Notes on Game Theory

European University Institute

Tuna Dökmeci*

Teaching Block II, 2021/2022

*These notes draw on David Levine's lectures. Other sources used are indicated in the main text.

Contents

Static Games of Complete Information

1 Strategic/Normal Form Games

Static games are games where players choose their plan of actions (strategies) once and for all. These decisions are made simultaneously, that is, in complete ignorance of the action the other players have taken. For this part of the class, we assume complete information such that all players know exactly which actions the other players can take, what the associated payoffs are, and know that the other players also have perfect information. Furthermore, players do not encounter each other again.

A classical example of such a game that you may know of is the prisoner's dilemma, described below.

Example 1.1 (Prisoner's Dilemma)**.** *There are two prisoners, and a crime was committed in the prison. They each face the same decision, the same set of actions that are available to them: they can* ***cooperate (C)*** *with each other by remaining silent, or* ***defect (D)****, by testifying that the other player committed the crime. If they both cooperate so that they are both silent, they only have to serve the rest of their time in prison, which is 2 years. If they both testify against each other, they will both receive an extra year, and serve 3 years. If one of them testifies against the other while the other remains silent, the one who cooperates serves 10 years, while the one who defected will* *be free to walk away. We can represent this game in the following compact form, which is the* ***strategic or normal form game****.*

Table 1.1: Prisoner's Dilemma

Prisoner 1 \ Prisoner 2	Cooperate	Defect
Cooperate	-2,-2	-10,0
Defect	0,-10	-3,-3

We can argue that the unique outcome of this game is that they both defect, and serve 3 years in prison. This is a strong prediction, because there is a strictly dominant strategy for each player. Why is that? Suppose that the other player is defecting, so that they blame me, I prefer to blame them and get -3 instead of -10. Suppose she cooperates so that she keeps silent, if I put the blame on her, I get to walk away, getting 0 instead of -2. Regardless of what the other player is doing, each of them is better off by defecting: defecting is a **strictly dominant strategy**, a concept we will develop further in detail.

Before we discuss how to solve these games in general, let us introduce some terminology to formally define some concepts we mentioned so far, and have a formal definition of what we mean by strategic/normal form games.

Definition 1.1 (Players)**.** *Any agent making a decision in a game is referred as a player. The set of players is denoted by $\mathcal{I}$, and the individual players are indexed by i. When talking about a player i, we refer to other players as opponents, and denote them with $-i$. All players seek to maximize their own payoffs in a game.*

Definition 1.2 (Actions)**.** *Each player has a set of actions that are available to them, denoted A_i.*

In the prisoner's dilemma, we have $A_1 = A_2 = \{\text{Cooperate}, \text{Defect}\}$.

Definition 1.3 (Strategies). *A strategy is a complete contingent plan or decision rule that specifies how the player will act in every possible circumstance in which she might be called upon to move. We call the entire set of strategies available to a player the* ***strategy space****, and denote it by* S_i.

In the static games where players take a single action, strategies are composed of one action so that the set of actions and strategy spaces coincide. In prisoner's dilemma, we have $S_1 = S_2 = A_1 = A_2 = \{\text{Cooperate}, \text{Defect}\}$. As we move to more complex games, we will make a clearer distinction between the two.

Definition 1.4 (Strategy profiles). *Strategy profiles are vectors whose each element represents a strategy of a player. A strategy profile is denoted* $s \in S = \times_i^N S_i$. S *is the Cartesian product of the strategy spaces of all players. An element* s *of* S *thus has the same dimension as the number of players. It is a list that specifies a strategy for each player.*

You can think of strategy profiles as possible outcomes of a game. A strategy profile in prisoner's dilemma is $s = \{\text{Defect}, \text{Defect}\}$. The entire strategy space is $S = \{\{\text{Cooperate}, \text{Cooperate}\}, \{\text{Cooperate}, \text{Defect}\}, \{\text{Defect}, \text{Cooperate}\}, \{\text{Defect}, \text{Defect}\}\}$. Also useful is the notation $s_{-i} \in S_{-i} = \times_{j \neq i}^N S_j$, strategy profile for all players except for player i. This will be useful while talking about a player's opponents' strategies. Notice also that we can write $S = (S_i, S_{-i}), s = (s_i, s_{-i})$.

Definition 1.5 (Payoff functions). *A payoff function for player* i*, denoted* $u_i(s)$ *determine the utility that player derives from an outcome of the game,* s.

For instance, for the prisoner's dilemma, we have that $u_1(\text{Cooperate}, \text{Cooperate}) = -2$.

With these definitions in mind, notice the information conveyed to us by the strategic or normal form representation of the prisoner's dilemma game: 1) the players, 2) the actions they can each take, 3) possible outcomes of the game, so the combination of different actions available to players, and 4) the payoffs they would receive for each combination of actions chosen. Further, implicit in this representation is the idea that the players choose their actions simultaneously, so the structure of the game.

Definition 1.6 (Strategic/normal form game). *Formally, a strategic or normal game consists of the following elements:*

1. *A finite set of N players,* $i \in \mathcal{I}$, $\{1, 2, ...N\}$,
2. *For each player* i*, a finite (pure) strategy space* S_i*, which is a list that consists of the actions available to them and gives rise to set of strategy profiles* S,
3. *For each player* i*, a payoff function* $u_i(s)$*, which represents the utility that Player* i *derives from the outcome of the game* s,

   and can be denoted by $G = \{\mathcal{I}, S, \{u_i(.)\}_{i\in\mathcal{I}}\}$.

