

Reward and Punishment

on the equilibria of Repeated Games.

8 January, 2013

by Kevin Hasker

1 Introduction

Of all the topics I teach, I think repeated games terrify the students the most. I find this rather odd, in fact downright mysterious, because these are the types of strategies students use the most in their personal interactions. These strategies are all based on *reward* and *punishment*. If you do the right thing you will be rewarded (or get a high payoff) later on, if you do the wrong thing then you will be punished (by getting a low payoff.)

Example 1 *When you were a child, did your parents require you to be back home by, say, 9 PM? If they did why did you follow their rules? Because you were afraid of punishment, no? They wouldn't let you go out at all if you didn't get back on time tonight.*

Example 2 *Guys, let me make a recommendation to you, if you have been dating the same girl for a long time be SURE to remember your anniversaries. Now, I know this is a somewhat hopeless task (one of my friends was upbraided for not remembering the day of his engagement) but do try. Why? Because if you do she will reward you enormously.*

Example 3 *What would you do if you got food poisoning at a restaurant? Or even received really bad service or food? You don't go back. Obviously, you say? Well, yes, repeated game strategies are part of the fabric of our life, we all know how to use them instinctively. You will punish the restaurant by never going back there again.*

Example 4 *Do you know what my real goal as a teacher is? Primarily it is to encourage you to learn. I just want you to work at trying to learn the material I put before you. I don't, actually, think that we should be grading based on natural ability at all. A great student to me is someone who works on the material every week. I firmly believe that if students do then they will do well, even if they don't get an A.*

So you may ask me, why do I have frequent quizzes? The answer is that I find that without any incentive students just wait for the midterm to study. Unfortunately then they try to cram over a month's learning into their head in two days, and it absolutely doesn't work. The intention of the quizzes is to give you an incentive to study as you go along.

In fact the only reason that teachers grade at all it so transform class into a finitely repeated game. If you are good (study) in the early periods of the class, you will do well on the exams. If you are not, then you will do less well. Thus

we get you to behave well early in the game so that in the end you will get a high payoff. Reward and Punishment.

All of these examples show how rewards and punishments are part of the fabric of our social life. One of the terrifying things about repeated games is that there are so many possible strategies, even equilibrium strategies, but they all are based on one basic, fundamental, concept:

If you do what you are supposed to you will be rewarded. If you do not then you will be punished.

Now I will formally introduce a repeated game, discuss the equilibrium concept and equilibrium in general, and then show some examples.

2 A Repeated Game

A *Repeated Game* is a stage game, which is just some normal form game, that is repeated T times. T may be finite or infinite. An important part of this game is that in period $t \in (1, 2, 3, \dots, T)$ the players know what happened at every $t' < t$. I.e. everyone knows the past. If T is finite, then the value from the repeated game is the sum of the utilities in the stage games. If T is infinite then it is the sum of payoffs except that payoffs t periods in the future are discounted by δ^{t-1} where $0 \leq \delta < 1$.

To give a little more precision let $G = (I, A, u)$ be a normal form game. I is the number of players in the game (in every example it will be two), A_i is the set of possible actions for $i \in (1, 2, 3, \dots, I)$, and $A = \times_{i=1}^I A_i$ is the set of action profiles. Let me give several examples of stage games just to help you understand this concept.

| | | | |
|----------|---|----------|------|
| | | Player 2 | |
| | | C | D |
| Player 1 | C | 5; 5 | 1; 7 |
| | D | 7; 1 | 3; 3 |

| | | | |
|-----------------|---|---------------|-------|
| | | Customer (P2) | |
| | | B | N |
| Restaurant (P1) | H | 2; 2 | -1; 0 |
| | L | 3; -1 | 0; 0 |

E1: Prisoner's Dilemma

E2: Quality Game

| | | | | |
|----------|---|----------|------|------|
| | | Player 2 | | |
| | | C | D | N |
| Player 1 | C | 5; 5 | 1; 7 | 0; 0 |
| | D | 7; 1 | 3; 3 | 0; 0 |
| | N | 0; 0 | 0; 0 | 0; 0 |

| | | | | |
|----------|---|----------|------|-------|
| | | Player 2 | | |
| | | L | C | R |
| Player 1 | U | 5; 5 | 5; 4 | 1; 12 |
| | M | 3; 2 | 0; 0 | 2; 11 |
| | D | 10; 0 | 6; 1 | 0; 0 |

E1a: Prisoner's Dilemma with Outside Option

E3: Abstract Game

So that's several examples of games that we might want to discuss. The first game (*E1*) is the classic Prisoner's Dilemma, this is the classic problem where

following short term incentives (dominant strategies in the stage game) results in the socially undesirable equilibrium ((D, D) is the equilibrium strategy profile). The hope is that by looking at the repeated game we can get a more optimistic outcome. The second game ($E2$) is the restaurant quality game. The customer would like it if the restaurant produces high quality food (H) but this has a higher cost for the restaurant and the sales price can not depend on the quality (for example, food poisoning won't become obvious until after you have paid the bill.) Again the Nash equilibrium of the stage game is socially undesirable, the customer does not go to the restaurant because it expects the restaurant to produce low quality. The third game ($E1a$) is the Prisoner's Dilemma with an outside option, a relatively innocuous variation on the standard game that has the feature that the game now has two Nash equilibria. The fourth game ($E3$) is a random abstract game that I will use for examples.

Now I want to be more precise about the value of a given sequence of actions. Let $\vec{a} = \{a_t\}_{t=1}^T$ be a sequence of action profiles. Or in other words each a_t is what each person in the stage game does. For example in the Prisoner's Dilemma a_t is either (C, C) , (C, D) , (D, C) , or (D, D) . Then if T is finite:

$$V_i(\vec{a}) = \sum_{t=1}^T u_i(a_t)$$

if T is infinite then:

$$V_i(\vec{a}) = \sum_{t=1}^{\infty} \delta^{t-1} u_i(a_t)$$

Again, let me give a few examples: Let $T = 3$ and in game $E1$ consider $\vec{a} = ((C, D), (D, C), (C, C))$. This means that in period one the players play (C, D) , in period two they play (D, C) , and in period three they play (C, C) . This is not an equilibrium but I just want to show you how to calculate the payoffs.

$$\begin{aligned} V_1(\vec{a}) &= u_1(C, D) + u_1(D, C) + u_1(C, C) \\ &= 7 + 1 + 3 = 11 \\ V_2(\vec{a}) &= u_2(C, D) + u_2(D, C) + u_2(C, C) \\ &= 1 + 7 + 3 = 11 . \end{aligned}$$

Quite simple, no? Now let's deal with the more complicated case, when T is infinite. Consider the sequence $\vec{a} = \{(C, D) \text{ in every period}\}$. In order to

calculate this sequence we're going to need to use a cool math trick.

$$V_1(\vec{a}) = u_1(C, D) + \delta u_1(C, D) + \delta^2 u_1(C, D) + \delta^3 u_1(C, D) + \dots \quad (1)$$

$$(1 - \delta) V_1(\vec{a}) = (1 - \delta) (u_1(C, D) + \delta u_1(C, D) + \delta^2 u_1(C, D) + \dots) \quad (2)$$

$$= (u_1(C, D) - \delta u_1(C, D)) + (\delta u_1(C, D) - \delta^2 u_1(C, D)) + (\delta^2 u_1(C, D) - \delta^3 u_1(C, D)) + \dots \quad (3)$$

$$= u_1(C, D) + (-\delta u_1(C, D) + \delta u_1(C, D)) \quad (4)$$

$$+ (-\delta^2 u_1(C, D) + \delta^2 u_1(C, D)) \quad (5)$$

$$+ (-\delta^3 u_1(C, D) + \delta^3 u_1(C, D)) + \dots$$

$$(1 - \delta) V_1(\vec{a}) = u_1(C, D) \quad (6)$$

$$V_1(\vec{a}) = \frac{1}{1 - \delta} u_1(C, D) = \frac{1}{1 - \delta} 7. \quad (7)$$

Now let me go through the steps here one by one. First of all equation 1 is just the value of this sequence written in the most obvious form. In the next line, equation 2, I multiply both sides by $(1 - \delta)$. Then I expand this multiplication in equation 3, and then I regroup the terms in equation 4. Notice what happens, the $-\delta$ times $u_1(C, D)$ cancels out the second period's payoff of $\delta u_1(C, D)$. Multiplying $-\delta$ by the second period's payoff cancels out the third period's payoff $\delta^2 u_1(C, D)$, and so on and so forth. Wow! Cool math trick, no? While we have to discount the future in an infinitely repeated game in order to make sure that values aren't infinite, the reason we use this form of discounting is almost definitely because of this cool math trick. There are many reasons to think this is not appropriate, even time additive utility is rightfully criticized, but lets face it: it's convenient. The result is equation 6. This is great except that the left hand side is not the value, it is something multiplied by the value, so we divide both sides by $(1 - \delta)$ in order to get the final result, which is equation 7.

There is another way to prove this, which will be of use in the future. This is the value function method, obviously given the path we are analyzing $V_1(\vec{a})$ will be the same no matter what period we do the calculation in, because the future will always be the same. Thus:

$$V_1(\vec{a}) = u_1(C, D) + \delta V_1(\vec{a})$$

$$(1 - \delta) V_1(\vec{a}) = u_1(C, D)$$

$$V_1(\vec{a}) = \frac{1}{1 - \delta} u_1(C, D) .$$

Quite a bit simpler, but it relies on properties of the value function and thus should be thought of as a second order method.

I did this using $u_1(C, D)$ just to emphasize the point that what the utility of the action pair is does not matter. Whatever it is the value of an infinite sequence of that action pair will just be the utility of that action pair divided by $1 - \delta$. From this work you can immediately see that:

$$V_2(\vec{a}) = \frac{1}{1 - \delta} u_2(C, D) = \frac{1}{1 - \delta} .$$

3 Equilibrium: a Discussion and the Super Strategies.

3.1 Definition

First of all it should be obvious that the equilibrium concept we want to use in these games is subgame perfect equilibrium. Let me give a definition as it applies to repeated games.

Definition 5 *A subgame perfect equilibrium in a repeated game is a strategy that is optimal from period t on no matter what has occurred in any period $t' < t$.*

Why is this such a good idea here? Well, what are we going to do? Basically we are going to use threats to keep people on their good behavior. Now, should we accept that something is an equilibrium if the threats are not believable? Let's consider... Say that your girlfriend told you that if you did not open every door for her she was going to kill herself. What would you do? I would either: a) laugh in her face, or b) check her into a loony bin (psychiatric hospital). Rewards and Punishments are going to be critical to our analysis. If the Rewards and Punishments are not credible then exactly why should we pay attention to them? If I told you I would give you a million TL if you got an A in my class, would that make you study harder? No, because frankly I don't have a million lira—and wouldn't give it to you if I did.

So basically this means that our rewards and punishments also have to be equilibria. If I promise you something I have to be willing and able to deliver on it. How are we going to do this? Well first of all our punishments will always be Nash equilibria of the stage game. That way all I have to pay attention to is my short term incentives, and in a Nash equilibrium these short term incentives are always aligned properly. In finitely repeated games are rewards are also going to have to be (different) Nash equilibria, again because that way our short term incentives are enough to guarantee that the reward is delivered.

Now a second issue, how many subgames will there be? To understand this issue it is best to think about the infinitely repeated game. Say that after two histories I face the same stream of action profiles in the future. In other words the future is the same after these two histories, but the past is different. Do I need to analyze both situations separately? No, because if the strategy works in one of them it must work in the other one. Thus the number of relevant subgames is the number of possible futures. Now remember that our punishments and rewards will all be static Nash equilibria, and in the super strategies once you start playing one of them you will never change what you are doing. This makes proving the strategy is subgame perfect in these subgames rather trivial. I myself will not always do it below, but on a quiz or exam you will always get points for doing it.

The formal proof is:

1. What happens in the future will not be affected by what happens today.

2. What happens today is a static Nash equilibrium.

Thus each player's best response is given by the static game, and everyone is behaving optimally. The reason I give points for this relatively trivial proof is because there can be situations (which we will discuss below) where these characteristics are not met, and this can make the strategy not an equilibrium.

A final thing which I will never make you prove but you should know about is the *one shot deviation principle*. This states that something is a Subgame Perfect equilibrium if and only if it is impervious to all one shot deviations—in other words changing the strategy in one period alone. This statement strongly relies on checking every subgame—which will be trivial with the super strategies but may not be in general.

3.2 Discussion

Now, on a fundamental level, what is it about repeated games that so confuses and frightens the students? After thinking about it for some time I think it is the fact that the equilibria we find are not, honestly, that closely related to the payoffs of the stage game. In all previous analysis on a basic level the payoffs of people have determined what can and can not be an equilibrium. This is utterly false in the repeated game, in the repeated game the entire point is to overcome these short term incentives to achieve long run goals. I.e. the short term payoffs do not matter! This is fairly disturbing. To understand repeated game equilibria you really need to think about one of the key alternative definitions of a Nash equilibrium. While I say Nash equilibrium it is really just any equilibrium concept.

