

Chapter 1

计量经济学的性质与经济数据



章节框架

- 了解什么是计量经济学以及计量经济学的基本目标
- 掌握经验经济分析的步骤（理解经济模型与计量模型的区别）
- 区分几种常见经济数据的结构
- 掌握因果关系和其他条件不变的概念

什么是计量经济学

- 什么是计量经济学？
 - 计量经济学：使用统计方法来分析经济数据
经济数据：股票收益率、收入、GDP等
 - 计量经济学一般分析非实验数据（nonexperimental data）
实验数据：在实验环境中获得，（在社会科学研究中）通常很难获取
- 计量分析的基本目标
 - 估计经济变量间的关系
 - 检验经济理论与假设
 - 预测经济变量
 - 评估政府政策或商业决策

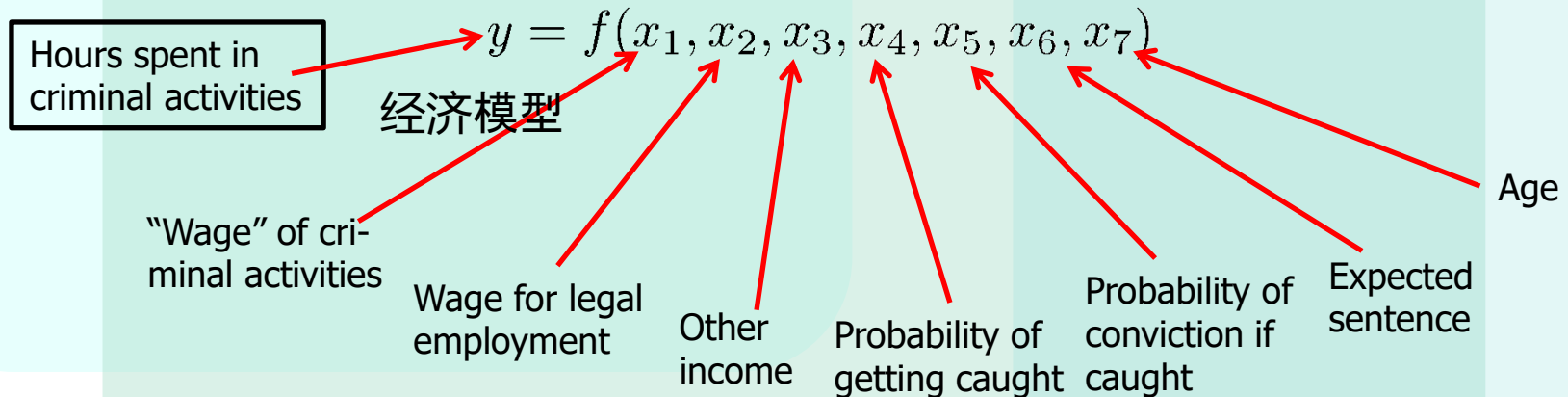
经验经济分析的步骤

- 经验分析（实证分析，empirical analysis）：利用数据来检验某个理论或估计某种关系。
- 主要步骤：
 - 1) 经济模型
 - 也许是微观或宏观理论模型
 - 例如一般均衡理论，消费者选择理论等
 - 建立经济变量间的关系
 - 例如定价等式、需求曲线等
 - 2) 计量模型
 - 解决经济模型在应用中的问题
 - 考虑不确定性