2 Solution Concepts

So far, we have described what a strategic/normal form game is. A game describes the players, the actions they can take, possible outcomes, the associated payoffs, and the rules of the game. As you may have noticed, what the players actually do is not a part of the game. For that, we turn to solution concepts. Solution concepts offer answers to the questions "How can we expect the players to behave? What are reasonable outcomes of the game?"

What we deem to be a reasonable solution depends on which solution concept we consider. You may have heard of Nash equilibrium, which is a solution concept. When we consider a specific solution concept, we require certain conditions, criteria on the strategy profiles. The solution of the game is the set of strategy profiles that survive these criteria. For static games, we consider three solution concepts, two of which related: 1) rationalizability and dominance, and 2) Nash equilibrium.

2.1 Dominance and Rationalizability

Dominance and rationalizability are two very closely related notions that allow us to derive predictions using only the structure of the game and the common knowledge of rationality. It is a very natural starting point to solve the games, as we ask ourselves what are the strategies that a rational player would definitely play, and what are those that they would never play. The underlying idea behind the two concepts is that the rationality of all players is common knowledge, so that players expect other players not to take actions that are **dominated** by other actions, and expect other players to reason the same way about all other players. We define these notions more formally, and introduce the two solution concepts.

2.1.1 Iterated Elimination of Strictly Dominated Strategies and Strict Dominance Equilibrium

The solution concept **iterated elimination of strictly dominated strategies** looks at the game from the point of view of an individual player i. We ask what the player would do without knowing anything about the actions of the opponents. The only assumption here is that Player i knows that her opponents are rational. To define a solution, she starts by eliminating the strategies that will always yield a lower payoff, regardless of what her opponents do. After these strategies are taken out, we are left with another game G^1. She knows that the other players are rational, so that they would do the same. We proceed to take out the strategies for the other players, which, given the remaining strategies of player i, always yield a lower payoff. We are now left with the game G^2. We repeat this exercise until there are no more strategies that can be taken out, the game G^∞ . By the iterated elimination of strictly dominated strategies, the remaining strategies can be reasonably played out in a game.

We now introduce some terminology related to this solution concept.

Definition 2.1 (Dominated strategies)**.** *A pure strategy s_i is said to be weakly (strongly/strictly) dominated for player i in game G if there exists another strategy s'_i, such that, for all $s_{-i} \in S_{-i}$, we have that*

$$u_i(s'_i, s_{-i}) \geq (>) u_i(s_i, s_{-i})$$

with strict inequality for at least one s_{-i}.

Definition 2.2 (Dominant strategies)**.** *A pure strategy s_i is said to be weakly (strongly/strictly) dominant for player i in game G if for all other strategies $s'_i \in S_i$, we have that*

$$u_i(s_i, s_{-i}) \geq (>) u_i(s'_i, s_{-i})$$

with strict inequality for at least one s_{-i}.

Taking another look at the prisoner's dilemma, we see that the strategy **Cooperate** always yields a lower payoff, so it is strictly dominated. Eliminating it for first Prisoner 1 and then Prisoner 2, we are left with the only strategy profile that survives the iterated elimination of strictly dominated strategies: $s = \{\text{Defect}, \text{Defect}\}$. Notice that Defect is a strictly dominant strategy.

Prisoner 1 \ Prisoner 2	Cooperate	Defect
Cooperate	-2,-2	-10,0
Defect	0,-10	-3,-3

When, as in prisoner's dilemma, the procedure of iterated elimination results in a single strategy left for all players, the strategy profile consisting of these strategies is called a **strict dominant strategy equilibrium**. Such a game is said to be **dominance solvable**.

Let us formalize the process defined in the example above.

Iterated Dominance:

Step 0: Define $S_i^0 = S_i$

Step 1: Define $S_i^1 = \{s_i \in S_i^0 | \nexists s_i' \in S_i^0 \ \textit{with} \ u_i(s_i', s_{-i}) > u_i(s_i, s_{-i}) \forall s_{-i} \in S_{-i}^0\}$

S_i^1 is the set of strategies which are not strictly dominated.

Step $k+1$: Define $S_i^{k+1} = \{s_i \in S_i^k | \nexists s_i' \in S_i^k \ \textit{with} \ u_i(s_i', s_{-i}) > u_i(s_i, s_{-i}) \forall s_{-i} \in S_{-i}^k\}$

S_i^{k+1} is the set of strategies which are still not strictly dominated when you know your opponent uses something in S_{-i}^k.

Clearly, this process defines an iterated deletion of dominated strategies. It requires player to be rational, to know that they are rational, etc.

Final step:

$$S_i^\infty = \cap_{k=0}^\infty S_i^k$$

Note that the process must eventually stop because each S_i is finite and the sets get only smaller.

Definition 2.3. *G is **solvable by pure strategy iterated strict dominance** if S_i^∞ contains a single element for each i.*

2.1.2 Iterated Elimination of Weakly Dominated Strategies and Weak Dominance Equilibrium

We could also argue that no player would play a strategy that yields the same payoff as some other strategy, and in some instances a lower payoff. We can do the same procedure of iterated elimination with weakly dominated strategies. It is however a less desirable procedure, because unlike the iterated elimination of strictly dominated strategies, the order of elimination may change the results. The next game provides an example.

Table 2.1: Osborne and Rubinstein 1994, p.63

Player 1 \ Player 2	L	R
T	1,1	0,0
M	1,1	2,1
B	0,0	2,1

Example 2.1.

