

经原小班第三次讨论课



起跑线

对应大课：微观经济学前沿



助教：房晨

2024年3月17日



Housekeeping

| 周序号 | 大课时间 | 小课备课时间 | 小课时间 | 小课备课内容 | 挑选案例助教 |
|------|----------|--------|----------|---|---------|
| 第一周 | 2月26/27日 | 2月27日 | 3月2/3日 | 生产要素市场（第18章） | 杜翔宇 黄嘉伦 |
| 第二周 | 3月4/5日 | 3月5日 | 3月9/10日 | 收入与歧视（第19章）；收入不平等与贫困（第19章） | 耿域新 孙昕知 |
| 第三周 | 3月11/12日 | 3月12日 | 3月16/17日 | 微观经济学前沿（第22章） | 宗晶莹 房晨 |
| 第四周 | 3月18/19日 | 3月19日 | 3月23/24日 | 一国收入的衡量（第23章） | 孙昕知 耿域新 |
| 第五周 | 3月25/26日 | 3月26日 | 3月30/31日 | 生活费用的衡量（第24章） | 房晨 程莘航 |
| 第六周 | 4月1/2日 | 4月2日 | (调休轮空) | 无课 | |
| 第七周 | 4月8/9日 | 4月9日 | 4月13/14日 | 生产与增长（第25章）；储蓄、投资和金融体系（第26章） | 宗晶莹 黄嘉伦 |
| 第八周 | 4月15/16日 | 4月16日 | 4月20/21日 | 失业（第28章） | 何禹桥 杜翔宇 |
| 第九周 | 4月22/23日 | 4月23日 | (调休轮空) | 无课 | |
| 第十周 | 4月29日/* | 4月30日 | 5月4/5日 | 货币制度（第29章） | 程莘航 房晨 |
| 第十一周 | 5月6/7日 | 5月7日 | 5月11/12日 | 货币增长与通货膨胀（第30章） | 孙昕知 何禹桥 |
| 第十二周 | 5月13/14日 | 5月14日 | 5月18/19日 | 开放经济宏观经济学的基本概念（第31章） | 程莘航 宗晶莹 |
| 第十三周 | 5月20/21日 | 5月21日 | 5月25/26日 | 开放经济的宏观经济理论（第32章） | 黄嘉伦 何禹桥 |
| 第十四周 | 5月27/28日 | 5月28日 | 6月1/2日 | 总需求与总供给（第33章） | 杜翔宇 耿域新 |
| 第十五周 | 6月3/4日 | 6月4日 | 6月8/9日 | 货币和财政政策对总需求的影响（第34章）；通货膨胀与失业之间的短期权衡取舍（第35章）；宏观经济政策的六个争论问题（第36章） | 何禹桥 房晨 |
| 第十六周 | */6月11日 | 6月11日 | (无) | 无课 | |

注：(1) 第十周周二（也即4月30日）校庆及劳动节调休，钟老师课堂轮空，董老师周一课堂正常
(2) 第十六周周一（也即6月10日）端午节调休，董老师课堂轮空，钟老师周二课堂正常
(3) 由于清明节调休，原排在4月5日（星期五）的课程调整到4月7日（星期日）进行
(4) 由于校庆及劳动节调休，原排在4月28日（星期日）的课程停上

Housekeeping

关于memo与课堂



- ✓ 如果没有批改, 请提醒助教
- ✓ 如果没有收到作业通知, 也请提醒助教
- ✓ 英文memo
 - ✓ 如果有理解错误请及时和我反馈
- ✓ 课上有问题随时提出
- ✓ 请带上名牌
- ✓ 尽量坐教室前排
 - ✓ 方便看白板演示
- ✓ 需要什么奖品可以提hh

◆ 请不要迟交memo!

上节课



- ✓ 生产要素的成本最小化计算?
 - ✓ C-D函数性质的数学推导
- ✓ 岳阳调研的结论与工作建议
 - ✓ 经济学原理与现实的鸿沟
- ✓ 残疾人面临的生存困境
- ✓ 中介的作用与意义
 - ✓ 减轻信息不对称现象
- ✓ 政策建议



New Session: One-Slide Show

本学期的全新环节：One-Slide Show

- 每周1~2位同学
- 时间在3分钟以内
- 课上展示1页PPT（可以超过1页，可以用英语）

□ 对于阅读材料的思考 😊

- 将memo内容以PPT形式呈现

□ 对于生活中经济学现象的思考 😊

- 排队占座、修自行车、食堂饭菜、游戏充卡

□ 分享经济学相关的阅读感悟 😊

- 相关书籍、公众号推文、新闻报道.....

➤ 提供物质奖励（本周纪念品：《好的经济学》）

➤ 不用写当周的memo

希望大家积极报名！！！

| 周次 | 日期 | 主题 | 7班 | 7班 | 8班 | 8班 |
|------|-------|---|-------------|-----|-----|-----|
| 第1周 | 3月3日 | 生产要素市场 | 例如: Charlie | | | |
| 第2周 | 3月10日 | 收入与歧视, 收入不平等与贫困 | | | 何泽宇 | 严嘉妮 |
| 第3周 | 3月17日 | 微观经济学前沿 | 罗家乐 | 刘凯欣 | 刘鲁宁 | 彭泽豪 |
| 第4周 | 3月24日 | 一国收入的衡量 | | | 董天一 | |
| 第5周 | 3月31日 | 生活费用的衡量 | | | 范圣烨 | |
| 第6周 | 调休 | | | | | |
| 第7周 | 4月14日 | 生产与增长, 储蓄、投资与金融体系 | | | | |
| 第8周 | 4月21日 | 失业 | | | 宫梓宁 | |
| 第9周 | 调休 | | | | | |
| 第10周 | 5月5日 | 货币制度 | | | 王宇轩 | |
| 第11周 | 5月12日 | 货币增长与通货膨胀 | | | | |
| 第12周 | 5月19日 | 开放经济宏观经济学的基本概念 | | | | |
| 第13周 | 5月26日 | 开放经济的宏观经济理论 | | | | |
| 第14周 | 6月2日 | 总需求与总供给 | | | | |
| 第15周 | 6月9日 | 货币和财政政策对总需求的影响; 通货膨胀与失业之间的短期权衡取舍; 宏观经济政策的六个争论问题 | | | | |
| 第16周 | 无课 | | | | | |

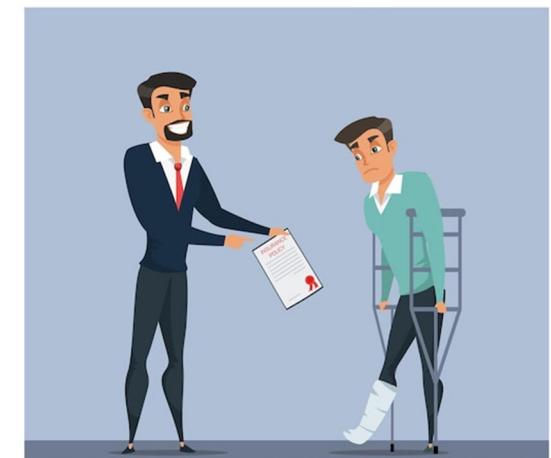
Basic Discussion

中介

- ✓ 对于中介产生了怎样的**需求**?
 - ✓ 残疾人群体
 - ✓ 企业
- ✓ 中介发挥的**作用**是什么?
 - ✓ “价格发现”
 - ✓ 消除信息差
- ✓ 为什么这个案例的中介**令人气愤**?
 - ✓ 克扣
 - ✓ 演唱会门票市场的黄牛本质上也属于中介
- ✓ 调节**供需关系**除了依靠中介还有哪些方式?
 - ✓ 政府
 - ✓ 公开透明的劳动力市场

歧视

- ✓ 为什么会出现**歧视**?
 - ✓ 价格歧视（盈利）、第一学历歧视（信号）
- ✓ **坏处**?
 - ✓ 公平，机会
- ✓ **好处**?
 - ✓ 利润，效率

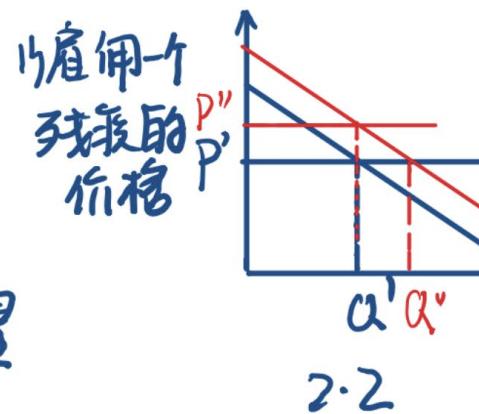
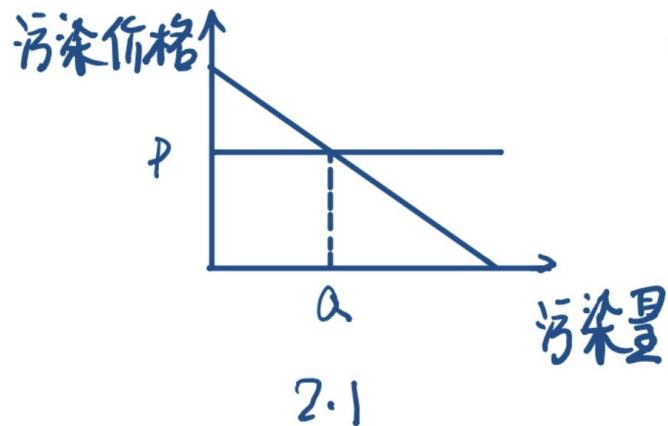