（经济模型的步骤通常被忽略）

经验经济分析的步骤

- 犯罪的经济模型(Becker (1968))
 - 建立基于效用最大化的犯罪行为等式



- 一般来说，函数 f 未进行任何设定；变量数据可获得？
- 在没有经济模型的情况下，这个方程本仍可以假设

经验经济分析的步骤

- 犯罪行为的计量模型
 - 函数形式必须设定
 - 变量可能必须用其他定量来近似

The diagram illustrates a regression model for criminal activity. The equation is: $crime = \beta_0 + \beta_1 wage_m + \beta_2 othinc + \beta_3 freqarr + \beta_4 freqconv + \beta_5 avgsen + \beta_6 age + u$. Red arrows point from descriptive labels to each term in the equation: 'Measure of criminal activity' points to 'crime'; 'Wage for legal employment' points to 'wage_m'; 'Other income' points to 'othinc'; 'Frequency of prior arrests' points to 'freqarr'; 'Frequency of conviction' points to 'freqconv'; 'Average sentence length after conviction' points to 'avgsen'; 'Age' points to 'age'; and 'Unobserved determinants of criminal activity' points to 'u'. A separate box lists examples for 'u': 'e.g. moral character, wage in criminal activity, family background ...'.

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

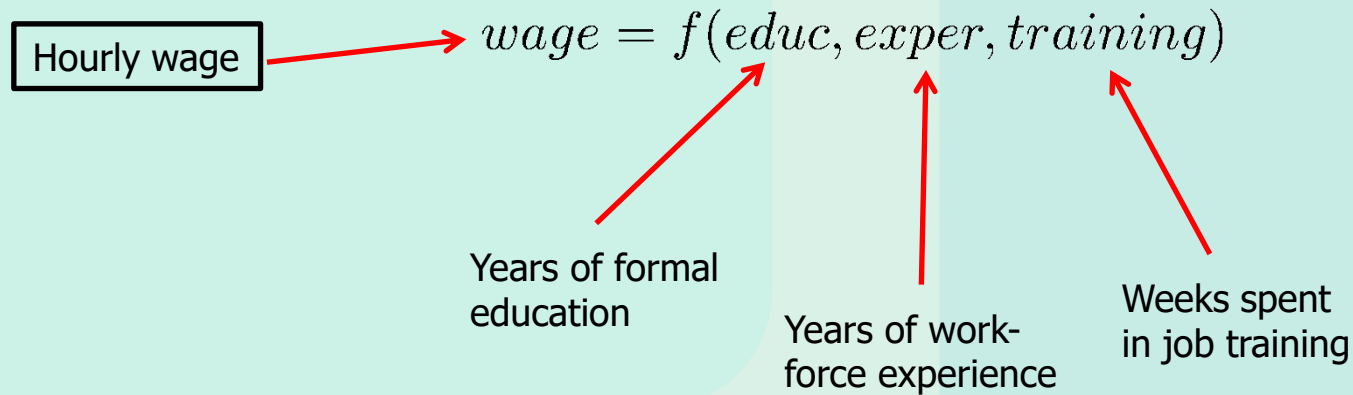
Labels and arrows:

- Measure of criminal activity → $crime$
- Wage for legal employment → $wage_m$
- Other income → $othinc$
- Frequency of prior arrests → $freqarr$
- Frequency of conviction → $freqconv$
- Average sentence length after conviction → $avgsen$
- Age → age
- Unobserved determinants of criminal activity → u

Examples for u : e.g. moral character, wage in criminal activity, family background ...

经验经济分析的步骤

- 工作培训与工人的生产力模型
 - 额外的培训会对工人的生产力产生何种影响？



- 不一定需要经济理论给出上面等式
- 其他的因素也许也是相关的

经验经济分析的步骤

- 工作培训与工人的生产力计量模型

$$wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 training + u$$

Hourly wage

Years of formal education

Years of work-force experience

Weeks spent in job training

Unobserved determinants of the wage

e.g. innate ability, quality of education, family background ...

- 绝大多数计量经济学模型都涉及对误差 u 的具体说明
- 计量经济模型可以用于假设检验
 - 例如，参数 β_3 代表“培训对工资的影响”
 - 这个影响有多大？是否异于0？

经济数据的结构

- 计量分析需要数据
- 不同种类的经济数据集
 - 横截面数据 (Cross-sectional data)
 - 时间序列数据 (Time series data)
 - 混合截面数据 (Pooled cross sections)
 - 面板/纵列数据 (Panel/Longitudinal data)
- 计量经济学方法取决于所用数据的性质
 - 使用不合适的方法将导致具有误导性的结果