Definition 6 *A (Nash) equilibrium is a self-enforcing social convention. It is something that you do because of what you expect others to do.*

The clearest and simplest example of this is the equilibria of the side of the road game.

| | | | |
|----------|---|----------|---------|
| | | Driver 2 | |
| | | L | R |
| Driver 1 | L | 1; 1 | -9; -10 |
| | R | -10; -9 | 0; 0 |

E4 : The Side of the Road Game

This game has two Nash equilibria, (L, L) and (R, R) . If you expect other driver's to drive on the right hand side, you want to drive on the right hand side too. If you expect them to drive on the left, then you want to drive on the left as well. Clearly you don't need to know what the law is in a given country to know which side of the road to drive on—all you need is to know what others are doing. The reason there is a law in most countries is merely to make it easy to know the equilibrium. If you go to England you know automatically to drive

on the left because it is the law, if you go the United States you know to drive on the right for the same reason.

Equilibria of repeated games are like this with a vengeance. Why do men usually open the door for women? Because if they do not the women they are with, and perhaps their male friends as well, will give them a hard time about their bad manners. That expectation of a small negative payoff is enough to cause most men to make the small exertion necessary to open the door. In the United States some women actually feel that it is an insult for men to open the door for them. They hold that it is treating them as unequal, and they want equality in all things. So if you interact long enough with those women then you would learn not to open doors. In other words, your behavior depends on society around you. If society expects one behavior then you will exhibit that behavior. If society expects a different one then you will follow the rules and exhibit whatever behavior society expects.

3.2.1 The Super Strategies

However at the same time this is a curse, it is also our savior. Because we are looking for equilibria that do not depend on the game we basically are looking for a one-size-fits-all solution. A strategy that will work no matter what the game is. And, in fact, we can find one such strategy that will work in finitely repeated games and one strategy that will work in infinitely repeated games. These are also the best strategies in many games, in other words if anything is an equilibrium one of these strategies is.

Now these strategy will require two, or maybe three, strategy profiles from the stage game. The first one is what we want these people to do, we write this as c . The second one is our punishment strategy, or p . The third (which we only need for finitely repeated games) is our reward strategy, we write this as r . Both r and p have to be Nash equilibria of the stage game. p should be one of the Pareto Worse equilibria (i.e. no other Nash equilibrium gives lower payoffs to both people). r should be one of the Pareto Efficient equilibria (i.e. no other Nash equilibrium gives higher payoffs to both people).

For example, in game $E1a$ (Prisoner's Dilemma with Outside Option) the obvious candidate for c is (C, C) . The punishment, p , should be (N, N) , since this has the worst payoff for both parties. The reward, r , should be (D, D) —I know that seems a little strange but it is the Pareto Efficient Nash equilibrium. Now for the strategies.

Finitely Repeated Games Super Strategy (Mark 1) .

1. In period 1 play c .
2. In period $t > 1$ if in period 1 people played c play r .
3. In period $t > 1$ otherwise play p .

So, in the first period we want you to do what society wants (play c). If you do that you are rewarded by playing r , if you do not then you are punished by

playing p . Why do you play c ? Because if you do not then you get punished, and if you do you get rewarded.

The surprising thing is that this strategy is almost the same in the infinitely repeated game. The only difference is now that cooperation (playing c) is its own reward. That is the only difference.

Infinitely Repeated Games Super Strategy (Mark 1) .

1. In period 1 play c .
2. In period $t > 1$ If you played c last period play c in this period.
3. In period $t > 1$ otherwise play p .

These strategies are not quite as general as I would like you to understand. To see the completely generalized strategies look in Section REF below. However these strategies are generally enough, and the more general strategies only are needed in cases where the stage game is not "cooperative." An example of this is game $E3$. Neither (D, C) or (M, R) is a punishment for both people, if we want people to play (T, L) then we need to use different punishments for different people. (For now don't worry about the reward will be.) Not much of a generalization, but a generalization that can strongly increase the power of what we can do. These are the Mark 2 strategies. In these games you play p^i if i deviated, where again p^i is a Nash equilibrium.

Finitely Repeated Games Super Strategy (Mark 2) .

1. In period 1 play c .
2. In period $t > 1$ if in period 1 c occurred play r .
3. In period $t > 1$ if player 1 was the first to deviate, play p^1 .
4. In period $t > 1$ otherwise play p^2 .

The equivalent Infinitely repeated game again just dispenses with playing r .

Infinitely Repeated Games Super Strategy (Mark 2) .

1. In period 1 play c .
2. In period $t > 1$ If you played c last period play c in this period.
3. In period $t > 1$ if player 1 was the first to deviate, play p^1 .
4. In period $t > 1$ otherwise play p^2 .

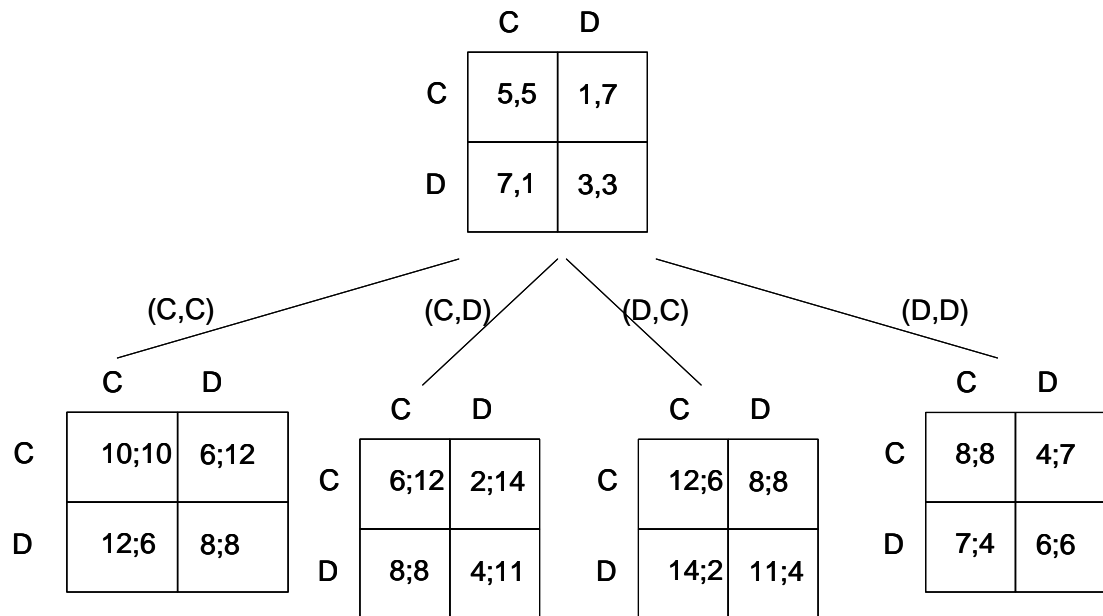
This simple modification can greatly increase the strength of our analysis, but it won't always. It is needed only if there is no unique worst Nash equilibrium, like game $E3$ above.

4 Two Period (Finite) Repeated Games.

Everything you need to know about finite repeated games can be learned from two period games. The only difference is that with more than two periods you can have bigger rewards and punishments. So let's look at the two period repeated game. To do this I want to use a short hand extensive form game tree. Each "decision node" in this tree is the stage game. Each "action" in this tree is an action profile (an action for each person).

Using this type of graph also makes it absolutely clear what a subgame perfect equilibrium is. In a sequential game it is an optimal decision at each decision node, in a repeated game it is an optimal decision in each stage game. The critical difference is that in a sequential game there is almost always only one optimal decision, in a repeated game there can be multiple Nash equilibria in a stage game—thus there can be more than one optimal decision.

Let's actually physically draw one of these game trees for the Prisoner's Dilemma (game *E1*).



Before analyzing this game let me just point out how complex even a two period finitely repeated game is. How many strategies are there in this game? Well each person has two actions in the first period, then depending on the action profile that happens in the first period they have two actions in the second period. There are four different possible action profiles in the first period, thus in the second period they have $2 * 4 = 8$ actions that need to be specified in a strategy, each of these can be paired with a different first period action, giving a total of 16 strategies for each player. OK, that's a pretty big number, but an equilibrium specifies a strategy for each player, so it is a strategy profile.

How many strategy profiles are there? Well with 16 strategies per player there are $16 * 16 = 256$ strategy profiles! And it will just get worse if we look at games with more than two actions, or more than two periods. Obviously this is beyond analyzing with any standard techniques, we will not be able to find the equilibrium by (for example) looking at a normal form game representation.

But we don't have to, we can use backward induction here and hopefully that will make things much simpler (just like in the sequential game.) So what can be the equilibrium in the second stage? You can easily check that no matter what happens in the first period D is still a dominant action for each player in the second period. This is because the optimal strategy does not depend on the absolute payoffs, rather the payoff differences, and these will always be determined only by the payoffs of the stage game. So *no matter what happens in the first period* the equilibrium outcome in the second period will be (D, D) . OK, now we go back to the first period, and like when we were using backward induction in sequential games we can combine the payoffs from the second stage into the first period's payoffs.

| | | | |
|----------|---|----------|-------|
| | | Player 2 | |
| | | C | D |
| Player 1 | C | 8; 8 | 3; 10 |
| | D | 10; 3 | 6; 6 |

Again you can quickly determine that the dominant strategy equilibrium is (D, D) . So the only subgame perfect equilibrium is: *play (D, D) no matter what happens.*

Weird. What happened? We analyzed the repeated game and gosh, it just turned into the static game—except now it happened more than once. What then is the point of analyzing the repeated game? Well, this is something of a strange result. It is not general, but understanding why it happens is critical to understanding how we *might* achieve something different in the repeated game.

So let's go through the logic, and think about things. First of all, obviously in the final period whatever happens must be a Nash equilibrium. Why is that? Well because there is no future so only short term incentives have any effect—thus we must be playing a Nash equilibrium. Now what if in the stage game there is only one Nash equilibrium? That means that *the same thing will happen in the final period no matter what happens earlier.* In other words the future is fixed, it is written in stone. In period two the (unique) Nash equilibrium will occur. So, now the future is fixed. No matter what happens in the first period what will happen in the second period is known. So what incentives should one pay attention to in the first period? Obviously the only incentives that are left are the incentives given by the stage game. The short term incentives must again dominate, and according to the short term incentives you only can do one thing, play the Nash equilibrium of the stage game.

Proposition 7 *In a finitely repeated game, if the stage game has only one Nash equilibrium then the only subgame perfect equilibrium of the repeated game is to always play that Nash equilibrium.*

Proof. (Two Period) In the second period no matter what has happened in the past the only incentives player have to respond to are their short term, stage game, incentives. Thus they must play a Nash equilibrium. If there is only one Nash equilibrium then they must play this Nash equilibrium no matter what happened in the first period.

Thus in the first period the future is fixed, no matter what happens in the first period in the second period both players will receive their Nash equilibrium payoffs. Thus the only incentives they have to pay attention to are their short term incentives, and the only thing to do is to play the Nash equilibrium of the stage game.

(Finite Periods) In the last period, T , like above they must play the Nash equilibrium of the stage game. In period $T - 1$ since the future is fixed again they must play the Nash equilibrium of the stage game, thus by induction they must always play the Nash equilibrium of the stage game. ■

Weird, huh? And isn't it fortunate that so many of the games we have looked at have only one Nash equilibrium. That's great, so now we can throw repeated games into the trash and go back to more interesting topics.

Yea, you wish. This argument depends critically on their only being one Nash equilibrium in the stage game, and honestly that is something of a freak occurrence. I know that a lot of the games I have shown you have only one Nash equilibrium, and that was on purpose. I wanted to simplify your life and my own. Frankly finding a game with only one Nash equilibrium in it is either due to: a) luck, b) the modelling decisions of the theorist. For example, you will agree that in most interactions you have the outside option of not interacting, right? This is a fundamental choice and it is a rare interaction where this is absolutely impossible. It may be undesirable, but it is possible. Now we go to game $E1a$, which I will rewrite below for clarity:

| | | Player 2 | | |
|----------|---|----------|------|------|
| | | C | D | N |
| Player 1 | C | 5; 5 | 1; 7 | 0; 0 |
| | D | 7; 1 | 3; 3 | 0; 0 |
| | N | 0; 0 | 0; 0 | 0; 0 |

E1a: Prisoner's Dilemma with
Outside Option

Notice that the outside option is very undesirable, even if the other player cheats you (plays D when you play C) it is better to interact. But our analysis will change dramatically when we consider the game with this outside option, because obviously (N, N) is a Nash equilibrium.

