Other Regarding Preferences

Mark Dean

Lecture Notes for Spring 2015 Behavioral Economics - Brown University

1 Lecture 1

We are now going to introduce two models of other regarding preferences, and think a little about their implications for various different games. We start with models of inequality aversion, as proposed by Fehr and Schmidt [1999] and Bolton and Ochenfels [2000], before discussing models of fairness and reciprocity [Rabin 1993] and interdependent preferences. We will then move on to some experimental evidence about these different models, and some 'real life' examples of fairness preferences in action.

1.1 Inequality Aversion

Probably the simplest model of other regarding preferences that goes beyond pure altruism is the inequality aversion model, popularized in the late 1990s by Fehr and Schmidt [1999] and Bolton and Ochenfels [2000]. These models modify the standard utility function in the following way. Let's consider a two player game (for example the ultimatum game), and think about an outcome of this game in which player 1 gets monetary outcome x_1 and player two gets outcome x_2. An inequality averse utility function assigns the following utility to each player

$$\begin{aligned} u_1(x_1, x_2) &= x_1 - \alpha \max\{x_2 - x_1, 0\} - \beta \max\{x_1 - x_2, 0\} \\ u_2(x_1, x_2) &= x_2 - \alpha \max\{x_1 - x_2, 0\} - \beta \max\{x_2 - x_1, 0\} \end{aligned}$$

where u_1 is the utility of player 1 and u_2 is the utility of player 2.

The inequality averse utility function has three parts (for simplicity we will talk through the

function of player 1). The first bit is just the standard utility of receiving the amount x. The second bit captures how much the player dislikes getting less than the other player. Note that, if player 1 is getting more than player 2 (i.e. player 1 is 'winning'), then this term disappears from the equation, as $\max\{x_2 - x_1, 0\} = 0$. However, if player 2 is getting more than player 1, then this term reduces player 1's utility by an amount governed by the parameter α. α therefore captures the degree to which the agent is inequality averse when they are the one getting screwed.

The third term captures how much the player dislikes getting more than the other player. This time, note that if player 1 is getting less than player 2 (i.e. player 1 is 'losing') then this term disappears (as $\max\{x_1 - x_2, 0\} = 0$). However, if player 1 is getting more than player 2, then this will reduce utility by an amount governed by the parameter β. β therefore captures the degree to which the player is a nice guy. Generally, α is restricted as being above 0, while β is restricted to being between 0 and $\min(\alpha, 1)$ (meaning that agents are assumed to dislike inequality at least as much when it is not in their favor). This leads to a utility function that has a 'kink' at x_2 (i.e. the amount the other player is getting). Below that point, the slope of the utility function is $(1 + \alpha)$, while above it, it is $(1 - \beta)$.

What would the inequality aversion model predict for the ultimatum game? As usual, we need to solve this game backwards, so we start off by thinking of the second player. Let's imagine that the pie is \$10, and player 1 has offered to keep x_1 and give $10 - x_1$. Player 2 has two choices. They can either accept the split, or reject it, in which case both players get 0. Clearly, the second option gives a utility of 0, as both players get the same, so the question is whether player 2 gets a utility of more than 0 from such an offer.

In order to plug in to the utility function above, we need to know whether $x_1 \geq 10 - x_1$ or $x_1 \leq 10 - x_1$ - i.e. whether player 1 has offered them less than or more than half the pie. Let's start with the former case. In this case

$$\begin{aligned} u_2(x_1, 10 - x_1) &= 10 - x_1 - \beta(10 - 2x_1) \\ &= (1 - \beta)(10 - x_1) + \beta x_1 \\ &\geq 0 \end{aligned}$$

So the second mover will always accept an offer over one half of the pie. What if the offer is less

than one half of the pie? Then we have

$$\begin{aligned} u_2(x_1, 10 - x_1) &= 10 - x_1 - \alpha(x_1 - (10 - x_1)) \\ &= (1 + \alpha)(10 - x_1) - \alpha x_1 \end{aligned}$$

