期 2 的收入加上他在时期 1 的储蓄本金及其利息。

现在假设消费者是借款者（资金借入者），因此他在时期 1 的消费大于时期 1 的收入。如果$c_1 > m_1$，他就是个借款者，在时期 2 他需要支付的利息为$r(c_1 - m_1)$。当然，他还需要归还借入的本金$(c_1 - m_1)$。这样，他的预算约束为

$$\begin{aligned} c_2 &= m_2 - r(c_1 - m_1) - r(c_1 - m_1) \\ &= m_2 + (1+r)(m_1 - c_1) \end{aligned}$$

这个式子正是我们前面得到的式子。如果$m_1 - c_1$为正，则消费者储蓄并赚取利息；如果$m_1 - c_1$为负，则消费者借入资金并支付利息。

如果$c_1 = m_1$，则必有$c_2 = m_2$，此时消费者既不借债也不放债。这种情形下，消费处于“波洛厄斯点（Polonius point）”[一]。

我们可以将消费者的上述预算约束变形从而得到其他两种形式的预算约束：

$$(1+r)c_1 + c_2 = (1+r)m_1 + m_2 \tag{10.2}$$

$$c_1 + \frac{c_2}{1+r} = m_1 + \frac{m_2}{1+r} \tag{10.3}$$

注意，上述两个式子都意味着

$$p_1x_1 + p_2x_2 = p_1m_1 + p_2m_2 .$$

在（10.2）式中，$p_1 = 1 + r$ 和 $p_2 = 1$；在（10.3）式中，$p_1 = 1$ 和 $p_2 = 1/(1+r)$。

我们说（10.2）式是以**终值**（future value）形式表达的预算约束，（10.3）式是以**现值**（present value）形式表达的预算约束。如此称呼的原因在于，（10.2）式中以远期消费的价格为 1，而在（10.3）式中以现期消费的价格为 1。换句话说，（10.2）式中时期 1 的价格是对于时期 2 价格的相对价格，（10.3）式正好相反。

现值和终值的几何图形解释见图 10.2. 两个时期资金禀赋的现值，是恰好能和资金禀赋产生同一预算集的时期 1 的资金数。这恰好是预算线的横截距，它表示时期 1 的最大可能消费量。由预算线可以计算出，这个最大消费量为$\bar{c}_1 = m_1 + m_2/(1+r)$，这是资金禀赋的现值。

[一]“不要借债，也不要放债；因为放债不仅血本无归而且会失去朋友，借债会使自己忘记勤俭。”（莎士比亚）哈姆雷特，第 1 幕第 3 场，波洛厄斯告诫儿子的话。

类似地，纵截距是时期 2 的最大可能消费量（$c_1=0$ 时）。从预算线也可以计算出它等于资金禀赋的终值 $\bar{c}_2=(1+r)m_1+m_2$。

跨期选择的预算约束表达式中，现值形式的表达式相对重要一些，因为它将终值折算为现值，而这正是我们看待问题的角度。

从终值形式的预算约束表达式也可以推导出现值形式的预算约束的表达式：（10.2）式两边同除以 $(1+r)$ 即可。既然消费者总能买得起 (m_1,m_2)，所以预算线通过 (m_1,m_2)；预算线的斜率为 $-(1+r)$。

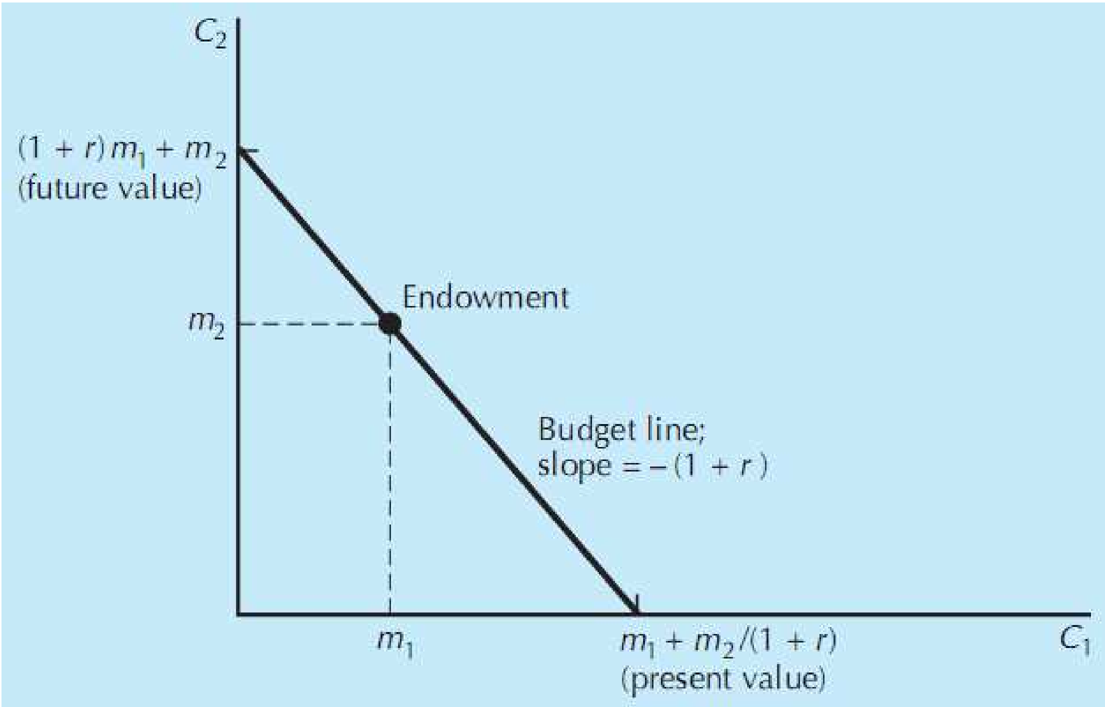


图 10.2：现值与终值。纵截距表示终值，横截距表示现值。

10.2 消费偏好

我们用无差异曲线分析消费者的偏好。无差异曲线的形状表明消费者对不同时期消费的偏好。例如，如果无差异曲线的斜率恒为 −1，则表明消费者对于是在今天消费还是在时期消费 2 消费无所谓。他在这两个时期的边际替代率为 −1。

1：1 完全互补情形的无差异曲线，表明 c_1 和 c_2 是 1:1 互补的，即消费者在今天和明天的消费数量是一样的。这种情形下，消费者不愿意用一期的消费替代另一期。

和往常一样，良好性状的偏好更为合理一些。消费者愿意用今天的消费替代明天或者用明天的消费代替今天的消费，替代率大小取决于消费者需求函数的具体形式。

在上述情形下，偏好为凸就比较自然，因为这意味着消费者更喜欢两个时期的“平均”

消费量，而不是今天多消费明天一点也不消费。

10.3 比较静态分析

给定消费者在两个时期内的预算约束和消费偏好，我们就可以研究他的最优选择(c_1,c_2)。如果他的消费满足$c_1<m_1$，则他是个**放款者**（lender），相反如果$c_1>m_1$，则他是个**借款者**（borrower）。图 10.3B 显示了消费者为放款者的情形。

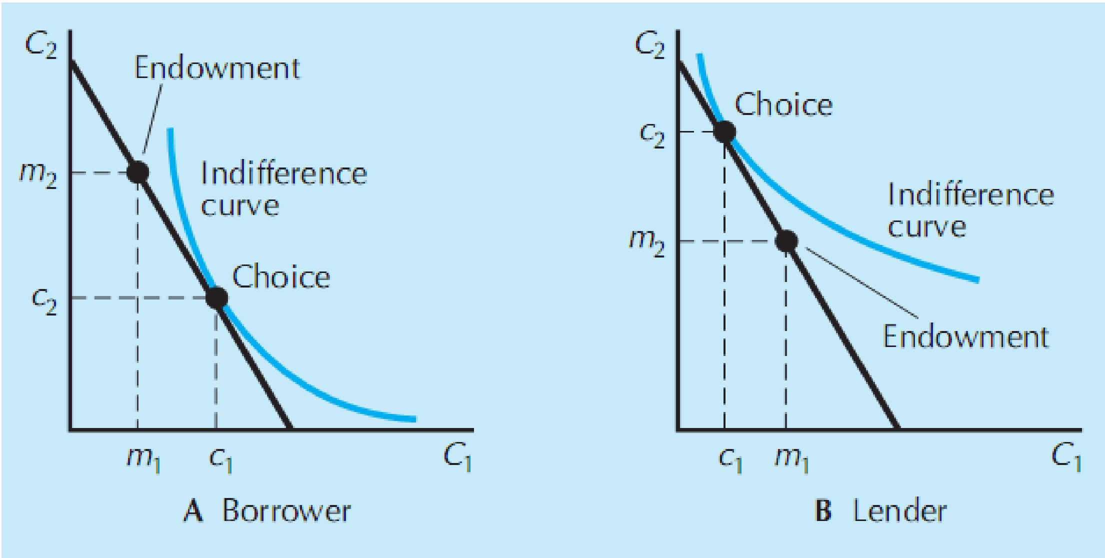


图 10.3：**借款者和放款者**。A 图显示的为借款者的情形，因为$c_1>m_1$；B 图显示的为放款者的情形，因为$c_1<m_1$。

下面分析消费者会对利率变化作出如何反应。从（10.1）式可知，利率增加会使无差异曲线转动到更陡峭的位置：因为利率增加时，对于给定的c_1减少量，消费者在下一期得到的消费量会更多。由于消费者总能买得起禀赋，因此这种无差异曲线的变动实际为绕着禀赋点转动。

我们也可以分析当利率变动时，消费者选择作为借款者或放款者的决策是否会变动。这样的情形有两种，到底为哪一种则要取决于消费者最初为借款者还是放款者。先假定消费者最初为放款者，则可以证明当利率增加时，消费者仍然会选择作为放款者。

我们借助图 10.4 说明这个结论。如果消费者最初为放款者，则他的消费束必然位于禀赋点的左侧。现在假设利率增加，那么消费者可能在禀赋点**右侧**消费吗？

答案为否，因为如果是这样，则违反了显示偏好原理：在原来预算集的情形下，消费者原本可以选择禀赋点右侧的消费束，但是他却放弃了；他选择的消费束是如图所示的消费束。既然在新预算线下，原来的消费束仍然可以买得起，新的最优消费束必然位于原预算集**之外**。这意味着新的最优选择点必然位于禀赋点的左侧。因此，当利率增加时，消费者必然

曹乾（东南大学 caoqianseu@163.com）

选择仍然作为放款者。

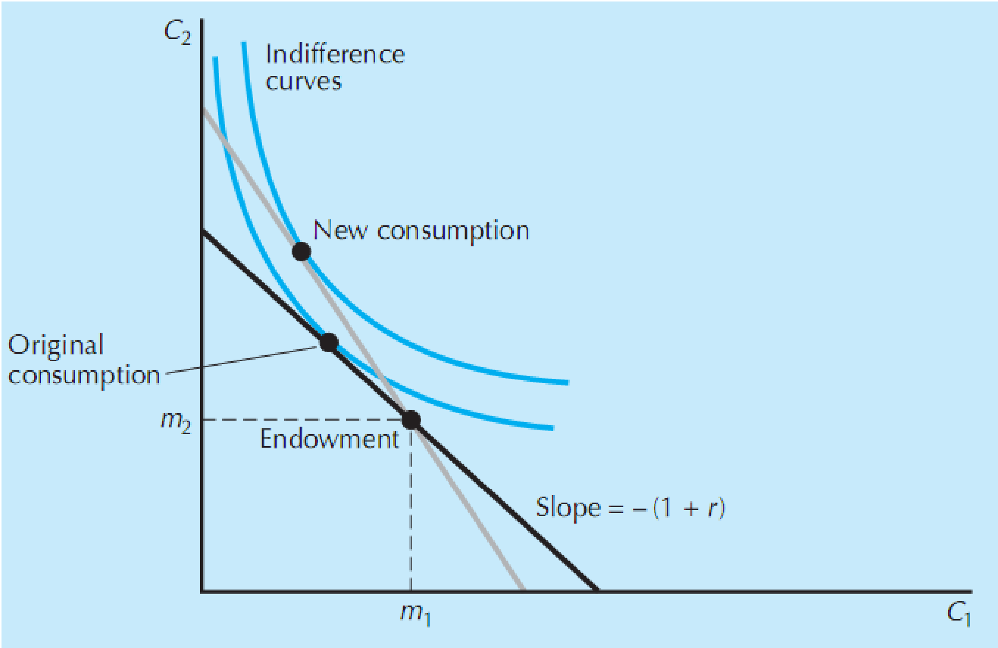


图 10.4：若某消费者是个放款者，在利率增加时，他必然选择继续作为放款者。利率增加时，预算线绕着禀赋点转动到更陡峭的位置；显示偏好原理表明新的消费束必然位于禀赋点的左侧。

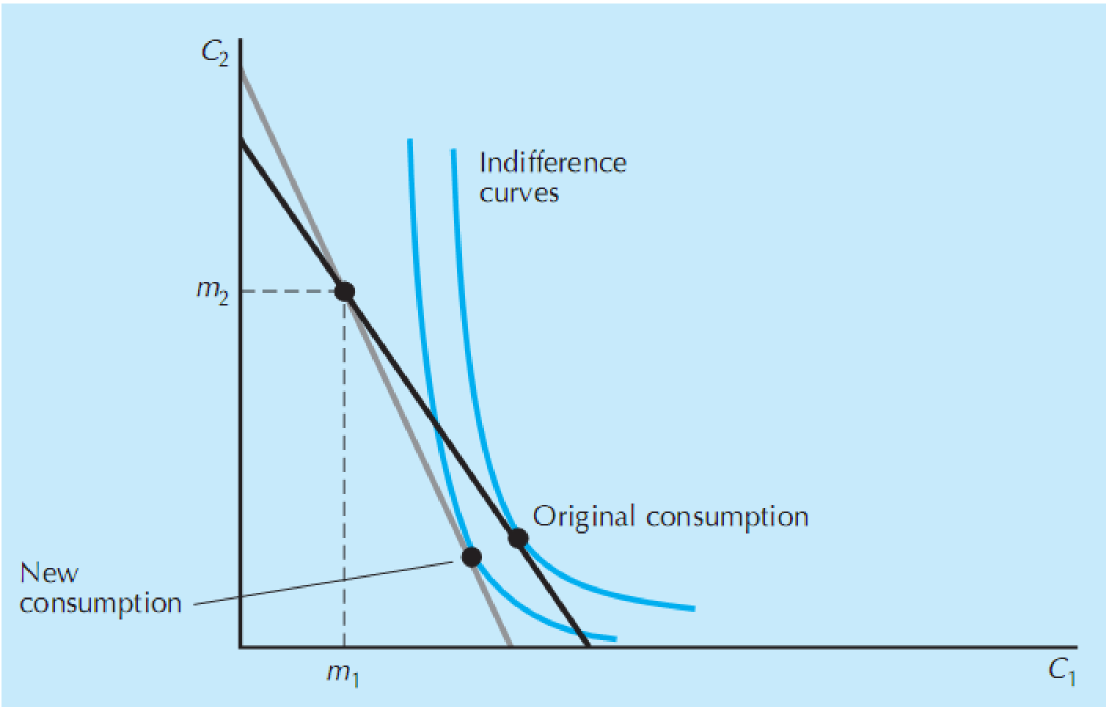


图 10.5：什么样的情形下，当利率增加时借款者的状况会变差？某借款者在利率增加后如果继续选择作为借款者，则他的状况会变差。

类似地，如果消费者最初为借款者，利率降低时，他必然选择继续借款。（请模仿图 10.4 画图分析一下此情形。）

因此，如果消费者最初为放款者，利率增加时，他会继续放款。如果消费者最初为借款者，利率降低时，他会择继续借款。另一方面，如果某消费者最初为放款者，利率降低时，他可能决定转变为借款者；类似地，利率增加可能诱使借款者变为放款者。但后面这两种情形无法使用显示偏好理论进行分析。

我们还可以使用显示偏好理论，分析当利率变化时消费者的福利变化。如果消费者最初为借款者，利率增加时，他决定继续作为借款者，则他的状况变差。可以用图 10.5 说明这个结论：如果消费者继续作为借款者，那么他的选择点必然位于原预算集内，而这样的选择点都是曾经被他放弃过的，这意味着他的状况必定变差。

10.4 斯勒茨基方程与跨期选择

我们可以象第 9 章一样，使用斯勒茨基方程，把由利率变动引起的需求变动分解为收入效应和替代效应。利率变动对每个时期的消费有何影响？

这种情形下使用终值形式的预算线进行分析会更容易一些。在这种形式的预算约束中，利率上升等价于时期 1 的消费价格提高了（与时期 2 的消费价格相比较）。写出斯勒茨基方程可得：

$$\begin{array}{cccc} \dfrac{\Delta c_1^t}{\Delta p_1} = & \dfrac{\Delta c_1^s}{\Delta p_1} + & (m_1 - c_1) & \dfrac{\Delta c_1^m}{\Delta m} . \\ (?) & (-) & (?) & (+) \end{array}$$

和以前一样，替代效应总是与价格变动方向相反。在该情形下，时期 1 的消费价格上升，因此替代效应表明消费者应该减少时期 1 的消费。所以在上式中，我们在替代效应下面标记了一个负号。假设消费品为正常商品，因此上式的最后一项（收入变动时消费如何变动）的符号为正。由此可见，整个表达式即总效应的符号取决于 $(m_1 - c_1)$ 的符号。如果消费者为借款者，则该项为负，从而总效应的符号必定为负。因为对于借款者来说，利率增加时他必然减少时期 1 的消费。

如果消费者为借款者，为什么利率上升的总效应的符号为负？当利率增加时，替代效应会使消费者减少时期 1 的消费。若该消费者为借款者，利率增加后意味着他在时期 2 偿还的利息增加。这种效应促使他减少借入的钱数，因此他在时期 1 的消费量会减少。

但是，如果消费者为放款者，则总效应的符号不明朗。总效应等于负的净效应与正的收入效应之和。从放款者的角度看，利率上升会使他的收入增加，因此时期 1 的消费量可能增加。

利率变动的效应并不神秘。和其他商品价格变动一样，利率（即货币的价格）变动也

曹乾（东南大学 caoqianseu@163.com）

会引起收入效应和替代效应。但是如果不借助斯勒茨基方程，我们就无法将各种效应隔离开，从而无法分解各种效应引起的变化。正是有了这个工具，我们的分析才如此简单。

10.5 通货膨胀

在以上的分析中，我们用于分析的商品没有指明具体名称，而是统一用“消费品”表示。我们断言，在时期 1 你放弃 Δc 单位的消费品，在时期 2 你可以购买 $(1+r)\Delta c$ 单位。但这种断言暗含着消费品“价格”固定不变的假设，即不存在通货膨胀或通货贬值。

将上述分析稍加变动即可用于分析通货膨胀的情形。现在假设消费品在不同时期的价格不同。为简单起见，通常将时期 1 的价格定为 1，而用 p_2 表示时期 2 的价格。同样出于方便的原因，通常用消费品的单位数衡量禀赋，所以禀赋在时期 2 的货币价值为 p_2m_2。因此，消费者在时期 2 可用于消费的钱数由下式给出

$$p_2c_2 = p_2m_2 + (1+r)(m_1 - c_1),$$

由上式解出 c_2 即可得到时期 2 的消费量

$$c_2 = m_2 + \frac{(1+r)}{p_2}(m_1 - c_1).$$

注意，这个式子非常类似我们前面给出的式子，不同之处在于此处我们用 $(1+r)/p_2$ 代替了 $1+r$。

下面我们将预算约束用通货膨胀率表达。通货膨胀率 π，是指价格上涨的比率。由于在前面我们令 $p_1 = 1$，我们有

$$p_2 = 1+\pi,$$

将上式代入时期 2 消费量的表达式，可得

$$c_2 = m_2 + \frac{1+r}{1+\pi}(m_1 - c_1)。$$

我们引入一个新的变量 ρ，用它表示**实际利率**（real rate of interest），将其定义为[一]

$$1+\rho = \frac{1+r}{1+\pi}$$

因此，预算约束变为

$$c_2 = m_2 + (1+\rho)(m_1 - c_1).$$

$1+\rho$ 表示如果你在时期 1 放弃一些**消费**，在时期 2 你可以得到多少额外的**消费**。这正是它

[一] ρ 为希腊字母，读作“row”.

为什么叫做实际利率的原因：它衡量你能得到多少额外的消费，而不是你能得到多少钱。

钱的利率称为**名义利率**（nominal rate of interest）。我们已知道，名义利率和实际利率的关系为

$$1+\rho=\frac{1+r}{1+\pi}$$

从上式可得到ρ的显性表达式

$$\rho=\frac{1+r}{1+\pi}-1=\frac{r-\pi}{1+\pi}.$$

从这个式子可计算出实际利率的确切数值，但是使用近似值更方便。如果通货膨胀率不大，上式的分母即$1+\pi$略微比 1 大。因此可用下式估计实际利率的近似值

$$\rho\approx r-\pi,$$

这就是说实际利率近似等于名义利率减去通货膨胀率。（符号≈表示“约等于”。）这个近似是合理的：如果名义利率等于 18%，但价格上涨了 10%，则实际利率（在当期放弃一些消费后在下一期能得到的额外消费量）大致等于 8%。

当然在制定消费计划时，我们通常关心未来的消费。一般来说，我们能知道下一期的名义利率，但不知道下一期的通货膨胀率。因此，实际利率通常用当前利率减去**预期**（expected）通货膨胀率来计算。由于人们对下一期的通货膨胀率的预期不同，他们估算出的实际利率不同。如果能够合理地预测通货膨胀率，那么他们估算出的实际利率不会存在大的差异。

10.6 现值：进一步研究

现在我们继续分析本章第一节中预算约束的两个表达式，即（10.2）式和（10.3）式：

$$(1+r)c_1+c_2=(1+r)m_1+m_2$$

$$c_1+\frac{c_2}{1+r}=m_1+\frac{m_2}{1+r}$$

请看上面两个式子的右端。我们说过，第一个式子中禀赋的价值是用终值衡量的，第二个式子则用现值衡量。

我们先分析终值的概念。如果借款和放款的利率都为r，今天的 1 元钱在下一期的终值是多少？答案是$(1+r)$元。即：利率为r时我们将 1 元钱存入银行，则在下一期可以得到$(1+r)$元。换句话说，下一期的$(1+r)$元等价于今天的 1 元，因为如果你今天借入 1 元钱，在下一期你要偿还$(1+r)$元。如果将下一期的 1 元作为计价物，则今天 1 元的价格即为$(1+r)$元。这一点可从上面第一个预算约束表达式看出，这个式子是用终值作为计价物，将

即第二期货币的价格定为 1，从而第一期货币的价格为$(1+r)$。

接下来我们分析现值。现值正好相反：一切东西的价值都用今天的钱来衡量。下一期的 1 元在今天值多少钱？答案为$1/(1+r)$元。这是因为利率为r时将$1/(1+r)$元钱存入银行，在下一期就可得到 1 元。下一期兑现的 1 元钱的**现值**为$1/(1+r)$元。

现值这个概念，让我们得到了跨期消费预算线的另外一种表达方法：**如果在某消费方案方案下，消费品的现值等于收入的现值，则该方案是可行的**。

现值的概念有一个重要的应用价值，这个应用价值与第 9 章中的下列结论密切相关：如果消费者能以不变价格自由买和卖，则消费者总是偏好价值更高的禀赋。在跨期选择的情形下，这个结论意味着：**如果消费者能以不变利率自由借和贷，则消费者总是偏好现值更高的收入方式**。

这个结论成立的原因和第 9 章中的上述结论成立的原因是一样的：禀赋价值增加后预算线向外移动。新的预算集包含原预算集，这就是说从图形上看，新预算集等于原预算集加上一块新的区域。经济学家有时会说现值高的禀赋**比**（dominate）现值低的禀赋**好**，因为在**任一**时期，销售现值高的禀赋所得的收入，都比销售现值低的禀赋所得收入高，从而消费者在前一种情形下的消费数量更大。

当然，如果一个禀赋的现值比另一个高，则终值也比另一个高。然而，使用现值衡量跨期消费情形下的货币禀赋的购买力更为方便，因此我们主要研究用现值衡量的情形。

10.7 若干时期的现值分析

下面我们分析一个三段时期的模型。假设在这三个时期，资金借和贷的利率都恒为r。因此，时期 2 的消费，若用时期 1 的消费衡量，则为$1/(1+r)$。

时期 3 的消费的价格是什么？如果我投资 1 元钱，则下一期将增长到$(1+r)$元；我把这笔钱再投资，第三期它将增长到$(1+r)^2$元。因此，如果今天投资$1/(1+r)^2$元，则在时期 3 可得到 1 元钱。因此，相对于时期 1 的消费，时期 3 的消费的价格为$1/(1+r)^2$。时期 3 消费 1 元，相当于在时期 1 消费$1/(1+r)^2$元。这意味着预算约束的表达式为

$$c_1+\frac{c_2}{1+r}+\frac{c_3}{(1+r)^2}=m_1+\frac{m_2}{1+r}+\frac{m_3}{(1+r)^2}.$$

这个式子很像我们前面已经介绍过的预算线，其中时期 t 的消费的价格若用今天的消费衡量，等于

$$p_t=\frac{1}{(1+r)^{t-1}}.$$

和以前一样，现值越高的禀赋，消费者越喜欢，因为禀赋价值增高后，预算线会向外移动。

我们是在利率恒定不变的假设前提下推导出上述预算线的表达式，但很容易将它推广到利率可变的情形。例如，假设时期 1 到时期 2 的利息率为 r_1，时期 2 到时期 3 的利息率为 r_2，因此时期 1 储蓄 1 元在时期 3 可以得到 $(1+r_1)(1+r_2)$ 元。因此，时期 3 的 1 元的现值为 $1/(1+r_1)(1+r_2)$ 元。这意味着预算线的表达式为

$$c_1+\frac{c_2}{1+r_1}+\frac{c_3}{(1+r_1)(1+r_2)}=m_1+\frac{m_2}{1+r_1}+\frac{m_3}{(1+r_1)(1+r_2)}.$$

处理这个式子并不难，但我们一般更乐意处理利率为既定常数的情形。

表 10 列出了未来第 T 年的 1 元钱在不同利率下的现值。你应该注意到了，随着利率“合理”增加，现值下降的速度非常快。例如，利率为 10%时，未来第 20 年的 1 元钱只相当于今天的 0.15 元。

表 10：**未来第 T 年 1 元钱的现值**

Rate	1	2	5	10	15	20	25	30
.05	.95	.91	.78	.61	.48	.37	.30	.23
.10	.91	.83	.62	.39	.24	.15	.09	.06
.15	.87	.76	.50	.25	.12	.06	.03	.02
.20	.83	.69	.40	.16	.06	.03	.01	.00

10.8 现值的用途

我们先给出一个重要的一般原理：**将未来货币流折算为今天现金的唯一正确方法就是使用现值**。这个原理可从现值的定义直接推知：现值衡量消费者货币禀赋的价值。只要消费者能以既定利率自由借和贷，那么禀赋的现值越高意味着各期的消费量也**越多**。不论你在各期的消费偏好如何，你都肯定喜欢现值更高的货币流，因为这会使得你在各期的消费量更多。

上述结论可用图 10.6 说明。在此图中，消费束 (m_1',m_2') 比原来的禀赋 (m_1,m_2) 差，因为它位于通过 (m_1,m_2) 的那条无差异曲线的下方。然而，如果消费者能以利率 r 借和贷，则他会偏好 (m_1',m_2') 胜于 (m_1,m_2)。原因在于，拥有禀赋 (m_1',m_2') 时他能够消费得起 (c_1,c_2)，(c_1,c_2) 显然比他当前的消费束好。

现值在评估不同类型投资产生的收入流时，非常有用。如果你想比较两个不同的投资项目，但它们产生的收入流不同，怎么知道哪个项目更好些？方法很简单，计算这两个项目的现值，选择现值大的那个项目即可。项目的现值越大，你越能消费更多的商品。

有时有必要以分期付款（即支付流）的形式购买收入流。例如，例如，你购买公寓时可以向银行抵借款，在随后几年内分期偿还。假设可用支付流 (P_1,P_2) 购买 (M_1,M_2) 这个收入流。为评估上述投资是否可行，我们比较这两个货币流的现值即可。如果

$$M_1 + \frac{M_2}{1+r} > P_1 + \frac{P_2}{1+r} \quad , \tag{10.4}$$

即如果收入流的现值大于支付流的现值，则投资可行，因为它可以增加我们禀赋的现值。

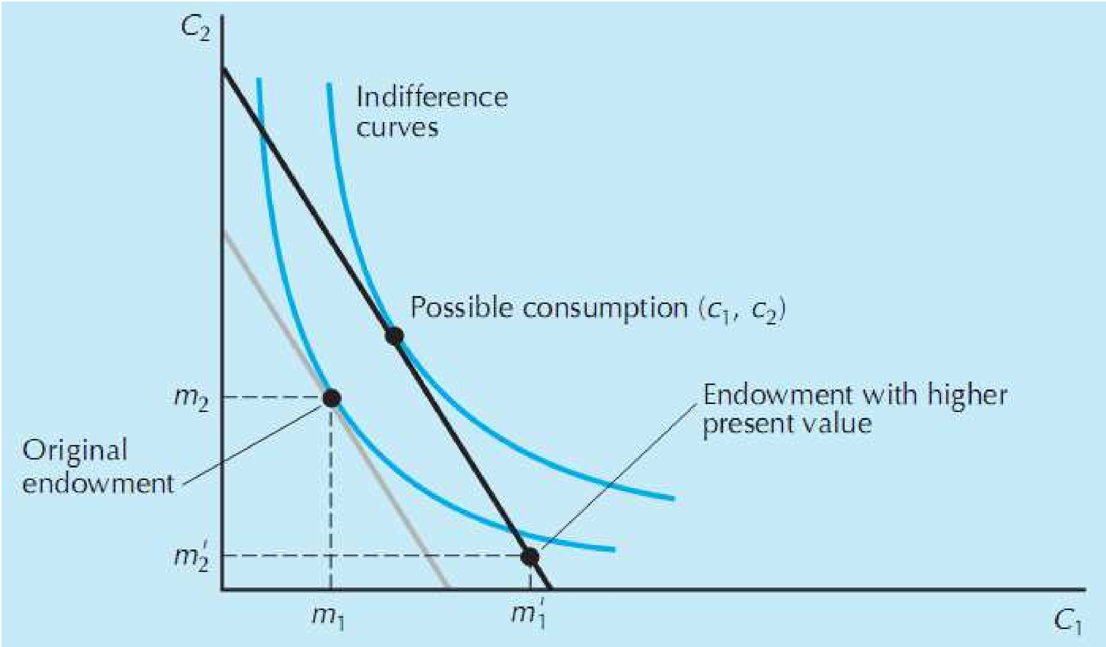


图 10.6：更高的现值。若消费者能按市场利率借和贷，则禀赋的现值越高，意味着他的消费可能性越大。

评估投资可行性时，还可以使用**净现值**（net present value）法，这种方法和现值方法的思想是一致的。为了计算净现值，我们先计算每个时期的**净现金流**，然后将其折算为现值。在本例中，净现金流为$(M_1 - P_1, M_2 - P_2)$，将其折算为现值可得净现值

$$NPV = M_1 - P_1 + \frac{M_2 - P_2}{1+r}.$$

例子：评估收入流

假设有两个投资项目 A 和 B。A 项目今年和明年产生的收入分别为 100 元、200 元；B 项目的今年和明年的收入分别为 0 元和 310 元。为简单起见，假设这两个项目的成本一样大。哪个项目更好一些？

答案取决于利率。如果利率为 0，将两年的收入相加即可得到答案。因为利率为 0 时，现值的计算可以归结为将收入相加。

利率为 0 时，A 和 B 收入的现值分别为 300 元和 310 元，显然 B 更好。但若利率比较

高，我们可能得到相反的答案。例如假设利息为 20%，则两个项目收入的现值分别为：

$$PV_A = 100 + \frac{200}{1.20} = 266.67$$

$$PV_A = 0 + \frac{310}{1.20} = 258.33.$$

现在 A 项目更好。由于 A 项目早期收入比较大，这意味着当利率较高时它的现值较大。

例子：信用卡的真实成本

使用信用卡借款代价昂贵：很多公司索要 15%-21%的利率。但是如果你知道它们是怎么计算借款费用的，你就会知道信用卡借款的真实利息率要远高于它们使用的利率。

假设在某月的第一天，某消费者使用信用卡购买了 2000 元的东西，假设利率为 1.5%。如果该月末，他全部还清了 2000 元，则不需要缴纳利息。如果他一分钱也没还，在下月月初他就需要缴纳 $2000 \times 0.015 = 30$ 元的利息。

如果消费者在借款当月的月末偿还了 1800 元，利息为多少？在本例中，消费者相当于只借了 200 元，因此利息为 3 元。然而很多信用卡公司索要的利息远高于 3 元。原因在于，很多信用卡公司是根据“月度平均欠款额”计算利息的，即使你偿还了部分欠款。在本例中，月度平均欠款额大约为 2000 元（$2000 \approx \frac{2000 \times 30 + 200 \times 1}{31} = 1942$）。因此尽管消费者只借款 200 元，但需要交纳大约 30 元的利息，这相等于月利率为 15%！

例子：延长版权保护期

美国议会根据宪法第 8 节第 1 款的规定授予相关人员专利和版权。这一款是这么说的：“为了促进科学和有用文艺作品的进展，作者和发明者对其作品和发明享有一定期限的独占权。”

但是，“一定期限”是什么意思？在美国，专利的保护期为 20 年，但版权的保护期要长得多。

美国议会于 1970 年通过了第一个版权法，此法规定版权的保护期为 14 年，期满后可以再延期 14 年。1831 年，版权保护期变为 28 年；在此基础上，1909 年议会规定，版权所有人有权选择是否再延期 28 年。1962 年，保护期变为 47 年；1978 年，变为 67 年。1967 年，版权保护期的定义为作者的寿命加上 50 年，若版权的使用权属于作者的雇主所有，则保护期为作者的寿命加上 75 年。1998 年索尼.博诺（Sonny Bono）版权保护延期法对保护期的定义为：作者的寿命加上 70 年；若版权的使用权属于作者的雇主所有，则此保护期为作者的寿命加上 75-95 年。

人们也许会质疑“作者的寿命加上 70 年”的保护期是不是太长了。你可能会问：1998 年版权延期法对作家创造作品有什么**额外的**激励（additional incentive）?

我们看一个简单的例子。假设利率为 7%。那么将保护期从 80 年增加到 100 年带来的现值增量，只占头 80 年现值的 0.33%左右。额外的 20 年对授予版权时的现值几乎没有影响，因为这些年份距离授予版权时点太远了。因此，延长版权保护期对作家创造作品最初激励几乎没有**额外的**影响。

既然如此，为何有人要游说议会改变版权的保护期呢？答案在于 1998 年的法案规定版权的延期是**追溯性的**（retroactive），这就使得保护期即将结束的作品又获得新的租赁期限[一]。

例如，民间盛传迪斯尼的游说最为激烈，因为第一部米奇老鼠电影《汽船威利》（Steamboat Willie）的保护期即将结束[二]。

这类追溯性的版权延期对于作者们来说毫无经济意义，因为真正重要的激励是他们创造作品时的激励。如果不存在这样的追溯性延期，任何人都懒得游说议会要求版权延期，因为额外保护期的经济价值很小。

10.9 债券

证券（securities）是承诺某种偿还方式的金融工具。由于人们想要的偿还方式很多，金融工具的种类也很多。人们在金融市场中买卖证券，相当于他们用某一时期的现金流交换另一时期的现金流。这些现金流通常用于某时期或另一时期的消费。

此处我们只分析一种特别的证券即**债券**（bond）。政府和公司都可发行债券。债券的本质是借钱的一种方式。借钱方，也就是债券发行者，承诺每期偿还既定的 x 元（**息票利息** coupon），直到某个时期 T（**到期日** maturity date），在到期日借钱方向债券持有人归还 F 元（债券**票面价值** face value）。

因此债券的支付流是 $(x, x, x, \ldots F)$。如果利率固定不变，债券的现值容易计算。计算公式为

$$PV = \frac{x}{1+r} + \frac{x}{(1+r)^2} + \ldots + \frac{F}{(1+r)^T}.$$

注意，如果利率上升债券的现值会下降。为什么？因为利率上升时，未来 1 元的现值将减少，债券支付流的现值也因此减少。

债券市场规模较大而且比较成熟。流通债券的市场价值将随利率的波动而波动，因为债

[一] 这对作品版权使用权属于作者的雇主情形有显著意义，因为在这种情形下，雇主可以继续享有独占的使用权。译者注。

[二] 这部电影于 1928 年上演，按 80 年保护期计算，保护期于 2008 年结束。延长 20 年的好处是显然的。译者注。

券带票的支付流的现值改变了。

一种非常有趣的债券是永久付息的债券。这样的债券称为统一公债（consols）或永续年金（perpetuities）。怎样计算此类债券的现值？假设该债券承诺永久性支付年息 x 元。计算它的现值需要计算下列无穷级数：

$$PV = \frac{x}{1+r} + \frac{x}{(1+r)^2} + ...$$

计算的技巧是把公因子 $1/(1+r)$ 提取出来，从而得到

$$PV = \frac{1}{1+r}[x + \frac{x}{(1+r)} + \frac{x}{(1+r)^2} ...].$$

注意到方括号内的项恰好等于 x 加上现值！将现值 PV 代入可解得 PV：

$$PV = \frac{1}{1+r}[x + PV] \Rightarrow PV = \frac{x}{r}.$$

这种方法并不难，但还有一种更简单的方法可以理解得到答案。如果利率恒为 r 而且你想永远获得 x 元，你需要投入多少钱（V）？你立即可写出方程

$$Vr = x,$$

这个式子是说 V 的利息一定等于 x。但这样一来，这样一笔投资的价值为

$$V = \frac{x}{r}.$$

因此，承诺永久支付 x 元的统一公债的现值必然等于 x/r。

对于统一公债可以直接看出：利率上升时债券价值下降。例如，假设统一公债发行时的利率为 10%。因此，如果它承诺永久支付 10 元的年息，它的现值为 100 元，因为 100 元可以产生 10 元的年息收入。

现在假设利率上升为 20%，上述债券的价值必然下降为 50 元，因为利率为 20%时，只要投入 50 元就可以获得 10 元的年息。

统一公债价值的计算公式可用于近似计算长期债券的价值。例如如果利率为 10%，30 年后的 1 元钱的现值约为 6 分钱（$\approx \frac{1}{(1+10\%)^{30}}$）。对于我们通常面临的利率水平来说，30 年已经够长了。

例子：分期偿还的贷款

假设你借了 1000 元并且承诺分 12 个月偿还，每月偿还 100 元。你支付的利率为多大？

乍看起来，似乎利率为 20%：你借款 1000 元，偿还了 1200 元。但这种分析是不对的，因为你并不是全年借款 1000 元。你借了一个月的 1000 元，在你偿还了 100 元后，你只借了 900 元，你本应该支付 900 元的月利息。你借了一个月的 900 元，接着又偿还了 100 元，以此类推。

我们想评估的货币流是

$$(1000,-100,-100,\ldots,-100).$$

借助计算器或者计算机，我们可以找到使上述货币流的现值为 0 的利率。该分期付款的利率高达 35%！

10.10 税收

在美国，利息收入按普通收入征税。这表明利息收入和劳动收入的税率是相同的。假设边际税率等级为 t，因此收入每增加 Δm，你的税负相应增加 $t\Delta m$。如果你投资 X 元，获得利息收入为 rX，但你要对这笔收入缴纳 trX 的税，现在只剩下 $(1-t)rX$ 的税后收入。我们将 $(1-t)r$ 称为**税后利率**（after-tax interest rate）。

如果你不是投出而是借入 X 元，结果会怎样？这种情形下，你需要支付利息 rX。在美国，某些类型的利息收入，可享受税收减免（tax deductible）。例如，抵押借款产生的利息享受税收减免，但普通消费贷款则不享受。另一方面，公司绝大部分的借款利息可从收入扣除。

如果某类型的利息可以税收减免，你可从你的其他收入中减去应支付的利息，仅按相减后所得的差额缴税。因此若你支付利息 rX 元，意味着你可以少缴 trX 元的税。这样，借款 X 元的总成本为 $rX-trX=(1-T)rX$。

因此对于税率等级相同的人来说，借款和贷款的税后利率是相同的。对储蓄征税会打击人们储蓄的积极性，但对借款补贴则鼓励人们积极借钱。

例子：奖学金和储蓄

美国的很多学生都可以获得某些类型的资金资助，以帮助他们支付学费。学生领取的资金资助数额取决于很多因素，但其中一个重要的因素是他的家庭支付学费的能力。美国大学入学考试委员会（College Entrance Examination Board, CEEB）负责计算学生家庭支付学费的能力，绝大多数大学采用了这种计算方法。

如果某个学生打算申请资金资助，他或他的家庭必须填写一份调查表，该调查表详细描述了他的家庭财务状况。CEEB 根据该学生父母的收入和财产信息，计算出“调整后的可支配收入”。父母需要根据调整后的可支配收入缴纳部分学费，一般为该收入的 22%-47%，具

体缴纳比例取决于调整后可支配收入的多少。在 1985 年，若父母的税前收入为 35,000 美元，则需要缴纳 7000 美元的学费。

随着父母收入和财产数额的上升，他们应缴纳学费的数额也上升，从而获得资助的数额也下降。CEEB 计算家庭收入的方法，实质上是对储蓄（用于为子女缴纳大学学费的储蓄）征税。马丁.菲尔德斯坦（美国国家经济研究局主席，哈佛大学经济学教授）曾计算过这种税收的规模[一]。

我们分析下面这样的情形：子女刚读大学时父母决定多储蓄 1 美元。在 6%的利率下，现在的 1 美元在 4 年后的终值为 1.26 美元。由于联邦政府和州政府都征收利息所得税，征税后只剩下 1.19 美元。然而，由于多储蓄了 1 美元相当于父母的财产增加了，他们子女在大学四年的每一年所获得的资助数额都下降。这种“教育税”使 1 美元在 4 年后的终值仅为 0.87 美元。这相当于对他们征收了 150%的收入税！

菲尔德斯坦还分析了子女即将考大学的中产阶级家庭的储蓄行为。他估计，由于联邦税、州税和“教育税”的综合影响，年收入 40,000 美元且有两个孩子即将读大学的家庭，要比同等收入的一般家庭少储蓄 50%。

10.11 税率选择

在上面分分析中，我们讨论了“利率”。在现实生活中，利率的种类很多：名义利率（nominal rates），实际利率，税前利率，税后利率，短期利率，长期利率，等等。在做现值分析时，到底应选择哪种利率？

回答这样的问题需要使用现值的原理。现值思想产生的原因，是我们希望将某个时点的资金换算为另一时点的等额资金。利率是投资的报酬，它使得我们能象上述一样将资金从一个时点转移到另一个时点。

如果有很多利率种类可用，我们在分析现值时，应该选择和我们研究的货币流性质类似的利率。如果对货币流不纳税，我们应该使用税后利率。如果货币流延续 30 年，我们应使用长期利率。如果货币流带有不确定性，我们应该使用带有类似不确定性的投资利率。（在后面章节我们会知道最后一句话是什么意思。）

利率衡量资金的**机会成本**（opportunity cost），即将资金用在其他用途的价值。因此，你应该将每个货币流与其他选择中最好的那个选择相比较，这个选择应与你的货币流具有类似的课税方式，风险和流动性的特征也应该类似。

[一]Martin Feldstein,“College Scholarship Rules and Private Savings”, American Economic Review, 85, 3 (June 1995).

总结

1.跨期消费情形下的预算线可表示为现值的形式，也可以表示为终值的形式。

2.我们在前面章节对于一般选择问题推导出的比较静态结果,同样适用于跨期消费问题。

3.实际利率衡量你现在放弃一些消费可以换得将来消费的数量。

4.若消费者能以既定利率自由借和贷，则他总是更喜欢现值更高的那个禀赋。

复习题

1.若利率为 20%，20 年后的 100 万元相当于今天的多少钱?

2.当利率增加时，跨期预算线变得更陡峭还是更平坦?

3.在跨期消费情形下，两商品为完全替代的假设还成立吗?

4.某消费者最初为放款人，利率下降后他仍然选择作为放款人，他的状况变好还是变坏了？如果利率下降后他决定变为借款人，他的状况变好还是变坏了?

5.若利率为 10%，一年后的 100 元钱的现值是多大？如果利率是 5%呢?

复习题答案

1.若利率为 20%，20 年后的 100 万元相当于今天的多少钱?

【复习内容】多时期情形下的现值

【参考答案】

$$PV=\frac{100}{(1+20\%)^{20}}=2.61(\text{万元})$$

或者根据教材表 10.1 进行粗略计算。从表 10.1 可知 20 年后的 1 元钱想当于今天的 0.03 元，因此 100 万元相当于今天的 3 万元。

曹乾（东南大学 caoqianseu@163.com）

2.当利率增加时，跨期预算线变得更陡峭还是更平坦？

【复习内容】跨期消费的预算线；利率变动对预算线的影响

【参考答案】

为简单起见，假设只有两期，容易写出预算线：

$$(1+r)c_1+c_2=(1+r)m_1+m_2$$

$$c_1+\frac{c_2}{1+r}=m_1+\frac{m_2}{1+r}$$

第一式是以终值形式表达的预算线，第二式是以现值形式表达的预算线，由于这两个式子是等价的，写出任何一个即可。

上述预算线的斜率（的绝对值）为 $(1+r)$，因此当利率增加时，斜率（绝对值）增大，预算线更陡峭。

3.在跨期消费情形下，两商品为完全替代的假设还成立吗？

【复习内容】完全替代；跨期消费

【参考答案】一般不成立。分析如下：

反证。为简单起见，假设两商品（即两个时期的消费 c_1, c_2 ）为 1:1 替代的，由于在完全替代的情形下，消费者只会消费价格便宜的那种商品。在跨期消费的情形下，由于一般来说 $r>0$，这意味着第 1 期的消费更贵，因此消费者在第 1 期的选择应该是不消费，而这是不可能的，因为不吃不喝他活不到第 2 期。

但这个结论并非绝对。因为如果 $r=0$ 时，即两个时期消费的价格相等时，完全替代的偏好仍然是可能的。请读者补充原因。

4.某消费者最初为放款人，利率下降后他仍然选择作为放款人，他的状况变好还是变坏了？如果利率下降后他决定变为借款人，他的状况变好还是变坏了？

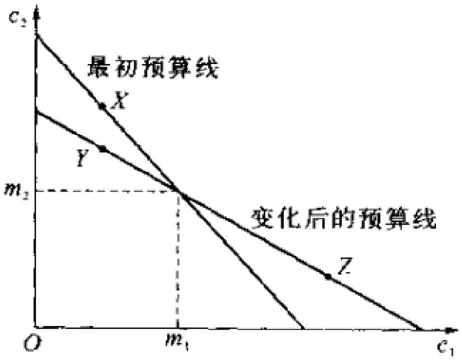
【复习内容】利率变动的比较静态分析

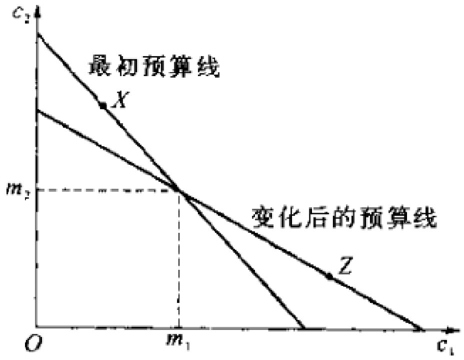
【参考答案】

第一问的答案很直观，他的状况必定变坏，除非他讨厌钱。第二问的答案需要稍微想一下。我们借助图形把这两个问题说一遍。

如果该消费者最初为放款人（左图），则他的选择一定位于 m_1 点的左边。不妨用 X 表示他的最初选择。当利率下降后，预算线变得更平坦。如果他还继续作为放款者，那么在新预算线上，他的选择也必定位于 m_1 点的左边，不妨假设他的新选择为 Y。

由于 Y 在原预算线下方，这意味着利率下降前，消费者原本可以选择 Y，但他没选 Y 而是选择了 X，根据显示偏好理论可知，消费者一定更喜欢 X。但在利率下降后，X 已经买不起，所以消费者的状况变差。





如果利率下降后，该消费者决定变为借款人（由图），这意味着他的选择 m_1 点的右边，不妨记为 Z。当消费者选择 X 时，他买不起 Z；当他选择 Z 时，他买不起 X。因此，此种情况下，我们无法比较 X 和 Z 哪个更好，也就是说消费者的状况变好还是变坏，我们无法判断。

5.若利率为 10%，一年后的 100 元钱的现值是多大？如果利率是 5%呢？

【复习内容】现值

【参考答案】

$PV_{10\%}=\dfrac{1}{1+10\%}=90.91$ 元；$PV_{5\%}=\dfrac{1}{1+5\%}=95.24$ 元。

曹乾（东南大学 caoqianseu@163.com）

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

11.资产市场（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

11 资产市场

资产（assets）是能提供长期服务流的商品。资产可提供消费服务流，例如住房，或者能提供购买消费品的货币流。提供货币流的资产称为金融资产(financial assets)。

上一章中的债券就是一种金融资产。债券提供的服务流就是利息收入流。其他类型的金融资产还有公司股票等，股票提供了不同方式的现金流。在本章，我们将分析：在资产提供的未来服务流完全可确定的情形下，资产市场有何种功能。

11.1 报酬率

在上述极端假设下，资产报酬率有个简单原理：如果资产提供的现金流没有不确定性，则各种资产的报酬率一定相同。理由很明显：如果某资产的报酬率高于另一资产，而且这两种资产除报酬率之外其余方面都相同，那么没人会购买报酬率较低的那种资产。因此均衡时，所有资产的报酬率必然相等。

下面我们分析这些报酬率的调整过程。假设资产 A 当前价格为 p_0，预计明天的价格为 p_1。每个人都确切知道该资产今天的价格和明天的价格。为简单起见，假设资产 A 在 0 期和 1 期之间不存在红利或者其他现金收入。再假设另外一项投资 B，持有期也分为 0 期和 1 期，但 B 资产在 0 期和 1 期之间有收益，报酬率为 r。现在考虑两种可能的投资方案：要么投资 A 项目 1 元钱，在下一期兑现；要么投资 B 资产 1 元钱，赚取 r 元的利息。

在 1 期期末，两种投资方案的价值各为多少？我们首先计算 1 元钱能买到的资产数量。令 x 表示购买到的资产数量，则有下式

$$p_0 x = 1 \text{或}$$

$$x = 1/p_0 .$$

由此可知价值 1 元钱的资产在下一期的终值(future value)为

$$FV = p_1 x = p_1 / p_0 .$$

另一方面，如果我们投资一元钱于 B 资产，在下一期将得到$1+r$元。如果 A 和 B 均衡时，则无论将一元钱投资于哪一项资产，在下一期这一元钱的价值都是相同的。因此，均衡条件为：

$$1+r = \frac{p_1}{p_0} .$$

如果上述等式不成立，这意味着什么？这意味着肯定存在着赚钱的方法。例如，如果

$$1+r > \frac{p_1}{p_0} ,$$

拥有 A 资产的人在第 1 期可以卖掉一单位 A，获得 p_0 元，然后投资于 B。在下一期，p_0 元变为 $p_0(1+r)$ 元，由上式可知，$p_0(1+r) > p_1$。这意味着在第二期，他们将有足够的资金来购买 A，这样他们又回到了起点，但是却有了多余的钱。

这类操作，即购买一些某类型的资产并且卖出一些另外类型的资产，以保证确定的收益，称为**无风险套利**(riskless arbitrage)或者简称**套利**。只要有人追逐“确定性的东西”，我们可以预期：运行良好的市场必然会快速消灭任何套利机会。因此，我们可以使用另外一种方式表达均衡条件，即<u>均衡时市场中不存在套利机会</u>。我们把这个条件称为**无套利条件**（no arbitrage condition）。

但是，套利是如何做到消灭不均衡的情形的？在上面的例子中，我们断言如果 $1+r > p_1/p_0$，则任何持有资产 A 的人都会希望在第一期将其卖掉，因为销售收入足够在第二期将该资产买回。问题是他们卖给谁？谁想买？市场上必定有很多人想把资产 A 卖掉，卖价为 p_0；但是在该价格下只有傻瓜才会购买。

这意味着供给超过了需求，因此价格会下降。下降的幅度为多大？价格会一直降到满足无套利条件即 $1+r = p_1/p_0$。

12.2 套利与现值

我们可以将上述无套利条件变形，从而得到下面更有用的表达形式

$$p_0 = \frac{p_1}{1+r}.$$

这个式子是说某资产的当前价格应该等于它的现值。上式在本质上就是对无套利条件的一种改写：以前我们用终值比较，现在转换为用现值进行比较。因此如果无套利条件得到满足，这意味着所有资产都是按它的现值出售的。只要资产的定价偏离了它的现值，就必定存在套利的方法。

12.3 不同资产差异的调整

无套利条件意味着两种资产提供的服务是相同的，除非纯粹货币方面的差异。如果不同资产提供的服务存在着某些特征上的差异，则我们在断言两资产的均衡回报率是相同之前，我们有必要调整这些差异。

例如，某资产可能比其他资产更容易出售。我们有时称这种资产比其他资产更有**流动性**（liquid）。在这种情形下，我们有必要调整资产回报率以使它反映资产出售的难易程度。例如，价值 100,000 元的房子的流动性可能比 100,000 元的国库券的流动性差。

类似地，某资产的风险可能比另外一项资产更低，该资产的回报率就有保障，而另外一项资产的回报率的保障性就低。我们将在 13 章分析如何调整风险差异。

此处，我们分析另外两种类型的调整。一种类型是某些资产的报酬中含有消费价值（consumption vale），需要调整；另外一种类型是某些资产的纳税特征不同，需要调整。

11.4 报酬中含有消费价值的资产

很多资产的报酬是以钱的形式。但是也有一些资产，它们的报酬是消费价值。最重要的例子是房子。如果你拥有你居住房子的产权，那么你不要租房。因此，拥有住房的部分“回报”是你有房住但不用交租金。或者，换一句话说，你把房租交给自己。后面这种解释听起来奇怪，但它包含这一个重要思想。

我们得承认，住自己的房子你没**直接**（explicit）向自己交租金，但是，你可以视为自己**隐性地**（implicitly）向自己交租金，这对分析有好处。你房子的**隐性租金率**（implicit rental rate）就是你租类似房子时需要支付的价格（比如每月每平方米多少钱）。或者等价地，隐性租金率是你在市场上把房子租给别人时你收取的价格。由于你选择“将你的房子租给自己”，你就放弃了向别人收取租金的机会，这就是房子自住的机会成本。

假设你房子的隐性租金是 T 元每年，那么你拥有住房的部分回报就是每年 T 元的隐性收入，因为如果你租住同样的房子需要向别人交纳 T 元，但由于房子是你自己的，相当于你每年收入 T 元。

但这部分隐性收入并不是你房子的全部回报。正如房地产开发商不厌其烦地对我们的劝说，房子也是一种**投资**。当你买房时你支付了一大笔钱，你当然期望这个投资能有货币回报，房子增值时你就赚钱了。某种资产的价值增加称为**增值**（appreciation）。

假设你预期一年后你的房子增值 A 元，这样拥有住房的全部回报是租金收入 T 加上投资回报 A。如果最初你买房花了 P 元，则你投资于房产的总报酬率 h

$$h=\frac{T+A}{P}=\frac{T}{P}+\frac{A}{P}.$$

上式是说，投资总报酬率等于消费报酬率（T/P）加上投资报酬率（A/P）。

我们用 r 表示其他金融资产的报酬率，则投资于房子的总报酬率，在均衡时，应该等于 r：

$$r=\frac{T+A}{P}.$$

请按下面的思路去想。在年初，你可以向银行储蓄 P 元从而赚取 rP 元；或者你可以将 P 元投资于房产，这样你可以节省 T 元的租金，而且在年末时你还可赚取 A 元。这两种投资项目的回报必须相等。因为，如果$T+A<rP$，你的明智做法是储蓄 P 元，用赚取的钱支付房租，这样在年末你还剩余的钱数为$rP-T$（注意$rP-T>A$）。如果$T+A>rP$，显然你会选择投资房产。（当然此处忽略了房地产经纪人的佣金以及和买卖相关的交易成本。）

由于房产的总报酬率（$\frac{T+A}{P}$）等于储蓄利率（r），房产投资总报酬率中的货币报酬率（A/P）通常会小于利率（r）。因此，提供消费价值回报的资产与纯粹金融资产相比，它的货币报酬率通常要低。这意味着，如果你购买消费品（例如住房、油画、珠宝首饰等）的

目的**仅仅**（solely）在于增值，这种做法并不可取。我们已说过原因——消费品的货币报酬率和消费报酬率之和等于利率，因此消费品的**货币报酬率**通常小于纯粹金融资产的货币报酬率（即利率）。另一方面，如果你很看重这些消费品的价值，或者出租这些消费品可以赚取很多的租金，那么购买消费品就是明智的做法，因为此时你在意的是这些消费品的**总报酬**而不仅仅是货币报酬。

11.5 资产报酬税

美国国内税务局（The Internal Revenue Service）对资产报酬征税时，将资产报酬分为两大类。一类是**股息或利息收入**（dividend or interest return）。这类报酬是在资产存续期内按年或者按月周期性支付的。股息或利息收入的税率是一种普通税率，和劳动收入的税率相同。

第二类报酬称为**资本收益**（capital gains）。当某项资产的售出价高于购买价时，就会产生资本收益。但是，只有当资产实际售出时，政府才会对资本收益征税。按照美国当前的税法，资本收益税的税率和普通收入的税率是相同的，但已有人提议对资本收益实施更优惠的税率。

有时人们认为，按普通收入的税率对资本收益征税是一种“中性（neutral）”政策。我们认为这种观点不对。理由有二。第一条理由是，资产收益税只有在资产实际被售出时才征收，但股息和利息收入税则按年征收。由于资本收益税的征收时间被延迟到资产出售后才征收，它的实际税率会**低于**普通收入税的税率。

第二条理由是，资本收益税是按照资产增值的**货币**额（卖价与买价之差）征税的。如果资产价值增加是由通货膨胀引起的，那么尽管该资产的**实际**价值没变，但消费者仍要纳税。例如，假设某人花 100 元购买了某资产，10 年后该资产的价值为 200 元。假设一般价格水平在这 10 年间也翻倍。那么，此人需要对这 100 元资本收益缴税，尽管该资产的实际价值未变。这种情形下，资本收益税的税率**高于**普通收入税的税率。上述两条理由阐述了两种效应，哪种效用更大一些是个有争议的问题。

除了股息税与资本收益税的差别之外，美国税法对其他资产报酬的征税政策也有差异。例如，美国联邦政府对地方政府债券（municipal bonds）即由城市政府或者州政府发行的债券不征税。对自有房屋的消费回报也不征税。而且，对自有房屋的部分资本回报也是不征税的。

由于不同资产的税收政策不同，因此为了比较报酬率，需要对无套利条件进行调整。假设某资产的税前报酬率为 r_b，另外一项资产免税，其回报率为 r_e。因此，如果某人同时持有这两项资产，而且他的收入所得税的税率为 t，则必然有

$$(1-t)r_b = r_e .$$

也就是说两种资产的税后报酬率必然相等。否则，个人就不会同时持有这两种资产，他必然选择持有税后报酬率较高的那种资产。当然，这里的讨论忽略了资产的其他方面的差异，比如流动性差异和风险差异，等等。

11.6 资产市场泡沫

假设你正打算购买一栋房子，你认为这栋房子明年这时候肯定值 22 万元。当前的利率（反映投资的机会成本）为 10%。此房子的公平价格（fair price）是它的现值 20 万元。

现在假设事情也有变数：尽管很多人认为此房子在明年能值 22 万元，但未必一定能值那么多。我们可以预期，由于存在着一定的购买风险，此房子的当前售价可能小于 20 万元。

假设一年之后，此房的价值为 24 万元，远超过人们的预期。尽管当前的利率为 10%，房子的价值上升了 20%。这种经历能让人们改变对房子未来价值的预期：也许下一年房价能涨 20%，甚至更多。

如果很多人都有这样的信念，那么他们现在就会哄抬房价——这有可能鼓励其他人对房屋市场做出更乐观的预期。我们在讨论价格调整时已经知道，人们预期某资产的收益率越高，越倾向于抬升它的价格。较高的资产价格会减少当前的需求，但是它也鼓励人们预期该资产的未来收益率较高。

第一种效应——高价格降低了需求——倾向于稳定价格。第二种效应——高价格导致人们预期该资产的未来价格更高——倾向于引起价格上升。

这是**资产泡沫**（asset bubble）的一个例子。在泡沫期，某资产的价格因为这种或那种原因上升，这导致人们预期它的价格在将来会升得更高。但是如果人们预期该资产的价格在未来会大幅上升，他们现在就会尽量多买，迫使价格升得更快。

金融市场可能存在着这样的泡沫，尤其是当参与者缺乏经验时。例如，2000-2001 年间，我们看到高科技股票急速上升，2005-2006 年间我们看到，美国大部分地区和很多其他国家的住房充斥着泡沫。

所有的泡沫最终都会破裂。价格急速下降，有些人持有资产的价值迅速缩水，它们的价值已远低于当初的购买价格。

避免泡沫的关键在于观察经济基本面。在美国住房泡沫形成的中期，住房的价格和相同住房的一年租金的比率，远远大于历史数据。这个差距大概反映了买者的预期：未来价格将上升。

类似地，房价中值（median）与收入中值的比率也达到历史最高水平。这两个指标预示着房价的高价格是无法持续的。

“这一次真得不同”的信念非常危险，尤其对于金融市场来说。

11.7 应用

无风险资产的报酬率必须相等，这个事实非常明显，也非常重要。它对资产市场的运行有令人惊讶的强大解释力。

不可再生的资源

我们研究不可再生的资源例如石油的市场均衡问题。假设石油市场是竞争性的，供给者众多，为简单起见，假设从地下开采石油的成本为零。石油价格在长期将会怎样变动？

可以证明，石油价格必须等于利率。为看清这一点，只要注意到地下的石油和其他资产一样，是一种资产。如果生产商愿意持有石油，那么石油提供的报酬必须和其他资产一样。令 p_{t+1} 和 p_t 分别表示时期 t+1 和时期 t 的价格，则石油市场的无套利条件为

$$p_{t+1} = (1+r)p_t .$$

论据可以归结为下面的简单思想：地下的石油类似于银行里的钱。如果银行里的钱的报酬率为 r，则地下石油的报酬率必然也为 r。如果地下石油的报酬率高于银行存款利率，没人立即开采石油，而是宁愿等到以后再来开采，因此当前石油价格被推高。相反，如果地下石油的报酬率低于银行存款利率，则油井所有者将会立即开采石油，目的在赶紧把钱存入银行，因此石油价格被压低。

上述论证说明了石油价格如何变动的。但是，油价本身的决定因素是什么？油价由石油的需求决定。我们用一个非常简单的需求模型进行分析。

假设石油的需求量固定在 D 桶每年，世界石油的总蕴藏量为 S 桶。这些石油一共可以使用 T=S/D 年。当石油被耗尽后，我们必须使用其他资源，比如液化煤气，假设液化煤气的生产成本为 C 元每桶，并假设它可以完全替代石油的用途。

那么，T 年之后，当石油快要耗竭时，石油价格为多少？显然，它的价格等于 C 元每桶，即等于它的完全替代品液化煤气的价格。这意味着今天的石油价格 p_0，必须按照利率 r 连续增长 T 年才能恰好等于 C。这样我们得到下面的式子

$$p_0(1+r)^T = C \text{ 或}$$

$$p_0 = \frac{C}{(1+r)^T} .$$

这个式子是说当前的油价是 C、r 和 T 的函数。现在我们可以问一些关于比较静态分析的问题。例如，意外发现了新油田会对油价有何影响？这意味着石油使用年数 T 增加，因此 $(1+r)^T$ 增加，所以油价 p_0 下降。由此可见，石油蕴藏量的增加必然会降低当前的油价。

如果技术取得了重大突破使得液化煤气的生产成本 C 降低，油价会如何变动？由上面的式子可以看出 p_0 必然下降。当液化煤气是石油的唯一并且完全的替代品时，油价必然等于液化煤气的价格。

何时砍伐森林？

假设我们用能得到的木材数衡量森林的规模，它是时间的函数即 F(t)。进一步假设：木材的价格是固定不变的；森林的增长速度一开始较快但逐渐减慢；木材市场是竞争性的。那

么，何时砍伐森林呢？

答案：当森林的增长率等于利率时。在此之前，森林的报酬率高于银行储蓄利率；在此之后，它又低于银行储蓄利率。因此砍伐森林的最佳时间就是它的增长率恰好和利率相等时。

我们可以将这个思想更正式地以现值方式表示，时期 T 砍伐森林所得现值为

$$PV = \frac{F(T)}{(1+r)^T}.$$

我们希望找到能使现值最大的选择变量 T，也就是说使森林的现值越大越好。如果我们选择的 T 值较小，森林的增长率会超过利率，这意味着现值将会增加，因此最好过一段时间再砍伐。另一方面，如果我们选择的 T 值过大，森林的增长率将会小于利率，因此现值会降低。所以，能使现值最大的 T 值，必然是它能正好使森林的增长率等于利率。

上述结论可用图 11.1 说明。在图 11.1A 中，我们画出了森林增长率以及储蓄一元钱的增长率。如果我们希望在将来的某个时点报酬最大，我们在每个时点都应该将钱投向报酬率最高的资产。当森林比较年青时，它是价值最高的资产。当森林成熟时，它的增长率下降，最终银行的报酬率更高。

对总财富的效应可用图 11.1B 说明。在 T 时点之前投资于森林，财富增长最快。过了 T 点之后，投资于银行，财富增长最快。因此，最优策略是在 T 点之前投资森林，T 点砍伐，接下来将钱投资于银行。

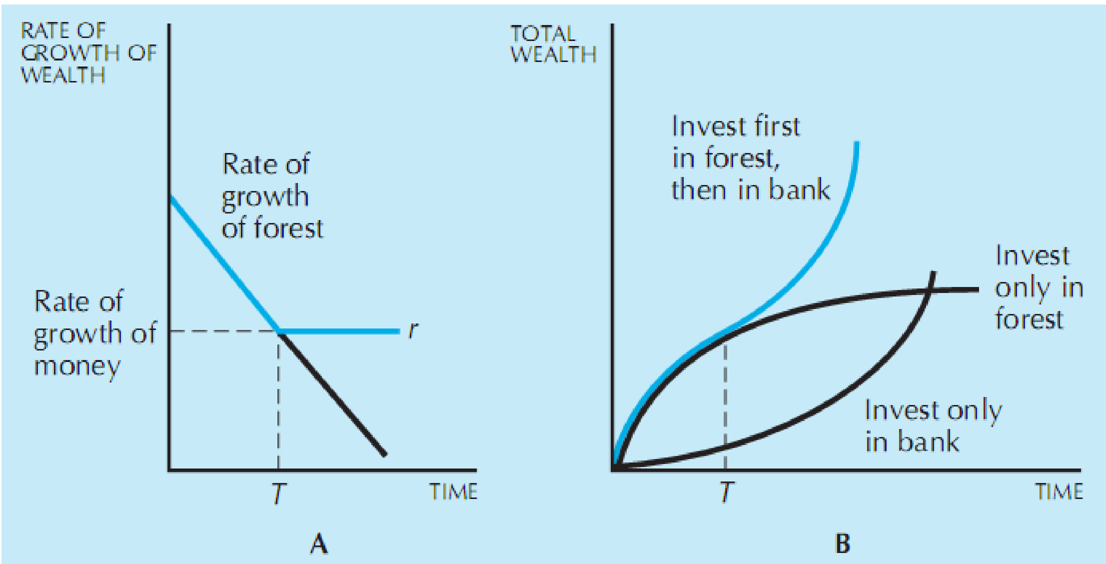


图 11.1：**砍伐森林**。砍伐森林的最佳时间是森林的增长率恰好等于利率时。

例子：海湾战争时的汽油价格

1990 年夏，伊拉克侵略科威特。联合国对此的反应是禁止从伊拉克进口石油。消息传出，世界市场的石油价格应声高涨。与此同时，美国加油站的汽油价格也显著上升。人们开始抱怨加油站大发“战争财”，晚间新闻节目也增加了若干时段轮番播报石油行业的消息。

人们觉得汽油价格高涨简直是无理取闹，他们说，高价石油从开采到运输至美国，再加上提炼成汽油，这至少要花 6 个星期的时间。因此他们认为，目前美国加油站的汽油是用低价汽油提炼的，现在却要提高价格，这分明是企图赚取“超额”利润。

你能象经济学家一样分析上观点吗？假设你有某种资产，比如储存在油罐里的汽油，它的当前价格为 1 美元每加仑。6 星期后你知道价格将上涨到 1.50 美元每加仑。你现在应如何为你的汽油定价？如果你要价低于 1.50 美元，你岂不是傻瓜？如果价格低于 1.50 美元，你干吗不让汽油在油罐内多呆 6 个星期，这样做你的状况会更好。显然，开采地下石油的跨期套利的思想同样可以应用于油罐内的汽油。

汽油公司并无过错，只有当明天汽油价格的（适度现值）等于今天的汽油价格时，它们才愿意在今天供应石油。

从福利的角度看，这个道理也能完全讲得通：如果汽油价格近期会升高，今天少消费点汽油难道不是理智的吗？汽油价格升高不仅鼓励人们采取节约措施，而且还反映了稀缺汽油真正的价格。

具有讽刺意味的是，两年之后，俄罗斯也出现了同样的现象。在向市场经济转型过程中，俄罗斯的石油价格为 3 美元每桶，而同期世界石油价格为 19 元每桶。石油生厂商预期政府很快就会允许石油涨价，因此他们尽可能地控制当前产量。正如俄罗斯某石油商所说，“你在纽约见过谁把一元钱只只卖 10 分钱的？”结果可想而知，每个加油站前都排着长长的队伍[一]。

11.8 金融机构

资产市场使人们有机会改变某时期内的消费模式。例如，有两个人 A 和 B，他们的财富禀赋不同。A 今天有 100 元，明天 0 元；而 B 今天有 0 元，明天有 100 元。他们有可能都喜欢今天和明天各有 50 元。他们可以通过交易实现这种消费模式：今天 A 给 B50 元；明天 B 给 A50 元。

在这个特别的例子中，利率为零：A 借给 B50 元，第二天只收回了 50 元。如果人们对今天和明天消费的偏好为凸，即使利率为零，他们也喜欢某时期的消费能平均一些，而不是在某时点把所有的东西都消费光。

我们可以运用上述思想分析其他资产禀赋模式。某个人的资产能产生稳定的收入流，但是他却偏偏喜欢能提供一次性总收入的资产，而另外一个人的资产能提供一次性的总收入，但他却喜欢稳定的收入流。例如，一个 20 岁的年青人希望有一笔收入好买房子，而另外一个 60 岁的老年人却希望退休后能有稳定的收入。显然，这两类人可以交换资产，从而达到了双赢的效果。

现代经济体系中的金融机构，为这样的交易提供了便利。在上例中，老年人可以将他的

[一] See Louis Uchitelle, "Russians Line Up for Gas as Refineries Sit on Cheap Oil," New York Times, July 12, 1992, page 4.

那笔钱存入银行，银行再把这笔资金贷给年青人。年青人按期偿还这笔抵押贷款，银行再向老年人支付利息。银行做这种业务是要赚取部分利润的，但是只要银行业是充分竞争的，这部分利润将会非常接近该业务的实际成本。

人们可以借助银行重新分配不同时期的消费，但银行不是具备这种功能的唯一金融机构。比如股票市场也具有这样的功能。假设某创业者创业成功。创业者需要金融家的支持，他们帮助创业者起步，他们会一直注入资金直到创业者的收入滚滚而来。一旦公司建立，公司所有者对公司未来产生的利润具有索取权：他们有权索要收入流。

但是，公司所有者可能希望将上述收入流一次性套现。在这种情形下，他们可以将公司在股票市场上出售给其他人。公司发行股份，你购买股份后就享有公司未来利润的索取权，这样，公司的原所有者就可以一次性套现。在这种方式下，交易双方都能做到重新分配不同时期的财富。

当然，除了银行和股票市场之外，还有很多种类的金融机构和市场，能为人们的跨期交易提供便利。但是如果买卖双方的人数如果不相等，比如，想将明天的消费出售的人数大于想要购买明天消费的人数，将会产生什么样的结果？和其他市场一样，如果某东西的供给超过了需求，价格将会下降。在这种情形下，明天消费的价格将下降。我们在前面已经知道明天消费的价格为

$$p=\frac{1}{1+r},$$

这个式子意味着利率必然会上升。利率上升会诱使人们多储蓄少消费，这样需求和供给最终相等。

附录

假设你对某资产进行投资，这种资产的利率为r，利息按年支付。如果你投资 1 元钱，T 年后你将拥有$(1+r)^T$元。假设利息是按月支付的，也就是说月利率为$r/12$，这样利息支付次数为 12T，因此，T 年后你将拥有$(1+r/12)^{12T}$元。如果利率是按天支付的，你拥有的钱数将为$(1+r/365)^{365T}$，以此类推。

一般来说，如果利息一年支付 n 次，T 年后你的钱数将变为$(1+r/n)^{nT}$元。你自然会问，如果利息是**连续**支付的，你将有多少钱？即，你的问题是求上述表达式在 n 趋向无穷大时的极限。可以证明[一]这个极限为

$$e^{rT}=\lim_{n\to\infty}(1+r/n)^{nT},$$

其中e等于 2.7183…，它是自然对数的底。

对于连续复利，使用上式计算会非常方便。例如，我们来验证课文中的一个结论：砍伐森林的最佳时间是当森林增长率等于利率时。由于森林在时点 T 的价值为 F(T)，此时砍伐森林的现值为

$$V(T)=\frac{F(T)}{e^{rT}}=e^{-rT}F(T).$$

为了求使现值最大的 T 值，可令上式对 T 求导，并且令求导后的表达式等于零，可得

$$V'(T)=e^{-rT}-re^{-rT}F(T)=0.\text{或}$$

$$F'(T)-rF(T)=0$$

整理可得

$$r=\frac{F'(T)}{F(T)}.$$

这个式子表明，最优 T 值满足下列条件：利率等于森林价值的增长率。

$e^{rT}=\lim_{n\to\infty}(1+r/n)^{nT}$ 的证明

证明由译者添加。这个结论的证明不难，但比较繁琐。这个证明可以转化为证明 $e=\lim_{n\to\infty}(1+1/n)^n$。因为在该式中令 $1/n=r/m$ 即可得到 $e^{rT}=\lim_{m\to\infty}(1+r/m)^{mT}$，这正是教材附录中的这个式子。为证明 $e=\lim_{n\to\infty}(1+1/n)^n$ 我们分以下几步：

1.准备知识

（1）数 e 的定义，$e=1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\ldots$，因此 e 代表 $\lim_{n\to\infty}S_n$，其中 $S_n=1++\frac{1}{1!}+\frac{1}{2!}+\ldots+\frac{1}{n!}$.

（2）二项式定理：$(a+b)^n=\sum_{k=0}^{n}\binom{n}{k}a^{n-k}b^k$

（3）定理：任何单调增大数列若有界，必收敛（即存在有限的极限）。

2.证明（1）中的 S_n 单调增且有界。

S_n 单调增一望即知。现证有界性。

$$\begin{aligned}S_n&=1+1+\frac{1}{2}+\frac{1}{2\cdot 3}+\frac{1}{2\cdot 3\cdot 4}+\ldots+\frac{1}{2\cdot 3\cdot 4\ldots\cdot n}\\&\le 1+1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\ldots+\frac{1}{2^{n-1}}\\&=1+\frac{1-1/2^n}{1-1/2}<3\end{aligned}$$

因此，S_n 单调增且有界，根据（3）可知，S_n 必收敛。

3.令 $T_n=(1+\frac{1}{n})^n$，现证 T_n 必收敛。

由（2）即二项式定理可知，

$$T_n=(1+\frac{1}{n})^n=1+n\frac{1}{n}+\frac{n(n-1)}{2!}\frac{1}{n^2}+...+\frac{n(n-1)(n-2)...1}{n!}\frac{1}{n^n}$$

$$=1+1+\frac{1}{2!}(1-\frac{1}{n})+...+\frac{1}{n!}(1-\frac{1}{n})(1-\frac{2}{n})...(1-\frac{n-1}{n})$$

对比T_n和 2 中S_n容易知道，$T_n\le S_n<3$，这说明T_n是有界的。T_n也是单调递增的（单调递增性从T_n的表达式一望便知）。因此，根据（3）可知，T_n必收敛，不妨令$\lim_{n\to\infty}T_n=T$。现在只要证明$T=\lim_{n\to\infty}S_n=e$即可。

4. 证明$T=\lim_{n\to\infty}S_n=e$。

为证明$T=e$，注意到，对于$m>n$，有

$$T_m>1+1+\frac{1}{2!}(1-\frac{1}{m})+...+\frac{1}{n!}(1-\frac{1}{m})(1-\frac{2}{m})...(1-\frac{n-1}{m})$$

如果我们固定n，让m持续增加以至无穷，则上式左端变为 T，右端变为S_n，

所以，$T>S_n$。因此，对于每个 n，都有$T\ge S_n\ge T_n$。

现在我们持续增加 n 以至无穷，因此 T 趋近于T_n。所以由$T_n\ge S_n\ge T_n$立即可知$T_n=S_n$，从而$T=\lim_{n\to\infty}S_n=e$。

证毕。■

总结

1.均衡时，所有能提供报酬的资产的报酬率必然相等。否则，就存在无风险套利的机会。

2.所有资产的报酬率必然相等，这意味着它们都是按现值出售的。

3.如果资产的征税方式不同，或者资产的风险特征不同，那么我们必须比较它们的税后报酬率，或者比较风险调整后的报酬率。

复习题

1.假设资产 A 在下一期能卖 11 元，而且已知与 A 类似资产的报酬率为 10%，A 的现值一定为多大？

2.一所房子，如果出租一年将收入 10,000 元，一年后若将房子卖掉，卖价为 110,000 元；但如果现在的售价为 100,000 元。如果你投资于该房产并按上述方式先租后卖，你的投资报酬率为多大？

3.某些类型的债券（例如，地方政府债券），它们的利息收入是不征税的。如果类似征

税债券的利率为 10%，利率收入的边际税率为 40%，那么不征税的债券的报酬率为多大？

4.假设人们对某种稀缺资源的需求固定不变，这种资源将在 10 年内耗尽。如果替代资源的价格为 40 元，利率为 10%，这种稀缺资源在今天的价格为多大？

复习题答案

1.假设资产 A 在下一期能卖 11 元，而且已知与 A 类似资产的报酬率为 10%，A 的现值一定为多大？

【复习内容】无套利条件；现值

【参考答案】

令 p_0 和 p_1 分别表示资产 A 在当期的价值（现值）和在下一期的价值，令 r 表示利率。由题意知，$p_1=11$，$r=10\%$，求 p_0.

无套利条件意味着 $1+r=p_1/p_0$，或者 $p_0=\dfrac{p_1}{1+r}$。后面这个式子是说某资产的当前价格应该等于它的现值。

所以，$p_0=\dfrac{p_1}{1+r}=\dfrac{11}{1+10\%}=10$（元），即 A 的现值应为 10 元。

2.一所房子，如果出租一年将收入 10,000 元，一年后若将房子卖掉，卖价为 110,000 元；但如果你现在的售价为 100,000 元。如果你投资于该房产并按上述方式先租后卖，你的投资报酬率为多大？

【复习内容】无套利条件；资产报酬率

【参考答案】

本题是说如果你现在投入 10 万元（$p_0=10$），1 年后你将拥有 12 万元（$p_1=12$）。

由无套利条件 $1+r=p_1/p_0$ 可知，$1+r=12/10$，解得 $r=20\%$.

3.某些类型的债券（例如，地方政府债券），它们的利息收入是不征税的。如果类似征税债券的利率为 10%，利率收入的边际税率为 40%，那么不征税的债券的报酬率为多大？

【复习内容】无套利条件；无套利条件的调整

【解题思路】

由于不同资产的税收政策不同，因此为了比较报酬率，需要对无套利条件进行调整。假设某资产的税前报酬率为 r_b，另外一项资产免税，其回报率为 r_e。因此，如果某人同时持有这两项资产，而且他的收入所得税的税率为 t，则必然有

$$(1-t)r_b = r_e.$$

也就是说两种资产的税后报酬率必然相等。否则，个人就不会同时持有这两种资产，他必然选择持有税后报酬率较高的那种资产。

【参考答案】

由题意知，$r_b = 10\%$，$t = 40\%$，求 r_e。

由调整后的无套利条件 $(1-t)r_b = r_e$ 可知 $r_e = (1-40\%)\cdot 10\% = 6\%$.

4.假设人们对某种稀缺资源的需求固定不变，这种资源将在 10 年内耗尽。到时如果替代资源的价格为 40 元，利率为 10%，这种稀缺资源在今天的价格为多大？

【复习内容】无套利条件

【参考答案】

10 年之后，当这种稀缺资源快要耗竭时，它的价格等于 40 元，即等于它的完全替代品的价格。这意味着今天这种资源的价格 p_0，必须按照利率 10%连续增长 10 年才能恰好等于 40 元。这样我们得到下面的式子

$$p_0(1+10\%)^{10} = 40$$

由此可得：

$$p_0 = \frac{40}{(1+10\%)^{10}} = 15.42 \text{（元）}。$$

因此，这种稀缺资源今天的价格应为 15.42 元。

曹乾（东南大学 caoqianseu@163.com）

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

12.不确定性（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

12 不确定性

不确定性是个无法更改的事实。人们无时无刻不面临风险，比如淋浴，步行过街或投资时都存在风险。某些金融制度例如保险市场和股票市场可以减少部分风险。我们将在下一章学习这些市场的功能，在本章我们将研究如果选择带有不确定性，人们将如何做出决策。

12.1 或有消费（contingent consumption）

我们已经学完了标准的消费者选择理论，现在将这一理论推广到不确定性环境下的消费选择。要问的第一个问题是，在不确定性的情形下，消费者选择的到底是什么“东西”？

在存在不确定性的情形下，我们认为消费者关心的是他得到不同消费束的**概率分布**（probability distribution）。概率分布通常由不同的结果以及每种结果的概率组成，在消费选择的例子中，这个结果就是不同的消费束。当某个消费者决定购买多少钱的汽车保险时，或者决定向股票市场投资多少钱时，他实际上就是对不同消费量的概率分布作出决策。

例如，假设你手头有 100 元，正考虑是否购买 13 号彩票。如果你购买 13 号而且开奖时抽奖机抽出了 13 号，你就获得 200 元。假设这个彩票要花 5 元钱。我们关注的结果有两个：抽奖机抽中 13 号和未抽中 13 号。

你的初始财富禀赋（不购买彩票时你的财富）分布为：100 元——中奖；100 元——未中奖。由于你没买彩票，13 号是否中奖和你的财富无关；但是如果你花了 5 元钱购买了 13 号彩票，你的财富分布为：295 元——中奖；95 元——未中奖。由于购买了彩票，不同情形下（中奖和未中奖）的财富概率改变了。下面我们更详细地分析这一点。

为了便于说明，我们仅限于分析货币赌博（monetary gambles）的情形。当然，我们关注钱是因为钱能买到消费品，因此我们最终关注的其实是消费选择。同样的理由适用于商品赌博，但货币赌博的情形更易于分析。还需要说明的是，我们分析的情形只涉及少数几个可能的结果，理由也是出于简单。

我们上面介绍的例子是博彩；下面我们将分析保险。假设某人的初始财产价值 35,000 元，但有可能损失 10,000 元。例如小偷偷了他的车，或者暴风雨摧毁了他的房子。假设损失发生的概率为 $p = 0.01$。此人财产的概率分布为：25,000 元——概率 1%；35,000 元——概率 99%。

购买保险则会改变上述概率分布。假设保险合同规定此人每缴纳 1 元保险费，在损失发生时可以获得 100 元的补偿。当然，不管损失是否发生，保险费都是要缴的。如果此人决定购买价值 10,000 元的保险，他要缴纳 100 元的保险费。这种情形中，在 1%的概率下他的财产为 34,900 元（=35,000 元初始财产-10,000 元损失+10,000 元保险公司补偿-100 元保险费），在 99%的概率下他的财产为 34,900 元（=35,000 元初始财产-100 元保险费）。因此不管风险

是否发生，他最终的财富都是相同的，都是 34,900 元。现在，保险充分补偿了他可能因风险而导致的损失。

一般来说，如果此人购买 K 元钱的保险，则需要缴纳保险费 γK，该情形下他面对的赌博是[一]：

概率 1%——财产 $(25{,}000 + K - \gamma K)$ 元

概率 99%——财产 $(35{,}000 - \gamma K)$ 元

此人将买多少钱的保险？答案取决于他的偏好。如果他很保守，他会买很多保险；如果他喜欢冒险，他可能一点也不买保险。正如人们对消费普通商品的偏好不同一样，人们对概率分布的偏好也不同。

事实上，在分析不确定性情形下的决策时，你可以把不同条件下的财产看成不同的商品。1000 元在遭受严重损失后，还能和 1000 元是同一个东西吗？显然不是。类似地，艳阳高照天气炎热条件下的冰淇淋甜筒，和阴雨绵绵寒冷彻骨条件下的冰淇淋甜筒也不是同一种商品。一般来说，“同一种商品”对某人的价值可能不同，这取决于此人在什么样的条件下得到这种商品。

可以将某种随机事件的结果看成不同的**自然状态**（states of nature）。上面的保险例子有两个自然状态：损失发生或者损失不发生。但一般来说有很多自然状态。于是我们可将**或有消费方案**（contingent consumption plan）定义为：在不同自然状态下的消费方案，即一个随机过程的不同结果。**或有**的意思是某事的发生带有前提条件，因此一个或有消费方案是指该消费方案取决于某些事件的结果。以购买保险为例，或有消费是用保险合同条款规定的：如果损失发生，你有多少钱；如果损失不发生，你有多少钱。消费有时也取决于天气条件，这种情形下，或有消费方案是指你在各种天气条件下（例如晴天与阴天）的消费。

就像消费者对不同消费束的存在偏好一样，消费者对不同或有消费方案也存在着偏好。例如，如果你厌恶风险，你当然更喜欢充分的保险保障。人们的选择决策反映了他们对于不同条件下的消费的偏好。因此，我们可以使用消费者选择理论分析这些选择。

如果将一个或有消费方案看成一个普通的消费束，我们就回到了前面几章的分析架构。我们可以将此时的偏好界定为对**不同或有消费方案**（即不同消费束）的偏好，而预算约束则给出了“交易条件（terms of trade）”。消费者必然在他能买得起的不同或有消费方案中，选择最好的。我们可以构建模型分析这个问题，就像我们构建消费者选择普通消费束的模型一样。

像以前一样，我们可以用无差异曲线分析消费者购买保险的行为。在前面我们已指出，自然状态有两个：损失发生以及损失不发生。或有消费是在上述不同自然状态下你相应拥有的钱数。将或有消费用图 12.1 表示。

[一] γ 为希腊字母，读作“gam-ma”.

你的或有消费的禀赋为：25,000 元——坏结果状态（损失发生）；35,000 元——好结果状态（损失不发生）。购买保险会改变这个禀赋点。如果你购买了价值 K 元的保险，相当于你放弃了好结果状态下 γK 元的消费可能性，以换取坏结果状态下的 $(K-\gamma K)$ 元的消费可能性。因此，在好结果状态下你损失的消费除以坏结果下你额外得到的消费，可得：

$$\frac{\Delta C_g}{\Delta C_b} = -\frac{\gamma K}{K-\gamma K} = -\frac{\gamma}{1-\gamma}.$$

这就是通过你的禀赋的那条预算线的斜率。它意味着你可以将好结果状态下消费的价格看为 $1-\gamma$，而将坏结果状态下消费的价格看为 γ。

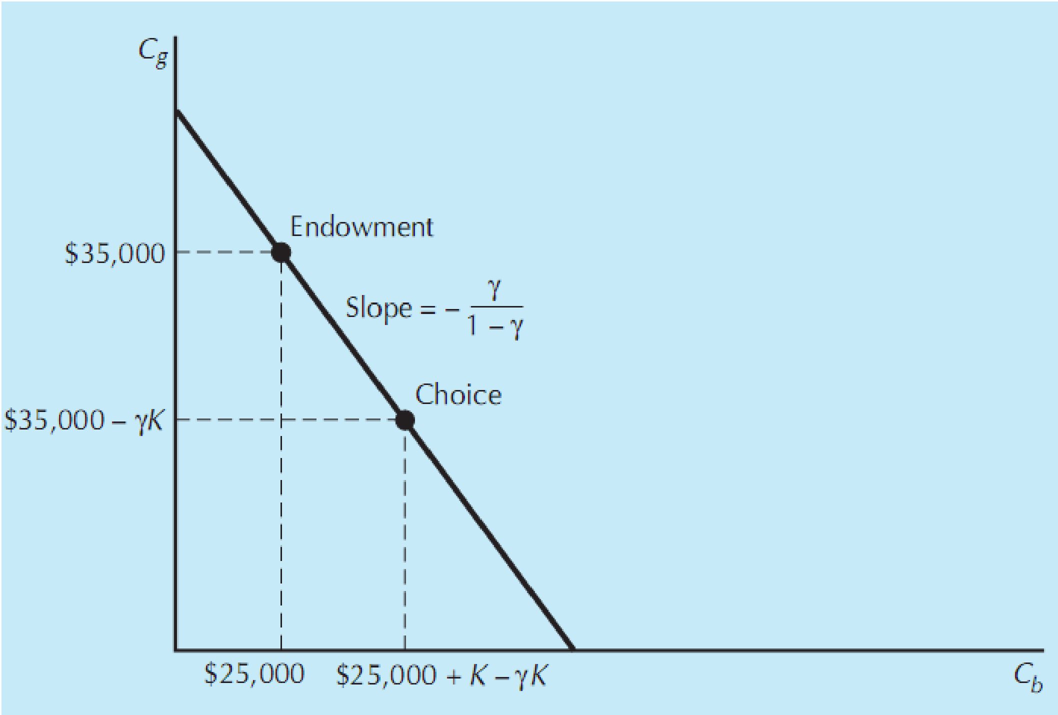


图 12.1：保险。上图画出了购买保险情形下的预算线。购买保险后，我们放弃了**好结果**状态下的一些消费 (C_g)，以换取**坏结果**状态下更多的消费 (C_b)。

我们可以画出代表消费者或有消费的预算线，此处假设无差异曲线为凸是比较自然的事：这表示消费者更喜欢在每个状态下消费量都不变，而不是在某个状态下多消费一些在另外的状态下少消费一些。

给定每种自然状态下消费的无差异曲线，我们可以分析消费者应购买多少保险。和以前一样，这可用相切条件描述：两种自然状态下的消费的边际替代率，应该等于这两种状态下消费的价格之比。

当然，只要我们有最优选择模型，我们就可以使用前面章节研发的工具进行分析。我们可以分析保险需求如何随保险价格变动而变动，也可以分析保险需求如何随消费者的财富状况变动而变动，等等。分析不确定性下的消费者行为，使用我们前面介绍的消费者行为理论已足够用了，分析方法就像分析确定性条件下消费者的行为一样。

例子：巨灾债券

我们已经知道，保险是转移财富的一种方法，它将财富从好的自然状态转移到坏的自然状态。保险交易涉及两个主体：保险买方和保险卖方。此处我们重点分析保险的销售方。

保险的销售可以分为两种类型：一是零售，即保险公司直接和终端消费者交易；二是批发，即保险公司将风险卖给其他保险公司。保险批发市场称为**再保险市场**（reinsurance market）。

一般来说，再保险市场依赖于诸如养老基金这样的大投资者，它们为风险提供了资金支持。然而，有些再保险公司却依靠个人大投资者。例如伦敦劳合社（Lloyd’s），它是世界最著名的一个再保险机构，一般使用个人投资者。

近来，再保险行业正试点发行**巨灾债券**（catastrophe bonds），据说，这种方式更能灵活地提供再保险。这些通常卖给大机构的债券，一般和诸如地震和飓风这类自然灾害捆绑在一起。

金融中介机构例如再保险公司或者投资银行，发行和某种可承保事件（比如赔款超过 50 亿美元的地震）挂钩的债券。如果地震没发生，投资者可获得丰厚的利率回报。但是，如果地震发生，赔款超过了债券规定的既定金额，则投资者血本无归。

巨灾债券有一些诱人的特征。它们能广泛分散风险，而且还可将债券无限细分，这样每个投资者只承担一小部分风险。购买债券的资金需要事先支付，因此对保险公司来说不存在违约风险。

从经济学的观点来看，“猫债券(cat bonds)”是一种**状态依赖证券**（state contingent security）[一]，也就是说，当且仅当某些特定事件发生时，这样的证券才支付报酬。状态依赖证券这个概念由诺贝尔奖获得者肯尼斯.J.阿罗首先提出，他在 1952 年发表的一篇论文中使用了这个概念。长期以来人们认为状态依赖证券只有理论意义，然而后来发现，所有种类的期权和其他金融衍生品都可以认为是状态依赖证券。现在，那些穿梭于华尔街金融市场的专家在创造新的衍生工具（如巨灾债券）时，都使用了这个已有 50 多年历史的科研成果。

12.2 效用函数与概率

如果消费者对不同环境中的消费偏好是理性的，我们就可以像前几章一样，用效用函数描述他的偏好。然而，此处我们考虑的是不确定性情形下的选择，这就对选择问题增添了新的形式。一般来说，消费者如何评价不同状态下的消费，取决于这些状态实际发生的概率。例如，考虑我打算用雨天的消费替代晴天的消费，替代比率显然和我认为下雨的概率有关。消费者对于不同自然状态下消费的偏好，取决于他认为这些状态发生的概率有多大。

[一] 猫债券（cat bonds）是巨灾债券（catastrophe bonds）的通俗叫法，原因在于 catastrophe 前三个字母正是 cat。译者注。

由于这个原因，我们将效用函数写为依赖于概率和消费水平的函数。假设我们考虑两种互不相容的状态，例如下雨和不下雨，损失和不损失，等等。令c_1和c_2分别表示状态 1 和状态 2 下的消费数量，令π_1和π_2分别表示状态 1 和状态 2 下损失实际发生的概率。

如果两种状态互不相容，这意味着只有其中一种状态发生，所以$\pi_2 = 1 - \pi_1$。但是我们一般还是写出二者的概率，目的只是对称好看。

定义了上述记号后，我们就可以写出状态 1 和状态 2 下消费的效用函数，$u(c_1, c_2, \pi_1, \pi_2)$。这个函数代表了消费者对每种状态下消费的偏好。

实例：效用函数的若干例子

在分析不确定性下的选择问题时，我们可以使用在前面几章介绍过的几乎所有的效用函数。一个漂亮的例子是完全替代。此处自然可以使用每种消费的概率作为权重。这样的效用函数的表达式为

$$u(c_1, c_2, \pi_1, \pi_2) = \pi_1 c_1 + \pi_2 c_2 .$$

在不确定的情形下，上述表达式称为**期望值**（expected value）。期望值是你能得到的平均消费水平。

我们也可以使用柯布-道格拉斯效用函数分析不确定性下的选择问题：

$$u(c_1, c_2, \pi, 1-\pi) = c_1^{\pi} c_2^{1-\pi} .$$

由上式可以看出，不同消费束组合的效用是以非线性消费方式表达的。

和以前一样，我们可以对上述效用函数进行单调变换，得到一个新的效用函数，但这个新函数和上述效用函数代表的偏好是相同的。使用对数形式的柯布-道格拉斯效用函数通常比较方便，它的表达式为

$$\ln u(c_1, c_2, \pi_1, \pi_2) = \pi_1 \ln c_1 + \pi_2 \ln c_2 .$$

12.3 期望效用

一种特别方便的效用函数是下面这样的函数

$$u(c_1, c_2, \pi_1, \pi_2) = \pi_1 v(c_1) + \pi_2 v(c_2) .$$

这个式子是说效用可以写成每种状态下的消费函数（$v(c_1)$和$v(c_2)$）的加权和，其中权重分别为每种状态发生的概率（π_1和π_2）。

我们在上一节已举了两个这样函数的例子。在完全替代或者称为期望值效用函数的表达式中$v(c) = c$。柯布-道格拉斯函数的原始形式不是期望值类型的函数，但是在取对数后，

它就变成了线性形式，其中 $v(c)=\ln c$ 。

如果其中一种状态肯定发生，比如 $\pi_1=1$ 时，$v(c_1)$ 就是状态 1 下消费的效用。类似地，如果 $\pi_2=1$，$v(c_1)$ 就是状态 2 下消费的效用。因此表达式

$$\pi_1 v(c_1)+\pi_2 v(c_2)$$

表示消费 (c_1,c_2) 的平均效用或期望效用。

由于这个原因，我们将上面的效用函数称为**期望效用函数**（expected utility function），或者有时称为**冯.诺依曼-摩根斯坦效用函数**（von Neumann-Morgenstern utility function）[一]。

当我们说消费者的效用可用期望效用函数表示时，或者说消费者的偏好具有期望效用的性质时，我们的意思是说，我们可以选择一个具有上述可加性形式的效用函数。当然我们也可以选择其他形式的效用函数；某个期望效用函数的任何单调变换和这个效用函数描述的偏好是相同的。但是可加性形式的效用函数使用起来非常方便。如果消费者的偏好可用 $\pi_1\ln c_1+\pi_2\ln c_2$ 表示，那么他的这种偏好也可用 $c_1^{\pi_1}c_2^{\pi_2}$ 表示。但是后面这种表达式不具有期望效用的性质，而前者具有。

另一方面，期望效用函数在经过某些类型的单调变换后，仍然具有期望效用的性质。如果函数 $v(u)$ 可以写成下列形式：$v(u)=au+b$（其中 $a>0$），则将函数 $v(u)$ 称为**正仿射变换**（positive affine transformation）。正仿射变换意味着将某函数 u 乘以一个正数，然后再加上一个常数。容易证明，如果你将某个期望效用函数进行正仿射变换，则这个正仿射函数不仅还代表着相同的偏好（这一点很明显，因为仿射变换只是一种特殊类型的单调变换），而且仍具有期望效用的性质。

经济学家说期望效用函数只能进行正仿射变换。这就是说，你可以对期望效用函数进行正仿射变换得到另外一个期望效用函数，这个新函数和原来的函数代表的偏好相同。但是如果你进行其他类型的单调变换，期望效用的性质就会被破坏掉。

12.4 为什么期望效用函数是合理的？

期望效用函数使用起来很方便，但它是否是合理的?为什么要认为对不确定性下选择的偏好具有期望效用函数这种特别的结构? 事实上，它的合理具有让人信服的原因。

随机选择的结果是在不同环境下消费的消费品，这个事实意味着最终**只有其中一种**结果会实际发生。你的房屋要么烧毁要么不烧毁；天气要么是阴天要么不是阴天。我们对选择问题的上述设定方式意味着很多可能结果中只有一种会实际发生，因此只有一种或有消费方案最终被实际实现。

[一] 约翰.冯.诺依曼是 20 世纪最厉害的数学家之一。他对物理学、计算机科学和经济理论也贡献了若干重要思想。奥斯卡.摩根斯坦是普林斯顿大学的经济学家，他和冯.诺依曼一起研发了数学博弈论。

这个结论的含义比较有趣。假设你正在考虑是否为你的房子购买明天的火灾保险。在做决策时，你会关注三种情形下的财富大小：你现在的财富（c_0）；若房屋烧毁后你的财富（c_1）；房屋没烧毁时的财富（c_2）。（当然，你真正关心的是三种结果中的消费可能性，但此处为简单起见，我们用财富作为消费的指标。）令π_1表示你房子被烧毁的概率，π_2表示未被烧毁的概率，则你对上述三种不同消费的偏好，通常可用效用函数$u(\pi_1,\pi_2,c_0,c_1,c_2)$表示。

假设我们正在考虑用现在的财富去交换其中一种可能结果——比如，我们现在愿意放弃多少钱，以换取若房屋烧毁时多得一点钱。于是，**这个决策和你在其他自然状态的消费量无关，即和若房屋未烧毁时你的消费量无关**。因为这个房屋要么烧毁要么未烧毁。如果它不幸烧毁，则额外财富的价值**不应该**取决于若它**未烧毁**时你拥有的财富量。过去的已经过去——因此，未发生的事情不应该影响**实际发生**结果中的消费的价值。

注意这是对消费者个人的偏好的**假设**。这个假设在某些情形下不成立。因为当人们在两件东西之间进行选择时，他们拥有的第三种东西的数量对他们上述选择决策很重要。例如，选择咖啡还是茶可能取决于你拥有多少奶油。注意这是因为你喜欢将咖啡中加入奶油。但是，如果奶油和咖啡不是放在一起消费的，比如你的选择是下面这样的：通过掷骰子的方式得到咖啡、茶和奶油中的**一种**，那么你可能得到的奶油数量不应该影响你对咖啡和茶的偏好。为什么?因为这三种东西你只能得到一种：如果你得到了奶油，那么你就不可能得到咖啡和茶。

因此，对于不确定性情形下的选择，各种结果在本质上是互相独立的，因为你不可能同时得到不同自然状态下的消费。人们在一种自然状态下作出的选择，应该和另外一种自然状态下作出的选择无关。这个假设称为**独立性假设**（independence assumption）。这个假设意味着或有消费的效用函数将采取比较特殊的形式：不同或有消费束之间必然可加。

也就是说，如果c_1,c_2和c_3是不同自然状态下的消费，π_1,π_2和π_3是这三种不同自然状态发生的概率，在独立性的假设条件下，效用函数必然采取下列形式

$$U(c_1,c_2,c_3)=\pi_1 u(c_1)+\pi_2 u(c_2)+\pi_3 u(c_3).$$

这就是我们前面称呼的期望效用函数。注意，期望效用函数的确满足下面的性质：两种商品的边际替代率和第三种商品的数量无关。比如，商品 1 和商品 2 的边际替代率为

$$MRS_{12}=\frac{\Delta U(c_1,c_2,c_3)/\Delta c_1}{\Delta U(c_1,c_2,c_3)/\Delta c_2}=-\frac{\pi_1\Delta u(c_1)/\Delta c_1}{\pi_2\Delta u(c_2)/\Delta c_2}.$$

由上式可知，商品 1 和商品 2 的边际替代率仅取决于商品 1 和商品 2 的数量比率，和你拥有的商品 3 的数量无关。

12.5 厌恶风险

在前面我们宣称，期望效用函数在分析不确定性情形下的选择问题时，具有某些方便的性质。在本节我们举例说明。

我们用期望效用函数分析一个简单的选择问题。假设某消费者现在有 10 元钱。他正考虑是否参加下面的赌博：50%的概率下可赚取 5 元，50%的概率下损失 5 元钱。他的最终财富因此是随机的：50%的概率下拥有 5 元钱，50%的概率下拥有 15 元钱。他财富的**期望价值**（expected value）是

$$\frac{1}{2}u(15元)+\frac{1}{2}u(5元).$$

用图 12.2 说明。财富的期望效用是两个数即$u(15元)$和$u(5元)$的平均值，在图形上以$.5u(5)+.5u(15)$表示。我们还用$u(10)$标记了财富期望价值的效用。注意，在该图中，期望财富的效用大于财富的期望效用，即

$$u(\frac{1}{2}\times 15+\frac{1}{2}\times 5)=u(10)>\frac{1}{2}u(15)+\frac{1}{2}u(5).$$

这种情形下，我们说消费者是**厌恶风险的**（risk averse），因为他更喜欢他财富的期望价值而不愿意去赌博。当然，消费者的偏好也可能出现下面的情形，即他偏好财产的随机分布胜于它的期望价值，这种情形下我们说消费者是**爱好风险者**（risk lover）。图 12.3 给出爱好风险者的一个例子。

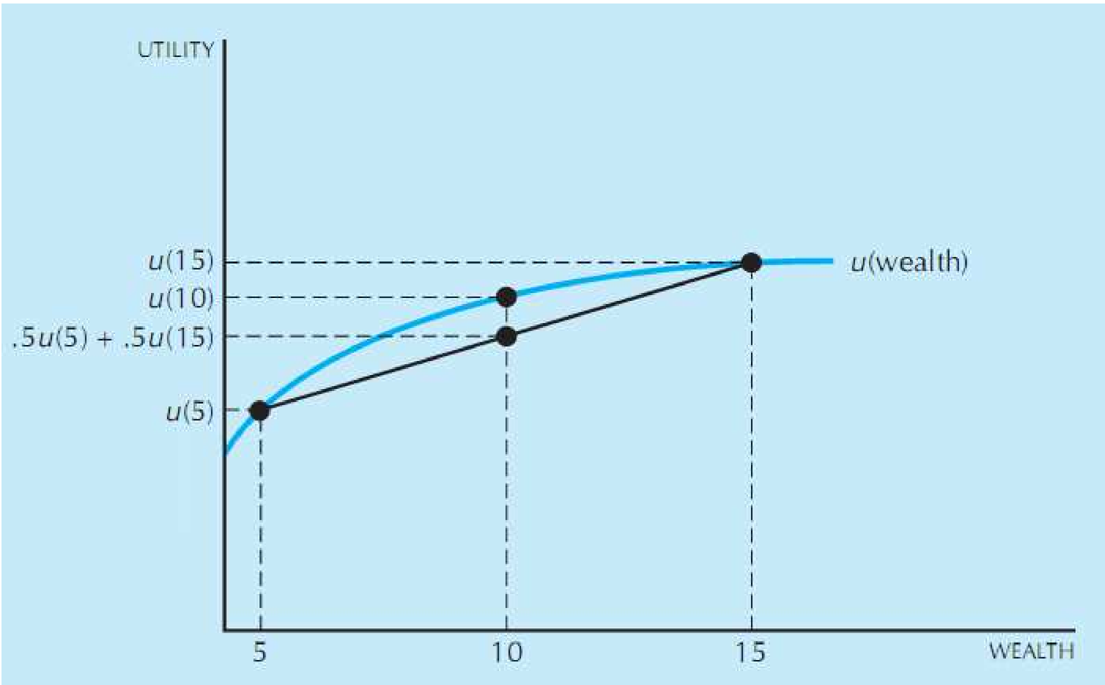


图 12.2：厌恶风险（risk averse）。对于厌恶风险的消费者来说，财富期望价值的效用$u(10)$，大于财富的期望效用$0.5u(5)+0.5u(15)$。

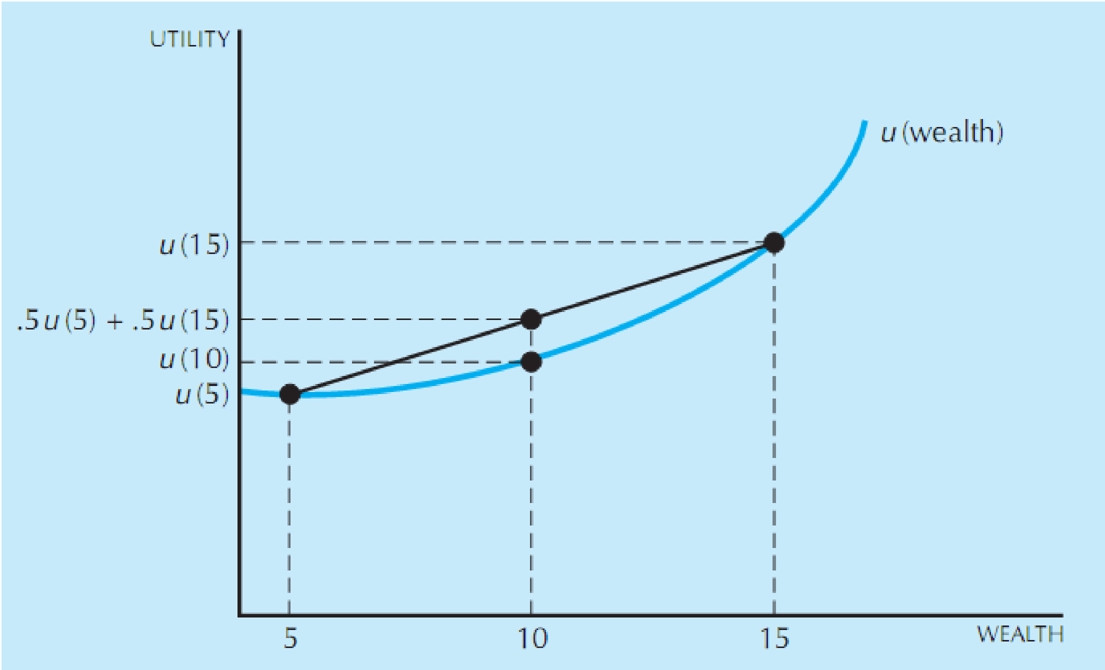


图 12.3: 爱好风险（risk loving）。对于爱好风险的消费者来说，财富的期望效用 $0.5u(5)+0.5u(15)$ 大于财富期望价值的效用 $u(10)$。

注意图 12.2 和图 12.3 的区别。厌恶风险的消费者的效用函数是**凹的**（concave）——它的斜率随着财富增加而逐渐减小（曲线越来越平缓）；爱好风险者的效用函数是**凸的**（convex）——它的斜率随着财富增加而逐渐增加（曲线越来越陡峭）。因此，效用函数的曲率（curvature）衡量消费者对待风险的态度。一般来说，效用函数越凹，消费者对风险越厌恶；效用函数越凸，消费者对风险越喜欢。

中间情形是线性效用函数。这种情形下消费者是**风险中性的**（risk neutral）：财富的期望效用等于财富期望价值的效用。在风险中性的情形下，消费者一点也不关心他财富的风险，他只关心财富的期望价值。

例子：保险的需求

我们用期望效用函数研究前面介绍过的保险需求。我们知道在那个例子中，消费者的财富为 35,000 元，可能遭受的损失为 10,000 元。损失发生的概率为 1%，购买 K 元的保险金额需要支付 γK 元的保险费。我们在前面已用无差异曲线分析过这个选择问题，在那里我们已经知道保险的最优选择由下列条件决定：两种结果（损失和不损失）下消费的边际替代率一定等于 $-\gamma/(1-\gamma)$。令 π 表示损失发生的概率，$1-\pi$ 表示损失不发生的概率。

令状态 1 表示损失不发生，因此此人在该状态的财富为

$$c_1 = 35,000 - \gamma K,$$

令状态 2 为损失发生，此状态下他的财富为

$$c_2 = 35,000 - 10,000 + K - \gamma K .$$

于是消费者对保险的最优选择，由两种结果下消费的边际替代率等于价格比率这个条件决定：

$$MRS = -\frac{\pi \Delta u(c_2)/\Delta c_2}{(1-\pi)\Delta u(c_1)/\Delta c_1} = -\frac{\gamma}{1-\gamma} . \tag{12.1}$$

现在我们从保险公司的角度看保险合同。风险发生的概率为π，此时它们需要赔付K元；风险不发生的概率为$1-\pi$，此时无需赔款。不管风险发不发生，它们铁定收集保费γK元。于是，保险公司的期望利润为

$$P = \gamma K - \pi K - (1-\pi)\cdot 0 = \gamma K - \pi K .$$

假设保险公司在该合同上恰好不亏不盈。也就是说保险公司的保费率是“公平的”。此处“公平”是指保险的期望价值恰好等于它的成本，即

$$P = \gamma K - \pi K = 0 ,$$

这意味着$\gamma = \pi$.

将$\gamma = \pi$代入（12.1）式可得

$$\frac{\pi \Delta u(c_2)/\Delta c_2}{(1-\pi)\Delta u(c_1)/\Delta c_1} = \frac{\pi}{1-\pi} .$$

删去上式左右两端的$\pi/1-\pi$并重新整理可知，保险最优数量必然满足

$$\frac{\Delta u(c_1)}{\Delta c_1} = \frac{\Delta u(c_2)}{\Delta c_2} . \tag{12.2}$$

这个式子表明：**损失发生时额外一元钱收入的边际效用，应该等于损失不发生时额外一元钱的效用**。

假设消费者是厌恶风险者，因此他的资金的边际效用随着资金数量的增加而下降。因此，如果$c_1 > c_2$，则在c_1处的边际效用将会小于c_2处的边际效用，反之亦然。而且，如果c_1处的边际效用等于c_2处的边际效用，正如（12.2）显示的那样，则一定有$c_1 = c_2$。将前面的c_1, c_2表达式代入$c_1 = c_2$可得

$$35,000 - \gamma K = 25,000 + K - \gamma K ,$$

这意味着 K=10,000 元。这表明如果保险的费率是“公平”的，厌恶风险者总是购买足额保险（full insurance）。

为什么？因为财富在每种状态下的效用，仅取决于消费者在该状态下拥有的财富总量，而不取决于在其他状态下的财富数量，因此，如果消费者在每种状态下拥有的财富量相等，

财富的边际效用必然也相等。

总结一下：如果某消费者是厌恶风险的又是追求期望效用最大化的，如果他有机会购买公平保险，那么他的最优选择就是购买足额保险。

12.6 多样化

现在我们转向另外一个涉及不确定性的问题——多样化的好处。假设你打算投资 100 元钱于两个不同的公司，一个生产太阳镜，一个生产雨衣。长期天气预报已经告诉你明年夏天晴阴天数各半。你怎么投资？

分散你的资金，两个公司各投入一些钱难道不是明智的做法吗？多样化投资会使投资报酬更为确定，因此深受厌恶风险者喜爱。

例如，假设雨衣公司和太阳镜公司的股份价格都为 10 元每份。如果夏季天天下雨，则雨衣公司的股份会值 20 元每份，太阳镜公司的股份值 5 元每份。如果夏季天天晴朗，结果正好颠倒过来：太阳镜公司的股份值 20 元每份，雨衣公司的股份值 5 元每份。如果你将 100 元全部投资于太阳镜公司，相当于你在赌博：50%概率下得到 200 元，50%概率下得到 50 元。如果你将 100 元全部投资于雨伞公司，结果是一样的：在两种情形下你的期望收益是 125 元。

如果你对上述两家公司各投入 50 元，结果会如何？那么，如果天气晴朗你可以从太阳镜公司得到 100 元，从雨衣公司得到 25 元。但是，如果是雨天，你从雨衣公司得到 100 元，从太阳镜公司得到 25 元。与孤注一掷相比，投资多样化使你减少了投资的整体风险，但仍保住了相同的期望收益（125 元）。

这个例子中的多样化非常简单，原因在于这两家公司的股份价值是负相关的——一个上升另一个就下降。类似于上例的资产组合非常重要，因为它们大幅度降低了风险。遗憾的是，这样的资产组合很难找到。大部分资产的价值运动方向是相同的：当通用公司的股票价格上升时，福特公司的股票和古德里奇（Goodrich）公司的股票也是上升的。但是，只要资产价格变动不是**完全**正相关的，投资多样化就会有好处。

12.7 风险分散

现在再回到保险的例子。在前面那个保险的例子中，我们考虑的情形是：某人有 35,000 元，风险发生时（概率 0.01）他会损失 10,000 元。假设有 1000 个这样的人，那么平均来说，有 10 人会遭受损失，因此每年损失总为为 100,000 元。这 1000 人中的每个人面临的**期望损失**（expected loss）都等于 0.01 乘以 10,000 元，即 100 元每年。假设任何人遭受损失的概率不影响其他人遭受损失的概率，也就是说，假设人们的风险是互相**独立的**（independent）。

这样，每个人的期望财富都等于 $0.99\times 35{,}000+0.01\times 25{,}000=34{,}900$ 元。但是每个

人也承担了大量风险：每个人都可能在 1%的概率下损失 10,000 元。

假设每个消费者决定分散（diversify）他面临的风险。他应该怎么做？答案：他可以把风险卖给其他人。假设这 1000 个消费者决定互相承保。如果任何人遭受 10,000 元的损失，比如他的房子被烧毁，这 1000 人每人都要给他 10 元。于是他的损失得到了全额补偿。其他的消费者内心也会平静，因为说不定哪天风险会降临到他的头上时，他也能得到补偿。这是一个**风险分散**（risk spreading）的例子：每个消费者将他自身的风险分散到其他消费者身上，因此减少了他自己承担风险的数量。

现在平均每年会有 10 栋房子被烧毁，因此每年这 1000 人每人平均付出 100 元。但这只是在平均意义上来说的。某些年份烧毁的房子数量可能是 12 栋，另外一些年份可能是 8 栋。因此，在任意一年中，每人实际付出 200 元以上的可能性很小。但是，可能性很小不等于说没有可能，万一某年烧毁 20 栋以上的房子怎么办？

这样小概率的风险仍有分散的方法。假设每个消费者同意每年实际支付 100 元，不管是否有房子被烧毁。这样他们就建立了一个火灾准备金（cash reserve fund），这些钱可以用于火灾倍增的年份。他们每年实际出资 100 元，平均来说这些钱已足够补偿遭受火灾损失的人。

你已经看到了，组建准备金意味着一个合作保险公司呼之欲出，因为保险公司的本质就是这样的。当然为了使它更像一个保险公司，我们可以增加一些特征，比如允许上述火灾准备金进行投资赚取红利等等。

12.8 股票市场的作用

从分散风险的角度来看，股票市场的作用类似于保险市场。我们在第 11 章已经知道，某企业股票上市后，该企业的原拥有者可以将未来多年的利润流一次性套现。当然，有了股票市场后，他们不必将所有财富维系在一家企业身上，这样做的风险很高；他们可将套现的钱用于分散投资。企业的原拥有者有动机发行自己公司的股份，因为这样做可将自己公司的风险分散到数量众多的股东身上。

类似地，该企业的新股东也可以使用股票市场分散他们的风险。如果你对你持有股份的那家公司的政策不满，比如你埋怨它太冒险或者埋怨它太保守，此时你可以卖掉这家公司的股份而买其他公司的股份即可。

在保险的例子中，个人通过购买保险的方式可将自身的风险降低为零。固定缴纳 100 元的保费，个人可以购买价值 10,000 元的足额保险，用来转嫁可能遭受的 10,000 元损失。风险能够降低到零的原因在于此处不存在总体风险（aggregate risk）[一]：如果损失发生概率为 1%，则 1000 人中平均有 10 个人可能遭受损失，只是我们不知道具体是谁遭受损失而已。

但在股票市场中，存在着总体风险即市场风险。某年份股票市场可能是牛市，另外一年

[一] 总体风险，又称市场风险或不可分散风险，指由于某些因素的影响，市场总体变坏。例如股票市场中的熊市。译者注。

却可能是熊市。总会有些投资者遭受损失。在股票市场上，不远承担风险的人可以将风险投资转嫁给（卖给）愿意承担风险的人。

当然，在拉斯维加斯这个赌城赌博的人可能喜欢风险，除此之外，大多数人都是厌恶风险的。因此股票市场使得风险从相对厌恶风险的人转移到相对喜欢风险的人手里，前提是喜欢风险的人能够获得充分的补偿。我们在下一章分析这种思想。

附录

我们用一个简单的问题说明期望效用最大化的一般原理。假设某消费者有 w 元，他打算投资 x 元于某项风险资产上。这种资产在“好”结果时的回报为 r_g 元，在“坏”结果时的回报为 r_b 元。你应该将 r_g 想像为正的回报，即资产价值增加；将 r_b 想像为负的回报，即资产价值下降。

因此该消费者在好结果和坏结果状态下的财富数量分别为

$$W_g = (w - x) + x(1 + r_g) = w + xr_g$$

$$W_b = (w - x) + x(1 + r_b) = w + xr_b .$$

假设好结果和坏结果发生概率分别为 π 和 $(1-\pi)$ 。那么若消费者决定投资 x 元时的期望效用为

$$EU(x) = \pi u(w + xr_g) + (1-\pi)u(w + xr_b) .$$

消费者希望选择 x 使得上述表达式的数值最大。

将上式对 x 求导数就可以发现效用值如何随 x 变化而变化的：

$$EU'(x) = \pi u'(w + xr_g)r_g + (1-\pi)u'(w + xr_b)r_b . \tag{12.3}$$

效用对 x 的二阶导数为

$$EU''(x) = \pi u''(w + xr_g)r_g^2 + (1-\pi)u''(w + xr_b)r_b^2 . \tag{12.4}$$

如果该消费者是厌恶风险的，他的效用函数将是凹的，这意味着对于任一财富量 w ,都有 $u''(w) < 0$ 。因此，期望效用对财富的二阶导数必然为负。期望效用函数是 x 的凹函数。

考虑投资一元钱于风险资产时期望效用的变化。这正是（12.3）式在 $x = 0$ 时的导数值：

$$\begin{aligned} EU'(0) &= \pi u'(w)r_g + (1-\pi)u'(w)r_b \\ &= u'(w)[\pi r_g + (1-\pi)r_b]. \end{aligned}$$

等式右端方括号内的式子是投资资产的**期望报酬**（expected return）。如果资产的期望报酬为负，则当投资一元于该资产时，期望效用必然下降。因为期望效用函数为凹，所以期望效用的二阶导数为负，这意味着随着投资额增加时，期望效用必然持续下降。

曹乾（东南大学 caoqianseu@163.com）

因此，我们已经发现如果某赌博的**期望报酬**为负，厌恶风险者在 $x^* = 0$ 处**期望效用**（expected utility）达到最大：对于这样注定会输的项目他一分钱也不会投入。

另一方面，如果资产的期望报酬为正，则投资从 0 增加到 x 时期望效用也会增加。因此不管他是如何厌恶风险，他都会投资一些钱于这个风险资产身上。

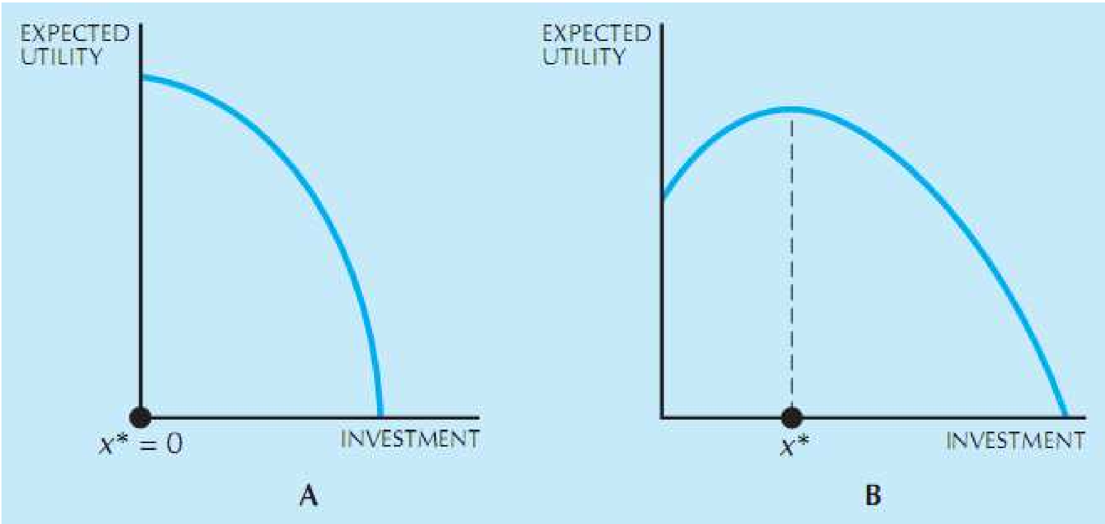


图 12.4：对于某风险资产应投资多少钱？在 A 图中，最优投资额为 0，而在 B 图中，最优投资额为正数。

期望效用是 x 的函数，这可用图 12.4 说明。在图 12.4A 中，期望报酬为负，最优选择为 $x^* = 0$。在图 12.4B 中，期望效用在某些范围中为正，因此消费者会投资 x^* 元，x^* 为正。

消费者最优投资额由下面的条件决定：期望效用函数对 x 的导数等于 0。由于对于凹的期望效用函数来说，它对 x 的二阶导数必然为负，因此这是个全局最大（global maximum）。

令（12.3）式等于 0 可得

$$EU'(x) = \pi u'(w + xr_g)r_g + (1-\pi)u'(w + xr_b)r_b = 0. \tag{12.5}$$

这个式子决定了该消费者的最优投资数量 x。

例子：征税对风险投资的影响

如果政府对投资报酬征税，会影响你对某风险资产的投资数额决策吗？如果个人缴税税率为 t，则税后报酬为 $(1-t)r_g$ 和 $(1-t)r_b$。因此，一阶条件决定了他的最优投资金额 x 为：

$$EU'(x) = \pi u'\left(w + x(1-t)r_g\right)(1-t)r_g + (1-\pi)u'\left(w + x(1-t)r_b\right)(1-t)r_b = 0.$$

上式左右两端同时除以$(1-t)$可得

$$EU'(x)=\pi u'\left(w+x(1-t)r_g\right)r_g+(1-\pi)u'\left(w+x(1-t)r_b\right)r_b=0 \qquad (12.6)$$

我们用x^*表示不征税（t=0）时的最优解，用$\hat{x}$表示征税情形下的最优解。x^*和$\hat{x}$有什么关系？

你的第一反应也许是认为$x^*>\hat{x}$，即认为征税会打击人们投资的积极性。但是这个观点是不正确的。对于上述风险资产征税实际上会**鼓励**人们的投资！

事实上，x^*和$\hat{x}$之间有确切的关系，即

$$\hat{x}=\frac{x^*}{1-t}.$$

只要注意到$\hat{x}$的这一数值满足征税情形下最优选择的一阶条件，你就可以证明上式。将$\hat{x}$代入（12.6）式可得

$$\begin{aligned}EU'(x)&=\pi u'\left(w+\frac{x^*}{1-t}(1-t)r_g\right)r_g+(1-\pi)u'\left(w+\frac{x^*}{1-t}(1-t)r_b\right)r_b\\&=\pi u'(w+x^*r_g)r_g+(1-\pi)u'(w+x^*r_b)r_b=0,\end{aligned}$$

其中最后一个等式可从x^*是不征税时的最优解这个事实推知。

此处发生了什么事？对风险资产的投资征税怎么会反而增加人们的投资金额？道理是这样的。征税后，投资者在好结果状态下得到的报酬减少，但他在**坏结果状态下的损失也减少了**。征税后，他只要将现在的投资额变为原投资额的$1/(1-t)$倍，则该投资的**税后**报酬恰好等于不征税时的投资报酬。征税减少了他的期望报酬，但同时也减少了他的风险：通过增加投资额的方式，投资者恰好可以得到原来的报酬数量，因此完全抵消了征税的影响。对风险资产投资征税，意味着当投资报酬为正时政府从你的投资收益中拿走了一部分，但是当投资报酬为负时，征税却相当于政府对你的损失进行了补偿。

总结

1.你可以把不同自然状态下的消费视为普通的消费品，这样你就可以使用前几章的分析工具研究不确定性情形下的选择问题。

2.然而，在刻画不确定性情形下的选择问题时，你需要使用具有特殊结构的效用函数。特别地，如果效用函数是概率的线性函数，你对赌博效用的赋值恰好就是不同结果（outcomes）的期望效用。

3.期望效用函数的曲率(curvature)描述消费者对待风险的态度。如果它是凹的，则消费者是厌恶风险者；如果它是凸的，则消费者是风险喜好者。

4.金融机构例如保险市场和股票市场，使消费者能够进行多样化投资和分散风险。

复习题

1.你的消费束如何才能位于图 12.1 禀赋点的左侧?

2.下列哪些效用函数具有期望效用函数的性质?

(a) $u(c_1,c_2,\pi_1,\pi_2)=a(\pi_1c_1+\pi_2c_2)$,

(b) $u(c_1,c_2,\pi_1,\pi_2)=\pi_1c_1+\pi_2c_2^2$,

(c) $u(c_1,c_2,\pi_1,\pi_2)=\pi_1\ln c_1+\ln\pi_2c_2+17$.

3.一个厌恶风险的人面临以下两项选择：一是 25%概率下得到 1000 元，75%概率下得到 100 元；二是确定得到 325 元。他应该选择哪一个?

4. 在上题中，如果第二项选择中的金额改为 320 元，他应该选择哪一个?

5. 画出描述下列行为的效用函数曲线：赌博额较小时是爱好风险的，赌博额较大时是厌恶风险的。

6.为什么同一个社区的人对于洪水损害比火灾损害更难相互承保?

复习题答案

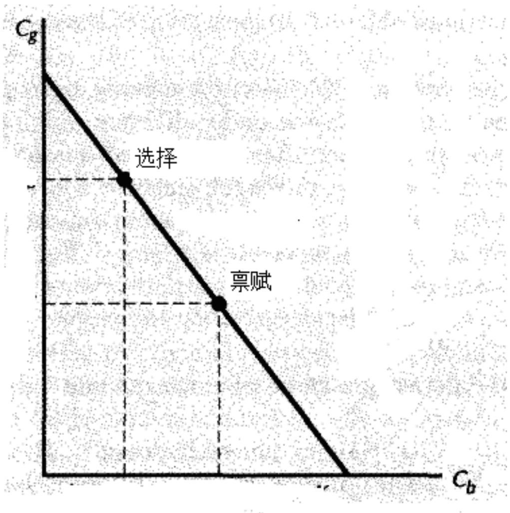
1.你的消费束如何才能位于图 12.1 禀赋点的左侧?

【复习内容】禀赋；转嫁风险

【参考答案】

如图所示。禀赋点表示（风险发生情形下的消费量，风险不发生情形下的消费量）。如果你的最优选择点在禀赋点左侧，见上图。这意味着，与禀赋相比，你在坏状态（风险发生）下的消费量 C_b 减少了；与禀赋相比，你在好状态（风险不发生）下的消费量 C_g 增加了。

曹乾（东南大学 caoqianseu@163.com）



怎样才能做到这一点？只要你是这种风险的购买者（或者说你销售承保这种风险的保险），那么就会出现上述情形（最优选择在禀赋点左侧）。为什么？因为比如别人转嫁价值1000 元的风险给你，你收取 100 元钱。如果风险不发生，相当于你的财富增加了 100 元，因此你的 C_g 加了；但是如果风险发生，你要赔偿别人 1000 元，因此你的 C_b 与禀赋相比大幅减少了。

2.下列哪些效用函数具有期望效用函数的性质？

(a) $u(c_1,c_2,\pi_1,\pi_2)=a(\pi_1 c_1+\pi_2 c_2)$,

(b) $u(c_1,c_2,\pi_1,\pi_2)=\pi_1 c_1+\pi_2 c_2^2$,

(c) $u(c_1,c_2,\pi_1,\pi_2)=\pi_1\ln c_1+\ln\pi_2 c_2+17$.

【复习内容】期望效用函数；正仿射变换。

【解题思路】

期望效用函数具有下面的形式：

$$u(c_1,c_2,\pi_1,\pi_2)=\pi_1 v(c_1)+\pi_2 v(c_2).$$

这就是说效用可以写成每种状态下的消费函数（$v(c_1)$ 和 $v(c_2)$）的加权和，其中权重分别为每种状态发生的概率（π_1 和 π_2）。

当我们说消费者的效用可用期望效用函数表示时，或者说消费者的偏好具有期望效用的性质时，我们的意思是说，我们可以选择一个具有上述可加性形式的效用函数。

需要注意，期望效用函数 u 在经过正仿射变换即 $w(u)=au+b$（其中 $a>0$）后，得到的这个正仿射函数 $w(u)$ 不仅还代表着相同的偏好，而且仍具有期望效用的性质。

还需要注意，正仿射变换只是正单调变换的一种特殊情形，但除了正放射变换之外，期望效用函数的其他单调变换都不再具有期望效用函数的性质。

【参考答案】

由以上解题思路可知，

（a）具有期望效用函数的性质，因为可将$u(c_1,c_2,\pi_1,\pi_2)=a(\pi_1c_1+\pi_2c_2)$看成由标准期望效用函数$z(c_1,c_2,\pi_1,\pi_2)=\pi_1c_1+\pi_2c_2$，经过正仿射变换$u(z)=az+b$而得到（只不过此处 b=0）。

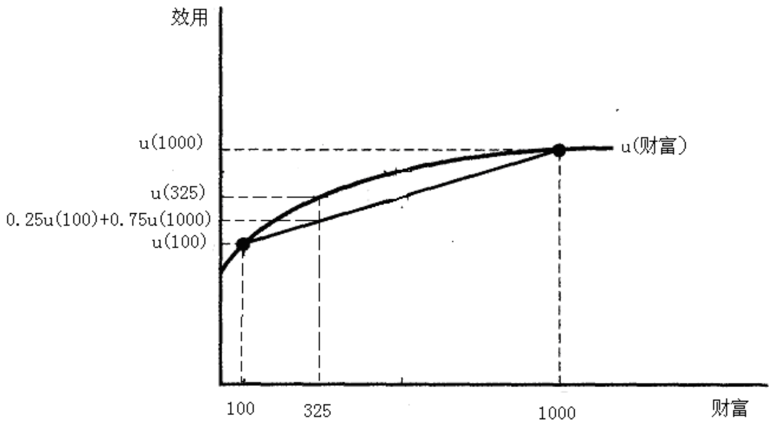
（b）不具有期望效用函数的性质，因为它不是标准的期望效用函数，也无法从某期望效用函数经过正仿射单调变换得到。请读者思考$u(c_1,c_2,\pi_1,\pi_2)=\pi_1c_1^2+\pi_2c_2^2$是不是具有期望效用函数性质？

（c）具有期望效用函数的性质，因为可将$u(c_1,c_2,\pi_1,\pi_2)=\pi_1\ln c_1+\ln\pi_2c_2+17$看成由标准效用函数$z(c_1,c_2,\pi_1,\pi_2)=\pi_1\ln c_1+\ln\pi_2c_2$，经过正仿射变换$u(z)=az+b$而得到，其中 a=1,b=17。

3.一个厌恶风险的人面临以下两项选择：一是 25%概率下得到 1000 元，75%概率下得到 100 元；二是确定得到 325 元。他应该选择哪一个？

【复习内容】

厌恶风险者；厌恶风险者的期望效用函数为凹。



【解题思路】

根据厌恶风险者的定义进行判断。厌恶风险者的效用函数具有下列特征：

$$u(\pi_1c_1+\pi_2c_2)>\pi_1u(c_1)+\pi_2u(c_2)$$

即对于厌恶风险者来说，期望财富的效用大于财富的期望效用。

从图形上看，厌恶风险者的期望效用函数为凹函数。

【参考答案】

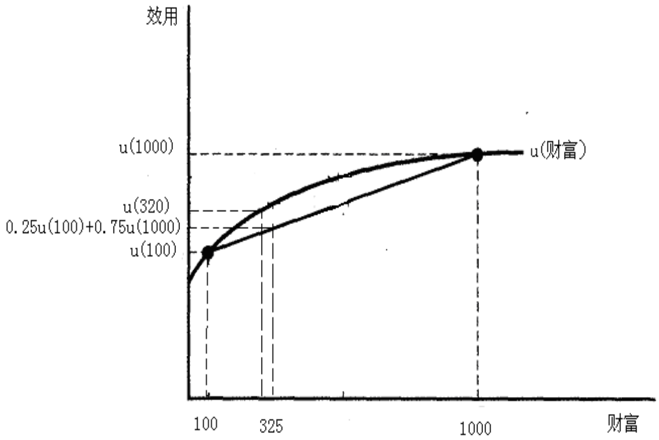
赌博时他的期望财富 $E_g = 0.25 \times 100 + 0.75 \times 1000 = 325$ 元；这与另外一种选项的财富数额（325 元）相等。但由于他是厌恶风险者，因此他的期望效用函数为凹的，见上图。此种情形下，必然有 $u(325) > 0.25u(100) + 0.75u(1000)$，所以他会选择确定情形下的 325 元而不会选择赌博。

4. 在上题中，如果第二项选择中的金额改为 320 元，他应该选择哪一个?

【复习内容】厌恶风险者；厌恶风险者的期望效用函数为凹函数

【解题思路】见上题。

【参考答案】



这种情形下，我们无法判断 u(320)和 $0.25u(100) + 0.75u(1000)$ 到底哪个更大。

如果消费者非常厌恶风险，那么从图形（如上图）上看他的期望效用函数会向下凹得更厉害，此时他会选择 320 元这种确定的报酬。

如果消费者厌恶风险，但不是非常厌恶（此时图形趋近于线性，即趋近于风险中性人的期望效用函数），这种情形的图形请读者自己补充。这种情形下，消费者会选择赌博。

因此，我们无法给出一个明确的答案，他到底选择哪个取决于他的期望效用函数凹的程度。

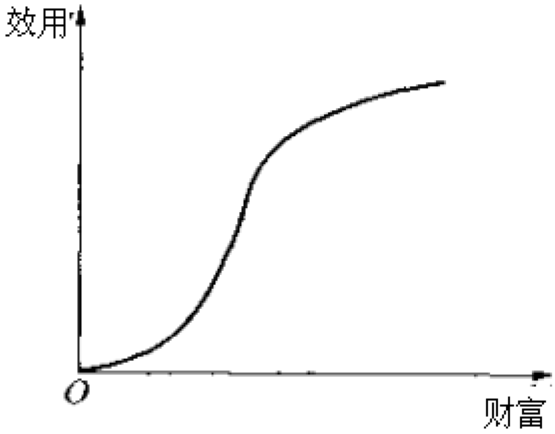
5. 画出描述下列行为的效用函数曲线：赌博额较小时是爱好风险的，赌博额较大时是厌恶风险的。

【复习内容】厌恶风险者；爱好风险者；厌恶风险者的期望效用函数是凹的；爱好风险者的期望效用函数是凸的。

【参考答案】

由于厌恶风险者的期望效用函数是凹的，爱好风险者的期望效用函数是凸的。因此根据

题目的要求，这样的图形容易画出，见下图。



在赌博数额较小时，他是喜好风险的，因此他的期望效用函数曲线是向下凸的；在赌博额较大时，他是厌恶风险的，因此他的期望效用函数曲线是向下凹的。

6.为什么同一个社区的人对于洪水损害比火灾损害更难相互承保？

【复习内容】互助保险；独立性假设

【参考答案】

互助保险有个重要的假设前提是独立性。以火灾为例，一家失火另一家未必失火，因此这两个事件是独立的。

然而在洪水灾害的情形下，独立性假设不成立，因为洪水一冲冲一片，所以洪水灾害很难相互承保。但这也不是绝对的，比如我们以沿江沿海地区为例，因为这些地区涉及很多省份，有可能今年某省份发生洪灾，另外的省份不发生。这种情形下，独立性的假设得以保证，这些省份可以互相承保洪水灾害。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

13.风险资产（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

13 风险资产

在上一章，我们分析了不确定性情形下的个人行为模型，以及保险市场和股票市场这两种经济制度的作用。在本章我们进一步分析股票市场如何分散风险的。为做此事，最好从一个简化的不确定性行为模型进行分析。

13.1 均值—方差效用

在上一章我们分析了不确定性情形下的选择问题，我们是用期望效用函数进行分析的。这样的问题还有另外一类分析方法，即用一些参数（parameters）描述选择的目标，然后将效用函数视为这些参数的函数。这类方法中最为流行的就是**均值—方差模型**（mean-variance model）。在均值—方差方法中，我们不再认为消费者的偏好取决于他的财富在每种可能结果上的整个概率分布，而是假设他的偏好可用几个关于他财富概率分布的统计量进行描述。

令随机变量w取值w_s的概率为π_s（其中$s=1,2,...,S$）。w概率分布的**均值**（mean）就是它的加权平均值：

$$\mu_w=\sum_{s=1}^{S}\pi_s w_s .$$

上式就是加权平均值的计算公式：每个结果w_s以它自身发生的概率π_s作为权重（即$\pi_s w_s$），然后全部相加[一]。

w概率分布的**方差**（variance）是$(w-\mu_w)^2$的加权平均值：

$$\sigma_w^2=\sum_{s=1}^{S}\pi_s(w-\mu_w)^2 .$$

方差衡量分布的“分散性”，因此可用来衡量风险。还有一种相近的衡量方法，称为**标准差**（standard deviation），用σ_w表示，它是方差的平方根：$\sigma_w=\sqrt{\sigma_w^2}$.

概率分布的均值衡量它的加权平均值，即这些分布围绕着的那个数值。概率分布的方差衡量分布的“分散性”，即这些分布离均值有多远。图 13.1 描述了不同均值和方差下的概率分布。

[一] μ和σ都是希腊字母，前者读作“mew”，后者读作“sig-ma”.

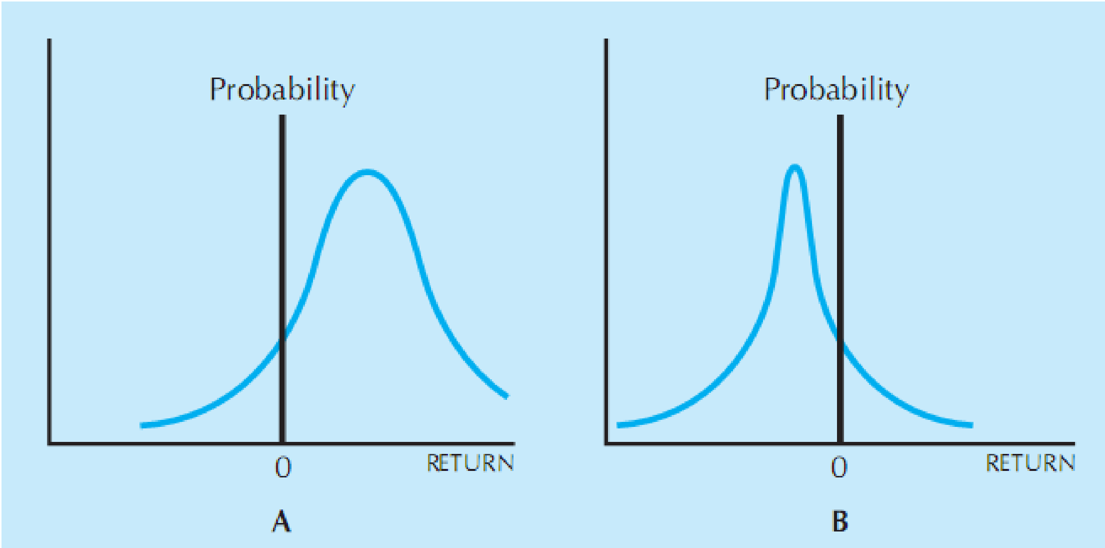


图 13.1：均值和方差。A 图的概率分布的均值为正，B 图的均值为负。A 图的分布比 B 图更“分散（spread out）”，这意味着 A 图的方差更大。

均值—方差模型假设概率分布效用（即投资者拥有财富 w_s 的概率为 π_s）可以表示为分布的均值和方差的函数即 $u(\mu_w,\sigma_w^2)$。或者，效用可以更方便地表示为分布的均值和标准差的函数 $u(\mu_w,\sigma_w)$。由于方差和标准差都可用于衡量财富分布的风险，在效用函数中使用哪一个都可以。

这个模型可以认为是简化版本的期望效用模型（我们在前几章已介绍过）。如果各种选择可以用均值和方差充分地刻画，那么均值和方差的效用函数 $u(\mu_w,\sigma_w^2)$ 对选择的排序，将和期望效用函数对这些选择的排序是一样的。而且，即使均值和方差不能充分地刻画概率分布，均值-方差模型也可以作为期望效用模型的合理近似，因此，均值-方差模型也是一种可行的方法。

我们自然可以假设：在其他条件相同的情况下，期望报酬越高越好；在其他条件相同的情况下，方差越小越好。这只是人们通常厌恶风险的另外一种表达方法。

我们使用均值-方差模型分析一个简单的组合投资问题。假设你可以投资两种不同的资产。一种资产是**无风险资产**（risk-free asset），这种资产的收益率恒为 r_f。国库券就非常类似无风险资产，因为无论何种情形发生，你都可以得到固定利率回报。

另外一种资产是**风险资产**（risky asset）。你可以将这种资产想象为你投资某大型共同基金，该基金主要做股票投资。如果股票市场行情好，你的投资回报就多。如果股票市场行情不好，你的投资回报就少。令 m_s 表示状态 s 发生时该资产的回报，令 π_s 表示状态 s 发生的概率。我们用 r_m 表示该风险资产的期望报酬，σ_m 表示报酬的标准差。

当然你没必要只选择其中一种资产进行投资；通常这两种资产你多少都会投资些。如果你投资风险资产的资金占你总资金的比例为 x，则你投资于无风险资产的资金比例为 $(1-x)$，你这个组合投资的期望报酬可由下式计算：

$$r_x = \sum_{s=1}^{S}(xm_s + (1-x)r_f)\pi_s$$
$$= x\sum_{s=1}^{S} m_s\pi_s + (1-x)r_f\sum_{s=1}^{S}\pi_s.$$

由于 $\sum_{s=1}^{S}\pi_s = 1$，我们有

$$r_x = xr_m + (1-x)r_f.$$

这样，投资组合的期望报酬等于这两种资产期望报酬的加权平均数。

你的投资组合报酬的方差为

$$\sigma_x^2 = \sum_{s=1}^{S}(xm_s + (1-x)r_f - r_x)^2\pi_s.$$

将 r_x 的表达式代入上式可得

$$\sigma_x^2 = \sum_{s=1}^{S}(xm_s - xr_m)^2\pi_s$$
$$= \sum_{s=1}^{S} x^2(m_s - r_m)^2\pi_s$$
$$= x^2\sigma_m^2.$$

因此，投资组合报酬的标准差为

$$\sigma_x = \sqrt{x^2\sigma_m^2} = x\sigma_m.$$

由于当风险资产的期望报酬小于无风险资产时，风险厌恶型的投资者决不会投资风险资产，因此我们自然可以假设 $r_m > r_f$。从我们上面推导出的期望报酬和方差（或标准差）的表达式可知，你如果加大风险资产的投资比例，你的期望报酬会增高，但是你的风险同样也增加。这个结论也可用图 13.2 说明。

如果你将所有的资金投于风险资产（$x=1$），那么你的期望报酬和标准差为 (r_m, σ_m)。如果你将所有资金投于无风险资产（$x=0$），则期望报酬和标准差为 $(r_f, 0)$。如果你两种资产各投资一些（x 介于 0 和 1 之间），那么这个投资组合将位于 $(r_f, 0)$ 点和 (r_m, σ_m) 点连线中间的某个位置。这条连线就是预算线，它描述了风险和报酬之间的市场交易，即高风险换取高报酬，低风险换取低报酬。

由于我们假设个人的偏好仅取他财富的均值和标准差，我们可以画出他的无差异曲线，这些曲线反映了消费者对风险和报酬的偏好。如果他是厌恶风险的，则：期望报酬越高，他

的状况越好；标准差越高，他的状况越差。这表明，标准差是对他来说是“厌恶品”。因此，无差异曲线的斜率为正，如图 13.2 表示。

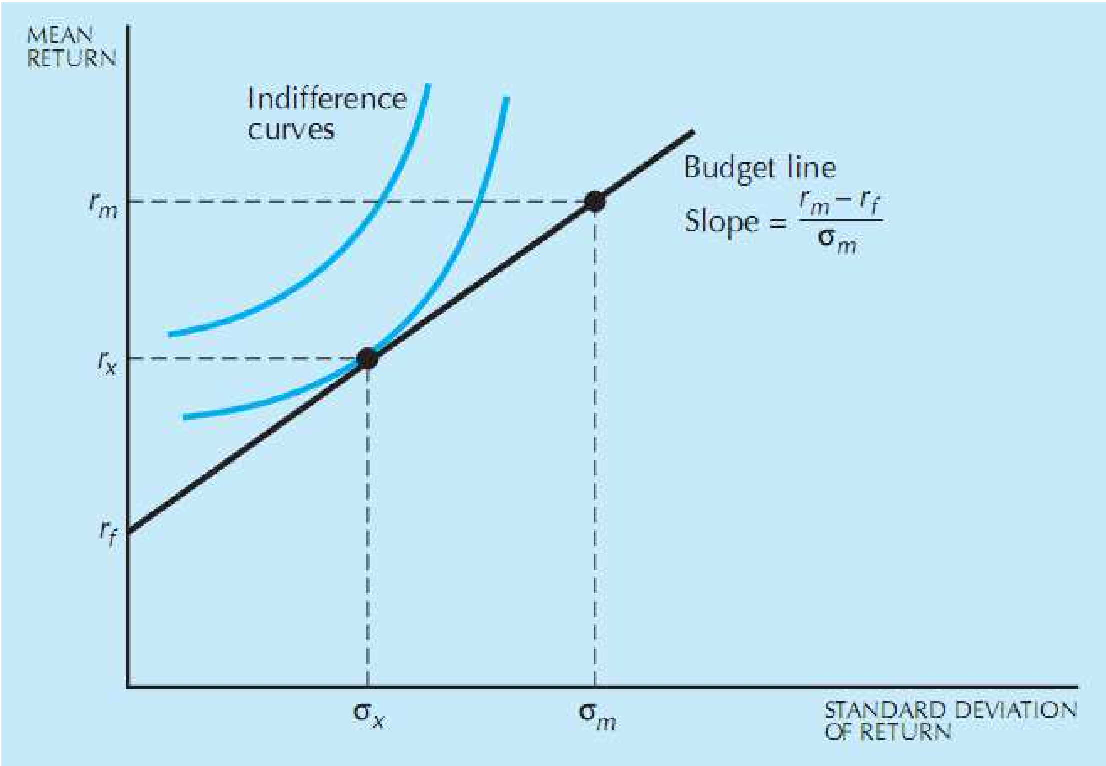


图 13.2：风险与报酬。预算线衡量期望收益增高时的成本（即风险），这个成本以标准差的增加量衡量。在最优选择处，无差异曲线必然和预算线相切。

在风险和报酬的最优选择处，无差异曲线的斜率必定等于预算线的斜率，请看如 13.2。我们将这条预算线的斜率称为**风险的价格**（price of risk），因为在做投资组合决策时，它衡量了风险和报酬是如何交易的。由图 13.2 可知，风险价格 p 的表达式为

$$p = \frac{r_m - r_f}{\sigma_m}. \tag{13.1}$$

因此，若无风险资产和风险资产的投资组合达到最优，则风险和报酬的边际替代率必然等于价格风险：

$$MRS = -\frac{\Delta U / \Delta \sigma}{\Delta U / \Delta \mu} = \frac{r_m - r_f}{\sigma_m}. \tag{13.2}$$

现在假设有很多人都投资于这两种资产。每个人的边际替代率都必须等于风险价格。这样，在均衡时，所有人的边际替代率都是相等的；如果人们有足够的机会相互交易风险，则每个人的风险均衡价格都会相等。在这个角度上，风险和一般商品没什么区别。

我们可以使用前几章开发的技术，分析当相关参数发生变动时最优投资组合是如何变

曹乾（东南大学 caoqianseu@163.com）

动的。我们在前面学过的正常商品、劣等商品、显示偏好等概念或工具，可以直接拿过来使用。例如，现在某人有机会投资风险资产 y，这个新资产的期望报酬为 r_y，标准差为 σ_y，如图 13.3 所示。

如果消费者投资于 x 和 y，消费者应如何选择？图 13.3 画出了原预算集和新预算集。注意：在原预算集内的风险和报酬的每个组合，在新预算线下都可以得到，这是因为新预算集包含原预算集。因此，x 和 y 的投资组合必然好于 x 和无风险资产的投资组合，因为消费者可以选择更好的投资组合。

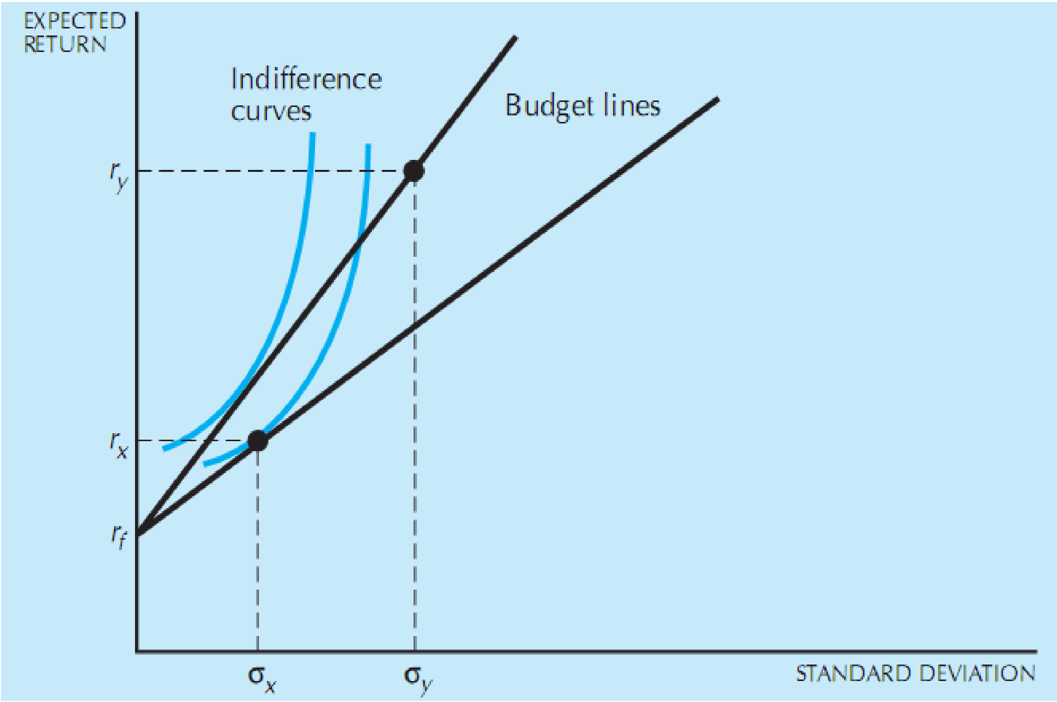


图 13.3: 对风险和报酬组合的偏好。x 和 y 的投资组合必然好于 x 和无风险资产的投资组合。

在上述结论中，要注意消费者能持有多少风险资产非常关键。如果规定消费者在 x 和 y 之间只能选择其一，结果就大不相同。在图 13.3 的情形中，消费者对 x 的偏好超过了 y，因为 x 位于更高的无差异曲线上，所以他会将全部资金都用于投资 x。但是如果允许他用无风险资产中和一部分风险，他会选择将无风险资产与 y 进行搭配，而不会选择将无风险资产与 x 进行搭配。

13.2 测量风险

在上面，我们已经有了描述风险价格的模型，但是如何测量某种资产的**风险大小**？你可能首先想到用该资产报酬的标准差测量风险。毕竟，我们前面假设效用取决于资产报酬的均值和标准差。

在上一节，由于风险资产只有一种，所以它的风险量就是标准差。但是，如果有很多种

风险资产，我们就不能再用标准差衡量某种资产的风险大小。

这是因为在这种情形下，消费者的效用取决于他持有的所有资产的均值和标准差，而不是其中某个资产的均值和标准差。此时，由于不同资产会**互相作用**，因此研究这种作用对总资产报酬的均值和标准差的影响就非常重要。和前面学过的思想一样，某资产的价值取决于它对总效用的边际影响，而不是取决于这种资产本身的价值。正如额外一杯咖啡的价值可能取决于你拥有的奶油数量一样，消费者对额外一单位某风险资产的支付意愿，取决于该资产和投资组合中其他资产的相互作用。

例如，假设你打算购买两种资产，而且你知道每种资产只有下面两种可能的结果：资产 A 的价值可能为 10 元，也可能为 -5 元；资产 B 的价值可能为 -5 元，也可能为 10 元。但是，如果资产 A 的价值为 10 元，资产 B 的价值就为 -5 元；反过来也是如此。也就是说，这两种资产的价值是**负相关**的（negatively correlated）:当一种资产价值较大时，另外一种资产的价值就较低。

假设每种资产上述两个结果的发生概率均为 50%，因此每种资产的平均价值都为 2.50 元。如果你是风险中立者而且规定你只能购买其中一种资产，你会花多少钱购买？答案为：对每种资产你都愿意最多支付 2.50 元，这恰好等于每种资产的期望价值。如果你是风险厌恶者，你的出价可能小于 2.50 元。

但是，如果你能购买这两种资产，收益是怎样的？如果上述两种资产，每种你都各买一份，则不管哪种结果发生，你都会得到 5 元，因为上面我们假设这两种资产价值是负相关的，即一种资产价值为 10 元，另外一种资产的价值只能为 -5 元。所以，如果允许你持有两种资产，你对这两种资产（都为一份）的出价为 5 元。

这个例子说明，某资产的价值通常取决于它和其它资产的相关性。负相关的资产即价值运动方向相反的资产，非常珍贵，因为它们能减少总体风险。一般来说，某种资产的价值更多地取决于它的报酬和其他资产的报酬的相关性，而不是取决于与它自身标准差的相关性。因此，某资产的风险大小取决于它和其他资产的相关性。

测度资产风险大小的一种简便方法，是将这种资产的风险与股票市场的总体风险（即市场风险）进行比较，这就是相对风险。某种股票的**相对风险**，是指相对于股票市场的风险来说，它的风险有多大。我们将这种股票的相对风险称为它的**贝塔值**（beta），并用希腊字母 β 表示。因此，如果我们用 i 表示某种股票，则 β_i 就是这种股票的相对风险。大致来说：

$$\beta_i = \frac{\text{资产}i\text{的风险}}{\text{股票市场的风险}}$$

如果某股票的 β 值等于 1，则它的风险和股票市场的总体风险一样大。如果股票市场的风险增加了 10%，这种股票的风险平均也增加 10%。如果某股票的 β 值小于 1，则股票市场风险增加 10%，这种股票的风险增加程度一般小于 10%。股票的 β 值可用统计方法估计，这

种方法大致来说就是测量某个变量对另外一个变量变动的敏感程度[一]。很多投资咨询机构都会帮你估算股票的β值。

13.3 合同对方不履行合约责任的风险

金融机构不仅借款给个人，彼此之间也会借款。一方总有可能无法偿还借款，这样的风险称为**合同对方不履行合约责任的风险**(counterparty risk)。

为了看清这种风险是如何产生的，假设有 3 个银行，分别称为 A, B 和 C。银行 A 欠 B10 亿元，银行 B 欠 C10 亿元，银行 C 欠 A10 亿元。现在假设银行 A 用光了钱，无法还债。银行 B 现在就有了 10 亿元坏账，可能因此无法偿还它欠银行 C 的钱，从而银行 C 有可能无法偿还它欠银行 A 的钱，这将银行 A 进一步推向了困境。这种效应称为**金融传染**（financial contagion）或**系统风险**（systemic risk）。这是美国金融机构在 2008 年秋季发生的危机事件的简化版本。

如何化解这种风险？其中一种方法是设立“最终贷款者”（lender of last resort），这个最终贷款者通常为中央银行，例如美国的联邦储备系统。银行 A 可向联邦储备申请 10 亿元的紧急借款。现在它有钱偿还银行 B，从而银行 B 可以偿还银行 C，最终银行 C 可以偿还银行 A。现在银行 A 有足够的钱归还从中央银行的借款。

当然，这是一个极其简化的例子。事实上，这三个银行之间不存在**净**债务。如果它们在一起比较各自的资产和负债，他们肯定会发现这个事实。然而，当资产和负债涉及成千上万个金融机构时，很难确定各自的净头寸（net positions），这就是为何有必要设立最后贷款者的原因。

13.4 风险资产的市场均衡

现在我们可以分析风险资产的市场均衡条件了。在 11 章我们已知道，无风险的条件下，所有资产的报酬率必须相等。此处我们有个类似的原理：所有资产，经过风险调整之后，它们的报酬率必须相等。

技巧在于如何调整风险。我们应该如何调整？我们已学过最优选择的知识，答案就隐藏在这些知识里面。回忆我们前面讲过的投资组合的最优选择问题，这种投资组合包括一种无风险资产和一种无风险资产。上述投资组合中的风险资产可以看成一种共同基金，即看成一种多元化的投资组合，这种组合包含很多风险资产。在本节，我们假设，这个投资组合包括**所有**的风险资产。

既然上述投资组合包括了所有的风险资产，那它和整个市场没什么区别，因此它的期望

[一] 如果你学过统计学，你就会知道股票 i 的β值的定义为：$\beta_i = \mathrm{cov}(\tilde{r}_i, \tilde{r}_m)/\mathrm{var}(\tilde{r}_m)$。此式中的分子是股票 i 的报酬与市场报酬之间的协方差，分母是市场报酬的方差。下标 m 表示市场。

报酬等于市场期望报酬r_m，它的标准差等于市场的标准差σ_m。假设无风险资产的报酬率为r_f。

（13.1）式表明风险价格 p 为

$$p = \frac{r_m - r_f}{\sigma_m}.$$

我们在前面已知道，风险资产 i 的相对风险为β_i，这个风险是相对于市场风险来说的。由相对风险β_i的定义可知，为了计算资产 i 的**总风险**（即绝对风险），必须将β_i乘以市场风险σ_m，因此资产i的总风险为$\beta_i\sigma_m$.

资产i的总风险的成本为多大？将$\beta_i\sigma_m$乘以风险价格即可得到这个成本。这个成本称为**风险调整**或风险校正（risk adjustment）：

$$\begin{aligned}\text{风险调整} &= \beta_i\sigma_m p \\ &= \beta_i\sigma_m\frac{r_m - r_f}{\sigma_m} \\ &= \beta_i(r_m - r_f)\end{aligned}$$

现在我们就可以说出风险资产的市场均衡条件了：均衡时，所有资产的经过风险调整的报酬率（risk-adjusted rate of return）必相等。这里的逻辑和第 12 章的逻辑是一样的：如果资产i经过风险调整的报酬率高于资产 j，则所有人都想持有资产i。因此，均衡时，经过风险调整的报酬率必然相等。

如果有两种资产i和 j，它们的期望报酬率分别为r_i和r_j，贝塔值分别为β_i和β_j，则在均衡时，必然有

$$r_i - \beta_i(r_m - r_f) = r_j - \beta_j(r_m - r_f).$$

这个式子是说，均衡时，这两种资产的经过风险调整的报酬率必须相等。上式中，风险调整$\beta_i(r_m - r_f)$和$\beta_j(r_m - r_f)$分别等于的各自资产的总风险乘以风险价格。

上述均衡条件还有另外一种表达方法。首先注意到，对于无风险资产来说，$\beta_f = 0$，因为根据定义，无风险资产的风险为零。因此，对于任何资产i，必有

$$r_i - \beta_i(r_m - r_f) = r_f - \beta_f(r_m - r_f) = r_f.$$

整理可得

$$r_i = r_f + \beta_i(r_m - r_f).$$

这个式子是说，任何资产的期望报酬率，必然等于无风险资产的报酬率r_f加上风险调整$\beta_i(r_m - r_f)$。上式右端第二项即$\beta_i(r_m - r_f)$，表示消费者持有风险资产时要求得到的额外报酬。上式其实是**资本资产定价模型**（Capital Asset Pricing Model, CAPM）中的一个主要

公式，资本资产定价模型在金融市场的研究中用途广泛。

13.5 如何调整报酬

在研究无风险的资产市场时，我们已说过资产的价格是如何调整最终等于报酬的。下面我们分析相同的调整过程。

根据上一节的模型可知，任何资产的期望报酬应该等于无风险报酬加上风险溢价（risk premium）：

$$r_i = r_f + \beta_i(r_m - r_f).$$

在图 13.4 中，我们画出了这条直线，其中横轴表示 β 值，纵轴表示期望报酬。我们前面已指出，所有资产的经过风险调整的报酬率必相等。因此，均衡时，所有资产都必然位于这条线上。这条线叫做**市场线**（market line）。

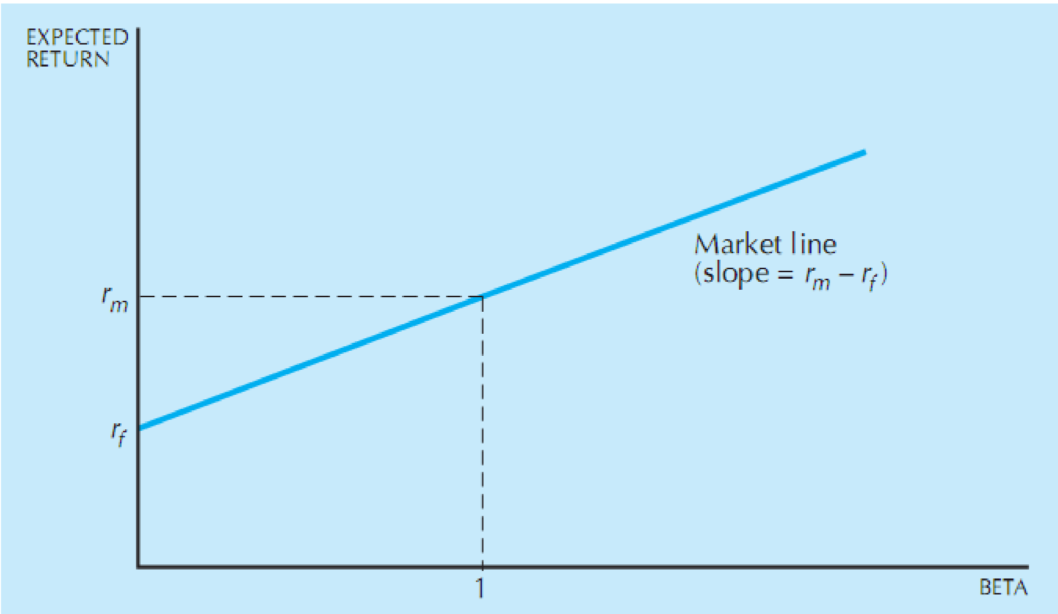


图 13.4：市场线。市场线衡量均衡时资产的期望报酬和风险（贝塔值）。

如果某些资产的期望报酬和贝塔值不在这条线上，结果会如何？

资产 i 的期望报酬等于它的期望价格变动量除以它的当前价格：

$$r_i = \frac{p_1 - p_0}{p_0}\text{的期望价值}.$$

这个定义和第 11 章的定义非常类似，只不过这里增加了“期望”这个词。由于资产将来的

 曹乾（东南大学 caoqianseu@163.com）

价格是不确定的，因此“期望”这个词不可省略。

假设你有幸发现了某个资产，它的经过风险调整后的期望报酬，高于无风险资产，即

$$r_i - \beta_i(r_m - r_f) > r_f .$$

那么，这个资产真是个好东西。原因就是因为它的经过调整后的期望报酬较高。

当人们发现了这样的资产，他们就希望购买。他们想持有或者买过来再卖。由于这种资产比现有资产更好，因此这种资产肯定有市场。

但是当人们争相购买这种资产时，该资产的价格 p_0 就会上升。这意味着期望报酬 $r_i = (p_1 - p_0) / p_0$ 会下降。下降多少？下降到正好使期望报酬率回落到市场线上。

因此，购买位于市场线上方的那些资产就能赚钱。为什么？因为当人们发现在风险一定的情况下，这种资产的报酬大于现有资产的报酬，他们就会争相购买，该资产的价格因此会上升。

当然，以上的分析有个假设前提：人们对各种风险资产的看法是一致的。如果他们对这些资产的期望报酬或风险看法不一，那么这个模型将会变得异常复杂。

例子：在险价值

有时人们对如何确定一组资产的风险感兴趣。例如，假设某家银行持有由若干种股票组成的投资组合。该银行想确定这个投资组合在某天的价值下降一百万元的概率。如果这个概率为 5%，我们就说该投资组合具有“100 万元在某天出险的概率为 5%”（one-day 5% **value at risk** of $1 million）的性质。通常**在险价值**（value at risk, VaR）的计算时段为一天或两周，使用的概率为 1%或 5%.

VaR 的理论思想比较吸引人。最大的难题是找到估计它的方法。但是，正如金融分析师菲利普.杰瑞恩（Phillippe Jorion）所说，“VaR 的最大好处是让人们用结构性的方法论来思考风险。完成计算 VaR 过程的金融机构被迫面对它们暴露的金融风险，并且建立合适的风险管理功能。因此，计算 VaR 的过程可能和计算出的数字结果同等重要。”

VaR 完全决定于投资组合价值的概率分布，而这个概率分布又取决于投资组合中各种资产的相关关系。通常，各种资产是正相关的，因此它们的价值会共同上升或下降。更糟糕的是，资产价格分布通常有“厚尾”，这意味着发生极端价格变动的概率相对较高。理想情形下，人们可以使用价格变动的长期历史数据来估计 VaR。然而，在实践中这难以做到，尤其对于新的外国资产来说。

在 2008 年秋季的美国金融危机中，很多金融机构发现它们的 VaR 估计方法存在着严重缺陷，因为资产价格下降的幅度远远超过了它们的预期。估计方法的缺陷部分源自下列事实：

金融机构的统计估计是基于非常小的样本之上的，而且这些样本是在经济稳定时期收集的。金融机构计算出的 VaR 低估了资产的真正风险。

例子：共同基金的评级

我们可以使用资本资产定价模型，比较不同资产的风险和报酬。一种比较流行的投资方式是共同基金。共同基金一般是大型机构，它从个人投资者手里吸收资金，再把这些钱用于买卖其他公司的股票，然后把在股市里赚取的利润返还给个人投资者。

共同基金的优点是专家理财，缺点是他们收取管理费。由于这些管理费通常不是特别高，所以大多数小投资者一般会购买共同基金。

但是你如何选取共同基金？你希望期望报酬越高越好，风险越低越好。问题是，为得到较高的期望报酬，你愿意忍受多大的风险？

你可以去查询不同基金的历史业绩，计算它们的平均年报酬率和贝塔值（风险）。由于我们这里没给出贝塔值的计算方法，所以你不会计算，这不要紧，因为你可以从某些书籍中查询这些共同基金的历史贝塔值。

如果你将期望报酬和相应的贝塔值画出来，你就得到了类似图 13.5 的图形[一]。注意，期望报酬越高的基金通常风险也越高，高期望报酬是对人们承担风险的补偿。

使用共同基金图你可以做一件比较有趣的事情，即比较下列两种投资方式的优劣：一是共同基金这种专家理财的投资方式；二是投资于**指数基金**（index fund）这类非常简单的投资方式。

股票市场指数的种类有好几种，比如道-琼斯工业平均指数、标准-普尔指数等等。股市指数通常是一组股票在某既定日期的平均报酬。例如，标准-普尔 500 指数是基于美国 500 只主要股票的平均业绩编制的。

指数基金也是一种共同基金，它持有一部分或全部的某指数所包含的股票（即指数成份股）。由该定义可知，你若购买了指数基金，你获得的报酬应大致等于指数成份股的平均报酬。这也意味着指数基金取得平均报酬这样的业绩并不困难，因此，指数基金的管理费比较低。由于指数基金持有的股票非常广泛，因此它的贝塔值非常接近于 1，也就是说它的风险和股票市场的总体风险差不多。

与典型共同基金相比，指数基金的业绩如何？注意，在比较它们的业绩时，你不仅应该比较报酬还应比较风险。做此事的一种方法是，画出标准-普尔指数基金的期望报酬和贝塔值，然后将它们与无风险报酬率连接起来，如图 13.5 所示。你可以得到这条线上的任何一

[一] 对共同基金的业绩的详细讨论可参见 Michael Jensen, "The Performance of Mutual Funds in the Period 1945-1964," Journal of Finance, 23 (May 1968), 389-416；更近年代的分析请见 Mark Grinblatt & Sheridan Titman "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings," The Journal of Business, 62 (July 1989), 393-416.

个风险-报酬组合，这取决于你对无风险资产和指数基金各投资了多少钱。

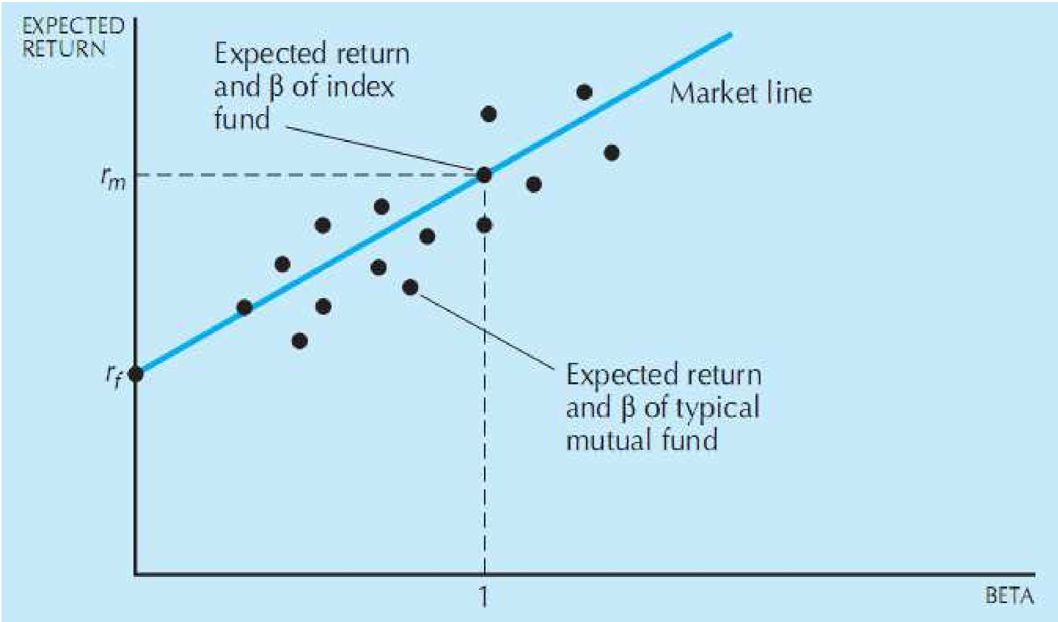


图 13.5：共同基金。将共同基金的投资报酬与市场线进行比较。

现在我们数一数位于这条线下方的共同基金的数量。这些共同基金提供的风险-报酬组合比指数基金和无风险资产投资组合的业绩差。数完了我们就知道，共同基金提供的绝大多数风险-投资组合位于这条线的下方。位于这条线上方的共同基金数量几乎没有。

从另外一个角度看，这种发现并不奇怪。股票市场是个高度竞争的市场。人们总是想尽一切办法试图找到价值被低估的股票，从而买进获利。这意味着，平均来说，所有股票的价格接近于它们各自的实际价值。如果情形的确如此，购买指数基金从而获得平均报酬，就是个相当合理的策略，因为该情形下你很难获得超过平均业绩的报酬。

总结

1.我们可以使用以前章节介绍过的预算集和无差异曲线工具，分析投资组合决策：风险资产和无风险资产应各投资多少钱。

2.风险和报酬的边际替代率必须等于预算线的斜率。这个斜率称为风险的价格。

3.某种资产的风险大小主要取决于它与其他资产的相关性。负相关的资产能帮助你减少投资组合的总体风险。

4.某种资产相对于市场总体风险的风险，称为该资产的贝塔值。

5.资产市场基本均衡条件是，所有资产的经过风险调整的报酬必须相等。

6. 合同对方不履行合约责任的风险是指其他交易方不付钱，这种风险也是一种重要的风险因素。

复习题

1.如果无风险的报酬率为 6%；某风险资产的报酬率为 9%，标准差为 3%；如果你只愿意接受标准差为 2%的风险，（1）计算你投资于该风险资产的资金比率；（2）计算你能实现的最大报酬率。

2.在上题中，风险价格为多大？

3.如果某股票的贝塔值为 1.5，市场报酬率为 10%，无风险报酬率为 5%，根据资本资产定价模型计算该股票的期望报酬率。如果该股票的期望价值为 100 元，计算该股票的今日价格。

复习题答案

1.如果无风险的报酬率为 6%；某风险资产的报酬率为 9%，标准差为 3%；如果你只愿意接受标准差为 2%的风险，（1）计算你投资于该风险资产的资金比率；（2）计算你能实现的最大报酬率。

【复习内容】均值-方差模型；投资组合报酬的均值和标准差的计算公式

令 r_f 表示无风险资产的报酬率，r_m 表示风险资产的期望报酬率，你投资风险资产和无风险资产的资金占你总资金的比例为分别 x 和 $(1-x)$。

投资组合的期望报酬等于这两种资产期望报酬的加权平均数：

$$r_x = xr_m + (1-x)r_f.$$

投资组合的方差和标准差分别为

$$\sigma_x^2 = x^2\sigma_m^2.$$

$$\sigma_x = \sqrt{x^2\sigma_m^2} = x\sigma_m.$$

【参考答案】

（1）由$\sigma_x=\sqrt{x^2\sigma_m^2}=x\sigma_m$可知，投资于风险资产的资金比例$x=\frac{\sigma_x}{\sigma_m}=\frac{2\%}{3\%}=\frac{2}{3}$；

（2）你能得到的最大报酬为投资组合的期望报酬

$$r_x=xr_m+(1-x)r_f=\frac{2}{3}\cdot 9\%+\frac{1}{3}\cdot 6\%=8\%.$$

2.在上题中，风险价格为多大？

【复习内容】风险价格的概念；风险价格的计算公式

【参考答案】

$$p=\frac{r_m-r_f}{\sigma_m}=\frac{9\%-6\%}{3\%}=1$$

3.如果某股票的贝塔值为 1.5，市场报酬率为 10%，无风险报酬率为 5%，根据资本资产定价模型计算该股票的期望报酬率。如果该股票的期望价值为 100 元，计算该股票的今日价格。

【复习内容】资本资产定价模型

【参考答案】

$r_i=r_f+\beta_i(r_m-r_f)=5\%+1.5(10\%-5\%)=12.5\%$

如果股票的期望价值为 100 元，则它的价格为：

$$\frac{100}{1+12.5\%}=88.89(\text{元})$$

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

14.消费者剩余（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

14 消费者剩余

在前面几章我们已经知道：如何从消费者不可观测的偏好或效用函数推导出他的需求函数。但在实践中，我们通常关心相反的问题——如何从观察到的消费者的需求行为估测他的偏好或效用。

事实上，我们在第 5 章和第 7 章已分析了这样的问题。在第 5 章，我们学习了如何从消费者需求的观测数据估计效用函数的参数。比如，在柯布-道格拉斯类型的偏好中，我们可以估计出描述消费者选择行为的效用函数。我们是如何做到这一点的？只要计算每种商品的支出占消费者收入的比例即可。根据推导出的效用函数，我们可以估计消费的变动。

在第 7 章，我们从消费者可观测到的选择行为入手，阐述了如何使用显示偏好这个工具还原消费者产生上述行为的潜在偏好。还原出的无差异曲线可用来估测消费变动。

在本章，我们介绍从可观测的需求行为推知消费者的效用的其他一些方法。尽管有些方法不像第 5 章和第 7 章的方法那样具有一般性，但以后你就会知道在本书后面的内容中，本章介绍的方法比较有用。

我们从一种特殊的需求行为入手分析，这种需求行为可以让我们比较容易地还原效用。然后，我们再分析偏好和需求行为更一般的情形。

14.1 离散商品的需求

第 6 章我们介绍过拟线性效用情形下的离散商品的需求问题，我们就从这个问题开始分析。假设效用函数的形式为 $v(x)+y$，并且商品 x 是离散的，也就是说只能以整数计算它的数量。令商品 y 表示花费在其他商品上的钱数，它的价格自然为1。令 p 表示商品 x 的价格。

在第 6 章我们已知道，此情形下消费者的行为可用保留价格进行描述，$r_1=v(1)-v(0)$，$r_2=v(2)-v(1)$，以次类推。保留价格和需求之间的关系非常简单：如果离散商品的需求量为 n 单位，则 $r_n \ge p \ge r_{n+1}$。

我们来验证一下上述结论。假设商品 x 的价格为 p 时，消费者选择消费 6 单位商品 x。由于这是该价格水平下他的最优选择，因此，消费 $(6, m-6p)$ 的效用必定大于或等于消费其他消费束 $(x, m-px)$ 的效用：

$$v(6)+m-6p \ge v(x)+m-px. \qquad (14.1)$$

特别地，当 $x=5$ 时，上述不等式也必须成立，因此

$$v(6)+m-6p \ge v(5)+m-5p.$$

整理即可得到

$$v(6)-v(5)=r_6 \geq p.$$

同理，当 $x=7$ 时，不等式（14.1）也必须成立，即

$$v(6)+m-6p \geq v(7)+m-7p,$$

整理可得

$$p \geq v(7)-v(6)=r_7.$$

上述分析过程表明，如果消费者需求 6 单位商品 x，则商品 x 的价格必定位于 r_6 和 r_7 之间。一般地，如果商品 x 在价格为 p 时的需求量为 n，则 $r_n \geq p \geq r_{n+1}$，这正是我们想证明的。根据这些保留价格包含的信息，我们就可以描述消费者的需求行为。这些保留价格在图形上呈现出“楼梯”的形状，如图 14.1 所示。这个楼梯图正是离散商品的需求曲线。

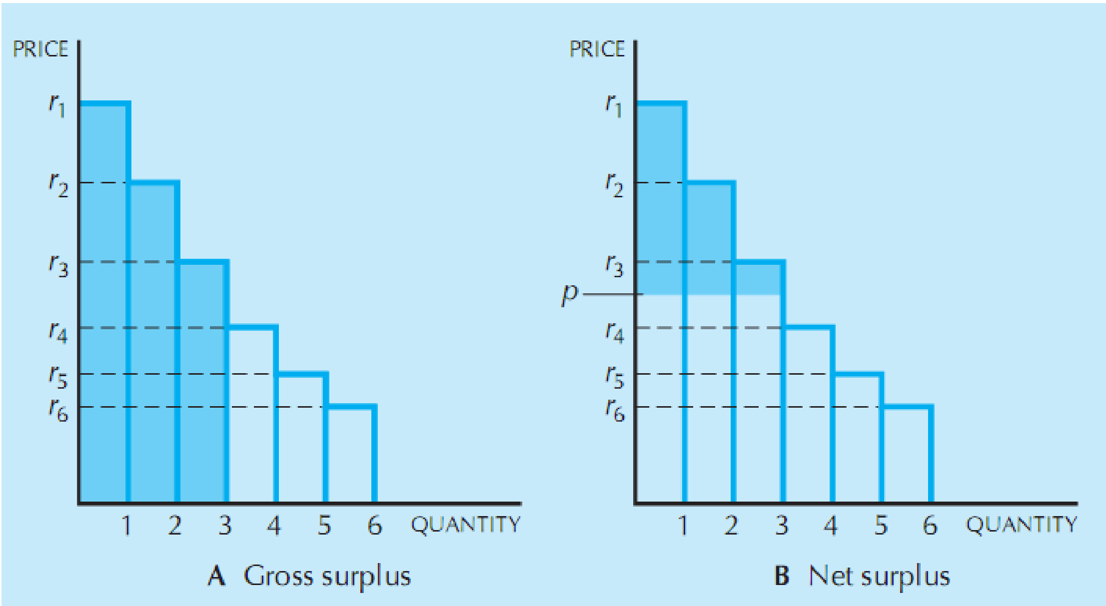


图 14.1：保留价格和消费者剩余。图 A 表示的是总收益（总消费者剩余），它是需求曲线以下区域的面积。总收益衡量消费者消费商品 x 的效用。图 B 表示的是净收益（消费者剩余）。净收益衡量消费者消费商品 x 和商品 y 的效用，其中商品 x 的价格 p 固定不变。

14.2 从需求构建效用

在上面，我们已经知道如何从保留价格或者效用函数构建需求曲线。但是，反过来，根据需求曲线能不能构建效用函数？可以，至少在拟线性效用的情形下是可以的。

在某种意义上来说，这种构建过程不过是平凡的算术运算。由于保留价格的定义为效用之差：

 曹乾（东南大学 caoqianseu@163.com）

$$r_1 = v(1) - v(0)$$

$$r_2 = v(2) - v(1)$$

$$r_3 = v(3) - v(2)$$

$$\vdots$$

例如，如果我们想计算$v(3)$，我们只要将上述一系列等式左右两端分别相加即可：

$$r_1 + r_2 + r_3 = v(3) - v(0).$$

为方便起见，假设消费零单位离散商品的效用为 0，即$v(0) = 0$，这样上式表明：$v(n)$等于前n个保留价格之和。

上述结论的几何解释很漂亮，请看图 14.1A。消费n单位离散商品的效用，恰好等于组成需求函数的前n个长条（bars）的面积。为什么？因为每个长条的长度等于与相应需求水平的保留价格，而长条的宽等于 1。这个面积有时称为消费既定数量某商品的**总收益**（gross benefit）或者**总消费者（的）剩余**（gross consumer's surplus）。

注意这只是消费商品 1 的效用。消费的最终效用取决于消费者对商品 1 和商品 2 的消费量。如果消费者选择消费n单位离散商品，那么他还剩下$m - pn$元钱去购买其他商品。在拟线性效用的情形下，他的总效用为：

$$v(n) + m - pn.$$

这个总效用同样可用面积进行解释：将图 14.1A 中的面积，减去花费在离散商品上的钱数，然后再加上m。

$v(n) - pn$称为**消费者（的）剩余**（consumer's surplus）或者**净消费者剩余**（net consumer's surplus）。它衡量某消费者消费n单位离散商品的净收益（net benefits）。为什么？因为他从n单位离散商品x身上得到的效用为$v(n)$，但他的花费为pn元，这意味着其他商品y的消费量因此减少了pn单位（因为y的价格为 1），在拟线性效用$u(x, y) = v(x) + y$的假设下，相当于效用损失了pn单位。因此他消费n单位离散商品的净效用（净收益）为$v(n) - pn$。

14.3 消费者剩余的其他解释方法

消费者剩余还有另外的一些解释方法。假设离散商品的价格为p。消费者认为第一单位该商品的价值为r_1元（保留价格），但他只花费p元就可以买到，因此购买第一单位后他得到的“剩余”为$r_1 - p$。他认为第二单位该商品的价值为r_2元，但他仍然只花费p元就可以买到，因此第二单位商品提供的“剩余”为$r_2 - p$。以此类推。如果我们把前n单位商品提供的剩余相加，就得到了消费者剩余：

曹乾（东南大学 caoqianseu@163.com）

$$CS = r_1 - p + r_2 - p + \cdots + r_n - p = r_1 + r_2 + \cdots + r_n - np.$$

由于保留价格之和就是消费该离散商品的效用，我们可以把上式写为

$$CS = v(n) - pn.$$

下面再介绍一种解释方法。假设某消费者消费 n 单位离散商品，因此花费 pn 元。如果我们想让他完全放弃这 n 单位商品的消费，需要补偿他多少钱？令 R 表示补偿资金总额，则 R 必须满足下式

$$v(0) + m + R = v(n) + m - pn.$$

由于我们假设 $v(0) = 0$，上式变为

$$R = v(n) - pn,$$

这正是消费者剩余。因此，消费者剩余衡量的是，如果让消费者放弃某商品的全部消费量，应该补偿他的钱数。但是需要注意，自始至终，我们都是在拟线性偏好的框架内分析消费者剩余的。

14.4 从消费者的剩余到消费者群体的剩余

直到目前，我们分析的都是一个消费者的情形，也就是说前面所说的消费者剩余(consume**r's** surplus)，是指某个消费者的剩余。如果涉及多个消费者，我们可以将每个消费者的剩余加总，这样就得到了**消费者群体的剩余**（consumer**s'** surplus）。注意这两个概念的区别：消费者的剩余指的是单个消费者的剩余；而消费者群体的剩余指的是若干消费者的剩余之和[一]。

正如消费者剩余可以衡量个体从交易中的收益大小，消费者群体的剩余可以衡量若干消费者在交易中的收益之和。

14.5 连续需求函数的近似

我们已经知道，某种离散商品需求曲线下方的面积，衡量消费这种商品的效用。我们可以将这个结论推广到连续商品的情形，方法是利用阶梯型需求曲线去近似（approximate）连续的需求曲线。因此，连续需求曲线下方的面积近似等于阶梯型需求曲线下方的面积。

图 14.2 说明了上面的思想。在本章附录，我们介绍如何使用微积分准确计算需求曲线

[一] 这两个概念在一般文献中都统称为“消费者（的）剩余”，因为可以根据上下文判断到底指的的是某个消费者的剩余还是消费者群体的剩余。比如在研究市场时，我们所称的消费者剩余显然指的是消费者群体的剩余。译者注。

下方的面积。

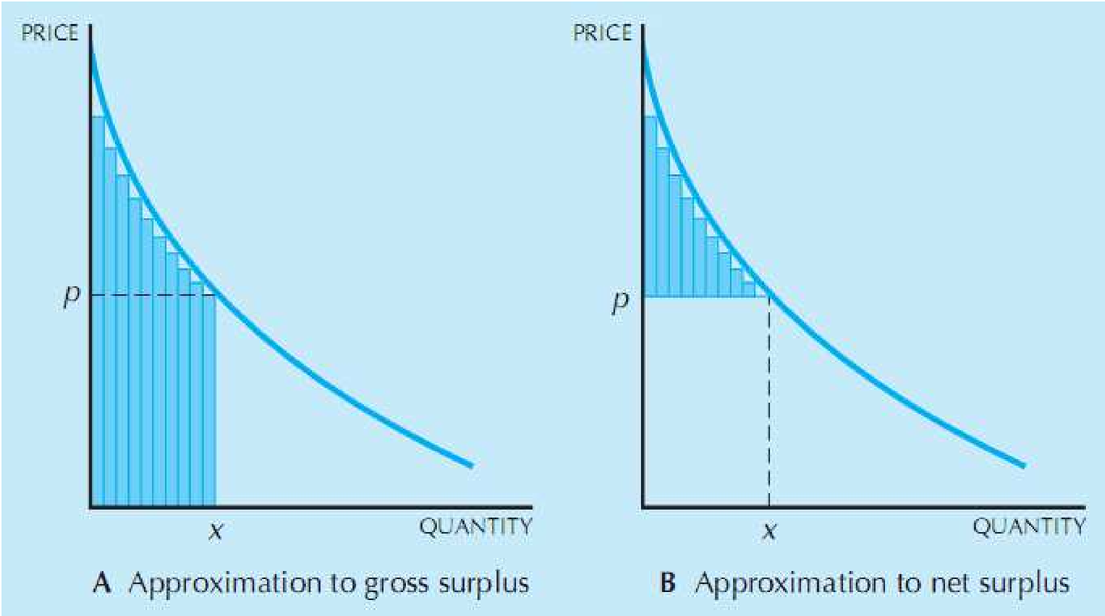


图 14.2：连续需求曲线的近似。连续需求曲线情形下的消费者剩余，可用离散需求曲线情形下的消费者剩余近似。

14.6 拟线性效用

有必要仔细想想拟线性效用在分析消费者剩余问题中的作用。一般来说，消费者愿意以多高的价格购买商品 1，取决于他还剩多少钱来购买其他商品。这意味着，商品 1 的保留价格通常取决于商品 2 的消费量。

但在拟线性效用 $u(x_1,x_2)=v(x_1)+x_2$ 这种特殊的情形下，商品 1 的保留价格和消费者花费在其他商品的钱数无关。在拟线性效用的情形下，经济学家称商品 1 的“收入效应为 0”或商品 1“不存在收入效应”，因为收入变动不会影响商品 1 的需求。这就是我们在前面几节为何能那样简易计算效用的原因。事实上，只有效用函数为拟线性时，你才能用需求曲线下方的面积衡量效用。

但是在某些情形下，可以用需求曲线下方的面积近似效用。比如，当收入变动时，如果某种商品的需求变化不大，也就是说收入效应较小时，消费者剩余的变动就是他效用变动的合理近似[一]。

14.7 对消费者剩余变动的解释

通常，我们并不特别关注的消费者剩余的绝对数值。由政策变化引起的消费者剩余的变

[一] 当然，消费者剩余的变动只是效用变动的一种表示方法，还有其它方法，比如消费者剩余的平方根的变动也能很好地表示效用变动。但是，人们通常使用消费者剩余作为效用的标准测度方法。

动，显然更有趣。例如，假设某商品的价格从 p' 变为 p''，消费者剩余如何变动？

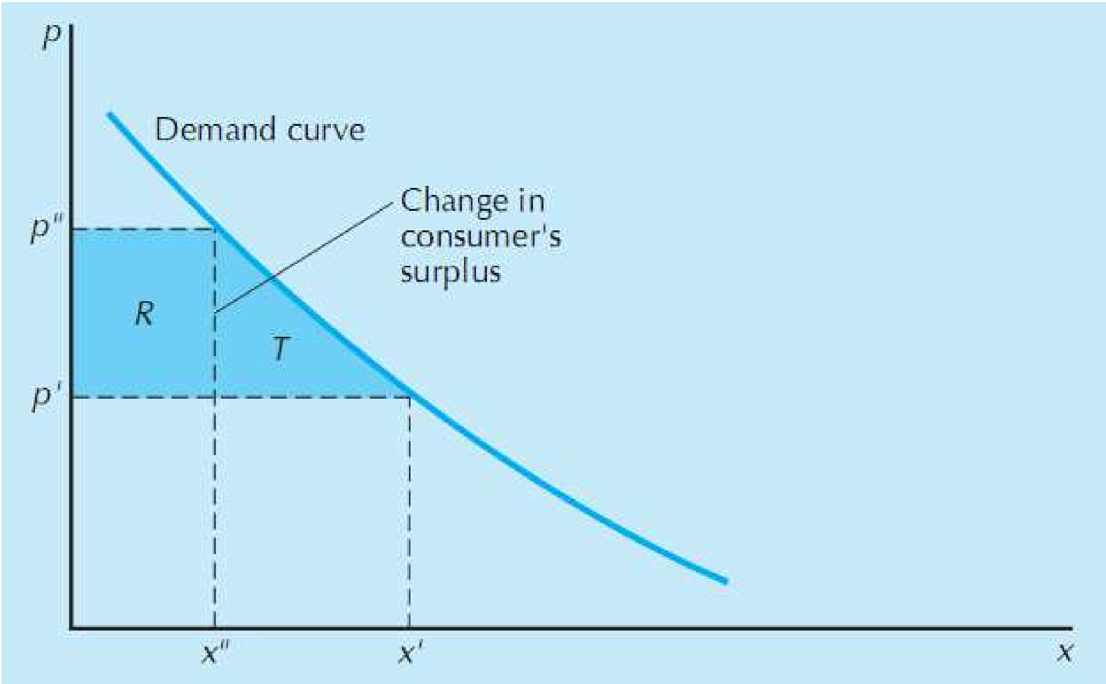


图 14.3：消费者剩余的变动。消费者剩余变动等于两个大致三角形的面积之差，因此它等于同中近似梯形的面积。

在 14.3 中，我们展示了价格上升引起的消费者剩余的减少。价格线 p' 以上需求曲线以下的区域大致呈现三角形的形状，价格线 p'' 以上需求曲线以下的区域也大致为三角形，消费者剩余的减少量等于这两个区域之差，因此，消费者剩余的减少量等于图中近似梯形的面积。这个梯形由两部分组成：矩形区域 R 和大致为三角形的区域 T 。

由图可知，价格为 p' 时的需求量为 x'，价格上升为 p'' 时的需求量等于 x'' 。由于价格上涨后，消费者消费 x'' 单位的商品，每单位商品都比原来多花费了 $(p''-p')$ 元，因此和原来相比，如果消费 x'' 单位，消费者总共多花费了 $(p''-p')x''$ 元，这个金额恰好就是矩形 R 的面积，因此矩形区域 R 衡量消费者剩余的部分损失，因为现在消费 x'' 单位要比原来多花钱。

但矩形区域 R 并不是全部的福利损失。由于商品价格上升后，消费者决定比原来少消费一些（消费量从 x' 减少至 x'' ）。三角形区域 T 衡量该商品**损失的**消费价值。消费者的福利总损失等于这两部分效应之和：矩形区域 R 衡量消费者因对继续消费的数量支付更多的钱而造成的损失；三角形区域 R 衡量消费量减少引起的损失。

例子：消费者剩余的变动

问题：假设线性需求曲线为 $D(p)=20-2p$.当价格由 2 上升为 3 时，消费者剩余变动

了多少？

答案：当 $p=2,\ \ D(2)=16$，当 $p=3, D(2)=14$。因此，我们需要计算梯形的面积，这个梯形的高为 1，两个底分别为 14 和 16。这等价于高为 1 长为 14 的矩形面积（14），加上高为 1 底为 2 的三角形的面积（1）。总面积因此为 15。

14.8 补偿变化和等价变化

在拟线性偏好的情形下，消费者剩余理论非常令人满意。即使效用不是拟线性的，在不少实际运用中，消费者剩余仍然可以很好地衡量消费者的福利。因为，与估计消费者剩余相比，估计需求曲线时的误差更大一些。

但对于某些应用来说，使用近似的方法达不到要求。在本节，我们介绍一种衡量“效用变动”的方法，这种方法没有使用消费者剩余工具。这种方法涉及两个问题。第一个问题是，在可以观测某消费者的选择行为时，我们如何估计效用函数。第二个问题是，我们如何用货币单位计量效用。

事实上，我们已经介绍过如何估计效用函数的问题。在第 6 章，我们介绍了如何估计柯布-道格拉斯效用函数。在那个例子中，因为我们注意到每种商品的支出份额是相对固定，因此我们知道这是柯布-道格拉斯效用函数，并且可以使用平均支出份额作为该函数的参数估计值。但是，如果消费者的需求行为没有呈现这样的特征，就不能再使用柯布-道格拉斯函数，而应该选择更复杂的效用函数，但是它们的原理是一样的：如果我们获得需求行为的足够观测数据，并且这种行为与效用最大化的特征一致，那么我们一般就能够估计出这个函数。

一旦我们估计出了效用函数，我们就可以使用它去分析价格和消费消费水平变化产生的影响。如果我们能做到这一点，我们就满足了。我们已经知道，偏好最为关键；任何效用函数，只要它能描述这种偏好，那么它们都是可行的，没有好坏之分。

然而，我们有时喜欢使用货币单位来衡量效用，因为这很方便。例如，我们应该补偿给某消费者多少钱，才能让他改变现有的消费模式？这样的问题在本质上涉及效用变动的测度问题，但是，显然这里的效用是以货币单位计量的。用货币计量效用有什么好处？

假设我们考虑的情形是图 14.4 所示的情形。此处，消费者最初面对的价格为 $(p_1^*,1)$，这时他消费的消费束为 (x_1^*,x_2^*)。现在商品 1 的价格由 p_1^* 上升为 $\hat{p}_1$，此时消费者的消费束为 $(\hat{x}_1,\hat{x}_2)$。价格上升对消费者造成多大损失？

回答上述问题的思路之一是，假设价格升高**之后**我们给予消费者一些补偿，让他的状况恰好和价格升高**之前**一样好，我们应该补偿他多少钱？从图形的角度，我们要做的事情是计算应该将新预算线向上移动多大距离，才能恰好与通过原消费束 (x_1^*,x_2^*) 的那条无差异曲线相切。使消费者恰好回到原来无差异曲线所必须的收入变动，称为**收入的补偿变化**（compensating variation in income），因为这种收入变动恰好补偿了价格变动给消费者造成

的影响。补偿变化衡量如果政府想恰好补偿因价格变动而给消费者造成的影响，政府应该补偿消费者多少钱。

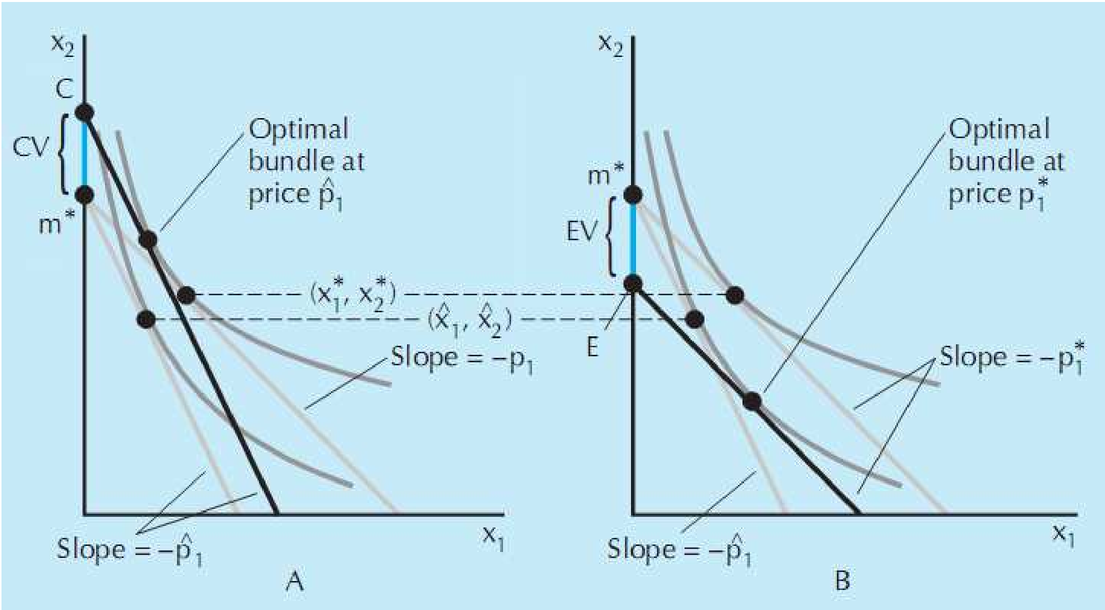


图 14.4：补偿变化和等价变化。A 图为补偿变化（Compensating Variation, CV），B 图为等价变化（Equivalent Variation, EV）。

测量价格变动对消费者效用影响（以货币计量效用）的另外一种方法，是计算在价格变动**之前**，我们应该从消费者手里拿走多少钱，才能让消费者的状况恰好和价格变动**之后**一样好。这种变化称为**收入的等价变化**（equivalent variation in income），因为这种收入变动引起的效用变动等于上述价格变动变动引起的效用变动。在图 14.4 中，这就是计算应该将原来的预算线向下移动多大距离才能恰好与通过新的消费束的那条无差异曲线相切。等价变化衡量消费者为了避免价格变动的影响，他愿意放弃的最大收入额。

一般来说，为了避免价格变动的影响，消费者愿意放弃的钱数，不等于为补偿价格变化而补贴给他的钱数。毕竟，在不同的价格水平下，一元钱对消费者的价值是不同的，这是因为在不同价格水平下，一元钱能购买到的消费数量是不同的。

从几何图形的角度看，补偿变化和等价变化，只是衡量两条无差异曲线之间距离的两种不同方法。在每种情形下我们对两条无差异曲线之间距离的测量方法是相同的，这就是测量它们切线之间的距离。一般来说，距离大小取决于它们切线的斜率，即取决于两种商品的价格比率（预算线的斜率）。

然而，在一种比较重要的情形下，补偿变化和等价变化是完全一样的，这种情形就是拟线性效用。在该情形下，无差异曲线互相平行，因此不管你在何处测量，两条既定的无差异曲线之间的距离是相同的，请见图 14.5。在拟线性效用的情形下，当计算价格变化对消费

者的效用影响时（以货币衡量），补偿变化、等价变化以及消费者剩余的变化这三个结果是一致的。

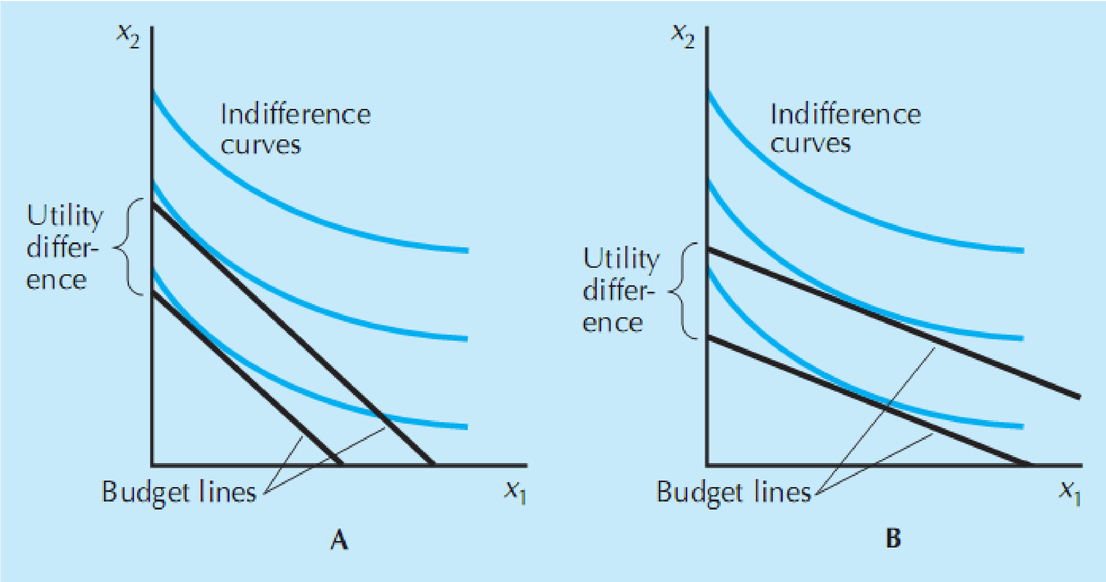


图 14.5：拟线性效用。在拟线性效用的情形下，两条既定无差异曲线之间的距离和预算线的斜率无关。这是由于无差异曲线是平行的。

例子：计算补偿变化和等价变化

假设某消费者的效用函数为 $u(x_1,x_2)=x_1^{1/2}x_2^{12}$，价格最初为（1, 1），他的收入为 100。现在商品 1 的价格上升为 2。请计算补偿变化和等价变化。

我们已知道柯布-道格拉斯效用函数的需求函数为

$$x_1=\frac{m}{2p_1}$$

$$x_2=\frac{m}{2p_2}.$$

使用上述两个式子，可以计算出消费者的需求从 $(x_1^*,x_2^*)=(50,50)$ 变为 $(\hat{x}_1,\hat{x}_2)=(25,50)$。

为了计算补偿变化，我们需要计算出在价格（2, 1）时应该拥有多少钱，才能让他的状况和（50, 50）一样好。如果价格为（2, 1）并且消费者的收入为 m，我们可以将该组价格和收入代入需求函数，代入后可以计算出，消费者的最优选择为消费束 $(m/4, m/2)$。令该消费束的效用值等于消费束（50, 50）的效用值，可得

$$(\frac{m}{4})^{1/2}(\frac{m}{2})^{1/2}=50^{1/2}\cdot 50^{1/2}$$

解出 m 可得

$$m = 100\sqrt{2} \approx 141.$$

因此，在价格变动后，为使消费者的状况和价格变动之前一样好，必须补偿他 141-100=41 元。

为了计算等价变化，我们需要计算在价格（1, 1）时，为了让消费着的状况和他消费（25, 50）的状况一样好，他应该拥有多少钱。令 m 表示这个钱数，根据上面的计算原理，可得

$$(\frac{m}{2})^{1/2}(\frac{m}{2})^{1/2} = 25^{1/2} \cdot 50^{1/2}$$

解出 m 可得

$$m = 50\sqrt{2} \approx 70.$$

因此，如果消费者在原来的价格水平下拥有 70 元，他的状况就会与新价格下他有 100 元钱的状况一样好。收入的等价变化因此大约等于 100-70=30 元。

例子：拟线性偏好情形下的补偿变化和等价变化

假设某消费的效用函数为拟线性的 $v(x_1) + x_2$.我们知道，在这种情形下，商品 1 的需求仅取决于自身的价格，因此我们将它写成 $x_1(p_1)$。假设价格从 p_1^* 变为 $\hat{p}_1$。请计算补偿变化和等价变化。

在价格 p_1^* 时，消费者的选择为 $x_1^* = x_1(p_1^*)$，此时他的效用为 $v(x_1^*) + m - p_1^* x_1^*$。在价格 $\hat{p}_1$ 时，消费者的选择为 $\hat{x}_1 = x_1(\hat{p}_1)$，此时他的效用为 $v(\hat{x}_1) + m - \hat{p}_1 \hat{x}_1$。

令 C 表示补偿变化。这是在价格变化后为使消费者的状况和价格变化之前一样好，他应该拥有的额外钱数。令价格变化前后的效用相等，可得

$$v(\hat{x}_1) + m + C - \hat{p}_1 \hat{x}_1 = v(x_1^*) + m - p_1^* x_1^*.$$

解出 C 可得

$$C = v(x_1^*) - v(\hat{x}_1) + \hat{p}_1 \hat{x}_1 - p_1^* x_1^*.$$

令 E 表示等价变化。这是在价格变化之前为使消费者的效用和价格变化之后的效用一样大，应该从消费者手里拿走的钱数。因此，E 应该满足下列等式

$$v(x_1^*) + m - E - p_1^* x_1^* = v(\hat{x}_1) + m - \hat{p}_1 \hat{x}_1.$$

解出 E 可得

$$E = v(x_1^*) - v(\hat{x}_1) + \hat{p}_1 \hat{x}_1 - p_1^* x_1^*.$$

注意，在拟线性效用的情形下，补偿变化和等价变化是相同的。而且，它们都等于（净）

消费者剩余的变动：

$$\begin{aligned}\Delta CS &= [v(x_1^*) - p_1^* x_1^*] - [v(\hat{x}_1) - \hat{p}_1 \hat{x}_1] \\ &= v(x_1^*) - v(\hat{x}_1) + \hat{p}_1 \hat{x}_1 - p_1^* x_1^*\end{aligned}.$$

14.9 生产者剩余

需求曲线衡量每个价格水平下某商品的需求数量；供给曲线（supply curve）衡量每个价格水平下某商品的供给数量。正如需求曲线下方的面积衡量由商品需求者享有的剩余一样，供给曲线上方的面积衡量商品供给者享有的剩余。

我们已将需求曲线下方的面积称为消费者剩余。类似地，供给曲线上方的面积称为**生产者（的）剩余**（producer's surplus）。消费者剩余和生产者这两个概念多少有些误导，因为这两个概念在本质上和谁消费以及谁生产无关[一]。使用“需求者剩余”和“供给者剩余”这样的称呼可能更好，但是为尊重传统，我们还是使用传统叫法。

假设我们已经有了某商品的供给曲线。这条曲线衡量每个价格水平下的该商品的供给数量。商品的供给者可能是拥有该商品的个人，或者是生产这种产品的企业。我们采用后面这种解释，这样就和传统的“生产者”概念相一致，我们在图 14.6 中画出了生产者的供给曲线。如果生产者在市场价格为 p^* 时的产量为 x^* 单位，他拥有多少剩余?

