

第15讲：数据与方法

吴要武

中国社会科学院人口与劳动经济研究所

2018.12.29

一、引言

- **1953年**，米尔顿·弗里德曼发表了著名的《实证经济学的方法论》。在经济思想史上，这是一篇具有里程碑意义的文献：任何科学是否可信，首先需要通过经验证据（数据）的检验。
- 这给纷争不休的各学派，提供了一个标准，大多数人接受。拒绝接受者，成为非主流，自动被淹没在科学进步的波涛之中。
- 在同一时期，社会科学开始进入“数据革命”时代，以密歇根大学为代表的学术机构，开始利用计算机（革命）带来的技术手段进步，收集处理大量的微观数据信息（谢宇，2006）。

- 经济学领域的数据，是对市场主体行为与选择结果的记录。
- 每个数据信息，都是理性决策的结果：或者是长期的，或者是短期的。
- 这些数据都可以看作形而上世界的规律性在形而下世界的投影。
- 数据是丰富的、变异的、生动的、甚至是令人困惑的，但我们都相信，它与形而上世界的规律性之间，有一种神秘的对应性。
- 数据是跨越形而上与形而下的桥梁，也是世界可知性的见证。

- “不要用汇总数据作分析”。违背了 Well-defined.
- 你在研究问题、使用数据时，需要经常思考这个词汇的准确含义。
- 微观数据，使你的研究根基建立在磐石上。
- 因为经济学模型里不可缺少了代理人。个人、家庭（追求最大效用）或公司（追求最大利润），他们在受约束条件下的选择，才是可以分析的。这必然是微观数据。

经济学重视因果关系

- 两个变量同时变动，这种共变的背后可能是因果，但更可能是相关。而相关不等于因果。
- 因果，既能解释，也能预测；相关，能预测，但不能解释。
- 举例子：公鸡先打鸣，太阳然后升起。
- 你不能说“公鸡的打鸣”是“太阳升起的原因”。
- 顾客：“窗外在下雨吗？”侍者：“在下雨，我看到路上的行人都打着伞”。

过程不是因果

- 在经验研究中，要熟悉事情发生的过程，在研究对象上成为专家。
- 但过程本身并不是因果关系。一个寿星总结出来的长寿经验，未必真是“原因”。
- 这也给理论的价值提供了一个间接标准：真正伟大的理论，通常是高度抽象的。
- 比如，边际效用原理；科斯定理.....

经验分析的规范

- 《应用经济计量学》是门核心课程，课程目标是：
- 1、帮助学生发现值得研究的重要（且有趣的）问题；
- 2、帮助学生运用《经济计量学》的基本知识和方法，解决这个重要问题，从而，培养和训练学生具有独立解决问题的能力。
- 3、帮助学生报告研究发现（以数据事实为原材料，撰写规范的经验论文或研究报告）。
- 这是斯坦福大学经济学系Nick Bloom在Econ103中确定的三个目标。

经验分析成为经济学研究的主流

- 理论分析：
- 弄清经济变量之间的因果关系和影响方向：
- **Levitt**:堕胎与犯罪率下降。
- **Nathan Nunn**: 相互不信任导致非洲的穷困。
- “太阳下面无新事”。单纯的理论分析越来越少，因为每一代人都都不乏聪明才干之士（如：凯恩斯），面对同样或者类似的问题，早就有人（哈耶克）作出了回答。
- 在今天的欧美经济学界，**95%**以上的文章是经验分析或以经验分析为主。
- **Stigler**在批评**1890-1950**年期间的效用理论，几乎没有取得有价值的进展（有了经验分析，很容易检验，就不会走这么长的弯路）。

二、经验分析的可信性

- 1990年以来，经济计量学走出了“模型推导”为核心的时代，而是揭开“黑箱”，尽可能弄清楚因果关系和准确的数量关系。
- 今天的一流学术期刊，经验分析是主流，而几乎所有的经验分析文献都是建立在可信性基础上。
- 学术潮流进入“可信性时代”时，我们就不要再滞留在模型推导的暗夜里。
- 最典型是时间序列数据，因为样本较小，还要处理序列相关，经过几轮推导之后，根本不再有什么经验含义。

