

社會偏好

Joseph Tao-yi Wang

3/20/2009

公平、互惠、信任

- 例子：最後通牒談判實驗
 - 下通牒的提議者：提出一個不接受就拉倒(take-it-or-leave-it)的方案
 - 接受通牒的回應者：選擇接受或不接受該方案
- 為什麼要在乎這個實驗的結果？(Why care?)
- 一般人常把公平和信任掛在嘴邊
- 但Stigler (1981)說：自利動機最後總是佔上風
- 實驗結果 = 把「報復性正義」標上價格

美國政治史上的例子

- 1787年在費城舉行制憲會議
- 「新加入的州，是否只能比我們(原本的十三州)次等？」
 - 喬治·梅森(George Mason)：「新的州民和我們一樣會為自己的州感到驕傲和自豪。因此，如果他們不能和我們享受同樣的權利，他們若非不願加入，就是加入後很快就會退出...」
- 是「害怕對方不接受」還是「追求公平」？
 - 這如何應用在台灣和中國的關係上？

人到底是不是「自利」的？

- 人是自利的(經原第一堂課)
- 哪「利他」怎麼解釋呢？
- 標準答案：你的效用函數包括你「朋友」的金錢報酬，因此利他就是自利(效用極大化)
- 為什麼接下來就不會在討論「利他」了？
 - 利他效用函數會讓福利經濟學第一定理不成立
- 但是，人們真的只在乎自己金錢的報酬嗎？

跟社會偏好有關的賽局

- 囚犯的兩難 (Prisoners' Dilemma, PD)
- 自願捐輸賽局(Public Goods Game, PG)
- 最後通牒談判(Ultimatum Game)
- 獨裁談判(Dictator Game)：回應者不能拒絕方案
- 互信賽局(Trust Game)：在獨裁談判之前加入一期，回應者在該期「投資」提議者決定餅的大小
 - 回應者先投資的數量代表「信任對方的程度」
 - 提議者回饋的數量代表「可被信任的程度」
- 蜈蚣賽局(Centipede)：多回合的互信賽局
- 禮尚往來賽局(Gift Exchange)：多回合互信賽局

囚犯的兩難(Prisoners' Dilemma)

	C	D
C	H, H	S, T
D	T, S	L, L

- 每個參與者選擇C或D
- 兩難的抉擇：
 - 雙方都合作(都選C)的話，對大家最好(Pareto)
 - 但是給定對方合作(C)，你背叛(D)比合作(C)更好
- 唯一的奈許均衡：(D, D)

($T > H > L > S$)

囚犯的兩難(Prisoners' Dilemma)

- 只做一次基準實驗(1-shot games Baseline) :
 - 選擇合作(C)的比例高達50%
- 改變報酬的影響 :
 - 降低背叛的好處或被背叛的損失(降低T或提高S)會促進合作
- 事先溝通(pre-play communication)會促進合作
- 隨機配對重複做幾次 :
 - 越來越多人背叛, 最後只剩下少數人還堅持合作



自願捐輸賽局(Public Goods Game)

- N 個參與者
- 每個人決定要從財產 e_i 當中捐出 c_i 來做公益
- 總捐款 $c_{all} =$ 所有 c_i 的加總
- 每人報酬 $= e_i - c_i + m * c_{all} / N$
 - 總捐款會被乘上 m 倍再平均分給每個人
- 很像是「囚犯的兩難」
 - 合作對大家都好; 但是各自有誘因要做沒公德心的事(搭便車)



自願捐輸賽局(Public Goods Game)

- 只做一次基準實驗(1-shot games Baseline) :
 - 平均捐出財產的50% (多半是全捐或全不捐)
- 改變報酬的影響 :
 - 改變 m (捐輸的邊際效果) 會促進捐輸(合作)
- 事先溝通(pre-play communication)會促進合作
- 隨機配對重複做幾次: 捐輸愈來愈少
- 懲罰效果: Fehr and Grachter (AER 2000)
 - 許多人都願意花錢懲罰捐得少的人, 雖然懲罰本身也是一種「自願捐輸」
- 合作似乎是互相的(Reciprocal)



Fehr and Grachter (AER 2000)

