

计量经济学的方法论意义

赵博

南开大学金融学院

2025 年 2 月 19 日

- 课程简介
- 经济学研究方法论 + 计量经济学在整个经济学研究中的位置
- 计量经济学简史 + 机器学习简史，二者区别和联系

课程考核方式

- 平时作业：30%；考试：70%
 - 作业各班不一样；考试题完全一样
- 教科书：Wooldridge, Jeffrey. Introductory Econometrics: A Modern Approach. 7th edition. Boston, MA: Cengage Learning, 2019.
- 参考书：
 - 墙裂推荐：Kennedy, Peter. A Guide to Econometrics. 6 edition. Malden, MA: Wiley-Blackwell, 2008.
 - 其他的好的参考书（你们不用看，知道一下就行）：
 - Hayashi (2000)
 - Wooldridge's Panel Data book
 - Bruce Hansen's book
 - Mostly Harmless Econometrics
 - 如果想搞计量经济学学术研究：先看概率统计机器学习
- 课程主页：https://abzhaobo.github.io/courses/econometrics_ug

一些可能产生的困惑

- 经济金融的学生一般需要研究生学位
- 需要很长时间的实习（半年甚至更长）
- 入学前：经济、金融赚钱，高大上。电视网络上首席经济学家侃侃而谈，没什么是他们不能分析的
- 入学后：宏微观、专业课学了有什么用？可以用作实际分析吗？首席经济学家们为什么不用模型？为什么不用计量经济学？仅仅是他们怕普通民众听不懂吗？
- 毕业论文：99% 的论文都用计量经济学看一个或者几个 x 是不是影响 y 。宏微观理论在哪里？需不需要？
- 机器学习 tmd 又是什么？现在都搞这些东西，我要去学吗？到底怎么就业？拼得过那些搞 AI 的吗？

我们探讨一下

我们认识世界的方式：科学/非科学

- 科学是什么？Karl Popper: Falsifiability (可证伪；允许检验)
- 科学：实证 (empiricism) + 奥卡姆剃刀原则 (Occam's Razor)
 - 物理学是科学 (弦理论?)；化学、生物——是；数学——不是
- 当理论/方法有清晰的实证检验方法，进步可见。贡献/成就也可见
 - 对已有数据的检验 (解释)
 - 对未来的预测 (预测，可重复)
- 当实证检验很困难时，怎么办？
 - 命题 1：明朝亡于党争 (历史；政治)
 - 命题 2：加入 WTO 是中国最近 20 多年经济增长的重要原因 (经济学)
 - 命题 3：美国 1933、1934 年的《证券法》、《证券交易法》显著稳定了金融市场 (金融学)

经济学研究方法举例：需求定律

需求定律：商品（东西）的价格与需求成反比：价格越高，需求越低；价格越低，需求越高

- 非数学推导的方式（古典经济学，Marshallian, Austrian, Chicago school, etc.）
- 数学推导的方式（当代经济学主流）
- 实证检验（计量经济学，当代经济学的最主流）
- 对各种研究方法的反思

研究生学费越来越贵，考研学生越来越多，是不是违反需求定律？

数学推导的方式

问题:

$$\max_{x,y} U(x, y) \quad \text{subject to} \quad p_x x + p_y y = I. \quad (1)$$

Lagrangian:

$$\mathcal{L} = U(x, y) + \lambda(I - p_x x - p_y y). \quad (2)$$

取偏导:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} = \frac{\partial U}{\partial x} - \lambda p_x = 0 & \Rightarrow \lambda = \frac{\frac{\partial U}{\partial x}}{p_x}, \\ \frac{\partial \mathcal{L}}{\partial y} = \frac{\partial U}{\partial y} - \lambda p_y = 0 & \Rightarrow \lambda = \frac{\frac{\partial U}{\partial y}}{p_y}. \end{aligned} \quad (3)$$

一阶条件:

$$\frac{\frac{\partial U}{\partial x}}{p_x} = \frac{\frac{\partial U}{\partial y}}{p_y} \Rightarrow \frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}} = \frac{p_x}{p_y}. \quad (4)$$

用 Cobb-Douglas 效用函数来具体观察

$$U(x, y) = x^\alpha y^\beta. \quad (5)$$

$$\frac{\partial U}{\partial x} = \alpha x^{\alpha-1} y^\beta \quad \text{and} \quad \frac{\partial U}{\partial y} = \beta x^\alpha y^{\beta-1}. \quad (6)$$