数据革命和经验分析

- 1931年，弗利希创办**Econometrica**，在发刊词里，就对这门学科的未来进行了前瞻，认为这是一门朝阳学科。
- 1932年成立的**Cowles Committee**，致力于在经济学、金融学、商业、产业组织以及技术等领域，鼓励和拓展逻辑学、数学和统计方法的应用。
- 二战中，科学技术取得了突飞猛进，成为决定战争胜负的重要因素。
- 社会科学也必须走向“准确区分”和“准确测量”。
- **What I can not create, I do not understand. Richard Feynman.**
- 自然科学家蔑视社会科学为伪科学（货拜族）。借用自然科学的术语，但没有自然科学的可重复、可检验、可复制。
- 当然，这些自然科学家低估了社会科学对准确区分的重视。

不同学派的论战推动了经验分析发展

- 论战目的是谁的解释更可信，谁提出的政策含义应该被接受。改造世界。
- 每个学派都自信满满。那就只能设定一个共同遵守的标准：通过经验证据的检验。
- 规则建立后，大家都遵守。失败者也承认公平。
- 比如，后来发现，否定凯恩斯学说的，大多数是那些自称“凯恩斯主义”的人，他们总也不能找到可信的证据，逐渐就放弃了，一旦他们在课堂上不再传授，那么，学生也就不再接受凯恩斯主义了。

凯恩斯学派与芝加哥学派的论战

- 二战结束，凯恩斯经济学登上了王座。
- 萨缪尔森以别人的诗句描绘了他们这一代青年经济学家的喜悦：
- “在那个早晨，活着已是幸运；年轻，则上了天堂”。
- 他在感叹自己遇到了好时候，在凯恩斯经济学里可以大展宏图。
- 然而，一个对手已经悄然站到他的面前：芝加哥学派顽强地、持续地质疑它：既否定他的理论，也否定他的政策建议。
- 如果经济学是科学，就不能面对对手的质疑无动于衷，而是证明：我是正确的。
- 科学的原则是，“谁主张，谁举证”。

凯恩斯经济学的根基与逻辑

- 凯恩斯经济学的基本出发地是：边际消费倾向递减；所以，随着收入增加，有效需求会变得越来越不足；那么，厂商的投资就因缺少需求而受到限制。
- 政府可以代替老百姓花钱，这个需求给厂商带来了机会。
- 而政府开支又有乘数效应，花少量的钱，可以推动更多的投资和消费。
- 这个过程持续下去，就能刺激经济从低谷中走出。
- 这一系列假设，都是可以检验的。

选择什么样的经验方程？

- 边际消费倾向递减吗？
- 如果没有边际消费倾向递减，凯恩斯理论大厦的根基（**Cornerstone**）就被抽去了。
- 经济学家利用住户调查（微观）数据，对消费函数进行了估计，模型设定为**OLS**，得到的结果是，**B=0.7**。
- 确实存在边际消费倾向递减。
- 然而，一些经济学家质疑这个结果：模型中的“收入”并不是外生的，而是与误差项相关。这时得到的结果是有偏差的。
- 必须改进估计方法，消除内生性偏差。

两阶段回归和固定效应模型

- 弗里德曼提出了持久收入假说。从经验数据的结果看，这个假说与经验证据一致。
- 采用两阶段估计，改进估计结果后，
B=0.92
- 如果真有边际消费倾向递减，一定是微不足道的。
- 芝加哥学派抽掉了凯恩斯学派根基上的第一块砖头。

- 1980年以后，凯恩斯学派的理论从现代教科书中逐步退出：
- 凯恩斯十字；
- IS-LM模型；
-
- 与此相反，芝加哥学派的理论则都成为教科书的标准内容：科斯定理；持久收入假说.....
- 萨缪尔森在《经济学》里，一再向对手让步，直到他去世。
- 新凯恩斯主义的代表人物，Mankiw承认，凯恩斯所有的理论都无法通过经验证据的检验。
- 经验证据成了最权威的裁判。
- 今天，宏观经济学也需要用微观数据来研究（Erik Hurst）。