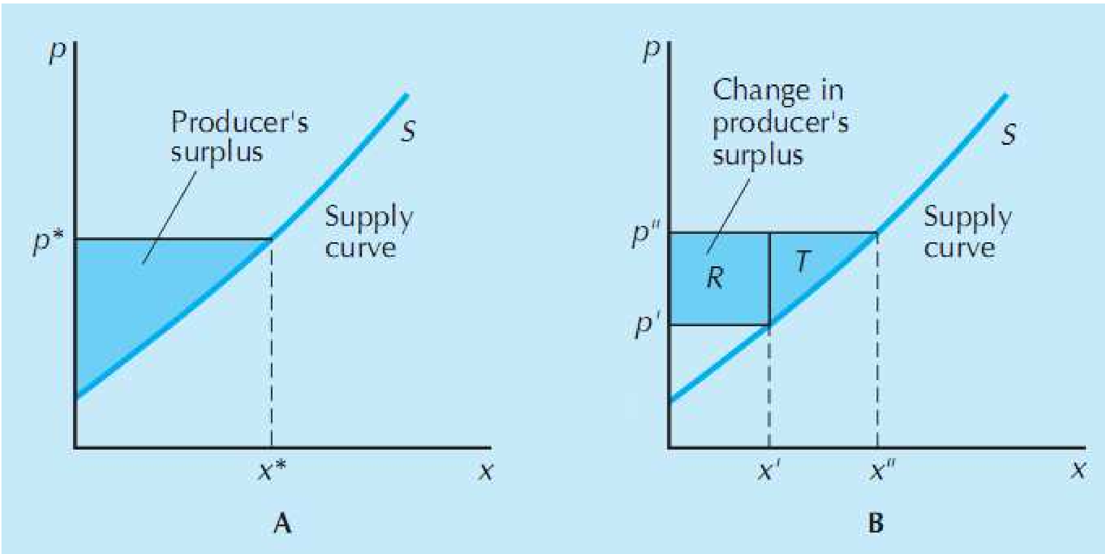


图 14.6：生产者剩余。A 图中的净消费者剩余等于供给曲线左上方和价格线 p^* 下方围成的三角形区域的面积；B 图中的生产者剩余的变动是梯形区域（矩形区域 R 与三角形区域 T 之和）。

使用反供给曲线（inverse supply curve） $p_s(x)$ 分析生产者剩余最为方便。这条曲线衡

[一] 作者的意思是说消费者未必是需求者，比如你买的苹果你家人吃，这里的剩余是指你的剩余。生产者剩余的意义类似理解。译者注。

量为了让生产者生产 x 单位的某种商品，价格应该为多大。

以离散商品的反供给曲线为例说明。在这种情形下，生产者愿意以价格 $p_s(1)$ 销售第一单位商品，但是他实际得到了 p^* 元（市场价格）。类似地，对于第二单位产品，他愿意按 $p_s(2)$ 的价格销售，但是他仍然得到了 p^* 元。以此类推。我们将知道生产者愿意以 $p_s(x^*)$ 的价格出售最后一单位商品。

销售 x^* 单位商品，生产者实际得到的销售收入减去他愿意得到的最低销售收入，这就是销售 x^* 单位商品的**生产者剩余**（producer's surplus）。这可以用图 14.6A 中的三角形区域表示。

正如消费者剩余的情形一样，我们可以计算当价格由 p' 上升为 p'' 时生产者剩余的变动。一般来说，生产者剩余的变动等于两个三角形区域的面积之差，因此通常呈现大致梯形的形状，如图 14.6B 所示。和消费者剩余的情形一样，这个梯形区域由两部分组成：矩形区域 R 和大致呈现三角形的区域 T。矩形区域 R 衡量按新价格 p'' 销售原来的销量得到的收益。三角形区域 T 衡量按新价格 p'' 销售额外数量商品得到的收益。这和前面分析过的消费者剩余变动非常类似。

尽管通常将这种变动称为生产者剩余的变动，在更深层的意义上，它实际代表的是消费者剩余的增加，当然这里的消费者是指拥有企业从而产生供给曲线的消费者。生产者剩余的概念和利润的概念密切相关，但我们要等到在分析企业的行为时再详细分析它们之间的关系。

14.10 收益-成本分析

我们可以使用消费者剩余这个工具计算各种经济政策的收益和成本。例如，我们可以分析最高限价（price ceiling）的影响。如图 14.7 所示。在没有干预的情形下，价格和销量分别为 p_0 和 q_0.

政府认为这个价格太高，因此采取最高限价的措施，假设政府限价为 p_c。在该情形下，供给者愿意供给的数量为 q_c，与原来相比，供给量下降，因此生产者剩余减少为图 14.7 的阴影区域 PS。

现在消费者只能得到 q_c 单位的商品，问题在于谁能得到这些商品？

一种假设是谁的支付意愿最高，谁就能得到商品。令 p_e 表示能诱使消费者需求 q_e 单位商品的价格，这种价格称为**有效价格**（effective price）。如果支付意愿高于 p_e 的每个消费者都能得到商品，则生产者剩余就是图 14.7 中的阴影区域。

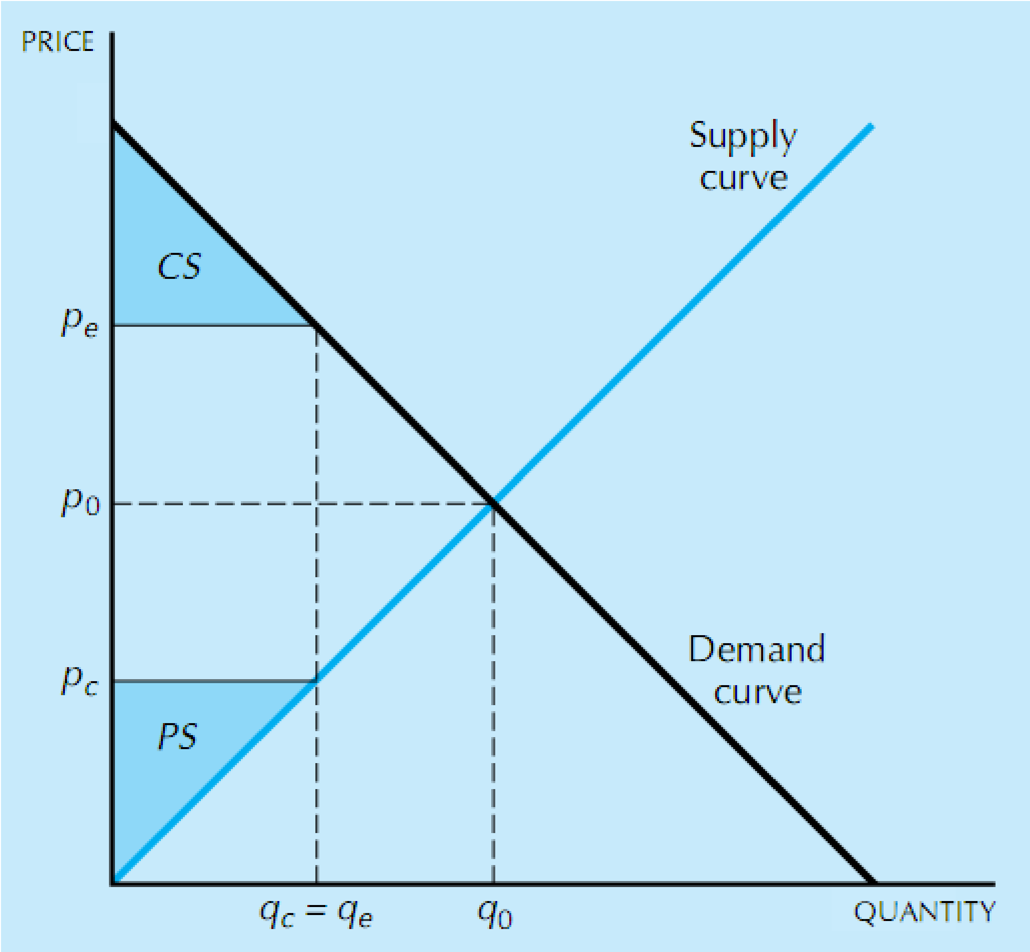


图 14.7：最高限价。将价格限制为 p_c，供给将因此减少为 q_c。因此，消费者剩余减少至 CS 区域，生产者剩余减少至 PS 区域。商品的有效价格 p_e 是使市场出清的价格。该图可用于分析配给制。配给情形下，配给券的价格为 $p_e - p_c$。

注意，消费者剩余的损失和生产者剩余的损失可分别用梯形表示。见图 14.7。这两部分损失怎么计算？先算出完全竞争市场中的消费者剩余与生产者剩余之和，以及最高限价情形下的消费者剩余与生产者剩余之和，然后再把二者相减即可。

这两部分损失就是最高限价政策造成的损失（成本），然而在绝大多数情况下，类似政策造成的实际损失远不止这些，因为在分析时，我们假设支付意愿最高的消费者得到了商品，这是个非常乐观的假设，实际情形很难做到这一点。所以上述损失只是损失的下限，实际损失通常不止这些。

配给制

我们可以使用图 14.7 分析配给制（rationing）造成的社会损失。在配给制的情形下，政府不再强行定价 p_c，而是派发配给券，使得消费者整体只能得到 q_c 单位的商品。为了购买一单位商品，单个消费者需要向卖方支付 p_c 元，并且出示配给券。

如果配给券可以买卖，那么配给券的价格为 $p_e - p_c$。这会使得商品的购买价格变为 p_e，这个价格恰好使市场出清（即市场均衡，也就是需求等于供给）。

11.4 计算收益和损失

如果我们已经估计出了某商品的市场需求曲线和供给曲线，那么一般就可以计算政府政策变化对消费者剩余造成的影响。例如，假设政府决定改变某商品的征税方法。这会改变商品的价格，因此会改变消费者的需求量。我们可以计算不同税收提案下的消费者剩余，比较哪种税收改革带来的损失最小。

这样的计算结果有助于评价各种征税方法的优劣，但是它有两个缺陷。第一，正如我们前面指出的一样，消费者剩余的计算方法只对某些特殊形式的偏好有效，即只对拟线性效用函数代表的偏好有效。我们在前面曾经指出，如果某些商品当收入变动时需求量变动很小，则可以使用拟线性效用去近似。但是，如果某些商品的消费量和收入密切相关，此时使用消费者剩余进行分析就不再合适。

第二，这种计算损失的方法实际上假设所有消费者是相同的，因此，使用这种方法估计出的某社会政策的“成本”，其实是某个虚拟的“具有代表性的消费者”的成本。很多时候我们不仅要知道人群的平均成本，而且要知道谁承担了成本。政策在政治上的成功或失败更多地取决于收益的**分布**和损失的**分布**，而不是平均收益或平均损失。

计算生产者剩余比较容易，但我们已经知道计算价格变动的补偿变化或等价变化也不难。如果我们已经估计出了每个家庭的需求函数，或者至少已估计出样本中具有代表性的家庭的需求函数，那么我们可以使用补偿变化或收入变化计算政策变动对每个家庭的影响。这样，我们就有办法衡量政策变动对每个家庭的“收益”或“成本”的影响。

默文.金，这位伦敦经济学院的经济学家，已经阐述过如何使用补偿变化或等价变化分析英国住房税改革的影响[一]。

默文.金首先分析了英国 5,895 户家庭的住房支出，并估计出了他们购买住房服务的需求函数。然后，他使用这个需求函数测算每个家庭的效用函数。最后，他使用估计出的效用函数计算在英国住房税变动时每个家庭的收益或者损失。他使用的方法和我们介绍的等价变化方法相似。当时英国住房税改革的基本特征是取消个人自有房屋的税收优惠以及提高共有住房的租金。由税收改革产生的税收收入将返还给每个家庭，返还金额与家庭收入成正比。默文.金分析了这种税收改革的影响。

默文.金发现在上述 5,895 户家庭中，有 4888 个家庭从中获得了收益。更重要的是，他明确指出哪些家庭在住房税的改革中损失最大。例如，他发现，94%的最高收入家庭从改革中获利，然而只有 58%的最低收入家庭从改革中获利。这类信息有助于政府在制定税收改

[一] Mervyn King，"Welfare Analysis of Tax Reforms Using Household Data," Journal of Public Economics, 21 (1983), 183-214.

革政策时，考虑税收对财富的分配效应，从而制定更好的税收改革政策。

附录

我们使用微积分严格地推导消费者剩余的计算方法。我们从拟线性效用最大化问题开始分析：

$$\max_{x,y} v(x)+y$$
$$\text{s.t.}\quad px+y=m.$$

从预算线表达式中解出 y，并将其代入目标函数可得：

$$\max_{x,y} v(x)+m-px$$

该问题的一阶条件为

$$v'(x)=p.$$

这意味着反需求函数 $p(x)$ 的定义为

$$p(x)=v'(x).$$

注意，这类似课文中对离散商品的分析：消费者愿意消费 x 单位商品时的价格恰好等于边际效用。

但是，由于反需求函数衡量效用的导数，我们可以将反需求函数积分从而得到效用函数。积分可得：

$$v(x)=v(x)-v(0)=\int_0^x v'(t)dt=\int_0^x p(t)dt.$$

因此，消费 x 单位商品的效用等于需求曲线下方的面积。

例子：若干需求函数

假设需求函数为线性，因此 $x(p)=a-bp$.则当价格从 p 变为 q 时，消费者剩余的变动等于

$$\int_p^q (a-bt)dt=at-b\frac{t^2}{2}\Big|_p^q=a(q-p)-b\frac{q^2-p^2}{2}.$$

再举一个常用需求函数的例子，我们将在下一章详细分析它。这种函数的形式为 $x(p)=Ap^{\varepsilon}$，其中 $\varepsilon<0$，A 为正的常数。当价格从 p 变为 q 时，消费者剩余的变动等于

$$\int_p^q At^{\varepsilon}dt = A\frac{t^{\varepsilon+1}}{\varepsilon+1}\Big|_p^q = A\frac{q^{\varepsilon+1}-p^{\varepsilon+1}}{\varepsilon+1}，其中 \varepsilon \neq -1。$$

当 $\varepsilon = -1$ 时，该需求函数为 $x(p) = A/p$，它和我们的老朋友柯布-道格拉斯需求 $x(p) = am/p$ 关系密切。 柯布-道格拉斯需求的消费者剩余变动为

$$\int_p^q \frac{am}{t}dt = am\ln t\Big|_p^q = am(\ln q - \ln p).$$

例子：补偿变化、等价变化和消费者剩余的变化

在课文中，我们计算了柯布-道格拉斯效用函数的补偿变化和等价变化，然后又计算了柯布-道格拉斯效用函数的消费者剩余的变动。这三种工具都可以描述价格变动对消费者效用的影响（用货币单位计量），我们来比较这三种工具。

假设商品 1 的价格从 1 变为 2, 3…但商品 2 的价格恒定为 1，消费者的收入恒定为 100。表 14.1 给出了柯布-道格拉斯效用函数 $u(x_1,x_2) = x_1^{1/10}x_2^{9/10}$ 的等价变化（EV）、补偿变化（CV）和消费者剩余的变动（CS）。

表 14.1：CV，CS 和 EV 的比较

p_1	CV	CS	EV
1	0.00	0.00	0.00
2	7.18	6.93	6.70
3	11.61	10.99	10.40
4	14.87	13.86	12.94
5	17.46	16.09	14.87

注意，消费者剩余的变化（CS）总是位于补偿变化（CV）和等价变化（EV）之间。这三个数之间的差值非常小。可以证明，上述两个结论在一般情形下也是成立的。请参见 Robert Willig, "Consumer's Surplus without Apology," American Economic Review, 66 (1976), 589-597.

总结

1.在拟线性效用的情形下，消费 n 单位离散商品的效用，正好等于前 n 个保留价格之和。

2.这个保留价格之和是消费 n 单位离散商品的总收益。如果我们从总收益中减去在该商

品上的支出，就得到了（净）消费者剩余。

3.某商品价格变动引起的消费者剩余变动，在图形上大致为梯形形状。消费者剩余的变动可以解释为价格变动引起的消费者的效用变动。

4.一般来说，我们可以使用收入的补偿变化和等价变化来测量价格变动引起的货币影响。

5.如果效用是拟线性的，补偿变化、收入变化和消费者剩余的变化都相等。即使效用不是线性的，我们也可以使用消费者剩余的变化去近似估计价格变动对消费者效用的影响。

6.生产者剩余是指，供给一定数量的商品时，生产者实际得到的销售收入减去他愿意得到的最低销售收入。

复习题

1.某种商品能在完全竞争的行业中生产，生产成本为 10 元每单位。市场里有 100 个消费者，每个消费者愿意花 12 元购买一单位商品，而且只购买一单位（多买对他们没有价值）。求均衡价格和均衡数量。现在如果政府对每单位商品征收 1 元钱的税收。求征税引起的无谓损失（deadweight loss）。

2.假设需求曲线为 $D(p)=10-p$.求消费 6 单位商品的总收益。

3.在上题中，如果价格从 4 上升为 6，消费者剩余变动了多大？

4.假设消费者消费 10 单位某离散商品，现在价格由 5 元每单位上升为 6 元每单位，在价格上升后，消费者仍然消费 10 单位。求价格变动引起的消费者剩余损失。

复习题答案

1.某种商品能在完全竞争的行业中生产，生产成本为 10 元每单位。假设有 100 个消费者，每个消费者愿意花 12 元购买一单位商品，而且只购买一单位（多买对他们没有价值）。求均衡价格和均衡数量。现在如果政府对每单位商品征收 1 元钱的税收。求征税引起的无谓损失（deadweight loss）。

【复习内容】消费者剩余的变动；生产者剩余的变动；完全竞争行业；税收转嫁；社会福利的变化

【复习提示】本题的涉及内容非常多，实际上该题多少已使用了后面章节（尤其是第 16 章）的知识，因此建议你在学完第 16 章后再回头做这个题目，到时就会豁然开朗。

【解题思路】

本题要把握住以下关键知识点。

一是完全竞争的行业。完全竞争行业的均衡价格等于单位产品的成本。其原因是这样的，只要市场价格大于成本，生产者就互相竞争直至价格回落到等于成本。

也就是说，完全竞争行业的供给曲线是水平的，这意味着在某个既定的价格水平上（价格等于单位成本），该行业将供给任意数量的产品，而在低于这个价格的水平上，该行业的供给的数量为 0。

由于供给曲线水平的，而且均衡价格等于单位产品的成本，因此这意味着**生产者剩余为 0**。

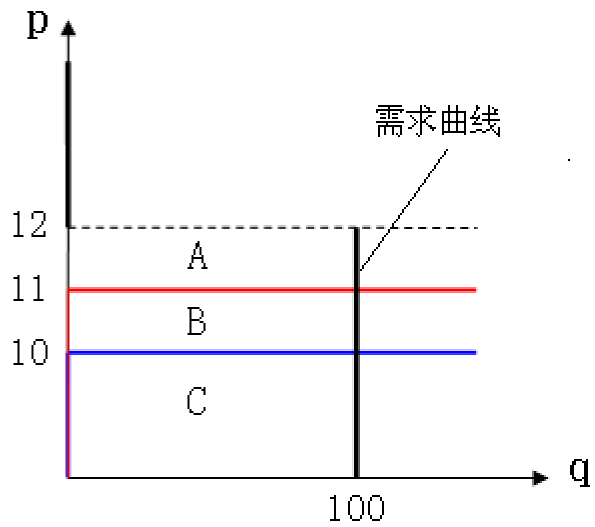
二是社会福利的变化。我们举个例子，如果将你的财富无偿转移给我，社会福利将怎样变化？答案是社会福利没有变化（变化量为 0），因为社会财富的总量没有变化（当然本例不涉及这些财富的生产成本，也就是说这些财富的生产成本变化也为 0）。这个思想对本题的解决非常关键，因为本题涉及了除消费者和生产者之外的第三方——政府。因此在考虑社会福利变化时，除了考虑消费者剩余变动、生产者剩余变动之外，还要考虑税收。

在本题中，由于题目告知该商品的生产行业是完全竞争的，因此，均衡价格等于生产成本，即均衡价格为 10 元。提醒注意，题目中的“消费者愿意支付…”指的就是消费者的支付意愿（保留价格），而不是实际支付的价格（市场价格）。

把握住以上知识点，这个复习题就不难解决。

【参考答案】

根据题意可知，供给曲线是水平的。征税之前，这条水平线为 P=10（如下图蓝线所示）。多说一句，实际上厂商的供给曲线是分段函数，从图形上可以看出，除了上述水平线外，如果价格低于 10，供给量为 0（纵轴上的蓝线部分）。



由题意可知，（反）需求曲线也是分段函数，当商品价格大于 12 时，由于该价格大于消费者的保留价格（12 元），此时需求量为 0，而如果当价格小等于 12 时，市场需求量为 100。（反）需求曲线在上图中，以黑色粗线表示。

由于需求曲线是垂直的，这意味着征税后，消费者将承担全部的税收（参见第 16 章）。因此征税后，均衡价格上升为 11 元。

下面我们开始分析社会福利变化。

首先看生产者剩余的变化。

回顾一下生产者剩余的概念。生产者剩余在图形上表现为，市场价格线以下供给曲线以上的区域的面积。

由于征税前，市场价格线和供给曲线重合，这意味着生产者剩余为 0；征税后，市场价格线（p=11）仍然和供给曲线重合，这意味着此时生产者剩余也为 0。所以，生产者剩余的变化=0-0=0。

接下来看消费者剩余的变化。

回顾一下消费者剩余的概念。消费者剩余在图形上表现为，需求曲线以下市场价格线以上的区域的面积。

征税前，从图形上明显可以看出，消费者剩余=A+B=$100\times(12-10)=200$（元）。

征税后，消费者剩余=A=$100\times(12-11)=100$（元）。

因此，消费者剩余变化=100-200=-100(元)。

从上面两个角度（生产者剩余变化和消费者剩余变化）可以看出，政府征税后，这二者损失了 100 元，因此能否说无谓损失为 100 元？不能！因为你忘了政府的状况变好了，政府税收增加了$100\times(11-10)=100$（元）。

换句话说，这相当于这 100 元从消费者的手里转移到政府的手里，从社会的角度看，这对社会福利没有影响。

因此，社会福利的变化=消费者剩余变化+生产者剩余变化+政府收入变化=0+（-100）+100=0。所以征税引起的无谓损失为 0。

2.假设需求曲线为 $D(p)=10-p$。求消费 6 单位商品的总收益。

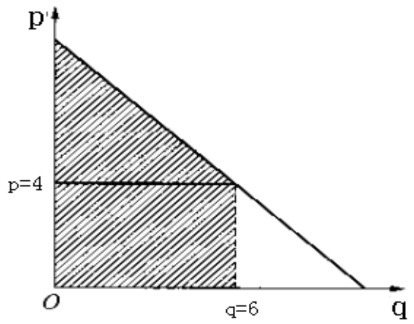
【复习内容】消费者的总收益（总消费者剩余）；总消费者剩余和净消费者剩余的区别

【解题思路与参考答案】

如下图所示，消费者消费某既定数量（此处 q=6）的总收益即总消费者剩余，在图形上

表现为需求曲线、q=6 这条垂线以及两个坐标周围成区域的面积。

而（净）消费者剩余----也就是我们通常所说的消费者剩余，是价格线（此处 p=4）以上和需求曲线以下区域的面积。

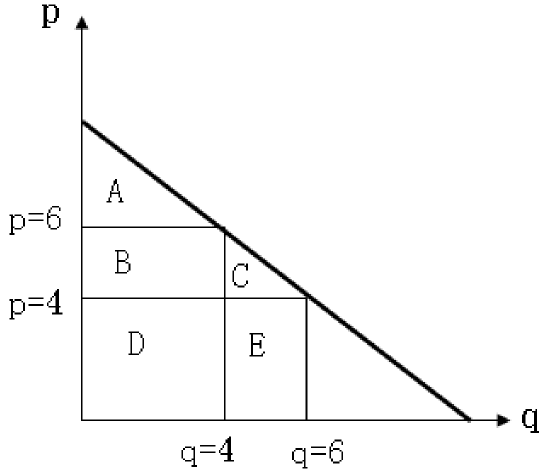


注意到总消费者剩余和净消费者剩余的区别，这个题目就容易解答。由于本题要求计算 q=6 时的总消费者剩余。将 q=6 代入需求曲线 $q=D(p)=10-p$ 可得 p=4，画出以上图形。从图形上可知，消费者消费 6 单位商品时的总收益（总消费者剩余）等于阴影区域（梯形）的面积，因此根据梯形面积的计算公式可知：

消费者消费 6 单位商品时的总收益（总消费者剩余）$=\frac{1}{2}\times(4+10)\times 6=42$。

3.在上题中，如果价格从 4 上升为 6，消费者剩余变动了多大？

【复习内容】（净）消费者剩余；（净）消费者剩余的变动



【参考答案】

解法一：几何图形法

P=4 时，将其代入需求函数，可知 q=6，此时在图形上（上图），消费者剩余=A+B+C。

P=6 时，将其代入需求函数，可知 q=4，此时消费者剩余=A。

因此，价格由 4 上升为 6 时，消费者剩余的变动=A-(A+B+C)=-(B+C)，在图形上，这相当于消费者剩余损失了梯形区域（B+C）面积那么大。

因此，消费者剩余变动=$\frac{1}{2}\times(4+6)\times(4-6)=-10$。

解法二：微积分

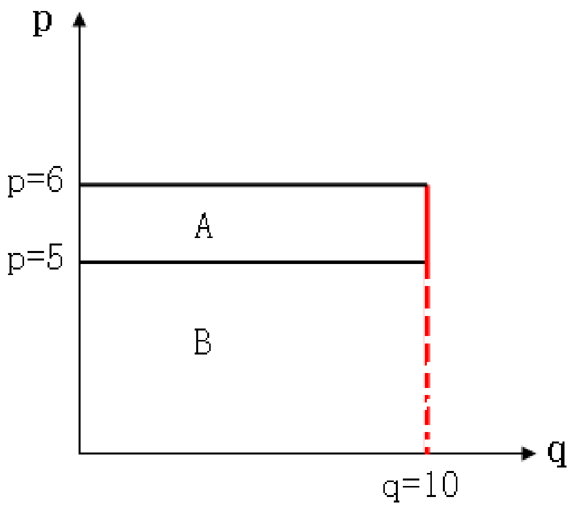
$$\int_p^{p'}(a-bt)dt=at-b\frac{t^2}{2}\Bigg|_p^{p'}=a(p'-p)-b\frac{p'^2-p^2}{2}$$

将相关数值代数上式得到的答案和解法一是一样的，结果都是-10.

4.假设消费者消费 10 单位某离散商品，现在价格由 5 元每单位上升为 6 元每单位，在价格上升后，消费者仍然消费 10 单位。求价格变动引起的消费者剩余损失。

【复习内容】离散商品的需求函数；消费者剩余变动

【参考答案】



根据题意，可以画出以上的图形。容易知道，价格从 5 上升为 6 但需求量不变（10）时，消费者剩余变动=A=$(5-6)\times10=-10$，因此消费者剩余损失了 10 单位。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

15.市场需求（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

15 市场需求

在前面几章，我们已知道如何建立个人消费选择模型。本章的任务是介绍如何将个人选择相加从而得到**市场需求**（market demand）。在推导出市场需求曲线之后，我们将分析它的性质，比如需求和销售收入的关系。

15.1 从个人需求到市场需求

令 $x_i^1(p_1,p_2,m_i)$ 表示消费者 i 对商品 1 的需求函数，$x_i^2(p_1,p_2,m_i)$ 表示消费者 i 对商品 2 的需求函数。假设有 n 个消费者。则商品 1 的**市场需求**（market demand），又称为商品 1 的**（加）总需求**（aggregate demand），是这些消费者个体需求的加和：

$$X^1(p_1,p_2,m_1,\ldots m_n)=\sum_{i=1}^{n}x_i^1(p_1,p_2,m_i).$$

商品 2 的市场需求可类似求得。

由于每个消费者对某商品的需求取决于所有商品（此处为两种商品）的价格以及他的收入，市场需求通常取决于所有商品的价格，也取决于收入的**分布**。然而有时，如果我们把市场需求想像为某个“具有代表性”的消费者的需求，而该消费者的收入等于所有消费者的收入之和，那么在分析时就非常方便。这种可加性的条件要求非常严格，在本书中我们不打算详细地讨论这些问题。

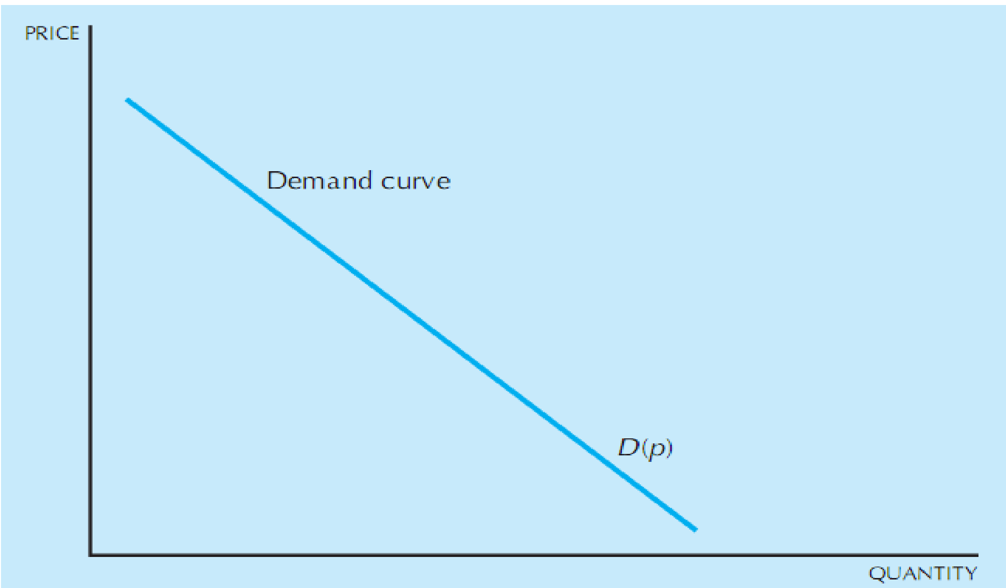


图 15.1：市场需求曲线。市场需求曲线是个人需求曲线的加总。

如果我们接受上述假设，即考察具有代表性消费者的需求，则市场需求函数的形式将为 $X^1(p_1, p_2, M)$，其中 M 为所有消费者的收入之和。在这个假设下，价格为 (p_1, p_2) 时的市场需求等同于该价格下拥有 M 元收入的**某个**消费者的需求。

保持所有货币收入和商品 2 的价格不变，我们可以说明商品 1 的需求和它的价格之间的关系，如图 15.1 所示。再次强调，这条曲线是维持其他商品价格和收入不变的前提下画出的。如果其他商品的价格或者收入变动，市场需求曲线就会移动。

例如，如果商品 1 和商品 2 是替代的，则我们知道商品 2 的价格上升会增加商品 1 的需求，不管此时商品 1 的价格为多大。这意味着商品 2 的价格上升会使商品 1 的市场需求曲线向外移动。类似地，如果商品 1 和 2 是互补的，商品 2 的价格上升会使商品 1 的市场需求曲线向内移动。

如果商品 1 对某消费者来说是正常商品，那么若增加他的货币收入但维持其他因素不变，则会增加他的需求，因此市场需求曲线会向外移动。如果我们接受具有代表性消费者的这个假设，则任何能使总收入增加的经济变动都会增加商品 1 的需求。

15.2 反需求函数

对于市场需求曲线，你可以说它表明需求量是价格的函数，你也可以说它表明价格是需求量的函数。当我们想强调后一种观点时，我们有时称其为**反需求函数**（inverse demand function）$P(X)$。这个函数衡量当需求量为 X 时商品 1 的价格应为多大。

在前面章节，我们知道某商品的价格衡量的是它与所有其他商品的边际替代率；也就是说，某商品的价格表示任何消费者购买额外一单位该商品时的支付意愿，这就是以前介绍过的边际支付意愿的概念。如果所有消费者面对的商品价格相同，则他们在最优选择处的边际替代率必然彼此相等。因此，反需求函数 $P(X)$ 衡量购买该商品的**每个**消费者的边际替代率，或边际支付意愿。

这种加总的操作如果用图形解释，则非常直观。注意，我们是把需求曲线或供给曲线**水平**加总的：给定任一价格水平，我们将所有消费者的需求量相加，当然，需求量要以横轴表示。

例子：“线性”需求曲线的加总

假设某消费者 1 的需求曲线为 $D_1(p) = 20 - p$，消费者 2 的需求曲线为 $D_2(p) = 10 - 2p$。市场需求的表达式是怎样的？此处我们得说说“线性”需求函数的意思。由于商品的数量为负通常没有意义，我们**实际上**是说，上述两个消费者的需求函数具有下列形式

$$D_1(p) = \max\{20 - p,\ 0\}$$

$$D_2(p) = \max\{10-2p,\ 0\}.$$

你已经看出来了，经济学家所谓的“线性”需求函数实际上并不是线性的！这两个需求曲线的加总看起像图 15.2。注意在 p=5 处，市场需求曲线出现弯曲。

图 15.2：两个“线性”需求函数的加总。由于“线性”需求曲线只有在需求量为正时才为线性，因此市场需求曲线通常会出现弯曲。

15.3 离散商品

我们已经知道，如果某种商品的数量是离散的，那么对于单个消费者来说，他对该商品的需求可用保留价格描述。此处，我们分析离散商品的市场需求曲线。为简单起见，我们仅分析单个消费者对该离散商品的需求量为 0 或 1 的情形。

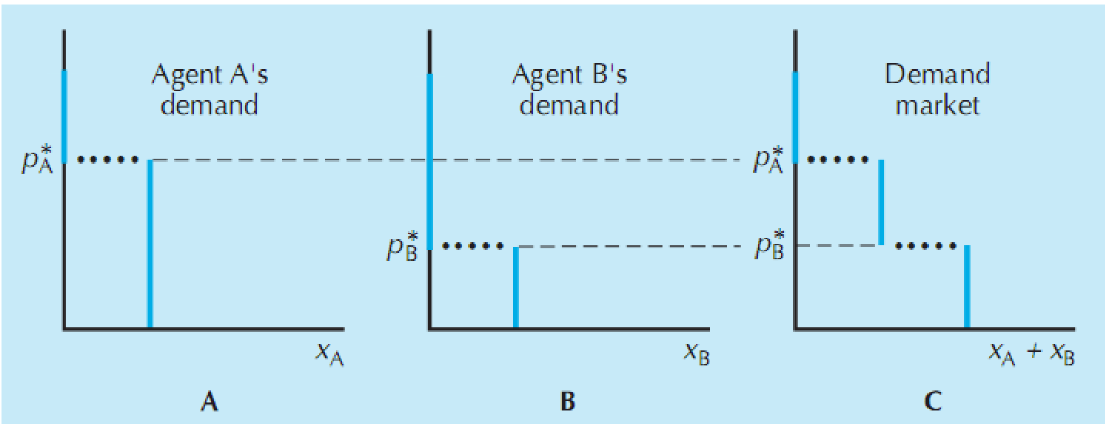


图 15.3：离散商品的市场需求。市场需求曲线是市场中所有消费者的需求曲线的加总，此处我们用两个消费者 A 和 B 代表所有消费者。

在这种情形下，单个消费者的需求完全可由他的保留价格描述：在该价格下，他恰好只愿意购买一单位商品。在图 15.3 中，我们画出了两个消费者 A 和 B 的需求曲线，以及市场需求曲线——上述两个需求曲线的水平加总。注意，在这种情形下，市场需求曲线必然“向下倾斜”，原因是市场价格下降会增加消费者的数量，这些消费者的保留价格都高于下降后的市场价格。

15.4 广延边际和集约边际

在前面几章，我们分析消费者的选择时，基本上分析的都是消费量为正的情形。当价格变动时，消费者决定多消费或少消费某种商品或者另外一种商品，但最终这两种商品的消费量还是正数。经济学家有时把这种变动称为在**集约边际**（intensive margin）上的调整。

在保留价格的模型中，消费者决定是否进入某种产品市场，也就是说消费者决定是否从不消费转为开始消费某商品。这种变动有时称为**广延边际**（extensive margin）上的调整。市场需求曲线的斜率受到这两种决策（集约边际上的调整和广延边际上的调整）的影响。

我们在前面章节已经知道，集约边际上的调整对于正常商品来说是相反方向的：当价格上升时需求量下降。广延边际上的调整也是相反方向的。因此市场需求曲线通常向下倾斜。

15.5 弹性

在第 6 章，我们已介绍过如何根据消费者潜在的偏好推导出需求函数。通常，我们也对下面的问题感兴趣：衡量需求对价格或收入变化的反应敏感性。你的第一感觉也许是使用需求函数的斜率衡量这种反应敏感性。毕竟，需求函数斜率的定义是需求量的变动除以价格的变动：

$$\text{需求函数的斜率}=\frac{\Delta q}{\Delta p},$$

这个式子当然是测量反应敏感性的一种方法。

它能测量反应的敏感性，但它存在着几个缺陷。最重要的缺陷是，需求函数的斜率取决于你使用的价格和需求量的计量单位。如果你用加仑（gallons）代替用夸脱(quarts)计量需求量，则斜率变为原来的 1/4[一]。如果每次都要说出计量单位，这显得很麻烦，自然你会想到使用无量纲（unitfree）的衡量方法。经济学家就是这么做的，这种方法叫**弹性**（elasticity）。

[一] 美制加仑和夸脱换算关系：1 加仑=4 夸脱。译者注。

需求的价格弹性（the price elasticity of demand）ε，定义为需求量的变化百分比除以价格变化百分比[二]。这种情形下，不论你是使用美元还是英镑计量价格，价格上升 10%，$\Delta p/p$ 是一样大的，所以使用百分比衡量增加幅度就使得弹性和计量单位无关。

用公式表示，弹性的定义为

$$\varepsilon=\frac{\frac{\Delta q}{q}}{\frac{\Delta p}{p}}.$$

整理即可得我们经常使用的表达式

$$\varepsilon=\frac{p}{q}\frac{\Delta q}{\Delta p}.$$

上式是说，弹性等于价格对需求量的比率与需求曲线斜率的乘积。在本章附录中，我们用需求函数的导数来描述弹性。如果你对微积分比较熟悉，你会发现使用导数计量弹性更为简便。

由于我们考察的商品一般为普通商品，也就是说需求曲线的斜率一般为负，所以需求弹性一般也为负。但是，每次都要说弹性是负的多少，比较麻烦。所以，有时称弹性是 2 或 3，而不是-2或-3。此时你需要注意的是，说弹性为 2 或 3 时，我们是指弹性的绝对值为 2 或 3，也就是说我们省略了弹性的负号。

弹性为负，在比较弹性大小时有时也容易出现混乱。比如，弹性-3比弹性-2大还是小？从数学的角度看，-3比-2小，但经济学家认为，弹性-3表示需求更具有弹性。这不难理解，因为负号只不过表示需求量变动方向与价格变动方向相反，数值（绝对值）才表示反应敏感程度。但是，为了避免这种混乱，你只要记得用弹性的绝对值比较弹性大小即可。

例子：线性需求曲线的弹性

请看图 15.4 所示的线性需求曲线$q=a-bp$。该需求曲线的斜率为$-b$，将其代入弹性计算公式，可得

$$\varepsilon=\frac{-bp}{q}=\frac{-bp}{a-bp}.$$

由上式可以知道：当 p=0 时，需求弹性等于 0；当 q=0 时，需求弹性等于$-\infty$。价格为多大时，需求弹性等于-1？

将$\varepsilon=-1$代入上式，可得

$$\frac{-bp}{a-bp}=-1.$$

[二] ε，希腊字母，读作“eps-i-lon”.

解出 p，可得

$$p=\frac{a}{2b},$$

在图 15.4 中可以看出，该价格恰好位于需求曲线的中点上。

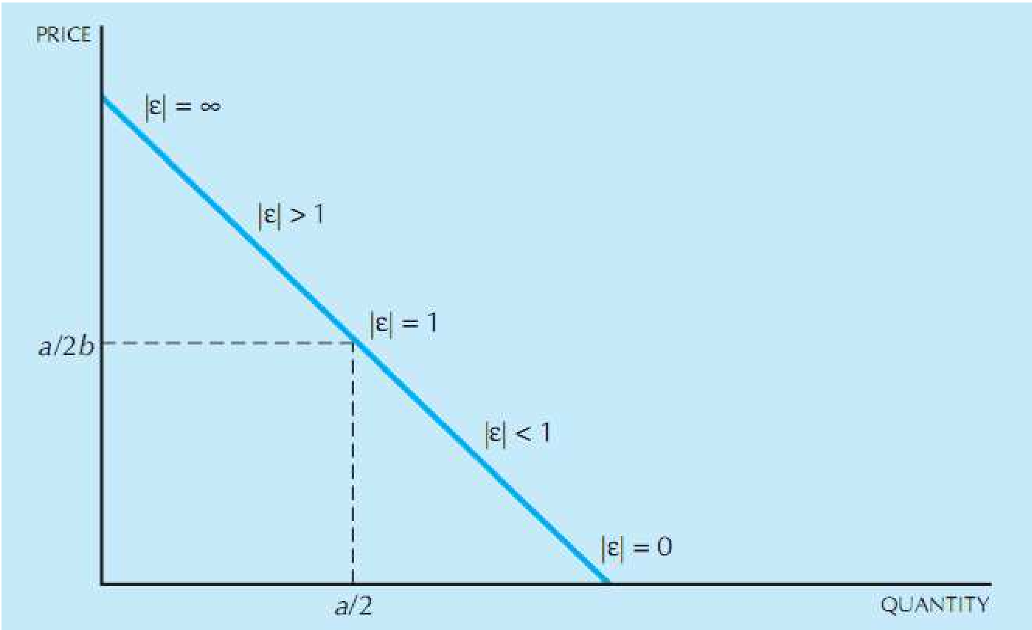


图 15.4：线性需求曲线的弹性。在纵截距处，弹性无穷大；在需求曲线的中点处，弹性等于 1；在横截距处，弹性为零。

15.6 弹性与需求

如果某商品的需求弹性绝对值大于 1，我们说该商品具有**弹性需求**（elastic demand）。如果弹性绝对值小于 1，称该商品具有**缺乏弹性的需求**（inelastic demand）。如果弹性恰好为 -1，称该商品具有**单位弹性需求**（unit elastic demand）。

弹性需求曲线是指需求量对价格变动非常敏感的需求曲线：如果价格上升 1%，需求量将下降超过 1%。所以，将弹性看成需求量对价格变化的反应敏感度，你就会很容易记住具有弹性和却反弹性这类说法的含义。

一般来说，某种商品的需求弹性在很大程度上取决于它有多少相近的替代品。举一个极端的例子，这个例子我们已非常熟悉，就是红蓝铅笔的例子。如果红蓝铅笔是完全替代的，而且消费者每种铅笔都购买一些，这意味着这两种铅笔的价格必须相同。现在假设红铅笔的价格上升，而蓝铅笔的价格不变时，红铅笔的需求将会有什么变化？显然，红铅笔的需求量

将降为零。由于红铅笔具有完全的替代品，所以红铅笔的需求弹性非常大。

若某商品的相近替代品有很多，则它的需求需求曲线对它自身的价格变动将非常敏感。另一方面，如果某种商品，我们找不到它的相近替代品，那么它的需求将非常缺乏弹性。

15.7 弹性与销售收入

销售收入（revenue）是某商品的价格与它的销售量的乘积。如果某商品的价格上升，则该商品的销售量就会下降，但是，销售收入有可能增加也有可能减少。具体为上述哪种情形，取决于需求对价格变动的敏感程度。如果价格上升时，需求大幅下降，则销售收入就会减少；如果价格上升时，需求下降有限，则销售收入就会增加。这说明，销售收入的变动方向与需求弹性有关。

价格弹性和销售收入变动这二者之间的确存在着某种关系。下面我们来说明这种关系。销售收入的定义为

$$R = pq$$

如果价格变动为 $p + \Delta p$，需求量相应变动为 $q + \Delta q$，则新的销售收入为

$$\begin{aligned} R' &= (p + \Delta p)(q + \Delta q) \\ &= pq + q\Delta p + p\Delta q + \Delta p\Delta q \end{aligned}.$$

从 R' 减去 R 可得

$$\Delta R = q\Delta p + p\Delta q + \Delta p\Delta q\,.$$

如果 Δp 和 Δq 很小，则上式右端中的最后一项（$\Delta p\Delta q$）可以忽略不计，这样上式只剩下

$$\Delta R = q\Delta p + p\Delta q\,.$$

这个式子是说，销售收入的变动，大致等于原来需求量乘以价格变动再加上原来的价格乘以需求变量变动。如果我们想要知道价格变动 1%销售收入的变动率，只要将上式除以 Δp 即可，从而得到下式：

$$\frac{\Delta R}{\Delta p} = q + p\frac{\Delta q}{\Delta p}\,.$$

我们可以用图 15.5 解释上述思想。销售收入等于图中方框的面积：价格乘以需求量。当价格上升时，我们在这个方框的上面再加上一个小矩形，这个小矩形的面积约为 $q\Delta p$，但是与此同时，我们还要在这个方框的右边再减去一个小矩形，这个小矩形的面积约为 $p\Delta q$。如果价格和需求量变化很小，上述对大矩形面积的加减正好可以得到上面的表达式。（注意，由图可以看出，我们还剩下一个更小的矩形，它位于方框的右上方，它的面积为 $\Delta p\Delta q$，由于和其他矩形面积相比，它的面积更小，所以我们忽略不计。

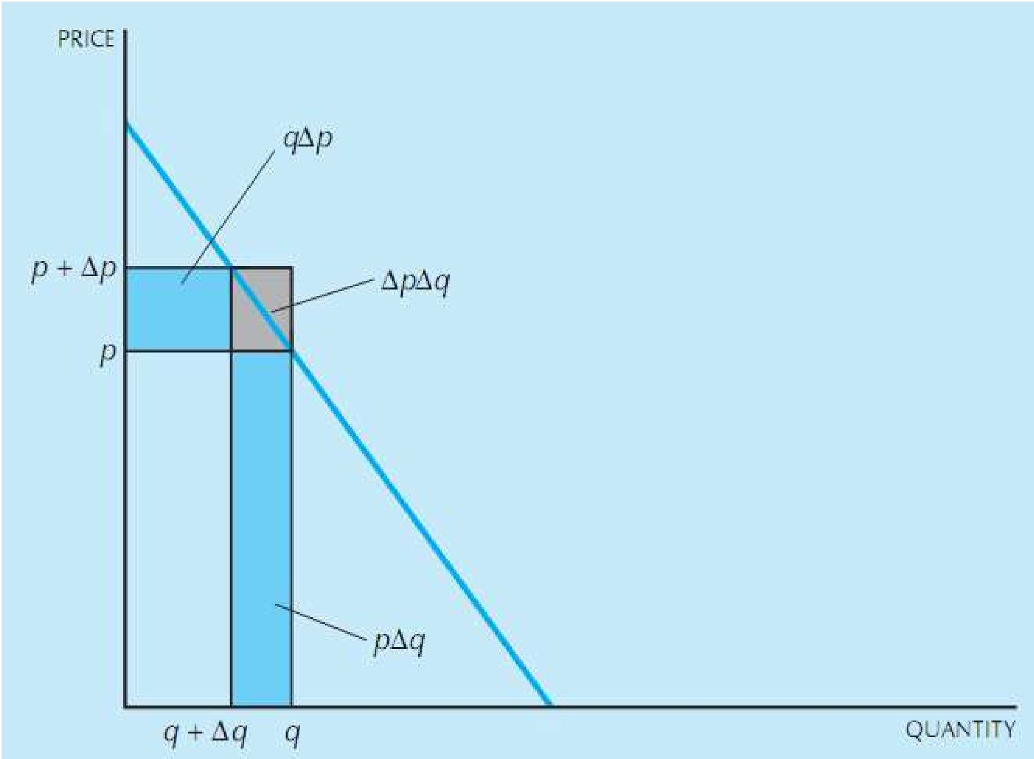


图 15.5：价格变动时销售收入如何变动。价格上升时，销售收入的变动量，等于方框上方的矩形面积（$q\Delta p$）减去方框右边的矩形的面积（$p\Delta q$）。

由上面的分析可知，价格上升时，销售收入既有增加项（$q\Delta p$）又有减少项（$p\Delta q$），那么何种条件下，这两个效应的最终结果为正？换句话说，什么条件下式才成立：

$$\frac{\Delta R}{\Delta p}=p\frac{\Delta q}{\Delta p}+q(p)>0\,.$$

整理可得

$$\frac{p}{q}\frac{\Delta q}{\Delta p}>-1.$$

上式左端是需求价格弹性$\varepsilon(p)$，它为负数。上式两端同乘以-1，不等式的方向因此改变，可得

$$|\varepsilon(p)|<1.$$

因此，如果需求价格弹性的绝对值小于 1，当价格上升时，销售收入也增加；类似地，如果需求价格弹性的绝对值大于 1，则当价格上升时，销售收入下降。

我们再用另外一种方法说明上述观点。首先将销售收入变动写成：

$$\Delta R = p\Delta q + q\Delta p > 0$$

整理可得

$$-\frac{p}{q}\frac{\Delta q}{\Delta p} = |\varepsilon(p)| < 1.$$

第三种方法是根据$\Delta R / \Delta p$的表达式，并将其整理为

$$\begin{aligned}\frac{\Delta R}{\Delta p} &= q + p\frac{\Delta q}{\Delta p} \\ &= q[1+\frac{p}{q}\frac{\Delta q}{\Delta p}] \\ &= q[1+\varepsilon(p)]\end{aligned}$$

由于需求价格弹性一般为负，我们也可以写成下列表达式

$$\frac{\Delta R}{\Delta p} = q[1-|\varepsilon(p)|].$$

借助上式，你很容易看出销售收入如何对价格变动作出响应的：如果需求价格弹性（绝对值）大于 1，则$\Delta R / \Delta p$必定为负。这意味着价格上升（$\Delta p > 0$）时，销售收入下降（$\Delta R < 0$）；如果需求价格弹性（绝对值）小于 1，则$\Delta R / \Delta p$必定为正。这意味着价格上升（$\Delta p > 0$）时，销售收入增加（$\Delta R > 0$）；如果需求价格弹性（绝对值）等于 1，则$\Delta R / \Delta p$必定为零。这意味着价格上升（$\Delta p > 0$）时，销售收入不变（$\Delta R = 0$）。价格下降的情形下类似的结论你可以自行推导。

这些借助数学推导出的结论其实容易记住，因为它的结论非常直观。如果需求对价格变动非常敏感，也就是说需求非常具有弹性，那么价格稍微上升将会使需求大幅下降，因此销售收入会下降。如果需求对价格变动不敏感，也就是说需求缺乏弹性，那么价格上升很大幅度也只能使需求小幅下降，因此销售收入会上升。上述两种情形的分界线是需求价格弹性等于-1，此时价格上升 1%，需求会下降 1%，因此销售收入不会变动。

例子：罢工和利润

1979 年，美国的联合农工工会（the United Farm Workers）发起罢工，表示对加利福尼亚州莴苣种植者的抗议。罢工的结果是莴苣产量下降了大约一半。但是，莴苣供应量的减少不可避免地导致莴苣价格上升。事实上，在罢工期间，莴苣的价格上升了大约 400%。由于产量减半，价格变为原来 4 倍，因此莴苣种植者的利润接近翻番！[一]

你可能会问为何莴苣种植者最终能化解罢工。答案涉及短期供给反应和长期供给反应。美国冬季消费的莴苣主要产自加州帝王谷（Imperial Valley）。在短期，当帝王谷的莴苣产量

[一] See Colin Carter, et. al., "Agricultural Labor Strikes and Farmers' Incomes," Economic Inquiry, 25, 1987,121-133.

急剧减少时，因为其他地区莴苣种植量很少，根本无法弥补市场莴苣供给大幅减少的局面，因此莴苣的价格急剧上升。在长期，比如罢工持续时间很长，其他地区的就会开始种植莴苣。这样以来，其他地区莴苣产量的增加会使价格下降到原来的正常水平，因此减少了帝王谷种植者的利润。

15.8 弹性不变的需求

什么样的需求曲线，它的需求价格弹性不变（恒为某常数）？在线性需求曲线上，需求弹性在 0 和无穷大之间变动，显然线性需求函数不是弹性不变的。

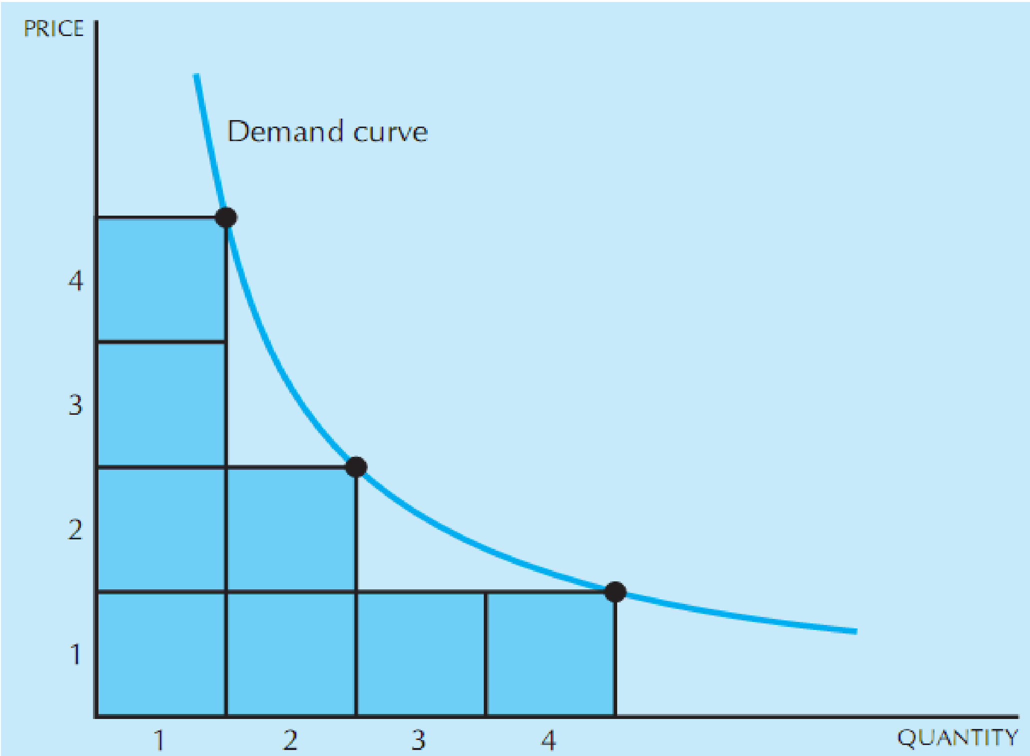


图 15.6：单位弹性的需求。在这条需求曲线上任何一点处，价格和需求量的乘积都为既定常数。因此该需求曲线的需求弹性恒为-1。

我们可以从上面销售收入和需求弹性关系的计算公式得到启示：如果价格为 p 时的需求弹性为 1，则当价格微小变动时，销售收入不变。因此，如果销售收入对于所有的价格变动都保持不变，我们就必定找到了一条弹性不变的需求曲线，这条需求曲线上任何一点的弹性都等于-1。

需求弹性恒为-1的需求弹性的表达式是什么样的？不难找到。由上面的分析可知，我

 曹乾（东南大学 caoqianseu@163.com）

们想让价格和需求量的关系满足

$$pq = \overline{R}\ ,$$

这意味着

$$q = \frac{\overline{R}}{p}$$

就是需求弹性恒为-1的需求曲线的表达式。这条需求曲线请见图 15.6。注意在需求曲线上，价格和需求量的乘积处处相等。

可以证明[㈠]，弹性恒为常数ε的需求曲线的一般表达式为

$$q = Ap^{\varepsilon}$$

上式中：A 为任意正的常数（因为需求量 q 不能为负）；ε为需求弹性，它一般为负。我们在后面会举例说明这样需求函数的作用。

将上式取自然对数，可以得到弹性不变需求函数的更简单的表达式

$$\ln q = \ln A + \varepsilon \ln p\ .$$

在上式中，$\ln q$是$\ln p$的线性函数。

15.9 弹性与边际收入

在 15.7 节，我们已知道当某商品的价格变化时，销售收入是如何变化的。但是我们通常也关注当需求量改变时，销售收入是如何变化的。这个问题显然对厂商的生产决策很有帮助。

在前面我们已看到如果价格和需求量变化很小，销售收入的变动为

$$\Delta R = p\Delta q + q\Delta p\ .$$

将上式两端同时除以Δq，我们就得到了**边际（销售）收入**（marginal revenue）的表达式:

$$MR = \frac{\Delta R}{\Delta q} = p + q\frac{\Delta p}{\Delta q}\ .$$

将上式整理变形最终可以得到一个有用的表达式，为了得到它，我们首先对上式变形

$$\frac{\Delta R}{\Delta q} = p[1 + \frac{q\Delta p}{p\Delta q}]\ .$$

㈠ 假设$\dfrac{dq/q}{dp/p} = \varepsilon(\text{常数})$，则$\dfrac{dq}{q} = \varepsilon\dfrac{dp}{p}$，该式两边取不定积分可得：$\displaystyle\int\frac{dq}{q} = \varepsilon\int\frac{dp}{p}$，由此可得$\ln q = \varepsilon\ln p + \ln A$，即$\ln q = \ln p^{\varepsilon} + \ln A$或$\ln q = \ln Ap^{\varepsilon}$从而可得$q = Ap^{\varepsilon}$。译者注。

括号内的第二项很像需求弹性的表达式，但它不是，它是需求弹性的倒数：

$$\frac{1}{\varepsilon}=\frac{1}{\frac{p\Delta q}{q\Delta p}}=\frac{q\Delta p}{p\Delta q}.$$

于是边际收入的表达式变为

$$\frac{\Delta R}{\Delta q}=p(q)[1+\frac{1}{\varepsilon(q)}].$$

这样我们最终得到了这个有用的表达式，注意在上式中，我们将 p 和 ε 分别写成 $p(q)$ 和 $\varepsilon(q)$ 的形式，只是提醒你价格和弹性通常取决于销量。

由于弹性通常为负，如果你喜欢把弹性看成它的绝对值，那么使用上式就容易出现错误。因此，有时为了避免出现混乱，我们有时也把上面的式子写为

$$\frac{\Delta R}{\Delta q}=p(q)[1-\frac{1}{|\varepsilon(q)|}].$$

这意味着，如果需求弹性等于 -1，那么边际收入就等于零，即销量增加不会改变销售收入。如果需求是缺乏弹性的，则 $|\varepsilon|$ 小于 1，即 $1/|\varepsilon|$ 大于 1，从而 $1-1/|\varepsilon|$ 为负，因此销量增加会使销售收入减少。

上述结论很直观。如果需求对价格变动不敏感，那么为增加一点销量你必须大幅降价，因此销售收入会下降。这个结论和我们前面的结论（价格变动时销售收入如何变动）是一致的，因为从需求函数 $p(q)$ 可知，数量增加则价格下降，反之亦然。

例子：定价决策

假设你要为你的产品定价，再假设你已经正确估计出该产品的需求曲线。如果你的定价目的是实现利润最大化，其中利润等于销售收入减去成本。沿着需求曲线，你会把价格定在什么位置上？你不会把价格定在需求弹性（绝对值）小于 1 处，因为这样的价格下，需求是缺乏弹性的。

为何不能把价格定在需求缺乏弹性处？我们反证一下。假设你这么做了。那么，你提高价格，你的销量因价格提高而下降。但是你的销售收入会增加，这是由于你将价格定在需求缺乏弹性处。既然你的销量下降了，你的总生产成本会下降，或者至少它不会增加。所以你的总利润必然会增加，这表明，把价格定在需求缺乏弹性处，不可能实现利润最大化。具体应该把价格定在何处，我们在以后章节再讲述。

15.10 边际收入曲线

在上一节，我们已知道边际收入的表达式为

$$\frac{\Delta R}{\Delta q}=p(q)+\frac{\Delta p(q)}{\Delta q}q$$

$$\text{或}\ \frac{\Delta R}{\Delta q}=p(q)[1-\frac{1}{|\varepsilon(q)|}].$$

边际收入曲线挺有用。根据第一式可知，当销量 q 为零时，边际收入恰好等于价格。销售第一单位商品得到的额外收入等于价格。但在此之后，由于 $\Delta q/\Delta p$ 为负，边际收入必然小于价格。

想想这是怎么回事。如果你正确预计到能以价格 p 销售 q 单位商品。现在你又想额外多卖一单位商品，你必须降价比如降为 p'。降价有两个效应：一是它将使原先预期销售 q 单位的销售收入下降；二是额外这一单位商品带给你额外的收入 p'。由于第一个效应的存在，因此你多卖一单位商品得到的额外收入必然小于 p'。

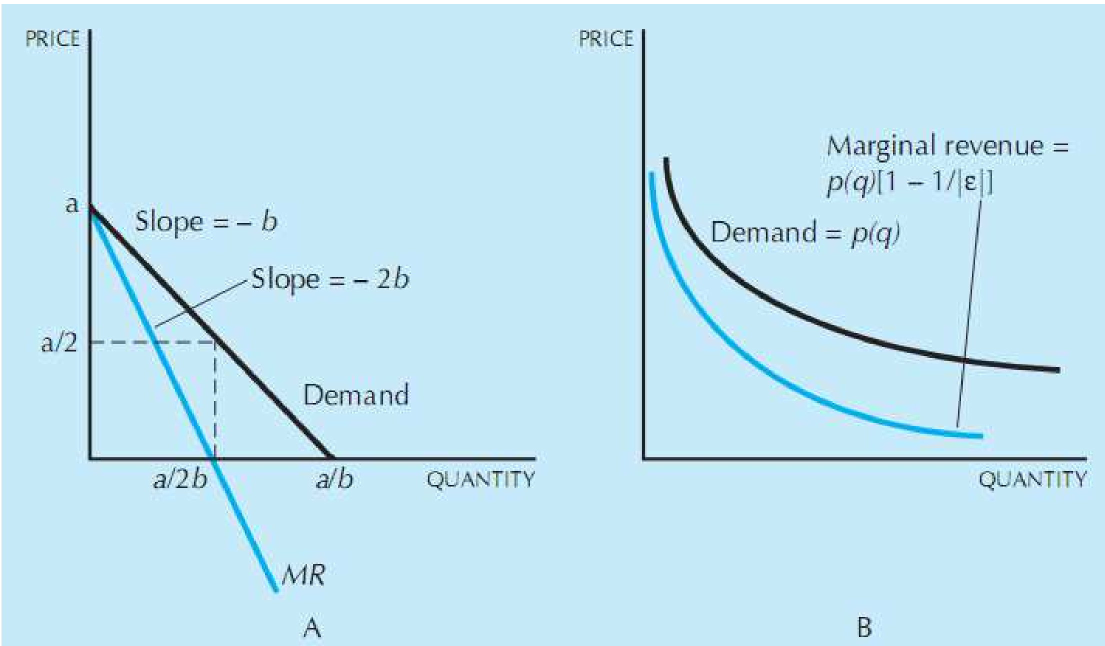


图 15.7：边际收入。（A）线性需求曲线的边际收入；（B）弹性不变的需求曲线的边际收入。

下面我们考虑特殊情形，首先分析线性（反）需求函数：

$$p(q)=a-bq.$$

容易看出它的斜率为常数：

$$\frac{\Delta p}{\Delta q}=-b.$$

因此边际收入的表达式变为

$$\begin{aligned}\frac{\Delta R}{\Delta q} &= p(q) + \frac{\Delta p(q)}{\Delta q}q \\ &= p(q) - bq \\ &= a - bq - bq \\ &= a - 2bq\end{aligned}$$

我们将这个边际收入曲线画在图 15.7A 中。边际收入曲线和需求曲线有相同的纵截距，但前者的斜率是后者的二倍。当 $q > a/2b$ 时，边际收入为负。需求量为 $a/2b$ 处，需求弹性为 -1。需求量大于 $a/2b$ 时，需求就会缺乏弹性，这意味着边际收入为负。

接下来我们再来看另外一种特殊情形，这就是弹性不变的需求曲线，产生的边际收入曲线（见图 15.7B）。如果需求弹性恒为 ε，即 $\varepsilon(q) = \varepsilon$，则边际收入曲线的表达式为

$$MR = p(q)[1 - \frac{1}{|\varepsilon|}].$$

由于括号内的项为常数，因此在图形上，任意给定的 q（q>0），边际收入与价格的比值恒为某既定常数。如果 $|\varepsilon|$=1，边际收入曲线就是横轴；如果 $|\varepsilon|$>1，边际收入曲线在反去求曲线的下方（见图 15.7B）；如果 $|\varepsilon|$<1，边际收入为负。

15.11 需求的收入弹性

我们已经知道需求价格弹性的定义

$$需求的价格弹性=\frac{需求量变动百分比}{价格变动百分比}.$$

这个定义让我们能够计算需求量对价格变动的反应，而且计算结果和使用何种单位计量价格和数量无关。

需求的收入弹性（income elasticity of demand）用来衡量需求对收入变动的反应；它的定义为

$$需求的收入弹性=\frac{需求量变动百分比}{收入变动百分比}.$$

回忆正常商品和劣等商品的概念。**正常商品**（normal goods）是指消费者收入增加时需求量也增加的商品，因此正常商品的需求收入弹性为正。**劣等商品**(inferior goods)是指消费者收入增加时需求量下降的商品。劣等商品的需求收入弹性为负。经济学家有时还使用**奢侈商品**（luxury goods）的概念。奢侈品是指需求收入弹性大于 1 的商品：收入增加 1%，奢侈品需求的增加幅度大于 1%。

然而根据一般的经验，需求的收入弹性一般在 1 附近变动。通过分析预算约束就会明白这一点。写出两种收入水平下的预算线：

$$p_1x_1' + p_2x_2' = m'$$
$$p_1x_1^0 + p_2x_2^0 = m^0.$$

第一式减去第二式并用符号 Δ 表示差值：

$$p_1\Delta x_1 + p_2\Delta x_2 = \Delta m .$$

将等式左边第一项乘以 x_1/x_1，左边第二项乘以 x_2/x_2（这样处理后上式显然仍成立）；然后再在等式左右两边同除以 m:

$$\frac{p_1x_1}{m}\frac{\Delta x_1}{x_1} + \frac{p_2x_2}{m}\frac{\Delta x_2}{x_2} = \frac{\Delta m}{m}.$$

最后，将上式左右两边同时除以 $\Delta m/m$，并且用 $s_i = p_ix_i/m$ 表示商品 i（i=1,2）的**支出份额**（expenditure share），这样我们就得到最后的式子，

$$s_1\frac{\Delta x_1/x_1}{\Delta m/m} + s_2\frac{\Delta x_2/x_2}{\Delta m/m} = 1.$$

这个式子是说，**需求收入弹性的加权平均值等于 1**，其中权重为商品的各自支出份额。如果这两个商品中有一种是奢侈品，由于奢侈品的需求收入弹性大于 1，因此另外一种商品的需求收入弹性必然小于 1，因此它们的需求收入弹性“平均值”大约为 1。

附录

需求的价格弹性可以使用导数定义如下

$$\varepsilon = \frac{p}{q}\frac{dq}{dp}.$$

在课文中我们断言，弹性不变的需求曲线的表达式为 $q = Ap^{\varepsilon}$。现在来验证一下，将此式对价格求导可得：

$$\frac{dq}{dp} = \varepsilon Ap^{\varepsilon-1}$$

将上式左右两端同乘以 p/q 可得：

$$\frac{p}{q}\frac{dq}{dp} = \frac{p}{Ap^{\varepsilon}}\varepsilon Ap^{\varepsilon-1} = \varepsilon.$$

所以我们的那个断言是正确的。

线性需求曲线的表达式为$q(p)=a-bp$。它在 p 点的需求弹性为

$$\varepsilon=\frac{p}{q}\frac{dq}{dp}=\frac{-bp}{a-bp}.$$

当 p 为零时，需求弹性为零。当q 为零时，需求弹性为无穷。

销售收入的表达式为$R(p)=pq(p)$。为了看清价格 p 变动时，销售收入如何变动，我们将销售收入对 p 求导：

$$R'(p)=pq'(p)+q(p).$$

假设价格 p 上升时，销售收入增加。那么必有

$$R'(p)=p\frac{dq}{dp}+q(p)>0.$$

整理可得

$$\varepsilon=\frac{p}{q}\frac{dq}{dp}>-1.$$

由于$dq/dp<0$，两端同乘以-1，可得

$$|\varepsilon|<1.$$

因此，如果价格上升时销售收入增加，我们必定位于需求曲线上缺乏弹性的那一段。

例子：拉弗曲线

在这一小节，我们将做一些简单的弹性计算，这些计算将用于分析一个非常具有政策意义的问题，即当税率改变时税收收入会如何变动。

假设我们想画出税收收入和税率的关系图。画图前简要分析一下。如果税率为零，则税收收入为零；如果税率为 1，没人会需求或供给征税的商品，因此税收收入也为零。因此，随着税率的增加，税收收入必然先增加但最终会下降。（当然，在税率从 0 到 1 逐渐增加时，税收收入可能上下波动多次，但为简单起见，我们不考虑这种可能性。）反应税收收入和税率关系的曲线称为**拉弗曲线**(Laffer curve)，如图 15.8 所示。

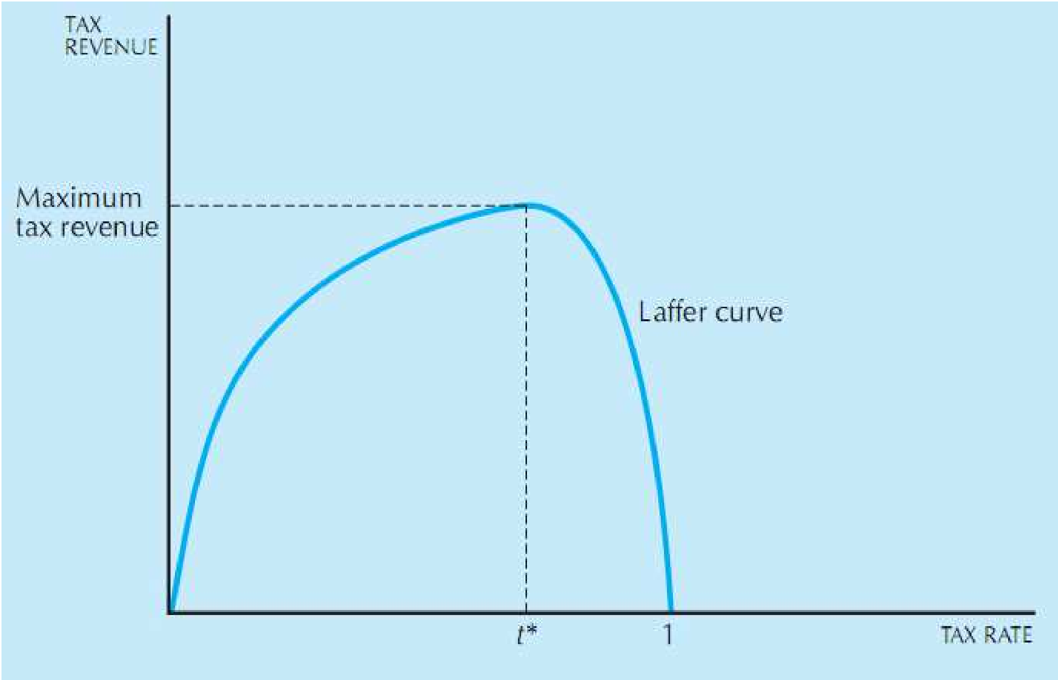


图 15.8：拉弗曲线。拉弗曲线是反应税率和税收收入关系的曲线。上图是拉弗曲线可能具有的一种形状，因为为简单起见，我们只考虑税收收入只有一个峰值的情形。

拉弗曲线表明当税率足够高时，如果税率再提高则将会减少税收收入。税率提高可能使商品供给大幅减少，从而导致税收收入下降。这称为拉弗效应，它是以美国经济学家拉弗的名字命名的，拉弗在 1980 年代初期提出这一图形后，风靡一时。据说拉弗曲线的好处是，如果你花半小时向某议员解释这个图形，他能谈论六个月。的确，在 1980 年代减税影响的政治辩论中，拉弗曲线立下了大功。上述论证有个关键点，即“税率足够高时”。那么，税率多高时拉弗效用才会起作用？

为了回答上述问题，我们分析下面这个简化的劳动市场模型。假设工资高于 $\overline{w}$ 时厂商对劳动的需求量为零，而如果工资等于 $\overline{w}$ 时，厂商对劳动的需求可为任意大。这意味着劳动的需求曲线是一条水平线，它和工资线 $w=\overline{w}$ 重合。假设劳动的供给曲线 $S(p)$ 向上倾斜。劳动市场的均衡情形请见图 15.9.

如果对劳动收入按税率 t 征税，则如果厂商向工人支付工资 $\overline{w}$，工人只得到 $w=(1-t)\overline{w}$。因此，劳动供给曲线向左移动，劳动供给量下降（见图 15.9）。征税降低了工人实际工资收入，打击了劳动供给。

于是，税收收入 T 可按下式计算

$$T=t\overline{w}S(w),$$

其中：$w=(1-t)\overline{w}$；$S(w)$ 为劳动供给。

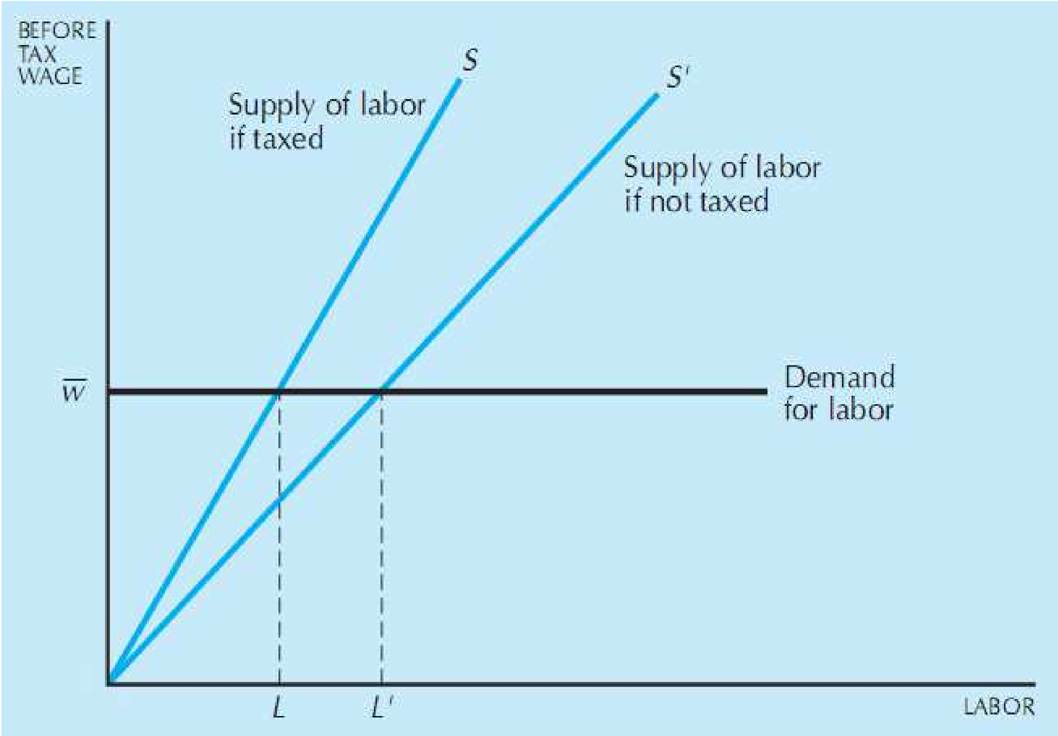


图 15.9：劳动市场。具有水平劳动需求曲线的劳动市场的均衡。当对劳动收入征税时，每个工资水平上的劳动供给都将减少。

为了看清税率变动时税收收入的变动，将上式对 t 求导，可得

$$\frac{dT}{dt}=[-t\frac{dS(w)}{dw}\overline{w}+S(w)]\overline{w} \tag{15.1}$$

（注意要使用求导的链式法则，以及注意 $dw/dt=-\overline{w}$ 。）

拉弗效应发生在当税率 t 增加但税收收入下降时，即（15.1）式为负时。显然这意味着劳动供给必定非常富有弹性，也就是说税率提高时劳动供给必定大幅下降。所以我们要求出能使（15.1）式为负的弹性值。

为了使（15.1）式为负，下列条件必须满足

$$-t\frac{dS(w)}{dw}\overline{w}+S(w)<0.$$

移项可得

$$t\frac{dS(w)}{dw}\overline{w}>S(w),$$

左右两端同除以 $tS(w)$ 可得

$$\frac{dS(w)}{dw}\frac{\overline{w}}{S(w)}>\frac{1}{t}.$$

左右两端同除以 $(1-t)$，并且根据 $w=(1-t)\overline{w}$ 求出 $\overline{w}$ 并代入，可得

$$\frac{dS(w)}{dw}\frac{w}{S(w)} > \frac{1-t}{t}.$$

上式左端是劳动供给弹性的表达式。因此，我们已证明了只有当劳动供给弹性大于 $(1-t)/t$ 时，拉弗效应才会起作用。

举个极端的例子说明。假设劳动收入的税率为 50%，则拉弗效应只有在劳动供给弹性大于 1 时才会发生。这意味着，如果工资下降 1%，劳动供给下降幅度将会大于 1%。这是一种非常敏感的反应。

计量经济学家经常估计劳动供给弹性，但他们报告的最大值也只为 0.2 左右。因此，对于美国的各种税率来说，拉弗效应几乎不可能发生。然而在其他国家比如瑞典，税率已很高，因此有证据表明拉弗现像可能已发生[一]。

例子：弹性的另外一种表示方法

我们介绍弹性的另外一种表示方法，有时使用这种表示方法比较方便。可以证明需求价格弹性可用下式表示

$$\frac{d\ln Q}{d\ln P}.$$

我们注意到

$$\begin{aligned}\frac{d\ln Q}{d\ln P} &= \frac{d\ln Q}{dQ}\frac{dQ}{d\ln P}\\ &= \frac{1}{Q}\frac{dQ}{d\ln p}.\end{aligned} \tag{15.2}$$

我们还注意到

$$\begin{aligned}\frac{dQ}{dP} &= \frac{dQ}{d\ln P}\frac{d\ln P}{dP}\\ &= \frac{dQ}{d\ln P}\frac{1}{P}.\end{aligned}$$

上式意味着

$$\frac{dQ}{d\ln P} = P\frac{dQ}{dP}.$$

将上式代入（15.2），可得

[一] See Charles E. Stuart, "Swedish Tax Rates, Labor Supply, and Tax Revenues," Journal of Political Economy, 89, 5 (October 1981), 1020-38.

$$\frac{d\ln Q}{d\ln P}=\frac{1}{Q}\frac{dQ}{dP}P=\varepsilon,$$

这正是我们想证明的。

因此，需求价格弹性衡量需求曲线的斜率，但要注意这里的价格和需求量都是对数形式。也就是说，此时需求价格弹性衡量 $\ln P$ 变动时 $\ln Q$ 是如何变动的。

总结

1.市场需求曲线只是所有个人需求曲线的加总。

2.某消费者的保留价格是指在该价格下，他恰好对买或不买某种商品无差异。

3.需求函数将需求量看成价格的函数。反需求函数将价格看成需求量的函数。给定一个需求函数，如果你能求出它的反函数，那么在图形上，它们是同一条曲线。

4.需求弹性衡量需求量对价格变动的反应敏感程度。它的正式定义为需求量变动百分比除以价格变动百分比。

5.如果某点处的需求价格弹性（绝对值）小于 1，则称在该点需求是缺乏弹性的。如果某点处的需求价格弹性（绝对值）大于 1，则称在该点需求是富有弹性的。如果某点处的需求价格弹性（绝对值）等于 1，我们说该点的需求是单位弹性。

6.如果需求在某点处是缺乏弹性的，则厂商增加销量会使销售收入减少。如果在某点处需求富有弹性，则厂商增加销量会使销售收入增加。

7.边际收入是某供给者增加销量得到的额外收入。边际收入和弹性之间的关系是 $MR=p[1+1/\varepsilon]=p[1-1/|\varepsilon|]$.

8.如果反需求函数是线性函数 $p(q)=a-bq$，则边际收入为 $MR(q)=a-2bq$。

9.需求的收入弹性衡量需求量对收入变动的反应程度。它的正式定义为需求量变动百分比除以收入变动百分比。

复习题

1.如果市场需求曲线 $D(p)=100-0.5p$，求反市场需求曲线。

2.一个吸毒成瘾的人对毒品的需求函数可能是非常缺乏弹性的，但毒品的市场需求函数

曹乾（东南大学 caoqianseu@163.com）

却有可能是非常富有弹性。请解释原因。

3.如果 $D(p)=12-2p$，收入最大化时价格为多大？

4.假设某商品的需求曲线为 $D(p)=100/p$，收入最大化时的价格是多大？

5.判断对错。在某两商品的模型中，如果一种商品为劣等品则另外一种商品必定为奢侈品。

复习题答案

1.如果市场需求曲线 $D(p)=100-0.5p$，求反市场需求曲线。

【复习内容】市场需求曲线和反市场需求曲线

对于市场需求曲线，你可以说它表明需求量是价格的函数，你也可以说它表明价格是需求量的函数。当我们想强调你的后一种观点时，我们有时将它称为反需求函数 $p(q)$。这个函数衡量当需求量为 q 时商品的价格应为多大。

【参考答案】

$$q=D(p)=100-0.5p \Leftrightarrow p(q)=200-2q 。$$

2.一个吸毒成瘾的人对毒品的需求函数可能是非常缺乏弹性的，但毒品的市场需求函数却有可能是非常富有弹性。请解释原因。

【复习内容】消费在广延边际上的调整与消费在集约边际上的调整

【参考答案】

消费在广延边际上的调整是指消费者决定是否进入市场进行消费的问题。注意这里一般是指新消费者（以前未消费这种商品）决定是否开始消费该商品，即他的消费“处女航”。消费在集约边际上的调整是指原消费者决定消费多少的问题。注意这里一般是指已经消费某商品的消费者对消费量的调整问题。

在这个题目中，毒品的市场需求就是消费在广延边际上和集约边际上的调整问题。例如，假设毒品价格上升，新消费者可能考虑是否进入毒品市场开始吸毒。但对于瘾君子来说，这不是吸不吸毒的问题，因为他已成瘾，他的问题是根据毒品价格变化选择吸多少，价格上升后部分瘾君子会少吸一些。因此对于瘾君子而言，这就是消费在集约边际上的调整。因此毒品的市场需求曲线是非常有弹性的。

3.如果$D(p)=12-2p$，收入最大化时价格为多大？

【复习内容】销售收入

$D(p)=12-2p \Rightarrow TR=pD(p) \Rightarrow TR=p(12-2p)=12p-2p^2$

令$\frac{dTR}{dp}=0 \Rightarrow 12-4p=0 \Rightarrow p=3$。

收入最大化时，价格为 3。

4.假设某商品的需求曲线为$D(p)=100/p$，收入最大化时的价格是多大？

【复习内容】销售收入

$D(p)=100/p \Rightarrow TR=pD(p) \Rightarrow TR=p(100/p)=100$

由此可以看出，销售收入和价格无关，因此任何大于零的价格都可以使收入最大化。

5.判断对错。在某两商品的模型中，如果一种商品为劣等品则另外一种商品必定为奢侈品。

【复习内容】劣等品和奢侈品的收入弹性

【参考答案】

正确。证明如下。

设两商品的需求为

$$x_1^*=x_1^*(p_1,p_2,m)$$

$$x_2^*=x_2^*(p_1,p_2,m)$$

显然以上两个数量满足预算线

$$p_1x_1+p_2x_2=m$$

即

$$p_1x_1^*(p_1,p_2,m)+p_2x_2^*(p_1,p_2,m)=m$$

上式两边对 m 求导可得

$$p_1\frac{\partial x_1^*(p_1,p_2,m)}{\partial m}+p_2\frac{\partial x_2^*(p_1,p_2,m)}{\partial m}=1$$

将上式左边第一项乘以$\frac{x_1}{m}\frac{m}{x_1}$，将左边第二项乘以$\frac{x_2}{m}\frac{m}{x_2}$，由于这两个乘子都等于 1，

所以等式仍成立即：

$$\frac{p_1x_1}{m}[\frac{\partial x_1^*(p_1,p_2,m)}{\partial m}\frac{m}{x_1}]+\frac{p_2x_2}{m}[\frac{\partial x_2^*(p_1,p_2,m)}{\partial m}\frac{m}{x_2}]=1$$

注意到等式左边括号内的两项分别为商品 1 和商品 2 的需求收入弹性，因此分别以 ε_1 和 ε_2 表示；再次，可以注意到 $\frac{p_1x_1}{m}$ 和 $\frac{p_2x_2}{m}$ 分别表示商品 1 和商品 2 的支出占总支出的比例，不妨分别记为 s_1 和 s_2，则上式变为

$$s_1\varepsilon_1+s_2\varepsilon_2=1.$$

这就是说，需求收入弹性的加权平均值等于 1，其中权重为商品的各自支出份额。如果这两个商品中有一种是劣等品，由于劣等品的需求收入弹性小于 1，因此另外一种商品的需求收入弹性必然大于 1，即另外一种商品是奢侈品。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

16.均衡（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

16 均衡

在前面章节，我们已知道如何根据偏好和价格信息构建个人需求曲线。在第 15 章，我们把这些个人需求曲线加总从而得到市场需求曲线。在本章，我们介绍如何使用这些市场需求曲线来决定均衡的市场价格。

我们在第 1 章曾说过，微观经济学基本原理有两条，一条是最优化原理，另外一条是均衡原理。直到目前，我们研究的都是最优化原理的例子：我们假设消费者在预算集内选择最优的消费束，然后研究根据这个假设能推导出什么结论。在后面章节，我们还会使用最优化这种分析方法研究企业的利润最大化行为。最后，我们将消费者行为和企业行为放在一起一起进行分析，因为他们在市场中会互相作用，所以我们要看看这种互相作用能导致什么样的均衡结果。

在详细分析市场均衡结果之前，我们有必要简单说明如何使用均衡分析这个工具，即分析价格如何变动才能使经济人（economic agents）的需求和供给决策相容（compatible）。我们已介绍过需求，在介绍如何使用均衡分析之前，我们需要先介绍供给，因为均衡分析涉及需求和供给这两个方面。

16.1 供给

事实上，我们已经看到过若干供给曲线的例子。例如，我们在第 1 章学过公寓的供给曲线，这条曲线是垂直的。在第 9 章，我们研究过消费者是否决定成为某种商品的净供给者或净需求者，我们还分析过劳动供给决策。

在上面的所有例子中，供给曲线都只不过是一种衡量方法：衡量在每种可能价格下消费者愿意供给商品的数量。事实上，这正是供给曲线的定义：对于任一价格水平 p，某商品的供给量。在下面几章，我们将介绍企业的供给行为。然而，对于我们的研究目的来说，在很多情形下，我们并不需要知道供给曲线或需求曲线是怎么得到，也就是说不必知道是哪种最优化行为产生了这两个曲线。对于很多问题来说，只要知道价格和商品数量的函数关系，即知道在任一价格水平下消费者愿意需求或供给的数量，就已经足够说明均衡分析这种重要的思想。

16.2 市场均衡

假设某商品的消费者有很多，给定他们各自的需求曲线，我们就可以把它们加总从而得到市场需求曲线。类似地，如果该商品的供给者也很多，并且假设他们能独立地作出供给决

策[一]，这样我们就可以把他们各自的供给曲线相加从而得到**市场供给曲线**（market supply curve）。

我们假设每个需求者和供给者都是既定价格的接受者，即假设个体不能影响价格，他们能做的事情就是对给定的市场价格作出最优的反应。如果某市场中的任何个体都不能影响价格即只能接受给定的价格，则这种市场称为**完全竞争市场**（competitive market）。

完全竞争市场的合理性在于，每个消费者的需求量或每个生产者的供应量，在市场总量中的份额都非常小，因此对市场价格的影响可以忽略不计。例如，当小麦的每个供给者在决定生产和供给多少时，他们都认为市场价格多少都和自己的生产决策无关。

尽管完全竞争市场的市场价格和任何一个个体的行为无关，但是所有个体的行为共同决定了市场价格。某商品的**均衡价格**（equilibrium price）是该商品的供给等于需求时的价格。在图形上，均衡价格是使需求曲线和供给曲线相交处对应的价格。

如果令 $D(p)$ 表示市场需求曲线，$S(p)$ 表示市场供给曲线，则均衡价格 p^* 是下列方程的解

$$D(p^*) = S(p^*).$$

上述方程的解 p^* 是市场需求等于市场供给时的价格。

为什么 p^* 是均衡价格？经济均衡（economic equilibrium）是下面这样的一种状态：所有个体都追求自身的行为最优而且每个人的行为和其他人的行为是相容的。在除了均衡价格之外的其他价格水平上，有些人的行为是不可行的，因此他们需要调整。因此，这样的价格不可能持续下去，因为至少有些人会有动机改变自己的行为。

需求曲线代表需求者的最优选择，供给曲线代表着供给者的最优选择，在均衡价格 p^* 处，供需相等意味着需求者和供给者的行为是相容的。除了均衡价格之外，在其他价格水平上，他们的行为不再相容。

例如，假设现行价格 $p' < p^*$，此时需求大于供给。那么，有些供给者将会发现他们可以提高售价，即按高于 p' 的价格卖给失望的需求者。当越来越多的供给者认识到这一点，市场价格将会被抬升到供需相等之处。

类似地，如果现行价格 $p' > p^*$，此时需求小于供给，则有些供给者的实际销量将小于预期销量。他们增加销量的唯一办法是降价。但是，如果所有供给者销售同样的商品，如果某个供给者降价，则其他供给者也会降价。因此超额供给压低了价格。只有使供需相等时的价格才是均衡价格，也只有这样的价格才能使市场均衡。

[一] 某商品供给者的决策通常是互相影响的。因此如果不假设他们彼此之间的独立性，就不能简单地将他们各自的供给曲线加总。译者注。

曹乾（东南大学 caoqianseu@163.com）

16.3 两种特殊情形

我们在此处介绍市场均衡的两种特殊情形，因为它们频繁出现。第一种情形是供给固定不变。此时，供给数量等于既定数量，它和价格无关；也就是说，供给曲线是垂直的。在这种情形下，均衡**数量**完全由供给曲线决定，均衡**价格**完全由需求曲线决定。

与上述情形恰好相反的情形是，供给曲线是完全水平的。如果某行业的供给曲线完全水平，这意味着该行业能够做到按照某不变价格供应任意数量的商品。这种情形下，均衡**价格**完全由供给曲线决定，均衡**数量**完全由需求曲线决定。

这两种情形请见图 16.1。在这两种特殊情形下，均衡价格和均衡数量的决定彼此独立；但在一般情形下，均衡价格和均衡数量是由需求曲线和供给曲线共同决定的。

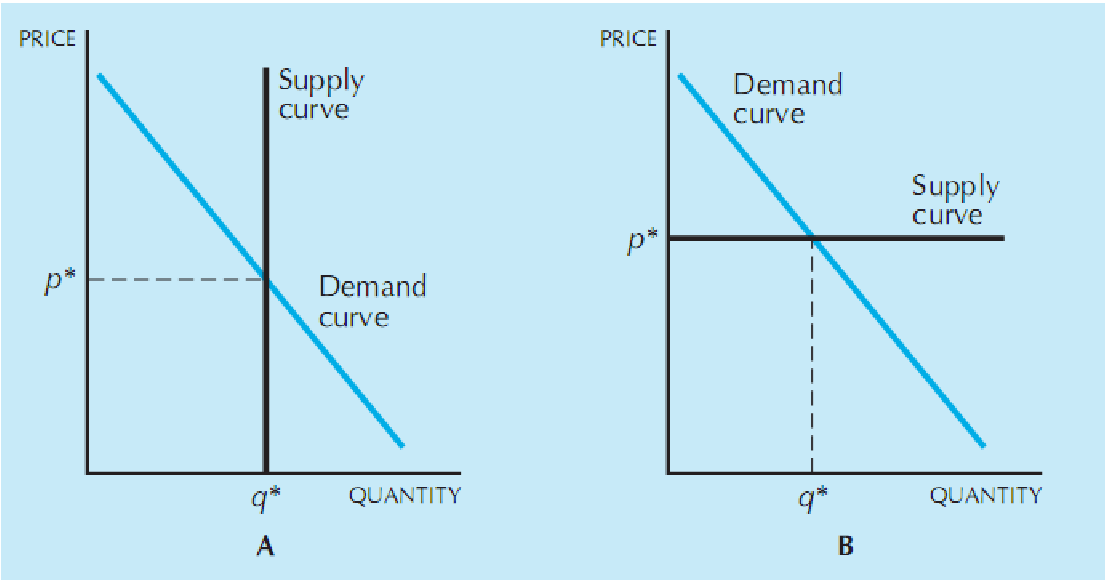


图 16.1：市场均衡的两种特殊情形。在 A 图中，供给曲线为一条垂线，此情形下均衡价格完全由需求曲线决定；在 B 图中，供给曲线为一条水平线，此情形下均衡价格完全由供给曲线决定。

16.4 反需求曲线和反供给曲线

我们可以从另外一个角度看待市场均衡，这种角度更有用。如前所述，个人需求曲线通常是将最优需求量看成价格的函数。但是，我们也可以把它看成反需求函数，反需求函数衡量为了得到既定数量的某种商品，需求者愿意支付的单位**价格**。类似地，我们也可以这样理解供给曲线。你可以认为个人供给曲线是将供给量看成价格的函数。但是，你也可以认为供给曲线衡量的是，为了供给既定数量的商品，价格必须为多少。

同样的思想也可用于**市场**需求曲线和**市场**供给曲线，解释当然也和上面类似。在这种架构内，你若想找到均衡价格，你必须找到某个数量，在这个数量上消费者愿意支付的价格，

恰好等于供给者为供给这个数量而必须得到的价格。

因此，如果令 $p_S(q)$ 表示反供给曲线，$p_D(q)$ 表示反需求曲线，则均衡由下列条件决定：

$$p_S(q^*) = p_D(q^*).$$

例子：线性曲线的均衡

假设需求曲线和供给曲线都是线性的：

$$D(p) = a - bp$$
$$S(p) = c + dp.$$

系数 (a,b,c,d) 是决定这两条曲线的截距和斜率的参数。均衡价格可以通过求解下列方程找到：

$$D(p) = a - bp = c + dp = S(p).$$

均衡价格为

$$p^* = \frac{a-c}{b+d}.$$

均衡的需求量（和均衡供给量）为

$$\begin{aligned} D(p^*) &= a - bp^* \\ &= a - b\frac{a-c}{b+d} = \frac{ad+bc}{b+d}. \end{aligned}$$

我们也可以使用反需求曲线和反供给曲线求解上述问题。首先我们需要找到反需求曲线。价格为多大时需求量为 q？用 q 替代 $D(p)$ 并解出 p，可得

$$q = a - bp,$$

因此，

$$p_D(q) = \frac{a-q}{b}.$$

类似地可以求出反供给曲线：

$$p_S(q) = \frac{q-c}{d}.$$

令需求价格等于供给价格，解出均衡数量：

$$p_D(q) = \frac{a-q}{b} = \frac{q-c}{d} = p_S(q)$$

$$q^* = \frac{ad+bc}{b+d}.$$

16.5 比较静态分析

无论你使用哪种方法——令需求量等于供给量或者令需求价格等于供给价格，都可以找到一个均衡状态。

在找到了这个均衡状态之后，我们可以分析当需求曲线和供给曲线变动时，该均衡状态是怎么变动的。例如，如果需求曲线向右平移——每个价格水平下需求量都增加了相同数额——均衡价格和均衡数量必然都会上升。另一方面，如果供给曲线向右平移，均衡数量增加，但均衡价格必然下降。

如果需求曲线和供给曲线都向右平移，结果会如何？如果是这样，那么均衡数量必然增加，但均衡价格是否上升并不明朗——它可能增加也可能减少，也就是说，如果需求曲线向右移动的距离比供给曲线大，则均衡价格会上升，如果如果需求曲线向右移动的距离比供给曲线小，则均衡价格会下降。

例子：需求曲线和供给曲线都移动

问题：考虑我们在第 1 章介绍过的公寓完全竞争市场。令公寓的均衡价格和均衡数量分别为 p^* 和 q^*。假设某开发商将 m 套公寓由租转为出售，出售给那些已经租住在这 m 套公寓的人。均衡价格（租金）将怎样变化？

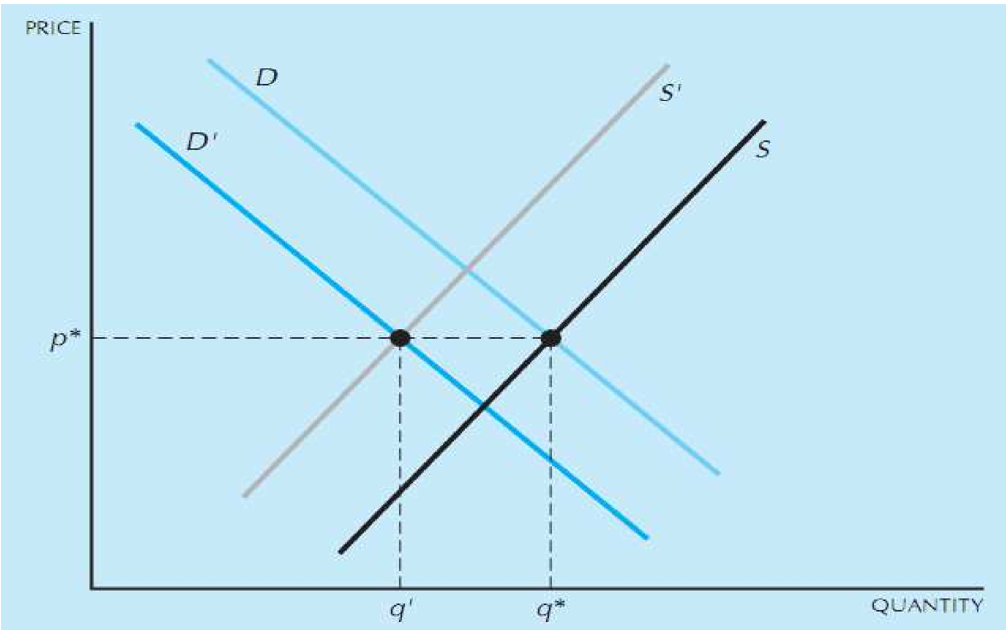


图 16.2：需求曲线和供给曲线都移动。需求曲线和供给曲线都向左平移，而且移动距离相同，这意味着新均衡价格和原均衡价格相同，都为 p^*。

答案：我们用图 16.2 表示上面的情形。需求曲线和供给曲线都向左平移，而且移动的距离**相同**。因此，均衡价格未变，均衡数量减少了 m 单位。

在代数上，均衡价格由下式决定

$$D(p)-m=S(p)-m,$$

新均衡价格显然和原均衡价格相同。

16.6 税收

征收前后市场变化的比较，为练习比较静态分析提供了很好的素材，这一分析结果也可用于指导经济政策。下面我们就来演练演练。

在做税收比较静态分析时，最重要的事情是要把握住下列要点，即对某商品征税时，市场存在**两个**价格：消费者支付的价格和供给者得到的价格。这两个价格不再相等，它们被税收分开了，也就是说单位税收等于需求价格和供给价格之差。

政府可能征收各种税收。此处我们分析两种税，从量税（quantity taxes）和从价税（value taxes or ad valorem taxes）。

从量税是按照商品的数量、容量等计量单位为标准计征的税收，从量税可以对买方征收也可以对卖方征收。例如，汽油税就是从量税的典型例子，大约为 12 美分/加仑。对于每加仑汽油，如果消费者支付的价格为 P_D=1.50 美元，则供给者得到的价格为 P_S=1.50-0.12=1.38（美元）。一般来说，如果某商品在销售时需要缴纳的税收为 t 美元/单位，则

$$P_D=P_S+t.$$

从价税以商品价格作为标准计征的税收，从价税率一般用商品价格的百分比表示。从量税可以对买方征收也可以对卖方征收。美国各州征收的销售税是从价税最常见的例子。如果你所在的州的销售税为 5%，则当你支付 1.05 美元（含税）购买某商品时，供给者得到 1 美元。一般来说，如果从价税的税率为 τ，则

$$P_D=(1+\tau)P_S.$$

下面我们将分析某商品征收从量税后对市场的影响。我们首先假设对供给者征税，例如汽油税。那么，**汽油的供给量取决于供给价格**，即取决于供给者在缴纳税收后实际得到的价格；**汽油的需求量取决于需求价格**，即取决于消费者支付的价格。供给者得到的价格等于消费者支付的价格减去单位税收。于是我们得到下面两个式子：

$$D(P_D)=S(P_S)$$
$$P_S=P_D-t$$

将第二式代入第一式可得均衡条件为：

$$D(P_D)=S(P_D-t).$$

或者先将第二式变形得到 $P_D=P_S+t$，然后将其代入第一式可得

$$D(P_S+t)=S(P_S).$$

以上两种方法都可行；在具体的问题中，哪种方法更方便就用哪种。

假设现在不对供给者征税而是对需求者征税。于是有

$$P_D-t=P_S.$$

这个式子是说需求者支付的价格减去税收等于供给者得到的价格。将其代入需求等于供给这个条件，可得

$$D(P_D)=S(P_D-t).$$

注意，这个式子与对供给者征税的那个式子相同。就需求者和供给者面对的均衡价格而言，谁支付税收并不重要，重要的是税收必须有人支付。

这并不难以理解。以前面所说的汽油税为例。在那里我们说过价格标签标注的价格是含税价。如果价格标签给出的仅是税前价格，消费者在购买时需要另外缴纳汽油税，此时汽油的需求量会改变吗？不会改变。因为前者是消费者将汽油税缴给供给者，然后由供给者代劳缴纳，而后者是消费者直接缴纳，显然这两种缴税方式的最终价格对于消费者来说是一样的。如果消费者知道商品的净价（不含税价），他不会在乎用哪种方式缴税。

我们还有一种更简单的解释方法，这就是使用反需求函数和反供给函数。令 q^* 表示均衡数量，则在该均衡数量水平处，需求价格**减去单位税收**应等于供给价格：

$$P_D(q^*)-t=P_S(q^*).$$

如果对供给者而不是消费者征税，则在 q^* 这个均衡数量处，供给价格**加上单位税收**应等于需求价格：

$$P_D(q^*)=P_S(q^*)+t.$$

但是由于上述两个式子是相同的，因此无论对哪一方征税，均衡价格和均衡数量都是相同的。

最后，我们用几何图形说明上述问题。最简单的方法是使用前面介绍的反需求曲线和反供给曲线。我们希望找到曲线 $P_D(q)-t$ 和曲线 $P_S(q)$ 相交处对应的数量 q，这就是均衡数量。为了找到这个数量 q，只要把需求曲线向下方移动 t 单位的距离，然后看看移动后的需求曲线和原来的供给曲线在什么地方相交即可。或者，我们可以找到使 $P_D(q)$ 和 $P_S(q)+t$ 相等时的 q。寻找该数量的方法是，将供给曲线向上移动，移动的距离等于单位税额 t。这两种方法都能找到均衡数量。由图 16.3 所示。

从图 16.3 我们可以容易地看出税收的定性效应（qualitative effects）：销量必定下降，消费者支付的价格必定上升，供给者得到的价格必定下降。

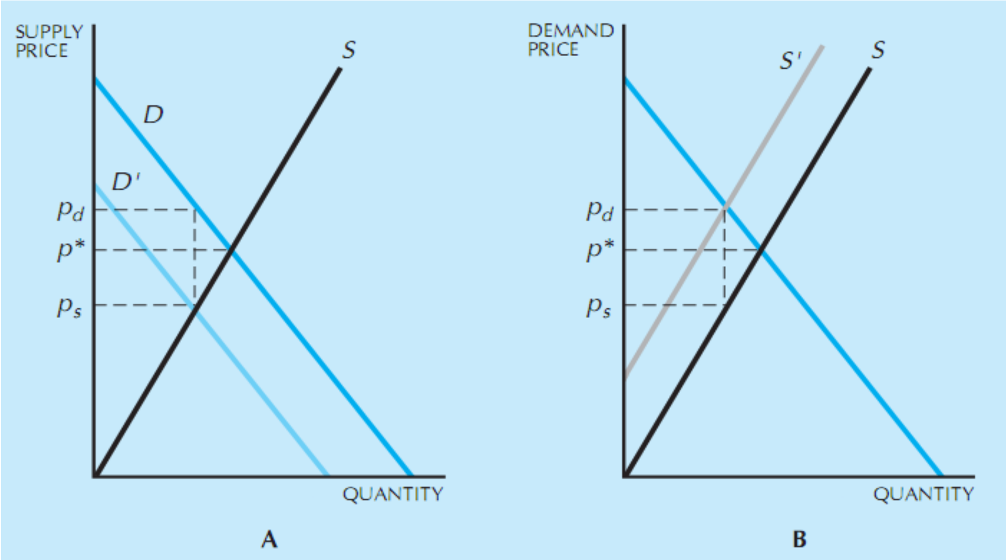


图 16.3：征税。为了分析税收的影响，我们可以将需求曲线向下移动（图 A），或者将供给曲线向上移动（图 B）。这两种方法中，消费者支付的价格是相等的，供给者得到的价格也是相等的。

我们还有一种方法来确定税收的影响，请看图 16.4。想一想市场均衡的定义。我们希望找到数量 q^*，使得当供给价格为 P_s 且需求价格为 $P_d = P_s + t$ 时，q^* 既是需求者的需求量又是供给者的供给量。可以用一条垂直线段代表税收 t，把这条线段沿着供给曲线滑动，直到它恰好接触需求曲线（图 16.4）。接触点对应的数量就是均衡数量！因为此时恰好有 $P_s(q^*) + t = P_d(q^*)$。

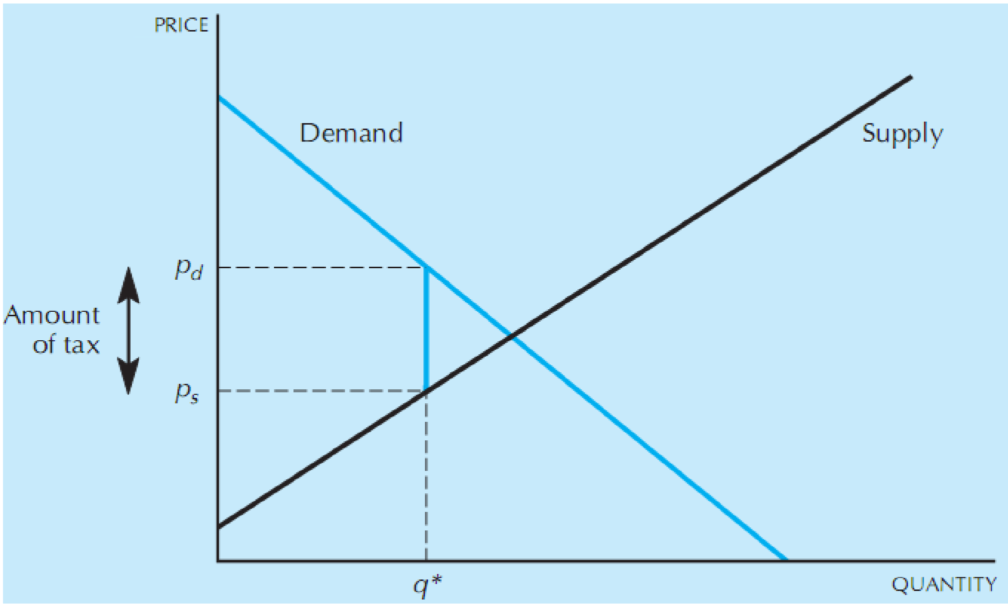


图 16.4：确定税收影响的另外一种方法。用一条垂直线段代表税收 t，把这条线段沿着供给曲线滑动，直到它恰好接触需求曲线，接触点对应的数量就是均衡数量。

例子：线性需求曲线和线性供给曲线情形下的征税

假设需求曲线和供给曲线都为线性的。如果对某商品征税，市场均衡由下列两个式子决定：

$$a - bp_D = c + dp_S .$$

$$p_D = p_S + t .$$

将第二式代入第一式可得

$$a - b(p_S + t) = c + dp_S$$

解出均衡供给价格 p_S^*：

$$p_S^* = \frac{a - c - bt}{b + d} .$$

由于均衡价格 p_D^* 等于 $p_S^* + t$，所以

$$\begin{aligned} p_D^* &= \frac{a - c - bt}{b + d} + t \\ &= \frac{a - c + dt}{b + d} \end{aligned}$$

注意，征税后，需求者支付的价格上升了，供给者得到的价格下降了。价格的变动量取决于需求曲线的斜率和供给曲线的斜率。

16.7 税收转嫁

你可能经常听说对生产者征税不会减少它的利润的说法，因为企业可以将税收转嫁给消费者。如前所述，我们不应该把税收看成是对企业征收还是对消费者征收，而应该看作对企业和消费者**之间的**交易征税。一般来说，征税会提高消费者支付的价格以及降低供给者得到的价格。税收转嫁程度因此取决于需求和供给的特征。

上述结论在极端的情形下更容易看得明白，这样的极端情形是供给曲线完全水平或者完全垂直时，前者又称为**完全弹性的**（perfectly elastic）供给，后者又称为**完全无弹性的**（perfectly inelastic）供给。

我们在前面已介绍过这两个极端的情形。如果某行业的供给曲线是水平的，这意味着该行业在某既定的价格水平处能提供它想提供的任何数量，但是如果市场价格低于上述价格水平，它的供给量为零。在这种情形下，价格完全由供给曲线决定，销量完全由需求曲线决定。如果某行业的供给曲线是垂直的，这意味着它供给的数量是固定不变的。此种情形下，均衡价格完全由需求曲线决定。

如果某商品市场的供给具有完全弹性，征税对这种市场有什么影响？如前所述，征税相当于供给曲线向上移动，移动的距离等于税额。如图 16.5A 所示。

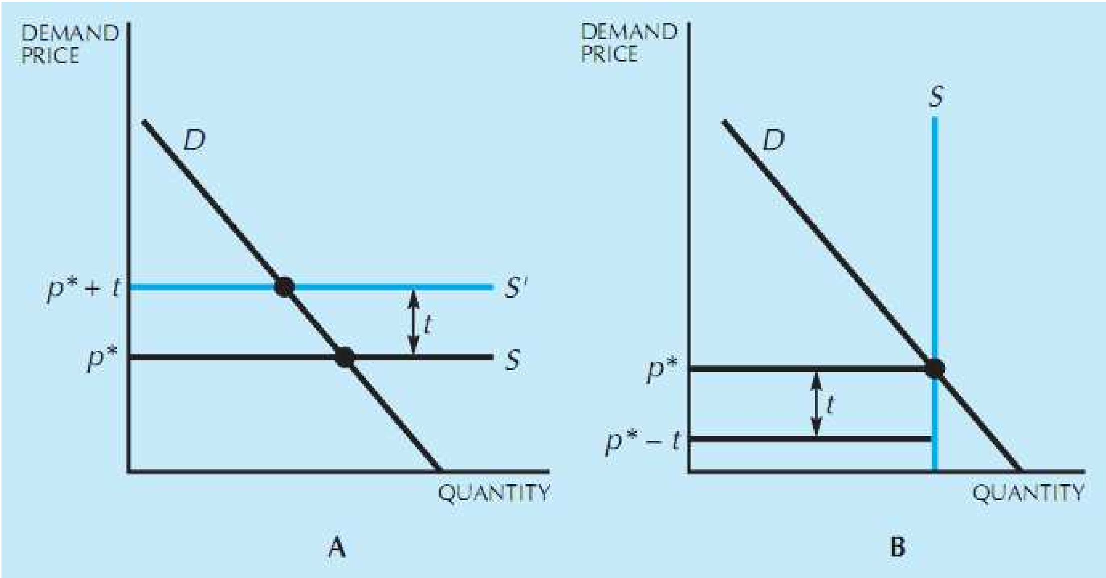


图 16.5：征税的特殊情形。A 图的供给曲线具有完全弹性，税收因此全部转嫁给消费者；B 图的供给曲线完全没有弹性，税收转嫁额为零。

在供给曲线具有完全弹性的情形下，容易看出消费者支付的价格上升了，上升额度恰好等于单位税收。供给价格和征税前的供给价格相等，消费者最终承担了全部税收。想一想水平供给曲线是什么意思，上述结论就不难理解。水平供给曲线意味着行业愿意在既定价格 p^* 处供给任何数量的商品，但是如果市场价格低于 p^*，行业供给数量为零。因此，均衡时供给量若为正，则供给者得到的价格必须等于 p^*。这实际上决定了均衡供给价格和均衡需求价格均为 p^*+t。

相反的情形即供给曲线为垂线时的情形见图 16.5B。在这种情形下，如果我们“将供给曲线向上移动”，则我们什么也没改变。供给曲线沿着自身向上移动，商品的供给数量没变，不管是否征税。在这种情形下，需求曲线决定了商品的均衡价格，消费者愿意对他们的需求量支付 p^* 的价格，不管是否征税。因此，他们最终支付的价格为 p^*，供给者最终得到的价格为 p^*-t。全部税收都由供给者承担。

这种情形通常让人感觉比较荒谬，因为按照常理，征税后，供给者生产成本会增加，因此他们会提高价格。但仔细想想就知道它并不荒谬。如果征税后供给者能够提价而且还能将他们固定的产量全部售出（见图 16.5B），那么在征税前他们早就会提价了，因为这样能更赚更多的钱！如果需求曲线不移动，提价的唯一方法是降低供给量。如果税收政策没有改变供给也没改变需求，它当然不会影响价格。

曹乾（东南大学 caoqianseu@163.com）

既然我们已理解了特殊情形，我们就要开始分析中间的情形，即供给曲线向上倾斜但不是完全垂直。在这种情形下，税收转嫁数额取决于供给曲线的相对倾斜程度（相对于需求曲线来说）。类比以上特殊情形可知，如果供给曲线接近水平，则几乎所有的税额都转嫁给了消费者；如果供给曲线接近垂直，则能够转嫁给消费者的数额会非常小。图 16.6 给出了几个例子。

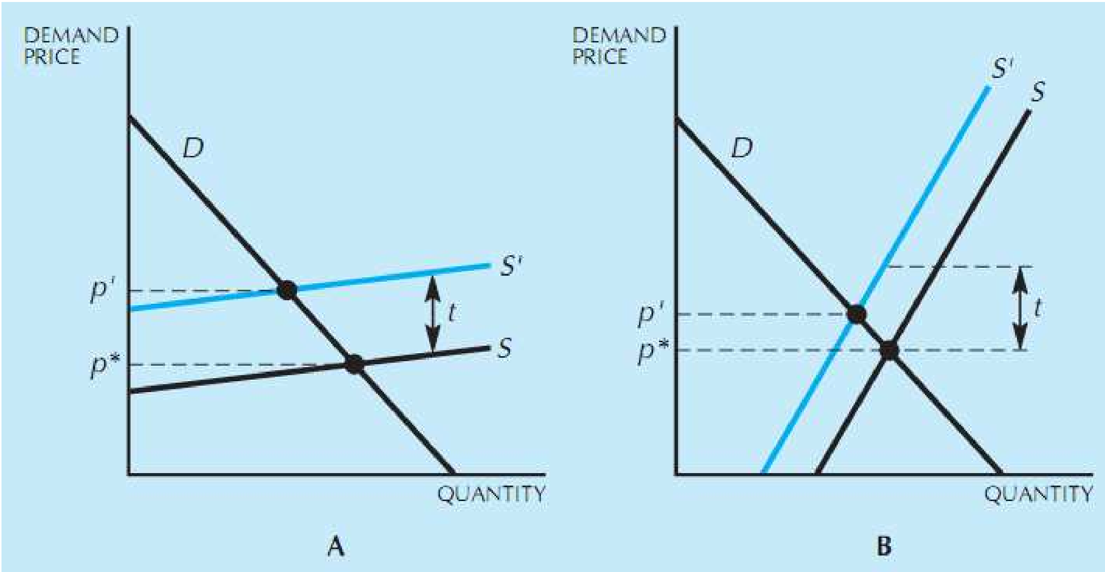


图 16.6：税收转嫁。A 图的供给曲线接近于水平，因此大部分税收可转嫁出去；B 图的供给曲线接近于垂直，因此能转嫁出去的税收非常小。

16.8 征税的净损失

我们已经知道，对某商品征税通常会增加需求者的支付价格，并且降低供给者得到的价格。这当然代表需求者和供给者的成本，但从经济学家的角度来看，征税的真实成本是它减少了产量。

损失的产量是征税的社会成本。我们用消费者剩余和生产者剩余这两个工具分析税收的社会成本，我们已在第 14 章介绍了这两个工具。请看图 16.7。这个图形画出了征税后（每单位商品征收 t 元）的均衡需求价格和均衡供给价格。

征税使产量减少，我们可用消费者剩余和生产者来估计社会损失。消费者剩余的损失为 A+B，生产者剩余的损失为 C+D。这些损失和第 14 章的分析类似。

由于我们想找到衡量征税的社会成本的表达式，似乎将区域（A+B）和区域（C+D）加在一起是可行的，这样我们就得到了该种商品的消费者和生产者的总损失。然而，我们还漏计了第三方，也就是政府。

政府获得了税收收入（revenue）。当然，由于政府使用税收收入服务消费者大众，因此消费者同样从征税中得到好处。但是我们还不能准确计算出消费者从征税中获得的好处到底

为多大，除非我们知道税收收入使用在哪些项目上。

我们假设税收收入全部返还给了消费者和生产者，或者假设政府使用征税收入提供服务，而且服务的总价值正好等于第一种假设中的资金总额。

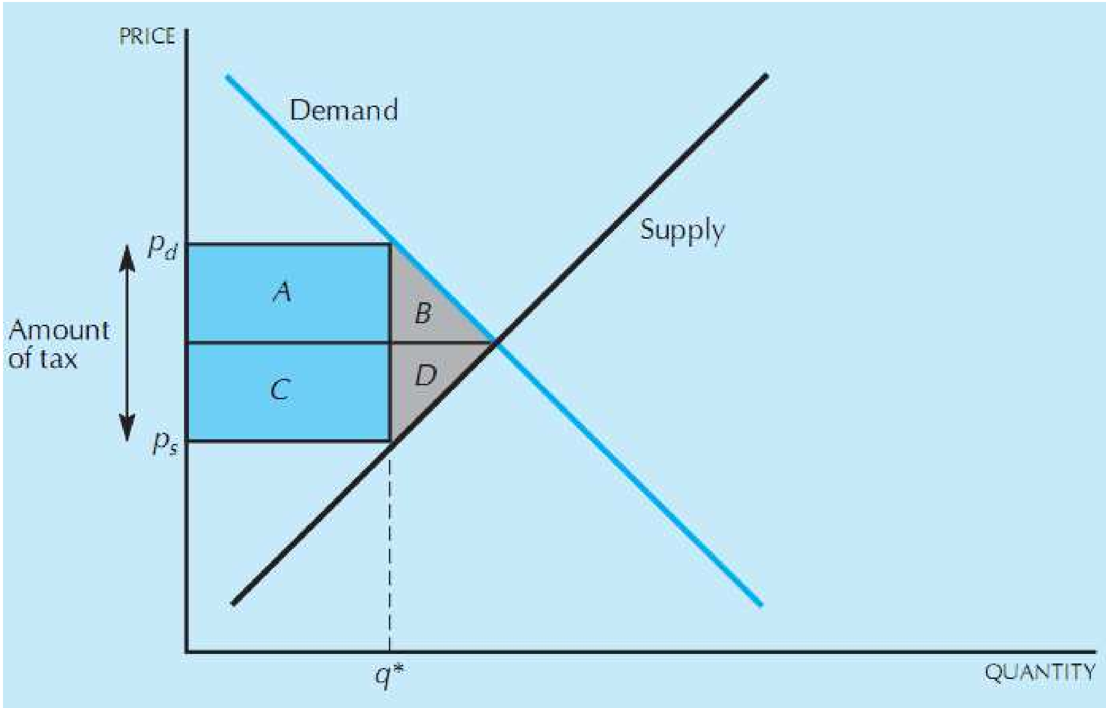


图 16.7：征税的净损失。征税引起的净损失等于区域（B+D）。

那么政府的净收益是区域（A+C），即税收总收入。由于生产者剩余和消费者剩余的损失之和为净成本，税收收入是政府的净收益，征税引起的净成本是这些区域的代数和：消费者剩余变动等于$-(A+B)$即消费者剩余损失了$(A+B)$；生产者剩余变动为$-(C+D)$，即生产者剩余损失了$(C+D)$；政府税收收入等于$+(A+C)$。

将这三项相加得到的净结果为$-(B+D)$。这个区域称为征税的**净损失**（deadweight loss）或者征税的**额外负担**（excess burden）。后面这个称呼更形象一些。

回忆消费者剩余损失的含义。它是指消费者为了少缴税而宁愿付出的代价，由上图可知，消费者宁愿付出的代价为$(A+B)$。类似地，生产者为了少缴税而宁愿付出的代价为$(C+D)$。这两个群体为了少缴税而愿意付出的代价总和为$(A+B+C+D)$，而政府税收收入只有$(A+C)$那么大。因此，征税造成的**额外负担**为$(B+D)$[一]。

[一] 税收是纳税人（消费者和生产者）的**负担**，此处纳税人承担的税负为（A+C），但是政府征税改变了他们的生产和交易行为——生产和交易得更少，他们**另外**损失了（B+D），所以这部分另外的损失称为**额外负担**。从这个意义上来说，“**deadweight loss**”是额外负担的同义语，可以翻译为:**净损失**；**纯粹损失**；**额外损失**；国内其他经济学教科书有时也使用“无谓损失”这样的称呼，也讲得通，因为这样的损失的确是“本来不应该发生的但却发生了”。译者注。

额外负担到底是怎么引起的？从根本上说额外负担是消费者和生产者的价值损失，而这是由征税后市场交易量减少引起的。政府不能对已不存在的交易征税[一]。因此，对已经减少的这部分交易，政府拿不到任何税收收入。从社会的观点来看，这是个纯粹的损失（pure loss），或说额外损失（deadweight loss）。

我们也可以从额外损失的定义直接推知额外损失的大小，即我们直接计算已损失的这部分产量的社会价值。怎么计算？假设我们一开始位于征税前的那个均衡点上，现在我们向左移动。在均衡点对应的交易量上，消费者愿意支付的价格正好等于供给者愿意销售的价格。此时，消费者剩余和生产者剩余之和最大（请见图 16.7）。如果我们向左稍微移动一点点，也就是说交易量开始减少，此时消费者剩余和生产者剩余开始减少，即开始出现社会损失，尽管数额很小。

交易量从均衡点向左移动，需求价格和供给价格出现了分离。需求价格衡量消费者对商品的支付意愿，供给价格衡量供给者愿意供给商品的价格。需求价格和供给价格之间的差额，衡量由于交易量减少了那一单位而引起的价值损失。如果我们征税后交易量减少引起的价值损失全部加起来，就得到了额外损失。

例子：贷款市场

市场中资金的借入量和贷出量在很大程度上取决于利率。利率就是借贷资金的价格。

令 $D(r)$ 表示借入方对资金的需求，$S(r)$ 表示贷出方对资金的供给。均衡利率 r^* 因此由供需相等这个条件决定：

$$D(r^*) = S(r^*). \tag{16.1}$$

现在假设政府对资金借贷开始征税。那么，它对均衡利率有何影响？

在美国，任何贷出资金赚取利息的个人都要缴纳利息所得税。如果每个人处于同一纳税等级，比如说税率为 t；那么资金贷出者的税后利率为 $(1-t)r$。因此征税后，资金贷出量取决于税后利率，此时资金供给函数为 $S((1-t)r)$。

另一方面，美国国内税法允许借入者在应纳税收入中将应支付的利息扣除，因此，如果资金借入者和贷出者所处的税率等级相同，则资金借入者支付的税后利率为 $(1-r)t$。所以，资金需求函数为 $D((1-t)r)$。征税后，利率由下列式子决定：

$$D((1-t)r') = S((1-t)r').$$

可以注意到如果 r^* 是（16.1）式的解，则 $r^* = (1-t)r'$ 必定是（16.2）式的解，因此

$$r^* = (1-t)r',$$

或

[一] 至少目前政府还对这样的事情无能无力，但他们正致力解决这样的问题。

$$r' = \frac{r^*}{1-t}.$$

上式是说，税后利率是税前利率的$1/(1-t)$倍。税后利率$(1-t)r'$等于r^*怎么理解？你可以理解成税前利率就是$(1-t)r'$，因为它是（16.1）式的解。

借助图 16.8，我们可以把这个问题说得更清楚一些。征收利息所得税将会使资金供给曲线向左转动，征税后供给曲线的斜率变为原来的$1/(1-t)$倍（$1/(1-t)>1$因为$t<1$）；类似地，征税后资金需求曲线向右转动，需求曲线的斜率变为原来的$1/(1-t)$倍。由于在任一价格水平处，供给曲线和需求曲线移动的幅度相同，但方向相反，因此二者变动的净结果为：资金交易量不会发生变化，但价格提高——均衡利率变为原来的$1/(1-t)$倍。

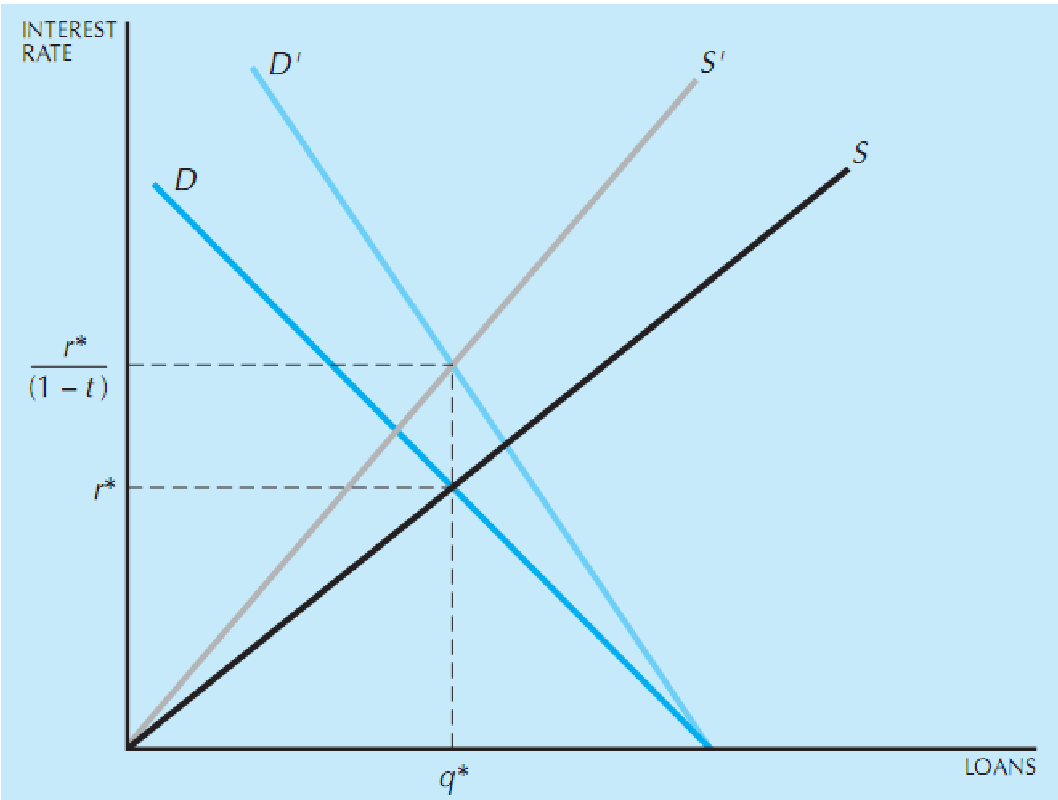


图 16.8：资金借贷市场的均衡。如果资金借入方和贷出方的税率相同，则税后利率变为原来的$1/(1-t)$倍，但资金交易量和原来相同。

我们也可以使用反需求曲线和反供给曲线分析这个问题。令$r_b(q)$表示资金借入方（borrows）的需求函数，这个函数是说为了让人们借入资金量为q，税后利率应为多大。类似地，令$r_l(q)$表示资金贷出方（lenders）的供给函数。于是，均衡交易量由下式给出：

$$r_b(q^*) = r_l(q^*). \tag{16.3}$$

现在假设政府开始征税。为了让这个问题更有意义，我们假设资金借入方和贷出方的税率等

级不同，即$t_b \neq t_l$。如果征税前市场利率为r，则借入方的税后利率为$(1-t_b)r$，他们决定借入的资金量由下式给出：

$$(1-t_b)r = r_b(q)$$

或者

$$r = \frac{r_b(q)}{1-t_b}. \tag{16.4}$$

类似地，资金贷出方的税后利率为$(1-t_l)r$，他们贷出资金数量决策由下式给出

$$(1-t_l)r = r_l(q)$$

或

$$r = \frac{r_l(q)}{1-t_l} \tag{16.5}$$

联立（16.4）式和（16.5）式可得到均衡条件：

$$r = \frac{r_b(\hat{q})}{1-t_b} = \frac{r_l(\hat{q})}{1-t_l}. \tag{16.6}$$

从上式容易看出，如果双方的税率等级相同即$t_b = t_l$，则$\hat{q} = q^*$。如果双方的税率等级不同，结果会如何？不难看出利息所得税对于资金借入方来说是补贴，对资金贷出方是税收，但这种税收的净效应如何？征税后若借入者面对的利率高于贷出者面对的利率，则税收的净效应为对**资金借贷行为**征税；但是，如果征税后若借入者面对的利率小于贷出者面对的利率，则税收的净效应为对资金借贷行为给与补贴[一]。改写（16.6）这个均衡条件的表达式，可得

$$r_b(\hat{q}) = \frac{1-t_b}{1-t_l} r_l(\hat{q}).$$

从上式可以看出，如果下式成立，资金借入者面对的利率将会高于贷出者的利率。

$$\frac{1-t_b}{1-t_l} > 1$$

这个式子意味着$t_l > t_b$。因此，如果资金贷出者的税率高于借入者，则利息所得税的净效应是对资金借贷行为征税，但是如果$t_l < t_b$，该税收的净效应是对资金借贷行为给与补贴。

[一] 理解这句话的关键就是，作出以下类比：资金就是一般商品，利率就是这种商品的价格，资金贷出者和借入者分别为这种商品的卖方和买方。不征税时，买方支付的价格和卖方得到的价格是统一的，但征税后出现了分离：若买方支付的价格高于卖方得到的价格，则征税的净效应就是对资金借贷这种**交易行为**征税；若买方支付的价格低于卖方得到的价格，则征税就是对这种交易行为进行补贴。译者注。

例子：食品补贴

19 世纪的英格兰，在收成不好的年份，富人会向穷人提供慈善资助，他们先把粮食全部买下，自己吃固定的数量，然后将剩下的粮食以他们购买价的一半卖给穷人。初看起来，穷人的状况会大幅改善，但仔细一想就不是那回事了。

要想使得穷人的状况变好，唯一的方法是最终他们能够吃到更多的粮食。但是既然富人自己吃固定的数量，对于某既定年份来说，这意味着穷人能吃到的粮食数量也是固定的。既然这样，你怎么能说穷人的状况因为富人的善举而变好了呢？！

事实上，穷人的状况未变好；不管富人是否行善，他们最终支付的粮食价格是一样的。为了分析原因，我们建立均衡模型。令 $D(p)$ 表示穷人的需求曲线；K 表示富人对粮食的需求数量，它是一个既定的常数；S 表示某收成不好年份的粮食供给量，显然，它也是一个既定的常数。由前面的假设可知，粮食的供给量 S 和富人对粮食的需求量 K 都是既定不变的常数。富人不行善时，均衡价格由总需求等于总供给决定：

$$D(p^*)+K=S.$$

富人行善时，均衡价格由下式决定

$$D(\hat{p}/2)+K=S.$$

现在你可以注意到：如果 p^* 是第一式的解，则 $\hat{p}=2p^*$ 是第二式的解。因此当富人以 $\hat{p}$ 的价格买光所有的粮食，并以 $\hat{p}/2$ 的价格卖给穷人时，市场价格将大幅上升，现在的市场价格是原来市场价格的 2 倍，因此穷人支付的价格和原来是一样的！

仔细想一下你就会知道上述结论并不奇怪。如果富人对粮食的需求和某年份的粮食产量都是固定不变的，那么穷人面对的均衡价格完全由穷人自身的需求曲线决定；穷人面对的粮食均衡价格是一样的，这和富人是否行善无关。

例子：伊拉克的补贴

政府会在某些情形下对居民进行补贴，这样的补贴显然是“具有正当理由的”，可是当上述情形不复存在时，要想取消这些补贴却非常困难。为什么？因为这些补贴培养出一大批支持政府的选民，如果你取消这些补贴，他们就不再支持你。这在任何国家都是一样的，但伊拉克的补贴表现得更为极端。以 2005 年为例，政府为燃料和食品提供的补贴将近占政府预算的三分之一[一]！

伊拉克几乎所有的政府预算来自石油出口。由于国内石油提炼能力非常有限，伊拉克以 30-35 美分/升的价格进口汽油，然后再以 1.5 美分/升的价格卖给国内居民。很大一部分进口的汽油在黑市上销售和走私到土耳其，土耳其的汽油价格大约为一美元每升。

[一] James Glanz, "Despite Crushing Costs, Iraqi Cabinet Lets Big Subsidies Stand," New York Times, August 11, 2005.

食品和燃油的补贴也很高。政客不愿意取消这些补贴，因为伊拉克的政治环境非常不稳定。也门国也有类似的补贴，但当这些补贴被取消时，也门国内发生了骚乱，很多人因此死去。世界银行的研究发现，伊拉克 GDP 一半以上的份额被用于补贴。伊拉克财政部长 Ali Abdulameer Allawi 曾经说道，“补贴已经达到令人发狂的程度。补贴以怪异的方式扭曲了经济，带来了最坏的经济激励。”

16.9 帕累托效率

帕累托效率（Pareto efficient）是指一种经济状态，在这种状态下已无法做到让任何人的状况进一步改善而又不损害其他人的状况。帕累托效率是个好东西——如果能找到让一部分状况变得更好，为何不这么做呢?——但效率不是经济政策的唯一目标。例如，效率通常和收入分配或经济公正（justice）无关。

然而，效率是一个重要的目标，因此有必要搞清完全竞争市场是否能实现帕累托效率。完全竞争市场，或者任何经济机制，必须决定两件事情。第一，应该生产多少；第二，谁得到这些产品。完全竞争市场根据以下信息决定产量的多少：与供给者供给产品要求得到的价格相比，消费者愿意支付的价格是更高还是更低。

请看图 16.9。在任何低于完全竞争产量 q^* 的产量水平处，消费者愿意为额外一单位产品支付的价格，高于生产者愿意供给额外一单位产品而索要的价格。

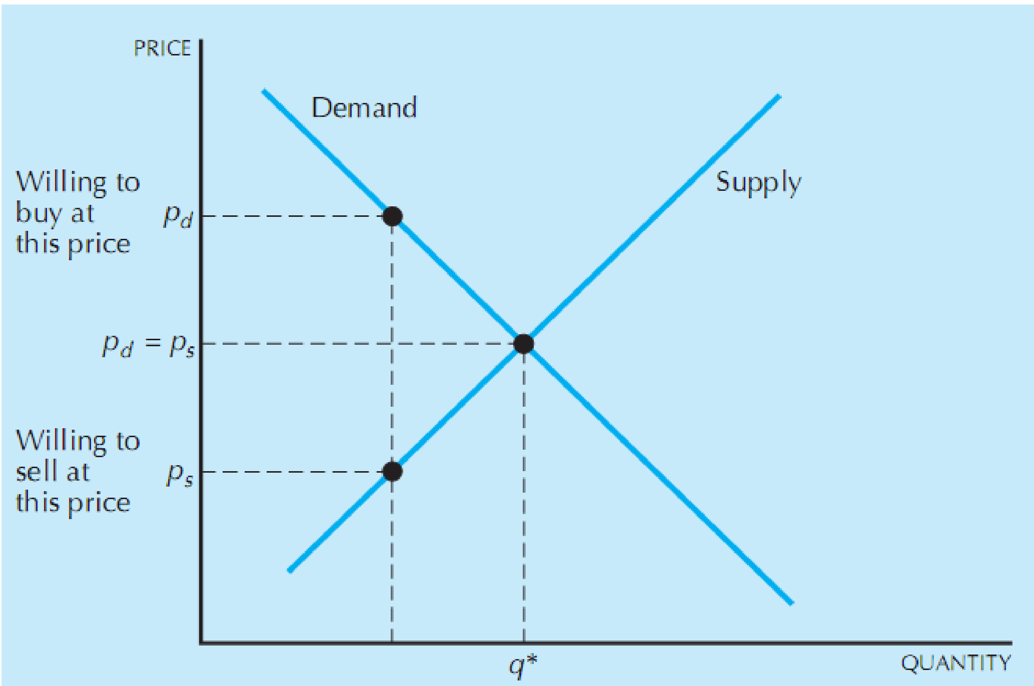


图 16.9：帕累托效率。完全竞争市场能够实现帕累托有效率的产量，因为在 q^* 处，消费者愿意支付的边际价格等于生产者供给的边际成本（额外供给一单位产品必须索要的价格补偿）。

在这样的产量水平上，如果生产者额外多生产一单位产品，如果这一单位产品的成交价位于需求价格和供给价格之间，那么生产者和消费者的状况都改善了。因此，任何低于均衡数量的产量不可能是帕累托有效率的，因为存在能使消费者和生产者状况都得到改善的方法。

类似地，在任何大于q^*的产量上，消费者对额外一单位产品愿意支付的价格，小于生产者额外生产一单位产品必须得到的价格。只有在市场均衡数量q^*处，才是帕累托有效率的，因为在这一产量上，消费者的边际支付意愿恰好等于生产者的边际生产成本（额外生产一单位产品必须得到的价格补偿）。

因此，完全竞争市场能生产帕累托有效率的产量。那么完全竞争市场如何分配产品的？在该市场机制下，任何人对某商品都支付相同的价格，因为该商品和"所有其他商品"的边际替代率必定等于该商品的价格。愿意支付该价格的人都能买到该商品，不愿意支付该价格的人都买不到该商品。

如果商品的分配方式不能做到该商品和"所有其他商品"的边际替代率相等，结果会如何？那么，至少有两人对商品的边际评价不同（即边际价值不同）。例如，一人认为商品的边际价值为 5 元，另外一人认为边际价值为 4 元。那么，如果这两个人交易，交易价格在 4 元到 5 元之间，则这两个人的状况都改善了。因此，任何分配方式只要它不能做到边际替代率相等，就必定不是帕累托有效率的。

例子：排队等候

一种常见的分配资源的方式是让人们排队等候，先到先得。在前面，我们使用边际支付意愿和边际生产成本（额外生产一单位产品所必须得到的价格补偿）相等这个工具分析了市场机制的效率，现在我们也可以使用这个工具分析排队等候这种资源分配方式的效率性。举一个具体的例子：假设你所在的大学准备分发篮球锦标赛的门票。门票免费，但需要排队领票，先到先得发完为止。

在这种情形下，门票的成本就是排队等候的时间成本。酷爱观看球赛的人将会在售票处门口熬夜等候以确保得到门票。对观看球赛无所谓的人，可能在售票处开门前几分钟才赶到，他们只是想碰碰运气看看还有没有剩下的票。人们对一张门票的支付意愿不再用金钱衡量，而是用等候的时间衡量，因为门票是按照等候的意愿分配的。

排队这种分派方式能实现帕累托有效率吗？问问你自己是否会发生以下的情形：排队领到票的人将票卖给那些不愿意排队的人。答案通常是肯定的，原因很简单——人们的支付意愿和等候意愿通常不同。如果有人愿意排队领票并且领到票后再卖给其他人，那么按照等候意愿分配门票的这种方式，没有将交易的好处取尽——因为在门票被领完后，还有人愿意交易。既然排队这种方式并没做到将交易的好处取尽，它就不可能是帕累托有效率的。

如果商品是按照支付意愿（即该商品的价格）分配的，那么需求者支付的资金将使供给

者受益。但是用排队的方式分配商品，排队等候的时间对任何人都没好处。等候时间对商品的需求者施加了成本，但对供给者却无任何好处。排队这种分配商品的方式是一种**净损失**（deadweight loss）——排队等候的人支付了“价格”，但是没有人从这些“价格”中受益。

总结

1.供给曲线衡量在每一价格水平上人们愿意供给某商品的数量多少。

2.均衡价格是指在该价格水平上，人们愿意供给的数量等于人们愿意需求的数量。

3.需求曲线和供给曲线移动时，均衡价格和均衡数量是如何变动的？——这个问题是一个比较静态分析问题。

4.对某商品征税时，通常存在两种价格：需求者支付的价格和供给者得到的价格。二者之差等于单位税额。

5.供给者能向消费者转嫁多少税收取决于供给曲线相对于需求曲线的倾斜程度。如果供给曲线是水平的，全部税收都转嫁给了消费者；如果供给曲线是垂直的，则转嫁给消费者的税收等于零。

6.税收的额外损失（deadweight loss）等于征税引起的消费者剩余的净损失（net loss）和生产者剩余的净损失之和。它衡量征税导致的产量损失的价值。

7.某经济状态是帕累托有效率的，如果已无法做到让某些人的状况改善而又不损害其他人的状况。

8.某商品市场的帕累托有效率产量是指需求曲线和供给曲线相交处对应的产量，因为只有在这一点上，需求者对额外一单位产品的支付意愿等于供给者愿意供给额外一单位产品而必须得到的价格。

复习题

1.如果某市场的供给曲线是水平的，那么补贴将有什么效应？如果供给曲线是垂直的呢？

2.假设需求曲线是垂直的但供给曲线是向上倾斜的，如果征税，那么谁是税收的最终承担者？

3.假设所有消费者都认为红铅笔和蓝铅笔是 1:1 完全替代的，假设红铅笔的供给曲线是

向上倾斜的。令红铅笔和蓝铅笔的价格分别为 p_r 和 p_b。如果政府只对红铅笔征税，结果将是怎样的？

4.美国石油需要量中大约有一半是进口的。假设外国石油生产者能够以 25 美元每桶的价格供给美国想要的任何数量。如果对美国对进口的国外石油征收 5 美元每桶的税，美国国产石油价格将会怎样变化？

5.假设供给曲线是垂直的，那么征税引起的净损失是多少？

6.考虑教材中资金借贷市场中征收利息所得税的情形。如果资金借入方和贷出方的税率相同，那么政府可以征收到多少税收收入？

7.在上题中，如果资金贷出方面对的税率小于资金借入方的，即 $t_l < t_b$，政府税收收入为正还是负？

复习题答案

1.如果某市场的供给曲线是水平的，那么补贴将有什么效应？如果供给曲线是垂直的呢？

【复习内容】特殊情形下（供给曲线水平时以及供给曲线为垂直时）的市场均衡；税收（补贴）的转嫁；比较静态分析

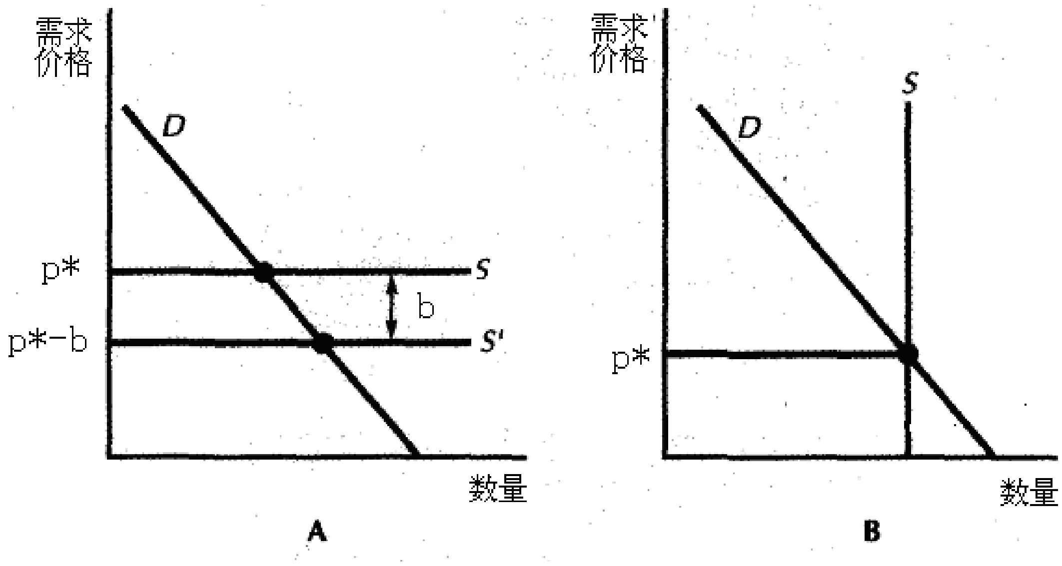
如果某行业的供给曲线是水平的，这意味着该行业在某既定的价格水平处能提供它想提供的任何数量，但是如果市场价格低于上述价格水平，它的供给量为零。在这种情形下，价格完全由供给曲线决定，销量完全由需求曲线决定。如果某行业的供给曲线是垂直的，这意味着它供给的数量是固定不变的。此种情形下，均衡价格完全由需求曲线决定。

税收（补贴）转嫁程度取决于供给曲线相对于需求曲线的斜率。

【参考答案】

假设政府向生产者给与补贴。其实补贴和税收类似，都是对交易行为进行的，因此涉及转嫁。为了分析的方便，我们假设给与生产者补贴，这样，补贴影响的是供给曲线的移动。当然你也可以假设补贴是补贴给消费者的，但该情形下，补贴影响的是需求曲线的移动。

在供给曲线水平（即具有完全弹性）的情形下，如果对生产者给与补贴，则供给曲线向下移动，则消费者支付的价格下降了，下降额度恰好等于单位补贴 b。供给价格和补贴前的供给价格相等，消费者最终获得了全部补贴（见下图 A）。



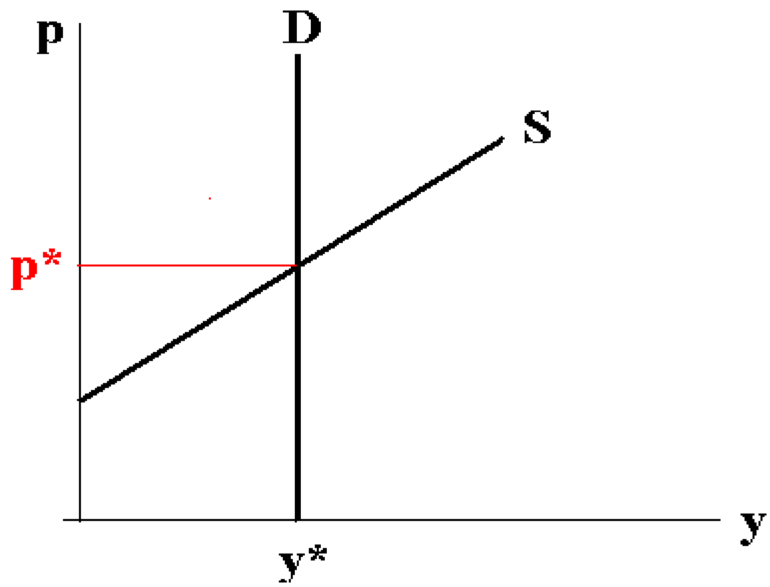
供给曲线垂直（即完全无弹性）的情形下，如果我们“将供给曲线向下移动”，则我们什么也没改变。供给曲线沿着自身向下移动，商品的供给数量没变。由上图 B 可知，政府给予生产者补贴后，消费者支付的价格仍然为 p^*，因此生产者（供给者）得到了全部补贴。

2.假设需求曲线是垂直的但供给曲线是向上倾斜的，如果征税，那么谁是税收的最终承担者？

【复习内容】市场均衡；税收转嫁；比较静态分析

税收转嫁程度取决于供给曲线和需求曲线的相对倾斜程度。谁越平坦（谁的弹性越大）意味着谁对税收的规避能力越强；谁越陡峭（谁的弹性越小）意味着谁的税收规避能力越弱。

【参考答案】



请见上图。由于需求曲线是垂直的，即需求曲线完全无弹性，这意味着需求者对税收的

规避能力为 0；而供给曲线是向上倾斜的，意味着供给者具有一定的税收规避能力。

为方便论述，假设政府对消费者征税，则需求曲线会向下移动。由于本题中的需求曲线为垂线，因此需求曲线向下移动意味着需求曲线“原地不动”。征税之前生产者得到的价格为 p*，征税之后，生产者得到的价格仍为 p*，这意味着税收完全由消费者承担。

3.假设所有消费者都认为红铅笔和蓝铅笔是 1:1 完全替代的，假设红铅笔的供给曲线是向上倾斜的。令红铅笔和蓝铅笔的价格分别为 p_r 和 p_b 。如果政府只对红铅笔征税，结果将是怎样的？

【复习内容】完全替代；需求具有完全弹性；征税的效应

【参考答案】

假设政府对每支红铅笔征税 t 元。由于红铅笔的供给曲线是向上倾斜的，也就是说供给具有一定的价格弹性，生产者通常承担部分税收，除非红铅笔的需求具有完全弹性（需求曲线水平），这种情形下生产者才会承担全部税收。

不妨假设消费者承担的税额为 at ，其中 a 为常数且 $a \in [0,\ 1]$

如果 $p_r + at < p_b$ ，则消费者仍然只会购买红铅笔；相反，如果 $p_r + at > p_b$ ，则消费者只会购买蓝铅笔，这时生产红铅笔的企业会退出市场或者改成生产蓝铅笔。

如果 $p_r + at = p_b$ ，那么消费者对红铅笔和蓝铅笔的需求是任意的，这意味着在这种情形下，红铅笔的需求曲线为一条水平线。既然是一条水平线，这说明此时红铅笔的需求具有完全弹性。因此，税收将由生产者完全承担，即 $a = 0$ 。

4.美国石油需要量中大约有一半是进口的。假设外国石油生产者能够以 25 美元每桶的价格供给美国想要的任何数量。如果对美国对进口的国外石油征收 5 美元每桶的税，美国国产石油价格将会怎样变化？

【复习内容】供给具有完全弹性；征税的效应

【参考答案】美国国产石油价格将为 30 美元/桶。

题目中“假设外国石油生产者能够以 25 美元每桶的价格供给美国想要的任何数量”这句话意味着，国外石油的供给曲线是一条水平线，即 p=25。

由于美国对石油的需求不可能具有完全弹性，最有可能是具有通常的负斜率，因此综合以上的供给曲线和需求曲线可知，政府对国外石油的征税 5 美元/桶，完全由美国国内的消费者承担。这意味着进口石油在美国的市场价格（含税价）位 30 美元/桶。

由于对于美国国内的消费者来说，进口石油和国产石油是完全替代的，因此国产石油价格必然也为 30 美元/桶。这是因为若国产石油价格低于 30 美元/桶，则消费者只会消费国产石油，因此将国内价格一直抬升到 30 美元为止；若美国国产石油价格高于 30 美元/桶，则

国内石油的需求量为零。

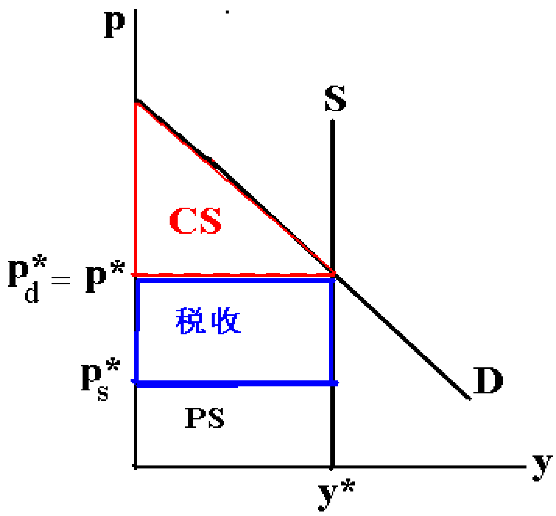
5.假设供给曲线是垂直的，那么征税引起的净损失是多少？

【复习内容】供给完全缺乏弹性；征税效应

【参考答案】征税引起的净损失为零。

由于供给曲线是垂直的，这意味着供给完全缺乏弹性，在需求曲线具有正常斜率的情况下，税收完全由生产者承担，税收没有影响企业的产量。既然没有影响企业的产量，因此征税引起的净损失为零。

如果你喜欢用总剩余的变动衡量税收的净损失，这种方法也可以。不妨假设政府对生产者征税，那么供给曲线会“向下”移动，由于供给曲线垂直，“向下”移动意味着供给曲线的位置根本没发生变化。



因此，征税之前和之后的均衡数量均为 y^*，征税之前，消费者支付的价格=生产者得到的价格= p^*；征税之后，消费者支付的价格和生产者支付的价格发生了分离，征税之后消费者支付的价格 p_d^* =征税之前消费者支付的价格，这意味着税收完全由生产者承担，如上图所示。

征税之后生产者得到的价格下降了，此时生产者得到的价格为 $p_s^* = p^* - t$，其中 t 为政府征收的单位税额。

征税前的总剩余=消费者剩余 CS（图中红色三角形）+生产者剩余（等于图中的税收矩形与 PS 矩形之和）

征税后的总剩余=消费者剩余 CS（图中红色三角形）+生产者剩余 PS（图中的 PS 矩形）

征税后的总剩余减少了图中税收矩形的面积，但这一部分被政府拿走了，相当于财富从生产者手里转移到政府手里，总财富未变。

因此，给定题目中的条件，征税引起的净损失为零。

6.考虑教材中资金借贷市场中征收利息所得税的情形。如果资金借入方和贷出方的税率相同，那么政府可以征收到多少税收收入？

【复习内容】利息所得税

贷出资金赚取利息的需要缴纳利息所得税，但资金借入者在应纳税收入中可将应支付的利息扣除。这意味着利息所得税对于资金贷出方是税收，但对于资金借入方来说是补贴。这种税收的净效应如何？征税后若借入者面对的利率高于贷出者面对的利率，则税收的净效应为对**资金借贷行为**征税；但是，如果征税后若借入者面对的利率小于贷出者面对的利率，则税收的净效应为对资金借贷行为给与补贴。

理解上一句话的关键就是，作出以下类比：资金就是一般商品，利率就是这种商品的价格，资金贷出者和借入者分别为这种商品的卖方和买方。不征税时，买方支付的价格和卖方得到的价格是统一的，但征税后出现了分离：若买方支付的价格高于卖方得到的价格，则征税的净效应就是对资金借贷这种**交易行为**征税；若买方支付的价格低于卖方得到的价格，则征税就是对这种交易行为进行补贴。

【参考答案】税收净收入为零。

贷出资金赚取利息的需要缴纳利息所得税，但资金借入者在应纳税收入中可将应支付的利息扣除。这意味着利息所得税对于资金贷出方是税收，但对于资金借入方来说是补贴。

如果资金贷出方和借入方的税率相等，这意味着政府从资金贷出方征得的税收，恰好全部补贴了资金借入方，因此税收净收入为零。

举个简单的例子更能看明白，假设资金借贷产生的利息为 100 元，也就是说资金借入方除了向贷出方归还本金外，还要偿还 100 元的利息。假设贷出方和借入方的税率 t 都为 10%。

为简单起见，假设除此之外资金贷出方没有其他应纳税收入，此时他需要缴纳利息所得说为$100\times10\%=10$元；为简单起见，假设资金借入方也没有其他应纳税收入，或者说其他应纳税收入为 0，那么借入方应缴纳的利息所得税=$(0-100)\times10\%=-10$元，这是由于按照税法规定，资金借入方可将利息在应纳税收入中扣除。由于借入方的纳税为负，这意味着他得到了正的补贴 10 元。所以政府对这笔交易的征税净收入=$10+(-10)=0$元.

7.在上题中，如果资金贷出方面对的税率小于资金借入方的，即$t_l<t_b$，政府税收收入为正还是负？

【复习内容】利息所得税

【参考答案】税收净收入为负。

贷出资金赚取利息的需要缴纳利息所得税，但资金借入者在应纳税收入中可将应支付的利息扣除。这意味着利息所得税对于资金贷出方是税收，但对于资金借入方来说是补贴。

如果资金贷出方的税率 t_l 小于资金借入方的税率 t_b，这意味着政府从贷出方征得的税收不足以对借入方的补贴，因此税收净收入为负。

仍以上题答案的例子为例。为完整起见，将上题答案中例子的主要信息拷贝如下：假设资金借贷产生的利息为 100 元，也就是说资金借入方除了向贷出方归还本金外，还要偿还 100 元的利息。假设贷出方的税率为 10%，借入方的税率为 15%。

为简单起见，假设除此之外资金贷出方没有其他应纳税收入，此时他需要缴纳利息所得说为 $100\times10\%=10$ 元；为简单起见，假设资金借入方也没有其他应纳税收入，或者说其他应纳税收入为 0，那么借入方应缴纳的利息所得税=$(0-100)\times15\%=-15$ 元，这是由于按照税法规定，资金借入方可将利息在应纳税收入中扣除。由于借入方的纳税为负，这意味着他得到了正的补贴 15 元。所以政府对这笔交易的征税净收入=$10+(-15)=-5$ 元.

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

17.拍卖（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

17 拍卖

拍卖是最古老的市场形式之一，其源头至少可以追溯到公元前 500 年。今天所有的商品，从二手电脑到鲜花，都可以使用拍卖方式销售。

1970 年代早期，由于 OPEC 石油卡特尔对石油提价，经济学家开始对拍卖感兴趣。美国沿海地区可能含有大量石油蕴藏，内务部决定使用拍卖方法出售这些地区的石油开采权。政府让经济学家帮忙设计拍卖方案，私有公司也让经济学家帮忙制定投标策略。这样，人们开始大力研究拍卖方案和投标策略。

近来，联邦通讯委员会决定拍卖部分无线电频谱，这些频谱段可用于手机、个人电子助手或者其他通讯设备。经济学家再一次扮演了主角，他们帮忙设计了拍卖方案和投标策略。这些拍卖被认为是成功的公共政策，因为截止到目前它给政府带来的收入已超过了 230 亿美元。

其他国家也使用拍卖方式将某些项目私有化。例如，澳大利亚卖掉了若干政府拥有的发电厂，新西兰将部分国有电话系统出售。

消费者导向的（consumer-oriented）拍卖在互联网上也有复兴的趋势。互联网上有很多拍卖，例如拍卖收藏品、计算机设备、旅行设备和其他东西。美国 Onsale 公司宣称自己是最大的网络拍卖公司，1997 年它拍卖了价值 4100 万美元的商品。

17.1 拍卖的分类

拍卖的经济学分类涉及两个方面：首先，拍卖的商品具有什么样的性质？其次，竞价规则是什么样的？从第一个方面即商品的性质来说，经济学家将拍卖分为**个人价值拍卖**（private-value auctions）和**共同价值拍卖**（common-value auctions）。

在个人价值拍卖中，每个竞买人对拍卖的商品的评价不同。一件特别的艺术品对某个收藏家来说可能价值 500 元，对另外一个人可能只值 200 元，甚至还有人认为它只值 50 元，具体价值多少取决于个人的偏好。在共同价值拍卖中，拍卖的商品对每个竞买人来说价值大致相同，尽管他们的具体评价存在着差异。例如前文提及的沿海石油开采权的拍卖就是共同价值拍卖：某个既定地带可能蕴含着一定量的石油，也可能没有。不同的石油公司对该地带石油蕴藏量的估计不同，因为它们的地质勘探结果通常不同，但是不管谁赢得拍卖石油的市场价值是相同的。

本章主要介绍个人价值拍卖，因为人们对这种拍卖最为熟悉。在本章的结尾处，我们

介绍共同价值拍卖的一些特征。

竞价规则

最为流行的拍卖方式称为**英式拍卖**（English auction）。拍卖人（auctioneer）从保留价格开始拍卖，这个价格是委托人（卖主）卖掉商品所愿意接受的最低价格[一]。不同投标人或称竞买人(bidders)连续报出更高的价格；一般来说，每次报价都必须比前一个报价至少高出一个既定的最小数额，这个数额称为**竞价增量**（bid increment）。当竞价人不愿意进一步提高报价时，拍卖的商品由报价最高的人获得。

另外一种拍卖方式称为**荷兰式拍卖**（Dutch auction），这是由于荷兰人最早使用这种方式销售奶酪和鲜花。在这种拍卖方式中，拍卖人从某高价处起拍，然后逐步降低价格，直到有人愿意购买。在荷式拍卖的实务中，“拍卖人”通常是类似钟表的机械设备，它有一个指针，随着拍卖的进行，指针逐渐指向越来越低的价值数字。荷式拍卖耗时较短，这是它的一个主要优点。

第三种拍卖方式叫做**密封拍卖**（sealed-bid auction）。在这种拍卖方式中，每个投标人将自己的报价写在纸条上并用信封密封起来。拍卖人将这些信封收集然后开封，报价最高的人得到了商品，然后他要如数交钱。如果拍卖的商品有保留价格，而且所有的报价都低于保留价格，则卖主将商品收回。

建筑工程中的招标通常使用密封拍卖方式。想建楼的人向若干建筑承包商发出投标邀请，承包商投标，中标者为报价最低者。

最后，我们介绍密封拍卖的一种变形，即**集邮者拍卖**（Philatelist auction）或者**维克里拍卖**（Vickrey auction）。第一个名字是因为这种拍卖最初是由集邮者使用的；第二个名字是为了纪念威廉.维克里（William Vickrey），他对拍卖的开创性研究帮助他获得了 1996 年度诺贝尔经济学奖。维克里拍卖的方式类似密封拍卖，但有一个重要区别：商品由报价最高的竞价人获得，但他只需要按**第二高**的报价支付。换句话说，报价最高的投标人得到了拍卖商品，但是他不需要按照他自身的报价支付，而是按照报价第二高的人的报价支付。尽管最初你可能觉得这种拍卖方式很奇怪，但是下面我们将发现，它具有一些很好的性质。

17.2 拍卖设计

假设我们有件商品要拍卖，n 个投标人对商品的评价分别为 $v_1, v_2, \ldots, v_n$。为简单起见，假设这些评价都是正的而且卖主的保留价格为零。我们的目的是设计一种拍卖方案将该商品卖出。

[一] 参见第 6 章关于“保留价格”的脚注。

这是经济机制设计（economic mechanism design）问题的一种特殊情形。在拍卖机制设计中我们自然必须考虑以下两个目标：

- 帕累托效率（Pareto efficiency）。设计出的拍卖要能实现帕累托有效率的结果。
- 利润最大化（Profit maximization）。设计出的拍卖要能为卖主带来最高的期望利润。

利润最大化目标简单易懂，但是此处的帕累托效率应怎么理解？不难看出，帕累托效率要求商品卖给**评价**最高的人。为了说明这一点，假设投标人 1 对商品的估计（v_1）最高，投标人 2 对商品的评价（v_2）略低。如果投标人 2 得到了商品，则存在让这两人的状况都变化的方法：投标人 2 以位于 v_1 和 v_2 之间的价格 p，将商品卖给投标人 1。由此可见，如果不是评价最高的人得到商品则不可能是帕累托有效率的。

如果卖主知道这 n 个投标人的个人价值 $v_1, v_2, ..., v_n$，则拍卖设计问题就非常简单。如果拍卖的目标是利润最大化，卖主只要将商品卖给评价最高的竞买人，而且卖主向他索要的价格恰好等于该竞买人的评价即可[一]。如果拍卖的目标是帕累托效率，评价最高的投标人仍然必须得到商品，但是他支付的价格可为他的评价和零之间的任何数值，因为剩余分配和帕累托效率无关[二]。

更有意义的情形是卖主不知道投标人个人价值的情形。在这种情形下，如何达到帕累托效率或利润最大化的目标。

首先考虑帕累托效率。不难看出英式拍卖能实现帕累托效率，因为评价最高的投标人最终得到了商品。我们需要花点时间考虑的是怎么确定该中标人支付的价格：**第二高**的**报价**，或许再加上最小竞价增量。

举个具体的例子说明。假设在某次拍卖中，最高评价为 100 元，第二高的评价为 80 元，竞价增量为 5 元。因此，评价 100 元的投标人愿意报价 85 元，但是评价 80 元的人不会愿意报价 85 元（因为 80 元是他的保留价格，即他愿意出的最高价格）。就象我们以前宣称的，评价最高的个人得到了商品，但是他实际支付的价格却是第二高的报价（也许再加上竞价增量）。（我们反复说“也许”，这是因为如果两个投标人的报价都是 80 元，则出现了平局，最终谁能得到商品取决于事先制定的打破平均的规则，比如如果出现平均则通过掷硬币来决定谁最终能得到商品。）

如果卖主不知道投标人的评价，怎么实现利润最大化？这要比分析帕累托效率要难。因为能否实现利润最大化取决于卖主对投标人评价的**认知**（beliefs）。为了说明这一点，假设

[一] 竞买人对拍卖商品的出价未必等于他对商品的评价（保留价格，即他能接受的最高价格）。事实上，竞买人的出价肯定小等于他自己的保留价格。但是如果卖主知道竞买人的保留价格，显然卖主会向竞买人索取该保留价。译者注。

[二] 假设某人对商品的评价最高，他认为该商品价值为 500 元，为简单起见假设其他人的评价均为 100 元且报价都为 0，那么此人无论是出价 1 元还是 10 元**都是**帕累托有效率的，因为这两种出价下他都能得到商品。如果他出价 1 元，他得到的消费者剩余为 499 元，卖主的生产者剩余为 1 元；如果他出价 10 元，他的消费者剩余为 490 元，卖主的生产者剩余为 10 元。两种情形下剩余总和均为 500 元，只不过剩余的分配方式不同。**帕累托效率仅取决于社会福利（此处为消费者剩余和生产者剩余）总和，而和具体的分配方式（谁得到剩余的多少）无关**。译者注。

曹乾（东南大学 caoqianseu@163.com）

有两个投标人，他们对商品的评价为：要么 10 元，要么 100 元。假设这两种情形发生的概率相同，因此，（投标人 1 的**评价**，投标人的**评价**）就有以下四种：（10, 10），（10，100），（100, 10），（100, 100）。最后，假设最小竞价增量为 1 元，而且如果出现平局则通过掷硬币的方式来解决。

我们先来分析一下，在上述四种评价情形中，谁会中标？在第一种情形和第四种情形中，都为平局，因此通过投掷硬币决定谁中标。在第二种情形和第三种情形下，评价最高(100元)的人，只要报价 11 元（等于第二高的价格 10 元再加上最小竞价增量 1 元）即可确保中标。因此，上述四种情形下，中标者的**报价**为（10,11,11,100），**评价**最高的人总会得到商品。卖主的期望收益为 33 元=$\frac{1}{4}(10+11+11+100)$。

卖方能得到更大的利润吗？当然，如果他事先设定了合适的保留价格。在这个例子中，能使利润最大化的保留价格为 100。回头看看上述四种报价情形，你就知道卖主按 100 元卖掉商品的概率为 3/4，卖不出去的概率为 1/4。这样，他的期望收益为 75 元，这显然高于不设定保留价格时的期望收益。

注意上述设定保留价格的策略不是帕累托有效率的，因为在四分之一的概率下无人得到商品。这和垄断造成的净损失类似，产生的原因也是相同的。

如果拍卖的目标是利润最大化，则设定保留价格是非常重要的。在 1990 年，新西兰政府使用了维克里拍卖方式，拍卖了部分无线电频谱，这些频谱段可用于收音机、电视机和手机。在其中一次拍卖中，中标者的报价为 10 万新西兰元，而第二高的报价仅为 6 新西兰元！这个拍卖也许能得到帕累托有效率的结果，但它不是收入最大化的！

我们已经看到，在英式拍卖中，如果保留价格设定为零，可以保证实现帕累托效率。那么，荷兰式拍卖能实现帕累托效率吗？答案是未必。为了看清这一点，假设有两个投标人，他们对商品的评价分别为 100 元和 80 元。如果评价最高的人（错误地！）认为第二高的评价为 70 元，那他的策略可能是等候，一直等到拍卖人喊价 75 元时才竞价。但，此时为时已晚，因为评价第二高的人已经以 80 元的价格得到了这件商品。一般来说，在荷兰式拍卖中，我们无法保证评价最高的人一定得到商品。

密封拍卖也无法保证做到帕累托效率。每个投标人的最优报价策略取决于他对其他人评价的认知。如果这些认知不准确，商品一般会被评价不是最高的人获得[一]。

最后，我们来看看维克里拍卖，它是密封拍卖的变种。和密封拍卖相同，维克里拍卖也是报价最高的人得到商品，但不同之处在于，维克里拍卖中，报价最高的人得到商品但他只要按第二高的报价支付。

[一] 另一方面，如果每个投标人对其他投标人的评价认知是准确的，而且假设每个投标人的投标行为都是最优的，那么上述各种拍卖方式的结果（谁最终得到商品）是相同的，并且均衡时期望价格也相等。更详细地分析参见：P. Milgrom, "Auctions and Bidding: a Primer," Journal of Economic Perspectives, 3(3), 1989, 3-22, and P. Klemperer, "Auction Theory: A Guide to the Literature," Economic Surueys, 13(3), 1999, 227—286.

首先，我们可注意到，**如果**每个人的报价都等于他自己的真实评价，则评价最高的人最终会得到商品，但他支付的价格等于第二高的评价。维克里拍卖在本质上和英式拍卖的结果相同(这取决于竞价增量。也就是说，只要竞价增量非常小，这二者的结果是完全相同的)。

但是，在维克里拍卖中如实报出你的真实评价是最优策略吗？我们已经知道，在标准的密封拍卖中，这通常不是最优策略。但维克里拍卖情形不同：如实报价符合每个投保人的利益，这多少有些出乎意料。

为了看清这一点，我们分析一个只有两个投标人的特殊情形。这两人的对商品的评价分别为v_1和v_2，他们在纸条上写下的报价分别为b_1和b_2。投标人 1 的期望收益（expected payoff）为：

$$\Pr ob(b_1 \geq b_2)[v_1 - b_2],$$

其中“Prob”表示概率（probalility）。

上式中的第一项是投标人报价在所有投标人的报价中为最高报价的概率；第二项是如果投标人中标他享有的消费者剩余。(如果$b_1 < b_2$，则投标人 1 得到的消费者剩余为零，因此无需考虑含有$\text{Prob}(b_1 \leq b_2)$的项)。

假设$v_1 > b_2$。因此投标人 1 希望使自己的中标概率尽可能大，所以他会将报价设定为等于自己的评价即$b_1 = v_1$。另一方面，假设$v_1 < b_2$，则他希望自己中标概率尽可能小，所以他会使得$b_1 = v_1$。在两种情形下，投标人 1 的最优报价策略就是令报价等于自己的真实评价！诚实是最好的策略…至少在维克里拍卖中是这样的！

维克里拍卖的一个有趣特征是，它在本质上可以实现与英式拍卖相同的结果，但与英式拍卖不同的是，它不存在投标人的相互影响。这也是为什么集邮者使用这种拍卖方式的原因。在邮市上他们通过英式拍卖销售邮票，而在散市后他们通过邮件进行密封拍卖。有人已经注意到，如果密封拍卖采用第二高报价的规则，则它的结果和英式拍卖类似。但是，对集邮者拍卖的成熟分析要归功于维克里，他不仅证明了说真话是最优的策略，而且证明了集邮者拍卖和英式拍卖是等价的。

17.3 拍卖的其他形式

人们一度认为维克里拍卖意义有限，但在网络拍卖大量涌现后人们改变了这一看法。世界上最大的在线拍卖商 eBay，宣称它的注册用户已达 0.3 亿，2000 年它的销售额为 50 亿美元[一]。

最初，eBay 组织的拍卖要持续几天甚至几个星期，用户很难连续监督拍卖过程。为了方便客户监督，eBay 引入了一种自动**报价代理人**（bidding agent）机制，他们将其称为**投标代理人**（proxy bidder）。对于某商品，客户告诉投标代理人两个价格：一是他愿意支付

[一] 2010 年,eBay 注册用户已近 2 亿，年交易额约在 100 亿美元。译者注。

的最高价格；二是初始报价。随着竞价进行，代理人会在必要时按最小竞价增量自动提高投标人的报价，当然这个价格不能高于投标人愿意出的最高价格。

本质上，这就是一种维克里拍卖：每个客户告知代理人他愿意支付的最高价格。在理论上，报价最高的投标人将获得商品，但他支付的价格却是第二高的报价（为避免出现平局，有时需要再加上一个最小竞价增量。）根据前文的分析，每个投标者都有动机告知代理人他对拍卖商品的真实评价。

在实务中，投标人的行为和维克里模型的预测结果稍微有所不同。通常，投标人会一直等到拍卖临近结束才开始报价。这种行为可能有两种不同的原因：一是不愿意在博弈中过早地显露兴趣；二是希望在只剩下少数投标人时获取更大的利益。然而，竞价代理人模型似乎能很好地为用户服务。维克里拍卖，人们一度认为它只有理论价值，现在却是世界最大在线拍卖商衷爱的方式。

拍卖实务中，有些拍卖形式更为奇特。例如，一种称为**双人支付拍卖**（escalation auction）的拍卖形式，在这种拍卖中，报价最高的人得到商品，但是报价最高的人**和**报价第二高的人都需要按照自己的报价支付。

例如，假设你使用这种拍卖方式拍卖 1 元钱硬币。通常一些人会报价 10 分钱或者 15 分钱，但是最终大多数人投标人放弃。当最高报价接近 1 元时，仍在坚持的投标人开始认识到自己面对的问题。如果某人报价 90 分，另一人报价 85 分，后者开始意识到如果他不再提高报价，那么他将支付 85 分钱但什么也得不到，如果他将报价提到高 95 分钱，他就可以得到那个 1 元钱硬币。

但是一旦他这么做了，报价 90 分的人也会同样推理。事实上，报价**超过** 1 元钱对他有利。例如，如果他报价 1.05 元（胜出），他会损失 5 分而不是 90 分钱！为 1 元钱最终报价 5 元或 6 元的情形也不罕见。

与双人支付拍卖多少相关的拍卖是**人人支付拍卖**（everyone pays auction）。例如某个不道德的政客宣称，一旦他当选他将为满足下列条件的项目投票：人人都要为他的竞选出钱，但是在他当选后他只会为出钱最多的那个人的项目投赞成票。这在本质上就是一种人人支付拍卖——人人都出钱，但是只有报价最高的投标者会得到他所想要的东西！

例子：eBay 上姗姗来迟的投标者

根据标准的拍卖理论，eBay 的投标代理人应该能诱使客户的报价等于他自身对商品的真实评价。正如维克里拍卖支出的，报价最高的人获得商品但只需要按照第二高的报价支付。但是，这在实践中却行不通。在很多拍卖业务中，投标人一直等到最后时刻才报价。一项研究表明，37%的拍卖在拍卖计时结束前一分钟内才有人报价，12%的拍卖在最后 10 秒内才有人报价。为何有那么多的“姗姗来迟的投标（late bids）”现象？

至少有两种理论能解释这种现象。Patrick Bajari 和 Ali Hortacsu 这两位投标专家认为，对于某些拍卖来说，人们不希望过早的投标，因为这样会抬升销售价格。eBay 通常公示投标人的身份和他们对拍卖商品的实际报价（不是最高价格）。如果你是邮票方面的专家，而且你在 eBay 上的网名路人皆知，你当然不会过早投标，因为这样会暴露你对某邮票有兴趣。

这种理论能够解释象邮票和硬币这类收藏品的拍卖现象，但是在一般商品例如计算机配件的拍卖中为何也会有姗姗来迟的投标者呢？A1 Roth 和 Axel Ockenfels 认为这是由于他们想避免价格大战。

假设你和另一人竞买一个 Pez 糖果盒，卖方的保留价格为 2 元。恰巧你们二人对该糖果盒的评价都为 10 元。如果你们俩都过早地投标——过早地表明你们的真正最高评价为 10 元。这是一个平局，假设你最终有幸胜出，你为这个糖果盒支付了多少钱？你需要按照第二高的报价支付，而这正是另一人的报价——10 元。没错，你是“胜出”了，但你的消费者剩余为零！

或者，假设你们一直等到拍卖快要结束前的几秒才投标，你们当然是向 eBay 投标代理人报出两个价格，一是最高价格 10 元；二是初始报价 2 元。（在 eBay,这被称为“猎杀（sniping）”）在这种情形下，通常会有一个人的投标手续没来得及完成，这样中标者最终支付的价格仅为卖方的保留价格，也就是 2 元。

在拍卖快结束前才投标，使得拍卖结果具有某些随机性。其中一个投标者获益很大，另一人获益为零。但是站在投标人的角度上，这种情形要比过早投标的结果要好，因为过早投标，其中一人支付的价格等于第二高的价格，另外一人获益仍为零。

在这个例子中，投标人姗姗来迟是一种“隐性合谋（implicit collusion）”。通过在拍卖即将结束前再投标让运气来说话，投标人最终获得的收益，平均来说要比过早投标好。

17.4 位置拍卖

位置拍卖（position auction）是一种拍卖位置的方法，例如线上位置或网页位置上的拍卖。它的特征是所有竞价人（登广告的人）对这些位置的排序是相同的，但是他们可能对这些位置的评价不同。每个人都认为排在前面的广告位置比排在后面的好，但是对于同样的位置，他们愿意支付的钱数是不同的。

位置拍卖最显著的例子，是搜索引擎提供者（例如 Google, Microsoft, Yahoo 等）拍卖广告。在这种情形下，所有打算登广告的人认为最高的位置最高，次高的位置第二好，依次类推。然而，这些人通常销售不同的产品，因此广告带给他们的期望利润是不同的。

下面我们描述这些在线广告拍卖的一个简化版本。尽管搜索引擎提供者拍卖方法略有差异，但下面的模型能描述它们的一般特征。

假设可以登广告的位置有 $s=1,\ldots,S$ 个。令 x_s 表示位置 s 上的广告预期能得到的点击次

数。假设广告位置按照各自能得到的点击次数多少排序，因此$x_1 > x_2 > ... > x_S$。

每个登广告的人都有每次点击的价值，这个价值和他能从广告访问者身上得到的期望利润相关。令v_s表示登广告的人认为位置s上的广告的每次点击所值的价值。

每个竞价人（登广告的人）说出自己的报价b_s，可以把它解释为他愿意为位置s支付的钱数。最好的位置（位置 1）由报价最高的人得到，第二好的位置（位置 2）由报价第二高的人得到，以此类推。

登广告人实际支付的钱数由仅次于他的报价的那个报价决定。这是维克里拍卖模型的变种，有时称为**广义第二价格拍卖**（Generalized Second Price auction, GSP）。

在 GSP 中，竞价人 1 支付的价格为每次点击b_2元，竞价人 2 支付的价格为每次点击b_3元，以次类推。为了看清 GSP 的合理之处，我们先想一想如果竞价人按照自己的**报价**支付，那么将会出现什么样的结果？答案是每个竞价人有动机降低自己的报价，只要这个报价能击败仅次于他的报价即可。而在 GSP 下，位置s的竞价人支付位置$(s+1)$竞价人的标价，每个竞价人最终支付的价格是获得该位置的必要最低报价。

由此可知，竞价人在位置s上的利润为$(v_s - b_{s+1})x_s$。这个式子是说竞价人的利润，等于位置s上的广告被点击的价值$v_s x_s$减去竞价人为此支付的成本$b_{s+1}x_s$。

这种拍卖（GSP）的均衡是什么？你可能禁不住从维克里拍卖进行外推，即你可能认为每个竞价人（登广告的人）的报价应该是自己的真实评价。如果拍卖的位置只有一个，这个推测是正确的，但不能进一步外推。

两个竞价人

首先考虑 2 个广告位置和 2 个竞价人的情形。假设报价最高的人得到的点击次数为x_1，他支付的价格为第二高报价b_2。报价第二高的竞价人得到位置 2，他支付的价格是他的保留价格r。

假设你是其中一个竞价人，你认为的价值为v，你的报价为b。如果$b > b_2$，你得到的利润为$(v-b_2)x_1$；如果$b \le b_2$，你得到的利润为$(v-b_2)x_2$。你的期望利润为

$$\textbf{Prob}(\mathrm{b} > \mathrm{b}_2)(v-b_2)x_1 + \left[1-\textbf{Prob}(\mathrm{b} > \mathrm{b}_2)\right](v-r)x_2.$$

整理你的期望利润表达式可得

$$(v-r)x_2 + \textbf{Prob}(\mathrm{b} > \mathrm{b}_2)[v(x_1-x_2) + rx_2 - b_2x_1] \tag{17.1}$$

注意，当（17.1）式方括号内的项为正时（即你的利润为正），你希望$(b > b_2)$的概率越大越好；当它为负时（即你遭受了损失），你希望$(b > b_2)$的概率越小越好。

然而，这不难做到。只要根据上述表达式选择报价即可:

$$bx_1 = v(x_1 - x_2) + rx_2.$$

现在容易验证，当$b > b_2$时，（17.1）式方括号内的项是正的；当$b \leq b_2$时，它是负的。因此，这个报价恰好是你想中标就能中标、想失败就能失败的报价。

注意，这个报价规则是一种占优策略：每个竞价人都希望按照这个表达式报价，不管其他竞价人如何报价。这当然意味着，报价最高的人获得最好的广告位置。

这个报价也容易解释。如果有 2 个竞价人和 2 个广告位置，报价第二高的竞价人总是得到第 2 块广告位置，他支付的钱数为rx_2。竞争是围绕报价最高的人得到的**额外**点击数展开的。评价最高的竞标人能得到这些点击，但是这个竞价人支付的钱数是能击败第二高的竞价人的必要的最低钱数。

在这个拍卖中，我们看到你不想报出每次点击对你的真实价值，但是你的确想让你的报价能反映**额外点击**（incremental clicks）对你的真实价值。

两个以上的竞价人

如果竞价人的数量超过了两个，将会发生什么样的结果？在这种情形下，通常不存在占优策略均衡，但是的确存在着均衡价格。我们分析 3 个广告位置和 3 个竞价人的情形。

广告位置 3 的竞价人支付的是他的保留价格r。均衡时，此竞价人不想得到广告位置 2，因此

$$(v_3 - r)x_3 \geq (v_3 - p_2)x_2\text{，或}$$

$$v_3(x_2 - x_3) \geq p_2x_2 - rx_3.$$

这个不等式是说，如果竞价人偏好位置 3 胜于位置 2，他从位置 2 得到的额外点击的价值必定小于这些额外点击的成本。

这个不等式给出了位置 2 成本的上限：

$$p_2x_2 \leq rv_3 + v_3(x_2 - x_3). \tag{17.2}$$

将同样的逻辑应用于位置 2 的竞价人，可得

$$p_1x_1 \leq p_2x_2 + v_2(x_1 - x_2). \tag{17.3}$$

将不等式（17.2）代入不等式（17.3）可得

$$p_1x_1 \leq rx_3 + v_3(x_2 - x_3) + v_2(x_1 - x_2). \tag{17.4}$$

这 3 块广告位置的拍卖收入为$p_1x_1 + p_2x_2 + p_3x_3$。将（17.2）加上（17.3）再加上位置 3 的收入，可得

$$R = v_2(x_1 - x_2) + 2v_3(x_2 - x_3) + 3rx_3.$$

到目前为止，我们已经分析了 3 个竞价人竞争 3 个广告位置。如果 4 个竞价人竞争 3 个广告位置，将会发生什么样的结果？在这种情形下，保留价格需要替换为第 4 个竞价人的评价。此处的逻辑是第 4 个竞价人愿意购买超过其评价的任何点击，这种情形和标准维克里拍卖是一样的。由此我们得到了收入表达式

$$R' = 3v_4x_3 + 2v_3(x_2 - x_3) + v_2(x_1 - x_2).$$

对于此表达式，我们指出几点。首先，在搜索引擎提供者的广告位置拍卖中，竞价人竞争的是额外的点击：如果你得到位置更好的广告位置，你能得到多少额外的点击量。齐次，点击量差距越大，收入越大。第三，当 $v_4 > r$ 时收入将更大。这只是说竞争倾向于抬升收入。

质量分数（Quality Scores）

在实践中，报价需要乘以**质量分**（quality score），这样就得到了拍卖排序分数。报价和质量分乘积最高的那个竞价人得到第一块（最好的）广告位置，排在第二的竞价人得到第二块广告位置，依次类推。每个竞价人支付的每次点击的价格，是为得到此位置必须支付的最低价。如果令 q_s 表示广告位置 s 的质量，竞价人的排序为 $b_1q_1 > b_2q_2 > b_3q_3 \dots$。

广告位置 1 的竞价人支付的价格正好足够得到该位置，因此 $p_1q_1 = b_2q_2$，或 $p_1 = b_2q_2 / q_1$。（为了打破可能出现的平局，可能需要适当的四舍五入。）

广告质量有几个组成成分。然而，最主要的成分通常是广告得到的历史点击率。这意味着广告排序基本取决于

$$\frac{\text{成本}}{\text{点击数}} \times \frac{\text{点击数}}{\text{观看数}} = \frac{\text{成本}}{\text{观看数}}.$$

因此，得到第一块广告位置的竞价人，是愿意支付每次观看（impression）价格最高而不是每次点击价格最高的人。

稍微想想，你就会明白其中的道理。假设一个竞价人愿意支付每次点击 10 元的价格，但是每天可能只得到一次点击。另外一个竞价人愿意支付每次点击 1 元的价格，但是每天可得到 100 次的点击，哪个竞价人应该得到最好的广告位置？

这种排序方法也能帮助消费者。如果两个竞价人的报价相同，但是消费者点击数量多的那个竞价人的广告将得到更好的位置。消费者对于最有用的广告将会“用鼠标点击投票”。

17.5 拍卖的问题

我们已经知道英式拍卖（或维克里拍卖）能实现帕累托结果。因此人们在分配资源时偏爱使用这样的拍卖方式。事实上，美国联邦通信委员会（FCC）对电视广播频谱拍卖时，基

本上使用的都是英式拍卖的变种。

但是英式拍卖并非完美无缺。它们容易导致串谋（collusion）。拍卖市场混合均衡的例子（详见第 24 章），表明了美国费城的古董商是如何在报价策略中串谋的。

投标人还可以通过各种方法操纵投标结果。在以上的分析中，我们假设一次报价实际上就是投标人的支付**承诺**。然而，有些拍卖允许投标人在投标结果宣布后退出。这样的拍卖制度容易使拍卖结果被操纵。例如 1993 年，澳大利亚政府使用标准的密封拍卖出售卫星电视服务许可证。其中一项许可证的获胜价格为 2.12 亿澳元，这是由 Ucom 报出的价格。当政府宣布 Ucom 公司中标后，Ucom 拒不支付资金，政府只好将许可证授予报价第二高的投标者——还是 Ucom 公司！他们再次违约；四个月之后，经过三番五次地违约，Ucom 最终为该许可证支付了 1.17 亿澳元，这比他们最初的报价节省了 0.95 亿澳元！许可证最终由报价最高的人按第二高的报价支付，基本实现了招标的目的，但是这个设计糟糕的拍卖方案的代价也是沉重的，它使得澳大利亚在拖延了至少一年后才引入付费电视[一]。

例子：让墙报价

操纵拍卖的一种常见方法是，拍卖品卖方自己进行虚假竞标，这种做法称为“让墙报价”（Taking Bids off the wall）[二]。这样的操纵也存在于在线拍卖中，尽管这种拍卖不涉及到墙。

根据近期报道[三]，某个纽约珠宝商在线销售大量的钻石、黄金和白金首饰。尽管这些商品在 eBay 上以无底价（保留价格）销售，但是这个商人将报价单分发给他的雇员，让他们填写报价，目的是抬升最终销售价格。根据此法律诉讼案可知，这些雇员在一年之内报价 232,000 次，将销售价格平均抬升了 20%。

当面对这些铁证时，为了解决诉讼，该珠宝商同意支付 400,000 美元的罚款。

17.6 胜利者的诅咒

现在我们开始分析共同价值拍卖（common-value auctions），在这种拍卖方式中，被拍卖的商品对于所有投标人来说，具有**相同**的价值。为了强调这个特征，我们将投标人 i 的（估计）价值写成 $v+\varepsilon_i$ 的形式，其中：v 是真实的共同价值；ε_i 是“误差项”，它和投标人 i 的

[一] 这个事件的细节以及美国波段拍卖设计中如何吸取教训从而加以改进的细节，请参见 John McMillan, "Selling Spectrum Rights," Journal of Economic Perspectives, 8(3), 145-152。这篇文章也介绍了本文前面提到的新西兰的拍卖案例。

[二] 这种虚假报价的叫法有很多，比如还有“让吊灯或椽报价（Chandelier or Rafter Bidding；taking bids from the chandelier or rafter）等。这是拍卖师代表拍卖品销售方的利益在关键时刻虚假报价，以造成需求旺盛的假像，目的是抬升报价从而抬升成交价格，这种做法一般是违法的。拍卖师报出这样的虚假价格时，通常看着墙或吊灯，观众无法知道到底有没有以及是哪个竞标人喊出的价格，所以戏称“让墙或吊灯报价”。译者注。

[三] Barnaby J. Feder, “Jeweler to Pay $400,000 in Online Auction Fraud Settlement,” New York Times, June 9, 2007。

估计值有关。

我们在上述框架内分析某密封拍卖。投标人 i 的报价应为多少？为了让你对这个问题有些直觉认识，我们看看如果每个投标人的报价等于各自的估计价值，结果将如何。在这种情形下，评价最高的人，即 ε_i 最大（$\varepsilon_{\max}$）的人得到了商品。但是，只要 $\varepsilon_{\max} > 0$，此人支付的价格就会高于商品的真实价值 v。这就是所谓的**胜利者的诅咒**（Winner's Curse）。如果你最终获胜，这是因为你过高估计了商品的价值。换句话说，你胜出只是因为你过于乐观！

在类似共同价值拍卖中，最优的报价策略是报价要低于你的估计价值，而且投标人越多，你的报价也应该越低。想想下面的情形：如果你在五个投标人中报价最高，你可能过于乐观了，但是如果你在二十个投标人报价最高，你一定是**超级**乐观。投标人越多，你就应该将拍卖商品的价值估计的越小。

1996 年 5 月，美国联邦通信委员会（FCC）拍卖个人通信使用的波段时，胜利者的诅咒出现了。此次拍卖中，报价最高的投标人是 NextWave 个人通信公司，它报价 42 亿美元购买 63 个许可证，将所有许可证都纳入囊中。然而，到了 1998 年 1 月，该公司发现无法支付账单时，只好申请破产法第 11 章的保护，以重组业务争取再度盈利。

17.7 稳定的婚姻问题

双边匹配模型（two-sided matching models）的例子有很多，在这样的例子中，消费者彼此匹配到一起。男女与婚姻介绍所（或媒人）的匹配、学生与大学的匹配、会员与姐妹会（sororities）的匹配[一]、实习医生与医院的匹配等都是这样的例子。

这种匹配问题有好的算法吗？“稳定的”结果总会存在吗？下面我们分析一个简单的匹配机制，这个机制（在精确定义的意义上）是稳定的[二]。

我们以舞伴匹配为例说明。有没有舞伴匹配的好方法？一种让人喜欢的标准是找到能产生“稳定”匹配的方法。这里的“稳定”是指没有哪一对人会互相更喜欢另一对中的人。换一种说法，如果某个男人认为另外一个女人比他当前的舞伴好，但该女不会想和他配对——她更喜欢当前的舞伴。

稳定的匹配总是存在的吗？如果是，怎样才能找到它？也许受连续剧和浪漫小说的影响，你可能认为答案为否。然而，的确总是存在稳定的匹配，而且也比较容易构造出。

最著名的算法是**延迟接受算法**（deferred acceptance algorithm），该算法的步骤如下[三]：

[一] 兄弟会和姐妹会（Fraternities and sororities），一般专指北美院校中的面对大学生的一种社团组织，入会以后不仅可以参加各种活动，还可能得到比较好的住房，比如一个兄弟会或姐妹会占有一个独立的洋楼，条件很好，但里面卧室有限，只有会员才能入住。译者注。

[二] 精确定义（precising definition）通过对常规意义强加限制来减少一个词项的模糊性。比如，“稳定”这个词，字典上的定义就是所谓常规的定义，但这里将它进一步细化，明确界定了它的意思。译者注。

[三] Gale, David, and Lloyd Shapley [1962], “College Admissions and the Stability of Marriage,” American Mathematical Monthly, 69, 9-15.

