Cops and Robbers

Recall the game of cops and robbers from class, shown in Figure 1.

	R	H
P	$2, -2$	$0, 1$
D	$-1, 3$	$1, 0$

Figure 1: The cops and robbers game from class

Check (easy): show that there is no pure-strategy Nash equilibrium in this game.

Last time, we saw that, together, there must always be an odd number of Nash equilibria and mixed-strategy Nash equilibria; importantly, this implies that there must always be at least one equilibrium, hence an equilibrium of some kind must exist. Since there is no pure-strategy Nash equilibrium of this game, we should look for a mixed-strategy Nash equilibrium.

Mixed-strategy Nash equilibrium

Suppose that the police **P**atrol with probability p_1, and the robbers **R**ob with probability p_2. We know that in mixed-strategy Nash equilibrium, each side's randomization must keep the other indifferent between its two strategies. Since the police's randomization must keep the robbers indifferent, we have

$$\begin{aligned} & u_2\left(p_1\mathrm{P} + (1-p_1)\mathrm{D}, \mathrm{R}\right) = u_2\left(p_1\mathrm{P} + (1-p_1)\mathrm{D}, \mathrm{H}\right) \\ \iff & p_1(-2) + (1-p_1)(3) = p_1(1) + (1-p_1)(0) \\ \iff & 3 = 6p_1 \quad \rightsquigarrow \quad \boxed{p_1 = \frac{1}{2}} \end{aligned}$$

Similarly, the robbers' randomization must keep the police indifferent, so we must have

$$\begin{aligned} & u_1\left(\mathrm{P}, p_2\mathrm{R} + (1-p_2)\mathrm{H}\right) = u_1\left(\mathrm{D}, p_2\mathrm{R} + (1-p_2)\mathrm{H}\right) \\ \iff & p_2(2) + (1-p_2)(0) = p_2(-1) + (1-p_2)(1) \\ \iff & 4p_2 = 1 \quad \rightsquigarrow \quad \boxed{p_2 = \frac{1}{4}} \end{aligned}$$

Therefore we have a mixed-strategy Nash equilibrium where the police **P**atrol with probability $p_1 = 1/2$ and the robbers **R**ob with probability $p_2 = 1/4$.

Updating payoffs

We now ask the question, what happens if we punish the robbers more for being caught red-handed? In particular, let's change their utility of being caught from -2 to -5. The new game is shown in Figure 2.

	R	H
P	$2, -5$	$0, 1$
D	$-1, 3$	$1, 0$

Figure 2: The cops and robbers game, modified to punish the robbers more if they get caught

Check (easy): show that there is *still* no pure-strategy Nash equilibrium in this game.

Let's keep the convention that the police **P**atrol with probability p_1 and the robbers **R**ob with probability p_2. Then we can repeat the above analysis to determine the probabilities which support a mixed-strategy Nash equilibrium.

$$\begin{aligned} & u_2\left(p_1\mathrm{P} + (1-p_1)\mathrm{D}, \mathrm{R}\right) = u_2\left(p_1\mathrm{P} + (1-p_1)\mathrm{D}, \mathrm{H}\right) \\ \iff \quad & p_1(-5) + (1-p_1)(3) = p_1(1) + (1-p_1)(0) \\ \iff \quad & 3 = 9p_1 \quad \rightsquigarrow \quad \boxed{p_1 = \frac{1}{3}} \end{aligned}$$

$$\begin{aligned} & u_1\left(\mathrm{P}, p_2\mathrm{R} + (1-p_2)\mathrm{H}\right) = u_1\left(\mathrm{D}, p_2\mathrm{R} + (1-p_2)\mathrm{H}\right) \\ \iff \quad & p_2(2) + (1-p_2)(0) = p_2(-1) + (1-p_2)(1) \\ \iff \quad & 4p_2 = 1 \quad \rightsquigarrow \quad \boxed{p_2 = \frac{1}{4}} \end{aligned}$$

In the new mixed-strategy Nash equilibrium, the police **P**atrol with probability $p_1 = 1/3$ and the robbers **R**ob with probability $p_2 = 1/4$.

Implications

Interestingly, by increasing punishment for the robbers we have not changed the probability with which they **R**ob, but we have changed the probability with which the police **P**atrol — although we've changed it the "wrong" way! More on this in a moment.