Basic Discussion

□ 负外部性

王宇轩



少雇佣一个残疾人的价格理解为企业对雇佣一个残疾人的评价 = 正常人生产价值 - 残疾人生产价值 (同等工作) 表现为高技术公司雇佣残疾人数量 司对雇佣残疾人的需求曲线和救助公司平均工资。

Basic Discussion

□ First-best solution

假如你进入了这样一个情景游戏，你在后疫情的经济下行时期运营一家企业，望着 Income Statement 中残保金支出一项你陷入了沉思：

刘祎旸

- A. 为残疾人提供就业岗位（企业人数*1.5%）
- B. 缴纳因人数不足的残保金（(企业人数*1.5%-实际提供岗位) *企业平均薪资）
- C. 选择提供“挂靠”的中介企业帮助完成残保指标

“挂靠”中介泛滥的现实告诉我们，不少企业选择了C。初读案例，我的首先反应是义愤填膺，且不论中介如何克扣残疾人工资、缺少对残疾人劳动权益的保障——供给根本来源于需求——这些钻政策空子的企业难道没有社会公德吗？冷静下来，感性总是先让我们看到那些被忽视的问题，而理性或许才更有助于问题的解决，不妨先试图理解企业的抉择，估测A、B项的机会成本。

Policy Discussion

针对以下若干政策建议，从**可行性**角度排序：

- ① 提高残疾人**最低工资标准** **(政府)**；
- ② 针对福利企业进行**税收减免** **(政府)**；
- ③ 根据残疾人的地区分布设定对应的比例 **(政府)**；
- ④ 求职阶段不涉及与**残疾**相关的问题 **(企业)**；
- ⑤ 在工作场所安装**无障碍设施** **(企业or政府)**；
- ⑥ 加大对残疾人**职业教育**的投入 **(政府)**

- 排序过程中你**考虑**了哪些因素？
- 如果从**政策效果**角度排序呢？
- 现实中的**评价标准**是什么？



Interesting MEMO

其中一个可能的方法是以政府为主导的产业创造，以台湾地区为例，公益彩券为当地政府的公营事业，残疾人士可透过向政府申请营业许可来开设彩券店铺，抑或以流动摊贩的形式进行零售，在此政策下，较有能力参与工作的残疾人士得到稳定的就业扶持，虽然不免还是存在弊端需要进一步限制，如家庭其他成员以残疾人士名义申请执照盈利等。此外，也有非政府主导的组织为其创造就业缺口，如喜憨儿基金会，透过体系化的培训计划，基金会使残疾人士参与烘焙工作，并由基金会负责市场贩售，在市场中以残疾人制作作为名义吸引民众购买商品。因此除在现有企业安排残疾岗位外，创造由残疾人构成的人资需求市场亦是一种解决方法。

孙湘旻



Interesting MEMO

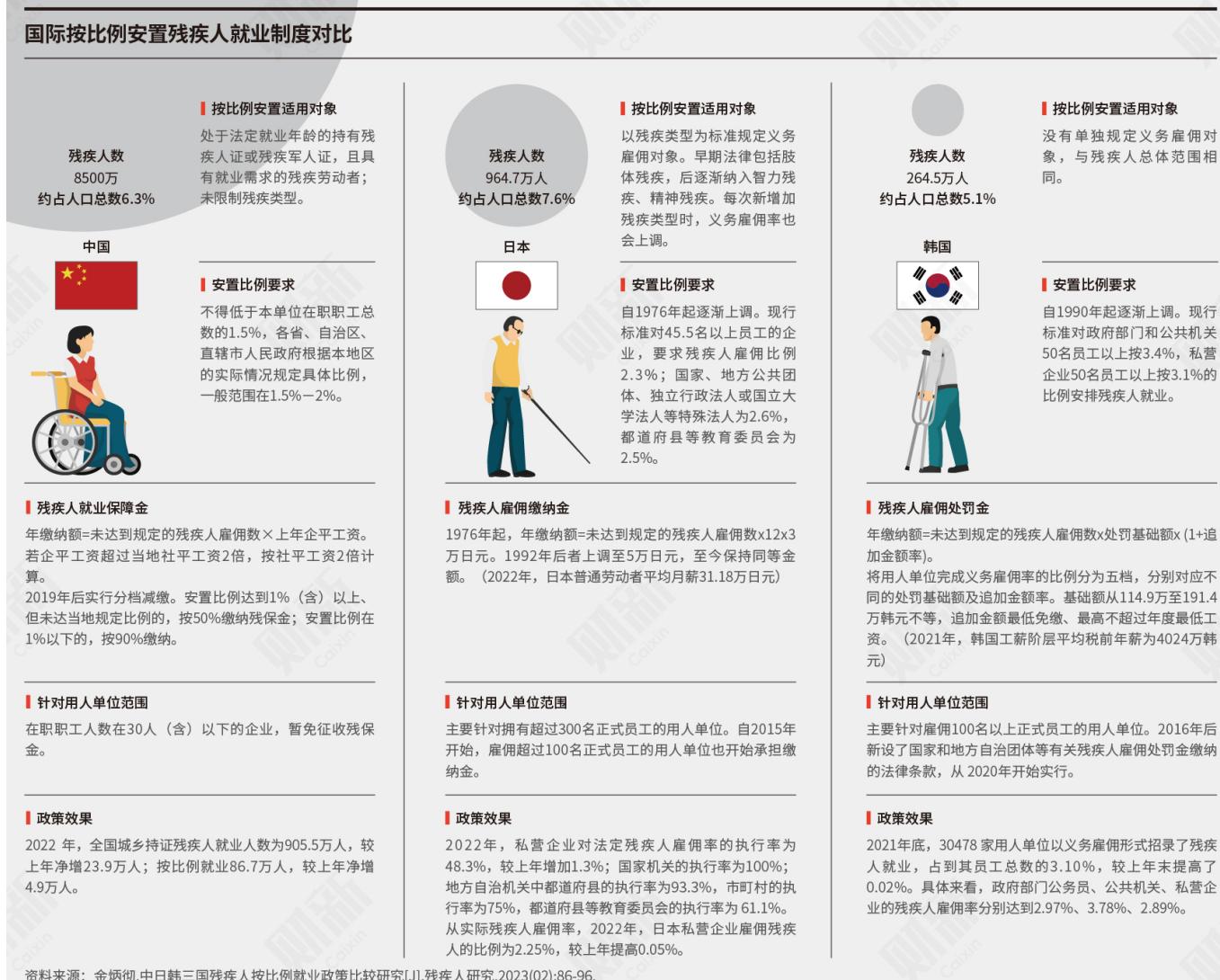
我看到这个问题的时候突然想到了一个很有趣的事情，如果残疾人就业的数据是可以通过“挂靠”这种“假就业”拉高的，那么如果是普通青年人的就业数据，是不是也是通过某种方式来拉高的？是所谓“灵活就业”吗？还是现在巧妙的“就业去向落实率”统计口径？在我国每周工作一小时以上就不算失业，按照最低工资换算每天收入不到三人民币，这样的统计数据真的合理吗？或许这有些偏题，但是我想如果完全正常的青年人就业数据就已经被这样虚假地拉高，那么残疾人就业数据就更不用说了——换句话说，我们在就业市场中又何尝不是一种“残疾人”，况且企业还不用为我们缴纳残保金。

魏冠曦



Extension

国际按比例安置残疾人就业制度对比



资料来源：金炳彻.中日韩三国残疾人按比例就业政策比较研究[J].残疾人研究,2023(02):86-96.