三、数据的产生过程

- 1、随机受控实验；
- 2、抽样调查；
- 3、经济或人口普查；
- 4、档案或资料记录：如贩奴贸易，（Nathan Nunn）；殖民地宗主国士兵的死亡率（AJR2001）；公司季报、年报等。
- 对数据质量的评估，需要一个“十字”结构：观察同一层面的多数据结果，收敛；观察具体数据经验上的操作过程，是否严格控制了关键环节以最接近于随机化。

经验数据的收集

- 1、政府相关部门公布：
 - 国家统计局公布年度、月度和季度的经济指标；
 - 民政部公布每个月全国各县最低生活保障人数、支出和支付水平的信息；
- 2、研究机构或个人自己抽样调查
 - 人口所所做的四轮城镇劳动力市场调查（**CULS**）。
 - 中科院地理所和北大教育财政科学研究所联合做的随机受控实验数据。
 - 北大中国经济研究中心的老年人调查（**CHARLS**）

数据特征与变异

- 1、截面数据
- 2、时间序列数据
- 3、混合截面数据
- 4、纵列数据（**Panel Data**）
- 不同的数据类型，决定了模型设定和所采取的经验策略。
- 因为不同的数据类型有不同的变异来源，需要不同的假定（**Assumption**）。

数据类型

- 1、观测性数据
- 经济（人口）普查或抽样调查；记录性数据。这是最常见到的数据。
- 2、实验性数据
- 随机受控实验。
- 3、模拟性数据
- 从数据产生过程看，前两类是客观数据；后者是人为操作数据。

如何处理数据？

- 首先要熟悉数据，才能运用数据。正像你的工具箱。寻找适用数据，对任何学者来说，都是最硬的约束。
- 怎么做到熟悉呢？研究抽样调查的技术文件，清楚数据来自哪个原始调查。要舍得花时间浸泡在数据中，清楚有哪些变量；作些描述性统计结果；观察异常值；甚至肉眼去查看一下数据库。
- 1、异常值的处理：
 - A、舍弃——取决于样本大小。大样本可以，小样本不能。
 - B、对异常值（通常是过大时）取对数（见AJR2001）。
- 2、序次变量的处理：
 - Nathan Nunn（2011）将这些序次分类数据直接处理为连续变量。
- 3、Beta 系数。为了使估计系数具有可比性，对变异进行正规化处理。自变量变动1个标准差，因变量会随之变动B个标准差。

数据要与经验事实联系起来

- 在“理论-方法-环境-写作”四大要素中，数据代表着环境，决定着方法。
- 数据结构还决定了模型设定。所以，在经验分析中，数据占据核心地位：“没有粘土难做砖”。
- 既然数据就是对经验事实的描述，那么，研究者就要把数据和经验事实时时联系起来。
- 如果你的经历丰富，对数据的结果会有良好的直觉，借助直觉，可以评估数据的质量。
- 经验分析总是从理论出发讲故事，而不是从数据出发。谁是头，谁是脚。别弄反了。
- 但无论是理论还是数据，都要经常与经验认知比对一下，看是否一致。不一致，肯定是哪个因素出了什么差错。创新的机会或许出现了。

数据质量

- 数据的可信性——只有可信的数据才有可信的经验研究。
- 最理想的数据莫过于随机受控试验产生，随机委派能确保内部有效性（但外部有效性上有麻烦）。
- 但大多数能得到的数据是观测性数据。抽样数据要正确反映总体。这是基本要求（**DGP**符合高斯-马尔科夫假定）。只有确保了随机性的抽样调查才能满足这个要求（理论上的**DGP**与经验上的数据产生过程结合起来）。
- 比如：抽样调查时富人低报收入；富人社区因为更难进入，在样本中代表性不足。
- 如果富人和穷人随机的“报告”收入，那么，衰减性偏差会带来教育回报率趋于0.
- 测量误差——如果存在测量误差且是系统性的，会扭曲估计结果；如果是非系统性的，会产生衰减性偏差。