An outcome that we might be interested in is the probability with which the robbers succeed in their gambit. For this to be the case, they must **R**ob while the police stay out and eat **D**onuts. Since the probability that the robbers **R**ob is p_2, the probability that the police eat **D**onuts is $1 - p_1$, and the randomizations are independent, the probability that the robbers successfully breach the target and make off with the goods is $(p_2)(1 - p_1)$.

In the first case, this is

$$p_2(1-p_1) = \frac{1}{4}\left(1 - \frac{1}{2}\right) = \frac{1}{8}$$

So roughly 12% of the time the robbers will successfully rob. Once we increase punishment, however, this probability changes to

$$p_2(1-p_1) = \frac{1}{4}\left(1 - \frac{1}{3}\right) = \frac{1}{6}$$

So roughly 17% of the time the robbers will successfully rob. In the face of increased punishment, this is both counterintuitive and really not good.

Intuition

We need to come up with a story explaining what's going on here, and in the end it comes down to the fact that we are presuming *equilibrium behavior*. That is, we are looking for strategies that support equilibrium. Here are the two facts that lead to the troublesome outcome above:

- The utility of the police has not been changed. Since it is the *robbers'* randomization which keeps the police indifferent, changing the robbers' utility will not affect the probability with which they must **R**ob *in equilibrium.* This in itself is a little counterintuitive, but is fairly reasonable once spelled out.

- Suppose that the strategy of the police remains fixed. Since we have increased the punishment for getting caught, **R**ob looks much less appealing to the robbers while staying **H**ome does not change. As they were previously exactly indifferent between the two actions, it must now be the case that staying **H**ome is strictly better than robbing, so the robbers will surely stay **H**ome. But if the robbers stay home, the police are happier to eat **D**onuts; and if the police are eating **D**onuts, the robbers are happier to **R**ob...

  This means that *in equilibrium* the police must make **R**obbing look more appealing! They will do this by decreasing the probability with which they patrol.

Consider this story: the government increases penalties for robbery on Monday. Since the robbers were perfectly indifferent before, now they won't **R**ob (yay, society!). On Tuesday, the police see that the robbers haven't robbed, so they decide to eat **D**onuts instead of patrolling. On Wednesday, the robbers see that the cops have been eating **D**onuts, so they **R**ob the target and make off with loot. On Thursday the police see that the target has been robbed, so they **P**atrol; the robbers see this, so don't **R**ob. On Friday, the police see that the target hasn't been robbed again, so they eat **D**onuts instead...

The only way out of this loop is to find the correct mixed-strategy Nash equilibrium. Here, this means that robbing must look *more* appealing than otherwise; since being caught is being punished more strictly, the only option is for the police to **P**atrol less often, which means that more successful robberies occur!

This brings into focus an important point: in game theory, we are concerned *only* with what the model on paper says. If these results seem counterintuitive and backwards (and the math checks out), it's a sign that either the model is ill-specified, or that our intuition is wrong. Not being an expert on the criminal justice system I'll withold judgment in this case, but my guess would be that we are not accounting for all of the important issues in the decision of the police to patrol or eat donuts, or in the decision of crooks to rob or stay home.

Discounting

It is a fact of life that people are impatient. Of course, people may be impatient to varying degrees and differently in different situations, but nonetheless anyone can tell you that a payoff today is better than a payoff tomorrow.[1] There are numerous ways of quantifying and codifying this, but the most mathematically useful[2] is the notion of *exponential discounting.*

Definition

Consider a game where time t passes, $t = 0, 1, 2, \ldots$. An agent has *discount rate* $\delta \in (0, 1)$ if utility u at time t' is equivalent to utility $\delta^{t'} u$ at time $t = 0$.

---

[1]There are of course some exceptions to this if we continue to equate utility and money: \$5000 on December 31 may be worth less than \$5000 on January 1, depending on your tax situation. However, if we consider *utility* rather than *dollars*, this confusion disappears. For now, we will continue to confuse money and utility and put aside little conundra like the previous.

[2]In particular, this form has the nice property of *time-consistent preferences*: what looks better today will also look better tomorrow. There are a handful of other schools of thought regarding how agents should discount future utility, but we will start with the simplest.

In particular, utility u one period in the future is worth δu right now. If s' is also a point in time and $t' > s'$, then utility u at time t' is equivalent to utility $\delta^{t'-s'}$ at time $t = s'$.[3]

This construction allows us to make statements about how agents will act today regarding choices they may have to make in the future. *This is distinct from sequential games*, although it is related: in a sequential game, agents act in turn but we don't consider more than a small amount of time as passing.[4] In games with discounting, we think of agents as, say, walking away from the game for a little while, then coming back at a later date. They are still engaged with the game at all points, but the outside world is continuing apace.