从这张表中：

- ✓ 中日韩在残疾人就业政策上有何差异？
- ✓ 中国有无可借鉴的地方？

Extension & Bonus Question

“贫困指标”

- 如何度量贫困？有怎样的指标？

$$\text{平均收入} \quad m = \frac{1}{n} \sum_i y_i$$

贫困线and贫困率

HC : 低于贫困线的人口
 p : 贫困线 y : 收入

$$\frac{HC}{n}$$

$$\frac{\sum_{y_i < p} (p - y_i)}{n \cdot m}$$

$$\frac{\sum_{y_i < p} (p - y_i)}{p \cdot HC}$$

国际上三种通行的贫困率度量指标

- 贫困人口比例 (HCR, head count ratio)
- 贫困差距比例 (PGR, poverty gap ratio)
- 收入差距比例 (IGR, income gap ratio)

在你看来：

- 三种指标应用范围如何？
- 三种指标各有怎样的优缺点？



Extension & Bonus Question

“贫困指标”

假设一个国家贫困线为1000元/年·人
 处于贫困线以下200人
 一组：人均收入为500元/年；
 另一组：人均收入800元/年
 两组各100人

国际上三种通行的贫困率度量指标

贫困人口比例 (HCR, head count ratio)

$$\frac{HC}{n}$$

贫困差距比例 (PGR, poverty gap ratio)

$$\frac{\sum_{y_i < p} (p - y_i)}{n \cdot m}$$

收入差距比例 (IGR, income gap ratio)

$$\frac{\sum_{y_i < p} (p - y_i)}{p \cdot HC}$$

假设政府有20000元扶贫预算，要决定如何分配给200个贫困人口

两种分配方式：

平均分配——每人100元

极端分配——只给第二组人，每人200元

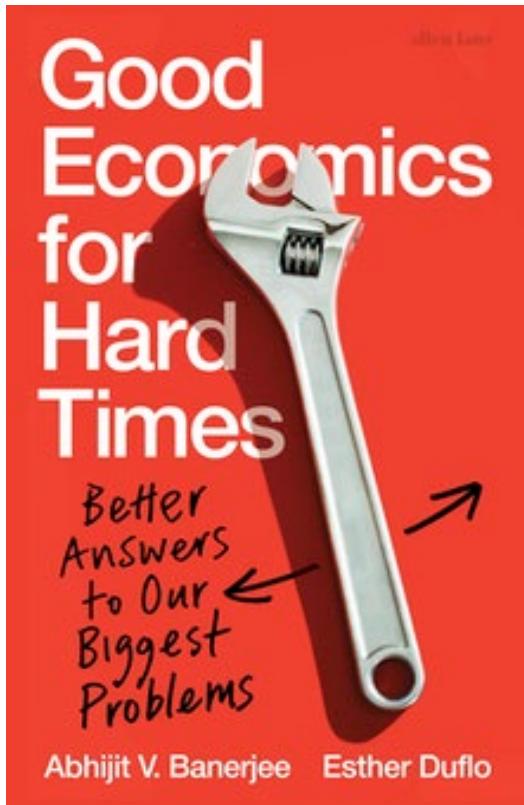
上述三种指标分别会得到怎样的扶贫政策含义 (policy implication) ?

HC：低于贫困线的人口

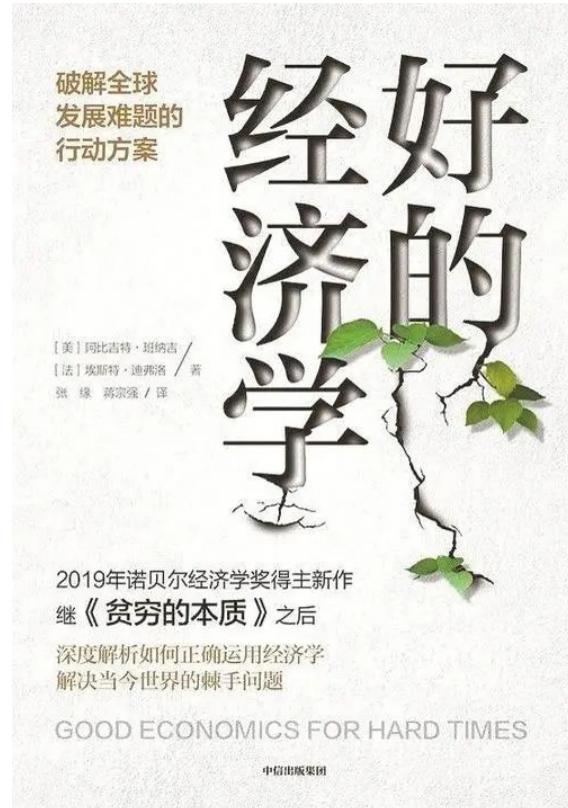
p：贫困线 y：收入

$$\text{平均收入 } m = \frac{1}{n} \sum_i y_i$$

About The Material



《好的经济学》（第九章：资助与关注-“起跑线”一节）



关键词：发展经济学、随机对照实验（RCT）

正文：

- ✓ 教育与发展的实验——“扶贫先扶智”

附录（财新周刊2019年10月16日）：

- ✓ 发展经济学的历史
- ✓ RCT的争议？
- ✓

经原知识点：微观经济学前沿

- 不对称信息
- 政治经济学
- 行为经济学

About The Material

落选材料

《财新周刊》>民生>正文

财新周刊 | 中产转向经济型托育，倒闭潮后如何满足需求

来源于《财新周刊》 2024年02月19日第07期

简 繁 听报道

托育遭遇寒冬，各地大力推进建设的普惠托位，能否显著提升送托率，引导行业复苏？



1月28日，北京，乐融儿童之家上地校区大门紧锁。图：受访者供图

《财新周刊》2024年第07期（02月19日）

托育行业：成本高，利润率低

服务**供给总量**与**实际需求**之间极不匹配

服务**供给结构**与**家庭诉求**之间极不匹配

服务**供给质量**与**家庭需求**之间极不匹配

解决**信息不对称**的问题

结构：

- ✓ 倒闭潮又起
- ✓ 如何完成4.5？
- ✓ 破解生源迷思
- ✓ 低价优质如何实现



Basic Discussion

四个实验

✓ 实验一：“向机会迁居”（美国） ○

- ✓ 实验内容：为严重贫困地区租住公共住房的家庭提供政府住房券——**一半随机迁入“好社区”**
- ✓ 实验结论：早期，女孩的精神状态和表现更好；20年后，更高的工资、更有可能上大学、更好的小区、更不可能成为单身母亲

✓ 实验二：班级规模实验（美国） ○

- ✓ 实验内容：将班级规模从**20~25人**减少到**12~17人**
- ✓ 实验结论：短期内提高分数，增加上大学的机会，未来生活更好

✓ 实验三：“全国启蒙”和“学前教育实验”（美国） ✗

- ✓ 实验结论：短期产生影响，但对考试的影响逐渐消失，甚至**逆转**
- ✓ 学前教育的有效性似乎与计划的质量有巨大关联（全天运作，家长参与，教师示范）

✓ 实验四：幼儿园数学课程（印度德里） ○

- ✓ 实验内容：随机选一些幼儿园**推广这套课程**
- ✓ 实验结论：玩数学游戏进步很快，充满热情，学到大量知识

Basic Discussion

优点 

✓ 提供微观视角

- ✓ 改变了发展经济学前辈们**指点江山**的风格
- ✓ 从更为微观的视角去**观察穷人**
- ✓ 微观理论基础，实证发展经济学

✓ 因果识别 (**Causal Inference**) 更准确

- ✓ 相比实验室里的实验：更接近真实环境，且更容易操控，更能满足研究者的需要
- ✓ 相比自然实验 (**Natural Experiment**)：更随机，避免内生性

✓ 得到**更新的洞见**（反思扶贫政策得失）

- ✓ 免费午餐和课本
- ✓ 驱虫药
- ✓ 疫苗接种
- ✓ 小额信贷

缺点 

✓ 研究的出发点

- ✓ 并非为了填补最紧迫的知识空白（能做实验的都是小问题）
- ✓ **助推 (nudge)**
- ✓ 难以推广到大规模，无法分析一般均衡影响

✓ 结论的**可靠性**

- ✓ **外部有效性差**（不同政府、制度、规模、国民）
- ✓ 难以解释**作用机制** (mechanism)

✓ **实验成本与可行性**

- ✓ 费钱费时费力
- ✓ 难重复

Interesting MEMO

自然科学&社会科学

最初认识随机对照实验，是在初中生物课上。以不太科学准确的语言概括而言，这种实验方法指：实验者想研究一个变量对研究对象的影响时，会将研究对象随机均分为两组，一组施加变量，标记为**实验组**，一组不加变量，标记为**对照组**。凭我有限的学习经验，总感觉这种方法和生物学特别适配，实际上这种印象也在上周“神奇的免疫”讨论课（没错，我就是一志愿抢上神免的幸运鹅）上得到了验证，吴老师说讲个笑话，生物学家想知道螃蟹的腿是否和听觉有关，就把螃蟹的腿卸掉，然后大喊一声“跑”，螃蟹没跑；把腿再按回去，再以同样的分贝大喊一声“跑”，螃蟹跑了，证明螃蟹腿确实和它的听觉有关——这就是生物学最基本的研究方法。（右图为课上讨论的2011诺贝尔文献的实验结果之一，看起来复杂其实只是数个RCT的叠加。）

