

计量经济学的性质 与经济数据

第 1 章讨论的是计量经济学的研究领域，并提出在应用计量经济方法时可能出现的一般问题。1.3 节考察了管理学、经济学和其他社会科学中所用数据集的种类。1.4 节直观地讨论了与社会科学中因果推断相联系的困难。

1.1 什么是计量经济学？

设想州政府雇用了你，让你对公共机构的工作培训项目进行效果评价。假设这个项目是培训工人在生产过程中使用计算机的各种方法。为期 20 周的培训都是在工人的非工作时间进行的。任何一个按小时计酬的生产工人都可以自愿参加全部或部分培训。你要决定培训项目对每个工人以后的小时工资有何影响。

现在假设你为一家投资银行工作，并研究对美国短期国债的各种不同投资战略的回报，看看它们是否与经济理论的含义相一致。

回答这种问题初看起来让人胆怯。此时，你对需要搜集的数据类型只有模糊的认识。在学完这本计量经济学入门教程后，你应该知道，如何用计量模型去规范地评价一个工作培训项目，或检验一个简单的经济理论。

计量经济学的发展，取决于估计经济关系、检验经济理论以及评价和实施政府与商业政策的统计方法的进步。计量经济学最常见的应用，就是对诸如利率、

通货膨胀率和国内生产总值等重要宏观经济变量的预测。尽管对经济指标的预测随处可见而又广为流传,但计量经济方法也可用于那些与宏观经济预测无关的经济领域。比如,我们可以研究政治竞选支出对投票结果的影响,还可以考虑教育领域学校支出对学生成绩的影响。此外,我们还将了解如何使用计量方法来预测经济时间序列。

由于计量经济学主要考虑在搜集和分析非实验经济数据时的固有问题,计量经济学已从数理统计分离出来并演化成一门独立学科。**非实验数据** (nonexperimental data) 并非从对个人、企业或经济系统中的某些部分的控制实验而得来。[非实验数据有时被称为**观测数据** (observational data) 或**回顾数据** (retrospective data), 以强调研究者只是被动的数据搜集者这一事实。] 自然科学中的**实验数据** (experimental data) 通常是在实验环境中获得的,但在社会科学中要得到这些实验数据则困难得多。虽然也可以设计一些社会实验,但为解决经济问题所需要实施的各种控制实验,要么行不通,要么代价高昂而让人望而却步,要么在道德上让人极为反感。在1.4节我们给出一些特殊的例子,来说明实验数据与非实验数据之间的区别。

当然,只要有可能,计量经济学家总会借用数理统计学家的一些方法。多元回归分析方法虽然在上述两个领域都是主要支柱,但其着眼点和解释可以极为不同。此外,经济学家已想出许多新方法,来处理经济数据的复杂性和检验经济理论所预测的结果。

1.2 经验经济分析的步骤

计量经济方法几乎在应用经济学的每一个分支中都相当重要。要么在我们有一个经济理论需要检验的时候,要么在我们的脑海中有一种关系,而在这一关系对商业决策或政策分析又相当重要的时候,我们便开始使用计量方法。**经验分析** (empirical analysis)* 就是利用数据来检验某个理论或估计某种关系。

应该如何构建一个经验经济分析呢?虽然看上去十分明显,但仍值得强调,任何一个经验分析的第一步,都是对所关心问题的详细阐述。问题可能涉及到对一个经济理论某特定方面的检验,或者对政府政策效果的检验。原则上,计量经济方法可用来回答诸多方面的问题。

在某些情形下,特别是涉及到对经济理论的检验时,就要构造一个规范的经济模型 (economic model)。经济模型总是由描述各种关系的数理方程构成。经济学家因建造模型来描述大量人类行为而声名远扬。例如在中级微观经济学中,个人在预算约束下的消费决策便可由一些数理模型来描述。这些模型背后的基本前提是效用最大化。个人选择能在资源约束条件下最大化其福利的假定,为我们创造一些简便

* 国内常称实证分析,但为了区别于 positive analysis,故译为经验分析。——译者注

的经济模型并做出一些明确的预测，提供了强有力的框架。在消费决策的背景下，效用最大化能导出一系列需求方程。在每个需求方程中，每种商品的需求量都取决于该商品的价格、其替代品和互补品的价格、消费者的收入和影响消费者个人喜好的一些个人特征。这些方程便形成对消费需求进行计量分析的基础。

经济学家还使用诸如效用最大化框架之类的基本分析工具，来解释那些初看起来具有非经济性质的行为。一个经典的例子就是，贝克尔（Becker，1968）针对犯罪行为所做的经济模型。

例 1.1

犯罪的经济模型

在一篇开创性的论文中，诺贝尔经济学奖得主加里·贝克尔（Gary Becker）系统地阐述了一个效用最大化框架，用以描述个人对犯罪行为的选择。虽然每一特定的犯罪都有明显的经济回报，但大多数犯罪行为也有其成本。犯罪的机会成本使得罪犯不能参加诸如合法就业之类的其他活动。此外，还存在与罪犯可能被抓住相联系的成本，以及罪犯被抓后，如果被证明有罪，与监禁相关的成本。从贝克尔的视角来看，决定进行非法活动的决策是资源配置的方式之一，并且是在充分考虑了各种可选择行为的成本和收益后决定的。