Back-and-forth ultimatum game

Let's generalize the ultimatum game slightly. Agents will split 1 unit of a perfectly-divisible good. Agent 1 will make the first offer x_1, then agent 2 can accept or reject. If agent 2 accepts, payoffs are $(1 - x_1, x_1)$ immediately; if agent 2 rejects, *time passes* and then agent 2 can respond with a counteroffer x_2. Upon receipt of this counteroffer, agent 1 can accept or reject; if agent 1 accepts, payoffs are $(x_2, 1 - x_2)$, and if agent 1 rejects payoffs are $(0, 0)$. This game is pictured in Figure 3.

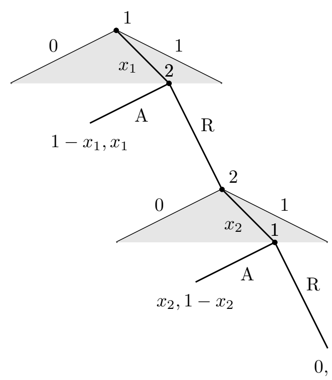

Figure 3: the twice-repeated ultimatum game, or, ultimatum game with counteroffers

This game clearly has a sequential flavor — one agent is acting, then another responds, etc. — so it is natural to look for subgame-perfect Nash equilibrium. Working backwards, we can see

(iv) In the last stage, agent 1 receives x_2 from **A**ccepting and 0 from **R**ejecting, therefore she should certainly **A**ccept any $x_2 > 0$. She is indifferent between **A**ccept and **R**eject when $x_2 = 0$; for mathematical niceness we will assume that she **A**ccepts.[5] Agent 1's strategy in this round is then (A).

(iii) In the second-to-last stage, agent 2 knows that agent 1 will **A**ccept any offer. Since he wants to maximize his own payoff, he will give her the minimum amount possible; as she will accept 0, he will offer her 0, leading to implied payoffs of $(0, 1)$.

[3]The discount rate is sometimes referred to as the *discount factor*. You should probably keep your ears peeled for anything involving the word "discount."

[4]Think about this: if you play rock-paper-scissors and your friend throws first, you don't need more than a fraction of a second to pick your move. In this sense, there are different phases to the game — your friend throws, then you throw — but no "real" time has passed.

[5]The reason for this is straightforward: suppose that agent 1 **R**ejects $x_2 = 0$. Agent 2 wants to maximize his payoff in the previous round, so he should offer the least amount possible. If he offers 0, agent 1 will **R**eject and so agent 2 will get 0; however, if he offers $\varepsilon > 0$ agent 1 will accept and he will get $1 - \varepsilon$. Therefore he should offer the *smallest possible* $\varepsilon > 0$. Since there is a continuum of numbers available, there is no such number! (for any $\varepsilon > 0$, there is a smaller $0 < \varepsilon' < \varepsilon$) This means that he has no optimal response to agent 1's strategy, and we are up a creek. The way around this is to assume that agent 1 **A**ccepts when $x_2 = 0$, which may as well be the case since she is indifferent between **A**ccepting and **R**ejecting.

(ii) In the third-to-last stage, agent 2 knows that if he **A**ccepts he can receive x_1 *today* and that if he **R**ejects he can receive 1 *tomorrow*. Remembering now that agents discount the future, we can see that 1 tomorrow is worth $1\delta = \delta$ today; hence his choice is between **A**ccepting and receiving x_1 today and **R**ejecting and receiving δ tomorrow. His optimal strategy is therefore

$$\begin{cases} \text{A} & \text{if } x_1 \geq \delta, \\ \text{R} & \text{otherwise.} \end{cases}$$

Notice that we again have the agent accepting where he is indifferent between accepting and rejecting. For more on this, see the previous footnote.

(i) In the fourth-to-last stage (the first stage), agent 1 knows that agent 2 will **A**ccept any offer $x_1 \geq \delta$. We need to consider two contingencies:

- If agent 1 can get agent 2 to **A**ccept, her payoff will be $1 - x_1$. In this case, she would like to minimize her offer to player 2 in order to keep more of the spoils for herself. She should therefore offer agent 2 the minimum amount that will lead to **A**cceptance, $x_1 = \delta$. Her payoff will be $1 - \delta$.