- Take T out, which is weakly dominated by M.
- Eliminate L, weakly dominated by R
- We are left with M and B for Player 1, and L for Player 2

Alternatively,

- Take B out, which is weakly dominated by M.
- Eliminate R, weakly dominated by L
- We are left with T and M for Player 1, and R for Player 2

Weak Dominance

While the procedure of iterated weak dominance may have undesirable results, weak dominant strategy equilibria deliver important results. Well known and important results include the second price auction and Bertrand Nash equilibrium.

2.1.3 Rationalizability

You may encounter the term rationalizability in this context. For this class, our focus will be mainly on dominant strategy equilibria, and iterated elimination of strict dominance. For rationalizability, you should know that it is a solution concept that will lead to the same strategies as the iterated elimination of strict dominance. If you wish, you can just read the definition of best responses, and skip this part, where we show why the two are equivalent.

Definition 2.4 (Best response)**.** *A strategy s_i is best response to the opponents' strategies s_{-i} if*

$$u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i}) \quad \forall s_i' \in S_i$$

and we write $BR_i(s_{-i}) = \{s_i\}$.

We use the set notation, as there may be multiple best responses to a strategy profile s_{-i}. Best responses are set-valued functions, so correspondences. They take the strategies of other players as inputs, and give us the set of strategies that are payoff maximizing given this input.

A strategy s_i is said to be **never best response** if there is no s_{-i} such that the best response would be s_i.

Rationalizability is a very related concept to, and is in fact the contrapositive of dominance. We ask ourselves the following question: What are all the equilibrium strategies that a rational player can play? The answer is that a rational player would only play those strategies that are best responses to some beliefs he may have of his opponents. Believing someone is rational is equivalent to thinking that whatever that person plays is a best response. So a rational player i would only play strategies that are best responses to the strategies that the other players play, which are in turn best responses to the strategies of the player i and so on.

The contrapositive of this statement is the following: a rational player should not play any strategy that is not a best response to some belief about his opponents' strategies: no player plays never best responses. Furthermore, since she knows that the opponents are rational, these beliefs are not arbitrary. In a two-person game, Player 1 would not expect Player 2 to play a strategy that would not be a best response to a strategy that Player 1 would play. Player 2 reasons the same way about Player 1, and so on.

You can see that the contrapositive of rationalizability is the same as that of elimination of dominated strategies. As such, the procedure of iterated elimination of dominated strategies will result in **rationalizable strategies**. This result holds only for finite strategic form games. Indeed, in finite strategic form games **never best responses** are equivalent to **strictly dominated strategies**. Same equivalence does not hold between best responses and strictly dominant strategies: a strategy that is not strictly dominant may be a best response to some strategy.

2.2 Nash Equilibrium

While iterated elimination of dominated strategies is an intuitive way to deliver predictions, in many games there are no strategies that are dominated. Nash equilibrium is by far the most commonly used solution concept used in game theory. It defines a steady state of that game. We do not discuss how the game reaches that steady state. We merely require all players to hold beliefs that are exactly correct on the equilibrium strategies of other players, and that these equilibrium strategies be payoff maximizing given the equilibrium strategies of the opponents. Simply put, everyone knows exactly what the others are doing, and given these strategies, there is no incentive to play another strategy for any of the players, i.e there are no **profitable deviations**.

Definition 2.5 (Pure strategy Nash equilibrium (PSNE))**.** *A strategy profile s^* is a pure strategy Nash equilibrium of the game $G = \{\mathcal{I}, S, \{u_i(.)\}_{i\in\mathcal{I}}\}$ if*

- $u_i(s_i^*, s_{-i}^*) \geq u_i(s_i', s_{-i}^*) \quad \forall s_i' \in S_i \quad \forall i \in \mathcal{I} \implies u_i(s^*) = max_{s_i' \in S_i} u_i(s_i', s_{-i}^*)$
- $\forall i \in \mathcal{I}$*, player* i *perfectly anticipates* s_{-i}*.*

Notice that the first criterion for Nash equilibrium is the same as the definition we had for best responses in Definition 2.4. We can thus express Nash equilibrium in terms of best responses.

Pure strategy Nash equilibria of a game is the set of strategy profiles that survives the criteria we defined in Definition 2.5, so we can write it as

$$s^{NE} = \{s \in S : s_i \in BR_i(s_{-i}) \quad \forall i \in \mathcal{I}\} \tag{1}$$

There are two ways of formally showing that a strategy profile $s = (s_i, s_{-i})$ is a Nash equilibrium.

1. Show that there are **no profitable deviations**: given s_{-i}, no strategy s_i' yields a higher payoff than s_i
2. Show that all s_i are best responses to opponents' strategies: that all $s_i \in s$ satisfies $BR_i(s_{-i}) = s_i$.

Doing only one would suffice.

Let's give an example to apply the notation we have seen so far.

Example 2.2 (Piatti e Fagotti or Canteen)**.** *On Fridays, Neus and Leon want to have lunch together. Neus has a preference for the canteen as she wants to have fish once a week. Leon prefers the Piatti e Fagotti schiacciata over the fish meal. They both prefer having lunch together to having any of the meals alone.*

Table 2.2: The Friday Lunch Dilemma: Piatti e Fagotti or Canteen

Neus \ Leon	PF	Canteen
PF	1,2	0,0
Canteen	0,0	2,1

Finding PSNE

An easy way to find pure strategy NE in a two-person game in matrix form:

1. For each column, find the rows that yield highest payoffs. You are effectively finding Player 1's best responses for each of Player 2's strategies.