在一般化的假定之下，我们便能推导出一个方程，把花在犯罪活动上的时间描述成各种影响因素的一个函数。我们可以把这个方程表示为

$$y = f(x_1, x_2, x_3, x_4, x_5, x_6, x_7) \quad (1.1)$$

其中，

y = 花在犯罪活动上的小时数，

x_1 = 从事犯罪活动每小时的“工资”，

x_2 = 合法就业的小时工资，

x_3 = 犯罪或就业之外的收入，

x_4 = 犯罪被抓住的概率，

x_5 = 犯罪被抓后，被证明有罪的概率，

x_6 = 被证明有罪后预期的刑罚，

x_7 = 年龄。

虽然还有其他因素通常会影响个人参与犯罪的决策，但上述因素从规范的经济分析来看可能具有代表性。如经济理论的惯常做法那样，我们未对式（1.1）中的函数 $f(\cdot)$ 进行任何设定。这个函数取决于一个很少有人知道的潜在效用函数。尽管如此，我们还是可以用经济理论（或反思）来预测每个变量对犯罪活动可能具有的影响。这正是对个人犯罪行为进行计量经济分析的基础。

虽然规范的经济建模有时候是经验分析的起点，但更普遍的情况是，经济理论的使用不是那么规范，甚至完全是依赖直觉。你可能也同意，方程（1.1）中出现的

犯罪行为的决定因素，从常识来看也是合情合理的；我们也许能直接得到这个方程，而不需要从效用最大化开始把它推导出来。尽管在有些情况下，规范的推导能提供直觉无法获得的视角，但这种观点也有其优点。

接下来这个例子中的方程，就是从多少有些不甚规范的推理中得到的。

例 1.2

工作培训与工人的生产力

现在考虑在 1.1 节之初提出的问题。一位劳动经济学家想考察工作培训对工人生产力的影响。在此情形下，几乎不需要什么规范的经济理论。基本的经济常识就足以使我们认识到，所受教育、工作经历和培训等因素会影响工人的生产力。此外，经济学家还清楚地知道，工人的工资与其生产力相称。这种简单的理由就使我们得到如下模型

$$wage = f(educ, exper, training) \quad (1.2)$$

其中，

$wage$ = 小时工资，

$educ$ = 接受正规教育的年限，

$exper$ = 工作年数，

$training$ = 花在工作培训上的周数。

同样，虽然也有其他因素通常会影响工资率，但式 (1.2) 还是刻画了这个问题的本质。

在我们设定一个经济模型之后，就需要把它变成所谓的计量模型 (econometric model)。既然我们在全书都要讨论计量模型，那么最好先了解一下计量模型和经济模型有何关系。以方程 (1.1) 为例，与经济分析不同，在进行计量经济分析之前，我们必须明确函数 $f(\cdot)$ 的形式。与式 (1.1) 相关的第二个问题是，对不能合理地观测到的变量该如何处理。比如，考虑一个人在进行犯罪活动时的工资。原则上，这个工资是清楚界定的，但对一个特定的人来说，这个工资是很难观测到的，甚至是不可能观测到的。虽然对某给定个人，诸如其被抓住的概率之类的变量也不能切实得到，但至少我们能找到相关的逮捕统计量，从而推导出一个能近似被抓住概率的变量。还有其他影响犯罪行为的因素，不要说观测，我们甚至连列出来都做不到，但我们多少都要对它们做出解释。

通过设定一个特定的计量经济模型，我们就解决了经济模型中内在的不确定性：

$$crime = \beta_0 + \beta_1 wage_m + \beta_2 othinc + \beta_3 freqarr + \beta_4 freqconv + \beta_5 avgsen + \beta_6 age + u \quad (1.3)$$

其中，

$crime$ = 参与犯罪活动频率的某种度量，

$wage_m$ = 在合法就业中所得到的工资，

$othinc$ = 通过其他途径得到的收入（如资产、继承等），
 $freqarr$ = 以前违法被抓住的概率（用来近似被捕概率），
 $freqconv$ = 被证明有罪的概率，
 $avg\ sen$ = 被证明有罪后判处监禁的平均时间长度。

对这些变量的选择，既以经济理论为依据，又有数据方面的考虑。 u 这一项则包括了不可观测的因素，诸如从事犯罪活动的工资、道德特征、家庭背景等，以及在度量犯罪活动和被捕概率等变量时的误差。虽然我们也可以在模型中加入家庭背景变量，如兄弟姐妹的个数、父母所受教育等，但我们仍不能完全消除 u 。实际上，对这个误差项（error term）或干扰项（disturbance term）的处理，可能是任何计量分析中最重要的内容。

常数 $\beta_0, \beta_1, \dots, \beta_6$ 都是这个计量模型的参数，它们描述了此模型中犯罪决定因素与犯罪之间关系的方向和强度。

对例 1.2 来说，一个完整的计量经济模型可能是

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 training + u \quad (1.4)$$

其中， u 这一项包含的因素有“天生能力”、教育质量、家庭背景等，以及能影响一个人工资的无数其他因素。如果我们专门考虑工作培训的影响，那 β_3 就是我们所关注的参数。

在多数情况下，计量经济分析是从对一个计量经济模型的设定开始的，而没有考虑模型构造的细节。我们一般遵循这一思路，主要原因是，对犯罪这种经济模型进行仔细推导，不仅消耗的时间过长，而且会把我们带到经济理论的某个特定而通常又极为困难的领域。在我们的例子中，经济上的逻辑将起作用，而且我们要把其背后的任何一种经济理论都放到对计量模型的设定中。而在犯罪一例的经济模型中，我们将从像方程（1.3）那样的一个计量模型出发，并以经济逻辑和常识作为选择变量的向导。尽管这一方法有失经济分析之丰富，但它总是被细心的研究者普遍而又有效地应用着。

一旦设定了一个像式（1.3）或式（1.4）那样的计量模型，我们所关心的各种假设便可用未知参数来表述。比如，在方程（1.3）中，我们可以假设合法就业的工资 $wage_m$ 对犯罪行为没有影响。在这个特定的计量模型背景下，这个假设就等价于 $\beta_1 = 0$ 。