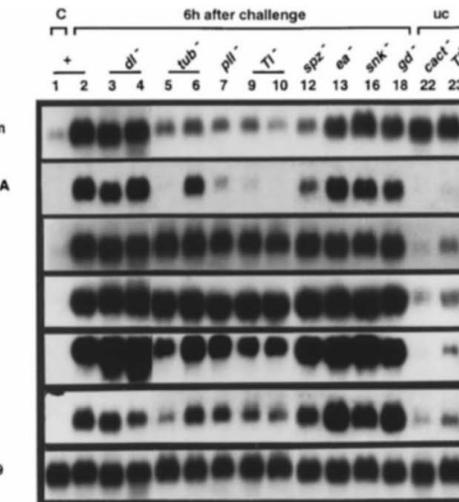


图 1 野生型和背腹突变型成虫抗菌基因的转录谱

该图显示的是基因转录的量。

横轴：C : Control, 1 道：未经感染刺激（无菌针蘸无菌 H₂O 直接刺扎果蝇），2-18 道：经感染刺激 6 小时后（将无菌针尖蘸菌液后刺扎果蝇），22-23 道：未经感染刺激。

+：野生型果蝇（即正常果蝇）。dl-、tub-、pII-、spz-、ea-、snk-、gd-：均为背腹形成通路上的基因，“-”代表该基因突变出现功能异常，其中 dl-、tub-、Ti 测试了两个不同的突变体。Cact-也是背腹形成通路上的基因，但它是抑制性基因，即 Cact-会导致该通路持续激活。Tl^o 是 Ti 的功能获得突变体，指该突变导致 Ti 基因持续激活。

纵轴：各个抗真菌肽和抗细菌肽的转录量。Drom : Drosomycin (果蝇霉素), CecA : cecropin, Dipt : diptericin, Drc : drosocin, Att : attacin, Def : defensin (防御素), rp49 : control。

Interesting MEMO

自然科学&社会科学

性”。“随机性”即抽奖式的客选取样本，以保证实验对象的特征符合一般正态分布，进而保证样本对整体的有效代表。生物学研究中，统计生物种群密度时采用的五点取样法和等距取样法即是“随机性”的典型体现。“对照性”则是在实验

王宇轩

RCT：从理解到实操？

随机对照实验的概念并不难理解，或者说，对于学过生物高中生来说，对照实验的要点已经深深刻在了 DNA 中（因为答题要写）。假如我们要研究光照对植物生长的影响，那么需要将若干生长状态类似的植物随机分成两组，一组放在光下，一组放在阴暗处，保证水肥湿度温度等等其他条件相同，最后检测和统计对照组和实验组的生长情况，得出结论：其他条件相同时，光照会促进植物生长，诸如此类……

纪佳彤

Interesting MEMO

The Abdul Latif Jameel Poverty Action Lab (J-PAL)



以下是额外查找的资料。

吴泓历

贫困行动实验室（Poverty Action Lab），也称为 J-PAL，是由 2019 年诺贝尔经济学奖得主 Abhijit Banerjee、Esther Duflo 和 Michael Kremer 等人创立的，专注于使用 RCT 方法研究减贫干预措施的效果。该实验室的研究遍及全球多个国家，涵盖教育、健康、金融等等。

贫困行动实验室的一项著名研究集中在印度，旨在评估通过提供辅导来提高学生的学习成绩。研究发现，针对成绩落后学生的补充辅导极大地提高了他们的测试成绩，揭示了针对性教育干预在提高教育成果方面的有效性。另一项研究关注于肯尼亚，评估免费提供蚊帐是否能有效减少疟疾的发病率。结果显示，免费分发蚊帐显著降低了疟疾病例，这一发现支持了公共健康干预措施的重要性和效果。

Interesting MEMO

RCT in Indonesia

One example is a cluster randomized controlled trial of a community-based initiative to reduce stunting in rural Indonesia, a study published by the National Institutes of Health.² The study involved 95 treatment and 95 control sub districts across three provinces. Despite a significant project investment and implementation, the results showed no significant impacts on stunting or other undernutrition outcomes one year after the project's completion. The study highlighted the challenges of implementing integrated interventions to reduce child stunting and emphasized the importance of considering implementation realities and best practices in project design. While some positive impacts were observed on proximal outcomes related to maternal and child nutrition, the overall findings underscored the complexities of addressing stunting effectively in practice.

曾静萱

Interesting MEMO

RCT的价值

在《贫穷的本质》一书中，主张 RCT 的班纳吉和迪弗洛夫妇提到了两个著名研究案例，充分体现了 RCT 方法的独特优势。其中一个例子，是在贫穷国家防治疟疾流行的政策选择。通过广泛使用杀虫剂处理过的蚊帐，可以有效降低疟疾感染率。然而在非洲发放的免费蚊帐，使用率很低，这造成一个重大的政策难题。有人提出假设，正因为蚊帐免费，人们使用率低，应当设置价格。而班纳吉通过很多村庄的对照，发现免费发放不是蚊帐使用率低的主要原因，而以补贴价卖出蚊帐，也不会提高它们的使用率。相反，以一个很小的奖励鼓励人们开始使用蚊帐，让他们体会到使用蚊帐的具体好处，在奖励结束后，蚊帐的使用率也不会下降。从统计学角度来说，班纳吉和迪弗洛的 RCT 方法提高了经济学理解特定经济现象时的内部效度，即 RCT 可以相对精确地测定特定的因素（蚊帐是否免费）和特定的现象（蚊帐使用率高低）之间的因果联系——这对检验特定建设项目的具体效果和设计新的社会政策格外有用。

洪念一

Interesting MEMO

RCT or 试点实验

政策制定作提前的准备实验。但弊端在于 RCT 一是无法有效理清某一经济学现象的因果关系逻辑，并且大多数时候会出现偏差甚至于反效果。并且开展 RCT 实验成本很高，不容易再次验证复现。因此，RCT 应该被用作为在理论逻辑较完善的宏观政策领域进行试点实验的一种方式，这样能够合理规避直接大规模推行政策所带来的风险性，并且应该应用于人数较多、范围较广的宏观范围，因为小规模样本通过 RCT 很容易由于变量很多不具有代表性而产生很大误差。**胡景添**



Interesting MEMO

实际上，让经济学更“科学”这个意图本身就是值得商榷的。面对贫困、教育等宏观议题，社会学、政治学等“非科学”体系的影响无法剥离。不能企图通过 RCT 去解释这个应该“以人为本”的社会。因此，经济学家可以调用经费通过 RCT 去研究贫困，但在调用之前，他们也应该想想，这些钱可以解决多少穷人的温饱问题。[王以予](#)

the scope and frequency of such studies. This financial barrier can restrict the range of questions explored through RCTs and also raises questions about the opportunity costs of investing heavily in such research.

[闵屹珊](#)

在 RCT 的背景下，机会成本可能体现在将有限的资源——如资金、时间和专业人力——投入到特定的研究上，而这些资源本可以用于其他可能在短期内产生更大社会效益的活动。例如，一个关于教育干预措施的 RCT 可能需要数百万美元的资金和几年的时间来完成，这些资源如果用于直接改善学校设施或提高教师工资，可能会立即对学生的学习环境产生积极影响。

在 RCT 中，当道德风险发生在当行为的后果部分或全部由其他人承担时，每个个体可能会采取更加冒险以及更激进的行为。这种情况可能出现在参与者知道自己正在参与一个实验时，他们的行为可能因此而改变，不再反映他们在日常生活中的真实行为。例如，参与一项旨在评估健康干预效果的 RCT 的个体，可能因为知道自己的健康状况被监测而采取更健康的生活方式，这种改变可能并不是干预本身的效果，而是参与者试图在实验中表现得更好的结果。

机会成本与道德风险

[林宜蓁](#)

Interesting MEMO

Sample Attrition

(1) 实验设计的“随机”是真随机吗？可以说随机抽取一部分人参与实验，但人不同于植物或者小动物，开展这种社会性实验的时候必须尊重被试的意愿，就像试吃新药，吃不起药的患者比起什么药都吃得起的患者机会成本更低，那么取样的随机能否消除隐隐干扰实验结果的贫富差距带来的影响呢？（只是举个例子，或许不太准确）对于一些时间跨度较长的跟踪实验，也可能会出现被试失联的情况，导致类似幸存者偏差的情况出现。（2）怎样量化不同因素的影响？怎样分析结果？在知乎看到这样一个表，感觉想要做好 RCT、分析好数据，需要相当的数学和专业的分析工具等等。

纪佳彤

Interesting MEMO

New Trend

在深入了解 RCTs 的过程中，我看到一个很有趣的例子：Daniel Zizumbo-Colunga 在他的论文中提到墨西哥的”有条件现金转移计划”在初期被认为对经济十分有利，但实际上根据他的分析和调查，这项计划从长期来看反而带来更多的社会暴力和动荡。⁴ 我们可以用一般均衡理论或溢出效应去解释该现象，换言之，当我们在设计和进行 RCTs 的过程中，往往会忽略很多外生变量，导致实验的效果并不持续，甚至最后会出现惊人逆转。