Now if x_1 is large enough (i.e. player 2 gets offered too little) then the utility off accepting the offer will be less than 0. Specifically, the offer will be rejected if

$$x_1 > \frac{(1 + \alpha)}{(1 + 2\alpha)} 10$$

Thus $\frac{(1+\alpha)}{(1+2\alpha)}$ represents the share of the pie that player 1 can take of the pie. If $\alpha = 0$, then they can take everything, as $\frac{(1+\alpha)}{(1+2\alpha)} = 1$. as α increases, then this number falls, reaching an asymptote at $\frac{1}{2}$.

So how will player 1 behave? Assuming that they know the preferences of player 2 (a big assumption, that we will come back to), we know that they can either offer to take an amount greater than $\frac{(1+\alpha)}{(1+2\alpha)}10$, and get rejected, or an amount less than or equal to $\frac{(1+\alpha)}{(1+2\alpha)}10$ and get accepted. Notice that, as $\frac{(1+\alpha)}{(1+2\alpha)}10 > 5$, player 1 can always offer an even split and get accepted. As the utility of an even split is 5, which is better than the zero they get when they are rejected, then they will always choose to make an offer that will be accepted. Notice as well that, the even split is always better than getting less than half the pie. Thus, player 1 will choose to take some amount between 5 and $\frac{(1+\alpha)}{(1+2\alpha)}10$. But how much?

Let x_1 be the amount that player 1 offers to take. As we have just demonstrated, x_1 will be at least 5, so the utility of making such an offer will be

$$u_1(x_1, x_2) = x_1 - \beta(2x_1 - 10)$$

Taking derivatives with respect to x_1, we see that

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

Thus, if $\beta < 0.5$, then $u_1(x_1, x_2)$ is increasing, and player 1 will choose to take as much as they can get - i.e. they will offer $\frac{(1+\alpha)}{(1+2\alpha)}10$. However, if $\beta > 0.5$, (i.e. the agent is very inequality averse), then $u_1(x_1, x_2)$ is actually decreasing in x_1, and player 1 will choose the midpoint: they actually

prefer this to receiving $\frac{(1+\alpha)}{(1+2\alpha)}10$. Thus, such an agent will offer a 50/50 split. If $\beta = 0.5$ then player 1 is indifferent to making any offer between 5 and $\frac{(1+\alpha)}{(1+2\alpha)}10$.

There is another way of thinking about player 1's behavior in this game. First, think about what they would choose if player 2 had no say - i.e. in the dictator game. Then think about what is the largest amount of the pie that player 2 will allow them to take. If player 1 can achieve their first best, then this is what they will choose. If not, they will choose the maximal amount player 2 will allow them to take. Thus, if player 1 is not very inequality averse (i.e. $\beta < 0.5$) then their first best is to take everything. However, generally, player 2 will not let them do this, so they take the most they can (i.e. $\frac{(1+\alpha)}{(1+2\alpha)}10$). However, if player 1 is quite inequality averse (i.e. $\beta > 0.5$) then their first best is a 50/50 split, which they can achieve, so this is what they will offer.

So, to a first approximation, this model predicts the play in the ultimatum game that we observe. If we assume that the population is heterogeneous in α and β (i.e people have different α's and β's) we would expect to see people make offers keeping at least half the pie, and with low offers rejected (by people who have higher than average alphas)

2 Lecture 2

2.1 Fairness and Reciprocity

Think about the following thought experiment. Take the standard ultimatum game, but restrict the strategies that player 1 can employ: specifically, imagine that the maximum that player 1 can offer player 2 is \$2. If you were player 2, would this affect how you would respond to receiving an offer of \$2? Not according to the Fehr-Schmidt model. In that model, the only thing that matters are the outcomes received by the two players. But for many people, it seems that the actions that a person takes matter in how I would like to treat them. Specifically, I might want to punish someone for giving me only \$2 when they could have given me \$5. However, if they had no choice but to give me \$2, then I do not feel the need to punish them: what matters is how the other person has treated me relative to how they could have treated me. This is the idea behind behind Mathew Rabin's 1993 paper "Incorporating Fairness into Game Theory and Economics.", which aimed to develop a model to capture the following insights:

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

In order to do this, we need some way of measuring how kind one player has been to the other. Unfortunately, we are going to need some notation for this. Let S_1 and S_2 be the set of strategies that player 1 and player 2 can pick, and let $\pi_i : S_1 \times S_2 \to \mathbb{R}$ represent the material payoff of player i. Now let's think about a player 1 who chooses an action a_1 when they think that the other player has chosen b_2 (b represents player 1's beliefs about the actions of player 2). We want to develop a kindness function

$$f_1(a_1, b_2)$$

which measures how kind player 1 thinks they are being by selecting a_1 when they think the other player has chosen b_2. We will illustrate Rabin's approach with the following example. Let's assume player 1 thinks player 2 has chosen b_2, and these are the possible actions they can take. The first

number in each cell is player 1's payoff and the second number player 2's payoff

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

Now we need to define the following concepts

- Let $\pi_2^h(b_2)$ be the highest payoff that player 1 could give player 2 (so in the case above this would be 9)
- Let $\pi_2^l(b_2)$ be the lowest payoff *amongst pareto efficient points* (i.e. points such that one player cannot be made better off without making another worse off) (in the example this would be 1)
- Let the equitable payoff be given by

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

  so in the above example, this would be 5

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

Using these concepts, Rabin defines the kindness of player 1 to player 2 as

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

So how to interpret this? The top bit says that player 1 is being 'kind' if they give player 2 more that the equitable split given what they believe about player 2 - i.e. if $\pi(a_1, b_2) > \pi_2^e(b_2)$. They are being unkind if less than the equal split is given. 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

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

This function captures how kindly player 1 thinks that they are treating player 2. However, the idea behind this model is one of reciprocity: player 1 wants to be kind to player two if they think player 2 is going to be kind to them. We therefore need to capture player one's beliefs about the fairness of player 2. We do this using the function

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

where b_2 is what player 1 believes player 2 with do, and c_1 is player 1's beliefs about what player 2 believes about the actions of player 1. Thus, this function measures how kind the action player 1 thinks that player 2 will take, given what they think player 2's beliefs are about their own actions.

Putting these elements together we get the utility function

$$\begin{aligned} & u_1(a_1, b_2, c_1) \\ = \; & \pi_1(a_1, b_2) \\ & + \bar{f}_2(b_2, c_1)(1 + f_1(a_1, b_2)) \end{aligned}$$

This is the expected payoff of an agent who takes action a_1, believing that player 2 will take action b_2 and that player 2 thinks that they will take action a_1. The first bit of which is just the standard utility function - i.e. the payoff that player 1 gets when they play a_1 and player 2 plays a_2. Notice that, in the second term, player 1 only gets to choose a_1, and so affects their own kindness $f_1(a_1, b_2)$. If player 1 thinks that player 2 is being kind (and so $\bar{f}_2(b_2, c_1) > 0$), then their utility is increasing in their own kindness. However, if the agent is being unkind (and so $\bar{f}_2(b_2, c_1) < 0$) then their payment is decreasing in their own kindness. - so they get positive utility from hurting the other player.

One thing that may be confusing you is fact that we use $1 + f_1(a_1, b_2)$ rather than just $f_1(a_1, b_2)$ in the second term. The first thing to note is that this doesn't affect the choices that player 1 makes (you should check to see that you understand this). The reason that Rabin does this is to ensure that, whenever player 2 treats player 1 unfairly, then player 1's payoff is below their material payoff (as $f_1(a_1, b_2)$ is bounded below by a half). It is not really clear why this matters.

