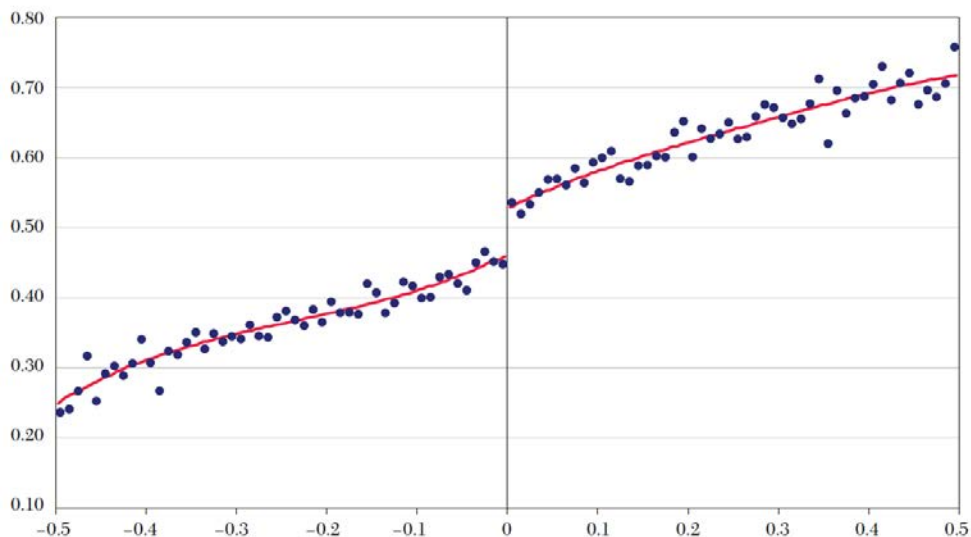


# 《因果推断与商业应用》期末考试 (总分 100 分)

November 22, 2022

## 一. 选择题 (2分\*5=10分)

1. 下面的因果推断事例哪些违反了一个好的分配机制应该具有的“非混杂性”? 【多选题】
  - A. 劳动培训是否增加就业机会: 某一城市组织的劳动培训规模过大, 影响了城市的劳动供给, 结果没参加培训的人受到负面影响, 导致因果关系不准确。
  - B. 疫苗是否有效: 幼儿园班级里面接受水痘疫苗的孩子很多, 导致有些家长认为没有必要打水痘疫苗了。
  - C. 劳动培训是否增加工资: 夕阳产业的就业者更加愿意进入培训机构, 而那些人的平均工资本来就不高。
  - D. 药物是否有效: 在药品的III期临床试验中, 有50个患者通过操作被确保分在治疗组。
  - E. 劳动培训是否增加就业: 有上进心的求职者(比如积极搜索相关信息)更容易进入培训机构, 而这些人本来就更更容易找到工作。
2. 我们想研究在位者效应 (incumbency effect) 是否存在。利用1946年-1998年的美国众议员选举数据, 我们有6558个数据点。图中的横轴是民主党的上次选举的得票净比例 (0表示得票率相同), 纵轴是民主党的本次选举的得票率。通过观察图中数据的特点来进行因果推断的方法叫做: 【单选题】
  - A. 双重差分法。
  - B. 断点回归法。
  - C. 倾向得分匹配法。
  - D. 合成控制法。



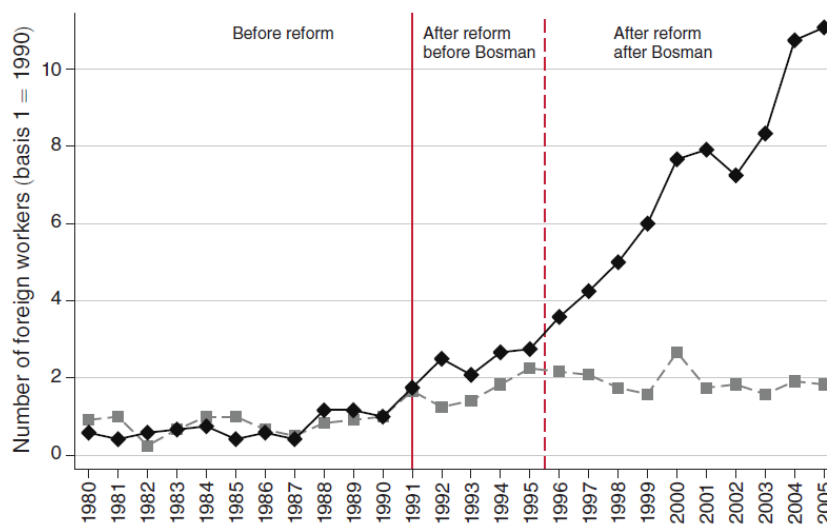
3. 1914年第一次世界大战爆发。1917年10月6号, 美国国会通过法律 (Trading with the Enemy Act) 允许美国企业对敌对国的战争相关专利可以自由使用。1918年美国国会进一步没收了敌对国的所有专利, 导致德国的专利可以让所有美国公司自由使用。为了研究这一法案对美国国内科技创新的影响, 有研究者以美国国内专利的增加