但我认为 RCTs 的未来仍然值得推广，不过可以做一些形式融合以提高其经济价值。其中一个新的趋势便是结合 structural modeling 和 reduced form models⁵，让实证研究基于一定的理论与预测，若预测的模型与实验不符，则可以更迅速地从模型中找到漏洞所在；且运用结构模型也能从宏观角度估算政策的实施成本与效益，从而更好地帮助实验结果落地。

严嘉妮

Interesting MEMO

Policy Implication

不难发现，RCT 得到的结论，诸如“提高学前教育、托儿服务、养老机构的质量等”此类的建议，实际人尽皆知；学者只是通过实验手段加以科学印证罢了。而当结论得出后，我们接下来要问的问题就是——“然后呢？”人们却只得哑口无言。实验很科学，愿景很美好，但在庞大、复杂、多变的现实面前，这些建议却只能在自我论证的构想中找到一席之地。

王以予

Interesting MEMO

“居高临下”？

RCT 作为一种实验方法，最多的应用领域就是发展经济学，其主要研究内容是贫困落后的农业国家或发展中国家如何实现工业化、摆脱贫困。就像在上次 memo 中提到的设计建设无障碍设施者往往是健全人，作为强势国、经济文化社会各方面都享有相对强权的发达国家去研究发展中国家和地区如何发展，本来就是一个无法落实、无法感同身受设身处地的问题。在上学期的学习中我同样有类似的疑问：经济学发源于资本主义经济更加发达的欧美，在学习和举例时，应用的大多数都是美国的事例。然而，在我跳出教材的情境试着解释现实生活时却遭遇了“货不对版”的问题。连中美之间的差异都如此明显，更不要谈西方经济学是否能适用于非洲或拉美某个情况复杂的发展落后地区了。

朱子宁

Interesting MEMO

不理想的RCT

泰国的 Puey Ungphakorn 经济研究所曾与政策评估和设计研究所 (RIPED) 和教育平等基金 (EEF) 一起启动了 RCT 人力资本研讨会系列。俄亥俄州立大学的约书亚·霍利教授以 “Labor force impacts of enhanced employment counseling in the United States: The limits of an encouragement RCT design” 为主题的演讲。他讲述了试图评估 “Comprehensive Case management and Employment program” (CCMEP) 的经验。一开始，该项目将评估每个年轻人是否需要帮助，而且需要在哪个方面。实验有超过 13,000 个样本，但 RCT 有错误。因为有一个治疗小组只注册了 4%，这可能是由推动引起的，选择不让人们看到项目的好处。此外，不应该注册的对照组来注册近 4%，这可能是由技术错误引起的，因为地方政府对信息系统有问题。后来，虽然他可以通过运用发展已久的行政数据与 quasi-experiment (建立独立变量和因变量之间的因果关系，可并没有随机分配) 来评估 CCMEP，可仍然有一些限度使得结果不是很理想。最终，霍利教授提出了几点建议。首先，如果条件不适宜，不要着急进行 RCT 实验。其次，进行小规模的 RCT 实验可以更有效地控制条件和保证质量。最后，数据库系统的开发和数据集成是非常重要的。

邱慧琳

Interesting MEMO

Online Survey

strategies for poverty alleviation and educational improvements, among others. A notable example of RCT application in economics is the JPAL's nationwide online survey in Indonesia during the Covid-19 pandemic, which aimed to understand how the lives of poor households were affected. This survey highlighted the adaptability and relevance of RCTs in providing timely and critical insights into socio-economic issues. It shows how RCTs can be used to gather valuable insights that inform policy decisions, even amidst a global crisis.

林佳敏

Interesting MEMO

Evidence-Based Policymaking

As a student from Indonesia, I know the importance of evidence-based policymaking in driving inclusive development in my country. Indonesia, like many other developing nations, deals with complex socioeconomic issues that require targeted interventions and resource allocation. For example, as Indonesia continues its journey towards sustainable development, integrating RCT methodologies into research projects and policy evaluations can enhance the effectiveness of interventions and contribute to building a more equitable and prosperous society. RCT offers a systematic approach to evaluating the effectiveness of policies and programs, providing policymakers with actionable insights into interventions that are most likely to generate positive outcomes for Indonesian communities.

李静佩

Interesting MEMO

RCT Design 1



虽然马来西亚政府通过各个媒介经常报道关于马来西亚的肥胖率，提醒大众注意饮食，但政府并未实际采取政策，解决此问题。

随机选择两个区域作为实验范围，分别为 A 及 B 区，并且采取盲目原则随机安排各区域的实验及杜绝实验信息传播出去，减少实验结果的偏倚和提高实验结果的可信度（因为马来西亚是个多元种族的国家，大众普遍对马来人有刻板印象，认为他们喜欢吃含糖量高的食物及饮料）：

黎润慧

| A 区 | B 区 |
|---|-----------------|
| <p>在每个超市售卖的包装饮料和甜品上贴上类似“红绿灯”的营养等级标签，其中：</p> <p>红色 标签表示 A 级（最不健康），</p> <p>黄色 标签表示 B 级（中等），</p> <p>绿色 标签表示 C 级（最健康），</p> <p>以指示其糖成分含量。同时，每个包装上将包含一个二维码，以便收集它们的销售结果。</p> | <p>不进行任何干预。</p> |

Interesting MEMO

RCT Design 2

Goal of the research

- Determine if ‘bad’ student’s study performance will be enhanced by presence ‘good’ students surrounding them

Participants

- 10th graders, students who just entered high school
- The top 25% of students and the bottom 25% in terms of school entering rank will be chosen

Further specifications about this study

- Find some random high school in China
- There should be around 400 new students entering the high school
- All of their rankings are obscured upon entering the school (parents teachers and students themselves would not know)
- The grade will be split into 10 classes

Interesting MEMO

RCT Design 3

The relationship between GPA score and acceptance rate for offers in the investment banking sector:

陈妍如

1. Participant Selection:

- Identify a diverse pool of candidates applying for positions in investment banking.
- Ensure candidates meet basic eligibility criteria, such as possessing relevant education and qualifications.

2. Random Assignment:

- Randomly allocate candidates into two groups: the treatment group and the control group.
- The treatment group comprises candidates with higher GPAs (e.g., GPA above a designated threshold).
- The control group consists of candidates with lower GPAs (e.g., GPA below the threshold).

3. Application Process:

- All candidates undergo a standardized application process, including submission of resumes, cover letters, and participation in interviews.
- Conduct the application process consistently across all candidates and administered by the same hiring team to minimize bias.

Interesting MEMO

RCT Design 4

One simple phenomenon that I believe can be researched through RCT is the impact of afternoon naps towards stress levels of students. A lot of students, including me and some of my friends take afternoon naps, especially after 8am classes. However, there are others who choose not to do this due to efficiency reasons, or they just simply don't have the time to. From this phenomenon, we may ask whether engaging in 1 hour of afternoon naps every day reduce stress levels compared to no naps with the same amount of overnight sleep.

First, we have to recruit some volunteers that currently all take a nap, or all don't take a nap. There can be no difference, as this will probably influence the result later on. Randomly assign participants to two groups - the nap group and the control group. Those in the nap group are made to take a 1 hour nap every day, and the other group must not take a nap. Say that this routine will go on for a few weeks to a few months, depending on the experimenter. After that is done, we then use validated questionnaires to enquire about the stress levels, and should use other, more objective, measures such as blood pressure, electroencephalograms, or other methods. By carefully designing the RCT and adhering to rigorous scientific standards, we can draw reliable and meaningful conclusions about the effect of daily exercise on stress levels. This includes picking the appropriate sample size, the discipline of the volunteers, as well as proper data cleaning and data analysis. Due to these not being done correctly, some RCTs may have invalid results.

Interesting MEMO

RCT Design 5

研究题目：大班课是否出勤对经济学原理课程成绩的影响

一、实验目的：

本实验旨在通过随机对照实验（RCT）方法，探究大班课是否出勤对经济学原理课程成绩的影响，以期为教育工作者提供有关课程管理和学生出勤率与学生成绩之间关系的实证依据。

二、实验设计

1. 实验对象：选取某大学经济学原理课程的一个大班作为实验对象。
2. 实验分组：将实验对象随机分为两组：出勤组（对照组）和缺勤组（实验组）。出勤组学生正常出勤，缺勤组学生则被随机安排部分课程缺勤。两组学生在其他方面保持相同的学习环境和条件。
3. 实验干预：随机安排缺勤组学生在部分课程缺勤。出勤组学生正常出勤。实验期间，需确保两组学生接受相同的课程内容、作业和考试。

秦子涵

Interesting MEMO

My RCT Study Proposal

RCT Design 6

As I'm deeply interested in the intersection of education and technology, I would propose an RCT to assess the impact of digital learning tools on educational outcomes in underserved communities. Recognizing the stark disparities in educational resources across different socio-economic backgrounds, I'm curious about the potential of technology to level the playing field. My proposed study would involve partnering with schools in low-income areas to randomly assign students to either a treatment group, which would receive access to state-of-the-art digital learning tools (like personalized learning software and online tutoring), or a control group that continues with the standard curriculum without additional digital resources.