As this is a game, we now need to define what we mean by an equilibrium. In general, and equilibrium in economics has two features:

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

The equilibrium of this game is no different

Definition 1 *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* $i = 1, 2$

Before moving on to the ultimatum game, lets think about how Rabin fairness plays out in some standard, static cases. First, lets think about what used to be called in less enlightened times, the battle of the sexes game

	O	B
O	$2x, x$	$0, 0$
B	$0, 0$	$x, 2x$

The idea being that a couple would both prefer to spend the evening together to spending it alone, but one party would prefer to spend it doing activity 0, while the other would prefer doing activity B. Under standard preferences, both O, O and B, B are equilibria of this game. Are they still equilibria under Rabin's preferences? To answer this question, we need to pose the following questions

1. If player 1 thought player 2 was going to play O, and thought that player 2 thought that they (player 1) would play O, would they prefer to play O or B?
2. If player 2 thought player 1 was going to play O, and thought that player 1 thought that they (player 2) would play O, would they prefer to play O or B?

Let's take the first of these questions. In order to figure this out, we need to figure out the relevant kindness functions. First, let's think about $\bar{f}_2(b_2, c_1)$. To calculate this, we need to figure out the following

- $\pi_1^h(O) = 2x$
- $\pi_1^l(O) = 2x$
- $\pi_1^e(O) = 2x$
- $\pi_1^{\min}(O) = 0$

Thus

$$\begin{aligned} \bar{f}_2(O, O) &= \frac{\pi_1(O, O) - \pi_1^e(O)}{\pi_1^h O) - \pi_1^{\min}(O)} \\ &= \frac{0}{2x} \\ &= 0 \end{aligned}$$

Thus, player 1 does not think that player 2 is either being fair or unfair. Their utility function therefore boils down to $u_1(a_1, b_2, c_1) = \pi_1(a_1, b_2)$, and so they would prefer to play O than B. A similar argument shows that player 2 would rather play O than B, so O, O is a valid equilibrium.

What about the case where player 2 plays B and player 1 plays O. Can this be an equilibrium? Obviously not under standard preferences, but what about under Rabin preferences. Well, let's once again ask whether, if player 1 thinks that player 2 is playing B, and thinks that player 2 believes that he (player 1) is playing O, would player 1 rather play O of B. Again, we have

- $\pi_1^h(O) = 2x$
- $\pi_1^l(O) = 2x$
- $\pi_1^e(O) = 2x$
- $\pi_1^{\min}(O) = 0$

but now, we have that

$$\begin{aligned}\bar{f}_2(O,B) &= \frac{\pi_1(O,B)-\pi_1^e(O)}{\pi_1^h(O)-\pi_1^{\min}(O)}\\ &= \frac{0-2x}{2x}\\ &= -1\end{aligned}$$

So player 1 thinks that player 2 is treating them badly. How will they respond? Well, their utility is now given by

$$u_1(a_1,B,O)) = \pi_1(a_1,B) - (1+f_1(a_1,B))$$

In order to calculate $f_1(a_1,B)$ we need to figure out

- $\pi_2^h(B) = 2x$
- $\pi_2^l(B) = 2x$
- $\pi_2^e(B) = 2x$
- $\pi_2^{\min}(B) = 0$

and so we have

$$f_1(a_1,B) = \frac{\pi_2(a_1,B)-2x}{2x}$$

So, the payoff for playing O is given by

$$\begin{aligned}u_1(O,B,O)) &= 0-(1-1)\\ &= 0\end{aligned}$$

While the payoff for playing B is given by

$$\begin{aligned}u_1(B,B,O)) &= x-(1-0)\\ &= x-1\end{aligned}$$

Thus, if x is small enough then player 1 will be prepared to play O in order to spite player 2 if they think player 2 is treating them unfairly (as the game is symmetric, the same argument can be

applied to player 2). The Rabin model supports a 'spiteful' equilibrium that the standard utility function does not. Note that, as x gets large, fairness concerns disappear (this will always be true, as the fairness functions are bounded independently of x). This is a feature of the Rabin model: as stakes get large, fairness concerns get small.