作为一个指标。如果化工部门的不同行业受德国专利影响的程度存在差异，利用不同行业在法案颁发前后的专利数量变化来进行政策分析，这种方法属于：【单选题】

- A. 双重差分法。
- B. 断点回归法。
- C. 倾向得分匹配法。
- D. 合成控制法。

4. 所得税的变化会对足球明星的转会产生巨大的影响，特别是1995年的博斯曼（Bosman）法案后欧盟球员在合同期满后可以自由转会。丹麦在1991年进行了税制改革，对高收入移民实施25%的平税，而不是68%的累进税，这对外籍球员的吸引力很大。如何评价这个政策对丹麦移民流入的影响？下图是以丹麦为实验组(黑色线)，以欧洲其他的一些国家（灰色线）作为控制组，横轴为时间，纵轴为移民数量（以1990年水平为基准）。根据图的特点，这种因果推断的方法属于：【单选题】

- A. 双重差分法。
- B. 断点回归法。
- C. 倾向得分匹配法。
- D. 合成控制法



5. 为了实施了一个培训项目，我们联系了800个人，并且收集了他们的个人相关信息。在招募过程中，只有其中200个人（实验组）报名参加了这个职业培训。通过观测这200人培训后的就业情况和工资，我们想知道这个培训项目的有效性。在进行控制组选取的时候，我们不是把所有剩下的600人作为控制组，而是先计算所有800人每人进入实验组的概率（根据每个人的年龄，以前收入，教育水平等因素来计算），然后每个实验组个体找到2个概率相近的控制组个体，最后进行比较。这种方法属于：【单选题】

- A. 双重差分法。
- B. 断点回归法。
- C. 倾向得分匹配法。
- D. 合成控制法

二. 简答题 (6分+6分+6分+7分+7分+8分=40分)

1. 我们想知道美国的私立中学和公立中学对学生数学成绩的影响（从而从学习成绩角度评价两种学校）。由于两校学生直接的成绩比较存在混杂因子（父母教育水平等），当地教育局实施了一次实验，进行随机抽签：当地所有 8 年级的学生参加抽奖，幸运者能够获得 3000 美元的私立学校学费补贴，得奖率是 10%。下图是 1 万名学生的抽奖结果：

		Public school	Private school
		$d_i=0$	$d_i=1$
Voucher loser	$z_i=0$	8000	1000
Voucher winner	$z_i=1$	800	200

如果抽到奖的学生平均成绩为 51.6 分，而没有抽到奖的学生平均成绩为 51.111 分，请问私立学校对学生成绩的作用提升有多大？请写下估计公式和计算结果（假设学习成绩为  $y$ ）。

2. 见下图。给定理论上的潜在结果和实际观测结果，请问利用“naïve” estimator 估计 ATE 的估计偏差是多少？

Unit	Potential outcome		Treatment effect	Treatment	Observed outcome
	If treated	If untreated			
$i$	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$	$D_i$	$Y_i$
1	4	2	2	1	4
2	6	3	3	1	6
3	3	1	2	1	3
4	3	1	2	0	1
5	8	4	4	0	4

3. 文章《Can emerging markets tilt global product design? Impacts of Chinese colorism on Hollywood castings》主要讲了中国电影市场的开放性政策对好莱坞电影中选角肤色的影响。假设 SHARELIGHT 是电影中浅肤色演员比例变量，DEXPOSURE 是指从电影特征中反映出该电影受中国市场欢迎的可能性，下面表格给出了简单 DID 的描述统计分析。请用表格的信息，估计电影市场开放带来的政策效应。【注：估计即可，不需要推断】