林佳敏

Interesting MEMO

A Conducted RCT

I am more than happy to show you an example of RCT I have conducted rather than conceiving an example of it. In 2019, I participated in a university hosted programme by Nanyang Technological University and later co-authored an international paper titled ‘Effect of Social Environment on Brand Recall in Sports Video Games’ in the International Journal of Gaming and Computer-Mediated Simulations (IJGCMS). In my research, the topic at hand is determining whether social communication between players playing sports video games will reduce the brand recall rate of advertisements displayed during the gameplay.

RCT is mainly used in clinical trial, however In social sciences, RCT is an experimental method of policy evaluation that enables understanding of the causal effects of policies with a high level of confidence. It involves assigning subjects in a fully randomised manner to either a treatment group that receives the policy being studied or a control group that does not.

刘煜鸿

In 1 of the studies of the research, 19 participants were randomly assigned to either Experimental Group 1 ($n = 9$) where they played against the computer i.e. single player or Experimental Group 2 ($n = 10$) where they played against each other i.e. multi-player. Participants in Experimental Group 1 thus played the game without communicating with other people while the participants in Experimental Group 2 were able to communicate with each other while playing the game. **In this case, Group 2 ‘received the policy treatment’ while Group 1 does not.**

Interesting MEMO

Perspective from Psychology

Psychology research investigates human behaviors on a much higher extent than numerous other social sciences. As we know, human behavior is often non-duplicable and context-dependent. When RCT is employed as its research method, the conclusions drawn lack external validity. For instance, a classic experiment by Tversky and Kahneman examined the impact of anchoring bias on decision making, with high implications regarding pricing decisions and marketing. They created a highly controlled environment, influencing participant behavior by simply presenting them with strips of paper and asking for predictions. The study was replicable and all in all supported the existence of anchoring bias, however critics still argued that the findings lacked applicability to real-world situations. Consequently, RCTs in psychology are seen as supporting hypotheses rather than providing absolute proof.

宫梓宁



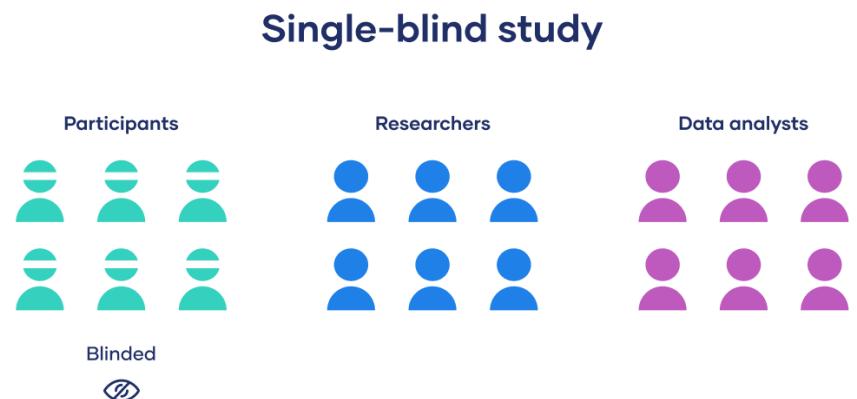
Extension 1: Blind Experiment

Example: Single-blind vaccine study

You have developed a new flu vaccine. In order to test the effectiveness of your new treatment, you run an experiment, giving half of your participants the flu vaccine and the other half a fake vaccine that will have no effect (to control for the **placebo effect**).

If participants in the control group realize they have received a fake vaccine and are not protected against the flu, they might modify their behavior in ways that lower their chances of becoming sick – frequently washing their hands, avoiding crowded areas, etc. This behavior could narrow the gap in sickness rates between the control group and the treatment group, thus making the vaccine seem less **effective** than it really is.

To prevent such an outcome, in a **single-blind** study, you hide from the participants which vaccine – real or fake – each of them received.



Extension 1: Blind Experiment

Example: Double-blind vaccine study

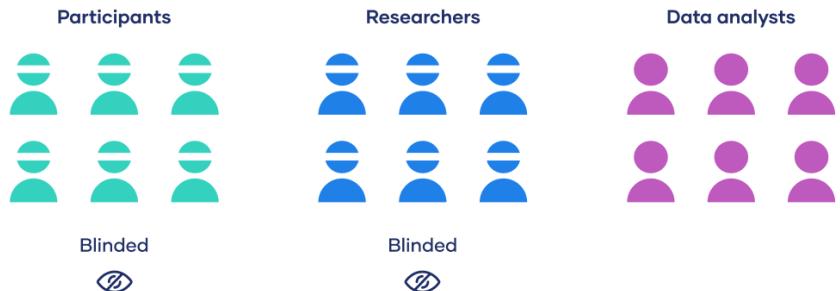
In the flu vaccine study that you are running, you have recruited several experimenters to administer your vaccine and measure the outcomes of your participants.

If these experimenters knew which vaccines were real and which were fake, they might accidentally reveal this information to the participants, thus influencing their behavior and indirectly the results.

They could even directly influence the results. For instance, if experimenters expect the vaccine to result in lower levels of flu symptoms, they might accidentally measure symptoms incorrectly, thus making the vaccine appear more effective than it really is.

To avoid this, you hide group assignments from both the participants and the experimenters giving the vaccines – a **double-blind** study.

Double-blind study





Extension 1: Blind Experiment

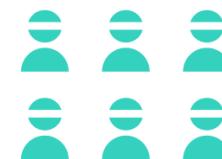
Example: Triple-blind vaccine study

In your vaccine study, you have also recruited assistants to analyze the data you gathered on flu infection rates. You decide to hide the group assignments from the participants, the people administering the experiment, and the people analyzing the data – a **triple-blind** study.

To achieve triple blinding, you assign each participant to group 1 or group 2, but do not inform the data analysts which number represents which group.

Triple-blind study

Participants



Researchers



Data analysts



Blinded



Blinded

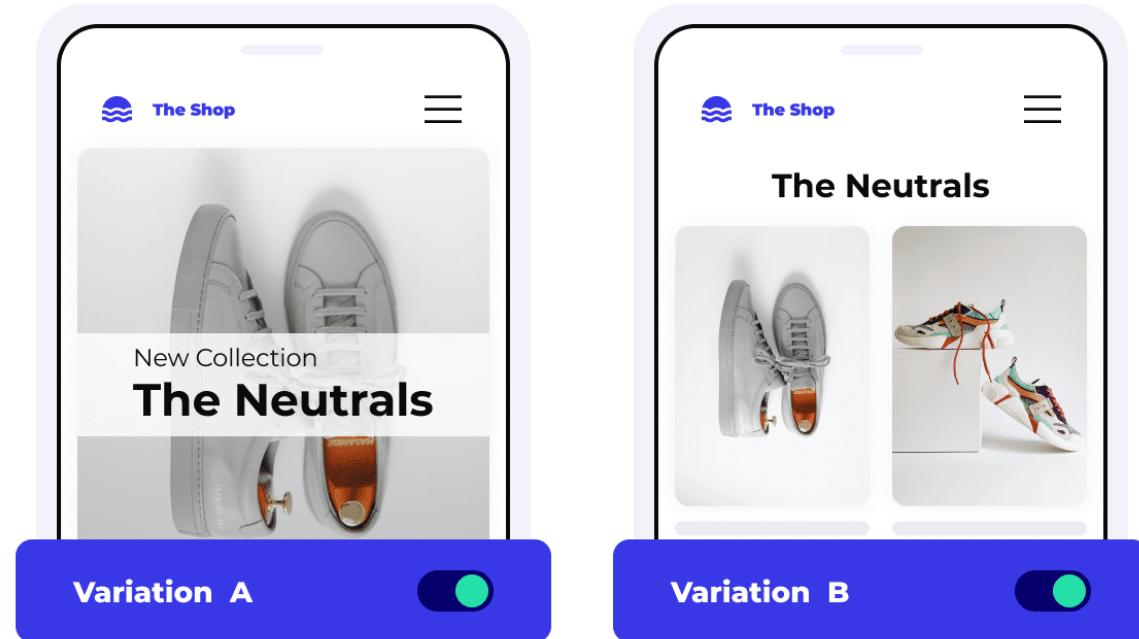
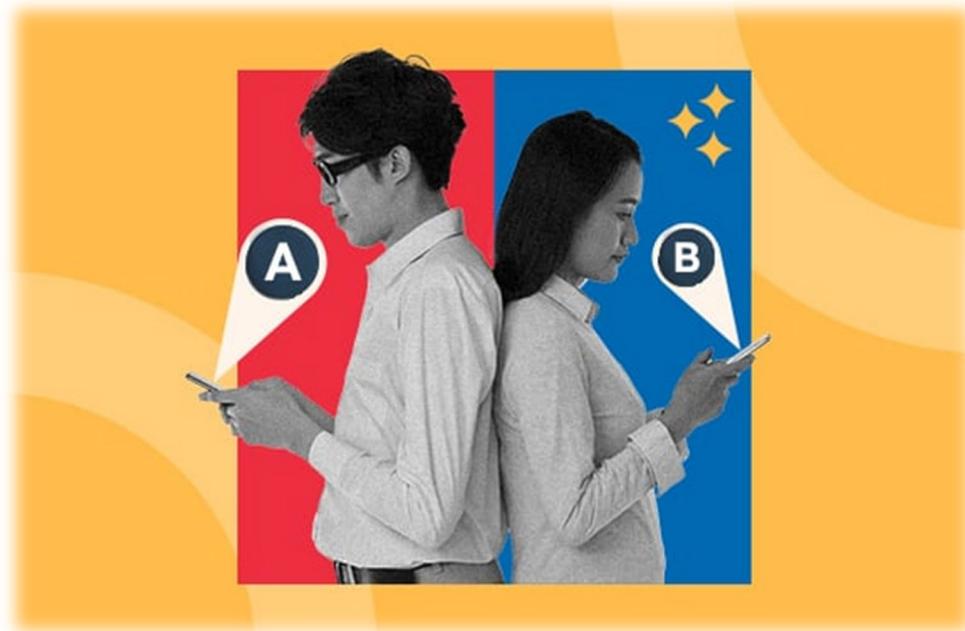


