

实证的经济分析方法

计量经济学的发展历史

赵博

南开大学金融学院

2026年6月1日

计量经济学的问题

- 经济理论：

价格上升，需求量下降

教育增加，收入可能增加

货币政策改变，产出和通胀会变

- 计量：

识别、估计、检验

- 统计形式 vs. 经济理论
- 回归系数的经济含义
- 从相关关系到因果解释

层次	问题
理论命题	一个关系在理论上是否成立
数量判断	这个关系有多大，在哪些人身上更强
因果判断	如果人为改变一个变量，另一个变量会怎样变

- 理论命题：概念、逻辑、模型
- 数量判断：数据和估计
- 因果判断：识别和反事实
- 计量经济学：主要处理数量判断和因果判断

- 统计形式：

$$Y_i = \alpha + \beta X_i + u_i$$

- β 的经济含义：

- 相关
- 预测
- 因果

- 需要说明：

- 样本来源
- 随机项内容
- X_i 与 u_i 的关系
- 外推对象

时期	代表	主要问题
1930 年代	Tinbergen, Keynes	经济数据能否支持统计推断
1940 年代	Haavelmo	经济理论如何写成概率模型
1950–1960 年代	Cowles Commission	联立方程和识别
1970 年代	Lucas, Sims	政策变化和宏观模型失灵
1980 年代以后	Granger, Engle, Rubin, Angrist, Imbens	时间序列和因果推断
当代	Heckman, Pearl, Athey 等	结构、设计、机器学习并存

起源：测量与质疑

- Ragnar Frisch 在 1920 年代使用 econometrics 这个词
- 三个组成：

经济理论 + 数学 + 统计测量
- 早期目标：
 - 给经济关系赋予数量
 - 描述商业周期
 - 改善预测
- 1933 年，Frisch 提出 impulse-propagation 思路
- 冲击来自外部，传播机制来自经济系统内部

- 1930 年代：宏观变量方程组

- 消费：

$$C_t = \alpha + \beta Y_t + \varepsilon_t$$

- 投资：

$$I_t = \gamma + \delta r_t + \eta_t$$

- 变量：

$$C, I, Y, r, P, W, \dots$$

- 商业周期理论：可估计方程

- Keynes 1939 年评论 Tinbergen 的商业周期研究
- 统计关系的根据：
 - 完备性：重要变量是否遗漏
 - 可测量性：预期、信心、制度变化
 - 独立性：经济变量相互决定
 - 稳定性：参数跨时期是否稳定
- 后续计量经济学反复处理这些问题

- 假设真实关系是：

$$Y_i = \alpha + \beta X_i + \gamma Z_i + e_i$$

- 但研究者估计：

$$Y_i = a + bX_i + u_i$$

- 如果 X_i 和 Z_i 相关， b 通常不是 β
- 例子：

$$\text{收入} = \alpha + \beta \text{教育} + u$$

- 能力、家庭背景、地区机会如果没有处理，教育系数就混入了这些因素

- 需求：

$$Q_t^d = a - bP_t + u_t$$

- 供给：

$$Q_t^s = c + dP_t + v_t$$

- 市场出清：

$$Q_t^d = Q_t^s = Q_t$$

- 观测到的是均衡点 (P_t, Q_t)
- Q_t 对 P_t 的简单回归：混合需求和供给

- 即使一段时期内有：

$$C_t = \alpha + \beta Y_t + \varepsilon_t$$

- 也要问：

β 在不同制度、不同预期、不同收入分布下是否一样？

- Keynes 对 Tinbergen 的批评里，已经有这个问题
- 后来的 Lucas Critique 把它变成宏观计量的中心问题

- Tinbergen：方程、估计、预测
- Keynes：统计推断的前提
- 缺口：

经济数据为什么可以用概率统计处理？

- 需要说明：
 - 数据生成过程
 - 随机项含义
 - 参数识别
 - 检验的概率含义

Haavelmo : 概率方法

- Haavelmo 1944 年发表 *The Probability Approach in Econometrics*

- 问题 :

经济理论怎样和随机数据发生关系 ?