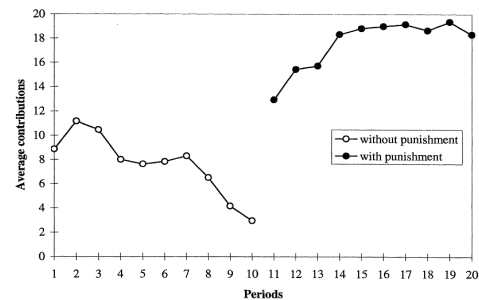


FIGURE 3B. AVERAGE CONTRIBUTIONS OVER TIME IN THE PARTNER-TREATMENT (SESSION 5)

利他偏好(Pure/Impure Altruism)

- 舉例來說: $U_i(X) = x_i + \alpha \cdot x_{-i}$
- 可以解釋PD實驗中(C,C)的結果
- 作業: 你可以用這種效用函數來解釋PG的結果(不論有或沒有懲罰)嗎?
- 別人的捐輸可以代替(crowd out)利他的捐輸
- 不能解釋 互利互惠(reciprocity)
 - 「那些對我好的人, 我也希望對他好; 那些對我不仁的人, 我也希望對他不義。」



不公厭惡: 不好意思—嫉妒 Inequality-Aversion: Guilty-Envy

- Fehr and Schmidt (1999)
- $$U_i(X) = x_i - \frac{\alpha}{n-1} \sum_{k \neq i} \max(x_i - x_k, 0) - \frac{\beta}{n-1} \sum_{k \neq i} \max(x_k - x_i, 0)$$
- 嫉妒比「不好意思」強烈 $0 \leq \beta \leq 1, \beta \leq \alpha$
 - 可用來解釋:
 - PD: 雙方都合作的結果(C,C)可以維持
 - PG: 每人捐款數目不同
 - 有些人捐0 vs. 有些人捐一部份
 - 懲罰PG中的「害群之馬」(不捐的人)



課堂實驗：自願捐輸賽局

	回合數1			回合數2		
集合總張數	13			5		
集合分配得分	2.6			1		
留下張數	0	1	2	0	1	2
累積組數	3	5	5	0	1	12
本回合分數	2.6	3.6	4.6	5.6	6.6	7.6
累積總數	3	5	5	3	6	4

PD和PG的結論

- 台灣這方面的專家：
 - PD: 楊春雷老師(中研院社科中心)
 - PG: 徐麗振老師(政大財政系)
- 這些結果是不是否定了賽局論？
 - 不完全。這些結果邀請經濟學家寫出新理論
- 行為賽局論正在發展的新理論：社會偏好(BGT, 2.8) 與有限層次策略性思考(BGT, Ch.5)
- PD/PG等實驗的問題：背叛是穩賺不賠的優勢策略
 - 無法區分「利他動機」與「(互利互惠下)有條件合作」

最後通牒談判(Ultimatum Game)

- 一個更恰當的賽局：最後通牒談判
 - 提議者：提出「不接受就拉倒(take-it-or-leave-it)」的方案
 - 回應者：接受或不接受該方案
- 基準實驗：只做一次、匿名、只詢問「實際反應」
 - 變型一：重複幾次，每回合隨機重新配對
 - 變型二：詢問完整策略(strategy method)——問回應者願意接受的最差方案(Minimum Acceptable Offer, MAO)
- 詢問完整策略vs.只詢問「實際反應」
 - 詢問完整策略是否太不自然？

最後通牒談判(Ultimatum Game)

- 主要結果(詳見BGT, Table 2.2, 2.3)
- 提議者
 - 中位數/眾數：40~50%
 - 平均：30-40%
 - 幾乎沒有人低於10%或高於50%
 - 為了公平還是害怕對方不接受？
- 回應者
 - 很少拒絕40~50%的方案
 - 50%拒絕低於20%的方案

不公厭惡：不好意思—嫉妒 Inequality-Aversion: Guilty-Envy

- Fehr and Schmidt (1999)

$$U_i(X) = x_i - \frac{\alpha}{n-1} \sum_{k \neq i} \max(x_k - x_i, 0) - \frac{\beta}{n-1} \sum_{k \neq i} \max(x_i - x_k, 0)$$