第 1 步：每个男人向他最喜欢的女人提出跳舞邀请。

第 2 步：每个女人将她得到的跳舞邀请（男人姓名）记录下来。

第 3 步：当所有男人都向他最喜欢的女人提出跳舞邀请之后，每个女人在她记录下来的男人名单中选择她最喜欢的，然后礼貌地拒绝名单中的所有其他男人。

第 4 步：被拒绝的男人按照他的女人排序名单中选择下一个女人，提出跳舞邀请。

第 5 步：重复第 2-5 步，直至所有女人都得到了跳舞邀请。

这个算法总能产生稳定的匹配结果。反证一下。假设某个男人偏爱某个女人胜于自己当前舞伴，那么他必然已经在邀请当前舞伴之前向这个女人提出过跳舞邀请。如果这个女人偏爱该男人胜于她当前的舞伴，她应该早就拒绝了她当前的舞伴而接受该男人。

站在男人的角度上，可以证明这个算法产生了最优稳定匹配结果，因为每个男人偏爱这个结果胜于任何其他稳定的匹配。当然，如果我们把问题中的男人和女人的角色互换一下，我们就得到对于女人来说的最优稳定匹配结果。

尽管我们这个例子稍微有些无聊，但是延迟接受算法广泛用于一些重要的领域，例如波士顿和纽约用此算法把学生和学校进行匹配。再比如，全国范围内的住院病人和医院的匹配、甚至器官捐献者和接受者的匹配等都是用的这个算法。

17.8 机制设计

我们在本章讨论的拍卖和双边匹配模型都是**经济机制**（economic mechanisms）的例子。经济机制的思想是定义能产生某种合意结果的“博弈”和“市场”。

例如，有个人想设计销售某幅油画的机制。很自然地，我们就能想到拍卖。但是，即使拍卖这种机制也有很多选择：目标应该是效率最大化吗（即，保证对油画评价最高的人能得到它）？或者，目标应该是即使油画存在卖不出去的风险，也要使油画卖方的期望收入最大化吗？

我们已经知道，拍卖类型有多种，每种都有优缺点。在某特定环境下，哪种拍卖是最好大的？

机制设计正好和**博弈论**（详见第 28 章）相反。在博弈问题中，博弈规则是给定的，我们要做的是在这些规则下找到可能的结果。在机制设计问题中，要达到的结果是给定的，我们要做的是设计能达到这个结果的博弈[一]。

机制设计问题不限于拍卖或匹配问题。它也包括**投票机制**（voting mechanisms）和**公**

[一] 2007 年的诺贝尔经济奖得主是下列三位经济学家：Leo Hurwicz, Roger Myerson 和 Eric Maskin，他们在经济机制设计问题上做出了杰出贡献。

曹乾（东南大学 caoqianseu@163.com）

共物品（详见第 36 章）的机制或外部性（详见第 33 章）机制。

在一般性机制中，我们考虑的情形是一些参与人（例如消费者或企业），每个参与人都有自己的私人信息。比如在拍卖中，这样的私人信息可能是每个人对拍卖品的评价。在涉及到企业的问题中，私人信息也许是它们各自的成本函数。

参与人将自己的部分私人信息告诉“中央”(center)，比如在拍卖中，这个中央就是拍卖师。中央分析这些信息并报出结果：谁得到了拍卖品；各个企业应该生产的产量；各个参与方的支付额或得到的报酬额，等等。

主要的设计决策是：1）应该向中央递交什么样的信息；2）中央应该使用什么规则来确定结果。问题的约束条件通常有两个：一是某种资源约束（例如拍卖品只有一件）；二是每个参与人都为自己的利益行事的约束。后面这个约束称为**激励相容约束**（incentive compatibility constraint）。

可能还存在着其他的约束。例如，我们可能希望人们自愿参与接受我们设计的机制，这就要求该机制提供给他们的收益不应小于不接受该机制时的收益。为简单起见，我们暂时忽略这种约束（详细内容将在第 37 章介绍）。

为了体验一下如何进行机制设计，我们考虑下面这样的简单问题：将某个不可分割的商品分配给两个人。令 $(x_1, x_2) = (1, 0)$ 表示参与人 1 得到了商品，$(x_1, x_2) = (0, 1)$ 表示参与人 2 得到了商品。令 p 表示该商品的价格，v_i 表示参与人 i 对该商品的评价。

价格 p 的约束条件是什么？假设参与人 1 对商品的评价最高，那么他向中央传递的信息应该是：这样的信息能使得他得到的收益，**不小于**若他传递的信息与参与人 2 相同时他得到的收益（参与人 2 的收益为零）。这就是说，

$$v_1 - p \geq 0.$$

出于同样的原因，参与人 2 从他向中央传递信息中的收益，必定不小于若他传递的信息与参与人 1 相同时他得到的收益（这导致参与人得到了商品）。这意味着，

$$0 \geq v_2 - p.$$

将这两个条件放在一起，可得 $v_1 \geq p \geq v_2$，这是说中央索要的价格必定介于最高评价和第二高评价之间。

为了确定中央必须索要的价格，我们必须考虑它的目标和信息。如果中央相信 v_2 无限接近于 v_1，而且它总是希望报价最高的参与人得到商品，那么中央必须将价格设定在 v_2 的水平上，即 $p = v_2$。

这就是前面介绍过的**维克里拍卖**（Vickrey auction），在这种拍卖中，每个竞标人报出自己的价格，商品由报价最高的人得到，但他支付的价格不是自己的报价（最高报价），而是第二高的报价。显然，对于我们此例中的问题来说，这种拍卖是个很好的机制。

总结

1.拍卖这种销售东西的方式已有几千年的历史。

2.如果每个投标人对拍卖商品的评价互不影响，则该拍卖为个人价值拍卖类型。如果商品的价值在本质上对每个人是相等的，则该拍卖为共同价值拍卖类型。

3.常见的拍卖方式有英国式拍卖、荷兰式拍卖、密封拍卖和维克里拍卖。

4.英国式拍卖和维克里拍卖都有个让人喜欢的特点，即它们的结果都是帕累托有效率的。

5.如果拍卖的目的是利润最大化，则必须设定保留价格。

6.尽管作为市场机制，拍卖优势明显，但是它们容易遭受串谋和其他策略行为的影响。

复习题

1.如果某拍卖是向收藏者出售一批古董被子。这是个人价值拍卖还是共同价值拍卖？

2.假设某拍卖中只有两个投标人，他们对拍卖商品的评价分别为 8 元和 10 元，假设最小竞价增量为 1 元。如果这个拍卖使用英国式拍卖而且要求利润最大化，那么卖主应将保留价格设定为多大？

3.假设我们有两本中级微观经济学的教科书，对三个（好学的）学生出售。我们如何使用密封拍卖的方式使得评价最高的两个学生最终得到教材?

4.分析课文中 Ucom 公司的例子。这个拍卖是有效率的吗？它做到了利润最大化吗？

5.一位博弈论学者将罐子塞满了一分钱硬币，在开学第一次课上他使用英国式拍卖将这罐硬币出售。这是个人价值拍卖还是共同价值拍卖？你认为投标胜出者最终能否盈利？

复习题答案

1.如果某拍卖是向收藏者出售一批古董被子。这是个人价值拍卖还是共同价值拍卖？

【复习内容】个人价值拍卖；共同价值拍卖

个人价值拍卖和共同价值拍卖是按商品的性质分类的。

在个人价值拍卖中，每个竞买人对拍卖的商品的评价不同，具体评价多少取决于个人的偏好。而且竞买人的估计彼此之间具有独立性。

在共同价值拍卖中，拍卖的商品对每个竞买人来说价值大致相同，尽管他们的具体评价存在着差异。

在图形上，如果我们计算出投标人对拍卖商品评价的均值，一般来说，个人价值拍卖中个人评价与均值之间的离散性很大；而共同价值拍卖中，个人估计与均值之间的离散性相对较小。

【参考答案】

非常有可能是个人价值拍卖。

个人价值拍卖和共同价值拍卖的区分有时比较微妙，一般来说，如果商品的同质性很强而且市场定价比较成熟（例如石油、粮食、金属等），则越有可能是共同价值拍卖；反之，则越有可能是个人价值拍卖，例如古董一般是个人价值拍卖。

由于这一批古董被子很可能没有成熟的市场定价，因此对被子的评价较多地取决于个人的偏好，而且某人的评价和其他人的评价之间一般不是互相影响的，所以非常有可能是个人价值拍卖。

2.假设某拍卖中只有两个投标人，他们对拍卖商品的评价均为：要么 8 元，要么 10 元，而且这两种情形概率相等。假设最小竞价增量为 1 元。如果这个拍卖使用英国式拍卖而且要求利润最大化，那么卖主应将保留价格设定为多大？

【复习内容】英国式拍卖；拍卖的利润最大化目标；

英国式拍卖：拍卖人从保留价格开始拍卖，这个价格是委托人（卖主）卖掉商品所愿意接受的最低价格。不同投标人或称竞买人连续报出更高的价格；一般来说，每次报价都必须比前一个报价至少高出一个既定的最小数额，这个数额称为竞价增量。当竞买人不愿意进一步提高报价时，拍卖的商品由报价最高的人获得。

拍卖的利润最大化目标：设计出的拍卖要能为卖主带来最高的期望利润。

【参考答案】

卖主不知道投标人对商品的估计，卖主如何设计保留价格才能使利润最大。在这里由于题目未告知卖主获得该商品的成本，不妨假设为常数 k 。这样，保留价格可能与这个成本有关，但在拍卖的情形中，宁愿把保留价格看作为卖主定价的一种策略。

因此，卖主的利润=期望销售收入 $-k$ 。这样问题可以转化为期望销售收入最大。

根据题意可知，**这两个投标人的个人价值（对商品的评价，注意不是报价）组合可能有以下四种：**（8,8）；（8，10）；（10,8）；（10,10）。

因为英式拍卖是从卖主设定的保留价格起拍的，但卖主不知道投标人的保留价格。

为了更清楚地看清英式拍卖的本质，我们假设卖主能够摸石头过河，即这两个投标人可以不厌其烦地陪着卖主（拍卖人）玩。

因此不妨假设卖主一开始玩得很猛，把保留价格设定为 50 元，即从 50 元起拍，但由于两个投标人的保留价格最高为 10 元（注意这对卖主来说是黑箱），因此必然无人报价，卖主的销售收入为 0。因为既然允许卖主试错，假设卖主再试一把，把保留价格定为 11 元，仍人无人报价，销售收入为 0。

现在卖主把保留价格设定为 10 元，从 10 元开始起拍。对照以上四种报价情形，可以知道，四种情形的胜报价分别为 0（无人胜出），10 元，10 元，10 元。因此，此时卖主的期望销售收入=$\frac{1}{4}(0+10+10+10)=7.5$元。

如果卖主把保留价格设定为 8 元以下（含 8 元）。对照以上四种报价情形，可以知道，这四种情形的报价分别为 8 元，9 元，9 元和 10 元。此时卖主期望销售收入=$\frac{1}{4}(8+9+9+10)=9$元。

如果卖主把保留价格设定为 8 元以上 10 元以下，对照以上四种报价情形可知，第一种情形下的销售收入为 0，虽然后面几种情形的销售收入有所增长，但不足以弥补第一种情形下的损失。

由此可以知道，如果卖主把保留价格设定为 8 元以下（含 8 元），则期望销售收入最大，从而期望利润最大。

最后多说一句，由于英式拍卖面对面竞争比较激烈，因此在实务中我们通常可以看到，对于一般价值不大的商品，一般不设底价（保留价格）。随着竞价的进行，自然是报价最高者得到商品。

注：我也参看了教材课后的参考答案，发现范里安提供的答案并不准确。除非把期望利润定义为“期望利润=期望销售收入-卖主的保留价格”，范式答案才成立。

3.假设我们有两本中级微观经济学的教科书，对三个（好学的）学生出售。我们如何使用密封拍卖的方式使得评价最高的两个学生最终得到教材?

【复习内容】密封拍卖；维多里拍卖的特征

密封拍卖（sealed-bid auction）：在这种拍卖方式中，每个投标人将自己的报价写在纸条上并用信封密封起来。拍卖人将这些信封收集然后开封，报价最高的人得到了商品，然后他要如数交钱。如果拍卖的商品有保留价格，而且所有的报价都低于保留价格，则卖主将商品收回。

集邮者拍卖（Philatelist auction）或者维克里拍卖（Vickrey auction）：维克里拍卖的方式类似密封拍卖，但有一个重要区别：商品由报价最高的竞价人获得，但他只需要按第二高的

报价支付。换句话说，报价最高的投标人得到了拍卖商品，但是他不需要按照他自身的报价支付，而是按照报价第二高的人的报价支付。

维克里拍卖的特征：如实报价（即各自的报价等于各自的评价）符合每个投保人的利益。

为了看清这一点，我们分析一个只有两个投标人的特殊情形。这两人的对商品的评价分别为 v_1 和 v_2，他们在纸条上写下的报价分别为 b_1 和 b_2。投标人 1 的期望收益（expected payoff）为：

$$\Pr ob(b_1 \geq b_2)[v_1 - b_2],$$

其中“Prob”表示概率（probalility）。

上式中的第一项是投标人报价在所有投标人的报价中为最高报价的概率；第二项是如果投标人中标他享有的消费者剩余。（如果 $b_1 < b_2$，则投标人 1 得到的消费者剩余为零，因此无需考虑含有 $\Pr ob(b_1 < b_2)$ 的项）。

假设 $v_1 > b_2$。因此投标人 1 希望使自己的中标概率尽可能大，所以他会将报价设定为等于自己的评价即 $b_1 = v_1$。另一方面，假设 $v_1 < b_2$，则他希望自己中标概率尽可能小，所以他会使得 $b_1 = v_1$。在两种情形下，投标人 1 的最优报价策略就是令报价等于自己的真实评价！诚实是最好的策略…至少在维克里拍卖中是这样的！

维克里拍卖的一个有趣特征是，它在本质上可以实现与英式拍卖相同的结果，但与英式拍卖不同的是，它不存在投标人的相互影响。维克里不仅证明了说真话是最优的策略，而且证明了集邮者拍卖和英式拍卖是等价的。

【参考答案】

根据题目的要求“评价最高的两个学生获得教材”可知，应该设计成维多里拍卖的形式。拍卖方案设计：三个学生各自独立写出自己的报价（注意未必是自己的评价），密封上交，并且事先宣告——报价位于前两名的学生得到教材，但他们只要按第三高的价格支付。

由维克里拍卖的特征可知，这种拍卖方式能使学生的报价逼近自己的评价，而且能保证评价最高的人得到教材。

当然为了防止学生串谋，卖主可以在拍卖方案中再设定保留价格。如果不设定，我们假设三个学生串谋：其中一名学生报价非常低，比如为 0。则前两名学生都以 0 价格得到了教材。

4.分析课文中 Ucom 公司的例子。这个拍卖是有效率的吗？它做到了利润最大化吗？

【复习内容】拍卖的设计；拍卖设计的帕累托效率目标

拍卖设计的帕累托效率要求商品卖给**评价**最高的人。为了说明这一点，假设投标人 1 对商品的估计（v_1）最高，投标人 2 对商品的评价（v_2）略低。如果投标人 2 得到了商品，则存在让这两人的状况都变化的方法：投标人 2 以位于 v_1 和 v_2 之间的价格 p，将商品卖给

投标人 1。由此可见，如果不是评价最高的人得到商品则不可能是帕累托有效率的。

【参考答案】

由课文这句话“许可证最终由报价最高的人按第二高的报价支付”，根据拍卖设计的帕累托效率目标——商品卖给评价最高的人，可知它是帕累托有效率的。

由课文这句话“但是这个设计糟糕的拍卖方案的代价也是沉重的，它使得澳大利亚在拖延了至少一年后才引入付费电视。”可知，如果考虑时间成本，它又是帕累托无效率的。因为设计良好的维克里拍卖或者英国式拍卖在很多时间内就能达到这个结果。

利润最大化的分析与上面类似，如果不考虑时间成本，是利润最大化；如果考虑时间成本则不是利润最大化。

5.一位博弈论学者将罐子塞满了一分钱硬币，在开学第一次课上他使用英国式拍卖将这罐硬币出售。这是个人价值拍卖还是共同价值拍卖？你认为投标胜出者最终能否盈利？

【复习内容】个人价值拍卖和共同价值拍卖；胜利者的诅咒

个人价值拍卖和共同价值拍卖的区别请见本章复习题第 1 题。

胜利者的诅咒：

共同价值拍卖方式中，被拍卖的商品对于所有投标人来说，具有相同的价值。为了强调这个特征，我们将投标人i的（估计）价值写成$v+\varepsilon_i$的形式，其中：v是真实的共同价值；ε_i是“误差项”，它和投标人i的估计值有关。

我们在上述框架内分析某密封拍卖。投标人i的报价应为多少？为了让你对这个问题有些直觉认识，我们看看如果每个投标人的报价等于各自的估计价值，结果将如何。在这种情形下，评价最高的人，即ε_i最大（$\varepsilon_{\max}$）的人得到了商品。但是，只要$\varepsilon_{\max}>0$，此人支付的价格就会高于商品的真实价值v。这就是所谓的胜利者的诅咒（Winner’s Curse）。如果你最终获胜，这是因为你过高估计了商品的价值。换句话说，你胜出只是因为你过于乐观。

【参考答案】

这很可能是一个共同价值拍卖，在共同价值拍卖中，拍卖的商品（此处为一罐一分钱的硬币）对每个竞买人来说价值大致相同，尽管他们的具体评价存在着差异。

如果每个投标人的报价等于各自的估计价值，结果将如何?在这种情形下，对这罐硬币评价最高的人，即ε_i最大（$\varepsilon_{\max}$）的人得到了它。但是，只要$\varepsilon_{\max}>0$，此人支付的价格就会高于商品的真实价值v。这就是胜利者的诅咒。因此，投标胜出者最终通常不能盈利。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

18.技术（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

18 技术

从本章开始，我们将学习企业行为。首先要做的事情是分析企业行为面临的约束。当企业在做生产决策时，它面对很多约束。这些约束来自它的消费者、它的竞争者以及自然。在本章，我们将考虑最后一种约束：自然条件约束。自然条件约束是说使用要素生产产品的可行方法并不多：只有少数类型的技术选择是可行的。这里我们将分析经济学家如何描述这些技术约束。

如果你已经理解了消费者理论，生产理论将会很简单，因为我们使用的分析工具是相同的。事实上，生产理论比消费理论更简单，因为生产过程的产出通常是可以观测到的，而消费的“产出”——效用，是不可直接观测到的。

18.1 投入和产出

生产中的投入称为**生产要素**（factors of production）。人们通常把生产要素大致分为几类：土地，劳动，资本和原材料。劳动、土地和原材料的意思很明显，但资本却可能是个新概念。**资本商品**（capital goods）是生产过程的投入品，它们本身就是商品，因为你在生产时需要在市场上购买这些投入品。资本品基本上是各种各样的机器：拖拉机，建筑物，计算机，等等。

有时人们把创业或企业运行所需要的资金也称为资本。为了不至于混淆，我们总是将这样的资本称为**金融资本**（financial capital），而把前面介绍的投入品称为资本商品或者**实物资本**（physical capital）。

我们通常用**流量**（flow）单位衡量投入和产出：每周多少劳动和每周多少机器运转时间，每周生产多少单位产品。

我们没必要经常使用投入和产出的上述分类方法。事实上，在研究技术问题时，大多数时候我们不会用到投入和产出的类别，只用到投入和产出的数量。

18.2 描述技术约束

自然条件对企业造成了**技术约束**(technological constraints)：在生产既定数量的产出时，只有某些投入组合可行，企业必须选择技术可行的生产方案。

描述可行生产方案最简单的方法是把这些方案列举出来。也就是说我们可以将技术可行的投入和产出的所有组合列举出来。技术可行的投入和产出的所有组合构成的集合称为**生产集**(production set)。

曹乾（东南大学 caoqianseu@163.com）

例如，假设我们只有一种投入和一种产出，分别以 x 和 y 衡量。则生产集的形状可能具有图 18.1 的形状。点 (x, y) 在生产集内的全部意思是说如果你有 x 单位的投入，则生产 y 单位的产出在技术上是可行的。生产集表明了某企业面对的**可能**技术选择。

只要生产要素不是免费的（需要花钱购买），则对于既定的投入水平，我们自然可以只研究**最大的可能产出**（maximum possible output）。这就是生产集的边界，如图 18.1 所示。描述这个生产集边界的函数称为**生产函数**（production function）。它衡量对于既定数量的投入你可以得到的最大可能产出。

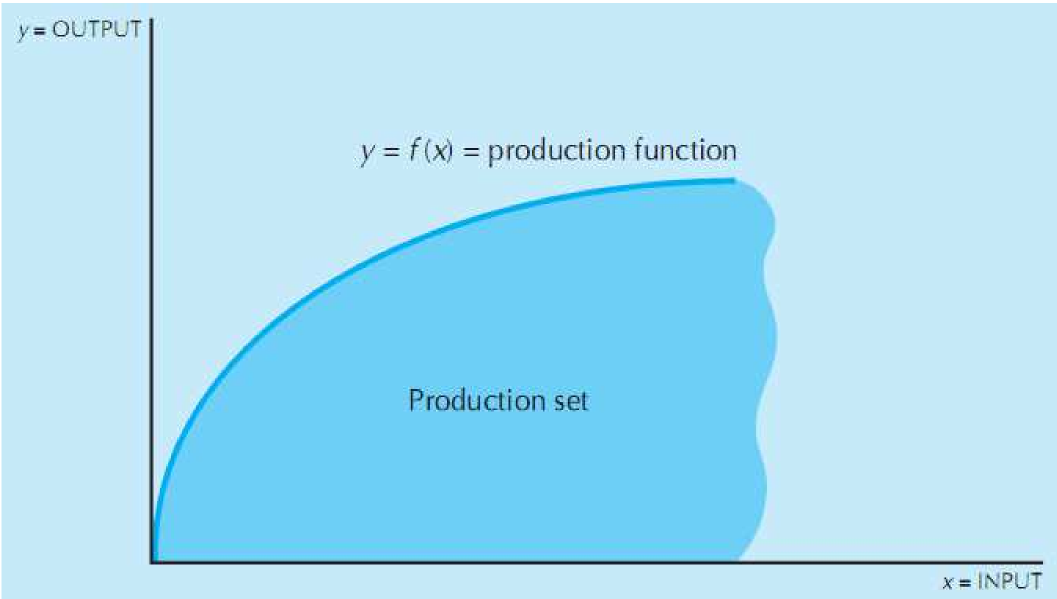


图 18.1：生产集。上图是生产集的一种可能形状。

当然，生产函数的概念完全可以应用于多种投入要素的情形。例如，如果我们考虑两种投入的情形，生产函数 $f(x_1, x_2)$ 衡量如果我们投入 x_1 单位要素 1 和 x_2 单位要素 2 所能得到的最大产出数量 y 。

在只有两种投入的情形下，通常用一种便利的方法描述生产关系，这就是**等产量线**（isoquant）。一条等产量线就是一个集合，这个集合由恰好足够生产给定数量产出的投入 1 和 2 的所有可能组合构成。

等产量线和无差异曲线类似。我们在前面已经知道，一条无差异曲线描述的是恰好足以产生某既定效用水平的不同消费束。但无差异曲线和等产量线有个重要区别。等产量线是用要素投入生产出的产量标记的，而不是以效用水平标记的。因此，等产量线的标记取决于技术，这种标记不具有效用标记的那种任意性。

18.3 技术的例子

既然我们已经熟悉了无差异曲线，因此很容易理解等产量线是如何运作的。我们先看看一些技术的例子和它们的等产量线。

固定比例

假设我们要在地上挖洞，挖一个洞的唯一方法是使用一个人和一把铲子。多余的铲子无用，多余的人也无用。因此所挖的洞数取决于人数和铲子数中最小的那个。我们将该类型的生产函数写为 $f(x_1, x_2) = \min\{x_1, x_2\}$。

等产量的形状如图 18.2 所示。注意这些等产量线和消费者理论的完全互补情形类似。

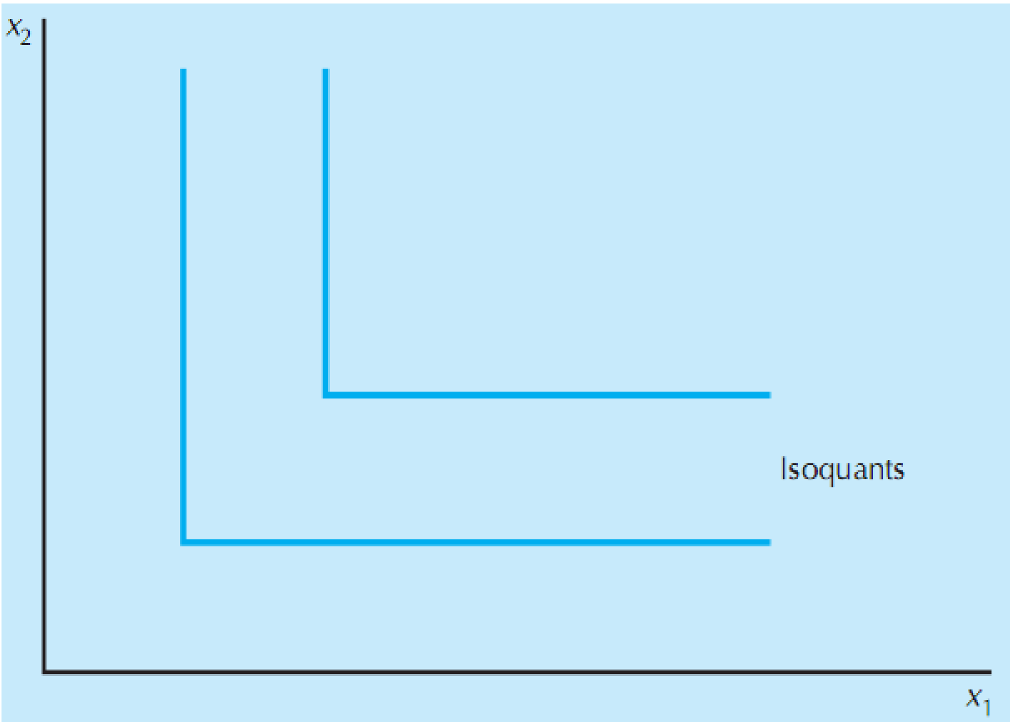


图 18.3：固定比例。固定比例情形下的等产量线。

完全替代

假设现在我们"生产"家庭作业，投入为红铅笔和蓝铅笔。生产出的家庭作业数量仅取决于铅笔的总数，因此我们将生产函数写为 $f(x_1, x_2) = x_1 + x_2$。这种情形下的等产量线类

曹乾（东南大学 caoqianseu@163.com）

似于消费者理论的完全替代，如图 18.3 所示。

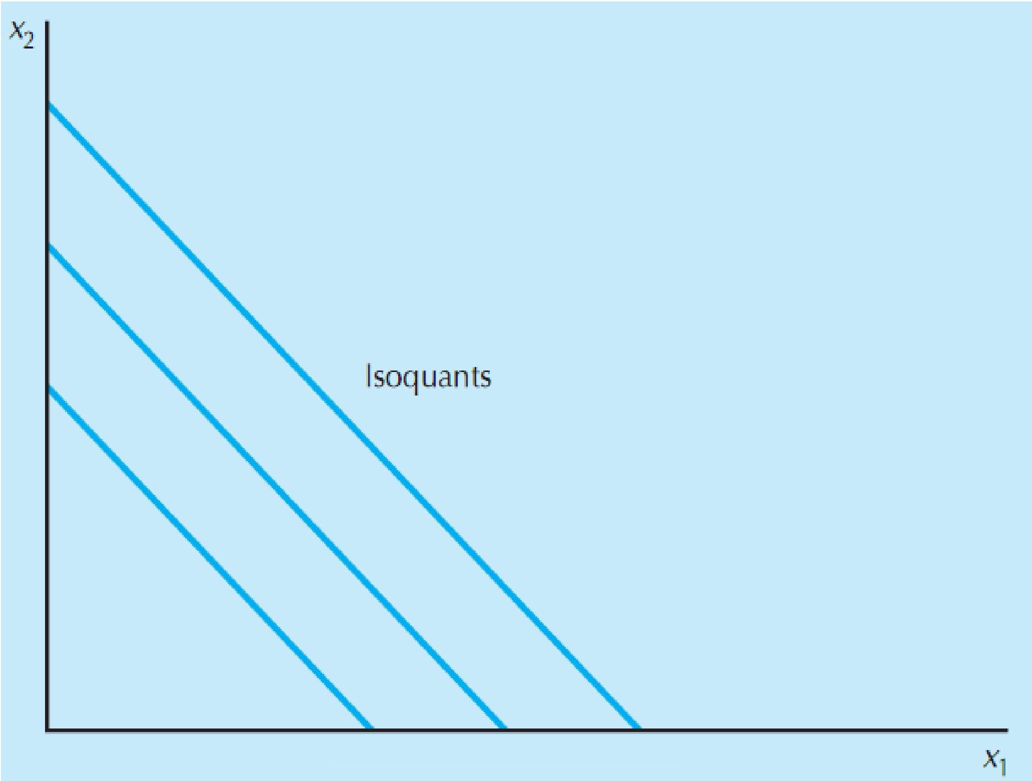


图 18.4：完全替代。完全替代情形下的等产量线。

柯布-道格拉斯

如果生产函数的形式为 $f(x_1,x_2)=Ax_1^a x_2^b$，那么我们称它是一个柯布-道格拉斯生产函数。它和我们以前研究过的柯布-道格拉斯偏好的形式是一样的。在效用函数中，数值的大小并不重要，因此我们将 A 设定为 1（即 A=1），而且通常设定 $a+b=1$。但是生产函数的数值大小非常重要，因此我们必须允许这些参数取任意数值。大致来说，参数 A 衡量生产规模：若每种投入各使用一单位将能得到多少产出。参数 a 和 b 衡量投入变化时产量如何变化。稍后我们将详细分析它们对产量的影响。在某些例子中，我们会令 A=1，目的在于简化计算。

柯布-道格拉斯等产量线和柯布-道格拉斯无差异曲线一样，它们的曲线形状是良好性状的；正如效用函数的情形一样，柯布-道格拉斯生产函数是良好性状等产量线的最简单例子。

18.4 技术的性质

和消费者情形类似，我们通常需要对技术的性质作出一些假设。首先，我们一般假设技术是**单调的**（monotonic）:如果你至少增加一种投入的数量，则产量至少和原来一样大。这

个性质有时称为**自由取舍**（free disposal）:如果企业追加要素投入不会增加额外的成本，则额外的要素投入不会引起企业的损失。

第二，我们通常假设技术是**凸的**（convex）。这意味着生产 y 单位的产量若有两种方法——(x_1, x_2) 和 (z_1, z_2)，则它们的加权平均也能生产**至少** y 单位的产量。

认为技术为凸的一种理由如下。假设你找到一种方法，这种方法能用 a_1 单位要素 1 和 a_2 单位要素 2 生产 1 单位的产量。假设你还有一种方法，这种法能用 b_1 单位要素 1 和 b_2 单位要素 2 生产 1 单位的产量。这两种生产方法称为**生产技术**（production techniques）。

进一步地，我们假设你可以按任意比例增加产量，因此 $(100a_1, 100a_2)$ 和 $(100b_1, 100b_2)$ 都能生产出 100 单位的产量。但是现在要注意，如果你有 $25a_1 + 75b_1$ 单位的要素 1 和 $25a_2 + 75b_2$ 的要素 2，你也能生产出 100 单位的产量：使用“a”技术生产 25 单位，使用“b”技术生产 75 单位。

上述情形可用图 18.4 表示。通过选择这两种技术的不同组合，你可以用很多种方法生产既定的产量。特别地，$(100a_1, 100a_2)$ 和 $(100b_1, 100b_2)$ 这两点连线之间的任何一个投入组合均可以生产出 100 单位的产量。

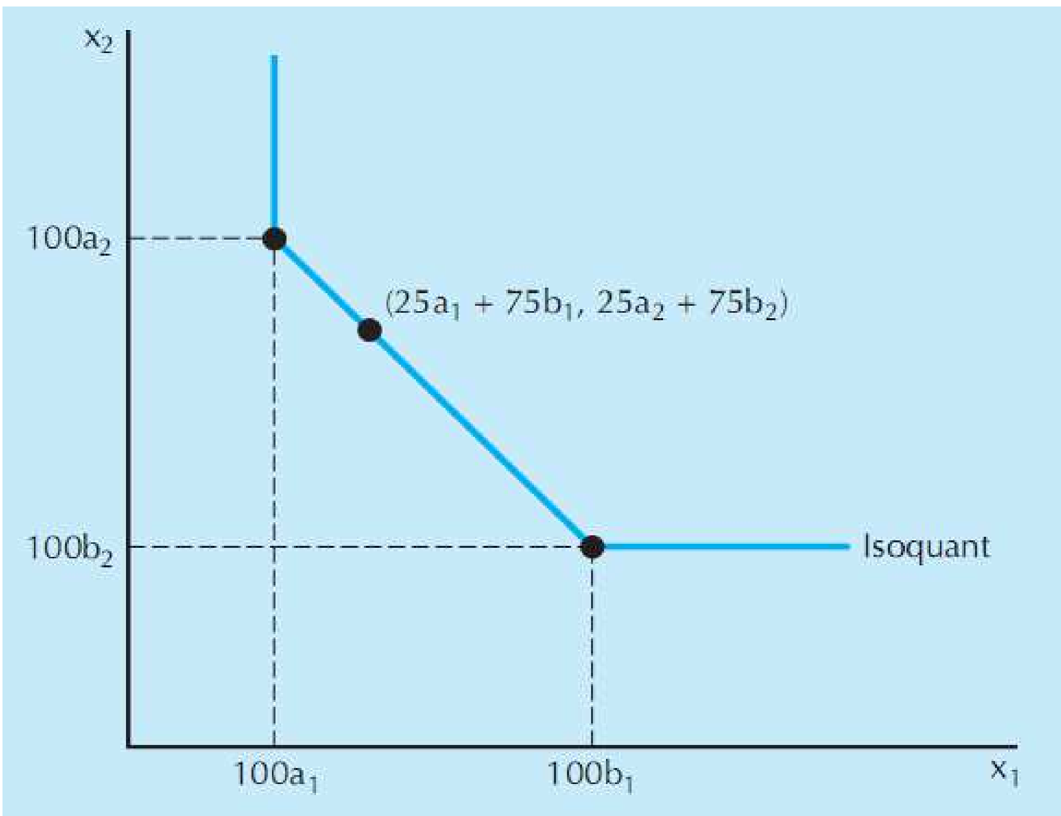


图 18.4：凸的技术。如果两种生产技术是互相独立的，则生产方案的加权平均也是可行的。因此等产量线的形状是凸的。

这种类型的技术具有下列特点：生产过程可以按任意比例增加和减少，而且不同生产过程之间彼此独立。在这种情形下，假设技术为凸是很自然的。

18.5 边际产量

假设我们在某点(x_1,x_2)生产，现在我们考虑稍微增加一些要素 1 的投入，但维持要素 2 的投入水平x_2不变。此时，每增加一单位要素 1，产量将增加多少?我们必须计算要素 1 投入量变动一单位导致的产量变化：

$$\frac{\Delta y}{\Delta x_1}=\frac{f(x_1+\Delta x_1,x_2)-f(x_1,x_2)}{\Delta x_1}.$$

我们将其称为**要素 1 的边际产量**（marginal product of factor 1）。要素 2 的边际产量可以类似地定义，我们用$MP_1(x_1,x_2)$和$MP_2(x_1,x_2)$分别表示要素 1 和 2 的边际产量。

有时，我们也不那么严格地使用边际产量的概念，认为它描述的是额外增加“一单位”要素 1 得到的额外产量。只要这“一单位”相对于要素 1 的使用量来说非常小，则这种观点没什么问题。但是，我们应该记住边际产量是一个**比率**：产量变动量与某种要素投入的变动量的比值。

边际产量的概念和消费者理论中的边际效用的概念类似，但有一个重要的区别：效用是序数性质的。这里我们分析的是实物产量：某要素的边际产量是一个具体的数值，这个数值原则上是可以观测到的。

18.6 技术替代率

假设我们在某点(x_1,x_2)生产，现在我们考虑稍微放弃一些要素 1 的投入（Δx_1），同时增加要素 2 的使用量（Δx_2），使得恰好还能生产相同的产量y。在这种情形下，Δx_2为多大？这正好是等产量线在点(x_1,x_2)的斜率；我们将其称为**技术替代率**（technical rate of substitution，TRS），记为$TRS(x_1,x_2)$。

技术替代率衡量两种要素之间的替代关系。它衡量企业为了维持产量不变，必须用一种要素替代另外一种要素的比率。

为了推导 TRS 的表达式，我们可以使用推导无差异曲线斜率的相同思想。考虑维持产量不变的要素 1 和 2 使用量的变化，可得：

$$\Delta y=MP_1(x_1,x_2)\Delta x_1+MP_2(x_1,x_2)\Delta x_2=0$$

由此可解得

$$TRS(x_1,x_2)=\frac{\Delta x_2}{\Delta x_1}=-\frac{MP_1(x_1,x_2)}{MP_2(x_1,x_2)}.$$

注意，它与边际替代率的定义非常类似。

18.7 边际产量递减

假设我们有一定数量的要素 1 和要素 2，我们考虑增加一点要素 1 但维持要素 2 的投入数量不变。那么，要素 1 的边际产量将怎样变化?

只要生产技术是单调的，则随着要素 1 的使用量增加，总产量也是不断增加的。但是我们可以预期总产量增加的速率是递减的。我们以一个具体的例子说明，即种地的例子。

使用一个人和一亩地可以生产 100 单位的玉米。现在我们增加一个人但维持土地的面积不变，我们可能得到 200 单位的玉米，在这种情形下，额外增加这个工人的边际产量为 100 单位。现在持续增加工人数量。每个工人都可以生产更多的产量，但最终一个额外增加的工人生产出的额外产量将小于 100 单位。在增加了 4 或 5 个工人后，每个工人带来的额外的产量可能下降到 90，80，70…甚至更少。如果我们在这一亩地上增加成百上千的工人，那么每个额外工人甚至可能造成产量下降！这和厨师做汤的道理是类似的，厨师多了事情反而做不好。

因此我们通常可以预期当某种要素投入量越来越多时，它的边际产量是递减的。这个规律称为**边际产量递减定律**（law of diminishing marginal product）。它并不是个“定律”，它只是大多数生产过程具有的常见特征。

需要强调的是边际产量递减规律仅适用于当**所有其他**要素投入量都不变的情形。例如，在上述种地的例子中，我们考虑的是仅改变劳动要素的投入量，但维持土地和原材料等要素的投入量不变。

18.8 技术替代率递减

与边际产量递减假设密切相关的另外一条关于技术的假设，是**技术替代率递减**（diminishing technical rate of substitution）。这就是说，当我们增加要素 1 的投入量，调整要素 2 的投入量从而使产量不变（还在原来的等产量线上），技术替代率是递减的。大致来说，TRS 递减假设，意味着当我们沿着等产量线向要素 1 增加的方向移动时，等产量线的斜率绝对值是递减的；但若我们沿着等产量线向要素 2 增加的方向移动时，等产量线的斜率绝对值是递增的。这表明等产量线具有和良好性状无差异曲线相同类型的凸的形状。

技术替代率递减假设和边际产量递减密切相关但并不完全相同。边际产量递减这个假设是说当某种要素数量增加但**维持其他要素数量不变**时，产量如何变化。TRS 递减这个假设则是说，当我们增加其中一种要素的数量，并且**减少另外一种要素数量从而使产量不变**时，这两种要素的边际产量（等产量线的斜率）是如何变化的。

18.9 长期和短期

我们现在回归技术的本原概念——技术是一系列可行的生产方案。我们想将生产方案划分为**立即**可行（immediately feasible）和**最终**可行（eventually feasible）。

在**短期**（short run），某些生产要素的数量固定在预先决定的水平。在上面种地的例子中，如果仅有一亩地，我们考虑的所有生产方案中的土地要素数量就是固定的。也许他能得到更多的土地，从而可以生产更多的玉米，但是在短期，他必须接受土地面积固定的这一束缚。

另外一方面，在长期，农民可以购买更多的土地，或者卖掉他现有的部分土地。他可以调整土地的面积以达到利润最大化。

经济学家对长期和短期的区分是这样的：在短期至少有一种生产要素的数量是固定的，例如，土地面积固定、工厂规模固定、机器数量固定等等。在长期，**所有**生产要素都是可变的。

此处不涉及具体的时间长短。长期和短期的区分取决于我们考虑的是什么类型的选择。在短期，至少某些要素的数量固定在给定的水平，但在长期，这些要素的数量都是可变的。

假设要素 2 的数量短期固定在 $\bar{x}_2$。则相关的短期生产函数为 $f(x_1, \bar{x}_2)$。我们可以画出产量和 x_1 之间的函数关系，如图 18.5 所示。

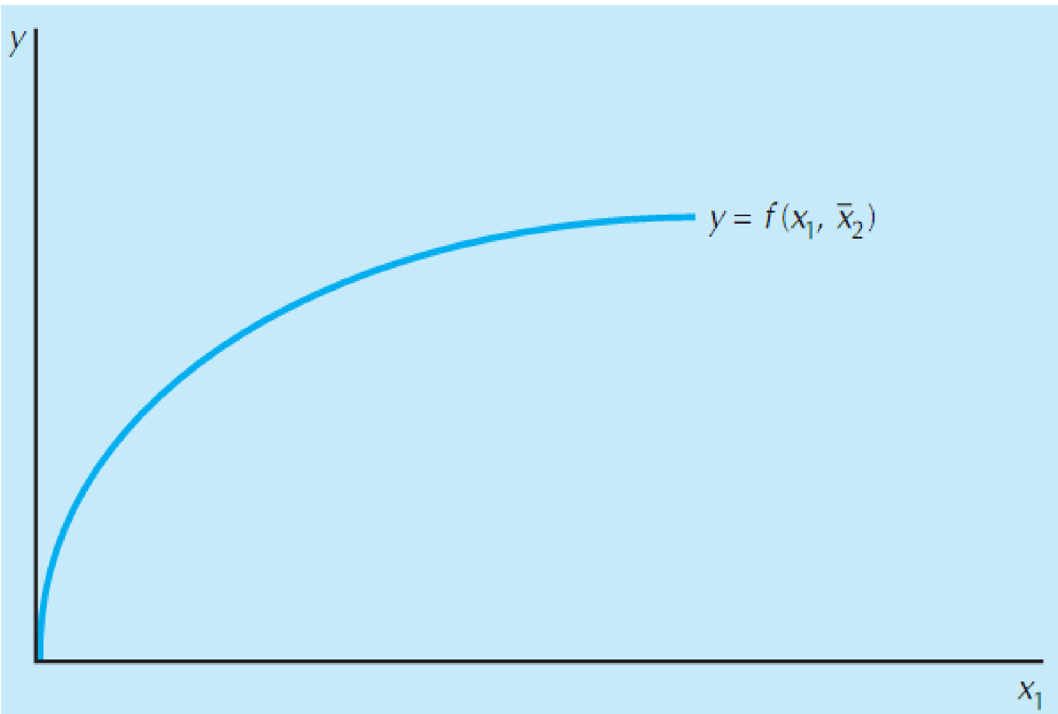


图 18.5：生产函数。这是短期生产函数的一种可能形状。

注意，随着要素 1 数量的增加，我们将短期生产函数画的越来越平坦。这是由于边际产

量递减规律在起作用。当然，也有可能出现下列现象——在一开始区间边际报酬是递增的，即随着要素 1 数量的增加，要素 1 的边际产量是递增的。在前面耕地的例子中，最初增加的几个工人可能使产量越来越多，因为他们可以有效地分工和合作。但是如果土地面积固定，劳动的边际产量最终是递减的。

18.10 规模报酬

现在我们考虑另外一种不同的实验。在该实验中，我们不是仅增加一种要素的投入数量并且维持其他要素投入水平不变，而是增加生产函数中所有要素的投入数量。换句话说，我们将所有要素的数量同时增加相同的幅度：例如要素 1 和 2 的投入量变为原来的 2 倍。

如果每种要素的数量变为原来的 2 倍，产量将为多大？产量非常有可能是原来的 2 倍。这种情形称为**规模报酬不变**（constant returns to scale）。用生产函数来表示，规模报酬不变意味着每种要素的投入量变为原来的 2 倍则产量也变为原来的 2 倍。在两种要素的情形下，我们将其表达为

$$2f(x_1,x_2)=f(2x_1,2x_2).$$

一般来说，如果我们将所有要素数量乘以某个正数t，规模报酬不变意味着产量也变为原来的t倍：

$$tf(x_1,x_2)=f(tx_1,tx_2).$$

出现上述结果的可能原因是：通常企业能够**复制**（replicate）它的生产过程。若企业拥有的要素数量变为原来的 2 倍，则它可以再建一个工厂，因此产量也为原来的 2 倍。如果要素数量为原来的 3 倍，它可以一共建 3 个完全一样的工厂，以此类推。

注意，完全有可能出现下列情形：某技术是规模报酬不变的但每种要素的边际产量又是递减的。**规模报酬**（returns to scale）描述的是如果你增加**所有**要素的投入，产量将会怎么变化；而边际产量递减描述的是当你增加**一种**要素投入但维持其他要素数量不变时，产量将会怎么变化。

由于企业能复制自己的工厂，规模报酬不变是最常见的情形，但这不等于说不存在其他的情形。例如，当我们将两种要素的数量都增加为原来的t倍时，我们得到的产量可能**大于**原来的t倍。这种情形称为**规模报酬递增**（increasing returns to scale）。在数学上，规模报酬递增意味着对于所有$t>1$均有

$$f(tx_1,tx_2)>tf(x_1,x_2).$$

能举一个规模报酬递增的技术的例子吗？输油管就是这样的。如果我们将油管的直径增加为原来的 2 倍，制造油管的材料需要量也为原来的 2 倍，但是油管的截面是原来的 4 倍。因此输油量大于原来的 2 倍。

(当然，我们不能无限增大油管的直径。如果持续增加油管的直径，则油管最终会被自己的重量压塌。规模报酬递增通常有自己的产量范围限制。)

还有一种情形是**规模报酬递减**（decreasing returns to scale），在该情形下对于所有 $t>1$ 均有：

$$f(tx_1,tx_2)<tf(x_1,x_2)$$

这种情形多少有些特别。如果要素投入都为原来的 2 倍，但产量小于原来的 2 倍，肯定是我们做错了什么事情。毕竟我们可以象以前一样复制工厂。

规模报酬递减现象最常见的原因是我们忘记了考虑某些生产要素。如果除了一种要素之外，其余要素的投入数量都为原来的 2 倍，这种情形并不是对原来生产过程的完全复制，因此产量不可能为原来产量的 2 倍。规模报酬递减实际上是个短期现象，即某些要素固定不变。

当然，某技术在不同产量水平可能出现不同规模报酬现象。当产量较低时，该技术可能是规模报酬递增的——所有要素变为原来的 t 倍，产量大于原来的 t 倍。后来，当产量较大时，该技术可能是规模报酬不变的——所有要素变为原来的 t 倍，产量等于原来的 t 倍。

例子：数据处理中心

数据处理中心是能容纳成千上万台计算机的建筑物。这些计算机用来执行诸如提供网页服务之类的任务。互联网公司，例如 Google, Yahoo，Microsoft, Amazon 等等在世界各地建立了几千个数据处理中心。

典型的数据中心由几百个用来支撑计算机母板的内存条（racks）组成。这些母板和你的台式计算机的母板是类似的。通常这些系统易于升级。因此，只要通过增加（或减少）计算机内存条，就能增加（或减少）数据处理中心的计算能力。

这样的复制意味着“计算服务”这种产品的生产函数实际上是规模报酬不变的：为了使产出翻番，你只要将所有投入要素翻番即可。

例子：完全复制！

英特尔（Intel）公司有几十家“绝妙的工厂”（fab plants）[一]，这些工厂生产、组装、分类和测试计算机集成电路片。

集成电路片的生产过程精致，英特尔发现难以在不同环境下进行质量管理。即使工厂设计的微小变动，例如清洁程序或冷却管长度的变动，都会对生产过程造成很大的影响。

[一] fab plants 是双关语，fab 的意思是“绝妙的”，而 fabricate（生产）前三个字母也是“fab”，因此这里指英特尔的生产工厂非常好。好在哪里？看完这个案例，你就知道了。译者注。

为了管理这些细微的影响，英特尔实施了名字叫做“完全复制！”的流程。英特尔对此的官方指导为：“任何可能影响生产过程的东西，或它是如何运转的，都必须复制到细微之处，除非不可能做到完全复制。若要变动，则该变动必须能够产生巨大的竞争利益。”

这意味着英特尔的一个工厂和另外一个工厂几乎完全相同，而且是故意这么做的。正如复制观点（replication argument）暗示的：对于英特尔公司来说，使产量翻倍的最容易方法，是尽可能完全地复制当前的运营程序。

总结

1.企业的技术约束可用生产集或生产函数描述。生产集刻画了技术可行的投入和产出的所有组合。生产函数则是说对于既定数量的投入，最大的产出为多少。

2.描述企业技术约束的另外一种方法是使用等产量线——描述能生产既定产量的投入的所有组合的曲线。

3.我们一般假设等产量线是凸且单调的，这一点类似消费者理论中的良好性状的偏好的假设。

4.边际产量衡量额外增加一单位某种要素的投入但维持其他要素投入量不变时，能得到的额外产量。我们通常假设当该要素投入量逐渐增加时，它的边际产量是递减的。

5.技术替代率（TRS）衡量等产量线在某点的斜率。我们通常假设沿着等产量线向 x_1 增加的方向移动时，TRS 是递减的——这是等产量线为凸的另外一种表达方法。

6.在短期，至少有一种要素的投入量是固定不变的，而在长期，所有要素的数量均可变。

7.规模报酬是指当我们改变生产规模时产量的变动方式。如果所有要素的投入量都变为原来的 t 倍，产量也变为原来的 t 倍，这种情形称为规模报酬不变。如果产量大于原来的 t 倍，则称为规模报酬递增；若产量小于原来的 t 倍，则称为规模报酬递减。

复习题

1.考虑生产函数 $f(x_1,x_2)=x_1^2x_2^2$。该生产函数是规模报酬不变、递增还是递减的？

2.考虑生产函数 $f(x_1,x_2)=4x_1^{1/2}x_2^{1/3}$。该生产函数是规模报酬不变、递增还是递减的？

3.柯布-道格拉斯生产函数为 $f(x_1,x_2)=Ax_1^ax_2^b$。可以证明该函数的规模报酬类型取决于 $(a+b)$ 数值的大小。$(a+b)$ 数值为多大时分别对应着规模报酬不变、递增还是递减？

4.两种生产要素 x_1 和 x_2 的技术替代率为 -4。如果你想继续生产原来的产量，而且你将

x_1 减少了 3 单位，那么你需要多少单位的 x_2？

5.对还是错？如果边际产量递减规律不成立，则我们用一个花盆就能种植出足够养活世界人口的粮食。

6.在某个生产过程中，有没可能出现某种要素的边际产量递减但规模报酬又是递增的现象？

复习题答案

1.考虑生产函数 $f(x_1,x_2)=x_1^2x_2^2$。该生产函数是规模报酬不变、递增还是递减的？

【复习内容】规模报酬不变、递增和递减

规模报酬不变：对于所有 $t>0$ 均有 $f(tx_1,tx_2)=tf(x_1,x_2)$.

规模报酬递增：对于所有 $t>1$ 均有 $f(tx_1,tx_2)>tf(x_1,x_2)$.

规模报酬递减：对于所有 $t>1$ 均有 $f(tx_1,tx_2)<tf(x_1,x_2)$.

注意第一种情形和后两种情形数学表达式中对 t 值要求的不同。

【参考答案】

$$f(tx_1,tx_2)=(t^2x_1^2)(t^2x_2^2)=t^4x_1^2x_2^2=t^4f(x_1,x_2)$$

对于任意 $t>1$，上式意味着 $f(tx_1,tx_2)=t^4f(x_1,x_2)>tf(x_1,x_2)$，根据定义可知，这是规模报酬递增的。

2.考虑生产函数 $f(x_1,x_2)=4x_1^{1/2}x_2^{1/3}$。该生产函数是规模报酬不变、递增还是递减的？

【复习内容】规模报酬不变、递增和递减

$$f(tx_1,tx_2)=4(t^{1/2}x_1^{1/2})(t^{1/3}x_2^{1/3})=t^{5/6}(4x_1^{1/2}x_2^{1/2})=t^{5/6}f(x_1,x_2)$$

对于任意 $t>1$，上式意味着 $f(tx_1,tx_2)=t^{5/6}f(x_1,x_2)<tf(x_1,x_2)$，根据定义可知，这是规模报酬递减的。

3.柯布-道格拉斯生产函数为 $f(x_1,x_2)=Ax_1^ax_2^b$。可以证明该函数的规模报酬类型取决于 $(a+b)$ 数值的大小。$(a+b)$ 数值为多大时分别对应着规模报酬不变、递增还是递减？

【复习内容】规模报酬不变、递增和递减

$$f(tx_1,tx_2)=A(t^a x_1^a)(t^b)x_2^b=t^{a+b}(Ax_1^a x_2^b)=t^{a+b}f(x_1,x_2)$$

显然，

若$a+b=1$，则对于任意$t>0$，上式意味着$f(tx_1,tx_2)=t^{a+b}f(x_1,x_2)=tf(x_1,x_2)$，根据定义可知，这是规模报酬不变。

若$a+b>1$，则对于任意$t>1$，上式意味着$f(tx_1,tx_2)=t^{a+b}f(x_1,x_2)>tf(x_1,x_2)$，根据定义可知，这是规模报酬递增。

若$a+b<1$，则对于任意$t>1$，上式意味着$f(tx_1,tx_2)=t^{a+b}f(x_1,x_2)<tf(x_1,x_2)$，根据定义可知，这是规模报酬递减。

4.两种生产要素x_1和x_2的技术替代率为-4。如果你想继续生产原来的产量，而且你将x_1减少了3单位，那么你需要多少单位的x_2？

【复习内容】技术替代率

技术替代率衡量两种要素之间的替代关系。它衡量企业为了维持产量不变，必须用一种要素替代另外一种要素的比率。

推导TRS的表达式：考虑维持产量不变的要素1和2使用量的变化，可得

$$\Delta y=MP_1(x_1,x_2)\Delta x_1+MP_2(x_1,x_2)\Delta x_2=0$$

由此可解得

$$TRS_{12}=\frac{\Delta x_2}{\Delta x_1}=-\frac{MP_1(x_1,x_2)}{MP_2(x_1,x_2)}.$$

【参考答案】

$$TRS_{12}=\frac{\Delta x_2}{\Delta x_1}\Rightarrow\Delta x_2=\Delta x_1\cdot TRS_{12}\Rightarrow\Delta x_2=(-3)(-4)=12$$

5.对还是错？如果边际产量递减规律不成立，则我们用一个花盆就能种植出足够养活世界人口的粮食。

【复习内容】

当某种要素投入量越来越多时，它的边际产量最终是递减的。这个规律称为边际产量递减定律。它并不是个“定律”，它只是大多数生产过程具有的常见特征。

边际产量递减最常见的一种解释是各种要素之间的搭配比例有一定的合适范围，超过这个范围边际产量就会递减。

需要强调的是边际产量递减规律仅适用于当所有其他要素投入量都不变的情形。例如，在上述种地的例子中，我们考虑的是仅改变劳动要素的投入量，但维持土地和原材料等要素的投入量不变。

【参考答案】

这种说法是正确的，因为如果边际产量递减不成立，则意味着边际产量不变或者递增。为论述简单起见，假设边际产量不变。

再假设在一个花盆种植粮食时，最初一单位劳动可以生产一单位粮食。由于边际产量不变，这意味着在这个花盆中连续追加 n 单位劳动可以生产 n 单位粮食。当 n 足够大时，不仅能够实现世界粮食现有的供给量，还能远远超过。

6.在某个生产过程中，有没可能出现某种要素的边际产量递减但规模报酬又是递增的现象？

【复习内容】边际产量；边际产量递减规律；规模报酬；规模报酬递增

（1）边际产量的概念

边际产量是生产函数对要素投入的偏导数。由于我们求某要素的边际产量时需要假定其他要素的投入量不变，因此**边际产量是个短期性质的概念**。

（2）边际产量递减

边际产量递减这个假设是说当某种要素数量增加但维持其他要素数量不变时，产量如何变化。具体地说，当某种要素投入量越来越多时，它的边际产量最终是递减的。这个规律称为边际产量递减定律。它并不是个“定律”，它只是大多数生产过程具有的常见特征。

由此可见，边际产量递减是个短期现象。

（3）规模报酬

规模报酬是指当我们改变生产规模时（所有要素的投入量变动幅度相同）产量的变动方式。如果所有要素的投入量都变为原来的 t 倍，产量也变为原来的 t 倍，这种情形称为规模报酬不变。如果产量大于原来的 t 倍，则称为规模报酬递增；若产量小于原来的 t 倍，则称为规模报酬递减。

由此可见，**规模报酬是个长期性质的概念，因为它涉及所有要素同时变动**。

【参考答案】

由上面的分析可知，边际产量是个短期性质的概念，而规模报酬是个长期性质的概念。因此二者属于不同范畴，当然可以出现某种要素的边际产量递减但规模报酬又是递增的现象。举一个具体的例子：

例如柯布-道格拉斯生产函数 $f(x_1,x_2)=x_1^{2/3}x_2^{2/3}$ 。

利用本章复习题第 3 题的结论，可知，该生产函数是规模报酬递增的，因为$\frac{2}{3}+\frac{2}{3}>1$。

要素 1 的边际产量 $MU_1=\frac{\partial f}{\partial x_1}=\frac{2}{3}x_2^{2/3}x_1^{-1/3}$，这个函数是个减函数，因为$\frac{\partial MU_1}{\partial x_1}=\frac{\partial^2 f}{\partial x_1^2}=-\frac{2}{9}x_2^{2/3}x_1^{-4/3}<0$，这就是说要素 1 的边际产量递减；同理可知要素 2 的边际产量也是递减的。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8^{th} Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

19.利润最大化（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

19 利润最大化

在上一章，我们讨论了如何描述企业面对的技术选择。在本章，我们分析企业如何选择产量以及它使用的生产方法。我们将使用利润最大化模型：企业选择生产方案使得利润最大化。

在本章，我们假设企业面对的投入品和产出品价格是固定的。我们曾指出过，经济学家将价格不受单个生产者影响的市场称为**完全竞争市场**（competitive market）。因此，在本章我们研究的是，在投入品和产出品的市场都是完全竞争市场的情形下，厂商如何实现利润最大化的问题。

19.1 利润

利润的定义是收入减去成本。假设某企业使用 m 种生产要素 $(x_1,\ldots,x_m)$ 生产 n 种产品 $(y_1,\ldots,y_n)$。令生产要素的价格为 $(w_1,\ldots,w_m)$，产品的价格为 $(p_1,\ldots,p_n)$。

企业得到的利润 π，可以表示为

$$\pi=\sum_{i=1}^{n}p_iy_i-\sum_{i=1}^{m}w_ix_i\ .$$

上式右侧第一项为收入，第二项为成本。

在成本表达式中，必须确保将企业使用的所有生产要素都包括进来，并且用市场价格计价。成本这一项通常很明显，但如果企业的所有人和经营者如果是同一人的话，则很有可能漏计了某些生产要素。

例如，如果某人在自己的企业工作，那么他自己的工作就是一种投入，需要包含在成本内。他的工资率就是他的劳动的市场价格——如果他在市场上出售自己的劳动所能得到的工资率。类似地，如果某个农民拥有土地并将土地用于耕种，则为了计算经济成本（economic cost），他的土地应该按市场价值估价。

这类经济成本通常称为**机会成本**（opportunity cost）。这个名字来源于如下的思想：例如，如果你把自己的劳动用于某种用途，那么你就放弃了将其使用在其他用途的机会。因此，损失的这部分工资是生产成本的一部分。农民耕种自己的土地也是类似的：农民本来可以将他的土地租赁给别人，但是他选择了自租从而放弃了租赁收入。损失的租金是他生产成本的一部分。

利润的经济学定义要求：所有投入品和产出品都应该按它们的机会成本计价。会计人员确定的利润未必正好等于经济利润，因为他们通常使用历史成本（historical costs）——某要素按最初购买的价格计算支出——而不是经济成本——某种要素如果按当前价格购买需要

支出多少。利润的定义有多种，但是我们坚持使用利润的经济学定义。

另外，如果你有时搞混了时间刻度，也容易出现错误。我们通常将要素投入看作**流量**（flows）。因此，我们通常说：每周多少劳动和每周多少机器小时生产每周多少产品。要素的价格要用这些流量的合适单位来衡量。工资自然用每小时多少元衡量。类似地，使用机器的代价是**租金价格**（rental rate）——在一定时期内你租用某机器而支付的费用。

在很多情形下，不存在租赁机器的成熟市场，因为企业通常购买它们的资本设备。在这种情形下，我们必须计算隐性租金费用——比较当初购买机器花了多少钱以及最后将机器卖出卖了多少钱。

19.2 企业的组织形式

在资本主义经济中，企业是由个人所拥有的。企业只是法人实体；企业的所有者对企业的行为负责，企业的所有者为该行为支付成本或者收获回报。

一般来说，企业按组织形式可以分为独资企业、合伙制企业或公司制企业。独资（proprietorship）企业为一人所拥有。合伙制（partnership）企业由两人或多人共同拥有。公司制（corporation）企业也是由多人共同拥有，但法律允许这样的企业与所有人相分离。因此，合伙制企业只有合伙人都存活而且同意继续经营时，企业才能延续下去。公司制企业的寿命则不受企业所有者的寿命所限。正是由于这个原因，大多数大型企业都是公司制企业。

这些不同类型的企业的所有者在企业经营方面的目标可能不同。独资企业或合伙企业的所有人通常直接参与企业的日常经营管理，所以他们能执行他们确定的目标。一般来说，企业所有者会对利润最大化感兴趣。但是，如果他们的目标是非盈利的，那么他们也可能一直坚持非盈利的目标。

公司制企业的所有人和经营者通常是分开的。因此所有权和控制权是分离的。企业的所有人必须为经营者设定经营目标，而且需要想方设法让经营者实际也在追求该目标。利润最大化再一次成为常见的目标。下面我们将看到，在某些条件下，这个目标将有可能引导经营者选取符合所有人利益的行为。

19.3 利润和股票市场价值

企业使用的生产过程通常要经历较长的时间。在时刻 t 投入的要素可能要在以后才能得到全部回报。例如，企业建立的工厂可能要持续 50 年或者 100 年。在这种情形下，在某时点的要素投入有助于未来的生产。

在上述情形下，我们必须估计一段时期内的成本流和收入流的价值。我们在第 10 章已知道，做此事的合适方法是使用现值的概念。当人们能在金融市场借贷时，可使用利率定义不同时期的消费价格。企业也可以进入这些金融市场，因此同样可用利率来评价投资决策。

考虑一个完全确定性的世界，在这样的世界，企业未来的利润流是公开的。利润流的现值就是**企业的现值**（present value of the firm）。这就是某人购买该企业愿意支付的资金数额。

我们在上面已指出，大多数大型企业是公司制企业，就是说这样的企业为多人所共有。公司发行的股票代表股东对公司所有权的拥有份额。在某些时间，公司会对股份派发股息，股息是公司利润的一部分。股票可以在股票市场买卖。股票价格代表人们期望从公司获得的股息流的现值。某企业的股票市场总值代表企业预期利润流的现值。因此，企业的目标——使企业利润流的现值最大化——也可以说成使股票市场价值最大化。在确定性的世界中，这两个目标是同一个东西。

企业所有者通常希望企业选择的生产方案能使该企业股票市场价值最大化，因为这样就能使他们持有的股份具有最大的价值。在第 10 章我们已经知道，无论消费者对不同时期的消费有什么样的偏好，他总是喜欢现值更高的禀赋。通过使股票市场价值最大化，企业就能使股东的预算集尽可能地大，因此这种做法代表了所有股东的最大利益。

如果企业的利润流具有不确定性，那么企业所有者要求经营者实现利润最大化就没有明确的意义。他们应该使期望利润最大化吗？他们应该使利润的期望效用最大化吗？经营者对待风险投资应持有什么样的态度？因此，在不确定性的情形下，很难说清利润最大化是什么意思。然而，即使在不确定性的情形下，使股票市场价值最大化仍然有意义。如果企业的经营者试图使得公司的股票价值尽可能地大，那么他们就会使公司所有人（所有股东）的状况尽可能地好。因此在几乎所有的经济环境中，使股票市场价值最大化对企业来说是一个定义清晰的目标。

尽管我们我们对现值和不确定性说了那么多，我们还是将我们的注意力放在更简单的利润最大化问题上，也就是说，我们主要分析时期为一期、产出只有一种且不具有不确定性的利润最大化问题。这事虽然简单但它仍能让我们具有深刻的见解，熟悉这种方法将有助于分析企业行为的更一般模型。利润最大化思想的大部分可以自然推广到更一般的模型。

19.4 企业的边界

企业经营者经常面对的一个问题是“自产还是购买”。也就是说，某企业应该使某些东西内部化还是从外部市场购买？这个问题非常广泛，因为它不仅可以指实物产品，也可以指各种服务。的确，在最宽泛的意义上，“自产还是购买”适用于企业的几乎所有决策。

公司内部应该设立或者提供以下的服务吗：食堂？看门人？复印？旅游咨询？显然，很多因素都会影响上述决策。一个重要因素是公司的规模。一个只有 12 个雇员的小型夫妻店一般不会提供内部食堂服务。但是它有可能将看门人服务外包，这取决于成本、产能和人员配置。

即使能够提供内部食堂的大型企业也未必真正设立食堂，这取决于员工能否很容易地找到吃饭的地方。如果企业坐落在大城市的市区，那么吃饭的地方会非常多；如果企业位于郊

区，员工用餐地点的选择就会较少。

一个关键性的问题是，这些商品或服务是由外部垄断市场还是竞争市场提供的。大体来说，企业经营者喜欢在完全竞争市场上购买商品或服务。次佳选择是从内部垄断者购买。最差的情形——从价格和服务质量角度来说——是从外部垄断企业购买。

以复印服务为例。最理想的情形是有很多复印店都想为你复印，这样你就能以较低的价格得到高质量的服务。如果你的大学比较大或者在市区，那么复印店就会很多。另一方面，位于郊区的较小的大学拥有的选择相对较少，价格通常也比较高。

商业活动也是如此。在高度竞争的环境下，消费者的选择就较多。相比而言，企业内部的复印部门相对不受欢迎。即使价格较低，但是服务质量较差。但是最差的选择肯定是向外部垄断者购买。企业内部垄断部门提供的服务，可能质量较差，但至少肥水没流外人田。

随着技术进步，原来很多企业内部化的商品或服务也随之外包。40 年前，企业内部提供很多服务，现则它们尽可能地将这些外包出去。餐饮服务、复印服务和看门人服务通常由专业化的外部企业提供。由于这些行业已经专业化，因此企业可以较低的价格获得更高质量的服务。

19.5 固定要素和可变要素

在某个既定的时间段，企业可能很难调整某些要素的投入数量。如果合同规定某企业使用既定数量的某些要素，那么该企业通常需要履行这样的义务。例如，房屋租赁合同可能规定企业在某时间段租借既定面积的厂房。如果企业某种要素的投入量是固定不变的，企业无法进行调整，则这种要素称为**固定要素**（fixed factor）；如果某种要素的数量可以调整，则称为**可变要素**（variable factor）。

在第 18 章我们已经知道，短期的概念是指在该时期内某些要素是固定要素——这些要素的投入量是固定不变的；而在长期，企业能够自由调整所有的生产要素：所有要素都是可变要素。

短期和长期的划分并没严格的界限。具体时间段的长短需要具体问题具体分析。重要的事情是，你要知道某些生产要素在短期内是不变的，但在长期它们是可变的。既然所有要素在长期都是可变的，企业可以自由决定使用零单位的投入从而产出也为零——即退出该行业。因此，在长期，企业的最小利润为零利润。

在短期，企业不得不继续使用某些生产要素，即使它决定生产零单位产品。因此，在短期企业的利润完全有可能为**负利润**。

由定义可知，对于固定要素来说，即使企业决定停产（产量为零）也必须为这些要素支付成本：如果企业长期租赁某厂房，它必须按期支付租金，这和它是否继续生产无关。但是还有一类生产要素，这类要素只有企业的产量为正时才需要支付成本。例如照明电。如果

企业停产，不需要提供照明；但是只要它生产，它必须购买一定数量的照明。

这样的要素称为**准固定要素**（quasi-fixed factors）。只要企业的产量为正（无论它生产多少单位产品），这些要素的使用量固定不变。区分固定要素和准固定要素有时会有助于分析企业的经济行为。

19.6 短期利润最大化

我们开始分析短期的利润最大化问题。假设要素 2 的数量固定在 $\bar{x}_2$，企业的生产函数为 $f(x_1, x_2)$，产品价格为 p，要素 1 和 2 的价格分别为 w_1 和 w_2。则企业的短期利润最大化问题可以写为

$$\max_{x_1} \quad pf(x_1, \bar{x}_2) - w_1 x_1 - w_2 \bar{x}_2 .$$

要素 1 的最优投入数量不难确定。

令 x_1^* 表示利润最大化时要素 1 的最优投入数量，则产品价格 p 乘以要素 1 的边际产量应该等于要素 1 的价格，即

$$pMP_1(x_1^*, \bar{x}_2) = w_1 .$$

换句话说，**某种要素的边际产量的价值应该等于该要素自身的价格**。

为了理解这个法则，考虑企业决定多投入一些要素 1。当要素 1 的投入量增加了 Δx_1，产量因此增加 $\Delta y = MP_1 \Delta x_1$，这些产量的价值为 $pMP_1 \Delta x_1$。但是这些边际产量的成本为 $w_1 \Delta x_1$。如果边际产量的价值大于成本，则**增加**要素 1 的投入会增加利润。如果边际产量的价值小于成本，则**减少**要素 1 的投入会增加利润。

如果企业的短期利润已达到最大，那么当我们增加或减少要素 1 的投入时，利润不会继续增加。这表明在利润最大化的投入和产出处，边际产量的价值 $pMP_1(x_1^*, \bar{x}_2)$ 应该等于要素的价格 w_1。

我们也可以使用图形推导这个结论。请看图 19.1。图中的曲线代表要素 2 的数量固定为 $\bar{x}_2$ 时的生产函数。令 y 表示企业的产量，利润表达式为

$$\pi = py - w_1 x_1 - w_2 \bar{x}_2 .$$

由此解出 y，即将产量 y 表达为 x_1 的函数

$$y = \frac{\pi}{p} + \frac{w_2}{p}\bar{x}_2 + \frac{w_1}{p}x_1 . \tag{19.1}$$

这个式子是**等利润线**（isoprofit lines）的表达式。它刻画的是能带来固定利润水平 π 的投入和产出的所有组合。当 π 变动时，我们得到了一系列互相平行的直线，它们的斜率都为 w_1 / p；它们的纵截距都为 $\pi / p + w_2 \bar{x}_2 / p$，纵截距衡量的是利润与固定成本的和。当然，

此处的利润和固定成本是经过价格调整后（除以 p ）的实际利润和实际固定成本。

固定成本是固定不变的，所以当我们从一条等利润线移动到另外一条等利润线时，唯一变动的是利润水平。因此，等利润线的纵截距越大，代表的利润水平越高。

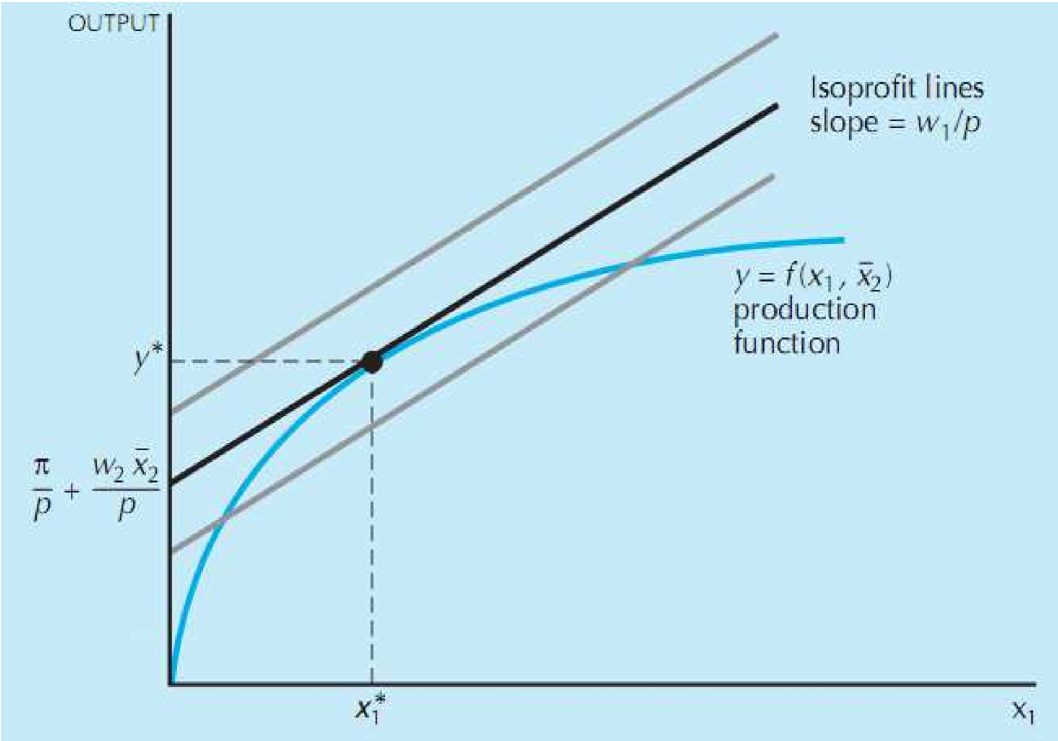


图 19.1：利润最大化。企业选择的投入和产出组合位于最高的等利润线上。在这种情形下，利润最大化的点为 (x_1^*, y^*) 。

因此，利润最大化问题转化为在生产函数曲线上找到一点，使得这点位于位置最高的等利润线上。图 19.1 画出了这样的点。和以前一样，可用相切条件刻画这样的点：生产函数的斜率应该等于等利润线的斜率。由于生产函数的斜率是边际产量，等利润线的斜率为 w_1/p，相切条件因此可以写为

$$MP_1 = \frac{w_1}{p},$$

这正好等价于我们前面推导出的条件。

19.7 比较静态分析

我们可用图 19.1 的图形，分析企业的投入和产出选择是怎样随着投入和产出的价格变化而变化的。这样我们就得到了分析企业行为的比较静态（comparative statics）方法。

例如：当我们改变要素 1 的价格 w_1 时，企业对要素 1 的最优选择将会如何变化？根据（19.1）式（等利润线的定义），我们看到当 w_1 上升时，等利润线会更陡峭，如图 19.2A 所

示。当等利润线更陡峭时，切点必然向左方移动。因此要素 1 的最优数量必然是下降的。这只是表明，当要素 1 的价格上升时，它的需求必然下降：要素需求曲线必定向下倾斜。

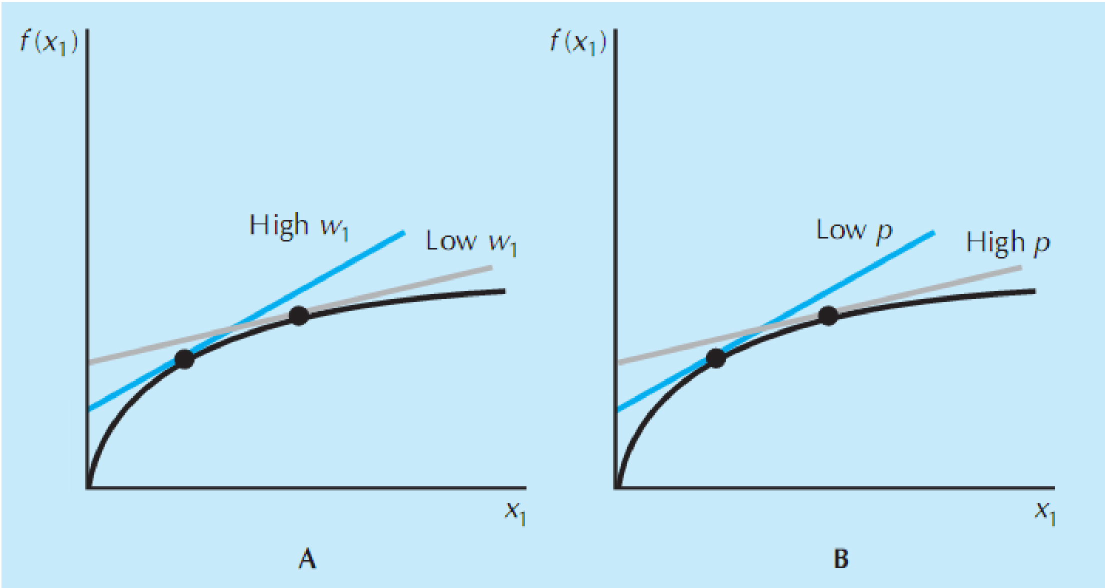


图 19.2：比较静态分析。A 图表示 w_1 上升减少了要素 1 的需求；B 图表示产品价格上升将会增加要素 1 的需求，从而增加了产品的供给。

类似地，如果产品价格下降，等利润线必然更陡峭，如图 19.2B 所示。与上一段的分析类似，要素 1 的最优数量将会下降。在短期，假设要素 1 的数量下降而且要素 2 的数量不变，则产品供给必定减少。这样我们就得到了另外一个比较静态结果：产品价格下降将会减少供给。换句话说，供给函数必然是向上倾斜的。

最后，我们分析要素 2 的价格变动时结果将如何。由于这是短期分析，要素 2 的价格变动不会改变企业对要素 2 的使用量——在短期，要素 2 的使用量固定在 $\overline{x}_2$ 水平。要素 2 的价格改变对等利润线的斜率无任何影响。因此，要素 1 的最优选择将不会改变，产品的供给也不会改变。唯一改变的是企业的利润。

19.8 长期利润最大化

在长期，企业可以自由选择所有要素的投入量。因此，长期利润最大化问题可以表述为

$$\max_{x_1, x_2} \; pf(x_1, x_2) - w_1x_1 - w_2x_2 .$$

这个表达式和短期利润最大化的表达式类似，区别在于长期中所有要素都是可变的。

描述最优选择的条件在本质上也和短期利润最大化的条件一样，但是现在我们必须把这一条件应用于**每种**要素身上。在前面我们已经知道，要素 1 的边际产量价值必须等于要素 1 的价格，不管要素 2 的使用量为多少。类似的条件必须对每种要素都成立。

$$pMP_1(x_1^*, x_2^*) = w_1$$

$$pMP_2(x_1^*, x_2^*) = w_2 .$$

如果企业已对要素 1 和 2 的投入量作出最优选择，则每种要素的边际产量价值必须等于各自的价格。在最优选择处，企业的利润不会随任何要素的投入量变动而增加。

上述结论的推理过程和短期利润最大化的推理相同。例如，如果要素 1 的边际产量价值大于要素 1 的价格，则稍微增加要素 1 的投入量，产量将会增加 MP_1 单位，销售收入因此增加了 pMP_1。如果这个产品的价值 pMP_1 大于生产它的要素成本，则增加这种要素的投入量就是值得的。

上述两个方程含有两个未知数 x_1^* 和 x_2^*。如果我们知道边际产量作为 x_1 和 x_2 函数表达式，我们就能求出每种要素的最优选择，它们是两种要素价格的函数。最终得到的方程就是**要素需求曲线**（factor demand curves）。

19.9 反要素需求曲线

企业的**要素需求曲线**（factor demand curves）衡量某要素的价格和该要素的利润最大化选择之间的关系。在前面我们已知道如何找到利润最大化的选择：对于任何价格 (p, w_1, w_2)，我们只要找到那些要素需求 (x_1^*, x_2^*)，使得每种要素的边际产量价值等于各自的价格即可。

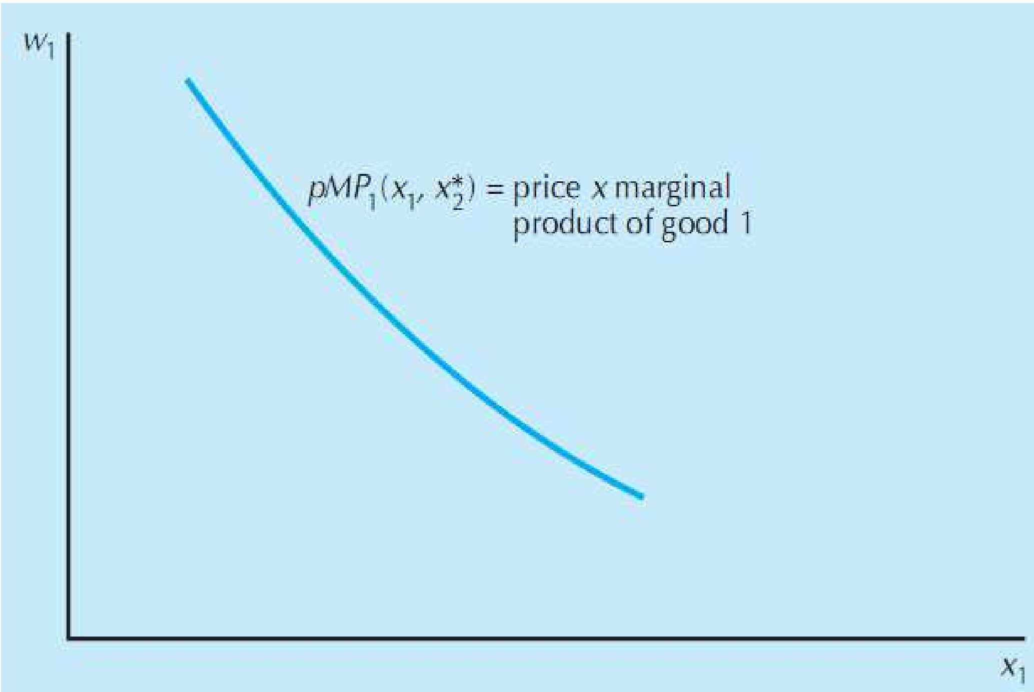


图 19.3：反要素需求曲线。该曲线衡量，在要素 2 投入数量固定在 x_2^* 的情形下，要素 1 的价格应为多大才能使要素 1 的需求为 x_1 单位。

反要素需求曲线（inverse factor demand curve）衡量的关系与上述相同，但站在不同的视角。它衡量要素需求量一定时要素价格必须为多大。给定要素 2 的最优选择，我们可以画出要素 1 的最优选择和它自身价格的关系图，如图 19.3 所示。它只是下列方程的图像

$$pMP_1(x_1, x_2^*) = w_1.$$

根据边际产量递减的假设，这条曲线是向下倾斜的。对于任何水平的 x_1，该曲线表示在要素 2 的数量固定在 x_2^* 的情形下，为了诱使企业需求 x_1 单位的要素 1，要素 1 的价格应该为多大。

19.10 利润最大化和规模报酬

竞争性企业的利润最大化和规模报酬之间存在着一个重要的关系。假设某企业已经选择出长期利润最大化的产出 $y^* = f(x_1^*, x_2^*)$，此时要素投入量为 (x_1^*, x_2^*)。

该企业的利润为

$$\pi^* = py^* - w_1x_1^* - w_2x_2^*.$$

假设该企业的生产函数是规模报酬不变的，而且假设均衡时它的利润为正。现在如果我们将所有要素的投入量变为原来的 2 倍，结果将如何？根据规模报酬不变的假设可知，产量也会变为原来的 2 倍。利润将如何变化？

不难看出它的利润也会变为原来的 2 倍。矛盾出现了！因为我们在前面已假设它的原来选择已实现利润最大化，所以利润就不可能再增大。矛盾出现的原因是我们假设企业的原来利润水平为正；如果原来的利润为零就不会再有矛盾，因为零的 2 倍仍然是零。

上述结论表明，如果某竞争性企业在所有产量水平上都是规模报酬不变的，那么它的唯一可能长期利润水平就是零利润。（当然，如果某企业的长期利润为负，它就应该退出该行业。）

很多人可能会对该结论大吃一惊。企业不是希望实现利润最大化吗，对吧？如果是这样，它们怎么在长期内只能取得零利润呢？

想一想企业试图无限扩大生产规模时的结果，你就不会吃惊了。结果可能有三种情形。第一种情形是，企业变得非常大，不能有效运行。这其实是说企业不能在任何产量上都是规模报酬不变的。最终，由于协调比较困难，企业会进入规模报酬递减的区域。

第二种情形是，企业变得非常大，因此控制了产品市场。在这种情形下，没有理由要求该企业是竞争性的，它不可能再接受给定的市场价格。相反，它会利用自己的规模影响市场价格。这样，该企业的行为就不再适合运用竞争性利润最大化模型进行分析。以后在我们分析垄断企业时，我们会介绍分析此类企业行为的合适模型。

第三种情形是，如果一家企业的生产技术是规模报酬不变的，而且它能取得正的利润

曹乾(东南大学 caoqianseu@163.com)

那么，任何可以使用该技术的其他企业也能取得正利润。但是如果所有企业的产量都扩大了，就会必然压低产品的价格，从而使该行业所有企业的利润都下降。

19.11 显示的盈利能力

当某利润最大化企业在做投入和产出决策时，它透露了两条信息：第一，它选中的投入和产出组合代表着一个**可行**生产方案；第二，这个选择比企业能选取但未选取的其他组合更具有盈利能力。下面我们将详细说明。

假设我们观察到企业在不同价格水平下做出的选择。在时期t，企业面对的价格为(p^t, w_1^t, w_2^t)，它的选择为(y^t, x_1^t, x_2^t)；在时期s，企业面对的价格为(p^s, w_1^s, w_2^s)，它的选择为(y^s, x_1^s, x_2^s)。如果企业在这两个时期内的生产函数不变，则我们有

$$p^t y^t - w_1^t x_1^t - w_2^t x_2^t \geq p^t y^s - w_1^t x_1^s - w_2^t x_2^s \tag{19.2}$$

$$p^s y^s - w_1^s x_1^s - w_2^s x_2^s \geq p^s y^t - w_1^s x_1^t - w_2^s x_2^t \tag{19.3}$$

也就是说，在t期价格下：与选取s期的方案能得到的利润相比，t期方案实现的利润更大。反之亦然。如果企业的行为违背了上述任何一个不等式，则它的行为就不可能时利润最大化的行为（前提是技术不变）。

因此，如果我们观察到在某两个时期下，企业的行为违背了上述不等式的任何一个，则我们可以断言，该企业至少在一个时期内的利润不是最大化的。上述两个不等式实际上是利润最大化行为的一个公理，我们将这个公理称为**利润最大化的弱公理**（Weak Axiom of Profit Maximization, WAPM）。

如果企业的选择决策满足利润最大化弱公理，则我们就可以推导出一个有用的比较静态结论，这个结论描述了价格变动时企业对要素的需求行为和产品的供给行为。将（19.3）式的两侧移项可得

$$-p^s y^t + w_1^s x_1^t + w_2^s x_2^t \geq -p^s y^s + w_1^s x_1^s + w_2^s x_2^s \tag{19.4}$$

将（19.4）式加到（19.2）式可得

$$\begin{aligned}&(p^t - p^s)y^t - (w_1^t - w_1^s)x_1^t - (w_2^t - w_2^s)x_2^t \\ \geq &(p^t - p^s)y^s - (w_1^t - w_1^s)x_1^s - (w_2^t - w_2^s)x_2^s\end{aligned} \tag{19.5}$$

整理可得

$$(p^t - p^s)(y^t - y^s) - (w_1^t - w_1^s)(x_1^t - x_1^s) - (w_2^t - w_2^s)(x_2^t - x_2^s) \geq 0 \tag{19.6}$$

最后定义：价格变化$\Delta p = p^t - p^s$；产量变化$\Delta y = y^t - y^s$等等，这样我们就可以得到更紧凑的（19.7）式

$$\Delta p\Delta y-\Delta w_1\Delta x_1-\Delta w_2\Delta x_2\geq 0 \tag{19.7}$$

这个式子就是我们想要的式子。它是说产品价格变动量乘以产量的变动量，然后再减去每种要素价格的变动量与该要素变动量的乘积，结果必然非负。注意，我们仅根据利润最大化的定义就推导出了这个式子。但是它包含了利润最大化选择决策的所有比较静态结果！

例如，假设产品价格变动，但所有要素的价格维持不变，即$\Delta w_1=\Delta w_2=0$。由此可知，（19.7）式变为

$$\Delta p\Delta y\geq 0$$

因此如果产品价格上升即$\Delta p>0$，则产量的变动必定非负即$\Delta y\geq 0$。这就是说一家竞争性企业的利润最大化供给曲线的斜率必然为正（或至少为零）。

类似地，如果产品价格和要素 2 的价格维持不变，仅要素 1 的价格变动，则（19.7）式变为

$$-\Delta w_1\Delta x_1\geq 0$$

即

$$\Delta w_1\Delta x_1\leq 0\,.$$

因此，如果要素 1 的价格上升即$\Delta w_1>0$，则（19.7）式意味着要素 1 的需求将下降（或者不变），所以$\Delta x_1\leq 0$。这表明要素需求是要素价格的减函数：要素需求曲线的斜率为负（或为 0）。

利润最大化弱公理中的两个简单不等式，以及由此得到的（19.7）式，是利润最大化企业必须遵守的限制条件。你自然会问，这些条件是利润最大化模型施加在企业行为上的全部限制吗？换句话说，如果我们观察到某企业的选择决策，而且这些选择满足利润最大化弱公理，我们能否估计出一种生产技术，使得这种技术产生的选择是利润最大化选择？可以证明答案是肯定的。图 19.4 显示了如何构建这样的生产技术。

为了借助图形说明这个结论，假设投入和产出分别只有一种。假设我们观测到某企业在t期和s期的选择分别为(p^t,w_1^t,y^t,x_1^t)和(p^s,w_1^s,y^s,x_1^s)。在这两个时期我们都能计算出利润π_t和π_s，并且可以画出产生这些利润的y和x_1的所有组合。

也就是说，我们画出两条等利润线

$$\pi_t=p^t y-w_1^t x_1$$

和

$$\pi_s=p^s y-w^s x_1\,.$$

位于t期等利润线上方的点的利润（按t期价格计算）都比π_t大；位于s期等利润线上方的点的利润（按s期价格计算）都比π_s大。利润最大化弱公理要求：t期的选择必须位于

s 期等利润线的下方，s 期的选择必须位于 t 期等利润线的下方。

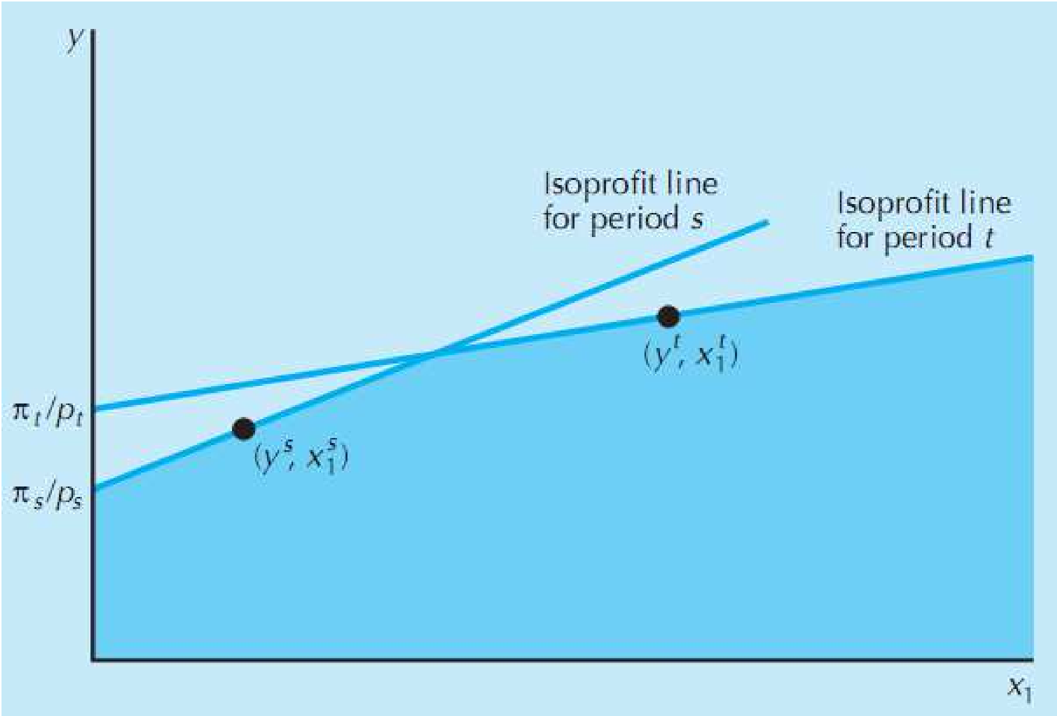


图 19.4：一种可行技术的构建。若观测到的选择是在每组价格下的最大利润选择，我们可以使用等利润线估计出产生这些选择的生产技术的形状。

如果这个条件得到满足，那么我们就能比较容易地构造一种生产技术，使得 (y^t, x_1^t) 和 (y^s, x_1^s) 都是利润最大化的选择点。这种技术（生产集）就是这两条等利润线下方的阴影区域[一]。请见图 19.5。这个区域中的任何一点产生的利润，无论按 t 期还是 s 期的价格计算，都小于（或等于）企业实际选中的点——(y^t, x_1^t) 和 (y^s, x_1^s) 能产生的利润。

现在证明我们构造的生产技术，能够生成观测到的企业选择（利润最大化选择）。这个证明在图形上一目了然。请见图 19.5。在价格 (p^t, w_1^t) 下，(y^t, x_1^t) 这个选择点位于最高的等利润线上，s 期的选择类似推理。

因此，当观测到的选择满足利润最大化弱公理时，我们就可以构造出一种生产技术，使得这种技术能够生成我们观测到的企业选择。在这个意义上，任何观测到的选择只要它们能满足利润最大化弱公理，则它们就是利润最大化的选择。随着我们的观测数据越多，我们对企业的生产函数估计越准确，如图 19.5 所示。

[一] 就是这两条等利润线与坐标轴围成区域的交集。注意，该交集（生产集）包含等利润线形成的边界。译者注。

我们估计出的这种生产函数可以用于预测企业在其他环境下的行为，或者用于其它经济分析。

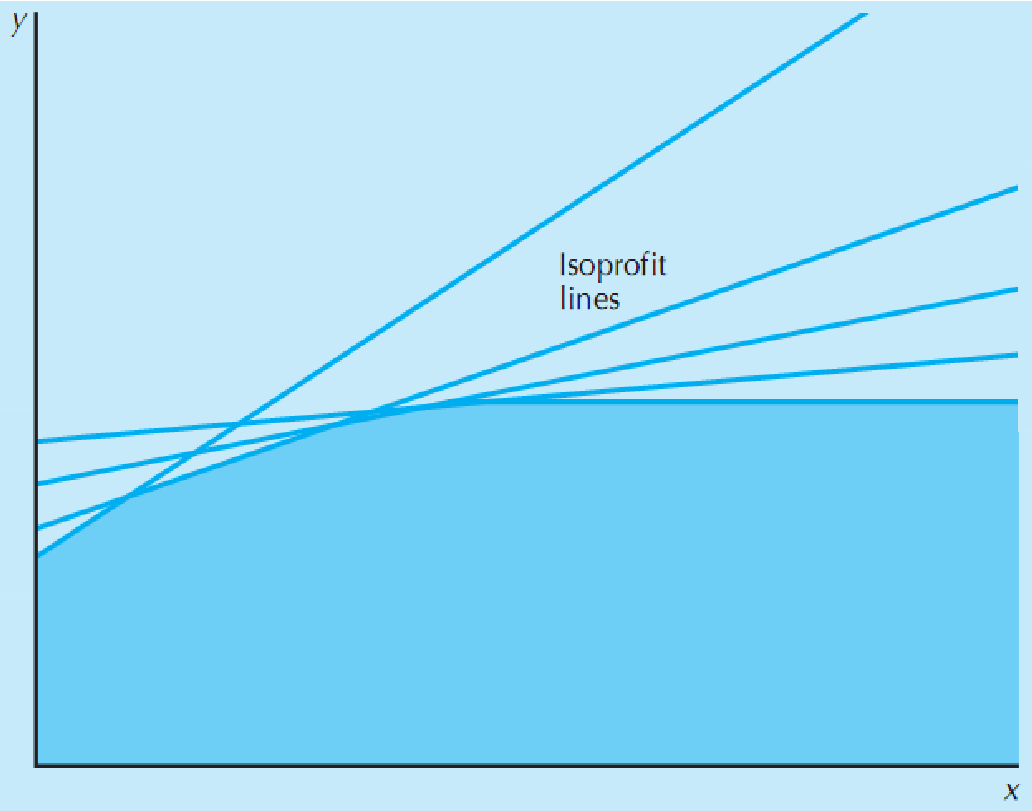


图 19.5：估计生产技术。当我们观测到的选择数据越多，我们对生产函数的估计就越准确。

例子：农场主对价格补贴的反应

当前，美国政府每年用于资助农场主的资金大约为 400 亿~600 亿美元。这笔资金中的很大一部分用于补贴各类农产品的生产，包括牛奶、小麦、玉米、大豆和棉花等。政府有时也试图减少或取消这些补贴。取消补贴将会使农场主得到的产品价格减少。

有时人们认为取消对牛奶之类产品的补贴不会减少牛奶的总供给，他们的理由是，补贴取消后，为了维持生活水平不变，奶农会**增加**奶牛的数量，从而增加牛奶的供给。

上述观点正确吗？不正确。如果奶农是追求利润最大化的，那么取消补贴后，牛奶的供给很可能减少。前文的比较静态结果告诉我们，利润最大化的内在逻辑**要求**：产品价格下降后，产品的供给必然因此减少。所以，如果 Δp 为负，则 Δy 必然也为负。

当然，家庭小农场的目标可能不是利润最大化，但是大型农业企业的目标却很可能是利润最大化的。因此，取消补贴后，奶农增加牛奶供给这种反常反应即使有，也不多见。因此，取消补贴后，牛奶供给很可能下降。

19.12 成本最小化

如果某企业是追求利润最大化的，而且如果他选择供给 y 单位产品，则它必须使得生产 y 单位产品的成本最小。如果成本没达到最小，这意味着还有其他更好的生产方法能这么多产品，这表明企业前一种生产方法没有实现利润最大化。

这个简单的结论在分析企业行为时非常有用。我们通常将利润最大化问题分解为两步：第一步，对于任何既定的产量 y，我们要找出使生产成本最小的方法；第二步，我们找出那种产量水平是利润最大化水平。我们将在下一章解决这个任务。

附录

企业的利润最大化问题为

$$\max_{x_1,x_2} \ pf(x_1,x_2)-w_1x_1-w_2x_2,$$

它的一阶条件为

$$p\frac{\partial f(x_1^*,x_2^*)}{\partial x_1}-w_1=0$$
$$p\frac{\partial f(x_1^*,x_2^*)}{\partial x_2}-w_2=0$$

这些条件和课文中得到的边际产量条件是一样的。下面我们看看对于柯布-道格拉斯生产函数来说，利润最大化行为是什么样子的。

假设柯布-道格拉斯函数为 $f(x_1,x_2)=x_1^a x_2^b$。则上述一阶条件变为

$$pax_1^{a-1}x_2^b-w_1=0$$
$$pbx_1^a x_2^{b-1}-w_2=0$$

将上面第一个式子乘以 x_1，将第二个式子乘以 x_2 可得

$$pax_1^a x_2^b-w_1x_1=0$$
$$pbx_1^a x_2^b-w_2x_2=0$$

使用 $y=x_1^a x_2^b$ 表示该企业的产量水平，我们可以将上面的表达式改写为

$$pay=w_1x_1$$
$$pby=w_2x_2$$

解出 x_1 和 x_2 可得

$$x_1^* = \frac{apy}{w_1}$$

$$x_2^* = \frac{bpy}{w_2}$$

由此，我们得到了要素 1 和 2 的需求函数，要素需求是最优产量选择 y 的函数。但是，我们还必须解出最优的产量。将最优要素需求量代入到柯布-道格拉斯生产函数可得

$$(\frac{apy}{w_1})^a(\frac{bpy}{w_2})^b = y.$$

将上式左侧中的 y 提取出来可得

$$(\frac{ap}{w_1})^a(\frac{bp}{w_2})^b y^{a+b} = y.$$

或

$$y = (\frac{ap}{w_1})^{\frac{a}{1-a-b}}(\frac{bp}{w_2})^{\frac{b}{1-a-b}}.$$

这样我们就得到了柯布-道格拉斯企业的供给函数。它和上面已经推导出来的要素需求函数一起，构成了利润最大化问题的完整解。

注意，如果企业的技术是规模报酬不变的（即 $a+b=1$ 时），供给函数无定义，因为此时上面供给函数右侧的两个指数的分母都等于 0。

只要产出和投入的价格和零利润一致，拥有柯布-道格拉斯技术的企业就不会关注具体的产量大小。这是由于此时多生产和少生产的利润都为零。

总结

1.利润等于收入减去成本。在这个定义中要记住，所有成本必须以合适的市场价格衡量。

2.生产要素如果它的使用量和产量无关，则称为固定要素；如果它的使用量随产量水平改变而改变，则称为可变要素。

3.在短期，某些生产要素的数量是事先确定的，因此无法改变；而在长期，所有要素均可自由变动。

4.如果企业是追求利润最大化的，则每种可变要素的边际产量价值必须等于该要素自身的价格。

5.利润最大化的逻辑意味着：一家竞争性企业的供给函数，必定是产品价格的增函数；该企业对每种要素的需求函数必然是该要素自身价格的减函数。

6.如果一家竞争性企业的生产技术是规模报酬不变的，那么它的长期最大利润必定为零。

复习题

1.在短期，如果固定要素的价格上升，利润将如何变动？

2.如果某企业的生产技术能做到对于任何产量水平都是规模报酬递增的。现在如果产品和所有要素的价格维持不变，但生产规模扩大为原来的 2 倍，它的利润将如何变化？

3. 如果某企业的生产技术能做到对于任何产量水平都是规模报酬递减的。现在该企业将自己一分为二，即设立两个规模相等的小企业，它的总利润将如何变化？

4.一个园丁大声说道："投入 1 元钱的种子，我就收获了 20 元产出！"某经济学家对此表示怀疑，他首先注意到这些产出只是价格便宜的南瓜，他还能指出这位园丁忽略了什么东西吗？

5.企业的利润最大化目标总是与企业股票市场价值最大化目标相一致吗?

6.如果 $pMP_1 > w_1$，为了增加利润，企业会增加还是减少要素 1 的投入量？

7.假设某企业追求短期利润最大化，而且已知 x_1 是可变要素，x_2 是固定要素。如果 x_2 的价格下降，x_1 的使用量将怎样变化？企业的利润水平将怎样变化？

8.选择题。一家追求利润最大化的竞争性企业，如果在长期均衡时利润为正，那么该企业（可能/不可能）使用的是规模报酬不变的技术。

复习题答案

1.在短期，如果固定要素的价格上升，利润将如何变动？

【复习内容】短期利润最大化；比较静态分析

【参考答案】

解法一:微积分

假设要素 2 是固定要素，它的数量固定在$\bar{x}_2$，企业的生产函数为$f(x_1,x_2)$，产品价格为p，要素 1 和 2 的价格分别为w_1和w_2。则企业的短期利润最大化问题可以写为

$$\max_{x_1} \quad pf(x_1,\bar{x}_2)-w_1x_1-w_2\bar{x}_2.$$

该问题的一阶条件为

$$pMP_1(x_1^*,\bar{x}_2)=w_1 \tag{1}$$

此时的x_1^*就是企业短期利润最大化时，要素 1 的最优使用量。

这样，企业的最大利润为

$$pf(x_1^*,\bar{x}_2)-w_1x_1^*-w_2\bar{x}_2 \tag{2}$$

由（1）式可知，要素 1 的最优数量取决于以下因素：产品价格p；要素 1 本身的价格w_1；要素 2 的数量$\bar{x}_2$以及生产函数的具体表达式。但是注意，**在短期，可变要素 1 的最优数量和固定要素 2 的价格无关**。

现在令要素 2（固定要素）的价格上升，仔细观察（2）式，由上面的结论可知，（2）式中的第一项$pf(x_1^*,\bar{x}_2)$和第二项$w_1x_1^*$在要素 2 价格上升后无任何变化。现在看第三项$w_2\bar{x}_2$，这一项中$\bar{x}_2$无法改变，w_2变大。因此，要素 2 价格上升后，企业的利润将减少。

解法二：几何法

假设要素 2 是固定要素，它的数量固定在$\bar{x}_2$，企业的生产函数为$f(x_1,x_2)$，产品价格为p，要素 1 和 2 的价格分别为w_1和w_2。

请看下图。图中的曲线代表要素 2 的数量固定为$\bar{x}_2$时的生产函数。令y表示企业的产量，利润表达式为

$$\pi=py-w_1x_1-w_2\bar{x}_2.$$

由此解出y，即将产量y表达为x_1的函数

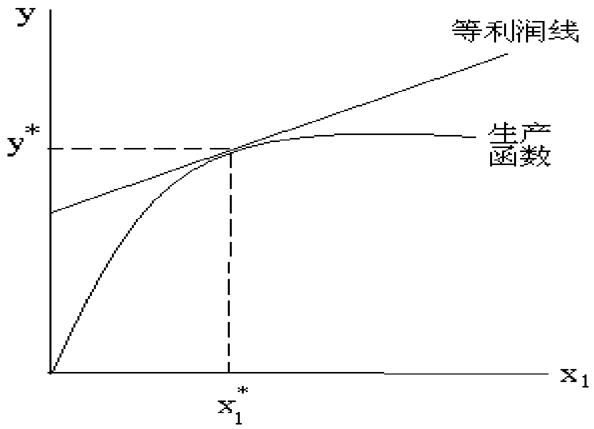
$$y=\frac{\pi}{p}+\frac{w_2}{p}\bar{x}_2+\frac{w_1}{p}x_1. \tag{3}$$

这个式子是等利润线的表达式。它刻画的是能带来固定利润水平π的投入和产出的所有组合。由（3）知，等利润线的斜率为w_1/p；纵截距都为$\pi/p+w_2\bar{x}_2/p$，纵截距衡量的是利润与固定成本的和。当然，此处的利润和固定成本是进过价格调整后（除以p）的实际利润和实际固定成本。

固定成本是固定的，因此当我们从一条等利润线移动到另外一条等利润线时，唯一变动的是利润水平。因此等利润线的纵截距越大，代表的利润水平越高。因此，利润最大化问题转化为在生产函数曲线上找到一点，使得这点位于位置最高的等利润线上。下图画出了这样的点。和以前一样，可用相切条件刻画这样的点：生产函数的斜率应该等于等利润线的斜率。

由于生产函数的斜率是边际产量，等利润线的斜率为 w_1/p，相切条件因此可以写为

$$MP_1(x_1^*, \bar{x}_2) = \frac{w_1}{p}\text{。}$$



企业选择的投入和产出组合位于最高的等利润线上。在这种情形下，利润最大化的点为 (x_1^*, y^*)。

我们分析要素 2（固定要素）的价格上升时的效应。由于这是短期分析，要素 2 的价格变动不会改变企业对要素 2 的使用量——在短期，要素 2 的使用量固定在 $\bar{x}_2$ 水平。要素 2 的价格改变对等利润线的斜率无任何影响，因为等利润线的斜率 w_1/p 和要素 2 的价格无关。因此，要素 1 的最优选择将不会改变，产品的供给数量也不会改变。唯一改变的是企业的利润。由于要素 2 的价格上升了，因此企业的利润减少了。

2.如果某企业的生产技术能做到对于任何产量水平都是规模报酬递增的。现在如果产品和所有要素的价格维持不变，但生产规模扩大为原来的 2 倍，它的利润将如何变化？

【复习内容】利润与规模报酬；规模报酬递增

【参考答案】

假设该企业的生产函数为 $y = f(x_1, x_2)$，要素投入量为 (x_1, x_2)。

该企业的利润为

$$\pi = pf(x_1, x_2) - w_1x_1 - w_2x_2\,.$$

由于该企业的生产技术是处处规模报酬递增的，又知道要素价格和产品价格都不变，因此对于任意 $t > 1$， $f(tx_1, tx_2) > tf(x_1, x_2)$，从而新的利润

$$\pi^* = pf(tx_1, tx_2) - w_1(tx_1) - w_2(tx_2) > tf(x_1, x_2) - tw_1x_1 - tw_2x_2 = t\pi$$

由题目知，$t = 2$，代入上式可得 $\pi^* > 2\pi$，所以新利润比原来利润的两倍还大。

 曹乾(东南大学 caoqianseu@163.com)

3. 如果某企业的生产技术能做到对于任何产量水平都是规模报酬递减的。现在该企业将自己一分为二，即设立两个规模相等的小企业，它的总利润将如何变化？

【复习内容】利润与规模报酬；规模报酬递减

【参考答案】总利润会增加

注意，我们从分拆后的小企业入手分析，假设一个小企业的生产函数为 $y = f(x_1, x_2)$，要素投入量为 (x_1, x_2)。

一个小企业的利润为

$$\pi = pf(x_1, x_2) - w_1 x_1 - w_2 x_2 .$$

由于该小企业的生产技术是处处规模报酬递减的，由出题人的意图可知要素价格和产品价格都不变，因此对于任意 $t > 1$，$f(tx_1, tx_2) < tf(x_1, x_2)$，从而新的利润

$$\pi^* = pf(tx_1, tx_2) - w_1(tx_1) - w_2(tx_2) < tf(x_1, x_2) - tw_1 x_1 - tw_2 x_2 = t\pi$$

由题目知，将这样的小企业再复制一个，即 $t = 2$ 时，上式意味着 $\pi^* < 2\pi$。也就是说两个小企业带来的利润小于原来利润的 2 倍。换个角度看，这正好是说将这家大企业分拆为两个小企业后，利润会增加。

4.一个园丁大声说道："只投入 1 元钱的种子，我就收获了 20 元产出！"某经济学家对此表示怀疑，他首先注意到这些产出只是价格便宜的南瓜，他还能指出这位园丁忽略了什么东西吗？

【复习内容】经济成本和会计成本的区别；机会成本

利润的定义是收入减去成本。必须确保将企业使用的所有生产要素都包括进来，用市场价格计价。

利润的经济学定义要求我们，所有投入品和产出品都应该按它们的机会成本计价。

【参考答案】

经济学家怀疑的原因是，此园丁的产出投入比高达 20:1。这对农业来说简直是不可能的。因此肯定是某些事情出现了问题。

这位园丁忽略了他种瓜的机会成本，准确地说，他没有将自身投入劳动的价值计为成本。因为他种瓜，那么他自己的劳动就是一种投入，需要包含在成本内。他的工资率就是他的劳动的市场价格——如果他在市场上出售自己的劳动所能得到的工资率。因此如果将此计入成本，则他的产出投入比没有那么大。

5.企业的利润最大化目标总是与企业股票市场价值最大化目标相一致吗?

【复习内容】利润和股票市场价值

【参考答案】通常不一致，因为企业的利润流具有不确定性。

大多数大型企业是公司制企业，就是说这样的企业为多人所共有。公司发行的股票代表对公司所有权的拥有份额。在某些时间，公司会对股份派发股息，股息是公司利润的一部分。股票可以在股票市场买卖。股票价格代表人们期望从公司获得的股息流的现值。某企业的股票市场总值代表企业预期利润流的现值。因此，**企业的目标——使企业利润流的现值最大化——也可以说成使股票市场价值最大化。在确定性的世界中，这两个目标是同一个东西**。

如果企业的利润流具有不确定性，那么企业所有者要求经营者实现利润最大化就没有明确的意义。他们应该使期望利润最大化吗？他们应该使利润的期望效用最大化吗？经营者对待风险投资应持有什么样的态度？因此，在不确定性的情形下，很难说清利润最大化是什么意思。然而，即使在不确定性的情形下，使股票市场价值最大化仍然有意义。如果企业的经营者试图使得公司的股票价值尽可能地大，那么他们就会使公司所有人（所有股东）的状况尽可能地好。因此在几乎所有的经济环境中，使股票市场价值最大化对企业来说是一个定义清晰的目标。

6.如果 $pMP_1 > w_1$，为了增加利润，企业会增加还是减少要素 1 的投入量?

【复习内容】利润最大化的均衡条件

企业的利润最大化问题为

$$\max_{x_1,x_2} \ pf(x_1,x_2)-w_1x_1-w_2x_2,$$

它的一阶条件为

$$p\frac{\partial f(x_1^*,x_2^*)}{\partial x_1}-w_1=0$$
$$p\frac{\partial f(x_1^*,x_2^*)}{\partial x_2}-w_2=0$$

即 $pMP_1=w_1$ 和 $pMP_2=w_2$。

【参考答案】

$$pMP_1>w_1 \Leftrightarrow pMP_1-w_1>0 \Leftrightarrow pMP_1\Delta x_1-w_1\Delta x_1>0$$

上面最后一个式子表示，稍微增加一些要素 1 的投入量利润会增加，因此企业会增加要素 1 的投入量。

7.假设某企业追求短期利润最大化，而且已知 x_1 是可变要素， x_2 是固定要素。如果 x_2 的价格下降， x_1 的使用量将怎样变化？企业的利润水平将怎样变化？

【复习内容】短期利润最大化；比较静态分析

【参考答案】

本题和第 1 题完全相同，请参考第 1 题的答案。

此处只给出最后的答案： x_1 的使用量不变，企业的利润将增加。

8.选择题。一家追求利润最大化的竞争性企业，如果在长期均衡时利润为正，那么该企业（可能/不可能）使用的是规模报酬不变的技术。

【复习内容】利润最大化与规模报酬

【参考答案】不可能

假设某企业已经选择出长期利润最大化的产出 $y^*=f(x_1^*,x_2^*)$ ，此时要素投入量为 (x_1^*,x_2^*) 。

该企业的利润为

$$\pi^*=py^*-w_1x_1^*-w_2x_2^*.$$

由题目知，该企业长期最大化利润为正。不妨假设该企业的生产技术是规模报酬不变的。现在如果我们将所有要素的投入量变为原来的 2 倍，结果将如何？根据规模报酬不变的假设可知，产量也会变为原来的 2 倍。利润也会变为原来的 2 倍。矛盾出现了！因为我们在前面已假设它的原来选择已实现利润最大化，这样利润就不可能再增大。矛盾出现的原因是我们假设企业的规模报酬不变。

由此可见，企业不可能使用的是规模报酬不变的技术。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:
A Modern Approach (8^{th} Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

20.成本最小化（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

20 成本最小化

我们的任务是研究竞争性和非竞争性市场环境下的企业的利润最大化行为。上一章我们分析了竞争性市场下企业的利润最大化行为，在那里我们使用的研究方法是直接分析利润最大化问题。

然而，我们还可以使用间接的方法进行研究，这样做的好处是能获得一些重要的思想。我们的策略是将利润最大化问题分解为两个部分。在第一部分，我们将分析下面这样的问题：给定任一产量水平，企业如何使成本最小；在第二部分，我们研究企业如何选择利润最大化的产量水平。在本章，我们分析第一部分——给定任一产量水平企业如何使成本最小。

20.1 成本最小化

假设我们有两种生产要素，价格分别为 w_1 和 w_2。对于既定的产量水平 y，我们希望找到成本最小的生产方法。令 x_1 和 x_2 分别表示两种要素的投入量，令 $f(x_1,x_2)$ 表示企业的生产函数，我们可以将此问题写为

$$\min_{x_1,x_2} \; w_1x_1 + w_2x_2$$

$$\text{使得 } f(x_1,x_2) = y.$$

我们再次强调：做此类分析时，你务必确保已经将**所有**生产成本囊括进来，还要确保所有数量都用统一的时间刻度衡量。

这个成本最小化的解——实现既定产量的最小成本——取决于 w_1，w_2 和 y，因此我们将其写为 $c(w_1,w_2,y)$。这个函数称为**成本函数**(cost function)，它很有用。成本函数 $c(w_1,w_2,y)$ 表示当要素价格为 (w_1,w_2) 时，生产 y 单位产品需要的**最小**成本。

为了理解成本函数，我们将企业的成本和技术约束画在同一张图上。等产量线表示企业的技术约束——能够生产出 y 单位产品的 x_1 和 x_2 的所有组合。

假设我们想画出投入的组合，使得这些投入组合的成本等于既定的成本水平 C。我们可以将其写为

$$w_1x_1 + w_2x_2 = C,$$

整理可得

$$x_2 = \frac{C}{w_2} - \frac{w_1}{w_2}x_1.$$

容易看出，这是一条直线，它的斜率为 $-w_1/w_2$，纵截距为 C/w_2。令 C 变动就可得到一系

曹乾（东南大学 caoqianseu@163.com）

列**等成本线**（isocost lines）。一条等成本线上的所有点的成本都是相同的，位置更高的等成本线表示成本更高。

因此，成本最小化问题可以改写为：找到等产量线上的一点，使得它位于位置最低的等成本线上。如图 20.1 所示。

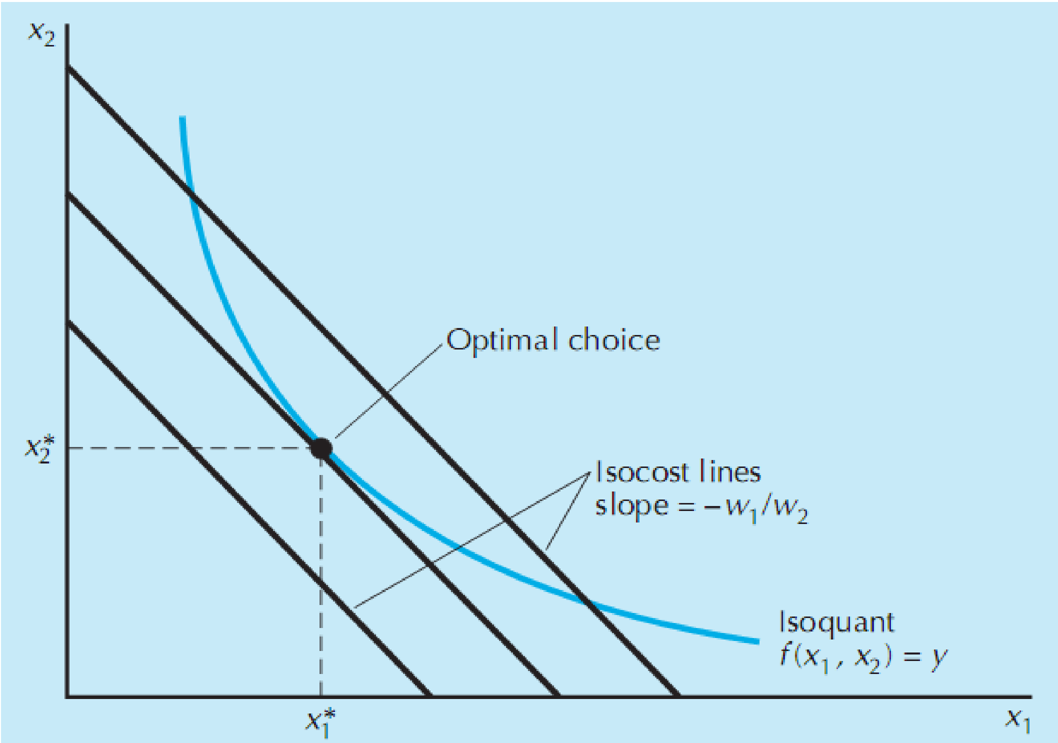


图 20.1：成本最小化。找到等产量线上的一点，使得它同时位于位置最低的等成本线上。这个点就是使生产成本最小的要素组合。

注意，如果最优解中每种要素的数量都大于零，而且如果等产量线是良好的平滑曲线[一]，那么成本最小化的点可用相切条件刻画：等产量线的斜率必定等于等成本线的斜率。或者，用第 18 章中学过的术语表达——**技术替代率（TRS）必定等于要素价格之比**：

$$-\frac{MP_1(x_1^*, x_2^*)}{MP_2(x_1^*, x_2^*)} = TRS(x_1^*, x_2^*) = -\frac{w_1}{w_2} \tag{20.1}$$

[如果最优解为角点解（boundary solution），即其中一种要素的使用量为零，则相切条件就不再适用。类似地，如果生产函数有“折弯”（kinks），则相切条件无意义，因为在折弯处导数不存在。这些例外和消费者理论中最优解的例外是一样的，因此，在本章我们不再强调这些情形。]

[一] 前半句话的意思是说最优解为内部解；后半句的意思是要求生产函数是可微分的。译者注。

用代数方法推导（20.1）式并不难。考虑生产模式的任何变动$(\Delta x_1, \Delta x_2)$，如果这些变动能使产量不变，则必然有

$$MP_1(x_1^*, x_2^*)\Delta x_1 + MP_2(x_1^*, x_2^*)\Delta x_2 = 0. \tag{20.2}$$

注意，Δx_1和Δx_2的符号必然相反；因为如果你增加要素 1 的投入量，那么你必须减少要素 2 的投入量才能使产量不变。

如果我们已达到成本最小，则该变动不可能降低成本，因此必有

$$w_1\Delta x_1 + w_2\Delta x_2 \geq 0 \tag{20.3}$$

现在考虑变动$(-\Delta x_1, -\Delta x_2)$，要求它也能使产量不变。与上面的理由一样，该变动也不可能降低成本。这意味着

$$-w_1\Delta x_1 - w_2\Delta x_2 \geq 0 \tag{20.4}$$

由（20.3）式和（20.4）式可知此时必有[一]

$$w_1\Delta x_1 + w_2\Delta x_2 = 0 \tag{20.5}$$

联立（20.2）式和（20.5）式，解出$\Delta x_1 / \Delta x_2$可得

$$\frac{\Delta x_2}{\Delta x_1} = -\frac{w_1}{w_2} = -\frac{MP_1(x_1^*, x_2^*)}{MP_2(x_1^*, x_2^*)},$$

这个式子正是前面用几何方法推导出的成本最小化的条件。

注意，图 20.1 和消费者选择理论中的最优解的图形有几分相似。尽管它们的最优解类似，但它们不是同一类问题。在消费者理论中，直线是预算线，消费者沿着预算线寻找他最喜欢的消费组合。在生产者理论中，等产量线是技术约束，生产者沿着等产量线寻找最优的要素投入组合。

使企业成本最小的要素选择，通常取决于所有要素的价格和企业想要生产的产量水平，所以我们可以将要素选择写为$x_1(w_1, w_2, y)$和$x_2(w_1, w_2, y)$。这称为**有附加条件的要素需求函数**（conditional factor demand functions）或**派生的要素需求函数**（derived factor demand functions）[二]。这些函数衡量在企业产量y既定的**条件**下，要素需求与要素价格及其产量之间的关系。

需要注意，**有附加条件的要素需求**和上一章介绍的利润最大化的要素需求之间的区别。有附加条件的要素需求是既定产量水平下的成本最小化的选择；而利润最大化的要素需求则是既定产品**价格**下的利润最大化的选择。

[一] 这个结论源自一个简单的事实：如果一个数和它的相反数都≥ 0，则这个数必为零。译者注。

[二] “有附加条件的要素需求函数”中的修饰语“有附加条件的”就是指“产量既定”，提醒你**产量是该要素需求函数的自变量**。起这样的名字只是为了和上一章中的“（利润最大化的）要素需求函数”区别开。译者注。

有附加条件的要素需求通常不可直接观测到；它是个假设出来的工具。它表示，如果企业希望使既定产量的成本最小，每种要素的最优投入量将是多少。有附加条件的要素需求是个有用的工具，我们可以使用它分析最有效率的生产方法问题，而用利润最大化要素需求去分析最优产量的决定问题。

例子：特别生产技术的成本最小化

假设生产技术中要素是完全互补的，因此 $f(x_1,x_2)=\min\{x_1,x_2\}$。现在如果我们想生产 y 单位产品，显然需要 y 单位的 x_1 和 y 单位的 x_2。

因此，最小的生产成本为

$$c(w_1,w_2,y)=w_1y+w_2y=(w_1+w_2)y.$$

如果生产技术是完全替代类型的，即 $f(x_1,x_2)=x_1+x_2$，结果是什么样的？既然要素 1 和 2 在生产过程中是完全替代的，显然哪种要素价格最低，企业就会使用哪种要素。因此，生产 y 单位产品的最小成本是 w_1y 和 w_2y 中最小的那个。换句话说：

$$c(w_1,w_2,y)=\min\{w_1y,\ w_2y\}=\min\{w_1,w_2\}y.$$

最后，看看柯布-道格拉斯技术，它的表达式为 $f(x_1,x_2)=x_1^ax_2^b$。在这种情形下，我们可以使用微积分推导出成本函数：

$$c(w_1,w_2,y)=Kw_1^{\frac{a}{a+b}}w_2^{\frac{b}{a+b}}y^{\frac{1}{a+b}},$$

其中 K 是一个取决于 a 和 b 的常数。具体推导过程请见附录。

20.2 显示的成本最小化

前面我们假设企业选择要素投入量使得成本最小化。根据这个假设可以推知：观测到的选择是如何随要素价格变动而变动的。

假设我们观测到两组价格 (w_1^t,w_2^t) 和 (w_1^s,w_2^s) 和相应的要素选择 (x_1^t,x_2^t) 和 (x_1^s,x_2^s)。假设这两个选择生产相同的产量 y。如果每种选择在相应价格下都是成本最小化的选择，则必然有

$$w_1^tx_1^t+w_2^tx_2^t\le w_1^tx_1^s+w_2^tx_2^s$$

和

$$w_1^sx_1^s+w_2^sx_2^s\le w_1^sx_1^t+w_2^sx_2^t.$$

如果企业在生产 y 单位产品时总是选择使成本最小的方法，则该企业在时期 t 和时期 s 的选择必然满足这些不等式。我们将这些不等式称为**成本最小化弱公理**（Weak Axiom of Cost

Minimization, WACM）。

将第二个式子改写为

$$-w_1^s x_1^t - w_2^s x_2^t \le -w_1^s x_1^s - w_2^s x_2^s$$

然后加到第一个式子得到

$$(w_1^t - w_1^s)x_1^t + (w_2^t - w_2^s)x_2^t \le (w_1^t - w_1^s)x_1^s + (w_2^t - w_2^s)x_2^s,$$

整理可得

$$(w_1^t - w_1^s)(x_1^t - x_1^s) + (w_2^t - w_2^s)(x_2^t - x_2^s) \le 0.$$

使用Δ符号表示要素需求量的变动和要素价格变动，可把上式写得更紧凑些

$$\Delta w_1 \Delta x_1 + \Delta w_2 \Delta x_2 \le 0.$$

我们只使用成本最小化行为的假设就推导出了上述不等式。它蕴含着对企业行为的限制——当要素价格改变但产品价格不变时企业行为应该怎样变化。

例如，若要素 1 的价格上升要素 2 的价格不变即$\Delta w_2 = 0$，因此上述不等式变为

$$\Delta w_1 \Delta x_1 \le 0.$$

如果要素 1 的价格上升，则这个不等式意味着要素 1 的需求必然下降；因此有附加条件的要素需求函数必定向下倾斜。

如果生产者问题中的参数发生变化，最小成本应该怎样变化？容易看出，如果任何一种要素价格上升则最小成本必上升：如果一种要素变得更贵而另外一种要素的价格不变，最小成本不可能下降，一般会上升；类似地，如果企业决定生产更多的产量而且所有要素价格维持不变，则企业的成本必然上升。

20.3 规模报酬和成本函数

在第 18 章我们介绍了生产函数的规模报酬概念。我们已知道：若对于所有$t>1$，$f(tx_1, tx_2)$大于、小于或者等于$tf(x_1, x_2)$，则该生产技术是规模报酬递增、递减或不变的。可以证明，生产函数的规模报酬类型和成本函数的行为之间存在着一个重要关系。

首先假设规模报酬不变。假若我们已经解决了生产 1 单位产量的成本最小化问题，因此我们得到了**单位成本函数**（unit cost function）——$c(w_1, w_2, 1)$。现在生产y单位产品的成本最小的方法是什么？答案很简单：将生产 1 单位产品的每种要素使用量乘以y。也就是说生产y单位产品的最小成本就是$c(w_1, w_2, 1)y$。由此可见，在规模报酬不变情形下，成本函数是产量的线性函数。

如果规模报酬递增，结果是什么样的？在这种情形下，成本增加幅度小于产量增加幅

曹乾（东南大学 caoqianseu@163.com）

度。如果企业决定生产的产量为原来的 2 倍，则成本小于原来的 2 倍，只要所有要素价格维持不变。这是规模报酬概念的自然应用：如果企业将要素投入变为原来的 2 倍，则产量大于原来的 2 倍。因此如果我们想生产的产量恰好为原来的 2 倍，规模报酬递增意味着要素投入小于原来的 2 倍就能做到这一点。

但是要素投入量变为原来的 2 倍，意味着成本恰好也为原来的 2 倍。因此，如果要素投入量小于原来的 2 倍，则成本也小于原来的 2 倍：这就是说成本增加幅度小于产量增加幅度。

类似地，如果生产技术是规模报酬递减的，成本增加幅度将大于产量增加幅度。如果产量变为原来的 2 倍，则成本会大于原来的 2 倍。

这些事实可用**平均成本函数**（average cost function）的行为表示。平均成本函数是产量为 y 时的每**单位**产品的成本。

$$AC(y)=\frac{c(w_1,w_2,y)}{y}.$$

如果技术是规模报酬不变的，我们在前面已知道该情形的成本函数形式为

$$c(w_1,w_2,y)=c(w_1,w_2,1)y$$

因此，平均成本函数为

$$AC(w_1,w_2,y)=\frac{c(w_1,w_2,y)}{y}=c(w_1,w_2,1).$$

也就是说，每单位产品的成本将是常数，无论企业想生产的产量为多大。

如果技术是规模报酬递增的，则成本上升幅度将小于产量增加幅度，因此平均成本将会产量的增加而下降。类似地，如果及时是规模报酬递减的，则平均成本将随产量的增加而增加。

我们在前面已经知道，一项既定技术可能同时含有规模报酬递增、不变或递减的**区域**——在不同产量水平上，产量的增长速度可能快于、等于或慢于企业经营规模的增长速度。类似地，在不同产量水平，成本增长速度可能快于、等于或慢于产量增长速度。这意味着在不同产量水平上，平均成本可能下降、保持不变或上升。在下一章我们将详细分析这些可能性。

从现在起，我们主要关注成本和产量之间的关系。在大多数情形下，我们假设要素价格固定在事先决定的水平上，因此可认为成本仅取决于企业的产量决策。所以，从此以后我们将生产函数视为产量这个唯一自变量的函数：$c(y)$。

20.4 长期成本和短期成本

我们将成本函数定义为生产既定产量时的最小成本。通常有必要区分以下两种成本：一是企业在能调整所有生产要素情形下的最小成本；二是企业只能调整部分生产要素情形下的最小成本。

我们将短期定义为某些生产要素使用量固定不变的时期。在长期，所有要素可以自由变动。**短期成本函数**（short-run cost function）定义为只能调整可变要素的投入量时，生产既定产量的最小成本；**长期成本函数**（long-run cost function）是指，在所有生产要素均可调整的情形下，生产既定产量的最小成本。

假设在短期，要素 2 的投入量固定为 $\bar{x}_2$，但在长期它可以自由变动。则短期成本函数的定义为

$$c_s(y,\bar{x}_2)=\min_{x_1} w_1x_1+w_2\bar{x}_2$$

$$\text{使得 } f(x_1,\bar{x}_2)=y.$$

注意，一般来说，在短期生产 y 单位产品的最小成本取决于固定要素的使用数量和成本大小。

在只有两种要素的情形下，这个最小化问题不难求解：我们找到满足 $f(x_1,\bar{x}_2)=y$ 的 x_1 的最小值即可。然而，如果短期内可变要素的种类很多，成本最小化问题就会涉及大量的计算。

要素 1 的短期需求函数是使成本最小的要素 1 的使用量。一般来说，除了取决于产量之外，它还取决于所有要素的价格和固定要素的使用量，因此我们将短期要素需求写为

$$x_1=x_1^s(w_1,w_2,\bar{x}_2,y)$$

$$x_2=\bar{x}_2$$

举个例子说明这两个式子的意思。例如，如果工厂规模在短期内既定（$x_2=\bar{x}_2$），则在任何给定的一组要素价格和产量水平 (w_1,w_2,y) 上，企业想要雇用工人的数量（x_1）通常取决于工厂规模（$\bar{x}_2$）。

注意，根据短期成本函数的定义

$$c_s(y,\bar{x}_2)=w_1x_1^s(w_1,w_2,\bar{x}_2,y)+w_2\bar{x}_2.$$

这个式子是说，产量为 y 时的最小成本取决于成本最小的要素投入选择决策。这个事实是由短期成本函数的定义推知，它很有用。

这个例子中，长期成本函数的定义为

$$c(y)=\min_{x_1,x_2} w_1x_1+w_2x_2$$

$$\text{使得 } f(x_1,x_2)=y.$$

此处两种要素都可以自由变动。长期成本仅取决于在要素价格给定情形下企业想要生产的产量。我们将长期成本函数写为 $c(y)$，将长期要素需求写为

$$x_1 = x_1(w_1, w_2, y)$$

$$x_2 = x_2(w_1, w_2, y).$$

我们也可以将长期成本函数写为

$$c(y) = w_1 x_1(w_1, w_2, y) + w_2 x_2(w_1, w_2, y).$$

和短期成本函数的定义类似，这个式子是说，长期中，产量为 y 时的最小成本等于企业使用成本最小化的要素选择所带来的成本。

短期成本函数和长期成本函数之间有一个有趣的关系，我们在下一章将使用这个关系。现在说说这个关系是什么。为简单起见，令要素价格固定在预先决定的水平上，将长期要素需求写为

$$x_1 = x_1(y)$$

$$x_2 = x_2(y).$$

于是，长期成本函数也可以写成

$$c(y) = c_s(y, x_2(y)).$$

为了看清这一点，请想想这个式子的意思。这个式子的左侧是企业产量为 y 时的长期成本，它是所有要素均可变时，生产 y 单位产品的最小成本。这个式子的右侧是企业产量为 y 时的短期成本，它是要素 2 的数量固定时，生产 y 单位产品的最小成本。现在，如果短期情形下生产 y 单位产品时要素 2 的使用量，恰好就是长期生产 y 单位产品时要素 2 的使用量。那么上式左右两侧显然相等。根据这个式子可知，可变要素的长期需求（成本最小化的要素选择）为

$$x_1(w_1, w_2, y) = x_1^s(w_1, w_2, x_2(y), y).$$

这个式子是说，长期内使成本最小化的可变要素的使用量，等于企业在短期内对可变要素的使用量——前提是固定要素的使用量恰好能使长期成本最小。

20.5 固定成本和准固定成本

在第 19 章，我们区分了固定要素和准固定要素。固定要素是指企业无论是否生产都必须支付成本的要素。准固定要素是指企业的产量为正时才需要支付成本的要素。

自然地，可以类似定义固定成本和准固定成本。**固定成本**（fixed costs）是为固定要素支出的成本：这样的成本和产量无关，无论企业是否生产都必须支付。**准固定成本**

（quasi-fixed costs）是企业为准固定要素支付的成本，这样的成本也和产量多少无关，但是只要企业的产量为正就必须支付这样的成本。

根据长期的定义，在长期不存在固定成本。然而在长期很可能存在准固定成本。只要企业在开始生产前需要支付资金，则准固定成本就产生了。

20.6 沉没成本

沉没成本是另外一种固定成本。我们举例说明沉没成本的含义。假设你决定承租一办公场所，租期一年。你承诺支付的月租金是一种固定成本，因为不论你的产量如何你必须支付这笔资金。现在假设你要对办公场所进行简单装修，粉刷一下并买些办公家具。粉刷成本是一种固定成本，但它同时也是一种**沉没成本**（sunk cost），因为这种成本在支出后就不能被收回。但是，购买家具的成本并非完全是沉没成本，因为你不想继续使用这些家具时可以将其卖掉。对于家具来说，沉没的成本是买进和卖出的**差价**。

为了更详细地说明，我们用数字举例。假设年初你借入 20,000 元资金，年利率为 10%。根据办公场所租赁合同，你事先为下一年支付了 12,000 元的租金。购买家具和粉刷费用分别为 6,000 元和 2,000 元。在年终，你偿还 20,000 元贷款本金和 2,000 元利息，并且你卖掉了家具，卖得 5,000 元。

你的总沉没成本包括：12,000 元的租金，2,000 元的利息，2,000 元的粉刷费用，以及 1,000 元的家具成本。注意，家具的沉没成本不是 6,000 元，因为你卖掉家具收回了 5,000 元。

沉没成本和可收回成本（recoverable costs）之间的差额可能非常大。例如，购买 5 辆小卡车花费了 10 万元，这笔钱听起来数额巨大，但是如果后来若能在旧车市场上将这些卡车卖掉，比如卖了 8 万元，那么沉没成本只有 2 万元。然而，如果你应顾客的要求为他的发明制造机器，比如你为此花费了 10 万元，但是顾客拒付资金，这种情形下，由于这种机器转售的价值为零，因此你的沉没成本就为 10 万元。

为了不至于算错，处理这类问题最好的方法是确保所有的支出都按流量计算：企业经营一年花费了多少钱？这样，你就不大可能忘记考虑资本设备的转售价值，也不大可能混淆沉没成本和可回收成本。

附录

我们使用第 5 章学习过的最优化技术，研究课文中的成本最小化问题。该问题是一个约束最小化问题，其形式为

$$\min_{x_1,x_2} \ w_1x_1 + w_2x_2$$

$$\text{使得 } f(x_1,x_2) = y .$$

我们已经知道求解此类问题的方法有好几种。一种方法是将约束条件代入目标函数。如果我们已经知道 $f(x_1,x_2)$ 的具体表达式，就可以使用这种方法。但对于一般情形（函数表达式未知）来说，我们一般不使用这种方法。

第二种方法是拉格朗日乘数法，这种方法好用。使用这种方法，我们先要建立拉格朗日方程

$$L = w_1x_1 + w_2x_2 - \lambda(f(x_1,x_2) - y)$$

对 x_1, x_2 和 λ 分别求导并相应令求出的导函数为零。由此我们就得到了一阶条件：

$$w_1 - \lambda\frac{\partial f(x_1,x_2)}{\partial x_1} = 0$$

$$w_2 - \lambda\frac{\partial f(x_1,x_2)}{\partial x_2} = 0$$

$$f(x_1,x_2) - y = 0 .$$

最后一个式子就是约束条件。我们可以将前两个式子整理并用第一式除以第二式可得

$$\frac{w_1}{w_2} = \frac{\partial f(x_1,x_2)/\partial x_1}{\partial f(x_1,x_2)/\partial x_2} .$$

注意到这个式子就是我们在课文中推导出来的一阶条件：技术替代率一定等于要素价格之比。

下面我们将这种方法应用于柯布-道格拉斯生产函数：

$$f(x_1,x_2) = x_1^a x_2^b .$$

它的成本最小化问题为

$$\min_{x_1,x_2} \ w_1x_1 + w_2x_2$$

$$\text{使得 } x_1^a x_2^b = y .$$

由于现在我们有了函数的具体表达式，因此我们可以使用将约束条件代入目标函数的方法，也可以使用拉格朗日方法。前一种方法需要将约束条件改写，即将 x_2 写成 x_1 的函数：

$$x_2 = (yx_1^{-a})^{1/b}$$

然后将这个式子代入目标函数，从而得到无约束的最小化问题

$$\min_{x_1} \ w_1x_1 + w_2(yx_1^{-a})^{1/b}$$

现在我们就可以对 x_1 求导，并令得到的导函数等于零。从这些方程可以解出 x_1， x_1 是 w_1, w_2 和 y 的函数，这样就得到了 x_1 的有附加条件的要素需求。这个过程不难做，但是有些繁琐，因此我们不给出具体的求解细节。

然而，我们给出拉格朗日方法求解的具体细节。三个一阶条件分别为：

$$w_1 = \lambda a x_1^{a-1} x_2^{b}$$

$$w_1 = \lambda b x_1^{a} x_2^{b-1}$$

$$y = x_1^a x_2^b .$$

将第一个式子乘以 x_1，将第二个式子乘以 x_2，可得

$$w_1 x_1 = \lambda a x_1^a x_2^b = \lambda a y$$

$$w_2 x_2 = \lambda b x_1^a x_2^b = \lambda b y$$

因此，

$$x_1 = \lambda \frac{ay}{w_1} \tag{20.6}$$

$$x_2 = \lambda \frac{by}{w_2} . \tag{20.7}$$

现在我们用第三个式子求出 λ。将（20.6）和（20.7）代入第三个一阶条件可得

$$(\lambda \frac{ay}{w_1})^a (\lambda \frac{by}{w_2})^b = y .$$

我们可用上式解出 λ，从而得到一个看上去比较“恐怖”的表达式

$$\lambda = (a^{-a} b^{-b} w_1^a w_2^b y^{1-a-b})^{\frac{1}{a+b}} ,$$

将其代入（20.6）和（20.7）就可以解出 x_1 和 x_2。这两个要素需求函数的表达式为

$$x_1(w_1, w_2, y) = (\frac{a}{b})^{\frac{b}{a+b}} w_1^{\frac{-b}{a+b}} w_2^{\frac{b}{a+b}} y^{\frac{1}{a+b}}$$

$$x_2(w_1, w_2, y) = (\frac{a}{b})^{-\frac{a}{a+b}} w_1^{\frac{a}{a+b}} w_2^{\frac{-a}{a+b}} y^{\frac{1}{a+b}} .$$

这样，我们就可以写出企业作出成本最小化选择时的成本函数

$$c(w_1, w_2, y) = w_1 x_1(w_1, w_2, y) + w_2 x_2(w_1, w_2, y) .$$

将前面的要素需求函数代入上式可得

$$c(w_1,w_2,y)=[(\frac{a}{b})^{\frac{b}{a+b}}+(\frac{a}{b})^{\frac{-a}{a+b}}]w_1^{\frac{a}{a+b}}w_2^{\frac{b}{a+b}}y^{\frac{1}{a+b}}.$$

(别担心，期末考试时我们不会考这个表达式。我们的目的只是想让你看清，如何使用拉格朗日乘数方法求解成本最小化问题的解。)

注意，当$a+b$小于、等于、大于 1 时，成本增加速度将大于、等于、小于产量增加速度。这是由于当$a+b$小于、等于、大于 1 时，柯布-道格拉斯技术相应表现为规模报酬递减、不变、递增。

总结

1.成本函数$c(w_1,w_2,y)$，衡量企业在既定要素价格情形下生产一定产量的最小成本。

2.企业的选择决策必须遵守成本最小化行为假设。特别地，有附加条件的要素需求函数向下倾斜，即斜率为负。

3.成本函数和技术的规模报酬类型关系密切。规模报酬递增意味着平均成本递减，规模报酬递减意味着平均成本递增，规模报酬不变意味着平均成本也不变。

4.沉没成本是指不可收回的成本。

复习题

1.证明利润最大化企业总是使成本最小。

2.若某企业在$MP_1/w_1>MP_2/w_2$处生产，企业怎样做才能降低成本但维持产量不变？

3.假设某成本最小化企业使用两种完全替代的生产要素。若这两种要素的价格相同，它们的有附加条件的要素需求是什么样的？

4.某成本最小化企业使用的一种生产要素是纸张，如果纸张的价格上升时，企业的反应是改变某些要素的需求但要维持产量不变。那么企业使用纸张的数量将如何变化？

5.若企业使用n种生产要素（$n>2$），对于某既定产出水平，要素价格变动（Δw_i）和要素需求变动（Δx_i）有什么样的关系？请根据显示成本最小化公理说明。

复习题答案

1.证明利润最大化企业总是使成本最小。

【复习内容】利润最大化；成本最小化；

【参考答案】

企业的利润恒等于收入减去成本，即$\pi \equiv py - c(y)$。现在我们要证明，如果y^*是该企业利润最大化问题的解，则必然意味着对于y^*这个产量来说，企业选择的生产方法必然是成本最小的。

反证一下。如若不然，则意味着对于y^*这个产量来说，仍存在使企业成本更小的方法。从上述利润恒等式可知，企业的利润会进一步增大。这样我们就得到了一个矛盾，因为我们已假设该企业的利润已实现最大化。

所以，利润最大化企业总是使成本最小。

2.若某企业在$MP_1/w_1 > MP_2/w_2$处生产，企业怎样做才能降低成本但维持产量不变？

【复习内容】成本最小化

成本最小化的条件为$MP_1/w_1 = MP_2/w_2$，此式可用几何法或代数法推导。我们使用代数法。

企业的成本最小化问题为

$$\min_{x_1,x_2} \ w_1x_1 + w_2x_2$$

$$\text{使得 } f(x_1,x_2) = y$$

建立拉格朗日方程

$$L = w_1x_1 + w_2x_2 - \lambda(f(x_1,x_2) - y)$$

对x_1, x_2和λ分别求导并相应令求出的导函数为零。由此我们就得到了一阶条件：

$$w_1 - \lambda\frac{\partial f(x_1,x_2)}{\partial x_1} = 0$$

$$w_2 - \lambda\frac{\partial f(x_1,x_2)}{\partial x_2} = 0$$

$$f(x_1,x_2) - y = 0\,.$$

最后一个式子就是约束条件。我们可以将前两个式子整理并用第一式除以第二式可得

$$\frac{w_1}{w_2}=\frac{\partial f(x_1,x_2)/\partial x_1}{\partial f(x_1,x_2)/\partial x_2}.$$

即 $$\frac{w_1}{w_2}=\frac{MP_1(x_1,x_2)}{MP_2(x_1,x_2)}$$

也即 $$\frac{MP_1(x_1,x_2)}{w_1}=\frac{MP_2(x_1,x_2)}{w_2}$$ 。

【参考答案】

由于成本最小化的条件为 $MP_1/w_1=MP_2/w_2$，而现在题目中已知的条件为 $MP_1/w_1>MP_2/w_2$。为了看清企业如何调整要素使用量，不妨假设 $w_1=w_2=1$。

立即由上式可得 $MP_1>MP_2$，这意味着要素 1 的边际产量大于要素 2 的边际产量，因此，在维持产量不变的情形下，企业增加要素 1 的使用量减少要素 2 的使用量，可使成本减小。直至 $MP_1=MP_2$ 时成本最小（注意我们的假设 $w_1=w_2=1$）。这其实仍然是说成本最小化的条件为 $MP_1/w_1=MP_2/w_2$。

3.假设某成本最小化企业使用两种完全替代的生产要素。若这两种要素的价格相同，它们的有附加条件的要素需求是什么样的？

【复习内容】有附加条件的要素需求；完全替代型生产技术

使企业成本最小的要素选择，通常取决于所有要素的价格和企业想要生产的产量水平，所以我们可以将要素选择写为 $x_1(w_1,w_2,y)$ 和 $x_2(w_1,w_2,y)$。这称为有附加条件的要素需求函数（conditional factor demand functions）或派生要素需求（derived factor demands）。这些函数衡量在企业产量 y 既定的条件下，要素需求与要素价格及其产量之间的关系。

需要注意，有附加条件的要素需求和利润最大化的要素需求之间的区别。有附加条件的要素需求是既定产量水平下的成本最小化的选择；而利润最大化的要素需求则是既定产品价格下的利润最大化的选择。

有附加条件的要素需求通常不可直接观测到；它是个假设出来的工具。有附加条件的要素需求表示，如果企业希望使既定产量的成本最小，每种要素的最优投入量是多少。有附加条件的要素需求是个有用的工具，我们可以使用它分析最有效率的生产方法问题，而用利润最大化要素需求去分析最优产量的决定问题。

【参考答案】

为简单起见，假设完全替代类型的技术 $f(x_1,x_2)=x_1+x_2$。此时，要素的条件需求函数是什么样的？

既然要素 1 和 2 在生产过程中是完全替代的，显然哪种要素价格最低，企业就会使用哪种要素。因此，生产 y 单位产品的最小成本是 $w_1 y$ 和 $w_2 y$ 中最小的那个。换句话说：

$$c(w_1, w_2, y) = \min\{w_1 y, w_2 y\} = \min\{w_1, w_2\} y$$

现在由于题目告知，两种要素的价格相等。即 $w_1 y = w_2 y$，由上式可知，企业在这种情形下，可以任意使用两种要素的数量，只要使得 $y = f(x_1, x_2) = x_1 + x_2$ 即可。

4.某成本最小化企业使用的一种生产要素是纸张，如果纸张的价格上升时，企业的反应是改变某些要素的需求但要维持产量不变。那么企业使用纸张的数量将如何变化？

【复习内容】显示的成本最小化；成本最小化弱公理

假设我们观测到两组价格 (w_1^t, w_2^t) 和 (w_1^s, w_2^s) 和相应的要素选择 (x_1^t, x_2^t) 和 (x_1^s, x_2^s)。假设这两个选择生产相同的产量 y。如果每种选择在相应价格下都是成本最小化的选择，则必然有

$$w_1^t x_1^t + w_2^t x_2^t \le w_1^t x_1^s + w_2^t x_2^s$$

和

$$w_1^s x_1^s + w_2^s x_2^s \le w_1^s x_1^t + w_2^s x_2^t.$$

如果企业在生产 y 单位产品时总是选择使成本最小的方法，则该企业在时期 t 和时期 s 的选择必然满足这些不等式。我们将这些不等式称为成本最小化弱公理（Weak Axiom of Cost Minimization, WACM）。

将第二个式子改写为

$$-w_1^s x_1^t - w_2^s x_2^t \le -w_1^s x_1^s - w_2^s x_2^s$$

然后加到第一个式子得到

$$(w_1^t - w_1^s)x_1^t + (w_2^t - w_2^s)x_2^t \le (w_1^t - w_1^s)x_1^s + (w_2^t - w_2^s)x_2^s,$$

整理可得

$$(w_1^t - w_1^s)(x_1^t - x_1^s) + (w_2^t - w_2^s)(x_2^t - x_2^s) \le 0.$$

使用 Δ 符号表示要素需求量的变动和要素价格变动，可把上式写得更紧凑些

$$\Delta w_1 \Delta x_1 + \Delta w_2 \Delta x_2 \le 0.$$

我们只使用成本最小化行为的假设就推导出了上述不等式。它蕴含着对企业行为的限制——当要素价格改变但产品价格不变时企业行为应该怎样变化。

【参考答案】

曹乾（东南大学 caoqianseu@163.com）

在使用两种要素进行生产的情形下，由成本最小化弱公理可知：

$$\Delta w_1 \Delta x_1 + \Delta w_2 \Delta x_2 \leq 0$$

若纸张（要素 1）的价格上升要素 2 的价格不变即 $\Delta w_2 = 0$，因此上述不等式变为

$$\Delta w_1 \Delta x_1 \leq 0 .$$

这表明，纸张价格上升，在维持产量不变时，企业对纸张的需要要么下降要么不变。

5.若企业使用 n 种生产要素（$n > 2$），对于某既定产出水平，要素价格变动（Δw_i）和要素需求变动（Δx_i）有什么样的关系？请根据显示成本最小化公理说明。

【复习内容】显示的成本最小化；成本最小化弱公理

【参考答案】

由成本最小化弱公理可知，在 n 种生产要素情形下有：

$$\Delta w_1 \Delta x_1 + \Delta w_2 \Delta x_2 + ... \Delta w_i \Delta x_i + ... + \Delta w_n \Delta x_n \leq 0$$

或者写成

$$\sum_{i=1}^{n} \Delta w_i \Delta x_i \leq 0$$

若其他要素价格不变，只有要素 i 价格变动，即 $\Delta w_j = 0$（$j \neq i$），则由上面的任何一个式子（最好用第一式，因为直观）可知，$\Delta w_i \Delta x_i \leq 0$.

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

21.成本曲线（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

21 成本曲线

在上一章我们分析了企业的成本最小化问题。在本章，我们将使用一种重要的几何工具——**成本曲线**（cost curve）继续分析这个问题。成本曲线可用于描述企业成本函数，在研究最优产出决策时，这个工具非常重要。

21.1 平均成本

考虑上一章描述的成本函数。成本函数$c(w_1, w_2, y)$表示，当要素价格为(w_1, w_2)时，企业生产 y 单位产品的最小成本。在本章我们将假定要素的价格是固定不变的，因此可将成本函数写为唯一一个自变量即产量 y 的函数：$c(y)$。

企业的某些成本和产量大小无关。我们在第 20 章已经知道，这样的成本是不变成本。不变成本是指无论企业产量为多大时必须支付的成本。例如，不管企业产量如何，企业都要偿还抵押借款的利息。

另一种成本随产量变动而变动：这样的成本是可变成本。企业的总成本（total costs）总是可以写为可变成本$c_v(y)$和不变成本 F 之和：

$$c(y) = c_v(y) + F\,.$$

平均成本函数（average cost function）衡量每单位产品的成本。**平均可变成本函数**（average variable cost function）衡量每单位产品的可变成本；**平均固定成本函数**（average fixed cost function）衡量每单位产品的固定成本。根据以上的式子：

$$AC(y) = \frac{c(y)}{y} = \frac{c_v(y)}{y} + \frac{F}{y} = AVC(y) + AFC(y)\,.$$

其中，$AVC(y)$表示平均可变成本，$AFC(y)$表示平均固定成本。这些函数是什么样子的？这些函数中最简单的当属平均固定成本函数：当 $y = 0$，它为无穷大；随着 y 增加，它开始逐渐下降，最终会趋近于 0。如图 21.1A 所示。

再来看看平均可变成本函数。我们从产量为零时开始分析，假设生产一单位产品，则 $y = 1$ 时的平均可变成本正好就是生产这一单位产品的可变成本。现在将产量增加到 2 单位。可以预期，在最坏的情形下，总可变成本变为生产 1 单位产品时的 2 倍，因此平均可变成本保持不变。当产量增加时，如果我们能使用更有效率的方法进行生产，那么平均可变成本甚至一开始下降。但是我们可以预期，平均可变成本最终会上升。为什么？因为如果有固定成本，它们最终会约束生产过程。

例如，假设我们租用规模既定的厂房，则租金就是固定成本。于是随着产量增加，平均可变成本——单位生产成本——可能暂时维持不变。但是当达到厂房的生产能力后，这些

成本会急剧上升，从而使得平均可变成本的曲线呈现图 21.1B 的形状。

平均成本曲线是平均固定成本曲线和平均可变成本曲线之和；因此它的形状为 U 型的，如图 21.C 所示。可变成本曲线最初会下降，这是由于平均可变成本下降引起的；但它最终会上升，这是由于平均可变成本上升所引起。这两种效应使得平均成本曲线为 U 型。

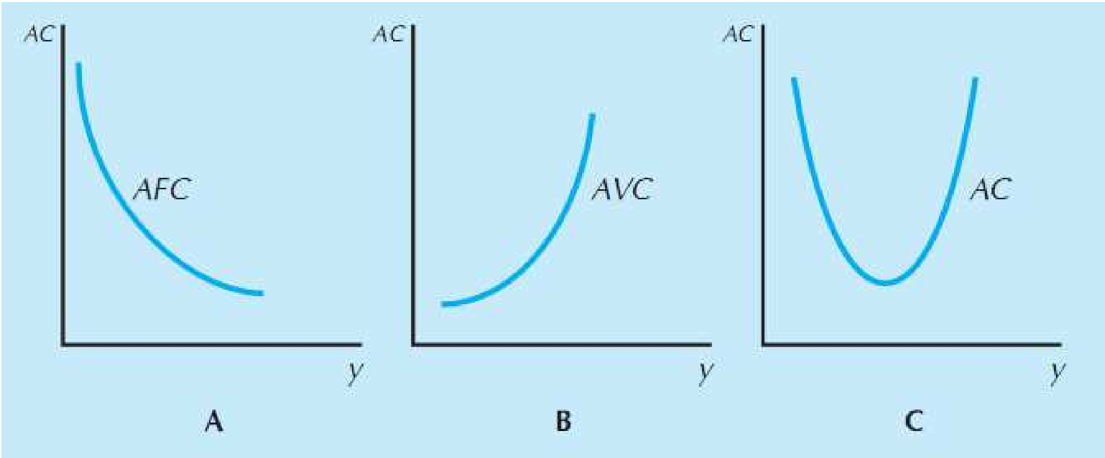


图 21.1：平均成本曲线的形状。（A）平均固定成本随产量的增加而下降；（B）平均可变成本随产量增加最终会上升；（C）由于（A）和（B）这两种效应，平均成本曲线为 U 型。

21.2 边际成本

我们还对一种成本曲线感兴趣，这就是**边际成本曲线**（marginal cost curve）。边际成本曲线衡量产量变动引起的成本变动。也就是说，对于任何给定的产量水平 y，我们想知道，如果产量变动 Δy，成本怎样变动。

$$MC(y)=\frac{\Delta c(y)}{\Delta y}=\frac{c(y+\Delta y)-c(y)}{\Delta y}.$$

边际成本也可以用可变成本函数定义：

$$MC(y)=\frac{\Delta c_v(y)}{\Delta y}=\frac{c_v(y+\Delta y)-c_v(y)}{\Delta y}.$$

这个式子等价于第一个式子，因为 $c(y)=c_v(y)+F$，而且固定成本 F 不随产量 y 变化而变化。

通常我们将 Δy 看成一单位产量，因此，边际成本表示额外多生产一单位产品所导致的成本变动。如果我们研究的是离散商品的生产，则生产 y 单位产品的边际成本为 $c(y)-c(y-1)$。这是边际成本计算的一种简单方法，但是有时它容易误导。因为边际成本衡量的是一种**变化率**：成本变动除以产量变动。如果产量变动正好为一单位，那么边际成本

似乎就是成本的变动，但它实际上是变动率。只不过此时恰好有$\Delta c/\Delta y=\Delta c/1=\Delta c$而已。

我们怎样画出边际成本曲线？首先，我们注意到下列事实。根据定义可知，当产量为零使，可变成本也为零。因此，对于第一单位产品

$$MC(1)=\frac{c_v(1)+F-c_v(0)-F}{1}=\frac{c_v(1)}{1}=AVC(1).$$

因此，第一单位产品的边际成本等于这一单位产品的平均可变成本。

现在假设产量处于**平均**可变成本下降的区域。那么，在该区域**边际**成本必定小于平均可变成本。这是因为你若想让平均数下降，新加进的数必须小于原平均数。

如果我们将不同产量上的平均成本用一系列数值表示。如果平均值下降，那么从某点增加额外一单位产量的成本，必定小于该点的平均值。要使平均值下降，新增的数值必须小于原来的平均值。

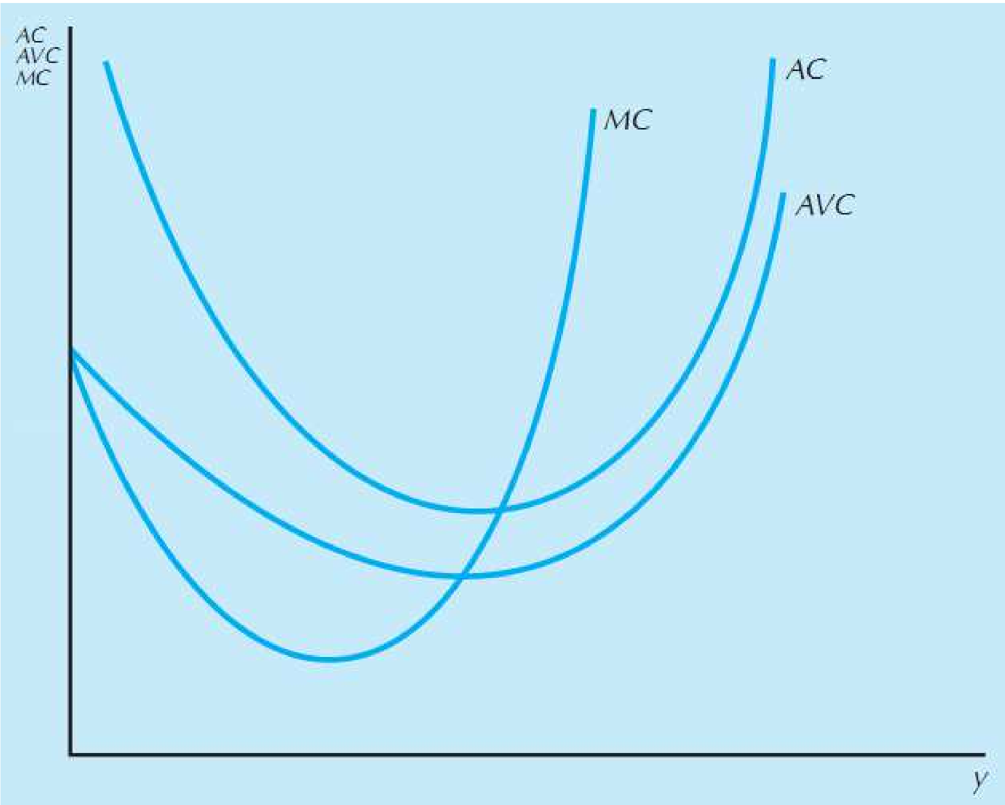


图 21.2：成本曲线。平均成本曲线（AC），平均可变成本曲线（AVC）和边际成本曲线(MC)。

类似地，如果产量处于平均可变成本上升的区域，则该区域内任何一点的边际成本必然大于该点的平均可变成本——较高的边际成本抬升了平均可变成本。

由上面的论述可知，在平均可变成本曲线最低点的左侧，边际成本曲线必然位于平均可变成本曲线的下方；在平均可变成本曲线最低点的右侧，边际成本曲线必然位于平均可变

成本曲线的上方。这意味着边际成本曲线必然穿过平均可变成本曲线的最低点。

上述论证过程完全适用于平均成本曲线和边际成本曲线的关系。如果平均成本下降，则边际成本必然小于平均成本；如果平均成本上升，则边际成本必然大于平均成本。根据这些结论，我们可以画出边际成本曲线的图形，如图 21.2 所示。

现在简要总结一下我们已经得到的重要知识点：

- 平均可变成本曲线可能一开始就下降，但未必一定如此。然而，它最终必然会上升，这是由于固定要素限制了生产过程。
- 平均成本曲线一开始会下降，因为平均固定成本下降；但是平均成本曲线最终也会上升，因为平均可变成本上升了。
- 在第一单位产量上，边际成本等于平均可变成本。
- 边际成本曲线穿过平均成本曲线的最低点，也穿过平均成本曲线的最低点。

21.3 边际成本和可变成本

各种成本曲线之间还存在着其他的一些关系。下面这条关系不是那么明显：可以证明产量为 y 时边际成本曲线下方区域的面积，等于产量为 y 时的可变成本。为什么？

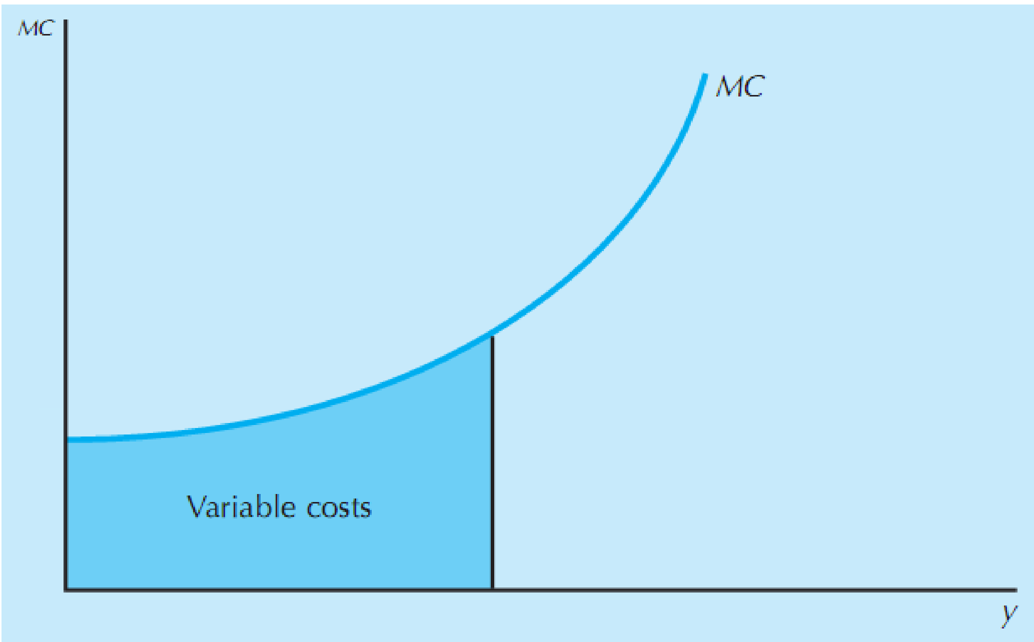


图 21.3：边际成本和可变成本。边际成本曲线下方区域的面积正好等于可变成本。

边际成本曲线衡量生产额外每单位产品的成本。如果我们将这些每单位产品的生产成本相加得到的和，正好等于总成本与固定成本之差。

我们先来看看在生产的产品为离散数量时，怎么严格证明这个结论。首先，我们注意

到

$$c_v(y)=[c_v(y)-c_v(y-1)]+[c_v(y-1)-c_v(y-2)]+...+[c_v(1)-c_v(0)]$$

这是对的，因为$c_v(0)=0$，所有中间的项都可消去；也就是说，第二项与第三项抵消，第四项与第五项抵消，以此类推。但是上式每个方括号的差值都是不同产量水平上的边际成本：

$$c_v(y)=MC(y-1)+MC(y-2)+...+MC(0)\,.$$

因此，上式中的每个方括号所代表的项都可以用一个矩形的面积表示，矩形的底为 1，高为$MC(y)$。将这些矩形面积的加和正好就是边际成本曲线下方的面积，如图 21.3 所示。

例子：具体的成本曲线

我们考察成本函数$c(y)=y^2+1$。由该成本函数可以推导出下列成本曲线：

- 可变成本：$c_v(y)=y^2$
- 固定成本：$c_f(y)=1$
- 平均可变成本：$AVC(y)=y^2/y=y$
- 平均可变成本：$AFC(y)=1/y$
- 平均成本：$AC(y)=\dfrac{y^2+1}{y}=y+\dfrac{1}{y}$
- 边际成本：$MC(y)=2y$

除了边际成本之外，其他成本表达式都很明显。当然，如果你学过微积分，边际成本的表达式也很明显。如果成本函数为$c(y)=y^2+F$，则边际成本函数为$MC(y)=2y$。如果你不知道它是怎么算出来的，就只好记住它，因为在习题中你要用到它。

这些成本曲线是什么样子的？最容易画出的曲线是平均可变成本曲线，在上例中它是一条斜率为 1 的直线。边际成本曲线也很容易画出，它是一条斜率为 2 的直线。

当平均成本等于边际成本时，平均成本曲线达到最低点，这就是说

$$y+\frac{1}{y}=2y\,,$$

由此可解出$y_{\min}=1$。当$y=1$时，平均成本和边际成本都等于 2。我们将这些成本曲线画在图 21.4 中。

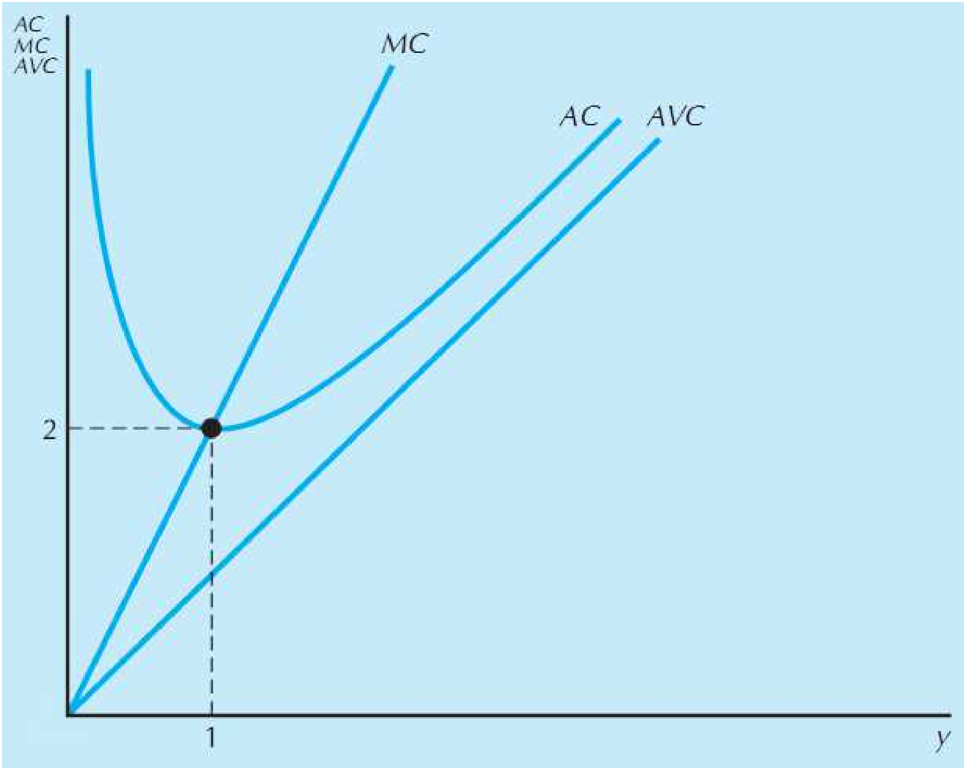


图 21.4：成本曲线。$c(y)=y^2+1$的成本曲线。

例子：两个工厂的边际成本曲线

假设你有两个工厂，它们的成本函数不同。假设工厂 1 和 2 的成本函数分别为$c_1(y_1)$和$c_2(y_2)$。你想以最小的成本生产y单位产品。一般来说，你会使用两个工厂各生产一些产量。问题是，两个工厂各应该生产多少？

建立最小化问题：

$$\min_{y_1,y_2} c_1(y_1)+c_2(y_2)$$

$$\text{使得 } y_1+y_2=m\,.$$

如何解这个问题？可以证明，在最优产量分配处，工厂 1 的边际成本必定等于工厂 2 的边际成本。为了证明这一点，假设边际成本不相等；则你肯定会将一些产量从边际成本较高的工厂转移到边际成本较低的工厂。如果产量分配已达到最优，那么将产量从一家工厂移动到另一家就不可能降低成本。

令 $c(y)$表示生产y单位产品时，成本最小的那种方法的生产函数，也就是说产量分配已达到最优时的成本函数。因此，无论你用工厂 1 还是工厂 2 生产额外一单位产品，它们的边际成本必定是相同的。

我们在图 21.5 中，画出了上述两个工厂的边际成本曲线——$MC_1(y_1)$ 和 $MC_2(y_2)$。两家工厂合在一起的边际成本曲线，就是这两个工厂各自边际成本曲线在水平方向上的加总，如图 21.5C 所示。

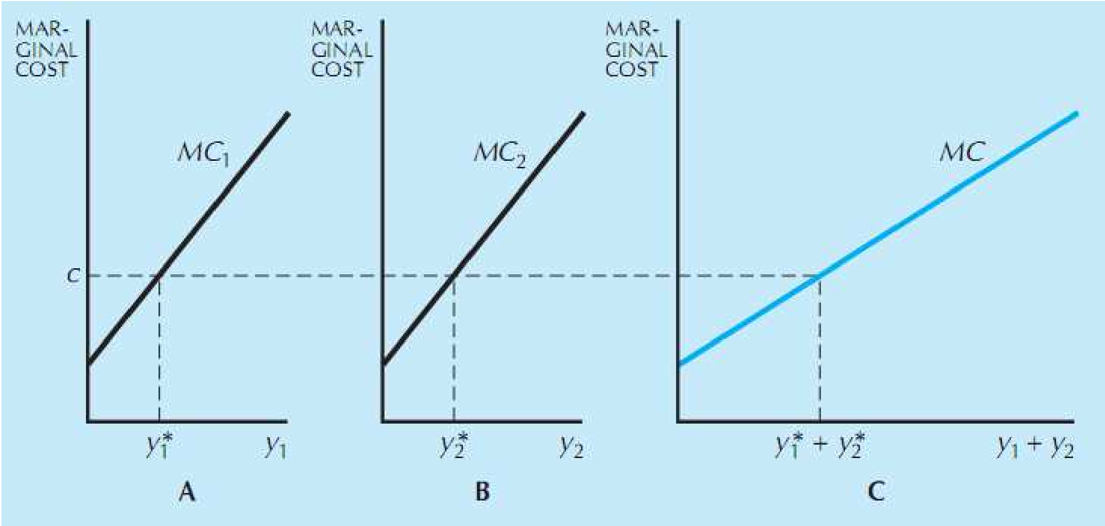


图 21.5：拥有两个工厂的企业的边际成本曲线。C 图中的总边际成本曲线是两个工厂的边际成本曲线（A 图和 B 图）在水平方向上相加而得到。

对于任何既定水平的边际成本，比如 c，两个工厂将分别生产 y_1^* 和 y_2^* 单位，使得 $MC_1(y_1^*) = MC_2(y_2^*) = c$，这样我们就生产出了 $y_1^* + y_2^*$ 单位产品。因此，在任一边际成本水平 c 上的产量，就是当两工厂各自边际成本均等于 c 时的产量之和：边际成本曲线在水平方向上相加。

21.4 在线拍卖的成本曲线

在第 17 章，我们分析了搜索引擎广告拍卖模型。在那里我们知道，当消费者在搜索引擎中输入问题时，该问题与登广告的人选择的关键词进行配对。关键词与问题匹配的那些登广告的人进行竞标。报价最高的人得到最好的广告位置，报价第二高的人得到第二好的广告位置，依次类推。在其他条件不变的情形下，广告位置越好，得到的点击数越多。

在以前的拍卖模型中，我们假设每个竞价人（登广告的人）可以为每个关键词选择一个单独的报价。在实践中，每个登广告的人在他参与的所有拍卖中都选择相同的报价。从竞价人的角度看，一次拍卖决定的价格并不重要。真正重要的是，他的广告得到的点击数 x 与这些点击的成本 $c(x)$ 之间的关系。

这就是我们已非常熟悉的总成本函数。当登广告的人知道成本函数时，他可以确定购买的点击数量。令 v 表示一次点击的价值，利润最大化问题为

$$\max_x vx - c(x).$$

我们已经知道，最优解可让价值等于边际成本。一旦登广告的人确定了利润最大化时的点击数量，他可以选择能产生这个点击量的报价。

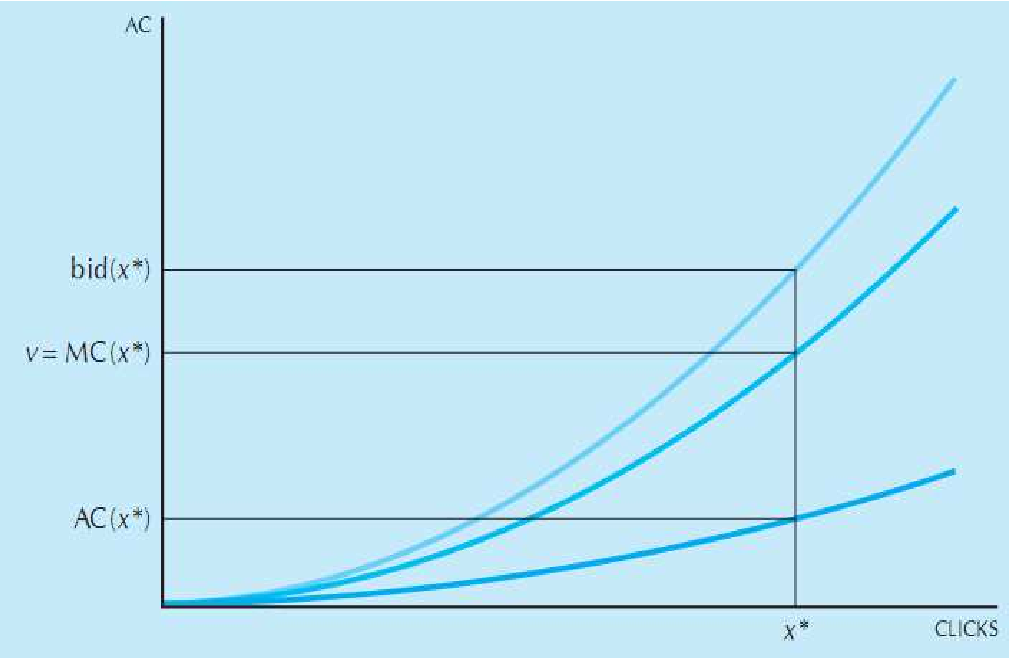


图 21.6：点击的成本曲线（click-cost curves）。能使利润最大化的点击量位于价值等于边际成本之处，它确定了合适的报价和每次点击的平均成本。

这个过程可用图 21.6 表示。在这个图中，平均成本曲线和边际成本曲线的画法和以前一样，我们只不过增加了一条新的曲线来表示报价。

登广告的人如何发现自己的成本曲线？一种方法是，他可以进行实验，在该实验中他报出不同的价格，并记录相应的点击数和成本。或者，搜索引擎公司可以根据拍卖信息代为估计成本函数。例如，假设我们想估计下列事件的后果：登广告人报出的每次点击的价格从 0.50 元上升到 0.80 元。搜索引擎公司可以查阅此人参与的所有拍卖，根据这些信息代替他估计他的广告位置如何变化以及他在新广告位置能得到多少额外的点击量。

21.5 长期成本

在前面的分析中，我们将固定成本看为企业短期内为固定要素支付的成本。但在长期，这些“固定”要素是可变的。

当然，长期情形下，仍然可能存在准固定要素（quasi-fixed factors）。也就是说企业的技术可能具有下列特征：只要产量为正就必须支付的某些成本。但在长期，不存在固定成本，因为企业总可以选择生产零单位产品，也就是说它可以选择退出它所在的行业。如果长期内确有准固定要素，则长期平均成本曲线的形状是 U 型的，这一点和短期平均成本曲线类似。

但在长期，根据长期的定义可知，企业总可以选择生产零单位，因此成本为零。

当然，这个长期有多长，则要取决于我们研究的具体问题。如果我们研究的固定要素是工厂规模，则长期是指企业要花多长时间才能改变工厂规模。如果固定要素是按照合同支付的工资，则长期是指企业要花多长时间才能改变雇员的数量。

不妨具体点，令固定要素就是指工厂规模，我们用 k 表示。企业的短期成本函数用 $c_s(y,k)$ 表示，其中 k 表示该企业的工厂规模，下标 s 表示“短期”。（此处，k 相当于 20 章中的 $\bar{x}_2$。）

任意给定一个产出水平，则存在生产该产量水平的最优工厂规模。我们用 $k(y)$ 表示该企业的工厂规模。注意，$k(y)$ 表示产量为 y 时的最优工厂规模。也就是说，该企业对工厂规模的条件需求是产量的函数。（当然，对工厂规模的条件需求还取决于工厂规模的价格以及其他生产要素的价格，但我们假设这些因素都不变。）于是，正如第 20 章一样，企业的长期成本函数可用 $c_s(y,k(y))$ 表示。它是指如果企业能将工厂规模调整到最优时，生产 y 单位产品的成本。企业的长期成本函数就是该企业工厂规模最优时的短期成本函数：

$$c(y)=c_s(y,k(y)).$$

企业长期成本函数的图形是什么样的？任意选取一产量水平 y^*，令产量为 y^* 时的最优工厂规模为 $k^*=k(y^*)$。若企业的厂房规模为 k^*，则其短期生产函数为 $c_s(y,k^*)$，长期成本函数为 $c(y)=c_s(y,k(y))$。这一点在上面已指出过。

现在，注意一个重要的事实：对于任一产量来说，短期生产成本必定不会小于长期成本。为什么？在短期，企业工厂规模固定，但在长期企业可以自由调整工厂规模。既然企业在长期可以任意选择工厂规模，它当然可以选择短期时的最优工厂规模 k^*，因此长期情形下生产 y 产品的成本 $c(y)$，必定不会大于 $c_s(y,k^*)$，即：对于所有产量水平 y，均有

$$c(y)\le c_s(y,k^*).$$

事实上，在某特定产量水平 y^* 上，有

$$c(y^*)=c_s(y^*,k^*).$$

为什么？因为产量为 y^* 时，最够工厂规模为 k^*。因此，产量为 y^* 时，长期成本等于短期成本。

如果短期成本总是大于长期成本，而且在某产量水平上二者相等，那么这表示短期平均成本和长期平均成本具有相同的性质：$AC(y)\le AC_s(y,k^*)$ 和 $AC(y^*)=AC_s(y^*,k^*)$。这意味着短期平均成本曲线必然位于长期平均成本曲线上方，而且它们在 y^* 点接触。因此，长期平均成本曲线（LAC）和短期平均成本曲线（SAC）必然在 y^* 点相切，如图 21.7 所示。

对于 y^* 之外的产量，分析类似。假设我们选取产量 $y_1,y_2,...,y_n$，相应的工厂规模为 $k_1=k(y_1),k_2=k(y_2),...,k_n=k(y_n)$。则我们可以得到图 21.8。图 21.8 是说，长期平均成

本曲线是短期平均成本曲线的下包络线（lower envelope）[一]。

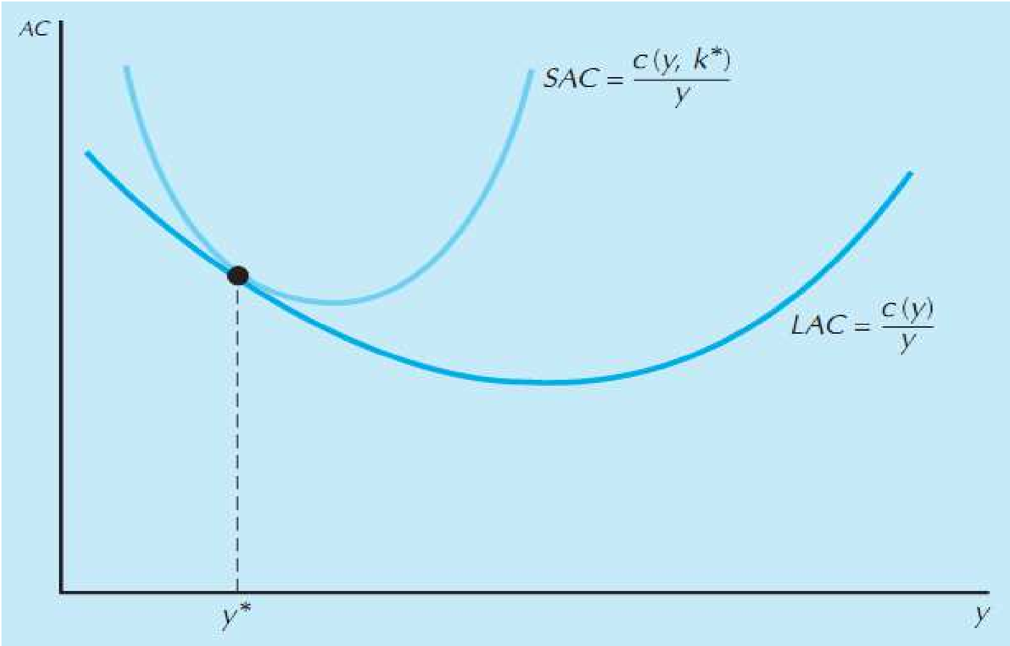


图 21.7：短期平均成本曲线和长期平均成本曲线。短期平均成本曲线必然和平均成本曲线相切。

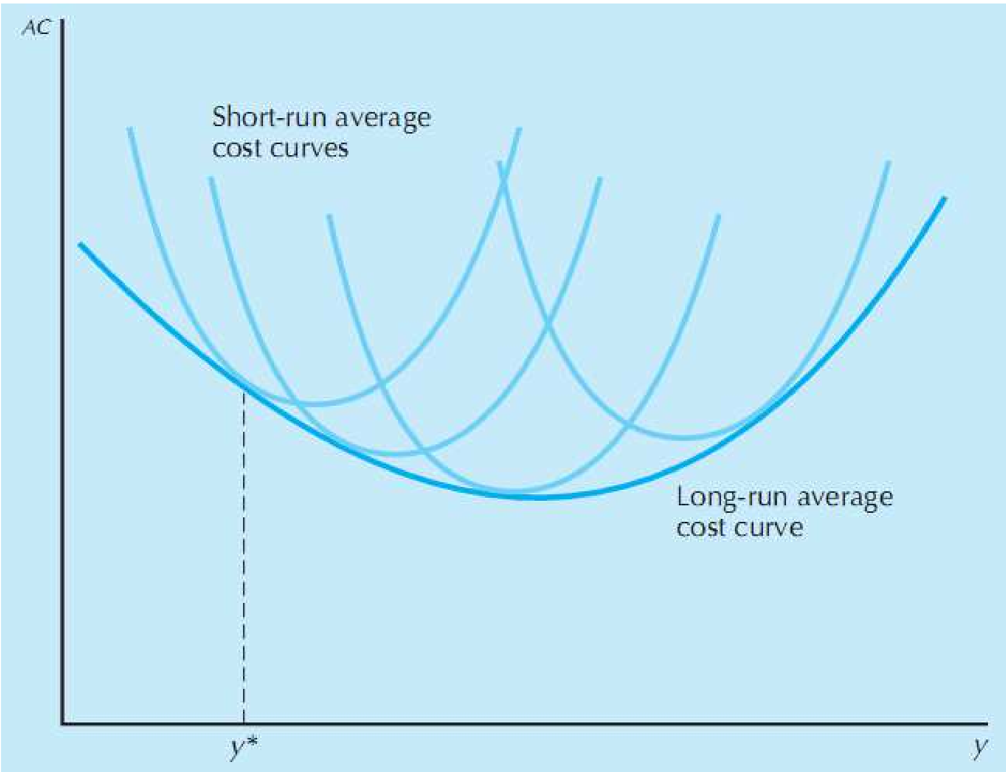


图 21.8：短期平均成本曲线和长期平均成本曲线。长期平均成本曲线是短期平均成本曲线的包络线。

[一] 在几何学中，某个曲线族的包络线是指与该曲线族的每条线都有至少一点相切的一条曲线。译者注。

21.6 工厂规模为离散的情形

在以上的分析中，我们隐含着一个假设，即工厂规模是连续的。因此，每个产量水平都有一个唯一的最优工厂规模。但是，如果工厂规模只有有限的几种可供选择，情形将是怎样的？

例如，假设，我们有四种工厂规模：k_1，k_2，k_3 和 k_4。我们将这四个工厂规模对应的四条不同平均成本曲线画在图 21.9 中。

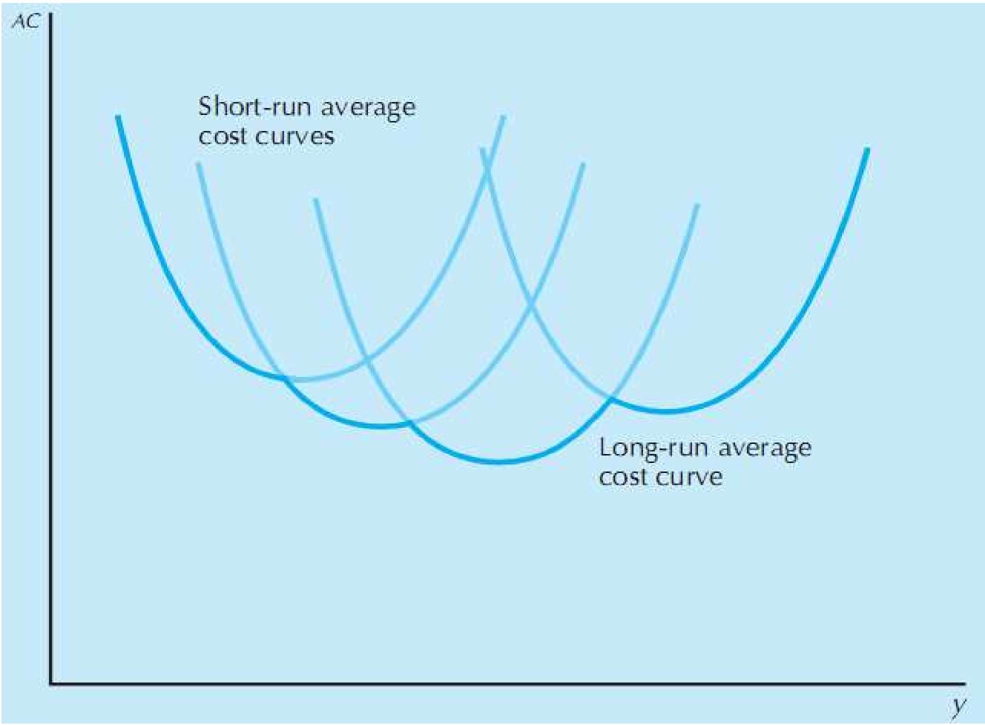


图 21.8：工厂规模为离散的情形。和工厂规模为连续的情形一样，长期平均成本曲线也是短期平均成本曲线的下包络线。

这种情形下，我们如何构建长期平均成本曲线？我们已知道，长期平均成本曲线是调整工厂规模 k 为最优时的成本曲线。知道了这一点，此事就不难做：既然只有四种不同的工厂规模，我们只要看看哪种工厂规模的成本最小，就选这种工厂规模即可。也就是说，对于任一产量水平 y，我们选取的工厂规模要能做到时生产成本最小。

因此，长期平均成本曲线将为短期平均成本曲线的下包络线，如图 21.9 所示。注意，该图与图 21.8 的意义在本质上是一样的：短期平均成本不会小于长期平均成本；在某特定产量上，长期平均成本和短期平均成本相等。我们已经知道，这样的特定产量是指企业对“固定”要素的长期需求恰好等于企业短期拥有的固定要素时的产量。

21.7 长期边际成本

在上一节我们已经知道，长期平均成本曲线是短期平均成本曲线的下包络线。这对边际成本意味着什么？我们首先分析工厂规模为离散的情形。在这种情形下，长期边际成本曲线由不同短期边际成本曲线适当的部分组成，如图 21.10 所示。对于每个产量水平，我们先找到应该在哪条短期平均成本曲线上生产，然后我们再找到这条平均成本曲线对应的边际成本曲线。

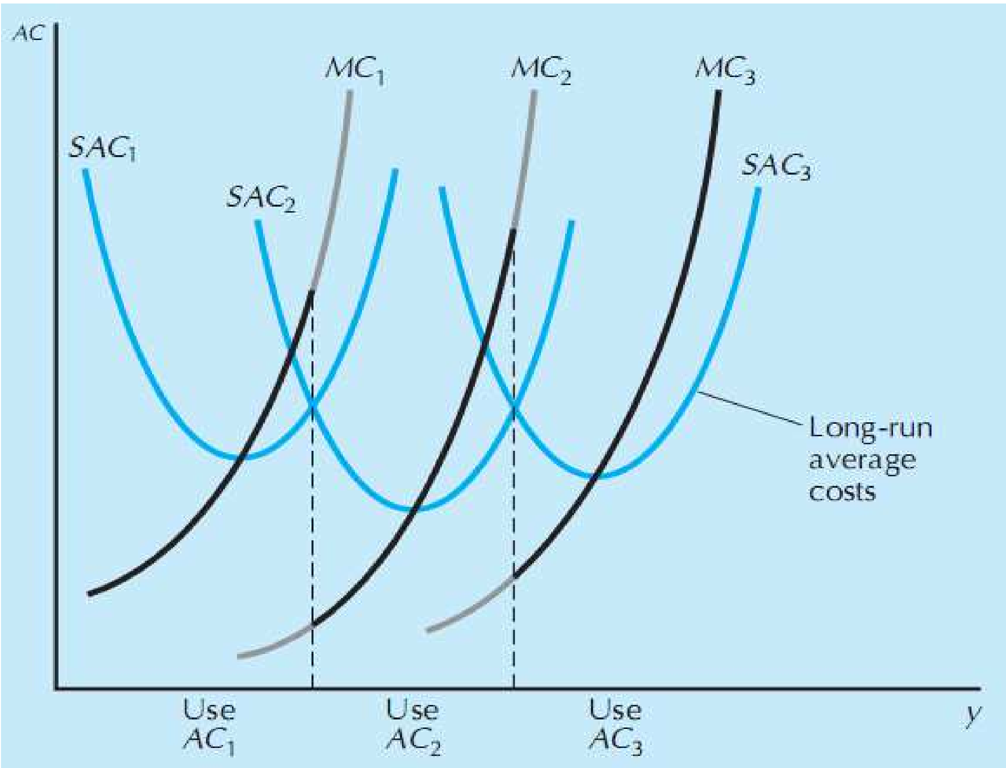


图 21.10：长期边际成本曲线。如果固定要素的数量是离散的，企业会选择使平均成本最小的固定要素。因此，长期边际成本曲线由若干短期边际成本曲线上的部分组成，这些短期边际成本曲线分别对应着固定要素的不同数量水平。

我们在图 21.10 中得出的结论，当然适用于工厂规模种类有任意多的情形，也就是说适用于固定要素的数量为连续的情形，这样我们就得到了图 21.11（见下一页）。任一产量水平 y 的长期边际成本，必然等于生产该产量的短期边际成本——前提是短期工厂规模恰好是生产 y 单位产品的最优工厂规模。

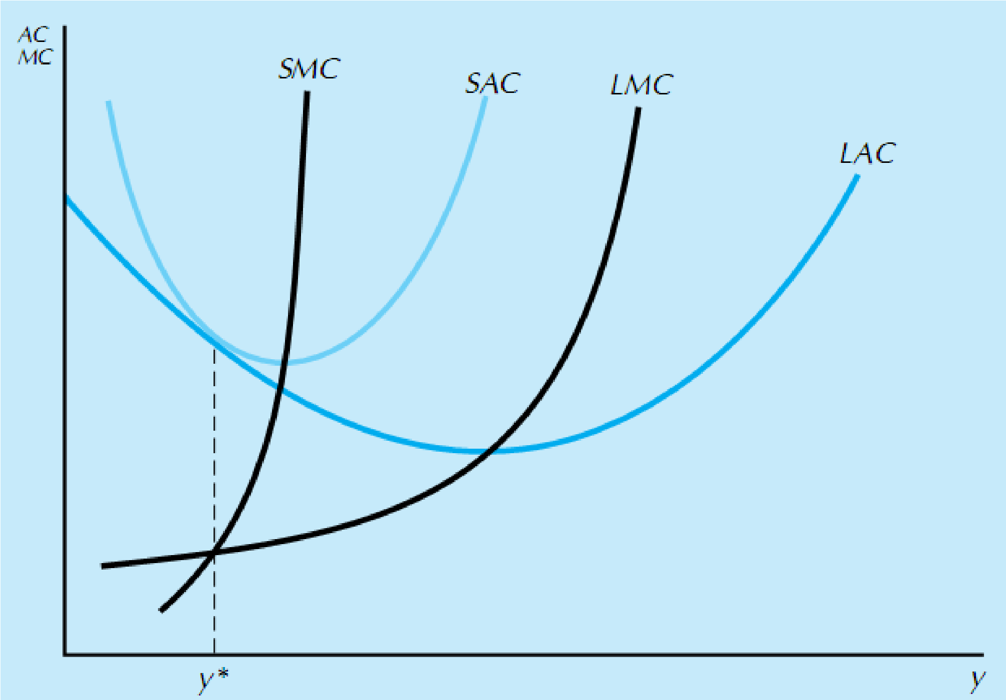


图 21.11：长期边际成本曲线。固定要素的数量为连续的情形下，长期边际成本和短期边际成本之间的关系。

附录

在正文中我们断言，对于第一单位产品来说，平均可变成本等于边际成本。若用微积分的语言表达，则为

$$\lim_{y\to 0}\frac{c_v(y)}{y}=\lim_{y\to 0}c'(y).$$

上式左侧在 $y=0$ 时没有定义。但它的极限有定义，我们可用洛必达法则（$L'\ H\hat{o}pital's$ rule）计算，这个法则是说分子分母都为 0 的分式的极限，可以先对分子分母先求导，然后再求极限。根据这一法则，我们有

$$\lim_{y\to 0}\frac{c_v(y)}{y}=\frac{\lim_{y\to 0}dc_v(y)/dy}{\lim_{y\to 0}dy/dy}=\frac{c'(0)}{1},$$

这样就证明了这个结论。

我们还断言，边际成本曲线下方区域的面积等于总可变成本。这个结论若使用微积分基本定理，则容易证明。由于

$$MC(y)=\frac{dc_v(y)}{dy},$$

而边际成本曲线下方区域的面积等于

$$\int_0^y \frac{dc_v(x)}{dx}dx = c_v(y) - c_v(0) = c_v(y).$$

长期边际成本曲线和短期边际成本曲线的关系在图形上简单明了，但这种关系的经济学意义是什么？微积分的论证可以帮助我们解决这一问题。这一证明比较简单。生产的边际成本就是由于产量变动导致的成本变动。在短期，我们必须维持工厂规模不变，而在长期我们可以自由调整工厂规模。因此，长期边际成本等于下列两部分之和：第一部分是，维持工厂规模不变时成本如何变动；第二部分是当工厂规模调整后，成本如何变动。但是，如果企业选择的工厂规模恰好是最优的，那么第二部分就为零。该情形下，长期边际成本和短期边际成本必然相等。

证明需要用到链式法则。根据课文中的定义：

$$c(y) \equiv c_s(y, k(y)).$$

对 y 求导可得

$$\frac{dc(y)}{dy} = \frac{\partial c_s(y,k)}{\partial y} + \frac{\partial c_s(y,k)}{\partial k}\frac{\partial k(y)}{\partial y}.$$

如果在上式中，令产量 y 等于某特定产量 y^*，并令 k 为对应于 y^* 的最优工厂规模 $k^* = k(y^*)$，则

$$\frac{\partial c_s(y^*,k^*)}{\partial k} = 0.$$

这是根据 k^* 是生产 y^* 时最优工厂规模的一阶条件（必要条件）得到的。因此前面那个式子右侧的第二项为零，剩下的只是短期边际成本：

$$\frac{dc(y^*)}{dy} = \frac{\partial c_s(y^*,k^*)}{\partial y}.$$

总结

1.平均成本等于平均可变成本与平均固定成本之和。平均固定成本总是随着产量的增加而下降，而平均可变成本则随着产量的增加最终会上升。这两种效应使得平均成本曲线为 U 型。

2.当平均成本曲线下降时，边际成本曲线位于平均成本曲线的下方；当平均成本曲线上

升时，边际成本曲线位于平均成本曲线的上方。因此，边际成本曲线必然穿过平均成本曲线的最低点，也就是说在这一点上，边际成本和平均成本必定相等。

3.边际成本曲线下方区域的面积等于总可变成本。

4.长期平均成本曲线是短期平均成本曲线的下包络线。

复习题

1.下列哪句话是正确的？

（1）平均固定成本绝不会随着产量的增加而上升；（2）平均总成本总是大于或等于平均可变成本；（3）当边际成本下降时，平均成本绝不会上升。

2.一家企业用两个工厂生产同一种产品。如果工厂 1 的边际成本大于工厂 2 的边际成本，那么企业怎样做才能减少成本，同时又维持产量水平不变？

3.对还是错？在长期，给定任一产量，企业总是选择在最优工厂规模的平均成本最小处生产。

复习题答案

1.下列哪句话是正确的？

（1）平均固定成本绝不会随着产量的增加而上升；（2）平均总成本总是大于或等于平均可变成本；（3）当边际成本下降时，平均成本绝不会上升。

【复习内容】各种成本概念之间的关系

【参考答案】

（1）对；（2）对；（3）对。

理由如下：

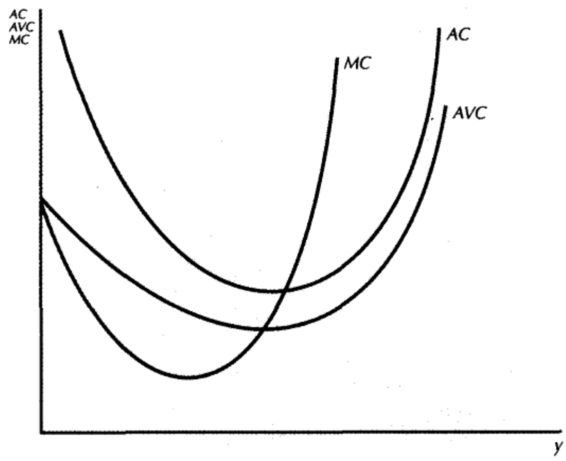
（1） $AFC(y)=F/y$，由于固定成本 F 固定不变，因此随着产量 y 的增加，AFC 必定下降；