- If agent 1 causes agent 2 to **R**eject, time will pass and agent 2 will offer her $x_2 = 0$ and she will **A**ccept. Her payoff in the future is then 0, hence her view of her payoff today is $0\delta = 0$.

Agent 1's choice is then between offering $x_1 = \delta$ and receiving $1 - \delta$, or offering $x_1 < \delta$ and receiving 0. Clearly $1 - \delta > 0$, so agent 1 should offer $x_1 = \delta$ to agent 2.

Thus using subgame perfection we can see that agent 1 will offer $x_1 = \delta$ in the first period, and agent 2 will accept. Payoffs are then $(1 - \delta, \delta)$. What is nice about this is that by simply repeating the situation but reversing the roles of the agents, we have given agent 2 significant bargaining power that simply did not exist in the one-period ultimatum game. This makes things look a lot more fair.

Question (easy): what happens when agents become more patient (e.g., as $\delta \to 1$)? In the limiting case ($\delta = 1$) can you come up with a straightforward reason for this prediction? (hint: do agents care about whether they receive payoffs today or tomorrow?)

Question (medium): what happens if agents have *outside options*? That is, what if when agent 1 **R**ejects in the final round, payoffs are (α_1, α_2) instead of $(0, 0)$? Assume that $\alpha_1 \geq \alpha_2 \geq 0$. What changes if instead $\alpha_2 > \alpha_1 \geq 0$?

Back-and-forth ultimatum game, redux

To more-robustly illustrate these concepts, let's link *four* copies of the ultimatum game rather than 2. That is, time now passes over four increments ($t = 0, 1, 2, 3$): agent 1 proposes x_1 and agent 2 responds ($t = 0$); conditionally, agent 2 proposes x_2 and agent 1 responds ($t = 1$); conditionally, agent 1 proposes x_3 and agent 2 responds ($t = 2$); finally, conditionally agent 2 proposes x_4 and agent 1 responds ($t = 3$). This is pictured in Figure 4.

Fortunately, subgame-perfect Nash equilibrium has a nice recursive structure. From our previous analysis, we can already see that at time $t = 2$ payoffs will be $(1 - \delta, \delta)$. We can then pick up where we left off:

$t = 1$, response Agent 1 knows that she can receive x_2 today if she **A**ccepts or $1 - \delta$ tomorrow if she **R**ejects. Since $1 - \delta$ tomorrow is worth $(1 - \delta)\delta$ today, she should **A**ccept if $x_2 \geq (1 - \delta)\delta$; her strategy is then

$$\begin{cases} \text{A} & \text{if } x_2 \geq (1 - \delta)\delta, \\ \text{R} & \text{otherwise.} \end{cases}$$

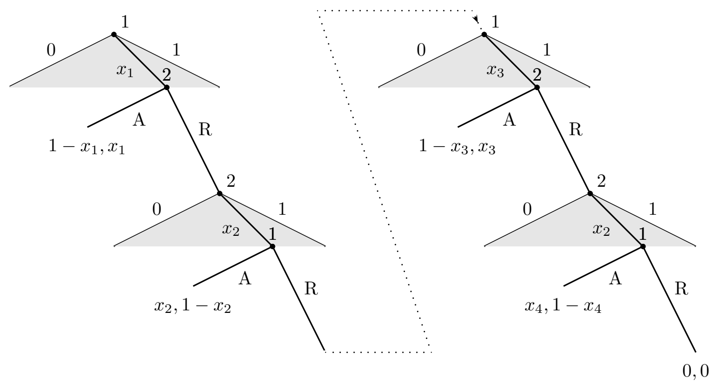

Figure 4: the four-times-repeated ultimatum game, or, ultimatum game with three counteroffers

$t = 1$, proposal Agent 2 knows that agent 1 will **A**ccept if $x_2 \geq (1-\delta)\delta$, hence if he wants her to **A**ccept he should offer the least amount possible, $x_2 = (1-\delta)\delta$. We need only check now that this is preferable to causing agent 1 to **R**eject.

Because we are looking at subgame-perfect Nash equilibrium and agents are rational, we know that if agent 1 **R**ejects she will offer δ tomorrow and agent 2 will **A**ccept, giving agent 2 a payoff of δ. This δ tomorrow is worth $\delta\delta = \delta^2$ today. If instead agent 2 offers $(1-\delta)\delta$ today, he receives an immediate payoff of[6]

$$1 - (1-\delta)\delta = 1 - \delta + \delta^2 > \delta^2$$

Therefore causing agent 1 to **A**ccept by offering $x_2 = (1-\delta)\delta$ is better than causing her to **R**eject. His strategy is therefore $x_2 = (1-\delta)\delta$.