一阶条件:

$$\frac{\alpha x^{\alpha-1} y^\beta}{p_x} = \frac{\beta x^\alpha y^{\beta-1}}{p_y} \Rightarrow y = \frac{\beta p_x}{\alpha p_y} x. \quad (7)$$

把 $y = \frac{\beta p_x}{\alpha p_y} x$ 代入约束条件:

$$p_x x + p_y \left(\frac{\beta p_x}{\alpha p_y} x \right) = I. \quad (8)$$

简化得:

$$x^* = \frac{I}{p_x \left(1 + \frac{\beta}{\alpha} \right)}. \quad (9)$$

类似可得: $y^* = \frac{I}{p_y \left(1 + \frac{\alpha}{\beta} \right)}.$

这套推导方式在《微观经济学》中构建

非数学推导的方式（反躬自省 + 演绎法）

- 假设 1: 主观价值论 (Subjective Value)。
 - 商品的价值是人主观判定的（也即，不是由其制造成本决定的，比如劳动价值论），其判定标准是商品对人的边际效用
 - 边际效用是人选择商品的依据
- 假设 2: 人们喜欢多种商品（人们通过机会成本的变化来考虑消费量） \Rightarrow 边际效用递减。消费商品 A 后，人得到满足，人们希望消费一些其他的商品（这些其他的商品是继续消费商品 A 的机会成本），下一个商品 A 的边际效用减少
- 推论：需求曲线的斜率是负向的：当人的消费量增加时，愿意为这个商品付出的价格减少。当一个商品的价格减少时，愿意买更多这个商品，因为消费其他商品的机会成本变小（也即可以消费更多其他商品）

这套推导方式被芝加哥学派、奥地利学派等应用

$$\begin{cases} q_t^d = \alpha_0 + \alpha_1 p_t + u_t & (\text{需求}) \\ q_t^s = \beta_0 + \beta_1 p_t + \nu_t & (\text{供给}) \\ q_t^d = q_t^s & (\text{均衡}) \end{cases} \quad (10)$$

令 $q_t \equiv q_t^d = q_t^s$,

$$\begin{cases} q_t = \alpha_0 + \alpha_1 p_t + u_t \\ q_t = \beta_0 + \beta_1 p_t + \nu_t \end{cases} \quad (11)$$

如果直接做回归, $q_t \xrightarrow{\text{OLS}} p_t$, 得到的是什么曲线? (OLS: 最小二乘法)

什么曲线也不是

把 (p_t, q_t) 看作未知（内生）， (u_t, v_t) 看作已知（外生），可求解

$$\begin{cases} p_t = p_t(u_t, v_t) = \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{v_t - u_t}{\alpha_1 - \beta_1} \\ q_t = q_t(u_t, v_t) = \frac{\alpha_1 \beta_0 - \alpha_0 \beta_1}{\alpha_1 - \beta_1} + \frac{\alpha_1 v_t - \beta_1 u_t}{\alpha_1 - \beta_1} \end{cases} \quad (12)$$

这里 $\text{Cov}(p_t, u_t) = -\frac{\text{Var}(u_t)}{\alpha_1 - \beta_1} > 0$. 大于 0 因为 $\alpha_1 < 0$, $\beta_1 > 0$

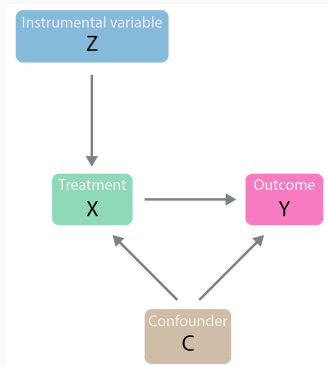
用 q_t 对 p_t 做回归，回归系数

$$\frac{\text{Cov}(p_t, q_t)}{\text{Var}(p_t)} = \frac{\text{Cov}(p_t, \alpha_0 + \alpha_1 p_t + u_t)}{\text{Var}(p_t)} = \alpha_1 + \underbrace{\frac{\text{Cov}(p_t, u_t)}{\text{Var}(p_t)}}_{>0}$$