(2) $AC(y)=\dfrac{c(y)}{y}=\dfrac{c_v(y)}{y}+\dfrac{F}{y}=AVC(y)+AFC(y)$。因此，平均总成本不可能小于平均

可变成本。

严格来说，只要存在固定成本，则 AFC 再小也不会为零，所以平均总成本会严格大于平均可变成本。但在长期，“固定”成本也成了可变成本，因此此时，平均总成本和平均可变成本相等。

（3）要使平均值下降，新增的数值（边际值）必须小于原来的平均值。因此，边际值下降必然拉低平均值。借助以下图形理解更为直观。



值得思考的是这样的问题：边际成本上升时，平均成本也上升吗？（答案：未必。这取决于边际成本是大于还是小于平均成本，若边际成本小于平均成本，即使边际成本上升（只要还小于平均成本）则平均成本仍然是下降的。）

2.一家企业用两个工厂生产同一种产品。如果工厂 1 的边际成本大于工厂 2 的边际成本，那么企业怎样做才能减少成本，同时又维持产量水平不变？

【复习内容】成本最小化条件

假设你有两个工厂，它们的成本函数不同。假设工厂 1 和 2 的成本函数分别为 $c_1(y_1)$ 和 $c_2(y_2)$。你想以最小的成本生产 y 单位产品。两个工厂各应该生产多少？

$$\min_{y_1,y_2} c_1(y_1)+c_2(y_2)$$

$$\text{使得 } y_1+y_2=m\,.$$

可以证明，在最优产量分配处，工厂 1 的边际成本必定等于工厂 2 的边际成本。为了证明这一点，假设边际成本不相等；则你肯定会将一些产量从边际成本较高的工厂转移到边际成本较低的工厂。如果产量分配已达到最优，那么将产量从一家工厂移动到另一家就不可能降低成本。

【参考答案】

根据以上知识可知，最优产量分配决策处，必须有 $MC_1(y_1^*)=MC_2(y_2^*)$。

由题意知，$MC_1(y_1)>MC_2(y_2)$，因此，此时应该将工厂 1 的产量减少，相应增加到

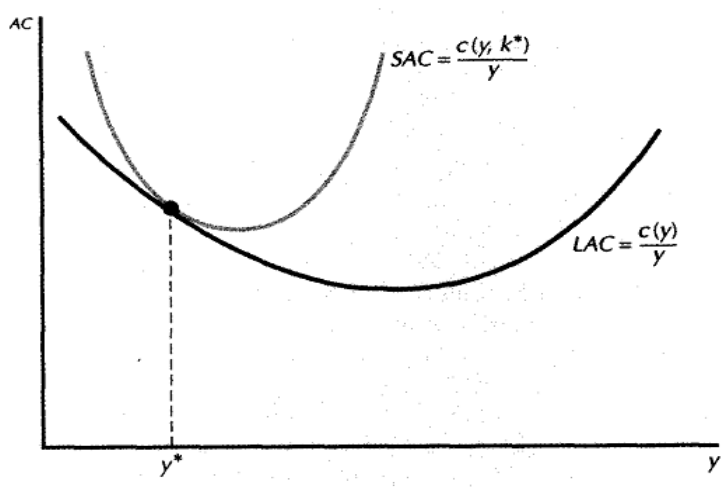
工厂 2，直到两个工厂的边际成本相等。

3.对还是错？在长期，给定任一产量，在长期，给定任一产量，企业总是选择在最优工厂规模的平均成本最小处生产。

【复习内容】长期平均成本和短期平均成本之间的关系

【参考答案】这种说法是错误的。

长期平均成本曲线是短期平均成本曲线的下包络线。如下图所示：



任给一产量 y^*（见上图），厂商无法在短期平均成本曲线 SAC 的最低点进行生产。事实上，在长期平均成本曲线下降阶段，长期成本曲线只能切于短期成本曲线的左侧，而无法切于短期成本曲线最低点。因此，题目中的说法错误。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

22.企业供给（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

22 企业供给

在本章，我们将学习如何从竞争性企业的成本函数推导出它的供给曲线。做此事，我们需要使用利润最大化模型。我们首先分析企业经营所处的市场环境。

22.1 市场环境

每个企业都面对着两个重要的决策：选择产量和选择产品价格。如果对利润最大化企业没有任何约束，它就可以随意定价和随意生产。但是，不受任何约束的企业显然是不存在的。一般来说，企业的经营活动面对着两种约束。

第一种约束是**技术约束**（technological constraints）。我们用生产函数表示这种约束。不是每种投入和产出的组合都是可行的，即使最贪求利润的企业也必须尊重现实世界的约束。我们已经讨论过如何用生产函数表示技术约束，我们也已知道技术如何约束如何变为由成本函数刻画的**经济约束**（economic constraints）。

第二种约束是**市场约束**（market constraint）。企业可以随意决定生产多少产品，也可以随心所欲地定价，但是产品的销量不是由它说了算，而是由消费者的购买意愿决定的。

如果企业的产品定价为 p 时，可以销售 x 单位产品，我们将这种价格和销量之间的关系称为**企业面对的需求曲线**（demand curve facing the firm）。

如果市场中只有一家企业，企业面对的需求曲线就非常容易描述：它就是前几章消费者理论中介绍过的市场需求曲线。因为市场需求曲线衡量在每个价格水平上人们想要购买的商品数量。因此，若市场上只有一家企业，则需求曲线就是这个企业面对的市场约束。

但是如果市场上还有其他企业，那么单个企业面临的约束就不再与上面相同。在这种情形下，每个企业在做价格和产量决策时，都要考虑（猜测）市场中**其他**企业的行为。

无论对于企业还是经济学家来说，这都不是个容易解决的问题。市场约束有多种可能性，我们将系统地进行分析。我们用**市场环境**（market environment）来描述企业在做价格和产量决策时对其他企业行为的反应方式。

在本章，我们将研究最简单的市场环境，即**纯粹竞争**（pure competition）或称为完全竞争。这种市场环境本身很有意义，而且它提供了其他市场环境参照的标准样本。下面我们首先给出纯粹竞争的经济学定义，然后我们再给出这种称呼的理由。

22.2 纯粹竞争

在外行人看来，“竞争”蕴含着积极且激烈对抗的意义。这也是为什么学生对竞争的经济学定义感到惊讶的原因——在经济学里，竞争竟然是如此被动的：**纯粹竞争的**（purely competitive）市场是指市场价格和每个企业的产量无关的市场。因此，在竞争性的市场中，每个企业只需要关心它自己想要生产的产量即可。无论它生产多少，都只能按照唯一的价格出售，这个唯一的价格就是当前的市场价格。

在什么样的市场环境下，纯粹竞争的假设才是合理的？比如假设某行业由生产相同产品的很多企业组成，每个企业占市场总量的份额很小。以小麦市场为例。小麦市场上有成千上万个麦农，即使是产量最大的农民，他的产量占市场总供给量的比重也是极其小的。在这种情形下，自然可以认为行业中的每个企业都把市场价格看作是预先决定的。某个麦农不必担心如何对小麦定价——如果他想卖小麦，他必须按当前的市场价格出售。他是个**价格接受者**（price taker）:对他来说市场价格是固定不变的；他关心的只是产量的多少。

因此，若市场上企业数量众多、规模很小而且生产相同的产品，我们可以把这些企业作为价格接受者。但是，在其他情形下，也有可能出现价格接受的行为。即使市场中只有少数几家企业，每个企业仍然认为市场价格不受它的控制。

我们考虑下面的情形，即某种易腐烂的商品供给量既定的情形：例如鲜鱼或插花用的鲜花的市场。即使市场中只有三四个企业，每个企业也只能认为**其他**企业的价格是既定的。如果商品同质，则谁的商品价格最低，消费者就愿意买谁的。因此，如果一家企业把价格定的最低，那么这个价格就是市场价格。如果**其他**的企业想卖产品，这些企业必须接受上述市场价格。因此，在这种情形下的竞争行为——市场价格不受你的控制——似乎也是合理的。

我们可以用图形描述某竞争性企业感知的价格和销量之间的关系，如图 22.1 所示。你可以看到，这个需求曲线非常简单。一个竞争性企业坚信，如果它的售价高于市场价格，则它的销量为零；相反，如果它的售价低于市场价格，那么它有多少产量就能卖掉多少，在这个较低的价格水平上，它就能获得全部的市场需求。

和以前一样，我们可以从两个角度分析需求曲线。如果将数量看成价格的函数，则这条曲线表示的是，当你的定价等于或者低于市场价格时你能卖掉的商品数量。如果将价格看成数量的函数，则这条曲线表示：不管你的销量如何，市场价格都和你的销量无关。

（当然，这里的“任何”数量不是指字面上的**任何**数量。这里的“任何”是说，对于你所希望销售的数量来说，价格和你的销量无关。例如，在鲜花的例子中，任何数量是指你手头存货以下的数量，在这个数量范围，价格和销量无关。）

一定注意区分“企业面对的需求曲线”和“市场需求曲线”。市场需求曲线衡量市场价格和市场总销量之间的关系；企业面对的需求曲线衡量市场价格和**某个特定企业**销量之间的关系。

市场需求曲线取决于消费者的行为。一家企业面临的需求曲线不仅取决于消费者的行

为，还取决于其他企业的行为。通常，如果市场里企业数量众多，而且每家企业面对的需求曲线都基本水平时，则可以认为是纯粹竞争的市场。但是，即使市场里只有两家企业，如果他们生产的是同质产品，而且一家企业坚持索要固定的价格，则另外一家企业面对的需求曲线就是图 22.1 所示的纯竞争需求曲线。因此，竞争性的行为模型的适用范围显然比你最初认为的要广泛。

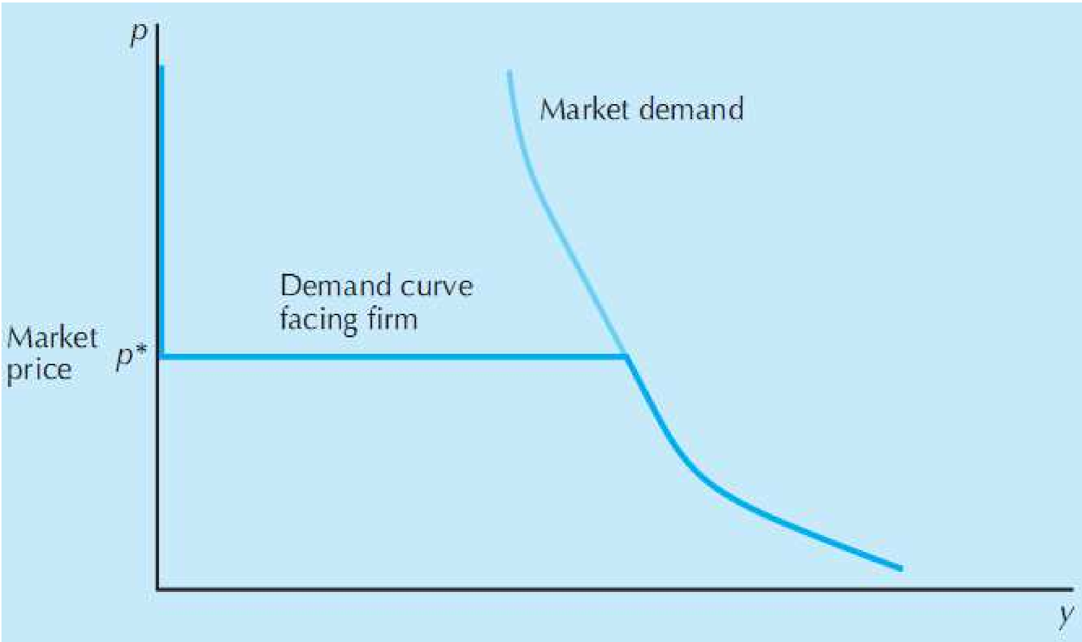


图 22.1：竞争性企业面对的需求曲线。若企业的定价等于市场价格，则它的需求曲线是一条水平线；如果它的定价高于市场价格，则它的销量为零；如果它的定价低于市场价格，则它面对的就是全部的市场需求。

22.3 竞争性企业的供给决策

下面我们使用成本曲线来推导竞争性企业的供给曲线。由定义可知，竞争性企业不会考虑它对市场价格的影响，即市场价格 p 是既定的。因此，竞争性企业面对的利润最大化问题为

$$\max_{y} \quad py - c(y).$$

这不过是说竞争性企业希望使得利润最大：使它的收入 py 和成本 $c(y)$ 之差最大。

竞争性企业会选择什么样的产量水平？答案：它会在边际收入等于边际成本处生产。边际收入等于边际成本是说，额外增加一单位产量所带来的**额外**收入恰好等于该单位产量产生的**额外**成本。如果企业选择的产量不满足上述条件，则改变产量就可以增加利润。

对于竞争性企业来说，边际收入就等于价格。为了看清这一点，问问你自己若企业产量

增加了 Δy，销售收入因此增加了多少。由于你还知道产品价格 p 是不变的，因此必有

$$\Delta R = p\Delta y .$$

因此，每单位产量的额外收入为：

$$\frac{\Delta R}{\Delta y} = p ,$$

这就是边际收入的表达式。

因此，竞争性企业选择的产量 y，要恰好使得在该产量上的边际成本等于产品的市场价格，即：

$$p = MC(y) .$$

对于既定的市场价格 p，我们想找到能使利润最大的产量水平。若在某产量 y 上，价格大于边际成本，则企业可以多生产一些，这样利润就会增加。其中的道理在于，价格大于边际成本意味着

$$p - \frac{\Delta c}{\Delta y} > 0 .$$

因此，增加产量 Δy 意味着

$$p\Delta y - \frac{\Delta c}{\Delta y}\Delta y > 0$$

化简可得

$$p\Delta y - \Delta c > 0 ,$$

这个式子是说，增加产量带来的额外收入大于因此增加的成本，所以利润必然增加。

当价格小于边际成本时，我们也可以做类似的论证。在这种情形下，减少产量将增加利润，这是由于减少产量后，与利润减少幅度相比，成本减少得更多。

因此，在最优产量水平上，价格必然等于边际成本。不管市场价格 p 为多大，企业选择的产量 y 都要满足 $p = MC(y)$。因此，**竞争性企业的边际成本曲线正是它的供给曲线**。换句话说，市场价格恰好等于边际成本——只要每个企业都在利润最大化产量水平上生产。

22.4 一种例外情形

上述结论未必那么准确，因为存在着让人讨厌的例外情形。一种例外情形是，满足价格等于边际成本条件的产量有好几个，如图 22.2 所示。由图可以看出，此时有两个产量都能使价格和边际成本相等。企业应该选择哪个产量？

曹乾（东南大学 caoqianseu@163.com）

答案不难看出。考虑第一个交点，此时边际成本曲线是向下倾斜的。现在如果产量稍微增加，边际成本会下降。这就是边际成本递减的含义。但是由于市场价格是固定不变的，因此利润必然上升。

这样我们就排除了边际成本曲线下降阶段的产量选择。在这个阶段，产量增加必然增加利润。竞争性企业的供给曲线必然位于向上倾斜的那部分边际成本曲线。这表示边际供给曲线本身必然向上倾斜。这也意味着，供给曲线不可能出现“吉芬商品”现象，也就是说不可能出现价格上升但产量递减的现象。

价格等于边际成本是利润最大化的**必要**条件，但它一般**不是充分**条件。这就是说，不能因为你找到了一个边际成本等于价格的点，你就断定它是利润最大化的点。但是如果你找到了利润最大化的点，那么必然有价格等于边际成本。

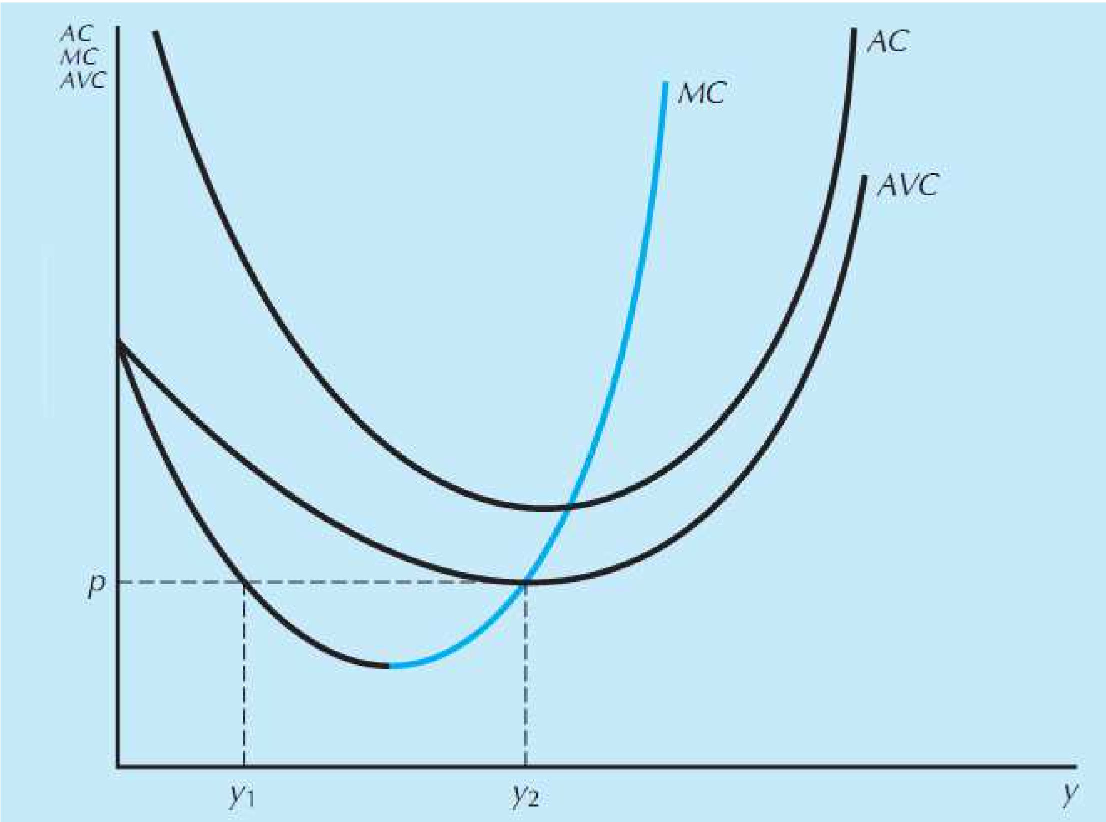


图 22.2：边际成本和供给。尽管满足价格等于边际成本条件的产量有两个，但是利润最大化的供给量只能位于边际成本曲线向上倾斜的那一部分。

22.5 另外一种例外情形

上述讨论隐含着一个假设，即生产比不生产好。但是，完全有可能出现企业不生产是最佳选择的情形。既然，生产零单位产品也是一种选择，因此我们有必要将利润最大化的候

选点与零产量进行比较。

在短期，如果企业决定生产零单位产品，那么它仍然必须支付固定成本 F 。因此，产量为零时的利润为 $-F$ 。如果企业生产的产量为 y ，则其利润为 $py-c_v(y)-F$ 。企业为何会选择停产？因为此时必然有

$$-F > py-c_v(y)-F\,,$$

也就是说，若产量为零时的“利润”（支付的固定成本），大于按价格等于边际成本条件生产所获得的利润，则停产就是最优选择。将上式整理可以得到所谓的**停产条件**（shutdown condition）：

$$AVC(y)=\frac{c_v(y)}{y}>p\,.$$

若平均可变成本大于市场价格，则企业的最佳选择是停产。这完全讲得通，因为如果不停产，那么销售 y 单位产品带来的收入 py ，还不足以补偿**可变**成本 $c_v(y)$ ，即 $py<c_v(y)$ 。在这种情形下，企业显然会选择停产。如果它不生产，它损失的是固定成本，但是如果生产它损失的更多。

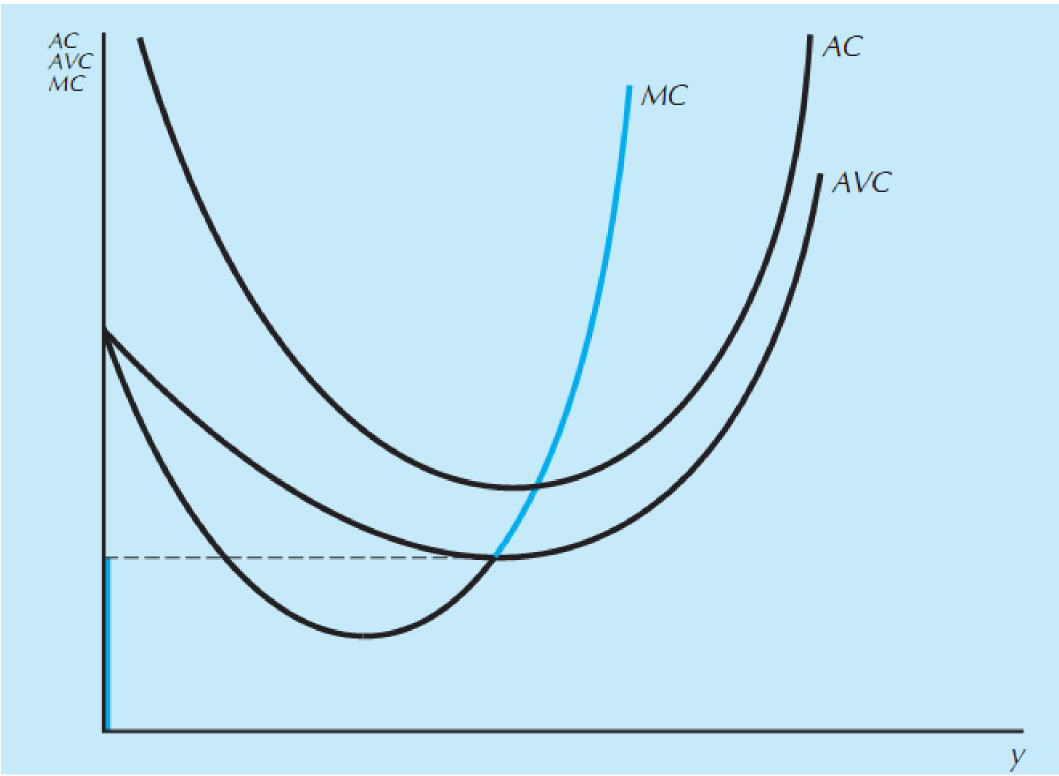


图 22.3：平均可变成本和供给。竞争性企业的供给曲线是位于平均成本曲线以上的那段向上倾斜的边际成本曲线。它不会在边际成本曲线位于平均可变成本曲线以下的那些点上生产，因为此时停产能得到更大的利润（其实是避免更多的损失）。

上述结论表明，供给曲线只能是边际成本曲线的一部分，准确地说，供给曲线是位于平均可变成本曲线上方的那部分边际成本曲线。如果价格等于边际成本的点位于平均可变成本曲线的下方，则企业的最优选择就是产量为零。

这样，我们就得到了竞争性企业的供给曲线，如图 22.3 所示。竞争性企业必然沿着边际成本曲线上的一段进行生产，该段边际成本曲线向上倾斜而且还要位于位于平均可变成本曲线以上。

例子：计算机操作系统的定价

计算机只有安装了操作系统才能运行，大多数硬件制造企业在销售计算机之前已在机子上安装了操作系统。在 1980 年代早期，IBM 的 PC 兼容机市场占有量很高，几家操作系统企业为了攀上 IBM 这棵大树，竞争得你死我活。当时最常见的做法是，计算机制造企业每安装一套操作系统，就要支付一套的钱。

微软公司独辟蹊径，它按硬件企业生产的计算机数量计算费用。微软公司将专利使用费定得非常低，因此对硬件企业非常由吸引力。

注意微软公司定价策略的聪明之处：一旦与硬件企业签订了合同，则在已生产出的计算机上安装操作系统的边际成本为零。但若安装竞争对手的操作系统，则每套成本约为 50~100 美元。硬件企业（最终是消费者）为微软的操作系统买单，但是由于微软的定价合同，使得它的操作系统比竞争对手更有竞争力。结果，微软的操作系统成为计算机的默认操作系统，占据了 90%以上的市场份额。

22.6 反供给曲线

我们已经知道，竞争性企业的供给曲线由价格等于边际成本这个条件决定。和以前一样，我们可以将价格和产量之间的关系用两种方法表达：一是将产量作为价格的函数，这也是我们通常的做法；二是将价格作为产量的函数，这就是反供给函数。使用后面这种方法，我们的眼界可能更宽阔些。

由于供给曲线上的每一点都满足价格等于边际成本的条件，我们可以将市场价格作为衡量行业中每个企业边际成本的一种方法。如果行业中一家企业的产量大而另一家企业的产量小，而且它们均实现了利润最大化，那么它们的**边际成本必然相等**。这是因为它们面对的市场价格是相等的。

反供给曲线的表达式为 $p = MC(y)$，也就是说价格是产量的函数。供给曲线的这种表达方法很有用。

22.7 利润和生产者剩余

现在，给定市场价格 p，我们就可以根据条件 $p = MC(y)$ 求解企业的最优产量选择。找出了最优产量，我们就可以计算出企业的利润。在图 22.4 中，矩形区域的面积表示总收入 p^*y^*。区域 $y^*AC(y^*)$ 是总成本，这是由于

$$yAC(y) = y\frac{c(y)}{y} = c(y).$$

利润等于这两个区域的面积之差。

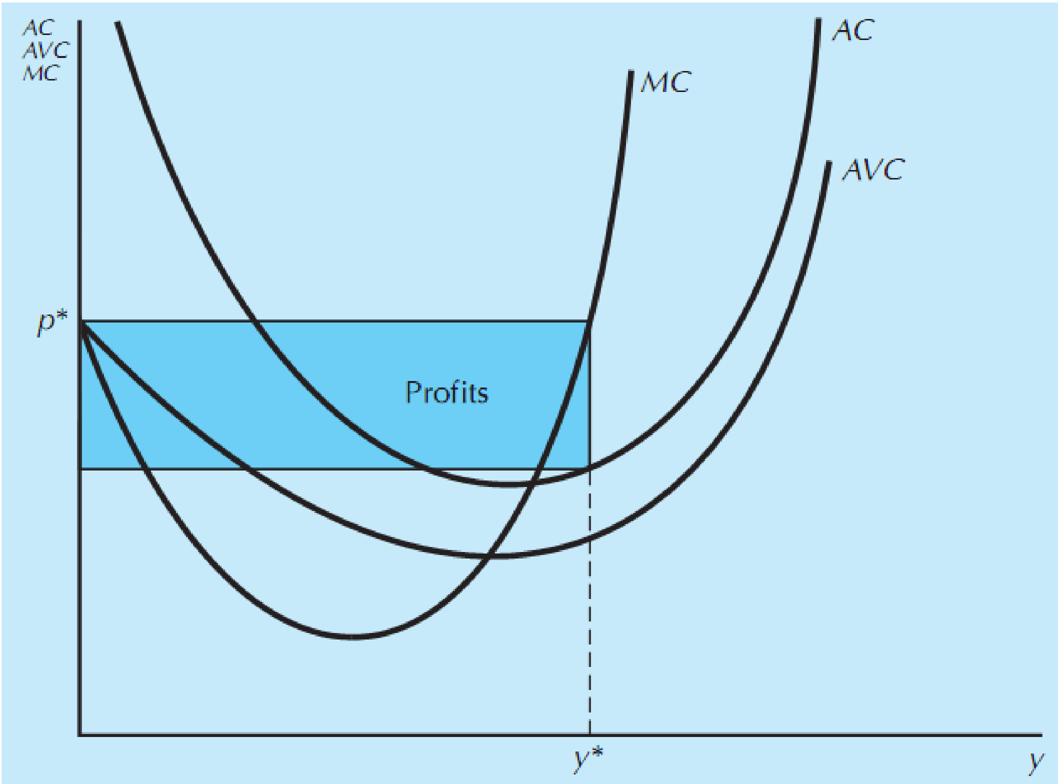


图 22.4：利润。利润等于总收入减去总成本。

请回忆一下我们在 14 章讨论过的生产者剩余。我们将生产者剩余定义为供给曲线左方的面积，这和消费者剩余是需求曲线左边的面积定义类似。可以证明，生产者剩余和利润有着密切的关系。更准确地说，生产者剩余等于收入减去可变成本，或者等价地，等于利润加上固定成本：

$$利润= py - c_v(y) - F$$

$$生产者剩余= py - c_v(y).$$

生产者剩余最直接的计算方法，是用收入矩形减去矩形 $y^*AVC(y^*)$，如图 22.5A 所示。

但是我们也可以使用边际成本曲线本身计算消费者剩余。

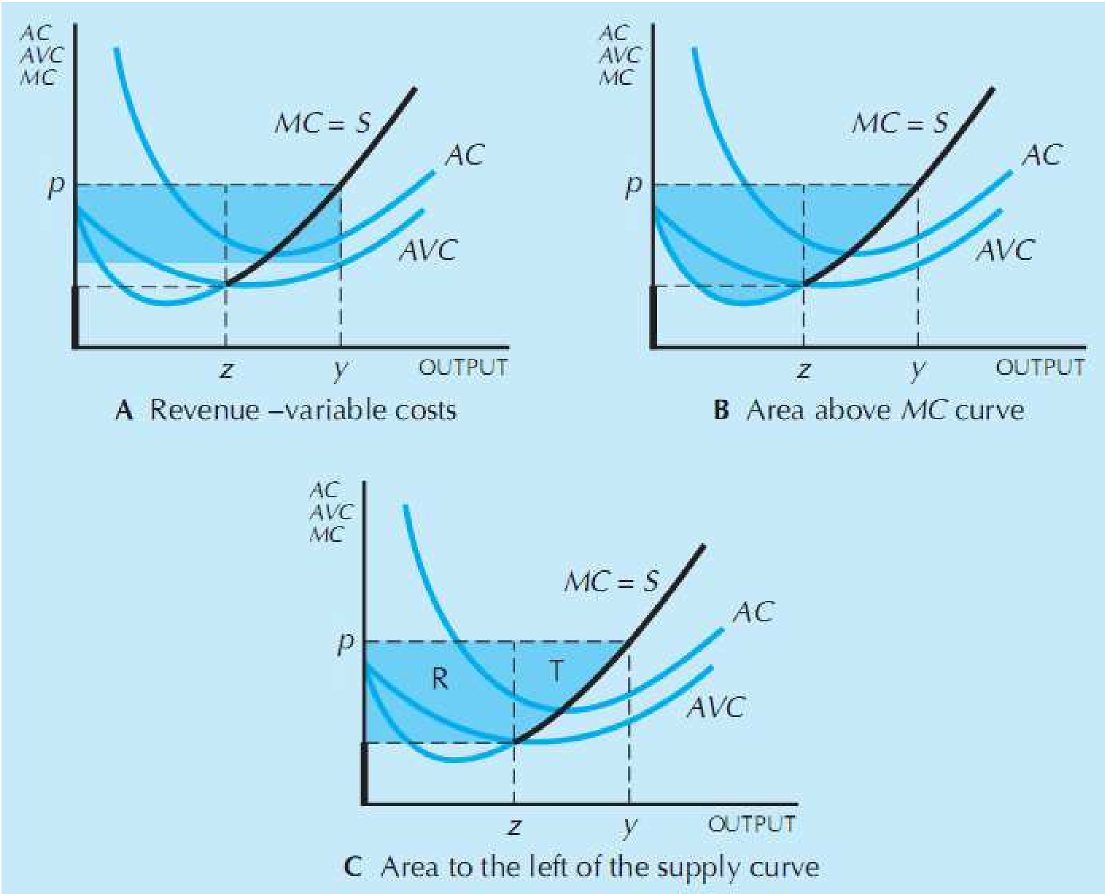


图 22.5：生产者剩余。生产者剩余有三种等价的计算方法，分别如 A 图、B 图和 C 图所示。A 图是用总收入减去总可变成本后的矩形面积表示；B 图用边际成本曲线上方的面积表示；C 图中，在产量 z 之前的生产者剩余为矩形面积 R，产量 z 之后 y 之前的生产者剩余用区域 T 表示，产量为 y 时的生产者剩余等于这两部分之和，即 R+T。

在 21 章我们已经知道边际成本曲线下方的面积衡量总可变成本。这是由于边际成本曲线下方的面积，等于生产第一单位产品的成本加上生产第二单位产品的成本，一直加下去。因此，为了得到生产者剩余，我们可以从收入矩形中减去边际成本曲线下方的面积，剩下的区域就是生产者剩余，如图 22.5B 所示。

最后，我们可以将前两种方法结合起来，这也是计算生产者剩余的一种方法。在边际成本等于平均可变成本之前的产量，我们使用第一种方法计算生产者剩余，这是个矩形区域。在此之后，我们使用第二种方法，即边际成本曲线上方的区域计算生产者剩余，如图 22.5C 所示。第三种方法最方便，因为它正好是供给曲线左侧的面积。注意，这种方法恰好和 14 章生产者剩余的定义是一致的。

我们很少对生产者剩余的**总额**感兴趣，更多时候我们关注的是生产者剩余的**变动**。当企业的产量从 y^* 变为 y' 时，由此引起的生产者剩余的变动通常为梯形形状，如图 22.6 所示。

注意，产量从 y^* 变为 y' 时，生产者剩余的变动额恰好等于利润的变动额，因为根据定义可知固定成本是固定不变的。因此，我们可以从边际成本曲线蕴含的信息，去推测产量变动对企业利润的影响，这种情形下我们根本无需使用平均成本曲线。

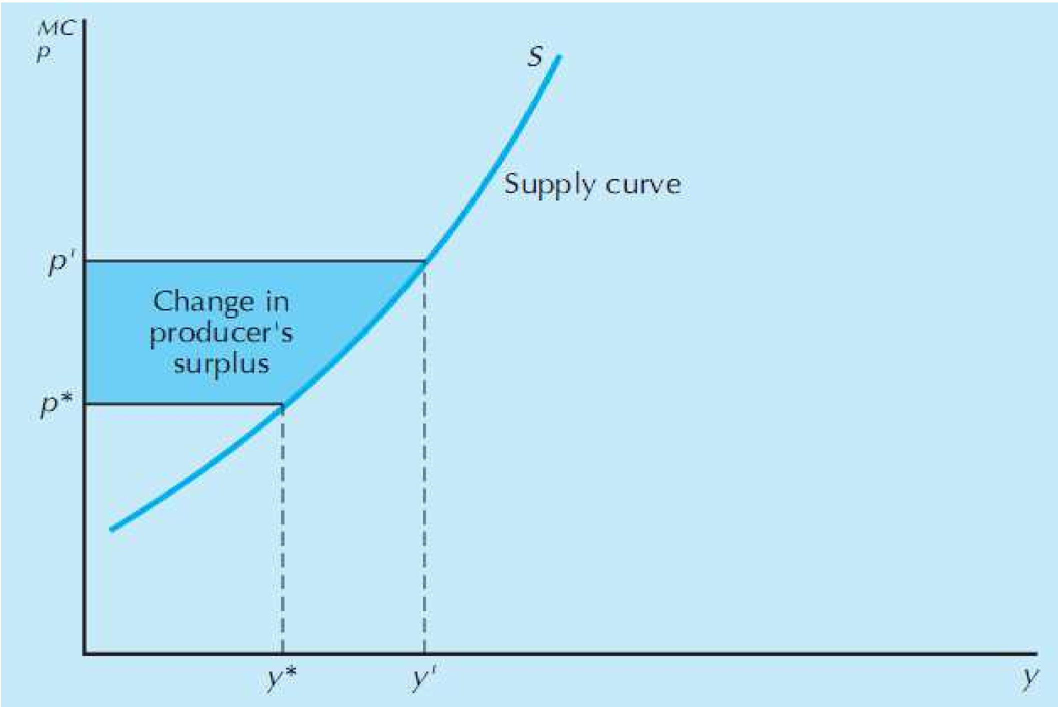


图 22.6：生产者剩余的变动。由于供给曲线和向上倾斜的那一段边际成本曲线重合，生产者剩余的变动通常为梯形的形状。

例子:具体生产函数的供给曲线

在上一章的成本函数 $c(y)=y^2+1$ 中，供给曲线是什么样子的？在该例中，边际成本曲线总是在平均可变成本曲线上方，而且它总是向上倾斜的，因此“价格等于边际成本”直接提供了供给曲线。将边际成本等于 $2y$ 代入这个条件，可得

$$p=2y.$$

这样我们就得到了反供给曲线，即将价格作为产量的函数。若将产量作为价格的函数，可得

$$S(p)=y=\frac{p}{2}.$$

如图 22.7 所示。

如果我们将此供给函数代入利润的定义，就可以计算出每个价格水平下的最大利润。

我们来计算一下：

$$\begin{aligned}\pi(p) &= py - c(y)\\ &= p\frac{p}{2} - (\frac{p}{2})^2 - 1\\ &= \frac{p^2}{4} - 1\end{aligned}$$

最大化利润和生产者剩余之间有什么关系？在图 22.7 中，我们可以看到生产者剩余——供给曲线左方且在价格 0 和价格 p 之间的面积——是一个底为 $y = p/2$、高为 p 的三角形。这个三角形的面积为

$$A = (\frac{1}{2})(\frac{p}{2})p = \frac{p^2}{4}.$$

将其余利润表达式比较，可知生产者剩余等于利润加上固定成本，这和我们在前面的断言一致。

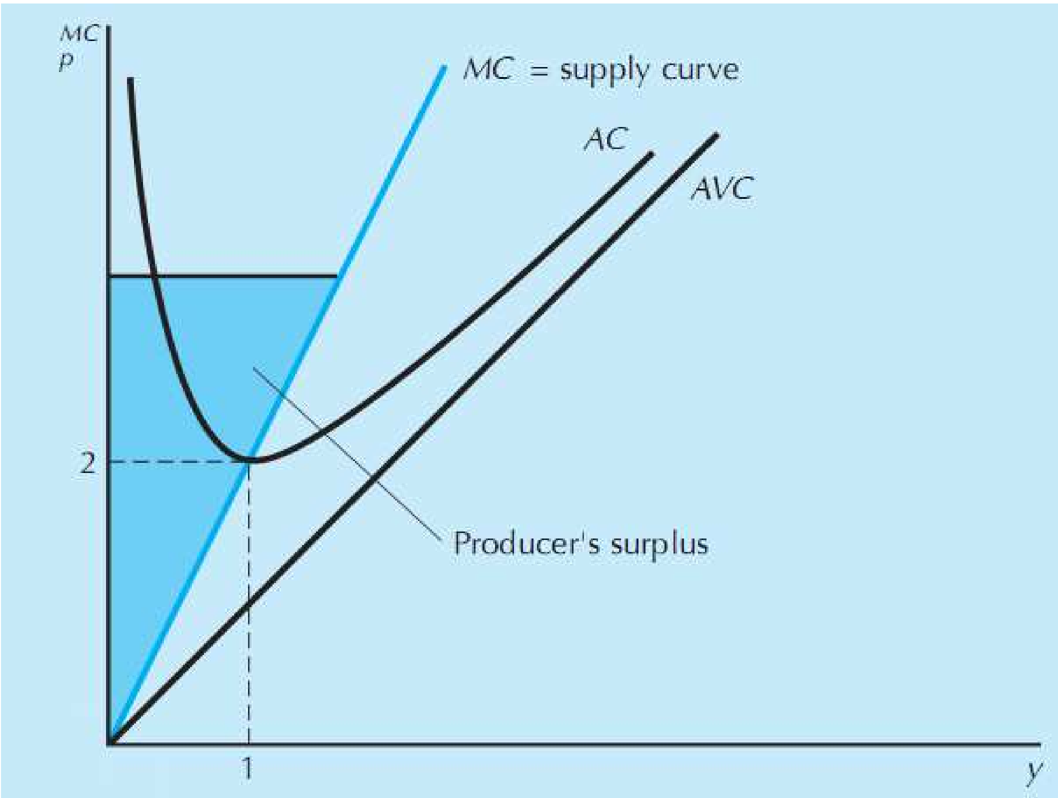


图 22.7：供给曲线的一个具体例子。成本函数 $c(y) = y^2 + 1$ 的供给曲线和生产者剩余。

22.8 企业的长期供给曲线

企业的长期供给曲线，衡量在工厂规模（或其他任意固定要素）可调整的情形下，企业如何选择最优的产量。也就是说，企业的长期供给曲线为

$$p = MC_l(y) = MC(y, k(y)).$$

该企业在工厂规模为 k 时的短期供给曲线为：

$$p = MC(y, k).$$

注意这两个表达式的区别。短期供给曲线涉及的是工厂规模 k 固定时的边际成本；而长期供给曲线涉及的是工厂规模 k 已调整到最优时的边际成本。

现在我们对短期和长期边际成本的关系有了一定的了解：短期边际成本和长期边际成本必然在某个产量（y^*）处相等，这种情形是一种巧合，也就是说企业在长期选择的最优工厂规模 k^*，恰好就是企业在短期就拥有的工厂规模。因此，在这种情形下，短期供给曲线和长期供给曲线必然交于 y^*，如图 22.8 所示。

在短期，企业的某些生产要素是固定不变的，无法进行调整；而在长期这些要素都是可变的。因此，当产品的价格变动时，企业在长期内调整的余地显然比短期要大。这意味着长期供给曲线对价格的反应更敏感，即长期供给曲线比短期供给曲线的弹性更大，如图 22.8 所示。

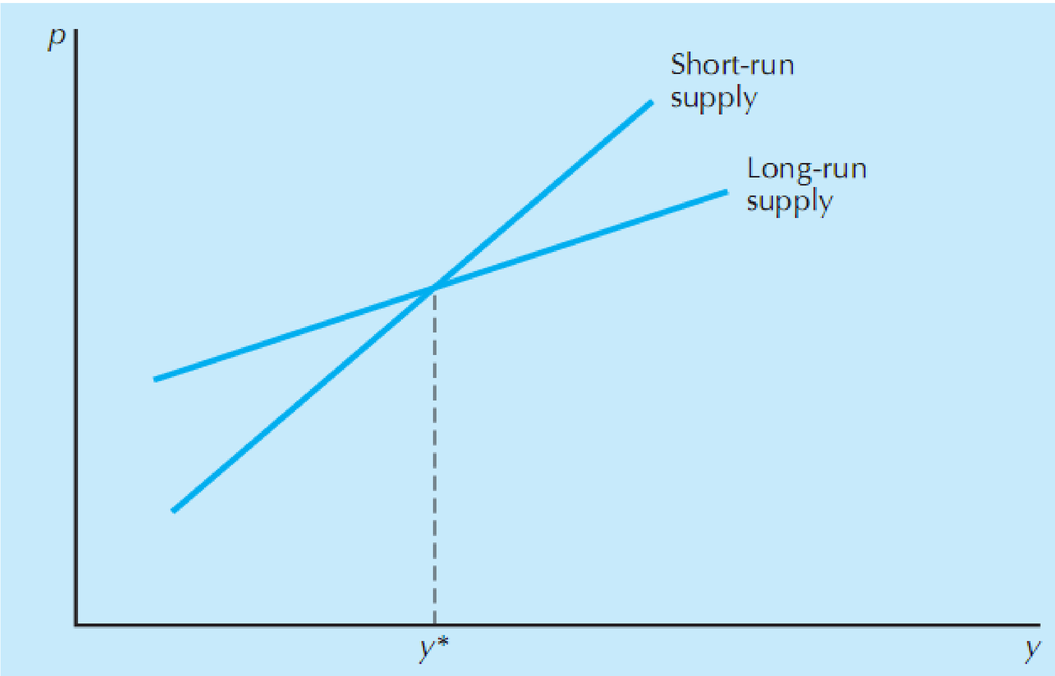


图 22.8：短期供给曲线和长期供给曲线。长期供给曲线通常比短期供给曲线更有弹性。

长期供给曲线还有什么其他特征？我们知道长期是指企业能够自由调整所有生产要素的时间段。企业的决策中包含着是否继续开业这种选择。由于在长期，企业总可以生产零单位产品，即退出它所在的行业，因此，企业在长期均衡时的利润必然非负：

$$py - c(y) \geq 0,$$

这意味着

$$p \geq \frac{c(y)}{y}.$$

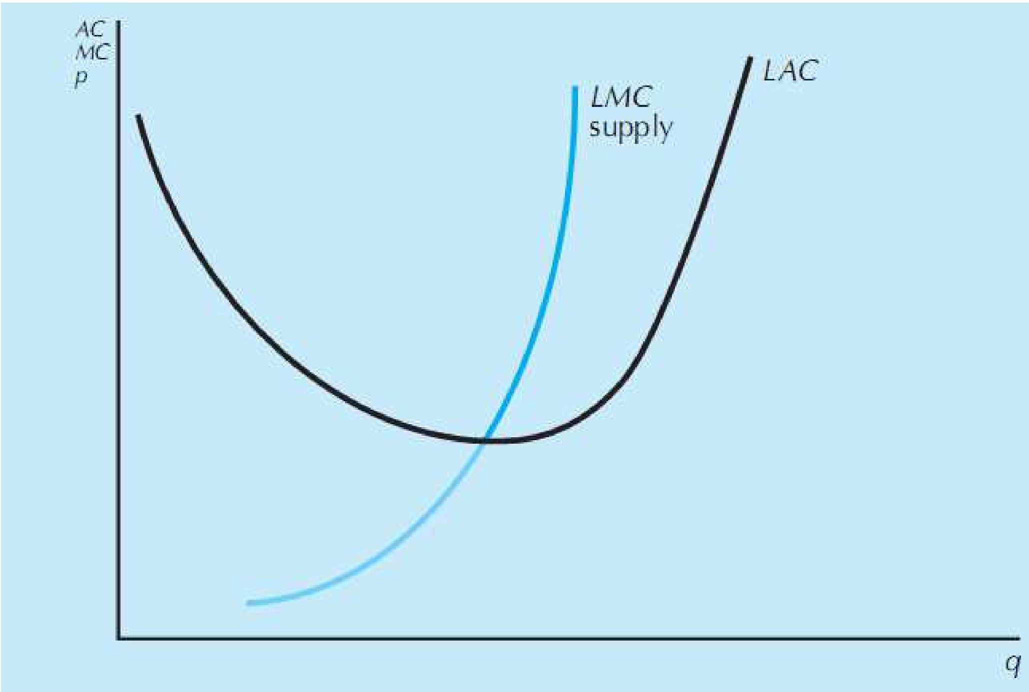


图 22.9：**长期供给曲线**。长期供给曲线是平均成本曲线上方的那一段边际成本曲线。

这就是说在长期，价格必然大于或等于平均成本。因此长期供给曲线是位于平均成本曲线上方的那一段边际成本曲线，这段边际成本曲线的斜率显然为正。如图 22.9 所示。

这一点短期供给曲线几乎完全相同，在长期，所有的成本都是可变成本，因此在短期要求价格高于平均可变成本，等价于在长期要求价格高于平均成本。

22.9 长期平均成本不变的情形

当企业的长期生产技术为规模报酬不变时，就会出现一个有趣的现象。此情形下的长期供给曲线就是长期平均成本曲线。这是因为，长期供给曲线是长期边际成本曲线(一部分)，但在规模报酬不变（从而平均成本不变）的情况下，长期边际成本曲线与长期平均成本曲线重合。因此，我们就得到了图 22.10 的情形。由图可知，长期供给曲线是一条纵截距为 $c_{\min}$ 的水平线，其中 $c_{\min}$ 表示不变的平均成本。

这条供给曲线表示：$p = c_{\min}$ 时，企业愿意供给任何数量的产品；$p > c_{\min}$ 时，企业愿

意供给任意大的产量；$p < c_{\min}$ 时，企业的供给量为零。理解上述结论需要回忆一下规模报酬不变情形下的复制企业论：比如所有要素变为原来的 2 倍，相当于又复制了一个原来的企业。

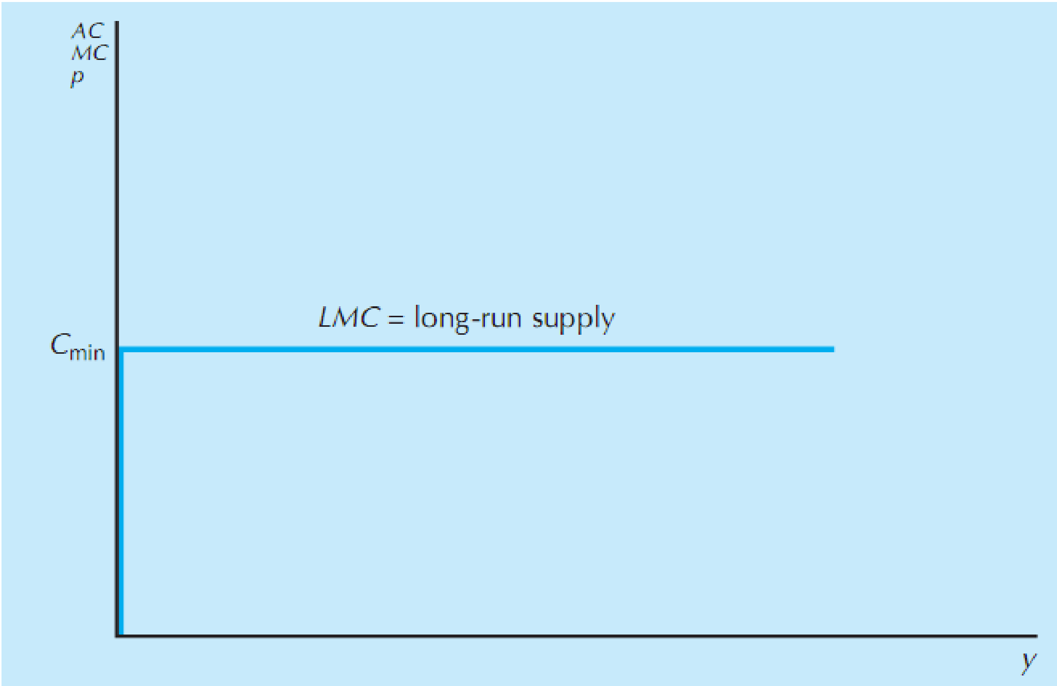


图 22.10：长期平均成本不变的情形。在平均成本不变的情形下，长期供给曲线是一条水平线。

对于此处的情形来说，规模报酬不变意味着，如果要素成本为 $c_{\min}$ 时你可以生产 1 单位产品，则如果要素投入量为原来的 n 倍，即要素成本为 $nc_{\min}$ 时，你可以生产 n 单位产品。因此，如果价格等于 $c_{\min}$，你愿意供给任何数量的产品；若价格大于 $c_{\min}$，你愿意供给的产品可能任意大。

另一方面，如果 $p < c_{\min}$，只要你一生产就意味着亏损，因此该情形下，你的最优选择就是供给量为零。 □

附录

若使用微积分，本章的讨论将非常简单。利润最大化问题是

$$\max_{y} \quad py - c(y)$$

$$\text{使得 } y \geq 0 .$$

最优供给 y^* 的一阶条件为

$$p - c'(y^*) = 0$$

二阶条件为

$$-c''(y^*) \leq 0 .$$

一阶条件是说价格等于边际成本；二阶条件是说边际成本曲线必然处于上升阶段。当然这两个条件是假设 $y^* > 0$ 时得到的。如果在 y^* 处，价格小于平均可变成本，则企业的最优选择是产量为零。为了确定竞争性企业的供给曲线，我们必须找到满足一阶条件和二阶条件的所有点，并且互相进行比较，然后再与 $y = 0$ 时的利润进行比较，从中选择利润最大的点即可。这即是利润最大化的供给。

总结

1.企业对其产品索要的价格和销售量之间的关系，称为企业面临的需求曲线。根据定义，竞争性企业面对的需求曲线是一条水平线，它的高度由市场价格决定，即由其他的企业索要的价格决定。

2.竞争性企业的（短期）供给曲线是它的（短期）边际成本曲线的一部分，准确地说是位于平均可变成本曲线上方的那部分边际成本曲线。

3.当市场价格由 p_1 变为 p_2 时，由此引起的生产者剩余变动，等于边际成本曲线左侧 p_1 和 p_2 这两条价格线之间区域的面积。这个面积同时衡量了企业利润的变动。

4.企业的长期供给曲线是长期边际成本曲线的一部分，准确地说，是位于长期平均成本曲线上方的那一段边际成本曲线。

复习题

1.某企业的成本函数为$c(y)=10y^2+1000$，求其供给曲线。

2.某企业的成本函数为$c(y)=10y^2+1000$，在产量为多大时它的平均成本为最小？

3.若供给曲线为$S(p)=100+20p$，求反供给曲线。

4.某企业的供给函数为$S(p)=4p$，它的固定成本为 100。若价格由 10 上升为 20，求利润的变动量。

5.若某企业的长期成本函数为$c(y)=y^2+1$，求它的长期供给曲线。

6.以下哪些是技术约束哪些是市场约束：生产要素的价格；市场上其他企业的数量；产量；在投入既定情形下能生产更多的能力。

7. 哪个假设能描述纯粹竞争市场的主要特征？

8.在纯粹竞争市场中，企业的边际收入总是等于什么？利润最大化的企业会选择什么样的产量？

9.如果平均可变成本大于市场价格，企业生产的产量为多大？如果没有固定成本，企业的产量为多大？

10.某竞争性企业即使亏损也要生产。会不会出现这种情形？为什么？

11.在完全竞争市场中，对于行业中的所有企业来说，市场价格和生产成本之间有什么关系？

复习题答案

1.某企业的成本函数为$c(y)=10y^2+1000$，求其供给曲线。

【复习内容】竞争企业的供给曲线

竞争企业的利润最大化问题是

$$\max_{y} \quad py-c(y)$$

$$\text{使得 } y\geq 0 .$$

最优供给y^*的一阶条件为

$$p-c'(y)=0\text{，即 } p=MC(y)$$

二阶条件为

$$-c''(y) \le 0.\text{即}\, c''(y) \ge 0$$

一阶条件是说价格等于边际成本；二阶条件是说边际成本曲线必然处于上升阶段。当然这两个条件是假设 $y^* > 0$ 时得到的。如果在 y^* 处，价格小于平均可变成本，则企业的最优选择是产量为零。因此，在求企业短期供给曲线时，必须验证 $p \ge AVC(y)$ 。

【参考答案】

竞争企业的供给曲线就是其边际成本曲线，准确地说是位于平均可变成本曲线上方的那一段边际成本曲线。

由 $c(y) = 10y^2 + 1000$ 可得边际成本曲线为 $MC(y) = 20y$ 。

竞争企业的利润最大化的一阶条件为 $p = MC(y)$

由上面两个式子可知企业的反供给曲线为 $p = 20y$ ，但还需要验证 $p \ge AVC(y)$ ，这其实是验证 $MC(y) \ge AVC(y)$ ，因为 $p = MC(y)$ 。如果 $MC(y) \ge AVC(y)$ ，就可以保证供给曲线边际成本曲线在成本曲线上方的那一段。

由 $c(y) = 10y^2 + 1000$ 可得平均可变成本曲线 $AVC(y) = \dfrac{TVC(y)}{y} = \dfrac{10y^2}{y} = 10y$ 。

因为 $p = 20y$ ，而 $AVC(y) = 10y$ ，所以 $p > AVC(y)$ 。

因此，可以确定企业的反供给曲线为 $p = 20y$ ，从而企业的供给曲线为 $y = p/20$ 。

2.某企业的成本函数为 $c(y) = 10y^2 + 1000$ ，在产量为多大时它的平均成本为最小？

【复习内容】平均成本

【参考答案】

由 $c(y) = 10y^2 + 1000$ 可知平均成本函数为 $AC(y) = \dfrac{c(y)}{y} = 10y + \dfrac{1000}{y}$

【解法一】

利用极值的一阶条件。

将平均成本函数对 y 求导并令其等于零，可得

$$\frac{dAC(y)}{dy} = 10 - \frac{1000}{y^2} = 0 \Rightarrow y = 10.$$

【解法二】

由平均成本和边际成本的关系可知，当边际成本等于平均成本时，平均成本达到最小值。

由 $c(y) = 10y^2 + 1000$ 可知边际成本函数为 $MC(y) = 20y$ 。

$MC(y)=AC(y)\Rightarrow 20y=10y+\frac{1000}{y}\Rightarrow y=10$。

3.若供给曲线为 $S(p)=100+20p$，求反供给曲线。

【复习内容】供给曲线和反供给曲线

供给曲线：产量是价格的函数；而反供给曲线：价格是产量的函数。

【参考答案】

令 $y=S(p)=100+20p$，可求出 $p=\frac{y-100}{20}$，这就是反供给曲线。

4.某企业的供给函数为 $S(p)=4p$，它的固定成本为 100。若价格由 10 上升为 20，求利润的变动量。

【复习内容】生产者剩余和利润之间的关系；生产者剩余变动量和利润变动量的关系

A 生产者剩余和利润之间的关系

生产者剩余等于收入减去可变成本，或者等价地，等于利润加上固定成本：

利润= $py-c_v(y)-F$

生产者剩余= $py-c_v(y)$.

生产者剩余= $\pi+F$，其中 π 表示利润

B 生产者剩余变动量和利润变动量的关系

由于生产者剩余等于利润加上固定成本，因此

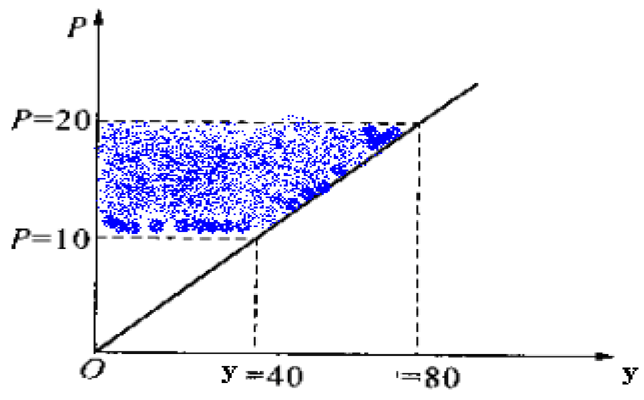
$\Delta PS=\Delta\pi$，这就是说生产者剩余变动量等于利润变动量。

【参考答案】

题目分析：由于题目要求我们求出利润变动量，但题目未告知成本函数，也就是说尽管我们可以求出收入，但因为缺乏成本信息，无法直接使用收入与成本之差这种利润计算方法。

但是，由于利润的变动等于生产者剩余变动，我们可以转而求生产者剩余变动。

由企业的供给函数 $S(p)=4p$ 和题目中的价格条件可以画出下图。当价格为 10 时，供给量为 40，当价格为 20 时供给量为 80.



我们知道生产者剩余是供给曲线以上价格线以下的面积。因此，从上图中可以看出，当价格从 10 上升为 20 时，生产者剩余增加了，增加的部分如阴影部分的面积所示。这是一个梯形。只要求出该梯形的面积，就可以计算出生产者剩余增加量，从而也就是利润增加量。

该梯形的面积= $\frac{1}{2}(80+40)\times(20-10)=600$.

因此，利润增加了 600.

5.若某企业的长期成本函数为 $c(y)=y^2+1$，求它的长期供给曲线。

【复习内容】竞争企业的长期供给曲线

企业的长期供给曲线是长期平均成本曲线最低点以上的那部分长期边际成本曲线。

这是因为根据竞争企业利润最大化的一阶条件可知： $p=MC(y)$ 。

另外，在长期中，如果企业亏损，那么企业就会退出市场。这意味着 $py-c(y)\geq 0$ ，这个式子两边同除以 y 可得 $p\geq AC(y)$ 。

所以在求企业长期供给曲线时，必须验证 $p\geq AC(y)$ 。

【参考答案】

【解法一】

由 $c(y)=y^2+1$ 可知企业的长期边际成本曲线 $MC(y)=2y$

由竞争企业利润最大化的一阶条件可知： $p=MC(y)$

因此，企业的长期反供给曲线为 $p=2y$ ，但需要验证 $p\geq AC(y)$ 。

由 $c(y)=y^2+1$ 可知企业的长期平均成本曲线为 $AC(y)=\frac{c(y)}{y}=\frac{y^2+1}{y}=y+\frac{1}{y}$

下面需要验证 $p=2y\geq AC(y)$，即验证 $2y\geq y+\frac{1}{y}$，由此解得 $y\geq 1$。

因此，企业的长期反供给曲线为 $p=2y$，其中 $y\geq 1$。

最后根据企业的反供给曲线求供给曲线，由 $y\geq 1$， $p=2y$ 可知， $p\geq 2$

因此，企业的长期供给曲线为 $y=\frac{p}{2}$，**其中** $p\geq 2$。

【解法二】

由 $c(y)=y^2+1$ 可知企业的长期边际成本曲线 $MC(y)=2y$

由竞争企业利润最大化的一阶条件可知： $p=MC(y)$

因此，**初步**确定企业的长期反供给曲线为 $p=2y$。

由于若在长期若 $p<\min AC(y)$，企业会退出市场，因此必须有 $p\geq \min AC(y)$。这样我们先把平均成本的最小值 $\min AC(y)$ 求出来。

由 $c(y)=y^2+1$ 可知企业的长期平均成本曲线为 $AC(y)=\frac{c(y)}{y}=\frac{y^2+1}{y}=y+\frac{1}{y}$

长期平均成本曲线的最小值必须满足一阶条件： $\frac{dAC(y)}{dy}=1-\frac{1}{y^2}=0$，由此可得 $y=1$，因此 $\min AC(y)=1+1=2$，所以必须有 $p\geq \min AC(y)=2$

所以，企业的长期供给曲线为 $y=\frac{p}{2}$，其中 $p\geq 2$。

6.以下哪些是技术约束哪些是市场约束：生产要素的价格；市场上其他企业的数量；产量；在投入既定情形下能生产更多的能力。

【复习内容】技术约束；市场约束

技术约束：通常用生产函数表示技术约束。不是每种投入和产出的组合都是可行的。

市场约束：企业可以随意决定生产多少产品，也可以随心所欲地定价，但是产品的销量不是由它说了算，而是由消费者的购买意愿决定的。

【参考答案】

A 生产要素的价格很可能是技术约束，但也有可能是市场约束；原因类似于 C。

B 市场上其他企业的数量为市场约束；

C 产量可能是技术约束，也可能是市场约束。说产量是技术约束是指，你是否有能力生产那么多；说产量是市场约束是指即使你有能力生产那么多，但是由于需求较小，你不会生

产那么多。

D 在投入既定情形下生产更多的能力是技术约束。

7. 哪个假设能描述纯粹竞争市场的主要特征？

【复习内容】纯粹竞争市场（即完全竞争市场）的特征

【参考答案】**纯粹竞争市场最主要的假设是市场中的每个企业都是价格接受者。**

在什么样的市场环境下，纯粹竞争的假设才是合理的？假设**某行业由生产相同产品的很多企业组成，每个企业占市场总量的份额很小**。比如小麦市场。正因为每个企业占市场总量的份额很小。自然可以认为行业中的每个企业都把市场价格看作是预先决定的。即**每个企业都是价格接受者**。

然而，即使市场中只有少数几家企业，每个企业仍然认为市场价格不受它的控制。我们考虑下面的情形，即某种易腐烂的商品供给量既定的情形：例如鲜鱼或插花用的鲜花的市场。即使市场中只有三四个企业，每个企业也只能认为**其他**企业的价格是既定的。如果商品同质，则谁的商品价格最低，消费者就愿意买谁的。因此，如果一家企业把价格定的最低，那么这个价格就是市场价格。如果**其他**的企业想卖产品，这些企业必须接受上述市场价格。

由以上的分析可知，**纯粹竞争市场最主要的假设是市场中的每个企业都是价格接受者。**

8.在纯粹竞争市场中，企业的边际收入总是等于什么？利润最大化的企业会选择什么样的产量？

【复习内容】竞争性企业的供给决策

【参考答案】

纯粹竞争企业的收入 $R(y)=py$，这个式子两边对 y 求导，可知 $MR(y)=p$，因此竞争企业的边际收入总等于市场价格。

企业的利润最大化问题为

$\max\limits_{y}\ \pi(y)=py-C(y)$，其中 $\pi(y),C(y)$ 分别表示企业的利润和成本。

利润最大化的一阶条件为 $\dfrac{d\pi(y)}{dy}=p-\dfrac{dC(y)}{dy}=0$，即 $p=MC(y)$。

这就是说利润最大化的企业选择的产量，应恰好使这个产量处的边际成本等于市场价格。

9.如果平均可变成本大于市场价格，企业生产的产量为多大？在上述情形下，如果没有固定

成本，企业的产量为多大？

【复习内容】竞争企业的短期供给决策

【参考答案】

若平均可变成本大于市场价格，则企业的产量为零。推理如下：

$AVC(y) > p \Rightarrow py - AVC(y)y < 0 \Rightarrow py - TVC(y) < 0$，这个式子是说企业的产量如果为正，那么企业不仅收不回固定成本，连可变成本也不能全部收回，在这种情形下，企业会停止营业，因此产量为零。

即使没有固定成本，上式也意味着可变成本不能全部收回，因此企业仍会停产，所以产量仍为零。

10.某竞争性企业即使亏损也要生产。会不会出现这种情形？为什么？

【复习内容】竞争性企业的短期供给决策

【参考答案】

如果市场价格满足下列条件 $AVC(y) < p < AC(y)$，那么企业仍会生产。

首先由 $p < AC(y)$ 可知该企业是亏损的，因为 $p < AC(y) \Rightarrow py < AC(y)y$，即收入小于成本，因此利润为负，这就是亏损。

但是该企业仍会继续生产，这是因为

$$AVC(y) < p \Rightarrow py - AVC(y)y > 0 \Rightarrow py - TVC(y) > 0$$

这个式子是说，如果企业继续生产，它不仅可以收回全部的可变陈本，而且可以收回部分固定成本，如果它停产的话，一点固定成本也收不回来，所以企业会继续生产。

11.在完全竞争市场中，对于行业中的所有企业来说，市场价格和生产成本之间有什么关系？

【复习内容】竞争企业的利润最大化的一阶条件

【参考答案】

详细过程请见本章第一题的【复习内容】，这里直接给出答案：在完全竞争市场中，对于行业中的所有企业来说，**市场价格应该等于边际成本**。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

23.行业供给（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

23 行业供给

我们已经如何知道根据企业的边际成本曲线推导它的供给曲线。但完全竞争市场中，通常有很多的企业，因此行业供给曲线是该行业所有企业的供给曲线之和。在本章，我们将分析行业供给曲线。

23.1 短期行业供给

我们首先分析行业中企业数量固定的情形，假设某行业的企业数目固定为n。令$S_i(p)$表示企业i的供给曲线，因此，**行业供给曲线**（industry supply curve）或称为**市场供给曲线**（market supply curve）为

$$S(p)=\sum_{i=1}^{n}S_i(p),$$

这个式子是说，行业供给曲线是单个企业供给曲线之和。在图形上，我们可以将每个价格水平上的所有企业的供给量相加，从而得到了一条水平加总的供给曲线，如图 23.1 所示。

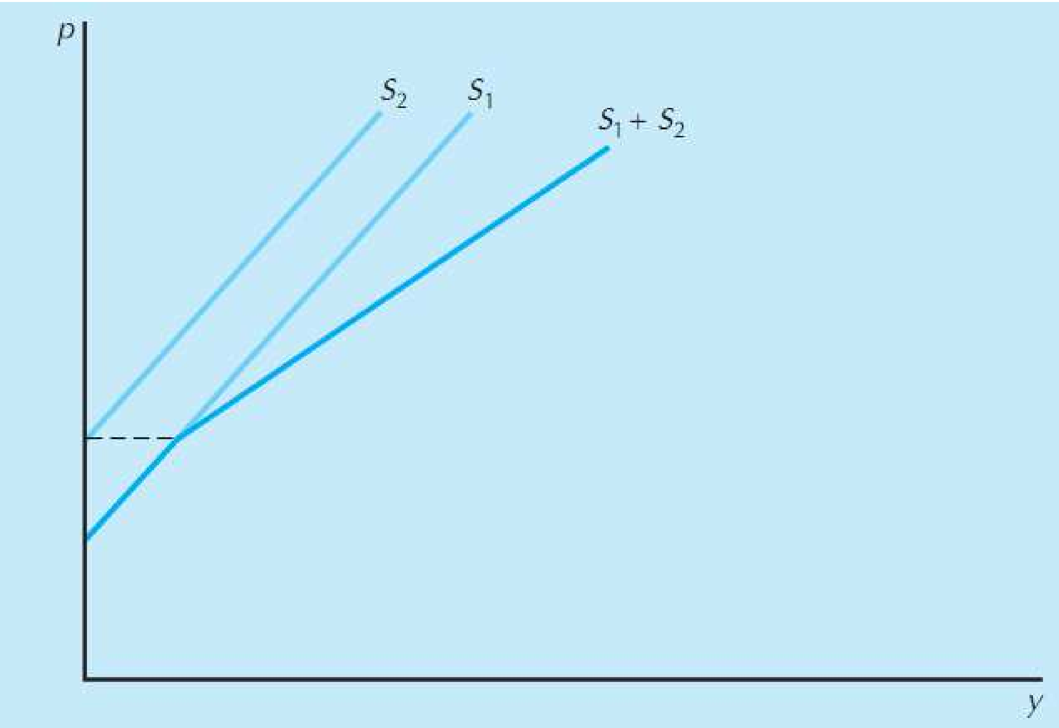


图 23.1：行业供给曲线。行业供给曲线（S_1+S_2）是单个企业供给曲线之和（S_1和S_2）。

曹乾（东南大学 caoqianseu@163.com）

23.2 短期中的行业均衡

为了找到行业在短期中的均衡，我们把上一节得到的市场供给曲线拿过来，找到它与市场需求的交点。这样，我们就得到了一个均衡价格 p^*。

在找到了均衡价格之后，我们再来分析单个企业，考察它们各自的产量和利润。我们构造三个企业（A、B 和 C）的短期均衡，如图 23.2 所示。在这个例子中，企业 A 的价格和产量组合位于平均成本曲线上。这意味着

$$p = \frac{c(y)}{y},$$

交叉相乘并整理可得

$$py - c(y) = 0.$$

因此 A 企业的利润为零。

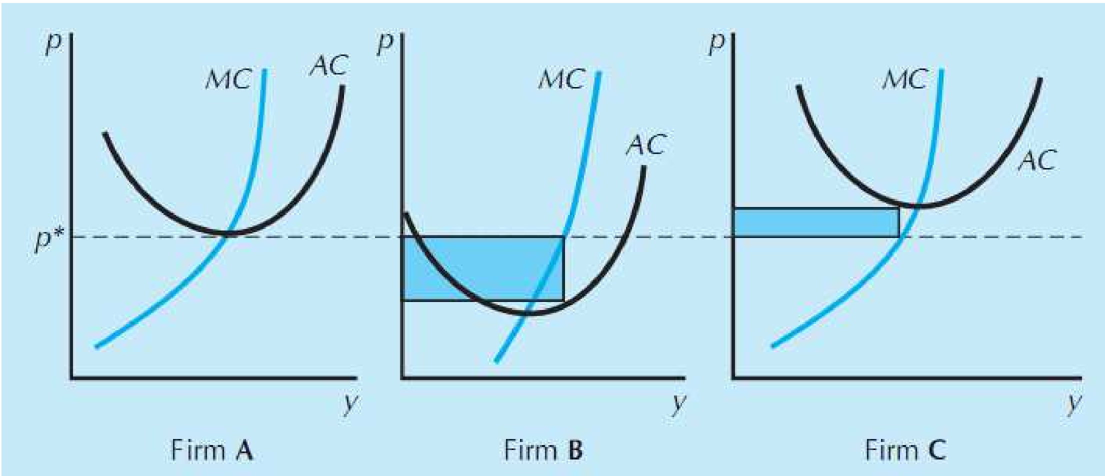


图 23.2：短期均衡。三个企业的短期均衡。均衡时，企业 A 的利润为零，企业 B 的利润为正，企业 C 的利润为负（即亏损）。

企业 B 在价格大于平均成本的点上生产：$p > c(y)/y$，这表示在短期均衡时，它的利润为正。企业 C 在价格小于平均成本的点上生产，因此它的利润为负，即发生了亏损。

一般来说，平均成本曲线上方的价格和产量组合代表着正的利润，平均成本曲线下方的价格和产量组合代表着负利润。在短期内，如果价格和产量组合位于平均可变成本曲线上方，即使企业的利润为负，它的最优选择也可能是继续营业而不是停产。这是因为，在这种情形下继续生产的利润比停产的利润更大（换句话说，生产时亏损更小）。为看清这一点，注意下列事实即可，$py - AVC(y)y - F = [p - AVC(y)]y - F > 0 - F = -F$。

23.3 长期中的行业均衡

在长期，企业可以调整它们的固定要素。它们可以选择工厂规模、资本设备或其他固定要素以使得长期利润最大。这只是说，它们从短期成本曲线移动到长期成本曲线，而这种变化，从分析的角度看，没有增添任何难度：我们使用由长期成本曲线决定的长期供给曲线进行分析，这和短期的情形非常类似。

但是，在长期可能还存在另外一种效应。如果某企业在长期是亏损的，它就没有必要继续呆在行业内，因此我们可以预期这样的企业会**退出**该行业，因为退出后，它可以将亏损降低为零。这种分析思路，从另外一个角度诠释了企业的长期供给曲线为什么只是边际成本曲线的一部分。我们在前面说过，企业的长期供给曲线是**等于**和**高于**平均成本曲线的那部分边际成本曲线——在这部分边际成本曲线上生产才能保证长期利润为非负。

类似地，如果一家企业的利润为正，我们可以预期这种信号会吸引其他企业**进入**该行业。这是因为，成本曲线包含了所有必要的生产要素的成本，而且这些要素的价格要按照市场价格（即它们的机会成本）计算。如果一家企业的长期利润为正，这意味着**任何**企业都可以进入市场，购买这些生产要素，按相同的成本生产相同数量的产品。

大多数竞争行业不存在新企业进入行业的限制；这种情形下，我们说该行业是**自由进出**（free entry）的。然而，某些行业却存在着**进入壁垒**（barriers to entry），例如政府颁发的许可证或者通过法律规定限制行业中企业的数量，就是一种进入壁垒。举个具体的例子，美国很多州对酒类零售行业实行管制，禁止自由进入。

这两种长期效应——获取固定资产以及进出行业的现象——是密切相关的。行业中现有的企业可以决定获取新的工厂或仓库以增加产量；或者一家新企业建立新工厂开始生产，从而进入该行业。这两种效应的唯一区别是谁拥有新的生产设施。

当然，随着企业进入或退出某行业，行业的总产量将会改变，从而使市场价格改变。这反过来又会影响企业的利润、影响企业进入或退出该行业的动机。自由进出行业的最终均衡是什么样子的？

假设某行业的长期成本函数 $c(y)$ 都是相同的。这样我们就可以计算出平均成本最小时对应的产量，用 y^* 表示这一产量。令 $p^* = c(y^*)/y^*$ 表示平均成本的最小值。这个最小值非常重要，因为它是市场能索要最低的价格而且还能做到让企业盈亏平衡。

这样，我们就可以画出企业可以自由进入行业情形下的行业供给曲线。如图 23.3 所示。该图显示了市场中企业的数量分别为 1~4 时的行业供给曲线。（我们只用 4 家企业的情形举例。现实中的竞争行业，企业的数量当然很多。）注意，既然所有企业都有相同的供给曲线，若市场中有 2 家企业时，市场供给量正好等于市场只有 1 家企业时供给量的 2 倍；若市场上有 3 家企业，市场供给量是只有 1 家企业时的 3 倍，以此类推。

现在，在图 23.3 中再增加量条线：一条是高度为 p^* 的水平线，这条线表示企业取得非负利润时要求的最低价格；另一条是市场需求曲线。分别考虑当企业数量 n =1~4 时，市场

供给曲线和市场需求曲线的交点。

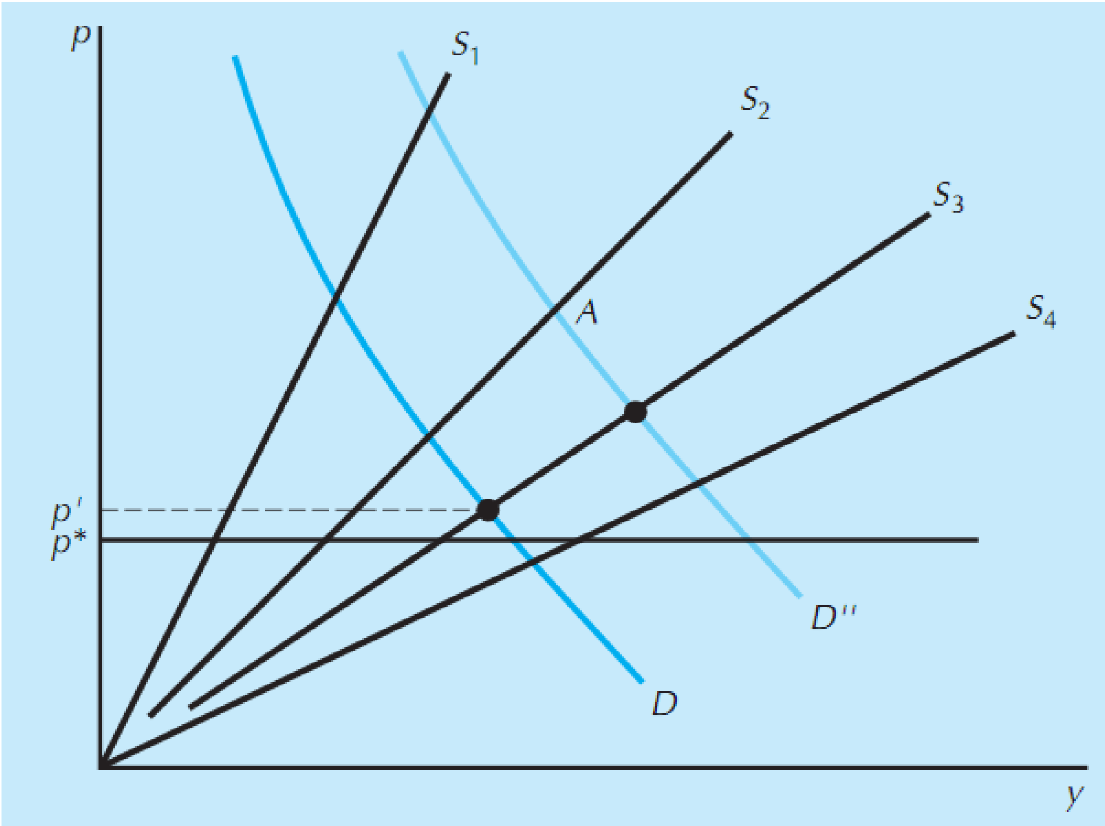


图 22.3：行业可以自由进出情形下的行业供给曲线。上图给出了企业数量为 1~4 家时的行业供给曲线。均衡价格 p' 为行业需求曲线与行业供给曲线最低的交点上，该最低交点仍要满足 $p' \geq p^*$。

如果利润为正时，企业进入该行业，那么我们最终关注的交点应放在**能使企业赚取非负利润的最低价格上**。在图 23.3 中，我们用 p' 表示这个价格水平，这正好对应着市场只有 3 家企业的情形。如果再有一家企业进入市场，则利润被压低为负利润。这种情形下，该竞争行业中能容纳的企业的最大数量为三家。

23.4 长期供给曲线

上一节我们使用的分析工具是，首先构造不同企业数量情形下的行业供给曲线，然后寻找能保证企业赚取非负利润的最大企业数目。这种方法非常严格也容易实施。但是我们还有一种简便的近似求解方法，这种方法得到的结果与正确答案非常接近。

这种方法就是使用类似图 23.3 中的 n 条行业供给曲线，直接构建出一条行业供给曲线。怎么做此事？首先注意到，我们可以排除供给曲线上所有低于 p^* 的点，因为企业在长期绝

对不会在这样的点上生产。但是，我们可以排除供给曲线上**某些高于** p^* 的点。为什么？分析如下。

我们一般假设市场需求曲线是向下倾斜的。最陡峭的需求曲线因此为直线。这意味着图 23.3 中类似 A 点的那些点，不可能出现在行业供给曲线上，因为任何穿过 A 点的需求曲线必然与多条行业供给曲线相交，如图 23.3 所示。由于我们关注尽可能低的均衡价格 p'（只要 p' 不小于企业平均成本最低点即可），因此我们最终关注的是穿过 A 点的需求曲线与企业数目更多时的供给曲线相交的情形。这样类似 A 点的那些点都可以被排除。

因此，我们可从每条供给曲线上删除一部分点，因为这些点不可能是长期均衡点。在前面我们说过，任何供给曲线上所有低于 p^* 的点都可以排除。现在我们把注意力放在排除供给曲线上某些高于 p^* 的点。

比如第一条供给曲线 S_1，即市场只有一家企业时的供给曲线，这条线上高于 p^* 的那些点中哪些是可以排除的？注意此时最关键的一条线是价格线 p^*。我们先找到这条价格线与第二条供给曲线 S_2（市场上有两家企业时）的交点，然后我们过这个交点作一条垂线，最后我们断言第一条供给曲线位于这条垂线右侧的点不可能是长期均衡点，也就是说这些点是可以排除的。类似地，第二条供给曲线 S_2 上，位于第三条供给曲线 S_3 与价格线 p^* 交点右侧的那些点都可以排除，依次类推。第 n 条供给曲线 S_n 上，位于第 $n+1$ 条供给曲线 S_{n+1} 与价格线 p^* 交点右侧的那些点，都不可能是长期均衡点，因此都可以排除。

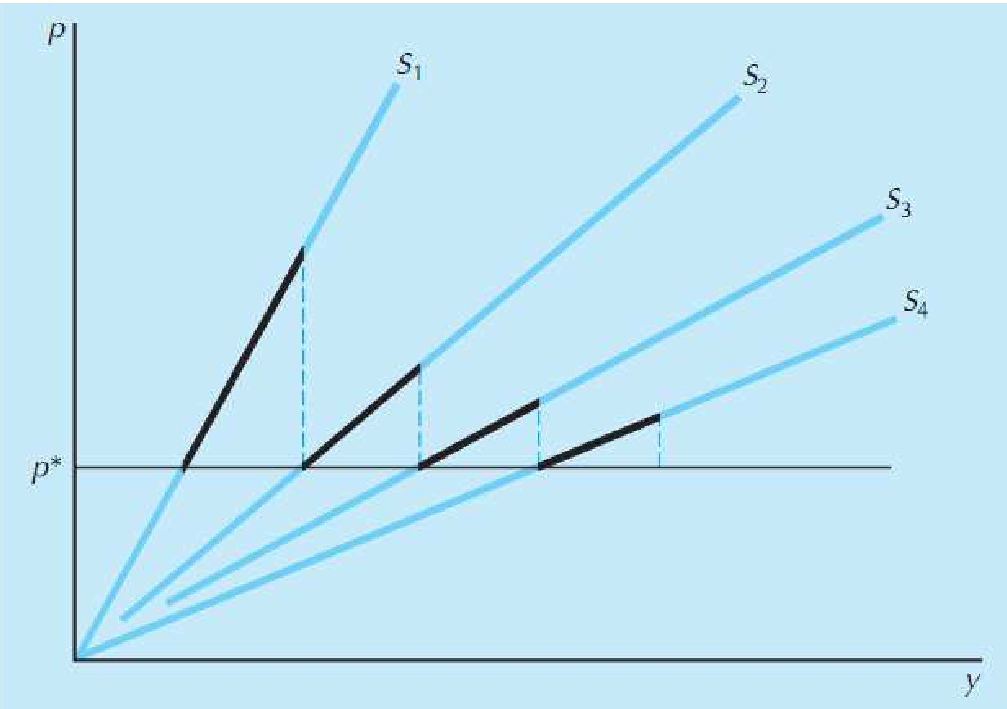


图 23.4：长期供给曲线。长期情形下，将每条供给曲线上不会与向下倾斜需求曲线相交的那些点删除，在图形中，就是将每条供给曲线位于虚垂线右侧的那些点删除，就得到了长期供给曲线。

如此删减以后剩下的那部分供给曲线，就是长期均衡可能出现的位置，我们将这部分供给曲线用黑色实线段表示，如图 23.4 所示。第 n 条黑色线段代表着市场有 n 家企业时，能实现长期均衡的价格和产量的所有组合。注意，随着市场中企业数量的增多，产量也越多，这些黑色实线段会越来越平缓。

为什么这些曲线会越来越平缓?思考一下吧。若市场中只有一家企业，并且市场价格上升了 Δp，它会多生产比如 Δy 单位的产品。若市场中有 n 家企业而且价格上升了 Δp，每家企业都会多生产 Δy 单位的产品，因此市场上的供应量增加了 $n\Delta y$。这表明，供给曲线会随着市场中企业数量的增多而变得越来越平缓，因为供给量对价格变动越来越敏感。

当我们终于找到市场能容纳的最大企业数量时，供给曲线必定会非常平缓。因为它太平缓了，所以我们可以认为它的斜率等于零，也就是说，将行业的长期供给曲线看成一条水平线，这条水平线的高度为平均成本曲线的最低点（即 $p^* = \min AC$）。如果在长期，市场中只有少数几家厂商，那么这种近似估计方法就不是一个好方法。但是你为什么要假设竞争行业只有少数几家企业呢？这样的假设一般也不合理。如果长期内行业中的企业数量很多，则长期均衡价格就不可能与平均成本最小值相差很大。如图 23.5 所示。

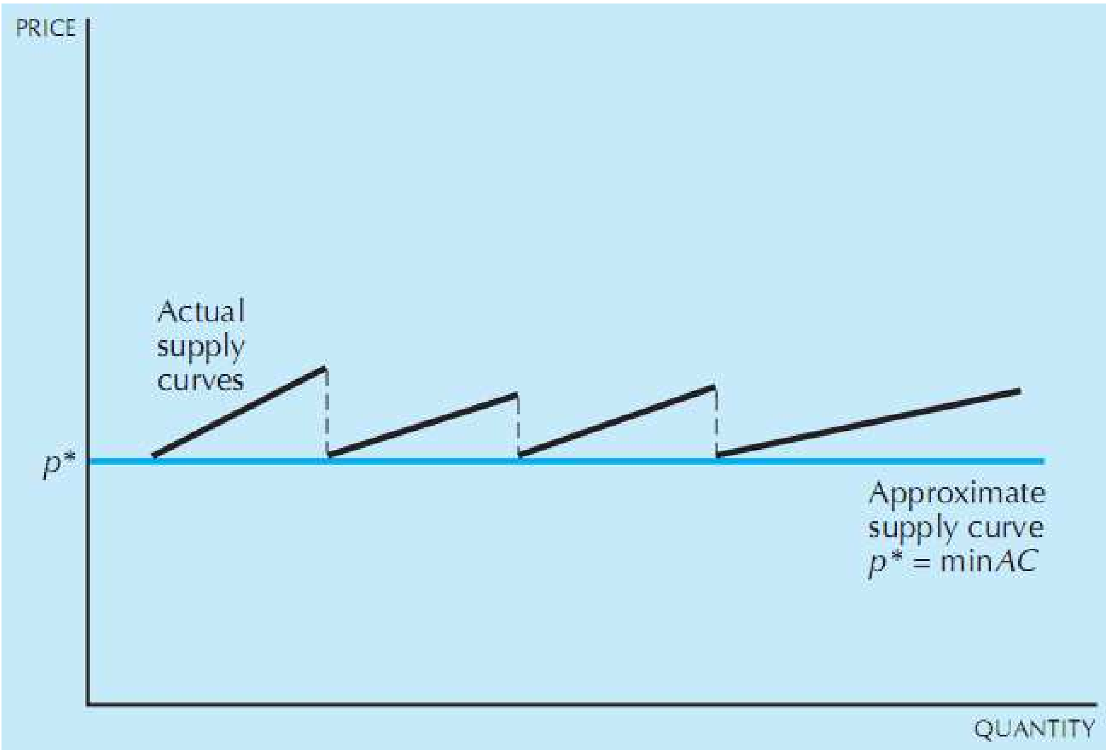


图 23.5：近似的长期供给曲线。长期供给曲线近似为一条水平线，它的高度为平均成本最小值，即 $p^* = \min AC$。

上述结论对于可以自由进出的竞争行业来说，具有重要的应用价值。根据这个结论我们可以断言，竞争行业中企业的长期利润近似为零。因为如果利润较高，就会吸引其他企业进入该行业，从而将利润压低以至趋近于零。

我们已经知道，利润等于收入减去成本。注意务必将所有成本都包含进来，也就是说计算经济成本时必须将所有生产要素按照市场价格计价。只要成本中包含了**所有**要素，而且均已合理计价，这种情形下，只要企业的利润为正，其他企业就会仿效。它们也会采用上述企业的生产技术，购买要素进行生产。这样每个企业的利润都会被压低。

如果某行业可以自由进出，则行业的长期平均成本曲线基本是一条水平线，这条水平线的高度等于平均成本的最小值，即 $p^* = \min AC$ 。这条曲线类似一家规模报酬不变企业的长期供给曲线。这不是偶然的。我们以前断言，规模报酬不变是个合理的假设，因为你的企业总可以复制它以前的行为。但是，别忘了，其他企业也可以复制你的企业！你将你的企业复制一次来扩大产量，这和别人复制你的企业后进入你所在的行业没有什么区别。因此，可以自由进出的竞争行业的长期供给曲线，和一家规模报酬不变企业的长期供给曲线的形状是一样的：都是一条水平线，而且水平线的高度都等于平均成本的最小值，即 $p^* = \min AC$ 。

例子：长期和短期情形下的税收

考虑一个可以自由进出的行业。假设最初该行业处于长期均衡状态，行业中的企业数量因此是固定的，并且每个企业的利润为零。如图 23.6 所示。在短期，由于企业的数量是固定不变的，行业供给曲线因此是向上倾斜的；但在长期，由于企业数量可变（因为企业可以进入或退出），行业的供给曲线是一条水平线，且其高度等于平均成本最小值（$p^* = \min AC$）。

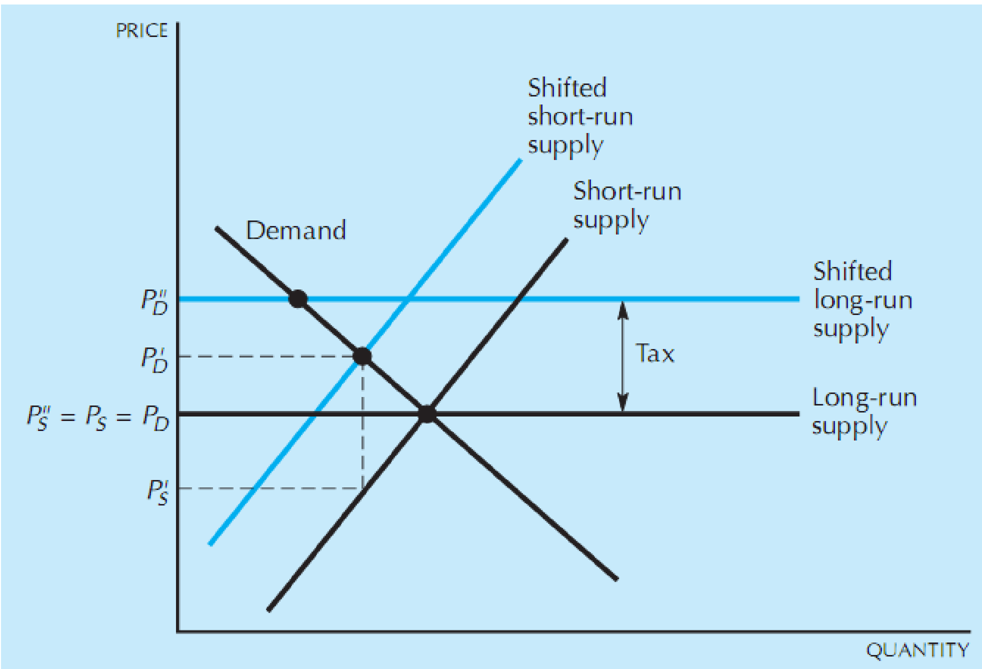


图 23.6：长期和短期情形下的税收。在短期，由于企业数量固定不变，行业供给曲线必然向上倾斜，因此部分税收由消费者承担，部分税收由生产者承担；在长期，行业供给曲线是水平的，因此税收全部由消费者承担.

现在，如果政府开始对该行业征税，这对该行业有何影响？我们用 16 章介绍过的图形分析方法：为了找到需求者支付的新价格，我们向上移动供给曲线，移动距离等于单位税额。

一般来说，征税后与征税前的价格相比，消费者会面对较高的价格，生产者将得到较低的价格。但是由于在征税前，生产者正好处于不亏不盈的状态；征税后生产者必将亏损。这些经济损失会迫使某些企业退出该行业。因此，行业的产品供给量将下降，从而消费者支付的价格将进一步上升。

在长期，行业将沿着水平的长期供给曲线供给产品。为了能沿着该曲线供给产品，企业得到的价格最低也不能低于平均成本的最小值，这和征税前它们能忍受的最低价格是一样的。因此，消费者面对的价格必然上升，而且上升幅度等于全部的单位税额。

在图 23.6 中，最初均衡时有 $P_D = P_S$。现在政府开始对生产者征税，这将使短期供给曲线向上移动，移动距离等于单位税额；在新的均衡状态下，需求者支付的价格上升为 P_D'，生产者得到的价格下降为 $P_S' = P_D' - t$。注意，这只是短期的情形，即行业中的企业数量固定不变。在长期，由于该行业可以自由进出，行业的**长期**供给曲线是一条水平线，其高度为 $P_D = P_S = \min AC$。因此，在长期，行业供给曲线向上移动意味着生产者将全部的税额转嫁给消费者。

稍微总结一下我们得到的结论：在一个可以自由进出的行业中，征税会使消费者支付的价格上升，上升额度小于单位税额，这是因为生产者也承担了部分单位税额。但在长期，征税会迫使某些企业退出该行业，因此供给减少，从而消费者承担了全部的税收。

23.5 零利润的含义

在一个自由进出的行业中，新进入的众多企业会将利润压低为零：只要利润为正，新的企业就有进入的激励，因为这样它可以获取某些利润。当利润为零时，这并不意味着该行业消失了；该行业还存在但它停止增长，因为新企业不再进入。

长期均衡时若利润为零，这意味着所有生产要素都已按市场价格得到补偿。例如，企业的所有者对该企业投入的劳动，或者他投入的资金，或者他对该企业的任何投入，都按市场价格得到了补偿。当然，其他生产要素也按市场价格得到了补偿。该企业仍在赚钱，只不过它赚的钱全部用于购买各种生产要素。每种生产要素在该行业得到的报酬和在其他行业一样，因此，此时不存在额外的报酬，即不存在纯利润，因此，新的生产要素也不会再流入该行业。但是，这些要素也没有理由离开该行业。行业长期均衡时利润为零，表明这样的行业是成熟的行业；它们不太可能被《商业周刊》选为封面故事，但它们是经济的支柱。

记住，经济利润等于收入减去经济成本，在经济成本中，所有要素必须按市场价格计价。市场价格衡量这些要素的机会成本——这些要素在其他地方能获得的报酬。如果收入超过了生产要素的成本，则意味着企业获得了纯经济利润（pure economic profit）。但是一旦有人能赚取纯经济利润，其他人就会试图进入该行业并且获取某些利润。在一个可以自由进出的竞

争行业中，正是这种追求利润的激励最终迫使利润为零。

在某些地方，人们会鄙视企业的逐利动机。但是如果你抛弃了感情色彩仅从经济学的观点来看，利润的确为人们提供了资源配置的正确信号。若某企业能赚取正的利润，这意味着人们对该企业产出的评价高于对其投入的评价。因此，从经济学的角度来说，这样的企业越多越好。

23.6 固定要素和经济租金

如果行业可以自由进出，则在长期，企业的利润会迫使降低为零。但是不是每个行业都是自由进出的。在某些行业，行业里的企业数量是固定不变的。

最常见的原因是某些要素的供给量是固定的。我们说过，在长期，企业可以买进或卖出固定要素。但有些要素对于**经济整体**（economy as a whole）是固定的，即使是在长期。

最明显的例子是资源提炼行业：石油提炼行业的一种必要投入就是石油，但是地下的石油就那么多。类似的例子是煤、天然气、贵金属或其他这样的资源。农业是另外一个好例子。因为适合耕种的土地面积是既定的。

一种更奇特的固定要素例子是天赋（talent）。只有部分人具有成为职业运动员或演员的天赋。这些行业也许是“自由进出”的，但是只有天赋很高的人才能进入！

某些情形下，固定要素的数量固定是由法律造成的，而不是自然界造成的。很多行业都有许可证的限制，法律可能规定了许可证的数量。很多城市的出租车行业就是使用许可证管制的。类似地还有卖酒的许可证等。

如果行业中的企业数量受到上述因素的限制，因此新企业不能自由进入，在这种情形下，该行业似乎有可能获得正的利润，因为**似乎**不存在迫使利润为零的经济力量。

上述“似乎”是错误的。因为确实存在一种经济力量可将利润压低为零。如果你看到某企业的长期利润为正，这很可能因为你没有将全部成本包含进来，准确地说，你没有准确估计妨碍企业进入市场的因素的市场价值。

记住成本的经济学定义：我们必须将每种生产要素的价格按照它的**市场价格**，即它的机会成本计算。如果我们将某个农民的种地收入减去种地成本后利润为正，我们很可能是因为忘记了扣除他的自有土地的成本。

假设在上个例子中，除了土地成本之外，我们考虑了其他的全部成本，最终计算出年利润为 π 元。那么，该农民的土地在市场上的价值如何？如果某人租赁该农民的土地一年，他应该支付多少租金？

答案是：他愿意支付的年租金为 π 元，这正是土地产生的“利润”。如果你租入该农民的土地，即使你不会种地，你也能“赚取” π 元——毕竟，我们对这个农民的劳动也是按市场价格计算的，这意味着你可以雇佣这样的农民让他替你耕种，你仍然能获得 π 元的利润。

曹乾（东南大学 caoqianseu@163.com）

因此，这块土地的市场价值——它的竞争性租金——正好是 π 元。种地的经济利润为零。

注意，这种确定租金率的程序和这块土地的历史成本没有任何关系。真正重要的不是你的买入价而是卖出价，这正好是机会成本的概念。

只要存在阻止企业进入某行业的固定要素，那么这种固定要素就必然存在着均衡租金率。即使存在着这样的固定要素，你也有办法进入市场，因为你可以买下行业中现有的某个企业。行业中的每个企业都可以将自己出售。注意，在这种情况下，如果企业选择不将自己卖掉，那么若将自己卖掉获得的收入就是企业继续经营的一种生产成本，这种机会成本在计算经济利润时必须加以考虑。

因此，在某种意义上，正是进入行业的这种**可能性**（possibility of entry）迫使利润为零。毕竟，进入某行业有两种途径：你可以成立一个新企业，或者你买下行业中现有的一家企业。如果新成立的企业能购买各种要素进行生产，而且还能获利，那么你会这么做。但是如果某些要素的供给量是固定的，那么潜在进入者对这些要素的竞争，会抬升这些要素的价格，直到利润为零。

例子：纽约市的出租车执照

1980 年代，纽约市出租车的营业执照的价格大约为 10 万美元。但在 1986 年出租车司机每周（50 小时）收入为 400 美元，相当于每小时不到 8 美元。纽约出租车协会认为这样的工资太低了，不可能吸引熟练司机加入该行业。因此他们认为应该提高出租车的乘车费用。

经济学家则认为，提高乘车费用不可能提高司机的实际收入；只会导致出租车营业执照价格上涨。我们使用该协会关于一辆出租车运营成本的数据，分析一下经济学家的观点对不对。在 1986 年，司机租赁一辆出租车的费用为：白班 55 美元；夜班 65 美元。司机在扣除租赁费、汽油费等成本后，一天大约可以净得 80 美元。

我们来看看拥有出租车运营执照的人能赚多少钱吧。假设一辆出租车一年可以出租 320 天，包括白班和夜班，这样出租执照的人一年的收入为 38,400 美元。一辆车一年的成本包括保险、折旧和维护费等等，这些成本加起来大约 21,100 美元。这样，一年的净利润大约为 17,300 美元。由于营业执照的成本为 10 万美元，这表明报酬率超过了 17%。

出租车的乘车费用上涨直接导致营业执照价格的上升。乘车费用的增加使得每年额外多收入了 10,000 美元，这导致营业执照的价格上升了 60,000 美元。出租车司机的工资率——由劳动市场决定——不会受到这种变化的影响[一]。

[一] 数字来源于 1986 年 8 月 17 日《纽约时报》的一篇文章，该文章未署名。

23.7 经济租金

上一节我们介绍的例子就是**经济租金**（economic rent）的例子。某生产要素的经济租金的定义是，企业购买该要素的实际支出金额减去要素供给者能接受的最低金额。

例如，考虑前面我们说过的开采石油的例子。为了能开采，你必须有一些劳动力、一些机械，更重要的是地下要有石油。假设从地下打出一桶油的成本为 1 美元。如果石油的价格大于 1 美元每桶，那么企业就会供应石油。但是现实中石油的价格远远高于 1 美元每桶。为什么？因为人们想要石油的原因有很多，所以他们愿意支付远高于生产成本的价格来得到石油。石油价格超出生产成本的那部分就是经济租金。

为什么企业不进入该行业？它们想进入，但是石油的储藏量是有限的。石油价格远高于成本的原因就是因为供给量有限。

现在考虑出租车执照。由于执照的本质就是纸张，因此成本几乎为零。但在纽约市一张出租车的营业执照可以卖到 10 万美元！为什么人们不能进入执照制造行业从而制造出更多的执照？答案是进入该行业是非法的——出租车营业执照的供给是由市政府控制的。

耕地是经济租金的另外一个例子。总体上说，耕地的数量是固定的。每英亩土地的价格无论是 0 元还是 1000 元，土地的供给量是不变的。因此总体上来说，支付给土地的资金就构成了经济租金。

从经济整体的角度看，耕地的价值由农产品的价格决定。但从单个农民的角度来看，他的土地价值是一种生产成本，是产品价格的组成部分。

我们用图 23.7 阐述经济租金。图中 AVC 表示**除了**土地之外的所有其他要素的平均成本曲线。（假设土地是唯一的固定要素。）如果农产品的价格为 p^*，那么土地贡献的“利润”是图中矩形的面积：这就是经济租金。经济租金是土地在完全竞争市场出租时获得的租金，它迫使利润为零。

包含土地价值在内的平均成本曲线，我们在图中用 AC 标记。如果我们正确衡量了土地的价值，耕种土地的经济利润应该恰好等于零。因为这块土地的均衡租金要保证支付地租后利润为零，因此必有

$$p^* y^* - c_v(y^*) - \text{租金} = 0$$

或

$$\text{租金} = p^* y^* - c_v(y^*) \tag{23.1}$$

这正好是我们介绍过的生产者剩余。的确，经济租金和生产者剩余是同一个概念，只是站在不同的角度看问题。因此，我们可用边际成本曲线左侧的面积来计算租金。

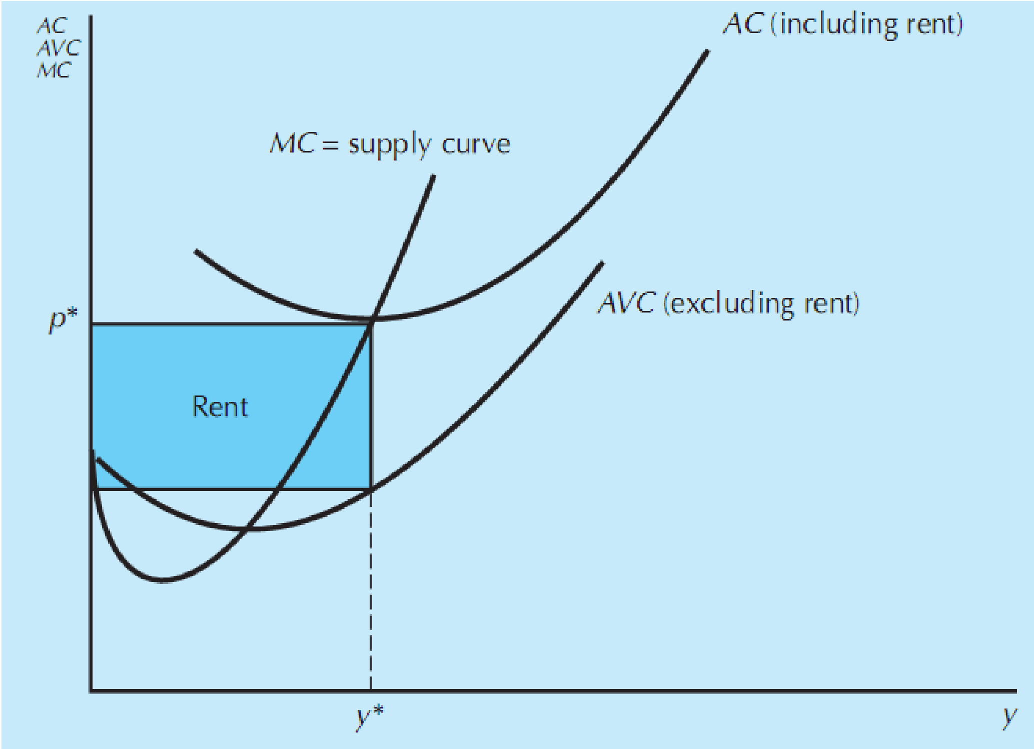


图 23.7：土地的经济租金。矩形面积代表土地的经济租金。

根据（23.1）式给出的经济租金的定义，现在容易看出我们前面说过的事实：是均衡价格决定了经济租金，而不是相反。企业沿着它的边际成本曲线供给产品——这和固定要素的支出无关。租金将调整到直至利润等于零。

23.8 租金率和要素的价格

由于我们用流量单位计算产量——单位时间内产量为多少，我们必须也用流量单位计算利润和租金——单位时间内多少钱。也正是这个原因，在上面的讨论中，我们使用的是年租金这样的称呼，这就是**租金率**（rental rates）。

如果土地或营业执照是出售而不是出租，那么它们的均衡价格等于租金流的现值。因为我们在前面章节说过，在竞争市场上，能产生收入流的资产应按它的净现值出售。

例子：卖酒许可证

在美国，每个州都有自己的酒类销售政策。有些州只存在一家垄断销售商；有些州则向想卖酒的人发放许可证。有些情形下，许可证是出售的；有些情形下，许可证的数量是固定不变的。例如，密歇根州规定，每 1500 个居民只发送一张卖酒许可证，而且这种许可证规定酒只能卖给在酒店里喝而不能将酒带出去的消费者。

每次联邦人口普查之后，密歇根州酒类管制委员会就会向人口增长的社区额外发放许可证。（然而，如果某社区的人口数下降，卖酒许可证并不收回。）这种情形实际上等于人为制造了许可证的稀缺性，从而使得人口增长快速的社区争相竞买许可证。例如，1983 年密歇根州安娜堡市已有 66 张卖酒许可证。1980 年人口普查之后，该城市将额外获得 6 张许可证。33 个申请人争相游说酒类管制委员会，希望得到许可证。当时，许可证的市场价格大约为 8 万美元。当地报纸曾报道“许可证的需求超过了供给。”价值为零的许可证（不过一张纸）卖到了 8 万元仍然供不应求，你不要奇怪。因为许可证被人为制造出了稀缺性。

人们三番五次地提议密歇根州应该放松对酒类的管制，应该发放更多的卖酒许可证。然而这些提议从未被采纳，因为很多政治团体强烈反对放松对酒类的管制。有些团体认为若放松管制就会影响到公共健康或宗教的教义。另外一些团体反对动机却耐人寻味。比如，反对最强烈的当属密歇根卖酒协会，这个协会代表该州已取得许可证的卖酒人的利益。你可能会认为放松酒类管制会对这些卖酒人有好处，因此他们的反对有些不可思议。但是，稍微思考一下你就会明白他们反对放松管制的真正原因：发放更多的许可证无疑会降低现有许可证的转售价值，从而使当前拥有许可证的人蒙受巨额资本损失。

23.9 租金政治

经济租金存在的最常见原因是对进入某行业的法律限制。我们在前文介绍了两个例子：出租车许可证和卖酒许可证。在这两种情形下，许可证的数量都是法律规定好了的，因此，法律限制了企业进入这些行业，从而导致了经济租金的产生。

假定纽约市政府希望增加市场内出租车的数量。这会对市场中已有出租车许可证的价值造成什么影响？显然，它们的价值会下降。许可证价值下降直接命中了该行业的经济要害，可以预期该行业必然会游说政府不要实施这样的方案。

联邦政府也会采取类似地措施，人为地限制某些产品的产量，从而产生了经济租金。例如联邦政府规定了哪些土地可以种植烟草。这些土地的价值就取决于烟草产品的需求。如果人们想取消这样的许可证制度，就必须与强大的游说力量做斗争。一旦政府创造了人为的稀缺，想要取消这种稀缺就很难。这种稀缺的既得利益者——已获得行业经营许可的人——会激烈地反对扩大行业的规模。

在法律禁入的行业中，已获得行业经营许可的人会竭力维护自己的有利地位。游说费、律师费、公共关系费等等都是很大的费用支出。从社会的观点看，这些费用代表的是纯粹的社会浪费。这些费用不是真正的生产成本；它们不会导致产量增加。游说和搞好公共关系只是决定了谁能得到这些既定利益。

例子：耕种的是土地还是政府？

美国对农业补贴的唯一好处是，它为经济学教科书提供了无穷的素材。每次农业项目改革都会产生新的问题。美国国家农业合作委员会的副会长特里.巴（Terry Bar）曾经说过，"如果你想找到农业补贴方案的漏洞，就把这个方案扔给农场主，没有人比他们更懂得钻政策的空子。"[一]

1996 年之前，美国农业补贴的基本做法是价格补贴：联邦政府对每种农产品都承诺一个保护价格，如果市场价格低于保护价格，联邦政府会补偿这部分差价。为了有资格享受这种补贴，农场主必须同意少耕种部分他自己的土地。

由这种方案的做法可知，大多数利益都被大农场主获得。一项研究表明，13%的联邦直接补流入了 1%的年销售额 50 万美元以上的大农场主。1985 年联邦食品安全法实施，该法案明显限制了对大农场主的补贴。大农场主的应对策略是将土地分块租给当地的投资者。这些投资者根据补贴方案量身定做了租种土地的策略：租种的土地面积要尽可能利用政府的补贴政策，但又达不到政策限制不能享受补贴的土地面积水平。投资者一旦获得了这些土地，它就向政府申请补贴，注意，此时联邦政府的补贴政策是，你有这样的土地但不能耕种才能获得补贴。政府的这种农业方案后来被人们称为"在政府身上耕种"（farming the government）。

研究表明，1985 年对大农场主补贴进行限制的法案，直接导致了额外 31,000 份补贴申请。政府为这些补贴支付的费用大约为 23 亿美元。

注意，1985 年这项法案的目标——限制对大农场主的补贴——并没有实现。当大农场主将土地租给小农场主，土地租金的价格取决于联邦政府补贴的程度。补贴越高，大农场主得到的均衡租金越高。这种补贴方案的大部分利益最终仍然为大农场主获得，因为土地租金的大小最终取决于土地能赚取的利润，不论你耕种的是土地还是耕种政府。

1996 年的农业法案承诺在 2002 年之前分阶段取消大部分农业补贴。然而，1998 年联邦预算又对农业补贴拨付了 60 多亿美元。这再次表明，要对政治和经济进行调和是多么艰难。

23.10 能源政策

我们用前面介绍的概念再分析一个例子，并以此作为本章的结束。

1974 年，石油输出国组织（OPEC）大幅提高了石油价格。国内不生产石油的国家对石油政策别无选择——石油的价格和石油制成品的价格只能上升。

当时美国国内石油消费量一半靠自产一半靠进口。议会认为，国内石油生产商不应该从高涨的石油价格中获取"暴利"（windfall profits）。暴利这个术语是指，利润增加的原因是由于外部事件引起的，而不是由于生产决策引起的。因此，议会设计了一种奇怪的方案，

[一] Quoted in William Robbins, "Limits on Subsidies to Big Farms Go Awry, Sending Costs Climbing," New York Times, June 15, 1987, Al.

试图降低石油制成品的价格。这些产品中最常见的就是汽油，因此我们分析这种方案对汽油市场的影响。

石油定价双轨制

议会采取的价格政策称为石油定价的双轨制（two-tiered）。定价方法如下。进口的石油按市场价格销售，但是国内生产的石油，即从 1974 年前已经存在的油井中开采的石油，只能按照原来的价格销售，即按照 OPEC 涨价之前的价格销售。大致来说，进口石油的售价约为 15 美元每桶，但国内生产的石油的售价约为 5 美元每桶。这种政策的思想是，如此定价可以使得石油的售价平均为 10 美元每桶，这显然有助于降低汽油的价格。

这种方案能奏效吗？我们从汽油生产者的角度进行分析。汽油供给曲线是什么样子的？为了回答这个问题，我们必须知道汽油边际成本是什么样子的。

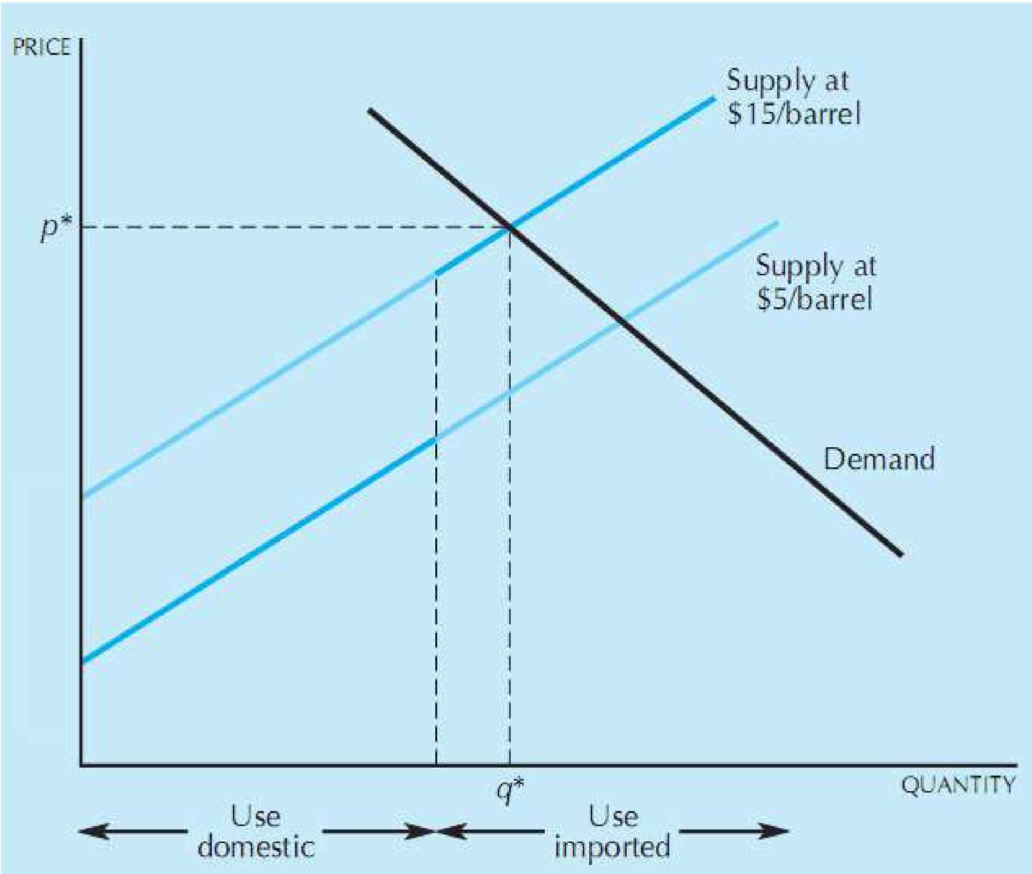


图 23.8：汽油的供给曲线。在石油定价双轨制的情形下，汽油的供给曲线将不再是连续的曲线，在国内石油耗尽的那个点上，供给曲线出现了跳跃，即从较低的供给曲线跳跃到较高的供给曲线。

如果你是个汽油提炼商，你会怎么做？显然你会尽可能地首先使用便宜的国内石油。只有国内石油使用完以后，你才会购买更贵的进口石油。因此，行业总边际成本曲线，即行业供给曲线的形状必然类似图 23.8 所示的形状。在国内石油用尽和开始使用进口汽油的那个分界点上，供给曲线出现了跳跃。在这个分界点之前，生产汽油使用的要素价格是国内石油的价格，在该点之后，要素的价格变为进口石油的价格。

图 23.8 画出了两条汽油供给曲线，第一条表示石油价格为 5 美元每桶时的汽油供给曲线；第二条表示石油的价格为 15 美元每桶时的汽油供给曲线。如果国内汽油的价格真为 5 美元每桶，进口的石油价格真为 15 元每桶，则汽油的供给曲线一开始将和第一条供给曲线重合，直到便宜的国内石油用尽；接下来，汽油供给曲线则和第二条供给曲线重合。

现在我们可以找一找均衡价格了，在图 23.8 中，这个价格就是汽油供给曲线和汽油需求曲线交点决定的价格。由图可知一个有趣的事实：在石油价格双轨制的情形下，汽油的均衡价格，恰好等于如果所有的石油均按进口石油的价格计价时的汽油均衡价格！汽油的价格由生产汽油的**边际成本**决定，而**边际成本**又由进口石油的价格决定。

如果你想一会，你就会明白上述结论是合理的。汽油公司按照市场可以承受的价格销售汽油。即使你有幸得到了便宜的国内石油，你也会按照其他汽油公司的售价出售汽油。

暂时假设所有汽油按统一价销售，这样均衡价格将为 p^*。现在政府介入规定每个汽油商使用的前 100 桶石油的价格下降。政府这种做法会影响汽油的供给决策吗？不可能！为了影响供给决策，政府必须改变汽油商在边际上的激励。让汽油价格降低的唯一可行方法是增加供给，这就是说政府必须降低汽油生产的边际成本。

石油定价双轨制的本质只是将利益从石油生产者手里转移到国内汽油商的手里。政府的这种做法使得国内石油商的收入减少了 10 美元每桶，因为按市场价格他们可以销售 15 美元每桶但政府限定只能售卖 5 元每桶。因此，他们的收入被政府转移到汽油商的手里。政府的定价政策对汽油的供给没有任何影响，因此它也不会影响到汽油的价格。

价格管制

市场力量很快让人们意识到它的存在。美国能源部意识到不能听任市场力量决定价格双轨制下石油的价格，因为市场力量会导致汽油的价格只有一种，即和不实行价格双轨制情形下的汽油价格是一样的。

因此，能源部对汽油的价格进行管制，要求每个汽油商按照汽油的生产成本定价，而汽油的生产成本又主要取决于汽油商购买石油的价格。

能否买到便宜的国内石油主要取决于汽油商所在的地区。由于德克萨斯州盛产石油，因此该地的汽油商能买到大量的便宜石油。由于价格管制，德克萨斯州汽油的价格相对很便宜。而在新英格兰地区，由于几乎所有的石油都依赖进口，该地区的汽油价格相当高。

当同一种商品有不同的价格时，企业自然会按最高的价格销售。能源部再次介入，阻止汽油从低价地区流向高价地区。这种政策造成了 1970 年代汽油短缺现象。某地区的汽油供给量总会枯竭，由于人为干预，即使你出价再高也买不到汽油。如果政府不对市场进行干预，绝不会出现这样的现象。汽油短缺完全是由政府同时实施价格双轨制和价格管制而引起的。

经济学家当时指出了政府的政策必然会导致汽油短缺的现象，但是没人理会。倒是汽油商对政府的大力游说起了作用。由于国内大部分石油是按长期合同销售的，因此某些汽油商能买到大量的便宜石油，但是另外一些汽油商只能购买昂贵的进口石油。自然地，这些汽油商会批评这种不公平的现象，他们游说政府的结果是，议会同意制定另外一种方案，旨在公平地分配便宜的国内石油。

授权方案

这种方案称为“授权方案”，其实施办法如下。汽油商每购买一桶昂贵的进口石油，他就可以得到一张优惠券，凭借优惠券就可以购买一定数量的便宜的国内石油。汽油商能购买到的国内石油数量取决于供给条件，但我们假设是一配一的：每购买一桶价格 15 美元的进口石油，他就有权购买一桶价格 5 美元的国内石油。

这种政策对石油的边际价格有何影响？现在石油的边际价格正好是国内石油价格和国外石油价格的加权平均数；如果这两种石油的数量是一配一的，那么石油的价格将为 10 美元每桶。它对汽油供给曲线的影响可用图 23.9 表示。

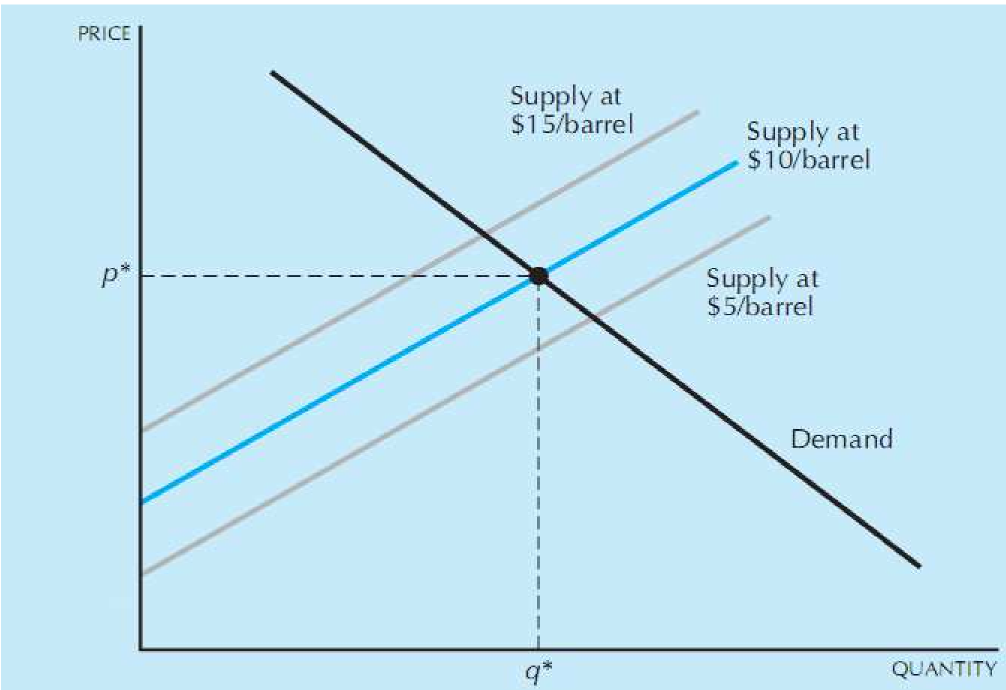


图 23.9：授权方案。在授权方案下，实际汽油的供给曲线将位于以下两条供给曲线之间：一条是若所有石油均按进口价格销售时的汽油的供给曲线；另一条是若所有石油均按国内石油价格销售时的汽油的供给曲线。

石油的边际成本确实下降了，这表示汽油的价格也下降了。但是我们要看看谁为政府的这种政策买单？国内石油生产者！美国按照 15 美元每桶的价格购入外国石油，但却自称购买价为 10 美元每桶。国内石油生产者本来可以按照世界石油价格（15 美元每桶）出售，但是却只能按照 5 美元每桶出售。这相当于美国为国外的石油生产者发放补贴，补贴的资金却要国内石油生产者支付。

政府最终也放弃了这种方案，政府转而对国内石油生产者征税，这样他们就不能因为 OPEC 的行为而获得暴利。当然，征税也会打击国内石油的生产，因此提高了汽油的价格，但在当时，这种方案是议会能够接受的。

23.11 烟尘排放税 Versus 排放上限与交易

一些气候学家担心全球变暖问题，他们督促政府实施减少烟尘排放的政策。从经济学的观点来看，有两种政策特别有趣：**烟尘排放税**（carbon taxes）和**排放上限与交易**（cap and trade）。

烟尘排放税是对烟尘排放征税，而排放上限与交易体系则是颁发烟尘排放许可证，这些许可证可以在市场中买卖。为了看清这两种政策的差异，我们分析一个简单的例子。

最优排放

我们首先分析降低既定排放量的成本最小方法问题。假设有两个企业，它们当前的烟尘排放量为 $(\bar{x}_1, \bar{x}_2)$。企业 i 降低 x_i 单位污染的成本为 $c_i(x_i)$。图 23.10 给出了这个成本函数的可能形状。

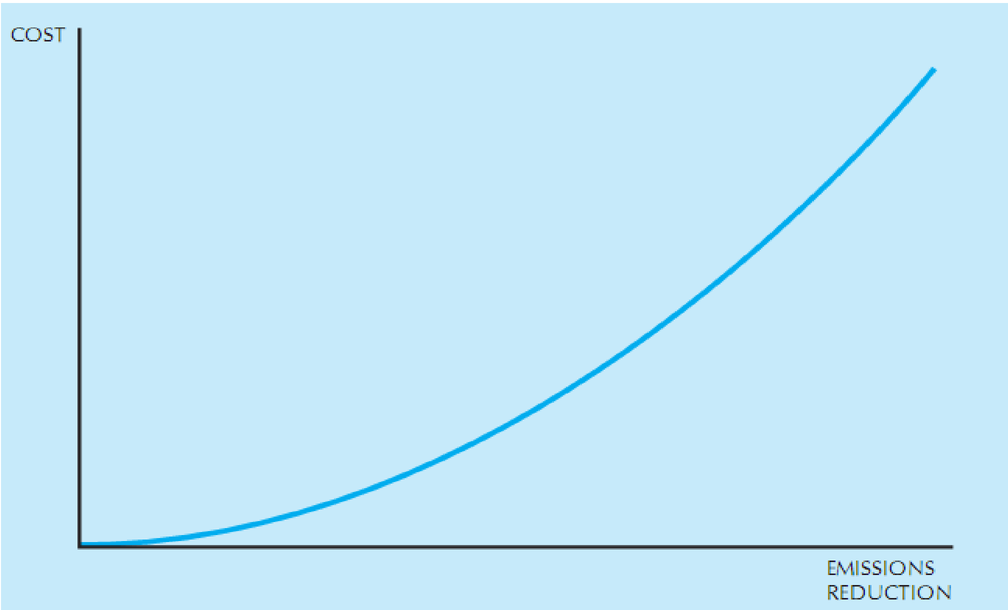


图 23.10：排污的成本函数。图中的曲线表示的是与降低排污量相关的成本。

政府的目标是以最小成本将排污量降低既定数量（T）。这个最小化问题可以写为

$$\min_{x_1,x_2} c_1(x_1)+c_2(x_2)$$
$$\text{s.t. } x_1+x_2=T.$$

如果政府知道企业降低排污的成本函数，那么在理论上，政府可以解出这个最优化问题，然后相应给每个企业分配具体的降污数量。然而，如果烟尘排放企业有成千上万家，这样的方法不可行。因此，需要找到能实现最优解的分权市场（decentralized market）的解决方法（一）。

我们分析一下这个最优化问题的结构。容易看出：在最优解处，每个企业减少污染的边际成本必定是相同的。否则让边际成本低的企业多排放、让边际成本高的企业少排放就是值得的。这个做法能以更低成本将排污量减少既定水平。

因此，我们得到了一个简单原则：在最优解处，每个企业减少排放的边际成本应该是相同的。在我们的两个企业的例子中，可用一个简单的图形找到这个最优解。令 $MC_1(x_1)$ 表示企业减少 x_1 单位污染的边际成本。假定降低排放的既定目标已实现，这样我们就可以将企业 2 减少排放的边际成本写成企业 1 减排量的函数：$MC_2(T-x_1)$。我们在图 23.11 中画出了这两个企业减排的边际成本曲线。这两条曲线的交点确定了如何将 T 单位减排量在两个企业中进行分配。

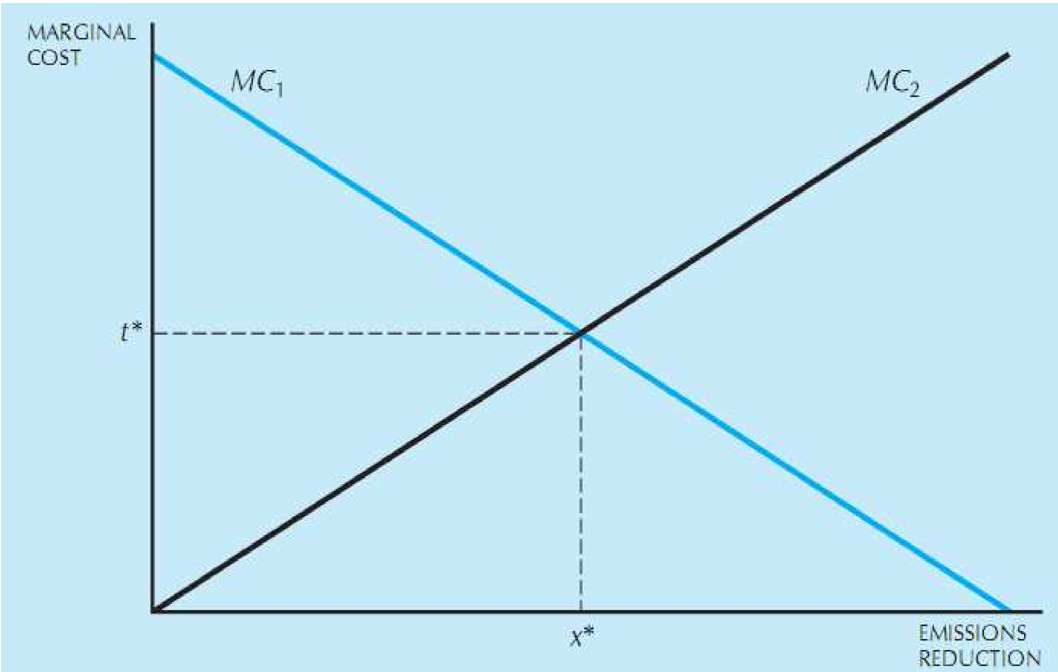


图 23.11：排污上限和交易市场中的均衡。点 t^* 给出了最优烟尘排放税和排放许可价格。

（一） 分权市场又译为“分散市场”，和计划市场（集权）相对，大致可以理解为自由竞争的市场。译者注。

烟尘排放税

我们也可以不直接求这个成本最小化问题的解，而是通过征税烟尘排放税方法得出的分散市场解。在这个架构内，政府规定烟尘排放税的税率为t。

如果企业 1 的初始排放量为$\bar{x}_1$，它要减少x_1单位的排放，因此它的最终排放量为$\bar{x}_1 - x_1$。如果每单位烟尘排放的税率为t，企业 1 需要缴纳的税额为$t(\bar{x}_1 - x_1)$。

在政府决定征收烟尘排放税之后，企业 1 选择的减排量（x_1）应该使它的总经营成本最小：减排的成本加上最终排放量需要缴纳的税额。这样就得到了下列成本最小化问题

$$\min_{x_1} c_1(x_1) + t(\bar{x}_1 - x_1).$$

显然，企业选择的减排量应能使得进一步减排的边际成本正好等于烟尘排放税的税率，即$t = MC_1(x_1).$

如果政府设定的烟尘排放税率为t^*，如图 23.11 所示，那么烟尘总的减排量将为既定目标数量T。因此，烟尘排放税提供了一种实现最优结果的分散市场方法。

排放上限和交易

假设政府不征收烟尘排放税，而是颁发**排放许可证**（emissions licenses），这种许可证可在市场上交易。持有许可证的企业可以排放一定数量的烟尘。政府选择颁发的许可证数量来实现既定的减排目标。

假设企业可在市场上购买排放量为x单位的许可证，每单位排放的价格为x元。企业 1 减排x_1单位烟尘的成本为$c_1(x_1) + p(\bar{x}_1 - x_1)$。显然，企业的选择要能使得每单位排污许可证价格等于边际成本，即$p = MC_1(x_1)$。也就是说，在它选择的排污水平处，减少一单位排污产生的成本正好等于不购买单位排污许可节省的成本。

因此，边际成本曲线表明排污量是每单位排污（许可证）价格的函数[一]。均衡价格是使得减排量等于既定目标量T时的价格。这个均衡价格等于最优烟尘税率t^*，请见图 23.11。

剩下的问题是政府如何分发排放许可证。一种方法是政府将许可证卖给企业。这种方法在本质上等同于烟尘排放税。政府可以选择一个价格水平，在此价格水平上有多少需求就卖多少许可证。或者政府可以选择既定的排污水平，然后将许可证拍卖，让企业们自己确定价格。这是“排放上限和交易”体系的一种类型。这两种政策都能得到相同的市场出清价格。

政府也可能根据某些程序分发许可证。这样的程序可能基于各种各样的标准，其中一个重要的标准可能是将这些珍贵的许可证分发给在政治立场上支持这些项目的企业。排污许可证可能根据下列客观标准分发：企业的雇员数量或者企业对某政治事业捐献资金数额等等。

[一] 排放量和减排量是一个硬币的两面，因此这句话也可以说成：减排量是每单位排污的价格的函数。译者注。

从经济学的观点看，政府在下列两种政策中选择哪一种并不重要，因为它们的结果在本质上是一样的：一是政府拥有许可证，然后将它们卖给企业（这基本等价于烟尘排放税体系）；二是政府将许可证分发给企业，但允许许可证在市场上买卖（这基本就是排污上限和交易的做法）。

如果政府选择的是排污上限和交易体系，企业就会想方设法花钱得到这些许可证。例如，他们会游说议会将许可证颁发给它们。这些游说费用应该作为这种政策体系的部分成本，正如我们以前讨论过的**寻租**（rent seeking）。当然，烟尘排放税体系也容易受到类似游说的影响。企业无疑会以各种理由寻求排污许可，但是经济学家认为，与排污上限和交易体系相比，烟尘排放税体系不大容易遭受政治上的操纵。

总结

1.某行业的短期供给曲线，是该行业内所有企业的短期供给曲线在水平方向上的加和。

2.某行业的长期供给曲线必须考虑企业进入和退出该行业的因素。

3.如果某行业允许自由进出，则长期均衡时行业内的企业数量，是保证企业能取得非负利润时行业能容纳的最大企业数量。这就是说，竞争性行业的长期供给曲线基本上是一条水平线，这条水平线的高度等于平均成本的最小值，即 $p^* = \min AC$ 。

4.如果存在阻止企业进入某盈利行业的生产要素，那么这些要素就能赚取经济租金。租金大小由该行业产品的价格决定。

复习题

1.若某行业由两家企业组成，且 $S_1(p) = p - 10$，$S_2(p) = p - 15$，那么价格为多大时，该行业供给曲线会出现折弯（kink）的现象?

2.在短期，香烟的需求是完全无弹性的。在长期，假设香烟的需求具有完全的弹性。现在若对香烟征税，那么它对消费者支付的价格有何影响？分别分析短期和长期的情形。

3.对还是错？大学校园附近的便利店价格较高，因为它们必须支付较高的租金。

4.对还是错？在长期行业均衡时，没有一家企业会亏损。

5.根据本章介绍的模型，说出是什么因素决定了进入或退出某行业企业的数量？

6.（选择）根据本章介绍的进入模型可知，若某行业中的企业数量越多，则行业长期供

给曲线（越陡峭，越平坦）。

7. 纽约市某出租车司机，经过对运营成本和劳动成本仔细分析后，认为在长期似乎也能够取得正利润。这违背了纯粹竞争模型吗？为什么？

复习题答案

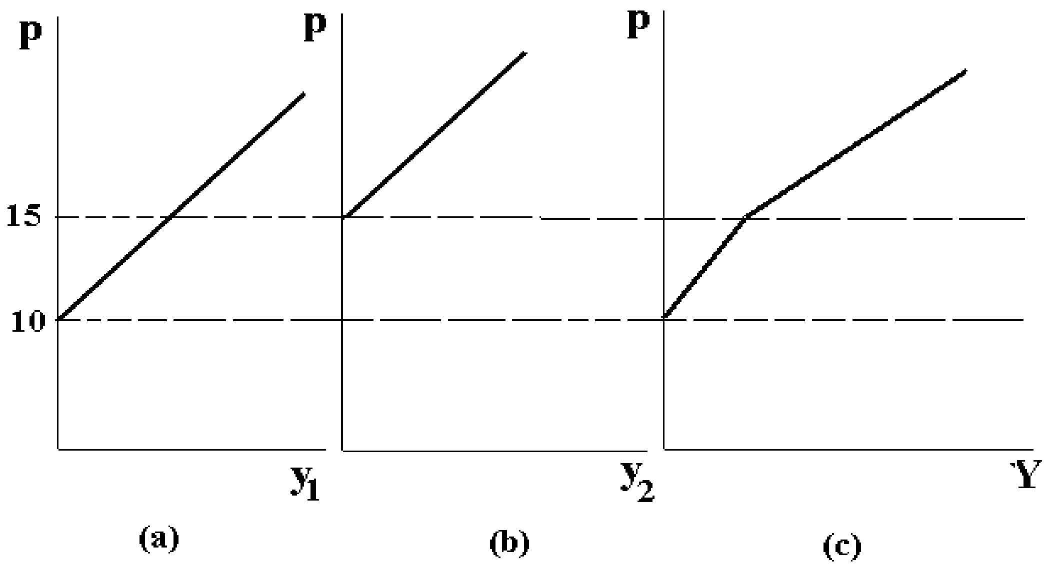
1.若某行业由两家企业组成，且 $S_1(p)=p-10$，$S_2(p)=p-15$，那么价格为多大时，该行业供给曲线会出现折弯（kink）的现象?

【复习内容】企业反供给曲线；行业反供给曲线

行业反供给曲线是所有企业反供给曲线在水平方向上的加总。

【参考答案】

由企业 1 的供给曲线 $S_1(p)=p-10$ 可知，当 $p\le 10$ 时，它的供给量为 0；当 $p>10$ 时，它进入行业开始供给；它的反供给曲线为 $p(y_1)=10+y_1$，如下图(a)所示。



由企业 1 的供给曲线 $S_2(p)=p-15$ 可知，当 $p\le 15$ 时，它的供给量为 0；当 $p>15$ 时，它进入行业开始供给；它的反供给曲线为 $p(y_2)=15+y_2$，如下图(b)所示。

由于行业反供给曲线是所有企业反供给曲线在水平方向上的加总，因此将企业 1 和 2 的反供给曲线加总后可得到行业的供给曲线，如图（c）所示。

从图（c）明显可以看出，当 $p=15$ 时，行业供给曲线出现了折弯。事实上，行业的供给曲线的表达式为

$$Y=\begin{cases}0 & 0\le p<10\\ p-10 & 10\ \le p<15\\ 2p-25 & p\ge 15\end{cases}\quad .$$

由此可见，即使每个企业的供给曲线是线性的，行业供给曲线也未必是线性的。

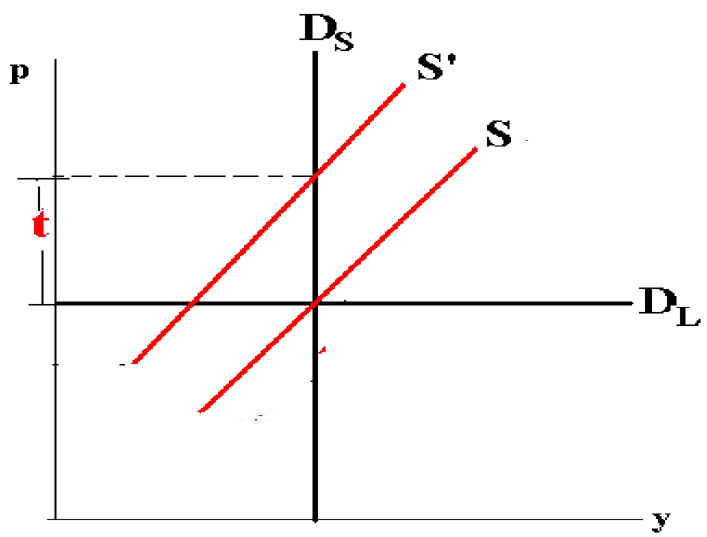
2.在短期，香烟的需求是完全无弹性的。在长期，假设香烟的需求具有完全的弹性。现在若对香烟征税，那么它对消费者支付的价格有何影响？分别分析短期和长期的情形。

【复习内容】长期和短期情形下的税收；税收分担

【参考答案】短期内，税收由消费者全部承担；长期中，税收由生产者全部承担。

为简单起见，我们假设政府对生产者征税。但是我们要指出，征税的本质是对交易行为征税，也就是说如果我们假设政府对消费者征税，得到的答案是完全一样的。

征税相当于增加了生产者的边际成本，因此，生产者的供给曲线会向左上方移动，即从曲线S移动到曲线S'。如下图所示。



我们先分析短期的情形。由于在短期，消费者对香烟的需求是完全无弹性的，这意味着短期需求曲线为一条垂线D_S。由上图可知，在这种情形下，消费者承担了全部单位税额t，因为征税后，消费者支付的价格上升幅度恰好等于t。

最后分析长期的情形。在长期中，消费者对香烟的需求具有完全弹性，这意味着长期需求曲线为一条水平线D_L。由上图可知，在征税后，消费者支付的价格没有发生变化，这意味着生产者承担了全部税收。

3.对还是错？大学校园附近的便利店价格较高，因为它们必须支付较高的租金。

【复习内容】固定要素和经济租金

【参考答案】错误。因为是价格决定了较高的租金，而不是相反。

我们用下图阐述经济租金。图中 AVC 表示**除了**便利店店铺之外的所有其他要素的平均成本曲线。（假设店铺是唯一的固定要素。）如果产品的价格为 p^*，那么店铺贡献的“利润”是图中矩形的面积：这就是经济租金。经济租金是店铺在完全竞争市场出租时获得的租金，它迫使利润为零。

包含店铺价值在内的平均成本曲线，我们在图中用 AC 标记。如果我们正确衡量了店铺的价值，便利店的经济利润应该恰好等于零。因为便利店的均衡租金要保证支付地租后利润为零，因此必有

$$p^*y^* - c_v(y^*) - \text{租金} = 0$$

或

$$\text{租金} = p^*y^* - c_v(y^*) \qquad （1）$$

这正好是我们介绍过的生产者剩余。的确，经济租金和生产者剩余是同一个概念，只是站在不同的角度看问题。因此，我们可用边际成本曲线左侧的面积来计算租金。

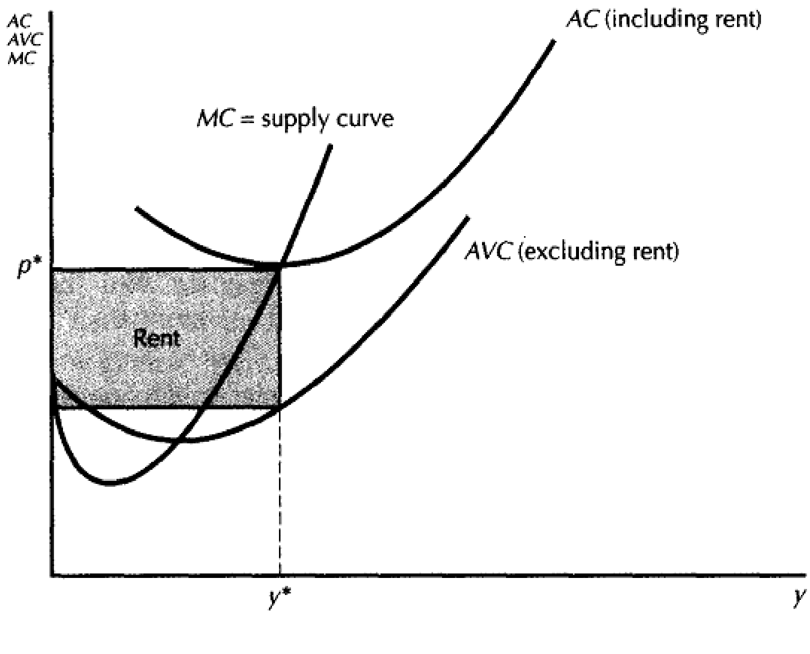


图 1：**便利店店铺的经济租金**。矩形面积代表店铺的经济租金。

根据（1）式给出的经济租金的定义，现在容易看出我们前面说过的事实：**是均衡价格决定了经济租金，而不是相反**。企业沿着它的边际成本曲线供给产品——这和固定要素的支出无关。租金将调整到直至利润等于零。

4.对还是错？在长期行业均衡时，没有一家企业会亏损。

【复习内容】长期行业均衡

【参考答案】正确。

在长期，企业可以自由进入或退出竞争行业。如果某企业在长期是亏损的，它就没有必要继续呆在行业内，因此我们可以预期这样的企业会退出该行业，因为退出后，它可以将亏损降低为零。所以，题目中的说法是正确的。

5.根据本章介绍的模型，说出是什么因素决定了进入或退出某行业企业的数量？

【复习内容】企业进入或退出行业的条件

【参考答案】

由长期行业均衡模型可知，新企业是根据行业内的原有企业是否盈利（利润是否为正）判断是否进入该行业的。如果行业内原有企业的利润为正，则意味着这个行业就有利可图，因此新企业就会进入该行业。

6.（选择）根据本章介绍的进入模型可知，若某行业中的企业数量越多，则行业长期供给曲线（越陡峭，越平坦）。

【复习内容】长期供给曲线

【参考答案】越平坦

若某行业中的企业数量越多，则行业长期供给曲线越平坦。这是由于如果市场中只有一家企业，并且市场价格上升了 Δp ，它会多生产比如 Δy 单位的产品。若市场中有 n 家企业而且价格上升了 Δp ，每家企业都会多生产 Δy 单位的产品，因此市场上的供应量增加了 $n\Delta y$ 。这表明，供给曲线会随着市场中企业数量的增多而变得越来越平缓，因为供给量对价格变动越来越敏感。

7. 纽约市某出租车司机，经过对运营成本和劳动成本仔细分析后，认为在长期似乎也能够取得正利润。这违背了纯粹竞争模型吗？为什么？

【复习内容】经济租金；竞争行业长期均衡利润为零

如果行业中的企业数量受到固定要素的限制，因此新企业不能自由进入，因此容易产生这样的错觉：该行业可能获得正的利润，因为似乎不存在迫使利润为零的经济力量。事实上，存在着一种经济力量可将利润压低为零。

曹乾（东南大学 caoqianseu@163.com）

如果你看到某企业的长期利润为正，这很可能因为你没有将全部成本包含进来，准确地说，你没有准确估计妨碍企业进入市场的因素的市场价值。

【参考答案】

没有违背纯粹竞争模型，在计算成本时，出租车司机忘记了营业执照的租金。

在计算成本时，必须注意将每种生产要素的价格按照它的**市场价格**，即它的机会成本计算。假设在题目中的例子，除了营业执照之外，我们考虑了其他的全部成本，最终计算出年利润为$\pi > 0$元。那么，该营业执照在市场上的价值如何？如果营业执照可以出租，比如出租一年，应该收取多少租金？

答案是：应收取的年租金为π元，这正是营业执照产生的“利润”。如果你租入出租车的营业执照，即使你不会开车，你也能“赚取”π元。因为我们对这个出租车司机的劳动也是按市场价格计算的，这意味着你可以雇佣这样的司机让他替你开车，你仍然能获得π元的利润。因此，这块营业执照的市场价值，也就是它的竞争性租金正好是π元。所以，出租车的经济利润为零。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 *2014* 版

24.垄断（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

24 垄断

在前几章，我们分析了竞争性行业的行为，如果行业中企业的规模普遍较小而且数量众多，那么这个行业就可能是竞争性行业。在本章，我们分析另外一个极端，即行业中只有一家企业的市场结构，这就是完全**垄断市场**（a monopoly）。

当某市场只有一家企业时，就不能再认为该企业是价格接受者。相反，垄断企业会认识到它对市场价格的影响力，因此它会选择能使利润最大的价格和产量组合。

当然，它也不能独立地选择价格和产出水平；对于任何给定的价格，垄断企业只能销售市场可以承受的产量。如果它定价高，那么它只能卖出很少的数量。消费者的需求行为将会约束垄断企业对价格和销量的选择。

我们可以认为垄断企业选择价格，令消费者根据这个价格选择购买数量；或者我们还可以认为垄断企业选择产量，令消费者决定对于这些产量他们愿意支付的价格。第一种方法可能更自然一些，但是第二种方法更便于分析。当然，如果计算时不出错的话，你就会发现这两种方法是等价的。

24.1 最大化利润

我们首先分析垄断企业的利润最大化问题。令 $p(y)$ 表示市场反需求曲线，$c(y)$ 表示垄断企业的成本函数，$r(y)=p(y)y$ 表示它的收入函数。垄断企业的利润最大化问题可以表示为

$$\max_{y} r(y)-c(y).$$

这个问题的最优条件非常直观：在最优产量水平上必有边际收入等于边际成本。如果边际收入小于边际成本，则企业应该减少产量，这么做收入会减少，但成本减少得更多，因此利润会增加。如果边际收入大于边际，则企业应该增加产量。只有边际收入等于边际成本时，企业才不会改变产量。

用代数的语言，我们可以将最大化条件写为

$$MR(y)=MC(y)$$

或

$$\frac{\Delta r}{\Delta y}=\frac{\Delta c}{\Delta y}.$$

竞争性企业在利润最大化时也必须 MR=MC 的条件，只不过该情形下，边际收入等于价格，因此上述条件简化为 P=MC。

对于垄断企业来说，边际收入稍微有些复杂。如果垄断企业决定将产量增加Δy，这将对收入产生两种效应。第一种效应是它卖出更多的产品，收入增加了$p\Delta y$。但是，第二种效应，垄断企业将价格压低了Δp，而且**所有**的产量都按压低后的价格$(p-\Delta p)$销售。

所以产量变动Δy对收入的总效应为

$$\Delta r = p\Delta y + y\Delta p,$$

因此，收入变动除以产量变动即边际收入为

$$\frac{\Delta r}{\Delta y} = p + \frac{\Delta p}{\Delta y}y.$$

(这个表达式就是我们在第 15 章推导出的边际收入表达式。你可以先复习一下然后再继续学习。)

另外一种推导方法，是可认为垄断企业同时作出产量和价格决策——当然，它的决策要受到需求曲线的约束。如果垄断企业想卖更多的产品它必须降低价格。但是降低价格意味着它的所有销量都要按照这个降低后的价格出售，新价格不只适用多卖的那部分产品。因此就得到了$y\Delta p$这一项。

在竞争情形下，如果某企业的定价小于其他企业的价格，则它会迅速占有整个市场。但在垄断情形下，垄断企业已经拥有了整个市场；当它降低价格时，它必须考虑降价对它的全部销量的影响。

根据第 15 章的讨论，我们可以将边际收入用弹性进行表达：

$$MR(y) = p(y)[1+\frac{1}{\varepsilon(y)}]$$

这样就可以将“边际收入等于边际成本”这个最优条件写为：

$$MR(y) = p(y)[1+\frac{1}{\varepsilon(y)}] = MC(y). \tag{24.1}$$

由于需求价格弹性一般为负，我们可以将上式写为

$$p(y)[1-\frac{1}{|\varepsilon(y)|}] = MC(y).$$

从这些式子我们容易看出它与竞争情形的联系：在竞争情形下，企业面对的是水平的需求曲线，该曲线的弹性为无穷。这表示$1/\varepsilon = 1/\infty = 0$，因此上式变为$p(y) = MC(y)$。即对于竞争企业来说，均衡条件是价格等于边际成本。

注意，垄断企业绝对不会在需求曲线**缺乏弹性**的那一段进行生产。因为若$|\varepsilon|<1$，则$1/|\varepsilon|>1$，此时边际收入为负，因此它不可能等于边际成本。想想需求曲线缺乏弹性的含义，

我们就能看清这一点：若$|\varepsilon|<1$，则减少销量可以增加收入，而且减少产量必然会减少总成本，因此利润必定上升。所以，任何一点只要$|\varepsilon|<1$，该点就不可能是垄断企业的利润最大化的选择点，因为此时减少销量可以增加利润。这也就是说，能够实现利润最大化的点必然满足$|\varepsilon|\geq 1$。

24.2 线性需求曲线和垄断

假设垄断企业面对的需求曲线为线性

$$p(y)=a-by.$$

则收入函数为

$$r(y)=p(y)y=ay-by^2,$$

边际收入函数因此为

$$MR(y)=a-2by.$$

(这个式子可从第 15 章结尾处的式子直接推出。使用微积分可以容易推出上式。如果你不懂微积分，就记住它，因为我们有时会用到它。)

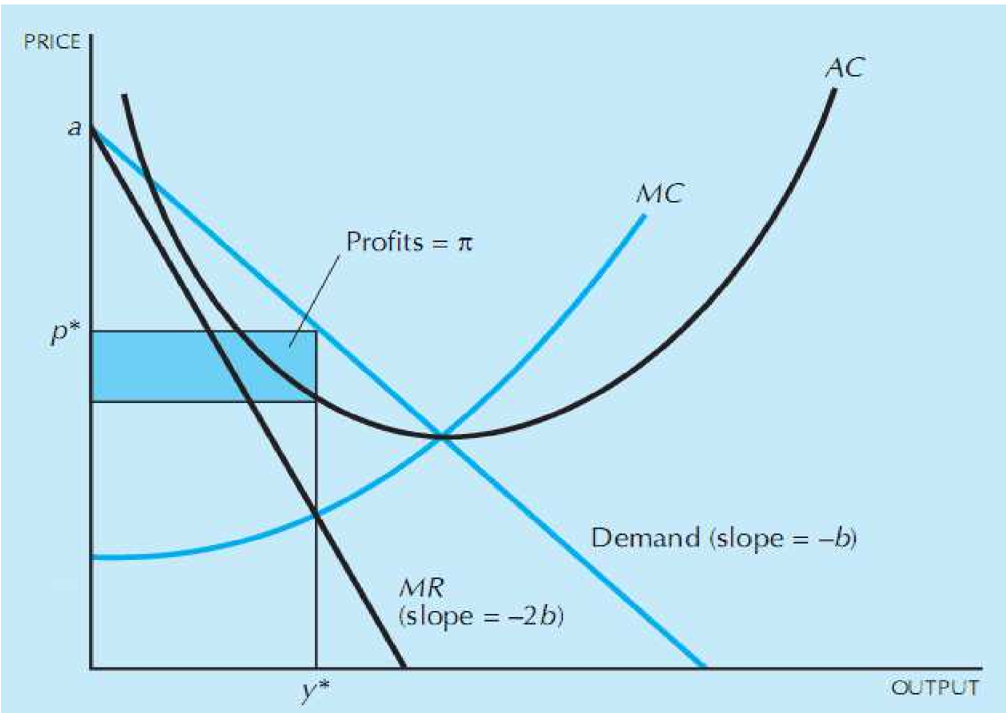


图 24.1：垄断企业需求曲线为线性的情形。垄断企业利润最大化的产量位于边际收入等于边际成本之处。

注意，边际收入函数和需求函数的纵截距相同，但前者的斜率是后者的二倍。这样，我们就很容易画出边际收入曲线。我们知道纵截距为 a。为了得到边际收入曲线的横截距，只要取需求曲线横截距的一半即可。将这两个截距点用直线连接起来就得到了边际收入曲线。我们在图 24.1 中画出了这样的需求曲线和边际收入曲线。

最优产量 y^* 位于边际收入曲线与边际成本曲线相交之处。垄断企业在这个产量 $p(y^*)$ 上将会索要最高的价格。在该产量上，垄断企业的收入为 $p(y^*)y^*$，成本为 $c(y^*)=AC(y^*)y^*$，二者相减即得到利润，如图 24.1 阴影区域所示。

24.3 加成定价

我们还有一种方法分析垄断企业的最优定价策略，这就是使用边际收入和弹性的关系表达式。将（24.1）式变形可得

$$p(y)=\frac{MC(y)}{1-1/|\varepsilon(y)|} \qquad （24.2）$$

这个式子表明市场价格是边际成本的一个加成价（a markup over marginal cost），其中加成数量取决于需求价格弹性。加成数为

$$\frac{1}{1-1/|\varepsilon(y)|}.$$

由于垄断企业总是在需求曲线具有弹性处生产，我们断定 $|\varepsilon|>1$，因此加成数大于 1。

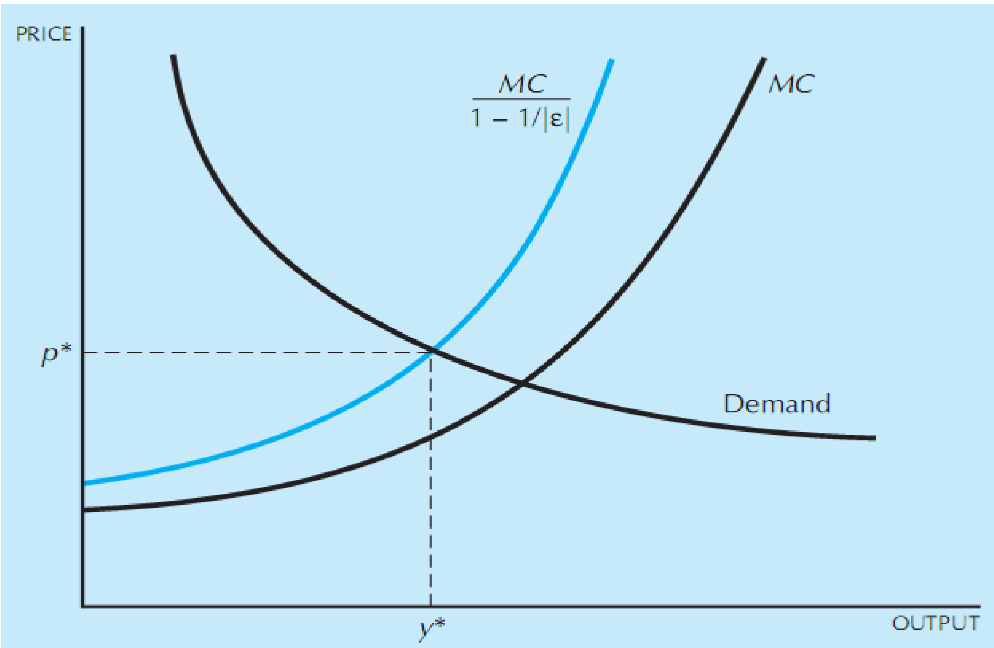


图 24.2：需求曲线的弹性固定不变的情形。曲线 $MC/(1-1/|\varepsilon(y)|)$ 与需求曲线的交点对应的产量就是最优产量。

在需求曲线的弹性固定不变的情形下，这个表达式更为简单，因为此时$\varepsilon(y)$是个常数。若面对的是这种需求曲线，垄断企业将按边际成本加成定价时，加成数是一个**常数**。如图24.2 所示。在该图中，曲线$MC/(1-1/|\varepsilon(y)|)$的高度是边际成本曲线高度的$1/(1-1/|\varepsilon(y)|)$倍；最优产量位于$p = MC/(1-1/|\varepsilon(y)|)$之处。

例子：税收对垄断企业的影响

如果某企业的边际成本固定不变，现在问：如果对它征收从量税，将会对它的价格有何影响？显然边际成本上升的幅度等于单位税额，但是它对市场价格有何影响呢？

我们首先考虑需求曲线为线性的情形，如图 24.3 所示。当边际成本曲线MC向上移动到$MC+t$（移动幅度等于单位税额t），边际收入和边际成本曲线的交点则向左移动。由于需求曲线的斜率是边际收入曲线斜率的一半，价格上升幅度等于单位税额的一半即$t/2$。

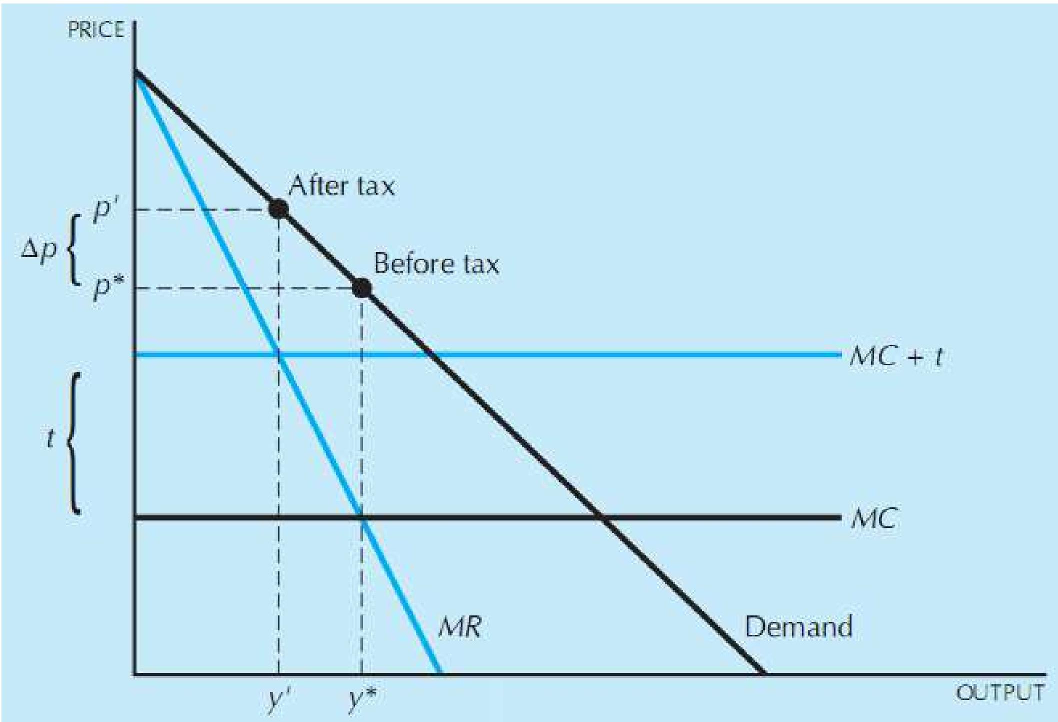


图 24.3：线性需求和税收。若某垄断企业的需求曲线为线性的，对其征税后，市场价格将上升，上升幅度等于单位税额的一半。

使用代数方法也容易看清楚这一点。征税后，最优条件变为边际收入等于边际成本加上单位数额，即

$$a-2by=c+t.$$

解出 y 可得

$$y=\frac{a-c-t}{2b}.$$

因此，产量的变动为

$$\frac{\Delta y}{\Delta t}=-\frac{1}{2b}.$$

需求曲线为

$$p(y)=a-by,$$

因此价格的变动等于 $-b$ 乘以产量的变动：

$$\frac{\Delta p}{\Delta t}=(-b)\times(-\frac{1}{2b})=\frac{1}{2}.$$

上式中出现 1/2 的原因是我们线性需求函数以及边际成本不变的假设。在这两个假设之下，价格上涨幅度小于单位税额。这个结论具有一般性吗？

答案是否定的——一般来说征税后，价格上涨幅度可能大于也可能小于单位税额。我们再来看看垄断企业需求曲线的弹性为固定不变时的情形。该情形下我们有

$$p=\frac{c+t}{1-1/|\varepsilon|},$$

因此

$$\frac{\Delta p}{\Delta t}=\frac{1}{1-1/|\varepsilon|},$$

该式显然大于 1。在这种情形下，价格上升幅度大于单位税额，这意味着消费者不仅承担了全部的税收，而且还要忍受额外的价格增量。

另外一种税是利润税。在这种情形下，垄断企业要将利润的一定百分比 τ 作为税收交给政府，企业面对的最大化问题为

$$\max_{y}\ (1-\tau)[p(y)y-c(y)].$$

但是能使利润 $p(y)y-c(y)$ 最大化的产量水平 y，同样也能使得 $(1-\tau)[p(y)y-c(y)]$ 最大。因此，利润税对垄断企业的产量决策没有影响。

24.4 垄断的低效率

竞争行业在价格等于边际成本之处生产。垄断行业在价格大于边际成本之处生产。因此，一般来说，与竞争相比，垄断行业的价格较高、产量较低。正因为此，对于同一个行业来说，垄断情形下消费者的状况通常比竞争情形下要差。

但是，出于同样的原因，企业在垄断情形下的状况要好于竞争情形时！如果我们同时考虑消费者和企业的状况，则很难判断出到底哪一种市场是“更好”的市场制度。回答这样的问题，似乎要对消费者和企业的相对福利作出价值判断。然而，如果我们仅站在效率的角度，那么我们就会反对垄断。

下面分析垄断的情形，如图 24.4 所示。假设我们不需花费成本就能让该企业变为竞争的企业，即命令它接受预先设定的市场价格。这样我们就得到了竞争的价格和产量(p_c, y_c)，其中下标c表示竞争。如果我们不强迫企业这么做，那么它就是个垄断企业，它会制定价格和产量决策以使利润最大化，垄断情形下的价格和产量为(p_m, y_m)，其中下标m表示垄断。

我们以前说过，一种经济制度如果它已无法做到在不损害其他人的利益情况下让任何人的状况变好，那么它就是帕累托有效率的（Pareto efficient）。垄断情形下的产量水平是帕累托有效率的吗？

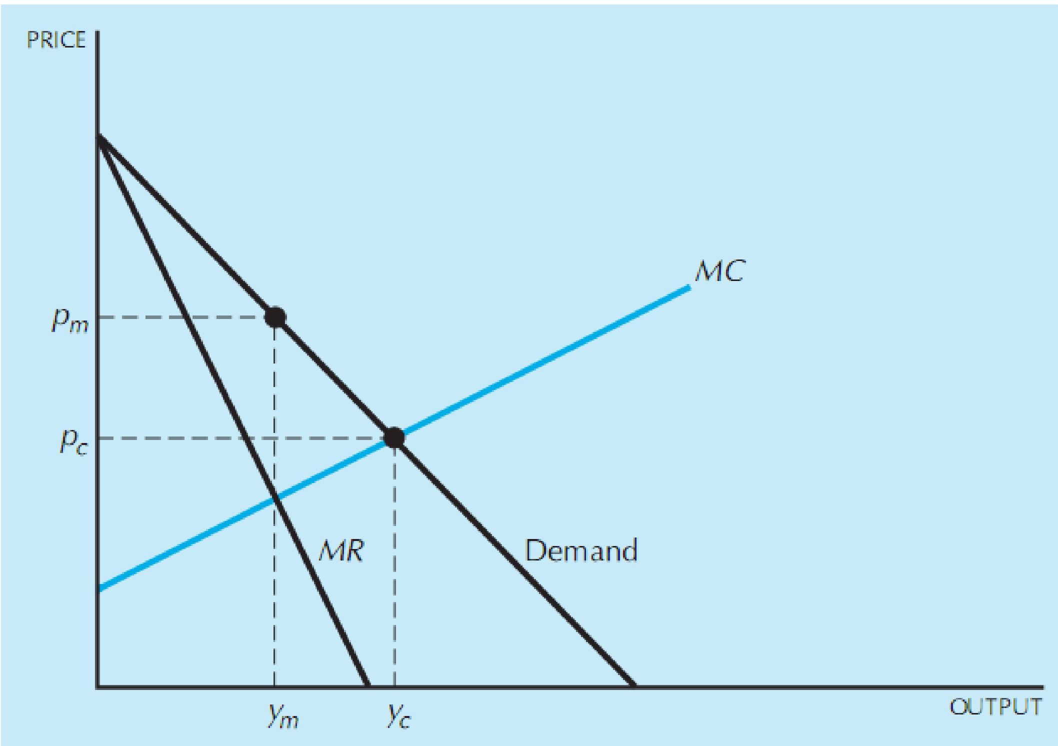


图 22.4：垄断的低效率。垄断企业的产量小于竞争时的产量，因此是帕累托低效率的。

由反需求曲线的定义可知，在每个产量水平上，$p(y)$ 衡量消费者愿意为额外一单位产品支付的价格。由于对于闭区间 $[y_m \quad y_c]$ 中的任何产量水平，$p(y)$ 都大于 $MC(y)$。这意味着对于这个区间的产量，消费者愿意支付高于边际成本的价格，以让企业多生产。显然，这就存在着帕累托改进（Pareto improvement）的余地！

例如，在垄断产量 y_m 处，由于 $p(y_m) > MC(y_m)$，这意味着如果企业多生产一单位产品，一定有人愿意支付高于边际成本的价格购买这件产品。如果企业真得生产出这一单位产品，而且以价格 p（该价格满足 $p(y_m) > p > MC(y_m)$）卖给该消费者。那么，该消费者的状况改善了，因为他原本愿意支付 $p(y_m)$ 元购买该件产品，但他实际支付价格为 p 元，而 p 小于 $p(y_m)$。类似地，垄断企业生产这件产品的成本为 $MC(y_m)$，但它却卖得了 p 元，而 p 打于 $MC(y_m)$。由于以前的产量仍按原来的价格销售，因此对原来的那部分利润没有影响。但是该企业再多销售一单位产品后，交易双方都得到了一些额外剩余，交易双方的状况都得到了改善，而且其他的人的状况没变差，这样我们就找到了一个帕累托改进。

垄断为什么会造成低效率？有必要分析一下。有效率的产量是使得消费者的边际支付意愿等于生产者的边际生产成本的那个产量。竞争企业会将自己的边际成本与消费者的支付意愿进行比较。但是，垄断企业还要考虑产量增加的另外一种效应，即对边际以内（inframarginal）[一] 那些产量销售收入的影响，这些边际以内的产量和效率无关。如果垄断企业不必降低它当前销售的所有其他边际以内那些产品的价格，那么它总是愿意以低于当前售价的价格销售额外一单位产品。

24.5 垄断的净损失

我们已经知道垄断是低效率的，我们还想知道它的效率低到什么程度。有没有办法衡量由于垄断造成的效率损失？如果消费者支付的价格为 p_m 而不是 p_c，我们就可以计算出消费者的损失——计算消费者剩余的变动。类似地，如果企业索要的价格为 p_m 而不是 p_c，我们可以用生产者剩余的变动衡量它的利润变动。

我们平等地对待企业（更准确地说是企业所有者）和消费者，这样我们自然地可将企业的利润和消费者剩余相加。企业的利润变动（等于企业的生产者剩余变动），衡量企业所有者为了得到垄断的高价格而愿意支付的代价；消费者剩余变动衡量因支付了垄断者索要的高价而必须给与他们多少补偿。因此，二者之差就可以衡量垄断的净收益（或净成本）。

由垄断产量变为竞争产量而引起的生产者剩余变动和消费者剩余变动，可用图 24.5 表

[一] Inframarginal 是指“边际以内”，它是相对于“在边际上”（the margin）而说的。例如：若某企业生产 10 单位产品，则**在边际上**通常是指第 11 单位产品，而**边际以内**是指 1~10 单位产品。一般来说，边际以内是个定性而不是定量的修饰语，因此在这个例子中边际以内不是一定指 10 单位产品。译者注。

示。由图可知，垄断者的生产者剩余面积减少了区域 A，这是因为原来的垄断产量要按较低价格出售。同时，生产者剩余面积又增加了区域 C，这是由于销售额外产量带来的利润。

消费者剩余的面积增加了区域 A，因为消费者能按较低的价格购买原来的消费量；消费者剩余的面积还增加了区域 B，因为他们从额外销量中得到了某些剩余。区域 A 从垄断企业转移到消费者手中——这一转移使企业的状况变坏、消费者的状况变好，但是总剩余没变。区域 B+C 表示剩余真正增加的部分，这个区域衡量的是消费者和生产者对额外产量的估价。

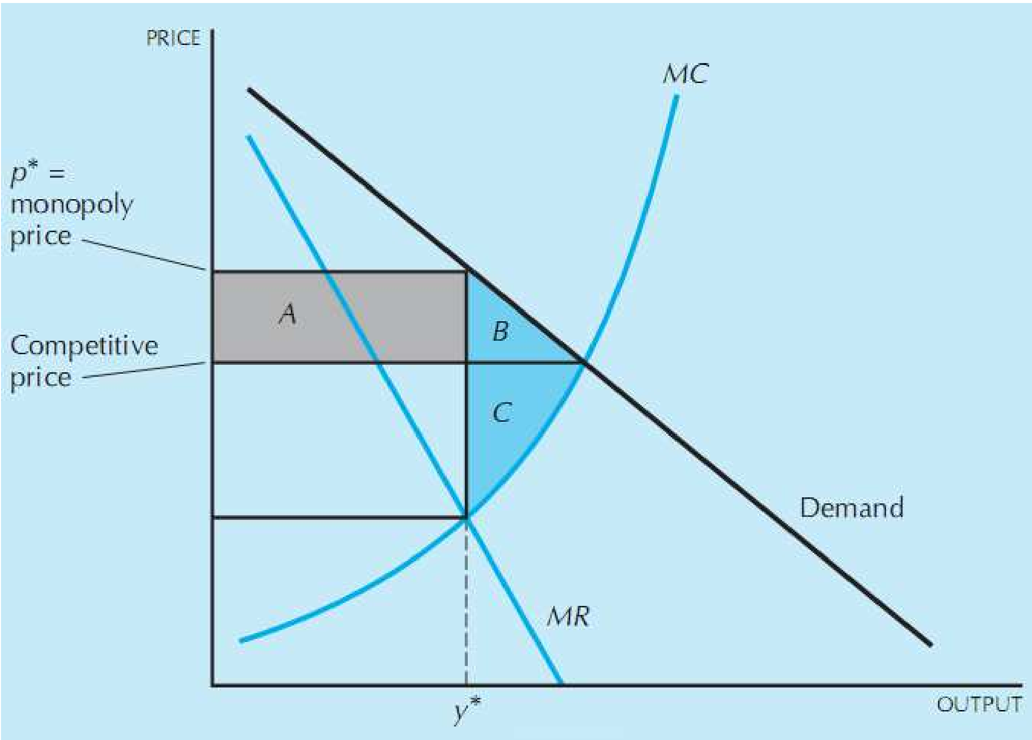


图 24.5：垄断的净损失。垄断引起的净损失为区域 B+C。

区域 B+C 称为由垄断造成的**净损失**（deadweight loss）。它衡量的是，与支付竞争价格相比，因支付垄断价格而造成的人们状况的恶化程度。垄断造成的净损失，和税收造成的净损失一样，都反映了因产量减少而损失的价值，这个价值是按照人们对单位产品的支付意愿衡量的。

为什么净损失衡量的是产量减少而损失的价值？为了看清这一事实，假设垄断企业在垄断产量的基础上额外增加一单位产量。这一单位产品的价值就是市场价格，而生产这一单位产品的成本就是边际成本。因此，额外多生产一单位产品的“社会价值”就是价格减去边际成本。假设在上面的基础上再多生产一单位产品，类似地，该单位产品的社会价值等于该单位产品的价格和边际成本之差。以此类推。当我们从垄断的产量水平移动到竞争的产量水平时，只要把需求曲线和边际成本曲线之间的距离“相加”，就得到了产量减少而损失的价值。

这两条曲线之间（从垄断产量到竞争产量）的区域，就是垄断造成的净损失。

例子：专利的最佳保护期限

专利让发明人享有对发明利益的独占权，当然专利有一定的保护期限限制。因此，专利提供了一种有限的垄断。保护专利的原因是鼓励发明创造。如果缺少专利保护制度，那么个人和企业可能不愿意投资于专利的研发，因为即使他们创造出来，也很快被竞争者仿效。

美国对专利的保护期限是 17 年。在保护期限内，专利持有人对他的发明享有独占权；过了该期限，任何人都可以使用这样的专利技术。专利保护期限越长，发明人从中获利越多，因此他们越有研发的干劲。然而，独占的时间越长，它造成的净损失越大。较长专利保护期的收益是它鼓励了创新；然而这却是以鼓励垄断为代价的。专利的“最佳”保护期限是使得这两种互相冲突的效应取得平衡的期限。

耶鲁大学学者威廉.诺德豪斯，曾分析过专利的最佳保护期限问题。他指出，最佳期限问题非常复杂，涉及未知关系众多。然而，通过简单的计算就可以发现，专利现有的保护期限和上面估计的利益和成本是否相一致。

诺德豪斯发现，对于普通的发明来说，17 年的保护期限能实现大约 90%的效率，也就是说，能够实现 90%的最大可能的消费者剩余。由这些计算结果可知，似乎没有迫切的理由去大幅修改现有的专利保护制度。

例子：专利丛林

对专利的知识产权进行保护有利于创新，但是这项权利容易被滥用。有些研究表明，由于知识产权保护已延伸到商业流程、软件和其他领域，专利的质量降低了。

我们可以认为一项专利有三个维度：长、宽和高。“长度”是指该专利适用的保护期限；“宽度”是专利内容涉及的范围；“高度”是专利的创新程度。不幸的是，只有长度容易量化。其他两个维度具有很大的主观性。

因为近年来获得专利非常容易，很多企业对自己从事的业务各个方面都申请了专利，从而形成了专利组合（patent portfolios）。想进入该行业的人很快就会发现，它要竞争的企业个个拥有大量专利，从而会产生身处**专利丛林**（patent thicket）的感觉。

即使经营良好的企业也会发现有必要获得投资组合。2004 年，微软公司向 InterTrust 技术公司支付了 4.4 亿美元，以购买该公司的一套和计算机安全相关的专利组合的使用权。同年，微软公司和 Sun Microsystems 公司签订了一项长达 10 年期限的合同，依据该合同微软向该公司支付了 9 亿美元以解决专利争端。2003-2004 年期间，微软获得的了 1000 多项专利。

为什么企业那么重视专利组合?对于象微软这样的大公司来说，专利组合的主要作用体现在当与其他公司签订专利交叉许可（cross-license）协议时能增加谈判的筹码[一]。

每个公司建立起来的专利丛林，宛如冷战时期美国和前苏联的原子弹战略。这两个国家都将足够的原子弹瞄准对方，如果谁敢率先发起攻击，大家都“同归于尽”。因此，哪一方都不敢冒险发动攻击。

专利丛林类似原子弹战略。如果 IBM 起诉惠普（HP）专利侵权，惠普也会拿出自己的专利反诉 IBM，说对方的某些技术也侵犯了它的利益。即使有些公司不是特别愿意申请专利，它们也必须这么做，因为只有这样，它们才能有足够的能力防御类似的诉讼。

专利丛林的“原子弹”按钮是“预先禁令”（preliminary injunction）[二]。在某些情形下，法官会责令被告停止出售涉嫌侵权的商品。有时，被告会因此付出惨重的代价。1986 年，由于法院下达预先禁令，柯达公司被迫完全关闭了自己的快速成像业务。最终，柯达因专利侵权不得不赔偿了 10 亿美元。

预先禁令停止生产的规定，对某些企业是巨大的威胁，对某些企业则毫无影响。因为后面的这种企业什么也不生产。例如 InterTrust 公司不出售任何东西，当然如果说它出售什么东西的话也可以，它出售的产品完全是专利技术。因此，InterTrust 公司可以以起诉对方侵权来威胁对方，但它却不怎么担心对方的反诉讼威胁。

例子：土豆供给的管理

每个人都熟悉石油输出国组织（OPEC），这个国际石油卡特尔试图通过规定每个成员国产油配额的方式，来影响石油价格。在美国，企业之间为抬升价格而进行的合作通常是违背反托拉斯法的，但也有一些例外行业。

最著名的例子是农业生产者。1922 年的《凯普-伏尔斯蒂德法》（Capper-Volstead Act）明企鹅之处农民不受联邦反托拉斯法约束。结果一些“农业营销委员会”应运而生，它们试图自发地管制农产品的供给。

例如，美国土豆种植者协会于 2005 年 3 月成立，它已与很多种植者签订了协议，他们的土豆种植面积占美国种植土豆总面积的 60%以上。2005 年，该协会宣布将土豆产量减少 680 万袋，每袋约 100 磅。根据《华尔街日报》报道，这意味着市场将减少 13 亿份炸薯条的供给。

[一]交叉许可是指交易各方将各自拥有的专利、专有技术的使用权相互许可使用，互为技术提供方和接受方。双方权利对等，一般不需支付使用费。译者注。

[二] 预先禁令，或称临时禁令，指起诉后、判决前由法院签发的禁令，禁止被告实施或继续某项行为。预先禁令一般在诉讼问题判决前做出，目的是避免某些不可挽回的损失。译者注。

曹乾（东南大学 caoqianseu@163.com）

24.6 自然垄断

我们已经知道，某行业帕累托有效率的产量出现在价格等于边际成本之处。但是，垄断生产者在边际收入等于边际成本之处生产，因此它的产量太少。勒令垄断企业消除低效率似乎并不难以做到，因为你只要将价格设定成等于它的边际成本，企业的利润最大化目标会帮你完成其他的事情。不幸的是，这种分析忘记了一个重要问题：在这样的价格水平上，垄断企业的利润可能会负。

我们用图 24.6 说明这样的情形。在图中，平均成本曲线的最低点位于需求曲线的右侧，需求曲线和边际成本曲线的交点位于平均成本曲线的下方。尽管产量 y_{MC} 是有效率的，但利润此时为负。如果你硬性规定它的产量水平必须为 y_{MC}，那么垄断企业宁可退出该行业。

这类情形通常在公用事业领域出现。以煤气公司为例。对于煤气公司来说，它要投入数额巨大的固定成本，例如铺设和维护输气管道的成本；但是，一旦这些管道铺设好以后，额外多输送一单位煤气的成本却很小，即边际成本很小。类似地，一家区域电话公司前期也要投入大量的固定成本，因为它要架设电话线和建立交换网络；然而当这一切做好以后，它提供额外一单位电话服务的边际成本却很小。如果高额的固定成本和较小的边际成本并存，你很容易得到图 24.6 的情形。这种情形称为**自然垄断**（natural monopoly）。

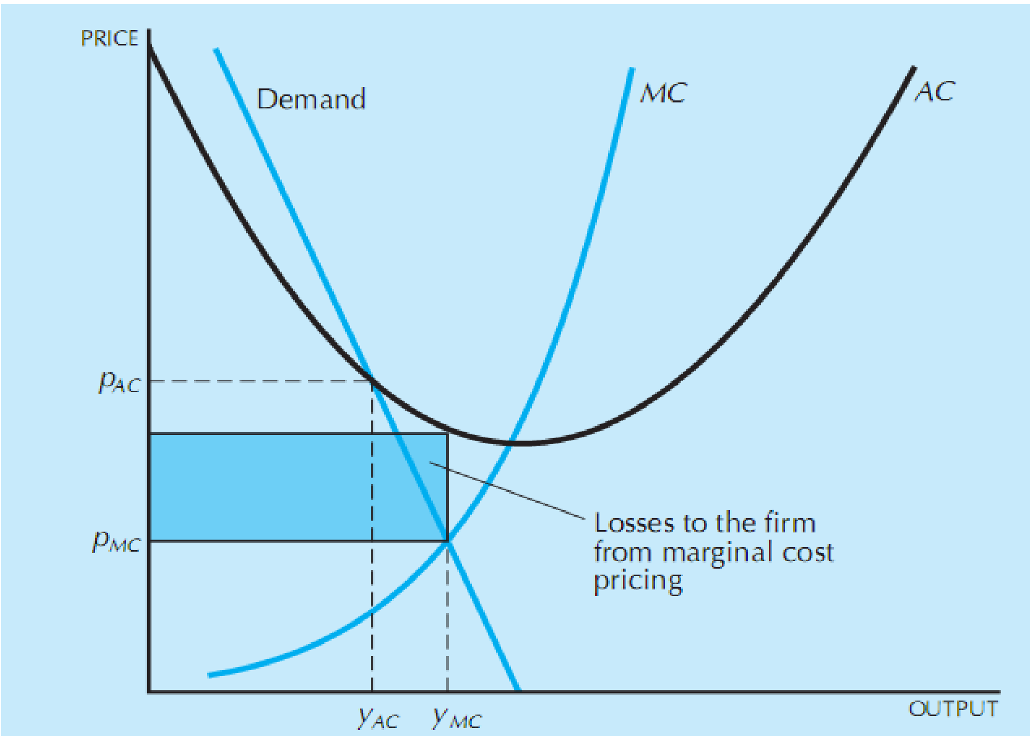


图 24.6：自然垄断。如果一个自然垄断企业在价格等于边际成本之处生产，那么，它的产量 y_{MC} 是有效率的，但是在此产量上它连成本都收不回来（即亏损）。如果规定它在价格等于平均成本之处生产，则它的产量 y_{AC} 恰好能够补偿总成本，但与有效率的产量相比，这个产量显然太低了。

自然垄断厂商的垄断价格决策会造成帕累托低效率，这和前面的分析是一样的。如果我们不喜欢这种低效率的结果，而强迫它按照竞争价格生产又行不通，因为这样做它只能取得负利润。还有其他的管制办法吗?世界上大部分国家都对自然垄断进行管制或者由政府亲自经营这样的企业。不同的国家有不同的方法。在某些国家，电话行业是由政府运行，而在其他国家，则由私营企业运行，但政府要对这样的企业进行管制。这两种方法各有优缺点。

我们先分析政府对自然垄断企业进行管制的情形。如果政府对受管制企业不给予补贴，则该企业必定要取得正利润，这意味着它必须在平均成本曲线上（on）或者在该曲线的上方(above)生产。如果它向所有愿意付费的消费者提供服务，那么它还必须在需求曲线上生产。因此，受管制企业的生产之处是类似图 24.6 中(p_{AC}, y_{AC})这样的点。由于在这一点上，企业产品价格等于平均成本，因此它能恰好将全部成本收回，但与有效率的产量相比，此时它的产量太低了。

政府通常采取上述这种办法为自然垄断企业制定合理的价格政策。政府制定的价格通常要考虑共用事业的性质。理想的定价方法是，制定的价格要能恰好使得企业不赚不赔，即让垄断企业在价格等于平均成本之处生产。为了合理定价，政府必须确定垄断企业的真实生产成本。这项任务政府通常交给共用事业委员会来做。他们调查垄断企业的生产成本，并计算出该企业真正的平均成本，然后将价格设定在平均成本水平上。（当然，其中一项成本是由于该企业对他人借入资金引起的，因此该企业必须向它的股东或其他债权人支付利息。）

美国对这样企业的管制权限通常下方在州政府和地方政府一级。一般来说，电力、天然气和电话服务由州政府管制，而有线电视等则通常由比州政府更低的地方政府管制。

自然垄断的另一种解决之道是政府的公有企业经营。在这种情形下，最优解是在价格等于边际成本之处生产，由于该产量会亏损，政府每年会向企业支付一笔补贴，以维持企业的继续经营。这是公共汽车和地铁这类地方公交系统通常采用的做法。补贴本身也许并不意味着企业的经营是低效率的，它只反映了这类公共事业要投入巨额固定成本。

但是，话又说回来，补贴可能正意味着低效率！要想确定公有垄断企业的生产成本，并不比确定受管制的私营垄断企业成本更容易。政府管制部门要频繁出席听证会，为成本的合理性辩驳，然而内部政府官僚却不吃这一套，他们会躲避这种细致的审查。在这个意义上，公有垄断企业比受管制的私营垄断企业对公众更不负责。

24.7 垄断的成因

如果我们得到某家企业的成本和需求数据信息，我们怎么判断它是一家竞争企业还是垄断企业?一般来说，判断依据是平均成本曲线和需求曲线的关系，其中决定行因素是**最小有效率的规模**（minimum efficient scale，MES），它是指相对于需求的规模来说能使平均成本最小的那个产量水平。

如图 24.7 所示，我们在图中分别画出了两种产品的平均成本曲线和市场需求曲线。在

第一种情形下，市场可以容纳很多企业，每个企业的要价都接近于 p^*，每个企业的规模都相对较小。在第二种市场，只有一家企业能取得正利润。我们可以预期第一个市场可能是竞争市场，第二个市场有可能是垄断市场。

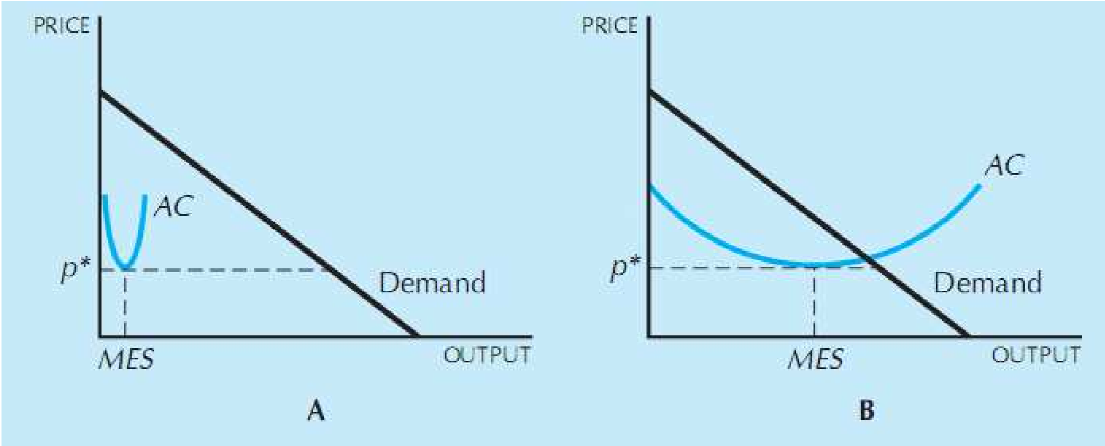


图 24.7：需求与最小有效率规模的比较。若需求相对于最小有效率规模来说很大，结果可能会出现竞争市场，如 (A)图所示；若需求相对较小，则可能出现垄断市场，如图（B）所示。

因此，平均成本曲线的形状（它由企业的生产技术决定），是判断一个市场是竞争还是垄断市场的重要依据。如果最小有效率的规模——使平均成本最小的产量——相对于市场的规模来说很小，我们可以预期这样的市场可能是竞争市场。

注意"**相对**"这个修饰语：最小有效率规模是**相对于**市场规模来说的，只有它相对较小时，我们才能判断该市场可能是竞争市场。我们不大可能改变最小有效率的规模——它是由生产技术决定的。但是经济政策会影响市场的规模。如果某个国家选择自由的对外贸易政策，由于国内的企业面对外国企业的竞争，那么国内企业影响价格的能力就较小。相反，如果一个国家限制自由对外贸易，因此市场规模仅限于国内市场，那么越有可能出现垄断现象。

如果某企业最小有效率的规模相对于市场规模来说较大，而且不大可能增加市场的规模，这样的企业很可能是垄断企业，因此政府需要对该企业进行管制或干预。当然，这样的管制和干预也是有成本的。管制委员会需要经费，而且企业为了达到管制委员的管制要求，也可能要花费大量的资金。从社会的观点来看，管制与否应该取决于垄断造成的净损失是否大于管制的成本。

造成垄断的第二个原因是，行业中的若干家不同企业可能会串谋（collude）——限制产量提高价格从而增加利润。当企业以这样的方式串谋，并企图减少产量和提高价格时，我们说这个行业出现了一个**卡特尔**（Cartel）。

卡特尔是非法的。美国司法部反托拉斯局和联邦贸易委员会促进竞争办事处，负责搜集企业的非竞争行为证据。如果政府证实一伙企业试图限制产量或从事某些其他的反竞争业

务，政府将责令这些企业缴纳巨额罚金。

另一方面，某行业出现主导企业（dominant firm）的原因可能纯粹出于历史偶然。如果一家企业率先进入某市场，那么它可能就具有足够的成本优势去阻止其他企业进入。例如，假设，若企业想进入某行业，则它需要购买大量的机械设备。那么，行业中已有的的已家企业，在某些条件下可能会威胁潜在的进入者：如果它们进入，则它会大幅降低价格。这样，这家企业可能最终变成市场中的主导企业。我们将在第 28 章研究阻止其他企业进入的定价策略。

例子：钻石恒久远

戴比尔斯（De Beers）钻石卡特尔，是由男爵欧内斯特.奥本海默（Ernest Oppenheimer）这位南非矿石经营者于 1930 年建立的。自此以后，它已成长为世界上最成功的一个卡特尔。戴比尔斯占有世界年钻石生产量的 80%以上，而且这种近乎垄断的地位已维系了几十年。这些年来，戴比尔斯已找到了控制钻石市场的几种方法。

首先，它储备了大量的种类不同的钻石。如果某个生产企业试图在卡特尔之外销售，戴比尔斯会在市场上大量抛售相同种类的钻石，从而惩罚了卡特尔的叛逃者。其次，卡特尔中的大企业的生产配额和它们的销量挂钩。因此，当市场不好时，每个企业的生产份额按比例下降，因此自动增加了稀缺性从而提高了价格。

第三，戴比尔斯既从事钻石矿的开采又从事钻石的批发业务。在批发市场上，钻石是按盒销售的，每个盒子都装着各种各样的钻石：买方要么整盒购买要么什么也买不到——它们不能挑选钻石。如果某类钻石的市场不好，戴比尔斯会减少盒子中这类钻石的数量，从而增加了稀缺性。

最后，戴比尔斯每年花费 1.1 亿美元做广告，因此它能影响消费者对钻石最终需求的潮流方向。而且它可以调整广告策略，比如鼓励人们购买某种相对稀缺的钻石[一]。

例子：拍卖市场上的竞买团伙

亚当.斯密曾经说过“同行的商人很少聚会，即使象消遣娱乐这样的活动也不例外，当他们聚会时，他们是在探讨如何对付公众或者如何提高价格。”拍卖市场上的串谋为斯密的结论提供了佐证。1988 年美国司法部控告了 12 名费城古董商，原因是他们“密谋如何对付公众”[二]，这违背了反托拉斯法。

法院控告这些古董商的理由，是他们在古董家具的拍卖涉嫌组织“竞买团伙”（bidding

[一] 钻石市场的简要描述请见"The cartel lives to face another threat," The Economist, January 10, 1987, 58-60. 更详细的分析可以参见 Edward J. Epstein, Cartel (New York: Putnam, 1978).

[二] 这个例子来自： Meg Cox, "At Many Auctions, Illegal Bidding Thrives As a Longtime Practice Among Dealers," Wall Street Journal, February 19, 1988.

rings, or pools）。竞买团伙委派其中一个成员竞买某项拍卖品。如果该成员成功购入该商品，竞买团伙随机会举办一个私人拍卖会，他们将这种拍卖称为“一招制敌”（knockout），在这种拍卖中，团伙成员再竞买该商品。这种做法使得团伙成员能以远低于单独竞价时的价格，获得拍卖品；通常，这种价格只比原来卖方索要的价格高出 50~100 个百分点。

这些交易商对美国司法部的诉讼很惊讶；他们认为他们的做法只是一种普通的商业合伙，而不认为是非法的。他们原本认为这样的做法是他们合作的传统方式；如果某交易商接到这样的合伙邀请，他甚至认为这是一种“声望的标志”。正如一位商人所说的，“我受邀入伙的那一天非常的美好。因为如果你不在团伙中，你就不是一个出色的商人。”这些商人非常天真，他们保留了内部拍卖的详细记录，后来成为司法部起诉他们的证据。

美国司法部认为“如果他们联合起来压低（卖方能得到的）价格，他们的行为就是违法的。”司法部的观点击败了商人的观点：12 个商人中有 11 人认罪，他们认缴了 1000 美元到 5000 美元不等的罚款，并接受了缓刑处理；另外一人不认罪，经过陪审团审判他被判定犯罪，处罚是 30 天的监禁和 30,000 美元的罚款。

例子：操纵存储设备器的价格

DRAM 器是指你计算机中的“动态随机存取存储” 器。由于这样的商品差别很小，因此 DRAM 器的市场是高度竞争的。然而有人宣称，几家 DRAM 生产企业密谋操纵价格，以高于完全竞争市场的价格销售这些产品。苹果、康柏（Compaq）、戴尔、Gateway、惠普和 IBM 这些计算机生产企业明显都受到了这种密谋策略的影响。

美国司法部于 2002 年开始调查这些传言。2004 年 9 月，一家叫做 Infineon 的德国 DRAM 制造企业承认了操纵价格的罪行，同意支付 1.6 亿美元的罚款。这是美国司法部反托拉斯局开出的第三高的罚单。

根据法院的文件，Infineon 公司被起诉的理由是：“和其他竞争者以开会、对话和通信方式，讨论对某些消费者应该索要什么样的价格水平；同意以某种价格水平向某些消费者销售 DRAM；和某些消费者共享 DRAM 的销售信息，目的在于控制和稳固协议价格。”

后来，Infineon 公司的四位管理人员被判坐牢和支付高额罚款。在写本书时，调查还在继续，预计还有其他的诉讼。反托拉斯局相当重视操纵价格的行为，任何公司和个人只要违法，后果将非常严重。

附录

定义收入函数 $r(y)=p(y)y$。则垄断企业的利润最大化问题为

$$\max_{y} \ r(y)-c(y).$$

该问题的一阶条件为

$$r'(y)-c'(y)=0,$$

这个式子意味着在最优产量上，边际收入应该等于边际成本。

对收入函数的表达式进行微分可得 $r'(y)=p(y)+p'(y)y$，将其代入上述一阶条件可得到垄断企业利润最大化问题一阶条件的另外一种表达方式

$$p(y)+p'(y)y=c'(y).$$

垄断企业利润最大化问题的二阶条件为

$$r''(y)-c''(y)\le 0.$$

这个式子意味着

$$c''(y)\ge r''(y).$$

这就是说边际成本的斜率应该大于或等于边际收入曲线的斜率。

总结

1.当某行业中只有一家企业时，我们说这是垄断行业。

2.垄断企业在边际收入等于边际成本之处生产。因此，垄断企业索要的价格是基于边际成本的加成价，加成幅度取决于需求弹性。

3.由于垄断企业的要价大于边际成本，它的产量是低效率的。低效率的大小可以用净损失衡量，即消费者剩余和生产者剩余的损失之和。

4.当某企业若生产有效率的产量会出现亏损，则这样的企业是自然垄断企业。很多共用事业都是自然垄断的，因此需要政府管制。

5.某行业是竞争的还是垄断的，部分取决于生产技术的性质。如果最小有效率的规模相对于市场需求规模来说比较大，则这种市场可能是垄断市场；如果最小有效率的规模相对需求来说比较小，则有可能是竞争市场。

复习题

1.据说海洛因的市场需求曲线几乎缺乏弹性。据说海洛因的供给由黑手党垄断。如果我们假设黑手党追求利润最大化，上述这两种说法是否矛盾？

2.某垄断企业面对的需求曲线为 $D(p)=100-2p$。它的成本函数为 $c(y)=2y$。它的最优产量和价格为多大?

3.某垄断企业面对的需求曲线为 $D(p)=10p^{-3}$。它的成本函数为 $c(y)=2y$。它的最优产量和价格为多大?

4.某垄断企业面对的需求曲线为 $D(p)=100/p$。它的成本函数为 $c(y)=y^2$。它的最优产量为多大?（该题要小心些。）

5.某垄断企业在 $|\varepsilon|=3$ 处生产。政府征收的从量税为 6 元/单位产品。若它面对的需求曲线是线性的，则价格将上升多少？

6.在上题中，若需求曲线的弹性固定不变，则价格将上升多少？

7.若某垄断企业面对的需求曲线的弹性恒为 2，求该企业在边际成本上的加成。

8.政府考虑对上题中企业的边际成本进行补贴。如果政府希望垄断企业的产量是社会最优产量，那么政府应该选择什么样的补贴水平？

9.用数学证明垄断企业的要价总是大于边际成本。

10.对还是错？对垄断企业征收从量税，总会使市场价格上升且上升幅度等于单位税额。

11.监管机构试图强迫垄断企业的定价等于完全竞争时的价格，这种情形下监管机构将面对什么样的问题？

12.什么样的经济条件和技术条件下更容易出现垄断的现象？

复习题答案

1.据说海洛因的市场需求曲线几乎缺乏弹性。据说海洛因的供给由黑手党垄断。如果我们假设黑手党追求利润最大化，上述这两种说法是否矛盾？

【复习内容】收入函数；垄断企业利润最大化一阶条件(边际收入等于边际成本)；需求价格弹性与边际收入的关系

（A）垄断企业利润最大化一阶条件(边际收入等于边际成本)

定义收入函数 $r(y)=p(y)y$，则垄断企业的利润最大化问题为

$$\max_y \ r(y)-c(y).$$

该问题的一阶条件为

$$r'(y)-c'(y)=0,$$

这个式子意味着在最优产量上，边际收入应该等于边际成本。

（B）需求价格弹性与边际收入的关系

收入函数 $r(y)=p(y)y$ 两端对 y 求导可得

$$r'(y)=p(y)+p'(y)y,$$

这就是边际收入 $MR(y)$ 的表达式。可对上式稍微变形

$$MR(y)=r'(y)=p(y)+p'(y)y \Leftrightarrow MR(y)=r'(y)=p(y)[1+\frac{dp(y)}{dy}\frac{y}{p(y)}]$$

$$\Leftrightarrow MR(y)=r'(y)=p(y)[1+\frac{1}{\varepsilon(y)}]$$

将其代入上述一阶条件可得到垄断企业利润最大化问题一阶条件的另外一种表达方式

$$p(y^*)[1+\frac{1}{\varepsilon(y^*)}]=MC(y^*)$$

上式左端是边际收入的另外一种表达。上式仍然是说在最优产量 y^* 处，边际收入等于边际成本。

由于需求价格弹性一般为负，我们可以将上式写为

$$p(y^*)[1-\frac{1}{|\varepsilon(y^*)|}]=MC(y^*)$$

【参考答案】

垄断企业利润最大化问题的一阶条件（必要条件）为

$$p(y^*)[1-\frac{1}{|\varepsilon(y^*)|}]=MC(y^*)$$

这就是说，在最优产量上，必有边际收入等于边际成本。

由上式可知，垄断企业绝对不会在需求曲线**缺乏弹性**的那一段进行生产。因为若 $|\varepsilon|<1$，则 $1/|\varepsilon|>1$，此时边际收入为负，因此它不可能等于边际成本。

因此，如果我们假设黑手党（垄断者）追求毒品经营的利润最大化，那它就绝对不会在需求缺乏弹性之处生产。因此题目中的两种说法是矛盾的。

曹乾（东南大学 caoqianseu@163.com）

2.某垄断企业面对的需求曲线为 $D(p)=100-2p$ 。它的成本函数为 $c(y)=2y$ 。它的最优产量和价格为多大?