2. For each row, find the columns that yield highest payoffs, finding best responses for Player 2.

3. The elements of the matrix with both strategies marked are the Nash equilibria of the game.

Table 2.3: Best responses for Neus

Neus \ Leon	PF	Canteen
PF	$1^*, 2$	0,0
Canteen	0,0	$2^*, 1$

Table 2.4: Adding best responses for Leon

Neus \ Leon	PF	Canteen
PF	$1^*, 2^*$	0,0
Canteen	0,0	$2^*, 1^*$

We end up with (PF, PF) and (C,C), and claim that these are the two PSNE of the game. Now that you have it, to show it formally, follow one of the two ways we described above: either show that there are no profitable deviations, or that all strategies in the profile are best responses to each other. To show that a profile is **not** a NE, finding **one** profitable deviation, or **one** strategy that is not a best response is enough.

- Is (PF, PF) a Nash equilibrium? Yes.

1. There are no profitable unilateral deviations for either player from (PF, PF), because:

$$u_N(PF, PF) \geq u_N(C, PF)$$
$$u_L(PF, PF) \geq u_L(PF, C)$$

2. $BR_N(PF) = \{PF\}, BR_L(PF) = \{PF\},.$

- Is (C, C) a Nash equilibrium? Yes.

1. There are no profitable unilateral deviations for either player from (C, C), because:

$$u_N(C, C) \geq u_N(PF, C)$$
$$u_L(C, C) \geq u_L(C, PF)$$

2. $BR_N(C) = \{C\}, BR_L(C) = \{C\}.$

- To show that (PF, PF) and (C, C), we need to show that no other strategy profile is a Nash Equilibrium.
  - (PF, C) cannot be a Nash Equilibrium, because
    1. $u_N(C, C) > u_N(PF, C)$: Player 1 has a profitable deviation.
    2. $PF \notin BR_C(C)$, Neus' best response to C is not PF, but C.
  - (C, PF) cannot be a Nash Equilibrium, because
    1. $u_N(PF, PF) > u_N(C, PF)$: Player 1 has a profitable deviation.
    2. $PF \notin BR_L(C)$, Leon's best response to C is not PF, but C.

Formally, we can write the set of (pure strategy) Nash equilibria as

$$s^* = \{(PF, PF), (C, C)\}^1$$

[1] Compare this expression with the definition of the set of Nash equilibria in Equation 1 and make sure you understand the notation, that you see why the pairs (PF, PF), (C,C) fulfill the criteria set in the definition of the set.

2.2.1 Mixed Strategies

[2]

Consider the following games:

- Penalty Shots:

		Pickford	
		L	R
Bonucci	L	-1,1	1,-1
	R	1,-1	-1,1

- Government Auditing Game:

		Player 2	
		Audit	Don't
Player 1	Honest	10,9	10,10
	Cheat	-35,10	15,5

- D-Day Invasion:

		Axis	
		Normandy	Calais
Allies	Normandy	-2,2	1,-1
	Calais	3,-3	-1,1

The common feature of these games is that if a player's strategy is known her opponent can take advantage. In Penalty Shots, if Pickford knows that Bonucci will shoot the ball to the left side, he will move to left. Bonucci's best response would then be to shoot the ball to the right side of the goal, in which case Pickford would move

[2]From Arda Gitmez' TA notes for Game Theory at MIT.

right, Bonucci responding by shooting to the left and so on. There are no PSNE. Check the non existence of the PSNE in the remaining games.

What might a player then do? One solution is playing a **mixed strategy**.

Let $G = (\mathcal{I}, S, u)$ be a finite normal form game.

Definition 2.6. *A **mixed strategy** σ_i for player i is a probability distribution on S_i. i.e. if S_i is finite, a mixed strategy is a function $\sigma_i : S_i \to \mathbb{R}^+$ s.t.*

$$\sum_{s_i \in S_i} \sigma_i(s_i) = 1.$$

Notation To describe mixed strategies:

1. In terms of functions: $\sigma_i(L) = \frac{1}{2}, \sigma_i(R) = \frac{1}{2}$
2. In terms of vectors: $(\sigma_i(s_{i1}), \sigma_i(s_{i2}), \ldots, \sigma_i(s_{iN}))$, e.g. $(\frac{1}{2}, \frac{1}{2})$. If you use this notation, make sure to specify which ordering you are following.
3. More informally: $\frac{1}{2}L + \frac{1}{2}R$, think of L and R as the degenerate probability distributions that play only L or R respectively—thus $\frac{1}{2}L + \frac{1}{2}R$ is also a probability distribution.

More notation:

- Write $\Sigma_i = \Delta(S_i)$ for the set of i's mixed strategies .
- Write Σ for $\Sigma_1 \times \Sigma_2 \times \ldots \times \Sigma_I$.
- A **mixed strategy profile** $\sigma \in \Sigma$ is an I-tuple $(\sigma_1, ..., \sigma_I)$ with $\sigma_i \in \Sigma_i$.