Blinded



Extension 2: A/B Test or Split Test

通过软件自身，让不同用户面前，对同一个软件功能展现两种不同的方式，进行效果对比获得用户反馈。



<https://www.kameleoon.com/de/ab-testing>

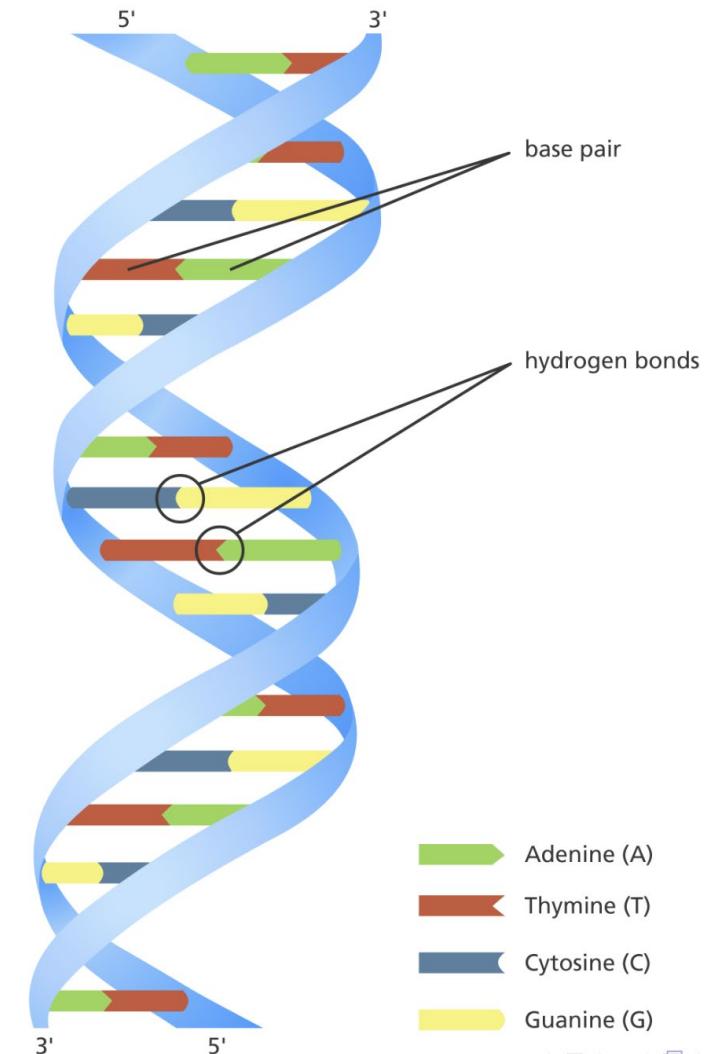
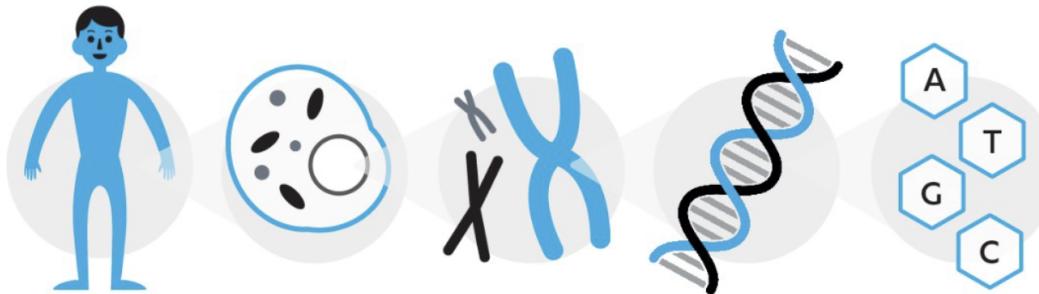
<https://blog.hubspot.com/marketing/how-to-do-a-b-testing>

Extension 3: Genoeconomics

DNA 由四种化学碱基组成：A（腺嘌呤）、G（鸟嘌呤）、C（胞嘧啶）和 T（胸腺嘧啶）

特定的 DNA 序列被称为基因

人类基因组包含大约 30 亿碱基对和 2-3 万个基因





Extension 3: Genoeconomics

孟德尔随机原理

“每个人的基因在形成时都是一个近似随机的过程，并且在个体出生前就已经决定，基因本身不会在出生后受到外在环境因素的影响改变”

因果推断的一个新思路：孟德尔随机法（Mendelian Randomization, 简称 MR）

最早由荷兰科学家 Martijn Katan 于 1986 年提出（Katan, 1986, Lancet）

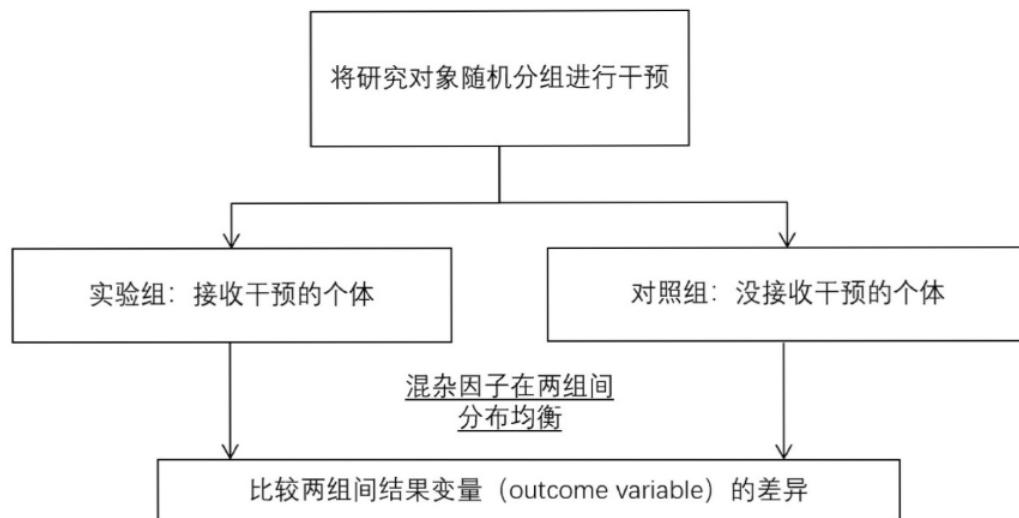
MR

根据孟德尔遗传定律，父母的 DNA 在形成生殖细胞并传给子女时会随机进行分配；因此，假如某个（或某些）基因决定了一种我们感兴趣的个体特征或疾病（比如身高、体重、BMI、糖尿病、高血压、阿尔兹海默症等等），那么子女是否携带这种（或这些）基因就可以被看作是一种天然的随机化过程。这一随机化过程与随机对照研究（Randomized Controlled Trial, 简称 RCT）的精髓十分相似，孟德尔随机法正是巧妙地利用了遗传基因的这一特质来进行因果推断。