经济数据的结构

- **横截面数据集 (Cross-sectional data sets)**
 - **特定时间点/特定时期内**个人、家庭、公司、城市、州、国家或其他单位的样本
 - 横截面观察值或多或少是独立的
 - 例如，从人群中进行纯随机抽样
 - 有时纯随机抽样会被违反，例如有人拒绝在调查中做出回应
 - 横截面数据一般应用于应用微观经济学中

经济数据的结构

- 工资与其他个体特征的横截面数据

TABLE 1.1 A Cross-Sectional Data Set on Wages and Other Individual Characteristics

obsno	wage	educ	exper	female	married
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

Indicator variables
(1 = yes, 0 = no)

Observation number

Hourly wage

Years of
education

Years of
experience

© Cengage Learning, 2016

经济数据的结构

- 经济增长率与国家特征的横截面数据

TABLE 1.2 A Data Set on Economic Growth Rates and Country Characteristics

obsno	country	gpcrgdp	govcons60	second60
1	Argentina	0.89	9	32
2	Austria	3.32	16	50
3	Belgium	2.56	13	69
4	Bolivia	1.24	18	12
.
.
.
61	Zimbabwe	2.30	17	6

Average growth rate of real
per capita GDP

Government consumption
as a percentage of GDP

Adult secondary
education rates

经济数据的结构

- 时间序列数据
 - 一个或多个变量的观测值是时变的 (over time)
 - 例如，股票价格、货币供给、CPI、GDP、年谋杀率、汽车销量等.....
 - 时间序列观测值一般是序列相关的
 - 观察结果的顺序传递了重要信息
 - 数据频率：日度、周度、月度、季度、年度.....
 - 时间序列的典型特征：趋势与季节性
 - 一般应用：应用宏观经济学与金融

经济数据的结构

- 工作与相关变量的时间序列

TABLE 1.3 Minimum Wage, Unemployment, and Related Data for Puerto Rico

obsno	year	avgmin	avgcov	prunemp	prgnp
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	1015.9
.
.
.
37	1986	3.35	58.1	18.9	4281.6
38	1987	3.35	58.2	16.8	4496.7

© Cengage Learning, 2016

Average minimum
wage for the given year

Average
coverage rate

Unemployment
rate

Gross national
product

经济数据的结构

- 混合横截面
 - 两个或多个横截面集中在一个数据集中
 - 横截面的收集是相互独立的
 - 混合横截面经常用于评估政策变化
 - 例子：
 - 评估房产税变化对房价的影响
 - 1993年房价随机抽样
 - 1995年房价的新随机样本
 - 比较政策前后 (1993: 改革前、1995: 改革后)

经济数据的结构

- 房屋价格的混合截面

TABLE 1.4 Pooled Cross Sections: Two Years of Housing Prices

obsno	year	hprice	proptax	sqrft	bdrms	bthrms
1	1993	85,500	42	1600	3	2.0
2	1993	67,300	36	1440	3	2.5
3	1993	134,000	38	2000	4	2.5
.
.
.
250	1993	243,600	41	2600	4	3.0
251	1995	65,000	16	1250	2	1.0
252	1995	182,400	20	2200	4	2.0
253	1995	97,500	15	1540	3	2.0
.
.
.
520	1995	57,200	16	1100	2	1.5

Property tax

Size of house
in square feet

Number of bedrooms

Number of bathrooms

Before reform

After reform

© Cengage Learning, 2016

经济数据的结构

- 面板或纵列数据 (**Panel or longitudinal data**)
 - 数据集中每个横截面单位为一个时间序列
 - 面板数据具有横截面和时间序列两个维度
 - 不同于混合截面，面板数据每个对象被跟踪观测一段特定时期
 - 面板数据可以用来解释不随时间变化的不可观测特征（固定效应）
 - 面板数据可以用来对滞后反应建模
 - 例子：
 - 城市犯罪统计；每个城市观测两年
 - 不随时间变化的城市不可见特征可能会被建模
 - 警察对犯罪率的影响可能表现出时滞