- 数据质量问题，成因而于抽样偏差和调查误差。对于大样本数据，从实践上看，主要是抽样调查过程未得到严格的控制与管理。
- 调查者会倾向于按照自己愿意的方式去收集数据（Cameron and Trivedi, 2005）。
- 数据质量差的后果：无法使用，无法公布。以此作预测，得出错误的结论（1932年美国总统选举，电话调查，预测胡佛胜选。少数富人才有电话，是胡佛的支持者）。
- 国家统计局作过很多调查，但从来不公布，一个重要原因是他们自己都不敢相信这个调查结果（2001年以来的千分之一人口调查数据的就业信息）。
- 调查过程的激励与管理：数据最终产生于抽样员、调查员和被调查者的接触性访谈。对基层抽样员和调查员，必须给予必要的管理和激励。
- 国家统计局最大的麻烦就是无法确保这个管理与激励。

- 对使用数据的学者来说，错误的的数据产生什么结果？
- 拿到数据后，通常会先观察一下描述性统计结果，与经验认知相比较。
- 很多人会用数据结果来修正自己的研究假说。
- 如果是错误的的数据和错误的结论，就可能反应到你的假说中去。
- 这样的结论很荒谬。但在我们身边却屡见不鲜。
- 有人警告，逻辑很危险。但我要警告：对数据，也要有风险意识。
- 一定要评估数据的质量！

样本大小

- 自变量要变：没有（外生）变异（**Variation**）就没有识别（**Identification**）——这是经济计量学中最大的真理。
- 小样本是个令人头疼的难题。
- 小样本条件下，自变量能产生的变异很有限，缺少足够变异，很难把**correlation or relevance** 变量的变异捕获（**Capture**）。
- 微观数据之所以受到研究者青睐，重要原因是样本大，变异丰富；允许研究者引入更多解释变量（自由度减少），结果更稳健；允许采用更高级的方法，比如**IV-GMM**方法。这些外生变量（**IV**）经过感兴趣（内生）变量的映射到（结构性方程中的）被解释变量后，变异更小。（**IV**估计的文氏图-必须是大样本）
- 计算机技术迅猛发展，给处理大样本数据、开发相关技术，提供了强大的技术支持。

测量误差与遗漏变量

- 这是微观经济学者最常遇到的难题。怎么解决：
- One strategy for dealing with omitted variables is to get data on variable and include it
- One strategy for dealing with measurement error variables is to get better-quality data
- These are good strategies but may be easier said than done
- IV offers another approach if instrument can be argued to be correlated with true value of variable of interest and not with measurement error/ omitted variable

拷打数据

- Data Mining。
- 面对结果不理想，学者几乎都有拷打数据的冲动。
- 经济学领域的拷打数据是必须避免的，这是个重要原则。因为，只要拷打，数据就会招供。
- 常见的拷打类型：不断的找借口加权；在模型设定时，没有根据的增加或删除解释变量。
- HWei混合三个微观数据。
- 然而，在一个科学巡查的法庭上，拷打得到的供词不被采信。
- 拷打数据的背后，是违背科学原则和学术道德。在各种压力下，一些学者难免有作弊的冲动。
- 拷打不是伪造。在有经验的学者面前，拷打的痕迹容易侦测到；而伪造数据则难以发现。
- 所以，学术界对拷打数据愿意给予更宽的尺度，但绝不谅解伪造数据。

一半的真理， 是更邪恶的谎言



警告： 不过分迷信大样本数据

- 高于数据的是常识。是公理。
- 大约在**2010**年， 人民银行对储户作了一个大样本调查（声称有几十万样本）， 目的是了解储户对存款利率（**2.25%**）的满意度。
- 结果显示， 超过**90%**的储户“满意”！
- 这不是调查， 是欺诈！
- 人为压低的利率和超过**3%**的通货膨胀， 实际利率为负值。人们竟然还会“满意”？！
- 主观故意可以扭曲大样本结果， 数据是可以伪造的。