【复习内容】垄断企业最优产量和价格决策

【参考答案】

由需求曲线可以求出反需求曲线：$D(p)=100-2p \Rightarrow p(y)=50-y/2$。

因此，由总收入函数表达式 $r(y)=p(y)y$ 可知，$r(y)=(50-y/2)y=50y-y^2/2$.

由总收入函数可求出边际收入（函数）：总收入函数对 y 求导即可。

$$r(y)=50y-y^2/2 \Rightarrow MR(y)=50-y$$

由成本函数可以求出边际成本（函数）：成本函数 $c(y)=2y$ 对 y 求导即可。

$$c(y)=2y \Rightarrow MC(y)=2$$

最后，因为垄断企业最优产量的一阶条件，即在该产量处，边际收入等于边际成本，因此令上述两式相等，即

MR(y)=MC(y) ，这就是 $50-y=2$ ，由此知 $y=48$ 。把它代入反需求函数 $p(y)=50-y/2$，可知 $p(48)=26$。

3.某垄断企业面对的需求曲线为 $D(p)=10p^{-3}$ 。它的成本函数为 $c(y)=2y$ 。它的最优产量和价格为多大?

【复习内容】垄断企业最优产量和价格决策；需求曲线的弹性固定不变的情形

【参考答案】

解法一：解题步骤类似上题

求反需求曲线：$D(p)=10p^{-3} \Rightarrow p(y)=\sqrt[3]{\frac{10}{y}}$

求收入函数：$r(y)=p(y)y=\sqrt[3]{\frac{10}{y}}y=\sqrt[3]{10y^2}$

求边际收入函数：$MR(y)=r'(y)=\frac{2}{3}\sqrt[3]{10}y^{-1/3}$

求边际成本函数：$MC(y)=2$

令边际收入等于边际成本，即 $\frac{2}{3}\sqrt[3]{10}y^{-1/3}=2$，从而可以解得 $y=\frac{10}{27}$。

将 $y=\dfrac{10}{27}$ 代入反需求函数可得 $p(\dfrac{10}{27})=3$。

解法二：直接利用公式 $p(y)[1-\dfrac{1}{|\varepsilon(y)|}]=MC(y)$

如果注意到该垄断企业的需求弹性是固定不变的，则利用上式求解会大幅简化计算。

由需求价格弹性计算公式可知，该企业的需求曲线的弹性恒为 -3，即 $\varepsilon(y)\equiv -3$。将其代入上式可得 $p=3$。

将 $p=3$ 代入需求曲线为 $D(p)=10p^{-3}$ 可得 $y=\dfrac{10}{27}$。

4.某垄断企业面对的需求曲线为 $D(p)=100/p$。它的成本函数为 $c(y)=y^2$。它的最优产量为多大?（该题要小心些。）

【复习内容】垄断企业最优产量和价格决策；需求曲线的弹性固定不变的情形

【参考答案】

垄断企业利润最大化的一阶条件为：

$$p(y)[1-\frac{1}{|\varepsilon(y)|}]=MC(y)$$

由题意可以注意到该企业需求弹性恒为 -1，即 $\varepsilon(y)\equiv -1$。因此，这种情形下上式左端等于 0，而右端恒大于零，这就是说边际收入恒小于边际成本。因此，该企业的最优产量为零。

5.某垄断企业在 $|\varepsilon|=3$ 处生产。政府征收的从量税为 6 元/单位产品。若它面对的需求曲线是线性的，则价格将上升多少？

【参考答案】

解法一：代数方法

令需求曲线为 $p(y)=a-by$，令边际成本为 c，则征收后的边际成本为 $c+t$。

则边际收入 $MR(y)=a-2by$，令其和征税后的边际成本相等，即 $a-2by=c+t$ 可得：

$$y=\frac{a-c-t}{2b}$$

因此，$\dfrac{dy}{dt}=-\dfrac{1}{2b}$，它表示征税对产量的影响。

将 $y=\dfrac{a-c-t}{2b}$ 代入需求函数 $p(y)=a-by$ 可得

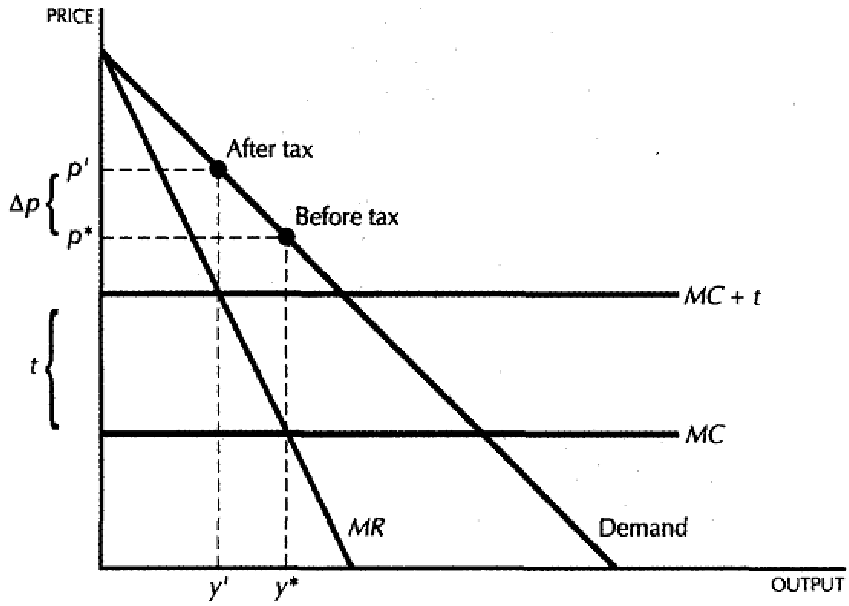
$$p(y)=\frac{a+c}{2}+\frac{t}{2}.$$

因此，$\frac{dp}{dt}=\frac{1}{2}$，它表示征税对价格的影响。

因此，$\frac{\Delta p}{\Delta t}=\frac{1}{2}\Rightarrow\Delta p=\frac{1}{2}t=\frac{1}{2}\times 6=3$（元）。

解法二：几何图形法

若需求曲线为线性，如下图所示。当边际成本曲线 MC 向上移动到 $MC+t$（移动幅度等于单位税额 t），边际收入和边际成本曲线的交点则向左移动。由于需求曲线的斜率是边际成本曲线斜率的一半，价格上升幅度等于单位税额的一半即 $t/2$。



根据上述结论可知，政府征收的从量税为 6 元/单位产品时，价格上升了 3 元。

6.在上题中，若需求曲线的弹性固定不变，则价格将上升多少？

【复习内容】基于边际成本的加成定价法；征税对垄断企业的影响

垄断厂商利润最大化的一阶条件（边际收益等于边际成本）为

$$p(y)[1-\frac{1}{|\varepsilon(y)|}]=MC(y)$$

将其变形可得

$$p(y)=\frac{MC(y)}{1-1/|\varepsilon(y)|}$$

这个式子表明市场价格是边际成本的一个加成价（a markup over marginal cost）。

加成幅度（加价额）$= p(y)-MC(y)=\dfrac{MC(y)}{1-1/|\varepsilon(y)|}-MC(y)$

【参考答案】

$$p(y)=\frac{MC(y)}{1-1/|\varepsilon(y)|}\Rightarrow \Delta p(y)=\frac{\Delta MC(y)}{1-1/|\varepsilon(y)|}\Rightarrow \Delta p(y)=\frac{6}{1-1/3}=9\text{（元）}$$

上式中边际成本变动是由政府征税引起的，此时边际成本变动额等于单位税额即 6 元。

7.若某垄断企业面对的需求曲线的弹性恒为 2，求该企业在边际成本上的加成。

【复习内容】基于边际成本的加成定价法；征税对垄断企业的影响

【参考答案】

$$p(y^*)=\frac{MC(y^*)}{1-1/|\varepsilon(y^*)|}\Rightarrow p(y^*)=\frac{MC(y^*)}{1-1/2}=2MC(y^*)$$

加成幅度$=p(y^*)-MC(y^*)=MC(y^*)$，因此企业在边际成本上的加成幅度等于边际成本。

8.政府考虑对上题中企业的边际成本进行补贴。如果政府希望垄断企业的产量是社会最优产量，那么政府应该选择什么样的补贴水平？

【复习内容】完全竞争情形下的产量；垄断情形下的产量

【参考答案】

社会最优产量通常以完全竞争情形下的产量表示，在完全竞争情形下，利润最大化的一阶条件为：价格等于边际成本，即

$$p(y^*)=MC(y^*)$$

其中 y^* 表示政府未介入时垄断企业的最优产量。

在上题中，企业的需求价格弹性恒为 2，所以它的边际收入

$$MR(y)=p(y)[1-\frac{1}{|\varepsilon(y)|}]\Rightarrow MR(y)=p(y)[1-\frac{1}{2}]=\frac{1}{2}p(y)$$

注意，$MR(y)=p(y)[1-\dfrac{1}{|\varepsilon(y)|}]$这个表达式的普实性：一是它适用于所有类型的企业，不局限于垄断企业；二是它使用于一切产量。也正因为此，在上式中，我们没对 y 施加任何限制（即没有用上标或下标表示它是一个特殊值）。

假设政府给该企业的补贴为 t 元/单位产品，才能使得它愿意生产社会最优产量 y'。那么，该企业的最优选择应满足

$$MR(y')=\frac{1}{2}p(y')=MC(y')-t$$

注意，政府的补贴相当于减少了企业的边际成本。

因此由上式可得政府的补贴：

$$\frac{1}{2}p(y')=MC(y')-t\Rightarrow t=MC(y')-\frac{1}{2}p(y')\Rightarrow t=MC(y')-\frac{1}{2}MC(y')=\frac{1}{2}MC(y')$$

上式最后一步是由于，在完全竞争产量 y' 处有 $p(y')=MC(y')$。因此，政府给予企业的补贴要恰好是其在完全竞争产量处生产成本的一半，简单地说，政府补贴力度为边际成本的50%。

9.用数学证明垄断企业的要价总是大于边际成本。

【复习内容】垄断企业最优产量和价格决策

具体内容请参考第 1 题。

【参考答案】

垄断企业利润最大化问题一阶条件的另外一种表达式为

$$p(y^*)[1+\frac{1}{\varepsilon(y^*)}]=MC(y^*)$$

上式左端是边际收入的另外一种表达。上式仍然是说在最优产量 y^* 处，边际收入等于边际成本。

由于需求价格弹性一般为负，我们可以将上式写为

$$p(y^*)[1-\frac{1}{|\varepsilon(y^*)|}]=MC(y^*)$$

将上式变形可得

$$p(y^*)=\frac{MC(y^*)}{1-1/|\varepsilon(y^*)|}$$

由于垄断企业总是在需求曲线具有弹性处生产，即 $|\varepsilon(y^*)|>1$，因此由上式可知 $p(y^*)>MC(y^*)$，证毕。

10.对还是错？对垄断企业征收从量税，总会使市场价格上升且上升幅度等于单位税额。

【复习内容】垄断企业最优产量和价格决策；征税对垄断企业的影响

首先注意一个问题，垄断企业没有供给曲线。我们以前说过，完全竞争企业的供给曲线是其边际成本曲线的一部分，准确地说是位于平均可变成本曲线最小值以上的那部分边际成本曲线。然而，垄断企业的价格和边际收入不相等，因此价格和边际成本不相等。正因为此，垄断企业对价格变动的反应不是沿着它的边际成本曲线变动。对于垄断企业来说，在价格较高时它未必提供更多的产量，在价格较低时，它未必提供较低的产量。

因此我们以前用过的使用供给曲线和需求曲线分析税收的影响在此处不适用。怎么办？只能使用垄断厂商利润最大化的一阶条件进行分析。

【参考答案】

令 y^* 分别表示征税前和征税后垄断企业的最优产量，则此时必有

$$p(y^*)[1-\frac{1}{|\varepsilon(y^*)|}]=MC(y^*)$$

这个式子是说，在最优产量 y^* 处，边际收入（上式左端）等于边际成本。

令 t 表示从量税的单位税额，相当于在任何产量上边际成本都增加了 t 元，此时必有

$$p(y^*)[1-\frac{1}{|\varepsilon(y^*)|}]<MC(y^*)+t\,.$$

上式是说，边际收入小于边际成本，因此垄断企业会减少产量。由于我们假设市场需求曲线是向下倾斜的，即价格和需求量（产量）成反方向变动关系，因此政府征收从量税后，市场需求曲线不会变动，但由于企业产量减少，即需求量减少，因此价格必然上升。

那么，价格上升幅度是否正好等于单位税额 t 呢？答案是否定的。为了看清楚这一点，令 y^* 和 y' 分别表示征税前和征税后垄断企业的最优产量；令 t 表示从量税的单位税额。

则

$$p(y^*)=\frac{MC(y^*)}{1-1/|\varepsilon(y^*)|}$$

$$p(y')=\frac{MC(y')+t}{1-1/|\varepsilon(y')|}$$

用第二式减第一式可得到征税后价格上升幅度：

$$\Delta p=p(y')-p(y^*)=\frac{MC(y')+t}{1-1/|\varepsilon(y')|}-\frac{MC(y^*)}{1-1/|\varepsilon(y^*)|}$$

由这个式子可以看出，对于垄断企业来说，政府征收从量税后对价格的影响比较复杂，价格上升幅度可能大于、等于或者小于单位税额。

11.监管机构试图强迫垄断企业的定价等于完全竞争时的价格，这种情形下监管机构将面对哪些问题？

【参考答案】垄断企业的管制

会面对一系列的问题：包括确定企业的实际边际成本；确保能为所有消费者服务；以及垄断厂商在新价格下不会亏损，否则它宁可退出该行业。

12.什么样的经济条件和技术条件下更容易出现垄断的现象？

【复习内容】垄断的成因

【参考答案】

第一，如果某企业最小有效率的规模相对于市场规模来说较大，而且不大可能增加市场的规模，这样的企业很可能是垄断企业。这样的情形一般具有以下特征：高额的固定成本和较小的边际成本，这样的情形下容易出现垄断的现象。

第二，行业中的若干家不同企业越容易串谋（collude），即限制产量提高价格从而增加利润。而且相关机构对这样的现象不进行打击的话，也越有可能出现垄断的现象。

第三，由政府政策造成的垄断，比如有些行业政府以法规的形式规定只允许一家企业经营，则容易出现垄断。

第三，纯粹出于历史偶然，一家企业控制了某行业最重要的生产要素，这样也会出现垄断的现象。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

25.垄断行为（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

25 垄断行为

在某竞争市场上，通常有若干家企业销售同样的产品。如果任何一家企业试图提高价格，消费者会抛弃该企业而转向它的竞争者。在垄断市场上，销售某产品的企业只有一家。当垄断企业提高价格时，消费者的数量会减少，但不会减少为零。

在现实世界中，大多数行业位于竞争和垄断这两个极端之间。如果某小镇上的一家加油站提高了汽油的价格，它会失去大多数消费者，因此可以认为该企业是竞争企业。如果该小镇上的某个饭店也提高了价格，但只失去了部分消费者，则该饭店具有部分垄断力量。

如果某个企业具有某种程度的垄断力量，那么它的选择余地比完全竞争企业要大。例如，与完全竞争企业相比，它可以使用比较复杂的定价和营销策略。或者，它可以实施差异化策略，即将自己的产品与竞争对手的产品区分开来，这样它就能进一步提高自己的市场力量（market power）。在本章我们将研究企业如何提高和利用市场力量。

25.1 价格歧视

我们以前说过，垄断企业生产的产量是低效率的，因为在这个产量的基础上再额外多生产一单位产品，消费者对于这一单位产品的支付意愿高于企业的边际成本。但是垄断企业不愿意多生产，因为这样会做会降低全部产量的价格。

但是，如果垄断企业能按不同价格销售不同单位的产品，那么它会愿意多生产。按不同价格销售不同单位产品的做法称为**价格歧视**（price discrimination）。价格歧视通常有三种：

第一级价格歧视（first-degree price discrimination）是指垄断企业按不同价格出售不同产量，而且这些价格可能因人而异。这种价格歧视有时又称为**完全价格歧视**（perfect price discrimination）。

第二级价格歧视（second-degree price discrimination）是指，垄断企业按不同价格出售不同产量，但是购买相同数量的每个人支付的价格是相同的。因此，价格按购买数量制定，而不是因人而异。最常见的情形是大宗购买时可以享受折扣。

第三级价格歧视（Third-degree price discrimination）是指垄断企业的销售价格因人而异，但对于同一个人来说，每单位产品的售价是相同的。这种价格歧视最常见。例如，对老年人打折，对学生打折等等。

我们将逐一分析这些价格歧视，看看它们是如何运行的。

曹乾（东南大学 caoqianseu@163.com）

25.2 第一级价格歧视

在第一级价格歧视或完全价格歧视下，垄断企业将每单位产品都卖给支付意愿最高而且愿意购买该单位产品的消费者。

请看图 25.1。这个图给出了两个消费者对某种商品的个人需求曲线。这两个人的需求曲线都是商品为离散情形下的需求曲线，也就是说消费者购买的数量只能为整数，需求曲线中的每个“台阶”表示消费者对额外一单位产品支付意愿的变化。在图中，我们还画出了垄断企业的（固定不变的）边际成本曲线。

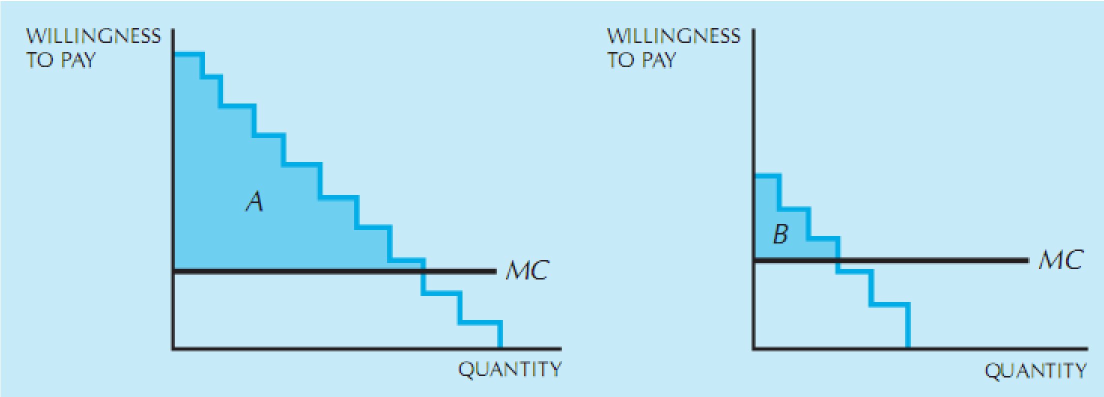


图 25.1：第一级价格歧视。图中画出了两个消费者对某商品的需求曲线以及生产者的固定边际成本曲线。生产者按它能索要的最高价格消费每单位产品，因此实现了利润最大化。

能够实行完全价格歧视的企业，它销售每单位产品的价格都是它能索要的最高价格，也就是每个消费者的保留价格。由于每单位商品都是按照消费者对这单位商品的保留价格销售的，这个市场产生的消费者剩余为零；所有的消费者剩余都被生产者攫取。在图 25.1 中，黑色区域表示垄断生产者获得的**生产者剩余**。在完全竞争市场上，这些区域代表的是**消费者剩余**，但在完全价格歧视的情形下，垄断企业能够占有这部分剩余。

由于生产者能得到市场中的所有剩余，它就会使得该剩余尽可能地大。换句话说，生产者的目标是最大化它的利润（生产者剩余），当然前提是消费者正好还愿意购买它的商品。这意味着垄断企业的产量是帕累托有效率的，因为已无法使得消费者和生产者的状况更好：生产者的利润不可能进一步增加，因为它已做到了利润最大化；在不减少生产者利润的前提下，消费者剩余也不可能进一步增加。

如果我们用平滑的曲线近似图 25.1，我们就得到了图 25.2。由图可知，实行完全价格歧视的垄断者必然在价格等于边际成本之处生产：若价格大于边际成本，这意味着若它再额外多生产一单位产品，有人会以高于边际成本的价格购买这单位产品。因此，垄断者会将这单位产品生产出来，并按照该人的保留价格卖给他，从而增加了利润。

由此可见，在完全价格歧视情形下，生产者剩余和消费者剩余之和，等于完全竞争市场情形下的这两种剩余之和，而且均已达到最大。然而，在完全价格歧视情形下，生产者最终占有了市场产生的所有**剩余**。

上面我们对第一级价格歧视的解释是，垄断企业将每单位商品都按照它能索要的最高价格销售。但是我们还可以这样解释：垄断企业按照某个价格销售数量固定的产品，这个价格具有“要买就买、不买请走人”的性质。在图 25.2 所示的情形下，垄断者会向消费者 1 销售 x_1^0 单位产品，销售的总价格（即价格与销量的乘积）等于消费者 1 需求曲线下方阴影区域 A 的面积。垄断者向消费者 2 销售 x_2^0 单位产品，销售的总价格（即价格与销量的乘积）等于消费者 2 需求曲线下方阴影区域 B 的面积。和图 25.1 一样，最终每个消费者的剩余都为零，整个剩余 A+B 最终都被垄断生产者所占有。

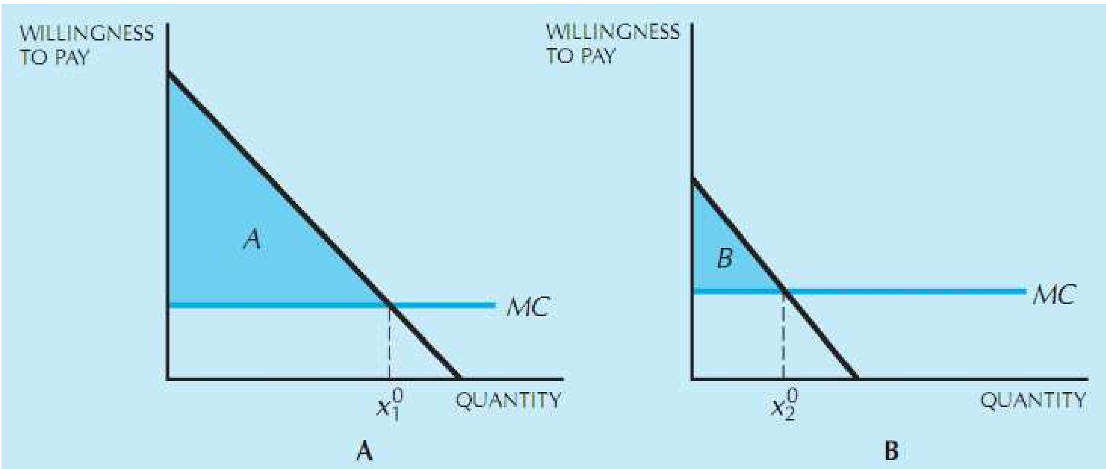


图 25.2：平滑需求曲线情形下的第一级价格歧视。图中画出了两个消费者对某商品的平滑需求曲线，以及垄断生产者的固定不变的边际成本曲线。生产者在价格等于边际成本之处生产，并在此处实现了利润最大化，这一点和完全竞争市场的情形是一样的。

完全价格歧视是一个理想的概念，正如“完全”这个词暗示的。但它在理论上是有趣的，因为它表明，除了完全市场之外，还有一种资源分配机制能够实现帕累托有效率。现实世界中很难找到符合完全价格歧视的例子。比较近似的例子有，小镇上的一个医生，他能按照病人的支付能力索要不同的价格。

例子：现实中的第一级价格歧视

我们在前面说过，第一级价格歧视基本上是一个理论上的概念。在现实世界中很难找到对每个消费者索要不同价格的例子。比较接近的例子是讨价还价的情形，例如汽车销售或者古董销售。但是这些还不是最好的例子。

西南航空公司最近引入了一种称为“Ding”的系统，目的在于实施接近于第一级价格歧视的定价策略[一]。这个系统聪明地利用了 Internet。用户在自己的计算机下载安装 Ding 程序后，航空公司会向该客户定期发送优惠机票价格。在你看到机票价格时会听到“Ding”的响声，该程序因此得名。根据一项分析，由 Ding 系统报出的价格要比同类机票价格低大概 30%。

但是这样的低价能持续下去吗?也许有人会使用这个系统报出较高的机票价格。然而，这似乎不可能，因为航空业竞争激烈。如果 Ding 系统的机票价格爬升，那么你可以容易地回到以前的传统购票方式。

25.3 第二级价格歧视

第二级价格歧视有时又叫作**非线性定价**（nonlinear pricing），这是由于每单位商品的价格不是固定不变的，而是取决于你购买的数量。这种形式的价格歧视通常用于共用事业的定价；例如每单位电力的价格通常取决于购买量。在其他行业，大宗购买商品时有时能享受到折扣优惠。

我们考虑图 25.2 的情形。垄断者愿意向消费者 1 销售 x_1^0 单位产品，销售总价格为“A+成本”；它愿意向消费者 2 销售 x_2^0 单位产品，销售总价格为“B+成本”。为了得到准确的价格，垄断者必须**知道**消费者的需求曲线；也就是说，垄断者必须知道每个人的确切支付意愿。即使垄断者大体知道每种人群的支付意愿，例如，大学生对电影的支付意愿小于高收入的年青小伙，但是这两类人排队买票时，你如何区分他们？

类似地，机票销售中介知道商务人士比一般旅客对机票的支付意愿高，但是它通常难以判断某个人是商务人士还是普通旅客。如果不穿灰色法兰绒套装（flannel suit）而是穿着百慕大短裤（Bermuda shorts）能省 500 美元旅行费，那么企业的职业装标准将立即改变！

第一级价格歧视存在的问题（见图 25.2）是：具有较高支付意愿的消费者 1，可以**伪装**成支付意愿较低的消费者 2。卖方没办法区分他们。

一种解决之道是向市场提供两种不同的价格-数量服务包(package)。一种针对高需求的消费者，另外一种针对低需求的消费者。垄断者通常能构建出不同的价格-数量服务包，诱使消费者恰好选择本来就为他们设计的服务包；用经济学的行话来说，垄断者构建出来的服务包使得消费者有**自我选择**(self select)的激励。

为了看清这种方法是如何运行的，我们用图 25.3 进行分析。图 25.3 中的需求曲线类型和图 25.2 相同，但现在我们把两个消费者的需求曲线放在同一张图中。为简单起见，令边际成本为零。

[一] See Christopher Elliott, "Your Very Own Personal Air Fare," New York Times, August 9, 2005.

和以前一样，垄断者愿意以价格 A 供给 x_1^0 单位产品，愿意以价格 A+B+C 供给 x_2^0 单位产品。这样垄断者就能占有所有的剩余，利润也尽可能地大。不幸地是，这些价格-数量组合与自我选择不相容。高需求的消费者会发现它的最优选择是选择数量 x_1^0，支付的价格为 A；这样他就能得到部分消费者剩余，即图中的区域 B 面积；而如果他选择的数量为 x_2^0，他的消费者剩余将为零。显然，选择 x_1^0 时的消费者状况好于选择 x_2^0。

垄断者可以调整销售策略，以价格 A 供给 x_1^0 单位产品，但供给 x_2^0 单位产品的价格变为 A+C。在这种情形下，高需求的消费者会发现他的最优选择是选择 x_2^0，这样他的总消费者剩余为 A+B+C。他向垄断者支付 A+C，从而给消费者 2 留下的净消费者剩余为区域 B，这和消费者 1 自己购买 x_1^0 单位的净消费者是一样的。这样的做法会让垄断者得到比只提供一种价格-数量包更多的利润。

但上面的策略仍有改进的余地。垄断者还可以优化销售策略增加利润。假设垄断者对低需求消费者不再提供价格为 A、数量为 x_1^0 的服务包，而是以比 A 稍低的价格提供比 x_1^0 稍小的数量。这样的做法会让垄断者从消费者 1 身上得到的利润减少了，利润减少额为图 25.3B 中的黑色小三角形区域。但是注意，此时由于消费者 1 的服务包对消费者 2 的吸引力下降了，垄断者对高需求的消费者 2 的 x_2^0 数量可以索要更高的价格！总结一下垄断者此时的销售策略：稍微减少 x_1^0 的数量，垄断者的利润减少了黑色小三角形的面积，但是却可从消费者 2 身上赚取更多的利润（利润增加额以黑色小三角形加上灰色梯形区域）。

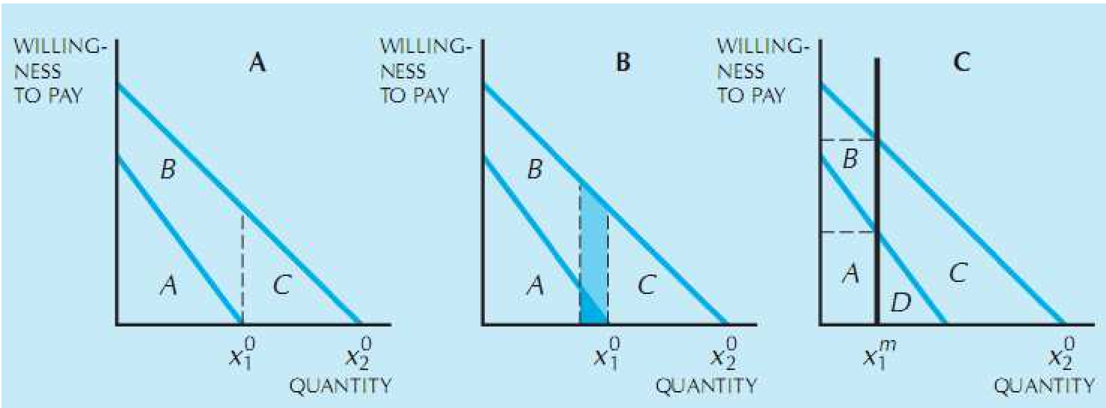


图 25.3：第二级价格歧视。上图中的三个图都画出了两个消费者的需求曲线；假设生产者的边际成本为 0。A 图表示自我选择的问题。B 图表示如果垄断者降低供给消费者 1 的数量，将会产生什么样的结果；C 图表示利润最大化的解。

由上面的分析思路可知，垄断者会一直减少提供给消费者 1 的数量，直至他从消费者 1 身上损失的利润恰好等于从消费者 2 身上多赚的利润。如图 25.3C 所示，在这一点上，由减少 x_1^0 而引起的边际收入等于减少 x_1^0 引起的边际成本。消费者 1 选择 x_1^m，支付的价格为 A；消费者 2 选择 x_2^0，支付的价格为 A+C+D。最终，消费者 1 的消费者剩余为零，消费者 2 的消费者剩余为区域 B 的面积，这恰好等于他选择 x_1^m 时能得到的消费者剩余。

在实践中，为了激励消费者自我选择，垄断者通常不是调整产品的数量而是调整产品的**质量**。图 25.3 中的产品数量可以看成质量，分析方法和结果是一样的。一般来说，垄断者希望降低它在低端市场提供的产品质量，目的在于保护高端市场的销量。如果没有高端市场，低端市场的消费者将得到质量较高的产品，但是最终他们的消费者剩余仍为零。如果没有低端消费者，高端消费者的消费者剩余也为零，因此存在低端消费者对于高端消费者来说是有利的。这是因为，垄断者为了阻止高端消费者选择原本为低端消费者提供的产品，它必须降低对高端消费者的要价。

例子：飞机票价的价格歧视

航空业在价格歧视方面做得非常成功（尽管航空业更喜欢使用“收入管理”这个词）。上面介绍的模型可以用于分析航空业面对的问题：航空业的客户基本上可以分类两类：商务旅行者和个人旅行者，这两类客户的支付意愿大不相同。尽管在美国市场上航空公司数量不少，但是如果具体到某两个城市之间往返业务来说，则一般只有一两家航空公司运营。因此，这样的航空公司在制定价格方面有很大的自由度。

我们已经知道，如果垄断者面对两类顾客时，它的最优定价策略是，对高支付意愿的顾客索要较高的价格；对低支付意愿的顾客索要较低的价格——但是，提供给这类顾客的产品质量也较低。降低低端顾客的产品质量的目的在于阻止高端顾客购买价格较低的产品。

航空公司的销售策略是对商务旅行者提供“无限制的票价”（unrestricted fare），对个人旅行者提供“有限制的票价”（restricted fare）。有限制的票价通常要求提前购买、深夜航班以及其他的限制等等。这些限制的目的在于将这两类客户区分开来，因为个人旅行者对价格更敏感。通过提供“低质量”的产品——有限制的票价——航空公司可以对商务旅行者索要较高的价格。

这样的销售策略即使站在社会的角度上，也是有价值的。这是因为，如果生产者无法进行价格歧视，那么它的最优销售策略就是**只**向高需求的消费者销售。

航空公司还可以采取另外一种歧视策略，即将机舱分为头等舱和经济舱。头等舱的旅客支付的价格较高，但是他们能享受更好的服务：更大的空间、更好的食物和更体贴的服务。经济舱的旅客享受的上述服务都打了“折扣”。几百年来，这种质量歧视已成为运输业的特征。来看看 Emile Dupuit (19 世纪的一位法国经济学家)对铁路乘务定价的评价：

铁路公司在火车上安装敞篷车厢并装备木质座位的原因，不是由于公司舍不得花几千法郎，去为三级车厢加上一个顶棚或者为木质座位添置椅垫...铁路公司真正的目的是阻止能买得起二级车厢票价的旅客购买三级车厢的车票。这种做法伤害了穷人的自尊心，但是公司的出发点并不是想伤害穷人，而在于威胁富人...出于同样的原因，这些对三级车厢旅客非常残忍、对二级车厢旅客非常吝啬的铁路公司，对于一级车厢的旅客却非常慷慨。他们拒绝

了穷人的基本需求，却为富人提供了那么多的服务。[一]

现在如果你再坐经济舱旅行时，你心里会舒服一些了吧，要知道 19 世纪法国的火车更不舒服。

例子：处方药的价格

抗抑郁药 Zoloft 一个月的剂量，在奥地利的售价为 29.74 美元，在卢森堡为 32.91 美元，在墨西哥为 40.97 美元，而在美国的售价为 64.67 美元。价格为何有那么大的差异？药品生产者，和其他企业一样，索要市场能承认的价格水平。相对贫穷的国家支付能力低于较富裕的国家，因此在穷国的药价相对要低。

但这并不是故事的全部，还有其他因素会影响药价。各国对药品的讨价还价能力也大不相同。加拿大，由于实行的是国家健康方案，它的药价比美国低，因为美国没有类似的全民健康方案。

有人曾经建议过，应该强迫药品企业在世界范围内索要同一价格。暂且不管怎么执行这种建议的问题，我们此处只关注这种建议的结果。世界范围内药品价格会因此更低还是更高？

答案取决于市场的相对规模。治疗疟疾的药品的需求主要在比较贫穷的国家。如果实施同一价，则药品公司愿意将这类药品按较低的价格销售。但是对于某些疾病（尤其是富裕国家居民易患的疾病）的治疗药品，在实施同一价后，药品公司会索要较高的价格，从而使贫穷地区更难以承受这样药品的价格。

一般来说，如果企业不实行歧视定价而是采取同一价，则它会提高某些产品的价格、降低某些产品的价格，从而使得某些消费者的状况变好、另外一些消费者的状况变坏。在某些情形下，如果强迫企业接受同一价，它宁愿退出该地区的市场从而使该地区的供给量为零。

25.4 第三级价格歧视

我们在前面说过，第三级价格歧视是垄断者对不同的消费者群体索要不同的价格，但是对于同一种消费群体中的每个人，它索要的价格是相同的。第三级价格歧视是最常见的价格歧视。类似的例子是电影票价对学生实行折扣，或者某些药店对老年人实行折扣价。垄断者如何在每个市场上决定最优的价格？

假设垄断者能够有效区分两类消费者，并对每类消费者群体索要不同的价格。再假设每个市场上的消费者都不能将商品转卖。令 $p_1(y_1)$ 和 $p_2(y_2)$ 分别表示消费者群体 1 和 2 的反

[一] 参见 R. B. Ekelund 的译文： "Price Discrimination and Product Differentiation in Economic Theory: An Early Analysis," Quarterly Journal of Economics, 84 (1970), 268-78.

需求曲线，令$c(y_1+y_2)$表示垄断者的成本函数。则垄断者的利润最大化问题为

$$\max_{y_1,y_2}\ p_1(y_1)y_1+p_2(y_2)y_2-c(y_1+y_2).$$

最优解必然为

$$MR_1(y_1)=MC(y_1+y_2)$$

$$MR_2(y_2)=MC(y_1+y_2).$$

也即是说，额外多生产一单位产品的边际成本必须等于每个市场的边际收入。如果市场 1 的边际收入大于边际成本，则应该扩大市场 1 的销量；市场 2 的情形类似推理。由于每个市场的边际成本是相等的，这当然表示每个市场的边际收入也必须相等。因此，一件商品无论在市场 1 还是市场 2 销售，它所带来的收入一定是相同的。

我们可以使用边际收入和需求价格弹性之间的关系，以及边际收入等于边际成本的关系，将利润最大化的一阶条件写为

$$p_1(y_1)[1-\frac{1}{|\varepsilon_1(y_1)|}]=MC(y_1+y_2)$$

$$p_2(y_2)[1-\frac{1}{|\varepsilon_2(y_2)|}]=MC(y_1+y_2)$$

其中$\varepsilon_1(y_1)$和$\varepsilon_2(y_2)$分别表示两个市场在最优销量处的需求价格弹性。

现在注意下列的关系。若$p_1>p_2$，则我们必有

$$1-\frac{1}{|\varepsilon_1(y_1)|}<1-\frac{1}{|\varepsilon_2(y_2)|}$$

这意味着

$$\frac{1}{|\varepsilon_1(y_1)|}>\frac{1}{|\varepsilon_2(y_2)|}$$

或

$$|\varepsilon_1(y_1)|<|\varepsilon_2(y_2)|.$$

因此，价格较高的那个市场的需求价格弹性必然较小。稍一思索你就会明白其中的道理。弹性需求是指对价格敏感的需求。因此，实行价格歧视的企业，对价格敏感的消费者群体制定的价格较低，而对价格相对不敏感的群体制定相对较高的价格。这样，企业就实现了总利润的最大化。

我们在前面说过对老年人打折和对学生打折，是第三级价格歧视较好的例子。现在我们看看企业为什么对这样的消费者群体打折。相对于一般消费者来说，这两类群体很有可能对价格比较敏感，因此在某个价格范围具有较大的弹性。所以，追求利润最大化的企业会对

这类消费者打折。

例子：线性需求曲线

下面我们分析垄断企业在两个市场上的需求曲线都是线性的情形，市场 1 和 2 的需求曲线分别为 $x_1 = a - bp_1$ 和 $x_2 = c - dp_2$。为简单起见，假设边际成本为 0。如果该企业能实行价格歧视，则它在每个市场上都是在边际收入等于边际成本（为零）之处生产，即在每条需求曲线的中点之处生产，此时产量分别为 $x_1^* = a/2$ 和 $x_2^* = c/2$；价格分别为 $p_1^* = a/2b$ 和 $p_2^* = c/2d$。

假设可以迫使该企业在两个市场上的售价相同。则它面对的需求曲线为

$$x = (a+c) - (b+d)p$$

它会在这条需求曲线的中点处进行生产，此时产量和价格分别为 $x^* = (a+c)/2$ 和 $p^* = (a+c)/2(b+d)$。注意，无论能否实行价格歧视，该企业的产量是相同的（这种情形只是线性需求函数的特征，在其他情形下一般不成立。）

然而，上述结论有一个重要的例外。我们的假设是当消费者选择最优的单一价格时，它在两个市场上的销量均为正。但是完全有可能出现下面的情形，即在利润最大化的价格水平上，垄断者只对其中一个市场进行销售，如图 25.4 所示。

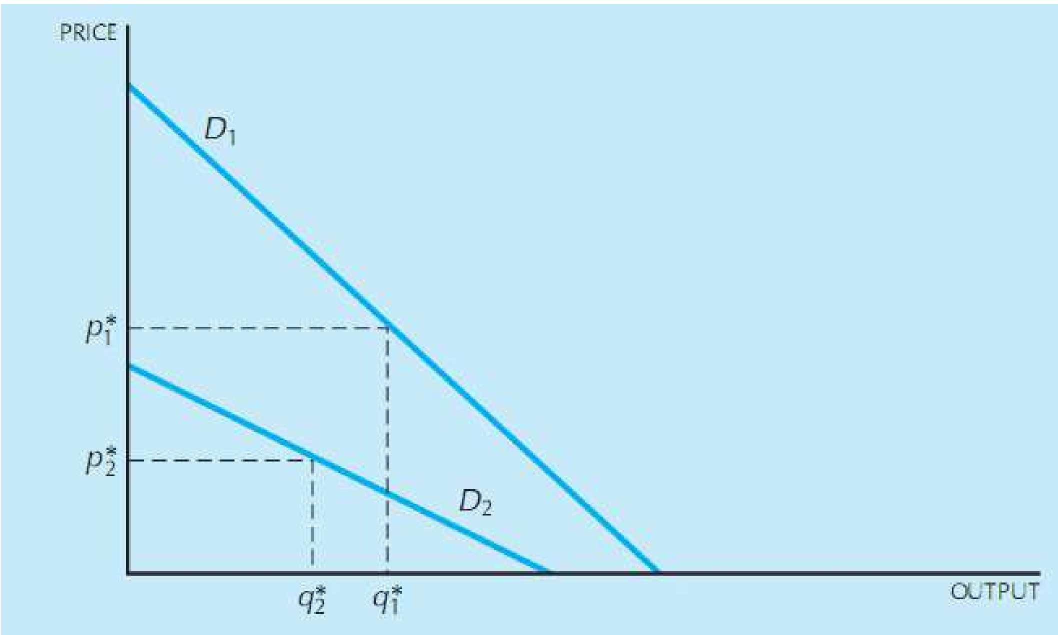


图 25.4：线性需求情形下的价格歧视。若垄断者对两个市场只能索要同一价格，则它会索要价格 p_1^*，而且只向市场 1 销售。但是如果允许价格歧视，它也会向市场 2 销售，市场 2 的销售价格为 p_2^*。

由于我们假设两条需求曲线都是线性的，而且还假设边际成本为零，因此垄断者会在需求价格弹性等于-1之处生产，而这正是需求曲线的中点。因此，价格 p_1^* 是利润最大化价格——降低该价格会减少从市场 1 得到的收入。如果市场 2 的需求很小，垄断者可能不想为了在市场 2 上扩大销量而继续降低价格：最终它只会向市场 1 这个较大的市场销售。

在这种情形下，允许价格歧视无疑会增加总产量，因为垄断者此时会发现若向两个市场都销售，它可以对每个市场索要不同的价格。

例子：计算最优价格歧视

假设某垄断者面临的两个市场需求曲线分别为

$$D_1(p_1) = 100 - p_1$$

$$D_2(p_2) = 100 - 2p_2$$

假设垄断者的边际成本固定为 20 元/单位产品。如果它能实施价格歧视，那么为了实现利润最大化，它在每个市场上应该索要多大的价格？如果它不能实施价格歧视而只能实行同一价格，它索要的价格又为多大？

为了求解价格歧视下的定价问题，我们首先计算反需求曲线：

$$p_1(y_1) = 100 - y_1$$

$$p_2(y_2) = 50 - y_2/2.$$

令每个市场的边际收入和边际成本相等，这样我们就得到了两个方程：

$$100 - 2y_1 = 20$$

$$50 - y_2 = 20.$$

由上面的方程可以解得 $y_1^* = 40$ 和 $y_2^* = 30$。代入反需求函数可得 $p_1^* = 60$ 和 $p_2^* = 35$.

现在，如果该垄断者在两个市场上只能索要同一价格，我们首先计算总需求：

$$D(p) = D_1(p_1) + D_2(p_2) = 200 - 3p.$$

它的反需求曲线为

$$p(y) = \frac{200}{3} - \frac{y}{3}.$$

由边际收入等于边际成本可得

$$\frac{200}{3} - \frac{y}{3} = 20,$$

从而可以解得 $y^* = 70$ 和 $p^* = 43\frac{1}{3}$.

为了与上一小节的分析相一致，有必要检验一下这个价格是否在每个市场上的需求都是非负的。可以验证，结果是成立的。

例子：学术期刊的价格歧视

学术期刊的作用是提供书面学术交流的场所。这些杂志的销售方式是由图书馆和学者个人订阅。通常图书馆和个人订阅价格是不同的。一般来说，我们可以预期图书馆的需求比个人需求更缺乏弹性。经济学家的研究结论证实了这个观点。他们发现图书馆的订阅价格要远高于个人订阅价格。具体来说，图书馆的订阅价格是个人订阅价格的 3 到 4 倍。

最近，一些出版企业开始根据地域的不同进行价格歧视。1984 年，当美国美元的价格空前高于英国英镑的价格时，很多英国出版企业开始对美国和欧洲的订阅者索要不同的价格。可以预期，美国的需求更缺乏弹性一些。由于按照当时的汇率计算，英国杂志的美元售价非常低，因此 10%的美元价格上涨引起的需求下降百分比，小于同幅度提高英镑价格时引起的需求下降幅度。因此，追求利润最大化的出版企业，当然会提高需求相对缺乏弹性的美国的订阅价格。根据 1984 年的一项研究，北美图书馆（North American Libraries）的订阅价格比英国图书馆的订阅价格高出了 67%，比世界其他国家高出了 34%[一]。

通过分析价格上升模式，可以发现价格歧视的更多证据。根据密歇根大学图书馆的一项研究，“…出版企业认真制定了新的定价策略。在图书馆使用率和定价差额之间…似乎存在着直接的相关关系。使用率越大，价格越高。”[二]

到了 1986 年，汇率开始有利于英镑，这样以美元标价的期刊的价格大幅度升高。随着价格的上涨，有些人开始强烈抵制订阅这些期刊。密歇根大学的上述研究在结尾处指出：“我们可以预期，对某产品具有垄断力量的销售者，将会根据需求决定索要的价格。大学作为消费者必须决定的是，是否继续为这样的同质产品支付比英国大学高出 114%的价格。”

25.5 捆绑销售

企业通常选择**捆绑销售**（sell goods in bundles）的方式销售商品：相关的产品组成产品包，一起销售。最明显的例子是软件包，有时称为“软件套装”。这样的软件包可能包含几种不同的软件工具——文字处理软件、电子表格软件和报告制作软件——它们被打包销售。另外一个例子是杂志：一本杂志包含若干篇文章，这些文章原则上是可以单独出售的。类似地，杂志的销售方式主要是订阅，这正是将各期杂志捆绑在一起销售。

[一] Hamaker, C. and Astle, D., "Recent Pricing Patterns in British Journal Publishing," Library Acquisitions: Practice and Theory, 8, 4 (Spring 1984), 225-32.
[二] The study was conducted by Robert Houbeck for the University of Michigan Library, and published in Vol. 2, No. 1 of the University Library Update, April 1986.

捆绑销售的原因可能是为了节省成本：将几篇文章装订在一起销售，通常要比单独出售便宜。或者被捆绑销售的商品之间存在着互补性：软件套装中的软件协调性很好，而如果你买现成的单个软件安装，可能和其他软件存在着不兼容的现象。

企业选择捆绑销售的方式也可能涉及消费者行为的因素。我们以一个简单的例子说明。假设消费者群体有两类，软件也有两种：文字处理软件和电子表格软件。消费者群体 A 对文字处理软件和电子表格软件的支付意愿分别为 120 元和 100 元。消费者群体 B 对这两种软件的支付意愿正好和 A 相反。我们用表 25.1 将这些信息总结如下：

Type of consumer	Word processor	Spreadsheet
Type A consumers	120	100
Type B consumers	100	120

表 25.1：不同消费者群体对软件的支付意愿

假设你来销售这些软件。为简单起见，假设边际成本可以忽略不计，因此你只要关注收入最大化即可。而且，我们做个保守的假设，即人们对每个软件包的支付意愿，恰好等于他们这两个软件各自支付意愿的加和。

现在考虑下面两种不同营销策略的利润。首先，假设你将每个软件单独销售。收入最大化的策略是将每个软件的价格定为 100 元。如果你这么做，你可以卖掉两份文字处理软件和两份电子表格软件，总收入为 400 元。

但是如果你将这两种软件捆绑销售，结果会如何？在这种情形下，每个软件包将销售 220 元，这样你得到的收入为 440 元。捆绑销售的策略显然更具有吸引力！

其中的道理是什么？回忆如果你向不同的人销售某商品时，价格取决于他们中**最低的**那个支付意愿。个人的支付意愿越分散，销售既定数量时的产品定价就越低。而在捆绑销售的情形下，你降低了消费者支付意愿的分散性，也就是说你可以对捆绑在一起的商品索要更高的价格。

例子：软件套装

微软、莲花(Lotus)和其他软件制造企业，都已经对他们的大部分软件进行捆绑销售。例如，1993 年微软将电子表格软件、文字处理软件、报告制作软件和数据库打包成“微软办公系统”包，建议零售价为 750 美元。（打折后的实际零售价大约为 450 美元。）如果这些软件单独购买，则总共要花费 1565 美元！莲花的“Smart 套装”价格基本也是 750 美元，但是如果单独购买，则需要支出 1730 美元。

作者 Steve Lohr 在纽约时报 1993 年 10 月 15 日发表的一篇文章支出，微软公司的应用软件大约有 50%都是捆绑销售的，这种方式的年销售收入超过了 10 亿美元。

这些软件套装可用捆绑销售模型解释。人们对软件的偏好通常是异质化的。有些人每天都要使用文字处理系统但是只会偶尔使用电子表格软件。如果你希望大量销售电子表格软件，那么你制定的价格应该能吸引这样偶尔使用的客户。类似地情形也适用于文字处理软件：**边际**客户（marginal user）的支付意愿决定了产品的市场价格。将两种产品捆绑起来销售，减少了支付意愿的分散性，因此可以使总利润增加。

当然上述因素不是企业将产品进行捆绑销售的全部原因，还有其他原因也在起作用。软件套装中的软件要能协调运行；在这个意义上，套装中的软件是互补的。而且，单个软件的成功与否取决于有多少消费者在使用它，将软件捆绑销售有助于巩固市场份额。我们将在后面的章节分析**网络外部性**（network externalities）的现象。

25.6 两部收费制

下面我们分析游乐园的拥有人对游乐服务的价格制定问题。他们可以对门票制定一个价格，对乘坐摩天轮制定另外一个价格。如果他们的目的是利润最大化，那么他们应该如何设定这两个价格？注意，门票需求和乘坐摩天轮的需求是相关的：门票价格将取决于人们对乘坐摩天轮的支付意愿。这种两部分定价的方案称为**两部收费制**（two-part tariffs）。美国经济学家沃尔特.奥在他的一篇经典文献《迪斯尼乐园的两难选择：米老鼠垄断的两部收费制》提出了这种定价策略[一]。

两部收费制的例子随处可见：宝丽来（Polaroid）的相机和胶卷分别定价。人们是否购买相机的决策大概取决于胶卷的价格。剃须刀生产企业对剃须器和刀片分别定价——他们对刀片的定价会影响剃须器的需求，反过来说也可以。

我们使用沃尔特.奥最初使用的例子，即迪斯尼两难选择的例子进行分析[二]。为简单起见，作出以下假设：首先假设迪斯尼乐园只有一种摩天轮；其次假设人们去迪斯尼乐园就是为了乘坐摩天轮的；最后我们假设每个人对乘坐摩天轮的偏好是一样的[三]。

在图 25.5 中，我们画出了需求曲线和摩天轮服务的（固定不变的）边际成本曲线。和以前一样，需求曲线是向下倾斜的——若迪斯尼对乘坐摩天轮的定价很高，人们对此的需求就会减少。假设他们的定价为 p^*，如图 25.5 所示，需求因此减少为 x^*。如果乘坐摩天轮的价格为 p^*，此时应将门票价格定为多少？

[一]See the classic article by Walter Oi, “A Disneyland Dilemma: Two-Part Tariffs for a Mickey Mouse Monopoly,” Quarterly Journal of Economics, 85 (1971), 77-96. （以下为译者注）：沃尔特.奥（Walter Oi）出生于 1929 年，他是美国罗彻斯特大学杰出的经济学教授、美国文理科学院院士，自入大学后，他逐渐失明无法读书，1956 年他完全失明。但他一直在与命运做抗争，成就了非凡的事业。这篇文献就是他在完全失明后写作的。

[二] 这个两难选择是说，迪斯尼是应该只收高额门票费而不对乘坐摩天轮收费，还是不收门票费只对乘坐摩天轮收取较高的费用？译者注。

[三] 最后一个假设比较微妙，在这个假设之下，所有消费者的需求需求是一样的，或者说给定某价格水平，他们的需求价格弹性是相同的。译者注。

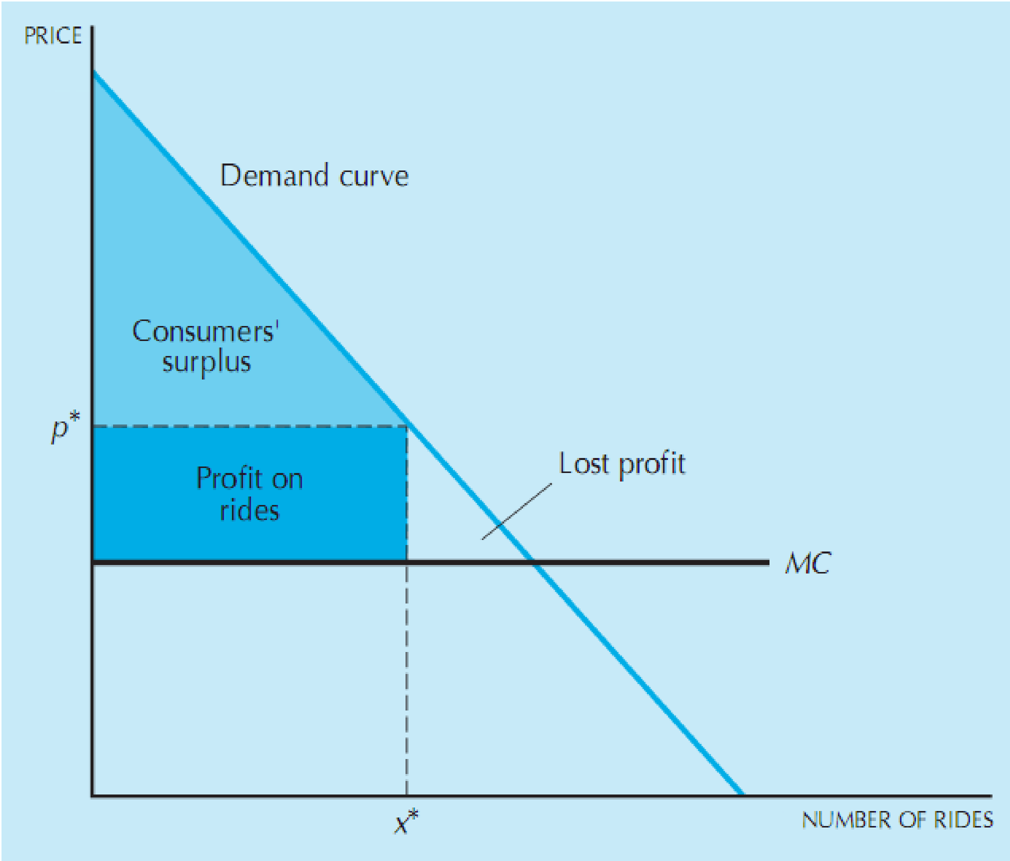


图 25.5：迪斯尼的两难选择。若迪斯尼将乘坐摩天轮（riders）的价格定为 p^*，则需求量为 x^*。他们能对门票索要的价格可用图中的消费者剩余衡量。由图看出，当迪斯尼将乘坐摩天轮价格设定为等于边际成本时，它的利润就实现了最大化。因为此时，它的利润为需求曲线以下边际成本 MC 曲线以上的那个大三角形区域，显然此时利润最大。

消费者对 x^* 单位的摩天轮服务的支付意愿可用消费者剩余衡量。因此，迪斯尼对门票能索要的最高价格为图 25.5 中标有“消费者剩余”的那个黑色三角形区域。垄断者迪斯尼的总利润等于这个区域加上摩天轮项目的利润 $(p^* - MC)x^*$。

不难看出当价格等于边际成本时总利润最大，这个价格使得消费者剩余和生产者剩余最大。由于迪斯尼能够占有消费者的剩余，因此它的最优定价策略是：将乘坐摩天轮的价格设定为等于摩天轮运营的边际成本；将门票价格设定为等于最终的消费者剩余。由图 25.5 可知，将乘坐摩天轮的价格设定为等于边际成本后的消费者剩余，等于需求曲线以下边际成本 MC 曲线以上的那个大三角形区域。如果迪斯尼将这部分消费者剩余用门票价格的形式全部占有，显然利润最大。

事实上，这正是迪斯尼乐园和其他游乐园采取的定价策略。游客仅园要缴纳门票费，但是乐园内的游乐项目却是完全免费的。这表明，摩天轮这类游乐项目的边际成本可能小于

对它们单独收费时所产生的交易成本（transactions cost）[一]。

25.7 垄断竞争

在前面，我们将垄断行业定义为只有唯一一个企业的行业。但是我们对行业到底由什么组成的还有些模糊。行业的一种定义是，由生产既定产品的所有企业组成的组织。但是产品又是指得什么？毕竟，生产可口可乐（Coca-Cola）的企业只有一家，这是否意味着这个企业就是垄断企业？

答案显然是否定的。可口可乐公司还必须与其它软饮料公司竞争。我们应该将行业看成某类企业的集合，这类企业生产的产品在消费者的眼里是相近的替代品。行业中的每个企业生产的产品都是独一无二的，因为它们产品的品牌都是不同的。但是在消费者眼里，这些品牌不同的产品在某种程度上都是互相替代的。

尽管某个企业对它的注册商标和品牌具有法定的独占权，因此其它企业不能生产与它**完全一样**的产品，但是其它企业通常可以生产**类似**的产品。从某家具体企业的角度来说，它在制定产量和价格决策时，必须考虑它的竞争对手的生产决策。

因此，某企业的需求曲线通常取决于生产类似产品的其他企业的产量和价格决策。该企业需求曲线的斜率，取决于其他企业的产品和它产品的相似程度。如果行业中有很多企业生产同质产品，那么任何一家企业的需求曲线基本都是水平的。每个企业都必须按照其他企业的要价销售商品。这种情形下，如果任何一家企业擅自提高价格，该企业都将很快失去所有消费者。

另外一方面，如果一家企业对某种特别产品的销售具有独占权，那么它提价时不会失去所有消费者。部分但不是全部消费者会购买它的竞争对手的产品。它失去的消费者的数量取决于消费者认为这些产品的相似程度有多大——也就是说，取决于该企业的需求曲线的弹性。

如果行业中的一家企业可以从销售产品中获得利润，而且不允许其他企业完全仿制该产品，那么其他企业仍然会发现，如果进入该行业并且生产与上述企业产品类似但有区别的产品的话，仍然有利可图。经济学家将这种现象称为**产品差异化**（product differentiation）——每家企业试图将自己的产品与行业中其他企业的产品区别开来。一家企业若越能成功地做到产品差异化，它所具有的垄断力量越大——也就是说，它的产品需求曲线越缺乏弹性。例如，考虑软饮料行业。在该行业中，有很多企业生产类似但不是完全同质的产品。每个产品都有自己的消费者，因此每个产品都有某种程度的市场力量。

[一] 简单地说，交易成本就是指当交易行为发生时，随同产生的信息搜寻、条件谈判与交易实施等的各项成本。在这里，作者是说，如果对摩天轮这类游乐项目单独收费，则可能需要添置收费系统或收费工作人员等，由此产生的成本大于不单独收费时摩天轮运营的边际成本。因此，如果将乘坐摩天轮的价格设定于等于边际成本的话，那么对摩天轮单独收费就是得不偿失的（因为价格小于交易成本）。译者注。

上面描述的行业结构既有竞争的因素又有垄断的因素；因此，有时将这种行业称为**垄断竞争**（monopolistic competition）。这种行业结构是垄断的，因为每个企业产品的需求曲线都是向下倾斜的。因此它有某些市场力量，即它可以指定自己产品的价格，而不是象竞争企业一样只能接受市场价格。另外一方面，这些企业必须在价格和产品方面竞争消费者。而且，新企业进入垄断竞争行业没有限制。在这些方面，这种行业又类似竞争性行业。

垄断竞争可能是现实世界中最普遍的行业结构形式。不幸的是，这种行业结构也最难以分析。完全垄断和完全竞争这两种极端情形的分析要简单得多，我们可以使用这两种行业结构初步近似更复杂的垄断竞争模型。在一个比较详细的垄断竞争模型中，结论主要取决于具体的产品和技术情况，也取决于企业可以使用的策略性选择。完全竞争和完全垄断的市场结构，我们可以进行抽象地分析，但是对垄断竞争进行抽象分析通常是不合理的。相反，我们需要仔细检查特定垄断竞争行业的制度细节。在下面两张，我们将介绍经济学家对策略性选择的一些分析方法，但是对垄断竞争行业的仔细分析却需要等到更高级的课程。

然而，我们可以分析垄断竞争行业自由进入的这一特征。当进入某特定产品行业的企业越来越多，行业内原有企业的需求曲线将会发生怎样的变化？首先，我们可以预期，需求曲线将向内移动，因为当更多的企业进入行业后，在每个价格水平下，该企业的销量都会减少。其次，我们可以预期，某个既定企业的需求曲线将变得更为平缓，因为新企业进入后，更多的企业出生产更多的类似产品。因此生产类似产品的新企业进入行业之后，行业内原有企业的需求曲线将向左移动并且变得更平缓。

如果企业预期到可以赚取利润就持续进入行业的话，那么均衡时必有：

1.每个企业的价格-销量组合都在它自己的需求曲线上。

2.给定每个企业的需求曲线，它们均已实现利润最大化。

3.自由进入迫使每个企业的利润为零。

这三个事实使得需求曲线和平均成本曲线呈现特殊的关系：需求曲线和平均成本曲线必定相切。

我们用图 25.6 说明这一点。事实 1 说明企业的产量-价格组合必然位于需求曲线上的某一点，事实 3 说明该产量-价格组合必然也位于平均成本曲线上。因此企业的生产之处必然位于这两条曲线的接触点上。那么，需求曲线能穿过平均成本曲线吗？不能，因为这样的话，需求曲线上的某些点就位于平均成本曲线的上方，而这又意味着这些点能产生正的利润。但是事实 2 指出，企业的最大化利润是零利润。因此，需求曲线必然和平均成本曲线相切。

理解这一点的另外一种方法是，问问你自己，在图 25.6 中，如果企业索要的价格不是盈亏平衡价格（即 $p \neq AC$）时，将会出现什么样的结果？这种情形下，不管企业索要的价格是高于还是低于盈亏平衡价格，企业都会出现亏损。这是因为只有在盈亏平衡价格-产量组合这一点上，这两条曲线才相切。在其他的价格-产量点上，需求需求都在平均成本曲线的下方，即 $p < AC$，从而利润 $p \times y - AC \times y < 0$。但是在盈亏平衡价格上，企业的利润

为零。因此，盈亏平衡价格是利润最大化价格。

对于垄断竞争均衡，我们有必要指出两点。首先，尽管利润为零，这种情形仍然是帕累托低效率的。利润和效率问题无关，也就是说我们不能以利润为零为否作为判断效率的依据。事实上，在垄断竞争均衡时，均衡价格必然大于**边际成本**，这是因为边际成本曲线穿过平均成本曲线的最低点，而由图 25.6 可看出，在均衡处，均衡价格是大于边际成本的。但是我们知道，如果价格大于边际成本，继续增加产量就是一种帕累托改进，因此垄断竞争均衡状态不是帕累托有效率的。

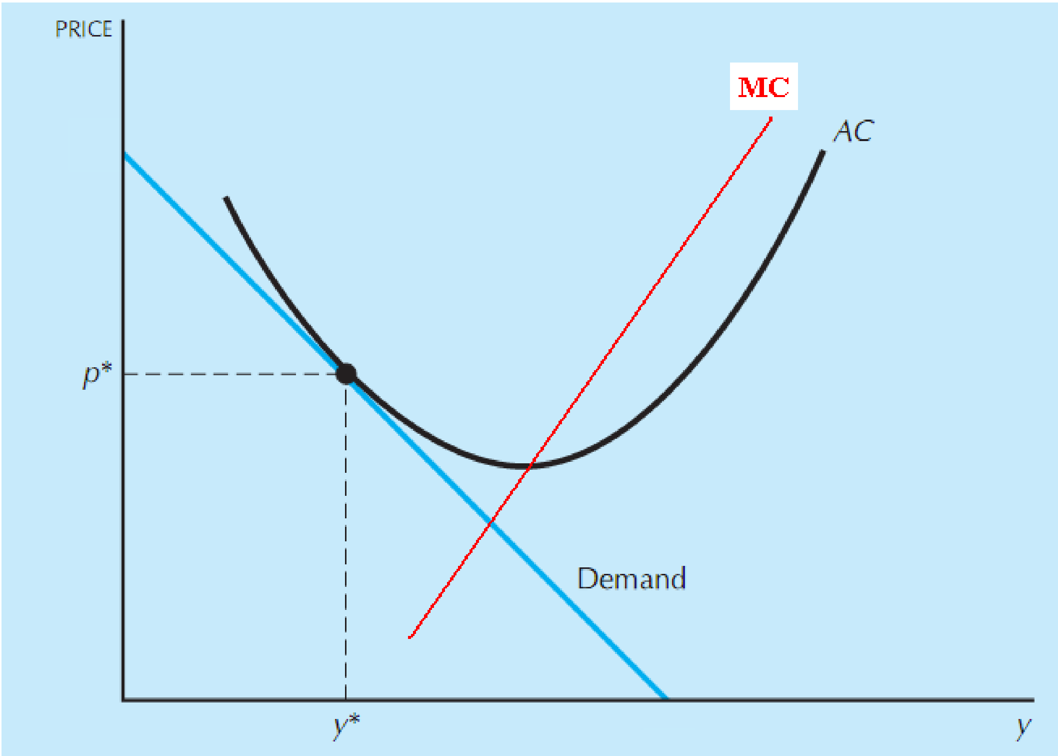


图 25.6：垄断竞争。在垄断竞争行业均衡时，若利润为零，则需求曲线和平均成本曲线必定相切。【注】图中的边际成本曲线为译者所加，目的在于分析垄断竞争均衡不是帕累托有效率的。添加了边际成本曲线后，在图形上明显可以看出，在均衡处，均衡价格大于边际成本。因此，此时增加产量就是一种帕累托改进，这也意味着垄断竞争均衡不是帕累托有效率的。

其次，从图 25.6 明显可以看出，企业通常在平均成本曲线最低点左侧生产。这种情形又是被解释为垄断竞争存在“**过剩的生产能力**”（excess capacity）。如果市场上企业数量比均衡时企业数量少，那么每个企业都可以在相对更有效率的规模上生产，这对消费者是有利的。然而，如果这样，这也意味着市场上的商品种类也减少了，这对消费者是不利的。到底哪种效应起主导作用，难以给出明确答案。

25.8 产品差异化的一个区位模型

在大西洋市，沿着海滩有一条用木板搭建的小路。在这条小路上，有些冰淇淋小贩推着手推车叫卖。如果允许一个小贩在这条小路上的任何位置销售，那么他应该选择什么样的位置？[一]

假设消费者在海滩上均匀分布。从社会的观点来看，让小贩定位是有价值的，因为我们使所有消费者购买冰淇淋时所走的路程总和最小。不难看出，这个最佳的位置就是这条小路长度的中间。

现在假设允许两个小贩售卖冰淇淋。假设我们固定冰淇淋的价格，而让小贩关注的问题仅为选择何处才能使消费者所走路程之和最小。如果每个消费者选择的是离他最近的小贩，我们应该将一个小贩放在小路长度的 1/4 处，将另外一个小贩放在小路长度的 3/4 处。位于小路中点的消费者对选择哪个小贩是无差异的，因为他们离这两个小贩的距离是一样的；每个小贩的市场份额都是 50%。（如图 25.7 所示。）

但是这两个小贩能呆在我们指定的上述位置上吗？比如你是小贩 L。如果你向右稍微移动一点，你不会失去原有的消费者，不仅如此，你还能从另外一个小贩的手里抢来部分消费者。向右移动，对于你左边的消费者来说，你仍然是离他们最近的小贩，但此时你却离你右边的消费者更近了，因此你的市场份额和利润都增加了。

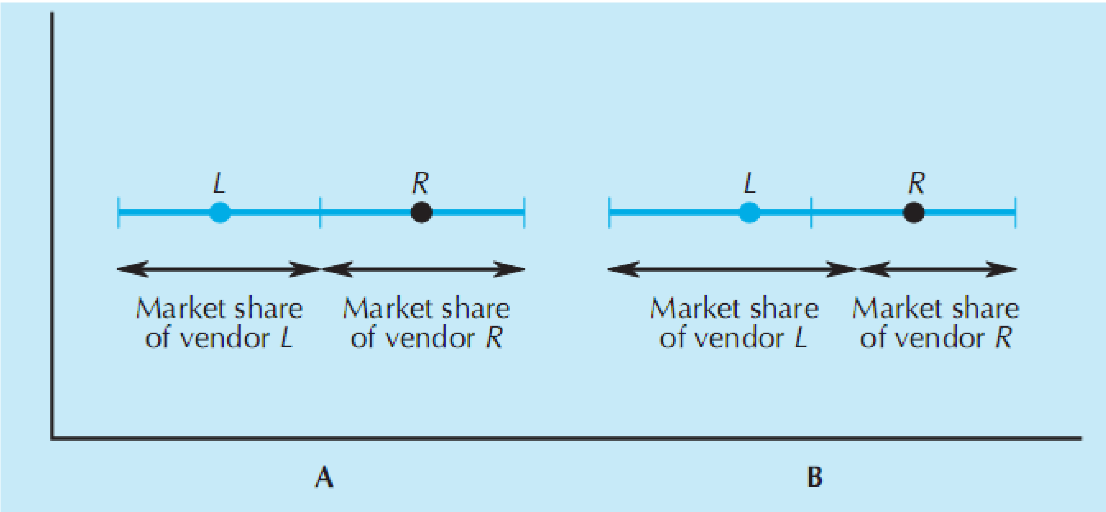


图 25.7：区位的竞争。A 图表明社会最优模式；小贩 L 和小贩 R 分别位于滨海小路长度的 1/4 处和 3/4 处。但是每个小贩都会发现向中点移动对他自己有利。唯一的均衡位置是两个小贩都位于这条小路的中点上，如图 B 所示。

但是小贩 R 可以类似推理。他向左移动就会夺走小贩 L 的部分消费者，而且他原有的

[一]这个例子来源于下面的经典模型：Harold Hotelling，“Stability in Competition，” Economic Journal, March 1929.

消费者一个也不会失去。这表明，社会最优位置模式无法实现均衡。唯一的均衡之处是两个小贩都在这条滨海小路的中点上，如图 25.7B 所示。在这种情形下，对消费者的争夺最终导致了一个缺乏效率的位置模式。

其他产品差异问题也可以使用这个定位模型进行分析。比如两家广播电台选择音乐类型的问题。音乐的两个极端是古典音乐和重金属摇滚。每个听众会选择最符合他偏好的广播电台。如果古典音乐电台稍微移向偏好范围的中间，它不会失去古典乐的听众，但会赢得部分品味一般的听众。单是摇滚乐电台出于同样的理由也会向中间移动。均衡时，两家电台播放的音乐类型相同，因此喜欢古典乐和喜欢摇滚乐的听众对这两个电台都不满意！

25.9 产品差异化

上述区位模型表明垄断竞争会导致产品差异化不足：每个企业都想让自己的产品尽量类似另外企业的产品，以便夺取其他企业的消费者。的确，我们认为如果某个市场中的产品过于类似，则这不是理想的情形。

但是，垄断竞争未必一定导致产品化差异化不足。事实上，它还有可能导致产品产异化过剩。假设滨海小路**非常**长，那么每个小贩都乐于呆在小路的两侧。如果他们的市场不重合，那么向中间位置移动不会增加利润。在这种情形下，垄断者没有模仿别人的激励，因此，它们会尽量使得产品差异化。

我们也可以建立导致产品差异化**过度**的垄断竞争模型。在这样的模型中，每个企业都试图让消费者相信它们的产品和竞争对手的产品是不同的，它们这样做的目的是创造某种程度的市场力量。如果某个企业成功做到让消费者相信自己的产品没有相近的替代品，那么它就可以对消费者索要较高的价格。

这会导致每个企业致力于建立与众不同的品牌。例如，洗衣皂是一种非常标准化的商品。但是很多生产企业投入巨资做广告，宣称如果你选择它的产品而不是竞争对手的产品，你的衣服将会洗得更干净、味道更好闻，甚至还会拥有更美满的婚姻和更幸福的人生！这种“产品定位”（product positioning）类似于小贩之间保持足够长的距离，以避免面对面的竞争。

有人曾经批评说，这种产品差异化的过度投资造成了大量浪费。在某些情形下，这种批评也许是对的，但是，过度差异化可能只是鼓励企业为消费者提供多样化产品的结果。因此，很难评价这两种效应到底哪一个起主导作用。

25.10 更多小贩

我们已经知道，如果两个小贩的市场区域重合，而且每个小贩的卖价相同，那么最终它们都会选择小路的中间位置。如果有更多的小贩参与竞争，结果将如何？

如果小贩数量众多，问题将变得比较复杂。显然，3 个小贩的情形最为简单。在这种情

形下，将会产生一个非常奇特的结果：不存在均衡位置！为了看清这一点，请看图 25.8.如果小路上有 3 个小贩，则其中一个小贩必然位于另外两个小贩之间。和前面的分析一样，外侧的 2 个小贩向中间的小贩移动，对这 2 个小贩有利，因为这样做的话，它们的原来消费者不会损失，而且还能抢夺中间小贩的消费者。但是如果这两个小贩距离中间小贩很近，那么中间的小贩会迅速移动到右侧小贩的右边或左侧小贩的左边，因为这样可以抢夺他们的消费者。无论在什么位置上，总会有人移动。

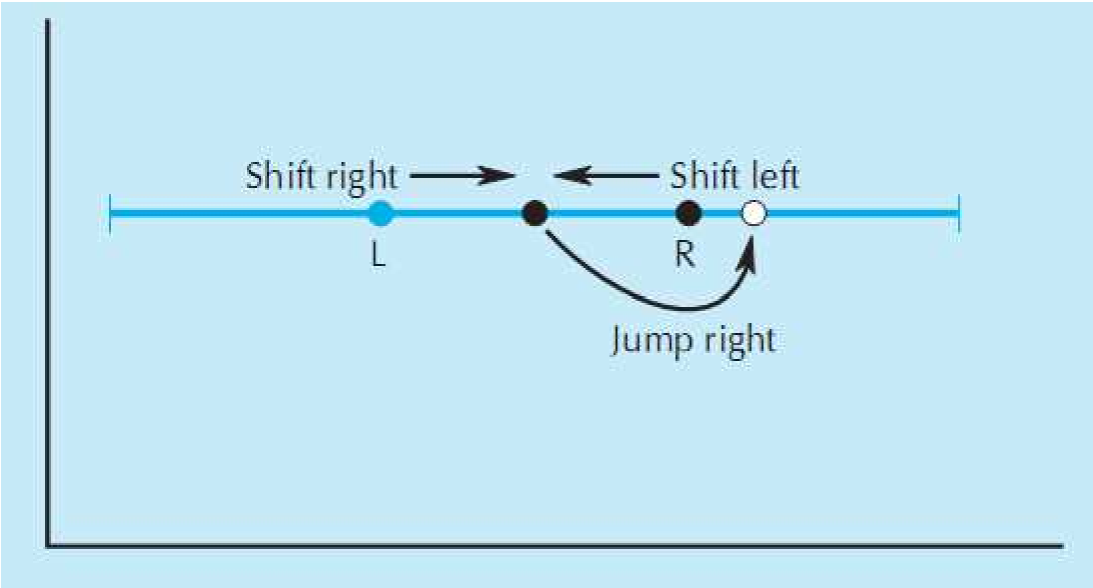


图 25.8：不存在均衡。在 3 个小贩的区位模型中，不存在纯策略均衡，因为给定任意的位置，总会有小贩会移动。

幸运的是，这种“反常”的结果只会出现在小贩数量为 3 个的情形下。如果小贩的数量为 4 个或者更多，则通常会有均衡位置。

总结

1.垄断企业通常有实施某种价格歧视的激励。

2.完全价格歧视是指企业对每个消费者索要一个不同的价格，这个价格通常具有“要买就买、不买请走人”的性质。这种情形下的产量水平是有效率的。

3.如果某企业能对两个不同的市场索要不同的价格，那么它总是向需求缺乏弹性的市场索要较高的价格，向需求富有弹性的市场索要较低的价格。

4.如果消费者能实施两部收费策略，而且消费者是同质的，那么它通常会将价格设定为

等于边际成本，从而全部利润都来源于门票。

5.垄断竞争的行业结构，是指在这种行业中存在产品差异化，因此每个企业都有某种程度的垄断力量，但是由于这种市场能够自由进入，从而利润为零。

6.垄断竞争一般会导致产品差异化不足或者产品差异化过度。

复习题

1.垄断企业会不会生产帕累托有效率的产量？

2.假设某垄断企业向两类消费者群体销售产品，这两类消费者群体的需求价格弹性固定不变，分别为ε_1和ε_2。企业的边际生产成本固定为c，该企业应该向两类消费者各索要多大的价格？

3.假设游乐园的园主，能对摩天轮服务实施完全价格歧视。假设摩天轮的运营的边际成本为零，而且假设消费者的偏好是相同的。那么这个园主会选择下列哪种定价策略：一是对乘坐摩天轮收费但门票免费；而是收取门票费用，但对乘坐摩天轮免费？

4.南加州的居民在去迪斯尼游玩时，门票可以享受打折优惠（在门口时你告诉他们你的邮政编码）。这是哪种价格歧视？你能据此推测出南加州乐园对迪斯尼游玩的需求是具有弹性的还是缺乏弹性的？

复习题答案

1.垄断企业会不会生产帕累托有效率的产量？

【复习内容】

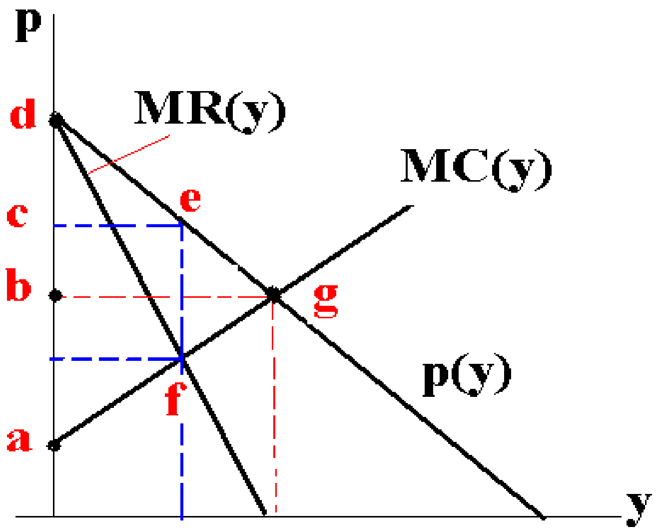
垄断的无效率；垄断企业的完全价格歧视

【参考答案】

垄断企业的产量是无效率的，但是如果垄断企业能够实施完全价格歧视，则它的产量是有效率的。分析如下：

有效率的产量要求$p(y)=MC(y)$，如下图所示。此时的产量是需求曲线$p(y)$和边际

成本曲线 $MC(y)$ 交点 g 对应的产量。此时价格在 b 点。此时消费者剩余为三角形 bdg，生产者剩余为三角形 abg，因此总剩余为三角形 adg。



垄断企业根据 $MR(y)=MC(y)$ 的原则进行生产，垄断产量是边际收入曲线 $MR(y)$ 和边际成本曲线 $MC(y)$ 交点 f 对应的产量。此时的价格在 c 点。此时的消费者剩余为三角形 cde，生产者剩余为梯形 acef。因此，总剩余为梯形 adef。

对照有效率的产量和垄断产量时的总剩余可知，垄断引起的净损失为三角形 efg。所以，垄断是无效率的。其实仅根据垄断产量小于有效率的产量即可知道垄断是无效率的。

但是，如果垄断企业能够实施完全价格歧视策略，则它的产量不再是 f 点对应的垄断产量，而是 g 点对应的有效率产量。这是由于 f 点对应的价格 e 大于此时的边际成本，而垄断企业又能做到完全价格歧视，因此在 f 点对应的产量处，如果它额外多生产一单位，它索要的价格高于边际成本。这种情形下，垄断企业会一直增加产量，直至价格等于边际成本，在图形上，它产量是需求曲线 $p(y)$ 和边际成本曲线 $MC(y)$ 交点 g 对应的产量。

由此可见，实施完全价格歧视的垄断企业的产量等于有效率的产量，这意味着这种情形下，它的产量是有效率的。如果从总剩余的角度分析也可以，此时，总剩余为为三角形 adg，和有效率的产量的总剩余相等。

所以说，垄断企业的产量是无效率的，但是如果垄断企业能够实施完全价格歧视，则它的产量是有效率的。

2.假设某垄断企业向两类消费者群体销售产品，这两类消费者群体的需求价格弹性固定不变，分别为 ε_1 和 ε_2。企业的边际生产成本固定为 c，该企业应该向两类消费者各索要多大的价格？

【复习内容】第三级价格歧视。

第三级价格歧视是垄断者对不同的消费者群体索要不同的价格，但是对于同一种消费群体中的每个人，它索要的价格是相同的。

假设垄断者能够有效区分两类消费者，并对每类消费者群体索要不同的价格。再假设每个市场上的消费者都不能将商品转卖。令 $p_1(y_1)$ 和 $p_2(y_2)$ 分别表示消费者群体 1 和 2 的反需求曲线，令 $c(y_1+y_2)$ 表示垄断者的成本函数。则垄断者的利润最大化问题为

$$\max_{y_1,y_2} \ p_1(y_1)y_1 + p_2(y_2)y_2 - c(y_1+y_2).$$

这个最大化问题的一阶条件为

$$\frac{\partial p_1(y_1)y_1}{\partial y_1} - \frac{\partial c(y_1+y_2)}{\partial y_1} = 0$$

$$\frac{\partial p_1(y_1)y_1}{\partial y_2} - \frac{\partial c(y_1+y_2)}{\partial y_2} = 0$$

可将这两个式子变形为

$$\frac{\partial p_1(y_1)y_1}{\partial y_1} - \frac{\partial c(y_1+y_2)}{\partial (y_1+y_2)}\frac{\partial (y_1+y_2)}{\partial y_1} = 0$$

$$\frac{\partial p_1(y_1)y_1}{\partial y_2} - \frac{\partial c(y_1+y_2)}{\partial (y_1+y_2)}\frac{\partial (y_1+y_2)}{\partial y_2} = 0$$

由于 $\frac{\partial (y_1+y_2)}{\partial y_1} = \frac{\partial y_1}{\partial y_1} + \frac{\partial y_2}{\partial y_1} = 1+0=1$；类似地，$\frac{\partial (y_1+y_2)}{\partial y_2} = 1$，所以上面两个式子又可变为

$$\frac{\partial p_1(y_1)y_1}{\partial y_1} - \frac{\partial c(y_1+y_2)}{\partial (y_1+y_2)} = 0$$

$$\frac{\partial p_1(y_1)y_1}{\partial y_2} - \frac{\partial c(y_1+y_2)}{\partial (y_1+y_2)} = 0$$

上面第一个式子也就是 $MR_1(y_1) - MC(y_1+y_2) = 0$，即

$$MR_1(y_1) = MC(y_1+y_2)$$

类似地，$MR_2(y_2) = MC(y_1+y_2)$

也即是说，额外多生产一单位产品的边际成本必须等于每个市场的边际收入。

【参考答案】

第三价格歧视的必要条件为

$$MR_1(y_1) = MC(y_1+y_2)；\quad MR_2(y_2) = MC(y_1+y_2).$$

也即是说，额外多生产一单位产品的边际成本必须等于每个市场的边际收入。

使用边际收入和需求价格弹性之间的关系，以及边际收入等于边际成本的关系，将利润最大化的一阶条件写为

$$p_1(y_1)[1-\frac{1}{|\varepsilon_1(y_1)|}]=MC(y_1+y_2)\text{；}\qquad p_2(y_2)[1-\frac{1}{|\varepsilon_2(y_2)|}]=MC(y_1+y_2)$$

将题目中的相关条件分别代入上面的两个式子可得

$$p_1(y_1)[1-\frac{1}{|\varepsilon_1|}]=c\text{；}\qquad p_2(y_2)[1-\frac{1}{|\varepsilon_2|}]=c$$

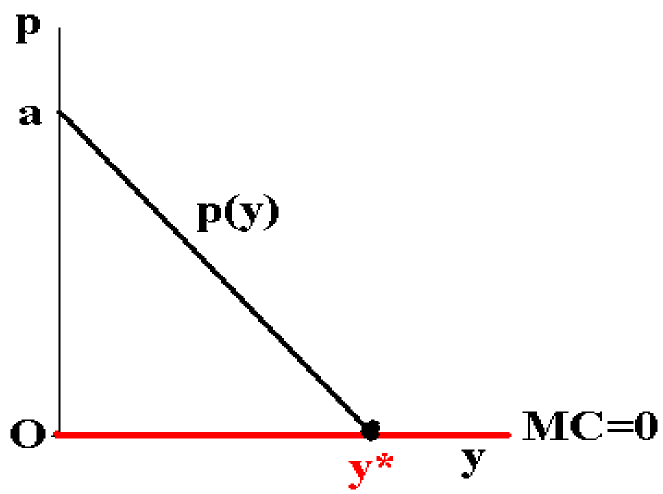
由此可得，

$$p_1(y_1)=\frac{c}{[1-\frac{1}{|\varepsilon_1|}]}\text{；}\qquad p_2(y_2)=\frac{c}{[1-\frac{1}{|\varepsilon_2|}]}\text{。}$$

3.假设游乐园的园主，能对摩天轮服务实施完全价格歧视。假设摩天轮的运营的边际成本为零，而且假设消费者的偏好是相同的。那么这个园主会选择下列哪种定价策略：一是对乘坐摩天轮收费但门票免费；而是收取门票费用，但对乘坐摩天轮免费？

【复习内容】两部收费制(two-part tariffs)

【参考答案】



首先假设园主对乘坐摩天轮收费，但对门票免费的情形。由于园主能够实施完全价格歧视策略，因此根据完全价格歧视理论可知，园主选择的最优产量是 y^* 使得

$$p(y^*)=MC(y^*).$$

由于题目设定 MC=0，因此最优产量位于需求曲线 p(y)和边际成本曲线 MC=0 的交点，此时的产量即为最优产量 y^*。这种情形下，园主可以占有所有的消费者剩余，此时原主的销售收入等于消费者剩余，即等于图中的三角形 Oa y^* 面积。假设园主的固定成本为 FC，

又因为 MC=0，所以园主的总成本就是 FC，因此他的利润=三角形 Oa y^* 面积—FC。

现在再来看若园主收取门票但对乘坐摩天轮免费的情形。这种情形下，如果园主将门票价格设定为等于三角形 Oa y^* 的面积，那么他的利润=三角形 Oa y^* 面积—FC。

由此可见，在这两种情形下，他的利润是一样的，因此他在这两种收费方式之间无差异。

4.南加州的居民在去迪斯尼游玩时，门票可以享受打折优惠（在门口时你告诉他们你的邮政编码）。这是哪种价格歧视？你能据此推测出南加州乐园对迪斯尼游玩的需求是具有弹性的还是缺乏弹性的？

【复习内容】第三级价格歧视

【参考答案】

首先来解释题目，迪斯尼之所以对南加州的居民实行价格优惠，显然是由于南加州居民对这位“邻居”的游乐内容已比较熟悉，因此他们具有较高的需求价格弹性。

而第三级价格歧视是垄断者对不同的消费者群体索要不同的价格，但是对于同一种消费群体中的每个人，它索要的价格是相同的。

由此可见，迪斯尼的这种价格歧视策略是第三级价格歧视。

曹乾（东南大学 caoqianseu@163.com）

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

26.要素市场（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

26 要素市场

我们在第 19 章已初步分析过要素需求，但是在那里我们考虑的情形是，企业面对的是竞争的产品市场和竞争的要素市场。既然我们已经研究过了垄断行为，我们就可以分析其他情形下的要素需求行为。例如，如果某企业在它的产品市场为垄断者，它的要素需求将是什么样的？或者如果某企业是某些要素的唯一需求者，结果又将如何？在本章我们就来分析这些问题以及一些相关的问题。

26.1 在产品市场上是垄断的

当企业决定能使利润最大化的要素需求时，它选择的要素数量要能使得稍微多使用一点该种要素带来的边际收入，等于多使用这点要素产生的边际成本。这可由标准的逻辑推出：如果某行为的边际收入不等于它的边际成本，那么企业改变行为就是有利的。

这个一般的原则有不同的版本，这取决于我们对企业运营环境的假设。例如，假设企业在它的产品市场是垄断的。为简单起见，假设只有一种生产要素，因此可将生产函数写为 $y=f(x)$。企业得到的收入取决于它的产量，因此 $R(y)=p(y)y$，其中 $p(y)$ 为反需求函数。下面我们分析要素数量的边际增加是如何影响企业的收入的。

假设我们稍微增加一些要素的投入 Δx，这将会导致产量的微小变动 Δy。产量增加量与要素增加量的比率称为该要素的**边际产品**（marginal product）：

$$MP_x=\frac{\Delta y}{\Delta x}=\frac{f(x+\Delta x)-f(x)}{\Delta x}. \tag{26.1}$$

边际产品的增加将会导致收入变动。收入的相对变动称为**边际收入**（marginal revenue）：

$$MP_y=\frac{\Delta R}{\Delta y}=\frac{R(y+\Delta y)-R(y)}{\Delta y} \tag{26.2}$$

要素投入边际增加对收入的效应称为**边际产品收入**（marginal revenue product）。由该定义以及（26.1）和（26.2）式可知

$$\begin{aligned}MRP_y&=\frac{\Delta R}{\Delta x}=\frac{\Delta R}{\Delta y}\frac{\Delta y}{\Delta x}\\&=MR_y\times MP_y.\end{aligned}$$

我们可以使用边际收入的表达式将上式改写为

$$\begin{aligned} MRP_y &= [p(y) + \frac{\Delta p}{\Delta y} y] \times MP_y \\ &= p(y)[1 + \frac{1}{\varepsilon}]MP_x \\ &= p(y)[1 - \frac{1}{|\varepsilon|}]MP_x. \end{aligned}$$

第一个表达式中的边际收入是边际收入的常用表达式。第二个和第三个表达式中的边际收入是使用弹性形式表达，这一点我们已经在第 15 章分析过了。

现在容易看出这个式子如何扩展了第 19 章对竞争情形的分析。竞争市场中的单个企业的需求曲线的弹性是无穷的；因此，竞争企业的边际收入正好等于价格。因此，竞争市场上，要素的“边际产品收入”正好等于该要素的**边际产品价值**（value of the marginal product）pMP_x。

(垄断情形下的)边际产品收入与边际产品价值相比哪个大？由于需求曲线的斜率非正，边际产品收入总是小于或等于边际产品价值：

$$MRP_y = p[1 - \frac{1}{|\varepsilon|}]MP_x \le pMP_x.$$

只要需求函数不是完全具有弹性的，MRP_x 总是严格小于 pMP_x。这个结论意味着在某要素的任何使用数量水平上，垄断企业额外一单位的边际价值总是小于完全竞争企业的。在本节剩下的部分，我们假设研究的就是生产企业拥有垄断力量的情形。

如果你初次见到这个结论，你可能会觉得会和以前的结论矛盾，因为你已知道垄断企业的利润大于竞争企业——这意味着垄断企业的总的要素投入比完全竞争企业的总的要素投入“更有价值”。

这种“矛盾”的解决之道是注意总价值和边际价值的区别。垄断企业的**总**要素投入比完全竞争企业的总要素投入的确更有价值，因为从这些要素身上，垄断企业比竞争企业能赚取更多的利润。然而，在**给定**的产量水平上，要素投入增加将会使产量增加，从而使垄断企业能索要的**价格**降低。但是，这种情形对于完全竞争企业来说却是不成立的，因为它的产量增加不会改变它能索要的价格。因此，竞争企业的**边际**要素投入比垄断企业的边际要素投入更有价值。

由于在短期，与竞争企业相比，垄断企业的边际要素投入能带来的价值更小，因此垄断企业通常会使用较少的要素数量。这一般来说是正确的：垄断企业增加利润的方式是减少产量，因此和竞争企业相比，它使用的要素数量较少。

为了确定垄断企业使用多少要素，我们必须比较额外增加一单位要素能带来的边际收入，和这单位要素的边际成本哪个更大。假设企业面对的要素市场上是完全竞争的，因此，给定要素价格 w，该企业会使用尽可能多的要素数量。在这种情形下，假设竞争企业的要素使用量为 x_c，其中下标表示“竞争”，则 x_c 需要满足

$$pMP(x_c) = w.$$

另一方面，垄断企业使用的要素数量 x_m 满足

$$MRP(x_m) = w.$$

我们用图 26.1 进行分析。由于 $MRP(x_c) < pMP(x)$，满足 $MRP(x_m) = w$ 的点 x_m，必然位于满足 $pMP(x_c) = w$ 的点 x_c 的左侧。因此，垄断企业的要素使用量小于竞争企业的要素使用量。

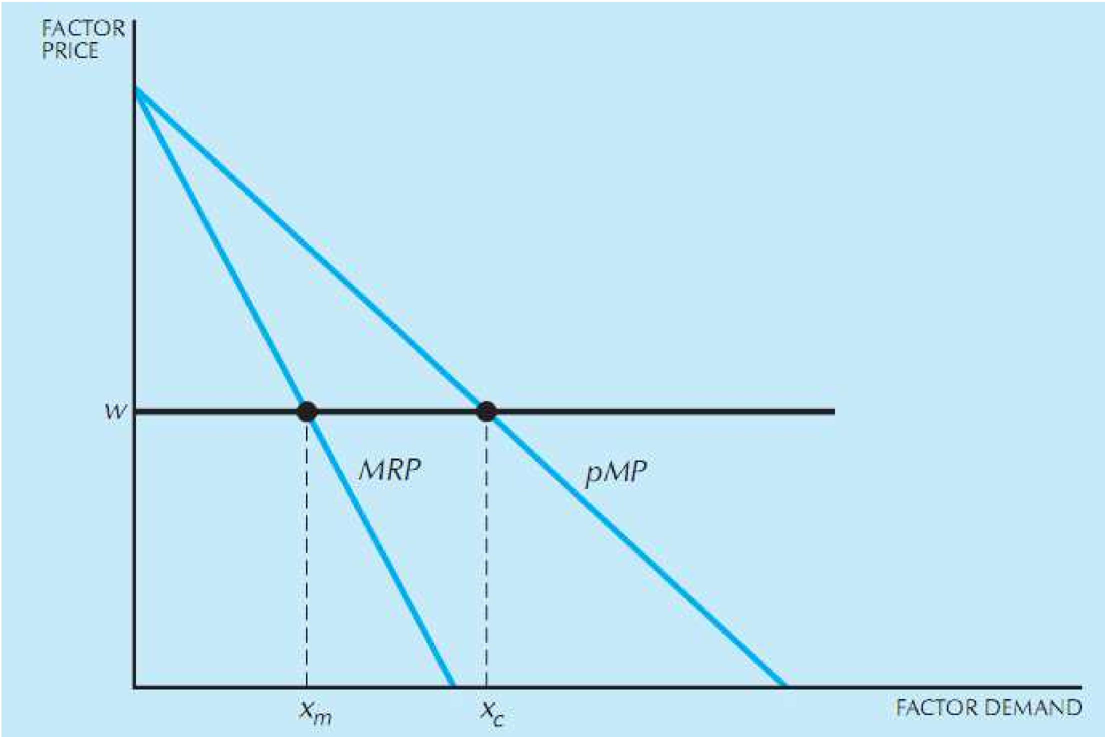


图 26.1: 垄断企业的要素需求。由于边际产品收入曲线（MRP）位于边际产品价值曲线（pMP）的下方，因此，垄断企业的要素需求量会小于它若是竞争企业时的要素需求量。

26.2 买方垄断

垄断（monopoly）是某商品市场上**卖方**只有一家，而**买方垄断**（monopsony）是指买方只有一个。对买方垄断的分析类似于对垄断的分析。为简单起见，假设买方垄断企业在产品市场上是完全竞争者。

和前面一样，假设该企业使用一种要素生产一种产品，生产函数为 $y = f(x)$。然而，与前面的分析不一致的地方是，这样的企业在要素市场上垄断者，它能认识到它对要素的需求会影响到它对这种要素支付的价格。

我们用（反）供给曲线 $w(x)$ 表示这种关系。该函数是说如果企业想使用 x 单位的要素，它必须支付的价格为 $w(x)$。假设 $w(x)$ 为增函数：x 要素的使用量越多，企业支付的单位价格越高。

根据定义，竞争的要素市场市场中的企业面对的要素供给曲线是水平的：在现行要素价格下，它想购买多少就购买多少。买方垄断企业面对的供给曲线则是向上倾斜的：它购买的要素越多，它支付的要素价格越高。竞争的要素市场中的企业是**价格接受者**（price taker），而买方垄断企业是要素的**价格制定者**（price maker）。

买方垄断企业面对的利润最大化问题为

$$\max_{x} \; pf(x) - w(x)x \,.$$

利润最大化的条件是额外多使用一单位要素带来的边际收入，等于这单位要素的边际成本。既然竞争的产品市场的边际收入为 pMP_x。边际成本为多大？

多使用 Δx 单位要素引起的成本的总变动为

$$\Delta c = w\Delta x + x\Delta w \,,$$

因此每单位成本的变动为

$$\frac{\Delta c}{\Delta x} = MC_x = w + \frac{\Delta w}{\Delta x} x \,.$$

这个表达式的解释类似于边际收入的解释：当企业使用的要素增加 Δx 时，它必须多支付 $w\Delta x$ 元。但是，要素使用量增加会使该要素价格上升 Δw 元，因此原来购买的 x 单位要素按新价格支付，需要多支付 $x\Delta w$ 元。

我们也可以把额外多使用一单位要素的边际成本写为

$$\begin{aligned} MC_x &= w[1 + \frac{x}{w}\frac{\Delta w}{\Delta x}] \\ &= w[1 + \frac{1}{\eta}] \end{aligned}$$

其中 η 是要素的供给价格弹性。由于供给弹性通常向上倾斜，η 一般为正。如果供给曲线具有**完全的**弹性，那么 η 为无穷大，这种情形下企业面对的是竞争的要素市场。注意这些结论和垄断情形的相似性。

下面我们分析买方垄断企业面对的要素供给曲线为线性的情形。反供给曲线的形式为

$$w(x) = a + bx \,,$$

因此总成本为

$$C(x) = w(x)x = ax + bx^2 \,,$$

多使用一单位要素产生的边际成本为

$$MC_x(x) = a + 2bx\,.$$

买方垄断企业的利润最大化问题的解可用图 26.2 说明。我们首先找到边际产品价值等于边际成本时的要素使用量 x^*，然后找到这个产量对应的价格。

由于使用额外一单位要素的边际成本大于该要素的价格，该要素的价格会小于若要素市场是竞争市场时的价格。和垄断的情形一样，买方垄断企业也在帕累托低效率的点经营。只不过现在这种低效率是在要素市场中发生的，而不是在产品市场中发生的。

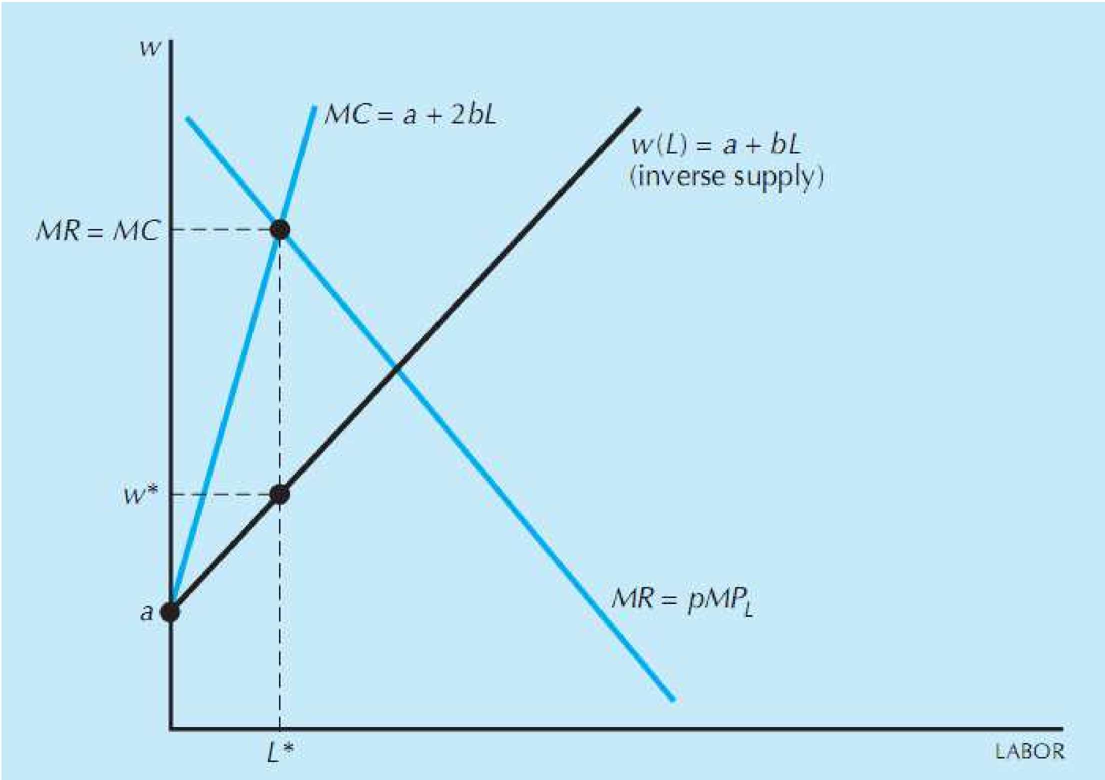


图 26.2：买方垄断。买方垄断企业使用要素的最优数量，是使得额外多使用一单位要素带来的边际收入，等于这单位要素的边际成本。

例子：最低工资

假设劳动市场是完全竞争的，并且政府设定的最低工资高于现行的市场均衡工资。由于在均衡工资处，需求等于供给，若最低工资高于均衡工资，则在该最低工资处，劳动的供给会大于劳动的需求。如图 26.3A 所示。如果劳动市场是买方垄断的，则结果就会大不相同。在这种情形下，若规定最低工资，则劳动者的就业可能**增加**。如图 26.3B 所示。

如果政府将最低工资设定为等于竞争市场的现行工资，那么“买方垄断企业”现在认

为它可以按照不变的工资 w_c 雇佣劳动者。由于现在它面对的工资率和它雇佣的劳动者的数量无关，它会增加雇佣劳动者的数量直至劳动的边际产品价值等于 w_c 。也就是说，这种情形下，它雇佣的劳动者的数量和若它面对的是竞争市场情形下雇佣的数量一样多。

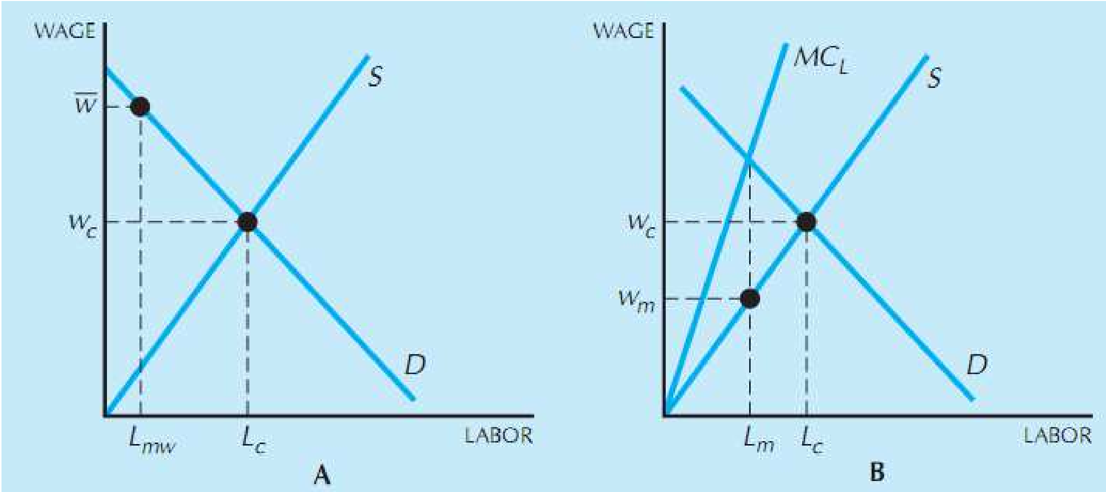


图 26.3：最低工资。A 图表示在竞争的劳动市场中，最低工资的效应。当工资为竞争性工资 w_c 时，劳动者的就业量为 L_c 。若规定了最低工资 $\overline{w}$ ，则就业量为 L_{mw} 。B 图表示在买方垄断的劳动市场中，最低工资的效应。在买方垄断的情形下，工资 w_m 和就业量 L_m 小于竞争劳动市场的工资和就业量。若最低工资设定为 w_c ，则就业量会增加到 L_c 。

在买方垄断的情形下，政府设定最低工资(wage floor)，这类似于在垄断情形下，对企业的产品设定最高卖价（price ceiling）。这两种政策都能使企业的行为类似竞争企业的行为。

26.3 上游垄断和下游垄断

在前面，我们已分析了两种不完全竞争的要素市场的情形：一种是企业在产品市场上作为卖方是垄断的，但在要素市场上作为买方是竞争的；另外一种情形是企业在产品市场上作为卖方是竞争的，但在要素市场上作为买方是垄断的。除了这两种情形外，还有其他的情形。比如，企业在要素市场上要面对一个垄断的要素卖方。或者，它的产品为某企业独家买断。当然还有其他情形，但是我们没必要仔细分析每一种情形，因为这样会出现重复。然而，我们将分析一种有趣的市场结构，在这种市场上，垄断企业的产品被另外一家垄断企业用作生产要素。

假设一家垄断企业 A 生产了 x 单位的产品，再假设它的边际生产成本恒为 x 。它将这些产品卖给另外一家垄断企业 B，卖价为 k 元/单位产品，企业 B 将这些产品当作生产要素使用。我们将企业 A 称为**上游垄断企业**（upstream monopolist），将企业 B 称为**下游垄断企业**（downstream monopolist）。下游垄断企业使用这些要素 x 生产出 y 单位产品，它的生产函数为 $y=f(x)$ 。然后，这些产品在垄断市场上销售，该市场的反需求函数为 $p(y)$ 。为

了简要说明这个例子，我们使用线性的反需求函数 $p(y)=a-by$ 说明。

为简单起见，假设生产函数为 $y=x$，也就是说对于每单位 x 要素的投入，都可以产出一单位商品 y。进一步假设下游垄断企业的成本，仅为按价格 k 购买上游企业产品的货款，除此之外，没有其他成本。

为了看清这种市场是怎么运行的，我们从下游企业开始分析。下游企业的利润最大化问题为

$$\max_{y}\ p(y)y-ky=[a-by]y-ky.$$

令边际收入等于边际成本，可得

$$a-2by=k\text{，}$$

由此可得

$$y=\frac{a-k}{2b}.$$

根据生产函数 $y=x$ 可知，下游企业生产 y 单位产品，需要向上游企业购买 x 单位要素，因此由上式可得到下游企业的要素需求函数

$$x=\frac{a-k}{2b}. \tag{26.3}$$

这个函数表示要素价格 k 和下游企业要素需求量之间的关系。

现在分析上游企业的问题。假设它知道上述过程，而且假设若它制定要素的不同价格 k，它能决定应该销售多少 x 要素，这就是（26.3）表示的要素需求函数。上游企业的利润最大化问题是决定 x 要素的最优销量。

很容易求得这个最优数量。从（26.3）式解出 k，它是 x 的函数，可得：

$$k=a-2bx.$$

这个要素需求函数相伴的边际收入为

$$MR=a-4bx.$$

令边际收入等于边际成本，可得

$$a-4bx=c\text{，}$$

或

$$x=\frac{a-c}{4b}.$$

由于生产函数为 $y=x$，因此企业最终产品的产量为

$$y=\frac{a-c}{4b} \qquad (26.4)$$

现在，我们要将这个产量和一个一体化的垄断企业的产量进行比较。假设上游企业和下游企业合并成一个一体化企业，这样我们就有了一家垄断企业而不再是两家，它的产品反需求函数为 $p=a-by$，它的边际成产成本恒为常数 c。令边际收入和边际成本相等可得

$$a-2by=c,$$

由此可得利润最大化的产量为

$$y=\frac{a-c}{2b}.$$

比较（26.4）式和（26.5）式可知，一体化的垄断企业的产量是非一体化垄断企业产量的 2 倍。

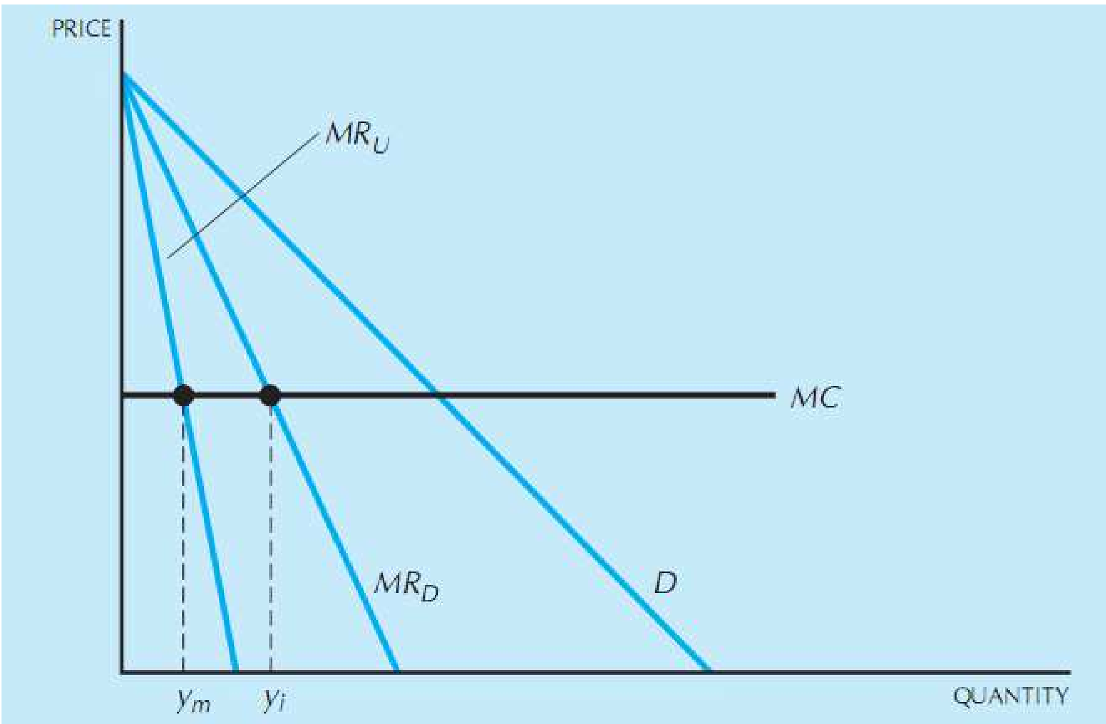


图 26.4：上游垄断和下游垄断。下游垄断企业面对的（反）需求函数为 $p(y)$，和该需求曲线相伴的边际收入为 $MR_D(y)$。这个边际收入曲线又是上游垄断企业面对的需求曲线，该需求曲线相伴的边际收入为 $MR_U(y)$。一体化垄断企业的产量为 y_i^*；非一体化垄断企业的产量为 y_m^*。

我们用图 26.4 说明这种情形。下游企业面对的最终需求曲线为 $p(y)$，和该需求函数相伴的边际收入曲线，同时又是上游企业面对的需求函数。和这个需求函数相伴的边际收入曲

线的斜率因此为下游企业最终需求曲线 $p(y)$ 斜率的 4 倍，这就是为什么该市场的产量是一体化市场产量的 1/2 的原因。

当然，上游企业边际收入曲线 $MR_U(y)$ 的斜率是下游企业需求曲线斜率的 4 倍的原因是我们使用了线性需求函数。然而，不难看出，一体化垄断企业的产量总是大于非一体化的上游-下游这样一对垄断企业的产量。在后面一种情形下，上游企业将价格设定为大于边际成本，下游企业也将价格设定为大于边际成本。这是**双重加价**（double markup）。这个价格，不仅从社会角度看太高，而且从使总利润最大化的角度看也太高！如果这两个垄断企业合并，价格就会下降，利润就会上升。

附录

我们可以使用求导数的链式法则（chain rule）计算边际产品收入。令 $y=f(x)$ 表示生产函数， $p(y)$ 为反需求函数。则收入是要素使用量的函数：

$$R(x)=p(f(x))f(x).$$

上式对 x 求导可得

$$\begin{aligned}\frac{dR(x)}{dx}&=p(y)f'(x)+f(x)p'(y)f'(x)\\&=[p(y)+p'(y)y]f'(x)\\&=MR\times MP.\end{aligned}$$

我们分析下面这种企业的行为，该企业在产品市场上是竞争者，在要素市场上是买方垄断者。令 $w(x)$ 表示要素的反供给函数，那么该企业的利润最大化问题为

$$\max_x \quad pf(x)-w(x)x.$$

该问题的一阶条件为

$$pf'(x)=w(x)+w'(x)x=w(x)[1+\frac{x}{w}\frac{dw}{dx}]=w(x)[1+\frac{1}{\eta}].$$

由于要素供给曲线向上倾斜，上式右端的表达式即 $w(x)[1+(1/\eta)]$ 大于 w 。因此，买方垄断企业使用的要素数量，会比竞争企业使用的要素少。

曹乾（东南大学 caoqianseu@163.com）

总结

1.追求利润最大化的企业总是希望它的每一项行动的边际收入，等于该行动的边际成本。

2.在垄断的情形下，增加要素使用量带来的边际收入称为边际产品收入。

3.对于垄断企业来说，它的边际产品收入（MRP）总是小于边际产品价值（VMP），这是由于增加产量带来的边际收入总是小于产品的价格。

4.和垄断市场上只有一个卖方类似，买方垄断市场上只有一个买方。

5.对于买方垄断企业来说，某种要素相伴的边际成本曲线比该要素供给曲线陡峭。

6.因此，买方垄断企业的产量是低效率的，因为这个产量小于若该企业为竞争企业的产量。

7.若某上游垄断企业将产品卖给一个下游企业，该下游企业将这些产品作为生产要素使用。那么下游企业的产品价格将会很高，原因在于双重加价（double markup）。

复习题

1.我们知道，垄断企业决不会在产品需求缺乏弹性之处生产。买方垄断企业会在要素供给缺乏弹性之处生产吗？

2.在课文最低工资的例子中，如果劳动市场是买方垄断的，而且假设政府设定的最低工资高于竞争工资，那么会有什么样的结果？

3.在课文中的上游和下游垄断企业的例子中，我们推导出了总产量的表达式。均衡价格 p 和 k 的表达式是什么？

复习题答案

1.我们知道，垄断企业决不会在产品需求缺乏弹性之处生产。买方垄断企业会在要素供给缺乏弹性之处生产吗？

【复习内容】垄断企业的利润最大化问题；买方垄断企业的利润最大化问题

先来回顾一下垄断企业为什么不会在产品需求缺乏弹性之处进行生产。

定义收入函数$r(y)=p(y)y$，则垄断企业的利润最大化问题为

$$\max_{y}\quad p(y)y-c(y)$$

该问题的一阶条件（目标函数对y求导，并令其等于0）为

$$\begin{aligned}&MR(y)=p(y)+p'(y)y=MC(y)\\&\Leftrightarrow p(y)[1+\frac{dp(y)}{dy}\frac{y}{p(y)}]=MC(y)\\&\Leftrightarrow p(y)[1+\frac{1}{\varepsilon(y)}]=MC(y)\end{aligned}$$

由于需求价格弹性一般为负，我们可以将上式写为

$$p(y)[1-\frac{1}{|\varepsilon(y)|}]=MC(y)$$

假设垄断企业会在产品需求缺乏弹性之处生产，即$|\varepsilon(y)|$，则上式左侧小于零，而上式右侧即边际成本不可能为负，所以垄断企业不可能产品缺乏弹性之处生产。

对于买方垄断企业（假设它在产品市场上是竞争者，在要素市场上是买方垄断者），注意到**要素供给的弹性是正的**，因此可以猜想买方垄断企业能在要素供给缺乏弹性之处进行生产。下面我们严格证明这个结论。

【参考答案】

为简单起见，假设买方垄断企业在产品市场上是竞争者，在要素市场上是买方垄断者。令$w(x)$表示要素的反供给函数，那么该企业的利润最大化问题为

$$\max_{x}\quad pf(x)-w(x)x.$$

该问题的一阶条件为

$$pf'(x) = w(x) + w'(x)x = w(x)[1 + \frac{x}{w}\frac{dw}{dx}] = w(x)[1 + \frac{1}{\eta}].$$

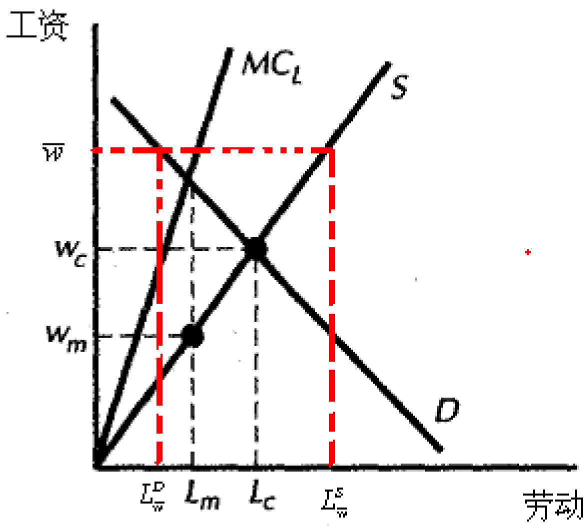
$$即\ pMP_x = w(x)[1 + \frac{1}{\eta}].$$

最后一个式子仍然是说是使用要素带来的边际收入（左侧）等于边际成本（右侧）。由该式可以看出，由于要素供给弹性$\eta > 0$，因此它不会出现垄断企业若在产品需求缺乏弹性处生产将导致边际成本为负的现象。所以说买方垄断厂商会在要素的任何供给弹性水平上进行生产，对供给弹性的大小没有要求。

2.在课文最低工资的例子中，如果劳动市场是买方垄断的，而且假设政府设定的最低工资高于竞争工资，那么会有什么样的结果？

【复习内容】买方垄断厂商使用生产要素的原则；最低工资

【参考答案】如下图所示。



当工资为竞争性工资w_c时（其中下标 c 表示竞争），劳动者的就业量为L_c。在买方垄断的情形下，工资w_m和就业量L_m小于竞争劳动市场的工资和就业量（其中下标 m 表示买方垄断）。若最低工资设定为w_c，则就业量会增加到L_c。

但是，如果政府规定的工资$\overline{w} > w_c$，那么由上图可以看出，买方垄断对劳动的需求为L_w^D，但劳动的供给为L_w^S（其中上标 D 和 S 分别表示需求和供给）。由此可见，此时会出现失业的现象。

3.在课文中的上游和下游垄断企业的例子中，我们推导出了总产量的表达式。均衡价格 p 和 k 的表达式是什么？

【复习内容】上游垄断和下游垄断

【参考答案】

在课文的例子中，我们已经解出了上游垄断企业的最优产量：

$$x=\frac{a-c}{4b}$$

以及下游垄断企业的最优产量

$$y=\frac{a-c}{4b}.$$

根据课文中的例子，我们可知下游企业产品需求函数为 $p=a-by$，所以将第二式 $y=\frac{a-c}{4b}$ 代入该需求函数可得 $p=a-by=a-b\times\frac{a-c}{4b}=\frac{3a+c}{4}$。

在课文中的例子中，我们还知道 $k=a-2bx$，因此将前面的第一式即 $x=\frac{a-c}{4b}$ 代入，可解得 $k=\frac{a+c}{2}$.

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

27.寡头垄断（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

27 寡头垄断

我们已经分析了两种重要的市场结构形式：纯粹竞争和纯粹垄断。在纯粹竞争的市场上，企业数量众多而且单个企业规模很小，而纯粹垄断市场上只有一家企业而且企业规模很大。然而，现实世界中的市场结构多数位于这两个极端之间。通常，市场中有一定数量的竞争者，但还没多到认为自己对市场价格无影响的程度。这种类型的市场称为**寡头（垄断）**（oligopoly）。

第 24 章分析的垄断竞争模型是寡头垄断的一种特殊形式，在那一章我们强调的是产品差异化和市场进入的问题。然而，我们在本章分析的寡头垄断模型，则侧重于市场中企业的策略性互动（strategic interactions）问题。

寡头垄断的模型有好几种，这是因为在寡头垄断的环境下企业的行为有多种。由于在现实世界中寡头企业的行为有多种不同的模式，因此不可能存在一种普适性的模型。我们想要的是一些可能的行为模式指南，以及通过什么样的特征去判断该使用哪种模型。

为简单起见，我们仅限于分析市场只有两家企业的情形，这种情形称为**双头（垄断）**（duopoly）。使用双头进行分析的好处是，我们既能得到参与策略性互动的企业的许多重要特征，又避免了使用多家企业进行分析时涉及的繁杂符号的难题。当然，我们仅限于分析双头生产的产品为同质产品的情形，这样做的好处是避免了分析产品差异化的问题，从而可以集中分析企业间的策略性互动行为。

27.1 策略选择

如果市场上只有两家企业，而且生产的产品相同，那么我们对四个变量感兴趣：每家企业索要的价格和每家企业生产的产量。

当一家企业在制定价格和产量决策时，它可能已知道另外一家企业的决策。如果一家企业率先作出价格决策，我们将其称为**价格领导者**（price leader），而将另外一家企业称为**价格追随者**（price follower）。类似地，如果一个企业率先作出产量决策，那么它就是**产量领导者**（quantity leader），另外一个企业就是**产量追随者**(quantity follower)。这些情形下的策略性互动形成了一个**序贯博弈**（sequential game）[一]。

另外一方面，在一个企业制定价格决策时，它可能不知道另外一个企业的价格决策。在这种情形下，为了制定合理的决策，它必须猜测另外一个企业决策。这就是**同时博弈**（simultaneous game）。这也有两种可能性：两个企业同时选择价格或者同时选择产量。

[一] 我们将在下一章详细介绍博弈论（game theory），但似乎在这一章介绍这些具体的例子是合适的。

上述分类方法给出了四种可能性：产量领导，价格领导，同时决定产量，同时决定价格。每种类型的互动都会产生一组不同的策略问题。

我们还将分析另一类互动，即企业不再以上述这种或者那种的形式相互竞争，而是选择了**合谋**（collude）。在这种情形下，两个企业联合起来共同决定价格和数量，目的是使它们的利润之和最大化。这种合谋叫做**合作博弈**（cooperative game）。

例子：价格匹配[一]

我们经常看到零售企业在广告中宣称自己的产品价格全城“最低”（“meet or beat” any price）。这些广告语一般被视为激烈竞争市场的信号。然而，这样的报价也可以作为阻碍竞争的方法。

假设某个小镇有两家轮胎商店，按所在位置不同分别称为东边轮胎商店和西边轮胎商店。假设它们在广告中对相同品牌的轮胎的报价相同，都为每个 50 元。

如果东边商店在广告中的报价将为每个 45 元，而西边商店仍报价 50 元，那么我们可以预期小镇西边的某些消费者愿意多跑几分钟路程，到东边商店购买，因为每个轮胎可以少花 5 元钱。由于价格低，东边商店的销量将增加。若增加的销量足够大、足以抵消降价的效应，那么它的利润会增加。

这在本质上仍然是价格竞争：如果消费者对价格非常敏感，那么降价商店的销量会大增，利润也会增加。

假设西边商店在广告中仍然报价 50 元，即它不明确降价，但它增加了一个承诺：它的价格将与任何更低的价格匹配（match）。现在如果东边商店降价将会发生什么样的后果？

在这种情形下，小镇西边的顾客只要向西边商店出示东边商店的广告，西边商店将按东边商店的价格卖给他们轮胎。结果，东边商店降低价格却没吸引到新的顾客。事实上，它损失了部分销售收入，因为它的销量没变但价格降低了。

结论：零售企业通过提供价格匹配承诺的方式，大大打击了竞争对手降低价格的想法。

27.2 数量领导

在数量领导的情形下，一个企业率先作出产量决策，这个模型有时叫作**斯坦科尔伯格模型**（Stackelberg model），以纪念第一位系统研究领导者-追随者互动问题的经济学家[二]。

[一] 术语“price matching”较难找到对应的中文翻译，因为 match 一词在这里同时含有“竞争”、“相等”两个词义，并且隐含着“搜索”的意思。姑且翻译为“价格匹配”。中文中的“匹配”一词虽然也含有“竞争”、“相符”等含义，但个人感觉“匹配”没有“math”微妙。译者注。

[二] Heinrich von Stackelberg 是一位德国经济学家，他于 1934 年出版了颇具影响力的市场组织著作《市场组织形式与均衡》。

经济学家通常用斯坦科尔伯格模型，描述行业中有一家企业是主导企业或自然领导者的情形。例如，IBM 通常被认为是计算机行业的主导企业。人们经常可以观测到计算机行业的行为特征是，小企业在 IBM 宣布新产品的决策之后，相应调整自己的生产决策。在这种情形下，我们可以将 IBM 作为斯坦科尔伯格模型中的领导者，行业中的其他企业作为追随者。

现在我们开始分析这个理论模型的细节。假设企业 1 是领导者，它选择的产量为 y_1。企业 2 对此的反应为选择产量 y_2。每个企业指导市场的均衡价格取决于总产量。我们使用反需求函数 $p(Y)$ 将均衡价格表示为行业产量的函数，$Y = y_1 + y_2$。

领导者选择什么样的产量才能使自身的利润最大化？答案取决于领导者如何认为追随者对它的决策的反应。领导者很可能预期到：给定领导者自身的决策，追随者的目的也是利润最大化。领导者为了确定自己的产量决策，它必须考虑追随者的利润最大化问题。

追随者的问题

追随者的利润最大化问题为

$$\max_{y_2} \quad p(y_1 + y_2)y_2 - c_2(y_2).$$

追随者的利润取决于领导者选择的产量，但从追随者的角度看，领导者的产量是预先决定好了的，也就是说领导者的产量已经选择完毕，追随者只要将领导者的产量视为常数即可。

追随者所选的产量要能使得它自身的边际收入等于边际成本：

$$MR_2 = p(y_1 + y_2) + \frac{\Delta p}{\Delta y_2} y_2 = MC_2 .$$

边际收入的解释方法和以前一样。当追随者增加产量，新增加的收入是按市场价格计算的。但是，产量增加也会使价格下降 Δp，这使得原来销量的收入按新价格计算，因此销售收入又减少了一部分。因此，产量变动引起的收入变动是上述两种效应之和。

需要注意的是追随者利润最大化决策取决于领导者的决策。我们将这种关系表示为

$$y_2 = f_2(y_1).$$

这个函数表明，追随者的利润最大化的产量是领导者选择的产量的函数。这个函数称为**反应函数**（reaction function），因为它告诉我们追随者如何对领导者的产量决策作出反应。

为简单起见，假设需求函数为线性，我们要从该需求函数推导出反应曲线。假设反需求函数为 $p(y_1 + y_2) = a - b(y_1 + y_2)$. 假设成本为零，目的仍然是为了简单。

则企业 2 的利润函数为

$$\pi_2(y_1 + y_2) = [a - b(y_1 + y_2)]y_2$$

或

$$\pi_2(y_1+y_2)=ay_2-by_1y_2+by_2^2.$$

我们使用这个表达式在图 27.1 中画出**等利润线**（isoprofit lines）。每一条等利润线表示，在这条线上的 y_1 和 y_2 的所有组合产生的利润是相等的。也就是说，等利润线是由满足下式的所有点 (y_1,y_2) 组成的

$$ay_2-by_1y_2+by_2^2=\bar{\pi}_2.$$

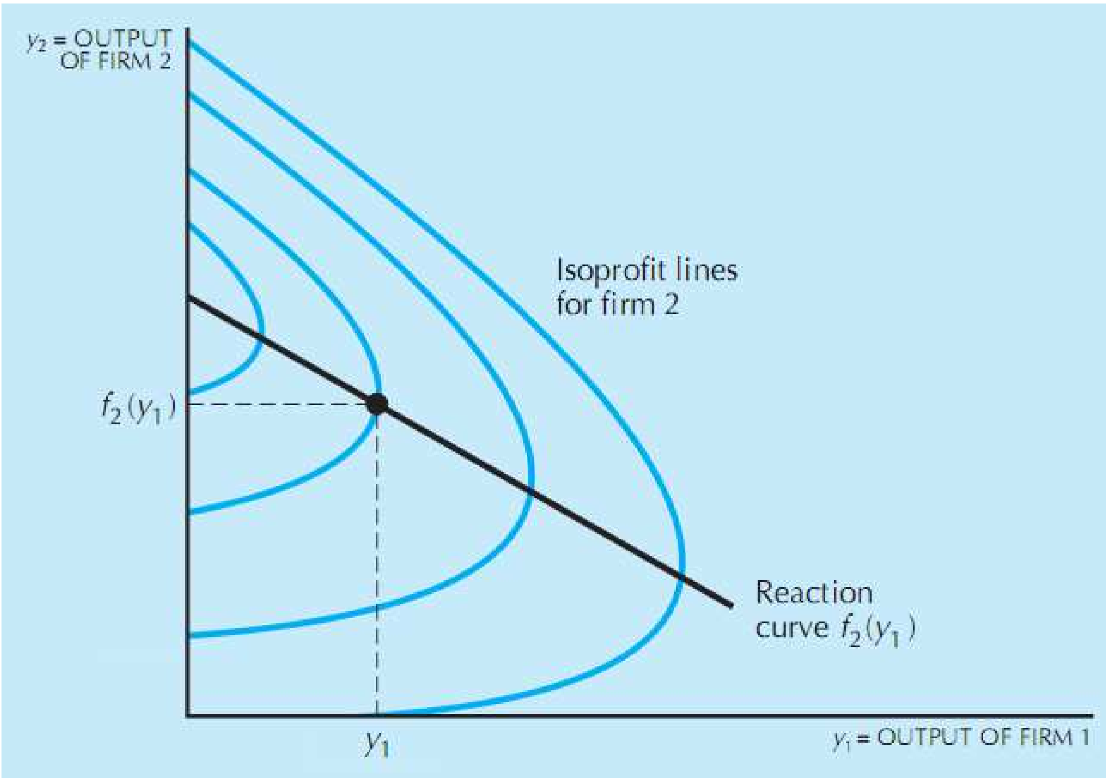


图 27.1：推导出反应曲线。反应曲线 $f_2(y_1)$ 表示，对于领导者（企业 1）选定的每个产量水平，追随者（企业 2）相应的利润最大化的产量为多少。对于 y_1 的每个选择，企业 2 选择的产量 $f_2(y_1)$ 应尽可能使它的等利润线尽可能地靠近左方。

注意，当我们从右方移向左方的等利润线时，企业 2 的利润会增加。这是因为如果我们把企业 2 的产量固定在某个水平上，随着企业 1 产量的减少，企业 2 的利润会增加。当企业 2 是垄断企业时，它会使它的利润尽可能地大；也就是说，当企业 1 选择的产量为零时，企业 2 的利润最大。

对于企业选择的每个产量水平，企业 2 都选择使利润尽可能大的产量。这表示，对于 y_1 的每个选择，企业 2 选择的产量应尽可能使它的等利润线尽可能地靠近左方。如图 27.1 所示。这一点满足相切条件：等利润线在最优选择处的切线必定是一条垂线。将这些切点连接

起来就得到了企业 2 的反应曲线 $f_2(y_1)$。

为了能使用代数的方法推知上面的结论，我们需要首先推出与企业 2 收入函数相伴的边际收入的表达式。由前面的企业 2 的收入函数(一)容易知道它的边际收入为

$$MR_2(y_1, y_2) = a - by_1 - 2by_2 .$$

（使用微积分不难推出这个式子，如果你没学过微积分，就相信这个式子吧。）由于前面我们假设成本为零，因此边际成本也为零。令边际收入等于边际成本（为零），可得

$$a - by_1 - 2by_2 = 0 ,$$

从该式可解出企业 2 的反应曲线：

$$y_2 = \frac{a - by_1}{2b} .$$

这条反应曲线为直线，如图 27.1 所示。

领导者的问题

我们已经分析了在给定领导者选择的产量水平的情形下，追随者如何选择自己的产量。现在我们转而分析领导者的利润最大化问题。

领导者很可能也意识到它的行为会影响到追随者的产量决策。我们可用反应函数 $f_2(y_1)$ 表示这种关系。因此，领导者在制定产量决策时，它应该认识到这一行为对追随者产生的影响。

因此，领导者的利润最大化问题为

$$\max_{y_1} \; p(y_1 + y_2)y_1 - c_1(y_1)$$

$$\text{使得 } y_2 = f_2(y_1) .$$

将第二个式子代入第一个式子可得

$$\max_{y_1} \; p\left(y_1 + f_2(y_1)\right)y_1 - c_1(y_1) .$$

注意，领导者会认识到当它选择产量 y_1 时，总产量将为 $y_1 + f_2(y_1)$：它自己的产量加上追随者的产量。

当领导者考虑改变产量时，它必须认识到这一决策会影响到追随者的产量决策。我们

(一) 在前面我们已经知道企业 2 的利润函数为 $\pi_2(y_1 + y_2) = ay_2 - by_1y_2 + by_2^2$，严格说来，这应是企业 2 的收入函数，只不过由于假设了成本为 0，才使得利润函数和收入函数的表达式相同。译者注。

使用前面的线性需求曲线分析这一问题。在前面我们已经知道反应函数为

$$f_2(y_1)=y_2=\frac{a-by_1}{2b}. \tag{27.1}$$

由于我们已假设边际成本为零，领导者的利润为

$$\pi_1(y_1+y_2)=p(y_1+y_2)y_1=ay_1-by_1^2-by_1y_2. \tag{27.2}$$

但是，追随者的产量 y_2，取决于领导者的产量决策，这可由反应函数 $y_2=f_2(y_1)$ 表示。

将（27.1）式代入（27.2）式可得

$$\begin{aligned}\pi_1(y_1,y_2)&=ay_1-by_1^2-by_1f_2(y_1)\\&=ay_1-by_1^2-by_1\frac{a-by_1}{2b}.\end{aligned}$$

化简这个表达式可得

$$\pi_1(y_1,y_2)=\frac{a}{2}y_1-\frac{b}{2}y_1^2.$$

与该函数相伴的边际收入为

$$MR=\frac{a}{2}-by_1.$$

令边际收入等于边际成本（在本例中我们已假设成本为零），由此可以解出

$$y_1^*=\frac{a}{2b}.$$

为了找到追随者的产量，我们将 y_1^* 代入反应函数可得

$$y_2^*=\frac{a-by_1^*}{2b}=\frac{a}{4b}.$$

将这两个企业的产量相加可得到行业的产量为 $y_2^*+y_2^*=3a/4b$.

斯坦科尔伯格模型的解可用图 27.2 中的等利润线进行说明。（我们还将用这个图分析 27.5 节中的古诺均衡。）在该图中，我们画出了两个企业的反应曲线以及企业 1 的等利润线。企业 1 等利润线的形状和企业 2 的类似；只是旋转了 90 度。企业 1 等利润线的位置越低，代表的利润越大，因为企业 2 产量下降时企业 1 的利润会增加。

企业 2 表现出的行为是追随者的行为，这表示它在它的反应曲线 $f_2(y_1)$ 上选择产量。因此，企业 1 在它自己的反应曲线上选择的产量，应能使得自己的利润尽可能地大。但是，

要注意，最大的利润位于它的反应曲线和**位置最低**的等利润线接触之处。如图 27.2 所示。由利润最大化的逻辑可知，反应曲线和这条等利润线必定在该点相切。

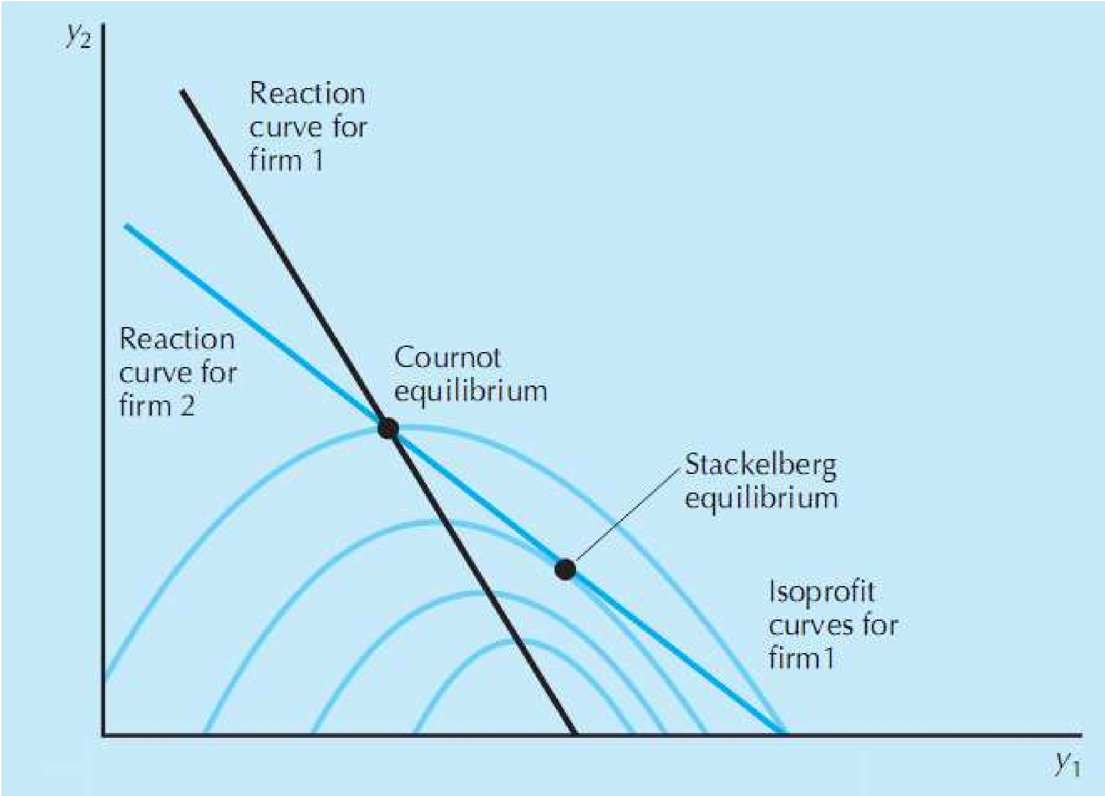


图 27.2：斯坦科尔伯格均衡。企业 1（领导者）选择的产量，位于企业 2 的反应曲线和企业 1 最低可能等利润线（代表利润尽可能高）相切之处，这一点能使企业 1 的利润尽可能地大。

27.3 价格领导

领导者也可以设定产量而是设定价格。为了合理决定如何设定价格，领导者必须预测追随者怎样反应。因此，我们必须首先分析追随者的利润最大化问题。

我们注意到，在均衡时，追随者设定的价格必定总是等于领导者设定的价格。这是因为我们假设两个企业销售的产品是相同的。如果其中一个企业索要的价格低于另外一个企业，那么所有消费者将偏好前者的产品，这样一来我们就无法得到两个企业销量都为正的均衡。

假设领导者将价格设定为 p 。我们假设追随者将该价格视为给定的，并相应选择利润最大化的产量。在本质上这和前面章节分析的竞争行为是一样的。在竞争模型中，每个企业都认为价格不受自己控制，因为它占市场份额非常小；在价格领导模型中，追随者也认为价格不受自己控制，因为价格已由领导者制定好。

追随者的利润最大化问题为

$$\max_{y_2} \quad py_2 - c_2(y_2).$$

这将导致我们已熟悉的条件，即追随者选择的产量要时价格等于边际成本。这个条件决定了追随者的供给曲线$S(p)$，我们在图 27.3 中画出了这条供给曲线。

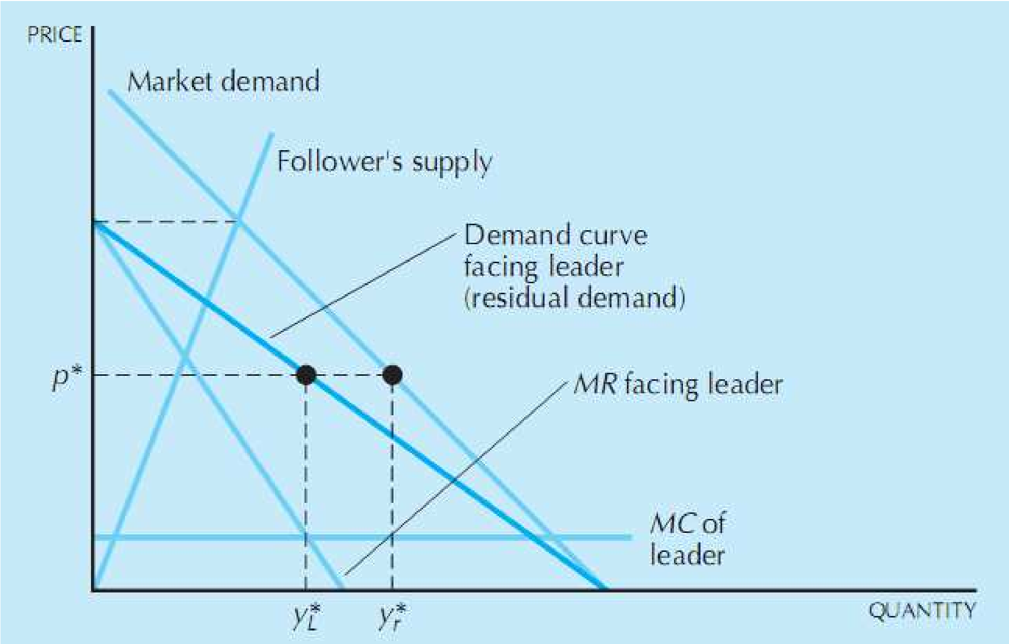


图 27.3：价格领导。领导者面对的需求曲线是市场需求曲线减去追随者的供给曲线。领导者根据边际收入等于边际成本这一条件找到最优供给量y_L^*，其中下标 L 表示领导者。市场的总给量为y_T^*，市场均衡价格为p^*。

现在开始分析领导者面对的问题。领导者认识到如果它设定了一个价格水平p，追随者的供给将为$S(p)$.这表示领导者的销量为$R(p)=D(p)-S(p)$.这条曲线称为领导者面对的**剩余需求曲线**（residual demand curve）。

假设领导者的边际成本恒为常数c，则对于任意的价格水平p，它可以实现的利润为

$$\pi_1(p)=(p-c)[D(p)-S(p)]=(p-c)R(p).$$

为了实现最大化利润，领导者选择的价格和产量组合，应满足边际收入等于边际成本这个条件。注意，它的边际收入应该是与**剩余**需求曲线相伴的边际收入，剩余需求曲线真正代表了在每个给定的价格水平上，领导者应该选择的产量水平。在图 27.3 中，剩余需求曲线是线性的；因此，与剩余需求曲线相伴的边际收入曲线与剩余需求曲线的纵截距相同，但斜率为剩余需求曲线的两倍。

我们举一个简单的代数例子说明。假设反需求曲线为$D(p)=a-bp$，追随者的成本函数为$c_2(y_2)=y_2^2/2$，领导者的成本函数为$c_1(y_1)=cy_1$.

对于任何价格水平 p，追随者在价格等于边际成本之处生产。若它的成本函数为 $c_2(y_2)=y_2^2/2$，则相应的边际成本曲线为 $MC_2(y_2)=y_2$。令价格等于边际成本可得

$$p=y_2.$$

由此可得出追随者的供给曲线 $y_2=S(p)=p$。

领导者面对的需求曲线，即剩余需求曲线因此为

$$R(p)=D(p)-S(p)=a-bp-p=a-(b+1)p.$$

从现在开始，这个问题类似于一个普通的垄断问题。从上式解出 p，它是领导者的产量 y_1 的函数，可得：

$$p=\frac{a}{b+1}-\frac{1}{b+1}y_1. \tag{27.3}$$

这是领导者面对的反需求函数，与该函数相伴的边际收入曲线的纵截距，与该反需求函数相同，但是斜率是反需求函数的 2 倍。这意味着边际收入曲线的表达式为

$$MR_1=\frac{a}{b+1}-\frac{2}{b+1}y_1.$$

令边际收入等于边际成本，可得

$$MR_1=\frac{a}{b+1}-\frac{2}{b+1}y_1=c=MC_1.$$

从该式可解出领导者利润最大化的产量

$$y_1^*=\frac{a-c(b+1)}{2}.$$

我们可以将该式代入（27.3）式解出均衡价格，但是这个解不是特别有趣，因为这个解本身没任何显著的特征，只是为了求解而求解。

27.4 价格领导和产量领导的比较

我们已经知道如何计算数量领导和价格领导情形下的均衡价格和均衡产量。每种模型决定了一种不同的均衡价格和产出组合；每种模型都有自己的适用环境。

企业制定产量决策的一种理解方式，是将其视为企业选择生产能力。当企业定产时，它实际上是在决定应向市场提供多少产品。如果一家企业能够率先进行生产能力投资，自然地我们可以将其作为产量领导者。

另一方面，假设在我们分析的市场中，生产能力的选择并不重要，但是其中一家企业在分发价目表。自然可以将该企业视为价格制定者。它的竞争对手可能将价目表中的价格最为给定的，因此相应作出自己的价格和供给决策。

是价格领导模型还是产量领导模型更合适一些，这个问题单纯从纯理论角度无法给出明确答案。我们必须观察企业实际是怎样决策的，这样才能选择最合适的模型。

27.5 同时决定产量

领导者-追随者模型的一个缺陷是它必须是非对称的(asymmetric)：一个企业比另外一个企业率先作出决策。在某些情形下，这是不合理的。例如，假设两个企业试图同时决定产量。在该情形下，为了制定合理决策，企业间必须互相预测对方的产量决策。

在本节我们将分析一个只有一个时期而不是多时期的模型，在这个模型中，每个企业必须预测对方选择的产量。给定它对其他企业产量的预测，每个企业选择自己的利润最大化产量。于是我们可以求解预测产量的均衡，即在均衡时，每个企业发现它对对方的预期是正确的。这种模型称为**古诺模型**（Cournot model），这是为了纪念 19 世纪法国数学家古诺，是他首先分析了这个问题[一]。

假设一开始企业 1 预期企业 2 的产量为 y_2^e（上标 e 表示预期）。如果企业 1 决定生产 y_1 单位产品，则它可以预期到总产量为 $Y = y_1 + y_2^e$，这个总产量将使市场价格为 $p(Y) = p(y_1 + y_2^e)$. 企业 1 的利润最大化问题为

$$\max_{y_1} \quad p(y_1 + y_2^e)y_1 - c(y_1).$$

对于企业 2 产量的任何一个预测 y_2^e，企业 1 都会相应作出产量的最优选择 y_1。我们将企业 2 的预期产量与企业 1 的最优产量选择之间的关系，用下式表示

$$y_1 = f(y_2^e).$$

这个函数就是我们前面分析过的反应函数。在前面，反应函数将追随者的产量看成领导者产量选择的函数。而在这里，反应函数将一个企业的最优产量决策看成它对另外一个企业决策**信念**（belief）的函数。由此可知，尽管在不同情形下，反应函数的解释不同，但是它们的数学定义是完全相同的。

类似地，我们可以推出企业 2 的反应函数:

$$y_2 = f(y_1^e).$$

这个式子是说，给定企业 2 对企业 1 产量 y_1^e 的预期，企业 2 选择的最优产量是多少。

现在，我们知道每个企业在选择自己的产量水平时，都是**假设**对方的产量为 y_1^e 或 y_2^e 的。但是要注意，y_1^e 或 y_2^e 只是预期值，一般来说它们和两企业的实际最优产量是不相等的，即 $y_1 \neq y_1^e$，$y_2 \neq y_2^e$。

[一] Augustin Cournot was born in 1801.　His book, Researches into the Mathematical Principles of the Theory of Wealth, was published in 1838.

我们最终想找到一个产量组合 (y_1^*, y_2^*) 使得：企业 1 在假定企业 2 的产量为 y_2^* 时，它选择最优产量就是 y_1^*；而且企业 2 在假定企业 1 的产量为 y_1^* 时，它选择的最优产量就是 y_2^*。也即是说，产量组合 (y_1^*, y_2^*) 满足

$$y_1^* = f_1(y_2^*)$$

$$y_2^* = f_2(y_1^*).$$

这样的一组产量组合叫做一个**古诺均衡**（Cournot equilibrium）。在古诺均衡中，给定每个企业对另外一个企业的预期产量，每个企业都实现了利润最大化；而且，这些预期产量在均衡时恰好就是实际最优产量：每个企业实际选择的最优产量就是另外一个企业对它估计的预期产量。在古诺均衡中，一旦企业发现另外一个企业实际选择的产量，它就不会改变自己的产量，因为此时改变产量已无法使利润更大，这正是均衡的含义。

举例说明。图 27.2 就存在着一个古诺均衡，这个古诺均衡就是两条反应曲线相交时对应的产量组合。在交点处，给定另外一个企业的实际选择，每个企业都找到了能使利润最大化的最优产量。

27.6 古诺均衡的一个例子

我们仍用前文的那个例子说明，即需求函数为线性而且边际成本为零。在这种情形下，我们知道企业 2 的反应函数为

$$y_2 = \frac{a - by_1^e}{2b}.$$

由于在这个例子中企业 1 和企业 2 是完全一样的，它的反应函数和企业 1 的形式是一样的：

$$y_1 = \frac{a - by_2^e}{2b}.$$

图 27.4 画出了这两条反应曲线。它们的交点就是古诺均衡。在交点处，给定企业对另外一个企业行为的预期，每个企业都作出了利润最大化的选择；而且每个企业**实际**选择的产量，恰好就是另外一个企业对它估计的预期产量。

为了使用代数方法计算出古诺均衡，我们找到点 (y_1, y_2)，在该点上，每个企业实际选择的产量正好就是另外一个企业对它产量的预期。令 $y_1 = y_1^e$ 以及 $y_2 = y_2^e$，这样我们就得到了下面的两元一次方程组：

$$y_1 = \frac{a - by_2}{2b}$$

$$y_2 = \frac{a - by_1}{2b}.$$

在这个例子中，两个企业是完全一样的，因此，在均衡时每个企业的产量是相同的。所以，我们可以将 $y_1 = y_2$ 代入上面的任何一条反应曲线，从而得到

$$y_1 = \frac{a - by_1}{2b}.$$

解出 y_1^*，可得

$$y_1^* = \frac{a}{3b}.$$

由于这两个企业是完全一样的，这意味着

$$y_2^* = \frac{a}{3b}.$$

因此，行业的总产量为

$$y_1^* + y_2^* = \frac{2a}{3b}.$$

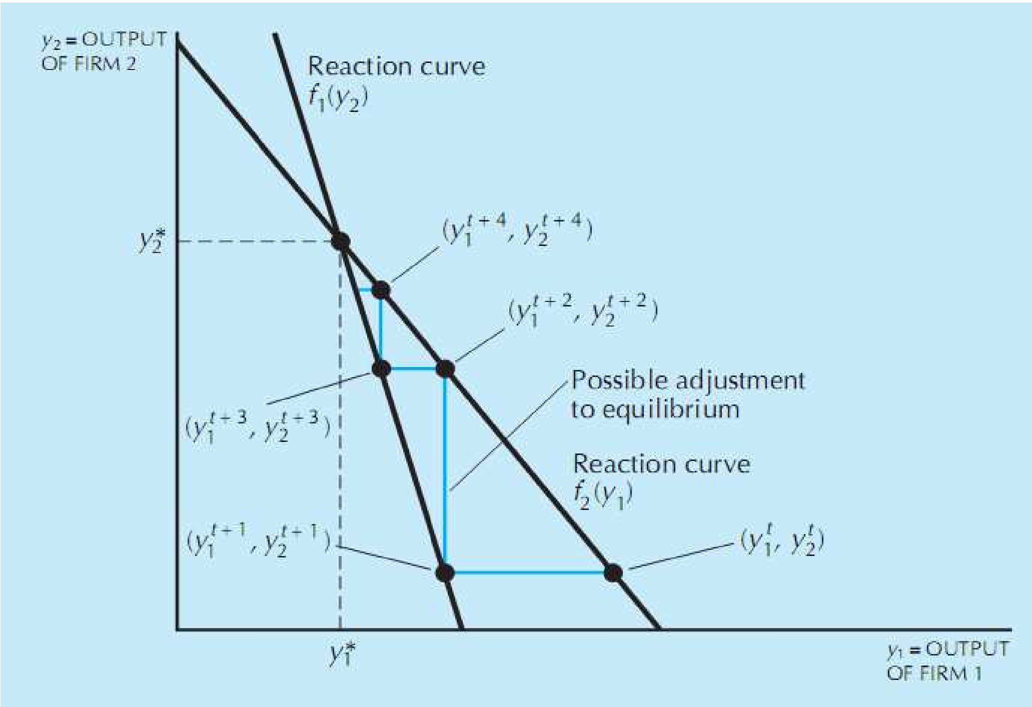


图 27.4：古诺均衡。古诺均衡时，给定企业对另外一个企业的产量预期，每个企业都实现了利润最大化。古诺均衡为 (y_1^*, y_2^*)，这正好是两条反应曲线的交点之处。

27.7 实现均衡的过程

我们可以用图 27.4 分析为了实现均衡，企业产量的调整过程。假设在时期 t 企业的产量为 (y_1^t, y_2^t)，这个产量组合未必是均衡产量。若企业 1 预期企业 2 的产量将继续维持在 y_2^t，

则在下一期，在企业 2 的预期产量为 y_2^t 时，企业 1 选择的利润最大化产量为 $f_1(y_2^t)$. 因此，企业 1 在 $t+1$ 期的产量选择为

$$y_1^{t+1} = f_1(y_2^t).$$

企业 2 当然可以进行类似推理，因此企业 2 在 $t+1$ 期的选择为

$$y_2^{t+1} = f_2(y_1^t).$$

这两个函数表明，当面对一个企业的产量选择时，另外一个企业如何调整自己的产量。图 27.4 给出了企业根据自己的反应曲线相应调整产量的过程。我们可以下面的方法解读这个图。假设我们从点 (y_1^t, y_2^t) 开始分析。给定企业 2 的产量水平，企业 1 的最优选择是在下一期生产 $y_1^{t+1} = f_1(y_2^t)$ 单位的产品。在图形上，这表现为企业 1 的产量水平向左移动，直到它到达它自己的反应曲线上。

若企业 2 预期企业 1 的产量继续维持在 y_1^{t+1} 的水平上，则它的最优反应是生产 y_2^{t+1} 单位产品。我们将企业 2 的产量向上方垂直移动，直到到达企业 2 的反应曲线，我们就找到了这一点。沿着图中的“阶梯”继续移动，就可以得到这两个企业的一系列产量选择。在本例中，这种调整过程向古诺均衡收敛。我们说在这种情形下的古诺均衡是一个**稳定均衡**（stable equilibrium）。

尽管这个调整过程在直觉上让人满意，但是它存在着一些缺陷。每个企业在制定产量决策时都假定对方的产量在本期和下一期是固定不变的，但是我们已经看到，这两个企业的产量都是不断变化的。只有在均衡状态下，企业对另外一家企业产量的预期才是正确的。正是由于这个原因，我们通常不考虑均衡是如何实现的这个问题，而只关注均衡时企业的行为。

27.8 多个企业的古诺均衡

现在我们假设古诺均衡的问题涉及多个企业，而不再是两个企业。在这种情形下，我们假设每个企业在制定产量决策时都必须预测行业中其他企业的产量决策，我们最终目的是找到均衡时各个企业的产量。

假设行业中有 n 个企业，令 $Y = y_1 + ... + y_n$ 表示行业总产量。则企业 i 的“边际收入等于边际成本的条件”为

$$p(Y) + \frac{\Delta p}{\Delta Y} y_i = MC(y_i).$$

对于上式的左端，我们将 $p(Y)$ 提取出来，并将左端的第二项乘以 Y/Y，可将上式变形为

$$p(Y)[1 + \frac{\Delta p}{\Delta Y}\frac{Y}{p(Y)}\frac{y_i}{Y}] = MC(y_i).$$

使用总需求曲线的弹性定义，并且令 $s_i = y_i / Y$ 表示企业 i 的产量占市场总产量的份额，上

式可以简化为

$$p(Y)[1-\frac{s_i}{|\varepsilon(Y)|}]=MC(y_i). \qquad (27.4)$$

我们也可以将这个表达式写为

$$p(Y)[1-\frac{1}{|\varepsilon(Y)|/s_i}]=MC(y_i).$$

除了多出了 s_i 一项外，这个式子很像垄断企业的利润最大化条件的表达式。我们可以将 $\varepsilon(Y)/s_i$ 看成企业 i 面对的需求曲线的弹性：它的市场份额越小，它面对的需求曲线的弹性越大。

如果它的市场份额为 1，即该企业是垄断企业，则它面对的需求曲线就是市场需求曲线，因此上述表达式简化为垄断企业利润最大化的最优条件。如果该企业的产量占市场总产量很小，它的市场份额基本等于零，那么它面对的需求曲线是水平的。因此，上述表达式简化为完全竞争者的利润最大化的条件，即价格等于边际成本。

这说明第 22 章介绍的竞争模型是合理的。如果行业中的企业数量众多，那么每个企业对市场价格的影响基本上可以忽略不计，这种情形下的古诺均衡在本质上就是完全竞争下的均衡。

27.9 同时制定价格

在上面的古诺模型中，我们假设企业选择的是产量，而让市场决定价格。另外一种方法是，可以认为企业制定价格，而让市场决定销量。这种模型称为**伯特兰竞争**（Bertrand competition）[一]。

当一个企业在选择价格时，它必须预测行业中其他企业制定的价格。和古诺均衡的情形一样，我们也想找到一个价格组合，使得在给定另外一个企业的决策时，这个价格组合就是企业的利润最大化的价格。

伯特兰均衡是什么样子的？当企业销售的产品是同质的，伯特兰均衡的结构非常简单。可以证明它是一个竞争均衡，即均衡时价格等于边际成本！下面我们证明此事。

首先可以注意到价格不可能小于边际成本，因为如果这样，任何一个企业都会减少产量从而增加利润。所以我们只需要考察价格大于边际成本的情形。假设两个企业的售价 $\hat{p}$ 大于边际成本。先分析企业 1 的处境。如果它稍微将价格降低 ε 但另外一个企业的价格仍为 $\hat{p}$，那么所有消费者都会更喜欢企业 1。企业 1 只是稍微降低了一点点价格，就可以夺取企业 2 的所有消费者。

[一] 约瑟夫.伯特兰（Joseph Bertrand），和古诺一样也是一位法国数学家，他在评价古诺模型的时提出了这一模型。

如果企业 1 坚信企业 2 索要的价格 $\hat{p}$（该价格大于边际成本），那么企业 1 就会将价格降低为 $\hat{p}-\varepsilon$。但是企业 2 也会这么推理！因此高于边际成本的任何价格水平都不可能是均衡价格；唯一的均衡就是竞争均衡。

当你第一次遇到这个结论的时候，你会觉得不可思议：如果市场上只有两个企业，这明显是寡头垄断，在这种情形下，我们怎么可能得的是均衡价格？如果我们将伯特兰看成竞价竞争模型，则它就合理了。假设一个企业在竞争消费者的业务中报出的价格高于边际成本。那么，另外一个企业总可以报出稍微低一些的价格进行竞争。由此可以推知，只有价格等于边际成本时，每个企业都不会再降价。

人们通常可以观察到，如果企业不能合谋，那么它们的竞争报价导致的价格，通常远远低于其他模型导致的价格。这种现象是伯特兰竞争模型的必然结果。

27.10 合谋

在前面各节介绍的模型中，我们都假设企业是独立作出决策的。但是如果企业合谋以便联合确定它们的产量，前面的这些模型不再合理。如果允许合谋，企业能挣得更多的利润：选择能使行业利润最大的产量，然后再把这些利润分配给参与合谋的成员。如果企业联合起来企图一起确定价格和产量以实现行业的利润最大，那么我们将这种组织称为一个**卡特尔**（cartel）。我们在第 24 章已指出，卡特尔是指一伙企业合谋，卡特尔的行为和一个完全垄断的企业是一样的，目的是实现卡特尔的利润（即成员利润之和）最大化。

因此，合谋的两个企业面对的利润最大化问题是选择产量 y_1 和 y_2，使得行业利润最大：

$$\max_{y_1, y_2} p(y_1+y_2)[y_1+y_2]-c_1(y_1)-c_2(y_2).$$

这一问题的最优条件为：

$$p(y_1^*+y_2^*)+\frac{\Delta p}{\Delta Y}[y_1^*+y_2^*]=MC_1(y_1^*)$$

$$p(y_1^*+y_2^*)+\frac{\Delta p}{\Delta Y}[y_1^*+y_2^*]=MC_2(y_2^*).$$

这些条件的解释是比较有趣的。当企业 1 考虑是否将产量增加 Δy_1 时，它需要考虑两种效应：由于销售更多产品带来的额外利润；由于价格下降导致的利润下降。但在合谋的情形下，在第二个效应中，它必须考虑价格降低对它自身产量的影响，还要考虑对另外一个企业产量的影响。因为在合谋的情形下，它关注的是行业利润最大，而不是自己的利润最大。

最优条件意味着，不论额外一单位产品由合谋的哪一个企业生产，它带来的边际收入必须相等。由此可推知，$MC_1(y_1^*)=MC_2(y_2^*)$，这就是说均衡时，两个企业的边际成本必须相等。如果其中一个企业具有成本优势，那么它的边际成本曲线必然位于另外一个企业边际成本曲线的下方，在均衡的卡特尔解中，这家具有成本优势的企业必然负责生产更多产量。

现实中卡特尔存在的问题是，卡特尔的成员有背叛的诱惑。例如，假设两个合谋的企业一开始生产的产量 (y_1^*, y_2^*) 能实现行业的利润最大，企业 1 考虑多生产 Δy_1 单位的产量。企业 1 由于多生产带来的额外利润为

$$\frac{\Delta\pi_1}{\Delta y_1} = p(y_1^* + y_2^*) + \frac{\Delta p}{\Delta Y} y_1^* - MC_1(y_1^*). \tag{27.5}$$

前面我们已经知道卡特尔解的最优条件为

$$p(y_1^* + y_2^*) + \frac{\Delta p}{\Delta Y} y_1^* + \frac{\Delta p}{\Delta Y} y_2^* - MC_1(y_1^*) = 0.$$

将这个式子重新整理可得

$$p(y_1^* + y_2^*) + \frac{\Delta p}{\Delta Y} y_1^* - MC_1(y_1^*) = -\frac{\Delta p}{\Delta Y} y_2^* > 0. \tag{27.6}$$

上式大于零是因为 $\Delta p / \Delta Y$ 为负，而这又是由于市场需求曲线的斜率为负。

对比（27.5）式和（27.6）式可知

$$\frac{\Delta\pi_1}{\Delta y_1} > 0.$$

因此，如果企业 1 坚信企业 2 会维持产量不变，那么它将认为增加自己的产量会增加自己的利润。在卡特尔解中，企业联合起来限制总产量以便不“破坏”市场。它们知道任何一个企业如果多上产将会对总利润带来什么样的影响。但是，如果每个企业坚信其它企业会维持约定的产量水平，那么每个企业都有单方面增加产量的诱惑，因为这样可以增加自己的利润。因此，在能使联合利润最大的产量水平上，每个企业如果单方面增加产量都会增加自己的利润，当然前提是每个企业预期其它企业会维持产量不变。

情形甚至可能更糟。如果企业 1 相信企业 2 会维持产量不变，因此它会增加产量以增加自己的利润。但是如果它相信企业 2 会增加产量，那么企业 1 会率先增加产量，使自己利润更大。

因此，为了维持有效的卡特尔，这些联合起来的企业需要找到能发现和惩罚欺骗的方法。如果它们无法观测到其它企业的产量，那么背叛的诱惑会导致卡特尔瓦解。我们稍后分析这一问题。

为了确保你理解了卡特尔的解，我们举个例子计算一下。我们使用前面古诺模型中引入的例子，即边际成本为零，需求函数为线性的情形。

行业总利润为

$$\pi(y_1, y_2) = [a - b(y_1 + y_2)](y_1 + y_2) = a(y_1 + y_2) - b(y_1 + y_2)^2,$$

因此，边际收入等于边际成本的条件为

$$a - 2b(y_1^* + y_2^*) = 0,$$

这意味着

$$y_1^* + y_2^* = \frac{a}{2b}.$$

由于边际成本为零，所以总产量可以在这两个企业间任意分配。因此，真正重要的是行业的总产量水平。

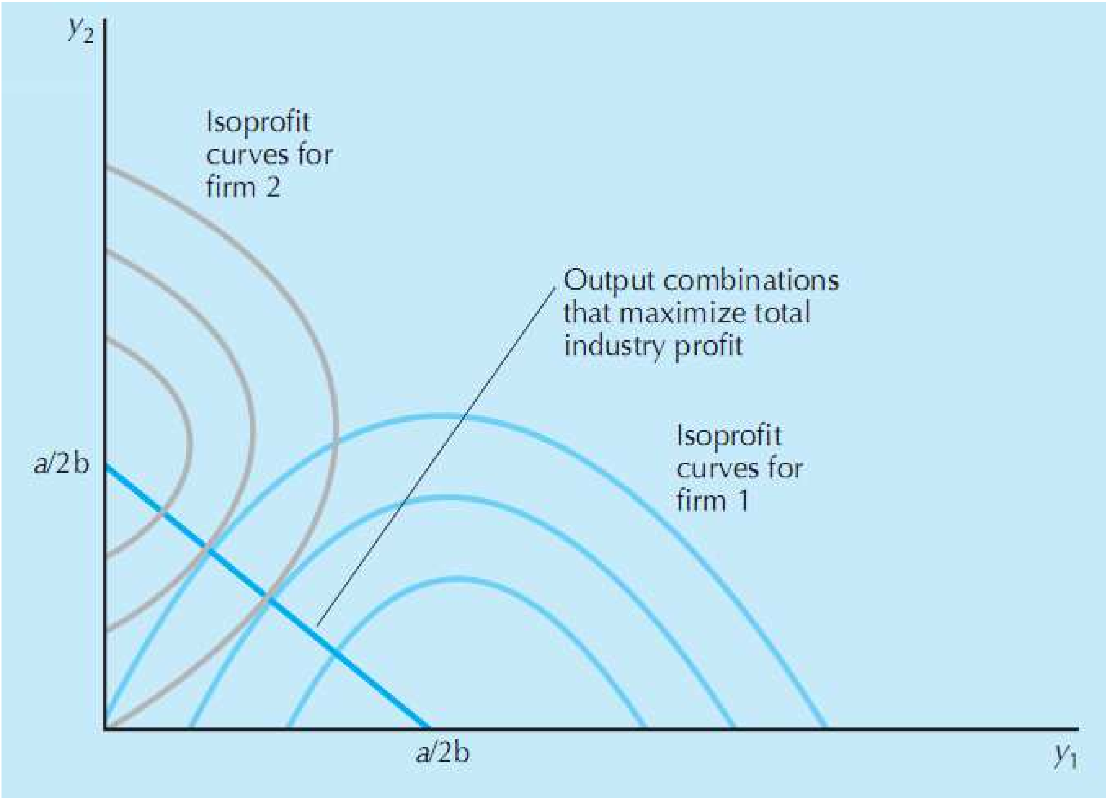


图 27.5：一个卡特尔。如果行业的利润已实现最大，那么此时每个企业的边际利润必然相等。这意味着在利润最大化的产量上，它们的等利润线必定相切。

我们可用图 27.5 说明这个解。在该图中，我们画出了每个企业的等利润线，而且重点强调了这两条等利润线的公共切点的轨迹——公共切线。为什么强调这条公共切线？由于卡特尔试图使得行业利润最大，由此可推知，任何一个企业在均衡时的边际利润必然相等。因为如果一个企业的边际利润更大，则若由该企业生产更多的产量，行业利润会更大。这反过来意味着每个企业的等利润线在均衡产量处的斜率是相等的；也就是说均衡时，一个企业的等利润线必然与另外一个企业的等利润线相切。因此，能使行业利润最大的产量组合——卡特尔解——必然位于这条公共切线上，如图 27.5 所示。

图 27.5 也说明了在卡特尔解之处存在着欺骗的诱惑。例如，假设两个企业一开始平分市场。现在想想如果企业 1 认为企业 2 会维持产量不变，将会产生什么样的结果？如果企业

增加产量但企业 2 维持产量不变，那么企业 1 将会移向位置更低的等利润线，由于对企业 1 来说，它的等利润线位置越低，代表的利润越大，这意味着企业 1 的利润会增加。这一点我们前面已使用代数方法说明过。如果一个企业认为另一个企业将维持产量不变，那么它就有增加产量的诱惑，因为这样做可以增加它自己的利润。

27.11 惩罚欺骗的策略

我们已经知道卡特尔是非常不稳定的，因为在行业利润最大的产量上，如果每个企业偷偷地继续增加产量，能使它自己的利润更大。如果卡特尔想成功运行，它必须找到能“稳定”企业行为的方法。一种方法是制定惩罚措施，即如果有企业背叛卡特尔协定，那么其他企业将惩罚这个背叛的企业。在本节，我们分析惩罚力度为多大时才能稳定卡特尔。

我们以由完全相同的两个企业形成的双头垄断行业为例进行说明。如果每个企业生产的产量是该垄断行业产量的一般，则总利润最大，每个企业得到的利润为 π_m。为了使这个结果稳定，一个企业对另外一个企业宣称：“如果你将产量维持在能使行业利润最大的产量水平，这没问题。但是如果我发现你的产量超过了这个约定的水平，我就会惩罚你——我会永远按照古诺产量水平生产。”这就是所谓的**惩罚策略**（punishment strategy）。

这类威胁何时才能足以稳定卡特尔？为了回答这个问题，我们必须分析，与合作相比，期盼的收益和成本。假设欺骗已发生，惩罚措施已实施。由于对古诺行为的最优反应仍是古诺行为（根据定义），这将导致每个企业每期的利润为 π_c。当然，古诺利润 π_c 小于科特尔情形下该企业的利润 π_m。

假设这两个企业一开始的产量是在合谋、垄断的产量水平。假若你是其中一个企业，你正在决定是否继续按照约定的份额生产。如果你生产的产量大于约定的份额，那么你得到的利润为 π_d，显然有 $\pi_d > \pi_m$。这是摆在卡特尔成员面前的标准诱惑，我们在前面已经分析过这一点：如果每个企业限制产量从而提升价格，则每个企业都有利用该较高价格的动机，通过增加产量就能增加自己的利润。

当然你可以增加产量，但是别忘了还有惩罚措施。如果按照约定生产卡特尔的产量，每个企业得到的稳定利润流为 π_m。从现在算起，这个利润流的现值为

$$\text{卡特尔行为的现值} = \pi_m + \frac{\pi_m}{r}.$$

如果企业生产的产量大于卡特尔约定中的它的产量，它可以得到利润 π_d，但只有一次。从此以后它必须接受卡特尔瓦解并且转向古诺行为的事实，这种情形下，它的利润流的现值为

$$\text{欺骗行为的现值} = \pi_d + \frac{\pi_c}{r}.$$

何时卡特尔行为的现值大于欺骗行为的现值？显然为下式成立时：

$$\pi_m + \frac{\pi_m}{r} > \pi_d + \frac{\pi_c}{r},$$

这个条件也可以写为

$$r < \frac{\pi_m - \pi_c}{\pi_d - \pi_m}.$$

注意，上式右端的分子为正，因为垄断利润大于古诺利润；分母也为正，因为欺骗时得到的利润大于垄断利润。

这个不等式是说，只要利率足够小，因此它预期将来的惩罚足够大，所以它会遵守卡特尔的约定，按份额进行生产。

这个模型的缺陷是，宣称永远实施古诺产量的这种威胁，企业并不相信。企业可能会相信其它企业会对它的欺骗行为进行惩罚，但宣称“永远”就不大让人信服了。一个更现实的模型是考虑较短时期内的惩罚，但是这样做会使模型更复杂。在下一章，我们将讨论某些“重复博弈”的模型，这些模型说明了一些可能的行为。

例子：价格匹配和价格竞争

我们已经知道，卡特尔的每个成员都有多生产的冲动。为了维持卡特尔成功运行，必须找到一些方法监视成员的行为，如果某个成员的行为偏离了卡特尔利润最大的产量，则需要对它进行惩罚。特别地，这意味着这些企业必须能够跟踪查找其他成员的价格和产量水平。

获得行业内其他企业价格信息的一种方法，是使用你的消费者监视其他企业。通常我们可以看到零售企业会宣称它们的价格是全城“最低价”。在某些情形下，这样的报价意味着该零售业竞争激烈。但在另外一些情形下，你也可以承诺“全城最低价”，通过这种方法就能收集其他企业的价格信息，以便维持卡特尔的稳定。

例如，假设两家企业约定，不管是明确约定还是默契地“约定”，对某类型冰箱的售价为 700 元。一家企业怎么发现另外一家企业背叛了约定，比如只卖 675 元？一种方法是你承诺全城最低价。如果你的价格不是最低的，消费者会向你报告，这样你就可以发现另外一个企业是否违背了合谋约定。

例子：自愿出口限制

1980 年代，日本汽车公司彼此约定实施“自愿出口限制”（voluntary export restraint，VER）。这就是说它们“自愿”减少对美国出口的汽车数量。一般美国人认为这代表美国在与日本贸易谈判中的取得了重大胜利。

但是如果你稍微想一下，你就会知道事情不是那回事。我们已经知道，寡头行业中企业面对的问题是，如何限制产量以提高价格和打击竞争。我们还知道，每个企业都有不遵守生

产协议的诱惑；每个卡特尔必须找到一种方法，去发现和阻止欺骗。如果第三方比如政府能充当侦探的角色，这对卡特尔来说无疑是种好事。美国政府就充当了日本汽车行业的侦探角色！

根据一项估计，1984 年美国进口的日本汽车比在自愿出口限制（VERs）实施之前贵了 2500 美元。而且，较高的汽车进口价格使得美国本土汽车公司的价格上涨了 1000 美元[一]。

由于汽车价格较高，1985~1986 年间，美国消费者为此多支付了大约 100 亿美元。这些钱直接流入日本汽车制造公司手里。日本汽车利用这些额外利润扩大生产能力，这使得这些公司能在以后年份以更低的成本生产新的汽车。自愿出口限制的确增加了美国居民的就业；然而，每增加一个就业，美国为此付出的代价大约为 16 万美元每年。

如果自愿出口限制的政策的目的，只是在于壮大美国汽车行业，那么有一种更为简单的方法：每进口一辆日本汽车，就征收 2500 元关税。这种方法的好处是，由于限制贸易而产生的收入落入美国政府的手中，而不是日本汽车制造行业。这样，1985~1986 年间流入日本的 100 亿美元就会为美国政府所用，它可以使用这些资金重新设计美国汽车行业方案，以便促进美国汽车行业长期健康发展。

27.12 各种解的比较

我们已分析了好几种寡头垄断行为的模型：产量领导（斯坦科尔伯格），价格领导，同时制定产量（古诺），同时制定价格（伯特兰）以及合谋。如何进行比较？

一般来说，合谋会导致最小的产量和最高的价格。伯特兰均衡，即竞争均衡的产量最高、价格最低。其他模型的结果介于这两个极端之间。

当然还存在其他类型的模型。例如，我们可以分析产品差异化的模型，即两个企业生产的产品不是完全替代的。或者，我们可以分析企业在较长时间内的一系列的选择问题。在这个模型中，一个企业在某时期做出的决策会影响另一个企业在下一期的决策。

在前面各节介绍的模型中，我们假设每个企业都知道行业中其他企业的需求函数和成本函数。在现实中，这些函数未必可知。每个企业在做决策时，必须估计竞争对手的需求函数和成本函数。经济学家已利用模型分析过这些现象，但是这些模型变得更为复杂。

总结

[一] Robert Crandall, “Import Quotas and the Automobile Industry: the Costs of Protectionism,” The Brookings Review, Summer, 1984.

1.寡头垄断行业的特征是市场中有一些企业但数量不是很多，而且这些企业认识到它们的策略彼此依赖。寡头企业的行为有多种方式，这取决于企业间互动的具体性质。

2.在产量领导（斯坦科尔伯格）模型中，一个企业率先作出产量决策而处于领导者的地位，其他企业追随。当领导者制定产量决策时，它必须考虑追随者如何作出反应。

3.在价格领导模型中，一个企业制定价格，其它企业按照该价格选择应出售多少单位的商品。价格领导者在制定决策时也必须考虑追随者的行为。

4.在古诺模型中，给定其它企业的选择，每个企业选择自己的产量使得它自己的利润最大化。在均衡时，每个企业发现它对某个企业预测的产量，就是这个企业的实际产量。

5.如果行业中企业数量众多，每个企业的市场份额很小，那么古诺均衡时的市场价格非常接近每个企业的边际成本，也就是说这个行业基本上是完全竞争的。

6.在伯特兰模型中，给定每个企业对另外企业选择价格的预期，它选择能使自己利润最大的价格。唯一的均衡价格是竞争均衡价格。

7.一个卡特尔由若干合谋的企业组成，它们合谋的目的是限制总产量以实现行业利润最大化。卡特尔通常是不稳定的，因为每个企业都有诱惑销售比约定销量更多的产品，当然前提是它相信其他企业不会做出反应。

复习题

1.假设有两个企业，这两个企业面对的需求函数为线性的，即 $p(Y)=a-bY$，再假设每个企业的边际生产成本都恒为常数 c。求古诺均衡的产量。

2.假设一个卡特尔中每个企业的边际成本都相等，而且恒为某个常数。如果该卡特尔已实现了行业利润的最大化，那么这意味着如何在这些企业中分配产量？

3.斯坦科尔伯格模型中的领导者能否获得比在古诺均衡时更低的利润？

4.假设有 n 个相同的企业实现了一个古诺均衡。证明市场需求曲线的弹性绝对值必定大于 $1/n$。（提示：在垄断的情形 $n=1$，这意味着垄断企业会在需求曲线具有弹性的那一段进行生产。使用我们证明垄断企业不会在需求缺乏弹性处生产的逻辑证明该问题。）

5.画出导致不稳定均衡的一组反应曲线。

6.寡头垄断企业能否生产帕累托有效率的产量？

曹乾（东南大学 caoqianseu@163.com）

复习题答案

1.假设有两个企业，这两个企业面对的需求函数为线性的，即 $p(Y)=a-bY$，再假设每个企业的边际生产成本都恒为常数 c。求古诺均衡的产量。

【复习内容】古诺均衡

我们只分析最简单的古诺模型，即只有一个时期而不是多时期的模型。在这个模型中，每个企业必须预测对方选择的产量。给定它对其他企业产量的预测，每个企业选择自己的利润最大化产量。于是我们可以求解预测产量的均衡，即在均衡时，每个企业发现它对对方的预期是正确的。

假设一开始企业 1 预期企业 2 的产量为 y_2^e（上标 e 表示预期）。如果企业 1 决定生产 y_1 单位产品，则它可以预期到总产量为 $Y=y_1+y_2^e$，这个总产量将使市场价格为 $p(Y)=p(y_1+y_2^e)$. 企业 1 的利润最大化问题为

$$\max_{y_1}\ p(y_1+y_2^e)y_1-c(y_1).$$

对于企业 2 产量的任何一个预测 y_2^e，企业 1 都会相应作出产量的最优选择 y_1。我们将企业 2 的预期产量与企业 1 的最优产量选择之间的关系，用下式表示

$$y_1=f(y_2^e).$$

这个函数就是我们前面分析过的反应函数。注意，这里的反应函数将一个企业的最优产量决策看成它对另外一个企业决策预期（belief）的函数。

类似地，我们可以推出企业 2 的反应函数:

$$y_2=f(y_1^e).$$

现在，我们知道每个企业在选择自己的产量水平时，都是**假设**对方的产量为 y_1^e 或 y_2^e 的。但是要注意，y_1^e 或 y_2^e 只是预期值，一般来说它们和两企业的实际最优产量是不相等的，即 $y_1\neq y_1^e$，$y_2\neq y_2^e$。

我们最终想找到一个产量组合 (y_1^*,y_2^*) 使得：企业 1 在假定企业 2 的产量为 y_2^* 时，它选择最优产量就是 y_1^*；而且企业 2 在假定企业 1 的产量为 y_1^* 时，它选择的最优产量就是 y_2^*。也即是说，产量组合 (y_1^*,y_2^*) 满足

$$y_1^*=f_1(y_2^*)$$

$$y_2^*=f_2(y_1^*).$$

这样的一组产量组合叫做一个古诺均衡。在古诺均衡中，给定每个企业对另外一个企业的预期产量，每个企业都实现了利润最大化；而且，这些预期产量在均衡时恰好就是实际最优产量：每个企业实际选择的最优产量就是另外一个企业对它估计的预期产量。

【参考答案】

假设在均衡时两个企业的产量分别为 y_1 和 y_2。

根据古诺均衡的两个特点（一是均衡时，每个企业选择的产量都使得自己的利润最大化；二是均衡时，每个企业实际选择的最优产量恰好等于另外一个企业对它估计的预期产量。）可知

$$y_1 = f_1(y_2)$$

$$y_2 = f_1(y_1).$$

这两个函数分别为企业 1 和 2 的反应函数，因此关键是求出反应函数。

企业 1 的利润最大化问题是

$$\max_{y_1} \quad p(y_1 + y_2)y_1 - c(y_1) = [a - b(y_1 + y_2)]y_1 - cy_1$$

这个最大化问题的一阶条件为

$$a - 2by_1 - by_2 - c = 0$$

整理可得企业 1 的反应函数

$$y_1 = f_1(y_2) = \frac{1}{2b}(a - c - by_2).$$

类似地可以求出企业 2 的反应函数

$$y_2 = f_1(y_1) = \frac{1}{2b}(a - c - by_1).$$

联立以上两个式子可得

$$y_1 = y_2 = \frac{a-c}{3b}.$$

因此，古诺均衡时每个行业的产量为 $y_1 = y_2 = \dfrac{a-c}{3b}$，行业产量为 $y_1 + y_2 = \dfrac{2(a-c)}{3b}$.

2.假设一个卡特尔中每个企业的边际成本都相等，而且恒为某个常数。如果该卡特尔已实现了行业利润的最大化，那么这意味着如何在这些企业中分配产量？

【复习内容】卡特尔的利润最大化决策

卡特尔的生产决策：选择能使行业利润最大的产量，然后再把这些利润分配给参与合谋的成员。卡特尔的行为和一个完全垄断的企业是一样的，目的是实现卡特尔的利润（即成员利润之和）最大化。

因此，合谋的两个企业面对的利润最大化问题是选择产量 y_1 和 y_2，使得行业利润最大：

$$\max_{y_1,y_2} p(y_1+y_2)[y_1+y_2]-c_1(y_1)-c_2(y_2).$$

这一问题的最优条件为：

$$p(y_1^*+y_2^*)+\frac{\Delta p}{\Delta Y}[y_1^*+y_2^*]=MC_1(y_1^*)$$

$$p(y_1^*+y_2^*)+\frac{\Delta p}{\Delta Y}[y_1^*+y_2^*]=MC_2(y_2^*).$$

这些条件的解释是比较有趣的。当企业 1 考虑是否将产量增加 Δy_1 时，它需要考虑两种效应：由于销售更多产品带来的额外利润；由于价格下降导致的利润下降。但在合谋的情形下，在第二个效应中，它必须考虑价格降低对它自身产量的影响，还要考虑对另外一个企业产量的影响。因为在合谋的情形下，它关注的是行业利润最大，而不是自己的利润最大。

最优条件意味着，不论额外一单位产品由合谋的哪一个企业生产，它带来的边际收入必须相等。由此可推知，$MC_1(y_1^*)=MC_2(y_2^*)$，这就是说均衡时，两个企业的边际成本必须相等。如果其中一个企业具有成本优势，那么它的边际成本曲线必然位于另外一个企业边际成本曲线的下方，在均衡的卡特尔解中，这家具有成本优势的企业必然负责生产更多产量。

【参考答案】

我们可以直接使用卡特尔均衡时的一个结论，以两个企业组成的卡特尔为例，在均衡产量 (y_1,y_2) 处，必有 $MC_1(y_1)=MC_2(y_2)$。否则，必然有 $MC_1(y_1)>MC_2(y_2)$ 或 $MC_1(y_1)<MC_2(y_2)$。前一个不等式意味着企业 2 具有成本优势，因此，让企业 2 多生产，利润会进一步增加，这和均衡时利润已实现最大相矛盾。

因此，卡特尔均衡时，必然有 $MC_1(y_1)=MC_2(y_2)$。

由于题目告知，形成卡特尔的这两个企业的边际成本是相等的，而且恒等于某常数。这意味着在均衡时，总产量可以任意在这两个企业间进行分配。因为无论怎样分配，都可以满足 $MC_1(y_1)=MC_2(y_2)$。

3.斯坦科尔伯格模型中的领导者能否获得比在古诺均衡时更低的利润？

【复习内容】斯坦科尔伯格模型；古诺模型

【参考答案】

在斯坦科尔伯格模型（产量领导模型）中，领导者和追随者的地位是不对称的，领导者

处于主导者地位。而在古诺模型中，企业间的地位是对称的，或称为“势均力敌”的。

这就是说，领导者的利润最差也是在它选择古诺均衡时的产量，也就是说领导者本来完全可以选择古诺均衡的产量水平，但它不会这么做，因为它是主导企业。所以斯坦科尔伯格模型中的领导者绝对不会获得比在古诺均衡时更低的利润。

4.假设有n个相同的企业实现了一个古诺均衡。证明市场需求曲线的弹性绝对值必定大于$1/n$。（提示：在垄断的情形$n=1$，这意味着垄断企业会在需求曲线具有弹性的那一段进行生产。使用我们证明垄断企业不会在需求缺乏弹性处生产的逻辑证明该问题。）

【复习内容】多个企业的古诺均衡

假设行业中有n个企业，令$Y=y_1+...+y_n$表示行业总产量。则企业i的“边际收入等于边际成本的条件”为

$$p(Y)+\frac{\Delta p}{\Delta Y}y_i=MC(y_i).$$

对于上式的左端，我们将$p(Y)$提取出来，并将左端的第二项乘以Y/Y，可将上式变形为

$$p(Y)[1+\frac{\Delta p}{\Delta Y}\frac{Y}{p(Y)}\frac{y_i}{Y}]=MC(y_i).$$

使用总需求曲线的弹性定义，并且令$s_i=y_i/Y$表示企业i的产量占市场总产量的份额，上式可以简化为

$$p(Y)[1-\frac{s_i}{|\varepsilon(Y)|}]=MC(y_i).$$

我们也可以将这个表达式写为

$$p(Y)[1-\frac{1}{|\varepsilon(Y)|/s_i}]=MC(y_i).$$

除了多出了s_i一项外，这个式子很像垄断企业的利润最大化条件的表达式。我们可以将$\varepsilon(Y)/s_i$看成企业i面对的需求曲线的弹性：它的市场份额越小，它面对的需求曲线的弹性越大。

【参考答案】

证明：多个企业实现古诺均衡时必有

$$p(Y)[1-\frac{1}{|\varepsilon(Y)|/s_i}]=MC(y_i).$$

其中，Y 表示行业产量，y_i表示企业i的产量，$\varepsilon(Y)$表示市场需求曲线的弹性，s_i表示企业i的市场份额。

由于题目告知，这 n 个企业是完全相同的，因此 $s_i = 1/n$，将其代入上式可得

$$p(Y)[1-\frac{1}{n|\varepsilon(Y)|}] = MC(y_i).$$

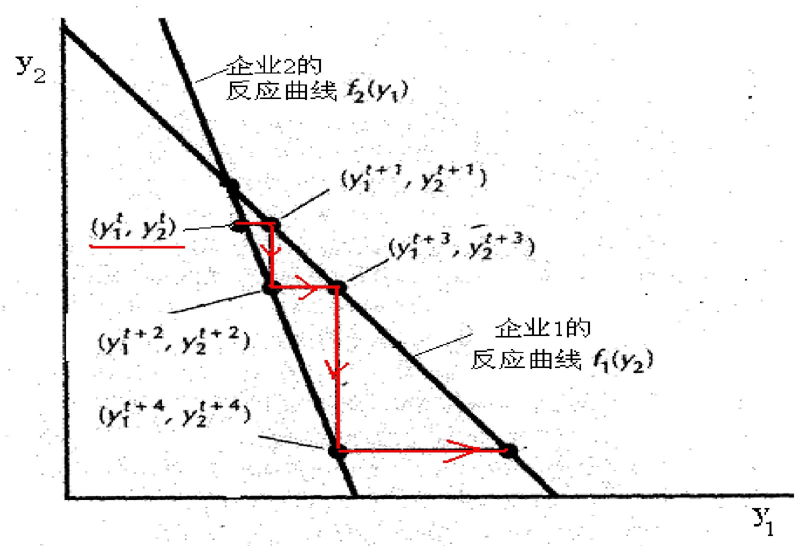
由于等式右端即 $MC(y_i)>0$，所以等式左端 $p(Y)[1-\frac{1}{n|\varepsilon(Y)|}]>0$，又因为 $p(Y)>0$，所以 $1-\frac{1}{n|\varepsilon(Y)|}>0$，这意味着 $|\varepsilon(Y)|<\frac{1}{n}$。证毕。■

5.画出导致不稳定均衡的一组反应曲线。

【复习内容】古诺均衡实现的调整过程

【参考答案】

这个图形和教材图 27.4 外观相似，但内涵不同。注意此处，**企业 2 的反应曲线比企业 1 的反应曲线陡峭**。



假设在时期 t 企业的产量为 (y_1^t, y_2^t)，这个产量组合未必是均衡产量。若企业 1 预期企业 2 的产量将继续维持在 y_2^t，则在下一期，在企业 2 的预期产量为 y_2^t 时，企业 1 选择的利润最大化产量为 $f_1(y_2^t)$.因此，企业 1 在 $t+1$ 期的产量选择为

$$y_1^{t+1} = f_1(y_2^t).$$

企业 2 当然可以进行类似推理，因此企业 2 在 $t+1$ 期的选择为

$$y_2^{t+1} = f_2(y_1^t).$$

假设我们从点 (y_1^t, y_2^t) 开始分析。给定企业 2 的产量水平，企业 1 的最优选择是在下一期生产 $y_1^{t+1} = f_1(y_2^t)$ 单位的产品。在图形上，这表现为企业 1 的产量**水平向右**移动，直到它

到达它自己的反应曲线上。

若企业 2 预期企业 1 的产量继续维持在 y_1^{t+1} 的水平上，则它的最优反应是生产 y_2^{t+1} 单位产品。我们将企业 2 的产量**向下方垂直移动**，直到到达企业 2 的反应曲线，我们就找到了这一点。沿着图中的“阶梯”继续移动，就可以得到这两个企业的一系列产量选择。在本例中，这种调整过程无法实现均衡。

6.寡头垄断企业能否生产帕累托有效率的产量？

【复习内容】寡头垄断行为的各种解释模型

【参考答案】

我们已分析了好几种寡头垄断行为的模型：产量领导（斯坦科尔伯格），价格领导，同时制定产量（古诺），同时制定价格（伯特兰）以及合谋。

一般来说，合谋会导致最小的产量和最高的价格。伯特兰均衡，即竞争均衡的产量最高、价格最低。其他模型的结果介于这两个极端之间。

我们已经知道，如果市场具有垄断因素（寡头垄断只是垄断的一种形态），产量一般不是帕累托有效率的。但是，要注意两个例外：

第一，如果寡头合谋形成卡特尔，由于卡特尔的行为类似于一个垄断企业，因此它的产量一般是无效率的，但是我们已经知道如果卡特尔能够实行完全价格歧视，它的产量就是有效率的。

第二，由于伯特兰模型（寡头同时制定价格模型）在均衡时就是一个竞争均衡，因此也是有效率的。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

28.博弈论（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

28 博弈理论

我们在上一章阐述的寡头理论，是企业间策略性互动的经典经济理论解释。但这只是冰山一角。经济行为人（agents）的策略性互动有多种方式，经济学家借助博弈理论（game theory）这个工具已研究了很多种策略性互动的行为。博弈理论关注的是策略性互动的一般分析。人们可使用博弈理论研究室内游戏（parlor games）、政治协商和经济行为[一]。在本章，我们将简要分析这一迷人的学科，目的是让你感受一下它是如何运行的，以及让你初步知道如何使用博弈理论分析寡头市场中的经济行为。

28.1 博弈的收益矩阵

策略性互动可能涉及很多选手和很多策略，但是我们仅限于分析两个选手之间的博弈，而且限于分析策略的数量有限的情形。这样做的好处是可以用收益矩阵（payoff matrix）描述博弈。最好举例进行分析。

假设两人玩一种简单的游戏。选手 A 在纸上写出“上”或“下”。与此同时，选手 B 独立地写出“左”或“右”。在两人写好后，经过分析，将他们的收益标记于表 28.1 中。若 A 选上且 B 选左，我们看矩阵的左上角的小方格。在该小方格中，A 的收益是第一个数，B 的收益是第二个数。类似地，如果 A 选下 B 选右，则 A 得到收益为 1，B 得到的收益为 0.