按定义，一项经验分析总需要数据。在搜集到相关变量的数据之后，便用计量方法来估计计量模型中的参数，并规范地检验所关心的假设。在某些情况下，计量模型还用于对理论的检验或对政策影响的研究。

由于数据搜集在经验工作中如此重要，所以 1.3 节将专门介绍我们可能会遇到的数据类型。

1.3 经济数据的结构

经济数据的类型各式各样。尽管我们可以在对许多不同的数据集不做任何修改的情况下，使用有些计量方法，但仍有必要对某些数据集的特殊性质进行阐释并加以利用。接下来，我们就描述在应用研究中可能会遇到的几种重要的数据结构。

□ 横截面数据

所谓**横截面数据集**（cross-sectional data set），就是在给定时点对个人、家庭、企业、城市、州、国家或一系列其他单位采集的样本所构成的数据集。有时，所有单位的数据并非完全对应于同一时间段。例如，几个家庭可能在一年中的不同星期被调查。在一个纯粹的横截面分析中，我们应该忽略数据搜集中细小的时间差别。如果一系列家庭都是在同一年度的不同星期被调查的，那我们仍视之为横截面数据集。

横截面数据的一个重要特征是，我们通常可以假定，它们是从样本背后的总体中通过**随机抽样**（random sampling）而得到的。例如，如果我们通过随机地从工人总体中抽取 500 人，并得到其有关工资、受教育程度、工作经历和其他特征方面的信息，那我们就得到所有工人构成的总体的一个随机样本。随机抽样是初级统计学教程中所讲授的抽样方案，而且它使得对横截面数据的分析大为简化。附录 C 对随机抽样进行了复习。

有时，以随机抽样作为对横截面数据的一个假定并不适当。例如，假设我们对研究影响家庭财富积累的因素感兴趣，虽然我们可以调查家庭的一个随机样本，但有些家庭可能拒绝报告其财富。比方说，如果越是富裕的家庭就越不愿意暴露其财富，那么由此得到的财富样本就不是由所有家庭构成的总体的一个随机样本。这是对样本选择问题的一个解释，我们还将第 17 章专门讨论这个高级专题。

当我们抽取的样本（特别是地理上的样本）相对总体而言太大时，可能会导致另一种偏离随机抽样的情况。这种情形中潜在的问题是，总体不够大，所以不能合理地假定观测值是独立抽取的。例如，如果我们想用工资率、能源价格、公司和财产税、所提供的服务、工人的质量及其他有关州的特征，来解释跨州的新兴商业活动，那么，邻近州之间的商业活动不可能相互独立。虽然事实表明，我们所讨论的计量方法在这种情形下确实能起作用，但有时这些方法还需要提炼。多数情况下，我们都放弃分析这种情形所引起的错综复杂的关系，而在随机抽样的框架下处理这些问题，尽管有时这样做在技术上是错误的，我们也只好如此。

横截面数据被广泛地应用于经济学和其他社会科学领域之中。在经济学中，横截面数据分析与应用经济领域有密切联系，如劳动经济学、州和地方公共财政学、产业组织理论、城市经济学、人口和健康经济学。对检验微观经济假设和评价经济政策而言，在某一给定时点上，有关个人、家庭、企业和城市的数据都是至关重要的。

用于计量经济分析的横截面数据可以在计算机上表示和存储。表 1.1 以缩略的形式给出了 1976 年 526 个工人的横截面数据集。（这是文件 WAGE1.RAW 中数据的一个子集。）变量包括 *wage*（美元/小时）、*educ*（受教育年数）、*exper*（潜在劳动经历的年数）、*female*（性别指标）和 *married*（婚姻状况）。后面两个变量在本质上是二值（0—1）变量，并表明个人的定性特征（这个人是否为女性；这个人是否结婚）。在第 7 章及以后，我们会详谈二值变量。

表 1.1 有关工资和其他个人特征的横截面数据集

<i>obsno</i>	<i>wage</i>	<i>educ</i>	<i>exper</i>	<i>female</i>	<i>married</i>
1	3.10	11	2	1	0
2	3.24	12	22	1	1
3	3.00	11	2	0	0
4	6.00	8	44	0	1
5	5.30	12	7	0	1
⋮	⋮	⋮	⋮	⋮	⋮
525	11.56	16	5	0	1
526	3.50	14	5	1	0

表 1.1 中的变量 *obsno* 是赋予样本中每个人的观测序号，与其他变量不同，它不是个人特征。所有的计量经济学和统计学软件包，对每个数据单位都会指定一个观测序号。凭直觉，对表 1.1 中的数据而言，哪个被标为观测 1，哪个被标为观测 2，等等，应该是没有区别的。数据排序不影响计量分析这一事实，是由随机抽样而得到横截面数据集的一个重要特征。

在横截面数据集中，不同的变量有时对应于不同的时期。例如，为了决定政府政策对长期经济增长的影响，经济学家研究了一定时期（如 1960 年至 1985 年）真实人均国内生产总值（GDP）的增长率与部分由 1960 年的政府政策（政府消费占 GDP 的百分比和成人中受过中级教育的百分比）所决定的变量之间的关系。这样一个数据集可由表 1.2 给出，这些正是 De Long and Summers（1991）对各国经济增长率的研究中所用到的数据集的一部分。

表 1.2 经济增长率和国家特征方面的数据集

<i>obsno</i>	<i>country</i>	<i>gpcrgdp</i>	<i>govcons60</i>	<i>second60</i>
1	阿根廷	0.89	9	32
2	澳大利亚	3.32	16	50
3	比利时	2.56	13	69
4	玻利维亚	1.24	18	12
⋮	⋮	⋮	⋮	⋮
61	津巴布韦	2.30	17	6