- 确定性方程 :

$$Y = f(X; \theta)$$

- 现实数据 :

$$Y_i \neq f(X_i; \theta)$$

- 偏离项的解释

- 统计模型：

$$Y_i = X_i' \beta + u_i$$

- 条件：

$$\mathbb{E}(u_i | X_i) = 0$$

- 违反条件：

$$\mathbb{E}(u_i | X_i) \neq 0$$

- 随机项内容：

- 未观察到的影响因素
- 测量误差
- 个体差异
- 模型没有写出的历史条件

- 随机变量：

$$W_i = (Y_i, X_i), \quad W_i \sim P$$

- 确定性理论：

$$Y = f(X; \theta)$$

- 分布限制：

$$\mathbb{E}(Y_i | X_i) = X_i' \beta$$

$$\mathbb{E}(u_i | X_i) = 0$$

- 估计和检验：在概率模型内定义

对 Keynes 批评的处理

Keynes 的问题	Haavelmo 之后的表述
变量不完整	未列入因素进入随机项，但必须说明它和解释变量的关系
变量难测量	区分理论变量、真实变量和观测变量
变量相互决定	写成联立方程系统，讨论识别和内生性
参数不稳定	稳定性成为模型假设，不能默认成立

- 问题保留
- 表述改变
- 从方法论争论进入统计模型

- 条件论证：

如果要从数据做统计推断，就要给出概率模型

- 工作规则：

- 样本空间
- 随机变量
- 参数
- 识别假设
- 估计
- 检验

- 未推出：

所有经济命题都必须检验

- 战后宏观预测和政策分析需求
- Neyman–Pearson 假设检验
- Cowles Commission：联立方程方法
- 1950–1960 年代：教科书化
- Keynes 1946 年去世：论战缺少继续者
- 可执行的研究程序

- 概率模型：假设
- u_i ：未解释部分
- 显著性：统计判断
- 机制：经济解释
- 样本内拟合：历史数据表现
- 政策外推：反事实问题
- 持续问题：

这个统计模型和现实生成过程之间的距离在哪里？

Cowles : 联立方程与识别

- Haavelmo 概率框架的技术展开
- 宏观经济：联立方程系统

$$Ay_t + Bx_t = \varepsilon_t$$

- y_t : 内生变量
- x_t : 外生变量或前定变量
- 目标：估计结构参数，用于政策模拟

- 结构型 structural form :

$$Ay_t + Bx_t = \varepsilon_t$$

- 方程对应理论关系
- 内生变量可以同时出现在方程两边
- 参数有经济含义

- 简化型 reduced form :

$$y_t = \Pi x_t + v_t$$

- 每个内生变量只由外生变量和扰动表示
- 更接近直接从数据中估计的关系

- 识别问题 :

Π 能否唯一反推出 A, B ?

结构型与简化型：供需例子

- 结构型：

$$Q_t = \alpha - \beta P_t + u_t \quad \text{需求}$$

$$Q_t = \gamma + \delta P_t + \lambda Z_t + v_t \quad \text{供给}$$

- Z_t ：只影响供给的外生变量，例如天气
- 两式相等，解出价格：

$$P_t = \frac{\alpha - \gamma - \lambda Z_t + u_t - v_t}{\beta + \delta}$$

- 再代回去，得到数量：

$$Q_t = \frac{\delta\alpha + \beta\gamma + \beta\lambda Z_t + \delta u_t + \beta v_t}{\beta + \delta}$$

简化型告诉我们什么

- 简化型把观测变量写成外生变量的函数：

$$P_t = \pi_{10} + \pi_{11}Z_t + \eta_{1t}$$

$$Q_t = \pi_{20} + \pi_{21}Z_t + \eta_{2t}$$

- 数据可以估计：

$$\pi_{10}, \pi_{11}, \pi_{20}, \pi_{21}$$

- 结构参数是：

$$\alpha, \beta, \gamma, \delta, \lambda$$

- reduced form 参数是结构参数的组合
- 识别：能否从这些组合中恢复出结构参数

- 需求：

$$Q_t^d = \alpha - \beta P_t + u_t$$

- 供给：

$$Q_t^s = \gamma + \delta P_t + v_t$$

- 出清：

$$Q_t^d = Q_t^s = Q_t$$

- 观测： (P_t, Q_t)