解决办法之一：工具变量

直觉: $y = \alpha_0 + \alpha_1 x + u$

- 要观察到 α_1 , 需要看到 x 变化时, y 怎么变。
- 然而, 若 x 与 u 相关, 则除了 x 变 $\rightarrow y$ 变, 同时还有 x 变 $\rightarrow u$ 变 $\rightarrow y$ 变
- y 的变化有两个来源, 这时 α_1 就被干扰了。
- 工具变量的想法: 找到一个 z , z 变 $\rightarrow x$ 变, 但 z 变 $\nrightarrow u$ 变
- 这样, x 的这种变化就不掺和 u 的变化, α_1 也就看见了。
- 总结: 1. z 变 $\rightarrow x$ 变; 2. z 变 $\nrightarrow u$ 变



比如：

- y_t 是咖啡豆产量
- z_t 是咖啡豆产地的气温
- 则 z_t 与供给相关（最好是高度相关，否则又会产生弱工具变量问题），但与需求的变动 (u_t) 不相关

回想这个案例：数学

- 亮点 1: 在给定的数学框架下，推导正确，则结论无可争议
- 亮点 2: 所有的定义都用数学表达，绝无歧义（但有时也许会被过度解读，比如金融里经常把“风险”表达为“收益率的标准差”）
- 局限 1: 得到的 x 和 p 之间的关系，被数学的假设限定死了。现实远远比这个复杂
- 局限 2: 新的见解是什么？（有时这被称为“so what?” test）

回想这个案例：反躬自省

- 亮点 1: 只要具备基本的思维能力和逻辑推理能力, 就能沿此观察到经济原理。有时这些原理会反直觉, 有时这些推导出来的结论会相当惊人。
- 局限 1: 我虽然知道是对的, 但无法量化“有多对”。有人认为: 无法量化, 就不是科学。(不是科学让人有些难受)
- 局限 2: 逻辑推导过程会不会产生错误? 推导的中间环节不需要某些成立的条件? 这些条件是什么? 有没有可能这个推导链条不是唯一的, 而是在某些时候有分叉? 比如往左往右有不同的概率
 - 这种基于基本假设往前推理的过程, 可以走多远?

回想这个案例：计量经济学

- 亮点：量化了模型，可以通过数据来观察。
- 但：如果你用了工具变量，想估计需求方程，得到的结果仍然是正的斜率。此时——
 - 你也许会认为这个工具变量还不够好
 - 你也许会认为数据收集得有问题
 - 你也许会认为计量模型还不够好，需要开发更好的计量模型工具
 -
 - 但一般而言，你一定不会认为需求定律错了！
- 换句话说，你通过反躬自省 + 逻辑推导得到的结论，无需经过检验！ 无论检验是什么结果，**你只想要做出你想要的结果**
- 也许你说，虽然我知道那个系数是负的，但负多少我不知道，可以去测一测
- 没错，但即便你测出来是 -10% ， -8% 好像也未必错。换一组数据，换一个人群，换一个时间，会不会是 -12% ？这个数就一直不变吗？

- 简单的统计图表
- 数理经济学的伴侣
 - 最开始的想法：检验经济学模型
 - 先贤们很快就意识到了“内生性问题”——永远笼罩在社会科学研究上空的阴影

Working (1927, QJE): 怎样估计供给曲线？怎样估计需求曲线？如果找到了价格、销售量数据，是否就能估计出来？
- 计量模型的建立 (Frisch, Tinbergen)
 - 难题：当时的先进科技（概率统计理论）能结合进经济理论中吗？怎么结合？
- 概率论的引入 (Haavelmo), 联立方程模型 (Cowles Commission)
- 简约模型 (Sims), 时间序列 (Engle, Granger)
- 因果推断 (Angrist, Imbens, Card)
 - 占领了经济学、金融学期刊的

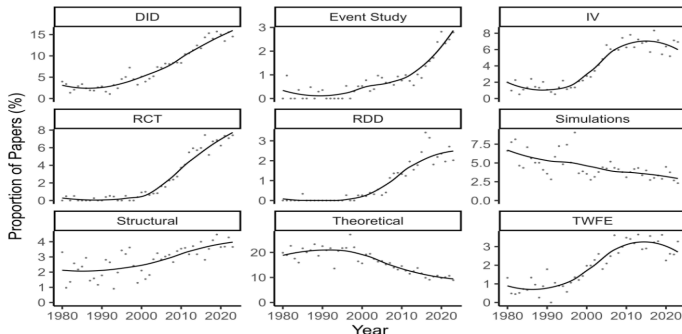
今天计量经济学的大致分类

- 因果分析 (Reduced form)。微观经济学 (劳动、教育)、公司金融常用。数据格式：横截面或者面板 (更常见)
 - Instrumental Variables (IV, 工具变量),
Difference-in-Difference (DID, 双重差分), Regression
Discontinuity Design (RDD, 断点回归)