变量 *gpcrgdp* 表示 1960 年至 1985 年间真实人均 GDP 的平均增长率。尽管 *govcons60* (政府消费占 GDP 的百分比) 和 *second60* (成人中受过中等教育的百分比) 都对应于 1960 年, 而 *gpcrgdp* 则是从 1960 年至 1985 年的平均增长率, 但把这些信息看成横截面数据集, 并不会导致任何特别的问题。虽然观测值以国家名称的字母排序, 但这种排序对以后的分析没有任何影响。

□ 时间序列数据

时间序列数据 (time series data) 集是由对一个或几个变量不同时间的观测值所构成。时间序列数据方面的例子包括股票价格、货币供给、消费者价格指数、国内生产总值、年度谋杀率和汽车销售数量。由于过去的事件可以影响到未来的事件, 而且行为滞后在社会科学中又相当普遍, 所以时间是时间序列数据集中的一个重要维度。与横截面数据的排序不同, 时间序列对观测值按时间先后排序, 这也传递了潜在的重要信息。

时间序列数据有一个关键的特征, 使得对它的分析比对横截面数据的分析更为困难, 这个特征就是如下事实, 即很少 (即使能够) 假设经济数据的观测独立于时间。多数经济及其他时间序列都与其近期历史相关 (通常是高度相关)。比如, 由于 GDP 的趋势从这个季度到下个季度保持着相当的稳定性, 所以对上一季度 GDP 的一些了解, 会告诉我们本季度 GDP 的可能范围。虽然多数计量程序既能用于横截面数据, 又能用于时间序列数据, 但在认为标准的计量方法适用之前, 要设定时间序列数据的计量模型, 还有更多工作要做。此外, 为了解释和利用经济时间序列的相互依赖性, 并解决某些经济变量常表现出清晰的时间趋势等问题, 对标准的计量方法还要进行改进和润色。

需要特别注意的是, 时间序列数据有另一个特征, 即数据搜集时的**数据频率** (data frequency), 最常见的频率是每天、每周、每月、每个季度和每年。股票价格以天为区间进行记录 (星期六和星期日除外)。美国的货币供给是逐月报告的。许多宏观经济序列都是按月列出, 如通货膨胀率和就业率。其他宏观序列报告得就没那么频繁, 比如国内生产总值就是典型地每三个月 (一个季度) 报告一次。其他的时间序列, 如美国各州的婴儿死亡率等, 则只有年度数据可供使用。

许多按周、按月、按季度报告的经济时间序列都表现出很强的季节类型, 这正是时间序列分析中的一个重要因素。比如, 各个月份的新住房动工数据之不同, 仅仅是因为气温条件的变化所致。在第 10 章中, 我们将会了解如何处理这种季节性时间序列。

表 1.3 包含的时间序列数据集, 来自 Castillo-Freeman and Freeman (1992) 研究波多黎各的最低工资条例影响的一篇论文。此数据集中最早的一年就是第一次观测, 最后一年则是最后一次观测。在用计量方法分析时间序列数据时, 数据应该按时间顺序排列。

表 1.3 波多黎各的最低工资、失业及相关数据

<i>obsno</i>	<i>year</i>	<i>avgmin</i>	<i>avgcov</i>	<i>unemp</i>	<i>gnp</i>
1	1950	0.20	20.1	15.4	878.7
2	1951	0.21	20.7	16.0	925.0
3	1952	0.23	22.6	14.8	1 015.9
⋮	⋮	⋮	⋮	⋮	⋮
37	1986	3.35	58.1	18.9	4 281.6
38	1987	3.35	58.2	16.8	4 496.7

变量 *avgmin* 表示当年的最低工资, *avgcov* 表示最低工资的覆盖率(最低工资法所涵盖的工人占工人总数的百分比), *unemp* 是失业率, 而 *gnp* 则是国民生产总值。我们将在以后研究最低工资对就业之影响的一个时间序列分析中使用这些数据。

□ 混合横截面数据

有些数据既有横截面数据的特点, 又有时间序列的特点。例如, 假设对美国的家庭进行了两次横截面数据的调查, 一次在 1985 年, 一次在 1990 年。在 1985 年, 对家庭的一个随机样本调查了工资、储蓄、家庭规模等变量。到了 1990 年, 用同样的调查问题又对家庭的一个新随机样本进行调查。为了扩大我们的样本容量, 可以将这两年的数据合并成一个混合横截面数据 (pooled cross section)。

把不同年份的横截面数据混合起来, 通常是分析一项新政府政策之影响的有效方法。其思想是, 搜集一个重要的政策变化之前和之后的数据。举例而言, 考虑在 1993 年和 1995 年都对住房价格搜集了数据集, 而在 1994 年则下调了财产税。假设我们在 1993 年有 250 个住房的数据, 在 1995 年有 270 个住房的数据。表 1.4 给出了存储这种数据的一种方式。

表 1.4 混合横截面数据: 两年的住房价格

<i>obsno</i>	<i>year</i>	<i>hprice</i>	<i>proptax</i>	<i>sqrft</i>	<i>bdrms</i>	<i>bthrms</i>
1	1993	85 500	42	1 600	3	2.0
2	1993	67 300	36	1 440	3	2.5
3	1993	134 000	38	2 000	4	2.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮
250	1993	243 600	41	2 600	4	3.0
251	1995	65 000	16	1 250	2	1.0
252	1995	182 400	20	2 200	4	2.0
253	1995	97 500	15	1 540	3	2.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
520	1995	57 200	16	1 100	2	1.5

观测 1 至 250 对应于 1993 年出售的住房,而观测 251 至 520 则对应于 1995 年出售的住房。尽管最终表明,我们如何存放数据的顺序并不重要,但明确每一观测的年份,则通常十分重要,这就是为什么我们把年份作为一个独立变量收入数据集。