- 嫉妒比「不好意思」強烈 $0 \leq \beta \leq 1, \beta \leq \alpha$
- 可用來解釋：

- 最後通牒談判：拒絕不公方案、提議公平方案
- 利他偏好(altruism)也可以這些行為嗎？
 - 不行！故利他偏好適用範圍較窄，不若不公厭惡(G-E)來得一般

ERC (Envy, Reciprocity, Competition)

不公厭惡II：嫉妒互惠競爭

- Bolton and Ockenfels (2000)

$$U_i(X) = U\left(x_i, \frac{x_i}{\sum_{k=1}^n x_k}\right)$$

- 在乎自己報酬的相對比例
- 不考慮個人際比較，只考慮與整體的比較
- 作業：這個模型能否解釋PD中的合作與最後通牒談判中拒絕不公方案、提議公平方案？
- 哪個賽局可以分辨本模型和G-E孰優孰劣？

獨裁分配賽局(Dictator Game)

- 一個不能被拒絕的最後通牒
 - 提議者：提出一個(自動通過的)分配方案(且回應者不得拒絕)
 - 可以區分公平和害怕方案被拒絕
- 實驗結果：(BGT, Table 2.4)
 - 給對方的比例比最後通牒談判低，但仍不是0
- 分配比最適對策(BR)要更優厚，顯示：
 - 利他與利害計算兩者並存
 - 提議者的認知過於「悲觀」

獨裁分配賽局：G-E模型

- G-E模型對獨裁分配賽局的預測：
 - 不是50-50對分，就是全部都不給
- 不符合獨裁分配賽局的實驗結果
 - 作業：只有極端角解的原因是效用函數為線性。把效用函數改成凸性(concave)可否解釋結果？
- 有哪些理論能夠解釋此一實驗結果？
 - ERC: 更複雜的「利他」效用
 - 害怕對方不接受 + 自動動機
- 作業：請用ERC或「害怕對方不接受」來解釋獨裁分配賽局的實驗結果

有沒有可能是其他因素造成的？

- X: 實驗方法論上的因素
 - 重複試驗、金額大小、匿名性、實驗者與回應者「雙盲」
- Y: 個人特質
 - 性別、種族、科系與年紀
 - 大腦、生理與外貌
- Z: 文化
- XX: 描述實驗的方式
 - 名稱(Labeling)與情境(Context)
- YY: 改變實驗內容
 - 加上其他決定或資訊(見下面)

X: 實驗方法論上的因素

- 重複試驗
 - 經驗對實驗結果影響不大(方案與拒絕情形只會稍微移往均衡)
 - 除非人群中存在一些「自利」電腦，只在乎自身得分
 - 有一點影響是因為情緒疲乏嗎？可以隔幾天再做看看
- 金額大小
 - 金額大增(幾個月薪水)對拒絕有些影響，對方案沒有影響
 - 有贊助的捐獻：1/2都不捐(自利)，1/3你捐我才捐(有條件合作者Leontief)，1/6捐(利他主義者utilitarian)
- 匿名性與實驗者/回應者「雙盲」設計
 - 除去要求效果：獨裁分配平均降為10%(一半給0)
 - 最後通牒談判結果幾乎沒變(Bolton and Zwick, GEB95)

Y: 個人特質

- 性別 - 沒有統一的「性別效果」
 - 女性較少拒絕方案
 - 「第三者可懲罰」獨裁分配賽局中對懲罰成本的計算較精明
- 種族 - 結果不多(因為政治正確的關係嗎?)
 - Ex: 信託賽局中白人男性並未回報亞裔的信任!(社會地位?)
- 學校科系 - 「經濟系效果」不同研究有不同的結果
- 年紀 - 自利 (<5) → 完全的對分(5-7) → 公平偏好(>7)
 - EX: 幼稚園學童接受1 penny的比例高達70% (vs. 30-60%)
- 大腦，生理與外貌
 - 勇猛：睾固酮濃度高者較常拒絕，但提出的方案也較寬大
 - 帥哥：面對帥哥，許多女性提議的分配願意給對方>50%

