Social Preferences

Mark Dean

Behavioral Economics Spring 2017

- So now we need a model that allows us to capture the fact that people have 'other regarding preferences'
  - Ultimatum game experiments offer a 'smoking gun'
  - But intuition (and other evidence) tells us it goes much further than that
- Starting point: What psychological processes do we think are important here
- We will focus on two
  - Inequality aversion
  - Fairness
- Notice: Altruism is also interesting but
  - Easier to fit into standard model
  - Can't explain ultimatum game results

Inequality Aversion

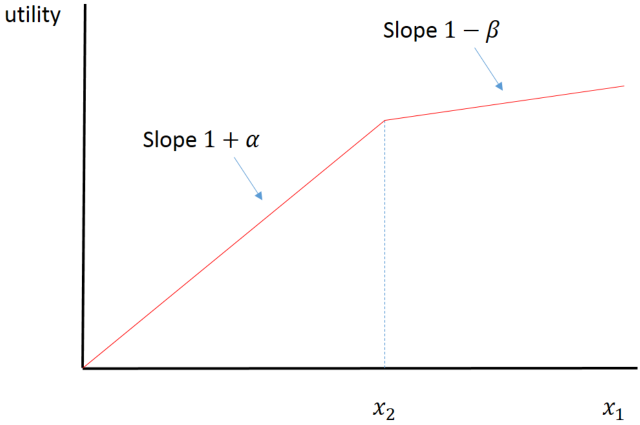
- One of the earliest and most influential models of other regarding preferences is that of **inequality aversion**
  - Fehr and Schmidt [1999]
  - Bolton and Ochenfels [2000]
- Basic idea is, well, people don't like inequality (!)
- Comes in two forms
  - Dislike of having more than other people
  - Dislike having less than other people

- Consider a game between two players
- What utility does player 1 get if they end up with x_1 and player 2 ends up with x_2?

$$u_1(x_1, x_2) = x_1 - \alpha \max\{x_2 - x_1, 0\} - \beta \max\{x_1 - x_2, 0\}$$

- Three parts
  1. Standard utility
  2. Dislike of having less that the other player
     - $\max\{x_2 - x_1, 0\} = 0$ if player 1 has more
  3. Dislike of having more than the other player
     - $\max\{x_1 - x_2, 0\} = 0$ if player 2 has more

Inequality Aversion



- Utility function has a kink at x_2
- β assumed to be less than 1

Inequality Aversion and the Ultimatum Game

- What does the inequality aversion model say about play in the ultimatum game?
- Assume pie is of size \$10
- What will player 2 do if player 1 offers to keep x and give player 2 $(10 - x)$
- Remember that player 2's utility is

$$u_2(x_1, x_2) = x_2 - \alpha \max\{x_1 - x_2, 0\} - \beta \max\{x_2 - x_1, 0\}$$

Player 2

- The choice is to reject the offer, in which case

$$x_1 = x_2 = 0$$

- or accept the offer, in which case

$$x_1 = x, \;\; x_2 = (10 - x)$$

- Utility of reject is obviously 0
- What is the utility of accept?

Player 2

- Depends on whether x is more or less than \$5
- If it is **less** than \$5, then player 2 is getting more than player 1
  - Utility of accepting is

$$\begin{aligned} & (10 - x) - \beta((10 - x) - x) \\ = \quad & (1 - \beta)(10 - x) + \beta x \geq 0 \end{aligned}$$

  - Will always accept such an offer

- If it is **more** than \$5 then player 2 is getting less than player 1
  - Utility of accepting such an offer is

$$\begin{aligned} & (10-x)-\alpha(x-(10-x)) \\ = \quad & (1+\alpha)(10-x)-\alpha x \end{aligned}$$

  - Will reject such an offer if

$$x > \frac{(1+\alpha)}{(1+2\alpha)}10$$

- This is the maximal share of the pie that player 1 can get
  - If $\alpha = 0$ then this is 10
  - As $\alpha \to \infty$, this fraction goes to $\frac{1}{2}10$

Player 1

- What about player 1?
- First, notice they can always guarantee themselves a payoff of 5
  - Offer $x = 5$
  - We know that this is accepted
- This means that they will never make an offer which is rejected
- So they will make an offer somewhere between 5 and $\frac{(1+\alpha)}{(1+2\alpha)}10$

Player 1

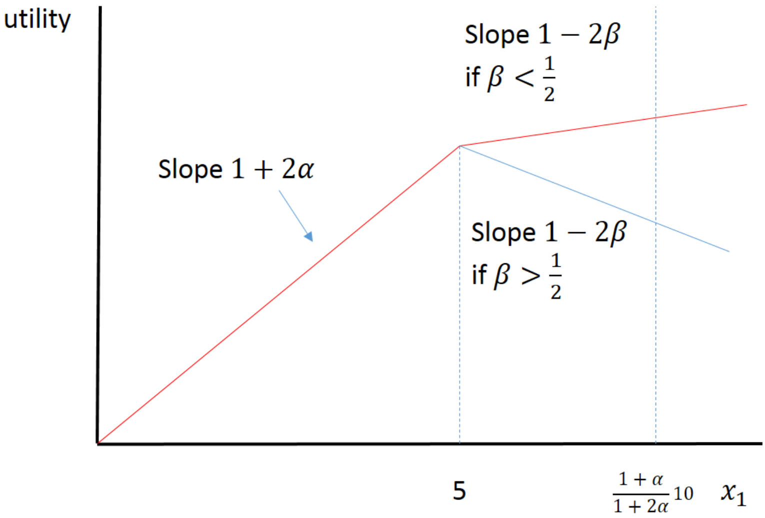
- Where depends on their utility function
  - In this range, Player 1 is getting more than player 2
  - Utility is given by