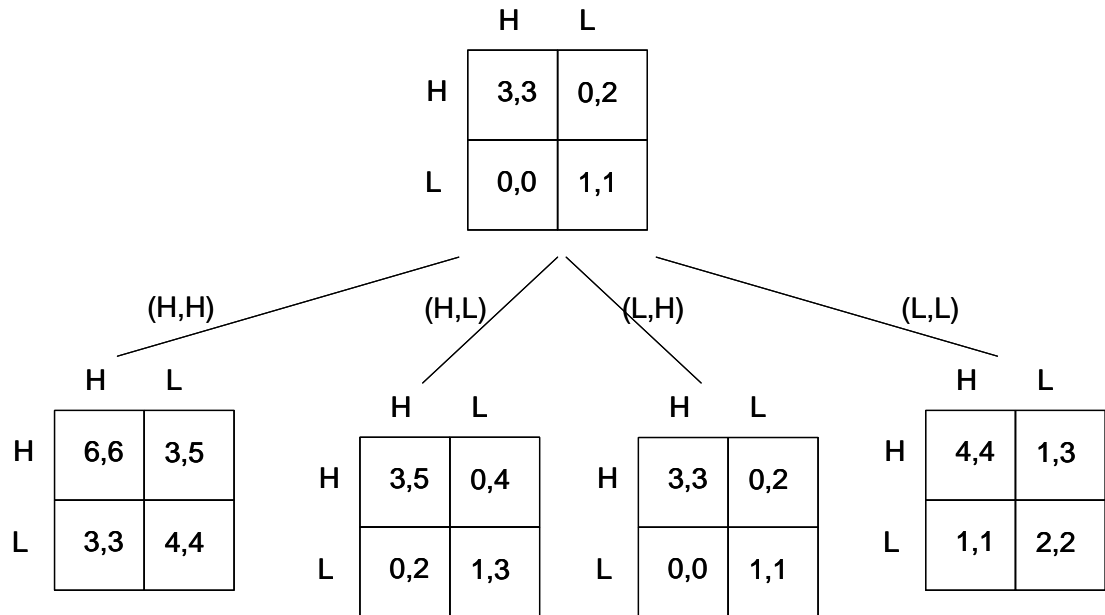
To understand this point lets first look at a game where we can actually look at the complete two period game without going crazy. In game $E1a$ there are nine possible action profiles in the first period. Do you really want to try and analyze a game with nine subgames? So let's look at a game with only two

actions for both players, the *High/Low Game*:

| | | | |
|----------|---|----------|------|
| | | Player 2 | |
| | | H | L |
| Player 1 | H | 3; 3 | 0; 2 |
| | L | 0; 0 | 1; 1 |

E4: The High/Low Game

In this game the Pareto Dominant payoff is also a Nash equilibrium of the stage game, so the optimal equilibrium from societies point of view is obviously to always play (H, H) . However there are other equilibria. Let's look at the two period repeated game:



What is the key difference here? Well, now, what happens in the second period *can* depend on what happens in the first period. If they do the "right thing" we play the (H, H) Nash equilibrium. If they do the "wrong thing" we play the (L, L) Nash equilibrium. So, we can use the Super Strategy (Mark 1) and we might be able to get something other than Nash equilibrium behavior in the first period.

- Finitely Repeated Games Super Strategy (Mark 1)**
1. In period 1 play c .
 2. In period 2 if in period 1 people played c play (H, H) .
 3. In period 2 otherwise play (L, L) .

Obviously this will work if c is either (H, H) or (L, L) , so let's look at the other options. What if c is (H, L) ? Can we get cooperation on playing (H, L) in the first period? What this effectively does is means that in the first period you play (H, L) you get an extra 3 points from playing (H, H) in the second period. If you play anything else you only get 1 point in the second period. So with these second period strategies the first period payoffs are:

| | | | | | | | |
|----------|---|--------------|--------------|---|----------|----------|------|
| | | Player 2 | | | | Player 2 | |
| | | H | L | = | | H | L |
| Player 1 | H | 3 + 1; 3 + 1 | 0 + 3; 2 + 3 | | Player 1 | 4; 4 | 3; 5 |
| | L | 0 + 1; 0 + 1 | 1 + 1; 1 + 1 | | | 1; 1 | 2; 2 |

And thus now H is the best response to L and L is the best response to H . In other words, yes we can. If we play (H, H) in the second period *only* when we play (H, L) in the first period then it is best to play (H, L) in the first period. This is an equilibrium.

Well, that methodology is fine, and for two period games it can generally be applied without any concerns, but it is a little awkward. The standard methodology is to compare the value of the "equilibrium" payoffs with the value of the best deviation.

$$\begin{aligned} V_1^* &= u_1(H, L) + u_1(H, H) \\ \hat{V}_1 &= u_1(L, L) + u_1(L, L) \end{aligned}$$

There are several ways to proceed here. The easiest is just to do the calculations, I call this the *value function method*:

$$\begin{aligned} V_1^* &= 0 + 3 = 3 \\ \hat{V}_1 &= 1 + 1 = 2 \end{aligned}$$

and verify that $V_1^* \geq \hat{V}_1$ so this strategy is optimal for player 1. But you can also do a little bit of analysis to make the benefits and costs of deviating more obvious. I call this the *difference method*.

$$\begin{aligned} V_1^* &\geq \hat{V}_1 \\ u_1(H, L) + u_1(H, H) &\geq u_1(L, L) + u_1(L, L) \\ u_1(H, H) - u_1(L, L) &\geq u_1(L, L) - u_1(H, L) \\ 2 &= 3 - 1 \geq 1 - 0 = 1 \end{aligned}$$

In this formulation, $u_1(L, L) - u_1(H, L)$ is *benefit of deviating*—the amount the short term payoffs have increased—while $u_1(H, H) - u_1(L, L)$ is the *cost of deviating*—the amount that this player will lose in the future.

We obviously also need to check this for player 2.

$$\begin{aligned} V_2^* &= u_2(H, L) + u_2(H, H) = 2 + 3 = 5 \\ \hat{V}_2 &= u_2(H, H) + u_2(L, L) = 3 + 1 = 4 \end{aligned}$$

$$\begin{aligned}
V_2^* &\geq \hat{V}_2 \\
u_2(H, L) + u_2(H, H) &\geq u_2(H, H) + u_2(L, L) \\
u_2(H, H) - u_2(L, L) &\geq u_2(H, H) - u_2(H, L) \\
2 = 3 - 1 &\geq 3 - 2 = 1
\end{aligned}$$

and using either method we arrive at the right answer, this strategy is also optimal for player 2, so it is an equilibrium.

Now, what about the strategy: (L, H) in the first period, if (L, H) in the first period (H, H) in the second, otherwise (L, L) . First let's do the normal form game analysis of this strategy:

| | | | | | | | |
|----------|---|--------------|--------------|---|----------|----------|------|
| | | Player 2 | | | | Player 2 | |
| | | H | L | = | | H | L |
| Player 1 | H | 3 + 1; 3 + 1 | 0 + 1; 2 + 1 | | Player 1 | 4; 4 | 1; 3 |
| | L | 0 + 3; 0 + 3 | 1 + 1; 1 + 1 | | | 3; 3 | 2; 2 |

Obviously this is not an equilibrium, player 1 wants to deviate to H . The value function method:

$$\begin{aligned}
V_1^* &= u_1(L, H) + u_1(H, H) = 0 + 3 = 3 \\
\hat{V}_1 &= u_1(H, H) + u_1(L, L) = 3 + 1 = 4,
\end{aligned}$$

the difference method:

$$\begin{aligned}
V_1^* &< \hat{V}_1 \\
u_1(L, H) + u_1(H, H) &< u_1(H, H) + u_1(L, L) \\
u_1(H, H) - u_1(L, L) &< u_1(H, H) - u_1(L, H) \\
2 = 3 - 1 &< 3 - 0 = 3
\end{aligned}$$

since the last line is true we know this is not an equilibrium. Notice one of the benefits of the difference method. The left hand side is the same for this strategy and the strategy " (H, L) in the first period, if (L, H) in the first period (H, H) in the second, otherwise (L, L) ." So all I need to do is analyze the right hand side, the benefit of deviating. This is one of the advantages of the difference method. By doing the calculations once I can see that the cost of deviating for both players is 2, thus if the benefit of deviating for a given strategy pair is lower than that it can be an equilibrium.

By the way, a word about complexity. Above I mentioned that there are 256 strategy profiles in this game (it has the same number of actions as the Prisoner's Dilemma, and thus the same number of strategies.) How many of them do you think are equilibria? Well first of all in an equilibrium what is expected to be played in the second period must be the same for both parties, thus there are 16 possible equilibrium action profiles in the second period, combined with the 4 possible action profiles for the first period we (only) have to consider $4 * 16 = 64$ strategy profiles. Furthermore we know that there are only two possible equilibrium action profiles for the second period, which can depend on

the actions in the first period, thus this means there are only $4 * 4 * 2 = 32$ possibilities that are really viable. If (H, H) is prescribed for the first period, then we can assign (H, H) or (L, L) to any of the four action profiles, thus there are 8 equilibrium strategy profiles with (H, H) being played in the first period. If (L, L) is prescribed in the first period and (H, H) in the second if (L, L) occurs then we can again assign the other three profiles at random, giving us another 6 equilibria. If (L, L) is expected to be played in both periods then we have to assign (L, L) if (H, L) or (L, H) occurs, thus we only have two equilibria like this. the same analysis holds if we expect (H, L) in the first period (note that this time we have to have (L, L) if (H, H) or (L, L) occurs), thus we have two more. So what's the total? Of the 32 strategy profiles that we recognize might be subgame perfect Equilibria 18 of them actually are equilibria.

Just to understand how complicated this would be in general, consider an arbitrary stage game with n static Nash equilibria. Then we can easily see that in the T period repeated game there are at least n^T subgame perfect equilibria—and that is only analyzing strategies where what happens depends only on the period.

This is the primary reason we use "super strategies." Obviously if I said "find all the equilibria" then the answer would be impossible to find.¹ All I can really ask is "in the T period finitely repeated game what action profiles can be a subgame perfect outcomes in the first period?"

Now let's return to a game we care about, the Prisoner's Dilemma with the outside option. The Super Strategy is:

Finitely Repeated Games Super Strategy (Mark 1) The strategy is:

1. In period 1 play c .
2. In period 2 if in period 1 people played c play (D, D) .
3. In period 2 otherwise play (N, N) .

Can we have $c = (C, C)$?

$$\begin{aligned} V_1^* &= u_1(C, C) + u_1(D, D) \\ \hat{V}_1 &= u_1(D, C) + u_1(N, N) \end{aligned}$$

$$V_1^* \geq \hat{V}_1$$

$$u_1(D, D) - u_1(N, N) = 3 - 0 = 3 \geq 2 = 7 - 5 = u_1(D, C) - u_1(C, C)$$

Obviously since this game is symmetric we don't need to check for person 2 independently, so yes this is an equilibrium.

What else can we support? Well we need a strategy pair where the benefit to deviating for both parties is less than 3. Obviously (D, D) or (N, N) will work since these are Nash equilibria of the stage game. (C, D) will work because

¹Or trivial, if the stage game has only one Nash equilibrium.

player 2 is best responding and player 1's benefit from deviating is only $2 = u_1(D, D) - u_1(C, D)$. Therefore obviously so will (D, C) . What about (C, N) or (N, C) ? Well consider (N, C) , obviously player 2 has no benefit from deviating, but player 1 can get 7 by playing D instead of N ($7 = u_1(D, C) - u_1(N, C)$). Thus neither of these are equilibria. What about (D, N) or (N, D) ? Like with (N, C) if they are supposed to play (N, D) then player 1's short run best response is to play D , this gives him a benefit to deviating of 3, which is exactly the cost. This is an equilibrium, but only weakly.

In other words, with the rather minor alteration to the Prisoner's Dilemma of allowing people not to interact we went from having one equilibrium to having many. We most certainly can get the socially appealing (C, C) , we can also get other things.

5 Infinitely Repeated Games

In an infinitely repeated game we no longer have a last period. Why does this matter? Because we can't use backward induction. But this problem also makes our lives easier, no final period means that cooperating can be its own reward. All we need is a punishment. Thus we can finally look at two of the most important games and find out when and how cooperation can be achieved. I will copy them here for your convenience.

| | | | | | | | |
|----------|---|----------|------|-----------------|---|---------------|-------|
| | | Player 2 | | | | Customer (P2) | |
| | | C | D | | | B | N |
| Player 1 | C | 5; 5 | 1; 7 | Restaurant (P1) | H | 2; 2 | -1; 0 |
| | D | 7; 1 | 3; 3 | | L | 3; -1 | 0; 0 |

E1: Prisoner's Dilemma

E2: Quality Game

Before we analyze these games let us first talk about δ . All of the results in infinitely repeated game are going to be "for δ close enough to 1 this is an equilibrium." So what exactly is δ ? Well there are several important ways to interpret it.

Interpretation of δ δ can be thought of as a combination of:

- Patience—if people are more patient they discount the future less, the future is relatively more important. So as patience increases δ should increase.

As an example if we are analyzing a firm then the proper value for δ is $\frac{1}{1+r}$, where r is the real interest rate. If r decreases this means that the real interest rate has fallen. The opportunity cost of investing money is the interest you could get from holding it for a period, so if r falls it is less important whether you get money today or tomorrow. In other words the firm is more patient.

- Frequency of Interaction—the more frequently you interact the more patient you are going to be. It is like the difference between being paid in a day and a year. If someone tells you that they will pay you tomorrow you probably won't care, but if they start talking about paying you back in a year you probably start demanding interest—more money to make up for the delay.
- Probability of Interacting again—instead of saying that you are going to interact an infinite number of times, say that you will interact again with probability δ . With probability $(1 - \delta)$ the interaction will terminate and you will never interact again. This causes a re-normalization of values, but essentially it is as if you are playing an infinitely repeated game. (This trick is often used by experimenters who want to analyze behavior in infinitely repeated games.)

Now let's show that we can finally get cooperating in these two paradigmatic games. Let us first focus on $E1$, the Prisoner's Dilemma. The Super Strategy for this game is:

- Infinitely Repeated Games Super Strategy (Mark 1)**
1. In period 1 play (C, C) .
 2. In period $t > 1$ If you played (C, C) last period play (C, C) in this period.
 3. In period $t > 1$ otherwise play (D, D) .

