Basic Principles of Experimental Design
經濟學實驗設計原理

Joseph Tao-yi Wang
Lecture 3, EE-BGT

Define Economic Experiment 定義經濟學實驗

- An Economic Experiment
- Constructs a controlled environment to
- observe how people make economic decisions under real incentives, to answer
- questions raised by the researcher, testing a hypothesis or which theory matches reality
  - 建構一個控制的環境，在有真實誘因的情況下，觀察人們如何做決定（經濟決策），為要回答研究者所提出的問題，檢驗哪個假說或理論比較符合現實。

Define Economic Experiment 定義經濟學實驗

- 4 Components of a Controlled Environment

1. Real Incentives:
   - Choices have real consequences (真實後果或誘因)
2. A Good Control Group
   - (對照組的設計)
3. Random Assignment
   - (隨機分組)
4. No deception
   - (完全不欺騙受試者)

Principles of Experimental Design 實驗設計原理

1. Control, Measure, or Assume (控制,測量或假設)
2. Instructions (實驗說明)
3. Anonymity (匿名性)
4. Matching Protocols & Reputation Building (配對方式與受試者信譽)
5. Incentives (金錢誘因)
6. Order Effects (不同實驗的先後次序)
7. Controlling Risk Tastes (控制風險偏好)
8. Within-Subject and Between-Subject Design (同一 vs. 不同受試者)
9. Experimetrics (實驗計量)
10. No Deception (不欺騙受試者)

Reference: BGT, A1.2

Control, Measure or Assume 控制/測量/假設

- Control (控制)
  - Taking an action to affect a variable's value (主動賦予該變因的參數值)
  - "Induced" value theory (賦予參數理論)
- Measurement (測量)
  - Measure the value of a variable (測量該變因的參數值)
  - Via various methods (see below) (用下頁所列不同方法)
- Assumption (假設)
  - Pseudo-control (直接假設該變因的參數值)
  - Accept a maintain hypothesis about the value of a variable

Control, Measure or Assume 控制/測量/假設

- Methods of Measurement (測量方法):
- Psychometric measures (surveys) (心理測驗/問卷)
- Risk-aversion measures (certainty equiv)(風險偏好)
- Probability judgments (scoring rules) (主觀機率判斷)
- Information acquisition (mouse/eye-tracking)
  - 資訊取得：滑鼠追蹤或眼球追蹤
- Psychophsiological measures (測量心理生理學上的反應)
  - fMRI (功能性磁振造影), GSR (皮膚電阻反應), PDR (瞳孔放大反應), EEG (頭皮腦電波), etc.

Instructions 實驗說明

- Tell subjects what they need to know (目的：告知所需資訊)
- Public Knowledge (公共知識):
  - Established by reading instructions out loud (來自公開朗讀說明)
- How much to reveal? (要告訴受試者多少？)
  - Entire payoff structure (default) (原則上要完整告知報酬如何決定)
  - Since we're not sure what subjects would think about what they are not told (如果不揭露所有資訊，我們不知道他們會怎麼想)
- Withhold some information: Study how people/markets learn under limited information
  - (不告知某些資訊：用以研究人們或市場如何在有限資訊下學習探索)

Anonymity 匿名性

- Who's Who? Subject behavior can change knowing opponent's identity due to
  - 見面三分情？受試者的行為會因為知道對方是誰而改變，因為
  - Appearance, gender, (長相外貌、性別)
  - Fear of retaliation, etc. (害怕被報復等等)
- Use the anonymity case as a benchmark
  - (因此，我們通常用匿名實驗來當作基準實驗)
  - Measure opponent characteristics (appearance) and compare to benchmark
    - 要研究這些變因時，可以測量對方特質(例如長相多好看)，同時把它們顯示給受試者看，然後把實驗結果和基準實驗作比較

Matching Protocol/Reputation 配對方式/信譽

- Random matching (random switch)
  - Empirically kills repeated game effects
    - 隨機配對：每回合重新洗牌，實證上可除去重複賽局效果
- Mean-matching (play with everyone)
    - 與所有人配對：每個決定都跟其他每個人各配對一次
- Other more strict matching protocols:(更嚴格配對方式)
- Non-repeat matching (meet only once)
    - 不重複配對：整個實驗中只跟每個對手配對一次，所需實驗參與者人數的魔術數字是2的n次方，因為能做2^n-1回合不重複配對
- Non-contagion matching (no chain-of-influence)

不污染配對：整場沒有「我跟你配、你再跟他配，最後我遇到他」的情形

Incentives 真實誘因

- Hypothetical vs. Real Money Decisions
  - This distinguishes economic & psychological experiments
    - 假設性問題 vs.「玩真的」：經濟學實驗和心理學實驗一大分野
- Assumption behind money payments:
  - "Everybody likes having more money and nobody gets tired of having more of it." 使用金錢報酬只需要一個行為假設：「每個人都喜歡獲得更多金錢，而且多多益善(沒有人會覺得拿太多)」
- Cost of deviation without real money is 0
    - 只是假設性問題的話，亂講亂選沒有成本
- Paying money reduces variation & outliers
    - 故金錢報酬會降低亂選的噪音，減少極端的特例

Incentives 真實誘因

- Pay Less vs. Pay More (付得少/多是否影響結果？)
  - Comparison studies not done often enough
  - Expensive to double/triple the payments
    - 目前有一些做這種比較的實驗，但還不夠多
    - 因為要把報酬乘上兩倍、三倍很貴
- Some experiments done in poor countries
    - (不過仍有些實驗會特別到貧窮的國家去做)
  - Vietnam (越南：試字率高、人民卻很窮、鮮少遷徙)
  - Few results that disconfirm theory have been overturned by paying more money
    - 這些少數重複實驗結果顯示：如果原本的實驗結果不支持理論預測，多付錢也大概不會讓實驗結果變得合乎理論的預測