对混合横截面数据的分析与对标准横截面数据的分析十分相似,不同之处在于,前者通常要对变量在不同时间的长期差异做出解释。实际上,除了能扩大样本容量之外,混合横截面分析通常是为了让我们看出一个基本关系如何随时间而变化。

□ 面板或纵列数据

面板数据 (panel data) (或纵列数据) 集,是由数据集中每个横截面单位的一个时间序列组成。举例而言,比方说我们对一系列个人的工资、受教育情况和就业历史跟踪了 10 年,或者我们对一系列企业诸如投资和财务数据等搜集了 5 年的信息。有些面板数据也可以以地理上的单位来搜集。比如,在 1980 年、1985 年和 1990 年对美国同一组县,分别搜集其人口迁移、税率、工资率、政府支出等数据。

面板数据有别于混合横截面数据的关键特征是,同一横截面数据的数据单位(上述例子中的个人、企业或县)都被跟踪了一段特定的时期。表 1.4 中的数据不能被看成面板数据集的原因是,在 1993 年和 1995 年出售的住房很可能不同;如果有一些住房相同,那些相同住房的数目可能太小,以至无足轻重。相比之下,表 1.5 则是美国 150 个城市在犯罪及相关统计量方面的一个两年面板数据集。

表 1.5 有几个有趣的特征。首先,每个城市都有一个编号,从 1 到 150。哪个城市被称为城市 1、城市 2 等并不重要。和纯粹的横截面数据一样,面板数据集中横截面单位的排序无关紧要。虽然我们可以用一个数字来代替城市名,但二者都保留通常很有好处。

表 1.5 城市犯罪统计量的一个两年面板数据集

<i>obsno</i>	<i>city</i>	<i>year</i>	<i>murders</i>	<i>population</i>	<i>unem</i>	<i>police</i>
1	1	1986	5	350 000	8.7	440
2	1	1990	8	359 200	7.2	471
3	2	1986	2	64 300	5.4	75
4	2	1990	1	65 100	5.5	75
⋮	⋮	⋮	⋮	⋮	⋮	⋮
297	149	1986	10	260 700	9.6	286
298	149	1990	6	245 000	9.8	334
299	150	1986	25	543 000	4.3	520
300	150	1990	32	546 200	5.2	493

第二个有用的方面是,城市 1 的两年数据占据了观测中的前两行。观测 3 和观测 4 对应于城市 2,等等。由于 150 个城市中的每一个都有两行数据,所以任何一个计量经济软件包都会把它看成是 300 个观测。这个数据集还可以被看成是,将 1986

年和1990年的两个横截面数据混合起来后,由在两个年份都出现的城市所组成。但如我们在第13和14章中将看到的那样,我们还可以利用面板数据结构来回答那些仅把它看成混合横截面数据所不能回答的问题。

在组织表1.5中的观测值时,我们将每个城市两年的数据依次相连,第一年的数据总是在第二年的数据之前。仅从实用的角度来看,这是组织面板数据集的最好方法。可以将这种组织数据的方法与表1.4中存储混合横截面数据的方法进行比较。简言之,像表1.5那样组织面板数据的原因在于,我们需要对每个城市两年的数据进行变换。

由于面板数据要求同一单位不同时期的重复观测,所以要得到面板数据(特别是那些个人、家庭和企业的数据库),比得到混合横截面数据更加困难。对同一观测单位观测一段时间,应该比横截面数据甚至混合横截面数据更有优越性。我们在本书中所强调的好处是,对同一单位的多次观测,使我们能控制个人、企业等观测单位的某些观测不到的特征。如我们将看到的那样,使用不止一次的观测,使我们在单纯横截面数据下很难做出的因果推断得以进行。面板数据的第二个优点是,它通常使我们能够研究决策行为或结果中滞后的重要性。由于预期许多经济政策在一段时间之后才产生影响,所以面板数据所反映的信息就更有意义。

多数本科水平的教材都不讨论面板数据的计量方法。但经济学家现在都认识到,没有面板数据,即便不是不可能,也很难对某些问题做出令人满意的回答。你将会看到,借助简单的面板数据分析(这种不比处理标准横截面数据集困难多少的方法),我们能取得相当可观的进展。

□ 对数据结构的评论

本书的第1篇首先考虑了对横截面数据的分析,因为这方面存在的概念和技术困难最少。同时,它也揭示了计量分析的绝大多数重要主题。在本书的剩余部分,我们还会用到横截面分析所提供的方法和深刻见解。

虽然对时间序列的计量分析也用到许多与横截面分析相同的工具,但由于许多经济时间序列的趋势性和高度持续性,对它的分析将更加复杂。过去用于说明计量方法可用于时间序列数据之方式的例子,现在普遍被认为是缺陷的。由于首先就使用这种例子只会使落后的计量实践进一步加深,所以这种做法没有意义。于是,我们把对时间序列的处理推后到第2篇,到那时还将引入有关趋势、持续性、动态性和季节性等问题。

在第3篇,我们将详尽分析混合横截面数据和面板数据。对独立混合横截面数据和简单面板数据的分析,基本上就是对纯粹横截面数据分析的直接推广。尽管如此,我们还是等到第13章再讨论这些专题。

1.4 计量经济分析中的因果关系和其他条件不变的概念

在多数对经济理论的检验中（当然包括对公共政策的评价），经济学家的目标就是要推定一个变量（比如受教育程度）对另一个变量（如犯罪率或工人的生产力）具有**因果效应**（causal effect）。虽然简单地发现两个或多个变量之间有某种联系很诱人，但除非能得到某种因果关系，否则这种联系很难令人信服。

其他条件不变（*ceteris paribus*）——意味着“其他（相关）因素保持不变”——的概念，在因果分析中有重要作用。虽然这种思想早已暗含在我们前面的某些讨论之中，特别是例 1.1 和例 1.2，但迄今为止，我们还没有明确提出。