		Player B	
		Left	Right
Player A	Top	1, 2	0, 1
	Bottom	2, 1	1, 0

表 28.1：一个博弈的收益矩阵

选手 A 有两个策略：上或下。这些策略可以代表类似“提高价格”或“降低价格”的

[一] 室内游戏（parlor games）是指一伙人在室内（indoors）参与的游戏。在维多利亚时代的英国和美国，室内游戏在中上流阶级非常盛行。译者注。

经济选择。或者它们可以代表类似“宣战”或“不宣战”的政治选择。博弈的收益矩阵表明了对于每个选定的策略组合，每个选手得到的收益。

这类博弈的结果是什么样的？表 28.1 表示的这种博弈，有一个很简单的解。从选手 A 的观点看，选择下总是比选择上更好，因为选择下的收益（2 或 1）总是大于选择上的相应收益（1 或 0）。类似地，对于 B 来说，选择左比总是比选择右更好，因为（2 或 1）相应比（1 或 0）大。因此，我们可以预期均衡策略是：A 选下，B 选左。

这种情形下，我们得到了一个占优势的策略或者简称**占优策略**（dominant strategy）[一]。每个选手只有唯一一个最优选择，不论对方怎么改变策略。例如，不论 B 怎么选择，若 A 选下，A 的收益总是大于选择上的收益，因此 A 自然会选择下。类似地，不论 A 怎么选择，B 选择左的收益更高。因此，这些选择比其他选择好，这样我们就得到了一个占优策略均衡解。

如果在一个博弈中，每个选手都有一个占优策略，我们可以预测占优策略组合就是该博弈的均衡结果。这是因为占优策略是指，不论对方如何选择，你选择的这个策略都是最优的。在这个例子中，我们可以预期均衡结果为：A 选下（均衡收益为 2），B 选左（均衡收益为 1）。

28.2 纳什均衡

占优策略均衡很好分析，可惜占优策略均衡不是那么常见。例如，表 28.2 表示的博弈不存在占优策略均衡解。在该博弈中，B 选左时，A 的收益为 2 或 0。B 选右时，A 的收益为 0 或 1。这表示 B 选左时，A 会选上；B 选右时，A 会选下。因此 A 的最优选择取决于他认为 B 会怎么选。

		Player B	
		Left	Right
Player A	Top	2, 1	0, 0
	Bottom	0, 0	1, 2

表 28.1：一个纳什均衡

然而，也许占优策略均衡要求太苛刻，因为它要求对于 B 的**所有**选择，A 的选择都是最优的。现在我们不这么要求，我们只要求对于 B 的**最优**选择来说，A 的选择是最优的即可。因为如果 B 是一个理性选手，他只会选择最优的策略。（当然，B 的最优策略也取决于 A 的选择！）

[一] 有时也翻译为“优势策略”，在翻译过程中，这两种译法我都使用了。译者注。

如果给定 B 的选择，A 的选择为最优，而且给定 A 的选择，B 的选择也为最优，那么我们将 A 和 B 此时选择的策略称为一个**纳什均衡**（Nash equilibrium）[一]。记住，当每个选手在选择自己的策略时，都不知道对方选择哪一个策略。但是他可以猜测对方选择的策略。一个纳什均衡可以看成一对预期选择，这样的选择要能使得一旦对方的选择展现后，选手都不再改变自己的行为。

在表 28.2 代表的博弈中，策略（上，左）是一个纳什均衡。为了证明这一点，先分析 B 的行为。假设 A 选择上，那么 B 的最优选择是选择左，这是因为他选择左的收益为 1，而选择右的收益为 0。再来分析 A 的行为。如果 B 选择左，那么 A 的最优选择是选择上，因为选择上的收益为 2，而选择下的收益为 0。

因此，如果 A 选择上，B 的最优选择是选择左；而如果 B 选择左，那么 A 的最优选择是选择上。这样我们就得到了一个纳什均衡：给定对方的选择，每个选手都作出了最优选择。

纳什均衡是上一章介绍的古诺均衡的一般形式。在古诺均衡中，选择为产量水平，每个企业在选择它的产量水平时，都假定对方的选择是既定的。每个企业在做选择时都假设对方选择原来的产量，也就是说按照以前选择的策略进行生产，在这种情形下它选择的产量应该使自己的利润最大化。给定对方的行为，每个企业的利润都实现了最大化，这就是一个古诺均衡。按照纳什均衡的定义，显然古诺均衡是一种纳什均衡。

纳什均衡的概念有一定的内在逻辑。不幸地是，该均衡也存在一些问题。首先，一个博弈可能有多个纳什均衡。事实上，在表 28.2 中，选择（下，右）也是一个纳什均衡。你可以按照我们上面介绍的推理方法进行分析。当然下面这种证明方法也可行。注意到这个博弈结构是对称的：B 在一种结果的收益，等于 A 在另一结果中的收益，因此我们证明了（上，左）是一个纳什均衡，这也意味着我们也证明了（下，右）也是一个纳什均衡。

		Player B	
		Left	Right
Player A	Top	0, 0	0, –1
	Bottom	1, 0	–1, 3

表 28.3：不存在（纯策略）纳什均衡解的一个博弈

纳什均衡概念的第二个问题是有些博弈不存在我们上面描述的纳什均衡解。例如，考虑

[一] 约翰.纳什是一位美国数学家，他在 1951 年提出了这个博弈理论中的基本概念。1994 年他和另两外博弈理论学者共同获得了诺贝尔经济学奖。2002 电影《美丽心灵》大致以纳什的生活为蓝本，该电影获得奥斯卡最佳电影奖。

表 28.3。该博弈不存在我们上面描述的那种纳什均衡解。如果 A 选择上，则 B 选择左。但是若 B 选择左，则 A 选择下。类似地，若 A 选择下，则 B 选择右。但是若 B 选择右，则 A 选择上。

28.3 混合策略

然而，如果我们扩大策略的定义，我们可以为表 28.3 的博弈找到一个新类型的纳什均衡解。我们在前面的思路其实是一直认为，每个选手在选择策略时能做到一劳永逸。也就是说，每个选手作出选择后就坚持这个选择。这种情形下，每个选手的策略都为一个**纯策略**（a pure strategy）。

另外一种思路是我们允许选手将他们的策略**随机化**，也就是说对每个选择都赋予一个概率值，而且按照这些概率选择策略。例如，A 以概率 50%选择上、以概率 50%选择下，而 B 以概率 50%选择左、以概率 50%选择右。这种情形下，每个选手的策略都为一个**混合策略**（a mixed strategy）。

如果 A 和 B 都采用上述混合策略，即每个选手以相等的概率选择他的两个策略中的一个，那么收益矩阵每个小方格中的收益，出现的概率都为 1/4。因此，A 的平均收益为 0，B 的平均收益为 1/2。

混合策略中的纳什均衡是指，均衡时，给定对方选择策略的概率，每个选手选择的含有概率的策略都是最优的。

可以证明对于本章分析的这类博弈，总是至少存在一个混合策略纳什均衡解。因为混合策略的纳什均衡解总是存在的，也因为该概念具有一定的内在合理性，所以它成为分析博弈行为的一个非常流行的工具。在表 28.3 的例子中，可以证明，如果 A 以 3/4 的概率选择上、以 1/4 的概率选择下，而且 B 以 1/2 的概率选择左、以 1/2 的概率选择右，那么这些策略就构成了一个纳什均衡。

例子：剪刀、石头和布

我们对混合策略说得已经够多了。现在来看一个重要例子，这就是我们都知道的游戏“剪刀石头和布。”在这个游戏中，每个选手同时选择出示拳头（石头）、手掌（布）或两个手指（剪刀）。游戏的规则为：石头砸烂剪刀，剪刀剪破布，布包住石头。

在人类历史上，该游戏百玩不厌。甚至还有一个称为 RPS 协会的专业团体，专门推广该游戏。它有自己的网站，它还提供了 2003 年在加拿大多伦多举行的锦标赛的纪录片。

当然，博弈论专家认识到这个游戏中的均衡策略是随机选择这三个选项中的一种。但是人类并不必然擅长选择完全随机的选项。如果你在某种程度上能预测到对手的选择，你在选择策略时将占有一定的优势。

纽约时报记者詹妮弗 8.李，曾经半开玩笑地说道，心理学是至高无上的。在她的文章中，她写道“大多数人在无防备的情况下，都有自己偏好的选择，这反映了他们的性格。“布”代表着优雅甚至被动的选择，因此文学人士和记者在玩这个游戏时一般会选择“布”。”[一]

经济学家在玩这个游戏时喜欢出哪一项呢？也许是剪刀，因为我们希望剪出影响人们行为的决定因素。经济学家出剪刀时，你是否应该出石头呢？也许，但是我并不总是出剪刀…

28.4 囚犯的两难问题

博弈纳什均衡解的另外一个问题是，它不必然导致帕累托有效率的结果。例如，考虑表 28.4 中的博弈。这个博弈称为囚犯的两难或囚犯的困境（prisoner’s dilemma）。最初这个模型是这样的：警察将合伙犯罪的两个人分别关押在单独的囚房内，分别审讯。每个罪犯的选择为：可以选择认罪，从而供出来了他的合伙人；也可以选择不认罪。如果只有一个罪犯认罪，那么他可以被释放，而另外一个罪犯会受到严惩——坐牢 6 个月。如果两个罪犯都不认罪，那么根据法律每人被监禁 1 个月；如果两个人都认罪，那么每个人被监禁 3 个月。这个博弈的收益矩阵可用表 28.4 表示。每个小方格中的元素表示每个罪犯对各种结果的效用评价。为简单起见，我们用负数表示他们的效用，这个效用取决于坐牢期限，时间越长，效用越小。

		Player B	
		Confess	Deny
Player A	Confess	−3, −3	0, −6
	Deny	−6, 0	−1, −1

图 28.4：囚犯的两难问题

我们先来分析 A 的选择。如果 B 选择否认，那么 A 最好的选择是认罪，因为这样 A 就会被释放。类似地，如果 B 选择认罪，那么 A 最好的选择也是认罪，因为这样 A 会被监禁 3 个月而不是 6 个月。因此，不论 B 怎么选择，A 最好的选择是认罪。

[一] Jennifer 8.Lee ,“Rock, Paper, Scissors: High Drama in the Tournament Ring", New York Times September 5, 2004.（以下为译者注）注意该记者的名字中含有数字“8”，据说她是个中国通，认为 8 这个数字很吉利，加 8 在名字中让她的名字很特别。

B 的选择可以类似推理，B 的最优选择也是认罪。因此，该博弈的唯一纳什均衡是两个罪犯都认罪。事实上，两个罪犯都认罪不仅是一个纳什均衡，而且是一个占优策略均衡，因为每个罪犯的最优选择和对方的选择无关。

但是如果他们咬紧牙关拒不认罪，那么他们的状况会变好！如果这两个人相信对方不会认罪，对方的确不会认罪，那么这种情形下，每个人的收益均为-1，这会使得每个人的状况变得更好。策略（否认，否认）是帕累托有效率的，因为已不存在能使者两个人的状况都变好的策略。策略（认罪，认罪）是帕累托无效率的。

问题在于这两个罪犯无法协调彼此的行为。如果他们彼此信任，则他们的状况都会变得更好。

囚犯两难模型可应用于广泛的经济和政治现象。例如军备控制问题。我们可以将囚犯困境中的策略“认罪”看为“使用新导弹”、将“不认罪”看成“不使用导弹”。注意该情形下表 28.4 表示的收益仍然是合理的。如果我的对手使用导弹，我当然希望是使用导弹，尽管我们双方最好的策略都是不使用导弹。但是，如果不能达成具有约束力的协议，我们双方都会使用导弹，结果我们的状况变差了。

卡特尔成员欺骗问题也是一个好例子。现在将策略认罪看成“生产比你的份额更多的产量”，将不认罪看成“坚持生产原来的份额”。如果你认为其他的企业将坚持它们各自的份额，那么多生产对你是有利的。如果你认为其他企业会多生产，那么你可能也多生产！

囚犯的难题问题让人们激烈辩论到底怎样进行博弈才是“正确的”，或者，更准确地说，参与博弈的合理方式是什么。答案似乎取决于你参与的是一次博弈还是无限次的博弈。

如果博弈只进行一次，欺骗的策略（在囚犯两难中是指认罪的策略）将是合理的。毕竟，不管其他人的策略如何，你选择这种策略都会让你的状况变好，而且你无法影响其他人的行为。

28.5 重复博弈

在上一节，选手只相遇一次而且他们也只参与一次囚犯两难博弈。但是，如果这些选手重复进行博弈，情形将会有所不同。在这种情形下，每个选手都可能想出新的策略。如果其他选手在某一轮博弈中选择的测策略为背叛，那么你可以在下一轮选择背叛的策略。因此，你的对手会因为“恶劣的”行为而“受到惩罚”。在一个重复博弈中，每个选手都有机会为自己塑造合作的声望，因此鼓励其他选手也这么做。

这种策略是否可行，取决于博弈是进行**既定**的次数（比如 10 次）还是进行**无限**次。

我们首先分析第一种情形，假设两个选手知道博弈将进行 10 次。该博弈的结果是什么？我们从第 10 轮开始分析。这是上述博弈的最后一轮。在这种情形下，似乎每个选手都会选择占优策略即背叛。毕竟，最后一轮博弈和只进行一次的博弈米什么区别，所以我们可以预

曹乾（东南大学 caoqianseu@163.com）

期这两种博弈的结果是一样的。

现在分析第 9 轮的结果。我们刚得出结论即每个选手都会在第 10 轮选择背叛的策略。既然这样，他们会在第 9 轮合作吗？不会。如果你选择合作，但是对方可能会利用你善良的天性，从而选择背叛。每个选手都可以同样推理，因此每个选手都会选择背叛。

现在考虑第 8 轮。如果某个选手在第 9 轮选择背叛…以此类推。如果选手都知道博弈进行的具体次数，那么每个选手在每一轮都会选择背叛。如果无法强制选手在最后一轮合作，自然也无法强制选手从第一轮直至最后一轮选择合作。

选手相互进行合作的原因，是希望将来进一步合作。但是这要求将来还有博弈的机会。由于在最后一轮，选手们都知道将来不可能再进行博弈，没有人会选择合作。既然这样，他们为什么在倒数第二轮合作吗?或者在倒数第三轮合作？以此类推——在囚犯困境博弈中，若选手都知道博弈的具体次数，那么合作解从最后一轮博弈将象多米诺骨牌一样开始“倒塌”，因此均衡解必然是选手都选择背叛。

但是如果博弈将重复无限次，那么你**的确**可以找到影响对手行为的方法：如果对手这一次不合作，那么下一次你可以拒绝合作。只要双方都非常看重将来的收益，将来不合作的这种威胁足以让人们选择帕累托有效率的策略，即都选择合作。

罗伯特.阿克塞罗德（Robert Axelrod）在进行了一系列实验后令人信服地证明了上述结论[一]。他恳请博弈论领域的几十位专家，向他提交他们认为的囚犯困境的最优策略，然后他在计算机上开展了“锦标赛”，让这些策略互相进行比赛。在计算机上，每种策略都要和其他每一种策略竞争，计算机实时记录博弈收益。

最终获胜的策略——收益最高的策略——竟然是一种最为简单的策略。这种策略叫做“以牙还牙”（tit for tat），它的运行方式如下。在第一轮，你合作(即选择不认罪的策略)。在以后的每一轮，如果你的对手在前一轮选择合作，你也选择合作。如果对方在上一轮选择背叛，你也选择背叛。换句话说，每个人的策略是选择对方在上一轮的策略。

以牙还牙策略收益最高，因为它对背叛行为立即实施惩罚措施。这种策略也是一种宽恕的策略：发现一次背叛，只惩罚一次。如果对方改邪归正开始合作，那么以牙还牙策略将以合作回报对方。在囚犯困境博弈将进行无限次的情形下，以牙还牙策略似乎是实现有效率结果的一种非常好的机制。

[一] Robert Axelrod is a political scientist from the University of Michigan. For an extended discussion, see his book The Evolution of Cooperation (New York: Basic Books, 1984).

28.6 实施卡特尔

在第 27 章，我们分析了双头垄断制定价格的博弈行为。在那一章我们断言，如果每个垄断企业能够选择价格，那么均衡结果将是竞争均衡。如果每个企业认为其他企业会保持价格固定不变，那么每个企业都会发现降价是有利可图的。这个结论只有在下列情形下才不会成立：每个企业的要价已是最低可能的价格，在 27 章的那个例子中，这个最低价格为零，因为我们假设边际成本为零。如果使用本章的术语表达，每个企业索要零价格是定价策略中的一个纳什均衡——但在第 27 章我们将其称为伯特兰均衡。

双头垄断的定价策略博弈，和囚犯的两难博弈具有同样的收益矩阵的结构。如果每个企业索要高价，那么每个企业都能得到更大的利润。这种情形就是它们合谋成卡特尔，并且坚持生产垄断产量。但是如果一个企业索要高价，另外一个企业稍微降低一点价格就是值得的，因为这样做可以夺取其他企业的市场，因此得到更大的利润。但是，如果两个企业都降低讲个，它们最终得到的利润都降低了。不论对方索要什么样的价格，你稍微降低一点价格总是有利可图的，当然前提是价格仍大等于边际成本。纳什均衡发生在每个企业索要最低可能的价格。

然而，如果博弈重复进行无限次，那么可能还有其它结果。假设你决定实施以牙还牙策略。如果另外一个企业这周降价，你可以在下周降价。如果每个选手知道对方都会以牙还牙，那么每个选手都不会降低价格，因为这样会引起价格大战，各个选手的利益都受损。因此，以牙还牙的潜在威胁，能够使得所有企业维持高价。

现实生活中的卡特尔有时会使用以牙还牙策略。例如，联合执行委员会是一个有名的卡特尔，它在 1800 年代后期负责制定美国铁路货运的价格。这个卡特尔形成于美国反垄断法规生效之前，当时它是完全合法的[一]。

这个卡特尔负责确定每个铁路公司货运的市场份额。每个企业独立制定自己的运费标准，该卡特尔记录每个铁路公司的货运数量。然而，在 1881、1884 和 1885 年间，有些公司认为其他成员公司偷偷降价来增加它们自身的市场份额，尽管所有公司事先约定不准降价。在这个时期，经常发生价格大战。当一个公司试图欺骗，所有其他公司都会降低价格以“惩罚”背叛者。这种以牙还牙策略显然能够保证卡特尔稳定运行一段时间。

例子：机票定价中的以牙还牙策略

机票定价为以牙还牙行为提供了一个有趣的例子。航空公司经常会提供这种或那种促销价格；航空业中的很多研究者认为，这些促销价格是用来向竞争对手发送信号，警告它们不要降低重要航线的机票价格。

美国某大型航空公司营销总监曾描述个一个案例：西北航空公司降低了从明尼阿波利

[一] For a detailed analysis, see Robert Porter, "A Study of Cartel Stability: the Joint Executive Committee, 1880-1886," The Bell Journal of Economics, 14, 2 (Autumn 1983), 301-25.

斯市（Minneapolis）到西海岸各个城市的夜间航班的价格，目的在于减少空座率。大陆航空公司则认为这种做法是在抢夺它的市场份额，因此宣布降低所有从明尼阿波利斯市到西北各城市的夜间航班价格。然而，大陆航空公司的降价只进行了一两天后就停止了。

西北航空将大陆航空这一行为解读为，大陆航空不想参与价格竞争，它的目的在于让西北航空停止夜间航班降价。但是西北航空公司决定向大陆航空发送自己的信号：它对从休斯顿到西海岸各个城市的航班都制定了一套便宜的价格，要知道休斯顿可是大陆航空的总部所在地！西北航空传递的信号想表明，它的降价措施是正当合理的，而大陆航空的反应是不恰当的。

所有这些降价活动持续时间都很短；这个特征似乎表明，降价行为的本意在于发出竞争的信号而不是争夺更大的市场份额。正如这位总监解释的，航空公司并不想提供含有价格适用期的机票，它们的目的是最终能使竞争活跃起来并且展开竞争。

双头垄断的航空市场上的潜规则似乎为：如果一家公司的机票价格高，我的机票价格也高；但是如果对方降低价格，那么我就会以牙还牙，我也降低价格。换句话说，两个企业都“遵守着一条重要原则”：以其人之道还治其人之身。这种报复措施使得机票价格高昂[㈠]。

28.7 序贯博弈

到目前为止，我们分析的博弈都有一个共同特征：选手都是同时行动的。但在很多情形下，其中一个选手可以率先行动，其他选手再做出反应。这样的博弈叫做**序贯博弈**（sequential game）。比如第 27 章介绍的斯坦科尔伯格模型就是这样的例子，在该模型中一个选手是领导者，另外一个选手是追随者。

下面我们分析这样的博弈。在第一轮，选手 A 率先进行选择，他可以选择上或下。选手 B 观察 A 的选择，并相应作出选择左或右的决策。该博弈的收益矩阵如表 28.5 所示。

注意，当这该博弈以表 28.5 这种形式表示时，它有两个纳什均衡解：（上，左）和（下，右）。然而，下面我们将证明其中一个均衡解是不合理的。收益矩阵隐藏了下列事实：一个选手可以再观测另外一个选手选择之后，再进行选择。在这种情形下，我们有必要用另外一种图形表示博弈的收益，这种图形能更好地反映该种类型博弈的非对称性质。

图 28.1 画出了这个博弈的**展开形**（extensive form）。展开形是博弈的一种表示方法，它能显示出选择的先后顺序。首先，A 必须选择上或下，然后 B 必须决定选择左还是右。但是在 B 做出决策时，他已经知道 A 选择了哪个策略。

[㈠] Facts taken from A. Nomani, "Fare Warning: How Airlines Trade Price Plans," Wall Street Journal, October 9, 1990, B1.

曹乾（东南大学 caoqianseu@163.com）

		Player B	
		Left	Right
Player A	Top	1, 9	1, 9
	Bottom	0, 0	2, 1

表 28.5：一个序贯博弈的收益矩阵

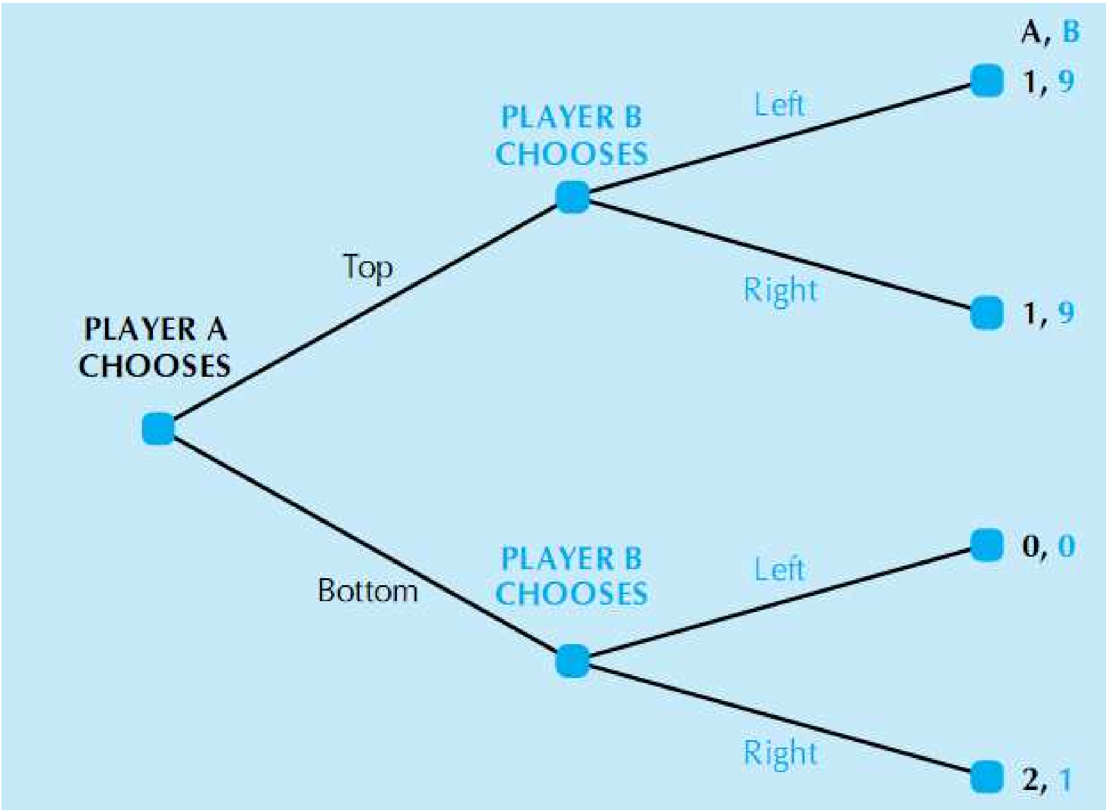


图 28.1：**博弈的展开形式**。展开形能表示出博弈中选手行动的先后顺序。

这种博弈的分析方法是从后（树根）向前（树枝）追溯。假设 A 已近作出了选择，我们处在该博弈树的一个分枝上。如果 A 选择上，则不管 B 怎么选择，A 和 B 的收益分别为 1 和 9，即图中的（1,9）。如果 A 选择下，则 B 合理的选择是选择右，因此收益为（2,1）。

现在分析 A 的初始选择。如果他选上，则结果为（1,9），因此他得到的收益为 1。但是，如果他选择下，则他得到的收益为 2。因此他会选择下。所以，该博弈的均衡解为（下，右），因此 A 的收益为 2，B 的收益为 1。

策略（上，左）不是该序贯博弈的合理均衡解。也就是说，由于这两个选手的行动有先后之分，这个策略集不是一个均衡。的确，如果 A 选择上，则 B 会选择左——但 A 不会傻到选择上！

 曹乾（东南大学 caoqianseu@163.com）

从 B 的角度来看，他相当不幸，因为他最终得到的收益为 1 而不是 9！他还有什么招数可使吗？

B 可以**威胁** A，即若 A 选下 B 就选左。如果 A 认为 B 真会这么做，那么他可能选择上。因为选择上他的收益为 1，而选择下——若 B 实施威胁计划——的收益为 0。

但 A 会相信 B 的威胁吗？毕竟一旦 A 做出了选择，就已无法反悔。B 的收益只能为 0 或 1，他很可能得到 1。除非 B 在某种程度上可以让 A 相信，他一定会实施威胁计划，即使自己利益受损也在所不惜，否则他只能得到较小的收益。

B 的问题是一旦 A 已经做出选择，A 期望 B 理性行事。如果 B **承诺**在 A 选择下时 B 会选择左，B 的状况会变好。

B 做出承诺的一种方式是让别人为他做出选择决策。例如，B 可以雇佣律师，让律师警告 A 如果 A 选择下则 B 必定选择左。如果 A 认识到这种警告的严重性，从他的角度看，结果将大不相同。如果他知道 B 对律师的指示，那么他知道如果他选择下，他最终的收益为 0。因此，他自然会选择上。在这种情形下，B **限定**了自己的策略，从而状况变得更好。

28.8 阻止进入的博弈（a game of entry deterrence）

我们在分析双头垄断时假设行业中的企业数目是固定不变的。但在很多情形下，新企业可能会进入该行业。当然，行业中原有的企业会想方设法阻止新企业进入。由于原有企业已在行业中，他们可以先发制人，因此在阻止竞争对手进入的博弈中具有先行选择策略的优势。

例如，假设某个垄断企业面对着另外一个企业进入行业的威胁。新企业（进入者）决定是否进入市场，原有企业决定是否降低价格作为回应。如果新企业决定不进入，它得到的收益为 1，原有企业得到的收益为 9。

如果新企业决定进入，那么它的收益取决于原有企业是否与它展开激烈竞争。如果企业进行竞争，那么我们假设两个选手的最终收益都为 0。另一方面，如果原有企业不进行竞争，我们假设进入者得到的收益为 2，原有企业得到的收益为 1。

注意，这正好是我们前面研究过的序贯博弈的结构，因此它的结构和图 28.1 是相同的。原有企业为 B，而潜在进入者为 A。策略上为不进入，策略下为进入。策略左为竞争，策略右为不竞争。我们已经知道，在这个博弈中，均衡结果是潜在进入者进入、原有企业不竞争。

原有企业的问题是它不可能事先承诺若其他企业进入他就会进行竞争。如果其他企业进入，损害已经造成，原有企业的理性行为是接受这一事实并且和平相处。然而如果潜在进入者认识到这一点，他自然会认为 B 的竞争威胁只是口号般的空话。

然而假设原有企业可以购买额外的生产能力，这样它就能以目前的边际成本生产更多的产量。当然，如果它仍然是垄断者，他不希望实际增加产量，因为原有垄断产量已实现了利

润最大化。

但是，如果其他企业进入，原有企业现在就能生产非常多的产量，因此可以与新进入者展开激烈的竞争。通过投资扩大额外产能，当其他企业试图进入时，它就可以降低成本打击进入者。假设如果原有企业购买额外产能而且选择竞争的话，那么他的收益为 2。这样博弈树 28.1 就变为了博弈树 28.2。

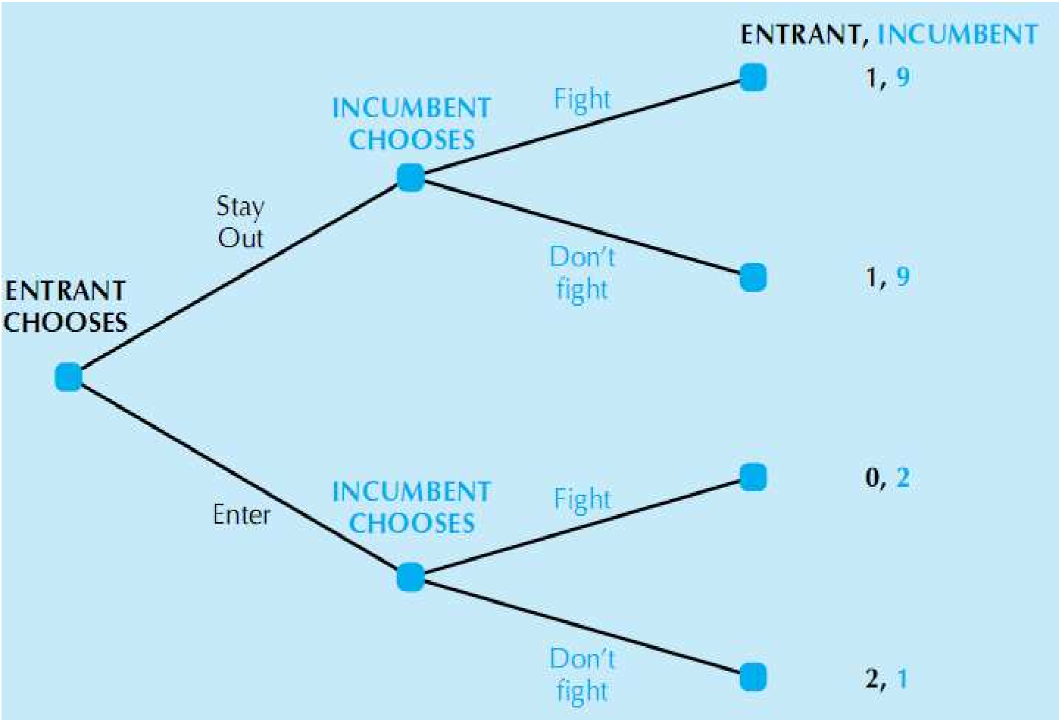


图 28.2：阻止进入的博弈模型。该图与 28.1 相比，收益发生了变化。

现在，由于增加了生产能力，竞争的威胁就是可信的。如果潜在进入者进入，若原有企业竞争他得到的收益为 2，若不竞争他的收益为 1；因此原有企业自然会选择竞争。潜在进入者如果进入得到的收益为 0，如果不进入得到的收益为 1，因此他自然会选择不进入。

但是，这意味着原有企业仍然是唯一的垄断者，它根本不会使用额外的生产能力！尽管如此，垄断则投资扩大产能还是值得的，因为在新企业试图进入市场时，它能够做到让对方相信它有能力对进入者进行打击。垄断者投资于“过剩”产能的目的，在于向潜在进入者发送信号：胆敢进入，必遭痛击。

总结

1.博弈的一种表示方法是，对选手的每个策略赋予相应收益。

2.一个占优策略均衡是指一组选择，这组选择具有下列特征：不论对方选择何种策略，每个选手的选择的策略都是最优的。

3.一个纳什均衡是指一组选择，对于这组选择：给定其它选手的选择，每个选手的选择都是最优的。

4.囚犯的两难是一种特殊的博弈，因为在该博弈中，选手选择的策略导致的是帕累托无效率的结果，而不是帕累托有效率的结果。

5.在序贯博弈中，选手选择的先后顺序非常重要。在这类博弈中，如果某个选手向其它选手事先承诺，他会沿着既定的路线进行博弈，那么它将处于有利地位。

复习题

1.在重复进行的囚犯两难博弈中，如果选手的策略都为以牙还牙。假设某个选手的本意是合作却不慎犯错——他背叛了。如果在下面回合的博弈中，这两个选手仍然以牙还牙，那么结果将如何？

2.占优策略均衡一定是纳什均衡吗？纳什均衡一定是占优策略均衡吗？

3.假设你的对手选择的不是他的纳什均衡策略，那么你还应该继续选择你的纳什均衡策略吗？

4.我们知道如果囚犯两难博弈只进行一次，那么它的结果是占优策略均衡，这一结果不是帕累托有效率的。如果两个罪犯在刑期结束即被释放后会报复对方。这样的行为将影响到该博弈的哪些方面？它能实现帕累托有效率的结果吗？

5.如果两个选手都知道他们之间的囚犯两难博弈将进行 100 万次，那么该博弈的占优纳什均衡策略是什么？如果你真得找到两个选手进行这样的实验，而且实验 100 万次，你能预测出他们会使用什么策略吗？

6.在教材图 28.1 表示的序贯博弈中，如果选手 B 而不是 A 先进行选择，请画出该新博弈的展开形。该博弈的均衡解是什么？选手 B 更喜欢自己先选择还是更喜欢让 A 先选择？

复习题答案

1.在重复进行的囚犯两难博弈中，如果选手的策略都为以牙还牙。假设某个选手的本意是合作却不慎犯错——他背叛了。如果在下面回合的博弈中，这两个选手仍然以牙还牙，那么结果将如何？

【复习内容】囚犯的两难博弈；以牙还牙策略

以牙还牙（tit for tat）策略运行方式如下：在第一轮，你合作。在以后的每一轮，如果你的对手在前一轮选择合作，你也选择合作。如果对方在上一轮选择背叛，你也选择背叛。换句话说，每个人的策略是选择对方在上一轮的策略。

以牙还牙策略收益最高，因为它对背叛行为立即实施惩罚措施。这种策略也是一种宽恕的策略：发现一次背叛，只惩罚一次。如果对方改邪归正开始合作，那么以牙还牙策略将以合作回报对方。在囚犯困境博弈将进行无限次的情形下，以牙还牙策略似乎是实现有效率结果的一种非常好的机制。

【参考答案】

在重复进行的囚犯两难博弈中，以牙还牙策略，简单地说是指，每个选手选择对方在上一轮的策略。

由题目可知，如果某选手 A 不慎背叛，即使不是出自其本意，但如果 B 无法了解这些信息，那么它会认为 A 的行为是真正的背叛。

按照以牙还牙的逻辑，B 在第二轮中的策略，他应选择 A 在上一轮中的策略，即选择背叛。这个信号会让 B 在下一轮中也选择背叛，以此类推。选手 A 和 B 不断地以背叛策略作为对对方背叛的反应。

这个例子说明，如果博弈中某个选手不慎犯错，他应该及时沟通，否则大家将一直背叛到底，这样的结果显然不是帕累托有效率的，也就是说在这种情形下，以牙还牙不再是一个很好的策略。

2.占优策略均衡一定是纳什均衡吗？纳什均衡一定是占优策略均衡吗？

【复习内容】占优策略均衡和纳什均衡

占优策略是指，不管对方选择哪个策略，你的最优选择是唯一的。也就是我们通常所说的“以不变应万变”，这里的“不变”的策略就是你的占优策略。如果均衡时每个对手选择的都是占优策略，那么该均衡就是占优策略均衡。

纳什均衡，简单地说，你的选择根据对手的选择相应调整，在均衡时，双方都不会再改

变策略。我们通常所说的"兵来将挡、水来土掩"就是纳什均衡的例子。

【参考答案】

占优策略均衡一定是纳什均衡，纳什均衡未必是占优策略均衡。

占优策略均衡要求对于 B 的**所有**选择，A 的选择都是最优的。而纳什均衡仅要求：对于 B 的**最优**选择来说，A 的选择是最优的即可。由于 A 和 B 的地位是对称的，你可以类似推理 B 的选择。由此可见，占优策略均衡是纳什均衡的一种。因为如果 A 选择的策略对 B 的**所有**策略来说都是最优的，那么显然 A 选择的策略对 B 的**最优**策略来说也是最优的。

纳什均衡未必是占优策略均衡，比如在"剪刀石头布"游戏中不存在占优策略均衡，但存在纳什（混合策略）均衡。既然我们已经举出了一个博弈是纳什均衡但不是占优策略均衡的例子，而且我们又知道占优策略均衡一定是纳什均衡，我们当然可以断言，纳什均衡未必是占优策略均衡。

3.假设你的对手选择的不是他的纳什均衡策略，那么你还应该继续选择你的纳什均衡策略吗？

【复习内容】纳什均衡策略

【参考答案】

你可能但一般不会继续选择纳什均衡策略。原因如下：

纳什均衡策略是指对方采用纳什均衡策略时，你选择的最优策略。典型的纳什均衡要求双方的决策是相互依赖的，你必须根据对手的选择相应出招。如果两个选手都是理性的，那么纳什均衡结果是"势均力敌的"，也就是说给定对方的最优选择，你的选择也是最优的。

比如足球比赛中你若是前锋，在与对方守门员的博弈中，若你和守门员都是理性的，那么你们的策略显然是相互依赖的。比如你踢向球门左方，守门员的最优选择就是扑向左方。我们假设不管什么原因，该守门员总是扑向左方，你自然会选择踢向右方。

这个例子说明，如果对方选择的不是纳什均衡策略，那么一般情形下你会有更好的选择，也就是说你不会继续选择纳什均衡策略。

但是，需要注意，由于占优策略均衡是一种比较特殊的纳什均衡，在这种情形下，不管对方怎么选择，你的策略都是不变的。因此，你会继续选择你的占优策略。

综合以上两种情形，可知答案为如果对方选择的不是纳什均衡策略，那么你可能但一般不会继续选择纳什均衡策略。

4.我们知道如果囚犯两难博弈只进行一次，那么它的结果是占优策略均衡，这一结果不是帕

累托有效率的。如果两个罪犯在刑期结束即被释放后会报复对方。这样的行为将影响到该博弈的哪些方面？它能实现帕累托有效率的结果吗？

【复习内容】囚犯两难博弈；重复博弈

在重复进行的囚犯两难博弈中，每个选手都可能想出新的策略。如果其他选手在某一轮博弈中选择的测策略为背叛，那么你可以在下一轮选择背叛的策略。因此，你的对手会因为“恶劣的”行为而“受到惩罚”。在一个重复博弈中，每个选手都有机会为自己塑造合作的声望，因此鼓励其他选手也这么做。

因此，面临报复的威胁时，参与博弈的选手都会重新思考和进行选择，在这种情形下，会改变博弈的收益，从而改变了博弈的结果。在该情形下，选手很可能选择合作，因此，产生了帕累托有效率的结果。

但是如果这种威胁并不可信，那么选手就不会选择合作，博弈结果和只进行一次的博弈结果是一样的，也就是说，每个选手都会背叛对方，从而结果不是帕累托有效率的。

5.如果两个选手都知道他们之间的囚犯两难博弈将进行 100 万次，那么该博弈的占优纳什均衡策略是什么？如果你真得找到两个选手进行这样的实验，而且实验 100 万次，你能预测出他们会使用什么策略吗？

【复习内容】囚徒两难博弈；重复博弈；以牙还牙策略

【参考答案】

该例子中，占优纳什均衡策略为两个选手在每一回合的博弈中都选择背叛。具体的分析思路是从后向前进行归纳。

选手相互进行合作的原因，是希望将来进一步合作。但是这要求将来还有博弈的机会。由于在最后一轮，选手们都知道将来不可能再进行博弈，没有人会选择合作。既然这样，他们为什么在倒数第二轮合作吗？或者在倒数第三轮合作？以此类推——在囚犯困境博弈中，若选手都知道博弈的具体次数为 100 万次，那么合作解从最后一轮博弈将象多米诺骨牌一样开始“倒塌”，因此均衡解必然是选手都选择背叛。

但在现实中，如果让选手博弈 100 万次，几乎是不可能完成的任务，因此可以视为无限次重复进行的博弈。这种情形下双方最好的选择都是以牙还牙。以牙还牙的威慑力可在很大程度上保证选手之间的合作。

因此，如果在现实中进行这样的实验，很可能出现的结果是两个选手彼此合作，因此结果为帕累托有效率的。

6.在教材图 28.1 表示的序贯博弈中，如果选手 B 而不是 A 先进行选择，请画出该新博弈的

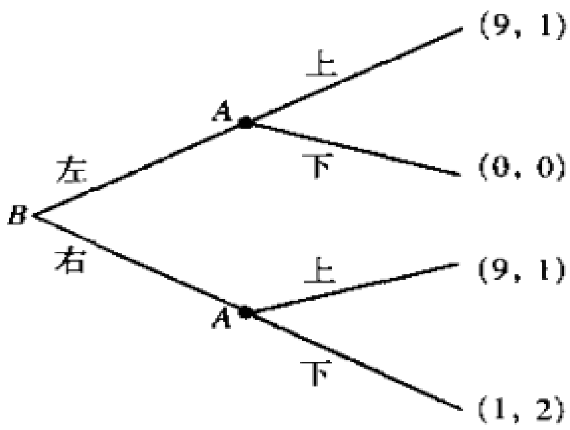
展开形。该博弈的均衡解是什么？选手 B 更喜欢自己先选择还是更喜欢让 A 先选择？

【复习内容】序贯博弈

序贯博弈是与同时博弈相对，是指选手行动有先后之分而不是同时决策。行动的先后顺序会影响博弈收益，从而影响博弈的结果。

由于收益矩阵通常不容易展示序贯博弈的结构，也就是说收益矩阵通常会将决策先后顺序信息隐藏起来，如果依据收益矩阵进行分析，非常容易出错。因此，在序贯博弈的情形下，我们通常用展开形即博弈树来表示博弈。

【参考答案】**该博弈的均衡解是（左，上）；选手 B 更喜欢自己先选择。**



图：选手 B 先行动时的博弈树（博弈展开形）

我们从博弈树的树根向树枝方向进行分析。如果 B 选择左，则 A 会选择上，此时收益为（9,1）。

如果 B 选择右，则 A 会选择下，此时收益为（1,2）。对比这两种情形下 B 的收益可知，B 会选择左，在 B 选择左的情形下，A 会选择上，因此均衡策略解为（左，上），此时收益为（9,1）。

对比教材中 A 先行动时最终均衡结果中 B 的收益（为 1）可知，B 会选择率先选择。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

29.博弈论应用（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

29 博弈论的应用

在上一章，我们介绍了博弈理论中的几个重要概念，并分别举例进行了分析。在本章，我们分析博弈理论中的四个重要问题——合作、竞争、共存和承诺——看看它们在各种策略互动中是如何运行的。

为了做此事，我们首先引入一个重要的分析工具，即**最优反应曲线**（best response curves）。我们可用这个工具求出博弈的均衡解。

29.1 最优反应曲线

我们以两人参与的博弈进行分析，假设你是其中一个选手。对于对方的任何选择，你的**最优反应**（best response）就是使的你的收益最大化。如果有若干个选择都能使你的收益最大，那么你的最优反应是这些选择的集合。

例如，考虑表 29.1 中的博弈，我们在上一章曾用该博弈说明纳什均衡的概念。如果位于列的选手 Column（以下简称 C）选择左，位于行的选手 Row（以下简称 R）的最优反应是选择上；若 C 选择右，则 R 的最优反应是选择下。类似地，若 R 选择上、下时，C 的最优反应分别为左、右。

		Column	
		Left	Right
Row	Top	2, 1	0, 0
	Bottom	0, 0	1, 2

表 29.1：一个简单的博弈。

我们将这些信息列于下表：

C 的选择：	左	右
R 的最优反应：	上	下
R 的选择：	上	下

C 的最优反应：　　左　　右

注意，若 C 认为 R 会选择上，则 C 会选择左；**而且**若 R 认为选择左，则 R 会选择上。因此（上，左）这对选择是相互一致（mutually consistent）的，因为这对选择是由每个选手对对方选择的最优反应组成的。

考虑一个更一般的两人博弈，其中 R 的选择为 $r_1,\ldots r_R$，C 的选择为 $c_1,\ldots c_C$。对于 R 的每个选择 r，令 $b_c(r)$ 表示 C 的一个最优反应；对于 C 的每个选择 c，令 $b_r(c)$ 表示 R 的一个最优反应。则一个纳什均衡为满足下列条件的一对策略 (r^*, c^*)：

$$c^* = b_c(r^*)$$

$$r^* = b_r(c^*).$$

选择的“相互一致性”其实就是指的是纳什均衡。若 R 预期 C 会选择左，则 R 会选择上，而且若 C 预期 R 选择上，则 C 会选择左。因此，纳什均衡时，选手之间的**信念**（beliefs）和**行动**（actions）实现了相互一致。

注意，有些情形下，其中一个选手可能对他的几个最优反应是无差异的。这也就是为什么我们只要求 c^* 是 C 的**一个**最优反应，r^* 是 R 的**一个**最优反应即可。如果每个选择的最优反应是唯一的，那么最优反应曲线可用最优反应**函数**表示。

这种看待纳什均衡概念的方式，使我们更容易明白纳什均衡只不过是第 27 章介绍的古诺模型的一般形式。在古诺模型中，选择变量是产量，它是个连续变量。古诺均衡具有下列性质：给定其它企业的选择，每个企业选择的是能使利润最大化的产量。

我们在第 27 章还介绍过伯特兰模型，该模型是价格策略的纳什均衡。给定其它企业的选择，每个企业选择能使利润最大化的价格。

这些例子表明最优反应曲线是以前模型的一般形式，借助最优反应曲线我们可以更容易求出纳什均衡解。这些性质使得最优反应曲线成为求解博弈均衡的非常有用的工具。

29.2 混合策略

下面我们用最优反应曲线分析表 29.2 所示的博弈。

		Ms. Column	
		Left	Right
Mr. Row	Top	2, 1	0, 0
	Bottom	0, 0	1, 2

表 29.2：求解纳什均衡。

我们对混合策略均衡和纯策略均衡同样感兴趣，因此，令r表示 R 选择上的概率，$1-r$表示他选择下的概率；类似地，令c表示 C 选择左的概率，$1-c$表示他选择右的概率。由纯策略的定义可知，当r和c等于 0 或 1 时，就是纯策略。

下面我们计算如果 R 以概率r选择上并且 C 以概率c选择左时，R 的期望收益。请看下表：

组合	概率	R 的收益
上，左	rc	2
下，左	$(1-r)c$	0
上，右	$r(1-c)$	0
下，右	$(1-r)(1-c)$	1

为了计算 R 的期望收益，我们将上表第 3 列的收益乘以以第 2 列的相应概率（权重），然后相加，可得

$$\text{R 的收益}=2rc+(1-r)(1-c)\text{，}$$

整理可得

$$\text{R 的收益}=2rc+1-r-c+rc\,.$$

现在假设 R 考虑将r提高Δr。收益将会怎样变化？

$$\text{R 的收益变化}=2c\Delta r-\Delta r+c\Delta r=(3c-1)\Delta r.$$

当$3c>1$时，上述表达式为正，当$3c<1$时为负。因此，当$c>1/3$时，R 会提高r；当$c<1/3$时，R 会降低r；当$c=1/3$时，$0\le r\le 1$中的任何r值对 R 都是无差异的。

类似地，C 的收益为

$$\text{C 的收益}=cr+2(1-c)(1-r)\,.$$

当c变动Δc时，C 的收益变化为

$$\text{C 的收益变化}=r\Delta c+2r\Delta c-2\Delta c=(3r-2)\Delta c.$$

因此，当$r>2/3$时，C 会提高c；当$r<2/3$时，C 会降低c；当$r=2/3$时，$0\le c\le 1$中的任何c值对 R 都是无差异的。

我们可以利用这些信息画出最优反应曲线。我们从 R 开始分析。若 C 选择$c=0$，R 会使r足够小，因此$r=0$是对$c=0$的最优反应。这个选择将一直为最优反应直到$c=1/3$时；$c=1/3$时，$0\le r\le 1$中的任何r值都是最优反应；当$c>1/3$时，R 的最优反应为$r=1$。

 曹乾（东南大学 caoqianseu@163.com）

这些曲线如图 29.1 所示。容易看出，它们在三个地方相交：$(0,0)$，$(2/3,1/3)$ 和 $(1,1)$，它们分别对应着该博弈的三个纳什均衡解。这三组策略中有两组是纯策略，一组是混合策略。

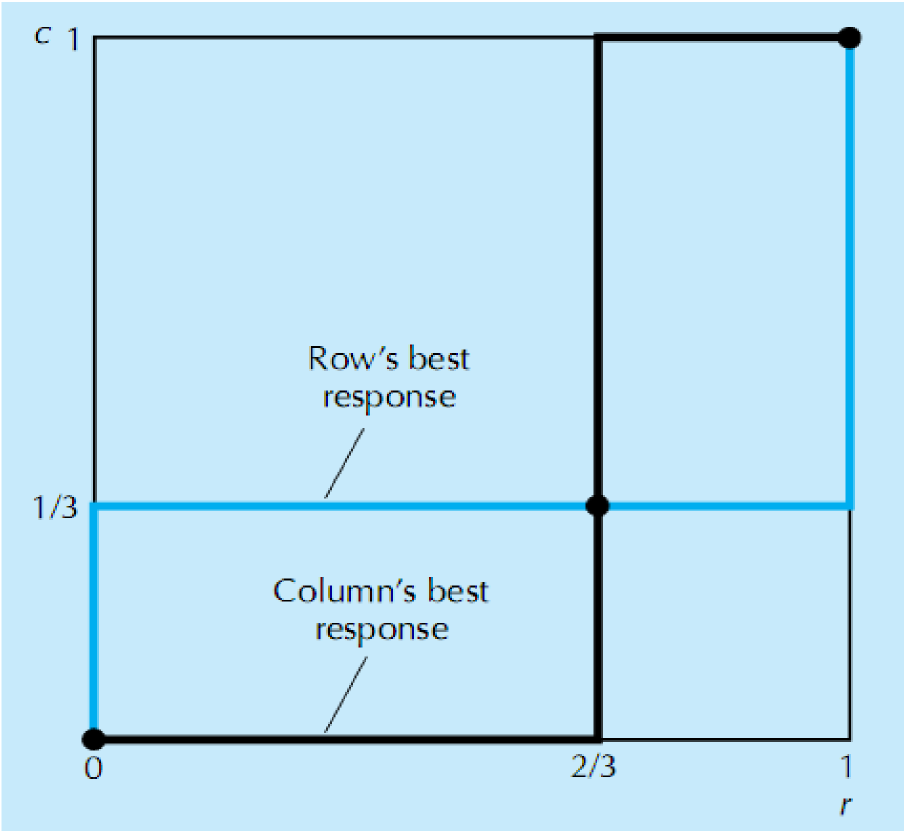


图 29.1：最优反应曲线。蓝色曲线表示选手 R 的最优反应曲线，黑色曲线表示选手 C 的最优反应曲线。这两条曲线的交点均为纳什均衡解。在这个例子中，有三个均衡解，其中两个是纯策略均衡解，一个是混合策略均衡解。

29.3 合作博弈

利用最优反应曲线，我们现在开始分析**合作博弈**（coordination games）。合作博弈是指当选手能调和他们的策略时，他们的总收益最大。在实践中，这类博弈的问题是要找到能使选手合作的机制。

性别大战

合作博弈中的一个经典例子就是所谓的性别大战（battle of the sexes）。在这个博弈中，一个男孩和一个女孩希望约会去看电影，但是没约好看哪一步电影。他们又没带手机，因此

无法商量看什么样的电影，只能猜测对方的偏好。

		Girl	
		Action	Art
Boy	Action	2, 1	0, 0
	Art	0, 0	1, 2

表 29.3：性别大战。

男孩想看新上映的动作片，而女孩则喜欢看文艺片，但是与不约会相比，他们宁愿看同样的电影。这些偏好的相应收益如表 29.3 所示。注意合作博弈的特征：合作时的收益比不合作时的收益高。

这个博弈的纳什均衡是什么？幸运的是，这个博弈恰好就是我们上一节用于介绍最优反应曲线的那个博弈。我们已经知道该博弈有三个均衡解：两人都选择动作片，两人都选择文艺片，或者每人以 2/3 的概率选择他（她）喜欢的电影。

由于这三个解都是可能的均衡解，因此仅靠上述信息我们很难确定他们具体会选哪一个。一般来说我们需要依靠博弈以外的因素来确定该博弈的解。比如，假设播放艺术片的影院离其中一人较近等。因此，这两个人可能认为一起看艺术片就是均衡的选择。

当所有选手有合理的理由认为其中一个均衡比其他均衡更“自然”，则这个均衡称为该博弈的一个**焦点**（a focal point）。

囚犯的两难

我们上一章分析的囚犯的两难博弈，也是一个合作博弈。我们已知道这个博弈说的是：每个囚犯都可以选择认罪从而供出了另一方，或者选择不认罪。该博弈的收益矩阵如表 29.4 所示。

		Player B	
		Confess	Deny
Player A	Confess	−3, −3	0, −6
	Deny	−6, 0	−1, −1

图 29.4：囚犯的两难

这个博弈的突出特征是认罪是占优策略，尽管合作（双方都不认罪）的总收益更大。合作能使得每个罪犯的收益最大，但问题是在只有一轮的博弈中，几乎找不到能让他们合作的方法。

囚徒两难的一种解决方法是拓展该博弈，比如增加新的选择。我们在上一章已知道，无限次重复的囚犯两难博弈，可以通过以牙还牙策略实现合作的结果。以牙还牙指的是在将来的行动中奖励合作惩罚背叛。该情形下，新增加的选择就是今天的不合作将会导致后来的惩罚。

另外一种“解决”囚犯的两难问题的方法，是选手之间可以签订合同彼此约束，和上面类似，这也是为原博弈增加备选策略的一种方法。例如，两个选手在合同中约定彼此坚持选择合作的策略。如果任何一方违约，则需要支付罚金或者受到其他惩罚。合同有助于实现博弈的各种结果，但是合同依赖于法律系统，即需要法律保证这些合同的约束力。商业合同的约束力通常可由法律保证，但是对于军事博弈或国际谈判这类情形，似乎没有合适的法律能够约束人们的行为。

保证博弈（assurance games）

1950 年代美国（U.S.）和前苏联（U.S.S.R）展开军事竞赛，每个国家都可以选择制造原子弹或者不制造。这些策略的收益如表 29.5 所示。双方最好的结果是不制造原子弹，此时收益为（4,4）。但是如果一个不制造而另外一个制造，那么制造者得到的收益为 3，而不制造者得到的收益为 1。双方都制造时的收益为（2,2）。

		U.S.S.R.	
		Refrain	Build
U.S.	Refrain	4, 4	1, 3
	Build	3, 1	2, 2

图 29.5：军事竞赛博弈

不难看出，该博弈有两个纯策略纳什均衡（不制造，不制造）和（制造，制造）。然而，（不制造，不制造）对于双方更有利。问题在于，哪一方都不知道对方的策略选择。在承诺不制造之前，每个国家都希望对方率先做出不生产的保证。

做出保证的一种方法是，其中一个国家率先做出表率，例如允许对它进行检查。注意，这种行为可以是单方面的，至少在它相信该博弈的收益时它会这么做。如果一个国家宣布它

停滞制造原子弹，并且能够向对方展示足够的停产证据，那么它就会相信对方也会停止制造，因为此时的收益(4,4)是最大的。

谁是懦夫？

我们分析的最后一种合作博弈叫懦夫博弈（chicken game）[一]，你在电影中经常会看到这样的场景：两个小伙（Row 和 Column，以下分别简称 R 和 C）分别从街道的两头，开着车笔直地向对方冲去。谁首先转弯谁就被讥笑为懦夫；如果两个人都不转弯，那么他们就会撞车。可能的收益见表 29.6。

这个博弈有两个纯策略纳什均衡（R 转弯，C 转弯）和（C 转弯，R 不转弯）。C 更喜欢第一个均衡，而 R 更喜欢第二个。注意该博弈和保证博弈的区别，在前面的那个保证博弈中，我们知道，两个选手的选择相同时（都制造原子弹或都不制造原子弹），它们各自的状况比其他选择更好。然而，在懦夫博弈中，两个选手的选择相同时（都不转弯或都转弯），它们的状况比其他选择更差。

每个选手都知道，如果他承诺绝不转弯，另外一个选手就会妥协。但是当然，每个选手也都知道撞车是很疯狂的事情。因此，其中一个选手如何实施他喜欢的均衡？

一种重要的策略是承诺。假设 R 夸张地将方向盘锁定从而只能直行。C 认识到 R 除了直行之外别无选择，所以 C 会选择转弯。当然，如果两个选手都锁定了方向盘，结果将是灾难性的撞车！

		Column: Swerve	Column: Straight
Row	Swerve	0, 0	−1, 1
	Straight	1, −1	−2, −2

图 29.6:懦夫博弈。

如何合作？

如果你是某个合作博弈的选手，你可能希望对方与你合作：在合作博弈中，你希望对方选择你们都偏爱的均衡；在性别大战博弈中，你希望对方选择你们其中一方偏爱的均衡；在囚徒的两难博弈中，你希望对方选择的不是实现均衡的策略（即希望对方不认罪）；在懦

[一] 国内教科书通常叫做“斗鸡博弈”，显然是误译。不过联系上下文可知，这个误译不是太离谱。

夫博弈中，你希望对方做出选择从而达到你喜欢的结果。

在保证博弈、性别大战博弈和懦夫博弈中，实现合作的方法是其中一个选手率先行动，并且他向对方承诺坚持某个既定的选择。对方于是可以观察他的行为，从而相应做出反应。在囚犯的两难博弈中，这种方法不可行：如果一个选手选择不认罪，则对方会选择认罪。“解决”囚犯的两难问题的主要方法是重复博弈和签订合同。

29.4 竞争博弈

竞争博弈是与合作博弈相反的另外一个极端。这种博弈就是有名的零和博弈（zero-sum games），这是因为在这种博弈中，一方的收益等于另一方的损失。

绝大多数体育比赛都是零和博弈：一个队伍得到一分，另外一个队伍就失去一分。这种博弈的竞争非常激烈，因为不同选手的利益正好相反。

我们以足球为例进行分析。选手 R 来踢罚球点球，选手 C 防守。R 可以踢向球门左方也可以踢向右方；为了扑出点球，C 可以扑向左方也可以扑向右方。

我们以期望得分（期望进球机率或期望扑出机率）表示这些策略的收益。显然如果 C 扑错了方向，R 的收益会更高。另外一方面，这个博弈可能不是完全对称的，因为 R 可能只擅长向某一个方向踢，C 也可能只擅长向某个方向扑。

假设 R 踢向左，若 C 扑向右时，R 进球的机率为 80%，但 C 也扑向左时，R 进球的机率只有 50%；R 踢向右，若 C 扑向左时，R 进球的机率为 90%，但 C 也扑向右时，R 进球的机率只有 20%。这些收益如表 29.7 所示。

		Column	
		Defend left	Defend right
Row	Kick left	50, –50	80, –80
	Kick right	90, –90	20, –20

表 29.7：足球赛中点球的博弈

注意，每个小方格中的两人收益之和为零，这表明选手得分正好完全相反。R 和 C 都希望使得自身的期望收益最大化，换句话说，就是使对方的收益最小化。

显然，如果 C 知道 R 的踢球方向，C 将具有明显优势。意识到这一点，R 会尽量让 C

猜测。具体地说，R 有时会踢向自己擅长的方向，有时踢向自己不擅长的方向。也就是说他的测了是一种**混合策略**（a mixed strategy）。

假设 R 以概率 p 踢向左，若 C 扑向右，则 R 的期望收益为 $50p+90(1-p)$，若 C 扑向右，则 R 的期望收益为 $80p+20(1-p)$。R 极力使自己的期望收益尽量大，C 则极力使 R 的期望收益尽量小。

例如，假设 R 选择踢向左的概率为 50%。如果 C 扑向左，R 的期望收益为 $50\times1/2+90\times1/2=70$，若 C 扑向右，则 R 的期望收益为 $80\times1/2+2\times1/2=50$.

C 当然也可以进行类似推理。若 C 相信 R 踢向左的概率为 50%，则 C 会扑向右，因为这样的选择能使 R 的期望收益最小（因此最大化 C 自己的期望收益）。

图 29.2 表明了 R 选择不同概率值 p 时的期望收益。画出这个图并不难，只要画出函数 $50p+90(1-p)$ 和 $80p+20(1-p)$ 的图形即可。由于这两个函数都是 p 的线性函数，因此它们的图形都是直线。

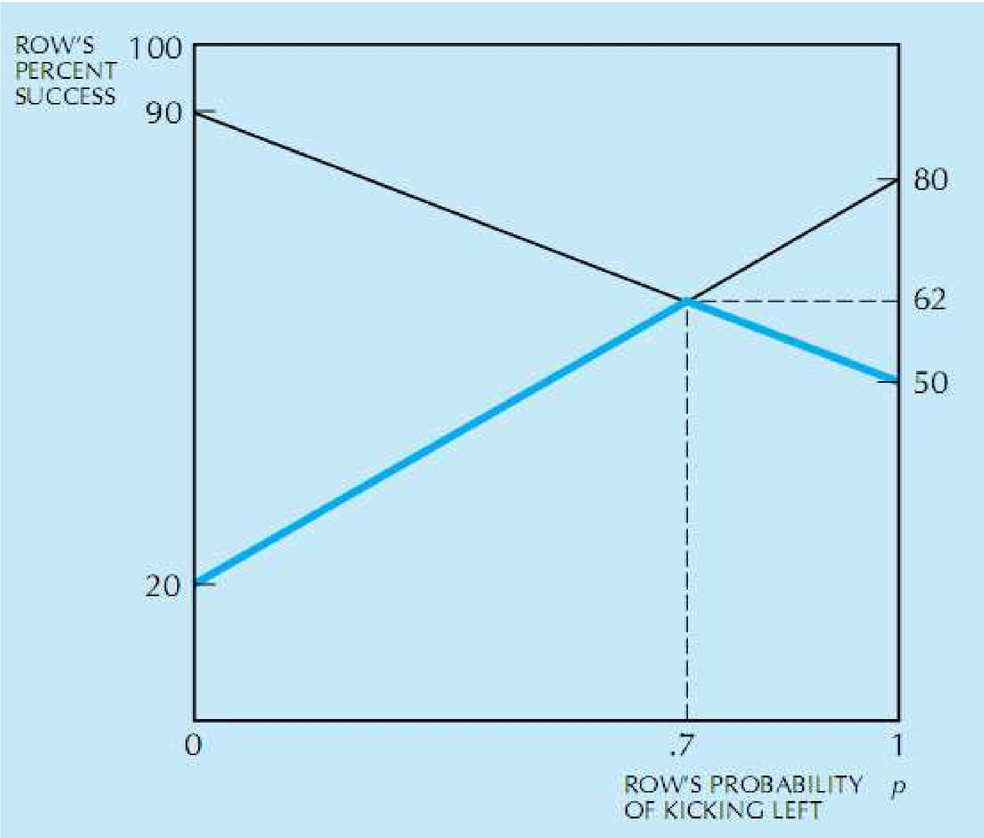


图 29.2：R 的策略。两条曲线表明，R 的期望收益是 p 的函数，其中 p 为他踢向左方的概率。不管 R 选择怎样的 p 值，C 都会极力使得 R 的收益最小。

R 知道 C 会极力使得 R 的期望收益最小。因此，对于任何概率水平 p，R 能获得的最大收益是他的两个策略的收益值中**最小**的那个。我们用黑色粗线段表示这些最小值。

这些最小收益中的最大值在哪个地方出现？显然，它位于黑色粗线段的顶点，或者说位于这两条直线相交之处。

我们可以使用代数方法求解此时的概率值 p，

$$50p + 90(1-p) = 80p + 20(1-p)$$

由此可得 $p = 0.7$.

因此，若 R 以 70%的概率踢向左，而且 C 以做出了最优的反应，则 R 的期望收益为 $50 \times 0.7 + 90 \times 0.3 = 62$。

C 的结果如何？我们可以对 C 的选择进行类似分析。假设 C 以概率 q 扑向左方，以概率 $1-q$ 扑向右方。因此，C 扑向左方时，R 的期望收益为 $50q + 80(1-q)$；C 扑向右方时，R 的期望收益为 $90q + 20(1-q)$。对于每个概率水平 q，C 都会极力使得 R 的收益**最小**，但是 C 也知道 R 会尽量使得这个收益**最大**。

因此，若 C 以概率 1/2 扑向左方：他知道若 R 踢向左，则 R 的期望收益为 $50 \times 1/2 + 80 \times 1/2 = 65$；若 R 踢向右，则 R 的期望收益为 $90 \times 1/2 + 20 \times 1/2 = 55$。在这种情形下，R 当然会选择踢向左方。

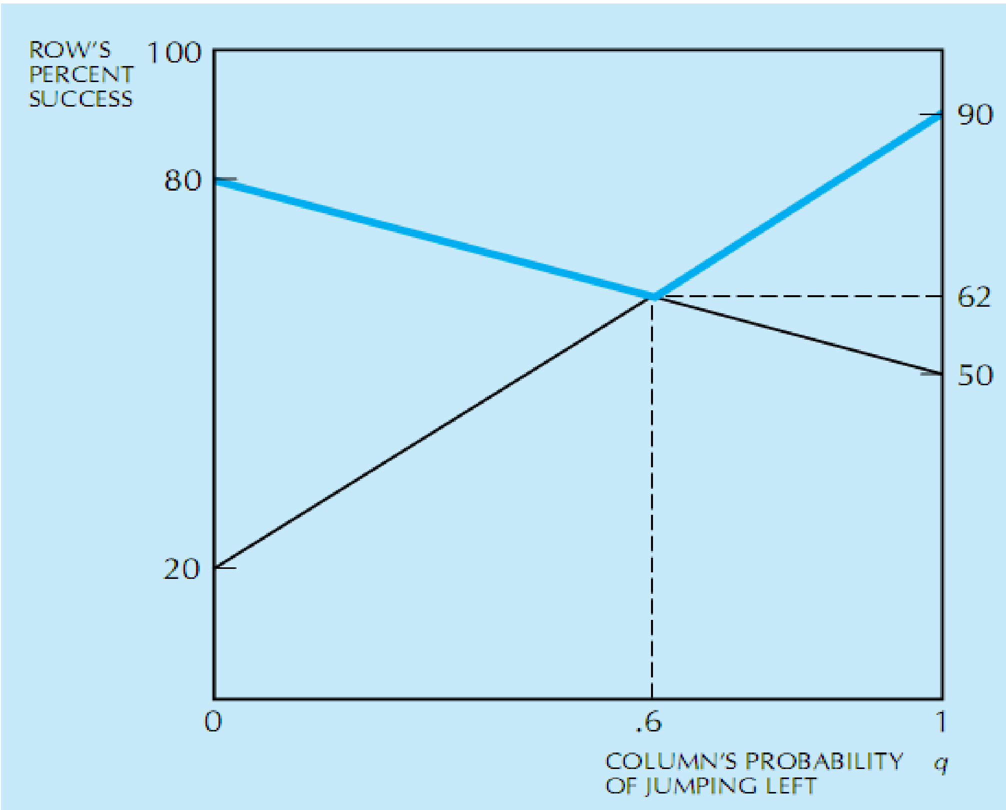


图 29.3：C 的策略。两条曲线表明，C 的期望收益是 q 的函数，其中 q 为他扑向左方的概率。不管 C 选择怎样的 q 值，R 都会极力使得他自己的收益最大。

我们可以画出这两条收益曲线，如图 29.3 所示，这个图和 29.2 类似。从 C 的观点来看，他关注的是这两条曲线的最大值，因为最大值反映了对于 C 所选择的概率 q，R 的最优反应。因此，我们用黑丝粗线段表示这些最大值。和前面一样，我们可以找到 C 的最优 q 值，该 q 值恰好使得 R 从他的最大收益集中选出了最大的那个。这个最大值发生在

$$50q + 80(1-q) = 90q + 20(1-q)$$

由此可得 $q = 0.6$

我们已经计算出了这两个选手 R 和 C 的均衡策略。R 应该以概率 0.7 踢向左，C 应该以概率 0.6 扑向左。此时，不管对方选择何种策略，R 和 C 的期望收益是相同的，因为我们计算这两个概率的方法就是令对方两个策略的期望收益相等。

因此，当 R 选择以概率 0.7 踢向左，则 C 在扑向左还是扑向右之间是无差异的，或者说此时 C 可以以任意概率 q 扑向左。特别低，C 最喜欢以概率 0.6 扑向左。

类似地，若 C 以概率 0.6 扑向左，则 R 在踢向左还是踢向右之间是无差异的，或者说他可以随便扑向左或右。特别地，R 最喜欢以概率 0.7 踢向左。因此，（R 以概率 0.7 踢向左，C 以概率 0.6 扑向左）是一个纳什均衡：给定对方的选择，每个选手选择的策略都是最优的。

在均衡时，R 进球的概率为 62%、不进球的概率为 38%。如果 C 的反应是最优的，这是 R 能实现的最好结果。

如果 C 的反应不是最优的，结果会如何？R 能实现更大的收益吗？为了回答这个问题，我们需要使用本章一开始介绍的最优反应曲线。我们已经知道，当 p 小于 0.7 时，C 会扑向左；当 p 大于 0.7 时，C 会扑向右。类似地，当 q 小于 0.6 时，R 会踢向左；当 q 大于 0.6 时，R 会踢向右。

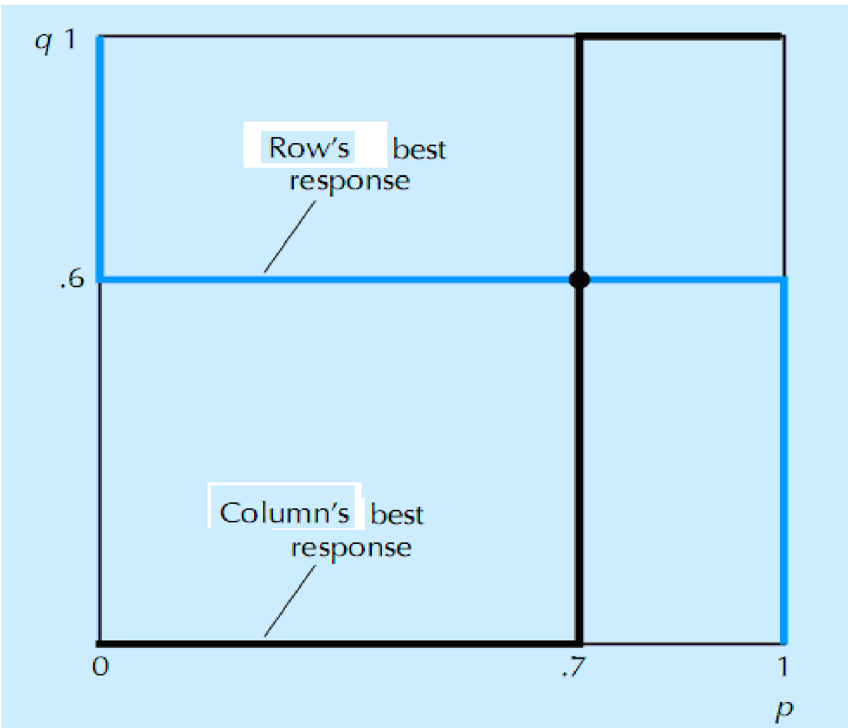


图 29.4：**最优反应曲线**。图中分别画出了 R 和 C 的最优反应曲线。R 的最优反应曲线是他踢向左的概率 p 的函数；C 的最优反应曲线是他扑向左的概率 q 的函数。

图 29.4 画出了 R 和 C 的最优反应曲线[一]。注意，这两条曲线在 $p=0.7$ 且 $q=0.6$ 时相交。最优反应曲线的优点在于，不论对方的选择是否是最优的，它都能告诉选手自己如何进行选择。对最优反应的最优反应，就是这两条曲线相交之处，在这一点上实现了纳什均衡。

29.5 共存博弈

在前面，我们将混合策略解释为选手随机选择策略。在罚点球的博弈中，若 R 的策略是以 0.7 的概率踢向左方、以 0.3 的概率踢向右方，则我们认为 R 将“混合使用他的策略”：70%的时间内踢向左方、30%的时间内踢向右方。

但是，还有另外一种解释。假设我们随机地选择一批罚球选手，而且假设这些选手中 70%的人总是踢向左方，30%的人总是踢向右方。那么，从守门员的角度来看，这相当于他面对这样一个罚球手：他以 70%的概率踢向左方、以 30%的概率踢向右方。

这种解释对于罚球博弈来说不太合理，但是对于动物行为来说则是合理的。此处的思想是，动物的各种行为来自遗传，能够生存下来的种群通常具有稳定的进化能力。近些年来，生物学家在分析动物行为时，博弈理论称为不可或缺的研究工具。

动物间相互作用的最著名博弈莫过于鹰鸽博弈（hawk-dove game）。这个博弈不是指鹰和鸽子之间的博弈（如果是，结果很容易预测），而是指的是具有两类行为的同一物种之间的博弈。

我们以野狗为例进行分析。当两只野狗同时遇到食物时，它们必须决定是争斗还是分享。争斗是鹰派的策略：一胜一输。分享是鸽派的策略：当对方也是鸽派时，分享是个好策略，但如果对方是鹰派，分享的提议就会被拒绝，鸽派将一无所得。

表 29.8 给出了可能的一组收益。

[一] 注意，英文原版书在图 29.4 中正好将 R 和 C 的反应曲线标注错了。我们此处修正了这个错误。

		Column: Hawk	Column: Dove
Row	Hawk	–2, –2	4, 0
	Dove	0, 4	2, 2

表 29.8：鹰鸽博弈。

如果两只野狗都愿意分享，则它们的收益为（2,2）。若一只猎狗想争斗，而另外一只想分享，则争斗者夺取全部食物。但是如果两只野狗都愿意争斗，那么它们将受伤严重。

显然，如果每只野狗都想争斗，则无法实现均衡，因为如果一只猎狗将争斗的策略改为分享，它的收益为 0 而不是-2。而且如果两只猎狗都愿意分享，也很难实现均衡，因为若一只猎狗将分享策略改为争斗，它将占有全部食物。因此，如果我们站在野狗种群的角度，那么均衡时，必然有些野狗是鹰派（争斗），另外一些野狗是鸽派（分享）。问题是它们之间的比例是多大？

假设鹰派野狗占野狗总数的比例为 p。因此，一只鹰派野狗遇到另外一只鹰派野狗的比例为 p，而遇到一只鸽派野狗的比例为$1-p$。鹰派野狗的期望收益为

$$H=-2p+4(1-p).$$

鸽派野狗的期望收益为

$$D=2(1-p).$$

假设哪一类的收益高，哪一类野狗的繁殖能力就越快，而且还会将这种倾向遗传给后代。那么，如果 $H>D$，那么种群中鹰派的比例就会上升；如果 $H<D$，那么鸽派的比例就会上升。

野狗种群中的唯一均衡出现在鹰派收益和鸽派收益相等时，即

$$H=-2p+4(1-p)=2(1-p)=D,$$

由此可解得 $p=1/2$。

我们已经知道均衡时，鹰派和鸽派的比例为 50%-50%。这个均衡在某种意义上是一个稳定的均衡吗？我们在图 29.5 中画出鹰派和鸽派的收益曲线，它们都是概率 p 的函数，其中 p 表示鹰派的比例。注意，当 $p>1/2$ 时，鹰派的收益小于鸽派的收益，因此我们可以预期，鸽派的繁殖速度更快，这使得我们又回到 50%-50%这个均衡比例。类似地，当 $p<1/2$

时，鹰派的收益大于鸽派的收益，使得鹰派的繁殖速度更快，这又回到上述均衡比例。

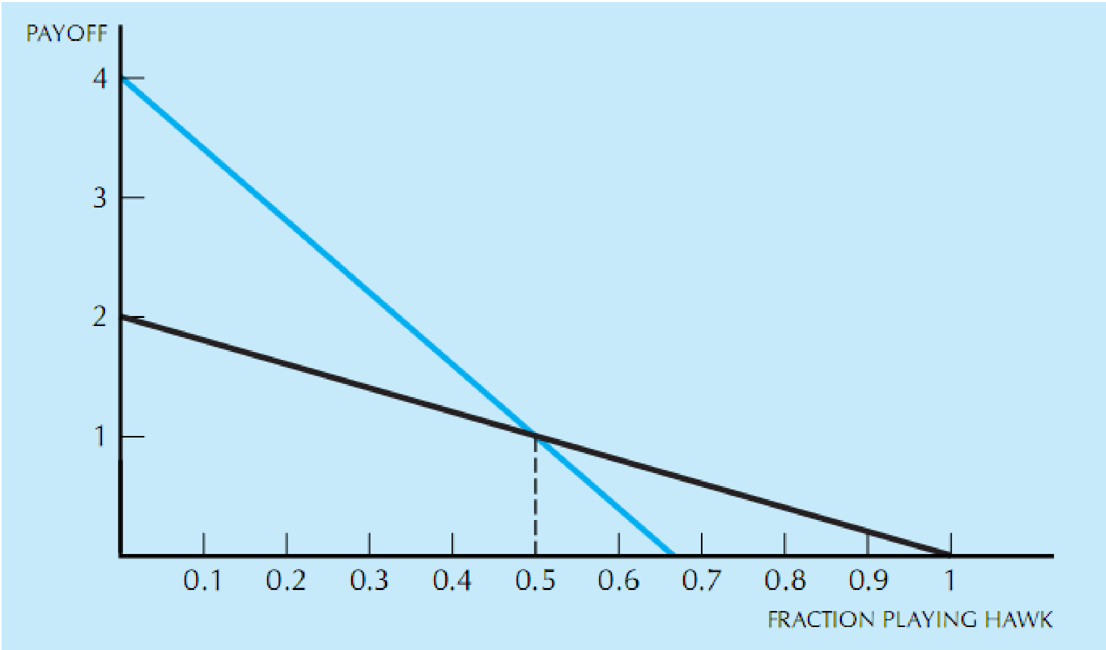


图 29.5：鹰鸽博弈中的收益。鹰派的收益曲线以红色直线表示，而鸽派的收益以黑色直线表示。当 $p>1/2$ 时，鹰派的收益小于鸽派的收益，反之则反是，这表明该均衡（0.5,0.5）是稳定的。

这个论证表明 $p=1/2$ 不仅是一个均衡解，而且在进化力的作用下，它也是一个稳定均衡解。这类分析引出了一个称为**进化稳定策略**（evolutionarily stable strategy, ESS）[一]。明显地，一个进化稳定策略就是一个纳什均衡，尽管我们是用种群而不是用个体推导出这个进化稳定策略的。

纳什均衡的概念旨在计算出下列情形下的均衡解：两个精于算计的理性人的博弈中，每个人试图找出自己的最优策略来应对对方的最优策略。而进化稳定博弈则是用来分析动物种群在进化力作用下的行为，哪一类（比如上例中的鹰派或鸽派）的收益更大，哪一类的繁殖速度将更快。尽管存在应用环境的差别，进化稳定博弈也是纳什均衡，这再次说明纳什均衡这个概念在博弈论中的地位有多么重要。

29.6 承诺博弈

前文介绍的合作博弈和竞争博弈的各个例子，都是**同时行动**（simultaneous moves）的博弈。在不知道对方正在或者已经选择出何种策略的情形下，每个选手必须做出自己的选择决策。的确，在合作博弈或竞争博弈中，如果一个选手知道对方的选择，那么博弈的结果将非常容易解得。

[一] See John Maynard Smith, Evolution and the Theory of Games, (Cambridge University Press, 1982).

在本节，我们将分析**序贯行动**（sequential moves），即选手行动顺序有先后之分而不再是同时。这类博弈中的一个重要策略称为**承诺**（commitment）。为了看清承诺是如何运行的，我们可以再回头看看上一章介绍的懦夫博弈。在懦夫博弈中，如果一个选手能强迫自己选择直行而不是转弯，另外一个选手的最优选择是转弯。在保证博弈中，如果一方能率先行动，那么博弈的结果将对双方选手都更有利。

注意，这个承诺必须是不可撤销（即不可反悔）的，而且必须是能被对方观测到。承诺本身就意味着不可撤销，否则还叫什么承诺。但是，承诺的可观测性非常重要，因为只有这样才能让对方改变行为。

青蛙和蝎子

我们从一个关于青蛙（frog）和蝎子(scorpion)的寓言故事开始分析。青蛙和蝎子站在河岸上，它们想找到过河的方法。“我想出了一个办法”，蝎子说，“我爬到你的背上，你游过河去。”青蛙回答说，“但是如果你蛰我，我怎么办？”蝎子说，“我为什么会蛰你呢？那样我们都会被淹死。”

青蛙发现这种方法可行，因此蝎子爬上了青蛙的后背，它们开始过河。到了河中央，也就是河水最深处，蝎子蛰了青蛙一下。青蛙疼得打了一个滚，叫道，“你为什么要蛰我？现在我们都要完蛋了！”“哎呀，”蝎子后悔地说，“这是我的天性。”青蛙和蝎子沉入了水底。

下面我们使用博弈论分析这个寓言故事。根据此故事，我们可以给出这个序贯博弈的收益，如图 29.6 所示。从博弈树的根部开始分析。如果青蛙拒绝了蝎子，它们的收益都为零。看图中蝎子不蛰青蛙的这根分支，此时青蛙因为做了件好事，得到的效用为 5，而蝎子因为过了河，得到的效用为 3。而若蝎子蛰青蛙时，青蛙得到的收益为-10，蝎子得到的效用为 5，因为蝎子从它的天性行为（蛰别人）中得到了满足。

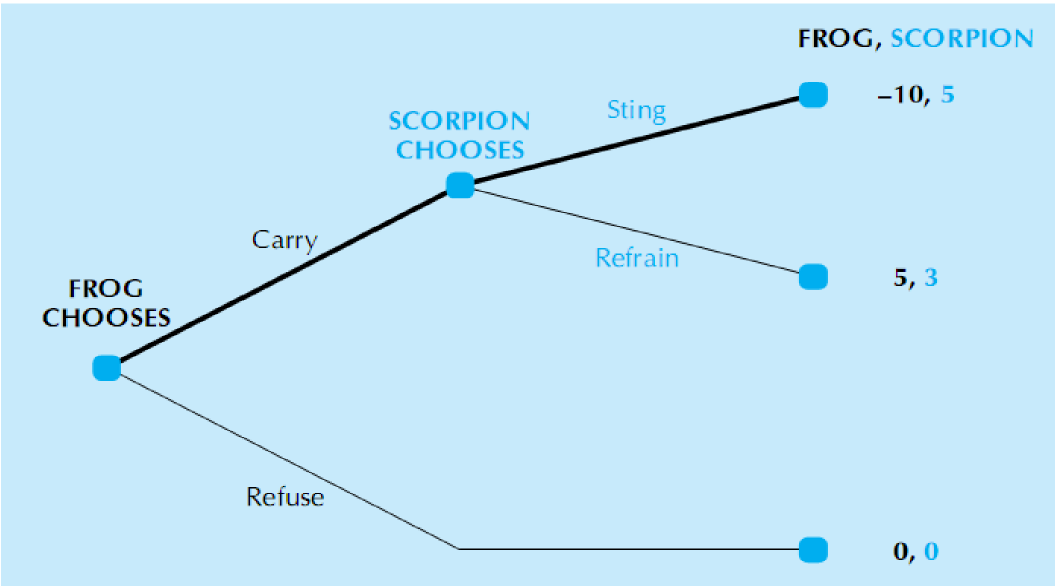


图 29.6：**青蛙和蝎子的博弈**。若青蛙选择背着蝎子过河，蝎子会选择蛰青蛙，它们都会被淹

死。

最好从该博弈的最后一步进行分析：蝎子可以选择蛰或不蛰。选择蛰的收益较高，因为蛰是蝎子的“天性”。因此，青蛙的理性选择是解决蝎子的提议。不幸的是，青蛙没有认识到蝎子的收益；显然，他错误地认为蝎子的收益为图 29.7 中的的情形。这个错误对于青蛙来说是致命的。

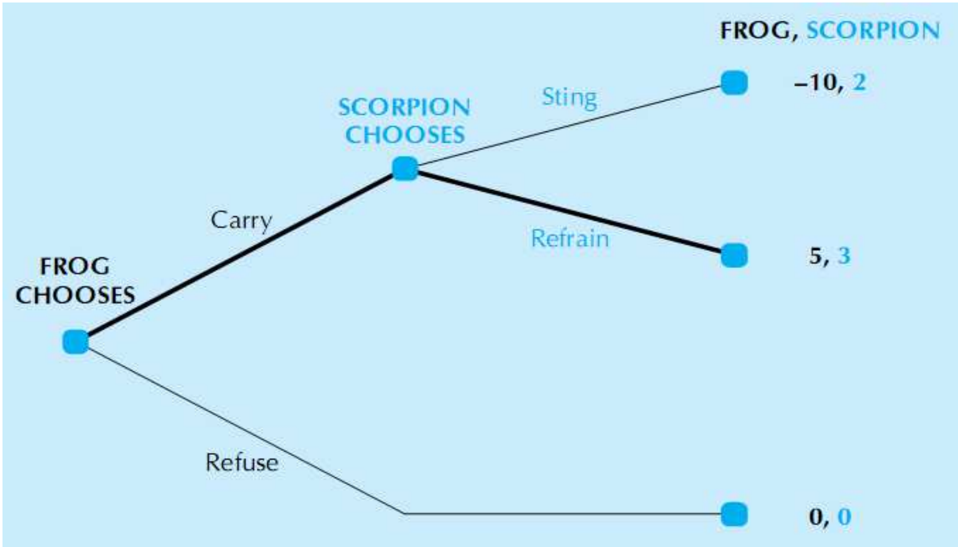


图 29.7：青蛙眼里的（错误的）博弈收益。在这样的博弈收益情形下，如果青蛙选择背着蝎子过河，蝎子不会蛰青蛙，它们因此能安全过河。

青蛙的明智选择是找到让蝎子做出承诺不蛰的方法。例如，将蝎子的尾刺绑住。或者它可以雇用青蛙杀手，如果蝎子不履行承诺，则青蛙杀手会对蝎子的家庭进行报复。不管选择什么样的方法，对于青蛙来说，最重要的事情是能改变蝎子的收益，即让蝎子为蛰这种行为付出更大的代价或者大幅度减少蛰的收益。

这些绑匪不太冷

在某些国家，绑架别人索要赎金是一种大生意。在哥伦比亚，据估计，每年发生 2000 例绑架勒索案。在前苏联，1992 年绑架案为 5 例，但到了 1999 年，该数字急剧增加到 105 例。很多受害者是来自西方发达国家的商人。有些国家，比如意大利制定法律禁止支付赎金。立法的原因是如果受害者家庭或者雇员承诺不支付赎金，那么绑匪就没有动机进行绑架。

问题当然在于，一旦绑架发生，受害者的家庭倾向于向绑匪支付赎金，尽管这么做是违法的。因此，对支付赎金的行为进行处罚，不是一种有效的承诺工具。

假设绑匪绑架了某个人质之后发现得不到赎金。那么，他们是否应该释放人质?人质当然会承诺不会揭发绑匪的身份。但是人质能遵守自己的承诺吗？一旦绑匪释放了人质，他没

有守诺的动机——每个人都希望绑匪得到惩罚。即使绑匪愿意释放人质，他们也不能这么做，因为这样做的后果是他们的身份就暴露了。

图 29.8 给出了某些可能的收益。如果绑匪杀掉人质，他内心会不安，因此得到的收益为-3。当然，人质的状况更糟，人质得到的收益为-10。如果绑匪释放人质，而且人质不揭发绑匪，那么人质的收益为 3，绑匪的收益为 5。但是，如果人质揭发了绑匪，人质得到的收益为 5，绑匪得到的收益为-5.

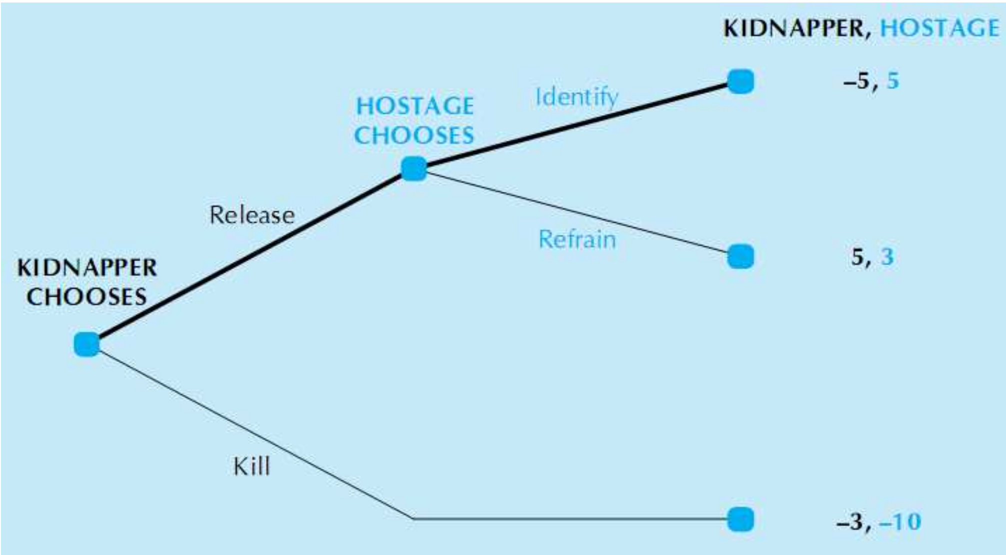


图 29.8：绑架博弈。绑匪可能愿意释放人质，但是如果他们这么做，人质在获得自由后会揭发绑匪的身份。

现在是人质面临承诺问题：他如何说服绑匪相信他不会违约，即不会揭发绑匪的身份？

人质需要找到能改变这个博弈收益的方法。具体地说，他要找到的方法要能做到：若他揭发绑匪，那么他将承担相应的代价。

马里兰大学有位经济学家叫做 Thomas Schelling，他长期致力于研究动态博弈中的策略分析。他的建议是，人质可以让绑匪拍摄他的不雅照片，然后把这些照片交给绑匪保存。这种策略有效地改变了博弈的收益，因为人质在获得自由后，他在考虑是否揭发绑匪时不得不顾虑那些不雅照片。

这种策略称为 “交换人质”(exchange of hostages)。在中世纪[一]，当两个国王希望保证他们之间的合同不被违约时，他们采取的策略通常是交换人质，而这些人质通常是他们的家庭成员。由于任何一个国王都不希望自己的家庭成员送命，因此每个国王都有遵守合同约定的激励。

在绑架博弈中，如果不雅照片流出，人质将会付出代价，因此这种策略保证了人质会

[一] 中世纪（Middle Ages）（约公元 476 年~公元 1453 年），是欧洲历史上的一个时期（主要是西欧）。译者注。

遵守约定，从而不会揭发绑匪的身份。

力大何时成为劣势?

我们下一个例子来自动物心理学领域。研究发现，猪群能迅速建立领导-被领导（dominance-subordinateness）关系，猪领导统治着身边的猪下属。

有些动物心理学家将两只猪放入一个长方形的猪圈内，其中一只猪是领导，另外一只猪是下属[二]。实验人员在猪圈的一端摆放了一个猪食操作杆，触动操作杆将有猪食流出，但是要注意，猪食槽却摆放在猪圈的另一端。实验人员关注的问题是：哪只猪会去触动操作杆，哪只猪能吃到食物?

		Dominant Pig	
		Don't press lever	Press lever
Subordinate Pig	Don't press lever	0, 0	4, 1
	Press lever	0, 5	2, 3

表 29.9：猪领导和猪下属之间的博弈。

实验的结果多少令人惊讶：猪领导触动操作杆，猪下属在猪食槽旁边等待食物流出。因此，猪下属得到了大部分食物，而猪领导在触动操作杆后拼命跑到猪食槽，却只能吃到很少一点残渣。表 29.9 给出了这个博弈的收益。

猪下属比较收益（0,4）和（0,2），断定：不触动操作杆明显好于触动。由于猪下属不去触动操作杆，猪领导别无选择只能亲自去触动操作杆。

如果猪领导不是将全部食物吃光，而是将部分食物奖励给触动操作杆的猪下属，那么它就能得到更好的结果。问题是，这两只猪之间没有合同约束，而且猪领导也很难做到不将食物吃光。

和前面绑架博弈一样，猪领导面临自己做出承诺的问题。如果它能承诺不将所有食物吃光，那么它的状况就会变好。

储蓄和社会保障

承诺问题并不局限于动物世界。它们也存在于现实中的经济政策。

[二] The original reference is Baldwin and Meese, "Social Behavior in Pigs Studied by Means of Operant Conditioning," (Animal Behavior, (1979)). I draw on the description of John Maynard Smith, Evolution and the Theory of Games (Cambridge University Press, 1982).

人们在退休后收入会减少，因此我们会预期人们会提前储蓄。然而，事实上尽管人们都口头上对提前储蓄赞不绝口，但是很少人真得去储蓄。人们不愿意储蓄的部分原因是，他们认为社会不会让他们挨饿，因此在将来，社会很有可能养活他们。

我们将这个问题表述为代际之间的博弈。假设老一代的策略有两个：储蓄或挥霍。年青一代的选择也有两个：赡养老一代和为自己退休进行储蓄。可能的收益矩阵如表 29.10 所示。

		Younger Generation	
		Support	Refrain
Older Generation	Save	3, –1	1, 0
	Squander	2, –1	–2, –2

表 29.10：储蓄的代际冲突

如果老年人自己储蓄而且年青人也会赡养他们，那么老年人的收益为 3，年青人的收益为-1。如果老粘人挥霍乱花而且年青人赡养他们，那么老年人的收益为 2，年青人的收益为-1。

如果年青人不赡养而且老年人自己储蓄，那么老年人的收益为 1，年青人的收益为 0. 最终，如果老年人挥霍而且年青人不赡养，那么各自的收益均为-2，因为这样老年人会挨饿，年青人因旁观而心里会不安，所以收益均为负。

不难看出，这个博弈有两个纳什均衡。如果老年人选择储蓄，那么年青人的最优选择就是不赡养。但是如果老年人选择挥霍，那么年青人的最优选择是赡养他们。当然，如果年青人选择赡养老年人，老年人的最优选择是挥霍！

然而，上面的分析忽略了这个博弈的时间结构：老年人为数不多的好处是能先变老，呵呵。如果我们画出博弈树，收益如图 29.9 所示。

如果老年人储蓄，年青人就会选择不赡养，因此老年人的收益为 1。如果老年人挥霍，而年青人不忍看他们到时挨饿而会赡养时，老年人的收益为 3。由于知道年青人会赡养，所以老年人的明智选择是挥霍。

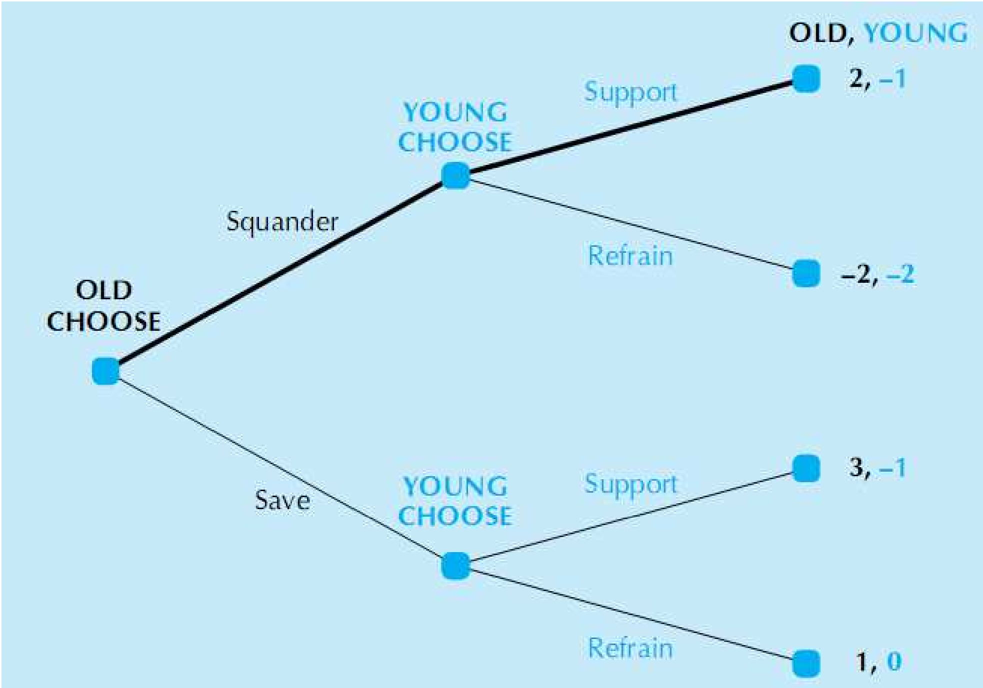


图 29.9：**储蓄博弈的展开形**。由于知道年青一代会赡养，老一代的最优选择是挥霍乱花。子博弈的完美均衡为（挥霍，赡养），均衡时收益为（2，-1）。

当然，大多数发达国家都实施了社会保障计划，强迫每一代人为将来的退休提前进行储蓄。

敲诈（hold up）

考虑下面的策略性互动。你雇佣了一个建筑商为你建造仓库。当建筑方案快要结束时，你认识到颜色不合适，因此你要求建筑商更换油漆，这样的花费非常小。建筑商则说："改颜色可以，请给我 1500 元。"

你认识到，你找到新油漆工会耽误工期，而且你非常喜欢新的颜色，因此，你在低声抱怨几句后，就支付了他 1500 元。恭喜你，你被敲诈了！

当然，在这类博弈中，建筑商也不是唯一的敲竹杠方。客户也可以拖延支付货款，从而"敲诈"建筑商，这当然会让建筑商叫苦不迭。

这个敲诈问题的博弈树如图 29.10 所示。我们假设客户为改换颜色的支付意愿为 1500 元，但是改换颜色的实际成本只有 200 元。我们从树枝方向进行分析，如果建筑商要价 1500 元，那么他的利润为 1300 元，客户的收益为零。

如果客户寻找另外一个油漆工，假设他向该油漆工支付 200 元，由于耽误工期造成的时间成本为 1400 元。他对新颜色的评价为 1500 元，但是这种情形下却支付了 1600 元，因此他的净收益为-100 元。

如果建筑商索要的价格等于更换颜色的实际成本 200 元，那么建筑商不赚不赔，收益为零，客户的净收益为 1300 元，因为他对更换颜色的评价为 1500 元，但却只要支付 200 元。

从图容易看到，建筑商的最优选择就是敲诈，而客户的最优选择是屈服让步。但是明智的客户会认识到，更换颜色这类要求在任何工程中都会出现。因此，客户不会愿意雇佣背负敲诈名声的建筑商，这对这样的建筑商也是不利的。

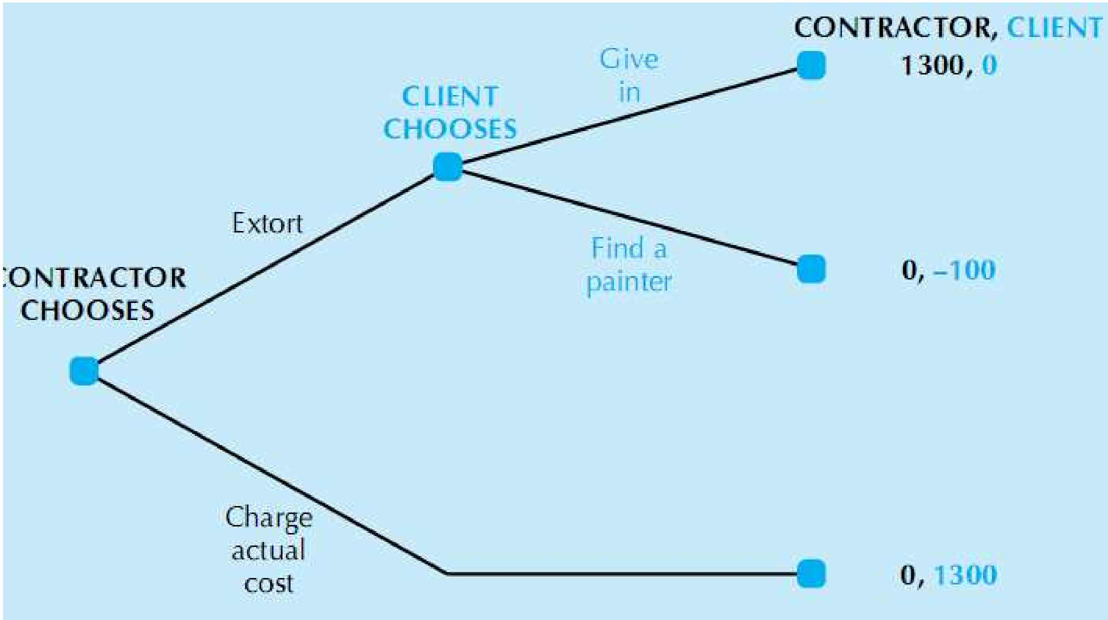


图 29.10：敲诈问题。由于客户别无选择，因此建筑商为更换颜色索要了较高的价格。

企业之间如何解决这类敲诈问题？基本的答案是签订合同。一般来说，建筑商会在合同中注明什么样的更换是允许的以及更换时成本是如何计算的。有时，在这类合同中还制定了仲裁条款或者其他解决争端的方式。制定合同需要花费大量的时间、精力和金钱，但是这是值得的，因为它可以防止这类敲诈问题的发生。

但是，签订合同不是敲诈问题的唯一解决方法。另外一种方法是承诺。例如，建筑商可能会缴纳保证金以保证按照工期施工。当然，双方一般还要约定什么样的情形才能算完工。

另外一个重要因素是名声（reputation）。当然，某个建筑商如果三番五次敲诈客户，那么他的名声就很差。客户不会雇佣这样的建筑商，当然也不会推荐其他客户雇佣。这个名声效应可用重复博弈进行分析，建筑商今天敲诈，将来就会付出代价。

29.7 讨价还价

经典的讨价还价问题（bargaining problem）是分钱问题。两个选手有一元钱，如何分?

如果只给你上面那么少的信息，你无法给出具体的答案，因为这些信息还不足以建立一个合理的模型。建立讨价还价模型的关键，是找到选手之间进行协商的机制。

例如，纳什讨价还价模型（Nash bargaining model），采用的是公理性的证明方法，也就是说，它首先假设合理的解应该具有什么样的性质，然后证明满足这些公理性假设的结果只有一个。最终结果取决于选手对风险的厌恶程度，以及如果没有讨价还价机制将会发生什么事情。不幸地是，这个模型的完整介绍超出了本书的范围。

还有一种模型，称为鲁宾斯坦讨价还价模型（Rubinstein bargaining model）。这个模型的思路是分析一系列选择决策，然后求解子博弈完美均衡（subgame perfect equilibrium）。幸运地是，这个模型的基本思想比较容易说明。

我们举个简单的例子说明。两个人，Alice 和 Bob（以下分别以 A 和 B 表示），共同分享 1 元钱。他们决定最多三天之内找到分钱的方法。第一天，A 提出一种分钱方法，B 可以接受，也可以拒绝。如果拒绝， B 要在第二天提出另外一种分钱方法。如果 A 拒绝了 B 的方法，那么 A 在第三天提出最终的分配方法。如果三天之内，他们无法达成分钱的一致意见，那么他们每个人都得到的钱数为零。

假设 A 和 B 的耐心程度不同：A 对未来收益的贴现率（折扣率）为 α 每天；B 对未来收益的贴现率（折扣率）为 β 每天。最后，我们假设如果一个选手对两种分配方法无差异时，他总是选择对方喜欢的那种分配方法。最后这个假设的思想是，对手可能只分配给另一方很小的钱数，从而使得这个选手严格偏好某个选择。由于这个钱数很小，我们可以近似看为零。可以证明这个讨价还价模型只有唯一一个子博弈完美均衡。

我们从第三天开始分析。由于这是博弈的最后一天。A 提出的分配方案一般为“要么接受，要么走人”。显然，A 的最优选择是分配给 B 能接受的尽量少的钱，根据前面的假设，这个钱数可以近似为零。因此，如果该博弈实际只进行三天，A 会得到 1 元，B 会得到零元（即任意小的钱数）。因为如果 B 拒绝的话，那么说明 A 和 B 最终未达成统一的分配意见，根据前面的假设可知，A 和 B 得到的钱数为零。但是如果 B 接受 A 的分配方案，那么他还能得到一个很小的钱数（尽管很小，还是比零好！）

现在回到第二天，这一天是 B 提出分钱方案。B 认识到如果 A 拒绝了这个方案，那么 A 在第三天就能确保得到 1 元。由于对于 A 来说，第三天的 1 元钱相当于第二天的 α 元。因此，如果 B 提出的方案中分配给 A 的钱数小于 α，那么 A 必定会拒绝 B 的方案。B 显然更喜欢此时的 $(1-\alpha)$ 元，而不是第三天的 0 元。因此 B 应该理性地分配 α 元给 A，A 当然会接受，因为这相当于第三天的 1 元钱。因此，如果该博弈在第二天结束，那么 A 得到 α 元，B 得到 $(1-\alpha)$ 元。

现在回到第一天。第一天是 A 提出分配方案。A 认识到如果 B 拒绝了他的方案，那么

曹乾（东南大学 caoqianseu@163.com）

B 在第二天可以得到$(1-\alpha)$元。因此，为了避免被拒绝，A 分配给 B 的钱数应该等于$(1-\alpha)$元在第一天的现值，即$\beta(1-\alpha)$。B 发现这个钱数（恰好）是他能接受的，因此博弈结束。该博弈的最终结果是，博弈在第一天结束，其中 A 得到的钱数为$1-\beta(1-\alpha)$，B 得到的钱数为$\beta(1-\alpha)$。

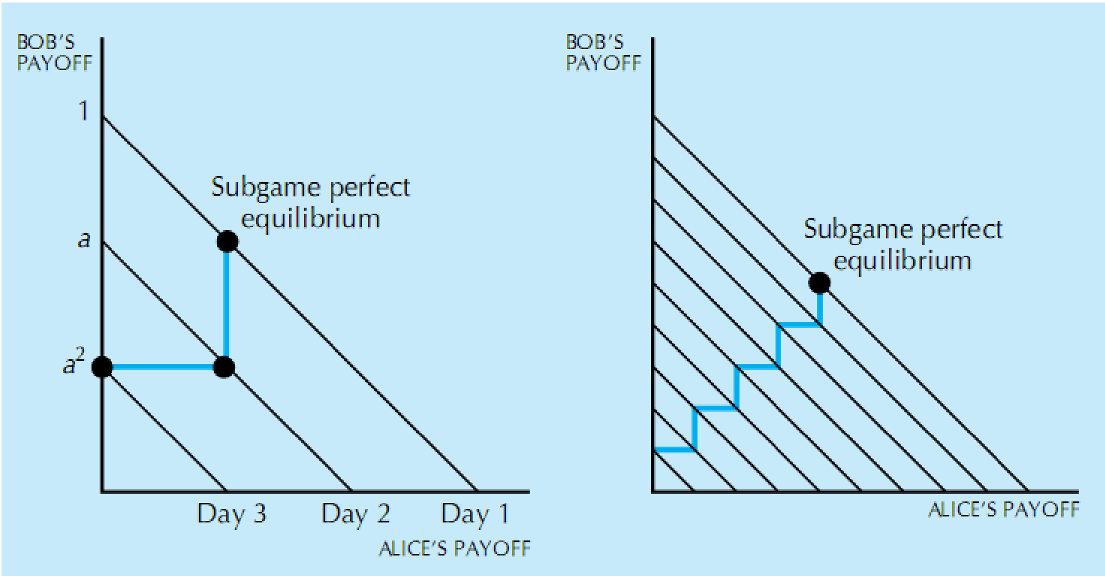


图 29.11：一个讨价还价模型。粗实线将子博弈的均衡均衡结果连在了一起。最外面那条线上的点是子博弈完美均衡。

图 29.11 中左侧的图形表示$\alpha=\beta<1$时的博弈过程。最外面的那条（45 度）直线表示的是，第一天的可能收益模式，即$x_A+x_B=1$。中间的那条直线表示，如果博弈在第二天结束，收益的现值为多少，即$x_A+x_B=\alpha(=\beta)$.最靠近原点的那条直线表示，如果博弈在第三天结束时，收益的现值为多少。这条线的表达式为$x_A+x_B=\alpha^2(=\beta^2)$.图中呈现直角形状的那条路径，表示的是每一天的最小可接受的分配钱数。图 29.11 中右侧的图表示当博弈的天数为很多天时的博弈过程。

我们自然地将博弈次数推广到无穷次，即分析在博弈次数为无穷时，结果是怎样的？可以证明，子博弈完美均衡时的分配额为

$$\text{A 获得}\ \frac{1-\beta}{1-\alpha\beta}$$

$$\text{B 获得}\ \frac{\beta(1-\alpha)}{1-\alpha\beta}$$

注意，如果$\alpha=1$且$\beta<1$，那么 A 得到了全部的钱数。

最后通牒博弈

鲁宾斯坦讨价还价模型的思想是如此优美，因此经济学家急忙做实验进行验证。他们发现，优美并不意味着准确。实验的结果也许意味着，缺乏经验的被实验人员（比如你让非经济专业的学生进行博弈），也许最多只能预测一两步。

而且，可能还有其他因素影响了实验结果。为了看清这一点，假设前面介绍的讨价还价模型只进行一天。A 和 B 仍然要共享 1 元钱。A 提出分配方案，如果 B 接受，博弈结束。问题是 A 应该提出怎样的分配方案？

如果根据单纯的理论分析，A 提出的方案应该为分给自己 99 分钱，分给 B 1 分钱。B 认识到 1 分钱也比 0 好，因此 B 会接受这个方案。A 高兴地回家了，他心里想他学过的经济学看来帮了大忙了。

不幸的是，结果并非如此。一个可能的结果是，由于 B 认为 1 分钱太少了，他会拒绝。因此 A 和 B 都一无所得。A 认识到这一点，会增大分配给 B 的钱数。在实际实验中，美国大学生分配给对方的钱数的平均值为 45 分钱，这种分配方案在大多数情形下都会被对方接受。

制定分配方案的一方是理性的，因为从实验中被拒绝的频率来看，分配 45 分钱给对方，已非常接近于自己最大的期望收益。与理论预期不同的是接受方的行为，因为在现实中，如果分配给他的钱数很小，他会拒绝，即使拒绝后他一分钱也得不到。

人们对理论结果和实验结果的差异提出了多种解释方法。一种观点是，如果分配给对方的钱数太小，则违背了人类行为的**社会准则**（social norms）。的确，经济学家发现，不同文化对最后通牒博弈的实验结果有显著影响。另外一种观点是，接收方拒绝较小的钱数，是对对方的报复，这样他能获得某种效用。毕竟报复的代价是自己损失了 1 分钱，但却让对方失去了 99 分钱，这样一比较，通过报复对方得到的满足感还是相当有诱惑力的。在下一章，我们还将详细分析最后通牒博弈。

总结

1.一个博弈选手的最优反应函数，是说他的最优选择是对方选择策略的函数。

2.两人博弈中的一个纳什均衡是指一对策略，这个**策略对是**由两个选手的最优策略组成，或

者说，给定一方的最优选择，另一方也相应作出了最优选择。

3.一个混合策略纳什均衡，是说均衡时的策略是随机的。

4.常见的合作博弈有以下几种：一是性别大战，其中两个选手的选择是相同的；二是囚犯的两难，其中占优策略使选手两败俱伤；三是保证博弈，其中，只要选手相信对方会合作，那么他也会选择合作；四是懦夫博弈，其中选手的最优策略是避免相同的选择。

5.在两人零和博弈中，一方的收益是另一方收益的相反数，因此收益之和为零，所以叫零和博弈。

6.进化博弈的结果在种群进化力的作用下是稳定的。

7.在序贯博弈中，选手的行动有先后之分、交替进行。因此，每个选手必须考虑对方对自己选择的反应。

8.在很多序贯博弈中，承诺问题是一个非常重要的问题。其中最为关键的是找到对选手守诺的约束方法。

复习题

1.在两人博弈的纳什均衡中，每个选手在制定最优选择时，考虑的最主要因素是什么？在占优策略均衡中，每个选手在制定最优选择时，考虑的最主要因素是什么？

2.复习教材中 C 和 R 在混合策略中的最优反应问题。他们的最优反应是否构成了最优反应函数？

3.（判断）在合作博弈中，如果两个选手作出了同样选择，那么他们对此结果都是满意的。

4.教材中断言，在足球博弈中，均衡时 R 得分的概率为 62%。这个结果是怎么计算出来的？

5.一个建筑商说，他打算“降低报价并在对方要求修改时将这个损失补偿回来”，他这句话的意思是什么?

复习题答案

1.在两人博弈的纳什均衡中，每个选手在制定最优选择时，考虑的最主要因素是什么？在占优策略均衡中，每个选手在制定最优选择时，考虑的最主要因素是什么？

【复习内容】纳什均衡；占优策略均衡

可以参考上一章（第 28 章）复习题第 2 题。

【参考答案】

在纳什均衡中，每个选手都对其他选手的最优反应作出了自己的最优反应；而在一个占优策略均衡中，每个选手的选择都是对其他选手所有选择的最优反应。

2.复习教材中 C 和 R 在混合策略中的最优反应问题。他们的最优反应是否构成了最优反应函数？

【复习内容】最优反应函数（曲线）

以两人参与的博弈进行分析，假设你是其中一个选手。对于对方的任何选择，你的最优反应（best response）就是使的你的收益最大化。如果有若干个选择都能使你的收益最大，那么你的最优反应是这些选择的集合。

考虑一个更一般的两人（C 与 R）的博弈，其中 R 的选择为 $r_1,\ldots r_R$，C 的选择为 $c_1,\ldots c_C$。对于 R 的每个选择 r，令 $b_c(r)$ 表示 C 的一个最优反应；对于 C 的每个选择 c，令 $b_r(c)$ 表示 R 的一个最优反应。则一个纳什均衡为满足下列条件的一对策略 (r^*,c^*)：

$$c^*=b_c(r^*)$$

$$r^*=b_r(c^*).$$

注意，有些情形下，其中一个选手可能对他的几个最优反应是无差异的。这也就是为什么我们只要求 c^* 是 C 的一个最优反应，r^* 是 R 的一个最优反应即可。如果每个选择的最优反应是唯一的，那么最优反应曲线可用最优反应函数表示。

【参考答案】

为了更清楚地理解这个问题，我们采用更一般也是最常用的纳什**混合策略**均衡解的求解方法。首先我们需要把两个选手的收益信息拷贝如下：

		Ms. Column	
		Left	Right
Mr. Row	Top	2, 1	0, 0
	Bottom	0, 0	1, 2

令 r 表示 R 选择上的概率，$1-r$ 表示他选择下的概率；类似地，令 c 表示 C 选择左的概率，$1-c$ 表示他选择右的概率。

R 选择的概率要能恰好使得 C 在选择左和选择右之间无差异，否则 C 会作出更有利自己的选择。这就是说对于 C 来说，有

$$1\times r+0\times(1-r)=0\times r+2\times(1-r)$$

由此可得$r=2/3$。

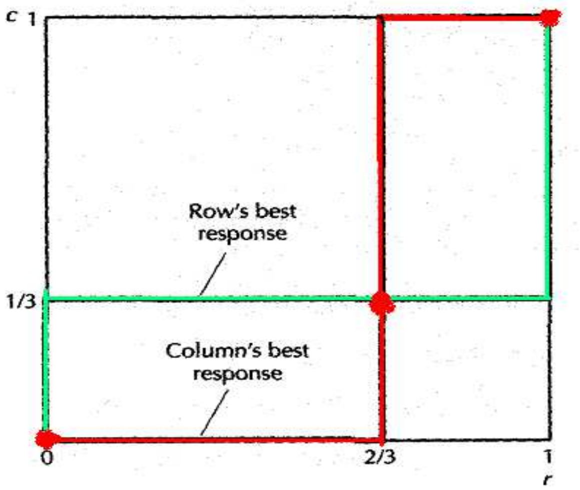
类似地，C 选择的概率要能使得 R 在选择上和选择下之间无差异，也就是说，对于 R 来说有

$$2\times c+0\times(1-c)=0\times c+1\times(1-c)$$

由此可得$c=1/3$。

也就是说 C 若以$1/3$的概率选择左，那么就可以使 R 可以**任意概率**选择上；换一句话说，当$c=1/3$时，R 的最优策略是无穷多的，从这个角度来说，这不是函数的概念，因为若是函数则要求 R 的最优策略只有一个。类似地，若 R 以$2/3$的概率选择上，那么就可以使 C 可以任意概率选择左，这同样也不是函数的概念。所以他们的最优选择不能构成最优反应函数。

下面我们用图形的方法说明个结论。在下图中，绿色曲线表示选手 R 的最优反应曲线，红色曲线表示选手 C 的最优反应曲线。这两条曲线的交点均为纳什均衡解。在这个例子中，有三个均衡解，其中两个是纯策略均衡解，一个是混合策略均衡解。



由图明显可以看出，C 若以$1/3$的概率选择左，那么 R 的最优反应曲线为一条水平线，也就是说 R 选择上的概率可以为 0 到 1 之间的任何数值，这显然不是函数的概念，因为函数要求函数值唯一。

3.（判断）在合作博弈中，如果两个选手作出了同样选择，那么他们对此结果都是满意的。

【复习内容】合作博弈

如果你是某个合作博弈的选手，你可能希望对方与你合作：在合作博弈中，你希望对方

选择你们都偏爱的均衡；在性别大战博弈中，你希望对方选择你们其中一方偏爱的均衡；在囚徒的两难博弈中，你希望对方选择的不是实现均衡的策略（即希望对方不认罪）；在懦夫博弈中，你希望对方做出选择从而达到你喜欢的结果。

在保证博弈、性别大战博弈和懦夫博弈中，实现合作的方法是其中一个选手率先行动，并且他向对方承诺坚持某个既定的选择。对方于是可以观察他的行为，从而相应做出反应。在囚犯的两难博弈中，这种方法不可行：如果一个选手选择不认罪，则对方会选择认罪。“解决”囚犯的两难问题的主要方法是重复博弈和签订合同。

【参考答案】错误，反例有懦夫博弈等。

由于合作博弈的类型很多，比如保证博弈、性别大战博弈、懦夫博弈和囚徒的困境都是合作博弈，因此笼统地说，如果两个选手作出了同样选择，那么他们对此结果都是满意的，那么这种说法是错误的。

例如，在懦夫博弈中，如果选手都将汽车径直开向对方，这是同样的选择，但是这种结果是他们都不想要的。所以题目中的结论是错误的。

4.教材中断言，在足球博弈中，均衡时 R 得分的概率为 62%。这个结果是怎么计算出来的？

【复习内容】竞争博弈（零和博弈）

【参考答案】

为了更清楚地说明问题，我们将选手的收益信息拷贝如下：

		Column	
		Defend left	Defend right
Row	Kick left	50, –50	80, –80
	Kick right	90, –90	20, –20

在教材中，我们已经知道这个博弈的纳什混合策略均衡解为（R 以概率 0.7 踢向左，C 以概率 0.6 扑向左）。

在均衡时，R 的收益是一种期望收益的形式，即他的收益=R 踢向左的概率×C 扑向左的概率×R 此时的收益+…+ R 踢向右的概率×C 扑向右的概率×R 此时的收益。用数值表示即为

$$(0.7\times0.6)\times50+(0.7\times0.4)\times80+(0.3\times0.6)\times90+(0.3\times0.4)\times20=62.$$

曹乾（东南大学 caoqianseu@163.com）

5.一个建筑商说，他打算“降低报价并在对方要求修改时将这个损失补偿回来”，他这句话的意思是什么?

【复习内容】

敲诈（hold up）问题

【参考答案】

这个建筑商的意思是说：他将出低价以确保赢得合同，但是如果客户在施工过程中需要修改，那么他就会对客户索要较高的价格。客户只能忍气吞声，因为在施工过程中的转换成本非常高。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

30.行为经济学（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

30 行为经济学

我们前面分析过的消费者选择模型简单且优美，它也是很多分析的合理起点。然而，它肯定不是故事的全部，在很多情形下，为了更准确地分析消费者行为，必须建立更深入的消费者行为模型。

行为经济学（behavioral economics）致力于研究消费者实际如何作出选择的。行为经济学使用了部分心理学的思想，借以预测消费者的选择行为。行为经济学中的很多预测结果和“理性”消费者的传统经济模型的结果不一致。

在本章，我们将介绍行为经济学家识别出的若干重要现象，并将行为理论的预测结果和本书前面章节的结果进行比较[一]。

30.1 消费者选择行为中的框架效应

在消费者行为的基本模型中，我们抽象地描述了消费者的选择：红铅笔和蓝铅笔，汉堡包和炸薯条等等。然而，在现实生活中，消费者的选择，受选择对象的具体提供方式的影响较大。

一条褪色的牛仔裤，你把它放在旧货店还是专卖店里，给消费者的感觉是不同的。购买股票组合的决策和出售股票组合的决策在感觉上也不一样，即使这两种组合包含的股票种类和数量是一样的。一家书店将某书标价 29.95 元能卖出几十本，但如果标价 29 元却只能卖出寥寥几本。

这些都是**框架效应**（framing effects）的例子，它们显然对消费者行为有强大的影响力。事实上，很多市场营销策略，恰恰是利用了消费者选择中存在的这些认知偏差。

疾病治疗困境

框架效应在涉及不确定性的选择中非常常见。例如，考虑下面的决策问题[二]：

某种严重的疾病威胁 600 个人的生命。你要在下列两种治疗方案 A 和 B 中作出选择，这两种方案的后果为：

方案 A：肯定可以挽救 200 人的生命。

方案 B：在 1/3 的概率下能挽救 600 人的生命，2/3 的概率下一个人也救不活。

[一] 在写本章时，我借鉴了由 Colin F. Camerer, George Loewenstein, 和 Matthew Rabin 撰写的《行为经济学进展》一书，普林斯顿大学出版社出版，2003.尤其借鉴了该书的导论部分。

[二] A. Tversky and D. Kahneman, 1981, "The framing of decisions and the psychology of choice," Science, 211, 453-458.

曹乾（东南大学 caoqianseu@163.com）

你会选择哪个方案？现在考虑下面两种方案的选择问题。

方案 C：有 400 人必定死亡。

方案 D：在 2/3 的概率下，600 人会死亡；在 1/3 的概率下无人死亡。

现在你会选择那一种治疗方案？**肯定性框架**（positive framing）给出的信息是能救活多少人，大多数人因此选择 A 而不是 B；但**否定性框架**（negative framing）给出的信息是不能救活多少人，大多人因此选择 D 而不是 C。但是要注意，A 和 C 的结果是完全一样的，B 和 D 的结果也是完全一样的。显然，将同一方案置于肯定性框架（能救活多少人），比将其置于否定性框架内（不能救活多少人）更受人们欢迎。

即使内行专家在面对这样的问题时也会落入陷阱。心理学家曾经让医生对上述问题做出选择，当问题是以肯定性框架给出时，72%的人选择了比较安全的治疗方案 A，而不是具有风险的方案 B；但是当同样的问题以否定性框架给出时，只有 22%的医生选择了具有风险的方案 C，而 72%的一声选择了安全的治疗方案 D。

尽管我们面对的问题通常不是生或死的选择问题，但是很多平凡的选择——例如股票的买或卖和上述问题具有相似性。对投资组合的理性选择，应该取决于该投资可能结果的评估，而不是你如何得到这些投资组合。

例如，假设有人给你 100 股 Concreteblocks.com 公司的股票（它的广告是“我们无偿赠送建筑材料，你只要支付包装费和运输费”）。如果有人以礼物形式送给你该公司的股票，你可能不愿意卖掉他们，尽管你从未想过购买该公司的股票。

人们通常不愿意卖掉亏本的股票，因为他们认为这些股票的价格还会“回升”。也许价格会上升，也许不会。但是最终你不应该让历史成本决定你的投资组合——你应该问你自己这样正确的问题：今天你是否拥有你想要的投资组合。如果不是，就卖掉它们。

锚定效应

上面虚构的 Concreteblocks.com 例子和所谓的**锚定效应**（anchoring effect）有关。锚定效应是说人们的选择会受到完全伪造的信息的影响。有一个经典的研究是这样的：比如你是被调查人员，实验者当着你的面转动标有不同数字的转盘，转盘停下来时指针会指向某个数字，然后实验者就该数字向你提出一些问题。比如，实验者让你回答，联合国中非洲国家的数量是大于还是小于指针指向的数字？

在你回答完以后，实验者还让你估计联合国中的非洲国家实际有多少个。尽管转盘上的数字是随机的，但研究表明，它对你的这一估计数字有显著的影响。

在类似的另外一项实验中，实验者给 MBA 学生一瓶昂贵的酒，然后问他们是否愿意出一定的资金购买这瓶酒，所出资金等于学生社会保障号（类似你的身份证号）最后两位数字。

例如如果你的社会保障号最后两位数字是 29，问题就是“你愿意出 29 元钱买这瓶酒吗？”

在学生回答完上述问题后，实验者又让他们回答，他们最多愿意出多少钱购买这瓶酒。研究表明，后面这个问题的答案受前一个问题的答案影响较大，也就是说每个人所愿意出的最高价明显受到他的社会保障号最后两位数字的影响。例如，社会保障号最后两位数字小于或等于 50 的学生报出的最高价的平均值为 11.62 元，但是社会保障号最后两位数字大于 50 的学生报出的最高价的平均值是 19.95 元。

也许你会说上面这几个例子更像实验室中的研究，但是有很多重大的经济决策也会受到框架效应的影响，也就是说，稍微改变一下问题的提问方式就会影响到人们的选择决策。

例如，人们对养老金方案的选择就是这样的[一]。

有些经济学家分析了三个企业（雇主）的数据，这三个雇主的雇员可以自动加入 401（k）养老金保险[二]。雇员也可以不参加，但是如果不参加他们必须在问卷上勾选“不参加”的选项。经济学家发现员工的自动参与率非常高，85%以上的员工自动选择了 401（K）方案中的默认参与选项。

这当然是个好消息，因为高参与率有助于养老基金的积累。但坏消息是，几乎所有自动参与的员工都选择了默认的投资选项，这些默认的投资项目通常是货币市场基金，它的回报率很低，员工每月储蓄率也很低。雇主制定的默认投资选项通常非常保守，因为这样可以减少投资收益下跌的风险，以及减少员工可能提出的诉讼。

随后，这些经济学家又分析了一个企业，这个企业员工不是自动加入 401（k）计划。它是这样规定的：新员工在进入公司的一个月之内，必须做出下列选择——要么加入 401（K）计划要么推迟加入。

显然这个企业没有提供“不加入”的默认选项，而且没提供加入回报率较低基金的默认选项，这种“积极的决策”方法将新员工加入 401（k）的参与率从 35%提高到了 70%。而且，加入 401(k)计划的员工大都选择了较高的储蓄率。

这个例子表明，对福利方案的精心设计可以显著改变员工对福利方案的选择结果，而且对他们的储蓄行为也有显著的影响。

[一] James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick, "For Better or for Worse: Default Effects and 401(k) Savings Behavior," NBER working paper, W8651, 2001.

[二] 401K 计划也称为 401K 条款，是指美国 1978 年《国内税收法》第 401 条 K 项的规定。该条款适用于私人公司，为雇主和雇员的养老金存款提供税收方面的优惠。按该计划，企业为员工设立专门的 401K 账户，员工每月从其工资中拿出一定比例的资金存入养老金账户，而企业一般也为员工缴纳一定比例的费用。员工自主选择证券组合进行投资，收益计入个人账户。员工退休时，可以选择一次性领取、分期领取和转为存款等方式使用。译者注。

托架效应（Bracketing）[一]

人们通常难以理解自己的行为，他们发现在不同的环境下要想预测自己的实际选择非常困难。

例如，一位营销学教授所做的实验是这样的[二]：在每周一次的营销学课堂中，他让学生从六种快餐中作出选择，学生选择好后，就可以在课下吃他选择好的快餐，快餐是免费的，实验连续进行三周（即三次）。 在一种方案中，学生必须提前选好三次快餐的种类；在另一种方案中，他们每次现场进行选择然后立即吃掉。

当学生必须提前作出选择时，他们通常选择多样化的快餐。事实上，在这种方案中，有64%的学生每周选择不同的快餐，而在其他组中只有 9%的人选择了不同的快餐。当要求人们立即作出选择时，他们通常喜欢多样化而不是单一性。但在实际选择时，他们选择的却是自己最喜欢的选项。我们都是有惯性的动物，即使在选择快餐时也是一样。

选择过多

常规理论认为选择多多益善。然而，这种断言忽略了选择的成本。在富裕的国家，消费者很容易迷失于过多的选择中，从而难以进行决策。

在一项实验中，两位市场营销研究人员在某超市中建立了两个果酱展示摊位[三]。一个摊位展示了 24 种风味的果酱，另一个摊位只展示了 6 种。前面这个摊位聚集了更多的消费者，但是后面这个摊位的实际销量却更多。选择更多似乎对消费者更有吸引力，但大摊位提供的过多选择似乎让消费者更难下决心购买。

两位行为金融方面的专家，怀疑人们的投资决策中也可能也存在着“选择过多”的问题。他们研究了加州大学洛杉矶校区教职工的退休金投资方案，具体地说，他们计算出了几种投资组合，其中一种是教职工自己设计的投资组合，这代表选择更多；另外一种是计算出了这些教职工投资组合的平均数。研究发现，教职工对于这两种投资组合的满意程度基本是一样的。人们在自己设计退休金投资组合时有很多的选择，但这样做似乎并没让他们感到更满意[四]。

[一] 在摄影中，Bracketing 通常是指对某同一对象连续拍摄多张，以反应对象的不同面貌。此处的意思应该借用了摄影中的这个术语，是指人们在面对一次性选择时，倾向于选择多样化。正如你旅游时，对某一山头连续拍摄多张照片的意思一样。但是要想概括翻译这个词比较困难，由于 bracket 的原意有“托架、支架”的意思，因此译者将此翻译为“托架效应”，类似于前面的“框架效应（framing effect）。

[二] I. Simonson, 1990, "The effect of purchase quantity and timing on variety-seeking behavior," Journal of Marketing Research, 17: 150-164.

[三] Sheena S. Iyengar and Mark R. Lepper, “When choice is demotivating: can one desire too much of a good thing?” Journal of Personality and Social Psychology, 2000.

[四] Shlomo Benartzi and Richard Thaler, “How Much Is Investor Autonomy Worth?”UCLA working paper, 2001.

构建偏好

我们如何解释这些案例?心理学家和行为经济学家认为偏好不是选择的向导;与此相反，他们认为，通过观察消费者的选择，可以“发现”偏好的一部分。

例如，你观察到某人在超市里拿起一个西红柿，放下，然后又将它拿起来。他想买还是不想买？这个价格-数量组合是可以接受的吗？当你观察到这样的行为，你看到的这个人正处于制定决策的“边际上”。或者，按照心理学家的解释，他在**发现**他的偏好。

常规理论认为消费者的偏好是已经存在的。因此，偏好**解释**了行为。心理学家则认为偏好是构建出来的——人们通过选择和消费的行为创造了偏好。

似乎心理学的模型对现实的解释力更好一些。然而，这两种观点并不是完全不相容的。我们已经知道，一旦人们发现了偏好，尽管这个过程有些神秘，偏好就成为了选择行为的组成部分。选择一旦做出，选择就锁定了决策。如果前文的那个消费者一旦买下了西红柿，而你又想从他手中购买的话，那么你所出的价格一般要高于他购买的价格。

30.2 不确定性

普通的选择已经非常复杂，但是不确定性环境下的选择更为复杂。我们已经知道，人们的决策可能取决于这些选项是如何阐述的（即框架效应等）。但是在该领域，行为还存在着很多其他的偏离现象。

小数法则

如果你学过统计学，你可能知道大数法则。这个数学原理大致是说，总体的一个大样本的平均数，趋近于该总体的平均值。

小数法则（Law of Small Numbers）是一个心理学命题，它的意思是说，小样本、尤其是人们亲身经历的小样本事件，对他们的行为影响更大[一]。

考虑下面的问题[二]：

“某个小镇只有两家医院。大医院每天有 45 个婴儿出生，小医院为 15 个。你已经知道，大约 50%的婴儿为男孩。然而，准确的百分比每天都变化。有时可能高于 50%，有时会低于 50%。在某一年间，每家医院都记录了男孩比例大于 60%的天数。你认为那家医院记录的天数更多？”

[一] The term originated with A.Tversky and D. Kahneman, 1971, "Belief in the law of small numbers," Psychological Bulletin,76, 2: 105-110. Much of the following discussion is based on a working paper by Matthew Rabin of the University of California at Berkeley entitled "Inference by Believers in the Law of Small Numbers."

[二] A. Tversky and D. Kahneman, 1982, "Judgments of and by Representativeness," in Judgment under Uncertainty: Heuristics and Biases, D. Kahneman, P. Slovic, and A. Tversky, Cambridge University Press, 84-98.

在对大学生所做的调查研究中，22%的受访者认为大医院记录的天数可能更多些，56%的受访者认为两个医院差不多一样多。只有 22%的受访者正确回答出了是小医院报告的天数多。

如果你对这个正确答案感到奇怪，不妨假设小医院每天出生 2 个婴儿，而大医院每天出生 100 个。在 25%（$=(1/2)\times(1/2)$）的时间里小医院出生的婴儿全部为男孩，然而大医院出生的婴儿全为男孩的概率极其小（$=1/2^{100}$）。

似乎人们预期这些样本和他们自己抽样时的分布是一样的。或者，换一句话说，人们低估了一个样本中个体实际大小的波动性。

一个相关的问题是，人们可能不知道随机性的具体含义。在一项实验中，实验人员要求受访者写出 150 次“随机”掷硬币结果的序列。受访者们写出的序列中大约有 15%在一行中含有三次硬币正面或反面，但这种模式按真正的随机计算其概率大约为 25%。受访者写出的序列中有 3%在一行含有 4 次正面或 4 次反面，但依据概率论计算这种模式的发生概率为 12%。

这个例子对博弈论应用有重要的启示意义。我们已经知道，在很多情形下，人们应该尽量使自己的策略选择随机化，以便不让竞争对手猜中。但是，心理学文献却显示，人们并不擅长随机化。另一方面，人们也不擅长于识别非随机的行为，至少对于没学过统计学的人们是这样的。因此，尽管混合策略均衡结果在数学上可以预测，但是对于参与博弈的人们来说，他们没有能力预测到这样的均衡结果。

有些经济学家研究了温布尔登网球比赛的决赛和半决赛结果。理想的情形是，选手应该随机变换发球区域以便让对手不能猜测到网球从哪个方向发过来。然而，即使是非常熟练的选手也不能很好地做到随机发球。根据这几个经济学家的研究[一]：

“我们的研究表明网球选手并不特别擅长随机打球：为了满足随机性，他们似乎有些过于频繁地变换发球区域。这和大量的心理学和经济学的实验结果是一致的，这些实验表明，当人们试图真正表现随机行为时，他们倾向于“矫枉过正”。”

资产整合和规避损失

在我们研究期望效用时，我们隐含地作出了下列假设：个人真正关心的是他最终拥有在各种可能结果下的财富总量，也就是财富的加权平均值，其中权重为相应结果的发生概率。这个假设称为**资产整合假设**（asset integration hypothesis）。

尽管大多数人认为这个假设比较合理，但想要将其付诸实践却很难（即使对于经济学家来说也是如此）。一般来说，人们倾向于规避过多的小风险但却接受过多的大风险。

假设你每年的收入为 10 万元。现在让你参加一个掷硬币的赌博，如果正面向上你得到

[一] M. Walker arid J. Wooders, 1999, "Minimax Play at Wimbledon," University of Arizona working paper.

14 元，如果反面向上你损失 10 元。这个赌博的期望值为 12 元，它占你年收入的比重是微不足道的。除非你对赌博行为会感到道德上的不安，你会认为这个赌博具有新引力，你很可能参加。然而，令人惊讶的是，现实调查研究表明相当多的人不会参加这个赌博。

这种**风险过度厌恶**（excess risk aversion）现象也存在于保险市场中，人们倾向于对各种小事件购买过度保险。例如，人们购买手机丢失保险，尽管他更换一个新手机的成本更低。人们还购买免赔额（deductibles）过低的汽车保险，但是他们不知道如果免赔额非常小，就几乎没有任何经济意义，因为只要有损失一般就会超过免赔额，这也说明人们是过度厌恶风险的。

一般来说，在做是否购买保险的决策时，你应该看看 “获胜机率”（house odds）或保险费率[一]。如果手机保险的保险费为 3 元/月或 36 元/年，而且新手机的价格为 180 元，那么保险的年费率为 36/180 或者 20%。如果你的手机丢失率大于 20%或者你没钱更换一个新手机时，购买保险才合算。

人们**厌恶风险**（risk averse），但是这种厌恶程度似乎小于他们**对损失的厌恶**(loss averse)程度。也就是说，人们似乎过度关注现状，即与将来相比，他们过分强调起点如何。

在一个重复多次的实验中，两个研究人员向一组实验对象随机发放咖啡杯，咖啡杯的发放数量等于实验对象的数量的一半[二]。

实验者要求得到咖啡杯的人报出他们愿意卖掉咖啡杯的价格。接着，实验者要求没得到咖啡杯的人报出咖啡杯的最高买价。由于咖啡杯是随机发放的，买价和卖价应该大致相等。然而，在该实验中，卖价的中位数为 5.79 元而买价的中位数为 2.25 元，存在显著差异。显然，与没得到咖啡杯的人相比，拥有咖啡杯的人不愿意放弃咖啡杯。人们的偏好似乎受到他们禀赋的影响，这一结论和标准的消费者理论结论不同。

类似的效应出现在所谓的**沉没成本错觉**（sunk cost fallacy）中。一旦你购买了某个东西，你支付的钱就“沉没”了，或者不可再挽回。因此，沉没成本不应该影响你的将来行为。

但是现实中，人们倾向于过度关注他们为某东西花了多少钱。研究人员发现，波士顿房主报出的卖房价和当初他们的购买价高度相关[三]。我们在前面曾经支出过，股票所有者非常不愿意卖掉亏本的股票，即使他们知道这些损失可以冲减他们的纳税额。

普通百姓容易受沉没成本错觉影响这一事实，比较有趣。然而更有趣的是专业人士一般不怎么受沉没成本错觉的影响。例如，前文指出的波士顿房价的例子中，研究人员发现，与购房自住的人相比，出于投资目的的购房者受沉没成本的影响较小。

[一] 此处的“house”意思同“Bingo”，原意是赌博胜利时发出的声音，引申为赌博获胜。House odds 原意为获胜机率，此处可翻译为保险费率。译者注。

[二] D. Kahneman, J. L. Kitsch, and R. Thaler, 1990, "Experimental tests of the endowment effect and the Coase theorem," Journal of Political Economy, 98, 1325-1348.

[三] David Genesove and Christopher Mayer, 2001, "Loss aversion and seller behavior: Evidence from the housing market," Quarterly Journal of Economics, 116, 4, 1233-1260.

类似地，金融中介咨询机构一般愿意出售已遭受损失的资产比如股票，尤其是这样做可以冲减它们的纳税额时。因此，人们咨询金融中介机构的原因之一，可能就是希望听取它的冷静的分析结论。

30.3 时间

正如涉及不确定性的行为容易受到各种反常行为的影响一样，涉及时间的行为也有很多反常性。

贴现

我们以时间贴现为例说明。经济学的标准模型——**指数化贴现**（exponential discounting）——假设人们对未来的贴现率是固定不变的。若$u(c)$代表今天消费的效用，则t年后的消费效用为$\delta^t u(c)$，其中δ为小于 1 的常数。

这样假设的目的是数学处理上比较简单，这也意味着存在着更好的贴现方法，它们能够好地拟合现实中的数据。

经济学家曾做过这样的实验：实验者拍卖债券，这些债券的兑现时期不同。研究发现，人们对未来收益的评价小于按指数贴现模型计算出来的受益值。另外一种贴现理论，即双曲线贴现（hyperbolic discounting），指出贴现因子（即贴现率）的形式不是δ^t而是$1/(1+kt)$.

指数化贴现最吸引人的特征是它表现出的消费者行为具有“**时间一致性**”（time consistent）。例如某个具有三期消费方案的消费者的效用函数为

$$u(c_1)+\delta u(c_2)+\delta^2 u(c_3).$$

时期 1 和 2 之间的边际替代率为

$$MRS_{12}=\frac{\delta MU(c_2)}{MU(c_1)},$$

时期 2 和 3 之间的边际替代率为

$$MRS_{23}=\frac{\delta^2 MU(c_3)}{\delta MU(c_2)}=\frac{\delta MU(c_3)}{MU(c_2)}.$$

最后一个式子表明，消费者愿意用时期 2 消费替换成时期 3 消费的比率，**等于**用时期 1 消费替换成时期 2 消费的比率。但是在双曲线贴现的情形下，这个结论不再成立。双曲线贴现的的贴现因子为$1/(1+kt)$，由此可见，t越大即时期距离现在越远，人们对效用的评价越低。

双曲线贴现情形下的消费者展现的行为具有**时间不一致性**：他可以在今天做出将来行

为的计划，但在将来的那个时间到来时，他又不想按原先的计划实施。例如，一对夫妻打算花费 5000 美元去欧洲旅游，而先不储蓄，他们打算在下一年夏天开始储蓄。但是到了下一年夏天，他们又决定去乘船旅游。

自我控制

和时间不一致性密切相关的一个问题，是**自我控制**（self-control）问题。每个人在不同程度上都面对着这样的问题。当我们站在浴室的磅秤上时，我们可能会宣称：需要控制食物的热量而且吃得更少。但是当我们面对佳肴时，我们的决心就烟消云散了。能自我控制的人通常苗条而且健康，显然我们大部分人都不是这样的人。

一个重要的问题是人们是否意识到很难进行自我控制。如果我知道我有做事拖沓的倾向，那么当重要的工作摆在面前时，我应该意识到需要立即动手去做。或者如果我喜欢过多承诺，那么也许我应该学会更多地说不。

但是，还有另外的可能性。如果我抵制不了明天多吃一块甜食的诱惑，我可能干脆今天就多吃一块。我们的肉体是脆弱的，但精神同样也是脆弱的。

自我控制难题的解决之道，是找到方法保证自己履行对未来行动的承诺。也就是说，你找到的这种方法要能使得你如果不履约，你就会付出更大的代价。例如，人们若对公众宣称他们未来行为的佳话，那么他们很少偏离目标行为。再比如，对于酒鬼子来说，如果他们吃了某种药片以后再喝酒的话，将病的非常厉害，他们还会喝酒吗？还有为减肥者设计的**承诺约束设备**（commitment devices）：将这种设备绑在胃部，就能减少暴食暴饮的可能性。

签订合同的目的在于约束合同当事人履行约定的将来义务，即使由于条件变换，这些义务不再吸引人也需要履约。类似地，人们可以雇佣别人监视自己的行为，如果他们的行为偏离了承诺，就需要付出代价，这其实是自己和自己签订合同。水疗减肥（dieting spas）、聘请体育锻炼教师或者教练都表明了人们“已购买了自我控制”。

例子：过分自信

自我控制的另外一种有趣的形式是过分自信或**自负**（overconfidence）。两位金融经济学家，Brad Barber 和 Terrance Odean，研究了拥有折扣经纪人账户的 66,456 个家庭的投资行为[一]。研究发现，频繁交易的家庭的报酬率为 11.3%，而不怎么频繁交易的家庭的报酬率为 18%。

研究者发现，性别是影响频繁交易与否的一个重要因素：男人比女人交易更频繁。心理

[一] 通常，人们在投资时需要首先在金融经纪公司（比如证券公司）开立账户，经济公司对你的投资要收取一定的中介费用。折扣经纪公司比一般经纪公司收取的中介费要低，折扣经纪公司账户的意思就不难理解了。译者注。

学家通常认为男人对自己的能力过度自信，而大部分女人则更现实一些。心理学家将男人的这种行为称为自私自利的归因倾向（self-serving attribution bias）。一般来说，男人（至少部分男人）认为他们的成功是自身能力的结果，而不是靠运气，因此变得过分自信。

男人的过分自信可从他们的交易行为上反应出来。在上例中，研究人员发现男人的交易次数比女人高出了 45%。男人的频繁交易导致了他们的收益率比女人低了一个百分点。正如这两位研究者指出的，“（频繁）交易会危及你的财富。”

30.4 策略性互动和社会准则

策略性互动是特别有趣的心理学行为（也许是社会学行为）。我们已经分析过博弈理论，该理论试图预测每个理性的选手是如何互动的。但是还有一门叫做**行为博弈理论**（behavioral game theory）的学科，该学科研究人们如何实际互动的。的确，实际行动与纯理论的分析相比，存在着很大的偏差。

最后通牒博弈

以**最后通牒博弈**（ultimatum game）为例。我们在上一章已详细讨论过这种博弈。我们已经知道，这个博弈有两个参与人，一个是提议者，一个是响应者。实验人员给提议者 10 元钱，让他提出一种分配方案，在他和响应者之间分配这 10 元钱。然后，实验人员将分配方案告诉响应者，看他愿意不愿意接受这个分配方案。如果他接受，分配就会按照提议者提出的分配方案进行；如果他不接受，那么这两个人啥也得不到。

我们首先分析完全理性的参与人会怎么做。一旦响应者知道了分配方案，他有一个占优策略：只要能得到钱不管得到多少，都会接受。毕竟，假设我给你两个选择：一是 10 分钱，二是零分钱。难道你选择的不是 10 分钱吗？

由于完全理性的响应者会接受大于零的任何钱数，提议者会给响应者尽可能小的钱数，比如说一分钱。因此，按照博弈论预测，分配结果将非常极端：提议者几乎得到了 10 元钱的全部。

但当实验人员在现实中具体实施这个游戏时，得到的结果和上述结果存在着很大的区别。实际上，当响应者认为分配方案不公平时，他会拒绝。提议者给响应者分配的比例小于 30%的情形下，拒绝概率超过了 50%。

当然，如果提议者认识到响应者会拒绝“不公平”的分配，提议者会理性地选择近于平均分配。在接近平均分配情形下，即提议者得到财富的 55%、响应者得到 45%情形下，拒绝概率大约为 16%。

有很多研究文献分析了参与人的特征是如何影响博弈结果的。其中一个例子是性别差异：似乎男人更倾向于接受更有利于自己的分配方案，尤其是当提议者为女人时。

文化差异的影响也是重要的。似乎某些文化更看重公平性，因此，在这样的文化背景下，不公平的分配方案遭受拒绝的概率更高[一]。非常有意思的事情是，提议者的分配方案受地区和文化的影响不大，但是响应者的拒绝概率却显著受这些因素的影响。用于分配的财富大小也很总要。如果分配的财富总额为 10 元，你可能不愿意接受分配给你的 1 元钱。但是如果分配的财富总额为 1000 元，你是否还拒绝分配给你的 100 元呢?显然，响应者一般不会拒绝较大的钱数。

人们还可以改变博弈方案的设计。其中一种方案称为**策略行动**（strategy method），在这种方案中，响应者在看到分配方案之前要先写下他能接受的最小分配数额。提议者知道响应者已提前作出了决策，但是他不知道响应者能接受的最小钱数是多少。这种实验设计提高了分配给响应者的钱数；也就是说，使得分配方案更公平。

公平

在最后通牒博弈中，有一种效应在起作用，这就是人们希望公平。大多数人似乎自然偏爱平均分配（或至少接近平均分配）。这不是个别现象，而是社会现象。即使人们是旁观者，他们也希望游戏规则要符合**公平准则**（fainess norms）。

例如，**带有惩罚的博弈**（punishment games）是最后通牒博弈的一个变种，它有三个参与方，即与传统博弈相比，它增添了第三方。第三方的作用是监督提议者或分配者提出的分配方案。如果第三方觉得这个分配方案不公平，他有权处罚分配方案的提议者，即减少提议者得到的钱数，但是要注意，第三方在行使处罚权时他自己也要付出代价，也就是说他自己也要出部分资金，这样设计的目的是保证第三方不滥用处罚权[二]。

实验人员发现，大约有 60%的第三方会处罚不公平的分配者。由此可见，似乎人类的本性就是讨厌不公平的行为，不论这个本性是天生的还是后天养成的。

的确，公平的社会准则在不同文化之间存在着差异；有些文化对公平评价很高，但有些文化对公平的要求没那么高。然而，不管文化背景如何，人们都倾向于惩罚不公平的现象。有人认为人们天性中就有偏爱“公平”结果的成分，原因也许在于，个人间只有公平地互相对待，才能有更多的生存和繁衍机会。

30.5 对行为经济学的评价

心理学家、营销专家和行为经济学家已经搜集了各种行为经济的案例，这些例子表明传统的选择理论是错误的，或者至少是不完整的。

[一] See Swee-Hoon Chuah, Robert Hoffman, Martin Jones, and Geoffrey Williams, "Do Cultures Clash? Evidence from Cross-National Ultimatum Game Experiments," Nottingham University Business School working paper.
[二] See Ernst Fehr and Urs Fischbacher, 2004, "Third-party punishment and social norms," Evolution and Human Behavior, 25, 63-87.

这些例子中可能有一些只是消费者的“错觉”（optical illusions）。例如，将选项置于不同的框架（frame）中，即以不同的方式提出选择问题，可能会影响消费者的选择决策。例如你画出同样的物体但衬以不同的背景，可能会影响人们对该物体大小和距离的判断。如果人们花点时间仔细思考如何选择，比如拿一把尺子量一量物体的大小或者距离，就能得到正确的答案。

尽管人们的行为不可能 100%符合即使是最简单的经济行为理论，人们也会说没有哪一个理论是 100%正确的。心理学家也指出，人们甚至不能理解简单的物理学原理。例如：如果在绳子的末端系上一块石头，在你头部上方绕上一圈然后松手，石头将向哪个方向飞去？

很多人回答说石头是射线一样向外飞去，他们不知道正确答案为石头将沿着圆周的切线轨迹飞出[一]。当然，人们的一生都是在物理世界中度过的，即使是这样他们还会偶尔错误地理解了物理原理，那么我们也不应该对人们错误理解经济世界而感到惊讶。

显然，我们对物理学的直觉理解，已经足够我们日常生活所需了。这些知识也足够满足业余或专业运动的需要：一个棒球选手可能无法描述出棒球的运动轨迹，尽管他可以打出漂亮的弧线球。类似地，我们可以认为，人们擅长于处理各种日常必须面对的决策，即使他们无法对这些日常决策进行抽象推理那有何妨？

对反常行为的另外一种解释是，市场倾向于奖励理性行为、惩罚不理性的行为。即使很多参与人行为不合理，那些行为合理的人才会对价格和结果产生最大的影响。这种观点似乎有一定的合理性。例如，在前面我们已经知道，出于投资目的购买房产的人比购房自住的人更少地受到沉没成本的影响。

另外，你可以向专家求助，让他们帮助你做出更好的决策。减肥咨询师和财务咨询师可以客观地建议你怎么吃和怎么投资。如果你担心这些建议过于宽松，你总可以求助更严格的专家。

我们再回到视觉错觉的例子中，我们使用尺子的原因是我们已学会不相信我们的眼睛。类似地，在做出重大决策时，最好咨询一下专家的客观观点。

总结

1.行为经济学关注的是消费者在现实生活中如何作出选择。

2.在很多情形下，消费者的实际行为，不同于使用理性消费者简单模型预测出的行为。

3.消费者的选择会受到问题提问方式的影响，这就是框架效应。

[一] See M. McCloskey, 1983, "Intuitive Physics," Scientific American, April, 114-123

4.在涉及不确定性的选择决策中，选择行为可能非常复杂。

5.消费者似乎偏好“公平”的分配方案，而且会惩罚行为不公平的人。

复习题

1.实验涉及人们购买彩票。实验人员告诉一组人他们有 55%的概率能中奖，但对另一组人说，他们有 45%的概率不能中奖。哪一组更有可能购买彩票？这种效应的名称是什么？

2.玛丽为她的家庭制定了一个星期的饮食计划，而弗瑞德则为他的家庭每天购买饮食。哪一个人更容易造成饮食多样化？这种效应叫什么名字？

3.假设你是一家中等规模公司的人力资源主管，你正考虑为你公司雇员的养老金投资提供投资方案，假设有两种方案，一种投资方案包含 10 种共同基金可供选择，另外一种包含 50 种共同基金可供选择，你认为哪种投资方案更好？

4.投掷硬币三次，三次全为正面向上的概率为多大？

5.约翰决定本周储蓄 5 元、下周储蓄 10 元。但当下周来临时，他决定只储蓄 8 元。这种行为的不一致性可用什么术语描述？

复习题答案

1.实验涉及人们购买彩票。实验人员告诉一组人他们有 55%的概率能中奖，但对另一组人说，他们有 45%的概率不能中奖。哪一组更有可能购买彩票？这种效应的名称是什么？

【复习内容】框架效应

框架效应（framing effect）是指选择问题以不同的方式进行表达时，会影响消费者的选择决策。一般来说，肯定性框架（positive framing）给出的信息正面信息，而否定性框架（negative framing）给出的信息是负面信息。由于人们是厌恶风险的，因此，肯定性框架给出的选择可以鼓励消费者进行选择。

【参考答案】

第一组人更有可能购买彩票，原因见上。这种效应的名称为框架效应。

2.玛丽为她的家庭制定了一个星期的饮食计划，而弗瑞德则为他的家庭每天购买饮食。哪一个人更容易造成饮食多样化？这种效应叫什么名字？

【复习内容】托架效应(bracketing)

在摄影中，Bracketing 通常是指对某同一对象连续拍摄多张，以反应对象的不同面貌。此处的意思应该借用了摄影中的这个术语，是指人们在面对一次性选择时，倾向于选择多样化。正如你旅游时，对某一山头连续拍摄多张照片的意思一样。但是要想概括翻译这个词比较困难，由于 bracket 的原意有“托架、支架”的意思，因此译者将此翻译为“托架效应”，类似于前面的“框架效应（framing effect）。

【参考答案】

托架效应（bracketing）是指，当人们可以提前但必须一次性作出选择时，他们通常选择多样化。但在实际选择时，人们选择的却是自己最喜欢东西，这是惯性，也就是实际选择通常不是多样化而是相对单一的。

由此可见，玛丽更容易造成多样性的饮食。这种效应名称叫做托架效应。

3.假设你是一家中等规模公司的人力资源主管，你正考虑为你公司雇员的养老金投资提供投资方案，假设有两种方案，一种投资方案包含 10 种共同基金，另外一种包含 50 种共同基金，你认为哪种投资渠道更好？

【复习内容】选择过多

【参考答案】

行为经济学认为选择是有成本的。因此，可供选择的东西并不是多多益善。尽管选择更多似乎对消费者更有吸引力，但消费者很容易迷失于过多的选择中，从而难以进行决策。

所以只提供 10 种共同基金让雇员进行选择效果可能更佳。

4.投掷硬币三次，三次全为正面向上的概率为多大？

【复习内容】概率学基本知识

【参考答案】

$$(\frac{1}{2})^3=\frac{1}{8}$$

5.约翰决定本周储蓄 5 元、下周储蓄 10 元。但当下周来临时，他决定只储蓄 8 元。这种行为的不一致性可用什么术语描述？

【复习内容】双曲线贴现

【参考答案】

双曲线贴现情形下的消费者展现的行为具有时间不一致性：他可以在今天做出将来行为的计划，但在将来的那个时间到来时，他又不想按原先的计划实施。

由此可见，约翰的行为是典型的双曲线贴现的行为，可用时间“不一致性”进行描述。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

31.交换（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

31 交换

直到现在，我们分析的市场都是孤立的单个市场。我们曾将某商品的需求函数和供给函数视为一元函数，自变量只有该种商品的价格，而没有考虑其他商品的价格。但一般来说，其他商品的价格会影响人们对某既定产品的需求和供给。当然，某商品的替代品和互补品的价格会影响这种商品的需求，而且更微妙的是，商品的售价影响卖方的收入，从而影响他们能够购买其他商品的数量。

直到现在，我们一直没考虑其他商品的价格对某既定商品市场均衡的影响。当我们讨论某既定商品的市场时，我们只分析了问题的一部分：商品的需求和供给如何受这种商品价格的影响。这种分析称为**局部均衡**（partial equilibrium）分析。

在这一章，我们将开始展开对**一般均衡**（general equilibrium）问题的研究：多种商品的价格是如何同时被决定的，为了分析这个问题，我们必须分析这些商品市场的需求和供给条件是如何相互作用的。正如你猜想的一样，这个问题比较复杂，我们必须进行一些简化。

首先，我们仅限于分析竞争市场的行为，因此每个消费者和生产者都是价格接受者，相应做出最优化决策。不完全竞争的一般均衡非常有趣，但是也难于分析，因此我们不考虑这种情形。

其次，假定商品数量和消费者数量尽可能地少：两种商品和两个消费者。通过这一简单模型，我们就可以描述很多有趣的现象，而且我们可以将我们得出的一般均衡结论推广到任意数量的商品和消费者的情形。使用两种商品和两个消费者的好处显然是便于分析。

第三，我们将一般均衡问题的分析分为两个阶段。第一阶段，我们分析的经济体具有下列特征：人们拥有的商品禀赋数量是固定不变的。我们主要分析人们如何交易这些商品的，不考虑商品的生产。这种情形自然可以称为**纯交换**（pure exchange）。第二阶段，在对纯交换市场有了清晰的理解之后，我们将分析一般均衡模型中的生产行为。

31.1 艾奇沃斯盒

我们使用一种比较方便的图形工具，即**艾奇沃斯盒**（Edgeworth box），分析两种商品在两个消费者之间的交换问题[一]。这个盒子的好处是，我们可以在同一张图中画出两个消费者的禀赋和他们各自的偏好，这样就可以分析交易过程的各种结果。为了理解艾奇沃斯盒的构造，有必要分析两人的无差异曲线和各自的禀赋。

令两个消费者分别为 A 和 B，两种商品分别为 1 和 2。令 $X_A=(x_A^1,x_A^2)$ 表示 A 的消费束，其中 x_A^1,x_A^2 分别表示 A 消费的商品 1 和 2 的数量。令 $X_B=(x_B^1,x_B^2)$ 表示 B 的消费束。

[一] 艾奇沃斯盒的命名是为了纪念英国经济学家 Francis Ysidro Edgeworth(1845-1926)，他首先使用了这个分析工具。

我们把 X_A 和 X_B 的任一组合 (X_A, X_B) 称为一个**配置**（allocation）。一个配置为**可行配置**（feasible allocation），若所消费的每种商品的总量等于该种商品的禀赋总量：

$$x_A^1 + x_B^1 = \omega_A^1 + \omega_B^1$$
$$x_A^2 + x_B^2 = \omega_A^2 + \omega_B^2.$$

初始禀赋配置(initial endowment allocation)——(ω_A^1, ω_A^2) 和 (ω_B^1, ω_B^2)——是指消费者在交易前就拥有的配置，也就是说，它是消费者带到市场的两种商品数量的组合。初始禀赋配置是一种可行配置，因为你将它们代入上式，上式显然是成立的。消费者在进行交易之后最终带回家的两种商品数量组合，称为**最终配置**（final allocation）

我们用图 31.1 中的艾奇沃斯盒来说明这些概念。我们首先用标准的消费者理论框架阐述消费者 A 的禀赋和偏好。还可以在坐标轴上相应标出经济体中的每种商品的总量，这个总量等于 A、B 拥有的每种商品的加和。由于我们只对这两个消费者的可行配置感兴趣，因此利用前面两个式子，可以画出艾奇沃斯盒。在这个盒子中我们可以画出 A 可以拥有商品束。

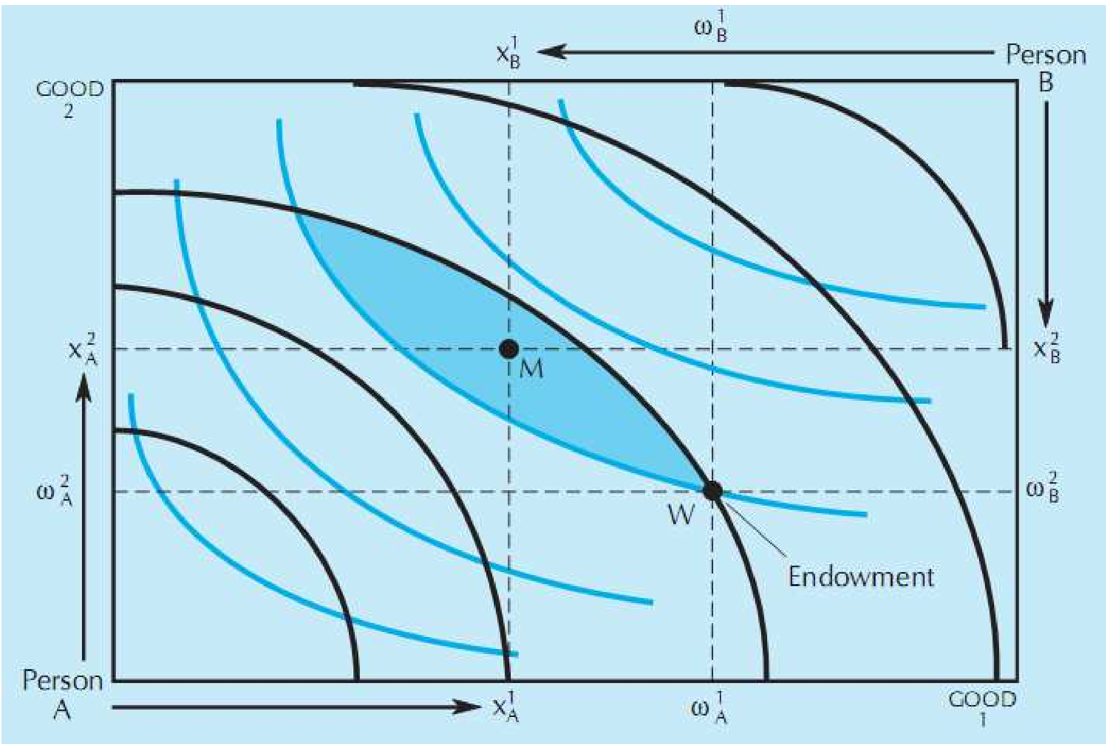


图 31.1：艾奇沃斯盒。这个盒子的长表示经济体内商品 1 的总量，宽表示商品 2 的总量。A 的消费选择从左下角开始衡量；B 的消费选择从右上角开始衡量。

注意，一旦我们画出 A 的一个商品束，那么这个商品束同时也代表 B 的一个商品束，但是此时要站在图形的右上角来看（即从 B 的原点看）。例如，如果商品 1 的总数为 10 单位，商品 2 的总数为 20 单位，而且假设 A 的消费束为（7,12），则 B 的消费束必定为（3,8）。它们在艾奇沃斯盒中为同一点。A 拥有的商品 1 的数量，可用从左下角 A 的原点开始沿着

横轴的线段的长度表示；B 拥有的商品 1 的数量则从右上角 B 的原点开始沿着横轴进行衡量。类似地，纵轴上线段的距离给出了 A 和 B 拥有的商品 2 的数量。因此，这个盒子内的每个点既表示 A 的消费束又表示 B 的消费束——只是从各自的原点开始衡量而已。这个经济体的所有可行配置都可用艾奇沃斯盒中的点表示。

我们可以象往常一样画出 A 的无差异曲线，但 B 的无差异曲线的形式稍有不同。你可以象往常一样画出 B 的无差异曲线，但把它旋转 90 度，就得到了图中所示的 B 的无差异曲线。如果我们从左下角 A 的原点开始，向右上方（含向上或向右）移动，我们就会到达另外一个配置点，这个配置点代表的效用更大。如果我们从右上角 B 的原点开始向左下方（含向下或向左）移动，我们也会到达效用更大的 B 的配置点。（将你的课本旋转 90 度后再看 B 的图形，你就会更清楚上述分析。）

利用艾奇沃斯盒可以画出两个消费者的各个可能的消费束——可行配置——以及两个消费者各自的偏好。由此可见，这个盒子对两个消费者的相关经济特征作出了比较全面的描述。

31.2 交易

既然我们已经画出了无差异曲线和禀赋，现在就可以分析交易如何进行了。我们从商品的初始禀赋W点开始分析，如图 31.1 所示。将你的注意力放在穿过点W的 A 和 B 的无差异曲线身上。位于 A 的这条无差异曲线**以上**区域中的任何一点，与 A 的禀赋点W相比，效用都更高，因此这个区域是 A 状况改善的区域。类似地，站在 **B 的原点**上看，与 B 的禀赋点相比，B 的状况改善的区域也位于 B 的那条穿过W点以上的区域。（如果你还不清楚，再次将你的课本旋转 90 度来看 B 的图，你就明白了。）

A 和 B 状况**都**改善的区域在哪里？答案显然是 A 的状况改善的区域与 B 的状况改善的区域的交集。这个交集就是图 31.1 中具有凸透镜形状的阴影区域。在他们讨价还价的过程中，这两个消费者很可能找到某种互利的交易——这个交易能使得他们移动至该凸透镜区域内部的某一点，比如点M。

点W移动到点M，意味着 A 放弃了$\left|x_A^1-\omega_A^1\right|$单位商品 1，以此交换到了$\left|x_A^2-\omega_A^2\right|$商品 2。所以 B 得到了$\left|x_B^1-\omega_B^1\right|$单位商品 1，放弃了$\left|x_B^2-\omega_B^2\right|$单位商品 2。

需要指出，点M这个配置并没什么特别之处。事实上，该凸透镜区域内部的任何一点都可以——因为该区域内的任何一点代表的效用都比原禀赋点高。所以我们只需要架设消费者在交易后达到了该区域内的**某一点**即可。

现在我们从点M开始重复与上述相同的分析。我们可以画出穿过点M的 A 和 B 的无差异曲线，构造出一个新的凸透镜形状的“互利区域”，想象进行交易后从点M移动到该区域中的另外一点N。以此类推…交易将继续进行下去，直至不存在对两方都有利的交易。这样的位置是什么样的？

31.3 帕累托有效率的配置

图 31.2 给出了答案。注意图中的点 M 以及穿过 M 点的 A 和 B 各自的无差异曲线。A 的这条无差异曲线以上的点组成的点集，与 B 的这条无差异曲线以上的点组成的点集（需要站在 B 的原点看）没有交集。也就是说，A 状况更好的区域和 B 状况更好的区域不相交。这意味着，从点 M 开始的任何移动，在使一方状况变好的同时，必然使另一方的状况变差。因此，此时已不存在互利的交易。在点 M 这样的配置点上，已没有彼此改善的交易。

象点 M 这样的配置称为一个帕累托有效率（Pareto efficient）)的配置。帕累托效率这个非常重要的概念有若干不同的表达方式:

1.已不存在让所有涉及的人状况更好的方法；

2.已不存在让某些人状况变好但又不使其他人状况变差的方法；

3.交易的所有收益（即好处）都已被取尽；

4.已不存在互利的交易，等等。

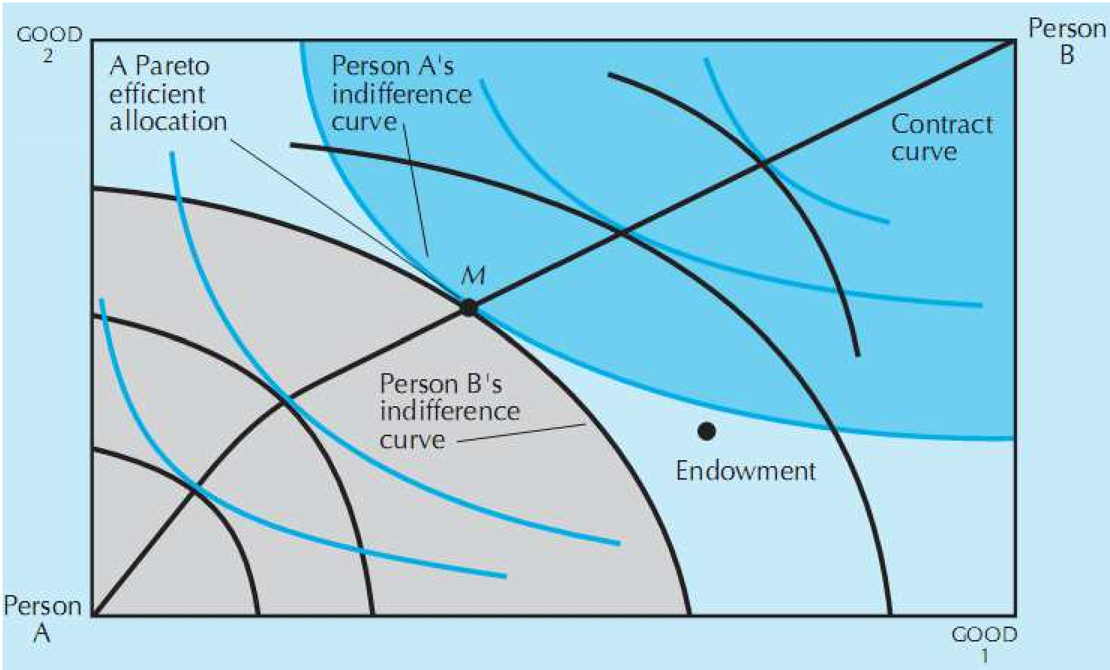


图 31.2：一个帕累托有效率的配置。在象点 M 这样的帕累托有效率的配置点上，给定另外一个人的无差异曲线，每个人都位于他自身的尽可能高的无差异曲线上。将这些有效率的配置点连接起来就得到了合同曲线。

事实上，我们在前面单个市场的情形下，已提到过几次帕累托有效率的概念：那时，我们说，单个市场的产量是帕累托有效率的，如果在该产量上，消费者的边际支付意愿等于供给者的边际销售意愿。在任何其他的产量上，如果上述两种意愿数量不相等，那么交易就能使得双方的状况变得更好。在本章，我们将更深入地分析帕累托有效率的概念，这涉及很多商品和很多交易者。

注意帕累托有效率配置点的几何特征：在艾奇沃斯盒内的任何一个帕累托有效率的配置点上，A 和 B 的无差异曲线必然相切。原因很简单。如果在盒内的某一点上，A 和 B 的无差异曲线不相切，那么它们必然相交，但是如果相交，则必然存在着互利的交易——因此，这样的点不可能是帕累托有效率的。（在盒子的边上也可能存在帕累托有效率的配置点——在这样的配置点上其中一个消费者的某种商品的消费量为零——这种情形下，他们的无差异曲线在该配置点上不相切。这些边界情形对当前的讨论并不重要。）

从相切条件可知，在这个艾奇沃斯盒中存在着很多帕累托有效率的配置点。事实上，给定 A 的任何一条无差异曲线，我们有可以按照下面方法很容易找到一个有效率的配置点：沿着 A 的这条无差异曲线移动，直到在这条无差异曲线上找到一点，使得这一点对 B 来说效用最大。这一点就是一个帕累托有效率的配置点，因此 A 和 B 的无差异曲线必然在这点相切。

艾奇沃斯盒中的所有帕累托有效率的配置点组成的点集，称为**帕累托集**（Pareto set），或者称为**合同曲线**（contract curve）[一]。后面这个名字来自于下列思想：交易的所有“最终合同”必然位于帕累托集上——否则这些交易不可能是最终的，因为如果不位于帕累托集上，说明仍存在着效率改进的余地，即仍存在着互利的交易。

通常，合同曲线将从 A 的原点开始、穿过艾奇沃斯盒的内部、一直延伸到 B 的原点，如图 31.2 所示（也就是说 A 和 B 的原点也是帕累托有效率的配置点，下面将说明原因）。如果我们从 A 的原点开始，此时 A 拥有的两种商品数量都为零，而 B 却拥有两种商品的全部数量。A 的原点是帕累托有效率的，因为 A 的状况改善的唯一方法是从 B 手中“夺走”某些商品[二]。当我们沿着合同曲线从 A 的原点开始向上移动，A 的状况越来越好，直到我们最终到达 B 的原点。注意 B 的原点也是帕累托有效率的，原因和 A 的原点有效率类似。

帕累托集给出了从盒内任何一点开始互利交易最终所能达到的所有可能结果。如果给定盒内任何一点（比如图 31.1 的 W 点）——每个消费者的初始禀赋——我们就可以找到每个消费者更偏好的帕累托子集（与他的初始禀赋相比）。这个子集位于图 31.1 所示的凸透镜形状的阴影区域内。该区域内的配置点，是从某个既定的初始禀赋点（比如 W 点）开始的互利交易所达到的可能结果。但是，帕累托集本身并不取决于初始禀赋。初始禀赋的作用是决定每种商品的总量，从而决定艾奇沃斯盒的大小。

31.4 市场交易

上面描述的交易过程的均衡——帕累托有效率的配置——非常重要，但是我们并没说清楚两个消费者最终会在**哪一个**有效率的配置点上。原因在于我们对交易过程描述得非常笼统。本质上，前文只是说，如果两个消费者所在的禀赋点不是有效率的，那么他们必然会向能使他们状况更好的**某个**配置点移动。

[一] 老一代经济学家经常翻译为“契约曲线”。也许他们更喜欢用“契约”这个词。但是翻译成契约曲线没有任何好处。因为现代法律将契约一律称为“合同”。译者注。

[二] 这就是说，此时已不存在让 A 状况变好但又不使 B 状况变差的方法，按照前文的帕累托有效率的定义，A 的原点显然是有效率的。译者注。

如果能确定某个**特定**的交易过程，我们就能更详细地描述均衡。下面我们来看一个类似竞争市场的交易过程。

假设我们有一个第三方，他自愿充当 A 和 B 的商品的“拍卖者”。拍卖者为商品 1 和 2 分别选取一个价格，然后告诉 A 和 B。他们就能按照该价格(p_1, p_2)计算出自己禀赋的价值，并决定按该价格购买每种商品的数量。

有必要指出，如果只有两个人进行交易，那么要求他们的行为是竞争的就不大合理。这种情形下，他们会对交易进行讨价还价。有一种方法能避开这个难题，这个方法就是把艾奇沃斯盒看成**两类**消费者在交易（平均交易量），但每类消费者的数量都很多。另一种方法是，我们直接指出虽然在两个人的情形下，要求他们竞争不太可能，但在很多人的情形下，就合理了，这也是我们真正关注的情形。

不论使用哪种方法，我们都知道如何分析这样的消费者选择问题，我们在第 5 章已介绍过分析消费者选择问题的分析方法。在图 31.3 中，我们画出了 A 和 B 的两个需求束。（注意，图 31.3 中的情形不是一个均衡状态，因为一个人的需求不等于另一个人的供给。）

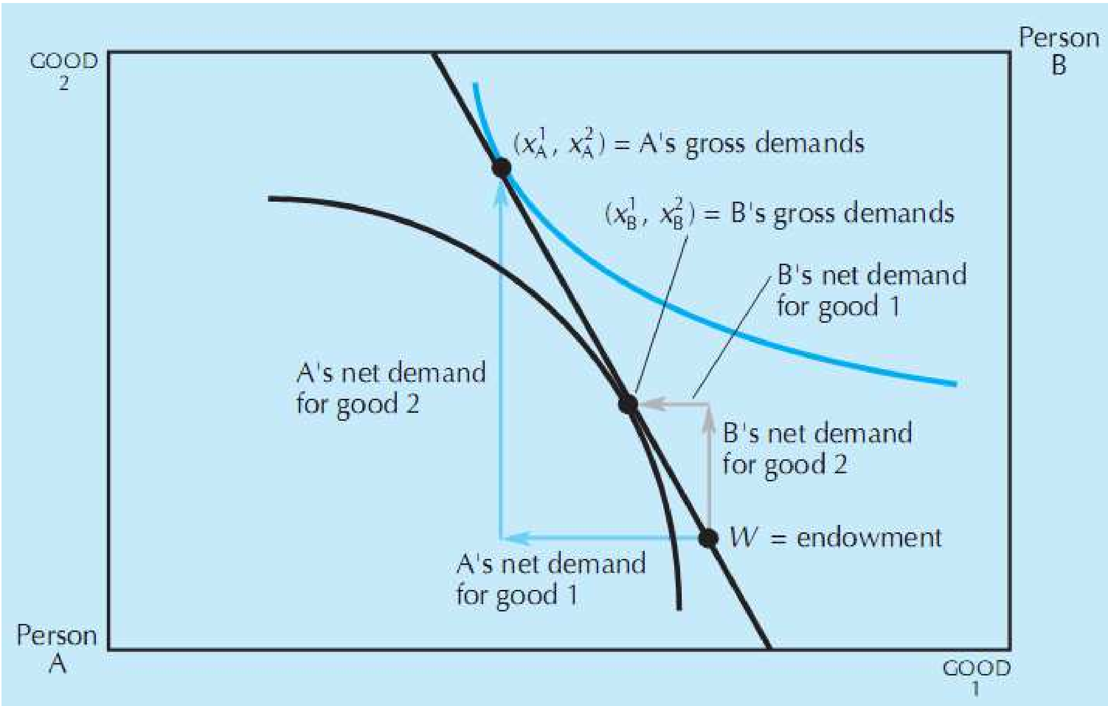


图 31.3：总需求和净需求。总需求是每个消费者想要消费的数量；净需求是消费者想要购买的数量。

此处的分析也要用到第 9 章介绍的两个“需求”概念。比如，A 对商品 1 的**总需求**（gross demand）是在当前价格下他想拥有的商品 1 的总数量。A 对商品 1 的**净需求**（net demand）是指 A 对商品 1 的总需求与他的商品 1 的初始禀赋之差。在一般均衡分析中，净需求有时也称为**额外需求**或**超额需求**（excess demand）。我们将 A 对商品 1 的额外需求记为e_A^1。根据定义，若 A 对商品 1 的总需求为x_A^1而且他的禀赋为ω_A^1时，我们有

$$e_A^1 = x_A^1 - \omega_A^1.$$

额外需求的概念可能比较自然，但是总需求的概念通常更有用。我们一般使用“需求”表示总需求，在可能出现混淆的情形下，我们才明确指出“净需求”或“额外需求”。

对于任意给定的价格(p_1, p_2)，不能保证供给一定等于需求——无论是从总需求还是净需求角度来说。如果以净需求衡量，这表示 A 想购买（或销售）的数量，不一定等于 B 想销售（或购买）的数量。如果以总需求衡量，这表示 A 和 B 想持有的商品总量不一定等于市场上的商品总量。事实上，图 31.3 所示的例子就是这样的。在这个例子中，A 和 B 不能完成他们想要的交易：市场不能出清，即需求不等于供给。

我们说这种情形下的市场处于**非均衡**（disequilibrium）状态。在这样的情形下，自然可以假设拍卖人将改变商品的价格。如果一种商品是超额需求的，拍卖人会提高这种商品的价格；如果另外一种商品是超额供给的，那么拍卖人会降低该商品的价格。

假设这种调整过程一直持续下去，直至每种商品的需求等于供给，此时的图形结构是什么样子的？

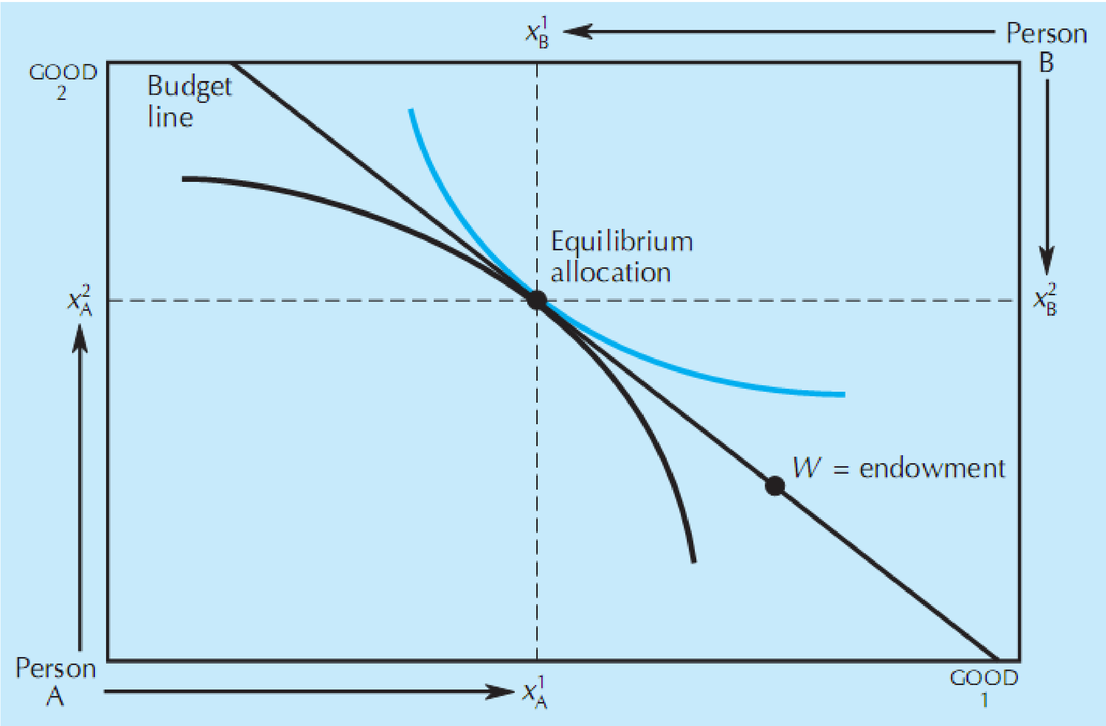


图 31.4: 艾奇沃斯盒中的均衡。在均衡时，每个消费者在他的预算集中选择最偏爱的消费束，他们的需求总和正好等于供给。

图 31.4 给出了答案。此处，A 希望购买的商品 1 的数量，正好等于 B 想出售的商品 1 的数量。商品 2 的情形也是类似的。换句话说，在当前价格水平下，每个人想要购买的每种商品数量之和，等于市场上可得到的每种商品的总量。我们说此时的市场处于一个**均衡**（equilibrium）状态。更准确地说，这种均衡叫作一个**市场均衡**（market equilibrium），或

竞争均衡（competitive equilibrium），或**瓦尔拉斯均衡**（Walrasian equilibrium）[一]。当然这些名字指的是同一件事：一组价格使得每个消费者选择他最偏爱的能买得起的消费束，并且所有消费者的选择是相容的，即每个市场的需求都等于供给。

我们知道，如果每个消费者选择他能买得起的最优消费束，那么他的边际替代率必定等于这两种商品的价格之比，于是，所有消费者的这两种商品的边际替代率是**相等**的。由图 31.4 可知，均衡的特点是每个消费者的无差异曲线与他的预算线相切。但是，既然每个消费者的预算线的斜率为 $-p_1/p_2$，这意味着两个消费者的无差异曲线必然是相切的。

31.5 均衡的代数表达方法

令 $x_A^1(p_1,p_2)$ 和 $x_B^1(p_1,p_2)$ 分别表示消费者 A 和 B 对商品 1 的需求函数，类似地可以定义商品 2 的需求函数，我们可以将均衡价格 (p_1^*,p_2^*) 定义为同时满足下列两个式子的价格：

$$x_A^1(p_1^*,p_2^*)+x_B^1(p_1^*,p_2^*)=\omega_A^1+\omega_B^1$$
$$x_A^2(p_1^*,p_2^*)+x_B^2(p_1^*,p_2^*)=\omega_A^2+\omega_B^2.$$

这两个式子是说，均衡时每种商品的总需求应该等于这种商品的总供给。这是均衡的第一种表示方法。

描述均衡的另外一种方法是，将上面两个式子变形整理，可得

$$[x_A^1(p_1^*,p_2^*)-\omega_A^1]+[x_B^1(p_1^*,p_2^*)-\omega_B^1]=0$$
$$[x_A^2(p_1^*,p_2^*)-\omega_A^2]+[x_B^2(p_1^*,p_2^*)-\omega_B^2]=0.$$

这两个式子是说，每个消费者对每种商品的**净需求之和**应该等于零。或者，换句话说，A 的净需求应该等于 B 的净供给，或 A 的净供给应该等于 B 的净需求。

描述均衡的第三种方法，需要使用**总超额需求函数**（aggregate excess demand function）的概念。下面我们推导出这种表示方法。首先令消费者 A 对商品 1 的净需求函数为

$$e_A^1(p_1,p_2)=x_A^1(p_1,p_2)-\omega_A^1$$

$e_B^1(p_1,p_2)$ 的定义类似。

函数 $e_A^1(p_1,p_2)$ 衡量消费者 A 的**净需求**或者说**超额需求**——他想消费商品 1 的数量与他的商品 1 的禀赋之差。下面我们将 A 和 B 对商品 1 的净需求加在一起，可得

$$\begin{aligned}z_1(p_1,p_2)&=e_A^1(p_1,p_2)+e_B^1(p_1,p_2)\\&=x_A^1(p_1,p_2)+x_B^1(p_1,p_2)-\omega_A^1-\omega_B^1\end{aligned}$$

我们将这个式子称为商品 1 的**总超额需求**（aggregate excess demand）。类似地可以得到商品 2 的总超额需求，我们用 $z_2(p_1,p_2)$ 表示。

[一] 里昂.瓦尔拉斯（Leon Walras, 1834-1910），是法国经济学家，任教于（瑞士）洛桑大学。他是一般均衡理论的开拓者之一。

于是，我们可用每种商品总需求等于零的说法来描述一组均衡(p_1^*, p_2^*)

$$z_1(p_1^*, p_2^*) = 0$$
$$z_2(p_1^*, p_2^*) = 0.$$

实际上，这个定义的要求过于严格了（stronger than necessary）[一]。可以证明，只要商品 1 的总超额需求为零，则商品 2 的总超额需求必然也为零。为了证明这个结论，我们有必要首先说明总超额需求函数的一个性质，这就是**瓦尔拉斯法则**（Walras's law）。

31.6 瓦尔拉斯法则

使用上一节引入的符号，可将瓦尔拉斯法则表达为

$$p_1 z_1(p_1, p_2) + p_2 z_2(p_1, p_2) \equiv 0.$$

这个式子是说，**总超额需求的价值之和恒等于零**。注意这个恒等式对于所有可能的价格均成立，并**不局限于**均衡价格。

为了看清这个结论，只要将两个消费者的预算线相加。首先考虑消费者 A。由于 A 对每种商品的需求满足他的预算线，我们有

$$p_1 x_A^1(p_1, p_2) + p_2 x_A^2(p_1, p_2) \equiv p_1 \omega_A^1 + p_2 \omega_A^2$$

或者

$$p_1[x_A^1(p_1, p_2) - \omega_A^1] + p_2[x_A^2(p_1, p_2) - \omega_A^2] \equiv 0$$
$$p_1 e_A^1(p_1, p_2) + p_2 e_A^2(p_1, p_2) \equiv 0.$$

这个式子是说消费者 A 的净需求的价值之和等于零。也就是说，A 想购买商品 1 数量的价值，加上他想购买商品 2 数量的价值必定等于零。(当然，他购买的一种商品数量必定为负，也就是说，他必须卖出这种商品才能购买另外一种商品。)

类似地，对于 B 也有这样的式子：

$$p_1[x_B^1(p_1, p_2) - \omega_B^1] + p_2[x_B^2(p_1, p_2) - \omega_B^2] \equiv 0$$
$$p_1 e_B^1(p_1, p_2) + p_2 e_B^2(p_1, p_2) \equiv 0$$

将 A 的式子和 B 的式子加在一起，并使用总超额需求的定义$z_1(p_1, p_2)$和$z_2(p_1, p_2)$，可得

$$p_1[e_A^1(p_1, p_2) + e_B^1(p_1, p_2)] + p_2[e_A^2(p_1, p_2) + e_B^2(p_1, p_2)] \equiv 0$$
$$p_1 z_1(p_1, p_2) + p_2 z_2(p_1, p_2) \equiv 0$$

现在我们可以看清瓦尔拉斯法则是怎么得到的了：因为每个消费者的超额需求价值等于零，即 $p_1 z_1(p_1, p_2) \equiv 0$ 且 $p_2 z_2(p_1, p_2) \equiv 0$，所以他们的超额需求价值之和也必定等于零，即 $p_1 z_1(p_1, p_2) + p_2 z_2(p_1, p_2) \equiv 0$。

[一] 这句话的意思是说，只要求上面的一个式子成立即可，而不要求两个式子同时成立。但是，我们即将看到，如果这两个式子中的一个成立，自动意味着两个式子同时成立。在这个意义上来说，这个定义又不是“过于严格”的。译者注。

下面我们就可以证明，如果一个市场的需求等于供给，那么另外一个市场的需求必定也等于供给。注意，瓦尔拉斯法则必须对所有的价格都成立，因为对于所有价格来说，每个消费者都必须满足他的预算约束。既然瓦尔拉斯法则对于所有价格都成立，那么对于能使商品 1 的超额需求等于零（$z_1(p_1^*,p_2^*)=0$）的价格(p_1^*,p_2^*)，它自然也成立。

根据瓦尔拉斯法则可知，对于价格(p_1^*,p_2^*)必然有

$$p_1^* z_1(p_1^*,p_2^*)+p_2^* z_2(p_1^*,p_2^*)\equiv 0.$$

由于$z_1(p_1^*,p_2^*)=0$，因此再根据上式可知，若$p_2>0$，则我们必然有

$$z_2(p_1^*,p_2^*)=0.$$

因此，正如前文指出的，如果我们找到一组价格(p_1^*,p_2^*)，使得商品 1 的需求等于供给，那么必然有：商品 2 的需求等于供给。或者，如果我们能找到一组价格使得商品 2 的需求等于供给，那么商品 1 的市场必然也处于均衡状态。

一般来说，如果有k种商品的市场，即k个市场，那么我们只需要找到一组价格使得$k-1$个市场均衡，那么第k个市场自动也处于均衡状态，这个结论可由瓦尔拉斯法则直接推出。

31.7 相对价格

正如上面指出的，瓦尔拉斯法则意味着在k种商品的一般均衡模型中，只有$k-1$个**独立的**方程：如果$k-1$个市场的需求都相应等于该市场的供给，那么第k个市场的需求必定等于供给。但是，如果有k种商品，则需要确定k个价格。问题是怎么样从$k-1$个方程中解出k个变量（价格）？

答案在于这个模型中实际只有$k-1$个**独立的**价格。我们在第 2 章已经知道，如果我们对预算线中的收入和所有价格都乘以一个正数t，那么预算集不会发生变动，因此需求束也不会发生变动。在一般均衡模型中，每个消费者的收入就是他的禀赋按照市场价格计算的价值。如果我们将所有价格乘以一个正数t，那么每个消费者的收入也将自动地乘以t。因此，如果我们找到一组均衡价格(p_1^*,p_2^*)，那么(tp_1^*,tp_2^*)也是均衡价格。

这意味着我们可以任意选取其中一种商品的价格，将其设定为常数。为方便起见，通常将这种商品的价格设定为 1，因此所有其他价格可以视为相对于这种商品来说的相对价格。我们在第 2 章已经知道，这样的价格称为**计价物价格**。如果我们选择商品 1 的价格作为计价物价格，则其他商品的相对价格可由原价格乘以$t=1/p_1$得到。

每个市场的需求等于供给的这一要求，只能决定均衡相对价格，这是因为将所有价格乘以一个正数不会影响任何人的需求和供给行为。

例子：均衡的代数实例

假设 A 的效用函数为柯布-道格拉斯类型（详见第 6 章），即$u_A(x_A^1,x_A^2)=(x_A^1)^a(x_A^2)^{1-a}$。消费者 B 的效用函数为$u_B(x_B^1,x_B^2)=(x_B^1)^b(x_B^2)^{1-b}$。由这两个效用函数可求出相应的需求函数：

$$x_A^1(p_1,p_2,m_A)=a\frac{m_A}{p_1}$$

$$x_A^2(p_1,p_2,m_A)=(1-a)\frac{m_A}{p_2}$$

$$x_B^1(p_1,p_2,m_B)=b\frac{m_B}{p_1}$$

$$x_B^2(p_1,p_2,m_B)=(1-b)\frac{m_B}{p_2}$$

其中a和b分别为 A 和 B 效用函数中的参数，m_A和m_B分别表示 A 和 B 的货币收入。

我们知道在均衡时，每个消费者的货币收入等于他的禀赋的价值：

$$m_A=p_1\omega_A^1+p_2\omega_A^2$$
$$m_B=p_1\omega_B^1+p_2\omega_B^2.$$

因此，这两种商品的总超额需求为

$$\begin{aligned}z_1(p_1,p_2)&=a\frac{m_A}{p_1}+b\frac{m_B}{p_1}-\omega_A^1-\omega_B^1\\&=a\frac{p_1\omega_A^1+p_2\omega_A^2}{p_1}+b\frac{p_1\omega_B^1+p_2\omega_B^2}{p_1}-\omega_A^1-\omega_B^1\end{aligned}$$

与

$$\begin{aligned}z_2(p_1,p_2)&=(1-a)\frac{m_A}{p_2}+(1-b)\frac{m_B}{p_2}-\omega_A^2-\omega_B^2\\&=(1-a)\frac{p_1\omega_A^1+p_2\omega_A^2}{p_2}+(1-b)\frac{p_1\omega_B^1+p_2\omega_B^2}{p_2}-\omega_A^2-\omega_B^2\end{aligned}$$

你应该验证一下，这些总超额需求函数满足瓦尔拉斯法则。

令p_2表示计价物价格，即令$p_2=1$，则上面的这两个式子变为

$$z_1(p_1,1)=a\frac{p_1\omega_A^1+\omega_A^2}{p_1}+b\frac{p_1\omega_B^1+\omega_B^2}{p_1}-\omega_A^1-\omega_B^1$$
$$z_2(p_1,1)=(1-a)(p_1\omega_A^1+\omega_A^2)+(1-b)(p_1\omega_B^1+\omega_B^2)-\omega_A^2-\omega_B^2.$$

现在我们有了商品 1 和 2 的各自超额需求函数的表达式$z_1(p_1,1)$和$z_2(p_1,1)$，每个式子是商品 1 相对价格p_1的函数。为了找到均衡价格，令每个式子等于零，分别解出p_1。根据瓦尔拉斯法则可知，无论从哪个式子求解p_1，得到的结果都是相同的。

均衡价格为

$$p_1^*=\frac{a\omega_A^2+b\omega_B^2}{(1-a)\omega_A^1+(1-b)\omega_B^1}.$$

(如果你有怀疑，可以将p_1^*代入需求等于供给表达式验证一下，看看是否成立。)

31.8 均衡的存在性问题

在上面的例子中，由于每个消费者的需求函数有具体的表达式，我们可以解出均衡价格。但是，一般来说，我们不知道每个消费者需求函数的具体表达式。因此，你也许会问，怎么样才能知道至少存在一组价格能使得每个市场的需求等于供给？这个问题称为**竞争均衡的存在性问题**（existence of a competitive equilibrium）。

竞争均衡存在性是非常重要的，因为它的作用是对前面章节各种模型的“一致性检验”。也就是说，如果这样的均衡通常不存在，那么我们建立的竞争均衡理论将是一堆废物。

早期的经济学家注意到，在k种商品的市场上，只有$k-1$种相对价格，每个市场需求等于供给也只涉及$k-1$个独立的方程。由于方程的数量等于未知数的数量，因此这些经济学家断言这k个方程必然有解。

但是，经济学家很快发现这样的结论是错误的。仅仅数一数方程的数量和未知数的数量，不足以证明均衡解是存在的。然而，我们可以用数学工具证明竞争均衡的存在性。人们发现在证明过程中，最重要的假设是总超额需求函数是**连续函数**。这意味着，大致来说，价格微小变动引起的总需求变动也是微小的：价格微小变动不会导致需求量出现大的跳跃。

总需求函数在什么样的条件下才是连续的？本质上有两类条件可以保证连续性。一种是每个消费者的需求函数是连续的，也就是说价格的微小变动只会导致需求量的微小变动。可以证明，这相当于要求每个消费者的偏好为凸的，我们在第 3 章已介绍过凸偏好的含义。另外一个条件更一般。即使消费者本身的需求函数是不连续的，但只要他们的需求量相对于市场规模来说非常小，总需求函数仍然是连续的。

后面这个条件比较好。毕竟，当市场上消费者数量众多，且每个消费者相对于市场规模来说比较小时，竞争行为的假设才比较合理。我们正是根据这个条件得到连续的总需求函数的。而在证明竞争均衡存在时，连续性又是必需的。由此可见，我们后面这个条件（消费者数量众多且每个消费者相对于市场规模来说较小），本来是对竞争市场的一种合理假设，但它却保证了均衡理论是成立的。这是多么美妙的事。

31.9 均衡和效率

我们已经分析了纯交换模型的市场交易，由此我们得到了一个具体的交易模型，我们可以将该模型与本章一开始讨论的一般交易模型进行比较。由此产生的一个问题是，竞争市场机制能否取尽交易的所有好处。在竞争均衡时，每个市场的需求都等于它自身的供给，这时候人们还想继续进行交易吗？换一种问法，即竞争市场均衡时是否为帕累托有效率的：人们在均衡价格交易之后，还想进一步进行交易吗？

通过研究图 31.4 可得到答案：竞争市场均衡时的配置是帕累托有效率的。证明过程如下：艾奇沃斯盒中的一个配置是帕累托有效率的，若消费者 A 偏好的消费束集合与 B 偏好的消费束集合不相交。但是在市场均衡处，A 更偏好的消费束集合必然位于他的预算线上方，B 的情形也是如此（但要从 B 的原点看）。因此，这两个人更偏好的消费束集合不相交。这意味着，与均衡配置点相比，不存在这两个人都更偏好的配置，因此竞争均衡为帕累托有效

率的。

31.10 效率的代数表达

我们也可以使用代数方法证明上一节的结论。使用反证法。假设某市场均衡不是帕累托有效率的，我们将证明这将导致逻辑上的矛盾。

说市场均衡不是帕累托有效率的，是说存在着其他可行配置$(y_A^1, y_A^2, y_B^1, y_B^1)$使得

$$y_A^1 + y_B^1 = \omega_A^1 + \omega_B^1 \tag{31.1}$$

$$y_A^2 + y_B^2 = \omega_A^2 + \omega_B^2 \tag{31.2}$$

并且

$$(y_A^1, y_A^2) \succ_A (x_A^1, x_A^2) \tag{31.3}$$

$$(y_B^1, y_B^2) \succ_B (x_B^1, x_B^2) \tag{31.4}$$

前两个式子是说y这个配置是可行的，后面两个式子是说A和B都偏好y配置胜于x配置。（$\succ_A$和$\succ_B$表示A和B的偏好。）

但是，根据假设，我们处于市场均衡状态，因此每个人购买的是他能买得起的最优消费束。如果(y_A^1, y_A^2)比A选择的消费束好，这意味着A必然买不起(y_A^1, y_A^2)。类似的结论对B也成立：

$$p_1 y_A^1 + p_2 y_A^2 > p_1\omega_A^1 + p_2\omega_A^2$$

$$p_1 y_B^1 + p_2 y_B^2 > p_1\omega_B^1 + p_2\omega_B^2.$$

现在将上面两个式子相加，可得

$$p_1(y_A^1 + y_B^1) + p_2(y_A^2 + y_B^2) > p_1(\omega_A^1 + \omega_B^1) + p_2(\omega_A^2 + \omega_B^2).$$

将（31.1）和（31.2）式代入上式可得

$$p_1(\omega_A^1 + \omega_B^1) + p_2(\omega_A^2 + \omega_B^2) > p_1(\omega_A^1 + \omega_B^1) + p_2(\omega_A^2 + \omega_B^2),$$

这个矛盾是由于我们假设市场均衡不是帕累托有效率的。因此，这个假设必定是错误的。由此可见，任何市场均衡都是帕累托有效率的：这个结论称为**福利经济学第一定理**(First Theorem of Welfare Economics)。

福利经济学第一定理保证了竞争市场能够取尽交易的所有好处：由一组竞争市场实现的一个均衡配置必定是帕累托有效率的。这样的配置可能没有让人喜欢的其它性质，但它一定是有效率的。

特别地，福利经济学第一定理一点也没涉及到经济利益的分配公平性问题。市场均衡可能不是“公平的”配置：如果在交易钱A拥有一切，那么交易后他仍拥有一切[一]。这样的配置是有效率的，但很可能不是公平的。但是，毕竟，效率毕竟能说明一些问题，而且它能让我们相信简单的市场机制能实现帕累托有效率的配置。

[一] 如果价格均为正，A不会和任何其他人进行交易，因此这种情形下，可以认为A和他自己交易。译者注。

例子：用艾奇沃斯盒分析垄断

为了更好地理解福利经济学第一定理，有必要考虑另外一种资源配置机制——这样的机制无法实现帕累托有效率的结果。例如，一个消费者是垄断者时就是这样的。现在假设不存在第三方拍卖人，而是由 A 向 B 报价，B 根据 A 报出的价格决定他自身的需求量。再假设 A 知道 B 的“需求曲线”，因此，给定 B 的需求行为，A 会尽力选择让他自己状况最好的价格。

为了更好地分析这个过程中的均衡，最好先回忆一下消费者**价格提供曲线**（price offer curve）的概念（详见第 6 章）。价格提供曲线代表了消费者在不同价格水平下的所有最优选择。B 的价格提供曲线表示在不同价格水平下，他购买的消费束；也就是说它刻画了 B 的需求行为。如果我们画出 B 的预算线，那么预算线和价格提供曲线的交点表示 B 的最优消费束。

因此，如果 A 选择的报价使得他自己的状况最好，他应该在 B 的价格提供曲线上找到一点，使得 A 的效用最大。这样的选择请见图 31.5.

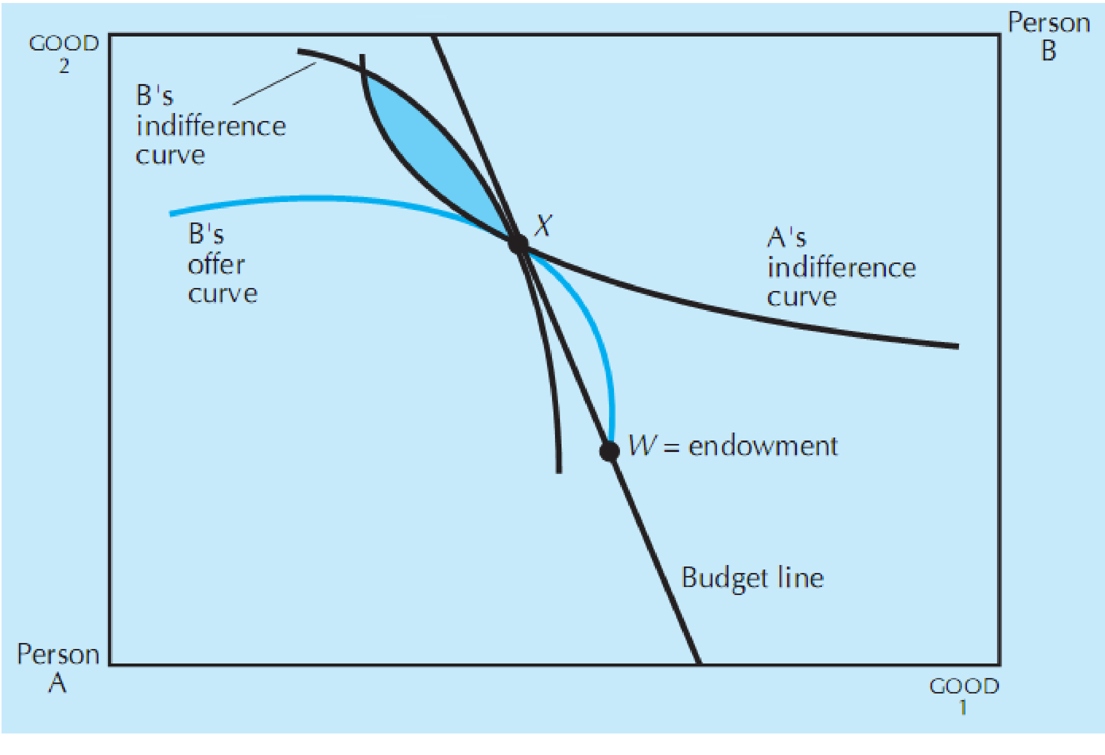


图 31.5：艾奇沃斯盒中的垄断。消费者 A 在 B 的价格提供曲线上选择一点，这一点要能使得他自己的效用最大。

这一最优选择可用通常的相切条件进行刻画：A 的无差异曲线和 B 的价格提供曲线相切。如果 B 的价格提供曲线与 A 的无差异曲线相交，则在 B 的价格提供曲线上存在着 A 更偏好的点，因此交点肯定不是 A 的最优点。

在找到了这样的点——图 31.5 中的 X 点——之后，我们可以画一条从禀赋点到 X 点的预算线。在与这条预算线相伴的价格水平上，B 会选择消费束 X，A 的状况达到了最好。

X 点是帕累托有效率的吗?一般来说答案是否定的。为了看清这一事实，只要注意 A 的无差异曲线不会和预算线相切于 X 点，因此 A 的无差异曲线不会和 B 的无差异曲线相切。A 的无差异曲线与 B 的**价格提供曲线**相切，这样它不可能再与 B 的无差异曲线相切。垄断配置是帕累托无效率的。

事实上，垄断配置无效率的原因，和第 24 章中的垄断的无效率的原因是一样的。在边际上，A 希望按照均衡价格多卖一些，但是如果他这么做，他必须对所有销量都降低价格，从而减少了他的全部边际内销售（inframarginal sales）[一] 得到的收入。

我们在第 25 章已经知道，一个实行完全价格歧视的垄断企业的产量是帕累托有效率的。这样的企业是指它能够将每单位商品都卖给出价最高的消费者。在艾奇沃斯盒中，完全价格歧视的垄断者的行为是什么样的?