- Write $u_i(\sigma_i, \sigma_{-i})$ for player i's expected payoff when she uses mixed strategy σ_i and all other players play as in σ_{-i},

$$u_i(\sigma_i, \sigma_{-i}) = \sum_{s_i \in S_i, s_{-i} \in S_{-i}} u_i(s_i, s_{-i})\sigma_i(s_i)\sigma_{-i}(s_{-i})$$

  - For example, in the penalty shots game could write: $u_1(\frac{2}{3}L + \frac{1}{3}R, L)$

Definition 2.7. *A strategy profile σ^* is a mixed strategy Nash equilibrium (MSNE) of the game $G = \{\mathcal{I}, \Sigma, \{u_i(.)\}_{i \in \mathcal{I}}\}$ if*

- $u_i(\sigma_i^*, \sigma_{-i}) \geq u_i({\sigma^*}'_i, \sigma_{-i}) \quad \forall \sigma_i' \in \Sigma_i \quad \forall i \in \mathcal{I}$
- *$\forall i \in \mathcal{I}$, player i perfectly anticipates σ_{-i}^*.*

You can also interpret this as the pure strategy Nash Equilibrium of a game where each agent's strategy set is the set of her mixed strategies.

Finding MSNE

Definition 2.8. *In a finite game, the support of a mixed strategy σ_i, $Supp(\sigma_i)$, is the set of pure strategies to which σ_i assigns positive probability.*

$$Supp(\sigma_i) = \{s_i \in S_i | \sigma_i(s_i) > 0\}$$

Proposition 2.1. *If σ^* is a NE, and $s_i', s_i'' \in Supp(\sigma_i^*)$, then*

$$u_i(s_i', \sigma_{-i}^*) = u_i(s_i'', \sigma_{-i}^*).$$

In words, players are indifferent between the strategies they employ with positive probability, given the mixed strategy of the opponents. We will talk about the intuition, but briefly, you should think that if a strategy gives a lower payoff than others, you are better off not playing it.

Let's find the MSNE of two of the games above to illustrate.

1. In **Penalty Shots**, we want both players to be indifferent between Left and Right. Denote with $\sigma_B(L)$ the probability that Bonucci shoots to left, and $\sigma_P(L)$ that Pickford leans to left.

   Indifference condition for Bonucci:

$$\begin{aligned} u_B(L, \sigma_P^*(L)) &= u_B(R, \sigma_P^*(L)) \\ \sigma_P^*(L)(-1) + (1 - \sigma_P^*(L))1 &= 1\sigma_P^*(L) + (1 - \sigma_P^*(L))(-1) \\ \sigma_P^*(L) &= \frac{1}{2} \end{aligned}$$

   Indifference condition for Pickford:

$$\begin{aligned} u_P(L, \sigma_B^*(L)) &= u_P(R, \sigma_B^*(L)) \\ 1\sigma_B^*(L) + (1 - \sigma_B^*(L))(-1) &= (-1)\sigma_B^*(L) + 1(1 - \sigma_B^*(L)) \\ \sigma_B^*(L) &= \frac{1}{2} \end{aligned}$$

   We can write the MSNE as $\sigma^* = (\sigma_B^*(L), \sigma_P^*(L)) = (\frac{1}{2}, \frac{1}{2})$.

2. In **D-Day Invasion**

   Indifference for **Allies**,

$$\begin{aligned} u_1(N, \sigma_2^*(N)) &= u_1(C, \sigma_2^*(N)) \\ \sigma_2^*(N)(-2) + (1 - \sigma_2^*(C))1 &= 3\sigma_2^*(N) + (1 - \sigma_2^*(N))(-1) \\ \sigma_2^*(N) &= \frac{2}{7} \end{aligned}$$

   Indifference condition for **Axis powers**:

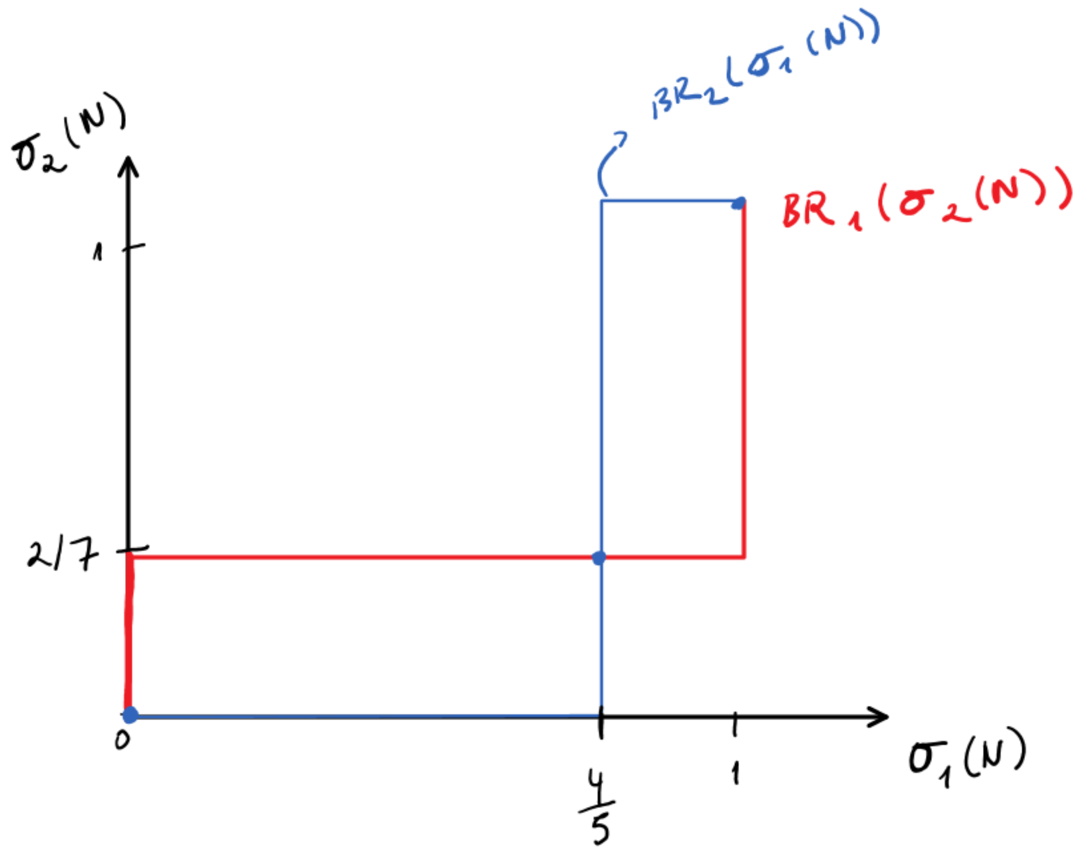
$$\begin{aligned} u_2(N, \sigma_1^*(N)) &= u_2(C, \sigma_1^*(N)) \\ 2\sigma_1^*(N) - 3(1 - \sigma_1^*(N)) &= (-1)\sigma_1^*(N) + 1(1 - \sigma_1^*(N)) \\ \sigma_1^*(N) &= \frac{4}{5} \end{aligned}$$