你可能记得，在初级经济学教程中，很多经济学问题都有其他条件不变的特征。比如，在分析消费需求时，我们想知道一种商品价格的变化对其需求量的影响，而让所有其他因素——譬如收入、其他商品的价格和个人偏好等——都保持不变。如果不保持其他因素不变，那我们就不能知道价格变化对需求量的因果效应。

保持其他因素不变的做法对政策分析也至关重要。在工作培训的例子（例 1.2）中，可能我们感兴趣的是，在其他因素（特别是受教育程度和工作经历）不变的情况下，增加一周工作培训对工资的影响。如果我们能保持所有其他相关因素不变，然后发现工作培训与工资之间有一种联系，那么我们就断定，工作培训对工人的生产力具有因果效应。尽管看上去相当简单，但即便在现阶段也应该清楚，除非在极为特殊的情形下，否则不可能如实地保持所有其他因素不变。多数经验研究中的一个关键问题是：对于做出一个因果推断而言，是否有足够多的其他因素被保持不变？对一项计量经济研究进行评价，都要提到这个问题。

在多数重大的应用中，能影响我们所关注变量——如犯罪活动或工资——的因素为数众多，而想要隔离任何一个特定的变量，看来都像是无谓的努力。但最终我们将看到，只要善加利用，用计量经济方法就可以模拟一个其他条件不变的实验。

就此看来，我们还不能解释如何使用计量经济方法去估计那种保持其他条件不变的影响，所以我们将考虑在做出经济学中的因果推断时可能出现的某些问题。在这种讨论中，我们不会使用任何方程。对于每一个例子，如果能进行适当的实验，推断因果关系的问题便不复存在。因此，最好描述一下如何才能构造一个这样的实验，并观察到，在多数情况下，要得到实验数据的想法都是不切实际的。但考虑一下，为什么可供使用的数据不具备一个实验数据集的重要特征，也颇有益处。

迄今为止，我们都只是靠直觉来理解诸如随机、独立和相关等术语，而这些在初级概率论与数理统计教程中都有详细介绍。（附录 B 回顾了这些概念。）我们先举一个例子来说明这些重要问题。

例 1.3

肥料对作物收成的影响

一些早期的计量研究 [例如 Griliches (1957)] 考虑了新肥料对谷物收成的影响。假设所考虑的作物是大豆。由于施肥量只是影响收成的因素之一——其他因素还包括降雨量、土地质量和攀附植物的出现等, 所以这个问题必须其他条件不变的情况下进行研究。决定施肥量对大豆收成之因果效应的途径之一, 是按如下步骤进行一项实验。选择几块面积为一英亩的土地, 再对每块土地施加不同数量的肥料, 然后度量每块土地的收成; 这就形成一个横截面数据。于是, 使用统计方法 (在第2章将介绍), 就可以测度收成与施肥量之间的关系。

如前所述, 这个实验看起来并不是很好, 因为我们并没提到, 在选择实验田时, 要保证除施肥量不同外其他方面都要完全一样。实际上, 选择这种实验田是不可行的: 诸如土地质量等某些其他因素甚至都不能被充分观测到。我们如何知道这个实验的结果可用来度量施肥量在其他条件不变的情况下的影响呢? 答案取决于如何选择施肥量的详细情况。如果在确定每一块土地的施肥量时, 施肥量的选择独立于影响收成的其他土地特征——即在决定施肥量时完全忽略土地的其他特征, 那么我们就可以达到目标。在第2章, 我们将证明这种观点的正确性。

对于应用经济学中推断因果关系时出现的困难, 下面这个例子更具代表性。

例 1.4

测度教育的回报

劳动经济学家和政策制定者长期以来都对“教育的回报”有浓厚兴趣。这个问题可以多少有些不太规范地这样提出: 如果从总体中选择一个人, 并让他或她多受一年教育, 那么, 他或她的工资会提高多少? 像前面的例子一样, 这也是一个其他条件不变的问题, 即意味着在保持所有其他条件不变的情况下, 对这个人增加一年的教育。

为了回答这个问题, 我们可以设想有一位社会计划者在设计一项实验, 这与农业研究者设计一项实验来估计施肥量的影响很相似。模仿例1.3中施肥量实验的一种方法是: 选择一群人, 随机地给每个人一定的教育程度 (给有些人8年级的教育, 给有些人高中教育等), 然后度量他们的工资 (假设这时每人都有工作)。这里的人就像上例中的实验田, 教育则起到肥料的作用, 而工资则类似于大豆的收成。像例1.3一样, 如果受教育水平的确定独立于影响生产力的其他特征 (如工作经历和天生能力), 那么忽略这些其他因素的分析将会得到有用的结论。同样, 第2章将花点工夫说明这个观点; 到目前为止, 我们只是在没有任何根据的情况下这么说。

与肥料—收成的例子不同, 例1.4中描述的实验是不可行的。不用说经济上的成本, 要对一群人随机地决定其受教育水平, 在道德上也明显成问题。作为一个逻辑

辑上的问题,如果一个人已有大学文凭,我们就不可能只给他8年级的教育。

尽管不能得到度量教育回报的实验数据,但我们肯定可以从工作人群总体中随机地抽取一大群人,并搜集其教育水平和工资的非实验数据,但这些数据所具有的特点,使它们很难用于估计教育在其他条件不变情况下的回报。人们选择其自身受教育的水平,因此受教育水平可能并不独立于所有那些影响工资的因素。这是大多数非实验数据集所共有的一个特点。