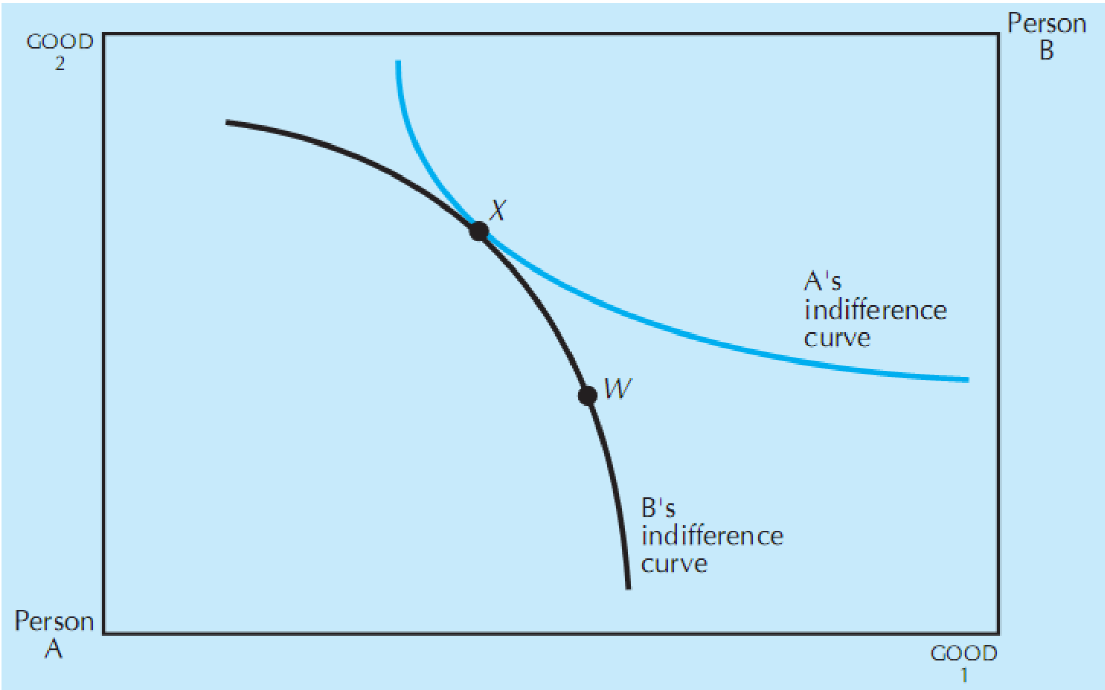


图 31.6：实施完全价格歧视的垄断者。A 选择点 X，这一点位于通过禀赋点的 B 的无差异曲线上，但也位于 A 的位置最高的无差异曲线上，也就是说 X 点位于 B 的无差异曲线与 A 的位置最高的无差异曲线的切点处，因此，这一点必然是帕累托有效率的。

图 31.6 给出了答案。我们从初始禀赋 W 开始分析，想象 A 卖给 B 的每单位商品 1 的价格都不同，也就是说 A 对某单位商品索要的价格恰好使得 B 在买和不买之间无差异，或者说 A 对每单位商品 1 索要的价格恰好等于 B 对这单位商品的支付意愿（保留价格）。因此 A 向 B 售出了第一单位商品之后，B 仍呆在通过 W 点的同一条无差异曲线上。于是，A 以

[一] 举个例子说明“边际内”的含义。例如若垄断厂商的利润最大化销售量为 10 单位，则“在边际上”意味着第 10 单位。如果他增加销量，比如决定销售 11 单位，则“边际内”是指 1~10 单位。译者注。

B 愿意支付的最高价格卖出第二单位商品 1。这意味着配置点会继续向左上方移动，但仍呆在 B 的通过 W 点的无差异曲线上。A 继续向 B 出售更多单位的商品，配置点会沿着 B 的这条无差异曲线一直向左上方移动，直到这个配置点也在 A 的位置最高的无差异曲线上，这就是图 31.6 中的 X 点。

容易看出点 X 这样的配置点必定是帕累托有效率的。给定 B 的无差异曲线，A 的状况（效用）实现了最大。在这样的配置点上，A 设法占有了 B 的全部消费者剩余：B 的状况和他的配置点的状况一样好。

垄断者的例子为福利经济学第一定理提供了两个典型案例。普通垄断者的例子，说明了垄断是一种导致无效率均衡的资源配置机制；而实行完全价格歧视垄断者的例子，则说明这样的资源配置机制是帕累托有效率的。

31.11 效率和均衡

福利经济学第一定理是说一组竞争市场的均衡是帕累托有效率的。反过来说对不对？给定一个帕累托有效率的配置，我们能否找到一组价格使得所有市场都实现均衡？可以证明，在附加某些条件后，答案是肯定的。论证过程如图 31.7 所示。

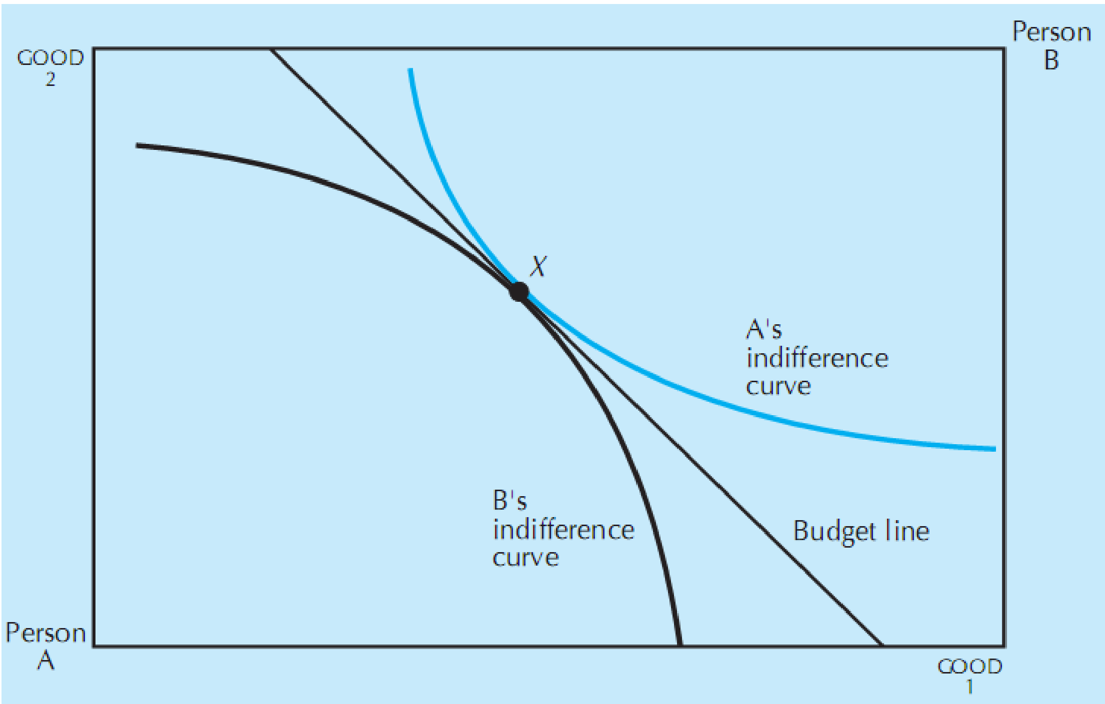


图 31.7：福利经济学第二定理。当偏好为凸时，帕累托有效率的配置，在某组价格水平上是一个均衡。

任选一个帕累托有效率的配置点，比如点 X。于是我们知道，与当前配置点相比，A 更偏好的配置点的集合，和 B 更偏好的配置点集合是不相交的。这当然意味着 A 和 B 的无差异曲线在点 X 相切。我们在图中还画出了这两条无差异曲线在 X 点的公切线。如图 31.7 所

示。

假设这条公切线代表这两个人的预算线。于是每个人在他的预算线上选择最优的消费束，由此产生的均衡将是原来的帕累托有效率配置点 X。

因为原来的配置 X 是有效率的，所以它自动决定了均衡价格。消费者从什么样的禀赋点开始交易能达到 X 这个配置点？能够产生图 31.7 中的那条预算线的任何禀赋——即位于这条预算线上的任何禀赋，在交易之后都能最终达到 X 点。

我们是否总能够构造出类似图 31.7 中的那条预算线？不幸地是，答案是否定的。图 31.8 给出了一个反例，也就是说在这个例子中我们构造不出这样的预算线。

在该图中，点 X 是帕累托有效率的，但不存在能使 A 和 B 都愿意在点 X 进行消费的价格。我们在图中也画出了一条最明显的候选预算线，但是在该预算线上，A 和 B 的最优需求不是同一个点。消费者 A 想要的消费束为 Y，但是消费者 B 想要的消费束为 X——在这样的价格下，需求不等于供给。

图 31.7 和图 31.8 的差别在于，图 31.7 中的偏好是凸的，而图 31.8 中的偏好是非凸的。如果 A 和 B 的偏好都是凸的，那么公切线只会与 A 和 B 的无差异曲线接触一次（即不可能有多个交点）。这样，我们就得到了福利经济学第二定理（the Second Theorem of Welfare Economics）：**如果所有消费者的偏好都是凸的，那么总存在着一组价格，使得每个帕累托有效率的配置是某个合理配置禀赋的市场均衡。**

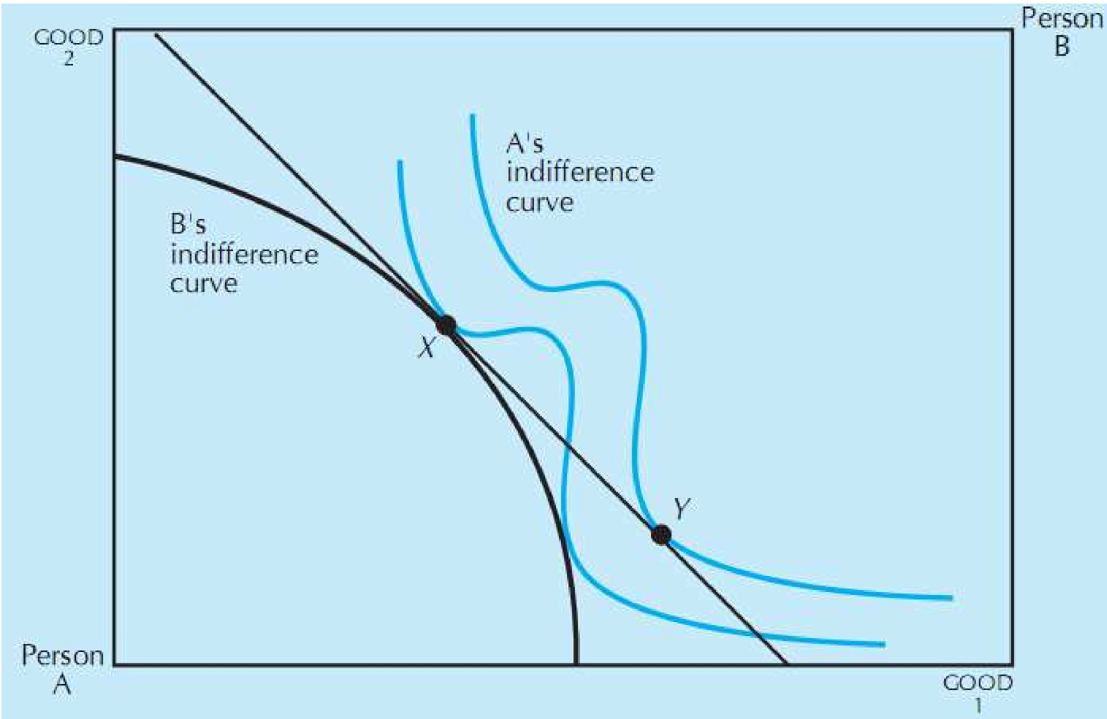


图 31.8：帕累托有效率配置不是均衡配置的情形。从这个图中，能够找到某个帕累托有效率的配置，比如点 X，但这样的配置无法通过竞争市场实现，因为消费者的偏好（在该图中为 A 的偏好）是非凸的。

证明过程和前面的几何图形推理是类似的。在某个帕累托有效率的配置点上，与该点相比，消费者 A 更偏好的消费束集合和 B 更偏好的消费束集合不相交。因此，如果这两个人的偏好都是凸的，那么我们就可以在这两个人更偏好的消费束集合之间画出一条直线，这条直线将 A 更偏好的消费束集和 B 更偏好的消费束集分开。这条线给出了两种商品的相对价格，任何能使 A 和 B 位于这条直线上的禀赋，都会导致最终的市场均衡，而且均衡位置就是原来的帕累托有效率配置。

31.12 第一福利定理的应用

福利经济学的两条定理，是经济学中最基本的结论之一。尽管我们是使用简单的艾奇沃斯盒说明这两个定理，但是对于含有任意多的消费者和商品的更复杂的模型来说，这两个定理也是成立的。福利定理在指导资源配置方法的设计方面具有深远的意义。

在本节我们只考虑第一定理。这条定理是说任何竞争均衡都是帕累托有效率的。这个定理的得出几乎不需要什么**外在**假设条件——它几乎完全是从定义推导出的。但是，它有几个隐含的假设。最主要的一个假设是消费者只关心他自己的商品消费，而不考虑其他人的消费情况。如果一个人的确关心另外一个人的消费，那么我们说这种情形存在着**消费外部性**（consumption externality）。我们将看到，如果存在消费外部性，那么竞争均衡未必是帕累托有效率的。

举个简单的例子说明，假设 A 关心 B 对雪茄的消费（抽烟）。那么这两个人在按照市场价格选择他自己的消费束后，我们没有理由认为这个结果是帕累托有效率的。因为，在这两个人购买能买得起的消费束之后，仍然存在让他们状况都变好的方法——例如，A 付钱给 B 让他少抽雪茄。我们将在第 33 章更详细地分析外部性。

第一福利定理中的另外一个重要假设是，消费者的行为真正为竞争性的。如果只有两个消费者，比如用艾奇沃斯盒分析的例子，那么他们不可能接受给定的价格。相反，这两个人可能会认识到他们的市场力量，并且试图运用他们的市场力量来改善自己的处境。竞争均衡的概念只有在下列情形下才是合理的：消费者的数量足够多，这样他们的行为才是竞争性的。

最后，第一福利定理只有竞争均衡确实存在时才有意义。正如我们在前文指出的，这种情形就是相对于市场规模来说，消费者是非常小的。

给定这些限制条件，第一福利定理是一个相当有力的结论：在自由市场上，每个人最求自身的效用最大化，最终实现了帕累托有效率的配置。

第一福利定理的重要性在于，它提供了一种实现帕累托有效率结果的一般机制：竞争市场。如果只有两个人进行交易，问题不大，因为这两个人完全可以聚在一起探讨可能的互利交易。但是如果交易者成千上万、甚至多达几百万，那么就必须对交易过程施加某种特别的结构。第一福利定理表明，竞争市场的特别结构能保证实现帕累托有效率的配置。

如果我们处理的问题是涉及很多人的资源分配问题。那么使用竞争市场的好处就非常明显，因为它能减少每个人需要掌握的信息，从而减少了每个人的交易成本。在竞争市场上，

一个消费者在做消费决策时，他仅仅需要知道他要消费的商品价格就可以了。消费者不需要知道这些商品是如何生产出来的，或者谁拥有这些商品的所有权，或市场中的商品是从哪里来的。如果每个消费者知道的仅是商品的价格，他能够确定他的需求，而且若运行良好的市场能确定竞争性的价格，我们必然能得到有效率的结果。因为竞争市场能够减少人们所必须掌握的信息，从而大幅减少了交易成本，所以它是深受人们欢迎的资源配置方法。

31.13 第二福利定理的应用

福利经济学第二定理断言，在某些条件下，每个帕累托有效率的配置都能通过竞争均衡实现。

这个结论的意义是什么？第二福利定理意味着分配问题和效率问题是可以分开的。你想要的任何帕累托有效率配置都可以通过市场机制实现。市场机制是在分配上是中性的（distributionally neutral）：无论你怎么界定什么样的福利分配方法才是好的或公平的，你都可以使用竞争市场实现你想要的结果。

价格在市场体系中发挥两种作用：一是**配置**作用（allocative role），二是**分配**作用（distributive role）。价格的配置作用表明了商品的相对稀缺性；价格的分配作用是指它决定了不同个人购买不同商品的数量。第二福利定理表明，价格的这两种作用是可以分开的：我们可以通过对商品再次分配的方法来决定每个人拥有的财富量（分配作用），然后使用价格表明商品的相对稀缺性（配置作用）。

人们在讨论政策时往往容易把价格的两种作用混为一谈。比如你可能经常听到：人们抱怨分配不公平，因此要对价格决策进行干预。然而这样的干预通常是不明智的。正如我们在前文指出的，实现资源有效率配置的简便方法，是让每个人面对他自身行为的真实社会成本，让他根据这些成本作出选择决策。因此，在完全竞争的市场中，消费者多消费还是少消费一些商品的边际决策，通常取决于商品的价格——它衡量所有其他人对该商品的边际评价。对效率的考虑本质上就是边际决策——每个人在做消费决策时，应该面对正确的边际权衡（marginal tradeoff）。

不同消费者应该消费多少的这类问题的决策，则是完全不同的问题。在竞争市场中，消费者消费商品的数量，取决于他的禀赋能卖多少钱。从纯理论的角度来看，国家完全可以用合理的方法转移购买力，即将禀赋在消费者中重新进行分配。

事实上，国家不需要转移实物禀赋，只需要转移禀赋的购买力即可。国家可以按照某消费者的禀赋价值进行征税，并将税款转移支付给另外一个消费者。如果税收是按照消费者的**禀赋**价值征收的，那么就不会存在效率损失。如果税收取决于消费者的**选择**(比如消费税)，那么税收会导致无效率，因为这种税收会影响人们的边际选择。

对禀赋征税的确会引起人们行为的改变。但是，根据第一福利定理，从任何初始禀赋开始的交易都能实现帕累托有效率的配置。因此，不管政府怎样对禀赋进行再次分配，由市场力量决定的均衡配置仍然是帕累托有效率的。

然而，在实施时还会存在一些问题。对消费者征收人头税（比如每人 5 元）是容易做到的。例如，我们可以对蓝眼睛的人征税，并将税收重新分配给棕色眼睛的人。只要眼睛的

颜色不能变动，就不会导致效率损失。或者我们可以对高智商的人征税，将税收重新分配给低智商的人。只要智商能够测定，这种税也不会有效率损失。

但是如果对商品禀赋征税，问题就出现了。我们应该如何衡量人们的商品禀赋？对于大多数人来说，他们拥有的主要禀赋就是他们的劳动力。人们的劳动禀赋由人们**能够**出售的劳动量而不是实际出售的劳动量组成。对人们实际出售的劳动进行征税，那么这种税就是一种**扭曲税**（distortionary tax）。因为这种税会扭曲人们的劳动供给决策，他们可能会较少劳动供给。但是，对劳动的潜在价值即劳动的禀赋进行征税，则不是扭曲的。根据定义，劳动的潜在价值是不会因为征税而改变的。对劳动的潜在价值进行征税，感觉起来并不难。但是实践起来却几乎不可能，原因在于这种做法是对**可能**出售的东西而不是已经出售的东西征税，如何衡量**可能**出售的东西是个难以解决的问题。

我们可以**想象出**一种征这种税的方法。假设在某个社会中，政府要求每人每周将其 10 小时劳动的收入上缴。这类税收和人们实际工作多长时间无关，它只取决于劳动的禀赋，而不取决于人们实际出售了多长时间的劳动。这类税收实质上是将每个人的劳动时间禀赋，部分转移到政府手中。政府可以使用这些收入来提供各种商品，或者它只是将这些税收转移给其他人。

根据福利经济学第二定理，这样的人头税不是扭曲的。在本质上，任何帕累托有效率的配置都可以通过这样的人头再分配加以实现。

但是，这样的税收体系的根本重建是不会得到人们的拥护的。大多数人的劳动供给决策对于工资率的变动并不敏感，因此对劳动征税引起的效率损失不会太大。但是第二福利定理传达的信息是重要的。价格应该用来反映相对稀缺性。财富在人头间的转移应该用于分配目标的调整。在很大程度上，这两种政策是可以分开的。

由于人们对财富的分配非常关心，这使得他们拥护各种各样的价格管制。例如，有人建议，老年人应该享受更便宜的电话服务，或者小用电户的电费应该比大用电户低。这类建议的细想基本是一致的：通过价格体系达到收入再分配的目的。

如果你仔细想想，你就会明白上述建议是严重缺乏效率的收入再分配方法。如果你想对收入再分配，为何不直接对收入再次进行分配呢？如果你多给某个人一元钱，那么他可以选择多消费任何他想多消费的商品，而不一定是得到补贴的那种商品。

附录

下面我们使用微积分推导出帕累托有效率配置的条件。根据定义可知，一个帕累托有效率配置，在给定其他人的效用水平时，使得每个人的效用尽可能地大。因此我们用 $\bar{u}$ 表示 B 的效用水平，我们的目的是让 A 的效用尽可能地大。

这个最大化问题为:

$$\max_{x_A^1, x_A^2, x_B^1, x_B^2} \quad u_A(x_A^1, x_A^2)$$

$$\text{使得} u_B(x_B^1, x_B^2) = \bar{u}$$

$$x_A^1 + x_B^1 = \omega^1$$

$$x_A^2 + x_B^2 = \omega^2.$$

此处$\omega^1 = x_A^1 + x_B^1$是市场中商品 1 的数量，$\omega^2 = x_A^2 + x_B^2$是市场中商品 2 的数量。这个最大化问题要求我们找到配置$(x_A^1, x_A^2, x_B^1, x_B^2)$，给定 B 的效用水平以及给定每种商品的总量等于禀赋总量的情形下，这个配置要能使得 A 的效用尽可能地大。

我们写出这个问题的拉格朗日函数

$$L = u_A(x_A^1, x_A^2) - \lambda(u_B(x_B^1, x_B^2) - \bar{u}) - \mu_1(x_A^1 + x_B^1 - \omega^1) - \mu_2(x_A^2 + x_B^2 - \omega^2).$$

其中，λ是效用约束条件的拉格朗日乘子，μ_1和μ_2是资源约束的拉格朗日乘子。当我们对每种商品分别求导时，我们就得到了最优解必须满足的四个一阶条件：

$$\frac{\partial L}{\partial x_A^1} = \frac{\partial u_A}{\partial x_A^1} - \mu_1 = 0$$

$$\frac{\partial L}{\partial x_A^2} = \frac{\partial u_A}{\partial x_A^2} - \mu_2 = 0$$

$$\frac{\partial L}{\partial x_B^1} = -\lambda \frac{\partial u_B}{\partial x_B^1} - \mu_1 = 0$$

$$\frac{\partial L}{\partial x_B^2} = -\lambda \frac{\partial u_B}{\partial x_B^2} - \mu_2 = 0$$

将以上式子变形整理，并用第一式除以第二式，用第三式除以第四式，可得

$$MRS_A = \frac{\partial u_A / \partial x_A^1}{\partial u_A / \partial x_A^2} = \frac{\mu_1}{\mu_2} \tag{31.5}$$

$$MRS_B = \frac{\partial u_B / \partial x_B^1}{\partial u_B / \partial x_B^2} = \frac{\mu_1}{\mu_2} \tag{31.6}$$

教材中已经给出了这两个式子的解释：在帕累托有效率的配置上，A 和 B 的边际替代率必然相等。否则，就存在让这两人状况都变好的交易。

回忆一下消费者最优选择必须满足的条件。如果消费者 A 在预算约束下实现了效用最大化，B 也是这样的，而且这两个消费者面对的商品 1 和 2 的价格是相同的，我们必然有：

$$\frac{\partial u_A / \partial x_A^1}{\partial u_A / \partial x_A^2} = \frac{p_1}{p_2} \tag{31.7}$$

$$\frac{\partial u_B / \partial x_B^1}{\partial u_B / \partial x_B^2} = \frac{p_1}{p_2} \tag{31.8}$$

注意这两个式子和前面的效率条件的相似性。效率条件中的拉格朗日乘子μ_1和μ_2，类似于

消费者选择条件中的 p_1 和 p_2。事实上，这类问题中的拉格朗日乘子有时候称为**影子价格**（shadow price）或**效率价格**（efficiency price）。

每个帕累托有效率的配置都必须满足类似（31.5）和（31.6）式的条件。每个竞争均衡都必须满足类似（31.7）和（31.8）式的条件。在市场环境下，刻画帕累托有效率的条件和刻画个人效用最大化的条件在本质上是一样的。

总结

1.一般均衡研究的是，经济整体如何调整到使所有市场同时实现需求等于供给。

2.艾奇沃斯盒是一种图形分析工具，可以用它分析两个消费者和两种商品的一般均衡问题。

3.如果某配置是帕累托有效率的，那么这意味着我们找不到可行的商品再配置的方法，使得所有消费者的状况不变但又让至少一个消费者的状况变好。

4.瓦尔拉斯法则表明，对于所有价格，都有总超额需求的价值之和等于零。

5.一个一般均衡配置是指，在这个配置中，每个人都已经从他能够买得起的消费束集中选择出他最偏好的消费束。

6.一般均衡价格体系决定的只是相对价格，而不是绝对价格。

7.如果价格变动时，每种商品需求的变动是连续的，那么总存在着一组价格，使得每个市场的需求等于供给；也就是说，实现了竞争均衡。

8.福利经济学第一定理是说，竞争均衡是帕累托有效率的。

9.福利经济学第二定理是说，只要消费者的偏好都是凸的，那么每个帕累托有效率配置都可以通过竞争均衡实现。

复习题

1.下列情形有没有可能发生：在某帕累托有效率的配置中，**某些人**的状况比在帕累托无效率的配置中更差。

2.下列情形有没有可能发生：在某帕累托有效率的配置中，**所有人**的状况比在帕累托无效率的配置中更差。

3.对还是错？如果我们以知道合同曲线，那么我们就可以知道任何交易的结果。

4.在帕累托有效率的配置上，某些人的状况还能进一步变好吗？

5.如果在 10 个市场中，已有 8 个市场的超额需求的价值之和等于零，那么剩下的 2 个

市场的结果一定是怎么样的？

复习题答案

1.下列情形有没有可能发生：在某帕累托有效率的配置中，某些人的状况比在帕累托无效率的配置中更差。

【复习内容】帕累托有效率配置的概念

A．对于一个帕累托有效率的配置来说，下列几种说法是等价的：

1.已不存在让所有涉及的人状况更好的方法；

2.已不存在让某些人状况变好但又不使其他人状况变差的方法；

3.交易的所有收益（即好处）都已被取尽；

4.已不存在互利的交易

B．帕累托效率和分配的公平性问题无关。例如，一个人拥有一切，其他人一无所有的配置是有效率的。但这一般来说是不公平的。

【参考答案】有可能发生。

一般来说，平均分配形成的配置是无效率的，但一个人拥有一切、其他人一无所有的配置是有效率的。我们根据这个结论说明题目中的情形是有可能发生的。在说明之前，首先要理解这个结论。

举例说明是最简单的说明方法。例如有 24 单位的香烟和 24 单位的糖果要在甲和乙之间进行分配，平均分配的结果就是甲和乙都分别有 12 单位的香烟和 12 单位的糖果。但这个结果一般来说是无效率的，因为甲和乙的偏好一般是不同的。例如甲愿意以 2 单位糖果换 1 单位香烟，而乙愿意以 2 单位香烟换 1 单位糖果，简单地说甲更喜欢抽烟、乙更喜欢吃糖。这样，仍存在着互利的交易，所以平均分配形成的配置是无效率的。

根据帕累托有效率的定义可知，一个人拥有一切、其他人一无所有是帕累托有效率的，因为此时不存在让所有人的状况都变好的方法。如果想让其他人的状况变好，势必需要损害拥有一切的这个人的利益。比如在上面的例子中，甲拥有 24 单位的香烟和 24 单位的糖果，乙一无所有。这种配置是帕累托有效率的。

比较这两种配置方法可知，乙的状况在平均分配时更好，但这个配置却是无效率的；乙在后一种配置方法中状况更差，但这个配置却是有效率的。因此“在某帕累托有效率的配置中，某些人的状况比在帕累托无效率的配置中更差”这种情形能够发生。

2.下列情形有没有可能发生：在某帕累托有效率的配置中，所有人的状况比在帕累托无效率的配置中更差。

【复习内容】帕累托有效率配置的概念

核心内容见第 1 题。

【参考答案】不可能发生。

反证法。假设能够出现下列情形：在某帕累托有效率的配置（不妨称为 A）中，**所有人**的状况比在帕累托无效率的配置（不妨称为 B）中更差。

这样我们就得到了一个矛盾。因为根据帕累托有效率配置的定义可知，在 A 这个配置中，已不存在让所有人的状况都变好的方法。但根据我们的假设，A 仍存在帕累托改进的余地。因为 B 就是 A 的一种改进方法。这说明我们的假设是错误的。

所以，不可能出现题目中所说的情形。

3.对还是错？如果我们以知道合同曲线，那么我们就可以知道任何交易的结果。

【复习内容】帕累托集；合同曲线

艾奇沃斯盒中的所有帕累托有效率的配置点组成的点集，称为**帕累托集**（Pareto set），或者称为**合同曲线**（contract curve）。后面这个名字来自于下列思想：交易的所有“最终合同”必然位于帕累托集上——否则这些交易不可能是最终的，因为如果不位于帕累托集上，说明仍存在着效率改进的余地，即仍存在着互利的交易。

【参考答案】错误。

根据合同曲线的定义可知，从艾奇沃斯盒内任何一点开始的互利交易最终会在这条合同曲线上，否则交易的好处还没取尽，交易方还会继续进行交易，直至到达合同曲线上的某一点。

但是，最终的结果究竟在合同曲线上的哪一个点上，仅根据题目给出的信息，我们是无法找到这个点的。一般来说，不仅需要知道禀赋在盒内的位置，还需要知道市场价格。

4.在帕累托有效率的配置上，某些人的状况还能进一步变好吗？

【复习内容】帕累托有效率配置的概念

核心内容请见本章复习题第 1 题。

【参考答案】

对于一个帕累托有效率的配置来说，已不存在让某些人状况变好但又不使其他人状况变差的方法。这也意味着在帕累托有效率的配置上可以让一些人的状况变好，但是前提是让其他人的状况变坏了。

比如在 A 配置中，甲拥有一切，乙一无所有。现在我们能让乙的状况变好——将甲的财富全部“掠夺”给乙，这样形成 B 配置：甲一无所有，乙拥有一切。最后多说一句，注意 A 配置和 B 配置都是帕累托有效率的。

5.如果在 10 个市场中，已有 8 个市场的超额需求的价值之和等于零，那么剩下的 2 个市场的结果一定是怎么样的?

【复习内容】瓦尔拉斯法则

瓦尔拉斯法则表明$\sum_{i=1}^{n} p_i z_i \equiv 0$，其中 z_i 表示第 i 种商品的总超额需求，即第 i 个市场的纵超额需求。这个式子是说，总超额需求的价值之和恒等于 0。

【参考答案】

不妨令剩下的市场为市场 9 和市场 10。

由瓦尔拉斯法则可知，$\sum_{i=1}^{10} p_i z_i \equiv 0$，即 $(\sum_{i=1}^{8} p_i z_i) + p_9 z_9 + p_{10} z_{10} \equiv 0$

由题目知$\sum_{i=1}^{8} p_i z_i = 0$，因此必然有 $p_9 z_9 + p_{10} z_{10} = 0$。也就是说剩下两个市场的总超额需求价值之和必定等于零。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

32.生产（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

32 生产

在上一章，我们介绍了纯交换经济的一般均衡模型，并相应讨论了每种商品数量既定时的资源配置问题。在本章，我们试图在一般均衡框架内加入生产因素。当存在生产因素时，商品数量不再是固定的而是随市场价格变化而变化的。

我们在上一章对交易行为的分析是在两个消费者两种商品的框架内进行的，如果你认为这个框架限制性较强，那么你能想象得出我们将在什么样的框架内分析生产行为吗？加入生产因素后，分析框架的最低要求是：一个消费者；一个企业；和两种商品。这种经济模型通常称为**鲁宾逊.克鲁索经济**（Robinson Crusoe economy），这个名字来自笛福（Defoe）小说《鲁宾逊漂流记》中主人公的名字。

32.1 鲁宾逊.克鲁索经济

在这种经济中，鲁宾逊身担二职：既是消费者又是生产者。他可以在沙滩上闲逛消费闲暇，他也可以作为生产者搜集椰子。他收集的椰子越多，他的食物就越多，但是他用于将皮肤晒成棕褐色的时间就越少。

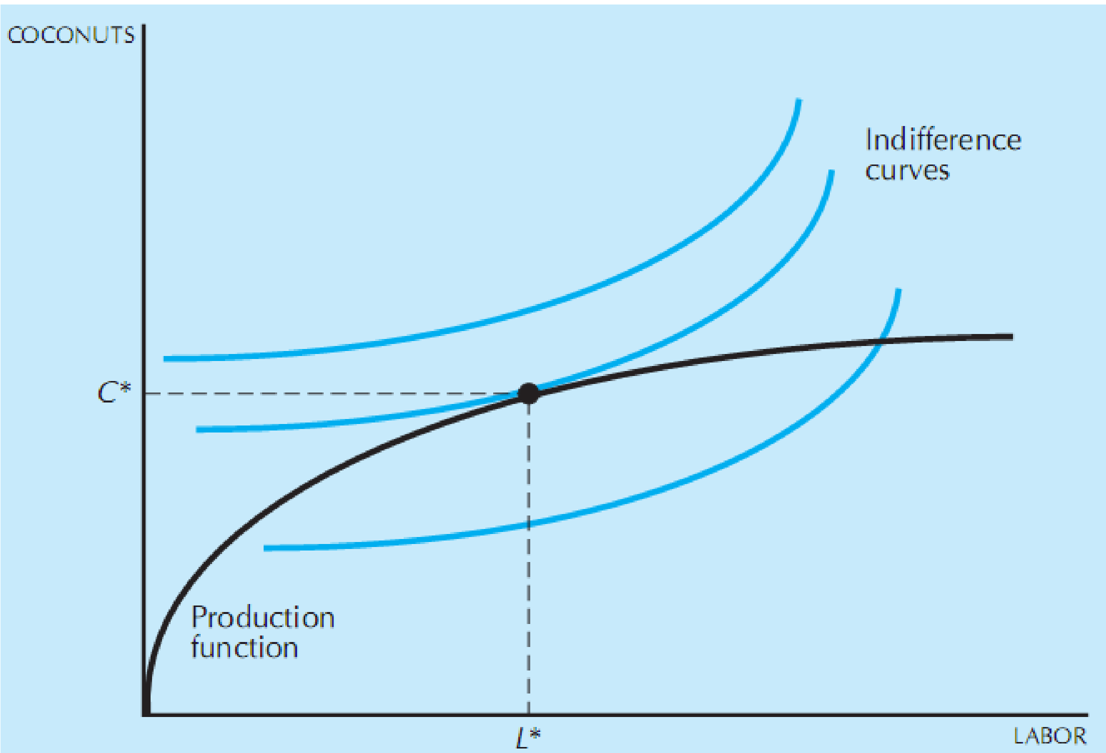


图 32.1：鲁宾逊.克鲁索经济。图中的无差异曲线表示鲁宾逊对椰子和闲暇的偏好。生产函数表示他工作时间和椰子产量之间的技术关系。

鲁宾逊对椰子和闲暇的偏好如图 32.1 所示。这个图和第 9 章中消费者对闲暇和消费的偏好的图是类似的，只不过此处横轴表示的是劳动而不是闲暇。到目前为止，尚未增加任何新内容。

现在我们在图中加入**生产函数**（production function）曲线，这条曲线表示鲁宾逊工作时间和椰子产量之间的关系。该曲线的形状通常如图 32.1 所示。也就是说，鲁宾逊工作时间越长，他收集的椰子越多；但是由于劳动的边际报酬递减，他的劳动的边际产量是递减的：随着劳动时间的增加，额外一单位劳动搜集的椰子数量是递减的。

鲁宾逊应该工作多长时间和消费多长时间（的闲暇）？为了回答这些问题，找出恰好与生产函数相切的那条位置最高的无差异曲线。在他收集椰子的技术一定时，切点就是他最偏好的劳动和消费组合。

在切点上，无差异曲线的斜率必定等于生产函数的斜率，推理如下：如果这两条曲线相交，那么肯定还存在着他更偏好的可行组合点。这意味着额外一小时劳动的边际产量，必定等于闲暇和椰子之间的边际替代率。如果边际产量大于边际替代率，则鲁宾逊应该稍微放弃一些闲暇，生产更多的椰子。如果边际产量小于边际替代率，鲁宾逊应该稍微减少工作时间。

32.2 克鲁索股份有限公司

到目前为止，我们只对以前的模型稍微做了一些扩展。现在对该模型加入一些新的特征。假设鲁宾逊厌倦了同时扮演生产者和消费者的角色，他决定隔天扮演一个角色：某一天他只作为生产者，另外一天他只作为消费者。为了协调这些行为，他决定建立劳动市场和椰子市场。

他还建立了一个公司，名字叫做克鲁索有限公司，他是这个公司的唯一股东。该公司根据市场上劳动和椰子的价格，决定雇佣劳动的数量和生产椰子的数量，假设该公司追求利润最大化。鲁宾逊此时的角色和任务如下：作为雇员，他要在该公司工作以获得收入；作为股东，他占有利润；作为消费者，他要决定购买公司产品（椰子）的数量。（显然这样的说法有些奇怪，但在一个荒岛上你还能怎么描述呢？）

为了记录他的这些交易，鲁宾逊发明了一种货币，他称之为“元”，他有些武断地将每个椰子的价格设定为一元。这样，椰子就成为了这个经济的计价物；我们在第 2 章已经知道，计价物是指价格被调整为一元的商品。由于椰子的价格已被标准化为 1，我们剩下的工作是确定工资率。鲁宾逊的工资率为多大时才能使得市场运行？

我们从两个角度分析这个问题。首先从克鲁索有限公司的角度，然后再从鲁宾逊作为消费者的角度进行分析。这样的讨论有时显得荒唐，但是在只有一个人的经济中，你应该忍受这样的分析。假设这个经济已运行一段时间并且处于均衡状态。均衡时，椰子的需求等于椰子的供给，劳动的需求等于劳动的供给。克鲁索有限公司和消费者都在自己的约束条件下作出最优选择决策。

32.3 企业

每天傍晚，克鲁索有限公司作出以下决策：雇佣多少劳动和生产多少椰子。给定椰子的价格 1 和劳动工资率 w，我们可以求解图 32.2 表示的企业利润最大化问题。我们首先考虑能产生既定利润水平 π 的椰子和劳动的所有组合。它的表达式为

$$\pi = C - wL .$$

其中，C 表示椰子数量，L 表示劳动数量。

解出 C，可得

$$C = \pi + wL .$$

这个式子就是第 19 章已介绍过的等利润线——能够产生既定利润 π 的劳动和椰子的所有组合。克鲁索有限公司选择在利润最大化的点进行生产。和往常一样，这意味着一个相切条件：生产函数的斜率即劳动的边际产量必定等于 w，如图 32.2 所示。

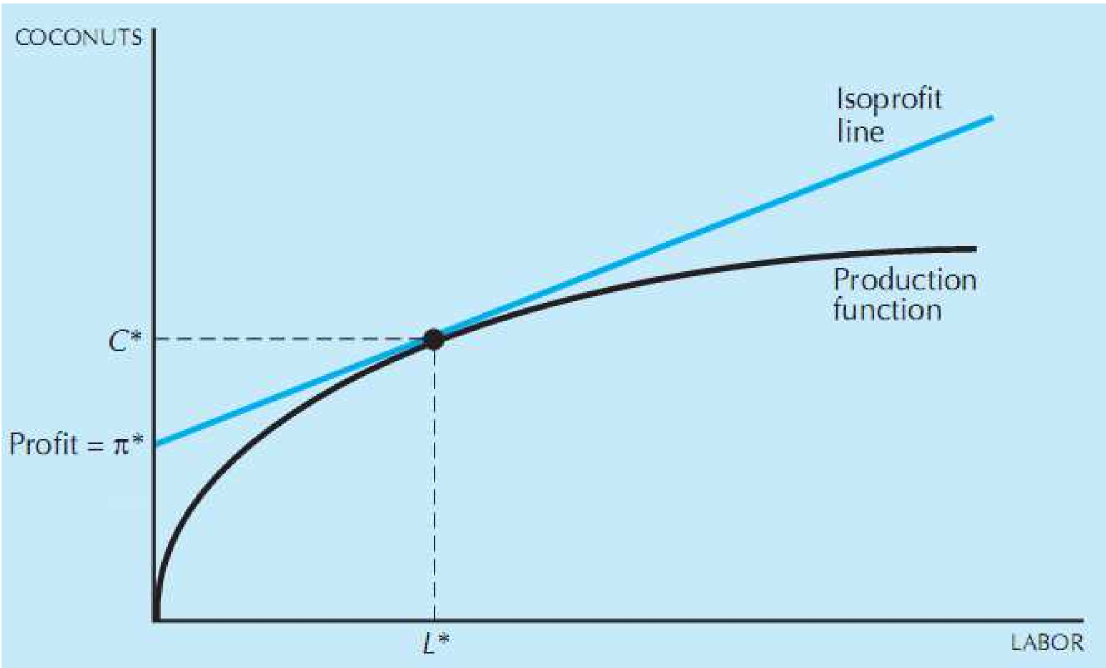


图 32.2：**利润最大化**。克鲁索有限公司选择能使利润最大化的生产方案。在最优点，生产函数曲线必定和一条等利润线相切。

因此，这条等利润线的纵截距表示最大的利润水平（以计价物椰子作为衡量单位）：如果鲁宾逊的利润为 π^* 元，他可以购买 π^* 单位椰子，因为椰子的价格为 1。在最优点就是这样的。克鲁索有限公司已经完成了自己的工作。给定工资率 w，它决定雇佣多少劳动，生产多少椰子，以及该方案能生产多少利润。因此，克鲁索有限公司宣布股息为 π^* 元，并将它邮寄给唯一的股东鲁宾逊。

32.4 鲁宾逊的问题

第二天鲁宾逊从睡梦中醒来后就收到了 π^* 元股息。他边吃着椰果早餐，边思考应该劳动多少消费多少。他可能想只消费自己的禀赋：将 π^* 元利润购买椰子以及消费他的闲暇禀赋。但是听着肚子咕咕叫让他不舒服，因此他也许应该多工作几个小时。于是鲁宾逊迈着沉重的步伐向克鲁索公司走去，开始收集椰子，这就是他的工作日所干的工作。

我们可用标准的无差异曲线分析方法描述鲁宾逊的劳动-消费选择问题。用横轴表示劳动时间、用纵轴表示椰子数量，我们可以画出他的无差异曲线，如图 32.3 所示。

由于劳动对于鲁宾逊来说是一种厌恶品，椰子则是他喜欢的商品，因此无差异曲线的斜率为正，如图所示。如果我们用 $\overline{L}$ 表示最长劳动时间，则从 $\overline{L}$ 到他实际选择的劳动时间之间的距离，表示他对闲暇的需求。这个图和第 9 章劳动供给的图形类似，只不过此处用横轴（正的方向）表示劳动时间。

图 32.3 还画出了鲁宾逊的预算线。这条预算线的斜率为 w 并且经过禀赋点 $(\pi^*,0)$ 。（鲁宾逊的禀赋为：劳动时间为零，椰子的数量为 π^* ，因为如果他不参加市场交易，这就是他的消费束。）给定工资率 w ，鲁宾逊选择劳动的最优时间以及椰子的最优消费量。在最优消费处，椰子和闲暇的边际替代率必定等于工资率，这一点和标准的消费者选择问题是一样的。

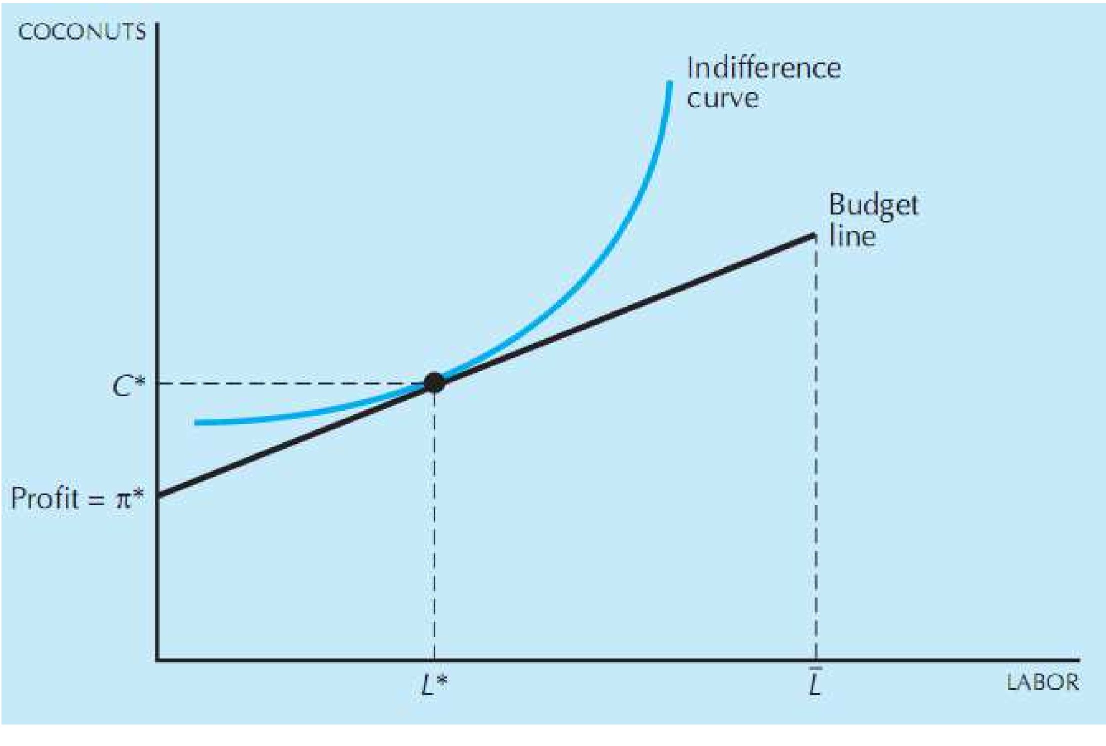


图 32.3：鲁宾逊的最大化问题。在椰子价格和工资率给定时，鲁宾逊作为消费者，他要决定最优工作时间和椰子的最优消费量。最优点位于无差异曲线和预算线的切点之处。

32.5 将两张图放在一起

现在我们将图 32.2 和图 32.3 画在同一个图中，就得到了图 32.4。看看发生了什么！鲁宾逊奇怪的行为竟然产生了正常的结果。他最终的消费量，恰好就是如果他同时做出生产和消费全部决策时的消费量。使用市场体系得到的结果，与直接选择消费和生产方案的结果是相同的。

由于闲暇和椰子的边际替代率等于工资率，劳动的边际产量等于工资率，这就意味着劳动和椰子的边际替代率等于劳动的边际产量，也就是说，无差异曲线的斜率和生产函数的斜率是相等的。

在只有一个人的经济中，使用市场有些愚蠢。为什么鲁宾逊要不嫌麻烦地将生产和消费决策一分为二？这不合理。但是在很多人参与的经济中，将生产和消费决策分开来考虑就不再令人感到奇怪。如果市场中的企业数量众多，问询每个消费者需要每种商品的数量的做法，行不通。在市场经济中，企业要做的事情只是观察市场中商品的价格，从而进行生产决策。这是因为，商品的价格，衡量消费者对**额外**一单位商品的支付意愿。在大多数场合下，企业面对的决策，是应该多生产一些还是少生产一些。

市场价格反映企业用于投入和产出的商品的边际价值。如果企业使用利润的变化来指导生产，其中利润是以市场价格衡量的，那么它们的决策将反映商品对消费者的边际价值。

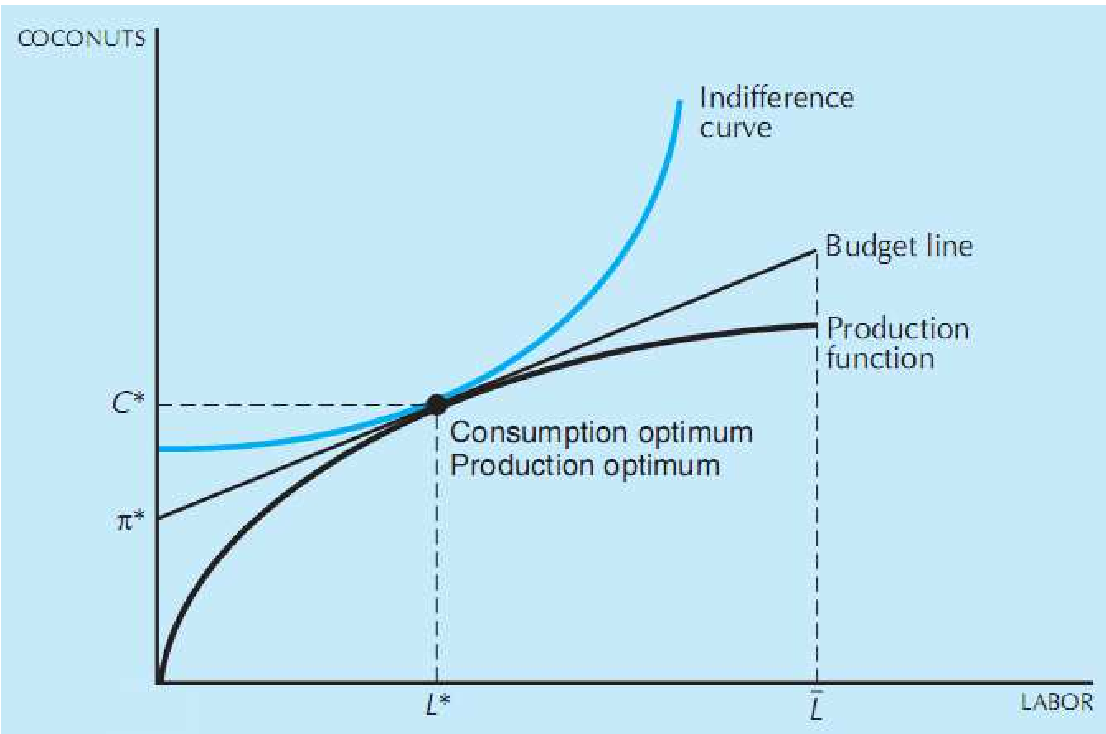


图 32.4: 消费和生产的同时均衡。鲁宾逊作为消费者对椰子的需求，等于克鲁索公司生产（供给）的椰子数量。

32.6 不同的生产技术

在上面的分析中，我们假设鲁宾逊使用的生产技术具有边际报酬递减的性质。由于劳动是生产的唯一投入要素，这等价于规模报酬递减。（如果投入要素多于一种，这个结论不一定正确。）

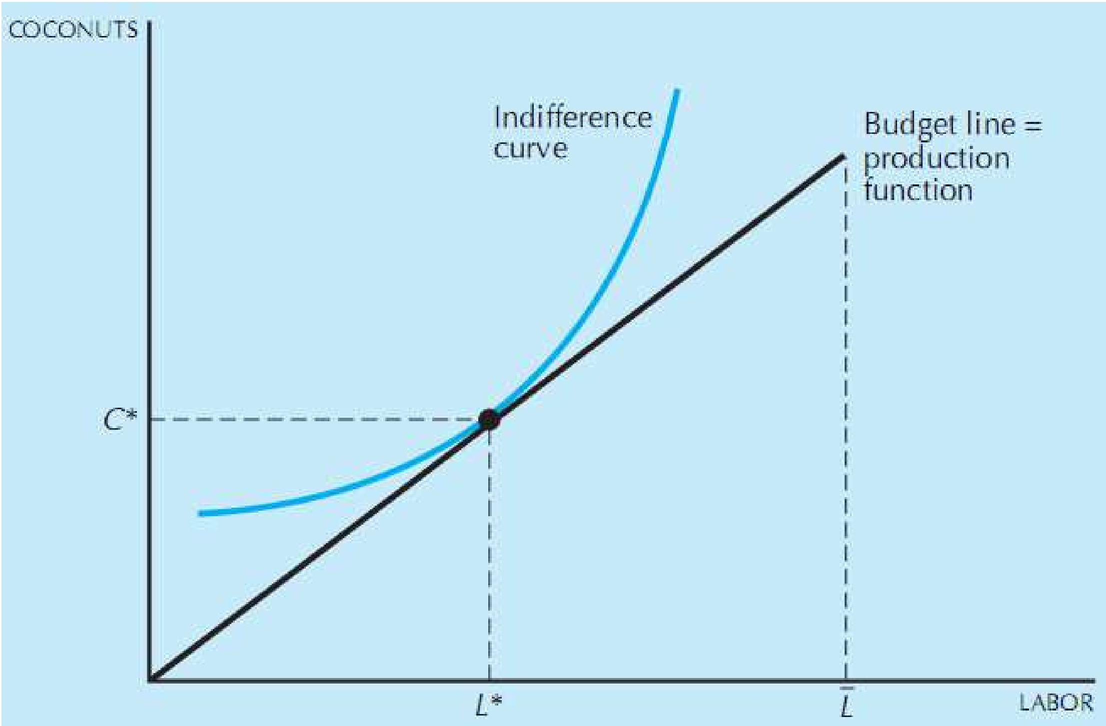


图 32.5：规模报酬不变。如果生产技术是规模报酬不变的，克鲁索公司的利润为零。

有必要考虑其他生产技术。例如，假设生产技术是规模报酬不变的。我们知道规模报酬不变是指，所有要素投入变为原来的 2 倍时，产量也变为原来的 2 倍。在只有一种投入要素的情形下，这意味着生产函数是经过原点的一条直线，如图 32.5 所示。

由于生产技术是规模报酬不变的，第 19 章的结论表明，竞争企业的唯一合理的生产之处是利润等于零的地方。这是因为如果利润大于零，则企业扩大产量无疑是有利的，如果利润小于零，则企业应该停产（产量为零）。

因此，鲁宾逊的禀赋为：零利润和他的初始劳动时间 $\bar{L}$ 。他的预算集和生产集重合，这种情形下，分析过程和结果和前面类似。

如果生产技术是规模报酬递增的，则情形就稍有不同，如图 32.6 所示。在这个简单的例子中，容易看出鲁宾逊对椰子和闲暇的最优选择。和以前一样，无差异曲线将和生产函数曲线相切。但在证明该点是利润最大化点时，问题就产生了。因为如果企业面对的价格等于

鲁宾逊的边际替代率，企业希望生产的产量就会大于鲁宾逊的需求量。

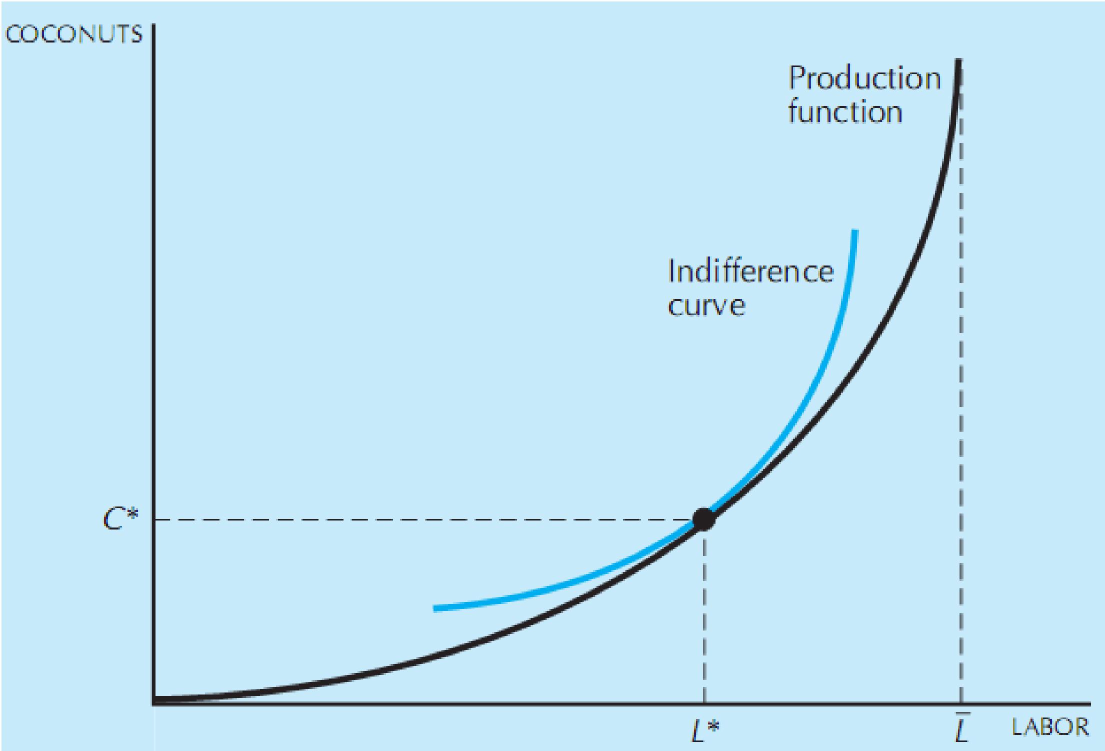


图 32.5:规模报酬递增。如果生产技术是规模报酬递增的，那么竞争市场无法实现帕累托有效率的结果。

如果在最优选择之处，企业的生产技术仍是规模报酬递增的，那么它的平均成本就会大于边际成本，这意味着企业的利润为负。追求利润最大化的目标会指引企业扩大产量，但这个产量和产品的需求量不相容，也和消费者提供的要素数量不相容。在这种情形下，**不会**存在这样的价格，它能使得消费者的效用最大化的需求等于企业利润最大化的产量。

规模报酬递增意味着生产集是**非凸的**（nonconvexity）。也就是说，在这种情形下，经济中椰子和劳动的可行的集合不是凸集。因此，图 32.6 中的无差异曲线和生产函数曲线在点 (L^*, C^*) 的公切线，不可能像图 32.4 一样将最优点和其他可行点分离开。

像这样的非凸性给竞争市场的运行造成了灾难性的后果。在竞争性市场上，消费者和企业只需要看一组相同的数字——不同商品的市场价格——确定他们的消费和生产决策。如果生产技术和偏好都是凸的，那么经济参与人在做决策时需要知道的仅仅是价格和边际替代率的关系，准确地说是在当前经济生产之处的价格和边际替代率的关系：价格告知了经济参与人在确定资源有效率配置时的一切信息。

但是，如果生产技术和/或偏好是非凸的，那么价格不能传递选择有效率配置的一切信息。在这种情形下，除了要知道当前运行状态时的生产函数和无差异曲线的斜率，也必须知

道其他状态时的生产函数和无差异曲线的斜率。

然而，这些结论仅适用于当规模报酬相对于市场规模来说比较大时。对于竞争市场来说，在较小产量区域内出现的规模报酬递增现象不会造成太大的麻烦。

32.7 生产和福利经济学第一定理

我们已经知道，在一个纯粹交换的经济中，竞争市场均衡是帕累托有效率的。这个事实称为福利经济学第一定理。对于生产行为来说，这个定理还成立吗？上面使用的图形分析方法不足以回答这样的问题，但是使用第 31 章代数论证的方法是可行的。可以证明，答案为肯定的：如果所有企业的行为是竞争的而且是追求利润最大化的，那么竞争均衡是帕累托有效率的。

和 31 章的第一定理一样，这里也需要提醒读者注意以下几点。首先，它和分配公平性问题无关。利润最大化仅能保证效率，而不是公平！第二，这一结果只有在竞争均衡确实存在时才有意义。特别地，这一定理不适用于在较大产量范围内存在规模报酬递增现象时。第三，这个定理隐含地假设任何一个企业的选择都不会影响其他企业的生产可能性。也就是说假设不存在**生产外部性**（production externalities）。类似地，该定理还要求企业的生产决策不会直接影响消费者的消费可能性；也就是说假设不存在着**消费外部性**(consumption externalities)。我们将在第 33 章介绍外部性的概念，在那一章我们将详细分析外部性对有效率配置的影响。

32.8 生产和福利经济学第二定理

在纯交换的经济中，只要消费者的偏好都是凸的，则每个帕累托有效率的配置都是一个竞争均衡。在涉及生产的经济中，上述结论仍然成立，但是现在我们不仅要求消费者的偏好都是凸的，还要要求企业的生产集都是凸的。正如前文指出的，这个要求有效排除了规模报酬递增的情形：如果企业在均衡产量上仍是规模报酬递增的，那么在均衡价格水平它们希望生产得更多。

然而，在规模报酬不变或递减的情形下，福利经济学第二定理运行良好。任何帕累托有效率的配置都能通过竞争市场实现。当然，一般来说，为了得到不同的帕累托有效率配置，需要将禀赋在消费者中重新进行分配。特别地，需要重新分配劳动禀赋得到的收入，以及企业的股份。正如上一章指出的，这类再分配实践起来非常困难。

32.9 生产可能性

我们已经知道在一种投入-一种产出的经济中，生产和消费的决策是如何作出的。在本节，我们试图将上述模型推广得到多种投入和产出的经济中。尽管我们的分析只涉及两种商

品，这些概念和结论可以自然地推广到很多种商品。

因此假设鲁宾逊还生产其他的商品，比如鱼。他可以将他的劳动时间用于收集椰子或者用于钓鱼。通过在这两种生产活动中分配不同的时间，鲁宾逊可以得到椰子和鱼的各种组合，如图 32.7 所示。这个集合称为**生产可能集**（production possibilities set），生产可能集的边界称为**生产可能边界**（production possibilities frontier）。注意生产可能集和生产函数的区别，生产函数描述的是投入和产出之间的关系，而生产可能集描述的只是可行的产出集。（在更高级的处理中，投入集和产出集都可以视为生产可能集的一部分，但是这样的处理在二维图形中不容易操作。）

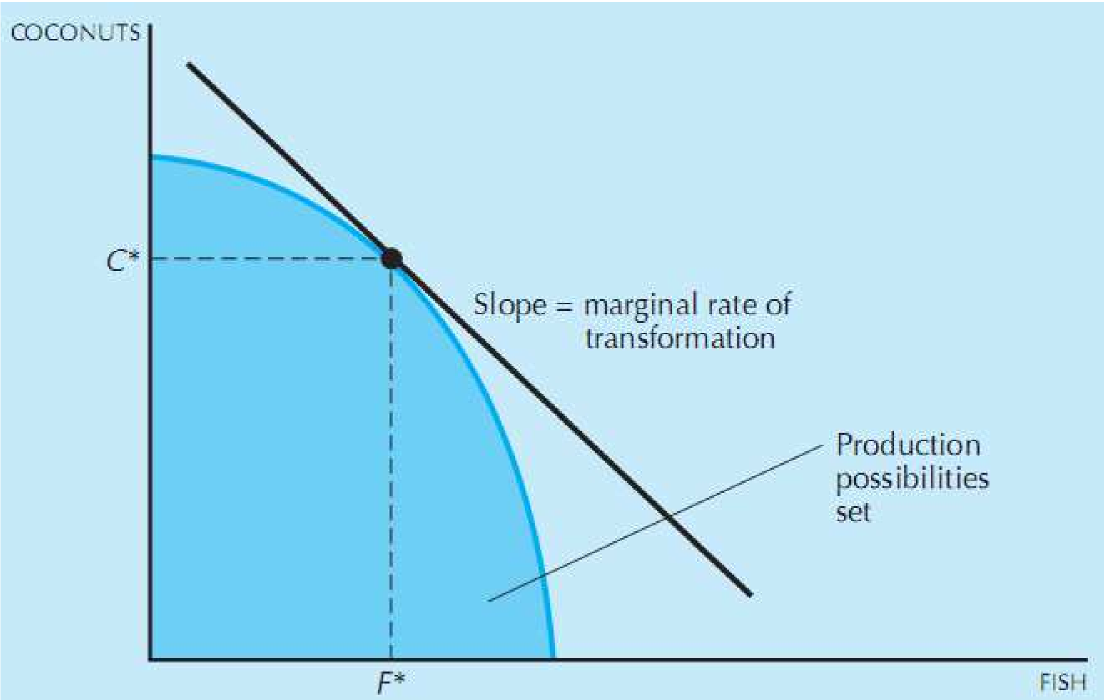


图 32.7：生产可能集。生产可能集衡量在技术和投入一定时，所有可行的产出集。

生产可能集的形状取决于潜在的技术性质。如果椰子和鱼的生产技术是规模报酬不变的，则生产可能集的形状非常简单。由于根据假设投入要素只有一种，即鲁宾逊的劳动，鱼和椰子的生产函数都是劳动的**线性**函数。

例如，假设鲁宾逊每小时的劳动可以生产 10 单位鱼或者 20 单位椰子。于是如果他用 L_f 小时生产鱼、用 L_c 小时生产椰子，他可以生产$10L_f$ 单位的鱼和 $20L_c$ 单位的椰子。假设鲁宾逊决定每天工作 10 个小时，则生产可能集由椰子（C）和鱼（F）的满足下列条件的所有组合构成：

$$F = 10L_f$$
$$C = 20L_c$$
$$L_f + L_c = 10.$$

前两个式子衡量的是生产关系，第三个式子表示的是资源约束。为了确定生产可能边界，从前两个式子解出 L_f 和 L_c，可得

$$L_f = \frac{F}{10}$$
$$L_c = \frac{C}{20}$$

现在把这两个式子加在一起，并利用 $L_f + L_c = 10$，可知

$$\frac{F}{10} + \frac{C}{20} = 10.$$

这个式子表示如果鲁宾逊每天工作 10 个小时，他能得到的鱼和椰子的所有组合。如图 32.8A 所示。

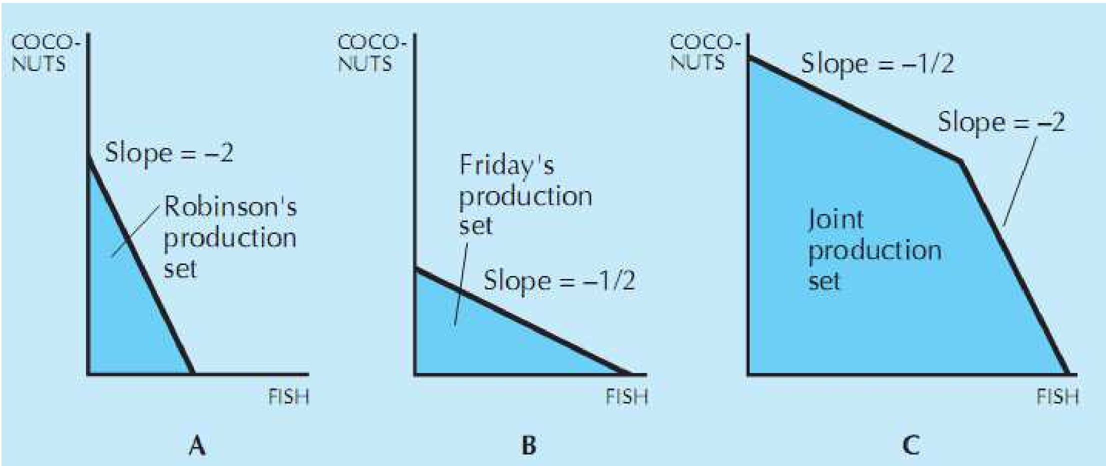


图 32.8：共同的生产可能集。鲁宾逊的生产可能集（图 A），星期五的生产可能集（图 B），以及这两人共同的生产可能集（图 C）。

这条生产可能边界的斜率称为**边际转换率**（marginal rate of transformation），它衡量如果鲁宾逊放弃一种商品，他能得到的另外一种商品的数量。如果鲁滨孙减少生产一单位鱼的时间，而将这个时间用于生产椰子，他可以多生产 2 单位椰子。换一种说法：如果鲁宾逊减少一小时生产鱼的时间，他少生产 10 单位鱼。但是如果他将这个时间用于生产椰子，他可以多生产 20 单位的椰子。因此，椰子和鱼的转换率为 2:1。

32.10 相对优势

上一节构造的生产可能集非常简单，这是因为我们假设鱼和椰子的生产方法分别只有一种。如果每种商品的生产方法有多种，结果将是怎样的?现在我们对这个孤岛经济上再增加一个人，他生产鱼和椰子的技术与鲁宾逊不同。

具体地说，我们将这个新人叫做星期五（Friday），假设他每小时可以生产 20 单位鱼或者 10 单位椰子。因此如果星期五工作 10 小时，他的生产可能集将由下列这些条件决定：

$$F = 20L_f$$
$$C = 10L_c$$
$$L_f + L_c = 10.$$

类似前面对鲁宾逊生产可能集的计算，我们可以求出星期五的生产可能集：

$$\frac{F}{20} + \frac{C}{10} = 10.$$

星期五的生产可能集如图 32.8B 所示。注意椰子和鱼的边际转换率为 $\Delta C / \Delta F = -1/2$，而鲁宾逊的这个边际转换率为 -2。星期五放弃每单位的椰子，他可以生产两单位的鱼；鲁宾逊放弃每单位鱼，他可以生产两单位的椰子。在这样的情形下，我们说星期五在鱼的生产上具有**比较优势**或**相对优势**（comparative advantage）[一]，而鲁宾逊则在椰子的生产上具有相对优势。在图 32.8 中，我们画出了三个生产可能集：A 图是鲁宾逊的生产可能集；B 图为星期五的生产可能集；C 图为这两个人共同的(joint)生产可能集——这两个人可能生产的每种产品的产量加和。

共同的生产可能集结合了这两个人的最优的生产可能集。如果两个人都生产椰子，可以得到 300 单位椰子，其中星期五生产了 100 单位、鲁宾逊生产了 200 单位。如果我们想得到更多的鱼，显然应该让更擅长生产鱼的人，即星期五，从椰子的生产中抽调出来，让他生产鱼。星期五每少生产一单位椰子，就可以多生产 2 单位的鱼。因此，共同的生产可能曲线的斜率为 $-1/2$，这正好等于星期五的边际转换率。

当星期五生产 200 单位鱼时，他一天的工作时间就用完了。如果我们想生产更多的鱼，我们必须使用鲁宾逊。从这一点起，共同的生产可能集的斜率变为 -2，因为这是沿着鲁宾逊的生产可能集进行生产的。最终，如果我们想生产尽可能多的鱼，那么鲁宾逊和星期五必须将生产时间全用于生产鱼，这样我们可以得到 300 单位的鱼，其中星期五生产了 200 单位，鲁宾逊生产了 100 单位。

由于这两个人在不同商品的生产上分别具有相对优势，因此共同的生产可能集就会出

[一] 相对优势是和绝对优势相对而言的。为了彻底理解这一对概念，举例如下。如果甲和乙两人单位时间内可以生产椰子的数量分别为 30 和 30，但和甲和乙两人单位时间生产鱼的数量分别为 10 和 15，则乙在鱼的生产上具有绝对优势，因为乙生产鱼的效率绝对比甲高；乙在鱼的生产上也具有比较优势，因为以椰子的生产作为参照，乙的转换率 15/30 大于甲的转换率 10/30。注意，甲在椰子的生产上具有比较优势，因为以鱼作为参照，甲的转换率 30/10 大于乙的转换率 30/15。译者注。

现一个“折弯”处（kink），如图 32.8 所示。在这个例子中，折弯处只有一个，这是由于生产方法只有两种，即鲁宾逊的生产方法和星期五的生产方法。如果生产方法非常多，那么生产可能边界将非常“圆滑”，如图 32.7 所示。

32.11 帕累托效率

在上两节，我们已经知道如何构建生产可能集，生产可能集描述的是经济整体能得到的所有可行消费束。本节我们将从这些可行消费束中选择出帕累托有效率的消费束。

令 (X^1, X^2) 表示总消费束（aggregate consumption bundles）。它表示可用于消费的商品 1 和 2 的总量分别为 X^1 和 X^2。在克鲁索-星期五经济中，两种商品分别为椰子和鱼，但是我们把它们抽象为 (X^1, X^2)，以便和第 31 章的分析类似。如果我们知道了每种商品的总数量，我们就可以画出艾奇沃斯盒，如图 32.9 所示。

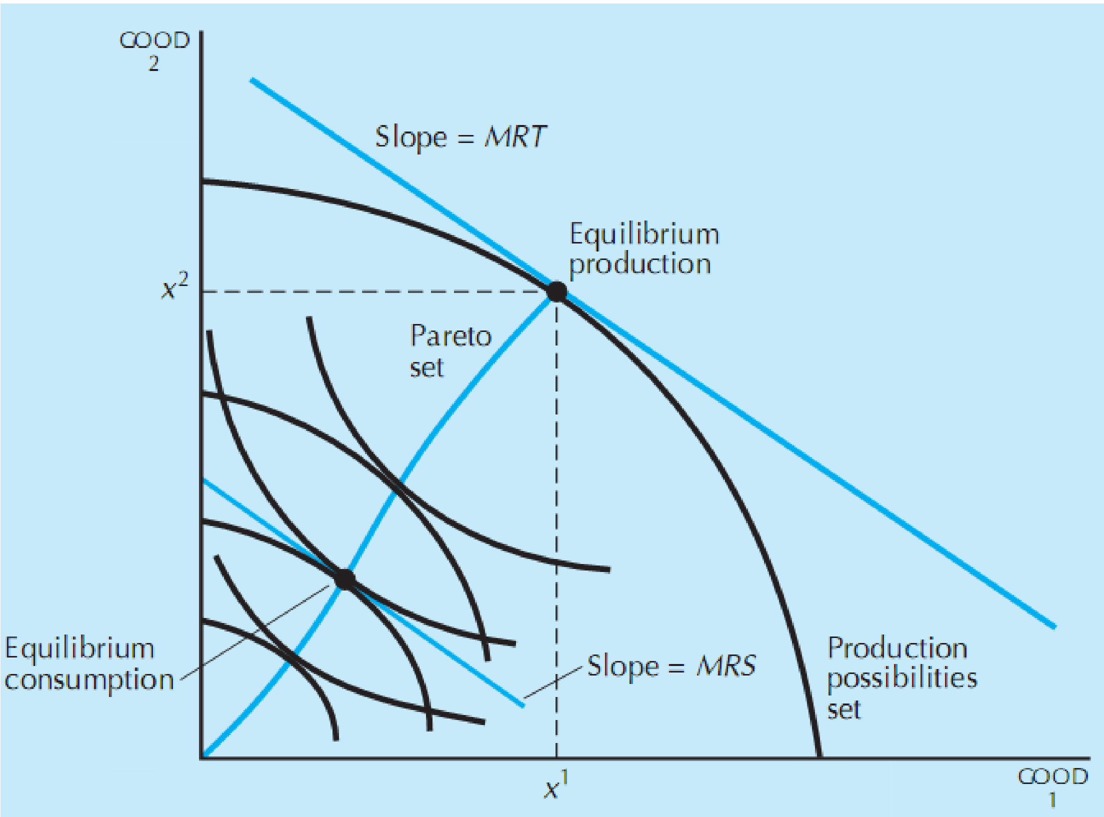


图 32.9：生产和艾奇沃斯盒。在生产可能边界上的任何一点，我们都可以画出一个艾奇沃斯盒，来说明可能的消费配置。

给定 (X^1, X^2)，帕累托有效率消费束的集合就和上一章的分析是一样的：帕累托有效率的消费束必然位于帕累托有效率集（即合同曲线）上，即两个人无差异曲线切点的轨迹，如图 32.9 所示。这些配置点都具有下面的特征：每个消费者的边际替代率，即每个消费者正好愿意交易的比率，必定等于另外一个消费者的。

如果仅考虑消费者的决策，也就是说在纯交换的经济中进行考察，那么这些配置都是帕累托有效率的。如果人们只是想用一种商品交换另外一种商品，那么帕累托集中的任何一个消费束都表示交易的所有好处都被取尽。但在具有生产因素的经济中，存在着另外一种“交换”商品的方法，即少生产一种商品并且多生产另外一种商品。

帕累托集描述在商品 1 和 2 的数量给定时能得到的帕累托有效率的消费束的集合，但是在含有生产因素的经济中，这些商品数量本身就可以从生产可能集中选择出。那么，从生产可能集中选择出什么样的商品组合才是帕累托有效率的选择？

我们思考一下边际替代率条件代表的逻辑。我们断言，在帕累托有效率的配置中，消费者 A 的边际替代率等于消费者 B 的边际替代率：每个消费者正好愿意按照该比率进行交易。如果他们的边际替代率不相等，那么就存在使得这两个人的状况都变好的交易。

我们已知道**边际转换率**（Marginal Rate of Transformation, MRT），衡量一种商品“转换”为另外一种商品的比率。当然，这不是说一种商品变成另外一种商品，而是说，人们改变要素的使用量从而少生产一种商品且多生产另外一种商品。

假设经济在下列位置运行：在该位置上，某个消费者的边际替代率**不等于**这两种商品的边际转换率。那么，这个位置不可能是帕累托有效率的。为什么？因为在这一点上，消费者愿意用商品 1 交换商品 2 的比率，不等于商品 1 转换为商品 2 的比率，在这种情形下，我们可以找到能使该消费者的状况变得更好的生产方案。

例如，假设该消费者的边际替代率为 1，也就是说他愿意用一单位商品 1 交换一单位商品 2。假设边际转换率为 2，也就是说少生产一单位商品 1，社会可以多生产两单位商品 2。显然少生产一单位商品 1 是合理的，因为这样可以多生产两单位商品 2。由于该消费者正好在一单位商品 1 和一单位商品 2 之间无差异，现在他用一单位商品 1 可交换两单位商品 2，因此，他的状况显然变好了。

当其中一个消费者的边际替代率不等于商品的边际转换率时，可作类似推理：重新安排消费和生产方案就能使得该消费者的状况变好。我们已经知道，帕累托有效率时，每个消费者的边际替代率都应该相等，上面的论断意味着每个消费者的边际替代率事实上必须等于边际转换率。

图 32.9 给出了一个帕累托有效率的配置。在这个配置点上，每个消费者的边际替代率都是相等的，这是因为他们的无差异曲线在艾奇沃斯盒内相切。而且在该配置点上，每个消费者的边际替代率必然等于边际转换率，即等于生产可能边界的斜率。

32.12 漂流者股份有限公司（Castaways, Inc.）

在上一节我们推导出了帕累托有效率的必要条件：每个消费者的边际替代率必定等于边际转换率。这意味着资源的任何分配方式，如果是帕累托有效率的，则必然要满足上述条件。前文指出，在一个竞争的经济中，如果企业是追求利润最大化的，而且消费者是追求效

用最大化的，那么它们就能导致一个帕累托有效率的配置。在本节，我们研究其中的机理。

现在，经济中有两个消费者：鲁宾逊和星期五。市场中有四种商品：两种生产要素（鲁宾逊的劳动和星期五的劳动）和两种最终商品（椰子和鱼）。假设鲁宾逊和星期五都是某公司的股东，我们将这个公司称为漂流者股份有限公司。当然，这个公司的雇员（生产者）只有两个人，他们同时又是消费者。但是，和前文一样，我们轮流考察这两种角色，并且假设这两个人无法看得更远。这样假设的原因是因为我们的目的在于，让读者理解一个**分权式**（decentralized）资源配置系统是如何运行的[一]：在这种配置系统中，每个人只决定他自身的决策，而不考虑经济整体的运行。

我们从漂流者公司开始分析，考虑它的利润最大化问题。该公司使用两种生产要素，即克鲁索的劳动（L_C）和星期五的劳动（L_F），生产两种产品——椰子 C 和鱼 F。假设椰子和鱼的价格分别为 p_C 和 p_F，克鲁索和星期五的劳动工资率分为为 w_C 和 w_F，则利润最大化问题为

$$\max_{C,F,L_F,L_C} p_C C + p_F F - w_C L_C - w_F L_F$$

这个最大化问题的约束条件为：生产可能集表示的技术约束。

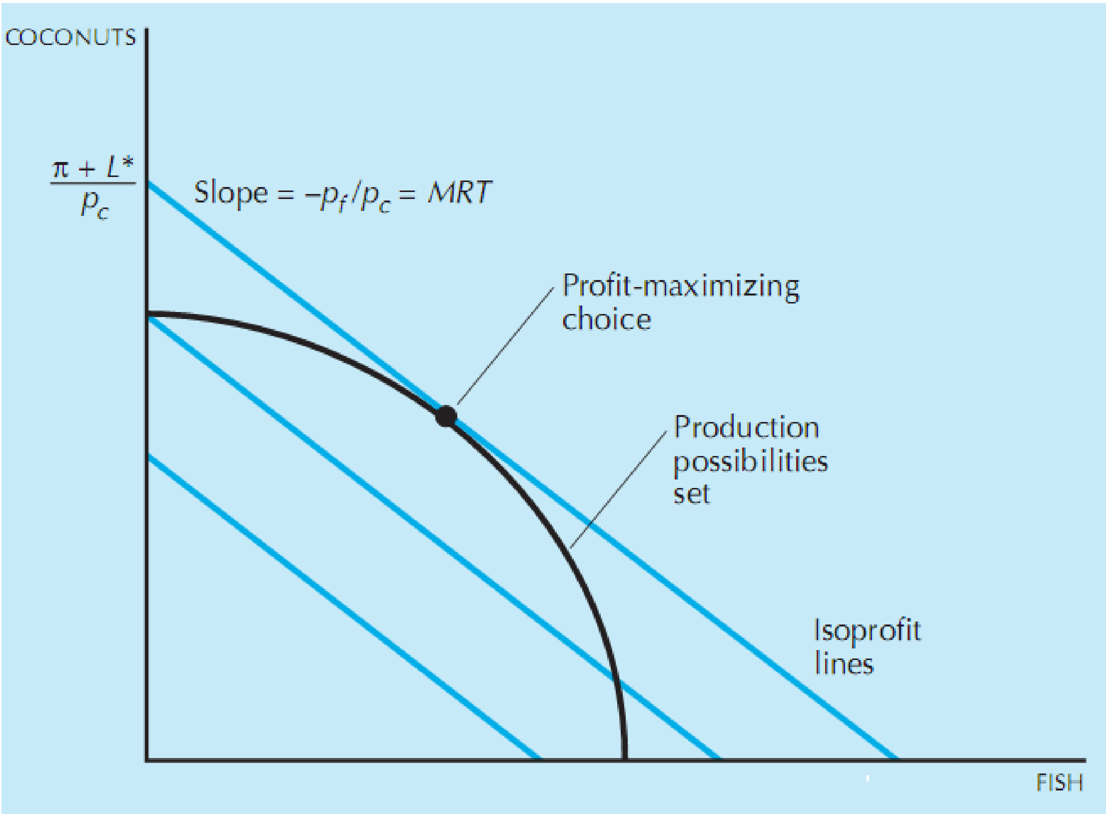


图 32.10：利润最大化。在利润最大化的点，边际转换率必定等于等利润线的斜率 $-p_F/p_C$。

[一] 分权式的资源配置方式是和中央集权式（计划经济）资源配置方式相对的。译者注。

假设公司发现均衡时最优劳动使用量分别为 L_C^* 和 L_F^*。此处我们关注的问题是，利润最大化是如何决定产品产量的。令 $L^* = w_C L_C^* + w_F L_F^*$ 表示生产所用的劳动的成本，并且将公司的利润 π 写为

$$\pi = p_C C + p_F F - L^* .$$

将上式变形，可得

$$C = \frac{\pi + L^*}{p_C} - \frac{p_F}{p_C} F .$$

这个式子表示的是公司的**等利润线**，如图 32.10 所示。等利润线的斜率为 $-p_F / p_C$，纵截距为 $(\pi + L^*)/p_C$。由于前面已假设 L^* 是既定的，纵截距越大的等利润线代表的利润越大。

如果公司试图实现利润最大化，它将在生产可能集中选择一点，使得通过这一点的等利润线的纵截距最大。到这一步，你应该清楚这意味着等利润线必定和生产可能边界相切；也就是说，在最优点处，生产可能边界的斜率（边际转换率），应该等于等利润线的斜率 $-p_F / p_C$：

$$MRT = -\frac{p_F}{p_C} .$$

在上面，我们使用一个企业说明利润最大化的问题，但所得到的结论适用于任何数量的企业：每个企业利润最大化时必然满足下列条件，即两商品之间的边际转换率等于这两种商品的价格比率。只要企业面对的两种商品的价格是相同的，即使它们的生产可能集差异很大，上述结论也依然成立。

这意味着均衡时两种商品的价格之比，可以衡量边际转换率——一种商品用另外一种商品表示的机会成本。如果你想多要一单位椰子，你必须放弃一些鱼。应该放弃多少鱼？只要看看椰子和鱼的价格比率[㈠]：这些经济变量的比率告诉我们技术上的权衡选择应该是什么样的。

32.13 鲁宾逊和星期五作为消费者

我们已经知道漂流者公司如何确定它的利润最大化生产方案。为了实施该方案，它必须雇佣一些劳动和产生一些利润。当它雇佣劳动时，它必须为劳动支付工资；当它赚得利润时，它对股东支付股息。这样漂流者公司赚得的利润都归还了鲁宾逊和星期五，无论是以工

[㈠] 用式子表示：$MRT_{12} = \frac{\Delta x_2}{\Delta x_1} = \frac{p_1}{p_2} \Rightarrow \Delta x_2 = \frac{p_1}{p_2} \Delta x_1$，其中商品 1 为椰子，商品 2 为鱼。当 Δx_1 =1，即为了多得到一单位椰子，需要放弃的鱼量为 $\Delta x_2 = \frac{p_1}{p_2}$。注意此处我们考虑的是边际转换率的绝对值。译者注。

资的形式还是以利润的形式。

由于该公司把所有收入都支付给了它的员工和股东，这意味着他们有足够的钱来购买公司的产品。这只是第 31 章介绍的瓦尔拉斯法则的一种变形：人们出售他们的禀赋获得收入，因此他们必然有足够的钱来购买这些禀赋。此处，人们获得收入的渠道有两种：一是出售自己的禀赋；二是从公司得到利润。由于在经济体系中，钱不会消失也不会增加，因此人们恰好有足够的钱来购买产品。

人们在公司拿到钱以后，将用钱做什么？和往常一样，他们用钱购买消费品。在两商品的价格为 p_F 和 p_C 时，每个人选择能买得起的最优商品束。我们已经知道，最优消费束需要满足以下条件，即两商品的边际替代率等于它们的价格之比。但是这个价格之比也等于边际转化率，这是由于企业实现了利润最大化。因此，帕累托有效率的必要条件得到满足：每个消费者的边际替代率等于边际转换率。

在这样的经济中，商品的价格是相对稀缺性的信号。它们表示的是技术的稀缺性——为了多生产一种商品，必须放弃的另外一种商品的数量；它们也表示消费的稀缺性——为了多消费一种商品，人们愿意放弃的另外一种商品的数量。

32.14 分权化的资源配置

克鲁索-星期五经济是一个非常简化的情形。在对经济建立更大的模型时，我们需要使用更复杂的数学。然而，尽管上述模型非常简单，它依然包含着一些有用的思想。

其中最重要的思想是，个人效用最大化的**私人**目标和资源有效使用的**社会**目标之间的关系。在某些条件下，每个人追求他自身的私人目标，最终实现了整体的帕累托有效率配置。而且，任何帕累托有效率的配置都能通过竞争市场实现，前提是初始禀赋（包括企业的产权）都能合理地再次分配。

竞争市场的最大优点是，每个消费者和每个企业只需要关注自身的最大化问题。联系企业和消费者的唯一中介就是商品的价格。由于价格是相对稀缺性的信号，在价格一定时，消费者和企业有足够的信息作出决策，这些决策实现了资源有效率的配置。在这种情形下，资源有效使用之类的社会问题，可以由分权化的个人解决。

每一个消费者都能解决他消费什么的问题。企业根据消费者消费商品的价格决定每种产品的产量。在作出生产决策时，企业受利润信号的指引。在这个意义上，利润是正确的向导。说一个生产方案是有利可图的，等于说人们的支付意愿高于商品的价格，因此企业自然应该扩大这样产品的产量。如果所有企业都是追求利润最大化的，而且所有消费者都是追求效用最大化的，那么竞争均衡必定是帕累托有效率的配置。

附录

下面我们使用微积分推导出含有生产的经济的帕累托有效率条件。令 X^1 和 X^2 既表示商品 1 和 2 的产量，又表示商品 1 和 2 的消费量，正如教材中指出的：

$$X^1 = X_A^1 + A_B^1$$

$$X^2 = X_A^2 + A_B^2.$$

我们首先需要找到一种方便的方法，来描述生产可能性边界—— X^1 和 X^2 的所有技术可行的组合。对我们的目的来说，最方便的方法是使用**转换函数**（transformation function）。转换函数是这两种商品的各自总量的函数 $T(X^1, X^2)$，该函数要求当且仅当

$$T(X^1, X^2) = 0 \text{ 时，}$$

组合 (X^1, X^2) 才位于生产可能边界上。

在对技术进行描述后，我们就可以计算出边际转换率：为了多生产商品 1，必须放弃商品 2 的数量。尽管“转换”这样的字眼可能会让人认为一种商品变成另外一种商品，但是这种理解是不对的。实际上，这里的“转换”是指，资源从商品 2 的生产转移到用于商品 1 的生产。由于我们减少了生产商品 2 的资源、增加了生产商品 1 的资源，我们从生产可能边界上的一点移动到另外一点。边际转换率正好就是生产可能边界的斜率 dX^2 / dX^1。

假设产量发生了微小变动即 (dX^1, dX^2)，假设变动后仍在生产可能边界上。因此，我们有

$$\frac{\partial T(X^1, X^2)}{\partial X^1} dX^1 + \frac{\partial T(X^1, X^2)}{\partial X^2} dX^2 = 0.$$

由此式可以解出边际转换率：

$$\frac{dX^2}{dX^1} = -\frac{\partial T / \partial X^1}{\partial T / \partial X^2}.$$

稍后我们将用到这个式子。

一个帕累托有效率的配置是指在这个配置上，给定其他人的效用，任何人的效用都达到了最大。在只有两个人的情形下，我们可以写出这个最大化的问题：

$$\max_{x_A^1, x_A^2, x_B^1, x_B^2} \quad u_A(x_A^1, x_A^2)$$

$$\text{使得} u_B(x_B^1, x_B^2) = \overline{u}\text{，} \quad T(X^1, X^2) = 0.$$

构造拉格朗日函数

$$L = u_A(x_A^1, x_A^2) - \lambda(u_B(x_B^1, x_B^2) - \overline{u}) - \mu(T(X^1, X^2) - 0)\text{，}$$

一阶条件为

$$\frac{\partial L}{\partial x_A^1}=\frac{\partial u_A}{\partial x_A^1}-\mu\frac{\partial T}{\partial X^1}=0$$

$$\frac{\partial L}{\partial x_A^2}=\frac{\partial u_A}{\partial x_A^2}-\mu\frac{\partial T}{\partial X^2}=0$$

$$\frac{\partial L}{\partial x_B^1}=-\lambda\frac{\partial u_B}{\partial x_B^1}-\mu\frac{\partial T}{\partial X^1}=0$$

$$\frac{\partial L}{\partial x_B^2}=-\lambda\frac{\partial u_B}{\partial x_B^2}-\mu\frac{\partial T}{\partial X^2}=0 .$$

将前两个式子整理并用第一个式子除以第二个式子可得

$$\frac{\partial u_A/\partial x_A^1}{\partial u_A/\partial x_A^2}=\frac{\partial T/\partial X^1}{\partial T/\partial X^2} .$$

类似地，用第三式除以第四式可得

$$\frac{\partial u_B/\partial x_B^1}{\partial u_B/\partial x_B^2}=\frac{\partial T/\partial X^1}{\partial T/\partial X^2} .$$

上面两个式子的左边我们并不陌生，它们就是边际替代率。右边是边际转换率。因此，这两个式子要求，每个人的边际替代率必须等于边际转换率：每个人愿意用一种商品替换另外一种商品的比率，必须等于将一种商品在生产上转换为另外一种商品的比率。

这个结果背后的直觉非常简单。假设某人的边际替代率不等于边际转换率，即用一种商品替换另外一种商品的比率，不等于将一种商品在生产上转换为另外一种商品的比率，但这意味着，我们可以找到某种方法使得他的效用增加而又不影响其他人的效用。

总结

1.一般均衡框架可以通过下列方式进行扩展，即允许竞争性的、追求利润最大化的企业生产用于交换的产品。

2.在某些条件下，我们能为经济中的所有投入品和产出品找到一组价格，在这组价格下，企业的利润最大化行为和消费者的效用最大化行为，能使得所有市场的需求等于供给，也就是说，竞争均衡是存在的。

3.在某些条件下，竞争均衡是帕累托有效率的：福利经济学第一定理在包含生产的经济

也成立。

4.如果假设生产集是凸的，那么福利经济学第二定理对于包含生产的经济也成立。

5.如果产品的生产是有效率的，那么两种商品之间的边际转换率，衡量经济为了得到额外一单位某种商品，必须放弃另外一种商品的数量。

6.帕累托效率要求每个消费者的边际替代率等于边际转换率。

7.竞争市场的优点是，它能将资源有效率配置这样的社会问题，交由分权化的企业和消费者进行决策。

复习题

1.椰子的竞争价格为每单位 6 元，鱼的竞争价格为每单位 3 元。如果社会想放弃一单位椰子，那么它能生产出多少单位的鱼？

2.如果图 32.2 中的企业决定支付较高的工资，结果将是怎么样的？

3.对于一个给定的经济来说，为什么说竞争均衡是好事？在什么样的情形下，竞争均衡是坏事？

4.如果鲁宾逊的椰子和鱼之间的边际替代率为-2，这两种商品的边际转换率为-1，如果他想增加效用，他应该怎么做？

5.假设鲁宾逊和星期五每人每天都需要 60 单位的鱼和 60 单位的椰子。使用教材给出的生产率，如果两个人不互相帮助，他们每天要劳动多长时间？假设他们决定合作，并用最有效的生产方式进行生产，他们每天需要劳动多长时间？劳动时间减少的经济学解释是什么？

复习题答案

1.椰子的竞争价格为每单位 6 元，鱼的竞争价格为每单位 3 元。如果社会想放弃一单位椰子，那么它能生产出多少单位的鱼？

【复习内容】边际转换率

A.边际转换率（Marginal Rate of Transformation, MRT），衡量一种商品“转换”为另外一种商品的比率。当然，这不是说一种商品变成另外一种商品，而是说，人们改变要素的使用量从而少生产一种商品且多生产另外一种商品。

B.均衡条件：每个企业利润最大化时必然满足下列条件，即两商品之间的边际转换率等于这两种商品的价格比率。

【参考答案】社会放弃 1 单位椰子可以生产 2 单位鱼。

令商品 1 表示椰子，商品 2 表示鱼，则：

$$MRT_{12}=\frac{\Delta x_2}{\Delta x_1}=-\frac{p_1}{p_2}\Rightarrow \Delta x_2=-\frac{p_1}{p_2}\Delta x_1\Rightarrow \Delta x_2=-\frac{6}{3}(-1)=2$$

2.如果图 32.2 中的企业决定支付较高的工资，结果将是怎么样的？

【复习内容】使用等利润线和生产函数求解利润最大化问题；消费者的劳动供给决策

A 使用等利润线和生产函数求解利润最大化问题

A1 等利润线的表达式

给定椰子的价格 1 和劳动工资率 w，首先考虑能产生既定利润水平 π 的椰子和劳动的所有组合。它的表达式为

$$\pi=C-wL\,.$$

其中，C 表示椰子数量，L 表示劳动数量。

解出 C，可得

$$C=\pi+wL\,.$$

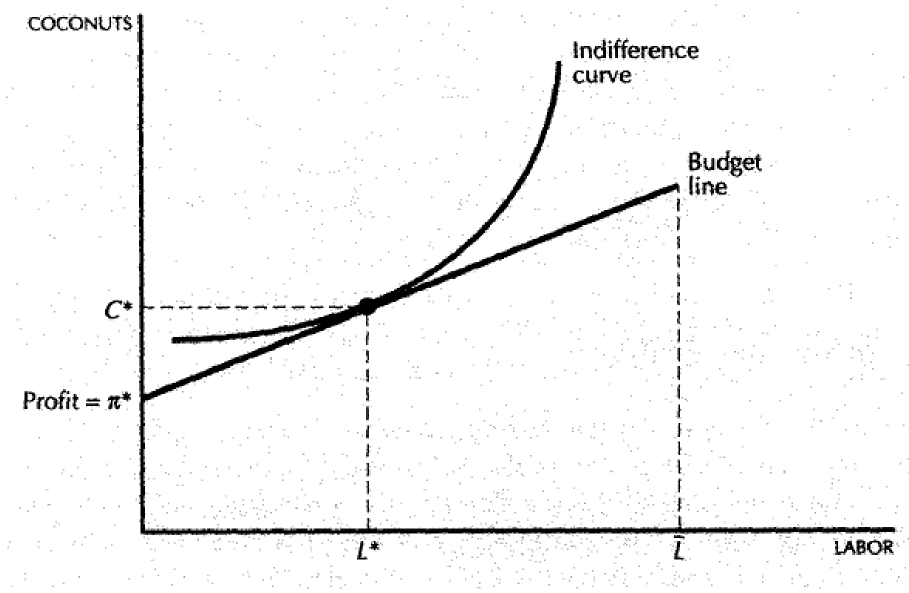
这个式子就是等利润线——能够产生既定利润 π 的劳动和椰子的所有组合。注意，在这种情形下，等利润线的斜率就是工资率 w。

A2 利润最大化的解

企业选择在利润最大化的点进行生产。和往常一样，这意味着一个相切条件：生产函数的斜率（即劳动的边际产量）必定等于等利润线的斜率 w。用式子表示即：$MP_L=w$。

B 消费者的劳动供给决策

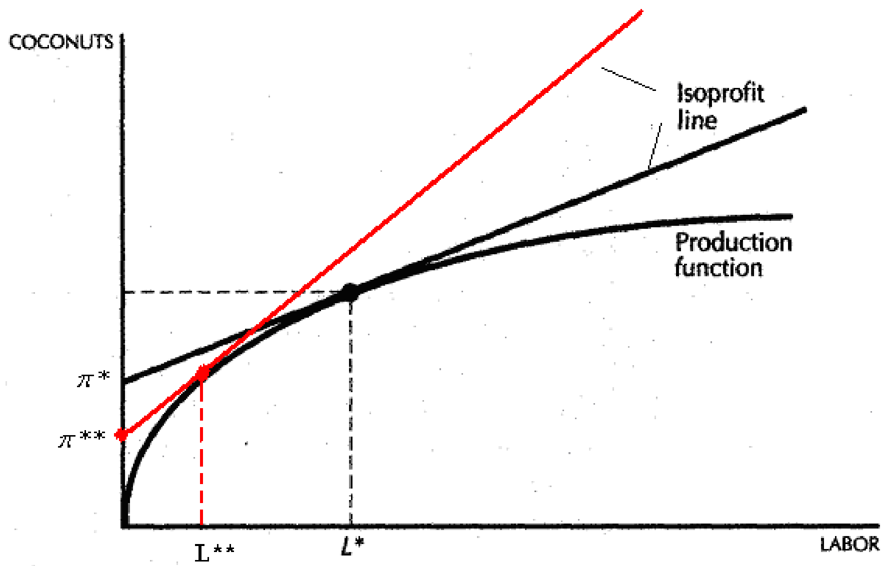
注意，在克鲁索一个人的经济中，给定椰子的价格和工资率，上述等利润线也是他的预算线。请复习图 32.3 的内容。这里只给出最简要的内容。



图：鲁宾逊的效用最大化问题。在椰子价格和工资率给定时，鲁宾逊作为消费者，他要决定最优工作时间和椰子的最优消费量。最优点位于无差异曲线和预算线的切点之处。

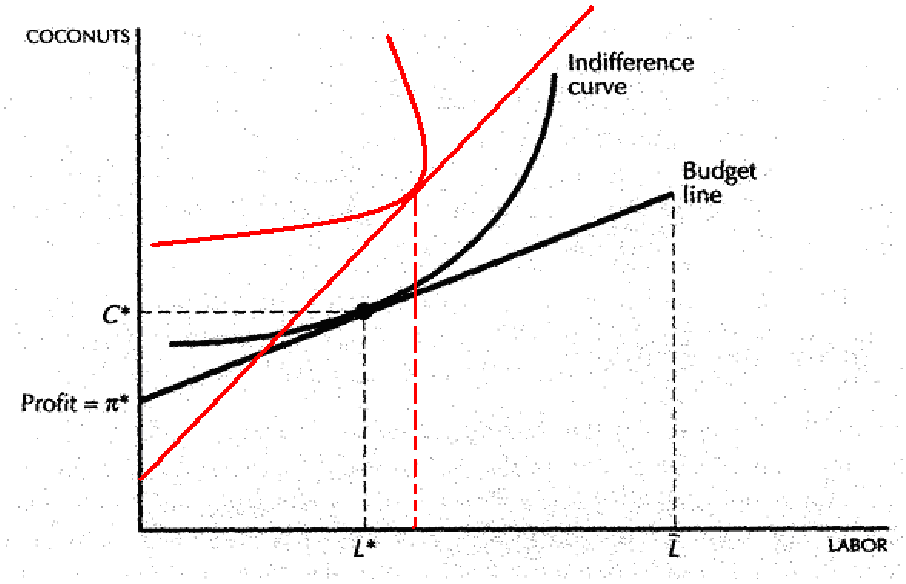
【参考答案】

如果企业提高工资率，即 w 变大，这意味着等利润线的斜率变大，换句话说，等利润线变得更陡峭。由于生产集是凸集（即生产函数曲线具有以下的形状），因此新的等利润线（红线）和生产函数曲线的切点，必然位于原利润线和生产函数曲线切点的左侧，如下图所示。因此企业提高工资率之后，企业的新利润 π^{**} 小于原利润 π^{*}，企业新的劳动需求 L^{**} 小于原劳动需求 L^{*}，如下图所示。



由于在克鲁索一个人的经济中，等利润线又是克鲁索的预算线，现在新的预算线变得更陡峭，这意味着新预算线将和更靠近左上方的无差异曲线相切，如下图所示。在这种情形下，克鲁索愿意供给更多的劳动量，大于原劳动供给量 L^{*}，从而也大于企业的新的劳动需求量 L^{**}，因此劳动的供给大于需求，劳动市场出现了不均衡。

克鲁索在新预算线的约束下愿意供给较多劳动的原因是，由于工资率的提高意味着闲暇的价格更高，因此他会用椰子替代一部分闲暇，从而闲暇的需求量减少，也就是劳动量增加。另外，由于在新的预算线下，他的收入减少，由于闲暇是正常商品，这也意味着闲暇的需求会减少，因此劳动的供给量也会增加。粗略地说，上述两种效应分别对应着替代效应和收入效应。



3.对于一个给定的经济来说，为什么说竞争均衡是好事？在什么样的情形下，竞争均衡是坏事？

【复习内容】福利经济学第一定理；财富分配的公平性

【参考答案】

福利经济学第一定理，粗略地说，是说竞争均衡是帕累托有效率的。如果仅站在效率的角度上来说，这当然是好事，因为这意味着经济中已不存在使任何人的状况改善而又不损害其他人利益的情形。

但是竞争通常意味着不均等，比如“一部分人先富起来”可以认为是竞争均衡的结果，但这样的结果从均等的意义上来说是不“公平”的，如果站在这个角度，竞争均衡就不是好事。因此社会通常又损害一部分的状况而让另一部分人的状况得到改善。

4.如果鲁宾逊的椰子和鱼之间的边际替代率为－2，这两种商品的边际转换率为－1，如果他想增加效用，他应该怎么做？

【复习内容】生产和消费同时均衡

以两个人（A 和 B）和两种商品（商品 1 和 2）的经济为例，生产和消费同时均衡的必

要条件为

$$MRS_{12}^{A}=MRS_{12}^{B}=MRT_{12}=\frac{\Delta x_2}{\Delta x_1}=-\frac{p_1}{p_2}$$

【参考答案】他将生产更多的鱼。

令商品 1 表示椰子，商品 2 表示鱼，由题意知

$$MRS_{12}=\frac{\Delta x_2}{\Delta x_1}=-\frac{MU_1}{MU_2}=-2\Rightarrow\frac{MU_1}{MU_2}=\frac{2}{1}$$

$$MRT_{12}=\frac{\Delta x_2}{\Delta x_1}=-1\Rightarrow\left|\frac{\Delta x_2}{\Delta x_1}\right|=\frac{1}{1}$$

第二个式子可以解读为，**一单位商品 1 和一单位商品 2 耗费的资源数量是一样的**；但是第一式却意味着商品 1 的（边际）效用更高，因此应该多生产商品 1，即多生产鱼。

5.假设鲁宾逊和星期五每人每天都需要 60 单位的鱼和 60 单位的椰子。使用教材给出的生产率，如果两个人不互相帮助，他们每天要劳动多长时间？假设他们决定合作，并用最有效的生产方式进行生产，他们每天需要劳动多长时间？劳动时间减少的经济学解释是什么？

【复习内容】比较优势

【参考答案】如果两人各干各的，那么每个人都需要工作 9 小时/天；如果合作，那么每个人只需要工作 6 小时/天。时间的减少在于比较优势发挥了作用。

A 如果不合作

根据教材可知鲁宾逊一小时可生产 10 单位的鱼或 20 单位的椰子，由于鲁宾逊每天需要 60 单位鱼和 60 单位的椰子，因此他每天必须工作的时间为 $\frac{60}{10}+\frac{60}{20}=9$ 小时。

星期五一小时可生产 20 单位的鱼或 10 单位的椰子，由于星期五每天需要 60 单位鱼和 60 单位的椰子，因此他每天必须工作的时间为 $\frac{60}{20}+\frac{60}{10}=9$ 小时。

B 如果合作

合作后，他们将采用最有效率的的生产方式进行生产。这里的“最有效率”可从两个等价的角度进行解读：一是生产既定产量所花费的时间最小；二是既定时间生产的产量最大。我们从第二个角度入手分析。

因为鲁宾逊在生产椰子上具有比较优势、星期五在生产鱼上具有优势，因此最有效率的生产方式就是鲁宾逊专门生产椰子、星期五专门生产鱼。这个结论符合直觉。但是为了让你彻底看清这个结论，我们用代数方法进行证明。

如果社会只生产一小时，而且鲁宾逊专门生产椰子、星期五专门生产鱼，则社会的产品为 20 单位的椰子和 20 单位的鱼，我们需要证明其他生产方法得到的产量都小于这个产量。

曹乾（东南大学 caoqianseu@163.com）

不妨设鲁宾逊生产椰子的时间为 a，则其生产鱼的时间为 1-a；类似地，设鲁宾逊生产椰子的时间为 b，则其生产鱼的时间为 1-b。那么我们只要证明在 $a \neq 1$ 且 $b \neq 0$ 时下列方程无正数解即可：

$$20a + 10b \geq 20$$

$$10(1-a) + 20(1-b) \geq 20$$

这个证明是简单的。它的确无解，你可以验证一下。

既然最有效率的生产方式就是鲁宾逊专门生产椰子、星期五专门生产鱼，那么生产 120 单位椰子需要花费鲁宾逊的时间为 120/20=6 小时；生产 120 单位的鱼需要花费星期五的时间为 120/20=6 小时。

C 时间减少的原因

由 B 的分析过程可知，时间减少的原因正是在于利用了比较优势（分工合作）。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

33.福利（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

33 福利

直到现在我们关注的重点是，评价经济配置的帕累托效率。但是还有其他重要的议题。必须记住帕累托效率和福利在人们间的分配问题毫无关系；将一切都分配给一个人，这个结果是帕累托有效率的。但是我们其他人可能认为这不是一个合理的分配。在本章，我们研究一些技术，并用这些技术分析福利的分配问题。

帕累托有效率本身就是一个人们想要的目标——如果存在能让一伙人状况更好，而又不损害其他人的方法，为什么不去做呢？但是帕累托有效率的配置有很多；社会应该选择其中哪一个配置？

本章的关键是**福利函数**（welfare function）的思想，使用福利函数可将不同消费者的效用"加总"。更一般地，我们可以使用福利函数对效用在消费者群体的不同分配方式进行排序。在应用这个概念之前，我们先分析怎样对个体的偏好进行"加总"，以构建某类"社会偏好"。

33.1 偏好的加总

首先回顾一下消费者偏好的内容。和往常一样，我们假设这些偏好是可传递的。以前，我们将消费者的偏好定义在他自身的消费束上，但是现在我们想扩展这个概念，认为每个消费者的偏好是针对商品在消费者之间分配的问题而说的。当然，这个概念也包含着下列的可能性：消费者可能并不关心其他人的消费束状况，这又回到了我们原来的假设。

我们用 x 表示某个特定的配置——每个人得到的每种商品的数量。于是给定两个配置，x 和 y，每个消费者可以判断他是否更偏好 x 而不是 y。

给定所有人的偏好，我们希望能找到一种方法将它们全部"加总"，从而形成一个**社会偏好**（social preference）。也就是说，如果我们知道所有个体是如何对各种配置结果进行排序的，那么我们想使用这些信息构建这些配置的社会排序方法。这是最广泛意义上的社会决策问题。我们先分析几个例子。

加总个人偏好的一种方法是使用某种投票机制。如果社会上大多数人偏好 x 胜于 y，我们可以认为 x 被"社会偏好"于 y。然而，投票方法存在着一个比较严重的问题——它可能不能产生一个传递性的社会偏好排序。例如，考虑表 33.1 所示的情形。

此处，我们列出了三个人对三个选择 x，y 和 z 的排序结果。注意**多数人**的偏好排序：$x \succ y$，$y \succ z$ 和 $z \succ x$。因此，通过多数人投票(majority vote)的方法来加总个人偏好的方法不可行，因为通过这样方法得到的社会偏好不是良好性状的偏好，这又是因为这些社会偏好不具有传递性。由于社会偏好不是传递的，因此我们无法从选择集 (x, y, z) 中选择出"最佳

的”选择。这种情形下，社会最终选择哪个结果取决于投票顺序。

Person A	Person B	Person C
x	y	z
y	z	x
z	x	y

表 33.1：多数人投票通常不能产生具有传递性的社会偏好

为了看清这一点，假设表 33.1 中的三个人先对 x 和 y 投票，然后再对胜出者与 y 进行投票。由于多数人偏好 x 胜于 y，x 暂时胜出。第二轮投票在 x 和 z 之间进行，由于多数人偏好 z 胜于 x，z 最终胜出。

但是如果这三个人先对 z 和 x 投票，然后再对胜出者和 y 进行投票，结果将如何？现在 z 在第一轮投票中胜出，但在第二轮，z 击败 y，从而最终结果为 z。最终哪个选择胜出取决于人们对这些选择的投票顺序。

另外一种投票方法称为排序打分-按总分排位的投票法（rank-order voting）。在这种方法中，每个人先按照自己的偏好对选择进行排序，然后对这些选择相应进行打分（赋值）：例如，最佳选择的分数为 1，次佳选择的分数为 2，以此类推。然后我们将每个人每种选择的分数相应进行加总，得到每一种选择的总分数，在这种情形下，哪一种选择的分数最低，哪一种选择就是社会最偏好的。

在表 33.2 中，我们给出了两个人对三种选择 x、y 和 z 的偏好排序。假设现在只有 x 和 y 两种选择，那么 A 对 x 的赋值为 1，B 对 x 的赋值为 2；A 对 y 的赋值为 2，B 对 y 的赋值为 1。由于 x 和 y 的总分数都为 3，因此 x 和 y 不分胜负。

Person A	Person B
x	y
y	z
z	x

表 33.2：人们对 x 和 y 的选择取决于 z .

现在我们将 z 引入投票中。A 对这三种选择的打分分别如下：x 为 1，y 为 2，z 为 3；B 的打分为：x 为 3，y 为 1，z 为 2。这种情形下，x 的总分为 4，y 的总分为 3，因此按

曹乾（东南大学 caoqianseu@163.com）

照这种投票方法，社会对 x 的偏好超过了 y ，因为我们在前面假设分数越低越好。

多数人投票法和排序打分-按总分排位投票法存在的问题是，狡猾的投票负责人可以操纵最终的结果。在多数人投票法中，投票负责人可以改变选项的投票顺序，从而得到他所想要的结果；而在排序打分-按总分排位投票法中，投票负责人可以引入新的选项，从而改变其他相关选项的最终排位。

人们自然会问：是否存在社会决策机制，即是否存在能将个人偏好加总的方法，而且要使得这种方法不能受到人为操纵的影响？能否找到对不具备传递性的偏好进行“加总”的方法？

我们先列出我们期望社会决策机制应该能做到的事情：

1.给定具有完备性、反身性和传递性的个人偏好的任一集合，社会决策机制应该能产生具有相同性质的社会偏好。

2.如果每个人对选择 x 的偏好超过 y ，那么社会的偏好应该将 x 排在 y 的前面；

3.人们对 x 和 y 的偏好应该仅取决于人们如何对 x 和 y 进行排序，而不应受到其他选择的影响。

这些要求似乎非常合理。然而却很难找到同时满足上述条件的社会决策机制。事实上，肯尼斯.阿罗已经证明了下面著名的结论[一]：

阿罗不可能定理（Arrow’s Impossibility Theorem）。**若一个社会决策机制满足上述性质 1,2 和 3，则它必定是独裁的：所有社会偏好都是独裁者一个人的偏好。**

阿罗的不可能定理让人们非常惊讶。它表明上述三条性质与民主是不相容的，尽管这三条性质非常合理而且令人喜欢。也就是说，不存在制定社会决策的“完美”方法。既然不存在将个人偏好“加总”为社会偏好的完美方法，如果我们又想找到加总的方法从而得到社会偏好，我们必须放弃阿罗定理中社会决策机制上述三个性质中的一个。

33.2 社会福利函数

如果我们想放弃社会福利函数上述三个性质中的一个，应该放弃哪一个？最有可能放弃性质 3。如果放弃了性质 3，前面介绍的排序打分-按总分排位投票法就可行了。

给定每个人 i 对各种配置方法的偏好，我们可以构建效用函数 $u_i(x)$ ，这些函数总结了个人的价值判断：个人 i 对 x 的偏好超过 y 当且仅当 $u_i(x) > u_i(y)$ 。当然，这些效用函数和其他效用函数一样——正单调变换后仍能保留偏好的原来排序（保序性）。因此，能代表个

[一] See Kenneth Arrow, Social Choice and Individual Values (New York: Wiley, 1963)。阿罗，斯坦福大学教授，因在该领域的工作而获得诺贝尔经济学奖。

人 i 的效用函数**不是唯一**的。

但是，出于方便起见，在我们找到这样的效用函数后，我们就一直用它。于是，从个人偏好得到社会偏好的一种方法是，将这些个人偏好加总，用加总得到的**和**表示社会偏好。也就是说，我们说 x 这个配置好于配置 y，若

$$\sum_{i=1}^{n} u_i(x) > \sum_{i=1}^{n} u_i(y),$$

其中 n 为社会中的人数。

这种方法是可行的，但是需要注意它也是非常主观的，因为我们对效用函数的选择是非常主观的。选择使用加和表示社会偏好也是主观的。当然，你完全可以使用下列方法中的任何一个来表示社会偏好：个人效用的加权平均数；个人效用的乘积；个人效用的平方和；等等。

我们需要对该“加总函数”作出一个假设，即它随每个人的效用增加而增加，这个假设比较合理。作出该假设的目的，是想保证若每个人偏好 x 胜过 y，则社会偏好必然也是偏好 x 胜过 y。

这类加总函数有个名字，叫做**社会福利函数**（social welfare function）。社会福利函数是个人效用函数的函数：$W(u_1(x),...,u_n(x))$。使用社会福利函数可以对不同配置进行排序。注意，社会福利函数仅取决于个人的偏好，而且它是每个人效用的增函数。

下面我们分析几个例子。一种特别的情形，就是前面说到的，社会效用函数是个人效用函数的加和：

$$W(u_1,...,u_n) = \sum_{i=1}^{n} u_i .$$

这个函数有时称为**古典效用主义福利函数**（classical utilitarian welfare function）或**边沁主义福利函数**（Benthamite welfare function）[一]。该函数的一般形式是**效用加权和福利函数**（weighted-sum-of-utilities welfare function）：

$$W(u_1,...,u_n) = \sum_{i=1}^{n} a_i u_i ,$$

其中，权重 a_i 表示第 i 个人在社会福利中的重要性。自然可假设 a_i 为正数。

另外一个有趣的效用函数为**极小极大**（即极大中的极小）**社会福利函数**或**罗尔斯社会福利函数**（minimax or Rawlsian social welfare function）：

$$W(u_1,...,u_n) = \min\{u_1,...,u_n\} .$$

[一] 杰里米.边沁（Jeremy Bentham，1748-1832），是道德哲学的效用主义学派的创始人，该学派认为最大的善是使得尽可能多的人实现最大幸福。

这个效用函数是说，某个配置的社会福利仅取决于状况最差的那个人的福利，即取决于效用最小的人的福利[一]。

上面介绍的每个福利函数都是可行的，都可以用来比较不同配置的评价。每个福利函数背后的伦理标准是不一样的。我们对福利函数结构的限制，暂时只要求它是每个人效用的增函数。

33.3 福利最大化

在找到福利函数之后，我们就可以分析福利最大化问题。令 x_i^j 表示第 i 个消费者拥有的商品 j 的数量，假设有 n 个消费者和 k 种商品。于是，配置 x 表示每个消费者拥有的每种商品的数量。

如果这 k 种商品的数量分别为 $X^1,...,X^k$，现在要将它们分配给上述 n 个消费者，则福利最大化问题为：

$$\max \ W(u_1(x),...,u_n(x))$$

$$\text{使得} \sum_{i=1}^{n} \max \ x_i^1 = X^1$$

$$\vdots$$

$$\sum_{i=1}^{n} \max \ x_i^k = X^k$$

因此，我们试图找到能使社会福利最大的可行配置。这样的配置应该具有什么样的性质？

我们首先指出，**使福利最大化的配置必然是帕累托有效率的配置**。证明很简单：假设使福利最大化的配置不是帕累托有效率的。既然不是有效率的，则意味着存在着帕累托改进的余地，因此总可以找到另一个配置，使得部分人的效用和前面的那个配置一样大，但其他人的效用更大。然而福利函数又是每个人效用的增函数，因此这个新配置的福利必然更大，但这和开始时假设原来的那个配置已实现福利最大化相矛盾。这样就证明了我们的断言：使福利最大化的配置必然是帕累托有效率的配置。

我们用图 33.1 说明这种情形，其中集合 U 表示两人情形下的可能效用集。这个集合称为效用可能集（utility possibilities set）。该集合的边界称为效用可能边界（utility possibilities frontier），效用可能边界是一个集合，它由与帕累托有效率的配置相伴的效用值

[一] 约翰.罗尔斯（John Rawls）是当代道德哲学家，哈佛大学教授，他认为这样的正义原则是合理的，即认为上述福利函数是合理的。

组成。如果某个配置在效用可能边界上，那么我们就无法找到能使这两个人的效用都更高的其他可行配置。

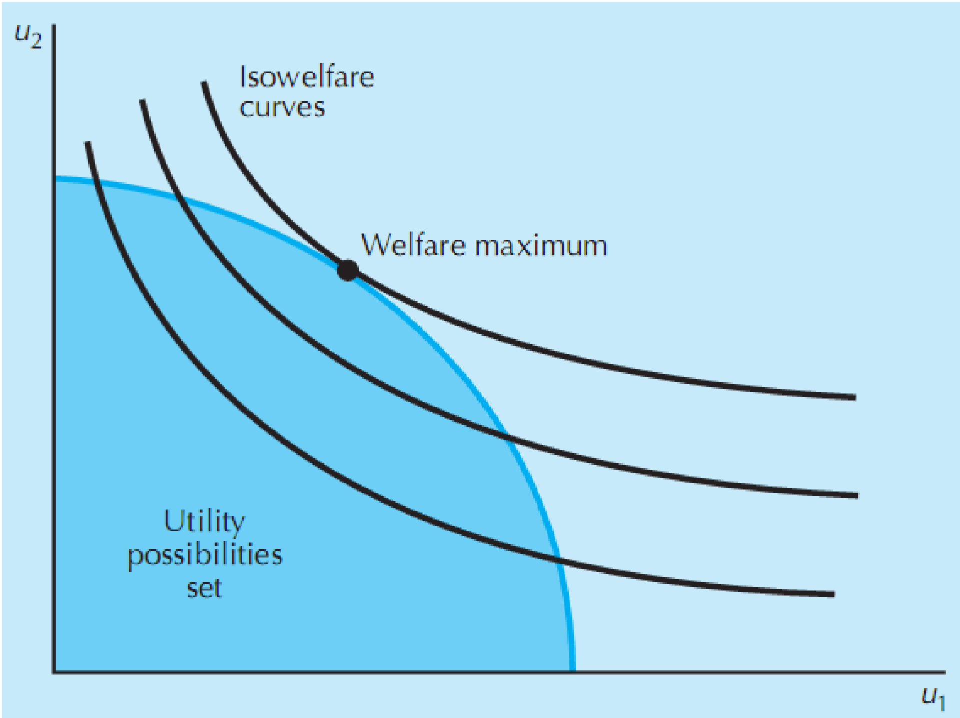


图 33.1：福利最大化。使福利函数达到极大值的配置必然是帕累托有效率的。

这个图中的“无差异曲线”称为**等福利曲线**（isowelfare curves），因为在同一条曲线上的这两个人的所有效用组合(u_1,u_2)，它们代表的福利值都是相等的。和往常一样，最优点可用相切条件进行刻画。但是对于我们的目的来说，我们关注的是，这个最大福利点是帕累托有效率的——它必然位于效用可能边界上。

根据这个图，我们还可以断言：**任一帕累托有效率的配置必然是某个福利函数的极大值点**。图 33.2 就是这样的例子。

在图 33.2 中我们我们选择了一个帕累托有效率的配置，并且找到了一组等福利线。由图可知，对于这组等福利线来说，该配置点的福利是最大的，因为尽管最外面一条等福利线代表的福利更大，但受效用可能集所限，无法达到该位置。事实上，我们可以进一步说，如果效用可能集为凸集（如图 33.2 所示），那么效用可能边界上的任何一点都是效用加权和福利函数的极大值点，如图 33.2 所示。因此，我们可以使用福利函数筛选出帕累托有效率的那些配置点：每个福利极大值点都是帕累托有效率的配置点，而且每个帕累托有效率的配置点都是福利极大值点。

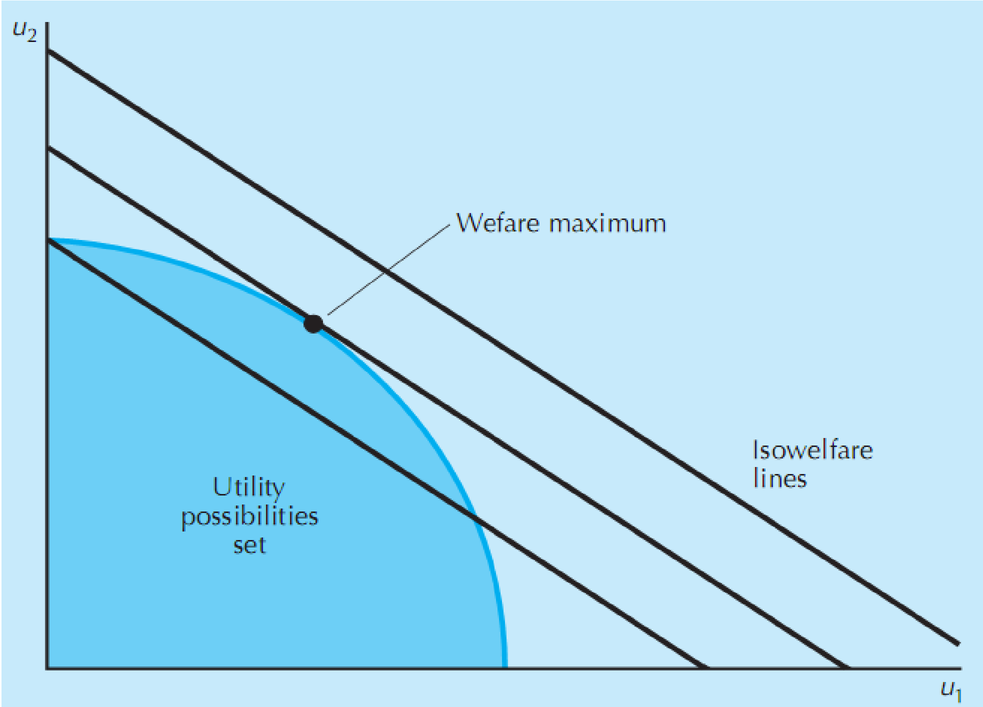


图 33.2：效用加权和福利函数的极大值。如果效用可能集为凸，则任何一个帕累托有效率的点都是效用加权和福利函数的极大值点。

33.4 个人主义的社会福利函数

直到现在，我们都将个人偏好看成个人对整个配置的偏好（即他关注所有人拥有的商品），而不是对他自己消费束的偏好。但是，我们在前面也说过，个人可能只关心他自己的消费束。在这种情形下，我们可用 x_i 表示第 i 个人的消费束，令 $u_i(x_i)$ 表示第 i 个人的效用水平，当然前提是我们已经找到了效用函数的表示方法。于是，社会福利函数的形式为

$$W = W(u_1(x_1),\ldots,u_n(x_n)).$$

这个社会福利函数是个人效用函数的直接函数，但它是个人消费束的间接函数。这种特别形式的福利函数称为**个人主义的福利函数**（individualistic welfare function）或**柏格森-萨缪尔森福利函数**（Bergson-Samuelson welfare function）[一]。

如果每个人的效用仅取决于他自己的消费，那么就不存在消费的外部性（consumption externalities）。因此，我们就可以直接运用第 31 章得到的结果，即帕累托有效率配置和市场均衡之间的关系：所有竞争均衡都是帕累托有效率的，而且，在凸性的假设下，所有帕累托有效率的配置也都是竞争均衡。

现在，我们可以将上述分类推进一步。给定上述帕累托效率和福利极大之间的关系，我们可以断言：所有的福利极大值点都是竞争均衡，而且所有的竞争均衡对于某个福利函数

[一] 艾布拉姆.柏格森（Abram Bergson）和保罗.萨缪尔森（Paul Samuelson）都是当代经济学家，他们在 1940 年代研究了这类福利函数的性质。萨缪尔森因为对经济学作出了很多贡献而获得了诺贝尔经济学奖。

来说都能是极大值点。

33.5 公平配置

在分析社会福利时，最常规的方法就是使用福利函数。正因为福利函数的这种常规性，因此可用它总结很多类型的道德判断（moral or ethical judgments）问题。但是，另一方面，福利函数一般不能用于决定哪种道德判断是合理的。

分析社会福利的另外一种方法是，首先界定所使用的具体道德判断标准，然后将其用于评价经济配置。**公平配置**（fair allocations）问题就是使用这种方法进行分析的。下面我们首先介绍什么样的商品束分配方法可能是公平的方法，然后再用经济分析方法研究这种公平配置方法的应用。

假设某几种商品的数量都是既定的，有 n 个人有权得到这些商品，而且这些人的索取权是相等的。如何公平地进行分配？在这种情形下，可能大多数人会选择将这些商品平均分配给这些人。由于我们假设这些人的索取权是相等的，你难道还能相处其他的分配方法？

平均分配思想有什么吸引人的地方？一是它具有**对称性**，每个人得到的每种商品数量是相等的；每个人都不偏爱其他人得到的商品束，因为他们得到的商品束是完全一样的。

不幸的是，平均分配**不一定**是帕累托有效率的。如果人们的偏好不同，在平均分配之后，他们通常会进行交易。假设这种交易发生，使得我们向一个帕累托有效率的配置移动。

由此产生的问题是：这个帕累托有效率的配置仍然是公平的吗？平等分配之后进行的交易过程是否仍然具有对称性？

答案为：不一定。例如，有三个人：A，B 和 C。A 和 B 的偏好相同，C 的偏好不同。我们从平均分配结果开始分析，假设 A 和 C 进行交易。那么，A 和 C 的状况通常会改善。但是 B 却没机会与 C 进行交易，因此 B 会**妒忌**（envy）A——也就是说，他偏好 A 的商品束胜过他自己的。尽管由于平均分配，A 和 B 得到的商品束是相同的，但是 A 却有幸进行了交易，这种行为破坏了原来分配结果的对称性。

这表明在平均分配后，如果人们进行交易，那么交易过程不一定保留了原来的对称性。我们因此会问：是否存在能保留对称性的配置？是否能找到某种方法，使得我们得到的配置既是帕累托有效率的又是公平的？

33.6 妒忌和平等

我们先分析一些概念。我们所说的“对称的”和“公平的”到底是什么意思？下面给出了一组可行的定义。

曹乾（东南大学 caoqianseu@163.com）

一个配置是**平等的**（equitable），若无人偏好其他人的商品束胜过自己的商品束。若第 i 人更喜欢第 j 人的商品束而不是他自己的，则 i **妒忌** j。若一个配置既是平等的又是帕累托有效率的，则它是一个**公平的**（fair）配置。

我们可以使用这些定义阐述前文对称性这个概念的思想。平均分配的配置的特点是，每个人都不妒忌其他人。但是其他一些配置也具有这种性质。

考虑图 33.3。为了判断任何一个配置（比如 F 点）是否为平等的，我们只要看看如果这两个人进行交易得到的配置在哪里。如果交换后的配置，位于通过原配置（F 点）的每个人的无差异曲线的“下方”，那么原配置（F 点）就是一个平等配置。（此处，“下方”表示每个人自己认为的下方；但是从我们的角度来看，交换后的配置点 S 位于两条无差异曲线之间。）[一]

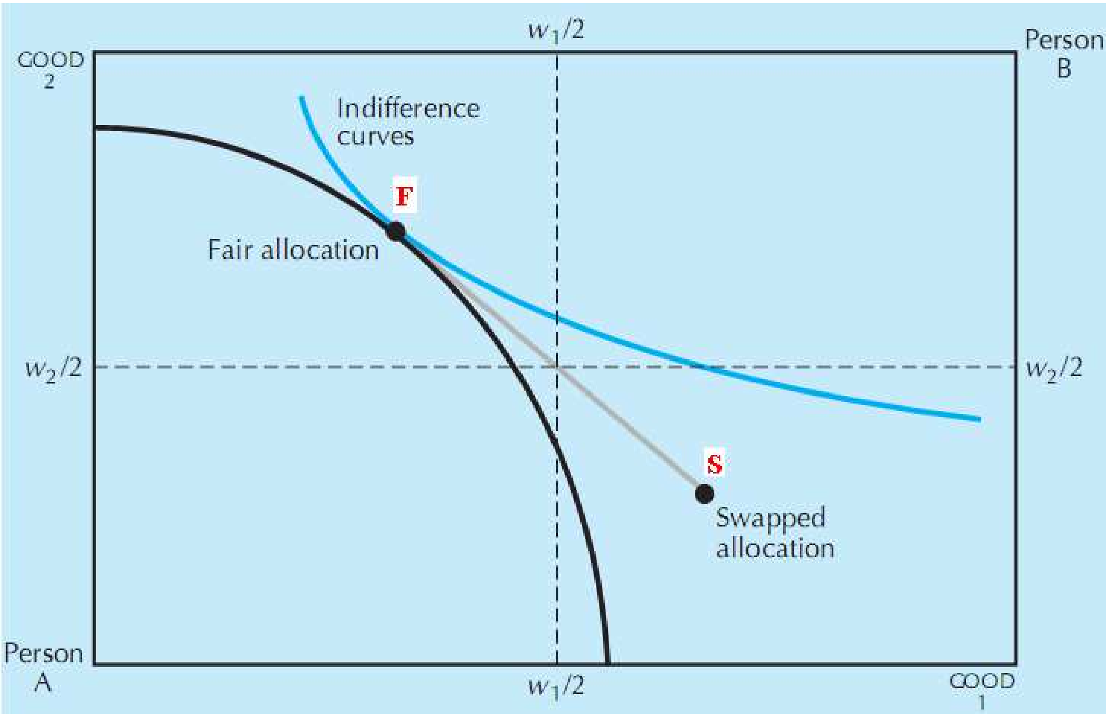


图 33.3：公平配置。埃奇沃思盒内的公平配置。每个人对公平配置（点 F）的偏好都胜过对交换后的配置（点 S）的偏好。

同时还要注意图 33.3 中的配置（F 点）也是帕累托有效率的配置。因此，它不仅是平等的（按照本节给出的定义），而且是有效率的。因此，根据前面的定义可知，它是一个公平的配置。这样的配置是偶然得到的吗？换一种问法，一般来说，公平配置是否存在？

可以证明，公平配置一般来说是存在的，证明此事很简单。和上一节一样，我们从平均分配配置点进行分析，让消费者进行交易从而实现一个帕累托有效率的配置。假设消费者

[一] 这个结论从相反的角度论证，可以看得更清楚。假设消费者从 S 点开始交易，最终达到 F 点。消费者交易的原因就是因为互相妒忌，即更喜欢对方的商品束。到达 F 点之后，双方不再彼此妒忌。因此 F 点是一个平等配置点。译者注。

在竞争市场中进行交易。交易后得到了一个新的配置，在这个配置中，每个消费者都按照均衡价格 (p_1, p_2) 购买他能买得起的最优商品束。根据第 31 章的知识可知，这样的配置必然是帕累托有效率的。

但是，它是否仍然是平等的？假设不是。因此其中一个消费者，比如 A 会妒忌 B。这表示 A 更喜欢 B 的商品束而不是自己的。即

$$(x_A^1, x_A^2) \prec_A (x_B^1, x_B^2).$$

但是，若 A 更喜欢 B 的商品束，而我们又知道他自己的消费束，是在价格 (p_1, p_2) 下他能买得起的最优商品束，这意味着 A 买不起 B 的消费束，即

$$p_1\omega_A^1 + p_2\omega_A^2 < p_1x_B^1 + p_2x_B^2.$$

这样我们就得到了矛盾！根据假设，A 和 B 开始交易之前所拥有的商品束是一样的，因为我们是从平均分配配置开始分析的。如果 A 买不起 B 的商品束，那么 B 买不起自己的商品束，显然这是不可能的。

因此，我们可以断言，在这种情形下 A 不可能妒忌 B。从平均分配配置开始交易而得到的竞争均衡，必然是公平配置。因此，市场机制能保留这样的平等性：若初始配置是平均分配配置，那么最终配置必定是公平的。

附录

此处我们使用个人主义的福利函数求解社会福利最大化的问题。使用第 32 章介绍过的转换函数来描述生产可能边界，我们可将福利最大化问题写成

$$\max_{x_A^1, x_A^2, x_B^1, x_B^2} \quad W(u_A(x_A^1, x_A^2), u_B(x_B^1, x_B^2))$$

$$\text{使得}\, T(X^1, X^2) = 0$$

其中，X^1, X^2 分别表示商品 1 和 2 的产量和消费量，并且假定产量和消费量相等。

该最大化问题的拉格朗日函数为

$$L = W(u_A(x_A^1, x_A^2), u_B(x_B^1, x_B^2)) - \lambda[T(X^1, X^2) - 0] = 0$$

对每个选择变量分别求导可得一阶条件：

$$\frac{\partial L}{\partial X_A^1} = \frac{\partial W}{\partial u_A}\frac{\partial u_A(x_A^1, x_A^2)}{\partial x_A^1} - \lambda\frac{\partial T(X^1, X^2)}{\partial X^1} = 0$$

$$\frac{\partial L}{\partial X_A^2} = \frac{\partial W}{\partial u_A}\frac{\partial u_A(x_A^1, x_A^2)}{\partial x_A^2} - \lambda\frac{\partial T(X^1, X^2)}{\partial X^2} = 0$$

$$\frac{\partial L}{\partial X_B^1}=\frac{\partial W}{\partial u_B}\frac{\partial u_B(x_B^1,x_B^2)}{\partial x_B^1}-\lambda\frac{\partial T(X^1,X^2)}{\partial X^1}=0$$

$$\frac{\partial L}{\partial X_B^2}=\frac{\partial W}{\partial u_B}\frac{\partial u_B(x_B^1,x_B^2)}{\partial x_B^2}-\lambda\frac{\partial T(X^1,X^2)}{\partial X^2}=0.$$

变形整理，并用第一式除以第二式，第三式除以第四式可得

$$\frac{\partial u_A/\partial x_A^1}{\partial u_A/\partial x_A^2}=\frac{\partial T/\partial X^1}{\partial T/\partial X^2}$$

$$\frac{\partial u_B/\partial x_B^1}{\partial u_B/\partial x_B^2}=\frac{\partial T/\partial X^1}{\partial T/\partial X^2}$$

注意，这两个式子正是我们在第 32 章附录中推导出来的式子。因此，福利最大化问题的一阶条件和帕累托有效率问题的一阶条件是相同的。

这显然不是偶然的。根据教材中的讨论可知，柏格森-萨缪尔森福利函数（即个人主义福利函数）极大值对应的配置点，是帕累托有效率的，而且每个帕累托有效率的配置都能使得某个福利函数达到极大值。因此，福利最大化和帕累托有效率配置必然有相同的一阶条件。

总结

1.阿罗不可能定理表明，不存在将个人偏好加总为社会偏好的理想方法。

2.尽管如此，经济学家仍然使用这种或那种福利函数分析配置的福利。

3.若福利函数是每个人效用的增函数，则福利极大值点必定是帕累托有效率的。而且，每个帕累托有效率的配置，都可以被认为是某个福利函数的极大值点。

4.对配置进行评价的另外一种方法，是使用公平分配的思想。这个思想强调配置的对称性。

5.即使初始配置是对称的，也不是每一种交易方式都能导致公平配置。然而，可以证明，市场机制能实现公平配置。

复习题

1.假设我们说，**只有**社会中的**每个人**都偏好配置 x 胜过配置 y 时，社会的偏好才是偏好 x 胜过 y。（这有时称为帕累托排序方法，因为它和帕累托效率的思想密切相关）。如果将这种说法作为制定社会决策的准则，有什么缺陷？

2.罗尔斯福利函数的值仅取决于福利最差的那个人的效用。与该福利函数相对的一种函数叫作“尼采”（Nietzschean）福利函数——其值仅取决于福利最好的那个人的效用。尼采福利函数的表达形式是什么样的？

3.假设效用可能集是一个凸集，而且假设每个消费者仅关注他自己的消费。如果我们采用尼采福利函数进行分析，那么什么样的配置能使得该函数达到极大值？

4.假设某个配置是帕累托有效率的，而且每个人仅关注他自己的消费。证明，在这种情形下，必然存在这样的人：他不妒忌其他任何人。请使用我们课文中对妒忌的定义。（这个题目需要思考一会，但这样做是值得的。）

5.制定投票日程的能力，是一种非常重要的能力。假设社会偏好将通过多数人投票方法决定，并且将备选项成对进行投票角逐。假设人们的偏好如教材图 30.1 所示，请制定投票日程从而让 y 最终胜出。再制定另外一个日程让 z 最终胜出。如果你具备这种制定投票日程的能力，那么你是利用了社会偏好的什么性质？

复习题答案

1.假设我们说，只有社会中的每个人都偏好配置 x 胜过配置 y 时，社会的偏好才是偏好 x 胜过 y。（这有时称为帕累托排序方法，因为它和帕累托效率的思想密切相关）。如果将这种说法作为制定社会决策的准则，有什么缺陷？

【复习内容】个人偏好加总；社会福利；不同配置的社会福利比较

【参考答案】

这种标准过于严格，在现实中几乎无法运用。因为这样的假设要求社会中的每个人的偏好都是一样的。

这直接导致帕累托有效率的两个不同配置无法进行比较。以两个人的社会为例进行说明。假设这两个人分别为 A 和 B。

假设配置 x 为（A 拥有一切，B 一无所有）；配置 y 为（A 一无所有，B 拥有一切）。这两个配置都是帕累托有效率的，因为按照帕累托效率的定义，这两种情形下，已无法找到让一人状况变好而又不损害另一方利益的方法。

显然，A 更喜欢 x，B 更喜欢 y。如果按照题目的要求，我们无法比较 x 和 y 这两个配置的福利大小。

2.罗尔斯福利函数的值仅取决于福利最差的那个人的效用。与该福利函数相对的一种函数叫作“尼采”（Nietzschean）福利函数——其值仅取决于福利最好的那个人的效用。尼采福利函数的表达形式是什么样的？

【复习内容】

罗尔斯社会福利函数，又叫极小极大（即极大中的极小）社会福利函数：

$$W(u_1,...,u_n)=\min\{u_1,...,u_n\}.$$

这个效用函数是说，某个配置的社会福利仅取决于状况最差的那个人的福利，即取决于效用最小的人的福利

【参考答案】

比照罗尔斯社会福利函数，容易写出尼采福利函数的表达式:

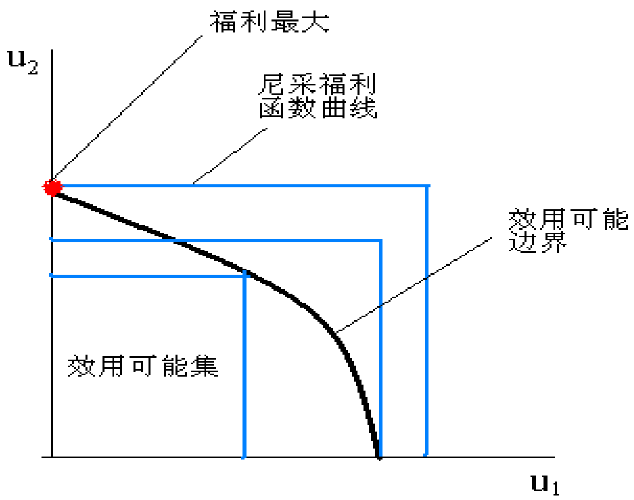
$$W(u_1,...,u_n)=\max\{u_1,...,u_n\}.$$

3.假设效用可能集是一个凸集，而且假设每个消费者仅关注他自己的消费。如果我们采用尼采福利函数进行分析，那么什么样的配置能使得该函数达到极大值？

【复习内容】凸集；尼采社会福利函数；社会福利最大化

【参考答案】

为简单起见，假设社会是两人社会，该两人分为为 1 和 2。由于题目告知，效用可能集为凸集、福利函数为尼采福利函数，因此可以相应画出效用可能集和三条尼采福利函数曲线。



图：效用可能集为凸且社会福利函数为尼采福利函数情形下的福利最大化

由于尼采福利函数只关注状况最好的那个人的效用，从图可以看出，这种情形下的福利最大化解是边界解，即位于纵轴上的那个红点，也就是说此时第 2 个人得到了全部财富。类似地，你也可以画出第 1 个人得到了全部财富的情形。

4.假设某个配置是帕累托有效率的，而且每个人仅关注他自己的消费。证明，在这种情形下，必然存在这样的人：他不妒忌其他任何人。请使用我们课文中对妒忌的定义。（这个题目需要思考一会，但这样做是值得的。）

【复习内容】帕累托有效率配置；妒忌的定义

若第 i 人更喜欢第 j 人的商品束而不是他自己的，则 i 妒忌 j。

【参考答案】

我们先分析两个人的社会。因为将这个情形下得到的结论推广到多人社会并不难。

使用反正法。即不存在这样的人。那么这意味着这两个人是互相妒忌的。根据妒忌的定义可知，这两个人必然会进行交易，从而两个人的状况比交易前更好。或者简单地说，既然妒忌意味着这两个人更喜欢对方的商品束，那么就让他们将两个商品束对换，这样他们的状况都改善了。

然而，这样我们就得到了矛盾，因为题目告知我们初始配置是帕累托有效率的，换句话说，初始时，每个人的效用都已最大。由此可知，题目中的结论是正确的。

5.制定投票日程的能力，是一种非常重要的能力。假设社会偏好将通过多数人投票方法决定，并且将备选项成对进行投票角逐。假设人们的偏好如教材图 30.1 所示，请制定投票日程从而让 y 最终胜出。再制定另外一个日程让 z 最终胜出。如果你具备这种制定投票日程的能力，那么你是利用了社会偏好的什么性质？

【复习内容】多数人投票；投票结果的操纵

【参考答案】

（1）让 y 最终胜出的日程安排：第一轮投票先在 x 和 z 之间进行，第二轮投票在胜出者和 y 之间进行。

（2）让 z 最终胜出的日程安排：第一轮先让 x 和 y 对决，然后胜出者再和 z 对决。

（3）你利用了社会偏好不可传递性的特点。

为方便说明，将教材中图 30.1 拷贝如下。

Person A	Person B	Person C
x	y	z
y	z	x
z	x	y

从教材中的例子我们可以知道，采用多数人投票的方法，对于**这个具体的例子**来说，如

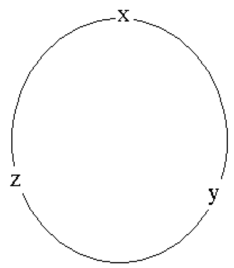
果想让哪个选项最终胜出，必须将这个选项安排在最后一轮进行投票。

因此，如果想让 y 最终胜出，那么对投票日程的安排是：第一轮投票先在 x 和 z 之间进行，第二轮投票在胜出者和 y 之间进行。我们来验证一下。根据多数人投票的规则可知，第一轮投票中，B 和 C 将票投给 z，而只有 A 投给 x，因此多数人选择了 z，z 在第一轮胜出。（这也正是多数人投票的含义）。

在第二轮投票中，z 将和 y 对决。按照上面的分析思路，可知在第二轮 y 胜出。因此最终的结果是 z。

类似地，如果想让 z 胜出，那么第一轮先让 x 和 y 对决，然后胜出者再和 z 对决。

你之所以能操作投票的最终结果，是因为你利用了社会偏好的不可传递性这个性质。因为在这个例子中，**多数人**的偏好排序：$x \succ y$, $y \succ z$ 和 $z \succ x$ 。如下图所示：



在这个图形中，哪一点是终点？这取决于你在什么位置将这个圆形砍为两半。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:
A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

34.外部性（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

34 外部性

如果某消费者直接关心其他人的生产或消费行为，则称这些行为具有**消费外部性**（consumption externality）。例如我对下列这些行为都有明确的偏好：凌晨 3 点邻居大声地放音乐、餐厅邻座的客人抽着廉价的雪茄或者当地汽车排放的尾气。这些都是**负的**消费外部性例子。另外一方面，我能从欣赏邻居的花园中得到快乐，这是**正的**消费外部性的例子。

类似地，如果某个企业的生产可能性受到另外一个企业或消费者行为的影响时，则称这样的行为具有**生产外部性**（production externality）。最好的例子是，果园旁边住着养蜂人，它们之间存在着相互的正的生产外部性，每一方的生产对另一方产量的影响都是正的。类似地，渔场会关心倾倒在捕鱼区的污染物的数量，因为这种行为显然会影响它的捕鱼量。这是负的生产外部性。

外部性的最重要特征是，人们关注某商品，但市场中又没有这种商品销售。市场中没有人销售凌晨 3 点钟播放的噪音、或者廉价雪茄飘出的厌恶，也没有人出售邻居的漂亮花园。正是由于不存在这些外部性的市场，才会出现问题。

直到目前，我们一直隐含地假设：每个人在做消费或生产决策时，根本不需要考虑他人的行为。消费者和生产者之间的所有互动都在市场中发生，所以经济人仅需要知道市场价格、消费量或产量即可。在本章，我们将放松这个假设，分析外部性的经济后果。

在前面几章，我们已经知道，在**不存在**外部性的情形下，市场机制能够实现帕累托有效率的配置。如果存在外部性，则市场不一定能导致帕累托有效率的供给。然而，某些社会制度，例如法律体系或者政府的干预，能够在某种程度上“模仿”市场机制，因此可以实现帕累托效率。在本章，我们将分析这些制度是如何运行的。

34.1 吸烟者和不吸烟者

我们用一个例子说明外部性问题的主要细节。例如，A 和 B 为室友，他们都对“钱”和“吸烟”有偏好。假设两人都喜欢钱，但 A 喜欢抽烟而 B 喜欢清新的空气。

使用艾奇沃斯盒表示这两人的消费可能性。横轴的长度表示两人拥有的总钱数，纵轴的高表示抽烟产生的总烟雾量。A 的效用随钱数以及吸烟量的增加而增加；B 的效用随着钱数增加而增加，但却随着吸烟量的增加而减少，因为 B 喜欢清新的空气。我们用 0 到 1 之间的数字表示房间内的烟雾浓度，其中 0 表示没有烟雾，而 1 表示房间内充满了烟雾。

这样，我们就得到了图 34.1。注意该图非常类似标准的艾奇沃斯盒，但是意思却不一样。A 喜欢抽烟但 B 却讨厌烟雾，因此如果 A 少抽烟，B 就移向位置更高的无差异曲线（从 B 的原点看）。一定注意横轴和纵轴衡量方法的区别之处。我们从左下方 A 的原点沿着横轴衡

量 A 的钱数，从右上方 B 的原点沿着横轴衡量 B 的钱数。但是烟雾浓度却是从左下方 A 的原点处沿着纵轴衡量的。这样差别就出来了：由于钱数能在两个人之间分配，因此我们需要衡量两个钱数，但是这两个人却必须共享一个烟雾数。

在普通的艾奇沃斯盒内，当 A 减少商品 2 的消费量时，B 的状况会变好，但这是由于 B 消费了更多的商品 2。在图 34.1 的艾奇沃斯盒中，当 A 减少商品 2 的消费（少抽烟）时，B 的状况也变好，但是原因却不同。在这个例子中，当 A 少抽烟时，B 的状况变好，原因在于两个人消费的烟雾量是一样的，但是 B 却厌恶烟雾。

我们已经说明了这两个人的消费可能性和他们的偏好。他们的禀赋是多少？假设一开始他们的钱数相同，比如每人都有 100 元，因此他们的禀赋位于图 34.1 虚垂线的某个位置。为了确定该禀赋的具体位置，我们必须确定烟雾/清新空气的初始“禀赋”。

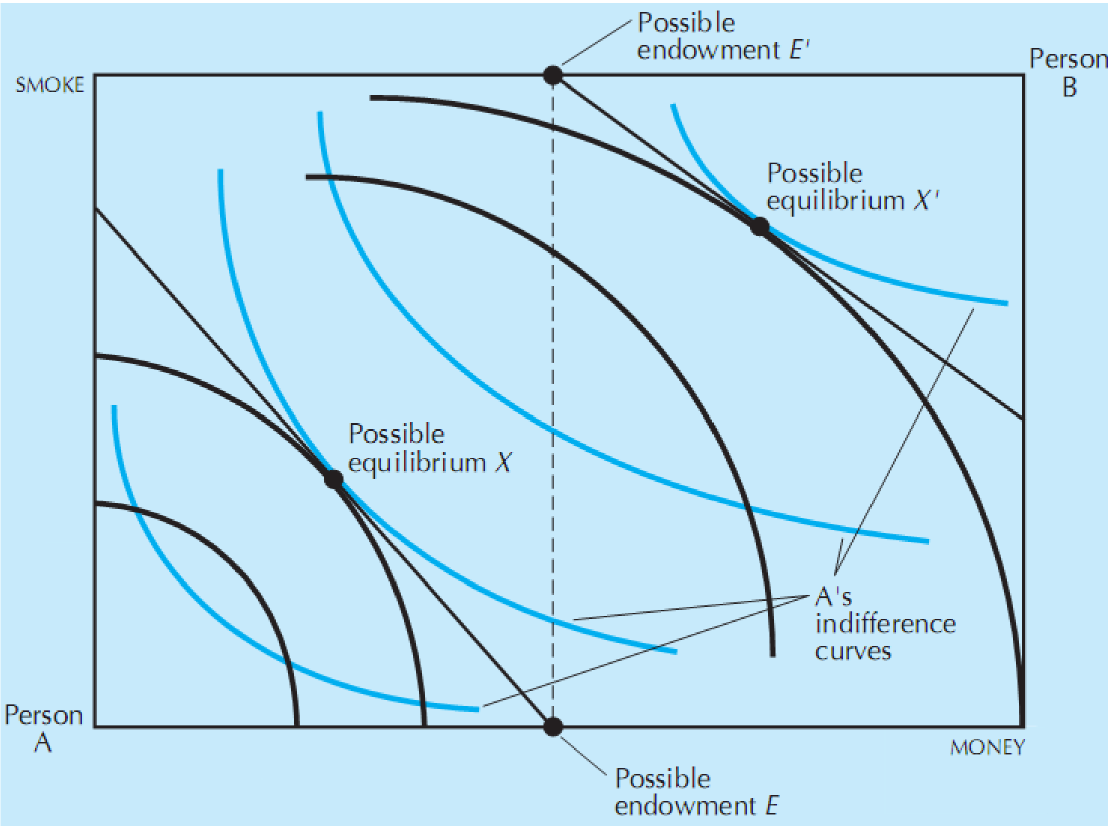


图 34.1：对钱和烟雾的偏好。A 喜欢抽烟但 B 却讨厌烟雾。最终均衡在什么位置取决于我们从哪个禀赋点开始。

这个问题的答案取决于吸烟者和不吸烟者的法定权利。也许 A 有权利想抽多少就抽多少，B 只能忍受；也许 B 有权利享受清新的空气。或者每个人的权利介于上述两个极端之间。

烟雾的初始禀赋取决于法律规定。这和普通类型商品的初始禀赋区别不大。说 A 的初始禀赋为 100 元，意味着他可以消费 100 元，或者他可以送人，或者他可以和别人交易。说某人“拥有”100 元或者“有权拥有”100 元，这是对财产权的法律定义。类似地，如果某人对清新空气享有产权，这表示他可以消费清新空气、或者将这项权利送人、或者将该权利卖给别人。在这个意义上，对清新空气拥有产权，和有权拥有 100 元没什么区别。

假设法律赋予 B 享有清新空气的权利。则初始禀赋为 E 点，如图 34.1 所示；E 点表示 A 的禀赋为（100,0），B 的禀赋为（100,0）。也就是说在交易之前，A 和 B 的钱的禀赋都为 100 元，烟雾的禀赋为 0（即全为清新空气）。

和以前一样，在不存在外部性的情形下，初始禀赋没有理由一定是帕累托有效率的。对清新空气拥有产权，意味着你可以用它交换你想要的商品，在本例中，就是交换钱。B 很有可能就这么做。图 34.1 中的 X 点就是这样的例子。

和以前一样，在帕累托有效率的配置点上，任何一个消费者的状况在不损害其他人的利益下，已无法变得更好。这样的配置点可用以下的相切条件刻画：A 的烟雾与金钱之间的边际替代率，等于 B 的。如图 34.1 所示。容易想象出 A 和 B 交易后达到这样有效率的配置点。事实上，由于 B 拥有清新空气，他可以“堕落”一下，“受贿”吸入烟雾。也就是说他卖出部分清新空气。

当然，法律也可以不赋予 B 享有清新空气的权利，比如法律赋予 B 有吸烟的权利，他想吸多少就吸多少。在这种情形下，B 可以向 A“行贿”让他减少吸烟量。B 有吸烟权时对应的禀赋为图 34.1 中的E'点。和以前一样，E'点通常不是帕累托有效率的，因此我们可以想象 A 和 B 交易后，达到一个双方都偏好的点比如X'点。

点X和点X'都是帕累托有效率的配置点，但是它们来自不同的初始禀赋。当然，吸烟者 A 在X'点的状况比在X点好，而不吸烟者 B 在X点的状况比在X'点好。这两个点的商品配置量不同，但在效率的意义上，它们同样令人满意。

事实上，我们没有必要仅关注这两个有效率的配置点。和往常一样，合同曲线上的任何一点，都是烟雾和金钱的有效率的配置点。如果 A 和 B 能够自由交易这两种商品，他们最终会在这条合同曲线上的某个点上。具体在哪个点取决于这两个人的商品财产权的状况，也取决于他们按什么机制进行交易。

他们可以采用的一种交易机制是价格机制。和以前一样，假设有个拍卖人喊价，然后询问每个人愿意按这个价格购买多少商品。如果初始禀赋中 A 有吸烟的权利，他可以考虑卖掉部分吸烟权给 B。类似地，如果 B 有享受清新空气的权利，他可以卖出部分权利给 A。

当拍卖人终于找到能使供给等于需求的一组价格时，问题就解决了：我们得到了一个帕累托有效率的结果。如果存在烟雾这种商品的市场，竞争均衡必然是帕累托有效率的均衡。而且，竞争价格将等于这两种商品的边际替代率，这和标准情形是一样的。

这类似于往常的艾奇沃斯盒分析，但是我们分析的框架稍微有些不同。如果某商品具

有外部性，而且它的产权是界定清晰的，那么不管是谁拥有产权，通过交易，他们都能达到帕累托有效率的配置。如果我们建立外部性这种“商品”的市场以促进交易，那么该市场也将运行良好。

如果财产权不明确，那么就会产生问题。若 A 认为他有权吸烟，而 B 认为他有权享受清新空气，难题就出现了。**现实中的外部性问题，通常是由于财产权界定不明确而引起的**。

我的邻居认为他有权凌晨 3 点吹小号，我则认为我有权享受安静的环境。一个企业认为它有权排放废气，我认为它没权这么做。在财产权界定不清晰的情形下，外部性的数量通常是无效率的，这意味着存在着改变外部性的数量的方法，从而让双方的状况变得更好。如果财产权界定清晰，人们就可以交易这些产生外部性的权利，正如他们交易普通商品的生产和消费权利一样。

34.2 拟线性偏好和科斯定理

我们在上面断言，只要财产权界定清晰，人们之间的交易就能导致外部性有效率的配置。一般来说，帕累托有效率配置中的外部性的数量，取决于财产权的归属。在上面两个室友的例子中，烟雾量取决于是吸烟者有吸烟的权利，还是不吸烟者有享受清新空气的权利。

但是，外部性的数量也可能和财产权的归属无关，这当然是一种特殊的情形。这就是：如果人们的偏好都是**拟线性**的(quasilinear)，那么在每个有效率的解中，外部性的数量是一样的。

我们用艾奇沃斯盒分析吸烟者和不吸烟者的行为。如图 34.2 所示。假设 A 和 B 的偏好都是拟线性的，即他们的效用函数都具有 $u(m,s)=am+v(s)$ 这样的形式，其中 m 表示金钱，s 表示烟雾。在这种情形下，A 的所有无差异曲线可以认为由一条无差异曲线平移后得到的。B 的无差异曲线也是这样的。因此，A 和 B 的无差异曲线的切点连线，也就是帕累托有效率配置集（或合同曲线），必定是一条水平线。这意味着，在每个帕累托有效率的配置上，烟雾的数量都是相等的[一]。不同帕累托有效率配置点的差别在于金钱数不同。

在某些条件下，具有外部性的商品的有效率数量和财产权的分配无关，这样的结果有时称为科斯定理（Coase Theorem）。然而，必须强调这个结果的前提条件非常特殊。拟线性偏好的假设意味着，对具有外部性商品的需求和收入的分配无关。因此，重新配置禀赋不会影响外部性有效率的数量。这个结果有时这样表达：如果不存在“收入效应”，则科斯定

[一] 必须注意 A 和 B 的偏好均为拟线性偏好是指，A 和 B 对烟雾的偏好不一定是线性的，但对金钱的偏好是线性的。我们在第 4 章和第 6 章已经知道，在这种情形下，A 对烟雾的需求和收入无关（收入效应为零），这就是说 A 对烟雾的需求量在不同收入水平下是唯一的，B 对烟雾的需求也是这样的，由于在 A 和 B 无差异曲线相切之处，A 和 B 对烟雾的需求量相同，这说明烟雾量是唯一的。译者注。

理成立[一]。

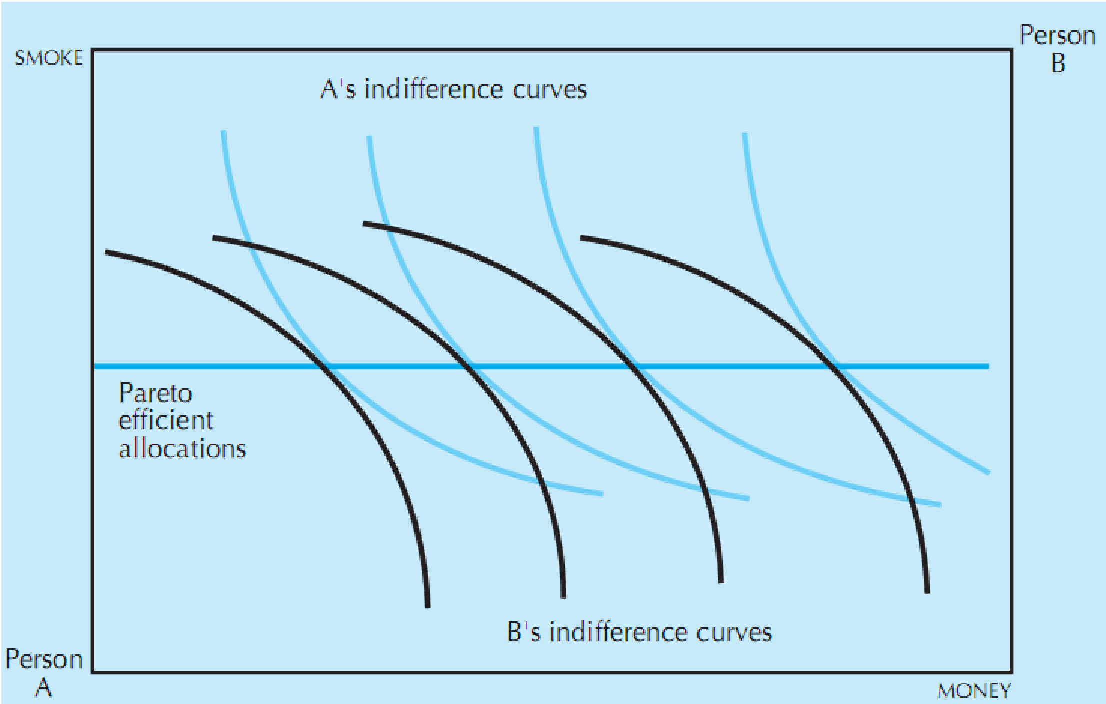


图 34.2：拟线性偏好和科斯定理。如果每个消费者的偏好都是拟线性的，因此各自的所有无差异曲线都是各自无差异曲线平移后得到的，帕累托有效率配置集是一条水平线。因此，在每个有效率的配置点上，外部性的数量是相同的，在这个例子中就是烟雾量相同。

在这种情形下，在帕累托有效率的配置中，外部性的数量是唯一的。在我们上面的例子中，在不同帕累托有效率配置中，两个消费者拥有的钱数不同；但是外部性的数量，即烟雾量却和财富的分配无关。

34.3 生产外部性

下面我们分析生产外部性的例子。钢铁厂 S 生产 s 单位的钢铁和 x 单位的污染物，这些污染物倒入附近的河中。渔场 F 位于这条河的下游，因此企业 S 的污染会对它的产量造成负的影响。

令钢铁厂 S 的成本函数为 $c_s(s,x)$，其中 s 是钢铁产量，x 是污染物的产量。渔场 F 的成本函数为 $c_f(f,x)$，其中 f 表示鱼的产量，x 是污染物得数量。注意，渔场 F 生产既定

[一] 罗纳德.科斯（Ronald Coase）是芝加哥大学法学院的荣誉退休教授。他的著名论文 “The Problem of Social Costs,” The Journal of Law & Economics,3(October 1960)，引起了很大的关注。人们对这篇论文给出了各种各样的解释。有些人认为科斯的意思是说，外部性的交易只有在交易成本为零的情形下才能实现帕累托有效率的结果，他们认为科斯的意思不是说这种结果和财产权的分配无关。科斯因为这项工作而于 1991 年获得诺贝尔经济学奖。

数量的鱼的成本，取决于钢铁厂 S 生产的污染数量。假设：污染**增加**了鱼的生产成本，即 $\Delta c_f / \Delta x > 0$；污染**降低**了钢铁厂生产钢铁的成本，即 $\Delta c_s / \Delta x \le 0$。后面这个假设是说，增加污染的数量会降低钢铁的生产成本，换句话说，减少污染会增加生产钢铁的成本，至少在某些产量范围内是这样的。

钢铁厂的利润最大化问题为：

$$\max_{s,x} \quad p_s s - c_s(s,x)$$

渔场的最大化问题为：

$$\max_{f} \quad p_f f - c_f(f,x).$$

注意，钢铁厂可以选择污染的产量，但是渔场必须接受这个污染量。

钢铁厂利润最大化问题的一阶条件为

$$p_s = \frac{\Delta c_s(s^*, x^*)}{\Delta s}$$

$$0 = \frac{\Delta c_s(s^*, x^*)}{\Delta x}$$

渔场的利润最大化问题的一阶条件为

$$p_f = \frac{\Delta c_f(f^*, x^*)}{\Delta f}$$

这些条件是说，在利润最大化的点上，每种产品（钢铁和污染）的价格应该等于它的边际成本。对于钢铁厂来说，它的一种产品是污染，根据假设，污染的价格为 0。因此，污染的最优条件是说一直生产污染直至污染的边际成本为 0。

容易看出这个例子中的外部性：渔场关心钢铁厂得污染数量，但是对此却毫无办法。钢铁厂在做利润最大化决策时仅关注钢铁的生产成本，它不考虑它强加在渔场身上的成本。渔场生产鱼的成本随着污染的增加而增加，这是钢铁生产的导致的部分**社会成本**（social cost），但是钢铁厂却不会考虑。一般来说，从社会的观点来看，可以预期钢铁厂生产的污染数量会**过多**，因为钢铁厂不会考虑污染对渔场造成的影响。

对于钢铁厂和渔场来说，什么样的生产计划是帕累托有效率的？有种简单的做法可以看清这个计划的样子。假设钢铁厂和渔场合并，合并后生产钢铁和鱼（可能还有污染）。这样，外部性就被消除了！因为生产的外部性是指一个企业的行为影响了另一个企业的生产可能性。如果只有一个企业，那么它在选择利润最大化的生产方案时，会仔细考虑不同“部门”间的相互影响。我们说，这种情形下，对财产权进行重新分配就将外部性**内部化**了

（internalized）。在合并之前，每个企业有权生产任意数量的钢铁、污染或者鱼，而不会考虑对方的行为。合并之后，新企业有权控制钢铁部门和渔场部门的生产。

合并后新企业的利润最大化问题为

$$\max_{s,f,x} \quad p_s s + p_f f - c_s(s,x) - c_f(f,x),$$

这个最大化问题的最优条件为

$$p_s = \frac{\Delta c_s(\hat{s},\hat{x})}{\Delta s}$$

$$p_f = \frac{\Delta c_f(\hat{f},\hat{x})}{\Delta f}$$

$$0 = \frac{\Delta c_s(\hat{s},\hat{x})}{\Delta x} + \frac{\Delta c_f(\hat{f},\hat{x})}{\Delta x}.$$

最重要的条件是最后一个式子。这个条件表明，合并后的企业会同时考虑，污染对钢铁厂边际成本的影响以及污染对渔场边际成本的影响。当钢铁部门决定生产多少污染时，它考虑到了该行为对渔场部门利润的影响；也就是说，它考虑到了自身生产的社会成本。

这意味着合并后企业将生产多少污染？在合并之前，两个企业是独立决策的，污染的最优数量由下列条件决定

$$\frac{\Delta c_s(s^*,x^*)}{\Delta x} = 0 \qquad (34.1)$$

也就是说，钢铁厂会一直生产污染，直至边际成本为零：

$$MC_S(s^*,x^*) = 0.$$

合并后的企业，污染的最优数量由下列条件决定

$$\frac{\Delta c_s(\hat{s},\hat{x})}{\Delta x} + \frac{\Delta c_f(\hat{f},\hat{x})}{\Delta x} = 0 \qquad (34.2)$$

也就是说，钢铁厂生产污染直至钢铁厂的边际成本和渔场的边际成本之和等于零。这个条件又可以写为

$$-\frac{\Delta c_s(\hat{s},\hat{x})}{\Delta x} = \frac{\Delta c_f(\hat{f},\hat{x})}{\Delta x} > 0 \qquad (34.3)$$

或者

$$-MC_S(\hat{s},\hat{x}) = MC_F(\hat{f},\hat{x}).$$

在最后一个式子中，$MC_F(\hat{f},\hat{x}) > 0$，这是因为污染增加会使生产既定数量的鱼的成本增加。因此，合并后的企业会在$-MC_S(\hat{s},\hat{x}) > 0$的地方进行生产；也就是说，与合并前

相比，它想生产的污染变少了。当钢铁厂考虑到了外部性的真正社会成本时，它将减少最优污染数量。

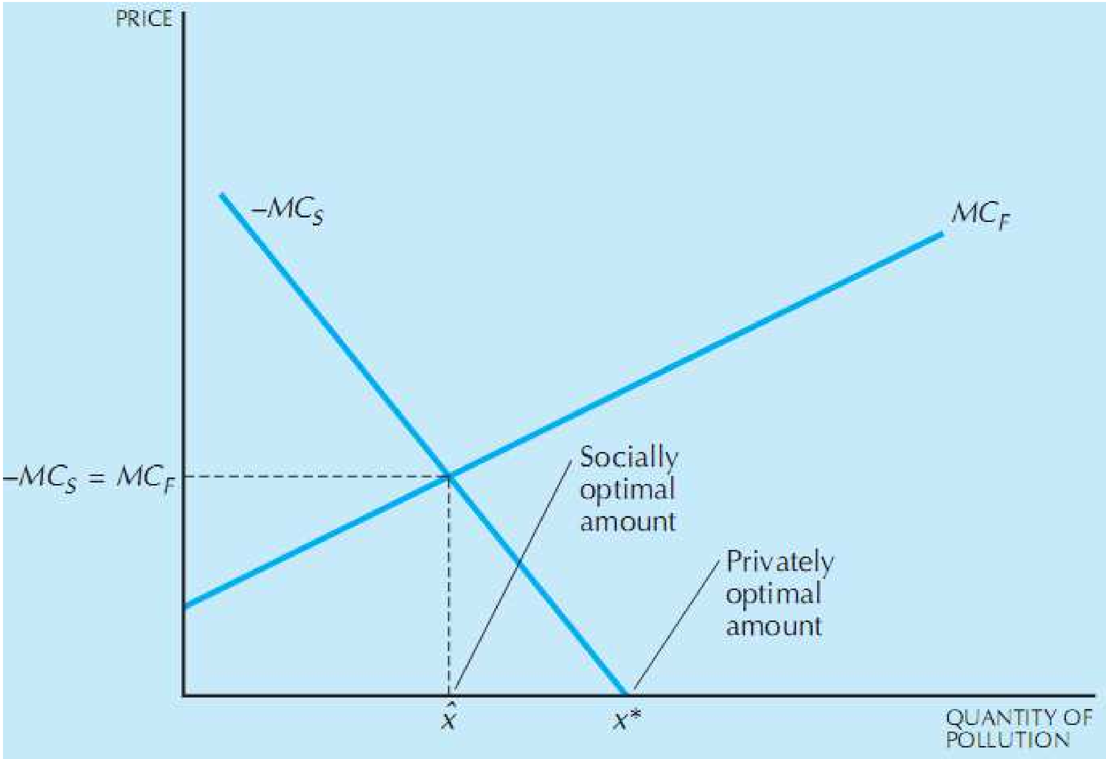


图 34.3：社会成本和私人成本。合并之前，钢铁厂在污染的边际成本等于零之处进行生产。但是帕累托有效率的污染水平要求它在价格等于社会成本之处进行生产，其中社会成本不仅包括其私人成本，还包括渔场承担的成本。

当钢铁厂仅考虑生产钢铁的私人成本（private costs）最小时，它会在污染的边际成本等于零之处进行生产，因为此时污染的私人成本最小；但是帕累托有效率的污染水平要求污染的社会成本最小。在帕累托有效率的污染水平处，两个企业的污染的边际成本之和必然等于零。

这个结论可用图 34.3 说明。在该图中，$-MC_S$ 衡量钢铁厂生产污染的边际成本。曲线 MC_F 衡量渔场承担的污染的边际成本。合并之前，追求利润最大化的钢铁厂会在污染的边际成本等于零之处进行生产。

但是，在帕累托有效率的污染水平上，钢铁厂生产的污染数量要使得污染的边际收入等于污染的边际社会成本，其中边际社会成本由污染对两个企业造成的边际成本组成。在帕累托有效率的污染产量处，钢铁厂愿意为额外一单位污染支付的补偿，应该等于这一单位污染的社会成本，这个社会成本不仅包括它的私人成本，还包括它对渔场造成的成本。

这和前几章得出的效率条件是完全一致的。在前几章，我们假设不存在外部性，因此私人成本和社会成本相等。在这样的情形下，市场将会决定每种商品有效率的数量。但如果私人成本不等于社会成本，市场本身可能不足以实现帕累托效率。

例子：排污许可

每个人都想要干净的环境…最好其他人出钱保持环境整洁。即使大家一致认为应该将污染控制在何种水平，仍有问题尚未解决。这个问题就是，为了实现这个目标应该采取什么样的最节省钱的方法？

以一氧化氮的排放为例。不同企业可能发现减少排放所花费的成本是不同的，有的高，有的低。应该要求这些企业减少的排放量相同吗？这个相同是应该按照排放废气的体积衡量，还是按照相同的百分比衡量，亦或者还有其他的标准？

我们来分析一个简单的经济模型。假设只有两个企业。规定企业 1 和 2 的排放定额分别为 x_1 和 x_2。为了实现各自的排放定额，企业 1 和 2 花费的成本分别为 $c_1(x_1)$ 和 $c_2(x_2)$。总排放量限定在既定目标水平 X 。在这个总排放量的约束下，实现目标排放量的成本最小化问题为：

$$\min_{x_1,x_2} c_1(x_1)+c_2(x_2)$$

$$\text{使得 } x_1+x_2=X\,.$$

由标准的经济分析可知，两个企业控制废气排放的边际成本必定相等。如果企业 1 控制废气排放的边际成本大于企业 2 的，那么降低企业 1 的排放定额增加企业 2 的排放定额，就能降低总成本。

我们如何才能实现这个结果？如果政府部门拥有所有企业排放成本的信息，那么他们就能够计算出合适的生产模式，并且规定所有企业执行，这样就能实现这个结果。问题是，搜集和更新这些信息的成本非常高，因此，最优解说起来简单做起来难！

很多经济学家认为，市场是实现帕累托有效率排放量的最优方式。南加州正在试点这样的市场，即允许企业在市场上交易排放定额。下面说说这个市场是如何运行的[一]。

政府对南加州 2700 家最大的污染企业事先指定排放限额。每个企业的排放限额按照该企业上一年的排放量确定，准确地说是上一年排放量的 92%。如果企业的实际排放量恰好达到限额，那么它不会受到处罚。然而，如果它的实际排放量低于它的限额，那么它可以将多余废气的“排放权”放到市场上进行交易。

假设某个企业的限额为每年排放 95 吨。如果它能做到实际只排放 90 吨，那么它可以将多余的 5 吨排放权卖给其他企业。每个企业都会比较：排放许可的市场价格与自己减少排放

[一] Richard Stevenson, "Trying a Market Approach to Smog," New York Times, March 25, 1992, C1.

引起的成本大小。从而决定是自己减少排放还是购买排放权。

有些企业能够比较轻松地减少废气排放，而有些企业却要花费很大的代价，前者可以将排放权出售给后者。均衡时，一吨废气排放权的价格，应该等于企业减少一吨废气所花费的成本。但这正是最优排放量的条件！排放权的市场自动实现了帕累托有效率的排放模式。

34.4 对最优条件的解释

上面得到的帕累托有效率条件有几种有用的解释。每种解释都对应着一种纠正生产外部性造成的效率损失的方法。

第一种解释是，钢铁厂面对的污染价格是错误的。如果仅从钢铁厂的角度看，它生产污染的成本为零。但这没有考虑到污染对渔场造成的成本。因此，如果能让钢铁厂面对污染的真正的社会成本，就能纠正这种情形。

对应的解决方法是，政府可以对钢铁厂的污染征税。假设每单位污染物的税额为 t 元。则钢铁厂的利润最大化问题变为

$$\max_{s,x} \quad p_s s - c_s(s,x) - tx.$$

这个问题的最优条件为

$$p_s - \frac{\Delta c_s(s,x)}{\Delta s} = 0$$

$$-\frac{\Delta c_s(s,x)}{\Delta x} - t = 0 .$$

将这两个条件与（34.3）比较，可知若令

$$t = \frac{\Delta c_f(\hat{f},\hat{x})}{\Delta x}$$

就会使得这些条件和前面的帕累托有效率的污染水平的条件相同。

这种税称为**庇古税**（Pigouvian tax）[二]。庇古税的问题在于，我们需要知道污染的最优数量，才能决定征多少税。但是如果我们能知道污染的最优水平，可以直接责令钢铁厂生产那么多污染，而不需要多此一举地去征税。

第二种解释是，造成问题的原因在于污染市场的缺失。也就是说，外部性问题产生的原因，是由于钢铁厂的“产品”——污染物——的价格为零。或者说，即使人们愿意花钱让钢铁厂减少污染也没有办法，因为不存在污染这种产品的市场。从社会的观点来看，污染物的价格应该为**负**。

[二]亚瑟.庇古（Arthur Pigou,1877-1959），剑桥大学的经济学家，他在《福利经济学》（The Economics of Welfare）一书中提出了庇古税的思想，这本书比较具有影响。

我们想象一下下面的情形：渔场对清洁的河水拥有产权，但是它可以将这个权利卖出，即允许钢铁厂污染。令 q 表示每单位污染的价格，x 表示钢铁厂生产的污染数量，那么钢铁厂的利润最大化问题为

$$\max_{s,x} \ p_s s - qx - c_s(s,x),$$

渔场的利润最大化问题为

$$\max_{f,x} \ p_f f + qx - c_f(f,x).$$

在钢铁厂的利润函数中，qx 前面有个负号，因为对于钢铁厂来说 qx 代表成本——钢铁厂必须购买生产 x 单位污染的权利。但在渔场的利润函数中，qx 的符号为正，因为渔场卖出了污染权得到了收入。

这两个企业的利润最大化条件为：

$$p_s = \frac{\Delta c_s(s,x)}{\Delta s} \tag{34.4}$$

$$q = -\frac{\Delta c_s(s,x)}{\Delta x} \tag{34.5}$$

$$p_f = \frac{\Delta c_f(f,x)}{\Delta f} \tag{34.6}$$

$$q = \frac{\Delta c_f(f,x)}{\Delta x} \tag{34.7}$$

因此，每个企业在选择应该出售或购买的污染数量时，都面对着自身行为的社会边际成本。如果污染的价格调整到需求等于供给，我们就得到了一个有效率的均衡，这和其他商品的情形是一样的。

注意，在最优解之处，（34.5）和（34.7）式意味着

$$-\frac{\Delta c_s(s,x)}{\Delta x} = \frac{\Delta c_f(f,x)}{\Delta x}.$$

这个式子是说，钢铁厂因减少污染而花费的成本，应该等于渔场从污染减少而得到的边际收益。如果这个条件未得到满足，那么污染水量就不是最优水平。这个条件当然和（34.3）式的条件相同。

在分析这个问题时，我们假设渔场对清洁河水拥有产权，而钢铁厂只能向它购买污染权。但是，事情也可能正好颠倒过来：钢铁厂有污染权，为了让钢铁厂少生产一些污染，渔场只能向它付钱。正如本章一开始的吸烟者和不吸烟者的例子一样，这种方法也能实现有效率的结果。事实上，这个结果和前面渔场有河水产权的结果是一样的，因为这种情形的最优条件正好也是（34.4）—（34.7）式。

为了看清这一点，假设钢铁厂有权生产污染的限额为 $\bar{x}$ ，但是渔场愿意出钱让钢铁厂减少污染。钢铁厂的利润最大化问题为

$$\max_{s,x} \quad p_s s + q(\bar{x} - x) - c_s(s, x) .$$

现在钢铁厂的收入来源有两个：一是它可以卖钢铁；二是它可以卖掉部分污染权。价格等于边际成本的条件变为

$$p_s - \frac{\Delta c_s(s,x)}{\Delta s} = 0 \qquad (34.8)$$

$$-q - \frac{\Delta c_s(s,x)}{\Delta x} = 0 \qquad (34.9)$$

渔场的利润最大化问题为

$$\max_{f,x} \quad p_f f - q(\bar{x} - x) - c_f(f, x) ,$$

这个问题的最优条件为

$$p_f - \frac{\Delta c_f(f,x)}{\Delta f} = 0 \qquad (34.10)$$

$$q - \frac{\Delta c_f(f,x)}{\Delta x} = 0 \qquad (34.11)$$

观察一下即可知道：（34.8）—（34.11）式恰好就是（34.4）—（34.7）式。在生产外部性的情形下，生产的最优模式和财产权的分配无关。当然，利润分配通常取决于财产权的分配。尽管社会结果和财产权的分配无关，但是出于各自利润的考虑，各个企业主对财产权分配的意见是不一样的，因为谁都想得到权利。

34.5 市场信号

最后我们来看看外部性的第三种解释，在某种程度上，这种解释最深入。在钢铁厂和渔场的情形中，如果两个企业可以合并，问题是它们为什么不合并呢？事实上，若你仔细一想，你就会明白这两个企业肯定有合并的激励：如果一个企业的行为影响到另外一个企业，则它们联合行动比各自行事的利润更大。**利润最大化目标本身就会促使生产外部性内部化**。

换一种说法：如果企业合作的共同利润大于不合作时的利润之和，则会有人购买这些企业，每个企业的购买价等于该企业利润流的现值，这样此人仍能获得额外利润。这个新的购买方可以是原有企业的企业主，或者是其他人。

市场本身会发送生产外部性内部化的信号，这是为什么人们很少能观测到这类生产外部性的一个原因。如果不同部门会影响彼此的利润，大多数企业就已经将这些部门的外部性内

部化。前面提及的果园和养蜂者的例子就是这样的。如果这两方没考虑到彼此间的相互影响，就会存在外部性…但是它们不会傻到认识不到这一点。至少有一方会认识到如果合作则能赚取更大的利润，合作的方式可以是签订互利的协议或者一方将自己的企业卖给另一方。事实上，果园通常自己养蜜蜂来给果树花粉受精。这种特别的外部性很容易内部化。

例子：蜜蜂和杏树

很多种果树都需要蜜蜂传授花粉，这样它们才能结果。根据 Carl Hayden 蜜蜂研究中心（位于亚利桑那州图森市）研究，在人类饮食中，约有 1/3 需要蜜蜂授粉。他们还发现，美国有 50 多种农作物需要蜜蜂授粉，这些作物的产值高达 200 多亿美元/年[一]。

有些果园主自己养蜜蜂；有些则依赖于邻居的蜜蜂或者野蜂。然而，正如外部性理论指明的，蜜蜂供应不足时，最自然的解决方法是通过蜜蜂服务市场解决。

例如，加利福尼亚州的杏树市场。加州的杏树大约有 53 万英亩，每年需要 100 万巢的蜜蜂进行授粉。但是加州当地只有 44 万巢蜜蜂。加州本地的蜜蜂不足以对这些杏树进行授粉。

解决的办法是从邻近的州引进蜜蜂。事实上，的确存在这样的市场，来自北达科他州、华盛顿州和科罗拉多州的养蜂人带着蜂巢涌进加州，从而弥补了加州本地蜂巢的不足。杏树种植者对这些养蜂人的回报是丰厚的：2004 年，蜜蜂授粉服务价格为 54 美元/蜂巢。

34.6 公地的悲剧

我们在上面认为，若财产权界定清晰，就不会存在生产外部性的问题。但是如果产权不清，则经济相互作用的结果无疑是无效率的。

在本节，我们将分析一个非常有名的无效率问题，这就是“公地的悲剧”（the tragedy of the commons）[二]。哈丁（Hardin）1968 年在《科学》上发表了《公地的悲剧》论文。我们使用他提出的例子即牧场进行分析，尽管也可以使用其他例子进行说明。

例如一个农业村庄，村民在一块公共牧场放牧奶牛。我们试图比较两种分配机制：一种是私有产权解，即这块牧场为某人所有，他可以决定应该放牧多少头奶牛；另外一种是牧场为村民共有，任何一个村民想放养多少就放养多少，没有限制。

假设买头奶牛要花费 a 元。一头奶牛的产奶量取决于公共牧场上其他奶牛的数量。令 $f(c)$ 表示若公共牧场上奶牛数量为 c 时，所有奶牛产奶总数的价值。因此，每头奶牛产奶量的价值为平均价值 $f(c)/c$。

如果要使所有村民的总财富最大化，应该在牧场上放牧多少头奶牛？为了使总财富最

[一] Anna Oberthur, "Almond Growers Face Need for Bees," Associated Press, February 29. 2004.
[二] G. Hardin, "The Tragedy of the Commons," Science, 1968, 1243-47

大，建立下面的问题

$$\max_{c} f(c)-ac.$$

现在你应该清楚，当一头奶牛的边际产量等于奶牛的边际成本 a 时[三]，奶牛数量最大：

$$MP(c^{*})=a.$$

如果奶牛的边际产量大于 a，应该多放牧；若小于 a，则应该少放牧。

如果某个村民对这块牧场拥有产权，那么他可以限制牧场上放养的牛数，这样就可以得到上面的解，即最优牛数 c^{*} 要满足 $MP(c^{*})=a$。因为在这种情形下，牧场主可以正好购买 c^{*} 头奶牛。

现在如果每个村民都有权在牧场上放养，结果会如何？每个村民是否放养一头奶牛的决策，取决于这头牛的边际产量是否大于奶牛的成本 a。假设现在牧场上的奶牛数为 c，则每头牛的产奶量为 $f(c)/c$。当某村民考虑是否增加一头奶牛时，他会做如下推理：总产奶量为 $f(c+1)$，奶牛总数为 $c+1$。因此，每头奶牛产生的收入为 $f(c+1)/(c+1)$。若 $f(c+1)/(c+1)>a$，则他会多放养一头奶牛，因为收入超过了成本。但是，其他村民也会这样推理。因此，村民的放牛总数会一直增加直至每头牛的平均产奶量等于 a。令此时的奶牛总数为 $\hat{c}$，由上面的分析可知，$\hat{c}$ 必然满足

$$\frac{f(\hat{c})}{\hat{c}}=a.$$

我们也可以使用另外一种方法得到上面这个式子。这个方法就是使用自由进入的思想。如果在公共牧场上放牧有钱可赚，村民会一直购买奶牛直至利润被压低为零时为止，即

$$f(\hat{c})-a\hat{c}=0,$$

将这个式子稍作变形就得到了前面那个式子。

当单个村民决定是否购买一头奶牛时，他考虑的是他能得到的额外价值 $f(c)/c$，并将其与购买奶牛的成本 a 进行比较。他这么比较没什么不可以，问题是他在计算时忘记了考虑以下事实：他多放养一头奶牛的行为将会降低牧场上**所有其他**奶牛的产奶量。由于他没有考虑他行为的**社会成本**，每个村民都可能犯这样的“错误”，因此，公共牧场上的奶牛数量太多了。（我们假设每个村民的奶牛数量，对于牧场上的奶牛总数来说是微不足道的。）

图 34.4 说明了这个结论。在该图中我们画出了一条向下倾斜的平均产量曲线，这个假设是合理的，因为随着奶牛数量的增加，每头奶牛的平均产奶量是下降的。

[三] 严格说来，应是边际产品价值等于 a，但是这里可以认为一单位牛奶的价格为一元，这样边际产品价值就和边际产量相等。译者注。

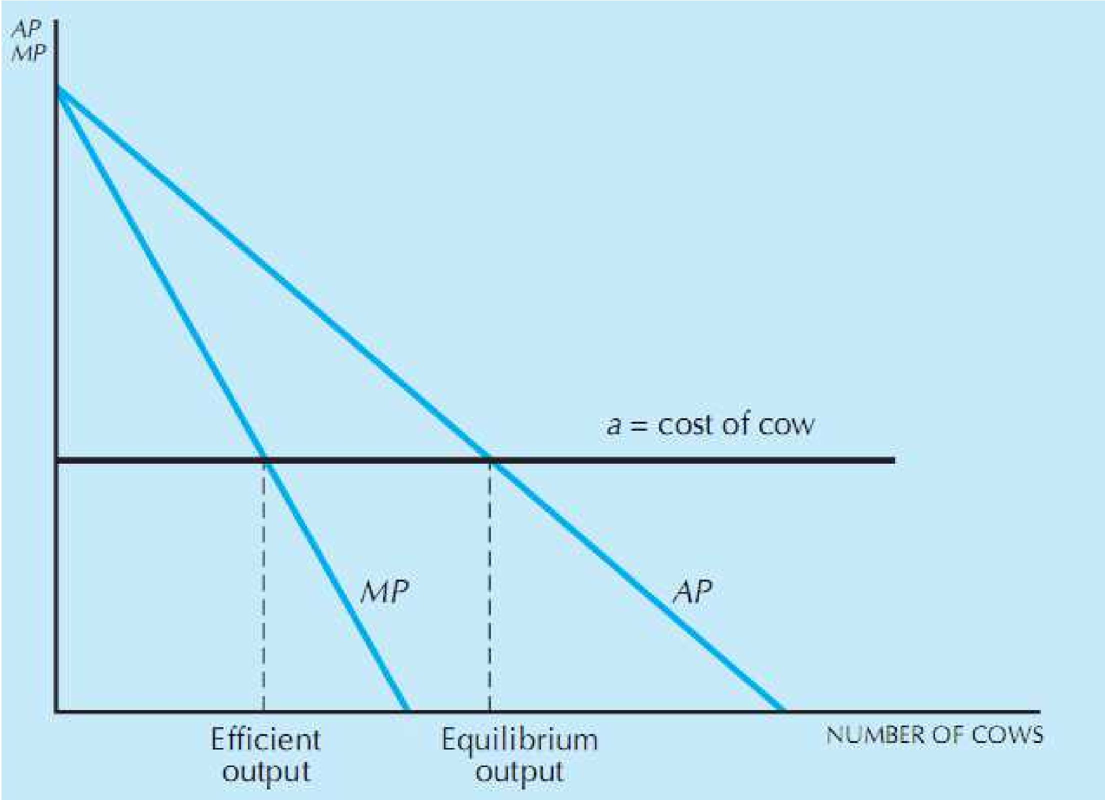


图 34.4：公地的悲剧。如果牧场私有，则牧场主选择的奶牛最优数量，恰好使一头奶牛的边际产量等于一头奶牛的购买成本。但若牧场公有，则放牧的奶牛数一直增加直至利润为零，此时就是过度放牧了。

由于边际产量是下降的，因此边际产量曲线必然位于平均产量曲线的下方。因此，由边际产量等于 a 得到的奶牛最优数量，必然小于根据平均产量等于 a 得到的奶牛数量。公共牧场最终会出现放牧过度的现象，这是由于限制牧场使用机制的缺失。

私有产权提供了这样的机制。事实上，我们已经知道，如果人们关注的东西为某人私有，这样他就可以禁止别人使用、更不用说过度使用，因此根据定义可知，这种情形下不存在外部性的问题。市场机制能实现帕累托有效率的结果。如果不能禁止别人使用某东西，那么这种情形就会导致无效率的结果，我们将在下一章详细分析这种问题。

当然，私有产权不是鼓励资源有效率利用的唯一社会机制。例如，政府可以制定规则限制公共牧场上放牛总数。如果法律能保证这样的规则施行，那么这种方法也能产生有效率的结果。然而，如果有关法律含糊不清或者缺少这样的法律，那么公地的悲剧现象很容易发生。国际水域的过度捕捞，以及由于过度打猎而使得若干动物物种灭绝的事实，都是令人沉痛的例子。

例子：过度捕捞

《纽约时报》有篇报道指出，“过度捕捞已经大幅度减少了鳕鱼、黑线鳕和鲽鱼的数量。几百年来，新英格兰地区的居民一直以这些鱼类为食。[一]”一位专家也指出，新英格兰[二]地区的渔民已经捕捉了鱼类存量的 50%~70%，是鱼类可持续数量的 2 倍多。

过度捕捞是公地悲剧问题的重要例子。每个渔民对鱼类总存量的影响几乎可以忽略不计，但是成千上万个渔民的影响累计起来，就导致了非常严重的后果：鱼类资源消耗严重。新英格兰渔民管理委员会颁布了禁入限令，试图缓解这个问题，它还要求渔民限制出海天数，增加渔网网眼的尺寸等。

如果这些保护措施实施，那么鱼类的数量很有可能至少 5 年内得以恢复。禁止过度捕捞能增加捕捞业的利润。但是，这样的措施无疑会使得行业中渔船的数量大幅减少，小渔民显然最不欢迎这样的措施，因为他们有可能被迫离开这个行业。

例子：新英格兰龙虾

有些捕捞业已实施严格规则，以避免过度捕捞。例如龙虾捕捞业。在这样的情形下，捕虾者小心作业唯恐受到处罚。例如，行业规定任何即将产卵的母虾、尺寸小于规定最小尺寸或者大于规定最大尺寸的龙虾，都必须放生。

母虾的卵可以生出更多的龙虾，“小虾米”长大后可以交尾再繁殖。把这两种虾放生容易理解。可是为什么要把大龙虾也放生呢？因为海洋生物学家认为，大龙虾的繁殖能力更强，而且繁殖出来的龙虾后代个体也大。如果捕虾者总是保留大龙虾，而将小龙虾放生的话，那么小龙虾将会将基因遗传给后代，从而导致龙虾的尺寸越来越小。

在龙虾方面，既有好消息又有坏消息。先说好消息。2003 年缅因州龙虾捕获量为 540 万磅，是 1945-1985 年平均捕捞量的 2.5 倍多。这表明该行业的细致管理已取得回报：龙虾数量显著上升。

然而，似乎保护措施不是导致这种现象的唯一因素。缅因海岸的其他物种数量也发生了显著的变化，例如海胆。有些观察家认为，这些变化是龙虾数量增加的主要原因[三]。

这就导致了坏消息。在缅因州南部的马萨诸塞州和纽约，龙虾的捕捉量急剧下降。没有人准确知道为何一个地区捕捉量那么大而其他地区那么少。具有讽刺意味的是，缅因州捕捉量大的原因，可能是由于该州对有鳕鱼以及海胆的捕捉量增加了，这些动物都吃小龙虾。马萨诸塞州的问题可能是由特殊因素造成的，例如该州海面曾发生了一次规模较大的石油泄漏，而且龙虾也爆发了甲壳毁坏病。另外一个原因是水温升高：那拉干湾区（Narragansett Bay）

[一] "Plenty of Fish in the Sea'? Not Anymore," New York Times, March 25, 1992, A15.
[二] 新英格兰（地区）包括美国的六个州，由北至南分别为：缅因州、新罕布什尔州、佛蒙特州、罗得岛州、康涅狄格州和马萨诸塞州（麻省）。译者注。
[三] See The Economist, "Claws!" August 19, 2004, and Cornelia Dean, "Lobster Boom and Bust," New York Times, August 9, 2004。

的水温在过去 20 年内已上升了 2 摄氏度。

生态可能非常复杂而且变化也非常迅速。避免过度捕捞的努力值得称赞，但是能不能奏效却不得而知。

34.7 汽车污染

前面已指出过，污染是外部性的一个重要例子。一个人开车通常会降低其他人呼吸的空气质量。似乎不受管制的自由市场，不可能产生帕累托有效率的污染数量。更有可能的是，如果消费者不承担污染的成本，产生的污染数量就会过多。

控制汽车污染数量的一种方法是，要求汽车排污量符合某些标准。事实上，这也是 1963 年清洁空气法案实施以来，美国抵制污染的主要做法。这个法案，准确地说，是该法案后来的修正案，对美国汽车制造业规定了排污标准。

劳伦斯.怀特分析了这一规定的收益与成本；下面的分析主要摘选自他的研究报告[一]。

怀特估计汽车安装控污设备的成本为 600 美元/辆，额外的维护成本为 180 美元/辆，为减少里程汽油使用量和使用无铅汽油的成本为 670 美元/辆。因此在车的寿命期限内因为满足排污标准而产生的总成本为 1450 美元/辆。（所有数字均按 1981 年美元计算。）

他认为目前的排污规定方法存在着几个问题。首先，该规定要求所有的汽车满足同样的标准（加州除外，加州对不同汽车施行不同排污标准）。这意味着只要你买车，不管你是住在高污染的地区还是低污染的地区，你都需要额外支付 1450 美元。美国国家科学院 1974 年的一项研究报告指出，63%的汽车不需要实施那么严格的排污标准。怀特指出，“大约有 2/3 的买车人…为不必要的排污系统支付巨额费用。”

第二，控污的绝大部分责任落在汽车制造企业的头上，而买车人承担的部分较少。车主几乎没有让控污设备运行的激励，除非他所在的州检查该设备是否正常运行。

更重要的是，车主没有节约行车费用的激励。在一些城市比如洛杉矶，污染相当严重，因此从经济意义上来说，有必要鼓励人们少开车。但是按照当前的规定，人们在洛杉矶每年驾车行驶 5000 英里，和他们在北达科他州每年行驶 2000 英里，支付的控污费用是相等的。这相当于激励人们在洛杉矶多开车。

解决污染问题的一种方法是收取**排污费**（effluent fees）。怀特指出，征收排污费需要对所有汽车进行年检，检查汽车行驶里程数从而估算出上一年汽车的排污量。于是，不同地区可以按照估算出的排污数对每辆汽车征收排污费。这种方法能使得驾车人面对排污的真实陈本，鼓励他们选择社会最优的排污量。

征收排污费可以鼓励车主寻找降低排污成本的方法，例如购买控污设备、改变驾车习惯

[一] See Lawrence White, The Regulation of Air Pollutant Emissions from Motor Vehicles (Washington, D.C.: American Enterprise Institute for Public Policy Research, 1982)。

以及更换汽车类型等。事实上，在污染比较严重的地区，征收排污费的做法比当前的排污规定要求更严格。只要制定合适的排污费征收标准，就能实现任何人们想要的控污水平…而且这种方法与当前的法律规定相比，更能节省费用。

当然，我们也没理由要求对低污染地区行驶的那批 2/3 车辆废除现行的法律规定。如果实施法律规定造成的成本，小于对里程器年检的成本，那么继续实施法律规定的做法就是合理的。控制汽车污染的合理方法，需要对这些方法的收益和成本进行分析，事实上，所有的社会政策在实施前都应该这么做。

总结

1.福利经济学第一定理表明，在不存在外部性的情形下，自由、竞争市场能实现帕累托有效率的结果。

2.然而，如果存在外部性，竞争市场实现的结果不可能是帕累托有效率的。

3.但是，在这种情形下，政府可以“模拟”市场的作用，即让个体面对他的行为的真正社会成本，而不仅是他的私人成本。

4.更重要的是，法律系统可以确保财产权明晰，从而使得个体进行交易达到帕累托有效率的结果。

5.如果偏好是拟线性的，消费外部性的有效率数量，和财产权分配问题无关。

6.生产外部性的解决方法包括：使用庇古税；为外部性建立市场；允许企业合并；或者允许个体以其他方式转让财产权。

7.公地的悲剧是指公有财产有被过度使用的倾向。这是一种最为常见的外部性。

复习题

1.对还是错？财产权的明确界定通常可以解决外部性问题。

2.对还是错？当偏好是拟线性时，财产最终分配结果和财产权的分配是无关的。

3.列举消费外部性为正的例子、为负的例子；列举生产外部性为正的例子、为负的例子。

4.假设政府希望对公地的使用进行控制，有哪些方法可以达到有效率的使用水平？

复习题答案

1.对还是错？财产权的明确界定通常可以解决外部性问题。

【复习内容】外部性；外部性的一种解决方法——明晰产权

【参考答案】这种说法是正确的。

现实中的外部性问题，通常是由于财产权界定不明确而引起的。理解这个结论只要注意以下两个事实即可：一是在一般商品（比如大白菜）市场上，由于谁拥有大白菜是明确的，因此竞争市场能实现帕累托有效率的结果；二是现实中存在着**过多**的污染（即污染数量不是有效率的），原因正是由于“污染”这种商品的产权界定不明确，如果界定明确，那么“污染”这种商品和大白菜就没本质区别。

在财产权界定不清晰的情形下，外部性的数量通常是无效率的，这意味着存在着改变外部性的数量的方法，从而让双方的状况变得更好。如果财产权界定清晰，人们就可以交易这些产生外部性的权利，正如他们交易普通商品的生产和消费权利一样。

这也就是说，如果某商品具有外部性，而且它的产权是界定清晰的，那么不管是谁拥有产权，通过交易，他们都能达到帕累托有效率的配置。因此，题目中的说法是正确的。

2.对还是错？当偏好是拟线性时，财产最终分配结果和财产权的分配是无关的。

【复习内容】拟线性偏好；科斯定理

【参考答案】这种说法是错误的。

拟线性偏好下，具有**外部性**的商品的有效率数量和财产权的分配无关（这样的结果有时称为科斯定理）。但是，必须强调的是具有外部性的那种商品才是这种情形。在两种商品的情形下，另外一种商品的最终分配结果显然和财产权分配有关。

例如，在吸烟者和不吸烟者的例子中，外部性的数量，即烟雾量却和财富的分配无关，但在不同帕累托有效率配置中，两个消费者拥有的钱数不同的。所以，题目中的说法是错误的。

3.列举消费外部性为正的例子、为负的例子；列举生产外部性为正的例子、为负的例子。

【复习内容】

正的消费外部性：对你来说，下列例子都是正的消费外部性：邻居美丽的花园；别人举

止得体；别人放着你喜欢的音乐；等等。负的消费外部性将上述例子逆反一下即可。

正的生产外部性：若两个企业生产的产品是互补的，则就是正的外部性；典型的例子是果园和养蜂人互为正的生产外部性；旅馆旁边有物美价廉的餐馆，它们也互为正的外部性；等等。

负的生产外部性：过度捕捞；过度打猎；过度污染；等等。

4.假设政府希望对公地的使用进行控制，有哪些方法可以达到有效率的使用水平？

【复习内容】公地的悲剧；公地悲剧问题的解决方法

【参考答案】

我们仍使用教材中公共牧场的例子来说明这个问题。

一种解决方法是将公地私有化。

私有化的方法也有几种：一是政府可以直接将该公地无偿送给某个人，比如送给你。由于此时产权私有，因此，你会限制牧场上放牧的牛数。你的利润最大化问题为

$$\max_c \ f(c)-ac\ .$$

奶牛最优数量c^*的决定条件：$MP(c^*)=a$。这个数量是帕累托有效率的。

二是将该公共牧场出售，比如以拍卖方式出售。假设这块地拍卖的最终价格为k元，则获得这块牧场的人的利润最大化问题为

$$\max_c \ f(c)-ac-k$$

容易看出，此时最优放牧数量仍为c^*，因为该最大化问题的一阶条件仍为$MP(c^*)=a$。

政府可以采用的第二种方法是，直接规定牧场上放牧的奶牛数量为c^*，其中$MP(c^*)=a$。

政府可以采用的第三种方法是出售放牧权。假设放牧每头牛权利的价格为t，则利润最大化问题为

$$\max_c \ f(c)-ac-tc$$

该利润最大化问题的条件为$MP(\hat{c})=a+t$，或者$t=MP(\hat{c})-a$。

政府可以采用的第四种方法是征税，假设对每头牛的征税税率为t，则此时的利润最大化问题和上述第三种方法是一样的，当然解也是一样的。

 曹乾（东南大学 caoqianseu@163.com）

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

35.信息技术（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

35 信息技术

最近 15 年里，社会经济中的最基本变化是**信息经济**（information economy）的出现。流行媒体充斥着关于计算机技术、网络和新软件进展的故事。当然，很多故事刊登在媒体的商业栏目，因为**技术**的革命也是**经济**的革命。

有些人甚至将信息革命比作第二次工业革命。正如工业革命改变了**商品**的生产、流通和消费方式，信息革命改变了**信息**的生产、流通和消费方式。

有些人因此断言这些完全新兴的技术，会导致与目前经济形式完全不同的经济形式。他们认为，比特和原子根本不同。比特的复制成本为零，而且可以光速在世界范围内传播，它们也不会变质。而由原子组成的实物商品不具备上述性质：这些商品的生产和运输都要花费成本，而且它们迟早会变质。

我们承认，比特的这些不寻常的性质会要求新的经济分析，但是我们不认为需要新**类型**的经济分析。毕竟，经济学是关于**人**而不是**商品**的学科。本书中的所有模型，都是关于人们如何决策以及他们之间是如何互动的。我们很少特别指明交易中涉及的商品名称，而通常是用商品 1 和 2 等这样笼统的名字。我们关注的东西归根结底还是下列内容：消费者的偏好、生产技术、市场的结构等等。在分析信息的市场是否能运行时，我们还是使用上面**这些**工具。

我们在本章的主要任务是，研究和信息革命相关的几个经济模型。第一个模型是网络经济模型；第二个是转换成本模型；第三个是信息商品的权限管理(rights management)模型。这些模型将让你知道，借助于经济分析的基本工具，我们不仅能理解原子世界，也可以理解比特世界。

35.1 系统竞争

信息技术一般用于各种**系统**（systems）之中。这些系统由不同的部件组成，不同部件通常由不同企业生产。单个部件用处不大，只有合起来才有用。离开了软件的硬件是废物，DVD 播放机若没有光盘就成了一种摆设，操作系统只有和应用软件联合起来才有用。没有网络服务器，只依靠网络浏览器，你无法上网。这些都是**互补品**（complements）的例子：一种商品的存在能显著提高另外一种商品的价值。

在分析消费者理论时，我们以左鞋和右鞋作为互补品的例子。上面的这些例子也同样极端：如果没有相应的驱动软件，世界上最好的计算机硬件也无法运行。硬件和鞋子的区别之处是，能应用于该硬件身上的软件越多，它的价值就越高。

这些部件的生产企业在竞争时，必须既关注它们的竞争者，又要关注它们的“互补者”（complementors）。苹果公司在制定竞争策略时，必须考虑它与软件开发企业之间的关系。

这使得信息技术行业中的竞争策略，在某种程度上不同于传统产业的竞争策略[一]。

35.2 互补品的问题

为了说明这些观点，我们举例说明：中央处理器（CPU）和操作系统（OS）。中央处理器是一块电路板，它是计算机的“大脑”。 CPU 生产企业中比较有名的是英特尔和摩托罗拉。操作系统是能让用户启动中央处理器功能的软件。苹果和微软都是生产操作系统的公司。一般来说，不同中央处理器要求的操作系统是不一样的。

从终端用户的角度来看，只有中央处理器和操作系统兼容时，才能使用。中央处理器和操作系统是互补的，就象左鞋和右鞋是互补的一样。

目前，最流行的中央处理器和操作系统，分别是由英特尔和微软制造的。这当然是两个不同的公司，它们分别对自己的产品独立定价。还有另外一种流行的中央处理器，叫作 PowerPC，它是由 IBM、摩托罗拉和苹果联合生产的。PowerPC 通常与苹果操作系统或 IBM 的 AIX 操作系统兼容。除了这些商业操作系统之外，还有一些免费的系统例如 BSD 和 GNU-Linux，它们都是由很多编程者自愿开发和提供的。

下面我们分析互补品生产企业的产品定价问题。两种互补商品定价的重要特征是，任何一种商品的价格取决于这两种商品的价格。若 p_1 是中央处理器的价格， p_2 是操作系统的价格，那么终端用户支付的价格取决于 p_1+p_2。当然，为了让系统运行，你还需要其他的部件，但这只是让你支付的钱数更多而已；为简单起见，我们只用中央处理器和操作系统这两个部件进行分析。

中央处理器的需求取决于整个系统的价格，因此可将其写为 $D(p_1+p_2)$。若令 c_1 表示一个中央处理器的边际成本， F 为固定成本，则中央处理器生产企业的利润最大化问题为

$$\max_{p_1}\,(p_1-c_1)D(p_1+p_2)-F_1.$$

类似地，操作系统生产企业的利润最大化问题为

$$\max_{p_2}\,(p_2-c_2)D(p_1+p_2)-F_2$$

为了分析这个问题，假设需求函数为线性的

$$D(p)=a-bp.$$

为简单起见，假设边际成本非常小因此可以忽略不计。那么中央处理器企业的利润最大化问题变为

$$\max_{p_1}\,p_1[a-b(p_1+p_2)]-F_1,$$

[一] See Shapiro, Carl and Hal R. Varian, Information Rules: A Strategic Guide to the Network Economy, Harvard Business School Press, 1998, for a guide to competitive strategy in IT industries.