书上的陈述 (我不能同意这种辩护):

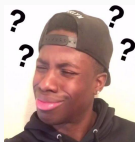
在多数情况下, 计量经济分析是从对一个计量经济模型的设定开始的, 而没有考虑模型构造的细节。我们一般遵循这一思路, 主要原因是, 对犯罪这种经济模型进行仔细推导, 不仅消耗的时间过长, 而且会把我们带到经济理论的某个特定而通常又极为困难的领域。在我们的例子中, 经济上的逻辑将起作用, 而且我们需要把其背后的任何一种经济理论都放到对计量模型的设定中。而在犯罪一例的经济模型中, 我们将从像 (1.3) 式那样的一个计量模型出发, 并以经济逻辑和常识作为选择变量的向导。尽管这一方法使经济分析的丰富性大打折扣, 但它总是被细心的研究者普遍而又有效地应用着。

Figure 4: Proliferation of Empirical Methods Over Time in NBER and CEPR Working Papers



Note: This figure shows the proliferation of key empirical methods used in NBER and CEPR working papers over time: Difference-in-Differences (DiD), Instrumental Variables (IV), Randomized Controlled Trials (RCTs), Regression Discontinuity Design (RDD), Two-Way Fixed Effects (TWFE), Structural Estimation, Event Studies, Simulations, and Theoretical/Non-Empirical research. Each panel represents the proportion of papers utilizing one of these methods per year, with the y-axis showing the proportion of total papers and the x-axis indicating the year of publication. The data covers all NBER and CEPR working papers from 1980 to 2023. DiD has seen a significant increase since the 1980s, rising from around 4% to over 15% of papers in recent years, reflecting its growing importance in empirical research. IV methods have also increased steadily from approximately 2% to over 6% over the same period. RCTs and RDDs, while starting from near zero in the 1980s, have grown to over 7% and 2% respectively in recent years, indicating the rising feasibility and acceptance of experimental and quasi-experimental designs in economics. Conversely, the use of theoretical and non-empirical research has declined significantly, from around 20% in 1980 to under 10% in 2023, suggesting a shift towards empirical analysis in the discipline. The use of simulations has decreased from over 6% in 1980 to around 2–4% in recent years. These trends highlight the increasing emphasis on credible identification strategies and the evolution of empirical methods in economics.

- “其他条件不变” 是什么???



- 最早开始提出这个概念的可能是李嘉图（David Ricardo）
- 很明显地在模仿自然科学的研究方法
- 现实中，社科领域几乎无法维持其他条件不变
- 我认为：机制分析——逻辑线条，更重要。但显然更难：更难被验证、更难被量化
- 折衷：因果分析的实证中，做机制分析的尝试

今天计量经济学的大致分类

- 时间序列（描述性、预测性——也即，不那么强调因果）。有时被冠以“金融计量经济学”。宏观经济学、金融资产定价常用
 - GARCH, VAR (Vector AutoRegression), Cointegration.....
- 其他：
 - 结构方程 (Structural form, 与 Reduced form 相对)。微观经济学 (Industrial Organization)
 - 贝叶斯 (Bayesian)
 - 机器学习
 -

例 1: 2024 年诺贝尔经济学奖

3 位 2024 年经济学诺奖获得者的代表作：

Acemoglu, D., S. Johnson, and J. A. Robinson. 2001. The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review* 91:1369–1401.

核心思想

- 主题：制度是否影响经济增长？
- 困难之处：因果关系（似乎）不明确
- 研究策略：利用殖民地定居者的死亡率差异，作为工具变量
 - (潜在) 定居者死亡率 → 是否定居 → 早期制度 → 现在制度 → 现在经济绩效

例 2: 2003 年诺贝尔经济学奖

Clive Granger, Robert Engle: Time series analysis

例 3: 1997 年诺贝尔经济学奖

- Black, F., & Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3), 637-654.
- 期权定价公式：通过基本的经济学原理（无套利原则），利用物理的热传导公式，给出了期权的理论合理定价
- 区别在哪里？在相对定价体系下，基准（Benchmark）是清楚的。股票期权的基准是股票价格。

AI（机器学习）简史

机器学习是怎么发展起来的，简要说一下历史

机器学习的发展大致可以分为以下几个阶段：

1. 早期理论奠基（1940s-1950s）

- 1943年，McCulloch 和 Pitts 提出了人工神经元模型，这是神经网络的基础。
- 1950年，图灵提出“图灵测试”，引发了对机器智能的思考。
- 1956年，达特茅斯会议标志着人工智能（AI）作为一个正式研究领域的诞生。