First of all let me point out that this is a very harsh strategy, another way of writing the same strategy is:

- Infinitely Repeated Games Super Strategy (Mark 1)**
1. In period 1 play (C, C) .
 2. In period $t > 1$ if there was (C, C) every period in the past, play (C, C) in this period.
 3. In period $t > 1$ otherwise play (D, D) .

Basically one mistake unleashes an infinite punishment. For this reason it is called the *trigger* strategy, it is like both parties are being held up at gunpoint. One mistake and they're dead. It is also called the *Grimm* strategy, after the Brother's Grimm, a pair of Germans who are famous for collecting and writing down fairy tales. It is supposed to be a "folk wisdom" strategy. There is no doubt that it is a very harsh strategy, and one that is not often used in practice. However it does have the benefit that it is easy to analyze, and that often it is the "best" strategy. In other words if this strategy is not an equilibrium then nothing is. So when will it be an equilibrium? First of all, the trivial case where you expect (D, D) forever. In this case since the current period's action will not affect the future and it is a Nash equilibrium, everyone will follow the strategy in this subgame. Now let $\overrightarrow{(C, C)}$ be "play (C, C) every period in the

future)" and $\overrightarrow{(D, D)}$ be "play (D, D) every period in the future." Then using what I showed above you can find out that:

$$\begin{aligned} V_1 \left(\overrightarrow{(C, C)} \right) &= \frac{1}{1-\delta} u_1(C, C) = \frac{1}{1-\delta} 5 \\ V_1 \left(\overrightarrow{(D, D)} \right) &= \frac{1}{1-\delta} u_1(D, D) = \frac{1}{1-\delta} 3 \end{aligned}$$

Thus

$$\begin{aligned} V_1^* &= u_1(C, C) + \delta V_1 \left(\overrightarrow{(C, C)} \right) \\ &= u_1(C, C) + \delta \left(\frac{1}{1-\delta} u_1(C, C) \right) \\ &= u_1(C, C) + \frac{\delta}{1-\delta} u_1(C, C) \\ \hat{V}_1 &= u_1(D, C) + \delta V_1 \left(\overrightarrow{(D, D)} \right) \\ &= u_1(D, C) + \frac{\delta}{1-\delta} u_1(D, D) \end{aligned}$$

Like before you can use the value function method to see whether this is an equilibrium, but I think it is better to use the difference method.

$$\begin{aligned} V_1^* &\geq \hat{V}_1 \\ u_1(C, C) + \frac{\delta}{1-\delta} u_1(C, C) &\geq u_1(D, C) + \frac{\delta}{1-\delta} u_1(D, D) \\ \frac{\delta}{1-\delta} [u_1(C, C) - u_1(D, D)] &\geq u_1(D, C) - u_1(C, C) \end{aligned}$$

and like before we can see that what matters is whether the benefit of deviating (the right hand side) is less than the cost (the left hand side). Notice that the left hand side is multiplied by $\frac{\delta}{1-\delta}$, thus as $\delta \rightarrow 1$ $\frac{\delta}{1-\delta} \rightarrow \infty$ and the left hand side will surely be bigger than the right hand side no matter how small the difference in payoffs between (C, C) and (D, D) . In this case δ can be really rather small.

$$\begin{aligned} \frac{\delta}{1-\delta} (5-3) &\geq 7-5 \\ \frac{\delta}{1-\delta} 2 &\geq 2 \\ \delta 2 &\geq 2(1-\delta) \\ \delta 2 &\geq 2-2\delta \\ \delta 4 &\geq 2 \\ \delta &\geq \frac{1}{2} \end{aligned}$$

Since this equilibrium is symmetric we don't need to independently check player 2, thus if $\delta \geq \frac{1}{2}$ then this is an equilibrium.

Let us do the same analysis in the quality game. The strategy we want to support is:

Infinitely Repeated Games Super Strategy (Mark 1) 1. In period 1 play (H, B) .

2. In period $t > 1$ If you played (H, B) last period play (H, B) in this period.

3. In period $t > 1$ otherwise play (L, N) .

We will first concern ourselves with player 1, the restaurant.

$$\begin{aligned} V_1^* &= u_1(H, B) + \delta V_1(\overrightarrow{(H, B)}) \\ V_1^* &= u_1(H, B) + \frac{\delta}{1-\delta} u_1(H, B) \\ \hat{V}_1 &= u_1(L, B) + \delta V_1(\overrightarrow{(L, N)}) \\ \hat{V}_1 &= u_1(L, B) + \frac{\delta}{1-\delta} u_1(L, N) \end{aligned}$$

$$\begin{aligned} V_1^* &\geq \hat{V}_1 \\ u_1(H, B) + \frac{\delta}{1-\delta} u_1(H, B) &\geq u_1(L, B) + \frac{\delta}{1-\delta} u_1(L, N) \\ \frac{\delta}{1-\delta} (u_1(H, B) - u_1(L, N)) &\geq u_1(L, B) - u_1(H, B) \\ \frac{\delta}{1-\delta} (2-0) &\geq 3-2 \\ \delta 2 &\geq 1-\delta \\ \delta 3 &\geq 1, \delta \geq \frac{1}{3} \end{aligned}$$

So this is optimal for player 1 if $\delta \geq \frac{1}{3}$. Now we also need to check the incentives for person 2.

$$\begin{aligned} V_2^* &= u_2(H, B) + \frac{\delta}{1-\delta} u_2(H, B) \\ \hat{V}_2 &= u_2(H, N) + \frac{\delta}{1-\delta} u_2(L, N) \\ u_2(H, B) + \frac{\delta}{1-\delta} u_2(H, B) &\geq u_2(H, N) + \frac{\delta}{1-\delta} u_2(L, N) \\ \frac{\delta}{1-\delta} (u_2(H, B) - u_2(L, N)) &\geq u_2(H, N) - u_2(H, B) \\ \frac{\delta}{1-\delta} (2-0) &\geq 0-2. \end{aligned}$$

Whoops, don't I feel stupid now, the right hand side is negative and the left hand side is positive, so this strategy is optimal for player 2 as long as $\delta \geq 0$. This is actually fairly obvious because a) he is always best responding to his opponents' action and b) he is better off at (H, B) than at (L, N) . (Notice we need both.) So this is an equilibrium if $\delta \geq \frac{1}{3}$.

5.0.2 A Forgiving Strategy:

The restaurant quality game is a good time to investigate an alternative strategy. One that does not always work but is more forgiving than the trigger strategy. So consider the following strategy:

- A forgiving strategy**
1. In period 1 play (H, B) .
 2. In period $t > 1$ If the strategy was followed last period play (H, B) in this period.
 3. In period $t > 1$ otherwise play (L, N) .

Now this strategy looks more complicated than the previous one, but it really isn't. Basically: "if this restaurant provides low quality I won't go back there once." This is, in reality, what we usually do with restaurants. If the problem was simply that the service or food was terrible we don't actually stay away from the restaurant forever. In fact even if we say we will we often forget it, and go back after a while. This is also probably the type of punishments your parents used on you. It certainly is the type I use on my children. Right now generally the worst punishment I ever impose is making a kid sit in his room for five minutes or not watch TV for a day.

In general, however, it is hard to establish when these strategies are equilibria, but for this one it is not too hard.

$$\begin{aligned} V_1^* &= u_1(H, B) + \delta u_1(H, B) + \delta^2 V_1 \left(\overrightarrow{(H, B)} \right) \\ \hat{V}_1 &= u_1(L, B) + \delta u_1(L, N) + \delta^2 V_1 \left(\overrightarrow{(H, B)} \right) \end{aligned}$$

When is this an equilibrium? When $V_1^* \geq \hat{V}_1$ or $V_1^* - \hat{V}_1 \geq 0$

$$\begin{aligned} V_1^* - \hat{V}_1 &= u_1(H, B) + \delta u_1(H, B) - (u_1(L, B) + \delta u_1(L, N)) \\ &= 2 + \delta 2 - (3 + \delta * 0) = 2\delta - 1 \geq 0 \\ \delta &\geq \frac{1}{2} \end{aligned}$$

so in this game, since the benefit of deviating is so small, going back once will work. Notice that the restaurant (player 1) has to be more patient, however. Also notice that if we even consider 2 period punishments then the math gets

insanely complicated:

$$\begin{aligned}
 V_1^* &= u_1(H, B) + \delta u_1(H, B) + \delta^2 u_1(H, B) + \delta^3 V_1 \left(\overrightarrow{(H, B)} \right) \\
 \hat{V}_1 &= u_1(L, B) + \delta u_1(L, N) + \delta^2 u_1(L, N) + \delta^3 V_1 \left(\overrightarrow{(H, B)} \right) \\
 V_1^* - \hat{V}_1 &= u_1(H, B) + \delta u_1(H, B) + \delta^2 u_1(H, B) - (u_1(L, B) + \delta u_1(L, N) + \delta^2 u_1(L, N)) \\
 &= 2 + \delta 2 + \delta^2 2 - (3 + \delta * 0 + \delta^2 * 0) = 2\delta^2 + 2\delta - 1
 \end{aligned}$$

and if you can solve for δ without using a calculator, well, more power to you. Since I am sitting at a computer I can tell you the answer is $\delta \sim 0.36603$, but there is no way you can solve a problem like this on a quiz or an exam.

Notice a bottom line comparative statistic that can be useful for you. Let T be the number of periods you don't go back to the restaurant, and $\delta^*(T)$ be the minimum required δ , then $\delta^*(1) = \frac{1}{2}$, $\delta^*(2) \sim \frac{2}{5}$, $\delta^*(\infty) = \frac{1}{3}$. In other words $\delta^*(T)$ is decreasing in T . Remembering that one way to interpret δ is the frequency of interaction, what this basically means is that if you go very often it may be enough not to go back once. If you interact less frequently then you shouldn't go back for more times, until in the extreme you shouldn't go back at all. Further notice that this is *the number of interactions you skip*, not the amount of time. Thus the amount of time you stay away from a restaurant should increase at more than a linear rate. Roughly a square rate would probably work. (If I usually go every week, I can skip them for a week. Then if I usually go once a month I should stay away for 4 months, etceteras.)

6 The Folk Theorems

A "Folk Theorem" is a result that is well known, but has never been proven in a body of theory. In economics these theorems all are variations on the following statement:

In the repeated game almost anything can be equilibrium behavior.

The Folk Theorem was first generally established in the 1980's for both the finite and the infinitely repeated game. What can "almost anything" be? Well let's establish a benchmark, the worst payoff a rational person can get. This is what is called the *minmax* or *minimax* payoff. Remember that $\Sigma_i = \Delta(A_i)$ is the set of mixed strategies in a normal form game, and $\sigma_i \in \Sigma_i$ is an arbitrary mixed strategy, finally let $\Sigma_{-i} = \times_{j \neq i} \Sigma_j$, then we can define:

$$\underline{u}_i = \min_{\sigma_{-i} \in \Sigma_{-i}} \max_{a_i \in A_i} u_i(a_i, \sigma_{-i}),$$

and let m^i be the strategy profile that achieves this minimax. Then a more precise way of writing the Folk Theorem is:

In the repeated game as long as everyone is at least getting an average payoff of \underline{u}_i then this can be an equilibrium.

We know that it can't be lower than this, because no matter what strategy his opponents use a player can always best respond each period, resulting in at least getting an average payoff of \underline{u}_i , so the key question is *can it be this bad?*

At this point I want you to notice how... Machiavellian this goal is. A natural *a-priori* motivation for analyzing repeated games is because in some games (like the Prisoner's Dilemma) the static Nash equilibrium is not Pareto Efficient. Thus you want to know "can Pareto Efficient outcomes be subgame perfect Equilibria?" But once you start analysis, you begin asking "can really bad outcomes be subgame perfect Equilibria?" First, let me point out to you that the two goals can be in congruence. Consider the divide the pie game², if the two players are a fat man and a thin man then it may be socially desirable to have the fat man get a very thin slice of pie—and what we want to know is how thin it can be. Secondly, answering the first question obviously answers the second as well. Can the Pareto Efficient strategy x be supported? Well yes, if $u_i(x) > \underline{u}_i$ for all parties. Otherwise it can not be (in the Pie game consider the very unequal example of person 1 getting the whole pie, this may not give high enough payoffs to both parties.)

I would also like to point out the basic dynamics of adding equilibria to a game. Consider an arbitrary game G , with a set of equilibria E^* . Now if we add an equilibria to G then it *may* Pareto dominate everything in E^* , it may be Pareto worse than everything in E^* , or it may be neither. When we analyze repeated games what we are doing can essentially be thought of as adding equilibria to G , and sometimes this will result in a Pareto dominant equilibria, but sometimes it will not. If you just spend your life focusing on the best things that happen, then you will often be surprised. Equilibria are Pareto Efficient only by coincidence, and just because there is a Pareto Efficient equilibria doesn't mean that it will be the equilibrium.

For example in the side of the road game I made the equilibrium (L, L) Pareto dominate (R, R) . I did this on purpose, if you are driving on the left hand side of the road then oncoming traffic is on the right hand side of the car, and since most of us are right handed we react more quickly to stimuli on our right hand side. It is, in fact, Pareto dominant to drive on the left hand side, but still in most of the world people drive on their right.