影响工资的因素之一就是工作经历。由于追求更多的教育通常需要推迟加入劳动力队伍的时间,所以那些受教育越多的人,工作经历就越短。因此,在一个有关工资和受教育水平的非实验数据集中,受教育水平可能与影响工资的一个关键变量呈负相关。还要知道,一个人越有天赋,就会选择接受越多的教育。由于更强的能力导致更高的工资,所以我们又发现受教育水平与影响工资的另一个重要因素呈相关关系。

在例1.4中省略了工作经历和个人能力等因素,在例1.3中也有相似的因素。工作经历通常易于度量,因此与降雨量相似。另一方面,能力则含糊不清而又难以量化;所以它与例1.3中的土地质量相似。正如本书中将要看到的那样,在估计教育这种变量在其他因素不变情况下的影响时,分离出像工作经历这种可观测度量的影响相对简单。我们还将发现,要分离像个人能力这种本来就难以观测的因素的影响,则相当成问题。公正地讲,计量经济方法的许多进展,都试图对计量模型中的这些无法观测因素进行处理。

例1.3和例1.4之间可以得到一个可比拟之处。试想,在例1.3中,施肥量并非完全随机决定。相反,选择施肥数量的助手认为,最好在较高质量的田里施更多的肥料。(尽管这些农业研究者不能对这些土地之间的差别充分量化,但他们应该大致知道哪一块土地的质量较好。)这种情形和例1.4中受教育水平与无法观测的个人能力相关的情况完全类似。因为较好的土地会导致较好的收成,而在较好的土地上又使用了较多的肥料,所以观察到的收成与施肥量之间的关系可能是有欺骗性的。

例 1.5

执法对城市犯罪活动的影响

如何最好地制止犯罪问题已伴随了我们一段时间,并将继续成为我们考虑的问题。这方面一个特别重要的问题是:更多警察出现在街上会制止犯罪吗?

这个其他条件不变情况下的问题很容易表述:如果随机地选择一个城市并增加10名警察,那么犯罪率会下降多少?表述这个问题的另一种方法是:如果两个城市各方面都相同,只是A市比B市多10名警察,那么这两个城市的犯罪率会有多大的差别呢?

要找到除警力规模外完全一样的两个社区,根本就不可能。幸运的是,计量经济分析并不要求这样。我们真正需要知道的是,我们所搜集到的有关社区犯罪水平和警力规模的数据,是否可被视为实验数据。我们确实可以设想一个真正的实验,其中有许多城市,而每个城市在下一年度将使用多少警力则完全由我们控制。

尽管可以利用政策来影响警力规模,但我们显然不能告诉每个城市该雇用多少警察。

如果（很可能）一个城市有关雇用多少警察的决策与另一个城市影响犯罪的因素相关，那么，这些数据就必须被看作是非实验性数据。实际上，看待这个问题的方法之一是，将一个城市对警力规模的选择与该城市的犯罪量看成同时决定的（simultaneously determined）变量。我们将在第16章清楚地解决这种问题。

我们前面讨论过的三个例子已探讨了各个层次（如个人或城市等层次）的横截面数据。在时间序列问题中进行因果推断时也会出现同样的障碍。

例 1.6

最低工资条例对失业的影响

一个重要但可能有争议的政策问题是，最低工资条例对各种工人群体的失业率有什么影响？尽管这个问题可以在各种数据背景（横截面数据、时间序列数据或面板数据）下进行研究，但通常都是使用时间序列来考察其总影响的。表1.3给出了失业率和最低工资方面时间序列数据集的一个例子。

标准的供求理论意味着，当最低工资被提高到市场出清工资之上时，劳动需求便会沿着劳动需求曲线向上移动，从而导致总就业减少。（劳动供给大于劳动需求。）为了定量分析这种影响，我们可以研究不同时间就业与最低工资之间的关系。除了处理时间序列数据可能引起某些特殊问题之外，在进行因果推断时可能还有一些问题。美国的最低工资并非在隔绝状态下确定，所以各种经济和政治力量对任一给定年份最终的最低工资都有影响。（一旦确定了一个最低工资，通常在几年内都不会改变，除非其相对于通货膨胀而被指数化。）因此，最低工资的大小很可能与影响就业水平的其他因素存有某种联系。

我们可以设想，美国政府在实施一项实验，以决定最低工资对就业的影响（以消除人们对低工资工人福利的担忧）。政府每年可以随机地设定最低工资，然后就业结果可以被制成表格。于是，使用相当简单的计量分析方法，就可以对如此得到的实验性时间序列数据进行分析。但这种方案无法描述最低工资是如何设定的。

如果我们能充分控制与就业相关的其他因素，那我们就仍有希望去估计最低工资在其他条件不变情况下对就业的影响。在这个意义上，这个问题与前面横截面数据的例子极为相似。

即便经济理论不能很自然地用因果关系进行描述，使用计量经济方法通常也能对它们的预测进行检验。如下即是这种方法之一例。

例 1.7

预期假说

来自金融经济学的预期假说认为，给定投资者在进行投资决策时可供使用的全部信息，任何两种投资的期望回报都相同。比如，考虑两个可能的投资，投资的时间长度为

三个月，并在同时进行：(1) 购买一份面值 10 000 美元三月期国债，其价格低于 10 000 美元；到了三个月，你就得到这 10 000 美元。(2) 购买一份六月期国债（价格低于 10 000 美元），到了三个月，把它作为一份三月期国债售出。虽然每种投资所需要的初始资本大致相当，但二者之间存在重大差别。对于第一项投资，由于你不仅知道三月期国债的面值，而且还知道其价格，所以你在购买时就确切地知道其回报。对第二项投资则不然：虽然你在购买时就知道六月期国债的价格，但你并不知道在三个月后的价格。因此，对于一个投资期限只有三个月的人来说，这种投资具有不确定性。