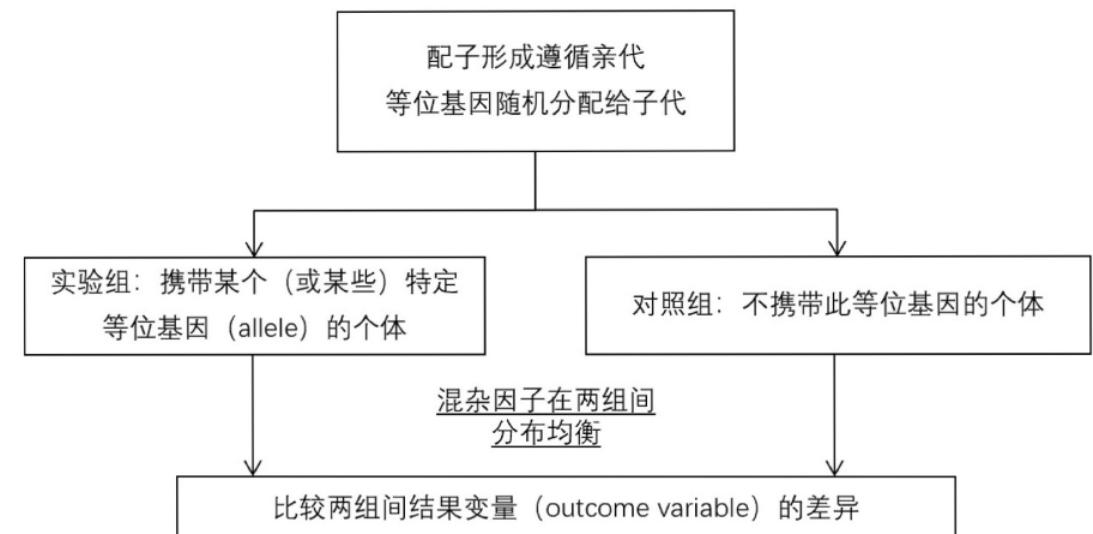
Extension 3: Genoeconomics

“最天然”的 RCT 设计

随机对照研究 (RCT) 原理:



孟德尔随机 (MR) 原理:



Opposite Voice

争议：公平、伦理与道德

4. 伦理性问题。从过往的 RCT 实验中可以发现，实验组的干预措施一般都是积极性的，比如给村里修路，或者帮助穷人孩子进入更好的学校等。经济学实验没有安慰剂，而措施都必须经过社会伦理性的审批，因为很多时候实验造成的影响将会是不可逆的。（这一点和药学实验很像。但是有伦理性审批也并不意味着实验不会对受试者造成负面影响，例如没有抽到入学抽奖的受试者可能要面临更大的心理压力等。）实验者也没有办法要求受试者顺从实验安排或者持续接受实验。

魏冠曦

ps：助教好～我对 RCTs 有个小疑惑，如果我们一般要进行的干预都是正面的，eg 给予更优质的教育机会，那么另一边的对照组的利益不是会受损吗？甚至对他们的整个人生都有负面影响。这样符合伦理道德吗？

严嘉妮

Opposite Voice

争议：公平、伦理与道德

RCTs itself can sometimes be unethical. Taking the same example from OHIE, it is important to note that participants of OHIE are chosen by the lottery method. Although this ensures that the data is random and ultimately unbiased, it does raises ethical questions in regards to equitable access to healthcare. It can be concerning that in OHIE, an individuals' access to healthcare can be determined by the lottery's probability, while there is a high probability that several individuals who are not chosen may have more serious health conditions that may require serious intervention and had to be subsidized by the Medicaid. Additionally, taking into account their low-income status, their unluckiness in the OHIE program means that they may not be able to afford healthcare independently, thereby posing higher risks of actually losing their lives or the lives of their loved ones. These possible incidents can be attributed towards the randomness in RCTs, in which the experimentation itself may cause harm towards a specific group of individuals.

邱子丹

Opposite Voice

争议：公平、伦理与道德

有效性弱。此外，在实现样本随机选取这一步时，对人的生活与命运随机的进行改变是否是一种居高临下的把玩与亵渎？而 RCT 实验本身也有其局限性。它往

王宇轩

Despite the technical advantages of RCTs, there are critics who raise valid ethical, practical, and theoretical concerns. Issues of ethics, especially those related to conducting experiments on vulnerable groups, attract considerable attention. The ethical dilemmas linked to the possibility of withholding beneficial interventions from control groups require a systematic and careful approach to the design and implementation of randomised controlled trials (RCTs), to ensure that the well-being of participants is of utmost importance. These ethical considerations emphasise the need for thorough ethical review processes and the creation of strong ethical guidelines to regulate the implementation of randomised controlled trials (RCTs), particularly in environments characterised by significant power imbalances and economic inequalities (Ravallion, 2020).

罗宗昕

Opposite Voice

争议：公平、伦理与道德



1843
MAGAZINE

How poor Kenyans became economists' guinea pigs

Randomised controlled trials have many problems. They may still be the best tool for solving poverty

The ubiquity of RCTs in the field of development economics, as well as in the daily life of impoverished Kenyans in the region, is prompting a growing number of people to voice concerns about the ethics and the methodology behind them. In 2018, an RCT run by two development economists, in partnership with the World Bank and the water authority in Nairobi, Kenya's capital, tracked what happened when water supply was cut off to households in several slum settlements where bills hadn't been paid. Researchers wanted to test whether landlords, who are responsible for settling the accounts, would become more likely to pay as a result, and whether residents would protest.

Extension & Bonus Question

通过抽奖，一部分贫民随机搬到“好社区”

所有中奖的贫民都会搬家吗？ Intention-to-treat



RCT的迷思

RCT (randomized control trial)

Randomized Encouragement Design

Z_i : Encouragement (0 or 1)

D_i : Actual Treatment (0 or 1)

How many types of outcomes?

$D_i(Z_i = ?) = ?$

Extension & Bonus Question

RCT的迷思

3) 试验难以实现被试者的绝对服从，被试者可能不会严格按照试验要求做。

例如在观察两地不同政策产生的结果的时候，某地政府可能会想尽量展现好的一面，从而破坏试验的客观性。

王陈以勒

Secondly, to account for potential non-compliance (students in the treatment group not using the digital tools as intended or students in the control group gaining access to similar resources elsewhere), I'd employ an intention-to-treat (ITT) analysis. This approach ensures that the comparison between groups reflects the impact of being offered the intervention, rather than just the effect on those who used it, preserving the benefits of randomization.

林佳敏

Extension & Bonus Question

RCT的迷思

- Four **principal strata** (or **compliance types**):
 - compliers: $D_i(1) = 1$ and $D_i(0) = 0$
 - non-compliers
 - always-takers: $D_i(1) = D_i(0) = 1$
 - never-takers: $D_i(1) = D_i(0) = 0$
 - defiers: $D_i(1) = 0$ and $D_i(0) = 1$

- Correspondence between observed and principal strata:

| | $Z_i = 1$ | $Z_i = 0$ |
|-----------|-----------------------|----------------------|
| $D_i = 1$ | Complier/Always-taker | Defier/Always-taker |
| $D_i = 0$ | Defier/Never-taker | Complier/Never-taker |

- Without further assumptions, compliance types cannot be identified from observed strata



Extension & Bonus Question

RCT的迷思

Problem 3 In a randomized encouragement design, let $Z_i \in \{0, 1\}$ be the encouragement, $D_i \in \{0, 1\}$ be the actual treatment, and Y_i be the observed outcome, $i = 1, \dots, 8$. Suppose that there are no defiers, and that Z_i can only affect Y_i through D_i .

- (a) (10') Estimate the proportions of always-takers, never-takers and compliers.



| i | Z_i | D_i | Y_i |
|-----|-------|-------|-------|
| 1 | 0 | 0 | 2 |
| 2 | 0 | 0 | 4 |
| 3 | 0 | 0 | 4 |
| 4 | 0 | 1 | 6 |
| 5 | 1 | 0 | 4 |
| 6 | 1 | 1 | 5 |
| 7 | 1 | 1 | 7 |
| 8 | 1 | 1 | 8 |

Extension & Bonus Question

RCT的迷思

在社会学实验中，还存在着一个和传统医学实验截然不同的问题，通俗的讲是幸存者偏差，即**选择性偏差**。从数学角度上来讲，选择性偏差出现在以下的过程当中：

我们失去了中途离开研究那部分被试的数据，但丢失的数据并不总是值得我们困扰的问题，我们是否需要关注这种情况的发生，取决于发生的原因：

“是否与 $Y(S)$ 相关？”

如果不相关 (uncorrelated)，那就不需额外关心，研究有效数据即可。

如果相关 (correlated)，那么独立性 (independence) 就被破坏了!!!

这种情况下我们就需要公式进行校正：

张子涵

$$E[Y^s | S = s, \text{observe}] = E[Y^s | \text{observe}]$$

After-Class Survey

- 请大家扫码填写匿名问卷，谢谢大家！



这文章也可参考



财新周刊 | 投资农村婴幼儿

文 | 财新周刊 范俏佳 汤涵
钰 为何需要全面支持农村
儿童早期发展的中国方案？



财新