A more complex example is the social norm that a woman is supposed to work at home. Before the industrial revolution there was a clear need to have some people specialize in housework. If you don't believe this try washing your own clothes for the next week, without any machines and (if you want to be a purest) only using hand soap. Scared? You should be. I'm not even going to encourage you to cook your own meals without a modern stove—I don't want to be responsible for any fires in the dorms. It was hard maintaining a household,

²Just to be specific, the strategies are (s_1, s_2) where $s_1 \in [0, 1]$ and $s_2 \in [0, 1]$ and if $s_1 + s_2 \leq 1$ $u_1(s_1, s_2) = s_1$, $u_2(s_1, s_2) = s_2$, if $s_1 + s_2 > 1$ then $u_1(s_1, s_2) = u_2(s_1, s_2) = 0$. You would have to add some strategy to make sure that the Nash equilibria are not $s_1 + s_2 = 1$.

and while it was not always a role forced on the woman the fact that she also gave birth to the children does suggest her for this role.

But then, why should this rule also be applied to the wealthy? The wives of aristocrats were generally not allowed to do anything but "proper womanly pursuits" which almost certainly precluded them becoming scientists or politicians or many other productive enterprises. Surely some test should have been applied to young girls, many of them would have definitely been more productive as scientists than as home makers. To prove this let me give one of the few examples of a woman who did manage to become a scientist. Marie (Skłodowska) Curie was born in 1867. She was the first person to get a Nobel prize in two sciences—Physics (1903, with colleagues) and Chemistry (1911, alone).³ Can you imagine how much further along the sciences would be today if the potential of all those women was properly channelled?⁴

And there is a second problem for a scientist, which is that with so many equilibria around what do you predict people will do? If there is only one equilibrium I can test it empirically and if it does not fit the data then my theory is wrong. But if anything can be an equilibrium... well. I don't know what to do. This impression, I should mention, is only partially correct. Many valuable papers have been written by economists looking at repeated game equilibria. The key thing about analyzing these things in the real world is that you must find the rewards and punishments. If you can identify them then you can explain why and how people are doing what they do, if you can not then you have to admit that your theory is wrong.

Now in our analysis (except the appendices) we will assume that m^i is a Nash equilibrium, but I want to be clear that this is an arbitrary assumption. I can easily find an intuitive game where it is not a Nash equilibrium, the punishment game:

| | | | |
|--------|----------|----------|----------|
| | | Child | |
| | | <i>O</i> | <i>B</i> |
| Parent | <i>N</i> | 5; 3 | 0; 6 |
| | <i>P</i> | 0; -3 | -5; 0 |

E5 : The Punishment Game.

The strategies for the child are to obey the rules or break the rules, notice (obviously) they have a dominant strategy to break the rules. The strategies for the parent are to punish or not, and notice they have a dominant strategy to not punish. Thus if both players play their dominant strategies the outcome is (*N*, *B*), which results in a spoiled child who gets away with everything they want. You don't want to go there, trust me.

These are the payoffs that your parents claimed they received when they punished you for breaking the rules as a small child, and if you've ever been

³My source for these claims is wikipedia, http://en.wikipedia.org/wiki/Marie_Curie .

⁴In all honesty I must admit that I do not have a tested alternative to the traditional division of labor. I am merely pointing out that to casual analysis the traditional method seems wrong.

around small children you will realize they were correct. Punishment always results in lots of crying, interpersonal battles to make sure the punishment is enforced, etceteras. Trust me, as a parent it is definitely a short term dominant strategy to not punish the child. Of course in the long run this is absolute suicide.

But this is beside the point. All we really care about is that the minimax strategy profile for the child is (P, B) , and this is not a Nash equilibrium. Notice something even more disturbing. The minmax strategy for the parent is (N, B) , which gives a utility of zero, thus playing the minimax strategy for the child gives the parent less than their minimax payoff. This makes analyzing these games rather complicated and we will only address them at the end of this handout (which you honestly don't need to read.) A key rule in a game like this is that breaking the rules **must** be forgiven after a while.

The minimax strategy may also be in mixed strategies. Now in this case I can't give you any intuitive game where this is true, I can only show you an abstract game based on matching Pennies.

| | | | | |
|----------|-----|----------|--------|------|
| | | Player 2 | | |
| | | H | T | O |
| Player 1 | H | 1; -2 | -1; -1 | 4; 0 |
| | T | -1; -1 | 1; -2 | 4; 0 |
| | O | -2; -1 | -2; -1 | 3; 3 |

E6 : A Mixed Minimax Game

In this game we assume that Player 2 has a disutility from gambling of -2 , so they have a dominant strategy of O . Player 1 likes to gamble, thus O is a dominated strategy. However you can take my word for it that the minimax strategy is for player 2 to play H half the time and T half the time. Notice that the best response to this strategy is for player 1 to play any mixture of H and T , but that without loss of generality we can assume he just plays H .

6.1 A Folk Theorem for Finitely Repeated Games

There are several ways to write this Theorem. The one that comes closest to the original statement is:

Theorem 8 *In a finitely repeated game, if the stage game has more than one Nash equilibrium then any strategy that gives an average payoff higher than that of the worst Nash equilibrium for each player can be supported as a subgame perfect equilibrium when T is very large.*

However this is not the Folk Theorem suggested by my super strategies. In those strategies I focus only on getting cooperation in period one. This is enough, because if T is large enough to get cooperation in period 1 then $T + 1$ is large enough to get cooperation in periods 1 and 2, and so on. Thus another way of writing this down is:

Theorem 9 *In a finitely repeated game, if the stage game has more than one Nash equilibrium then any pure strategy outcome in period 1 can be part of a subgame perfect equilibrium when T is very large.*

I will prove this Theorem, it is easy to do. Consider the super strategy mark 2, and let $\hat{u}_i = \max_{a_i \in A_i} u_i(a_i, c_{-i})$. Now first we need to prove that people are in equilibrium after the first period, this is easy to do but important. In these subgames what happens in the future does not depend on what happens today, and since they are supposed to play a static Nash equilibrium we can be sure their incentives are correct.

Then the value of the equilibrium path is:

$$V_i^* = u_i(c) + (T - 1) u_i(r)$$

the value of the optimal deviation is:

$$\hat{V}_i = \hat{u}_i + (T - 1) u_i(p^i)$$

and

$$\begin{aligned} V_i^* &\geq \hat{V}_i \\ u_i(c) + (T - 1) u_i(r) &\geq \hat{u}_i + (T - 1) u_i(p^i) \\ (T - 1) (u_i(r) - u_i(p^i)) &\geq \hat{u}_i - u_i(c) \end{aligned}$$

now we must have $\hat{u}_i - u_i(c) < \infty$, and by construction we have $u_i(r) - u_i(p^i) > 0$, thus as $T \rightarrow \infty$ this must be satisfied. In fact, given r , we can easily find the critical value for T for person i , T_i^* , we need:

$$\begin{aligned} (T - 1) (u_i(r) - u_i(p^i)) &\geq \max_{c \in A} \hat{u}_i - u_i(c) \\ T &\geq \frac{\max_{c \in A} \hat{u}_i - u_i(c)}{u_i(r) - u_i(p^i)} + 1 \end{aligned}$$

so $T_i^* = \frac{\max_{c \in A} \hat{u}_i - u_i(c)}{u_i(r) - u_i(p^i)} + 1$. Thus to get everyone to cooperate we need $T \geq T^*$
 $= \max_i T_i^* = \max_i \left(\frac{\max_{c \in A} \hat{u}_i - u_i(c)}{u_i(r) - u_i(p^i)} + 1 \right) = \max_i \frac{\max_{c \in A} \hat{u}_i - u_i(c)}{u_i(r) - u_i(p^i)} + 1$, and thus the optimal r is found by finding $\min_{r \in NE} \max_i \frac{\max_{c \in A} \hat{u}_i - u_i(c)}{u_i(r) - u_i(p^i)}$.

6.1.1 A Technical Aside: Games with Two Pure Strategy Nash Equilibria.

Now, one final point. What if the game has exactly two Nash equilibria? Or more specifically games like $E3$:

| | L | C | R |
|---|--------------------|--------------------|---------------------|
| U | 5; 5 | 5; 4 | 1; 12 ¹² |
| M | 3; 2 | 0; 0 | 2; 11 ² |
| D | 10; 0 ¹ | 6; 1 ¹² | 0; 0 |

where there are exactly two pure strategy Nash equilibria that are not Pareto ranked. Since I used three Nash equilibria in my proof it would seem this is not enough. However that is not correct. I said "Nash equilibria" not "Pure Strategy Nash equilibria" and if there are two Nash equilibria there is always a third one. For example in the game $E3$ there is a mixed strategy Nash equilibrium where player 1 plays U with probability $\frac{1}{9}$ and D with probability $\frac{8}{9}$, and player 2 plays both C and R with probability $\frac{1}{2}$. The utilities of this equilibria are $u_1 = 3$, $u_2 = \frac{4}{3}$, so this can be our r .

However there is a simpler method, called *correlation*. For an arbitrary set of action profiles $A = \times_i A_i$ let α_a be a mixture over A , $\sum_{a \in A} \alpha_a = 1$, $\alpha_a \geq 0$. Essentially what we are imagining is that first the players jointly roll a dice, and then based on the outcome of that dice roll they play some strategy. While this might seem a little artificial it is not that unreasonable of an assumption. For example it is almost equivalent to playing a cycle of strategies. (Play x today, y tomorrow, z the day after and then repeat.) It also can be interpreted as checking whether the day is sunny, who was first into the office, etceteras, and conditioning behavior on the outcome. But we are essentially allowing for it because it makes our life easy, and we like having a simple life.

Obviously in this game we want $\alpha_{D,C} = \alpha \geq 0$ and $\alpha_{U,R} = 1 - \alpha$, and we can actually precisely find the optimal α ,

$$\begin{aligned} \frac{\max_{c \in A} \hat{u}_1 - u_1(c)}{u_1(r) - u_1(p^1)} &= \frac{10 - 5}{6\alpha + (1 - \alpha)1 - 1} = \frac{1}{\alpha} \\ \frac{\max_{c \in A} \hat{u}_2 - u_2(c)}{u_2(r) - u_2(p^2)} &= \frac{11 - 0}{\alpha + (1 - \alpha)12 - 1} = \frac{1}{(1 - \alpha)} \end{aligned}$$

and the minimum is achieved when the right hand sides are equal:

$$\frac{1}{\alpha} = \frac{1}{(1 - \alpha)}$$

or $\alpha = \frac{1}{2}$. This tells us that $T^* = \frac{1}{\frac{1}{2}} + 1 = 3$, thus if $T \geq 3$ we are done. Of course I would never assign such a complex problem. I just wanted to lay out clearly how I would proceed.

6.2 The Folk Theorem in the Infinitely repeated Prisoner's Dilemma.

Now we want to do the infinitely repeated folk theorem, and I am first going to do this in the Prisoner's Dilemma because the precision of this example allows for a clearer analysis. The next section covers the most general theorem you are responsible for.

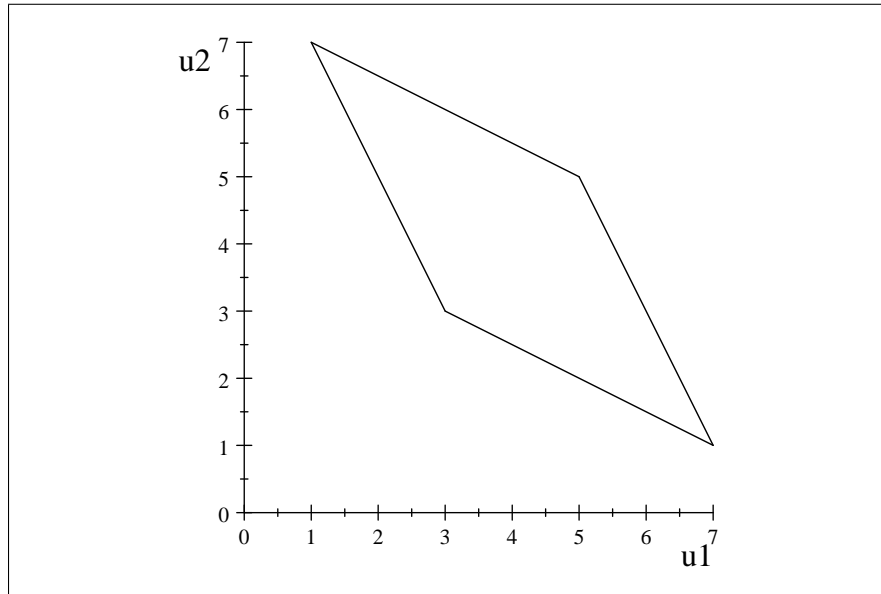
In order to show you the full folk theorem I am going to use correlated strategies. We will have people play the *correlated* strategy α . In the strategy α you play (C, C) with probability α_1 , (C, D) with probability α_2 , (D, C) with probability α_3 , and (D, D) with probability $1 - \alpha_1 - \alpha_2 - \alpha_3$. Your utility of

this strategy?

$$U_1(\alpha) = \alpha_1 u_1(C, C) + \alpha_2 u_1(C, D) + \alpha_3 u_1(D, C) + (1 - \alpha_1 - \alpha_2 - \alpha_3) u_1(D, D)$$

$$U_2(\alpha) = \alpha_1 u_2(C, C) + \alpha_2 u_2(C, D) + \alpha_3 u_2(D, C) + (1 - \alpha_1 - \alpha_2 - \alpha_3) u_2(D, D) .$$

Why do we use this silly trick? Because then by varying $\alpha_1, \alpha_2,$ and α_3 we can get any payoff that is a convex combination of the four payoffs of the original game. Looking at these payoffs in (u_1, u_2) space:



we can get anything in the diamond. Now lets consider the strategy:

1. In period 1 play α .
2. In period $t > 1$ If you played α last period play α in this period.
3. In period $t > 1$ otherwise play (D, D) .