$t = 0$, response Agent 2 knows that he can receive x_1 today if he **A**ccepts or $1-(1-\delta)\delta$ tomorrow if he **R**ejects. Since $1-(1-\delta)\delta)$ is worth $(1-(1-\delta)\delta)\delta)$ today, he should **A**ccept if $x_1 \geq (1-(1-\delta)\delta)\delta$; his strategy is then

$$\begin{cases} \text{A} & \text{if } x_1 \geq (1-(1-\delta)\delta)\delta, \\ \text{R} & \text{otherwise.} \end{cases}$$

$t = 0$, proposal In the initial round, agent 1 knows that agent 2 will **A**ccept if $x_1 \geq (1-(1-\delta)\delta)\delta$, hence if she wants him to **A**ccept she should offer the least amount possible, $x_1 = (1-(1-\delta)\delta)\delta)$. We need only check now that this is preferable to causing agent 2 to **R**eject.

Because we are looking at subgame-perfect Nash equilibrium and agents are rational, we know that if agent 2 **R**ejects he will offer $(1-\delta)\delta$ tomorrow and agent 1 will **A**ccept, giving agent 1 a payoff of $(1-\delta)\delta$. This $(1-\delta)\delta$ tomorrow is worth $(1-\delta)\delta\delta = (1-\delta)\delta^2$ today. If instead agent 1 offers $(1-(1-\delta)\delta)\delta$ today, she receives an immediate payoff of

$$1 - (1-(1-\delta)\delta)\delta = 1 - \delta + \delta^2 - \delta^3 > \delta^2 - \delta^3 = (1-\delta)\delta^2$$

Therefore causing agent 2 to **A**ccept by offering $x_1 = (1-(1-\delta)\delta)\delta$ is better than causing him to **R**eject. Her strategy is therefore $x_1 = (1-(1-\delta)\delta)\delta$.

Thus subgame perfection tells us that agent 1 proposes $x_1 = (1-(1-\delta)\delta)\delta$ in the first period, and agent 2 immediately accepts. Payoffs are therefore

$$(u_1, u_2) = \left(1 - \delta + \delta^2 - \delta^3, \delta - \delta^2 + \delta^3\right)$$

If you are noticing a pattern here, see the following question.

[6] Remember, $\delta < 1$.

Question (hard): what happens as the number of times the game is repeated goes to infinity? Assume that we are repeating the two-stage game above (i.e., the first mover is not the final responder). (hint [makes question **(medium)**]: if agent 2 will get v tomorrow, agent 1 will offer δv today and receive $1 - \delta v$ herself; therefore one period prior agent 2 should offer $\delta(1 - \delta v)$ and receive $1 - \delta(1 - \delta v)$ himself)

Question (medium): compare the 2- and 4-stage games above with the 3-stage (and, optionally, 5-stage) games in which agent 1 is both the initial proposer and the final responder. What does the extensive-form game tree look like? Is agent 1 better off or worse off?

Rubinstein bargaining

As a last iteration, what happens if bargaining may continue indefinitely? If you answered **Question (hard)** above, you have a preview of this outcome; still, there is a simpler way of analyzing the situation that may prove useful in the future.

Suppose that we still have two agents bargaining in the above fashion; we will generalize the model slightly by assuming that agent 1 has discount rate δ_1 and agent 2 has discount rate δ_2 — that is, the agents may vary in the extent to which they are patient. This variation will allow us to compare whether it is good to be more or less patient than your opponent. Although the "circular notation" is far from standard, Figure 5 gives a reasonable picture of the structure of the game.

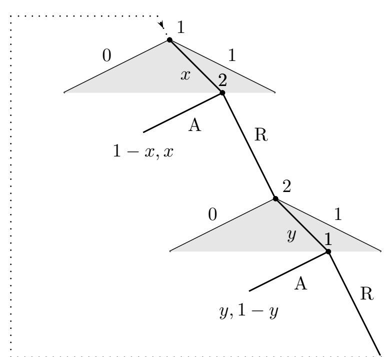

Figure 5: Rubinstein bargaining with two players

Now, the immediate trouble with repeating this process infinitely is that we can no longer use backward induction to find subgame-perfect Nash equilibrium: since there is no final stage of the game, there is nowhere to start the process of working backward! We are going to have to find a more suitable structure for analyzing behavior. Instead, let's think about what changes from $t = 0$ to $t = 2$.