Z: 文化

- 實驗方法論上的問題(與解決辦法):
- 金額多寡：用等量的購買力；N天的薪水
- 語言：「翻譯回英文(Back translation)」
- 實驗者效果：雙語實驗者(同一個人)
 - 每個人都去每個文化下作一次實驗
- 文化與其他因素混雜(Confound)：控制兩邊受試群體的個人特質(demographics)相仿並測量那些「無法控制的因素」

Z: 文化

- 秘魯的馬欽瓜那(Machiguenga)農夫
 - 提出的方案平均為26%，眾數為15%
 - 離群索居(social disconnect)，非親戚就沒名字
- Henrich et al. (2002): 研究20個文化群體
- 超平等(>50%)的方案(且對方會被拒絕！)
 - 巴拉圭的Ache獵人頭族和印尼Lamelara捕鯨族
 - 送禮的競爭(Competitive gift-giving)：接受超平等的方案表示你未來有義務回報，是一種侮辱

Z: 文化

- 兩個重要決定因素 ($R^2 = 0.68$):
 - 生活中合作活動的多寡(生產的規模經濟)
 - 市場整合的程度(degree of market integration)
- 合作活動愈多、市場整合愈高，愈有50-50對分的社會規範(norm)
 - 市場經濟和自私自利不見得相關！
- 這才是真正的「文化研究」！

XX: 描述實驗的方式

- 名稱(Labeling)與情境(Context)
 - 在市場上可以接受有人「自利」：把實驗描述成買賣雙方交易會讓最後通牒談判的方案降低10%，但不影響拒絕率
 - 「共有資源」產生共有意識：把實驗描述成討論如何分配共有資源會讓雙方都更「寬大」
 - 「提點(priming)」：實驗說明提醒「如果你是對方會如何做？」讓人更害怕對方不接受
- 關於Framing的普遍理論？有待來者！

YY: 改變實驗內容 加上其他決定或資訊(見下面)

- 身分、溝通、應得與否(Entitlement)
 - 知道對方是誰(學生、慈善機構)會讓獨裁分配者給更多
 - 贏得提議者的角色會讓所提方案降低10%
- 競爭壓力與外在選項(outside option)
 - 若「需要獲得X元才能進行下一回合」會提較不公之方案
 - $U(\text{拒絕}) = (2,3)$: 不同的「對分點」；造成更多拒絕
- 不曉得總分配金額
 - 較不願意拒絕，因為「可能」是公平的
- 多位回應者競逐單一提議者會讓方案趨近0，反之亦真
- 意圖：如果只能提議(8,2)或(10,0)，會接受(8,2)

社會偏好 (G-E, ERC)能否改變？

- 作業：請用G-E模型解釋為何最後通牒談判中，多人競相回應(提案)會讓方案趨近0%(100%)。
- 作業：Guth and Van Damme (1998)作以下實驗：
 - 提議者對兩位回應者提出方案(x, y, z)，但只有一位能回應
 - 有決定權的回應者只看到y或z(或都看到)，決定是否接受
- 實驗結果：有人會忽略沒有決定權的另一位回應者
 - 只看到y或都看到：方案y=30-40%，z=5-10%，拒絕率~5%
 - 只看到z：方案z=12-15%，其餘大都給x，拒絕率~5%
- 請用ERC解釋以上結果

ERC vs. Guilt-Envy

- 其他實驗：G-E解釋結果較ERC為佳
 - Charness and Rabin (2000)研究自願捐輸實驗發現人們還是在乎他人之間的公平與否
 - 除了相對差異，絕對差異還是有影響(見下面)
- 不過，兩個模型都假設效用可區分(separability)
 - 人的效用是否能夠不跟「怎麼走到這一步」有關？(關於「意圖」的實驗結果就不支持，但數學上比較好解)
- 而且兩個模型都不能解釋交易互惠(reciprocity)
 - 「對我好的人，我也要對他好。你對我不仁，我就對你不義。」

公平均衡(Fairness Equilibrium)

- 心理賽局(psychological games)
- Rabin (1993)討論同步賽局：我的決策是 a_1
- 我認知對方的決策是 b_2
- 我認知對方認知我的決策是 c_1
- 我對你的善意是：

$$f_1(a_1, b_2) = \frac{\pi_2(b_2, a_1) - \pi_2^{fair}(b_2)}{\pi_2^{max}(b_2) - \pi_2^{min}(b_2)}$$