   We can write the MSNE as $\sigma^* = (\sigma_1^*(N), \sigma_2^*(N)) = (\frac{4}{5}, \frac{2}{7})$.

Setting the indifference conditions is the most practical way of finding MSNE. After you do, you can write the best response correspondences as:

$$BR_1(\sigma_2(N)) \equiv \sigma_1^*(N) = \begin{cases} 0 & \text{if } \sigma_2(N) < \frac{2}{7} \\ \in (0,1) & \text{if } \sigma_2(N) = \frac{2}{7} \\ 1 & \text{if } \sigma_2(N) > \frac{2}{7} \end{cases} \tag{2}$$

$$BR_2(\sigma_1(N)) \equiv \sigma_2^*(N) = \begin{cases} 0 & \text{if } \sigma_1(N) < \frac{4}{5} \\ \in (0,1) & \text{if } \sigma_1(N) = \frac{4}{5} \\ 1 & \text{if } \sigma_1(N) > \frac{4}{5} \end{cases} \tag{3}$$



Intuition for Mixed Strategies

The notion that players are indifferent between the strategies, or that a player assigns probabilities to strategies such that the opponents are indifferent may not seem intuitive at first. Very simply put, if you are not indifferent, and are leaning towards one of the strategies, then you should rather play that strategy only. Likewise, if there is a strategy that yields you a lower payoff given the mixed strategy of your opponents, you should just not play that strategy at all, and assign zero probability. Mixing means that there are multiple pure strategies that you may end up playing given what the other player does. In that case, it would only make sense to do so if any one of them give you the same expect payoff.

You can think of playing a mixed strategy as deciding on a device, and following the signals that the device gives. Usually in game theory when we talk about devices, we mean a random variable whose potential values and distribution are known: coin toss, throwing a die, the weather, the lunch at the canteen, the first person you encounter in a given day. These are all random variables and can be used as devices.

For instance, in the D-Day Invasion game, imagine that the Allies throw a five-sided die, commit to invade Normandy if it lands on 4 or below, and to invade Calais if it is 5. Knowing what the Axis powers do, which is to send forces to Normandy with probability $\frac{2}{7}$, they should not be secretly wishing for the die to land on 5, or on a number below 4 at the moment of throwing the die. If they did, then they would directly play what they wish for. Mixing would then not be an equilibrium strategy, i.e there would be a profitable deviation. Even though invading Calais when Axis forces are in Normandy may yield a higher payoff, there is a lower chance that the Axis forces will actually be situated in Normandy. With the probabilities assigned by the opponent, a player becomes indifferent between the two strategies.

In order to explain how to solve for MSNE, we often say that the players choose their strategies so to make the other players indifferent. While helpful to solve for the MSNE, it may be confusing to think that a player wants to make her opponent indifferent. The players do not *choose to make the other player indifferent.* As long as they are mixing, they themselves do not care about what probability they are assigning. From the individual player's point of view, there is no reason to assign these particular probabilities *as long as they are mixing.* Mixed Nash equilibrium is a concept that is related to the entire game, not to individuals' rationality only: only with these probabilities are we at the steady state of the game. Once these probabilities are known to all players, no one can have a profitable deviation.

You can see in the graph of the best response correspondences that we derived in Equations 2 and 3. It allows you see that, at the probabilities that constitute the MSNE, players' best responses are the entire set of probabilities laying between 0 and 1, but only the intersection constitutes a steady state, so a MSNE.

Note!: PSNE are degenerate MSNE. So if you are asked to find MSNE, that includes finding mixed strategies where only one of the strategies is played, and with probability 1, i.e finding the PSNE. A finite normal form game can have both pure and mixed strategy equilibria. But it definitely will have at least one mixed strategy equilibrium, which may be degenerate, so a PSNE.

To gain further insight to why players play mixed strategies, and how we can interpret choosing probabilities, you can check section 3.2 of the handbook by Osborne and Rubinstein.

2.3 Discussion: Dominance, Rationalizability and Nash Equilibrium

While solving the game by dominance/rationalizability, we determined the set of strategies that can reasonably be expected to played in a game **without knowing what**

the opponents play. While doing so, we look at the game from the perspective of an individual player. With Nash equilibrium, we are more restrictive in that we look at the entire game and decide what strategies can be played together. We require **everyone to know what the opponents are playing**. In the former, we eliminate strategies, in the latter we eliminate strategy profiles. With NE, we are more restrictive, but this is precisely what allows us to restrict the potential outcomes of the game.