或

$$\max_{p_1} ap_1 - bp_1^2 - bp_1p_2 - F_1.$$

可以证明价格增加Δp_1引起的边际收入增加量为

$$(a - 2bp_1 - bp_2)\Delta p_1.$$

如果利润已实现了最大，则价格 p_1 提高而引起的收入变动必定为零：

$$a - 2bp_1 - bp_2 = 0.$$

由此可解得

$$p_1 = \frac{a - bp_2}{2b}.$$

用同样的方法，我们可以解出操作系统的利润最大化的价格：

$$p_2 = \frac{a - bp_1}{2b}.$$

注意，每个企业的价格的最优选择，取决于另外一个企业对它自己产品索要的价格。和往常一样，我们对纳什均衡感兴趣，均衡时，每个企业对另外一个企业行为的预期都得到了验证。

解上述两个未知变量的两个方程，可得

$$p_1 = p_2 = \frac{a}{3b}.$$

这个式子是说，若每个企业都是单方地、独立地制定自己产品的价格，则整个系统的价格为

$$p_1 + p_2 = \frac{2a}{3b}.$$

下面我们分析下面的实验。假设这两个企业合并成一个一体化的企业。现在这两种产品不再单独定价，而是对整个系统制定一个价格，我们用 p 表示该价格。此时的利润最大化问题为

$$\max_{p} p(a - bp)$$

价格增加Δp引起的边际收入增加量为

$$(a - 2bp)\Delta p.$$

令上式等于零，可以解出这个一体化企业对整个系统产品的定价

$$p = \frac{a}{2b}.$$

注意下列有趣的事实：一体化企业的利润最大化价格，小于两个独立企业利润最大化时的价格之和（$a/2b<2a/3b$）。由于整个系统的价格降低了，消费者能买得更多从而状况变好。而且，一体化企业的利润，大于合并钱两个独立企业均衡利润之和，企业的状况也变好了。因此，协调定价决策让交易双方的状况改善了！

可以证明下列一般事实：两个垄断企业的产品若是互补的，那么与独立制定自己的产品价格相比，企业合并后（总）价格将降低，但（总）利润会更大[一]。

上述结论符合直觉。若企业 1 考虑降低中央处理器的价格，则中央处理器和操作系统的需求会增加。但是这只考虑了降低价格对它自己利润的影响，而没有考虑其他企业的利润因此会增加。因此，与如果企业 1 关注的是联合利润最大的情形相比，它的降价幅度尚小。同样的分析也适用于企业 2。因此，两个企业独立定价时，从联合利润最大化与消费者剩余的角度来看，这个价格（之和）"太高"。

互补企业之间的关系

上述"将互补企业合并"的分析颇具煽动性，但是我们不应该直接跳到下列结论：将中央处理器生产企业和操作系统生产企业合并，是个不错的主意。这个结论其实是说，从联合利润最大化的角度来看，独立定价会使价格过高。但是，现实中的企业并不是处于完全独立或完全一体化这两个极端，处于中间情形的更多。

例如，其中一个企业可以与其他企业协商部件的价格，然后将所有部件整合一体后再销售。苹果公司或多或少就是这么做的。它从摩托罗拉公司大量购买 PowerPC 中央处理器，安装进自己生产的电脑，然后将操作系统和计算机打包销售给终端消费者。

出了上述方法外，收入分享方法也可用于解决系统的定价问题。波音（Boeing）公司和通用电气（GE）公司就是这么做的。波音生产机身，通用电气生产飞机发动机。终端用户当然想要包括机身和发动机的完整产品。如果这两家公司分别对自己的产品定价，那么价格之和就过高。因此，它们的实际做法是销售整机、分享收入。波音公司和终端客户谈判决定整机的销售价格，然后按照和通用电气的合同将整机销售收入分成一定比例给通用电气。通用电气当然愿意"坐享其成"。

在其他行业还存在着其他的定价方法。例如，前面提及的 DVD 产业。DVD 是一项非常成功的新产品，但是这个产业的运行比较复杂。DVD 播放器生产企业不愿意生产，除非它们相信市场上 DVD 光盘的供应充足；DVD 光盘生产企业也不愿意生产，除非它们相信市场上 DVD 播放器的供应充足。

除此之外，这两类企业还担心互补品的定价问题：如果市场上 DVD 播放器生产企业和 DVD 光盘生产企业的数量都很少，那么每类企业都想对自己的产品索要"过高"的价格，从而减少了行业的总利润，并且使得消费者的状况变坏。

[一] 奥古斯丁.古诺（Augustin Cournot）首先发现了这个显著的事实，我们已在第 27 章介绍过这个数学家。

索尼和飞利浦拥有 DVD 技术的主要专利，它们是怎么解决 DVD 行业的上述问题的？它们采取的措施是大幅降低专利使用费。它们还认识到要想使该行业加速发展，必须使得市场中有足够的竞争企业以便降低 DVD 产品的价格。这些策略当然是成功的，因为在成功行业中拥有较小的份额，显然要好于在根本不存在的行业中拥有较大的份额。

互补企业关系的另外一种模型，可以称为“将互补品普及化”（commoditize the complement）。回顾企业 1 的利润最大化问题：

$$\max_{p_1} \ p_1 D(p_1 + p_2) - F_1.$$

对于一组给定的价格 (p_1, p_2)，降低 p_1 能否增加企业 1 的销售收入要取决于需求价格弹性。但是降低 p_2 **总能**增加企业 1 的销售收入。问题是：企业 1 如何才能让企业 2 降低价格？

一种方法是极力使企业 2 面临更加激烈的竞争。此处可以使用的策略很多，具体使用哪一种要根据行业的性质确定。在技术密集型行业，标准化已称为一个重要的工具。例如，操作系统生产企业，希望促进硬件标准化。这样的策略不仅能使操作系统企业的生产变得更简单（无须再量体裁衣），也能保证硬件行业趋于高度竞争。竞争力会降低硬件的价格，从而降低了终端用户购买整个系统的价格，因此增加了操作系统的需求[一]。

例子：苹果公司的 iPod 和 iTunes

苹果公司的 iPod 音乐播放器非常流行。到 2009 年 1 月，苹果公司已经销售了 600 万首歌曲，占美国在线音乐销量的 70%左右和 88%的市场份额。

音乐播放器和音乐之间存在着明显的互补关系。经典的商业互补模型来自吉列公司（Gillette）：“免费送剃须刀架但卖刀片。”但在苹果公司的这个例子中，这个互补模型反过来了：苹果公司在此互补产品上的大部分利润来自销售 iPod，销售音乐得到的利润只占较小的比例。

这主要是因为这些音乐的知识产权不属于苹果公司，因此在 iTunes 上销售的音乐得到的收入，苹果公司必须与音乐的制作方分享。由于苹果公司主要从音乐播放器上赚钱，它希望音乐能便宜一些。然而由于音乐制作方主要靠歌曲赚钱，它们希望音乐能贵一些。这导致了苹果公司和音乐制作方之间存在着某些冲突。

原来，iTunes 上所有歌曲的售价都是 0.99 美元。有些音乐发行方认为新歌曲的价格应该更高一些。在多次讨价还价之后，苹果公司于 2009 年 3 月宣布了新的政策：新歌曲售价为 1.29 美元。这是一种差别定价策略，或者称为“版本控制”（versioning），这种做法在媒体市场比较流行。那些热爱音乐和缺乏耐心的人会支付较高的价格，而那些更具耐心的人会等着降价。

[一] See Brandenburger, Adam and Barry Nalebuff, Co-opetition, Doubleday, 1997 for further analysis of strategy for complementors.

例子：iPod 是谁生产的？

提示：不是苹果公司。事实上，iPods 是由几个亚洲国家的公司组装而成的，这些公司包括华硕、英资达、富士康等。

但这不是故事的全部。上述组装公司只是从其他公司买来零部件进行安装。iPod 有 451 个零部件，最近，一些经济学家试图追踪这些零部件的原产地[一]。

这几个经济学家分析的是储存量为 30G 的 iPod，它的零售价为 299 美元。这款 iPod 的最贵的部件是硬驱，它是由东芝生产，成本约为 73 美元。比较贵的零部件还有显示组件（约 20 美元）、视频/多媒体处理器芯片（8 美元）和控制器芯片（5 美元）。他们估计，由中国的公司最后组装只花费了 4 美元/台。

他们试图追踪 iPod 的主要部件的产地和每一步生产过程的附加值。在 iPod 的 299 美元零售价中，有 163 美元为美国公司和工人所占有。这 163 美元的进一步细分如下：批发和零售成本 75 美元、苹果公司 80 美元、其他美国零部件制造者 8 美元。在附加值中，日本贡献了 26 美元（主要是东芝的硬驱），而韩国的贡献不到 1 美元。

比较理想的是，每种零部件都是从成本最低的供给者手中买到的，在很大程度上，这些决策反映了不同供给者的比较优势。

尽管中国公司的组装的贡献只占 iPod 价值的 1%左右，每台 iPod 在中美双边贸易中贡献了 150 美元的美方逆差。这表明**双边**贸易逆差没有任何意义。中国在组装 iPod 时，大部分高价值的零部件需要先从其他国家进口。iPod 中价值最高的部分，即它的设计和工程技术，来自美国。

例子：关键词竞价广告和相关广告

Google 的广告项目中，有两个项目分别称为关键词竞价广告（Adwords）和感知广告（Adsense）。关键词广告是网络用户搜索查询时出现的广告。感知广告则根据网页内容显示广告。关键词广告显示的是“搜索导向的广告”，而感知广告显示的是“内容导向的广告”。

当网络用户在某个网站上点击内容导向的广告，广告客户按照广告被点击次数支付，每次点击的价格是由拍卖竞价决定的，这类似于第 17 章中介绍的情形。因广告被点击而产生的收入需要在网页发布者和 Google 之间按收入分成公式进行分享。因此，感知广告项目为网页发布者提供了一种赚取广告收入的方法，这种方法的好处之一是网页发布者不需要自己管理广告项目。

关键词广告和感知广告这两个项目之间具有很强的**互补性**（complementarity）。感知广告使得网页发布者发布内容可以赚钱，因此可以鼓励发布者发布内容。这意味着网络上有用

[一] Greg Linden, Kenneth L. Kraemer, and Jason Dedrick, “Who Captures Value in a Global Innovation Network,” Communications of the ACM, 52 (3), March 2009, 140–144.

的信息变得更多，因此能满足 Google 进行索引和搜索。由于 Google 创造了鼓励网页发布者发布内容的商业模式，它使得自己的搜索服务更有价值。

35.3 锁定

由于系统中的 IT 各种部件是协调运行的，转换（switch）任何一个部件通常需要转换其他部件。这意味着 IT 行业中，因此转换部件引起的**转换成本**（switching costs）可能非常大。例如，如果你想从一台基于 Macintosh 的计算机更换为基于 Windows 的计算机，几乎需要“从头换到脚”：你需要支付计算机得赢家成本，需要购买一整套新的软件，更重要的是，你需要重新学习一种新的系统。

当转换成本非常高时，用户发现自己被**锁定**（lock-in）了，也就是说用户处于这样的情形，当他想转换为一个新的系统时需要支付很高的成本，以至于几乎无法进行转换。这对于消费者显然是不利的，对于组成系统的部件生产企业来说却非常具有吸引力。由于被锁定的用户的需求非常**缺乏弹性**，生产企业（们）可以提高部件的价格，夺取消费者剩余。

当然，明智的消费者会极力避免被锁定，或者企业要想锁定他们也可以，但是至少他们会努力地讨价还价以争取部分补偿。即使消费者的讨价还价能力比较弱，系统销售者之间的竞争会压低**最初**的价格，因为被锁定的消费者后来能给企业提供稳定的收入流。

例如，在选择互联网服务商（Internet service provider ，ISP）的决策问题中，如果你已经选择了某个服务商，后来再转换为另外一个服务商时，就非常不方便，因为你要花费时间将你的新电子邮件（email）告诉你的亲朋好友以及同事，重新安装各种程序，等等。由于转换成本引起的垄断力，可以保证互联网服务商向你索要高于它的边际成本的价格，当然前提是你被它锁定了。但是这种效应的另一面是，由于互联网服务商们都知道被锁定的客户是价值很高的财富，因此它们事先就会为争夺客户展开竞争，例如提供各种优惠诱使消费者与其签约。

存在转换成本的竞争模型

如果存在转换成本，企业如何定价？下面我们分析这个问题。假设互联网服务商向一个客户提供服务的成本为 c 元每月。再假设市场为完全竞争的，即市场中这类服务商的数量很多，而且服务质量相仿。因此，如果不存在转换成本，则提供上网服务的价格就是 $p=c$。

现在做出以下假设：更换服务商引起的转换成本为 s 元；为吸引新客户，任何服务商现在对于第一个月给与 d 元的优惠，也就是说第一个月的服务价格为 $p-d$；在某个既定月份的开始，消费者考虑是否转换服务商。如果转换，它向服务商支付的价格为：第一个月为 $p-d$ 元，从第二个月起为 p 元。除此之外，他必须还要承担转换成本 s。如果不转换，那他必须一直向服务商支付 p 元每月（可以认为以前没有优惠活动）。

 曹乾（东南大学 caoqianseu@163.com）

消费者在什么样的条件下才会转换服务商？不难知道，如果下式成立则消费者会转换服务商:

$$(p-d)+\frac{p}{r}+s<p+\frac{p}{r}.$$

其中，r 为（月）利率。

上式是说，当支付给新服务商的钱数加上转换成本（上式左侧），小于向原服务商支付的钱数时，消费者就会转换。

但是，由于我们假设市场是完全竞争的，因此服务商之间的竞争，会使消费者在转换和不转换这两个选择之间无差异，这意味着

$$(p-d)+s=p.$$

由此可得 $d=s$，这就是说服务商提供的优惠额，恰好能补偿消费者的转换成本。

由于市场是完全竞争的，每个服务商的利润现值为零。每个服务商的利润现值，等于初始折扣加上未来利润的现值。令 r 表示（月）利率，由于 $d=s$，零利润的条件因此可以写为

$$(p-s)-c+\frac{p-c}{r}=0. \tag{35.1}$$

重新整理上式，可得到均衡价格的两种等价表达方法

$$p-c+\frac{p-c}{r}=s \tag{35.2}$$

$$p=c+\frac{r}{1+r}s. \tag{35.3}$$

（35.2）式表明，从消费者身上得到的未来利润的现值，正好等于消费者的转换成本。（35.3）式是说，服务价格等于在边际成本基础上的加价（markup），加价部分和转换成本成正比。

由此可见，在竞争模型引入转换成本后，**月度**服务价格大于边际成本，但是服务商对利润流的竞争会压低**初始**价格。实际上，服务商为了能获得未来加价部分，需要给予消费者的优惠额 $d=s$。

现实中，出了能从消费者得到每月的支付之外，很多互联网服务商还有其他收入渠道。例如，美国在线（American Online）公司收入中的很大一部分来自广告。因此，该公司为了能获得广告收入，采取的措施是，如果客户提前付款可以享受较大幅度的优惠，尽管服务价格甚至都有可能低于边际成本，也在所不惜。

我们很容易地就可以将这种效应加入上述模型。如果 a 是消费者每月产生的广告收入，零利润条件要求

$$(p-s)+a-c+\frac{p+a-c}{r}=0 \qquad (35.4)$$

解出 p 可得

$$p=c-a+\frac{r}{1+r}s.$$

这个式子表明，真正重要的是服务的**净成本** $c-a$，它等于服务的成本减去广告收入。

例子：账单在线支付

很多银行在提供账单支付服务时，索取的服务费很低或者干脆免费。有些银行甚至还向消费者支付一笔报酬，如果他们使用它们的在线账单支付服务的话。

为什么银行热衷于这样的事情？答案在于，银行发现了下列事实：如果消费者不怕麻烦安装某家银行的账单在线服务软件，那么他们一般就不会转换银行。根据一家美国银行的研究报告，消费者转换银行的频率下降了 80%[一]。

这样的事情不难明白，比如你一旦安装了这样的软件，你很难再放弃。银行为了吸引新客户，通常会将支票账户的利率提高 0.1%，但是如果你已经被一家银行锁定，这样的利率对于你来说没有什么吸引力。根据前面锁定的知识可知，银行投资于能产生转换成本的服务项目，对于它们来说是非常有利可图的。

例子：手机号码可携带

手机运营商曾经禁止个人在转换运营商后仍保留原来的手机号码。这种做法显著增加了手机用户的转换成本，因为你在转换运营商之后，需要将新号码告诉你的亲朋好友，比较麻烦。

正如本章模型指出的，消费者在面临较高的转换成本时，运营商可以对他们索要高价。但是，正因为此，运营商为了争取消费者而不得不进行更为激烈的竞争。市场上流行很多的招揽客户的花招，例如，如果你和某个运营商签约，你可以花很少钱的就可以买到一部手机，或者它干脆白送你一部手机；另外，运行商还提供“免费打电话某某分钟”、“延期缴费”、“手机亲情号码之间通话优惠”等等。

运营商在阻止号码携带（number portability）方面是齐心协力的，它们游说管理机构和议会维持现状，即消费者在转换运行商后禁止将原手机号码携带过去。

尽管缓慢但可以确定的是，潮流开始不利于手机运营商，因为消费者要求号码可携带。负责管理手机行业的美国联邦通讯委员会，开始暗示手机运行商应该考虑实施手机号码可携

[一] Michelle Higgins, "Banks Use Online Bill Payment In Effort to Lock In Customers," Wall Street Journal, September 4, 2002.

带的方法。

2003 年 6 月，Verizon 公司宣称它不再反对号码携带。它的决定可能出于两方面的考虑。第一，它知道如果仍反对客户携带号码，那么它就是在打一场注定失败的战争：手机号码迟早要允许携带；第二个因素可能更重要，因为在 Verizon 公司做出上述姿态之前，几项对消费者的调查都表明，该公司在客户满意度方面遥遥领先于对手。因此，如果它降低转换成本，它可能赢得更多的客户。事实上，似乎 Verizon 公司最终从该策略中获利匪浅。

这个小插曲对于制定商业策略的借鉴意义是：增加消费者转换成本的策略，只在一段时间内有价值。真正明智的运营商最终还是要在提高服务质量方面下功夫。

35.4 网络外部性

我们已在第 34 章介绍过外部性（externalities）的概念。我们已经知道这个概念是指，某人的消费直接影响其他人的效用的情形。**网络外部性**（Network externalities）是一种比较特殊的外部性，在这种情形下，一个人消费某商品的效用大小取决于消费这种商品的其他消费者的**数量**[一]。

以某消费者对传真机的需求为例。人们需要传真机是希望能够彼此通讯。如果其他人都没有传真机，你买传真机还有什么用？调制解调器（modems）也具有类似的性质：你的调制解调器只有在你想通讯的人也有这个设备时才有用。

互补品也具有网络外部性，虽然不是那么直接。如果某社区内的居民都没有影像播放机，谁会在这个区内开设影碟店？反过来，如果没有影碟店，你为什么购买影像播放机？在这种情形下，对影碟的需求取决于影碟播放机的数量，影碟播放机的需求取决于能得到的影碟的数量，这就是更一般形式的网络外部性。

35.5 存在网络外部性的市场

我们使用一个简单的需求和供给模型，建立网络外部性的模型。假设市场中某商品的需求者有 1000 人，将他们标记为 $v = 1,...,1000$ 。将 v 看成第 v 个人对该商品的**保留价格**，比如第 10 个人的保留价格为 10。因此，若该商品的价格为 p，那么保留价格大于或等于 p 的人数为 $1000 - p$ 。例如，若价格为 200 元，则支付意愿至少为 200 元的人数为 800 人，因此商品销售数量为 800 单位。这样，需求曲线的斜率必然为负。

但是，现在我们对该模型稍作变动。假设该商品具有网络外部性，比如传真机或电话。为简单起见，假设第 v 个人对该商品的评价为 vn，其中 n 为消费该商品的人数，也就是说共享该网络的人数。消费这种商品的人数越多，**每个人**越愿意花更多的钱购买该商品[二]。

[一] 更一般地，一个人的效用可能取决于其他人的**身份**；把这个因素加入模型中并不难。

[二] 其实应该将 n 解释为**预期**消费这种商品的人数，但是这种区别对此处的分析不重要。

此时，该模型的需求函数是什么样子的？

如果价格为 p，则存在这样的消费者，他在购买和不购买之间无差异。令 $\hat{v}$ 表示这个边际个人的标号。根据定义，他买不买该商品恰好对他是无差异的，因此，他对该商品的支付意愿等于商品的价格：

$$p = \hat{v}n. \tag{35.5}$$

由于这个“边际消费者”对买和不买是无差异的，因此对该商品的评价（v）高于 $\hat{v}$ 的那些消费者必定会购买该商品。这意味着购买该商品的人数为

$$n = 1000 - \hat{v}. \tag{35.6}$$

联立（35.5）和（35.6），我们得到了该市场的均衡条件：

$$p = n(1000 - n).$$

这个式子给出了商品价格和购买该商品的人数之间的关系。在这个意义上，它是一种需求函数；若有 n 个人购买商品，则边际消费者的支付意愿等于需求曲线的高。

然而，由图 35.1 所示的这条需求曲线的形状可知，它和标准需求需求的形状大不相同！如果共享网络的人数较少，则边际个人的支付意愿也较低，因为他能联系的人数较少。但是，如果共享网络的人数非常多，那么边际个人的支付意愿反而比较低，这是由于对该网络评价高的个人已经加入网络，这两种力量使得需求曲线呈驼峰的形状，如图 35.1 所示。

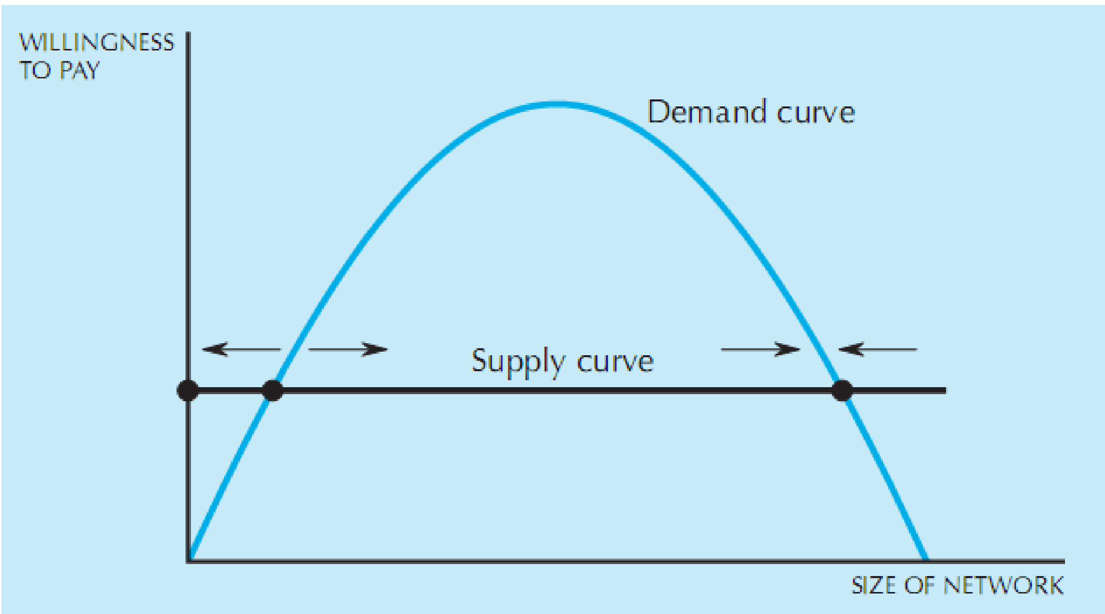


图 35.1：网络外部性。图中，需求曲线呈现驼峰形状，供给曲线为一条水平线。注意，需求曲线和供给曲线有三个交点。

我们已经知道了市场的需求，下面来看看供给。为简单起见，假设商品能以规模报酬不

变的技术供给。我们已经知道，这表示供给曲线是一条水平线，即 $p = AC$，其中 AC 表示平均成本。

注意，需求曲线和供给曲线有三个交点。$n^* = 0$ 时的均衡是一个低水平均衡，此时，没有人消费该商品（加入网络），因此无人愿意购买该商品。这个均衡可以称为“悲观预期”均衡（“pessimistic expectations” equilibrium）。

在中间那个均衡处，消费者的数量为正但是人数较少。这个均衡表明，人们认为网络不大，因此他们不愿意为加入网络而支付较高的价格，也正因为此，网络不够大。

最后的那个均衡处，人数很多，此时人数为 n_H。此时，价格较低，因为购买该商品的边际个人对该商品的评价不够，尽管市场已非常大。

35.6 市场的动态分析

上述三个均衡中的哪一个会发生？根据现有的模型信息，我们无法进行选择。在每个均衡处，需求都等于供给。然而，在模型中加入动态调整过程后就有可能找到哪个均衡更有可能发生。

假设当人们支付意愿大于商品的成本时，市场规模会扩大，反之，市场会萎缩。这个假设比较合理。在图形上，这个假设意味着，当需求曲线高于供给曲线时，商品数量上升，反之，商品数量下降。图 35.1 中的箭头表示了这个调整过程。

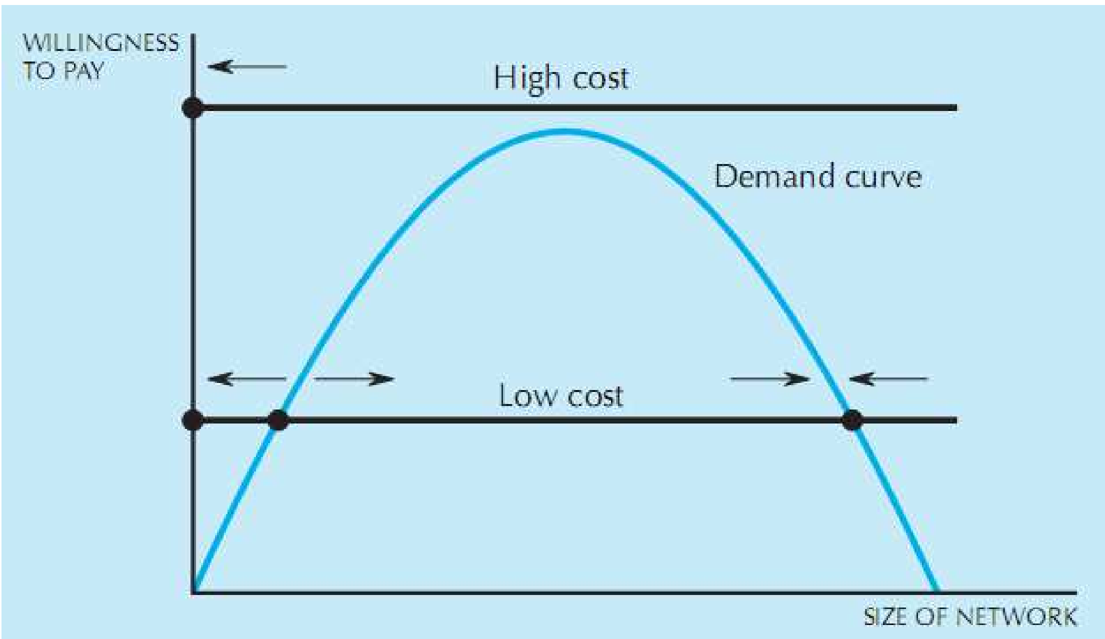


图 35.2：成本调整和网络外部性。当成本较高时，唯一的均衡意味着市场规模为零。当成本下降时，其他均衡有可能发生。

这个动态调整使我们得到了一些额外的信息。容易看出，低水平均衡（网络人数为零）和高水平均衡（网络人数很多）都是稳定均衡，但中间均衡是不稳定的。因此，最终均衡不可能是中间均衡。

现在只剩下两个可能的稳定均衡；我们如何判断出哪一个均衡会发生？一种方法是思考成本如何随时间变化。对于前文分析的传真、影像播放机和计算机网络等这类例子来说，自然可以认为，一开始成本较高，但是由于技术进步，成本会降低。这样的成本变化过程请见图 35.2。单位成本较高时，只有一个稳定均衡，即需求等于零的均衡。当成本大幅下降时，市场存在两个稳定均衡。

现在向系统中加入噪音（noise）。考虑这个噪音干扰的是 $n^* = 0$ 时加入网络的人数。这些干扰可能是随机的，也可能是商业策略的一部分——例如对新消费者给与的折扣或者其他促销措施。当成本越来越小时，干扰越有可能使得系统直接跳过不稳定均衡。当这种情形发生时，动态调整过程会促使系统到达高水平的稳定均衡。

图 35.3 给出了该商品消费者人数变动的可能路径。消费者人数一开始约等于零，随着时间推移，人数会出现一些较小的波动。随着成本逐渐降低，当降低到一定程度，就会发生质变，此时的消费者人数称为临界数量（critical mass）[一]。质变使得系统跳过低水平的均衡，并且快速达到高水平均衡。

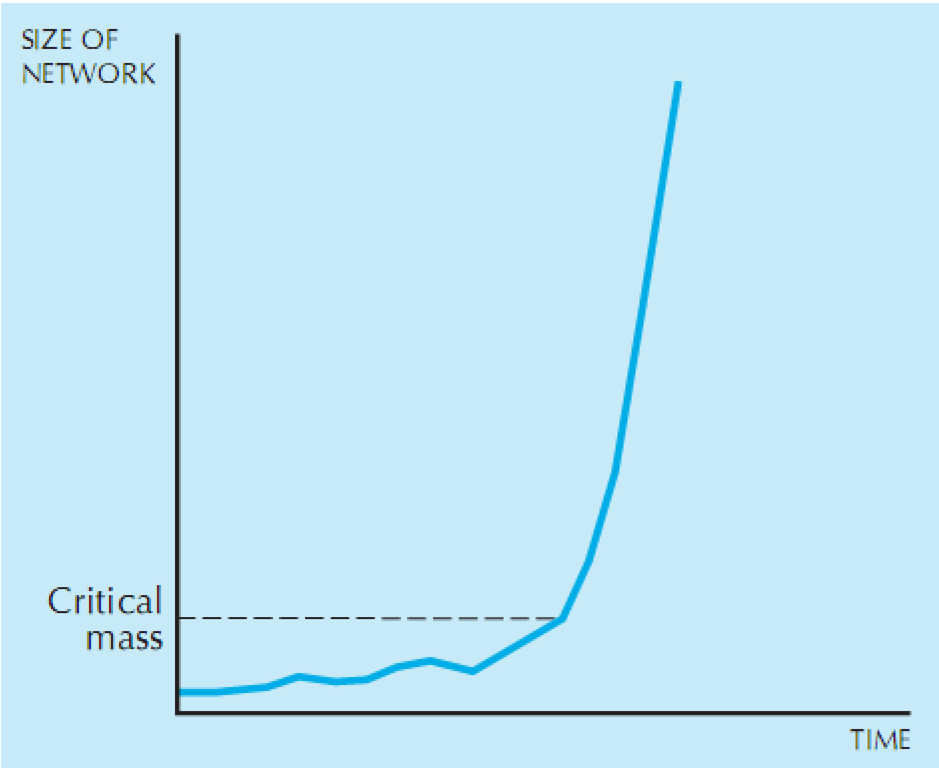


图 35.3：均衡的可能调整过程。加入网络中的人数一开始很小，随着成本下降人数开始增加。当达到临界数量后，网络开始急剧增长。

[一] 临界数量本是核反应的术语，是指能实现持续核裂变而需要的最少核裂变物质。译者注。

这类调整过程在现实中的一个例子，是传真机市场的变动。图 35.4 说明了 12 年间的传真机的价格和销售量[一]。

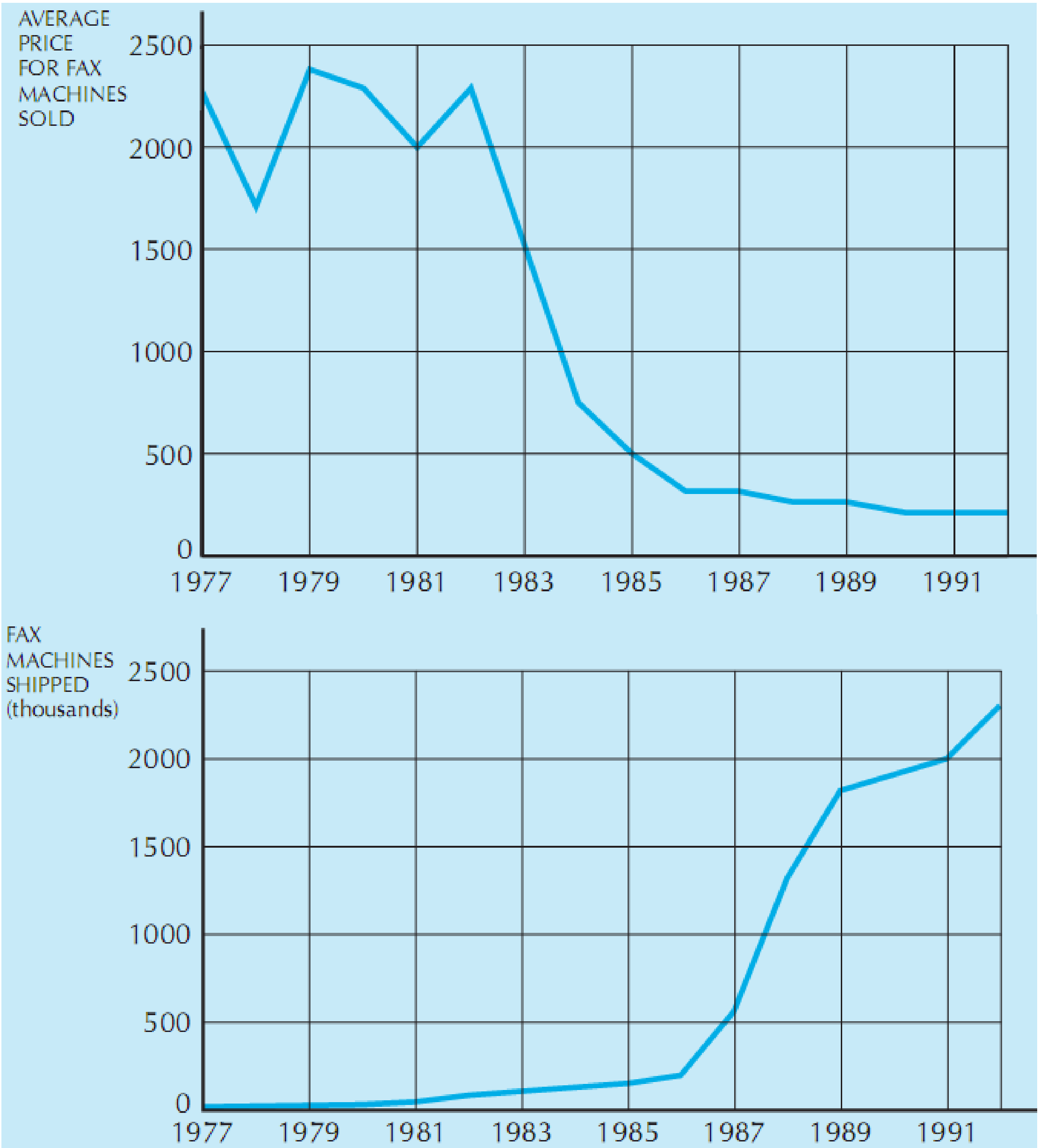


图 35.4：传真机的市场。在很长一段时间内，传真机的需求量很小，因为用户很少。在 1980 年代，当传真机价格大幅下降后，需求突然喷发。

[一] This diagram is taken from "Critical Mass and Network Size with Applications to the US Fax Market," by Nicholas Economides and Charles Himmelberg (Discussion Paper no. EC-95-11, Stern School of Business, NewYork University, 1995). See also Michael L. Katz and Carl Shapiro, "Systems Competition and Network effects," Journal of Economic Perspectives, 8 (1994), 93-116, for a nice overview of network externalities and their implications.

例子：计算机软件中的网络外部性

计算机软件的销售中自然存在着网络外部性。使用相同软件的用户可以方便地交流数据文件和其他小材料。因此，市场中谁的软件占有的份额越大，谁就越具有优势。正因为此，软件生产企业不遗余力地争夺市场。

类似的例子到处都是。例如，Adobe 公司耗费巨资研发了用于桌面出版的 PostScript 页面描述语言[一]。Adobe 认识到，没人愿意花费时间和金钱学习 PostScript 语言，除非它是“行业标准”。因此，为了创建 PostScript 解释程序的竞争市场，该企业故意允许它的竞争者“克隆”该语言。Adobe 公司的策略取得了成功：出现了若干个竞争企业，有一个企业甚至放弃了它自己的产品；PostScript 最终在桌面出版领域中得以广泛应用。Adobe 保留了几项专利，比如以低清晰度显示字体的技术等，这些专利使得该公司占据了高端市场。具有讽刺意味的是，Adobe 市场的成功却要归功于它鼓励竞争对手进入市场。

最近几年，很多软件公司纷纷仿效这种做法。Adobe 自己也放弃了几个软件产品的产权，例如 Adobe Acrobat 阅览器。1995 年美国发行的股票中有一只热门股票，这就是 Netscape 通讯公司的股票。该公司通过放弃主要产品的做法，成功获得了网络浏览器市场的最大份额。这个例子成为“每笔交易都亏损，却最终赚了钱”的最好案例[二]。

35.7 网络外部性的应用

上一节介绍的模型尽管简单，却仍提供了一些思想。例如，临界人数问题非常重要：如果一个用户的需求取决于其他用户的数量，那么在产品生命周期的早期阶段就应该激发人们对该产品的需求。现在，为了培育尚不存在的市场，很多企业愿意以较低的价格销售软件，或者提供通讯服务。这样的例子比比皆是。

当然，关键的问题是当市场多大时，它才能够自发壮大？仅依靠理论分析，我们无法给出答案。答案取决于商品的具体性质、消费者接受这种商品时面临的具体成本和收益等。

网络外部性的另外一个重要应用，体现在政府政策的作用。例如互联网。互联网起源于一些小实验室之间相互传输数据。在 1980 年代，美国国家自然基金委员会使用互联网将一些主要的大学连接起来，连接中介是分布在不同地点的 12 台大型计算机。这么做的初衷是让大学研究人员和计算机之间来回传输数据。但是通讯网络的基本特征是：一旦人们被连接在一起，那么人们两两就被连接在一起。因此，实验人员可以互相发送邮件，这和大型计算机无关。一旦加入网络的人数达到临界数量，网络对于新用户的价值就会急剧增加。大多数新用户对上述大型计算机中心根本不感兴趣，尽管这是提供网络的初衷所在。

[一] 你熟悉的 pdf 电子文档（Portable Document Format）应该归功于 Adobe 公司，由于 pdf 使用 Adobe 公司开发的 PostScript 页面描述语言，使得页面中的文字和图形的质量非常高。译者注。

[二] 类似的例子还有我国的奇虎公司，就是开发“360 安全卫士”的公司，该公司让用户免费使用一系列软件，成功占有了杀毒软件等市场的最大份额。最终该公司在美国上市。译者注。

例子：黄页

你熟悉的载有当地电话号码的黄页，代表着价值高达 140 亿美元的生意。十年以前，黄页业务主要由电话公司来做，它们占据了 95%的市场份额。现在，它们的份额只有 85%。

份额下降的原因在于竞争。近几年，几个小企业进入黄页市场，夺走了当地电话公司的部分业务。这不是个简单的事情，因为当地商业电话号码簿具有典型的网络外部性：消费者过去都使用当地电话公司提供的黄页号码簿，因此当地商人不得不在该黄页里做广告，原因是当地商人都在这里。

一个叫做黄本（Yellow Book）的公司，成功克服了网络外部性的效应，它采取的策略是大幅降低广告费、抢先出版电话号码簿。电话公司认为它们的黄页市场是安全的，对新进入者不屑一顾，等它们意识到问题的严重性时，为时已晚。在过去几年里，该行业的竞争逐渐升温。这个例子表明，即使是具有很强网络效应的行业，也难以避免竞争的压力，尤其是当行业中的原有企业过分自信时。

例子：无线电广告

1910 年，无线电这种“杀手应用程序”（killer app）被用于船与岸的通讯[一]。不幸的是，无线电通讯不是私密的，因为将接收器调整到合适频率的人都能听到这些信息。在某个时点上，大卫.索诺夫（David Sarnoff）认识到这个缺陷可能变成优点，他向人们提供“无线电音乐盒”，这种设备通过电波发送音乐。索诺夫的同事怀疑这种做法，他们说道：“无线电音乐盒毫无商业价值。谁会傻到为人人都能听到的音乐付费呢？”

索诺夫同事的质疑也不是毫无道理。尽管人们喜欢无线电广播，但是这个行业缺乏一种商业模式。他们如何赚钱？

《无线电世界》杂志举办了个竞赛，在这个竞赛中它提出了无线电广播的 5 种商业模式，人们要投票选出自己最喜欢的模式。这些商业模式有：

- 由一般税收赞助；
- 公众的捐款；
- 无线电硬件生产者资助无线电内容的生产；
- 插播广告；
- 通过征收真空管税（vacuum tube tax）来维持无线电内容的生产。

胜出者是最后一种商业模型：对真空管征税。不过另外几种商业模型至今仍在应用。英国广播公司（BBC）的广播和电视是由电视税供养的。美国国家公用无线电台（National Public Radio）是由公众的捐款维持的。然而，在大多数国家，插播广告是最流行的商业模式。

[一] 杀手应用程序原是信息领域中的一个术语，它是一个应用程序，要使用这个程序就必须去购买这个程序运行的系统。例如你要收听广播，你就要有接收器（收音机）。译者注。

1922 年，美国运营的无线电台有 30 家，卖出了 10 万台收音机。第二年，无线电台增加到 556 家，卖出了 50 万台收音机。收音机的春天到来了。

35.8 双边市场

双边市场（two-sided market）是一种特殊的网络效应。以蓝光 DVD 这样的新技术为例。我不在乎某些人拥有什么样的 DVD 播放机，因此不存在直接的网络效应，但存在一种**间接的**网络效应：蓝光 DVD 播放机的销量越大，市场供应的碟片越多；而且，市场供应的碟片越多，人们越倾向于购买蓝光 DVD 播放机。

你能想出很多类似的例子。例如新型信用卡：接受这种信用卡的商店越多，它对于消费者就越具有吸引力。另一方面，接受这种信用卡的消费者越多，它对商店越具有吸引力。

另外一个例子是 Adobe 的 PDF 平台。使用 PDF 阅览软件（Acrobat Reader）的人越多，图形设计者越愿意用这种格式发布内容，而且人们对 Acrobat Distiller 的需求越多，Acrobat Distiller 是用于创造 PDF 文档的软件。

最后这个例子说明了一个重要道理：Adobe 让人们免费使用一种产品（Reader），目的是为了诱使人们使用它的另外一种产品（Distiller）。这种商业模式和吉列公司的“免费送剃须刀架但卖刀片”是一样的，但是由于电子商品和互联网的结合让电子商品流通成本更低，这种策略变得非常流行。

再比如苹果公司卖得是 iPod 音乐播放器。它也在 iTunes 商店里销售可在 iPod 上播放的歌曲。根据行业报道，苹果在歌曲上挣钱有限——大部分利润都流入音乐歌曲制作方手里。然而，从苹果公司的观点来看，这是一种合理的商业模式：为了卖刀片（iPods），免费送剃须刀架（歌曲）。

双边市场的一个例子

为了分析双边市场，我们继续使用 34.5 节中的模型，但让它更一般化。

现在假设有两种商品。商品 1 的保留价格是 v_1，它呈现的价值（不同消费者对它的评价）为 $v_1 = 1,\ldots,1000.$ 类似地，商品 2 的保留价格呈现的价值为 $v_2 = 1,\ldots,1000.$

商品 1 的总价值取决于使用商品 2 的人数，商品 2 的总价值取决于使用商品 1 的人数，因此我们写为 $U_1 = v_1 n_2$ 和 $U_2 = v_2 n_1$。最后，假设商品 1 和 2 的价格（外生变量）为 p_1 和 p_2。（你可以将这两种商品的价格视为规模保持不变生产过程的成本。）

商品 1 的边际用户（marginal adopters）由 $\hat{v}_1 n_2 = p_1$ 决定，商品 2 的边际用户由 $\hat{v}_2 n_1 = p_2$ 决定。对商品 1 的评价大于 $\hat{v}_1$ 的消费者会购买商品 1，因此 $n_1 = 1000 - \hat{v}_1$。类似地，$n_2 = 1000 - \hat{v}_2$。

将这些式子列在一起，有

$$\hat{v}_1 n_2 = p_1$$
$$\hat{v}_2 n_1 = p_2$$
$$n_1 = 1000 - \hat{v}_1$$
$$n_2 = 1000 - \hat{v}_2$$

将（3）代入（1），将（4）代入（2）可得

$$(1000 - n_1) n_2 = p_1$$
$$(1000 - n_2) n_1 = p_2$$

首先，容易看出当 $n_1 = 0$ 且 $n_2 = 0$ 时，商品 1 和 2 的市场是均衡的。因为如果无人购买商品 1，商品 2 的价值为零；如果无人购买商品 2，商品 1 的价值为零。为了找到其它的解，我们画出这两个函数的图形。可能你已经猜到了，正如图 35.5 描述的一样，这类问题的解除了零解（原点）之外，通常还有两个解：一个解是低水平均衡，此时两种商品的销量都很小；另外一个解是高水平均衡，此时两种商品的销量都很大。

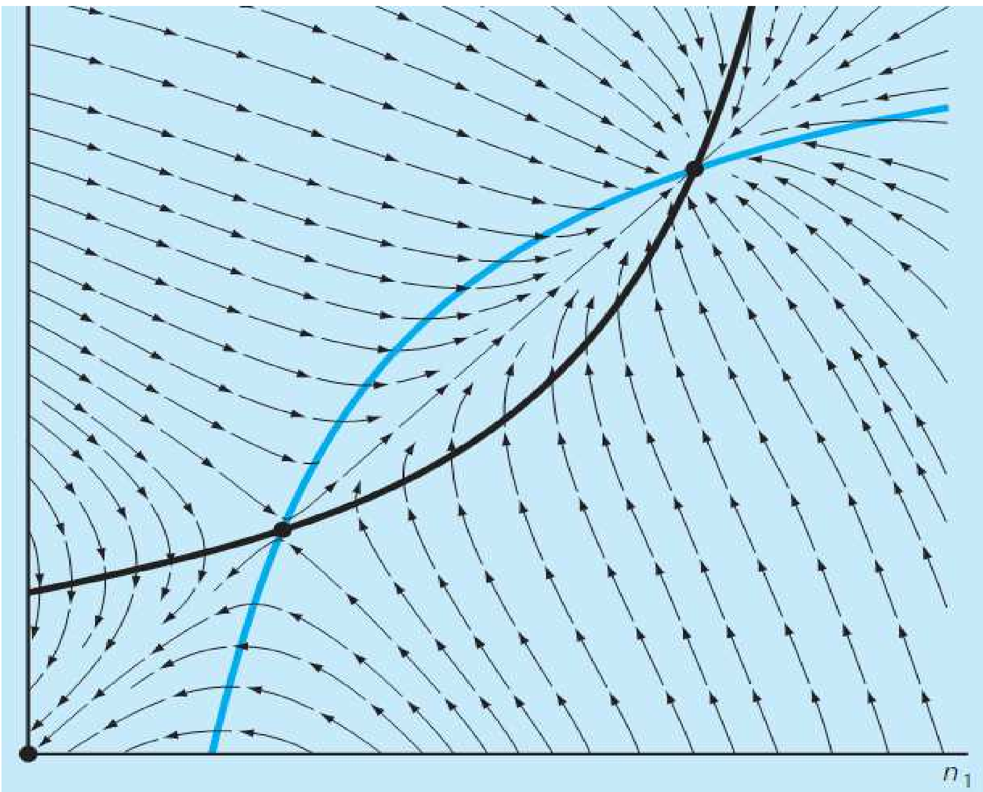


图 **35.5**：**双边市场的均衡**。双边市场通常有三个均衡解。

供给者面对的挑战是如何实现高水平的均衡。我们在前面已经知道，一种策略是对其中一种产品进行补贴。将其中一种产品以低于成本的价格销售是合理的，只要它能导致另外

一种产品的销量更大并赚取更多利润。

35.9 权限管理

近来，人们开始对知识产权（intellectual property，IP）的新商业模式感兴趣。知识产权的交易方式有多种：书籍可以直接出售或者可从图书馆借阅。影像带可以卖也可以租。有些软件授权特定的用户使用，有些软件则直接售出。共享软件比较特别，你愿意出多少钱就出多少钱，全凭自愿。

授权别人使用某知识产权时应该做出什么样的限制性规定，这是一个很重要的商业决策。你应该禁止别人拷贝吗？你是否应该鼓励用户之间分享新闻视频?你是直接卖给个人还是按区域授权？

使用一些经济学知识可以帮助我们理解这些问题。以某种纯粹电子商品为例，比如在线报纸，这样做的好处是它们的边际成本为零。我们首先分析某些默认条件下的行为。电子商品的卖方选择的价格要能使得利润最大化，由于价格是产量的函数，因此这意味着卖方也是在选择产量，卖方的利润最大化问题为：

$$\max_{y} \ p(y)y \tag{35.7}$$

这个问题的最优解为(p^*, y^*)。

现在卖方考虑放松限制条件：比如降免费试用期从一星期延长到一个月。这将对需求曲线产生两种效应。首先，它提高了产品对每个潜在用户的价值，需求曲线会向上移动；其次，它有可能减少销量，因为对于某些消费者来说，一个月的试用期已经够用了。

我们为上述情形建立模型：假设新的消费量为$Y = by$，其中$b > 1$；新的需求曲线为$P(Y) = ap(Y)$，其中$a > 1$。则新的利润最大化问题为：

$$\max_{Y} \ p(Y)y .$$

注意，我们将价格$p(Y)$乘以销量y而不是消费量Y。

将$Y = by$和$P(Y) = ap(Y)$代入上式，可得：

$$\max_{Y} \ ap(Y)\frac{Y}{b} = \max_{Y} \ \frac{a}{b} p(Y)Y = \frac{a}{b} \max_{Y} \ p(Y)Y .$$

这个最大化问题和（35.7）很像，除了目标函数之前有常数a/b之外。这不会影响到最优选择，因此我们可以断定$Y^* = y^*$。

通过这个简单的分析，可以得到以下几个结论：

- 商品的消费量Y^*和企业是否改变限制条件无关。
- 商品的产量y^*/b小于y^*。

- 利润可能上升，若$a/b>1$；利润也可能下降，若$a/b<1$。也就是说，如果从购买产品的消费者身上增加的收入，大于因消费者数量减少而造成的损失，则利润会增加；反之，利润会下降。

例子：影像带租赁

影像带租赁店可以制定租赁的限制条件。你持有影像带的时间越长，你得到的价值越大，因为你可以看更长的时间。但是这样做，租赁店的利润就会越少，因为它无法将录像带再租给别人。制定影像带最优租赁期限需要权衡这两种效应。

在实践中，这两种效应导致了产品差异化。新上线的影像带租赁期限较短，因为这样可以增加既定时期内的租赁次数，从而大幅增加利润。老影像带的租赁期限较长，因为这种影像带的租赁需求较小，即使顾客暂时租不到，引起的损失也较小。

35.10 共享知识产权

知识产权通常是共享的。例如，图书馆的好处是方便共享书籍。影像店帮助人们“分享”影像带，当然租赁者要支付一定费用。馆际互借使图书馆之间能共享图书。即使是教科书——比如你手头的教科书——也可以通过二手书市场分享。

出版界和图书馆界对共享的作用争议不休。图书馆馆长们已为馆际互借订立了非正式的“五次规则”：一本书最多借处五次，超过五次就要向出版商支付额外的特许使用金。出版商和作者通常对图书的二手市场不感兴趣。

数字信息的出现使得这种情形更为严重。数字信息可以被完全复制，几乎达到完全被“共享”的程度。近来，一位著名乡村音乐歌手发起运动抗议商店销售二手 CD。问题在于 CD 反复播放也不会变质，因此完全可能出现下列情形：某人买片新 CD、刻录后留下刻录盘，将买来的 CD 卖给二手 CD 商店。

下面我们建立描述这类共享现象的模型。我们首先分析不存在分享的情形。在这种情形下，影碟生产者选择产量y使利润最大化：

$$\max_{y}\ p(y)y-cy-F. \tag{35.8}$$

其中，$p(y)$为反需求函数，c为（固定不变的）边际成本，F为固定成本。令利润最大化的产量为y_n，下标n表示“不存在共享”。

现在假设允许影碟租赁。在这种情形下，你应该将影碟的**观看量**和影碟的**产量**区分开来。如果影碟的产量为y，而且每个影碟都被租赁k次，则影碟的观看量为$x=ky$。（为简单起见，假设每个影碟都被租赁k次。）

假定消费者在分享影碟时是人以群分的。最简单的假设是，按消费者对影碟评价高低

进行分类，评价较高的人在一起分享，评价较低的人在一起分享。也就是说，评价最高的前 k 个消费者分成一类；然后再向下排，再排出 k 个消费者，分成一类；依次类推（也可以做出其他假设，但是这种假设易于分析。）

如果影碟的产量为 y 个，则观看量为 $x=ky$，因此，边际消费者的支付意愿为 $p(x)=p(ky)$。然而，租赁影碟显然没有购买影碟观看起来方便，所以，租赁影碟会有“不方便的成本”，将这种交易成本记为 t，因此边际消费者的支付意愿变为 $p(x)-t$。

回顾前面，我们曾假设所有影碟都被 k 个消费者分享。因此，**影碟店**的支付意愿等于 k 乘以边际消费者的支付意愿。也即是说，如果影碟的产量为 y 个，则影碟店的支付意愿为

$$P(x)=k[p(ky)-t]. \tag{35.9}$$

(35.9)包含因分享而产生的两种重要效应：一是，边际支付意愿**下降**了，因为观看量大于产量；二是边际支付意愿也**上升**了，因为一片影碟可以在 k 个消费者之间分享。

影碟生产者的利润最大化问题因此变为

$$\max_y P(y)y-cy-F,$$

此式可以写为

$$\max_y k[p(ky)-t]y-cy-F,$$

或

$$\max_y p(ky)ky-(\frac{c}{k}+t)ky-F.$$

而 $x=ky$，我们因此可以将上式写为

$$\max_y p(x)x-(\frac{c}{k}+t)x-F.$$

注意，这个最大化问题和（35.8）相同，除了成本现在为 $(c/k+t)$ 而不再是 c 之外。

将上式与（35.8）式进行比较，可知：**当且仅当**

$$\frac{c}{k}+t<c \text{ 时}$$

允许租赁比不允许租赁的利润高。

将这个条件变形可得，

$$(\frac{k}{k+1})t<c.$$

当 k 较大时，上式左侧括号内的分数约等于 1。因此，这个条件是说，租赁的不方便成本 t 小于企业的边际生产成本 c 。

如果生产成本较大，而且租赁的不方便成本较小时，影碟生产企业应该生产影碟、以较高的价格卖给租赁店，然后让消费者租赁。另外一方面，如果租赁的不方便成本大于生产成本，那么生产企业应该禁止租赁：因为消费者租赁影碟非常不方便时，影碟租赁店就不会愿意花高价购买影碟。这种情形下，影碟生产企业将影碟直接卖给消费者，利润会更大。

例子：在线双边市场

互联网上的在线市场的例子并不罕见。例如，eBay 为收藏品的买卖双方提供了交易市场。比如你打算出售你珍藏的稀有硬币，你希望在具有很多潜在买者的市场上销售。类似地，如果你是个买者，你也希望到有若干卖者竞争的市场上购买。这种双边网络效应倾向于导致这样的市场只有一个。在最近几年，eBay 的业务范围已从收藏品市场扩展到很多商品的市场。

另外一组例子是网络社交网站，例如 Facebook, Myspace, LinkedIn 等等。人们希望到他们的朋友已经注册的网站注册。这又导致了网络效应：最大的网络吸引最多的新注册者。

Facebook 增速惊人。根据它的网站上披露的数字可知：Facebook 于 2004 年 2 月上线，到当年年底，活跃用户为 100 万人；到了 2009 年，活跃用户数已超过了 3 亿人，而且用户遍及世界各地。

总结

1.因为信息技术（不同部件）在系统内一起发挥作用，所以消费者转换任何一个部件都会造成很高的转换成本。

2.如果生产互补品的两个垄断企业协调定价，则互补品的总价格就会低于单独定价时的价格之和。

3.协调定价会使这两个垄断者的利润更大，也会使得消费者的状况更好。

4.协调的方法有多种，例如合并、协商、收入分成和使互补品普及化。

5.在一个锁定均衡中，第一期给与消费者的优惠额，将由后来的提价得到补偿。

6.网络外部性是指某人对某商品的支付意愿取决于消费该商品的其他消费者的数量。

7.网络外部性模型通常有多个均衡解。最终结果（最终均衡是哪一个）取决于行业的发展历程。

8.权限管理需要在下列两种效应之间做出权衡：一是价格上升带来的收入增加；二是销量减少带来的收入减少。

9.信息商品，例如书籍和影碟，通常可售可租。哪一种方式利润更大，取决于交易成本（租赁的不方便成本）与生产成本的大小关系。

复习题

1.如果消费者转换长途电话公司的转换成本大约为 50 元，那么为了获得新客户，长途电话公司愿意支付多少钱？

2.说明文字处理软件包如何表现出网络外部性的，即市场上有很多种文字处理软件包，消费者会选择哪一种呢？

3.假设生产影碟的边际成本为零，租赁影碟的不方便成本（即交易成本）。那么影碟生产企业应该选择将影碟直接卖给消费者，还是将影碟卖给租赁店然后让它租赁给消费者？

复习题答案

1.如果消费者转换长途电话公司的转换成本大约为 50 元，那么为了获得新客户，长途电话公司愿意支付多少钱？

【复习内容】存在交换成本时的竞争模型

为简单起见，做出如下假设：（1）市场是完全竞争的；（2）客户已被某个电话服务商锁定一段较长的时间，因此他每个月为公司贡献的收入流为 p 元；（3）任何一个服务商在第一个月索要的价格为 $p-d$ 元（即优惠 d 元），以后各月都按 p 元收取；（4）若转换服务商，转换成本为 s 元。

消费者在下列条件下会转换服务商：

$$(p-d)+\frac{p}{r}+s<p+\frac{p}{r}$$

其中，r 为（月）利率。

由于假设市场是完全竞争的，因此服务商之间的竞争，会使消费者在转换和不转换这两个选择之间无差异，这意味着

$$(p-d)+s=p.$$

由此可得 $d=s$ ，这就是说服务商提供的优惠额，恰好能补偿消费者的转换成本。

由于市场是完全竞争的，每个服务商的利润现值为零：

$$(p-s)-c+\frac{p-c}{r}=0.$$

整理可得：

$$p-c+\frac{p-c}{r}=s$$

上式表明，从消费者身上得到的未来利润的现值，正好等于消费者的转换成本。

【参考答案】

长途电话公司愿意支付 50 元。

由上面的分析可知，存在转换成本的情形下，公司从消费者身上得到的未来利润的现值，正好等于消费者的转换成本。所以，为了获得一个新客户，公司愿意支付的代价为 50 元。

2.说明文字处理软件包如何表现出网络外部性的，即市场上有很多种文字处理软件包，消费者会选择哪一种呢？

【复习内容】网络外部性

网络外部性是指，一个人消费某商品的效用大小取决于消费这种商品的其他消费者的数量。

【参考答案】

文字处理软件包的销售中存在着网络外部性。使用相同软件的用户可以方便地交流数据文件和其他小材料。因此，哪种软件包的使用者越多，消费者越愿意购买哪种软件包。

3.假设生产影碟的边际成本为零，租赁影碟的不方便成本（即交易成本）。那么影碟生产企业应该选择将影碟直接卖给消费者，还是将影碟卖给租赁店然后让它租赁给消费者？

【复习内容】知识产权共享

从教材中的推导可知，当且仅当 $\frac{c}{k}+t<c$ 时，允许租赁比不允许租赁的利润高。其中 c 为企业生产的边际成本，t 为租赁影碟与购买影碟相比而产生的不方便成本（交易成本）,k 为影碟的租赁次数。

将这个条件变形可得，$(\frac{k}{k+1})t < c$.当k较大时，上式左侧括号内的分数约等于 1。因此，这个式子最终变为$t < c$。

这就是说，当租赁的不方便成本t小于企业的边际生产成本c，企业应该允许租赁。

【参考答案】

由于两种方式利润相等，因此企业可以任意选择任何一种方式。

由前面复习内容可知，如果生产的边际成本较大，而且租赁的不方便成本较小时，影碟生产企业应该生产影碟、以较高的价格卖给租赁店，然后让消费者租赁。另外一方面，如果租赁的不方便成本大于生产成本，那么生产企业应该禁止租赁：因为消费者租赁影碟非常不方便时，影碟租赁店就不会愿意花高价购买影碟。这种情形下，影碟生产企业将影碟直接卖给消费者，利润会更大。

当然，如果边际成本和租赁影碟的不方便成本（交易成本）相等时，企业在上述两种方式之间无差异，即利润是相等的。

曹乾●经济学译丛精品系列

Intermediate Microeconomics:

A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

36.公共物品（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

36 公共物品

在第 34 章，我们认为对于某些类型的外部性，无效率的消除并不困难。例如，在两个人的消费外部性的例子中，你要做的事情仅是明确规定初始财产权。这样，这两个人就向交易其他商品一样，对产生外部性的权利进行交易。至于生产外部性，市场本身就提供了解决财产权问题的最有效方法，因为市场可传递利润信号。对于公共财产来说，将财产权赋予某些人就能消除无效率。

不幸的是，并非所有外部性都能通过这种方式解决。只要外部性涉及两个人以上，问题将变得非常困难。例如，上一章抽烟的例子涉及两个人，现在假设**三个**人同居一屋，一人抽烟两人不抽。烟雾对于两个非抽烟者来说是负外部性。

假设财产权明晰，比如非抽烟者有权享受清洁空气。同以前的例子一样，尽管他们拥有清洁空气的**权利**，他们也有出卖部分清洁空气以换取适当补偿的权利。但现在有个新问题——两个非吸烟者需要商讨决定烟雾浓度和补偿金额。

也许一个人比另一个人对烟雾更敏感，或者一人比另一人更富有。尽管他们的偏好和钱数都不同，他们还是要协商决定烟雾的数量。

不再考虑上述室友，我们考虑一国的所有居民。在该国应允许的污染数量为多少？如果你认为仅三个室友达成协议已够困难，想象一下几百万的居民达成协议的困难程度！

上述抽烟的外部性（涉及三个舍友）是**公共物品**（public good）的一个例子。公共物品是指对所有受影响的消费者必须提供相同数量的物品。在抽烟的例子中，抽烟产生的烟雾数量对于所有消费者是相同的，尽管每个人对此物品评价不同，但都必须面对同样的数量。

很多公共物品由政府提供。例如，大街和人行道由当地政府提供。每个城镇都有某些数量和质量的街道，每个人都可使用。国防是另外一个绝佳的例子，一国之内所有居民面对的是国防水平是同一的。居民对国防服务的评价可能不同，有些人希望多点有些人希望少点，但提供给他们的国防数量是相同的。

公共物品是消费外部性的一个特例：每个人必须消费相同数量的该种物品。这种外部性非常麻烦，因为经济学家热衷的分散市场解决方案（decentralized market solutions），对于公共物品的分配来说效果不佳。人们不能购买不同数量的国防，他们必须设法确定共同的数量。

我们首先研究的问题是公共物品理想数量的确定。接下来我们将探讨某些评价方法，以便使用这些方法对公共产品作出社会决策。

36.1 何种条件下提供公共物品？

我们以简单的例子开始分析。假设有两个室友，1 和 2。他们商讨决定是否购买电视。由于受公寓大小所限，电视只能放在客厅。两人都可看电视，因此电视是公共物品而不是私人物品。问题在于他们购买电视是否值得？

我们用 w_1 和 w_2 分别表示该二人的初始财富；g_1 和 g_2 表示为买电视各自所出资金；x_1 和 x_2 表示各自所剩资金，用于私人消费。预算约束为：

$$x_1 + g_1 = w_1$$
$$x_2 + g_2 = w_2.$$

再假设电视的成本为 c 元钱，因此为买得电视，二人所出资金之和应该至少为 c:

$$g_1 + g_2 \geq c\,.$$

该式表示了提供公共产品的可行技术（technology available）：该二人可获得电视，若他们合力支付了电视成本 c。

第 1 人的效用函数取决于他的私人消费 x_1 和是否得到电视这种公共物品。用 $u_1(x_1, G)$ 表示第 1 人的效用函数，其中 G 的取值为 1 或 0，分别表示有或无电视。第 2 人的效用函数为 $u_2(x_2, G)$。每个人的私人消费都有下标，分别表示各自的私人消费；但公共物品没有下标，因为它由二人共同“消费”。当然，这里的“消费”不是指把电视“吃掉”，而是指电视为二人提供的娱乐服务。

他们对电视服务的评价可能大不相同。我们可以询问每个人愿意出多少钱购买电视，借此估算他对电视的评价。做此事要用到第 15 章介绍的**保留价格**（reservation price）概念。

第 1 人的保留价格是指为购买电视他所愿意出的最大钱数，即它是指这样的价格 r_1，在此价格下，购得电视（花掉 r_1 元）和不买电视（持有 r_1 元），这两种选择的效用对他来说一样大。若第 1 人支付了保留价格买得电视，他还剩下 $w_1 - r_1$ 元钱用于私人消费。若不买电视，他有 w_1 元用于私人消费。因为上述两种选择提供给他的效用一样大，我们有：

$$u_1(w_1 - r_1, 1) = u_1(w_1, 0)\text{。}$$

这个等式给出了第 1 人保留价格的定义：为购买电视他愿意支付的最大金钱数。类似地，可以用等式定义第 2 人的保留价格。注意，每个人的保留价格通常取决于他的财富大小：个人**愿意**支付的最大数额在某种程度上依赖于他**能够**付得起多少钱。

回想一下：某种配置，若再无他法使两人的状况变得更好，则是帕累托有效率的；相反，若仍存在可使两人状态更好的方法，则目前的配置是帕累托**无效率**的。在帕累托无效率配置的情形，我们说仍存在**帕累托改进**（Pareto improvement）的可能。在上述电视的问题中，我们感兴趣的只有两种配置方式：一是不提供电视，这种配置采取的形式为 $(w_1, w_2, 0)$，即每个人将财富仅用于各自的私人消费；另外一种是提供电视这种公共物品，

该配置的形式为$(x_1, x_2, 1)$，其中

$$x_1 = w_1 - g_1$$

$$x_2 = w_2 - g_2 .$$

这两个等式是预算约束的变形。它们的意思是说，每个人的私人消费等于初始财富减去各自花在电视机上的钱。

在什么条件下应该提供电视？也就是说，在什么条件下可以找到一种支付方案(g_1, g_2)，使得在该方案下两人合力购买电视比不买电视要好？用经济学的语言表达，即何种条件下，供给电视是一种帕累托改进？

配置方式$(x_1, x_2, 1)$即是一种帕累托改进，若两人在购买电视后状况比不买电视要好。这意味着

$$u_1(w_1, 0) < u_1(x_1, 1)$$

$$u_2(w_2, 0) < u_2(x_2, 1) .$$

使用保留价格的定义以及预算约束可得

$$u_1(w_1 - r_1, 1) = u_1(w_1, 0) < u_1(x_1, 1) = u_1(w_1 - g_1, 1)$$

$$u_2(w_2 - r_2, 1) = u_2(w_2, 0) < u_2(x_2, 1) = u_2(w_2 - g_2, 1) .$$

看一下上述不等式的左右两端，并且记住私人消费增加则效用也增加，我们可以推知

$$w_1 - r_1 < w_1 - g_1$$

$$w_2 - r_2 < w_2 - g_2 .$$

这意味着

$$r_1 > g_1$$

$$r_2 > g_2 .$$

若配置方式$(w_1, w_2, 0)$是帕累托无效率的，上述条件必须满足，即为买电视各人实际支付的资金必须小于各自的支付意愿，这时购买电视才对他有利。因此，若保留价格（最大支付意愿）大于各人实际支付额，就存在帕累托改进余地。这两个条件显然为购买电视是帕累托改进的**必要条件**（necessary condition）。

若每个人的支付意愿大于实际支付额，则支付意愿**之和**必然大于电视的成本：

$$r_1 + r_2 > g_1 + g_2 = c \qquad (36.1)$$

这个条件，是提供电视为帕累托改进的**充分条件**（sufficient condition）。若该条件可满足，则存在某种支付方案，使得提供公共物品可以改善二人的状况。若$r_1 + r_2 \geq c$，则二人支付意愿之和至少和电视成本一样大，因此他们很容易找到一种支付方案(g_1, g_2)，使得

$r_1 \geq g_1$，$r_2 \geq g_2$ 以及 $g_1 + g_2 = c$。这个条件是如此简单，也许你会问为何要详细地推导它。好吧，我来告诉你这是因为这个条件有一些微妙之处。

首先，要注意到，提供公共物品为帕累托改进的条件，仅取决于每个人的支付**意愿**以及公共物品的成本。若保留价格之和超过了电视的成本，总存在某种支付方案，使得两人拥有电视比不拥有电视更好。

其次，提供公共物品是否为帕累托有效率，通常取决于财富的初始配置情况 (w_1, w_2)，因为保留价格通常取决于财富的分配状况。完全有可能存在某种财富分配状态 $r_1 + r_2 > c$，和另外的财富分配状态 $r_1 + r_2 < c$。

为了搞明白不同财富分配状态的后果，我们假设一人很喜欢电视另外一人无所谓。于是如果喜欢电视的人拥有所有财富，他愿意单独购买电视。此时，提供电视就是帕累托改进。但如果对电视无所谓的人拥有所有财富，则喜欢电视的人无力出钱购买电视，此时**不提供**电视是帕累托有效率的。

由此可见，是否应该提供公共物品取决于财富的分配状况。但在某些具体的情形下，公共物品的提供与财富分配状况无关。例如，假设两人的偏好都是拟线性的，即他们的效用函数的形式如下

$$u_1(x_1, G) = x_1 + v_1(G)$$

$$u_2(x_2, G) = x_2 + v_2(G).$$

其中 G 的取值为 1 或 0，取决于电视的有或无。为简单起见，假设 $v_1(0) = v_2(0) = 0$，即没有电视时，电视提供的效用为零[一].

在该情形下，保留价格的定义变为

$$u_1(w_1 - r_1, 1) = w_1 - r_1 + v_1(1) = u_1(w_1, 0) = w_1$$

$$u_2(w_2 - r_2, 1) = w_2 - r_2 + v_2(1) = u_2(w_2, 0) = w_2\text{。}$$

这意味着保留价格为

$$r_1 = v_1(1)$$

$$r_2 = v_2(1).$$

由此可见，保留价格和财富的数额无关，因此公共物品的最优供给数量和财富无关，至少在某些财富区间内情形如此[二]。

[一] 也许看电视的效用应该赋予负值。

[二] 对于这些财富区间还要进一步限制，因为我们必须要求 $r_1 \leq w_1$ 和 $r_2 \leq w_2$，即支付意愿不能大于支付能力。

36.2 私人提供公共物品

我们已经看到，若两人的支付意愿之和大于公共物品的成本，购买电视就是帕累托有效率的。这个条件是公共物品配置的效率条件，但这并不意味这他们会真正决定购买电视。他们是否决定购买电视取决于他们采用何种方法以达成共同决策。

若这两人合作并真实地回答他们对电视的评价，那么他们不难决定是否购买电视。但在某些情形下，他们没有激励如实报出各自对电视的评价。

例如，假设两人对电视的评价相等，并且每人的保留价格均大于电视的成本，即 $r_1 > c$ 和 $r_2 > c$ 。第 1 人可能认为，如果他说他对电视的评价为 0，另外一人无论如何都会购买电视。但第 2 人也会同样推理！你可以想象出这样的情形，即两人均拒绝合伙购买电视，因为每个人都期望对方首先沉不住气，从而单独购买电视。

在该情形中，经济学家说每个人都希望**搭**对方的**便车**（fee ride）:每个人都希望对方单独购买电视。因为只要有人购买电视，其他人都可以坐享其成，所以每个人都有尽可能少出钱的激励。

36.3 搭便车

搭便车的情形和第 28 章中介绍的囚徒困境类似，但并不完全相同。为了说明这一点，我们对上述电视问题构造一个数值例子。假设每人的财富均为 500 元，每人对电视的评价均为 100 元，电视的成本为 150 元。因为保留价格之和超过了成本，购买电视是帕累托有效率的。

假设一方无法禁止另一方看电视，而且购买电视与否的决策由各人独立作出。考虑其中一人即选手 A 的决策。如果他购买电视，其收益为 100 元成本为 150 元，净收益为－50 元。然而，如果选手 A 购买了电视，选手 B 可以免费观看，从而 B 的收益为 100 元。表 36.1 描述了该博弈的得益（payoffs）。

表 36.1：搭便车的博弈矩阵

		Player B	
		Buy	Don't buy
Player A	Buy	–50, –50	–50, 100
	Don't buy	100, –50	0, 0

该博弈的优势策略均衡（dominant strategy equilibrium）是两方都不买电视。若选手 A 购买，则搭便车对选手 B 有利：可以看电视但不用出钱买。若选手 A 决定不买，对 B 有利的策略也是不买。这和囚徒困境相似，但并非完全相同。在囚徒困境中，使选手效用之和最大的策略是双方作出相同的决策。在搭便车的博弈中，使选手效用之和最大的策略是仅让一方购买电视（且双方都可以看电视）。

若选手 A 买而且两人都可以看电视，我们可以构建一个帕累托改进，让 B 向 A“单方支付”即可。例如，若选手 B 给 A51 元，则 A 买电视会让两人的状况都变好。更一般地，只要选手 B 补偿给 A 的金额大于 50 元（但不大于 100 元），都是帕累托改进。

事实上，这也可能是实践中的做法：购买电视时每人多少出些钱。这个公共物品问题相对容易解决，但家务活分担的搭便车问题可能难以解决。例如，谁来打扫客厅？每个人都希望客厅干净也愿意尽责。但是每个人都想搭其他人的便车，最终没人打扫客厅，这就是客厅通常凌乱的原因。

如果事情涉及两个人以上，情形可能更糟，因为可以搭更多人的便车！从你个人的观点看，让其他人做事对你最有好处，但从社会的观点看，它是帕累托无效率的。

36.4 公共物品的不同水平

在上述的例子中，我们的决策是二选一的：要么提供电视要么不提供。但是当公共物品的**数量**可以选择时，搭便车的现象也会发生。假设，两个室友必须决定应该在电视上花多少钱。花钱越多，买到的电视质量越好。

和前文一样，令 x_1 和 x_2 表示每个人的私人消费，g_1 和 g_2 表示购买电视时每个人所出资金。现在令 G 表示他们所买电视的“质量”，令质量的成本函数为 $c(G)$，这表示若两人所买电视质量为 G，他们必须支付 $c(G)$ 元钱。

两人面对的约束是，他们花费在公共物品和私人消费的金钱总和，必须等于他们拥有的金钱数：

$$x_1 + x_2 + c(G) = w_1 + w_2 .$$

一个帕累托有效率的配置是，给定消费者 2 的效用水平，消费者 1 的效用达到最大的那个财富配置。若将消费者 2 的效用固定为 $\bar{u}_2$，该问题可以写为

$$\max_{x_1, x_2, G} \quad u_1(x_1, G)$$

使得：

$$u_2(x_2, G) = \bar{u}_2$$

$$x_1 + x_2 + c(G) = w_1 + w_2$$

可以证明该问题的最优化条件是，私人物品和公共物品边际替代率绝对值**之和**，等于额外供给一单位公共物品的边际成本：

$$|MRS_1| + |MRS_2| = MC(G).$$

或者，将边际替代率的定义写出，上式即为：

$$\left|\frac{\Delta x_1}{\Delta G}\right| + \left|\frac{\Delta x_2}{\Delta G}\right| = \frac{MU_G}{MU_{x_1}} + \frac{MU_G}{MU_{x_2}} = MC(G).$$

为了明白为何上式是正确的效率条件，我们使用惯用的技巧，想想如果违背了这个条件将出现什么样的情形。例如，假设边际替代率绝对值之和小于边际成本：比如令 $MC = 1, |MRS_1| = 1/4$ 和 $|MRS_2| = 1/2$。我们需要说明存在让这两人状况更好的方法。

给定上述边际替代率，对于第 1 人，减少 1 元钱的公共物品，他愿意接受 1/4 元钱的私人物品（因为两类物品的成本都是每单位 1 元钱）。类似地，减少第 2 人 1 元钱的公共物品，只要补偿他 1/2 元钱的私人物品，他也是愿意的。假设我们减少公共物品的数量并相应对这两人进行补偿。减少 1 单位公共物品需要补偿他们 3/4 元（=1/4+1/2）。但减少 1 单位公共物品节省了 1 元，补偿后我们还剩下 1/4 元，可以将这笔剩下的钱分配给这两个人，从而使他们的处境更好。

类似地，若边际替代率绝对值之和大于 1，可以增加公共物品的数量以改善该两人的处境。若 $|MRS_1| = 2/3$ 和 $|MRS_2| = 1/2$，这意味着为了多得到 1 单位公共物品，第 1 人愿意放弃 2/3 元的私人消费，第 2 人愿意放弃 1/2 元的私人消费。但若第 1 人放弃 2/3 单位和第 2 人放弃 1/2 单位，由此节省的钱已足够生产额外 1 单位公共物品，因为公共物品的边际成本为 1 元钱。将剩下的钱返还给这两个人，他们的状况因此变得更好。

这个帕累托有效率的条件是什么意思？一种解释方法是，将边际替代率视为人们对额外一单位公共物品的**边际**支付意愿。则效率条件仅仅是说边际支付意愿之和必须等于公共产品的边际成本。

对于离散商品来说，只有供给和不供给两种情形。因此，效率条件是边际支付意愿之和应至少等于公共物品的成本。对于涉及不同水平的公共物品来说，效率条件为，在公共物品最优供给数量之处，**边际**支付意愿应该**等于**公共物品的边际成本。若边际支付意愿之和大于边际成本，则应该多供给公共物品。

下面我们比较公共物品和私人物品的效率条件有何不同。对于私人物品，每个人的边际替代率必须等于边际成本；对于公共物品，边际替代率之和必须等于边际成本[一]。对于同一私人物品，每个人可以消费不同的数量，但他们对该物品的边际评价必须相同，否则他们就会进行交易。对于公共物品，每个人必须消费共同的数量，但他们对公共物品的边际评价

[一] 作者在此处有些不严谨，一般来说（两种商品均为“好的”商品），边际替代率为负。因此，本章中的边际替代率均应理解为边际替代率的绝对值，事实上本章中不少地方已使用边际替代率绝对值这种表达方法。

可以不同。

用图 36.1 来说明公共物品的效率条件。我们画出每个人的 MRS 曲线，将它们垂直相加得到 MRS 总和曲线。如图所示，公共物品有效率的配置将发生在 MRS 之和等于边际成本之处。

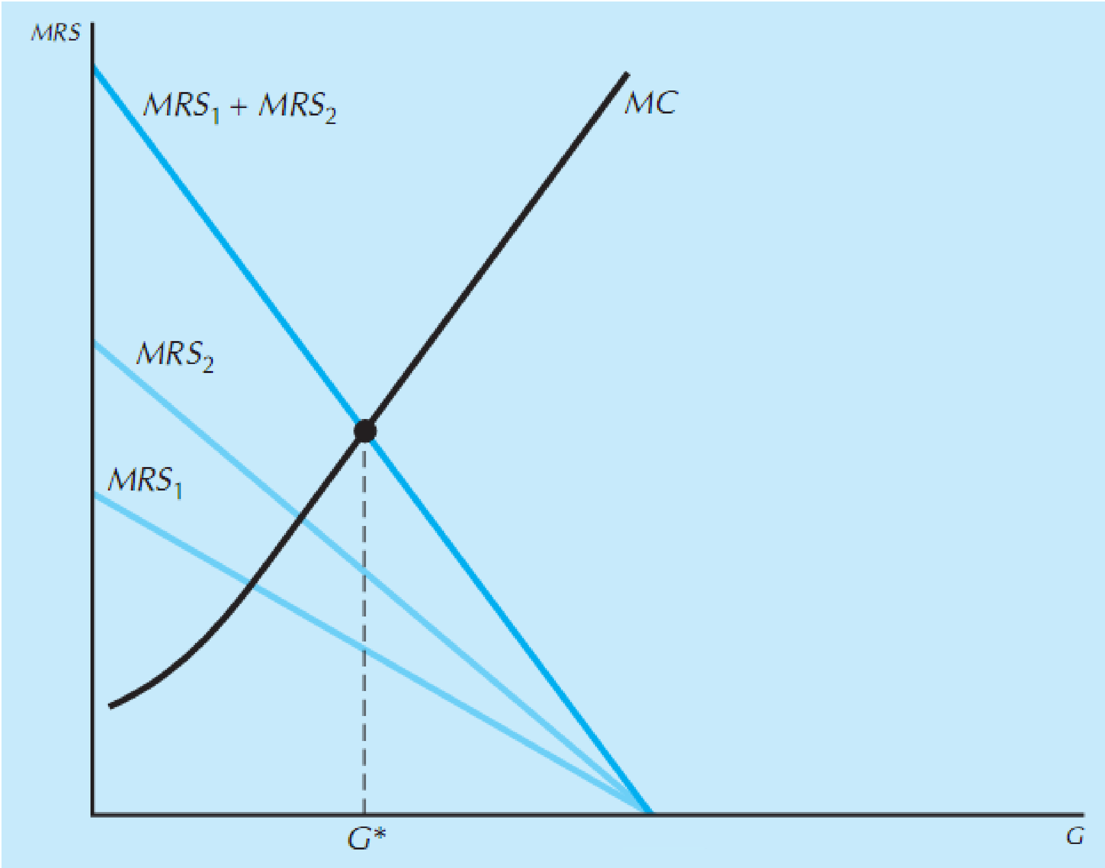


图 36. 1：确定有效率的公共物品数量。边际替代率之和必须等于边际成本。

36.5 拟线性偏好和公共物品

一般来说，私人物品的不同配置会导致公共物品的最优数量也不同。但是如果消费者的偏好是拟线性的，则无论私人物品如何配置，公共物品的最优数量只有一个。理解这个结论最简单的方法是构造一个代表拟线性偏好的效用函数。

我们在第 4 章已知道，拟线性偏好的效用函数具有下列形式：$u_i(x_i, G) = x_i + v_i(G)$，这意味着私人物品的边际效用恒为 1，因此私人物品和公共物品的边际替代率，即它们的边际效用之比，将仅取决于 G。对于前文两个人的例子，有：

$$|MRS_1| = \frac{\Delta u_1(x_1, G)/\Delta G}{\Delta u_1/\Delta x_1} = \frac{\Delta v_1(G)}{\Delta G}$$

$$|MRS_2| = \frac{\Delta u_2(x_2, G)/\Delta G}{\Delta u_2/\Delta x_2} = \frac{\Delta v_2(G)}{\Delta G}.$$

我们已知道帕累托有效率的公共物品水平需要满足下列条件

$$|MRS_1| + |MRS_2| = MC(G)\text{。}$$

使用拟线性效用中 MRS 的表达式，可将该条件写为

$$\frac{\Delta v_1(G)}{\Delta G} + \frac{\Delta v_2(G)}{\Delta G} = MC(G).$$

注意这个式子决定了 G 的数量而且该式不含有 x_1 和 x_2，因此公共物品最优供给数量是唯一的。

另外一种理解方法是考察无差异曲线的行为。在拟线性偏好中，所有的无差异曲线是无差异曲线互相平移得到的。这意味着，这些无差异曲线的斜率，即边际替代率，不会因为我们改变了私人物品的数量而改变。假设我们已经找了公共物品和私人物品的一个有效配置点，此时边际替代率 MRS 之和等于 MC(G)。现在如果我们从某人手里取走一些私人物品，将其给与另外一个人，这两条无差异曲线的斜率都不会改变，因此边际替代率 MRS 之和仍然等于 MC(G)，于是我们得到了另外一个帕累托有效率的配置。

在拟线性偏好的情形下，通过重新分配私人物品的方法，可以找到所有帕累托有效率的配置。公共物品的数量固定在唯一的一个效率水平。

例子：污染问题再思考

回忆第 34 章介绍的那个钢铁厂和渔场的例子。那时，我们的结论是污染的最优数量是将由钢铁厂和渔场承担的污染成本内部化。假设现在有两个渔场，钢铁厂的污染是个公共物品（称为公共厌恶品更合适）。

污染的最优数量涉及将三个厂商的利润之和最大化，即将污染的总社会成本最小化。正式地，令 $c_s(s,x)$ 表示钢铁厂生产 s 单位钢和 x 单位污染的成本；$c_f^1(f_1,x)$ 表示污染水平为 x 时，第 1 个渔场捕捉到 f_1 单位鱼的成本；$c_f^2(f_2,x)$ 表示污染水平为 x 时，第 2 个渔场捕捉到 f_2 单位鱼的成本。为计算污染的帕累托有效率的数量，我们最大化三个厂商的利润之和：

$$\max_{s,f_1,f_2,x} \quad p_s s + p_f f_1 + p_f f_2 - c_s(s,x) - c_f^1(f_1,x) - c_f^2(f_2,x)$$

我们感兴趣的效应是，污染增加对总利润的影响。增加污染降低了钢铁厂的成本但增加了每个渔场的成本，该利润最大化问题的最优条件为：

$$\frac{\Delta c_s(\hat{s},\hat{x})}{\Delta x} + \frac{\Delta c_f^1(\hat{f}_1,\hat{x})}{\Delta x} + \frac{\Delta c_f^2(\hat{f}_2,\hat{x})}{\Delta x} = 0.$$

这个条件表明，三个厂商的污染的边际成本**之和**应该等于零。正如公共消费品的例子一样，

公共物品帕累托有效率的水平，与所涉及行为人的边际收益之和或者边际成本之和有关。

36.6 搭便车问题

既然我们已经知道公共物品帕累托有效率配置是怎么一回事，我们将注意力转移到如何才能达到有效率的配置。对于没有外部性的私人物品，我们已知道市场机制可产生有效率的配置，那么市场能有效率地配置公共物品吗？

假设每个人都有一些禀赋，比如一些私人物品 w_i 。部分私人物品可用于自己消费，部分可以用于购买公共物品。令 x_1 表示第 1 人的私人消费， g_1 表示他购买的公共物品数量，类似地可以写出第 2 人的这两个数量。为简单起见假设 $c(G) \equiv G$ ，这意味着公共物品的边际成本恒等于 1。公共物品的总供给数量 $G = g_1 + g_2$ 。由于每个人对公共物品的关注是关注其总数量，因此第 i 人效用函数的形式为 $u_i(x_i, g_1 + g_2)$ =$u_i(x_i, G)$ 。

第 1 人在决定应出多少钱购买公共物品时，他必须预测第 2 个人所出资金的数额。此处我们采用第 28 章介绍的纳什均衡。假设第 2 人所出资金为 $\bar{g}_2$ ，并假设第 2 人同样也对第 1 人所出资金数额进行预测。我们要找到这样的一个均衡，即给定对方的行为，每个人所出资金的选择为最优。

因此，第 1 人的最大化问题为

$$\max_{x_1, g_1} \; u_1(x_1, g_1 + \bar{g}_2)$$

$$\text{使得 } x_1 + g_1 = w_1.$$

这个问题类似于通常的消费者最大化问题。因此最优条件是相同的，即若该两人都购买两种物品，则公共物品和私人物品的边际替代率，对每个人来说都应该等于 1：

$$|MRS_1| = 1$$

$$|MRS_2| = 1.$$

然而，此处分析应该谨慎。的确，如果第 2 人可购买任何数量的公共物品，他将一直购买直至边际替代率等于 1。但很容易发生下列事情，即第 2 人认为第 1 人购买的公共物品数量已足够多，因此他没必要再购买任何数量的该公共物品。

正式地，我们假设每个人为购买公共物品所出资金为正：一旦出资就不能收回。因此，对于每个人所出资金还有一个约束条件，即 $g_1 \geq 0$ 和 $g_1 \geq 0$ 。每个人只能决定他是否增加公共物品的数量。但情形很有可能是，一人认为另一人所购买的数量已足够，因此他不需要再出钱购买。

这样的情形可用图 36.2 表示。图中，横轴代表每个人的私人消费，纵轴代表每个人的公共物品消费。每个人的“禀赋”由两部分组成：一是他个人的财富 w_i ；二是由**对方**购买

的公共物品数量，这部分公共物品数量他不需要出钱购买。图 36.2A 表示只有第 1 人出钱购买公共物品的情形，因此 $g_1 = G$。如果第 1 人购买了 G 单位公共物品，则第 2 人的禀赋不仅包括他的个人财富 w_1，还包括 G 单位公共物品（因为无论他出钱购买与否，他都可以消费公共物品）。根据假设，第 2 人不能减少而只能增加公共物品的数量，所以他的预算线可用图 36.2B 中的粗线表示。如图所示，给定第 2 人无差异曲线的形状，第 2 人的最优选择是搭第 1 人公共物品的便车，并且消费自己的禀赋。

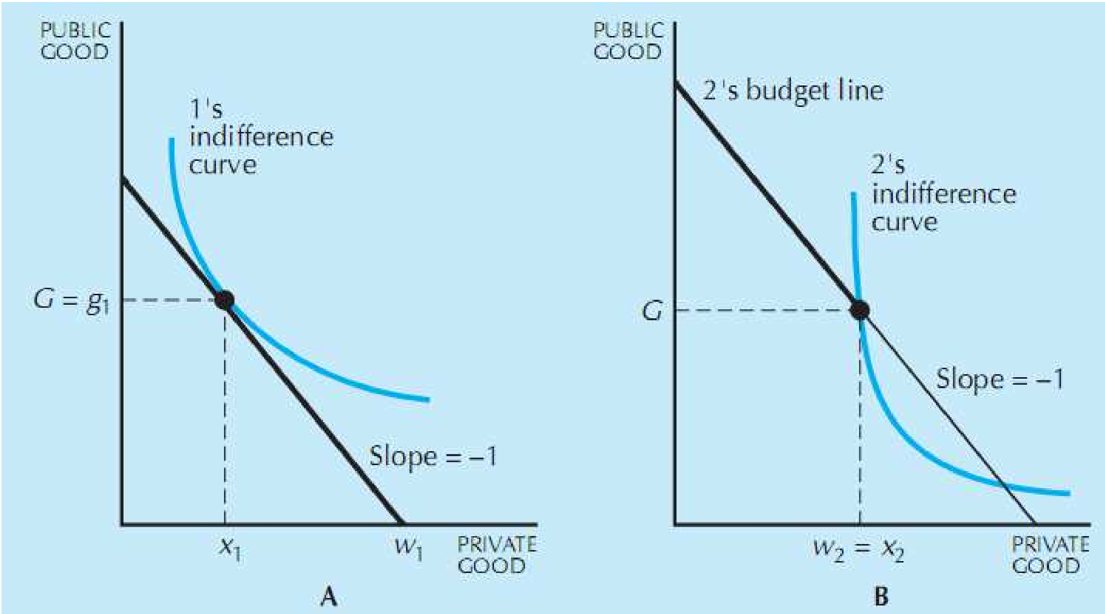


图 36.2：搭便车的问题。第 1 人购买公共物品而第 2 人搭便车。

在这个例子中，第 2 人搭乘了第 1 人公共物品的便车。由于每个人消费的公共物品数量必须相同，任何一人供给公共物品将有可能减少其他人的供给。因此，与有效率的公共物品供给数量相比，自愿均衡时的公共物品供给数量通常太少。

36.7 与私人物品相比较

在私人物品的分析中，我们已说明，完全竞争市场这种特别的社会制度，能实现私人物品的帕累托有效率的配置。每个消费者自己决定购买各种商品的数量，将会导致帕累托有效率的消费结构。该分析的一个主要假设是个人消费不影响其他人的效用，即不存在消费的外部性。因此，每个人最优化其消费行为就能实现某种社会最优。

公共物品的情形完全与此不同。该情形下，每个人的效用不可避免地联系在一起，因为每个人消费公共物品的数量相同。因此，由市场供给公共物品不大可能导致帕累托有效率的结果。

的确，我们通常使用不同的社会制度来确定公共物品的供给数量。有时人们使用命令

机制（command mechanism），即由一人或一小伙人确定应由大众供给的各种公共物品的数量。有时人们使用**投票表决系统**（voting system），即个人投票决定公共物品的供给。对于自由市场，你曾质疑过它的效率。你完全有理由对投票以及其他用于决策的社会制度，提出类似的疑问：它们能够实现公共物品帕累托有效率的配置吗？公共物品的任何一个帕累托有效率配置都能凭借这样的机制实现吗？对于上述问题的完整分析超出了本书的范围，但我们能够简要说明这些方法的运行机制。

36.8 投票表决

由私人供给公共物品通常行不通，但存在几种其他的可用于社会选择的机制。民主国家最常见的一种机制就是投票表决。我们来分析这种机制如何用于公共物品供给的决策问题。

两个人之间的投票表决比较无趣，因此我们假设有 n 个消费者。为避免出现投票结果平局的现象，假设 n 为奇数。假设消费者投票决定某种公共物品的数量，比如国防支出水平。每个消费者都有他最偏好的国防支出水平，他对其他国防支出水平的评价，取决于这些水平离他最爱的水平的远近。

使用投票方法来决定社会结果，存在几个问题。第一个问题我们已在第 33 章分析过。假设我们考虑三个支出水平，A，B 和 C。完全有可能出现下列情形：多数消费者选 A 不选 B，多数消费者选 B 不选 C…多数消费者选 C 不选 A！

用第 33 章中的术语来说，这些消费者的社会偏好不是传递的。这表明公共物品水平的投票结果不是定义清晰的（well defined），因为任何一个支出水平都会击败其他水平。如果允许一个社会对某一问题多次投票，各个选项可能出现“循环”。如果只投票一次，则结果依赖于选项出现的顺序。

如果先在 A 和 B 之间投票，然后再在 A 和 C 之间投票，C 最终胜出。但若先在 C 和 A 之间投票，然后是 C 和 B 之间投票，B 最终胜出。只要选择好投票顺序，你可以让这三个选项中的任何一个胜出！

上述“投票难题”比较烦人。你自然会问，对偏好加以何种限制就能解决这个难题？即什么样的偏好形式能保证不会出现上述“循环”问题？

我们用图 36.3 中的曲线来描述第 i 个消费者的偏好，曲线的高度表示他对公共物品不同支出水平的评价或者净效用。术语“净效用”用在这里是合适的，因为每个人既关心公共物品的水平，又关心他必须为公共物品出多少钱。更高的支出水平代表更多的公共物品，也代表更多的税收（用于购买更多的公共物品）。因此下列假设是合理的：公共物品支出的净效用一开始会升高，这是因为人们从公共物品中得到了好处；但净效用最终会下降，这是由于供给公共物品的成本增加。

对此类偏好的一种限制是令它们为**单峰的**（single-peaked），这意味着偏好的形状必须

为图 36.3A 描述的那样，而不是图 36.3B。对于单峰偏好，支出水平的净效用先是升高直至达到最偏好点，然后下降，请看图 36.3A。它绝不会象图 36.3B 描述的那样，即绝不会先升高，再下降，然后再升高。

如果每个人的偏好都是单峰的，那么可以证明，多数投票（majority vote）显示的偏好绝不会出现前文描述的非传递问题。若暂时接受这个结论，我们问个问题，如果每个人都有单峰偏好，那么应该选择什么样的支出水平？答案是**中间支出水平**(median expenditure)，在该支出水平上，一半的人想支出更多，一半人想支出更少。这个结果符合直觉：如果超过一半的人希望公共物品支出更多，他们会为更高支出水平的选项投票，因此唯一可能的投票均衡结果是，希望增加支出的投票和希望减少支出的投票的票数恰好相等。

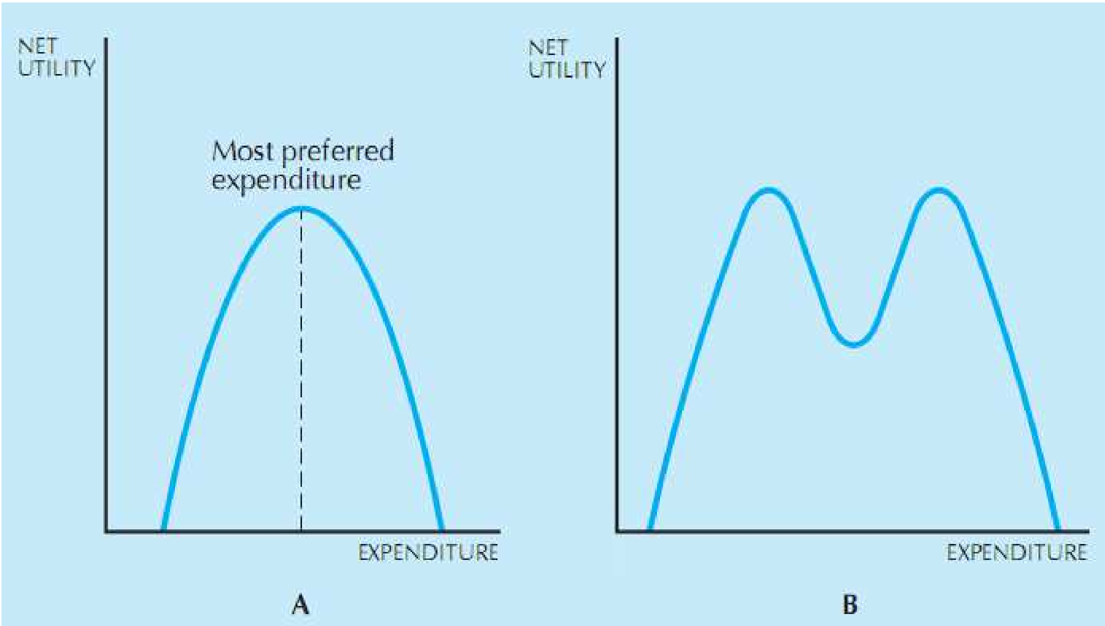


图 36.3：偏好的形状。图 A 表示单峰偏好，图 B 表示多峰偏好。

公共物品的这个水平有效率吗？答案通常为否。中间结果仅仅表明一半的人想要更多，一半的人想要更少；它丝毫未说明人们到底想要**增加多少支出**。因为效率问题需要考虑增加多少支出这类信息，投票表决通常不能达到有效率的结果。

而且，即使人们的偏好是单峰的（从而投票能产生一个合理的结果），人们在投票时也可以故意报出错误的偏好。因此为了最终投票结果，人们有激励隐藏真实的偏好而报出虚假的偏好。

例子：操纵议事日程

我们已看到，多轮投票的最终结果取决于投票顺序。有经验的政客深谙此道。在美国议会，对某议案的修改要先于该议案进行投票表决，这就提供了影响立法程序的一个常用方法。

1956 年，参议院要对要求联邦政府资助学校建设的议案进行投票。一位代表提出修改意见，即要求联邦政府只资助拥有综合学校[一]（integrated schools）的州。三个代表团对该问题持有不同观点，这三个团体人数大体相当。

- 共和党。他们反对联邦政府资助教育，但他们对修改的议案的偏好超过原议案。他们对选项的排序为：无议案，修改后的议案，原议案。
- 北方民主党。他们希望联邦政府资助教育并且赞成资助综合学校，因此他们对选项的排列顺序为：修改后的议案，原议案，无议案。
- 南方民主党。这个团体希望对教育给与联邦资助，但由于南方的学校不是综合性的，若修改后的议案胜出，他们得不到任何资助。因此他们的排序为原议案，无议案，修改后的议案。

对修改后的议案投票表决，共和党和北方民主党占多数，因此修改后的议案击败了原议案。在对修改后的议案表决时，共和党和南方民主党占多数，因此修改后的议案被淘汰。尽管原议案在修改前受到了多数人的拥护！

36.9 维克里-克拉克-格罗夫机制

下面我们在非常一般的架构（framework）内思考公共物品问题。我们的目的是选择某个结果（例如，是否提供路灯），以便使得相关参与人的效用之和最大。难题在于确定这些个人效用函数是什么样的，因为消费者可能缺乏恰当的激励来如实报告自己的评价。

在最简单的情形下，选择可能是 0-1 决策：如果 $x=1$，则设立路灯；如果 $x=0$，则不设路灯。在更一般的情形下，人们选择的是某种物品的提供数量——路灯的数量、亮度或位置。我们用 x 表示可能的选择，不管这些选择是什么。假设有 n 个参与人，令 $u_i(x)$ 表示参与人 i 的效用。目标是选择 x 使所有参与人的效用之和 $\sum_i u_i(x)$ 最大。

如果决策制定者知道每个人的效用函数，此事并不难。不幸的是，在任何实际情形下，决策制定者不知道这些效用函数。另外，我们已经知道，这些人也有误报自己真实效用函数的激励。

多少有些令人惊奇的是，有种聪明的方法能让所有人如实报告自己的评价并实现有效率的结果。这种经济机制称为**维克里-克拉克-格罗夫机制**（**V**ickrey-**C**larke-**G**roves mechanism）或**VCG 机制**（VCG mechanism）。

[一] 综合学校是指允许学生、教师等来源于不同种族的学校。译者注。

格罗夫机制

我们分两步介绍 VCG 机制。首先，我们讨论格罗夫机制（Groves mechanism）。

1.政策制定者（the center）让每个人 i 报告自己对 x 单位公共物品的支付意愿。我们用 $r_i(x)$ 表示这些个人对 x 单位公共物品的支付意愿。

2.政策制定者选择能使个人报告的效用之和 $R=\sum_{i=1}^{n} r_i(x)$ 最大的公共物品水平 x^*。

3.每个人 i 得到一个单方支付，这个单方支付等于按照步骤 2 确定的那个 x^* 计算出的所有其他人报告的效用之和。我们用 $R_i=\sum_{j\neq i} r_j(x^*)$ 表示这个单方支付。

可以证明，在这种机制中，每个人的优势策略（dominant strategy）是报告自己的真实效用函数。为了看清原因，考虑个人 i 得到的总收益，这个总收益等于他的效用加上他得到的单方支付

$$u_i(x)+\sum_{j\neq i} r_j(x).$$

注意：个人 i 关心自己的真实效用函数，但是他得到的单方支付取决于所有其他人报告出的效用函数。

个人 i 知道决策制定者在最大化所有人的效用之和时，使用的是他报告出的效用函数，因此效用之和为：

$$r_i(x)+\sum_{j\neq i} r_j(x).$$

然而，个人 i 希望政策制定者能让他自己的真实效用与他得到的单方支付之和最大，

$$u_i(x)+\sum_{j\neq i} r_j(x).$$

个人 i 的下列策略能确保决策制定者使得上式最大：报告自己真实的效用，即令 $r_i(x)=u_i(x)$。

格罗夫机制的本质是，在所有人之间将“个人外部性内部化”。它使得每个人面对自己报告出的效用对所有其他人施加的成本和收益。每个人都想报告自己的真实效用，因为只有这样才能使他的总收益（他的真实效用与他得到的单方支付之和）最大。

VCG 机制

格罗夫机制存在的问题是潜在成本巨大：决策制定者必须向每个人进行单方支付，这个单方支付等于所有其他人报告的效用之和。如何才能降低单方支付的数额？

一个重要的结论是，我们可以对每个人“征税”，只要这个税独立于个人的选择即可。

如果此税独立于个人 i 的选择，那么它不会影响他的选择[一]。我们的任务是找到一种税，让它能确保决策制定者得到的净支付为非负。因此，决策制定者总有足够的钱数来支付公共物品。

一种非常方便的做法是，对个人 i 征收的税收额等于他**不在场**时每个人报告效用之和的**最大值**。令 W_i 表示这个最大值，则 $W_i = \max\limits_z \sum\limits_{j \neq i} r_j(z)$。

于是，个人 i 承担的净税收为[二]

$$W_i - R_i = \max_z \sum_{j \neq i} r_j(z) - \sum_{j \neq i} r_j(x)$$

注意，这个数要么为正要么为零。为什么？因为 $n-1$ 个人报告的效用之和的**最大值**必定大于此和的所有其他值。

我们此处计算的是如果个人 i 不在场时的结果与他在场时的结果之差。因此，它衡量的是个人 i 对其他人造成的成本。只要让 i 承担他对其他人施加的成本，他有适当的激励来报告自己真实的效用。

现在我们可以完成 VCG 机制的描述了。VCG 机制的第 1 步和第 2 步和格罗夫机制相同，然后使用下列步骤替代格罗夫机制中的第 3 步。

3.政策制定者还计算出如果个人 i（其中 $i = 1, 2, ..., n$）不在场时的 $n-1$ 个人报告的效用之和的最大值。令 W_i 表示这个最大值。

4.每个人 i 缴纳的税收等于 $W_i - R_i$。

36.10 VCG 的例子

上一节的内容有些抽象，所以我们举几个具体的例子说明一下。

维克里拍卖

我们分析的第一个例子是**维克里拍卖**（Vickrey auction），我们在第 17 章已经介绍过。这里的结果非常简单：哪个人应该得到拍卖品。令 $v_1 > v_2$ 表示两个竞标人对拍卖品的真实评价，$r_1 > r_2$ 表示他们的报价。

[一] 假设每个人的效用函数都是拟线性的，这个假设非常重要。
[二] 原书的表达式有误，我已修正了这个错误。译者注。

如果个人 1 在场，他得到的效用为v_1；如果个人 1 不在场，个人 2 得到拍卖品，因此个人 1 的收益为$v_1 - r_2$。无论如何，个人 2 的收益为零。每个人都有报告自己真实评价的激励，因此我们得到了最优结果[一]。

克拉克-格罗夫机制

第二个例子是表 36.1 描述的公共物品问题。和那个例子一样，假设有两个室友正在决定是否购买一台电视。令c_i表示如果购买电视他将出的钱数。由于电视的价格为 150 元，因此，我们必有$c_1 + c_2 = 150$。

根据 VCG 机制，每个人报告他对电视的评价，用r_i表示。如果$r_1 + r_2 > 150$，那么他们会购买电视，每个人按照 VCG 付钱。令$x = 1$表示购买电视（电视的数量为 1），$x = 0$表示不买电视。

在使用 VCG 机制之前，我们首先假设如果我们使用的是下列原始的机制，将会发生什么情形。这个原始的机制为：让每个人报告自己对电视的评价，如果他们的报告评价之和大于电视价格，则购买。

假设个人 1 的评价大于他自己的出钱份额，因此$v_1 - c_1 > 0$。于是个人 1 可能报告说他对电视的评价为百万元；这将保证他们购买电视，因为他想看电视。另一方面，如果$v_1 - c_1 < 0$，他可能报告说他对电视的评价为负的百万元。

问题在于每个人都是独立决策的，他不会考虑其他人对公共物品的评价。每个人都有激励夸大或贬低自己报告出的评价。

下面我们看看 VCG 机制如何解决这个问题。个人 1 的净收益为他的效用减去他缴纳的 VCG 税收。

个人 1 从电视中得到的效用为$(v_1 - c_1)x$，他缴纳的税收为$\max_y (r_2 - c_2)y - (r_2 - c_2)x$，所以他的净收益为

$$(v_1 - c_1)x - \left[\max_y (r_2 - c_2)y - (r_2 - c_2)x\right]，即$$

$$(v_1 - c_1)x + (r_2 - c_2)x - \max_y (r_2 - c_2)y$$

在第二式中，第一项为个人 1 从电视中得到的效用：他对电视的评价减去他支付的钱数。第二项是他的室友报告的净效用。最后一项是如果个人 1 不在场时，他的室友（个人 2）能得到的最大效用。由于个人 1 不能影响最后一项，所以暂时忽略这一项。

[一] 理解这一段内容的关键是维克里拍卖中的中标者支付的价格等于他造成的外部性。在这个例子中，如果竞标人 1 不参与竞标，则竞标人 2 以r_2的价格得到了拍卖物，也就是说竞标人 1 对竞标人 2 造成的外部性为r_2元，所以竞标人支付的价格为r_2元（他造成的外部性），根据我们在外部性章节所学知识可知，拍卖的最终结果显然是有效率的。译者注。

重新整理前两项，我们得到个人 1 的收益为

$$[(v_1 + r_2) - (c_1 + c_2)]x.$$

如果上式为正，则个人 1 如果报告 $r_1 = v_1$ 就能保证电视被购买，因为这两个人报告的评价大于总成本。如果上式为负，个人 1 如果报告 $r_1 = v_1$ 就能保证电视不被购买。在这两个方面，个人 1 报告真实的评价都是最优的策略。类似的结论也适用于个人 2。如果这两个人报告各自的真实评价，则只有当 $v_1 + v_2 > 150$ 时，电视才会被购买，这是最优的。

注意：个人 i 只有改变了社会决策才必须缴税。在这种情形下，我们说他是关键人（pivotal）。关键人缴纳的税收就是他对其他人造成的成本。

36.11 VCG 机制存在的问题

VCG 机制能使参与人报告自己的真实评价，并且能导致最优的公共物品提供水平。然而，它也存在着问题。

第一个问题是，它只适用于拟线性偏好。这是因为如果你的偏好不是拟线性的，那么你拥有的钱数将会影响你对公共产品的需求，而公共产品的最优数量通常只有一个[一]。

第二个问题是，VCG 机制不能真正实现帕累托有效率的结果。公共产品的水平是最优的，但私人消费水平却可能比最优水平低。其中的原因在于征税。我们已经知道，为了让人们面对正确的激励，关键人必须实际缴纳税款，该税款要能反映他对其他人造成的损害。但是这些税收又不能交给参与公共产品决策的任何人，因为如果这样势必会影响他们的决策。税收必须从这个决策系统消失。这正是问题所在——如果关键人必须缴纳税收，那么私人消费就会比不缴税的情形低，因此是帕累托无效率的。

然而，只有某个人是关键人时他才需要缴税。如果参与人很多，那么任何一个人充当关键人的概率可能性不大，因此征税的可能性也不大。

VCG 的第三个问题是它容易遭受合谋的控制。例如，考虑上面描述的公共物品问题。假设有 3 个室友参与电视的拍卖，但是其中有 2 人合谋。合谋者商定每个人对电视均报出 100 万元的净收益。这保证电视被购买，但是由于哪个人也不是关键人（即，任何一个合谋者都没改变社会决策），因此没有人缴税。

最后一个问题是关于 VCG 机制固有的公平和效率的权衡问题。由于支付方案必须事先制定，可能出现下列情形：尽管提供的公共产品数量是帕累托有效率的，但某些人的状况却可能因此变坏。提供公共产品是帕累托改进的说法，意味着存在某种支付方案使得每个人在拥有公共产品时的状况，好于不拥有公共产品时的状况。但是，这并不意味着对于任意支付方案，每个人的状况都会变好。VCG 机制能做到的是，如果存在公共物品时每个人的状况都会变好，那么就应该提供公共物品；但这并不意味着提供公共物品会让每个人的状况都真

[一] 在研究克拉克税问题时，通常假设社会只有一种私人产品和一种公共物品。消费者 i 的效用函数为 $u_i = x_i + F_i(Y)$，其中 x_i 为此人的私人消费，Y 为公共物品的数量。这种效用函数显然为拟线性的，它的好处是公共物品的收入效应为零，因此公共物品的数量是唯一的。译者注。

正变好。

如果我们能找到某种支付方案，使得这种方案不仅能确定是否提供公共产品，它还是帕累托有效率的方法，也就是说这种方案能使得每个人的状况都变好，这样的事情当然很美好。遗憾的是，似乎不存在这样的方案。

附录

帕累托有效率的公共产品的数量是如何确定的？下面我们求解这个最大化问题

$$\max_{x_1,x_2,G} \; u_1(x_1,G)$$

$$\text{使得} u_2(x_2,G)=\bar{u}_2$$

$$x_1+x_2+c(G)=w_1+w_2$$

构造拉格朗日函数：

$$L=u_1(x_1,G)-\lambda[u_2(x_2,G)-\bar{u}_2]-\mu[x_1+x_2+c(G)-w_1-w_2]$$

分别对 x_1,x_2,G 求导，令分别令它们等于零可得

$$\frac{\partial L}{\partial x_1}=\frac{\partial u_1(x_1,G)}{\partial x_1}-\mu=0$$

$$\frac{\partial L}{\partial x_2}=-\lambda\frac{\partial u_2(x_2,G)}{\partial x_2}-\mu=0$$

$$\frac{\partial L}{\partial G}=\frac{\partial u_1(x_1,G)}{\partial G}-\lambda\frac{\partial u_2(x_2,G)}{\partial G}-\mu\frac{\partial c(G)}{\partial G}=0.$$

如果我们将第三个式子除以 μ 并整理，可得

$$\frac{1}{\mu}\frac{\partial u_1(x_1,G)}{\partial G}-\frac{\lambda}{\mu}\frac{\partial u_2(x_2,G)}{\partial G}=\frac{\partial c(G)}{\partial G}. \tag{36.2}$$

现在从第一个式子解出 μ，可得

$$\mu=\frac{\partial u_1(x_1,G)}{\partial x_1}$$

从第二个式子解出 μ/λ，可得

$$\frac{\mu}{\lambda}=\frac{\partial u_2(x_2,G)}{\partial x_2}$$

将上面两个式子代入（32.6）式可得

$$\frac{\partial u_1(x_1,G)/\partial G}{\partial u_1(x_1,G)/\partial x_1}+\frac{\partial u_2(x_2,G)/\partial G}{\partial u_2(x_2,G)/\partial x_2}=\frac{\partial c(G)}{\partial G}$$

上面这个式子正是教材中得到的式子：

$$MRS_1 + MRS_2 = MC(G).$$

总结

1.对于公共产品来说，每个人“消费”的数量都是相同的，例如国防、空气污染等。

2.如果公共产品的决策是“二选一的”——要么提供固定数量的公共产品，要么不提供——那么，提供公共产品为帕累托有效率的必要和充分条件是，人们的支付意愿（保留价格）之和大于或等于公共产品的成本。

3.如果公共产品可以按照可变数量提供，那么公共产品数量为帕累托有效率的必要条件是，人们的边际支付意愿（边际替代率）之和等于公共产品的边际成本。

4.搭便车的问题是指，每个人希望其他人提供公共物品，这样他可以免费享用。一般来说，纯粹的个人主义机制不能产生公共产品的最优数量，原因就是因为存在搭便车问题。

5.在公共产品供给的决策问题上，人们已提出了各种各样的集体决策方法，例如命令机制（command mechanism）、投票以及 VCG 机制等。

复习题

1.假设某个街道上住着 10 个人，每个人都愿意为额外一盏路灯支付 2 元，不管实际已有多少路灯。如果路灯的成本函数为 $c(G) = G^2$，请计算帕累托有效率的路灯数量。

2.（第 7 版习题）在某商品的拍卖活动中，人们轮流报价。每一次报价都要比前一次报价至少高一元钱，该商品将卖给报价最高的人。如果该商品对于第 i 人的价值为 v_i，最终胜出的报价为多少？哪一个人将得到该商品？

3.（第 7 版习题）在一次密封拍卖活动中，有 n 个竞标人。令 v_i 表示拍卖品对第 i 人的价值。证明，如果该商品以第二高的报价卖给报价最高的人，那么每个竞标人都会说真话，即会真实报价。

复习题答案

1.假设某个街道上住着 10 个人，每个人都愿意为额外一盏路灯支付 2 元，不管实际已有多少路灯。如果路灯的成本函数为 $c(G)=G^2$，请计算帕累托有效率的路灯数量。

【复习内容】公共物品的最优数量

假设社会只有一种私人产品和一种公共产品。消费者 i 的效用函数为 $u_i=u_i(x_i,G)$，其中 x_i 为此人的私人消费，G 为公共产品的数量。假设有 n 个人参与公共物品决策，那么公共物品的数量若为最优时，必须满足的条件是：

$$\sum_{i=1}^{n}\left|MRS_{Gx_i}^{i}\right|=\sum_{i=1}^{n}\left|\frac{\Delta x_i}{\Delta G}\right|=\sum_{i=1}^{n}\frac{MU_G}{MU_{x_i}}=MC(G)$$

【参考答案】

由于每个人都愿意为额外一盏路灯支付 2 元，即 $\left|\dfrac{\Delta x_i}{\Delta G}\right|=\dfrac{2}{1}=2$

由路灯的成本函数 $c(G)=G^2$ 可知其边际成本：$MC(G)=2G$

$$\sum_{i=1}^{10}\left|\frac{\Delta x_i}{\Delta G}\right|=MC(G)\Rightarrow 10\times 2=2G\Rightarrow G=10\text{（盏）}$$

因此，路灯的最优数量为 10 盏。

2.在某商品的拍卖活动中，人们轮流报价。每一次报价都要比前一次报价至少高一元钱，该商品将卖给报价最高的人。如果该商品对于第 i 人的价值为 v_i，最终胜出的报价为多少？哪一个人将得到该商品？

【复习内容】拍卖

【参考答案】

为论述方便，假设第 h 人认为该商品的价值最高，为 v_h。

无法具体确定最终胜出的报价，但它一般不会不等于 v_h，它最有可能等于第二高的报价加上一元，这样它就成了最高报价，最终胜出。谁出这样的价格谁就得到了商品。

例如，第 h 人认为该商品的价值最高（v_h），不妨令 v_h 为 1000 元，假设最终只剩下两个人角逐。比如对方喊出了 888 元的价格，那么这个人可以报价 889 元，如果对方不再报价，那么最终此人获得拍卖品，但是要注意他支付的价格（889 元）小于他的保留价格（1000 元），而是等于第二高的报价加上 1 元。

3.在一次密封拍卖活动中，有 n 个竞标人。令 v_i 表示拍卖品对第 i 人的价值。证明，如果该

商品以第二高的报价卖给报价最高的人，那么每个竞标人都会说真话，即会真实报价。

【复习内容】密封拍卖；克拉克税

【参考答案】

举例说明。假设你认为拍卖品的价值为 1000 元。现在你在纸上写下一个报价比如 800 元，现在你考虑是否将报价提高到 1000 元以上。(你这么做的原因是希望你的报价最高，你知道如果你中标的话，你只需要支付第二高的价格，而不是你的报价)。

如果你原先的报价（800 元）已是最高价格，那么你提高报价的决策不会影响到你中标的概率。如果你估计原先的报价（800 元）不是最高价格，因此决定报价 1500 元，你估计这个新的报价可能是所有竞标人中的最高报价。如果第二高的报价是 1200 元，那么显然是你中标，你需要支付 1200 元，而这个价格比你认为的价值（1000 元）高，因此你遭受损失。因此你不会报出比 1000 元更高的价格。

那么你会报出比 1000 元低的价格吗？比如你报价 900 元。但是如果有人报出 950 元的价格，假设这一价格为最高报价，那么他只需要支付 900 元就得到了拍卖品，但是你要知道这件拍卖品在你心目中值 1000 元。因此，你也不会报出比 1000 元更低的价格。

所以，最终你会如实报出价格——1000 元。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

37.不对称信息（含全部习题详细解答）

曹乾 译

（东南大学 caoqianseu@163.com）

37 不对称信息

直到目前，我们对市场的研究还未涉及到由信息差异引起的问题，因为我们一直假设买卖双方都完全了解市场中出售的商品。若商品的质量容易验证，则该假设合理。若辨别商品质量高低的成本不大，那么商品价格就会调整，以反映它们之间的质量差异。

但是若获得质量信息的成本较大，再假设买卖双方对交易的商品有相同信息就不合理了。现实世界中无疑存在很多这样的市场：获取商品准确信息的成本很高，甚至不可能获取。

一个明显的例子是劳动市场。在前面介绍的简单模型中，劳动是同质商品——人们的劳动“种类”相同，每个人在单位时间内提供相同数量的成果（effort）。这样的模型太简化了！在现实中，企业很难确定雇员们的生产能力。

信息成本高昂并不是劳动市场所特有的问题。类似的问题在消费品市场中也存在。当消费者购买了一辆二手车，他很难确定这车是好车还是孬车（lemon）。相比之下，二手车的卖方对车的质量很了解。我们将看到，这种**不对称信息**（asymmetric information）对市场运行效率将造成严重影响。

37.1 “柠檬”市场

我们看一个市场模型，在该模型中需求者和供给者对出售的商品拥有不同信息[一]。

考虑一个二手车市场：有 100 人想卖车，100 人想买车。每个人都知道市场中好车（plum）和孬车（lemon）各有 50 辆[二]。当前的车主知道车的质量，但潜在的购买者不知道车是好车还是孬车。

孬车车主的保留价格为 1000 元，好车车主的保留价格 2000 元。购买者对好车的保留价格 2400 元，对孬车的保留价格为 1200 元。

如果车的质量容易验证，市场就没什么问题。孬车的成交价在 1000 元 和 1200 元之间，好车的成交价在 2000 元和 2400 元之间。但若购买者**不能**观知车的质量，市场将会如何？

这种情形下，购买者只能猜测车的价值。为简单起见，假设购车者的猜测如下：某车为好车或孬车的可能性各占一半。在此情形下，典型购买者将只愿意支付车的期望价值（expected value）。使用上面提及的数值可以算出购买者愿意支付的价格为 $\frac{1}{2}\times1200+\frac{1}{2}\times2400=1800$ 元。

[一] 指出存在这种问题市场的第一篇论文是 George Akerlof, “The Market for Lemons: Quality Uncertainty and the Market Mechanism,” The Quarterly Journal of Economics, 84, 1970, pp. 488-500. 因为这项工作，他获得了 2001 年诺贝尔经济学奖。

[二] 在俚语中，lemon 是指质量低劣的东西，因此这里指孬车，而 plum 的意思正好相反，这里指好车。

可是在该价格水平下谁愿意卖车？孬车车主愿意，好车车主不愿意——根据假设，好车车主需要至少 2000 元才卖车。买者对一辆“平均”质量二手车的出价小于好车车主的保留价格。在 1800 元的价格水平，只有孬车出售。

若买者确知他将买到孬车，他绝不愿意支付 1800 元！实际上，该市场的均衡价将是 1000 元和 1200 元之间的某个价位。在该价格范围内，只有孬车车主愿意卖车，买者因此（正确地）预测到他将买到孬车。在这个市场中，一辆**好车**也不会卖掉！即使买者对好车的出价超过了车主愿意卖车价，这样的交易也不会发生。

二手车市场失灵的原因值得深思。问题在于好车卖者和孬车卖者之间存在着外部性。当个人决定出售孬车时，他影响了买者对市场中平均车辆质量的认知，这降低了购买者对平均车辆的意愿支付价，因此伤害了那些想卖好车的人。正是这种外部性导致了市场失灵。

最有可能被叫卖的车辆，恰好是人们想要摆脱的孬车。正是这种可着劲叫卖的行为，向潜在购买者传递了车辆质量的信息。如果市场中充斥着低质量产品，则高质量产品很难被卖掉。

37.2 质量选择（Quality Choice）

在“柠檬”模型中好车孬车的数量都是既定的。此处我们改变一下上述模型，让生产者来决定产品的质量。我们将说明在这个简单的市场中，均衡质量是如何被决定的。

假设每个消费者都想买且只买一把伞，伞的质量有两种。消费者认为高质量伞的价值 14 元，低质量伞价值为 8 元。购买时无法分别伞的质量高低，只有在经历几次暴风雨后才能知晓。假设某些伞厂生产高质量伞，另外一些生产低质量伞。再假设两种雨伞的生产成本均为 11.50 元，而且雨伞行业是完全竞争的。请问雨伞的均衡质量是怎样的？

假设消费者根据已售出的雨伞**平均**质量，来判断市场中雨伞的质量，就象上一节孬车市场的情形一样。若高质量雨伞的比例为 q，则消费者对伞的愿意支付价格为 $p = 14q + 8(1- q)$。

需要考虑三种情形。

只有低质量制造者生产雨伞。在该情形消费者对一把平均质量的雨伞只愿意支付 8 元，但伞的生产成本为 11.50 元，市场无伞可卖。

只有高质量制造者生产雨伞。该情形下生产者的竞争会降低伞的价格，直至降到价格等于边际成本 11.50 元。消费者愿意对每把伞支付 14 元，因此他们能得到一些消费者剩余。

两类制造者都生产雨伞。在该情形，竞争导致伞价为 11.50 元。因此，市场中平均质量的伞在消费者眼里必须至少值 11.50 元。这表示我们必有

$$14q + 8(1-q) \geq 11.50$$

满足该不等式的最小 q 值为 $q = 7/12$. 这表示若高质量企业的比例为 7/12 时，消费者恰好愿

意为每把伞支付 11.50 元。

图 37.1 描述了均衡时高质量企业的比例是如何决定的。横轴表示高质量企业的比例 q。纵轴表示，若高质量伞的比例为 q 时消费者对伞的意愿支付价格。企业愿意以 11.50 元的价格供给两种质量的雨伞，因此，供给条件可用图形中刻度为 11.50 元的水平线表示。

只有在$14q+8(1-q)\geq 11.50$时，消费者才愿意买伞；这个条件的边界在图形中以虚线表示。均衡时 q 值介于 7/12 和 1 之间[一]。

在该市场中均衡价格为 11.50 元，但消费者对平均质量伞的评价介于 11.50 元和 14 元之间，具体数值取决于高质量企业的比例。

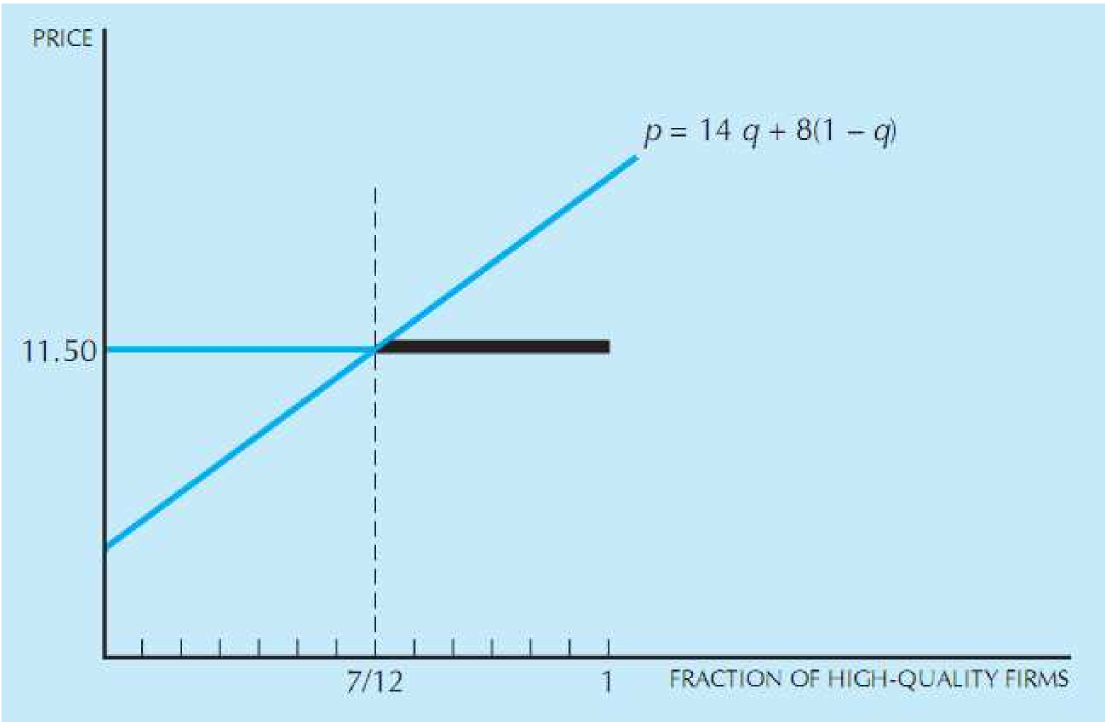


图 37.1： 均衡质量。水平线代表供给条件：市场愿意以价格 11.50 元供给任何质量的雨伞；斜线代表需求条件：若平均质量越高，则消费者越愿意多付钱。若高质量企业的比例至少为 7/12，则市场处于均衡中。

然而从社会的观点来看，所有这些均衡不是等价的。由于假设完全竞争和边际成本不变，企业的生产者剩余为零，因此我们只需要研究消费者剩余即可。此处容易看到，伞的平均质量越高，消费者的状况越好。从消费者的观点来看，最好的均衡是只有高质量的伞被生产。

[一] q 的取值范围为闭区间即 $q\in[\frac{7}{12}\ \ 1]$.在本章中除非特别说明，类似“介于…和…之间”的表达都是指闭区间。译者注。

选择质量（choosing the quality）

现在我们稍微改变一下这个模型。假设每个生产者能够选择自己生产的雨伞质量，并且生产每把高质量和低质量雨伞的成本分别为 11.50 元和 11 元。在这种情形下，将会发生什么样的结果？

假设选择生产高质量雨伞的生产者占所有雨伞生产者的比例为 q，其中 0<q<1。考虑其中一个雨伞的生产者。如果它的行为是竞争的，而且它相信它对雨伞市场的价格**和**质量的影响都可忽略不计，那么它将总是选择只生产低质量的雨伞。由于根据假设，这个生产者的市场份额微小，它可以忽略它对市场价格的影响，因此会选择生产利润更大的产品（低质量雨伞）。

但是每个生产者都会进行同样的推理，因此只会生产低质量的雨伞。但是消费者只愿意花 8 元钱购买低质量的雨伞，因此不存在均衡。如果你硬要用均衡表示，那么唯一的均衡就是**两种**雨伞的产量都为零！生产低质量雨伞的可能性摧毁了**两种**雨伞的市场。

37.3 逆向选择

上一节介绍的情形是**逆向选择**（adverse selection）的一个例子。在我们刚刚分析的模型中，由于获得信息的成本很高，低质量的商品将高质量的商品从市场中挤出（crowd out）。正如我们所看到的，这个逆向选择是如此严重以至于它可能完全摧毁市场。让我们看看逆向选择的一些其他例子。

首先考虑来自保险市场的例子。假设保险公司想提供自行车失窃保险。他们经过详细的市场调研后发现，失窃率在不同社区之间变化很大。在某些地区自行车被盗窃的概率很大，某些地区则相反。若保险公司提供保险的价格是基于**平均**失窃率，你认为后果将如何？

答案：保险公司可能很快破产！想一想，在该价格水平下谁会购买保险？安全社区的人们不会买，他们不需要很多保险。相反，高失窃率社区的人们会买，他们才是需要保险的人。

但是，这意味着大部分保险索赔是由高风险地区的人们提出的。基于失窃**平均**概率的费率不能反映保险公司实际遭受索赔的情况。保险公司得到的将不是消费者的无偏选择（unbiased selection），而是**逆向**选择。事实上，“逆向选择”这个词就是保险行业首先使用的，用来描述我们刚刚谈到的这类问题。

这将导致下列情形发生：为了保本，保险公司必然会根据“最坏情形”的估计相应调整费率，保险价格上升，从而那些风险较小但又不能忽略不计的消费者不愿意购买保险。

健康保险业有类似的问题，即保险公司不能根据人群的平均发病率来确定保险费率。他们只能根据高风险购买者团体的**平均**发病率来确定费率。但最想购买健康保险的人，可能是最需要保险的，因此费率必须反映这种差异。

在该情形下，也有办法让每个人的状况变好，这个办法就是根据人群平均风险确定保险

费率，然后**强制**每个人购买。高风险的人们处境变好，是因为该费率低于他们面对的实际风险水平；低风险的人们处境也变好，因为该费率低于**只有**高风险人群购买时的费率水平。

在上述情形中，强制购买方案使市场达到均衡，这一点会让大部分经济学家感到吃惊。我们通常认为“选择越多越好”，因此，限制选择也能导致帕累托改进的情形真得特殊。但应该强调，这种出乎意料的结果，是由于低风险和高风险人群之间存在着外部性。

事实上，存在着矫正这种市场失效的社会制度。通常的情形是，作为一种额外福利，雇主会向雇员提供健康保险。保险公司按照所有雇员的平均风险厘定费率，雇主向保险公司保证**所有**雇员全部参加保险，这样可消除逆向选择。

37.4 道德风险

保险行业另外一种有趣的问题是道德风险问题。该术语有些特别，但这种现象并不难以描述。再次以自行车失窃保险为例，为简单起见，假设所有消费者失窃概率相同，因此不存在逆向选择问题。可是，失窃概率会受车主**行为**的影响。

例如，若车主懒得锁车或者使用不坚固的锁，则自行车很容易被偷。在其他类别的保险中，也存在类似的问题。例如，在健康保险中，若消费者的生活方式健康，则需要保险的可能性较小。我们将影响某事件发生概率的行为称作**预防**(taking care)。

在确定费率时，保险公司必须将消费者采取适度预防数量的激励考虑在内。如果不存在保险，消费者有最大限度的预防激励。若买不到自行车失窃保险，则所有车主将使用昂贵的大锁。该情形下，个人承担了其行为的所有成本，因此他会“投资”于预防，直到预防的边际收益等于预防的边际成本时为止。

但若消费者能买到保险，则车被偷时个人遭受的损失要小得多。毕竟，在车被偷后，该人只要向保险公司报案即可，他将获得保险公司的赔偿重新买辆车。在极端情形下，即保险公司为失窃的车全额赔偿时，个人将毫无预防动机。我们将这种预防激励的缺失称作**道德风险**（moral hazard）。

注意此处涉及到的权衡：保险的保障程度太低意味着人们承担的风险很大；保障过度则又会导致人们预防数量不足。

若预防的数量可知，则不存在问题。保险公司可根据预防的数量确定费率。在现实生活中，保险公司通常制定不同的费率，办公地点有自动喷水灭火设备的公司享受低费率。健康保险公司也通常对吸烟者和非吸烟者制定不同的费率。在这些情形中，保险公司根据消费者不同的行为选择（从而影响损失概率），实施歧视定价。

可是，保险公司无法获知被保险人的所有相关行为。因此，此处又涉及到了前面提及的权衡问题：足额保险（full insurance）意味着预防数量太少，因为个人面对的不是他行为的全部成本。

据此推测一下保险公司将提供何种保险合同？一般来说，保险公司不想提供“足额（complete）”保险，他们总是希望消费者面对部分风险。这就是为何大部分保险单中含有“扣除额(deductible)”的规定，这个金额是被保险方在任何索赔下必须自负的额度。让消费者负担部分损失，保险公司可以保证消费者总有激励进行某种程度的预防。若保险公司能验证预防数量，则保险公司愿意承保足额保险。可是事实上因为预防数量的选择权在消费者，保险公司观测不到预防水平，所以保险公司不会允许消费者无限度地投保。

与标准市场分析相比，上述情形再次出乎意料。竞争市场的交易数量通常由需求等于供给的条件决定，即由购买的边际意愿等于销售的边际意愿所决定。在道德风险的情形中，市场均衡具有下列性质：每个消费者想多买保险；若消费者能持续地进行相同数量的预防，保险公司也想多卖保险…但该交易不会发生，因为若消费者能多买保险，则他们会理性地减少预防数量！

37.5 道德风险与逆向选择

道德风险是指，市场一方无法观测到另一方行为的情形。因此，它有时也被称作隐藏行为（hidden action）问题。

逆向选择则是指，市场一方无法观测到另一方的商品的 “种类”或者质量的情形。因此，它有时被称为隐藏信息（hidden information）问题。

具有隐藏行为的市场的均衡问题，通常涉及某种形式的定额配给（rationing）制——企业可以供给的更多，但他们不愿意这么做，因为这样会改变消费者的激励。具有隐藏信息的市场在均衡时，交易量通常很小，因为在“好”和“坏”商品类型中存在着外部性。

这种市场的均衡结果是无效率（inefficient）的，但如此断言需要谨慎。要问的问题是“该均衡相对于什么样的均衡来说是无效的？” 该均衡相对于完全信息时的均衡是无效率的。但这个结论对政策的制定几乎没有帮助：若行业中的企业发现收集更多的信息成本太高，政府也会发现同样的问题。

真正要问的问题是，即使政府收集信息成本和企业一样高昂，那么政府对市场进行某种干预是否会增进效率？

在隐藏行为的情形下，答案为“否”。若政府不能观察到消费者的预防程度，它不会比保险公司做得更好。当然，当然政府拥有保险公司所不具备的其他工具——政府可强行规定一既定水平的预防程度，对于没适当预防的人们，政府进行刑事处罚。但若政府只能调整价格和产量，那它不可能比自由市场做得更好。

类似的问题也存在于隐藏信息的情形。我们已看到，若政府可以强制所有不同风险等级的人群购买保险，则可能会使每个人的状况得以改善。从表面上看，这是政府干预的一个好例子。稍微深入地想，就会发现政府干预也有成本。通过政府法令实施的经济决策未必比私

营企业的决策更有成本效果（cost-effective）。所以，即使政府干预**会**增加社会福利，也不意味着这些干预必须实施！

此外，逆向选择问题有可能通过纯粹市场途径解决。例如，我们已看到雇主给雇员提供健康保险，可以帮助消除逆向选择问题。

37.6 发送信号

回忆二手车市场的模型：车主知道质量，而购买者只能猜测。我们已知道此信息不对称会导致市场出现问题；在某些情形下，逆向选择会使市场交易量太小。

然而这个故事还没结束。好车车主会向潜在消费者传递他有好车的事实，这对他有好处。他可能会采取不同的行为，向潜在买主**发送**车的质量**信号**（signal）。

一种可行的信号是好车车主提供**保证书**（warranty），比如，保证若车是孬车则车主将向买方赔偿某一商定的金额。好车车主可以提供这样的保证而孬车车主不能。好车车主使用这种方法传递他们有好车的信号。

在这种情形，发送信号可改善市场运行状况。通过提供保证书这种发送信号方式，好车车主将自己与孬车车主区分开来。但也存在发送信号使市场功能更糟的情形。

我们研究一个教育市场模型，该模型最初由迈克尔.斯彭斯（Michael Spence）提出[一]，此处我们做了简化处理。假设有两类员工：不能干的和能干的。不能干和能干员工的边际产量分别为 a_1 和 a_2，且 $a_2>a_1$。再假设能干和不能干员工的比例分别为 b 和 1－b.

为简单起见，我们假设生产函数为线性的，这样 L_1 个不能干的员工和 L_2 个能干员工的总产量为 $a_1 L_1 + a_2 L_2$. 再假设劳动市场是竞争的。如果员工的种类容易观知，则企业为员工提供的工资如下：对于不能干的员工，$w_1 = a_1$；对于能干的，$w_2 = a_2$。也就是说，每个员工按照他的边际产量获得相应的工资，这样我们就可得到有效率的均衡。

但若企业不能观知边际产量，情形会如何？若企业不能区分员工的种类，则它会提供平均工资，$w=(1-b)a_1+ba_2$。只要能干和不能干的员工都愿意接受该平均工资，而且努力干活，则不会出现逆向选择的问题。该情形与企业能观知员工类型的情形相比，产量有差异吗？利润有差异呢？在上述线性生产函数模型的假设下，两种情形的产量相等，利润也相等。

然而，假若员工可获得某些信号，从而将两类员工区分开。例如，员工可获得教育。令 e_1 和 e_2 分别表示不能干员工和能干员工获得的教育数量。再假设两类员工获得教育的成本不同，c_1e_1 和 c_2e_2 分别表示不能干员工和能干员工获得教育的总成本。这些成本不仅包括上学成本（用金钱衡量），还包括机会成本、必要努力（effort required）的成本，等等。

我们需要考虑两个决定。员工必须决定受教育的数量，企业必须决定对不同受教育程度

[一] Michael Spence, Market Signaling (Cambridge, Mass: Harvard University Press,1974).

的员工支付多少工资。我们做个极端的假设，假设教育程度对员工的生产能力毫无影响。当然，这个假设在现实生活中是不正确的，尤其对于学习经济学课程来说是不正确的，但这个假设可简化模型。

该模型的均衡本质主要取决于获得教育的成本。假设$c_2 < c_1$，即能干员工获得教育的边际成本小于不能干的。令 e*表示满足下列不等式的受教育水平：

$$\frac{a_2 - a_1}{c_1} < e^* < \frac{a_2 - a_1}{c_2}$$

由于我们已假设$a_2 > a_1$以及$c_2 < c_1$，这样的 e*肯定存在。

接下来考虑以下的决策集：所有能干员工获得的教育数量都为 e*，所有不能干员工获得教育数量都为 0；企业为受教育水平为 e*的每个员工支付工资数额为a_2，为低于该教育水平的员工支付工资为a_1。注意，员工对受教育数量的选择，传递了他是能干还是不能干员工的信号。

但是，这样能达到均衡吗？每个人是否有改变自己行为的激励呢？企业按照每个员工的边际产量支付相应的工资，所以企业没有改变自己行为的动机。唯一的问题在于，员工面对不同的工资时，他们的行为是理性的吗？

若不能干员工购买的教育数量为 e*，这是否对他们自身有利？若这么做，每个不能干员工的工资增加了$a_2 - a_1$，教育成本为$c_1 e^*$。收益在下列条件下将小于成本：

$$a_2 - a_1 < c_1 e^*.$$

根据前面对 e*的定义（即 e*的取值范围）可知，上述条件自然成立。因此，不能干员工发现选择零教育水平是最佳的。

若能干员工获得的教育数量为 e*，这真对他们有利吗？收益超过成本的条件为：

$$a_2 - a_1 > c_2 e^*.$$

同样，根据前面对 e*的定义可知，上述条件自然也成立。

因此，这种模式的工资安排的确可实现均衡：若每个能干员工选择的教育水平为 e*，不能干员工选择的教育水平为 0，则无人有理由改变自己的行为。由于我们假设教育成本存在差异，处于均衡中的员工的受教育水平，就是一个信号，这个信号传递了生产能力存在差异的信息。这种信号传递的均衡有时称作**分离均衡**（separating equilibrium），因为这种均衡涉及每个员工自行选择，从而将自身与其他类别的员工区分开。

另外一种均衡称为**联合均衡**（pooling equilibrium），在这种均衡中，每个员工的选择是**相同的**。例如，假设$c_2 > c_1$，即能干员工的教育成本高于不能干员工的教育成本。在该情形下，可以证明，唯一的均衡发生在员工按平均生产能力获得工资[一]，因此无人有动机发

[一] 此处提供一个“证明”，如果能干员工如果发现获得教育的成本超过了收益（工资增长），则他会放弃受

送信号。

分离均衡比较有趣，因为从社会的观点看它是无效率的。每个能干的员工发现花钱买信号对自己有利，尽管这并不能提高他的生产能力。能干员工想要获得信号的原因，不是在于它能增加生产能力，而是在于将自身与不能干员工区分开。（分离）信号传递均衡时的产量和无信号传递时的产量是完全相同的。在该模型中，从社会观点来看，获得信号纯粹是一种浪费。

这种无效率的本质值得思考。和以前的例子一样，无效率是因为外部性。若按员工的**平均**产量支付工资，则能干员工会受到伤害，因为存在不能干员工。因此，他们有投资于信号的激励，以将自身与不能干员工区分开。这种投资的回报是私人收益，但社会收益为 0.

当然，发送信号并非总是导致无效率。一些类型的信号，例如前面提及的二手车的质量保证书，对交易是有帮助的。在该例中，有信号的均衡比无信号的均衡更好。因此，信号传递可使事情更好或者更糟，这就要具体情况具体分析。

案例："羊皮"效应（文凭效应）

前文描述的教育信号模型比较极端，因为它假设教育对生产能力无影响，也就是说它认为在学校里度过的岁月仅是个人某些既定能力(fixed ability)的反映。这显然是不合理的：一个受过 11 年教育水平的学生比受过 10 年教育水平的学生更有生产能力，因为在额外的一年中，他多学了一些有用的技术。受教育的部分回报可能源于信号传递（signaling），部分回报可能来自在学校种获得了有用的技术。我们怎样区分这两部分回报？

一些劳动经济学家研究过教育回报，他们曾观察到以下具有启示性的事实：中学毕业生的收入远远高于仅读过 3 年中学的那些人。一项研究发现，与仅读过一年中学从而未毕业的学生相比，中学毕业可使收入增加 5 到 6 倍。收入的陡然跳跃现象也发生在大学毕业生身上。根据一项估算，完成 16 年教育的经济回报是完成 15 年教育回报的 3 倍[二]。

若教育能传授生产技巧，我们可以预期，完成 11 年教育的人的收入比完成 10 年教育的要高。真正令人惊讶的事是中学毕业生的收入存在着跳跃增长。经济学家将这种现象称为**羊皮效应**（sheepskin effect），因为文凭在过去是写在羊皮上的。收入跳跃增长的原因大概是中学毕业是某种信号。但它是什么的信号？在前文的教育信号模型中，教育成就是能力的信号。那么，能力是中学毕业的信号吗？或者它代表着其他？

安德鲁.外斯（Andrew　Weiss）这位波士顿大学的经济学家，试图回答这些问题[一]。他研究了一组数据，这组数据刻画了工人是如何装配设备的，从而可以估算刚参加工作的员

教育，即教育水平和不能干员工一样，都是 0。这种情形下，企业只能按员工的平均生产能力支付工资。译者注。

[二] Thomas Hungerford and Gary Solon, "Sheepskin Effects in the Returns to Education," Review of Economics and Statistics, 69, 1987, 175–77.

[一] "High School Graduation, Performance and Wages," Journal of Political Economy, 96, 4, 1988, 785–820.

工在第一个月的产量。他发现教育对产量的影响很小：中学教育每增加一年，产量仅增加1.3%.而且中学毕业的员工的产量与未毕业员工的产量大致相同。显然，教育对员工最初生产能力增加所起的作用很小。

于是他研究了另外一些数据，这些数据描述了不同职业员工的各种特征。他发现中学毕业员工的缺勤率明显低于中学未毕业的员工。中学毕业的员工工资较高的原因似乎在于，他们的产量高，但这种产量高是因为他们工作时间长以及缺勤较少。这表明，信号传递模型的确可以帮助我们研究现实世界的劳动市场。然而，教育成就的实际信号传递比我们简单模型的描述要复杂得多。

37.7 激励

我们转向一个稍微不同的主题，即开始研究**激励制度**（incentive system）。你会发现，该主题的讨论自然涉及信息不对称问题。但从完全信息开始研究有好处。

激励制度设计的中心问题是“我如何能让某人为我做某事？”让我们在具体背景下提出该问题。假设你有块地但你不能亲身耕作，于是你想雇佣某人代劳。你应该建立何种补偿制度？

一种方案是不管他生产多少，你都付他一笔数量固定的报酬。该情形下他没有激励去努力工作。好的激励方案通常将报酬与产量挂钩。激励制度设计的问题在于，准确测定报酬对产量的敏感度。

令 x 表示工人的“努力”程度，且令 $y = f(x)$ 表示产量；为简单起见假设产品价格为1，因此 y 还表示这些产量的价值。令 $s(y)$ 表示支付给工人的报酬，若他生产的产量价值为 y 。估计你愿意选取函数 $s(y)$ 来最大化你的利润 $y - s(y)$ 。

你面对的约束条件是什么？要回答它，应从工人的角度看问题。

假设工人发现努力工作有代价，用 $c(x)$ 表示努力程度为 x 的成本。再假设该成本函数有通常的形状：总成本和边际成本都随努力程度增加而增加。若工人选择努力程度为 x，则其效用为 $s(y) - c(x) = s(f(x)) - c(x)$ 。工人还有其他方法获得某种效用 L，该效用可能来自于其他的活，或者一点活也不干。

激励方案设计最重要的事，是使得工人干此工作得到的效用，至少与他从别处得到的效用一样大。这就给出了**参与约束**（participation constraint）：

$$s(f(x)) - c(x) \geq \overline{u} .$$

给定该约束条件，我们可以确定工人的产量。你想诱使工人选择努力程度为 x，该努力程度既要让工人愿意为你工作，又能带给你最大的剩余。

曹乾（东南大学 caoqianseu@163.com）

$$\max_{x} \quad f(x)-s(f(x))$$

$$使得 s(f(x))-c(x)\geq \bar{u}$$

通常，你希望工人选择的努力程度恰好满足约束条件，即 $s(f(x))-c(x)=\bar{u}$ 。将该式带入目标函数，即可得到无约束的最大化问题（unconstrained maximization problem）：

$$\max_{x} \quad f(x)-c(x)-\bar{u}$$

该最大化问题容易解决，只要选择 x^* 使得边际产量等于边际成本即可：

$$MP(x^*)=MC(x^*)\text{。}$$

x 的其他数值若不能使边际产量和边际成本相等，利润就不是最大。

这个最优解即 $x=x^*$ 表明雇主想要工人达到的努力程度；我们要问问他要付多少报酬才能使工人这么卖力。也就是说，什么样的函数 $s(y)$ 才能诱使工人选择 x^* 这个最优数量的努力程度？

你要想工人选择上述最优努力程度，你就要让他这么做对他有利，即你必须设计激励方案 $s(y)$，使得工人努力程度为 x^* 时他的效用最大。于是我们得到约束条件：

$$s(f(x^*)-c(x^*)\geq s(f(x)-c(x) \qquad 对所有 x 成立。$$

这个约束条件称为**激励相容约束**（incentive compatibility constraint）。该条件表明工人努力程度为 x^* 时的效用大于其他努力程度的效用。

因此，激励方案必须满足两个条件：一是它提供给工人的效用应等于 $\bar{u}$ ；二是它要使得在努力程度为 x^* 时边际产量等于边际成本。下列几种方案都可以满足这两个条件。

租金。雇主将土地租赁给工人，收取租金 R，因此工人在支付租金 R 后拥有所有产出。对于该方案，若工人使 $s(f(x))-c(x)=f(x)-R-c(x)$ 最大化，他选择的努力程度将为 x^*，此时 $MP(x^*)=MC(x^*)$，这正是雇主想要的努力程度。地租 R 可从参与约束条件解得。因为工人得到的总效用应为 $\bar{u}$ ，则有

$$f(x^*)-c(x^*)-R=\bar{u}$$

由此得 $R=f(x^*)-c(x^*)-\bar{u}$.

雇佣劳动（wage labor）。在这种方案中，工人报酬由两部分组成：一是数额固定的底薪 k ；二是按努力程度计酬，每单位努力的价格为 w 。即该激励报酬的形式如下：

$$s(x)=wx+k\text{ 。}$$

工资率 w 等于工人在最优努力程度 x^* 时的边际产量 $MP(x^*)$ ；常数 k 的选取，要使得工人为雇主工作与在其他地方工作相比，恰好无差异。即，该常数的选取应恰好满足参与约束。

于是，最大化 $s(f(x))-c(x)$ 的问题变为

$$\max_{x} \quad wx + k - c(x)$$

这意味着工人选择努力程度 x 以使得其边际成本等于工资率 $w = MC(x)$。因为工资率等于 $MP(x^*)$，这表明工人**最优**的选择，是选择 x^* 的努力程度使得 $MP(x^*) = MC(x^*)$，而这正是雇主想要的。

要么接受要么放弃（Take-it-or-leave-it）。在该方案中，若工人努力程度为 x^* 则雇主支付报酬 B^*；对于其他努力水平，工人报酬为 0. 报酬 B^* 由参与约束所决定，即 $B^* - c(x^*) = \bar{u}$，由此可得 $B^* = c(x^*) + \bar{u}$。若工人的努力程度为 x 且 $x \neq x^*$，则他的效用为 $-c(x)$；若其努力程度为 x^*，则效用为 $\bar{u}$。因此，工人的最优选择是使 $x = x^*$。

就我们的分析而言，上述方案是等价的：在每个方案中，工人得到的效用均为 $\bar{u}$；每个方案都可激励工人选择最优的努力程度 x^*。在这个意义上，上述方案没有好坏之分。

如果所有上述方案都是最优的，非最优方案是什么样子的？下面给出个例子。

佃农分成制（Sharecropping）。在佃农制方案中，工人和雇主按既定的百分比分成。假设工人的份额采取下列形式：$s(x) = \alpha f(x) + F$，其中 F 为常数，$\alpha < 1$。就我们考虑的问题而言，这***不是***一个有效率的方案。原因不难看清。工人的最大化问题为

$$\max_{x} \quad \alpha f(x) + F - c(x)。$$

这意味着工人选择的努力程度为 $\hat{x}$，此时有

$$\alpha MP(\hat{x}) = MC(\hat{x})。$$

这一努力水平显然没有满足效率条件 $MP(x) = MC(x)$。

我们总结一下上述分析思路。为了设计一个有效率的激励方案，必须保证作出努力程度决策的人对产品具有**剩余索取权**（residual claimant）。雇主使得自己状况尽量好的方法是，设法保证雇员能够生产最优产量。由此可推知，该激励方案要能做到雇员的边际收益等于他的边际成本。

例子：对公司的投票权

一般来说，公司的股份持有人对公司的管理事宜拥有投票权，但是该公司债券的持有人则不拥有这种权利。为什么？答案是股份持有人和债券持有人的收益结构不同。如果某公司某年的利润为 X 元，债券持有人对这些利润享有首先索取权，剩下的利润再分配给股份持有人。如果债券持有人拿走了 B 元，则分配给股份持有人的利润总额为（X－B）元。这使得股份持有人成为剩余索取人，因此他们有激励使得 X 尽量大。但是债券持有人的激励在于保证 X 大于或等于 B，因为他们最多也只能拿走 B 元。因此，给与股份持有人决策权通常能导致更大的利润。

例子：中国的经济改革

1979 年之前，中国的农村公社是按照标准的马克思主义路线组织的。公社大致估算每个社员对公社收入的贡献，然后相应支付工资。公社土地的 5%被留作自留地，但是农民不能到城市里销售他们自留地生产的产品。所有交易均通过高度管制的政府市场进行。

1978 年底，中国中央政府对农业结构进行重大改革，这项改革称为“联产承包责任制”。在这种制度中，政府先规定一个既定的定额，比如每亩地上缴 X 公斤，剩下的部分由承包土地的家庭获得，这部分产品可以在自由市场上销售。政府取消了对自留地的现值，增加了用于私人耕种的土地数量。到 1984 年底，97%的农民都在这种责任制下工作。

注意，这种责任制的结构非常类似上面介绍的最优激励方案：每个家庭上缴给公社的产品数量是固定的，剩下的部分都由家庭占有。因此，对家庭生产的**边际**激励，在经济上是合适的。

这种新制度对农业产量的效应是非常显著的：1978 年到 1984 年间，中国的农业产量增加幅度超过了 61%！但是，我们不能将这种全部增长全部归功于上述激励制度；在这项改革进行的同时，政府改变了对农产品价格的管制方法，甚至还允许某些农产品的价格由自由市场决定。

三位经济学家试图将中国农业产量增长的效应分为两个部分：一是联产承包责任制；二是农产品价格变化。他们发现产量增长的 3/4 归功于承包责任制，只有 1/4 的增长是由价格变化引起的[一]。

37.8 信息不对称

上面的分析提供了一些如何运用不同类型激励方案的思想。例如，它表明将土地租给农民比分成制更好。但是这个证明有些武断了。如果我们的分析与现实世界相符，那么可以预期在农业中将普遍采用租赁制或绩效工资制，而不是采用佃农分成制。

显然这是不对的。世界上某些地方，佃农分成制已使用了几千年，因此，这种方法可能能够满足某种需要。那么，我们的模型中忽略了什么因素？

根据本节的标题，也许你已经猜出了答案：我们忽略了不完全信息的问题。在前面的模型中，我们假设雇主能够完全观察到雇员的努力程度。在很多情形下，这是不可能做到的。雇主最多能观察到努力程度的某些**信号**，例如产量发生了变化。雇员的产量部分取决于他的努力程度，部分取决于天气，还有可能取决于投入品的质量，以及其他很多因素。由于存在着这么多“噪音”，雇主按照产量支付给雇员的报酬，一般不等于按努力程度支付的报酬。

这在本质上是一个信息不对称问题：雇员可以选择自己的努力程度，但是雇主却不能完全观测到雇员的努力程度。雇主只能按照产量估计雇员的努力程度，因此，在设计最优激

[一] J. McMillan, J. Whnlley, and L. Zhu, "The Impact of China's Economic Reforms on Agricultural productivity Growth." Journal of Political Economy 97, 4, 1989. 781-807.

励方案时，必须考虑到这样的猜测问题。

现在我们分析上述四种激励方案。如果努力程度和产量不是完全相关的，它们还能发挥作用吗？

租金。如果企业将技术租赁给工人，那么工人在支付既定的租赁费后，就可以占有剩下的产品收入。如果产量具有随机性，这意味着工人必须独立承担随机因素引起的所有风险。如果工人比雇主更厌恶风险（很有可能就是这样的），那么这种激励方案就是无效率的。一般来说，工人为了获得风险更小的收入流，他愿意放弃一部分剩余利润。

雇佣劳动。这种激励方案的问题在于，它要求获知准确的劳动**投入量**。工资应该按照投入生产的劳动量计算，而不是按照员工呆在企业的时间计算。如果雇主不能观测到准确的劳动投入量，那么他就不可能实施这种激励方案。

要么接受要么放弃。如果这种激励方案是按照劳动投入量计算报酬的，那么它存在的问题和雇佣劳动方案的问题是一样的。如果它是按照产出量计算报酬，那么这又涉及到工人承担所有风险的问题。在这种情形下，即使工人的产量稍微小于“目标产量”，那么他也是没完成任务，从而得到的报酬为零。

佃农分成制。这种激励方案可行。工人的收入部分取决于可观测到的产量，但是工人和雇主共同分担产量波动的风险。这使得工人有多生产的激励，而且不必承担全部风险。

在我们的模型中引入信息不对称因素后，不同激励方案的可行性发生了重大变化。如果雇主不能观测到员工的努力程度，那么雇佣劳动这种方案就是不可行的。在租金和要么接受要么放弃这两种方案中，员工承担的风险太大。佃农分成制折中了上述两类极端方案：它使得工人有多生产的激励，但是又不让工人承担全部风险。

例子：监督成本

雇主很难观测到雇员工作的努力程度。例如，雇员在某个 24 小时营业的便利店工作。雇主不在场时，他如何知道雇员的工作绩效？即使雇主能够观测到雇员的实物产出（查看货架上的存货或者收银机记录的销售额），他也很难观测到雇员的服务态度。

前东欧某些共产主义国家的服务，可能是世界上最差的服务：你向营业员打招呼时，迎接你的不是笑脸而是怒容。然而，一个匈牙利企业家 Gabor Varszegi，通过提供高质量的照片洗印服务，却在布达佩斯成功赚取百万身家[一]。

Varszegi 说他的经商生涯可以追溯到 1960 年代，那时他弹电吉他并管理着一个摇滚乐队。“那时候，”他说道，“东欧的唯一一种商人就是摇滚乐手。”1985 年他开始在匈牙利经营“一小时即取”的相片洗印服务，当时国营店却要顾客等待一个月才能取到照片。

[一] See Steven Greenhouse, " A New Formula in Hungary: Speed Service and Grow Rich," New York Times, June 5, 1990, A1.

Varszegi 在雇佣员工时坚持两条原则：一是他从来不雇佣曾在国营店工作过的员工；二是他支付给雇员的工资是市场的四倍。这样的做法从监督成本角度来看是非常合理的。由于他的每个连锁店的员工数量很少，如果对员工的行为进行监督，那么监督成本将非常高。如果解雇员工时，员工的损失非常小，那么员工很有可能偷懒。Varszegi 给员工很高工资的目的是让员工自己监督自己，因为如果被接管，他们遭受的损失非常大，这样的做法使得监督成本大幅度下降。

例子：格莱珉银行（The Grameen Bank）

孟加拉国的一位银行家向借款人索要的年利率高达 150%。任何美国银行家都会喜欢这么高的报酬率：为什么花期银行不在孟加拉营业？这个问题其实是自问自答：花期银行可能不如那位银行家做得好。乡村银行的放款人在小额借款方面具有比较优势，原因可能有以下几个方面：

- 乡村银行在处理小额借款方面更有效率；
- 乡村银行比国外银行更容易获得当地客户信誉信息资料；
- 乡村银行在监督还款流程方面处于有利地位，从而保证了还款。

由于乡村银行在处理规模报酬、逆向选择和道德风险三个方面具有优势，因此它在当地的信贷市场中占有垄断地位。

这样的垄断者对于象孟加拉这样的不发达国家来说，是非常不利的。由于年还款利率高达 150%，导致农民不能开展很多能盈利的项目。如果降低还款利率，人们的借款就会增加，投资就会相应增加，这样就会改善他们的生活水平。

默罕穆德.于纳斯（Muhammad Yunas），这位曾在美国受训的经济学家建起了一家银行——格莱珉银行（乡村银行），巧妙地解决了上述三个难题。在格莱珉方案中，具有不同创业项目的人聚在一起，共同申请一笔贷款。如果贷款获得批准，该小组的两个成员可以得到贷款，开始他们的投资活动。如果他们能还款，那么该小组的另外另个成员可以得到贷款。如果他们也能还款，那么该小组的最后一位成员，也就是组长，将得到贷款。

格莱珉银行成功地解决了上述三个难题。因为团队的信誉影响团队成员能否得到贷款，所以他们在选择伙伴时就会非常谨慎。由于只有在其他成员投资成功并还款后，团队的某些成员才能得到贷款，因此他们非常愿意互相帮助、共享技能。最后，选择放款对象以及监管还款进程都是由农民自身完成的，而不是由银行工作人员直接完成。

格莱珉银行大获成功。它每月放款 475,000 笔，平均每笔 70 美元。它的呆账回收率为 98%，而孟加拉一般银行的呆账回收率只有 30%-40%。由于团体责任项目在鼓励投资方面取得了成功，北美和南美的一些贫困地区也采用了这样的项目。

总结

1.信息不完全和信息不对称能导致市场均衡的性质变得非常不同。

2.逆向选择是指市场一方的产品类型或质量是不可观测到的，因此另一方必须根据对方的行为猜测产品的类型或质量。

3.如果市场中存在逆向选择的现象，那么交易量通常太小。在这种情形下如果强迫人们进行交易，就能使他们的状况变得更好。

4.道德风险是指市场一方不能观测到另一方行为的情形。

5.发送信号是指，当逆向选择或道德风险存在时，有些人会致力于发送信号，从而将他们与其他人区别开。

6.对发送信号进行投资可能对个人有利，对社会来说却是浪费。另一方面，对发送信号投资可能有助于解决信息不对称的问题。

7.有效率的激励方案（可以完全观测到努力程度）让雇员变成剩余索取人。这表明雇员将使得他们的边际收益等于边际成本。

8.但是如果信息是不完全的，那么上面的结论就不再成立。一般来说，如果一种激励方案既能做到风险共享又能做到提供激励，那么它就是合理的方案。

复习题

1.考虑教材中介绍的二手车市场的模型。市场均衡时的交易能实现的最大消费者剩余是多少？

2.在上一题的模型中，如果将买方随机指派给卖方进行交易，【而且假设卖方也不知道二手车的质量】。这样的交易能产生的消费者剩余为多大？与上一题的方法相比，哪一种方法产生的消费者剩余更大？

3.【假设某员工的生产函数为 $y = f(x) = x$，其中 x 为其努力程度，并且假设产品价格为 1】，该员工生产 x 单位的产品的成本为 $c(x) = x^2/2$。如果他在别的地方工作，他能得到的效用水平 $\bar{u} = 0$。最优雇佣劳动激励方案 $s(x)$ 应为多少？

4.给定上一题的条件，该员工对租赁生产技术的支付意愿为多大？

5.在第 3 题中，如果工人在其他地方工作能得到的效用水平 $\bar{u} = 1$，那么答案又是什么样的？

曹乾（东南大学 caoqianseu@163.com）

复习题答案

1.考虑教材中介绍的二手车市场的模型。市场均衡时的交易能实现的最大消费者剩余是多少？

【复习内容】信息不对称；逆向选择

先来回顾一下这个模型：一个二手车市场，有 100 人想卖车，100 人想买车。每个人都知道市场中好车（plum）和孬车（lemon）各有 50 辆。当前的车主知道车的质量，但潜在的购买者不知道车是好车还是孬车。孬车车主的保留价格为 1000 元，好车车主的保留价格 2000 元。购买者对好车的保留价格 2400 元，对孬车的保留价格为 1200 元。

【参考答案】

由于卖方知道二手车的质量，但买方不知道，因此买方对车的价值只能猜测。由于好车或孬车的可能性各占一半，则典型购买者将只愿意支付车的期望价值：0.5×1200+0.5×2400=1800 元。

因为这一价格低于好车车主的保留价格（2000 元），但高于孬车车主的保留价格（1000 元）因此只有孬车卖出。

然而，若买者确知他将买到孬车，他绝不愿意支付 1800 元，实际上，该市场的均衡价（此时就是孬车的均衡价）将是 1000 元和 1200 元之间的某个价位，具体价格取决于买卖双方的讨价还价能力。

所以，市场均衡时的消费者剩余位于下列两个数之间：

（1000-1000）×50=0（元）

（1200-1000）×50=10000（元）。

其中 50 是指孬车的数量，因为在这种情形下只有孬车才能成交，好车的成交量为 0。

也就是说，消费者剩余$\in[0\ \ 10000]$，显然消费者剩余的最大值为 10000 元。

2.在上一题的模型中，如果将买方随机指派给卖方进行交易，【而且假设卖方也不知道二手车的质量】。这样的交易能产生的消费者剩余为多大？与上一题的方法相比，哪一种方法产生的消费者剩余更大？

【重要提醒】

注意，题目中【】号的内容，即“而且假设卖方也不知道二手车的质量”这一句话为译

者本人所加，而不是作者范里安的原话。

但这绝对不是多此一举。如果不加上这一句话，那么这个题目和第一题是一样的。从作者范里安的本意（参考他提供的简要答案）可知，加上这一句话是很有必要的。

【复习内容】信息不对称；逆向选择

首先需要指出，如果卖方也不知道二手车的质量，那么这种情形就不是信息不对称，换句话说，这种情形是**信息对称**的，因为信息对称未必要求信息完全，只要双方的信息量对等就可以了。因此，在这种情形下，不存在逆向选择的问题。可以预期孬车和好车都可以成交。

【参考答案】

如果将买方随机指派给卖方进行交易，而且卖方也不知道二手车的质量，那么在这种情形下：

买方对一辆二手车的平均支付意愿：0.5×1200+0.5×2400=1800 元；

卖方对这辆二手车的平均销售意愿：0.5×1000+0.5×2000=1500 元。

因此，在这种情形下，一辆车的成交价在 1500 元到 1800 元之间，具体价格取决于双方的讨价还价能力。但是，要注意，在这种情形下，孬车和好车都可以成交，因此汽车的交易量为 100 辆。

与上题计算方法类似，

消费者剩余的最小值=（1500-1500）×100=0（元）；

消费者剩余的最大值=（1800-1500）×100=30000（元）。

所以，比较这一题和上一题的消费者剩余的最大值可知，这种方法产生的最大消费者剩余较大。

3.【假设某员工的生产函数为 $y=f(x)=x$，其中 x 为其努力程度，并且假设产品价格为 1】该员工生产 x 单位的产品的成本为 $c(x)=x^2/2$。如果他在别的地方工作，他能得到的效用水平 $\bar{u}=0$。最优雇佣劳动激励方案 $s(x)$ 应为多少？

【重要提醒】题目中【】号内的文字，即“假设某员工的生产函数为 $y=f(x)=x$，其中 x 为其努力程度，并且假设产品价格为 1”为译者所加，否则本题无法求出最终结果。

【复习内容】最优激励制度的设计；最优雇佣劳动激励方案

在雇佣劳动这种激励方案中，工人报酬由两部分组成：一是数额固定的底薪 k；二是按努力程度计酬，每单位努力的价格为 w。即该激励报酬的形式如下：

$$s(x)=wx+k\ 。$$

工资率 w 等于工人在最优努力程度 x^* 时的边际产量 $MP(x^*)$；常数 k 的选取，要使得工人

为雇主工作与在其他地方工作相比，恰好无差异。即，该常数的选取应恰好满足参与约束。

于是，最大化 $s(f(x))-c(x)$ 的问题变为

$$\max_x \ wx+k-c(x)$$

这意味着工人选择努力程度 x 以使得其边际成本等于工资率 $w=MC(x)$ 。因为工资率等于 $MP(x^*)$，这表明工人最优的选择，是选择 x^* 的努力程度使得 $MP(x^*)=MC(x^*)$，而这正是雇主想要的。

【参考答案】最优激励方案应为 $s(x)=x-\frac{1}{2}$ 。

最优雇佣劳动这种激励方案的形式为： $s(x)=wx+k$ ，这是雇主支付给雇员的报酬。我们的问题是要把 w 和 k 求出。

先看雇员的利润最大化问题为

$$\max_x \ wx+k-c(x)$$

这个最大化问题的一阶条件(目标函数 $wx+k-c(x)$ 对 x 求导，并令其等于零)为：

$$w=MC(x).$$

题目中已告知雇员的成本函数 $c(x)=x^2/2$，因此可求出边际成本函数：

$$MC(x)=x.$$

由以上两个式子可知：

$$w=x. \tag{1}$$

下面我们来看雇主的利润最大化问题

$$\max_x \quad f(x)-(wx+k)$$

$$使得\ wx+k-c(x)=\overline{u}$$

将约束条件 $w(x)+k-c(x)=\overline{u}$ 代入目标函数 $f(x)-(w(x)+k)$ 可得到如下的无约束最大化问题：

$$\max_x \quad f(x)-c(x)-\overline{u}$$

将题目给定的条件（ $f(x)=x$； $c(x)=x^2/2$； $\overline{u}=0$ ）代入这个最大化问题，可得

$$\max_x \quad x-x^2/2$$

这个最大化问题的一阶条件为：$1-x=0$，所以:

$$x=1 \tag{2}$$

由（1）式和（2）式可知：

$$w=1 \tag{3}$$

当$x=1$时，雇员的成本$c(x)=x^2/2=1/2$。因此：

$$wx+k-c(x)=\bar{u}\Rightarrow 1\times 1+k-\frac{1}{2}=0\Rightarrow k=-\frac{1}{2} \tag{4}$$

由（3）和（4）可知，最优激励方案应为$s(x)=wx+k=x-\dfrac{1}{2}$。

4.给定上一题的条件，该员工对租赁生产技术的支付意愿为多大？

【复习内容】最优激励方案的设计；租金方案

在租金这种激励方案中，雇主将土地租赁给工人，收取租金R，因此工人在支付租金R后拥有所有产出。对于该方案，若工人使$s(f(x))-c(x)=f(x)-R-c(x)$最大化，他选择的努力程度将为x^*,此时$MP(x^*)=MC(x^*)$，这正是雇主想要的努力程度。地租R可从参与约束条件解得。因为工人得到的总效用应为$\bar{u}$，则有

$$f(x^*)-c(x^*)-R=\bar{u}$$

由此得$R=f(x^*)-c(x^*)-\bar{u}$.

【参考答案】

由上面的分析可知，雇主收取的租金应为：

$R=f(x^*)-c(x^*)-\bar{u}$，其中x^*为雇员最优努力程度　（1）

而在上题中，我们已知：$x^*=1$；$f(x^*)=x^*=1$；$c(x^*)=1/2$；$\bar{u}=0$。

将这些条件代入（1）式可知，$R=1-1/2-0=1/2$。

5.在第3题中，如果工人在其他地方工作能得到的效用水平$\bar{u}=1$，那么答案又是什么样的？

【复习内容】最优激励方案的设计

【参考答案】$s(x)=x+\dfrac{1}{2}$

雇员的利润最大化问题及其解请参考第1题。下面从雇主的利润最大化问题开始分析：

$$\max_{x} \quad f(x)-c(x)-\bar{u}$$

将题目给定的条件（$f(x)=x$；$c(x)=x^2/2$；$\bar{u}=1$）代入这个最大化问题，可得

$$\max_{x} \quad x-x^2/2-1$$

这个最大化问题的一阶条件为：$1-x=0$，所以:

$$x=1 \tag{1}$$

由于我们在雇员的利润最大化问题中已得到$w=x$，所以有：

$$w=1 \tag{2}$$

当$x=1$时，雇员的成本$c(x)=x^2/2=1/2$。因此：

$$wx+k-c(x)=\bar{u}\Rightarrow 1\times 1+k-\frac{1}{2}=1\Rightarrow k=\frac{1}{2} \tag{3}$$

由（2）和（3）可知，最优激励方案应为$s(x)=wx+k=x+\dfrac{1}{2}$。

曹乾●经济学译丛精品系列

Intermediate Microeconomics: A Modern Approach (8th Edition)

Hal R. Varian

范里安

中级微观经济学：现代方法（第 8 版）

完美中文翻译 2014 版

38.数学知识回顾

曹乾 译

（东南大学 caoqianseu@163.com）

附录

在此附录中，我们简要复习教材中使用到的数学概念。这些材料旨在让你熟悉这些术语的意思。显然它不是数学讲义。因为这里给出的定义通常是最简单的，而不是最严格的。

A.1 函数

函数（function）描述的是数字之间关系的规则。对于每个数字 x，函数根据某个规则赋予**唯一**数字 y。因此，函数可通过描述规则进行说明，例如“任取一数，将其平方”，或者“任取一数，将其乘以 2”等等。我们可以将这两个函数分别写为 $y = x^2$， $y = 2x$。函数有时也称为**变换**（transformations）。

我们经常想说明某个变量 y 取决于另外一个变量 x，但是我们又不知道这两个变量之间的具体的代数关系。在这种情形下，我们写为 $y = f(x)$，它的意思是说变量 y 根据规则 f 取决于 x。

给定函数 $y = f(x)$， x 通常称为**自变量**（independent variable）， y 通常称为**因变量**（dependent variable）。意思是说， x 自主变化，但 y 的值**取决于** x 的值。

另外，某个变量 y 通常取决于**若干个**其他变量 x_1, x_2 等等，因此我们写为 $y = f(x_1, x_2)$。意思是说，这两个自变量共同决定了 y 的值。

A.2 函数图

函数图用图形描述函数的行为。图 A.1 给出了两个函数图。在数学中，自变量通常画在横轴上，因变量画在纵轴上。因此，函数表示的是自变量和因变量之间的关系。

然而，在经济学中，函数图通常将自变量画在纵轴上，因变量画在横轴上。例如，需求函数通常将价格放在纵轴上，将需求量放在横轴上。

A.3 函数的性质

连续函数（continuous function）是可以一笔画成的曲线：连续函数的图形不存在跳跃。

非连续函数如图（1）所示。**平滑函数**（smooth function）是没有“拐折”（kinks）或尖角（corners）的曲线[一]。非平滑函数如图（2）和（3）所示。**单调函数**（monotonic function）是指总是上升或总是降低的曲线[二]；**正单调函数**（positive monotonic function）的函数值总是随着 x 的增加而增加，而**负单调函数**（negative monotonic function）的函数值总是随着 x 的增加而下降[三]。

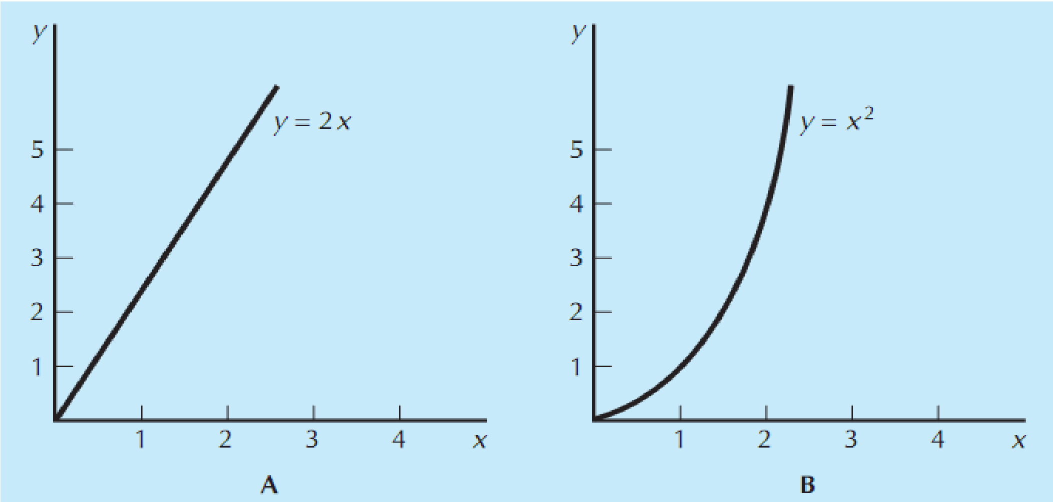
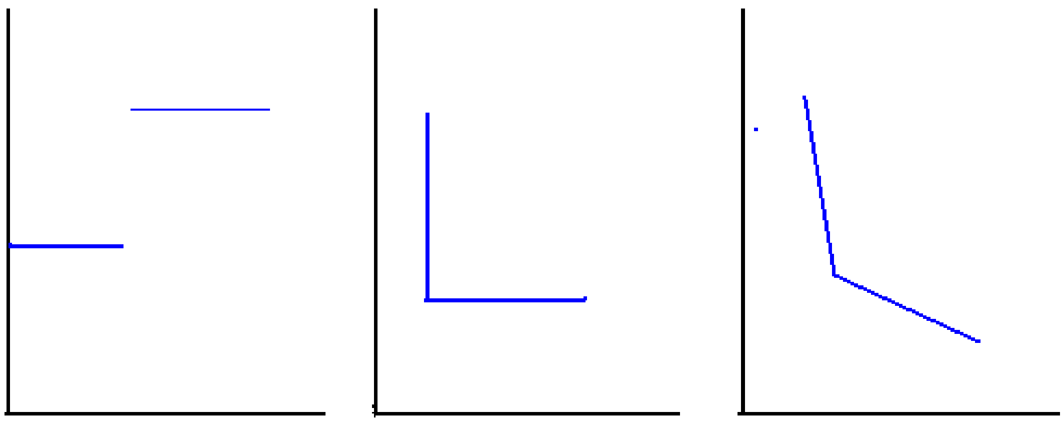


图 A.1：**函数图**。图 A 画出的是 $y=2x$ 的图；图 B 画出的是 $y=x^2$ 的图。



图：非连续函数与非平滑函数。图（1）的函数值出现了跳跃（一笔画不成此函数图），是非连续函数；图（2）和（3）的函数图中有拐折（或称尖角），是非平滑函数。[注：这组图为译者所加。]

[一] 平滑函数是可微分的函数。另，单调函数也是可微函数。这些都是你在高等数学中学过的。译者注。
[二] 单调递增函数和单调递减函数统称为单调函数,注意单调函数包括所谓的常数函数例如 $y=6$。译者注。
[三] 正单调函数即严格递增函数，负单调函数即严格递减函数。译者注。

A.4 反函数

我们已经知道函数具有下列性质：对于每个 x 值，都有唯一的 y 值与其相对应；还知道，严格单调函数的图形总是一直上升（或一直下降）。这意味着，对于任何一个严格单调函数来说，对于每个 y 值，都有唯一的 x 值与之相对应。

我们将具有上述性质的函数称为**反函数**（inverse function）。如果 y 是 x 的函数，如何求这个函数的反函数？只要我们将 x 表示为 y 的函数即可。如果 $y = 2x$，那么它的反函数为 $x = y/2$。如果 $y = x^2$，则它没有反函数。因为给定任何 y，$x = \sqrt{y}$ 和 $x = -\sqrt{y}$ 的平方都等于 y。因此，对于每个 y 值，不存在和它对应的唯一的 x 值，不符合反函数的定义。

A.5 方程和恒等式

方程问的是一个函数何时等于某个既定的数值。下面几个等式都是方程：

$$2x = 8$$
$$x^2 = 9$$
$$f(x) = 0.$$

方程的**解**（solution）是能满足该方程的 x 值。上面第一个方程的解为 $x = 4$；第二个方程有两个解：$x = 3$ 和 $x = -3$。第三个方程只是个一般方程（方程的一般表达式），只有知道 f 代表的具体规则，我们才能求出它的解，但是我们可以用 x^* 表示它的解。这只是说，x^* 是使得 $f(x^*) = 0$ 的数。我们说 x^* **满足**方程 $f(x) = 0$。

恒等式（identity）表示的是变量之间的一种特殊关系：**任意**给定变量的值，等式恒成立。下列两个式子都是恒等式：

$$(x+y)^2 \equiv x^2 + 2xy + y^2$$
$$2(x+1) \equiv 2x + 2.$$

上面的式子中有个特殊的符号≡，它表示对于变量的**所有**值，式子的左端都等于右端。方程和恒等式的区别在于，方程只对于变量的**某个（些）**值成立，而恒等式对于变量的所有值都是成立的。恒等式的成立与否通常取决于式中的变量是怎么定义的[一]。

A.6 线性函数

线性函数（linear function）是具有下列形式的函数

$$y = ax + b,$$

[一] 例如恒等式 $\sin^2 x + \cos^2 x \equiv 1$ 取决于 $\sin x$ 和 $\cos x$ 的定义。译者注。

其中 a 和 b 是常数。下列函数都是线性函数

$$y = 2x + 3$$
$$y = x - 99.$$

严格说来，形如 $y = ax + b$ 的函数应该称为**仿射函数**（affine function），只有形如 $y = ax$ 的函数才称为线性函数。然而，我们不严格区别这两类函数，都视为线性函数。

线性函数也可以用隐性形式表示，例如 $ax + by = c$。在这种情形下，我们通常求出 y（将其表示为 x 的函数），这样就得到了“标准”形式的线性方程：

$$y = \frac{c}{b} - \frac{a}{b}x.$$

A.7 变化量和变化率

符号 Δx 读作“x 的变化量”。注意它的意思不是 Δ 乘以 x。如果 x 从 x^* 变为 x^{**}，则 x 的变化量可以记为

$$\Delta x = x^{**} - x^*.$$

我们也可以将上式写为 $x^{**} = x^* + \Delta x$，用来表示 x^{**} 是 x 与 Δx 之和。

Δx 通常指的是 x 的**微小**变化。因此，我们有时也将 Δx 称为**边际变化**（marginal change）。

变化率（rate of change）是两个变化量的比率。如果 y 是 x 的函数，即 $y = f(x)$，则 y 关于 x 的变动率为

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

线性函数具有下列性质：y 关于 x 的变动率是常数。为了看清这一点，设 $y = a + bx$, 则

$$\frac{\Delta y}{\Delta x} = \frac{a + b(x + \Delta x) - a - bx}{\Delta x} = \frac{b\Delta x}{\Delta x} = b.$$

对于非线性函数来说，函数的变动率取决于 x。例如，对于函数 $y = x^2$ 来说：

$$\frac{\Delta y}{\Delta x} = \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x} = 2x + \Delta x.$$

由上式可以看出，变动率取决于 x 的值与 x 的变化量 Δx。但是如果我们考虑的是 x 的微小变动，Δx 将趋近于零，因此 y 关于 x 的变动率将趋近于 $2x$。

A.8 斜率与截距

函数的变化率在图形上就是这个函数的斜率。在图 A.2A 中，我们画出了线性函数 $y=-2x+4$ 的图。这个函数的**纵截距**（vertical intercept）是当 $x=0$ 时的 y 值，因此它的纵截距为 $y=4$。这个函数的**横截距**（horizontal intercept）是当 $y=0$ 时的 x 值，所以横截距为 $x=2$。该函数的斜率是 y 关于 x 的变动率，即 $\Delta y/\Delta x$，因此它的斜率为 -2。

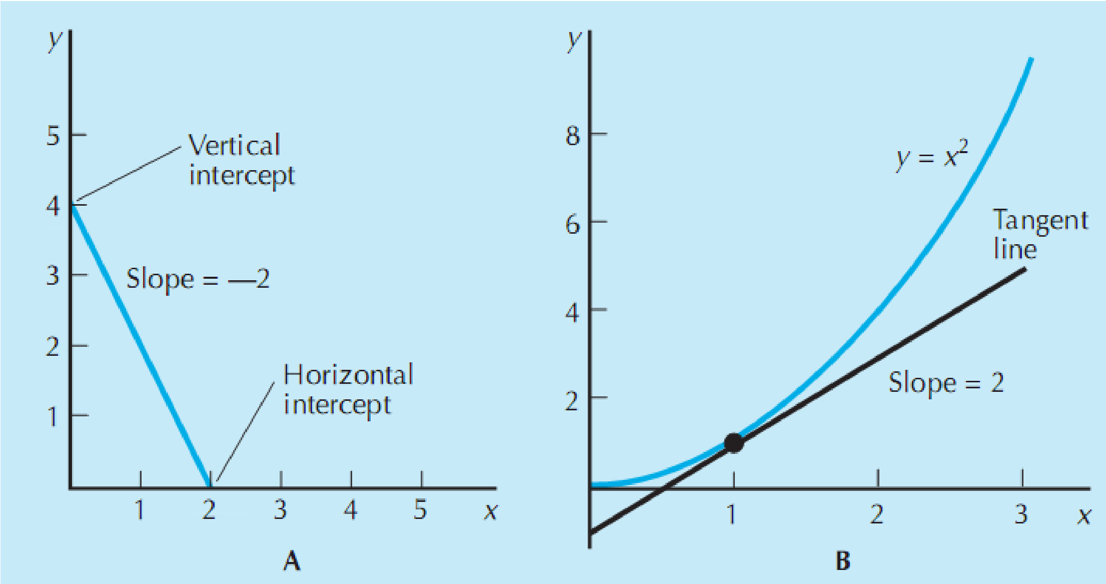


图 A.2A：斜率与截距。图 A 画出的是函数 $y=-2x+4$ 的图形；图 B 是函数 $y=x^2$ 的图形。

一般来说，如果线性函数的表达式为 $y=ax+b$，纵截距为 $y^*=b$，横截距为 $x^*=-b/a$。如果线性函数的表达式为

$$a_1x_1+a_2x_2=c,$$

它的横截距是当 $x_2=0$ 时的 x_1 值，所以横截距为 $x_1^*=c/a_1$；纵截距是当 $x_1=0$ 时的 x_2 值，因此纵截距为 $x_2^*=c/a_2$。该函数的斜率为 $-a_1/a_2$.

非线性函数的斜率随着 x 的变化而变化，它在 x 点的**切线**（tangent）是线性函数，二者斜率相同。在图 A.2B 中，我们画出了函数 $y=x^2$ 的图形以及它在 $x=1$ 点的切线。

如果当 x 增加时 y 也增加，则 Δy 与 Δx 的符号相同，因此函数的斜率为正。如果当 x 增加时 y 降低，或当 x 下降时 y 增加，则 Δy 与 Δx 的符号相反，因此函数的斜率为负。

A.9 绝对值和对数

某个数的绝对值（absolute value）可用下列函数定义：

$$f(x)=\begin{cases} x\ , & x\geq 0\text{时} \\ -x\ , & x<0\text{时} \end{cases}$$

因此，非负数的绝对值就是它本身，负数的绝对值是此数的相反数。绝对值通常记为$|x|$。

x的（自然）**对数**是关于x的一种特殊函数，记为$y=\ln x$或$y=\ln(x)$。对数函数是唯一具有下列性质的函数：对于所有正数x和y，有

$$\ln(xy)=\ln(x)+\ln(y).$$

以及

$$\ln(e)=1$$

（在最后这个式子中，e是自然对数的基，它等于 2.7183…）。用文字表述，两个数乘积的对数的计算方法是，先对每个数取对数然后再相加。从这个性质容易推出对数的另外一个重要性质：

$$\ln(x^y)=y\ln(x).$$

它是说：x的y次幂的对数等于y与$\ln(x)$的乘积。

A.10 导数

函数$y=f(x)$的**导数**（derivative）的定义为

$$\frac{df(x)}{dx}=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)}{\Delta x}.$$

也就是说，导数是当Δx趋于零时，y关于x变化率的极限。导数准确描述了“y关于x变化率的极限”这个意思。$f(x)$关于x的导数有时也用$f'(x)$表示。

我们已经知道，线性函数$y=ax+b$的变化率是常数。因此，线性函数的导数

$$\frac{df(x)}{dx}=a.$$

我们还知道对于非线性函数$y=f(x)$来说，y关于x变化率通常取决于x。例如，在前面的$f(x)=x^2$例子中，$\Delta y/\Delta x=2x+\Delta x.$ 使用导数的定义可知，

$$\frac{df(x)}{dx}=\lim_{\Delta x\to 0}(2x+\Delta x)=2x.$$

因此，x^2关于x的导数为$2x$。

如果$y=\ln(x)$，用更高级的方法可以计算出

$$\frac{df(x)}{dx}=\frac{1}{x}.$$

A.11 二阶导数

某函数的**二阶导数**（second derivative）是该函数导数的导数。对于 $y = f(x)$， $f(x)$ 关于 x 的二阶导数记为 $d^2 f(x)/dx^2$ 或 $f''(x)$。我们知道

$$\frac{df(2x)}{dx} = 2$$
$$\frac{df(x^2)}{dx} = 2x.$$

因此，

$$\frac{d^2 f(2x)}{dx^2} = \frac{d(2)}{dx} = 0$$
$$\frac{d^2 f(x^2)}{dx^2} = \frac{d(2x)}{dx} = 2.$$

二阶导数衡量函数的曲率（curvature）。某函数在某个点上的二阶导数若为负，则它在此点附近是凹的（concave）；它的斜率是递减的。某函数在某个点上的二阶导数若为正，则它在此点附近是凸的（convex）;它的斜率是递增的。某函数在某个点上的二阶导数若为零，则它在该点附近是水平的。

A.12 微分的积法则和链式法则

假设 $g(x)$ 和 $h(x)$ 都是 x 的函数。定义 $f(x)$ 为上述两个函数的乘积，即 $f(x) = g(x)h(x).$ 则 $f(x)$ 的导数为

$$\frac{df(x)}{dx} = g(x)\frac{dh(x)}{dx} + h(x)\frac{dg(x)}{dx}.$$

给定函数 $y = g(x)$ 和 $z = h(y)$，可以构造**复合函数**（composite function）

$$f(x) = h(g(x)).$$

例如，若 $g(x) = x^2$ 和 $h(y) = 2y + 3$，则复合函数为

$$f(x) = 2x^2 + 3.$$

链式法则（chain rule）是说复合函数 $f(x)$ 关于 x 的导数的计算方法为

$$\frac{df(x)}{dx} = \frac{dh(y)}{dy}\frac{dg(x)}{dx}.$$

在我们的例子中， $dh(y)/dy = 2$ ， $dg(x)/dx = 2x$ ， 因此根据链式法则可知， $df(x)/dx = 2\times 2x = 4x$。验证一下：这正是函数 $f(x) = 2x^2 + 3$ 关于 x 的导数。

A.13 偏导数

假设 y 取决于 x_1 和 x_2，因此 $y=f(x_1,x_2)$。于是 $f(x_1,x_2)$ 关于 x_1 的**偏导数**（partial derivative）的定义为

$$\frac{\partial f(x_1,x_2)}{\partial x_1}=\lim_{\Delta x_1\to 0}\frac{f(x_1+\Delta x_1,x_2)-f(x_1,x_2)}{\Delta x_1}.$$

$f(x_1,x_2)$ 关于 x_1 的偏导数正是维持 x_2 不变（视为常数）时该函数关于 x_1 的导数。类似地，$f(x_1,x_2)$ 关于 x_2 的偏导数为

$$\frac{\partial f(x_1,x_2)}{\partial x_2}=\lim_{\Delta x_2\to 0}\frac{f(x_1,x_2+\Delta x_2)-f(x_1,x_2)}{\Delta x_2}.$$

偏导数和一般导数具有完全相同的性质；唯一的区别是名字不同，这么做的原因是为了保护初学者（这些人未见过符号 ∂）。

特别地，偏导数也适用链式法则，但在形式上变得更复杂。假设 x_1 和 x_2 都取决于某个变量 t，我们定义函数 $g(t)$

$$g(t)=f(x_1(t),x_2(t)).$$

$g(t)$ 关于 t 的导数为

$$\frac{dg(t)}{dt}=\frac{\partial f(x_1,x_2)}{\partial x_1}\frac{dx_1(t)}{dt}+\frac{\partial f(x_1,x_2)}{\partial x_2}\frac{dx_2(t)}{dt}.$$

当 t 变化时，它同时影响 $x_1(t)$ 和 $x_2(t)$。因此，我们需要计算 $f(x_1,x_2)$ 关于 x_1 和 x_2 的导数。

A.14 最优化

如果 $y=f(x)$，若对于所有 x 都有 $f(x^*)\geq f(x)$，则我们说 $f(x)$ 在 x^* 处达到了**最大值**（maximum）。可以证明，若 $f(x)$ 是平滑函数且在 x^* 处达到最大值，则

$$\frac{df(x^*)}{dx}=0$$

$$\frac{d^2f(x^*)}{dx^2}\leq 0.$$

这两个式子分别称为最大值的**一阶条件**（first-order condition）和**二阶条件**（second-order condition）。一阶条件是说，函数在 x^* 处是水平的，而二阶条件是说函数在 x^* 附近是凹的。显然，如果 x^* 真能使 $f(x)$ 最大，这两个条件必须都满足。

若对于所有 x 都有 $f(x^*)\leq f(x)$，则我们说 $f(x)$ 在 x^* 处达到了**最小值**（minimum）。

可以证明，若 $f(x)$ 是平滑函数且在 x^* 处达到最大值，则

$$\frac{df(x^*)}{dx}=0$$

$$\frac{d^2f(x^*)}{dx^2}\geq 0.$$

一阶条件再次表明，函数在 x^* 处是水平的，而二阶条件是说函数在 x^* 附近是凸的。

若 $y=f(x_1,x_2)$ 是平滑函数且在点 (x_1^*,x_2^*) 处达到最大值或最小值，则我们必须满足

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_1}=0$$

$$\frac{\partial f(x_1^*,x_2^*)}{\partial x_2}=0\,.$$

这些式子称为**一阶条件**。这个问题也有二阶条件，但更难以描述。

A.15 约束最优化

我们常常希望找到某函数在带有约束条件时的最大值或最小值。最优化问题的典型形式是

$$\begin{aligned}&\max\ f(x_1,x_2)\\&s.t\ \ g(x_1,x_2)=c.\end{aligned}$$

它的意思是说：找到 (x_1^*,x_2^*)，使得对于满足 $g(x_1,x_2)=c$ 的所有 x_1 和 x_2，都有 $f(x_1^*,x_2^*)\geq f(x_1,x_2)$。

在上面的表达式中，$f(x_1,x_2)$ 称为**目标函数**（objective function），$g(x_1,x_2)=c$ 称为**约束条件**（constraint）。$s.t$ 是 “such that”的缩写，意思是 “使得”，表示它后面紧跟着约束条件。这类约束最优化问题的解法请见第 5 章的附录。

THIS IS THE END OF THE TEXTBOOK.