经验层面，不可能有完美的证据

- 经济学家大胆使用各种证据，对数据的有限性，要有充分的认识。
- 经验分析，很大程度上是“就汤下面”，不可能有完美的证据。
- 如果有人声称自己的证据是完美的，很可能是伪造的。
- 我们可以做各种改进去接近完美，但对仍然无法实现的部分，要坦然接受。

四、经验分析的方法

- 经验分析要以研究设计为基础（Design based research）。Identification Strategy=Research Design.(AP, 2010)
- 经验分析的价值在于弄清变量间的因果关系并对数量关系进行准确区分。
- 因果关系的判断来自理论、经验、常识、专家意见等。
- 以随机试验为例：“干预-传递-影响-结果”。这是一个逻辑上内在一致的框架。
- “投入-机制-影响——长期目标”。前三个要素是逻辑框架；第四个要素是干预的结果。
- 以罗斯高的试验为例：为贫困农村小学生配发复合维生素—减少贫血发生率—提高学习成绩—积累人力资本和提高收入—更长期的缓解贫困和促进增长。

经验分析的识别

- 满足经验分析可信性的方法，包括随机受控实验和自然实验。
- 1、随机受控实验（RCT）
- 2、自然实验
 - ——工具变量（Heart）
 - ——双差分（Arrow）
 - ——断点回归（RCT的“表亲”）
- 3、匹配估计——PSM, SCM近似于自然实验。

- 数据自身不会作因果推断。
 - 经济计量学不是变戏法，而是对经济学原理及其推论的科学性，进行严格的检验。
 - 所以，在对自己作专业培训时，要相信经济学理论和方法的科学性。不要热衷于变戏法。
-
- 因果关系来自哪里？
 - 来自学者的头脑中，来自经济学理论的训练和推论，来自学者的经验认知，来自专家意见或者生活常识。
 - **Nancy Cartwright.** 因果与证据。

模型设定

- 理论与经验数据，分处在河的两岸，模型设定是将两者联系起来的桥梁。
- 数据结构决定了模型设定。截面数据和面板数据设定的模型不同，因为自变量变异的来源和机制不一样。
- 模型设定探讨因变量与自变量之间的关系。理论帮助你思考，哪些自变量会是真正的解释变量。但经验和学习，则深化这个技巧：这些自变量对因变量的影响是否线性？是否应该有交互项。
- 有了模型设定上的细致思考，不仅能避免拷打数据，还使经验分析变得可信。

从理论到经验（方程）

$$U = pf(N_M + N_F) - w_M N_M - w_F N_F - dN_F$$

- 研究厂商支付的工资性别差异时，这是一个经常见到的理论方程。
- 如何将这个理论方程转化为经验方程呢？
- 思考 $f(\cdot)$ 的经验形式。
- 参数还是非参数？线性还是非线性方程？需要交互项吗？
- 别忽略了天天与之打交道的 $f(\cdot)$ 。
- 理论或假说，和经验证据之间，并没有严谨的逻辑通道。

形而上与形而下的世界

- 观念的世界，存在于我们的头脑中，那是完美的，简单的，稳定的，长期的。
- 现象的世界，则是复杂的，变动的，有限的、临时的。
- 科学研究的任务，是将两个世界连接起来，对应起来。这是经验研究面临的持续的挑战。
- 对经济学研究来说，也同样如此：在形而上世界里构建的理论、模型——通常是逻辑链条，需要经验世界的证据提供一致的、对应的证据——证据链条。
- 一篇经验分析文章中，体现着这两个并行的链条。
- 爱因斯坦，弗里德曼，Maximilian Kasy，方法论。

结论与评述

- 学术规范要遵守，不要偏离了学术主流。
- 在本科阶段，要学好经济学理论和方法，接受规范的训练，掌握可信性革命范畴下的研究框架，使用微观数据做研究。
- 只要指导路径正确，学生努力，本科生也可以做出很好的研究。

- 推荐文献:
- Angrist and Pischke, 2010. The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics.
- Journal of Economic Perspectives—
Volume 24, Number 2—Spring 2010—
Pages 3–30