To illustrate this point, consider again the PF or Canteen game,

Table 2.5: The Friday Lunch Dilemma: Piatti e Fagotti or Canteen

Neus \ Leon	PF	Canteen
PF	1,2	0,0
Canteen	0,0	2,1

Notice that there are no strictly or weakly dominated strategies for either player. If we were to solve this game by dominance or rationalizability, a reasonable outcome of the game would be {PF, C} since there is no reason to argue that a rational player would not play PF or C. In fact, all outcomes are rationalizable/cannot be eliminated by dominance: {PF, C}, {PF, PF}, {C, C} and {C, PF}. {PF, C} and {C, PF} are however not Nash equilibria as we already argued above.

The following relationships exist between dominance and Nash equilibrium.

Proposition 2.2 (Strict dominance and Nash equilibrium)**.** *If s^* is a pure strategy Nash equilibrium of game G, then it $s^* \in S_1^\infty \times S_2^\infty \times ...S_N^\infty = S^\infty$.*

What this proposition tells us is that if a strategy equilibrium is a Nash equilibrium of the game, it cannot be eliminated by iterated elimination of strict dominance. You are asked to prove the proposition in Problem Set 1, Question 2.

Proposition 2.3 (Weak dominance and Nash equilibrium 1)**.** *If s^* is a strategy that* *survived iterated elimination of weakly dominated strategies, it is a Nash equilibrium of the original game.*

The proof is left as an exercise in Q2 of PS1.

Proposition 2.4 (Weak dominance and Nash equilibrium 2)**.** *A Nash equilibrium of game G can be eliminated in the process of removal by iterated weak dominance.*

An example would suffice to prove this proposition, which you are asked to do in Q2 of PS1. You can also check it with the Example 2.1. Try find a different game for the PS.

2.4 Refinement: Trembling Hand Perfect Equilibrium

Often, we encounter games where we have a multitude of Nash equilibria. In these cases, it may be useful to further eliminate some of the potential outcomes using refinements. One of such refinements is the *trembling hand perfection.*

Trembling Hand Perfect Equilibrium (THPE) requires Nash equilibrium strategies to be robust to small perturbations, mistakes; to hand trembling. An equilibrium strategy profile is robust to these errors, i.e **trembling hand perfect**, if all players would choose to play their strategies even if other players make small mistakes.

Before we formally define THPE, let us give a motivating example[3].

Example 2.3 (Cuban missile crisis game)**.** *Consider the Cuban missile crisis game. There are two players: the US and the USSR. The USSR has to decide whether to station missiles in Cuba (S) or not (N). If the USSR chooses N, the US does nothing. If the USSR chooses S, US has to decide if to start a nuclear war (W) or forfeit (F).*

[3]from Omer Tamuz lecture notes

Table 2.6: Cuban missile crisis game

USSR \ US	W	F
S	-1000, -1000	1,-1
N	0,0	0,0

In this example, there are two Nash equilibria: (N,W), and (S,F). Consider the (N,W) equilibrium. Suppose there is a small probability ε that USSR's hand trembles, and that they play S. For the US, playing W yields $-1000\varepsilon + (1-\varepsilon)0 = -1000\varepsilon$. Playing F yields $-\varepsilon$. For any arbitrary $\varepsilon > 0$, playing F is a best response for the US: $-\varepsilon > -1000\varepsilon \quad \forall \varepsilon > 0$. We say that (N,W) is not trembling hand perfect. It is not robust to a tiny probability of the opponent making a mistake, no matter how small that probability actually is.

Now let us formally define THPE.

Definition 2.9 (Completely/fully mixed strategy profile)**.** *A strategy σ_i is said to be completely or fully mixed if it assigns positive probability to all pure strategies, i.e, $supp(\sigma_i) = S_i$.*

Definition 2.10 (Trembling hand perfect strategy profile 1)**.** *A strategy profile σ is trembling hand perfect if there exists a sequence $(\sigma^k)_{k=0}^{\infty}$ of fully mixed strategies which converges to σ, i.e $(\sigma^k)_{k=0}^{\infty} \to \sigma$, such that for each player i, σ_i is a best response to σ_i^k for all k.*

Let us break down Definition 2.10 to gain intuition to THPE.

- We have a Nash equilibrium strategy profile $\sigma = (\sigma_i, \sigma_{-i})$. It can be a pure or a mixed strategy profile.
- A mistake is playing anything else than the NE strategy, σ_i.

- We require the strategies profiles to be fully mixed so to include the probability of mistake/error/irrationality: any strategy can be played
- You can think of the elements in the sequence $(\sigma^k)_{k=0}^{\infty}$ as the times that this game has been played (by different players each time): as k is 1, no one has ever seen this game being played, as it tends to infinity, it has been played many times by a large sample of players.
- σ is the steady state strategy profile to which the sequence limits, i.e converges to.
- Probability of making a mistake decreases as the sequence converges, so as k tends to infinity, and becomes 0 at the limit.
- We want to ensure that even when there is a small probability that the opponents make mistakes, the players find it optimal to play their equilibrium strategy.
- A strategy profile is trembling hand perfect if we can find at least one sequence which converges to it even if there may be small perturbations, small probabilities of irrationality. If we cannot find at least one, the strategy profile could only be an outcome in a world with perfect rationality, perfect knowledge of equilibria. It means that even when no one has encountered this game being played before, players must directly play the Nash equilibrium. It is not robust to the possibility of small errors: it is not trembling hand perfect.