$$\begin{aligned} u_1(x,(10-x)) &= x-\beta(x-(10-x)) \\ &= x-\beta(2x-10) \end{aligned}$$

  - Taking derivatives WRT x gives

$$\frac{\partial u_1(x_1,x_2)}{\partial x} = 1-2\beta$$

- If $\beta < \frac{1}{2}$ utility is increasing in x, will take the maximum amount they can: $\frac{(1+\alpha)}{(1+2\alpha)}10$
- If $\beta > \frac{1}{2}$ utility is decreasing in x, will take the 50/50 split


utility
Slope $1 - 2\beta$
if $\beta < \frac{1}{2}$
Slope $1 + 2\alpha$
Slope $1 - 2\beta$
if $\beta > \frac{1}{2}$
5
$\frac{1+\alpha}{1+2\alpha}10$
x_1

- The Fehr-Schmidt model provides one mechanism by which people may have social preferences
  - Inequality aversion
- However this is not the only possibility
- Consider this thought experiment
  - Take the standard ultimatum game
  - But now restrict the strategy space of player 1 so that the **maximum** they can offer player 2 is $2
  - How would you respond to an offer of $2 as player 2?

- The Fehr-Schmidt model says that if you rejected \$2 in the original game, you must also reject it in this game
  - The only thing that matters is **outcomes**
- However, you may think that this is not reasonable
  - In the first game rejected \$2 because player 1 was being unfair
  - In the second game they were not being unfair, so you would accept it
- This intuition was formalized in a model by Rabin [1993]
  - The details of which are a bit hairy
  - Will try to give you the intuition

- Two key ideas

1. People are willing to sacrifice their own payoff to help those that they think have been kind to them
2. The are prepared to give up their own payoff to punish those that they think have been unkind

- i.e. this is a model of **fairness** and **reciprocity**
- In order to operationalize this we need some way of measuring how kind one player is being to another

- Let
  - S_1 be the set of strategies that player 1 can choose from
  - S_2 be the set of strategies that player 2 can choose from
  - $\pi_1(s_1, s_2)$ the (materiel) payoff from player 1 if strategies s_1 and s_2 are played
  - $\pi_2(s_1, s_2)$ the (materiel) payoff from player 2 if strategies s_1 and s_2 are played
- We want to develop a kindness function

$$f_1(a_1, b_1)$$

  - How kind does player 1 think they are being if they play a_1, and they think that player 2 will play b_1

Player 1's actions	b_2
a_1^1	$3, 9$
a_1^2	$4, 5$
a_1^3	$7, 1$
a_1^4	$-1, -1$

- Note
  - a_1^1 gives player 2 the highest possible payoff
  - a_1^4 is Pareto dominated
  - a_1^3 gives player 2 the lowest possible payoff ignoring pareto dominated options
- How would you measure fairness?

Fairness Function

- $\pi_2^h(b_2)$ be the highest payoff that player 1 could give player 2
  - In this example 9
- $\pi_2^l(b_2)$ be the lowest payoff *amongst pareto efficient points*
  - In this example 1
- The equitable payoff is given by

$$\pi_2^e(b_2) = \frac{\pi_2^h(b_2) + \pi_2^l(b_2)}{2}$$

  - In this example 5
- let $\pi_2^{\min}(b_2)$ be the worst possible outcome for player 2
  - in our example -1

Fairness Function

- Rabin defines the kindness of player 1 to player 2 as

$$\begin{aligned} f_1(a_1, b_2) &= \frac{\pi_2(a_1, b_2) - \pi_2^e(b_2)}{\pi_2^h(b_2) - \pi_2^{\min}(b_2)} \text{ if } \pi_2^h(b_2) \neq \pi_2^{\min}(b_2) \\ &= 0 \text{ otherwise} \end{aligned}$$

  - Player 1 is being 'kind' if they give player 2 more that the equitable split given what they believe about player 2
  - The degree of kindness is scaled by the range of possible outcomes that player 2 could have received.

- In our example, a_1^1 would be a kind act, as

$$f_1(a_1, b_2) = \frac{9-5}{9-(-1)} = 0.4$$

Fairness Function

- So we now have a way to capture fairness
- But we also want to capture **reciprocity**
  - P1 wants to be kind to P2 if they think P2 has treated them kindly
  - P1 wants to be nasty to P2 if they think that P2 has treated them badly
- We use

$$
\begin{aligned}
\bar{f}_2(b_2, c_1) &= \frac{\pi_1(c_1, b_2) - \pi_1^e(c_1)}{\pi_1^h(c_1) - \pi_1^{\min}(c_1)} \text{ if } \pi_1^h(c_1) \neq \pi_1^{\min}(c_1) \\
&= 0 \text{ otherwise}
\end{aligned}
$$

- To capture P1's beliefs about how kind they think P2 is being to them
  - b_2 is the action they think P2 is taking
  - c_1 is what P1 thinks P2 thinks P1 is playing (!)