Now let's think about how to use Rabin preferences in order to analyze the ultimatum game. To do so, we are going to begin by thinking about player 2 (the receiver). The first thing to note is that player 2 will always accept the highest offer that player 1 can make. To see this, let say that the two players are bargaining over a pie of size p, and the highest fraction that the sender is allowed to offer the receiver is m (i.e. they have to keep m for themselves). First, let's ask if it is an equilibrium for the receiver to accept the offer m. To check this, lets assume that player 1 has offered m assuming that player 2 will accept, and see whether player 2 has any incentive to deviate. Remember that the utility of accepting such an offer is given by

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

where S is the general strategy of player 2. Remember that this strategy has to specify what happens as a result of every offer. At this stage we don't care what this strategy entails for other offers, only that the offer m will be accepted. Note that it has to be the case that $\bar{f}_1(m, S) \geq 0$, as mp is the best payoff it is possible for the sender to give the receiver. Notice also that $f_2(A, m) \geq 0$, as accepting gives the sender $(1 - m)p \geq 0$, which is what the sender would get under a rejection. 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 follows from the fact that $f_2(R, m) \leq 0$. Thus, it is always an equilibrium for the receiver to accept the maximal offer.

Can it also be a subgame perfect equilibrium to reject the maximal offer? While it is (I think) possible for any arbitrary strategy for the receiver, it requres player 2 to act in a peverse way, by

rejecting higher offers and accepting lower offers. If we rule this out (by assuming that the receiver has a cutoff, above which they accept allo offers), then it cannot be an equilibrium to reject the maximal offer. To see this, note that

$$u_2(R, m, S') = \pi_2(R, m) + \bar{f}_1(m, S') f_2(R, m)$$

Now S' is again some arbitrary strategy for the receiver, but note that if we assume that they have a cutoff strategy, and are rejecting m, then they must reject *all* offers. This means that $\bar{f}_1(m, S') = 0$, as there is no way the sender can give the receiver anything other than 0 (as the receiver rejects all offers). Thus we have

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

We derive the complete solution to the ultimaitum game in the appendix, because it is a little complicated. The key point for now is that the Rabin model predicts that receiver behavior may depend on the action set of the sender. For example, we can show that, for a pie of size 1, if the sender can make any possible proposal, then the receiver will reject any offer of 0.2.To see this, assume that the sender acceps any offer greater than z, where $z \leq 0.\dot{2}$ Then, in order to calculate the kindness of player 1 of offering an 0.2, we need to figure out

- $\pi_2^h(z) = 1$
- $\pi_2^l(z) = z$
- $\pi_2^e(z) = \frac{(1+z)}{2}$
- $\pi_2^{\min}(z) = 0$

Thus, the kindess of an offer 0.2 is given by

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

The question is whether it is better to accept or reject this offer. Accepting the offer gives fairness 0, so has utility 0.2 Rejecting the offer has fairness -1, and so gives utility $-\left(0.2-\frac{(1+z)}{2}\right)$. The question is thereofe whether

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

Which it is. Thus, the receiver will reject such an offer. Yet, as we have shown, if the maximal offer that the sender can make is 0.2, then the receiver will accept it.

This is the key difference between inequality aversion models and fairness baseed models. We will now have a quick looka t some of the relevant data to see what happens when people are confronted by this very experiment

3 Appendix

We will now characterize the equilibrium of the standard ultimaitum game under Rabin preferences. We start with the behavior of the receiver. Again, we assume that player 2 has a cutoff strategy z, such that they will reject any offer below z and accept any offer above z. The strategy we will use to determine player 2's strategy is to find a z such that, if this were player 2's cutoff, then they would in fact want to reject offers if and only if they were below z.