- 需求和供给同时移动：散点图不等于需求曲线

- Z_t : 天气
- 作用 : 影响供给 , 不直接影响需求
- 供给 :

$$Q_t^s = \gamma + \delta P_t + \lambda Z_t + v_t$$

- 相关性 :
- 外生性 :

$$\text{Cov}(Z_t, P_t) \neq 0$$

$$\text{Cov}(Z_t, u_t) = 0$$

- 贡献：
 - 明确了内生性
 - 明确了识别问题
 - 把结构模型和统计推断结合起来
 - 发展出 2SLS、LIML、FIML、3SLS 等方法
- 问题：
 - 大型方程组需要很多排除限制
 - 这些限制常常很难从理论上充分说明
 - 系统越大，设定错误越难避免
 - 政策环境改变时，参数可能改变

1970 年代：宏观计量的危机

- 1960 年代：预测和政策模拟
- 1970 年代：滞胀、石油冲击、菲利普斯曲线不稳定
- 暴露问题：

过去样本中的关系，不能直接用于新政策环境
- 两个反应：

Lucas Critique
Sims 的 VAR

- 经验关系：

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

- 政策模拟：改变 x_t ，预测 y_t
- Lucas 的问题：

β 可能取决于政策规则本身

$$\beta = \beta(\lambda)$$

- λ 是政策规则或制度环境
- 政策规则改变后，人的预期和行为也会改变

- 历史上观察到：

$$\pi_t \uparrow \Rightarrow u_t \downarrow$$

- 政策想法：更高通胀，更低失业
- 预期调整：工资和价格反应
- 原关系消失
- 不是估计误差
- 政策改变行为规则

- 宏观模型：
 - 个体最优化
 - 理性预期
 - 技术、偏好等较深层参数
 - 明确的政策规则
- 后续：RBC、DSGE
- 新问题：
 - 代表性主体是否合适
 - 理性预期是否是好的经验描述
 - “深层参数”是否真的稳定
 - 冲击的经济含义是否清楚

- 诊断：大型联立方程模型限制过多
- VAR：

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t$$

- y_t 可以包括：

$$(Y_t, \pi_t, i_t, M_t, \dots)'$$

- 每个变量：由所有变量滞后值解释
- 动态相关
- 结构解释

- 残差 ε_t ：统计冲击
- “货币政策冲击”：需要识别
- 常见识别：
 - Cholesky 排序
 - 短期约束
 - 长期约束
 - 符号约束
- 理论假设：从结构方程转向冲击识别

- 预测概念
- X 的过去值改善对 Y 的预测：

$$\mathbb{E}(Y_t | Y_{t-1}, X_{t-1}) \neq \mathbb{E}(Y_t | Y_{t-1})$$

- 不等于真实因果机制
- 例子：天气预报 Granger-causes 下雨吗？
- 术语：谨慎使用

时间序列与金融计量

- 宏观变量：趋势
- 互不相关的趋势变量：高 R^2
- 回归：

$$Y_t = \alpha + \beta X_t + u_t$$

- 非平稳序列：普通 t 检验失效
- 问题：时间序列的概率性质

- 一个简单过程：

$$Y_t = \rho Y_{t-1} + \varepsilon_t$$

- $|\rho| < 1$ ：围绕稳定水平波动
- $\rho = 1$ ：

$$Y_t = Y_{t-1} + \varepsilon_t$$

- 冲击具有永久影响
- 宏观变量：可能接近随机游走

- Y_t 和 X_t 都有单位根
- 线性组合：

$$Y_t - \beta X_t$$

- 平稳：协整
- 经济含义：

变量各自漂移，但存在长期关系

- 例子：

消费和收入

价格水平和货币供应量

- 长期关系 + 短期调整

- 长期关系：

$$Y_t = \beta X_t + e_t$$

- 误差修正：

$$\Delta Y_t = \alpha(Y_{t-1} - \beta X_{t-1}) + \gamma \Delta X_t + \varepsilon_t$$