- We can now write down the Rabin fairness utility function

$$u_1(a_1, b_2, c_1) = \pi_1(a_1, b_1) + \bar{f}_2(b_2, c_1) f_1(a_1, b_2)$$

  - First bit is standard utility
  - Second bit is fairness utility
  - Payoff increasing in f_1 if $\bar{f}_2(b_2, c_1) > 0$ (i.e. P2 is being fair)
  - Payoff decreasing in f_1 if $\bar{f}_2(b_2, c_1) < 0$ (i.e. P2 is being unfair)

- In order to predict what happens in the game we need a concept of **equilibrium**

1. Players are doing the best thing, given their beliefs
2. Their beliefs are correct, given their information

Definition

An equilibrium of a Rabin Fairness game is a set of actions a_1, a_2, first order beliefs b_1, b_2 and second order beliefs c_1, c_2 such that

1. $a_i = \arg\max_{a_i \in S_i} u_i(a_i, b_j, c_i)$ for $i = 1, 2$ $j = 1, 2$, $i \neq j$
2. $a_i = b_i = c_i$ for $i = 1, 2$

Fairness in the Ultimatum Game

- What does this mean for behavior in the ultimatum game?
- First thing to note is that P2 will always accept P1's offer if it is the highest offer they can make
  - Let p be the size of the pie
  - Let m be the maximum that player 1 is allowed to offer
- We want to check whether it is an equilibrium for P2 to accept m
  - Assume that m is being offered and that player 2 will accept
  - See if there is any benefit to deviating

Fairness in the Ultimatum Game

$$u_2(A, m, T) = \pi_2(A, m) + \bar{f}_1(m, T) f_2(A, m)$$

- Where
  - A is the strategy 'accept'
  - m is the offer of P1
  - T is the minimum amount that P2 would accept
- Notice that
  - $\bar{f}_1(m, T) \geq 0$ as pm is the most P1 can give P2
  - $f_2(A, m) \geq 0$, as accepting gives P1 $(1 - m)p \geq 0$, which is what they would get if P2 rejects

- Thus we have

$$
\begin{aligned}
& u_2(A, m, S) \\
= \;& \pi_2(A, m) + \bar{f}_1(m, S) f_2(A, m) \\
\geq \;& \pi_2(A, m) \\
> \;& 0 \\
\geq \;& \pi_2(R, m) + \bar{f}_1(m, S) f_2(R, m) \\
= \;& u_2(R, m, S)
\end{aligned}
$$

- Where the last inequality comes from the fact that $f_2(R, m) \leq 0$

Fairness in the Ultimatum Game

- However, the same offer might be rejected if P1 could have offered the whole pie
- Assume
  - $p = 1$
  - $m = 1$
- We can show that a receiver will reject an offer of 0.2
- Assume that receiver would accept any offer greater than $z \leq 0.2$
- We can show that this is not an equilibrium

Fairness in the Ultimatum Game

- What is the kindness of offering 0.2 in this situation?
  - $\pi_2^h(z) = 1$
  - $\pi_2^l(z) = z$
  - $\pi_2^e(z) = \frac{(1+z)}{2}$
  - $\pi_2^{\min}(z) = 0$

- Thus, the kindness of an offer 0.2 is given by

$$\begin{aligned}\bar{f}_1(x,z) &= \frac{\pi_2(x,z)-\pi_2^e(z)}{\pi_2^h(z)-\pi_2^{\min}(z)} \\ &= \frac{0.2-\frac{(1+z)}{2}}{1} \\ &= 0.2-\frac{(1+z)}{2}<0\end{aligned}$$

- Is it better to accept or reject that offer?
- Accepting has a fairness of 0, as it is the only pareto dominated option
  - Utility is therefore 0.2
- What is the fairness of rejecting?
  - $\pi_1^h(0.2) = 0.8$
  - $\pi_1^l(0.2) = 0.8$
  - $\pi_1^e(z) = 0.8$
  - $\pi_1^{\min}(z) = 0$
- Fairness is -1

- Utility of rejecting is therefore

$$\begin{aligned} & \pi_1(a_1, b_1) + \bar{f}_2(b_2, c_1) f_1(a_1, b_2) \\ = & -\left(0.2 - \frac{(1+z)}{2}\right) \end{aligned}$$

- Rejecting is better if

$$\begin{aligned} 0.2 & \leq -0.2 + \frac{(1+z)}{2} \\ & \Rightarrow 0.8 \leq 1 + z \\ & \Rightarrow -0.2 \leq z \end{aligned}$$

- Which it is

Take Home

- The Rabin model of fairness can allow for a 20% offer to be
  - Accepted if that is the most that P1 could offer
  - Rejected otherwise
- This is not something that the Fehr-Schmidt model can allow for