Incentives 真實誘因

- Flat Maximum Critique (「作什麼都沒差」)
  - Is it worthwhile (high stakes) to think hard?
    - 想得深入一點到底值不值得?
    - (重賞之下必有勇夫? 沒有勇夫表示賞金不夠?)
  - EX: Costless to deviate from (1/3, 1/3, 1/3) in rock-paper-scissors (例如: 「剪刀石頭布」不按照均衡策略(1/3, 1/3, 1/3)來做沒有損失，因為不管出什麼報酬都相同)
- No ideal solution yet... (尚未有很好的解決辦法，不過...)
  - Design steep marginal incentives
    - 可以設計得讓「邊際誘因」很高
  - Modest effect on high stakes anyway
    - 高獎金對結果的影響雖不是0，但也沒那麼大

Order Effects 不同實驗的先後次序

- AB: A came first; B came second
  - Is this why we see different behavior?
  - AB即「先做實驗A，再做實驗B」
  - 行為的改變是否只是因為先後次序不同？
- Try BA and include order dummies in analysis
  - 必須也做BA(次序反過來)，然後在資料分析的時候加上次序的虛擬變數(dummies)來檢驗是否有影響
- What if ABC?
- ACB/BAC/BCA/CBA/CAB or simplify design
  - 有ABC怎麼辦？
  - 做ACB/BAC/BCA/CBA/CAB或簡化實驗設計

Control Risk Preferences 控制風險偏好

- Binary Lottery Procedure: (發樂透彩券當報酬)
  - Widely used to control risk preferences, but not much evidence that it works (控制風險偏好的常用辦法，但沒有太多證據顯示有效，且需假設人們能把複合樂透簡化為簡單樂透)
- Alternatives: (其他辦法)
- Assume risk neutrality (假設受試者是風險中立)
- Measure risk preferences (測量風險偏好)
  - Holt and Laury (2002) or Tanaka et al. (2010)
  - Choi et al. (2007); Andreoni and Sprenger (2012)
  - DOSE: Wang et al. (AER-R&R)

Within-subject vs. Between-subject 同一|不同

- Within-subject Design (「比較同一受試者」的設計)
  - Same subject observed in various treatments
  - Pro: More statistically powerful
  - Con: Possible demand effect
  - 觀察同一受試者在不同實驗環境中的反應。優點：統計上解釋力強(可做paired t-test)；缺點：可能有「要求效果」
- Between-subject Design (「比較不同受試者」的設計)
  - Different subjects observed in each treatment
  - Norm in experimental economics
  - Con: "Impossible" for fMRI or eyetracking
  - 觀察不同受試者在各自的實驗環境中的反應。實驗經濟學界一般的「標準做法」；缺點：幾乎不能做fMRI或眼球追蹤的實驗

Experimetrics 實驗計量

- Econometrics customized for experiments
- Just like Econometrics is
  - Statistics customized for economics
    - 特別為實驗經濟學開發的計量方法，正如計量經濟學是特別為經濟學開發的統計方法
- Bottom line: Use all econometrics feasible to get the most out of your (experimental) data
    - (良心建議：請充分利用所有可能的計量工具來分析實驗資料)
- Experimental Design and Experimetrics are sometimes substitutes
  - But complement each other in a good paper!
    - 實驗設計和實驗計量有時可以互相替代，因為有好的設計，簡單的敘述統計可能就夠了。但兩者相輔相成能產生最好的論文

List of Experimetrics 實驗計量方法列表

- Mann-Whitney-Wilcoxon Test(s) vs. T-test
  - Non-parametric test similar to (un-)paired t-test
- Regression (with random effects)
- Maximum Likelihood Estimations (最大概似估計)
  - Level-k, Cognitive Hierarchy models
  - Learning (學習理論): EWA, Reinforcement
  - Quantal Response Equilibrium (手滑反應均衡)
  - Simulate (模擬) and Estimate (估計)
- Out-of-sample Predictions (預測新的資料)
- Markov-switching (Eyetracking), SPM (fMRI)

No Deception 不欺騙受試者

- Experimental Economists do not deceive their subjects
  - (實驗經濟學家從不欺騙他們的受試者)
- This creates credibility (「徙木立信」)
  - Makes monetary payments "real" (因而相信真的有金錢報酬)
- And avoids anticipation/strategic responses
  - Differs from psychologists
  - 避免預期會被騙時的鬥智反應
  - 這和心理學家不同(他們只要事後有說明即可)
- Can achieve most goals with better design (except)
  - How can we study the effect of deception?
  - 所謂「需要欺騙的實驗」大部分都能設計另一個「不需欺騙的實驗」來達到同樣目的 (除了研究人們被實驗者欺騙時的反應)

Conclusion: The Gold Standards 設計十誡

1. Control, Measure, or Assume (控制,測量或假設)
2. Instructions (實驗說明)
3. Anonymity (匿名性)
4. Matching Protocols & Reputation Building (配對方式與受試者信譽)
5. Incentives (金錢誘因)
6. Order Effects (不同實驗的先後次序)
7. Controlling Risk Tastes (控制風險偏好)
8. Within-Subject and Between-Subject Design (同一 vs. 不同受試者)
9. Experimetrics (實驗計量)
10. No Deception (不欺騙受試者)

Reference: BGT, A1.2