- $Y_{t-1} - \beta X_{t-1}$ ：上期偏离
- α ：调整速度
- 长期均衡 + 短期波动

- 金融收益率：均值难预测
- 波动率：持续性
- ARCH：

$$r_t = \mu + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

- GARCH：

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

- 建模对象：从均值到方差

- 动态关系：适合
- 宏观样本：通常不长
- 制度变化、结构断点
- 预测表现
- 机制解释
- 政策反事实

预测、解释、政策反事实

因果推断

- 估计：

$$Y_i = \alpha + \beta D_i + X_i' \gamma + u_i$$

- β ：相关、预测、因果
- 因果解释：

D_i 的变化从哪里来

- 处理组选择问题
- 识别先于估计

- 对每个个体 i , 定义:

$$Y_i(1), \quad Y_i(0)$$

- 个体因果效应:

$$Y_i(1) - Y_i(0)$$

- 观测结果:

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

- 缺失反事实
- 因果推断: 构造反事实

潜在结果框架 (Potential Outcomes, Rubin's causal model)

- 对每个个体 i ，定义潜在结果 (Potential Outcomes)：

$$Y_i(1) \quad (\text{或 } Y_{1i}), \quad Y_i(0) \quad (\text{或 } Y_{0i})$$

- 无论个体实际是否接受处理，两类潜在结果在假想中均存在。
- 个体因果效应 (Individual Causal Effect)：

$$Y_i(1) - Y_i(0)$$

- 实际观测结果 (Observed Outcome)：

$$Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

或写成线性组合形式： $Y_i = Y_i(0) + [Y_i(1) - Y_i(0)]D_i$ ，其中 $D_i \in \{0, 1\}$ 为处理变量。

- 因果推断的核心挑战：缺失反事实 (Missing Counterfactuals)。推断的本质在于如何合理构造反事实。

观测差异的分解 (Naive Comparison)

- 简单比较两组的观测均值差异，通常无法直接得到因果效应：

$$E[Y_i | D_i = 1] - E[Y_j | D_j = 0] = \underbrace{E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 1]}_{\text{ATT (处理组平均因果效应)}} + \underbrace{E[Y_i(0) | D_i = 1] - E[Y_j(0) | D_j = 0]}_{\text{Selection Bias (选择偏差)}}$$

- 典型示例：去医院对健康的影响
 - 显然 $E[Y_i(0) | D_i = 1] < E[Y_i(0) | D_i = 0]$ （去医院的人本身基础健康状况较差）。
 - 这导致选择偏差为负，从而低估了医院对患者的真正医疗作用（ATT）。

随机化实验 (Randomization)

- 完全随机化分配保证了处理变量 D_i 独立于潜在结果：

$$(Y_i(1), Y_i(0)) \perp\!\!\!\perp D_i$$

- 此时选择偏差被消除：

$$E[Y_i(0) | D_i = 1] = E[Y_i(0) | D_i = 0] = E[Y_i(0)]$$

- 观测差异直接等于平均因果效应 (ATE)：

$$\begin{aligned} E[Y_i | D_i = 1] - E[Y_i | D_i = 0] &= E[Y_i(1) | D_i = 1] - E[Y_i(0) | D_i = 0] \\ &= E[Y_i(1)] - E[Y_i(0)] \\ &= E[Y_i(1) - Y_i(0)] \quad (\equiv \text{ATE}) \end{aligned}$$

回归表示 (Regression Representation)

- 设定线性回归模型：

$$Y_i = \alpha + \beta D_i + \epsilon_i$$

- 假定常数效应 (Constant Treatment Effect) : $Y_i(1) - Y_i(0) = \beta$ 。
 - 对照观测结果公式，可推得误差项为： $\epsilon_i = Y_i(0) - E[Y_i(0)]$ 。
- 条件期望相减，考察 OLS 估计量的因果含义：

$$E[Y_i | D_i = 1] - E[Y_i | D_i = 0] = \beta + \underbrace{\left(E[\epsilon_i | D_i = 1] - E[\epsilon_i | D_i = 0] \right)}_{\text{Selection Bias}}$$