**Table 1.** Average SHARELIGHT Values

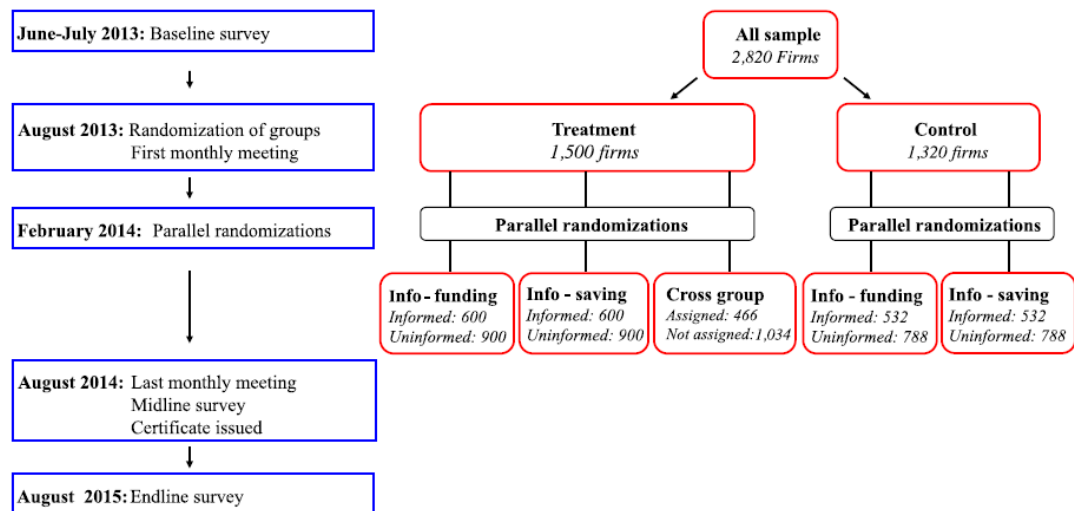
	DEXPOSURE = 0	DEXPOSURE = 1
PRE ( $t \leq 2013$ )	0.65 (0.32)	0.65 (0.31)
POST ( $t \geq 2014$ )	0.66 (0.32)	0.72 (0.31)

Note. Standard deviations are in parentheses.

- 请描述模糊断点回归 (Fuzzy RDD) 和工具变量法 (IV) 的关系【注，可以用因果图辅助】。
- 请简述倾向得分匹配 (PSM) 方法的基本步骤和优点。
- 给定有向无环图 (DAG) 中的后门理论，请证明 DAG 中的前门理论。

### 三. 应用题 (5分\*10=50分)

A 《Interfirm Relationships and Business Performance》文章中，作者用随机试验检验了企业主之间社交的作用。



A1 一个好的分配机制应该具有哪些特点？针对此项随机试验，这些特点如何满足？

A2 在南昌市进行的此项随机试验，为研究会面频率的影响，在常规小组会议外，设置了一次“跨组”小组会议（2014年2月，Cross group），即随机选择实验组中的466名企业主管，每10名主管组成一个“跨组”小组，总共43组，只在2014年2月见面一次。如果我们要比较常规小组和“跨组”小组对企业影响的比较，请问常规小组中的组员伙伴能进入同一个“跨组”小组吗？解释原因。

A3 本文的实证策略和结论有哪些潜在缺陷？

B 《Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program》文章中，作者研究了烟草政策的有效性。

B1 为什么要采用合成控制法？

B2 简要描述本文应用合成控制法的基本设计框架。

B3 如何在合成控制法下进行统计推断？

C 《Elite Recruitment and Political Stability: The Impact of the Abolition of China's Civil Service Exam》文章中，作者利用双重差分法对科举制度取消的政治后果进行了研究。

C1 请简要描述这篇文章双重差分方法的思路。

C2 写下基本回归方程（假设1905年的政策为 Post，配额为 Quotai，其他变量可以记为 controlit），指出我们感兴趣的参数。

C3 有观点认为，秀才配额多的地方革命次数多，是因为那个地方本来更加具有反叛精神。文章中是如何排除这个可能性的？

C4 取消科举后，地方上原有的政治关系网络也会影响后续的精英选拔，文章用工具变量法对这个潜在的遗漏变量问题进行了校正。其中一个工具变量是府的河流数目。请阐述河流数目可以作为工具变量的两个条件。