经济数据的结构

- 城市犯罪统计的两年面板数据

TABLE 1.5 A Two-Year Panel Data Set on City Crime Statistics

obsno	city	year	murders	population	unem	police
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

Each city has two time series observations

Number of police in 1986

Number of police in 1990

© Cengage Learning, 2016

因果关系和其他条件不变

- 因果关系(**causality**)与其他条件不变(**ceteris paribus**)

x 对 y 具有因果影响的定义为：

“How does variable y change if variable x is changed but all other relevant factors are held constant”

- 绝大多数经济模型都是其他条件不变问题

例如：商品价格对需求的影响等

- 保持其他条件不变对政策分析也很重要

- 例如：最低工资对就业的影响等

- 其他条件不变下的因果关系检验理论上可以在实验中实现

因果关系和其他条件不变

- 肥料对作物产量的因果效应
 - “如果增加土壤施肥量，大豆产量会增加多少”
 - 隐含假设：影响作物产量的所有其他因素，如土地质量、降雨量、寄生虫的存在等，都是固定不变的
- 实验：
 - 选择几块一英亩的土地；在不同的地块上**随机**分配不同数量的肥料；比较收成
 - 试验之所以有效，是因为施肥量与影响作物产量的其他因素无关

因果关系和其他条件不变

- 测量教育回报率
 - “如果一个人从人群中被选中，并被给予额外一年的教育，他或她的工资会增加多少？”
 - 隐含假设：所有其他影响工资的因素，如经验、家庭背景、智力等，都是固定不变的
- 实验：
 - 选择一群人；随机给他们分配不同数量的教育；比较工资
 - 没有进行随机分配会导致的问题：教育程度与影响工资的其他因素有关（例如智力）

因果关系和其他条件不变

- **执法对城市犯罪水平的影响**
 - “如果随机选择一个城市，并增加10名警察，其犯罪率会下降多少？”
 - 或者：“如果两个城市在所有方面都相同，只是A市的警察人数比B市多10人，那么这两个城市的犯罪率会相差多少？”
 - 隐含假设：所有其他影响犯罪水平的因素，如人均收入等，都是固定的或者相等的
- **实验：**
 - **随机**分配警察到很多城市中
 - 实际上，警察人数将由犯罪率决定

因果关系和其他条件不变

- 最低工资对失业率的影响
 - “如果最低工资增加了一定数量（其他因素不变），失业率会增加多少（如果有）？”
 - 隐含假设：所有其他影响失业率的因素，如劳动力总供给等，都是固定的或者相等的
- 实验：
 - 政府每年**随机**选择最低工资，并观察失业结果
 - 这个实验会奏效，因为最低工资水平和决定失业的其他因素无关
 - 实际上，最低工资水平将取决于政治和经济因素，这些因素也会影响失业率

因果关系和其他条件不变

- **总结：**虽然理论上可以通过实验实现其他条件不变，但可能因为成本高昂、伦理道德问题等原因难以实现。
- 计量分析可以利用非实验数据模拟一个其他条件不变的实验。
- 经济理论并不总是以因果效应的形式表述的
 - 例如，期望假说认为长期利率等于期望短期利率的连续复利

$$(1+r_{lt})^n = (1+r_{year1}^e)(1+r_{year2}^e)\dots(1+r_{yearn}^e)$$

- 例如，三个月期国库券的利率应等于六个月期国库券前三个月的预期利率；这可以用计量经济学的方法来检验



小结

- 计量经济学使用统计方法分析经济数据，从而估计经济变量的关系，以及检验经济理论
- 虽然计量经济学模型有时可以从经济模型推导出来，但通常情况下，计量模型都是以非规范的经济推理和直觉为基础
- 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
- 在研究两个变量之间的因果关系时，需要其他条件不变：所有其他的相关因素都必须固定不变。



思考

- 面板数据中是否普遍存在相关性？
- 有没有我们介绍的四种数据类型之外的数据存在？
- 计算两个变量观测数据的相关系数是否可以反映因果关系？