- 结论：若 D_i 与误差项 ϵ_i 相关（存在选择偏差），则普通最小二乘法（OLS）得到的回归系数将不等于真正的因果效应 β 。

- 处理组 T ，对照组 C
- 处理前和处理后都有观测

$$\hat{\tau}_{DID} = (\bar{Y}_{T,after} - \bar{Y}_{T,before}) - (\bar{Y}_{C,after} - \bar{Y}_{C,before})$$

- 关键假设：无政策时，两组趋势相同

parallel trends

- 处理后数据：不能证明平行趋势
- 处理前趋势、制度背景、安慰剂检验

- 处理由阈值决定：

$$D_i = 1\{X_i \geq c\}$$

- 阈值附近：近似可比
- 断点跳跃：

$$\tau = \lim_{x \downarrow c} \mathbb{E}[Y_i | X_i = x] - \lim_{x \uparrow c} \mathbb{E}[Y_i | X_i = x]$$

- 关键假设：潜在结果在断点处连续
- 主要限制：断点附近的局部效应

方法	利用什么变化	最重要的问题
IV	工具变量造成的处理变化	工具变量是否只通过处理影响结果
DID	处理组和对照组的处理变化差异	平行趋势是否可信
RDD	阈值两侧的处理差异	阈值附近是否可比，是否存在操纵

- 共同点：识别假设前置
- 代价：结论更局部

- 优点：
 - 因果问题更明确
 - 识别假设更透明
 - 研究更重视制度细节
- 代价：
 - 估计对象常常很局部
 - 容易只回答小问题
 - 理论解释可能变弱
 - 外推到其他时期、地区和制度时要小心

当代格局

结构方法 vs. 设计方法

	结构方法	设计方法
问题	机制是什么，换政策会怎样	某个处理在某个环境下有没有效果
优点	能做反事实和机制分解	识别假设通常更清楚
风险	模型设定强	外部有效性有限
代表	Cowles, Heckman, BLP, DSGE	IV, DID, RDD, RCT

- 政策外推：结构方法
- 具体政策评估：设计方法
- 现实研究：二者经常结合

- 公式：

$$p(\theta | data) \propto p(data | \theta)p(\theta)$$

- 先验：

$$p(\theta)$$

- 似然：

$$p(data | \theta)$$

- 在宏观中常见：

- Bayesian VAR
- Bayesian DSGE

- 优点：复杂模型、参数不确定性
- 问题：先验选择、先验敏感性

- 机器学习：预测
- 计量经济学：因果参数、统计推断
- 交叉方向：
 - 用机器学习处理高维控制变量
 - 估计异质性处理效应

- 例子：

$$\tau(x) = \mathbb{E}[Y(1) - Y(0) \mid X = x]$$

- 识别问题：仍需经济和制度论证
- 主要作用：函数估计、预测、高维控制

- 规格搜索：
 - 换样本
 - 换控制变量
 - 换函数形式
 - 换标准误
 - 换处理窗口
- 显著性和发表偏误
- 数据和代码公开
- 预分析计划
- 复制研究
- 研究过程的可审计性

- Tinbergen：把经济理论写成可估计方程
- Keynes：追问统计推断的前提
- Haavelmo：用概率模型重写这个问题
- Cowles：把识别和联立方程变成标准方法
- Lucas：政策会改变行为，参数未必稳定
- Sims：减少大模型中的任意结构限制
- Rubin–Angrist–Imbens：把因果效应和识别策略放到中心

- 哪些经济命题不需要经验检验？
- 哪些经济命题必须通过数据判断？
- 一个局部因果效应是否足以指导政策？
- 如果模型机制清楚但预测不好，应该怎样评价？
- 如果模型预测好但机制不清楚，应该怎样评价？

- Frisch (1926, 1933); Tinbergen (1936, 1939); Keynes (1939)
- Haavelmo (1944), *The Probability Approach in Econometrics*
- Koopmans (1950), Cowles Commission Monograph 10
- Lucas (1976), "Econometric Policy Evaluation: A Critique"
- Sims (1980), "Macroeconomics and Reality"
- Engle (1982); Engle and Granger (1987)
- Rubin (1974); Rosenbaum and Rubin (1983)
- Imbens and Angrist (1994); Angrist and Pischke (2009, 2010)
- Leamer (1983); Heckman (2000); Pearl (2009)