这两种投资的实际回报通常存有差异。根据预期假说，给定投资时的全部信息，第二项投资的预期收益应该与购买三月期国债的收益相同。像我们在第 11 章将看到的那样，这个理论相当容易检验。

小结

在导论中，我们已经讨论了计量经济分析的目的和研究范围。所有的应用经济学领域都使用计量经济学来检验经济理论，指导政府和私人政策制定者，并预测经济时间序列。虽然计量经济模型有时也可以从规范的经济模型推导出来，但在另一些情况下，计量经济模型都是以非规范的经济推理和直觉为基础。任何一个计量经济分析的目的，都是估计模型中的参数，并检验对这些参数的假设；参数的数值和符号，决定了经济理论的有效性和某项政策的效果。

横截面数据、时间序列数据、混合横截面数据和面板数据，是应用计量经济学中最经常使用的数据结构类型。由于大多数经济时间序列都存在跨时相关关系，所以对涉及时间维的数据集（如时间序列和面板数据），都需要特别处理。另外，诸如趋势性和季节性等问题，也只会在对时间序列数据分析中出现，横截面数据则不存在这种问题。

在 1.4 节，我们讨论了其他条件不变和因果推断的概念。在多数情况下，社会科学中的假设都具有“其他条件不变”的特点：在研究两个变量之间的关系时，所有其他的相关因素都必须固定不变。因为社会科学中所搜集到的多数数据都具有非实验特征，所以发现其中的因果关系极具挑战性。

关键术语

因果效应

其他条件不变

横截面数据集

数据频率

计量经济模型

经济模型

经验分析

实验数据

非实验数据

观测数据

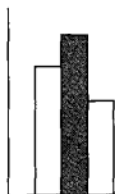
面板数据

混合横截面数据

随机抽样

回顾数据

时间序列数据



简单回归模型可以用来研究两个变量之间的关系。出于我们将会看到的原因，简单回归模型要作为经验分析的一般工具还存在着局限性。不过有时把它作为一个经验工具还是很合适的。学会解释简单回归模型，对于我们接下来几章要学习的多元回归模型，无疑也是很好的练习。

2.1 简单回归模型的定义

许多应用计量经济分析都是从如下假设前提开始的： y 和 x 是两个代表某个总体的变量，我们感兴趣的是“用 x 来解释 y ”，或“研究 y 如何随 x 而变化”。我们在第1章中讨论了一些例子，其中包括： y 是大豆的产出， x 是施肥量； y 是每小时的工资， x 是受教育的年数； y 是社区的犯罪率， x 是警察的数量。

在写出用 x 解释 y 的模型时，我们要面临三个问题。首先，既然两个变量之间没有一个确切的关系，那么我们应该如何考虑其他影响 y 的因素呢？第二， y 和 x 的函数关系是怎样的呢？第三，我们何以确定我们在其他条件不变的情况下刻画了 y 和 x 之间的关系（如果这是一个理想目标的话）呢？

我们可以通过写出关于 y 和 x 的一个方程来消除这些疑惑。一个简单的方程是：

$$y = \beta_0 + \beta_1 x + u \quad (2.1)$$

假定方程(2.1)在我们所关注的总体中成立,它便定义了一个简单线性回归模型(simple linear regression model)。因为它把两个变量 x 和 y 联系起来,所以又把它叫做两变量或者双变量线性回归模型。我们现在来讨论等式(2.1)中每个量的含义。[“回归”一词的起源,对于大部分现代计量经济学的应用来说并非特别重要,所以,我们在此不予讨论。参见 Stigler (1986) 对回归分析充满魅力的历史所做的介绍。]

若通过方程(2.1)联系起来,则变量 y 和 x 就有许多可以互换的不同名称。 y 被称为**因变量**(dependent variable)、**被解释变量**(explained variable)、**响应变量**(response variable)、**被预测变量**(predicted variable)或者**回归子**(regressand)。 x 则被称为**自变量**(independent variable)、**解释变量**(explanatory variable)、**控制变量**(control variable)、**预测变量**(predictor variable)或者**回归元**(regressor)。[x 还被称为**协变量**(covariate)。]“因变量”和“自变量”两词在计量经济学中使用较多,但要注意这里所说的“自变”(independent)与统计学概念里面随机变量之间的“独立”有所不同(见附录 B)。

“被解释”和“解释”变量这两个词可能是最具描述性的。“响应”和“控制”在实验性科学中运用最多,其中 x 变量在实验者的控制之下。我们不使用“被预测变量”和“预测变量”这样的字眼,尽管有时候在一些纯粹预测而非因果推断的应用研究中会看到它们。在表 2.1 中,我们总结了简单回归中的术语。

表 2.1 简单回归的术语

y	x
因变量	自变量
被解释变量	解释变量
响应变量	控制变量
被预测变量	预测变量
回归子	回归元

变量 u 被称为关系式中的**误差项**(error term)或者**干扰项**(disturbance),表示除 x 之外其他影响 y 的因素。简单回归分析有效地把除 x 之外其他所有影响 y 的因素都看成无法观测的因素。你也可以把 u 看作是“观测不到”的因素。

方程(2.1)还表现出 y 和 x 之间的函数关系。若 u 中的其他因素都保持不变,于是 u 的变化为零,即 $\Delta u = 0$,则 x 对 y 具有线性影响,其表述如下:

$$\text{若 } \Delta u = 0, \text{ 则 } \Delta y = \beta_1 \Delta x \quad (2.2)$$

因此, y 的变化量无非就是 β_1 与 x 的变化量之积。这就是说,保持 u 中其他因素不变, β_1 就是 y 和 x 的关系式中的**斜率参数**(slope parameter);在应用经济学中,它是人们研究的主要兴趣所在。有时被称作常数项的**截距参数**(intercept parameter)的 β_0 也有其作用,但很少被当作分析的核心。