When will this be an equilibrium for some $\delta < 1$?

$$V_1^* = U_1(\alpha) + \frac{\delta}{1-\delta} U_1(\alpha)$$

$$\hat{V}_1 = U_1(D, \alpha) + \frac{\delta}{1-\delta} U_1(D, D)$$

$$U_1(\alpha) + \frac{\delta}{1-\delta} U_1(\alpha) \geq U_1(D, \alpha) + \frac{\delta}{1-\delta} U_1(D, D)$$

$$\frac{\delta}{1-\delta} (U_1(\alpha) - U_1(D, D)) \geq U_1(D, \alpha) - U_1(\alpha)$$

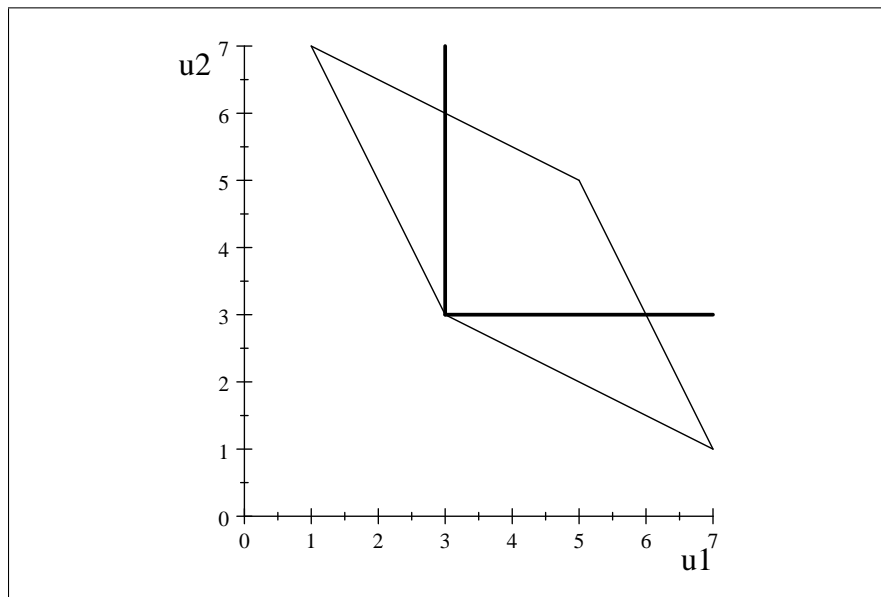
Now let us look at this equation. The right hand side is a fixed positive number. If $U_1(\alpha) - U_1(D, D) > 0$ then $\frac{\delta}{1-\delta}(U_1(\alpha) - U_1(D, D)) \rightarrow \infty$ as $\delta \rightarrow 1$. Thus all we need is $U_1(\alpha) > U_1(D, D)$ and there will be some δ high enough so that this is an equilibrium. For player 2 the key equation is obviously:

$$\frac{\delta}{1-\delta}(U_2(\alpha) - U_2(D, D)) \geq U_2(D, \alpha) - U_2(\alpha)$$

and by a similar argument what we need is $U_1(\alpha) > U_1(D, D)$ and $U_2(\alpha) > U_2(D, D)$.

Theorem 10 *In the Prisoner's Dilemma any correlated strategy α with $U_1(\alpha) > U_1(D, D)$ and $U_2(\alpha) > U_2(D, D)$ can be an equilibrium if δ is high enough.*

Graphically the set of possible equilibria is anything to the upper right of the dark line in the picture below.



6.3 A Folk Theorem in Infinitely Repeated Games

Now we will just consider an arbitrary game and continue to use the correlated strategy α , the strategies we will use will be the Super Strategies Mark II:

Infinitely Repeated Games Super Strategy (Mark 2) .

1. In period 1 play α .
2. In period $t > 1$ If you played α last period play α in this period.
3. In period $t > 1$ if player $i < I$ was the first to deviate, play p^i .

4. In period $t > 1$ otherwise play p^I .

Where each p^i is a Nash equilibrium of the stage game. Notice that we will allow $p^i = p^j$, in fact it won't change our analysis at all. The folk theorem we will prove is:

Theorem 11 *If for all i , $u_i(\alpha) > u_i(p^i)$ then there is a critical δ^* such that if $\delta \geq \delta^*$ there is a subgame perfect equilibrium where players expect to play α forever.*

First, like always, in the subgames that are not the equilibrium path players will follow the strategy because what happens today will not affect the future and what is supposed to happen today is a static Nash equilibrium. Now we just need to show that playing α forever is an equilibrium for high enough δ .

$$V_i^* = u_i(\alpha) + \frac{\delta}{1-\delta} u_i(\alpha) .$$

Like before, let $\hat{u}_i = \max_{a_i \in A_i} u_i(a_i, \alpha_{-i})$ then:

$$\hat{V}_i = \hat{u}_i + \frac{\delta}{1-\delta} u_i(p^i)$$

$$\begin{aligned} V_i^* &\geq \hat{V}_i \\ u_i(\alpha) + \frac{\delta}{1-\delta} u_i(\alpha) &\geq \hat{u}_i + \frac{\delta}{1-\delta} u_i(p^i) \\ \frac{\delta}{1-\delta} [u_i(\alpha) - u_i(p^i)] &\geq \hat{u}_i - u_i(\alpha) \end{aligned}$$

and we notice that $u_i(\alpha) - u_i(p^i) > 0$ by assumption, and $\hat{u}_i - u_i(\alpha) < \infty$, thus there is a critical δ_i^* such that if $\delta \geq \delta_i^*$ then is true. To be precise $\delta_i^* = \frac{\hat{u}_i - u_i(\alpha)}{u_i(\alpha) - u_i(p^i)}$, and $\delta^* = \max_i \delta_i^* = \max_i \frac{\hat{u}_i - u_i(\alpha)}{u_i(\alpha) - u_i(p^i)}$.

7 Some Interesting Strategies in Infinitely Repeated Games

From this point on this is actually just supplementary reading. I may cover something from what follows in class (in which case it is required) but generally you aren't responsible for this material. In this section I want to discuss some interesting strategies in Infinitely Repeated Games.

7.1 What's wrong with Tit-for-Tat

There are generally thought to be two paradigmatic strategies in infinitely repeated games. The Grim strategies and Tit-for-Tat. In the Tit-for-Tat strat-

egy you do today what your opponent did yesterday. Let us consider this strategy in a different Prisoner's Dilemma.

| | | | |
|----------|---|----------|------|
| | | Player 2 | |
| | | C | D |
| Player 1 | C | 6; 6 | 1; 7 |
| | D | 7; 1 | 3; 3 |

E1b : Prisoner's Dilemma with High $u(C, C)$.

The only difference is that I have increased the payoff of (C, C) by one for each person.

Now first of all I want you to really think about what will happen after player 1 deviates once. Don't just assume it is like you think it's supposed to be (you tattled me so I'm going to play D forever) think about what the strategy actually says. In the first period player 2 is going to play D because that's what player 1 did last time, if player 1 follows the strategy he's supposed to play C . Next period it will reverse. Wild, huh? Not at all what you intended to write down, huh? I guess you should be more careful next time. So... let's get precise about these payoffs. Let $V_1^*|_{1 \text{ dev., odd}}$ be the continuation payoff player 1 should expect in odd periods when he has deviated, and $V_1^*|_{1 \text{ dev., even}}$ be the continuation payoff player 1 should expect in even periods. I will use the value function to calculate these values, using the cute trick that after two periods we return to the same continuation value:

$$\begin{aligned}
 V_1^*|_{1 \text{ dev., odd}} &= u_1(C, D) + \delta u_1(D, C) + \delta^2 V_1^*|_{1 \text{ dev., odd}} \\
 (1 - \delta^2) V_1^*|_{1 \text{ dev., odd}} &= u_1(C, D) + \delta u_1(D, C) \\
 V_1^*|_{1 \text{ dev., odd}} &= \frac{u_1(C, D) + \delta u_1(D, C)}{1 - \delta^2} \\
 V_1^*|_{1 \text{ dev., odd}} &= \frac{1 + 7\delta}{1 - \delta^2},
 \end{aligned}$$

and

$$\begin{aligned}
 V_1^*|_{1 \text{ dev., even}} &= u_1(D, C) + \delta u_1(C, D) + \delta^2 V_1^*|_{1 \text{ dev., even}} \\
 V_1^*|_{1 \text{ dev., even}} &= \frac{7 + \delta}{1 - \delta^2}.
 \end{aligned}$$

Now, consider that player 1 is supposed to play D today, then what happens if he deviates by playing C ? Afterwards both players will start playing (C, C) forever. Gosh, that sounds pretty good. So maybe it will be better than the strategies prescribed path.

$$\hat{V}_1|_{1 \text{ dev., even}} = \frac{1}{1 - \delta} u_1(C, C) = \frac{1}{1 - \delta} 6$$

$$\begin{aligned}
V_1^*|_{1 \text{ dev., even}} &\geq \hat{V}_1|_{1 \text{ dev., even}} \\
\frac{7+\delta}{1-\delta^2} &\geq \frac{1}{1-\delta}6 \\
7+\delta &\geq \frac{1-\delta^2}{1-\delta}6 \\
7+\delta &\geq 6+6\delta \\
\delta &\leq \frac{1}{5}
\end{aligned}$$

Uhh ohh, this is only going to work if δ is **small**, not a good sign. OK, but we can proceed, we now know this won't work if people are too patient, but perhaps its still an optimal strategy with "strangers" like the Grimm strategy. What about the other case, where player 1 is supposed to play C today? Again with one more deviation he can switch to (D, D) forever strategy:

$$\hat{V}_1|_{1 \text{ dev., odd}} = \frac{1}{1-\delta}u_1(D, D) = \frac{1}{1-\delta}3$$

$$\begin{aligned}
V_1^*|_{1 \text{ dev., odd}} &\geq \hat{V}_1|_{1 \text{ dev., odd}} \\
\frac{1+7\delta}{1-\delta^2} &\geq \frac{1}{1-\delta}3 \\
1+7\delta &\geq \frac{1-\delta^2}{1-\delta}3 \\
1+7\delta &\geq 3+3\delta \\
\delta &\geq \frac{1}{2}.
\end{aligned}$$

OK, now we know that this strategy is never an equilibrium. I should point out that these results depend critically on the payoffs. You can show that if $u(C, C) = (5, 5)$ it will work only when $\delta = \frac{1}{2}$, and if $u(C, C) = (4, 4)$ then it will work if $\delta \geq \frac{1}{2}$. Thus these strategies work in some games, but not all. Not very interesting in my opinion. Basically when people think of the *tit-for-tat* strategy what they really mean is: "If you treat me well today then I will treat you well tomorrow, otherwise I won't trust you." This is the Grimm strategy, not the standard way of writing tit-for-tat.

7.2 Allowing for Reputation Rebuilding in the Quality Game.

In the quality game we have a one-sided incentive problem, the customer is always best responding given his beliefs about the restaurant. Thus it is only natural to think that a one-sided incentive system should be enough to keep the restaurant in line. This is actually not true, because like in the tit-for-tat strategy patient restaurants will want to rebuild their reputation.