At both points in time, agent 1 is making a proposal to agent 2, and agent 2 is responding. At both points in time, the game can continue infinitely after: by reaching time $t = 2$, the agents have chewed up two rounds, but two off of infinity is still infinity. In this sense, the game at time $t = 2$ looks *exactly* like the game at $t = 0$ except that two units of time have passed.

It stands to reason that we can think of the game in the following manner: prior to play, agents 1 and 2 have respective utilities u_1 and u_2 that they anticipate receiving. If agent 2 rejects agent 1's initial offer and agent 1 rejects agent 2's counteroffer, we have returned to a phase in which agent 1 is making a proposal. At this $t = 2$, we are again at a point where the game looks the same as at $t = 0$, hence *both players should anticipate obtaining utilities* u_1 *and* u_2, respectively; pulling back to $t = 1$, if agent 1 rejects agent 2's counteroffer, payoffs are $(\delta_1 u_1, \delta_2 u_2)$. This allows us to specify the game as in Figure 6.

We can now use our standard subgame-perfection logic to deduce what should happen. Supposing that agent

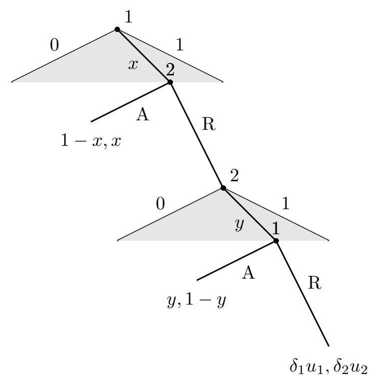

Figure 6: "Rubinstein bargaining," terminated ad hoc with presumed payoffs

2 makes offer y at time $t = 1$, agent 1 can obtain y today by accepting or u_1 tomorrow by rejecting, worth $\delta_1 u_1$ today. Thus she should accept 2's offer if $y \geq \delta_1 u_1$, and in turn agent 2 should offer $y^* = u_1$.

If agent 1 receives utility u_1, agent 2 must receive utility $u_2 = 1 - u_1$. Hence at time $t = 0$, agent 2 is choosing between accepting x today and rejecting in favor of receiving $1 - \delta_1 u_1$ tomorrow, worth $(1 - \delta_1 u_1)\delta_2$ today. It follows that agent 2 should accept agent 1's offer if $x \geq (1 - \delta_1 u_1)\delta_2$. Looking to minimize her offer to agent 2, agent 1 should offer agent 2 $x^* = (1 - \delta_1 u_1)\delta_2$; since agent 2 will accept such an offer, agent 1's payoff is $1 - (1 - \delta_1 u_1)\delta_2$.

By assumption, agent 1's payoff is u_1; hence we have

$$\begin{aligned} & & u_1 &= 1 - (1 - \delta_1 u_1)\delta_2 \\ &\iff & u_1 - \delta_1\delta_2 u_1 &= 1 - \delta_2 \\ &\iff & u_1 &= \frac{1 - \delta_2}{1 - \delta_1\delta_2} \end{aligned}$$

Since agent 2 immediately accepts agent 1's offer, agent 1's payoff must be $u_1 = 1 - x^*$ and agent 2's payoff must be $u_2 = 1 - u_1 = x^*$. Thus we know that agent 1 offers

$$x^* = 1 - \frac{1 - \delta_2}{1 - \delta_1\delta_2} = \frac{\delta_2(1 - \delta_1)}{1 - \delta_1\delta_2}$$

Agent 2 accepts immediately.

The leap we had to make here was presuming that players know the value of the game before they play; however, given an equilibrium players can always assess their expected utility from playing according to the specified strategies. Once we accept that players know the value of the game, we have all the information we need to find this particular equilibrium.

Question (easy): which player gets greater utility, the more- or less-patient player? Does it depend on who makes the first proposal?

Question (easy): what is the outcome if players are equally patient? Does this align with your answer to **Question (hard)** above?

Question (medium): considering the previous footnote, what is agent 1's optimal strategy if agent 2 threatens to offer 0 at every point in time? Can this constitute an equilibrium? If so, what are payoffs and why is this unappealing? What "nice" assumptions given in the footnote are violated?