In order to do so, assume that player 2 has a cutoff z (which player 1 knows), and that player 1 offers x. Furthermore, lets assume that the total value of the pie is p. How kind is player 1 being? Well we can see that

- $\pi_2^h(z) = p$
- $\pi_2^l(z) = zp$
- $\pi_2^e(z) = \frac{(1+z)p}{2}$
- $\pi_2^{\min}(z) = 0$

Thus the kindness of player 1 towards the second mover (if the offer is above z) is given by

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

We need to figure out how kind player 2 is being towards player 1 if they accept or reject, To do so, we need to know that

- $\pi_1^h(x) = (1-x)p$
- $\pi_1^l(x) = (1-x)p$
- $\pi_1^e(x) = (1-x)p$
- $\pi_1^{\min}(z) = 0$

For simplicity, we are going to write the utility function as

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

where A indicates acceptance. Thus, the utility of accepting an offer is given by

$$\begin{aligned} u_2(A, x, z) &= xp + \left(x - \frac{(1+z)}{2} \right) 0 \\ &= xp \end{aligned}$$

while the utility of rejecting is

$$u_2(R, x, z) = 0 - \left(x - \frac{(1+z)}{2} \right)$$

So, the receiver wants to accept if

$$\begin{aligned} xp &\geq -\left(x - \frac{(1+z)}{2} \right) \\ &\Rightarrow (2p+2)x \geq 1 + z \end{aligned}$$

For this to be an equilibrium, it has to be the case that the decision maker wants to accept if x is above z. Thus, it has to be the case that, for $x = z$, the above expression holds, and so

$$\begin{aligned} (2p+2)z &\geq 1 + z \\ &\Rightarrow z \geq \frac{1}{2p+1} \end{aligned}$$

Thus, it must be the case that player 2's cutoff is above $\frac{1}{2p+1}$. If not, then there will be some values of $x > z$ that player 2 would like to reject. Notice that this cutoff goes to zero as the pie gets big, and goes to 1 as the pie gets very small.

It also has to be the case that the decision maker wants to reject when the offer is below z. When doing so, the sender is making an offer that they know that player 2 will reject. Therefore, their kindness is given by

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

And player two has the choice of accepting and getting utility xp, or rejecting and getting utility $\frac{(1+z)}{2}$. Thus, the decision maker will reject if

$$\frac{(1+z)}{2} \geq xp$$

Again, we need it to be the case that, when $x = z$, the decision maker is happy rejecting, and so we have that

$$\begin{aligned} \frac{(1+z)}{2} &\geq zp \\ &\Rightarrow \frac{1}{2p-1} \geq z \end{aligned}$$

Notice that we aready showed that it has to be an euqilibrium to accept the maximal offer that player 1 can make, which in this case is 1. Thus, any cuttoff that satisfies

$$\min\left(1, \frac{1}{2p-1}\right) \geq z \geq \frac{1}{2p+1}$$

is an equilibrium strategy for the receiver. So, for example, of the size of the pie is 1, then any cutoff between $\frac{1}{3}$ and 1 will work.

What about the proposer? Consider a proposer facing a receiver with a cutoff z. If they make an offer x above z, then the receiver will accept, which, gives them a fairness of 0. Thus the utility of any such offer is equal to the materiel payoff - $(1-x)p$. Of such offers, the proposer would therefore like to take the largest slice of the pie they can - i.e. set $x = z$.

Would they ever prefer to make an offer of less than z are get rejected? In this case, the fairness of the responder's behavior is given by -1 (as they are giving the worst possible outcome to the sender). To calculate the fairness of the sender, we need to calculate

- $\pi_2^h(z) = 1$
- $\pi_2^l(z) = z$
- $\pi_1^e(x) = \frac{1+z}{2}$
- $\pi_1^{\min}(z) = 0$

Thus, the fairness of player 1 making an offer x below z is equal to $-\frac{1+z}{2}$. Thus the question is whether this is better or worse that the best payoff they can get from getting an offer accepted, which is $(1-z)p$. In other words we need to know whether

$$\frac{1+z}{2} \geq (1-z)p$$

This will be true if

$$z \geq \frac{1}{1+2p}$$

If we compare this to the conditions above, we see that there is one cutoff that supports cooperation: $z = \frac{1}{1+2p}$