2. 符号主义与早期算法（1950s-1970s）

- 1957年，Rosenblatt 提出了感知机（Perceptron），是最早的神经网络模型之一。
- 1967年，最近邻算法（Nearest Neighbor）被提出，为模式识别奠定基础。
- 由于感知机无法解决非线性问题（如 XOR 问题），机器学习一度陷入停滞。

3. 统计学习崛起（1980s-1990s）

- 1986年，Hinton 等人提出反向传播（Backpropagation），推动了多层神经网络的发展。
- 1995年，Vapnik 提出了支持向量机（SVM），为机器学习提供了强大的数学工具。
- 统计学习理论逐渐取代符号主义，决策树、贝叶斯方法等得到广泛应用。

4. 数据驱动与深度学习复兴（2000s-2010s）

- 2006年，Hinton 提出深度信念网络（DBN），标志着深度学习的复兴。
- 2012年，AlexNet 在 ImageNet 比赛中取得突破性成绩，使**卷积神经网络（CNN）**广受关注。
- 2014年，Goodfellow 等人提出生成对抗网络（GANs），推动了生成模型的发展。
- 2017年，Google 提出Transformer，成为自然语言处理（NLP）的核心架构。

5. 大模型时代（2020s-至今）

- 2020年，OpenAI 发布GPT-3，标志着大规模预训练语言模型的突破。
- 2021年，Google DeepMind 提出的AlphaFold 解决了蛋白质折叠问题。
- 2022年，ChatGPT 推广，使大规模对话模型进入大众视野。
- 未来，机器学习正向**更强大的通用人工智能（AGI）**方向发展，同时强调可解释性、安全性和公平性。

机器学习和计量、统计的关系

- 统计学习 (Statistical Learning) 和机器学习 (Machine Learning) 在很长时间里、很多人认为 (包括机器学习大牛, 比如 Trevor Hastie, Robert Tibshirani) 是同一个东西
- Tom Sargent (2011 经济学诺奖)
<http://www.elecfans.com/d/748807.html>

我怎么讲到这儿来了呢? 人工智能首先是一些很华丽的辞藻。人工智能其实就是统计学, 只不过用了一个很华丽的辞藻, 其实就是统计学。好多的公式都非常老, 但是我们说所有的人工智能利用的都是统计学来解决问题。有一些新发展, 过去二三十年, 今天我们统计学的完成质量更高, 首先电脑运算速度更快, 有新的算法, 好多是源于物理学发展, 还有统计力学等等, 这是过去3到4年从物理学拿来, 加速了我们做系列统计学方面的进步。

今天的机器学习和统计学 (计量经济学) 到底有什么区别?

- 方法论: 机器学习关注预测, 也就是可验证性。(统计学上的平稳性)

- 量化投资：预测性问题
- 计量经济学：解释性问题。直接应用难，学术界在努力地拼接

Applied Causal Inference Powered by ML and AI

Victor Chernozhukov*

Christian Hansen†

Nathan Kallus‡

Martin Spindler§

Vasilis Syrgkanis¶

February 5, 2025

Publisher: Online

Version 0.1.1

公众号 · 计量经济圈

文科类的学科，需要好的提问、对的研究方法

问题的大致分类（二者可重叠）

- 解释性
- 预测性
 - 任何理论实际上都不能脱离“预测性”。即便是研究历史：“以史为鉴，可以知兴替”
 - 但是，对机制的研究才能让人真正理解模型的正确预测从何而来。
 - 我不能同意：assumptions need not be "realistic" to serve as scientific hypotheses; they merely need to make significant predictions. —Milton Friedman

以上的分析也部分说明了：

- 业界的经济学（叶檀、任泽平……）和学界经济学为什么存在如此大的区别？（这种区别几乎不存在于理工科类专业，比如，计算机、物理、人工智能）
- 找工作为什么需要这么长时间的实习
- 为什么当今的经济学研究会趋向于碎片化、甚至于无聊化
 - 基金经理颜值 -> 基金业绩
<https://www.yicai.com/news/101994712.html>

问好的问题!!! 用合适的模型!!!

The cost of computing has dropped exponentially, but the cost of thinking is what it always was. That is why we see so many articles with so many regressions and so little thought.

—Zvi Griliches