So formally a strategy like this for the customer is: *B if the restaurant chose H last time or $t = 1$, N else.* The game we will consider is:

| | | | |
|-----------------|---|---------------|---------|
| | | Customer (P2) | |
| | | B | N |
| Restaurant (P1) | H | $\pi; 2$ | $-1; 0$ |
| | L | $3; -1$ | $0; 0$ |

E2a : Quality Game with Different Profits

The only difference between this and the previous game is now the profits of the restaurant if they produce high quality is some $3 > \pi > 0$. Now analyzing this strategy is a little harder than previously, because we aren't specify the restaurants optimal strategy—we just let it best respond to the customer's strategy. However we want them to choose H forever, so a benchmark is what these payoffs are:

$$V_1^* = \frac{1}{1-\delta} \pi .$$

Now if they choose L given they are expected to choose H there payoffs must be written as:

$$\hat{V}_1 = u_1(L, B) + \delta V_1(A_{2,t+1} = N)$$

because we can't be sure what they will do when $A_{2,t+1} = N$. They have two choices, first they can just produce low quality—and if they choose to do it once they will always choose to do it—second they can *rebuild their reputation* by producing high quality. If they do this the customer will trust them again and they can conceivably get V_1^* from the second period on, thus:

$$V_1(A_{2,t+1} = N) \in \left\{ \begin{array}{l} u_1(H, N) + \delta V_1^* = -1 + \frac{\delta}{1-\delta} \pi \\ \frac{1}{1-\delta} 0 \end{array} \right.$$

and it will be equal to whichever one is higher (we will assume H if they are indifferent). Thus we need to know when:

$$\begin{aligned} -1 + \frac{\delta}{1-\delta} \pi &\geq 0 \\ \delta &\geq 1 - \delta \\ \delta(1 + \pi) &\geq 1 \\ \delta &\geq \frac{1}{\pi + 1} \end{aligned}$$

Thus if $\delta \leq \frac{1}{\pi+1}$

$$\hat{V}_1 = 3 \leq \frac{1}{1-\delta} \pi = V_1^*$$

$$\begin{aligned} 1 - \delta &\leq \frac{\pi}{3} \\ 1 - \frac{\pi}{3} &\leq \delta \end{aligned}$$

So this will work when $1 - \frac{\pi}{3} \leq \delta \leq \frac{1}{\pi+1}$, which requires that $\pi \geq 2$. If $\delta \geq \frac{1}{1+\pi}$ it is useful to write:

$$\begin{aligned} V_1^* &= \pi + \delta\pi + \delta^2 V_1^* \\ \hat{V}_1 &= 3 - \delta + \delta^2 V_1^* \end{aligned}$$

$$\begin{aligned} V_1^* &\geq \hat{V}_1 \\ \pi + \delta\pi + \delta^2 V_1^* &\geq 3 - \delta + \delta^2 V_1^* \\ \pi + \delta\pi &\geq 3 - \delta \\ \delta(\pi + 1) &\geq 3 - \pi \\ \delta &\geq \frac{3 - \pi}{\pi + 1} \end{aligned}$$

since we need $\delta < 1$ this requires that $\pi > 1$. Notice that if $\pi > \frac{3}{2}$ one can establish that a one period punishment will work:

$$\begin{aligned} \hat{V}_1 &= 3 + \delta 0 + \delta^2 V_1^* \\ V_1^* &= \pi + \delta\pi + \delta V_1^* \end{aligned}$$

$$\begin{aligned} V_1^* &\geq \hat{V}_1 \\ \delta &\geq \frac{3 - \pi}{\pi} < 1 \\ \pi &> \frac{3}{2} \end{aligned}$$

Thus this strategy is stronger than a one period punishment, but not strong enough to always work. Now since these are the actual strategies that we use in the real world it is worthwhile to think about them a little bit more carefully.

First of all, are the restaurants payoffs reasonable? Yes, from my experience working at a restaurant trust me it does cost you to prepare for customers that don't show up, and of course these costs are higher in the quality of the food you produce. Of course you could just shut down, but then essentially the restaurant is choosing to follow a trigger strategy. That's very nice if it's the only profitable strategy, but in reality restaurants do not do this. Thus they do prepare for customers that don't show up, even when they are being punished. (Not being able to serve customers will generally unleash even harsher punishments.)

Secondly, do restaurants usually switch to lower quality when their business drops? (In this strategy that means they're being punished.) I would say not, in fact I think most restaurateurs would tell you that they work even harder to maintain quality standards when business is low. Why? Well they'll tell you its because they need to work to establish a good reputation (just like in our strategy). Think about how many new restaurants you've gone to and then when you've gone back in a year or so you've thought you got lower quality food or service. Most restaurants do this, and those that do not are truly the stars of the restaurant world.

Finally, how do customers observe the restaurant is producing high quality when the customer does not buy? Well obviously it is from word of mouth—their friends still go to the restaurant. And this is the key to why restaurants are disciplined even though they may not change the quality of good they provide. We don't trust word of mouth as much as we do our own experience, and we should not. People have different standards, and in the case of food poisoning observing low quality is random anyway. So while we all follow strategies like this, it takes many "observations" of H for us to start using the B strategy again. More formally, let β be our belief about the probability the restaurant is producing high quality. Then B is optimal if $\beta 2 + (1 - \beta)(-1) \geq 0$ or $\beta \geq \frac{1}{2}$. So the strategy we actually use is: *Given a history h if $\beta(h) \geq \frac{1}{2}$ we take the action B , otherwise we take the action N .* To make this an equilibrium we have to have: *if $A_{1,t-1} = L$ then $\beta(h) < \frac{1}{2}$, and we have to have at least T observations of H after an observation of L to have $\beta(h) \geq \frac{1}{2}$.* Then this results in a T period punishment strategy where the restaurant *must* maintain high quality to rebuild it's reputation.

7.3 Forgiveness using Stochastic Punishments

A mathematically more elegant way to analyze forgiving strategies is: p with probability ρ if last period c did not occur, c otherwise. Where p is the punishment action profile in Super Strategy Mark I, and c is the cooperative strategy. These strategies are *Markov* because they depend only on the state of the world in the last period (the strategy profile played). It should be obvious that if $\rho = 1$ then these are the Grimm strategies, and for strategy that punishes for only T periods there is an equivalent ρ such that the expected payoffs are equivalent.

Let $-c$ be the state where something other than c occurred last period, then with these strategies:

$$\begin{aligned}
 V_i^* &= u_i(c) + \delta V_i^* \\
 V_1(-c) &= \rho u_i(p) + (1 - \rho) V_i^* \\
 \hat{V}_1 &= \hat{u}_i + \delta V_1(-c) \\
 &= \hat{u}_i + \delta \rho u_i(p) + \delta (1 - \rho) V_i^* \\
 V_i^* - \hat{V}_1 &= u_i(c) - \hat{u}_i - \delta \rho u_i(p) + \delta \rho V_i^* \geq 0 \\
 \delta \rho (V_i^* - u_i(p)) &\geq \hat{u}_i - u_i(c) \\
 \delta \rho \left(\frac{1}{1 - \delta} u_i(c) - u_i(p) \right) &\geq \hat{u}_i - u_i(c)
 \end{aligned}$$

This will obviously work for any $\rho > 0$ as long as δ is high enough, and indeed we can solve for δ_i^* if we normalize payoffs so that $u_i(p) = 0$. (This can be done

without loss of generality). Then:

$$\begin{aligned}\frac{\delta}{1-\delta} &\geq \frac{\hat{u}_i - u_i(c)}{\rho u_i(c)} \\ \delta &\geq \frac{\hat{u}_i - u_i(c)}{\hat{u}_i - (1-\rho)u_i(c)}\end{aligned}$$

and $\delta_i^* = \frac{\hat{u}_i - u_i(c)}{\hat{u}_i - (1-\rho)u_i(c)}$. Much more convenient than analyzing a T period finite punishment, no?

It is also more reasonable. Consider for example the quality game. We all might *intend* to not go to a restaurant for the next ten periods, but then we might be in a situation where we have to eat at the restaurant despite our intentions. Alternatively we might just end up not eating there for twenty periods because, again, we just weren't in a situation where eating at that restaurant was optimal. We use stochastic punishments, though they don't usually have this sort of elegant mathematical formula, and its nice to see that we can use them to analyze forgiveness.

7.4 Several Strategies in the Repeated Bertrand Oligopoly

The Bertrand Oligopoly... well you know I don't like it. The idea that my demand shrinks to zero if I charge one kurus more than you is just... ridiculous. But as I've said before we often use it because it is mathematically simple, and in a repeated game it is extremely mathematically simple. Indeed, we will be able to prove a folk theorem for this game without any sweat, and surprisingly enough for reasonable levels of the market price the critical δ will not depend on the price, but rather only the number of firms in the industry.

So, the basic game is each firm has a marginal cost of c , and chooses a price $p_i \geq 0$. Let p be the vector of prices, and J be the set of firms in the industry then:

$$d_i(p) = \begin{cases} 0 & \text{if } p_i > \min_{j \neq i} p_j \\ \frac{\gamma_i}{\sum_{j \in M} \gamma_j} D(p_i) & \text{if } p_i = \min_{j \neq i} p_j \quad \text{and} \quad M = \{j \in J | p_j = p_i\} \\ D(p_i) & \text{if } p_i < \min_{j \neq i} p_j \end{cases}$$

where $\gamma_i > 0$ and $\sum_{j \in J} \gamma_j = 1$. We will of course analyze the case where $M = J$, or the entire industry charges the same price, in this case $\frac{\gamma_i}{\sum_{j \in M} \gamma_j} = \gamma_i$. It is simple to show that in the Nash equilibrium $p_i = c$ for all i .

The first issue that we will need to resolve is what will be the profit if all firms j have $p_j = p \geq c$ and i charges $p - \varepsilon$? These profits for a fixed ε are:

$$\begin{aligned}D(p - \varepsilon)(p - \varepsilon - c) &= D(p - \varepsilon)(p - c) - \varepsilon D(p - \varepsilon) \\ &= D(p)(p - c) - D(p)(p - c) + D(p - \varepsilon)(p - c) - \varepsilon D(p - \varepsilon) \\ &= D(p)(p - c) + (D(p - \varepsilon) - D(p))(p - c) - \varepsilon D(p - \varepsilon)\end{aligned}$$

Writing $\pi = D(p)(p - c)$, then these profits are:

$$D(p - \varepsilon)(p - \varepsilon - c) = \pi - \mu$$

where $\mu = (D(p - \varepsilon) - D(p))(p - c) - \varepsilon D(p - \varepsilon)$, notice that $\mu \rightarrow 0$ as $\varepsilon \rightarrow 0$, and one can establish that if p is less than the monopoly price then $\frac{\partial \mu}{\partial \varepsilon} < 0$. Thus we want to analyze this for very small ε , and in this case:

$$\lim_{\varepsilon \rightarrow 0} D(p - \varepsilon)(p - \varepsilon - c) = \pi$$

so we will assume that in this case the firm will get π . Now we need to consider what will happen if p is strictly higher than the monopoly price. Then what the deviating firm will do in this case is just charge the monopoly price, which we will denote by the profits π^m .

7.5 Equilibria with the Super Strategy (Mark I)

So the obvious strategy to follow in this game is: *if $t = 1$ or p was charged by everyone last period then set $p_i = p$, otherwise $p_i = c$* . Notice that the profits firm i will get will be $\gamma_i D(p)(p - c) = \gamma_i \pi$, and first assume that $p \leq p^m$ —the monopoly price. In this case:

$$\begin{aligned} V_i^* &= \frac{1}{1 - \delta} \gamma_i \pi \\ \hat{V}_i &= \pi + \frac{\delta}{1 - \delta} 0 \\ V_i^* &\geq \hat{V}_i \\ \frac{1}{1 - \delta} \gamma_i \pi &\geq \pi \\ \gamma_i &\geq 1 - \delta \\ \delta &\geq 1 - \gamma_i \end{aligned}$$

thus $\delta_i^* = 1 - \gamma_i$. Now notice that the γ_i 's can be a subject of negotiation, firms can argue about how to split demand since customers don't care, and that $\min_i \gamma_i \leq \frac{1}{|J|}$ because otherwise we would have $\sum_i \gamma_i \geq |J| \min_i \gamma_i > |J| \frac{1}{|J|} = 1$ if that is true. Thus cooperation is easiest to achieve when $\gamma_i = \frac{1}{|J|}$ for all firms, thus the equilibrium exists if $\delta \geq \delta^* = 1 - \frac{1}{|J|}$. Now let us consider the perverse case where the industry is charging more than the monopoly price, in this case:

$$\begin{aligned} \hat{V}_i &= \pi^m + \frac{\delta}{1 - \delta} 0 \\ V_i^* &\geq \hat{V}_i \\ \frac{1}{1 - \delta} \gamma_i \pi &\geq \pi^m \\ \gamma_i \frac{\pi}{\pi^m} &\geq 1 - \delta \\ \delta &\geq 1 - \gamma_i \frac{\pi}{\pi^m} . \end{aligned}$$

Thus the result is:

Proposition 12 *If $p \leq p_m$ then there is a subgame perfect equilibrium where all firms charge p on the equilibrium path if $\delta \geq 1 - \frac{1}{|\mathcal{J}|}$. This is independent of p . If $p > p_m$ then a subgame perfect equilibrium where all firms charge p on the equilibrium path if $\delta \geq 1 - \frac{1}{|\mathcal{J}|} \frac{\pi(p)}{\pi^m}$.*

Wild, huh? For the prices we think they may want to consider cooperation is completely independent of the level of profits. Only if price gets ridiculously high does the critical δ^* increase in the price (because $\pi(p)$ will decrease in the price.) What a wild result, I don't know any other game like it. Essentially the full folk theorem independent of the payoffs.

From now on we will always assume that $\gamma_i = \frac{1}{|\mathcal{J}|}$ and $p \leq p^m$ because we have other things to talk about.

7.5.1 Delayed Punishments

In most industries firms can not change prices over night. Among other things they sell to middle men, who then sell their goods at a retail level, so it takes time for a price change to percolate down. So how does this affect analysis? Well this means that they can change price at most K periods after they realize someone is undercutting them. In this case:

$$\hat{V}_i = \pi + \delta\pi + \delta^2\pi + \dots + \delta^{K-1}\pi = \sum_{t=1}^K \delta^{t-1}\pi = \frac{1 - \delta^K}{1 - \delta}\pi$$

In equilibrium we must have:

$$\begin{aligned} V_i^* &\geq \hat{V}_i \\ \frac{1}{1 - \delta} \frac{1}{|\mathcal{J}|} \pi &\geq \frac{1 - \delta^K}{1 - \delta} \pi \\ \delta^K &\geq 1 - \frac{1}{|\mathcal{J}|} \\ \delta &\geq \left(1 - \frac{1}{|\mathcal{J}|}\right)^{\frac{1}{K}}. \end{aligned}$$

Thus it will be harder to support cooperation, but still it just requires that δ is high enough—and again it is independent of the profits. However we need to point out that one usual way of getting cooperation will not work. Usually we can just say "well, interaction is very frequent so δ is nearly one." However if interaction is more frequent logically K has to be larger, after all if we accept it takes firms a meaningful amount of calendar time to react to other firms cheating then if they "interact" every day as opposed to every month obviously K should be 30 times as large. This is a general limit of the Folk Theorem, if reaction takes a real amount of time (which it always does) then having interaction be more and more frequent will just result in larger and larger payoffs to cheating, counteracting the usual result.