公平均衡(Fairness Equilibrium)

- 我對你的善意是：

$$f_1(a_1, b_2) = \frac{\pi_2(b_2, a_1) - \pi_2^{fair}(b_2)}{\pi_2^{max}(b_2) - \pi_2^{min}(b_2)}$$

- 我認知你對我的善意是：

$$\tilde{f}_2(b_2, c_1) = \frac{\pi_1(c_1, b_2) - \pi_1^{fair}(c_1)}{\pi_1^{max}(c_1) - \pi_1^{min}(c_1)}$$

公平均衡(Fairness Equilibrium)

- 我的(社會)偏好：

$$U_1(a_1, b_2, c_1) = \pi_1(a_1, b_2) + \alpha \cdot \tilde{f}_2(b_2, c_1) + \alpha \cdot \tilde{f}_2(b_2, c_1) \cdot f_1(a_1, b_2)$$

- 理性預期： $a_1 = b_2 = c_1$
- 例 1: 囚犯的兩難
- 例 2: 對撞賽局(Chicken Game)
- 延伸到非同步賽局Falk and Fischbacher (1998)

ERC, Guilty-Envy vs. Fairness Eq.

Offer	Accept	Reject	Reject %	ERC	G-E	Fairness Eq.
Equal	5,5	0.5, 0.5				
Unequal	8,2	0.8, 0.2	0.38	None	Some	Some
Equal	5,5	3,3				
Unequal	8,2	6,0	0.19	None	None	Some

眾多社會偏好研究給我們的啟示

- 已經有太多實驗結果了...
 - 有沒有一個大一統的理論來解釋所有結果呢？
- 至少知道用盡辦法也無法否定社會偏好...
 - 人們並非完全「自利」的
- 實驗方法：可觀摩他們多小心執行實驗
- 哪一點讓你覺得這個實驗很有趣？
 - 你如何將之融入你自己的實驗設計？

互信(Trust)

- 我如何知道你會履行你的承諾？
 - 法律契約，第三方保證
 - 家族壓力，暴力威脅
- 這些成本都很高，但是互信的成本較低
 - 淳樸鄉下夜不閉戶
 - 東京機場的失物招領 (72%物歸原主)
 - 公司寧可裁員也不願減薪

互信(Trust)

- 互信(一種社會資本)能夠解釋經濟成長
- Putnam (1995)
 - "Since trust is so central to a theory of social capital, it would be desirable to have strong behavioral indicators of trends in social trust or misanthropy. I have discovered no such behavioral measures."
 - 既然互信對社會資本理論這麼重要，如果有一個行為指標該有多好？可惜我還沒有看過。

互信投資(Trust Game)

- 委託人決定要投資受託人多少錢
 - 投資金額會乘上 m 倍
- 受託人決定要回報委託人多少錢
- 如果是你，你會投資多少？回報多少？
- 這個賽局可供測量社會資本：
 - 信任：投資金額
 - 可信賴程度：回報金額

互信投資(Trust Game)

- Berg, Dickhaut and McCabe (1995)
- 雙盲實驗，初始金額10元美金，乘數 $m=2$
- 實驗結果：委託人平均投資初始金額的50%
- 受託人平均回報投資金額的95%
- 重複實驗：(許多文章)
 - 投資比例40-70%
 - 回報比例110-150% (最低是腐敗的肯亞：55%)
- 例外：Van Huyck, Battalio Walters (95/01)
 - 農民(幾乎不投資) vs. 獨裁的地主(幾乎全拿)

互信投資：人們為何彼此信任？

- 「可信賴」是因為互惠還是單純的利他？
- 利他：可用獨裁分配賽局的方案來測量
- 互惠：「互信投資中的回報(佔總獲利金額)比例」減去「獨裁分配中給對方的比例」
- Dufwenberg and Gneezy (2000):
 - 30% vs. 33% (統計上不顯著) - 只是單純利他？
- Cox (1999):
 - 差額只有10% (統計上顯著，但經濟上很小)

互信投資：直接與間接的互信

- 改變回報時的配對方式：
 - 兩人固定配對 → 四人交叉回報 → 全體混合回報
- Buchan, Croson and Dawes (2000)
 - 信任：64% → 48% → 39%
 - 可信賴程度：35% → 19% → 20%
- Dufwenberg et al. (2000)
 - 信任：60% → 53%
 - 可信賴程度：28% → 37%