7.5.2 Stochastic Rewards (and Stochastic Punishments).

One very common fact about repeated games is that sometimes people make mistakes, and more often other people think that someone did something wrong when actually it was just some freak of nature. This was actually a topic of close study in repeated oligopoly, more specifically Porter, 1983. "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886" *Bell Journal of Economics* pp 301-314. Since it is such a real issue of concern for us in general we will analyze it in this environment.

In that paper Porter studied an interesting moment in Economic History. Prior to anti-trust laws in the United States Chicago Railroad firms entered into an explicit cartel agreement. However the demand for these firms was random, their primary market was shipping goods east and their primary competition was barges on the Great Lakes. If the Great Lakes thawed unexpectedly then their demand and profits would drop sharply. These low profits would trigger periods of price wars between the firms. Formally we assume that firms do not observe the prices of competitors, rather only their own profits. If $\pi_{it} = 0$, then they have to assume someone cheated and start punishing. Notice that since now punishment periods will occur even if everyone follows the rule you want to use forgiving strategies—the harsher the punishments are the lower your profits.

This model was originally written and estimated assuming Cournot competition, however when Rob Porter taught us this model at Northwestern he said he wished he had explained the model using Bertrand competition and finite punishments. This was before the realization of how convenient stochastic punishment strategies are in analysis, and thus I will use stochastic punishments. Thus the strategy firms will use is: $p_{it} = c$ with probability ρ if $\pi_{i,t-1} = 0$, $p_{it} = p$ otherwise. The demand curve will now be given by:

$$d_i(p) = \begin{cases} \frac{1}{|M|} D(p_i) & \text{with probability } q \text{ if } p_i = \min_{j \neq i} p_j \quad \text{and} \quad M = \{j \in J | p_j = p_i\} \\ D(p_i) & \text{with probability } q \text{ if } p_i < \min_{j \neq i} p_j \\ 0 & \text{otherwise} \end{cases}$$

Now we can immediately start calculating the value functions:

$$\begin{aligned} V^*(\pi) &= q \left(\frac{1}{|J|} \pi + \delta V^*(\pi) \right) + (1-q)(0 + \delta V^*(0)) . \\ V^*(0) &= \rho((0 + \delta V^*(0))) + (1-\rho)V^*(\pi) \\ \hat{V} &= \pi + \delta V^*(0) \end{aligned}$$

Notice we now have two equations in two unknowns ($V^*(\pi)$ and $V^*(0)$), and we have to solve for the value functions simultaneously.

$$\begin{aligned} V^*(\pi) &= \frac{q}{(1-\delta q)} \frac{1}{|J|} \pi + \frac{(1-q)\delta}{(1-\delta q)} V^*(0) \\ V^*(0) &= \frac{(1-\rho)}{(1-\rho\delta)} V^*(\pi) \end{aligned}$$

Notice this rather nice relationship between $V^*(0)$ and $V^*(\pi)$.

$$V^*(\pi) = \frac{q}{(1-\delta q)} \frac{1}{|J|} \pi + \frac{(1-q)\delta}{(1-\delta q)} \frac{(1-\rho)}{(1-\rho\delta)} V^*(\pi)$$

$$V^*(\pi) = \frac{\frac{q}{(1-\delta q)}}{1 - \frac{(1-q)\delta}{(1-\delta q)} \frac{(1-\rho)}{(1-\rho\delta)}} \frac{1}{|J|} \pi$$

at this point we have to say "gulp" and hit the simplify key in Scientific Workplace.

$$V^*(\pi) = \frac{q}{1-\delta} \frac{1-\delta\rho}{1-q\delta\rho} \frac{1}{|J|} \pi$$

$$V^*(0) = \frac{q}{1-\delta} \frac{1-\rho}{1-q\delta\rho} \frac{1}{|J|} \pi$$

Relatively speaking, these are actually rather simple functions. Given these values:

$$\hat{V} = \pi + \delta \frac{q}{1-\delta} \frac{1-\rho}{1-q\delta\rho} \frac{1}{|J|} \pi$$

$$V^*(\pi) \geq \hat{V}$$

$$(1-\delta\rho) q \frac{1}{|J|} \geq (1-\delta)(1-q\delta\rho) + \delta(1-\rho) a \frac{1}{|J|}$$

$$q(1-\delta) \frac{1}{|J|} \geq (1-\delta)(1-q\delta\rho)$$

$$\delta\rho \geq \frac{1}{q} - \frac{1}{|J|}$$

While the math was rather complicated the result is elegant. It should not be surprising at all that it will now be harder to get cooperation for low q , and this is exactly what we find. We must have $\frac{1}{q} - \frac{1}{|J|} < 1$, or $q > \frac{|J|}{|J|+1}$, and then the condition must hold when $\rho = 1$, so we must have $\delta \geq \delta^* = \frac{1}{q} - \frac{1}{|J|}$. One interesting exercise in this analysis is that we can now optimize our profits over ρ , assuming the conditions on q and δ are met. One can easily show that $\frac{\partial V^*(\pi)}{\partial \rho} < 0$ so the optimal ρ is the minimal one that will make cooperation optimal, this is $\rho = \frac{1}{\delta} \left(\frac{1}{q} - \frac{1}{|J|} \right)$.

8 Towards the Full Folk Theorem in Two Player Games.

In two player games we can define the *mutual minimax* strategies as $m = (m_1^2, m_2^1)$, in these strategies player 1 is holding player 2 down to his lowest

possible value, and player 2 is holding player 1 down to his lowest possible value. Considering the Punishment game:

| | | | |
|--------|----------|----------|----------|
| | | Child | |
| | | <i>O</i> | <i>B</i> |
| Parent | <i>N</i> | 5; 3 | 0; 6 |
| | <i>P</i> | 0; -3 | -5; 0 |

E5 : The Punishment Game

then this strategy is (P, B) . Notice that while $\underline{u}_1 = \max_{a_1 \in (N, P)} u_1(a_1, B) = 0$ that $u_1(m) = u_1(P, B) = -5$. Thus when using the mutual minimax strategies players may be getting strictly less than their minimax payoffs. This means that we have to have forgiveness built into the strategies.

Throughout this section I will make the simplifying normalization that $\underline{u}_i = 0$. This can be done without loss of generality and makes presentation simpler. I will prove the folk theorem assuming $m \in A$ —or that both players are playing pure strategies—and then briefly discuss what you have to do if players use mixed strategies.

8.1 The Full Pure Strategies Folk Theorem.

The strategy we will use is the stochastic punishments super strategy: m with probability ρ if last period c did not occur, c otherwise.

$$\begin{aligned} V_i^*(c) &= \frac{1}{1-\delta} u_i(c) \\ V_i^*(-c) &= \rho(u_i(m) + \delta V_i^*(-c)) + (1-\rho)V_i^*(c) \\ V_i^*(-c) &= \frac{\rho}{1-\delta\rho} u_i(m) + \frac{1-\rho}{1-\delta\rho} \frac{1}{1-\delta} u_i(c) \end{aligned}$$

Now, however, we need to check for subgame perfection in both subgames $(-c$ and $c)$. Thus we have a $\hat{V}(c)$ and a $\hat{V}(-c)$.

$$\begin{aligned} \hat{V}(c) &= \hat{u}_i + \delta V_i^*(-c) \\ \hat{V}(-c) &= 0 + \delta V_i^*(-c) \end{aligned}$$

Notice that if you are playing m your optimal deviation is to m_i^i , and this will give you a payoff of zero.

$$\begin{aligned} V_i^*(c) &\geq \hat{V}(c) \\ \left(1 - \frac{1-\rho}{1-\delta\rho}\right) \frac{1}{1-\delta} u_i(c) &\geq \hat{u}_i + \frac{\delta\rho}{1-\delta\rho} u_i(m) \\ \frac{\rho}{1-\delta\rho} u_i(c) &\geq \hat{u}_i + \frac{\delta\rho}{1-\delta\rho} u_i(m) \\ \frac{\rho}{1-\delta\rho} (u_i(c) - \delta u_i(m)) &\geq \hat{u}_i \end{aligned} \tag{8}$$

Now to analyze this condition we can drive $\delta \rightarrow 1$ and see that we have:

$$\frac{\rho}{1-\rho} (u_i(c) - u_i(m)) \geq \hat{u}_i$$

thus if $u_i(c) - u_i(m) > 0$ and ρ is high enough this will be satisfied. To find the condition on ρ we need to look at the subgame $-c$.

$$\begin{aligned} V_i^*(-c) &\geq \hat{V}(-c) \\ V_i^*(-c) &\geq \delta V_i^*(-c) \\ (1-\delta)V_i^*(-c) &\geq 0 \end{aligned}$$

ad $V_i^*(-c) \geq 0$:

$$\rho u_i(m) + \frac{1-\rho}{1-\delta} u_i(c) \geq 0$$

$$\begin{aligned} \frac{1-\rho}{1-\delta} u_i(c) &\geq \rho(-u_i(m)) \\ \frac{1-\rho}{\rho} &\geq (1-\delta) \frac{(-u_i(m))}{u_i(c)} \end{aligned}$$

remembering that $u_i(m) \leq 0$ or that $-u_i(m) \geq 0$ this requires that:

$$\frac{u_i(c)}{(1-\delta)(-u_i(m)) + u_i(c)} \geq \rho \tag{9}$$

Thus we can see that as $\delta \rightarrow 1$ ρ can go to one, and both conditions can be vacuously satisfied. To find the critical δ_i^* we will let $\rho = \bar{\rho} = \frac{u_i(c)}{(1-\delta)(-u_i(m)) + u_i(c)}$, and substitute this into equation 8.

$$\begin{aligned} \frac{\frac{u_i(c)}{(1-\delta)(-u_i(m)) + u_i(c)}}{1-\delta \left(\frac{u_i(c)}{(1-\delta)(-u_i(m)) + u_i(c)} \right)} (u_i(c) - \delta u_i(m)) &\geq \hat{u}_i \\ \frac{u_i(c)}{(1-\delta)(-u_i(m)) + u_i(c) - \delta u_i(c)} (u_i(c) - \delta u_i(m)) &\geq \hat{u}_i \\ \frac{u_i(c)}{(1-\delta)(u_i(c) - u_i(m))} (u_i(c) - \delta u_i(m)) &\geq \hat{u}_i \\ \frac{1}{(1-\delta)} u_i(c) \frac{u_i(c) - \delta u_i(m)}{u_i(c) - u_i(m)} &\geq \hat{u}_i \end{aligned}$$

and this gives us an implicit solution for δ^* . Thus if $\delta \geq \delta^*$ and $\rho \leq \bar{\rho}(\delta)$ there is an equilibrium where players play c .

8.2 Now we're Really Going too Far: What if the Mutual Minimax is in Mixed Strategies?

At this point we really need to get into the nitty gritty of how you play a mixed strategy. Is it:

1. You decide on the probabilities you are going to play an action, and then let someone else do the randomization and choose your action.
2. You do the randomization then based on the outcome of the randomization you choose your action.

Formally we have always assumed that it is the first one, because otherwise it is hard to conceptualize what we mean by an expected utility. But the joy of it is that in a Mixed Strategy Nash equilibrium either method will work. This has stopped us from worrying about this too much. However in a repeated game strategy players may be using a mixed strategy which is not a Nash equilibrium. For example consider the game:

| | | | | |
|----------|----------|----------|----------|----------|
| | | Player 2 | | |
| | | <i>H</i> | <i>T</i> | <i>O</i> |
| Player 1 | <i>H</i> | 1; -2 | -1; -1 | 4; 0 |
| | <i>T</i> | -1; -1 | 1; -2 | 4; 0 |
| | <i>O</i> | -2; -1 | -2; -1 | 3; 3 |

E6 : A Mixed Minimax Game

In this game if player 2 has to minimax player 1 they should play *H* half the time and *T* half the time. Now if player 1 is always playing *H* player 2 could just play *T* all the time and observationally speaking we couldn't tell if he was using the mixed strategy or just playing *T*. Thus this would be an undetectable deviation. What is a theorist to do? Well he could pay player 2 back every time he plays *H*. He obviously can't do it in the current period, so he has to do it in the future. And the payment has to be *precisely* right, or otherwise player 2 will play *H* all the time. This solution was first presented in Fudenberg and Maskin (1986, "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information." *Econometrica*, pp. 533–554) and has not been improved on since. That is the first general folk theorem for repeated games and a seminal paper. They had to address this point, but I'm not very comfortable with the only feasible solution.

The main reason I am not comfortable is because it is obviously not *generic*—or it will not work if we are slightly wrong about either a player's discount factor or their payoffs in the stage game.⁵ In general all of the equilibria I have discussed are very strong. If players are patient enough these equilibria will withstand any refinement in the literature, and they certainly are generic. These equilibria don't even withstand that minimal test.

⁵Normal mixed strategy Nash equilibria do, this has been carefully proven.