互信投資：其他延伸實驗

- 過去紀錄對連續互信投資的影響：
 - 選擇是否投資250 (需花成本150)
 - 看得到對方過去六回合的投資紀錄
- Seinen and Schram (1999)
 - 25% → 70% (顯示過去紀錄)

互信投資：其他延伸實驗

- 多回合互信投資
 - 就像是蜈蚣賽局：但是最後節點是(0,0)
 - 所以自利者無法模仿好人每回合都回報(pass)
- Ho and Weigelt (Management Sci. 2005)
 - 四回合；每回報一次金額加倍；詢問完整策略
 - 30% (50%) 玩家1(2)第一回合就留下95%不投資
- Rapoport et al. (GEB 2003):
 - 三個人；九回合；最後金額可高達1,500元美金
 - 1/3的實驗在前兩回合就結束了

職場上的互信：禮尚往來？

- Fehr et al.; cf. Fehr and Gächter (JEP 2000)
- 受試者扮演八個員工，六家公司
- 公司提供薪資 w 給員工(建議努力程度 e')
- 如果接受，員工選擇努力程度 e
- 報酬：公司的金錢報酬是 $(q-w)e$
 - 員工的金錢報酬是 $w-c(e)$, $c(\cdot)$ convex on 0.1~1
- 如果你是公司(或員工)，你會如何選擇？

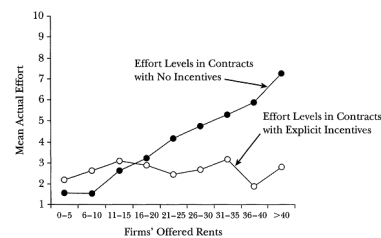
職場上的互信：禮尚往來？

- 數學賽局論的預測：
 - 員工會選擇最低的努力程度
 - 公司會給予最低的薪資
- 禮尚往來 Gift Exchange (Akerlof, 1982)
 - 面對高薪，員工會死心塌地為公司賣命
- 效率性工資 Efficient Wages
 - 如果員工偷懶被抓到就拿不到高薪產生的工作租 (job rent) $w-c(e')$
- 在實務上很難區分以上兩個理論...

職場上的互信：禮尚往來？

- Fehr and Gächter (JEP 2000) 支持禮尚往來

Figure 3
Actual Effort-Rent Relation in the Absence and Presence of Explicit Performance Incentives



Source: Fehr and Gächter (forthcoming).

職場上的互信：禮尚往來？

- 其他文章發現的禮尚往來程度比較少
- Gneezy and List (2006): 在實務上剛開始有禮尚往來，但是幾個小時後效果就消失了
 - "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments," *Econometrica*, 74(5), 1365-1384.
 - "We experimented with the individual-specific variables found to be important in Landry [et al.] (2006) and found that their inclusion does not change the qualitative insights."

互信投資：巨人的肩膀在何處？

- 禮尚往來在勞動市場到底多常見(robust)？
- 這仍是(實驗與實證上)正在研究的問題
- 問題：互信從何而來？
- Kosfeld, Heinrichs, Zak, Fischbacher and Fehr (2005), "Oxytocin increases Trust in Humans," *Nature* 435, 2 June 2005, 673-676.

結論

- 人們是否對誘因作出反應？
 - 是！但是不只是對金錢誘因，還有其他誘因！
- 外在的(貨幣)誘因：金錢報酬
- 內在的誘因：公平、利他等等
- 已經有太多關於社會偏好的實驗了！
 - 請不要矇著眼睛就提議作另一個社會偏好的實驗！
 - 要先看文獻！(BGT, ch.2, MGS, ch.12-14)
- 有沒有一個大一統的理論可以來解釋這麼多實驗的結果(並且作出新的預測)？

Public Goods Game Results

	回合數1			回合數2		
集合總張數	13			5		
集合分配得分	2.6			1		
留下張數	0	1	2	0	1	2
累積組數	3	5	5	0	1	12
本回合分數	2.6	3.6	4.6	5.6	6.6	7.6
累積總數	3	5	5	3	6	4