What do these sequences look like in practice?

- Take $\sigma = (\sigma_{USSR}(N), \sigma_{US}(W)) = (1, 1)$
- A sequence converging to σ is $(\sigma^k)_{k=0}^{\infty} = (\sigma_{USSR}^k(N), \sigma_{US}^k(W))_{k=0}^{\infty}) = ((1-\varepsilon^k), (1-\varepsilon^k))_{k=0}^{\infty}$ for $\varepsilon < 1$
- Another one could be $((1 - \frac{1}{5k}), (1 - \frac{1}{10k}))_{k=0}^{\infty}$

- See that they converge to (1,1) as k tends to infinity: possibility of error is larger at the beginning, and it approaches to 0.

In the Example 2.3, it may seem like we did not look at any sequence to show that (N,W) is not trembling hand perfect. In two player games where both players have two actions, the two are equivalent. For more complicated games, you should use the sequences to determine THP.

How about the mixed strategy NE? We have the following two propositions that have important implications.

Proposition 2.5 (Trembling Hand Perfection, Mixed Strategies and Weak Domination)**.** *A strategy profile in a finite two-player strategic game is a trembling hand perfect equilibrium if and only if it is a mixed strategy Nash equilibrium and the strategy of neither player is weakly dominated.*

The intuition is simple for why the strategy of neither player should not be weakly dominated: a weakly dominated strategy is never a best response to a vector of completely mixed strategy profiles of the opponents by definition.

What this proposition tells us is that if you found a mixed strategy equilibrium that gives no probability to a weakly dominated strategy for any of the players, it is trembling hand perfect. **Any mixed NE that assigns zero probability to weakly dominated strategies is THP.** Recall that pure strategies are degenerate mixed strategies, and you should see why (S,F) is THP, and (N,W) is not. **You can use this proposition in the problem sets and in the exam.** You then see if a strategy profile is THP or not, and can then come up with any sequence that proves your point.

Proposition 2.6 (Existence of Trembling Hand Perfect Equilibrium)**.** *Every finite strategic game has a trembling hand perfect equilibrium.*

2.5 Correlated Equilibrium

[4]

Consider the Canteen or PF game. There exists a mixed equilibrium in this game: $\sigma^{MSNE} = (\frac{1}{3}PF + \frac{2}{3}C, \frac{2}{3}PF + \frac{1}{3}C)$. *(Verify this yourself).*

Table 2.7: The Friday Lunch Dilemma: Piatti e Fagotti or Canteen

Neus \ Leon	PF	Canteen
PF	1,2	0,0
Canteen	0,0	2,1

The equilibrium seems to be suboptimal: 5 out of 9 times, so more than half of the time, the two players end up going to different places, obtaining 0. In expectation, the payoffs are $\frac{7}{9}$ for both players. This seems particularly unsatisfying since if they manage to meet up, at the very least they get 1. Is there a way in which we can obtain better expected payoffs for both players?

Correlated equilibrium is an answer to this question. In a correlated equilibrium, players are allowed to meet before playing the game, and come up with a *device* or a *random variable* that will send (potentially less than perfectly) correlated signals to the players.

In this game for example, imagine we were living in a city where the weather being sunny or rainy is equally likely, even though we thankfully do not. In this case, weather is a random variable, taking value "sunny" with probability $\frac{1}{2}$, and "rainy" with probability $\frac{1}{2}$ as well. The players could then meet up before playing the game, and commit to playing PF if it is sunny, and C if it is rainy. Each player knows that the other is playing by this device, and given the information they have, there would be no reason to deviate. Each player obtains $\frac{3}{2}$.

[4] Osborne and Rubinstein 3.3

This is a very simple example where the signals that the players observe are perfectly correlated. In general, correlated equilibrium allows for less than perfectly correlated signals. In these cases, I do not know exactly which signal the other player observes, but my signal tells me something about the signal they receive.

Here, you may notice that the concepts we use are very close to those for the mixed strategy equilibrium. Indeed, these two equilibria are closely related. For MSNE, we talked about each player conditioning her action on the outcome of a random variable. Think: two players are in their rooms tossing a coin, throwing a dice, being in different cities and observing the weather etc. In a correlated equilibrium, the players condition their actions on the same random variable whose outcome they may or may not perfectly observe. Think: two players play the game in different but close cities. The weather in one city that a player observes allows me to infer further information on the weather in the other city.

Before we formalise the correlated equilibrium, let us introduce some definitions.

Definition 2.11 (State Space). *A state space is the set of all possible events, and is denoted* Ω. *We write* $\Omega = \{\omega_1, \omega_2, ..., \omega_l\}$.

In the weather example, we have $\Omega = \{\text{sunny}, \text{rainy}\}$. The sum of the probabilities of each event ω, denoted $\pi(\omega_l)$, is 1.

Definition 2.12 (Information partition). *An information partition* $\mathcal{P}$ *of state space* Ω *is a collection of subsets* $P^i \in \Omega$ *so that every element* $\omega \in \Omega$ *occurs in exactly one of these subsets.*

Let's give an example to apply this notation.

Example 2.4 (Dice as randomization device). *Two players coordinate on an outcome using a die as a randomization device. The first player is almost blind and cannot recognize between 1,2 and 3, but can tell if one of those numbers is drawn that is smaller* *or equal than 3, and the same for the numbers greater than 3. The second player is tipsy and can only tell apart 1 and 6, and confounds all numbers in-between.*

We have

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$\pi(\omega_i) = \frac{1}{6} \quad \forall i = \{1, .., 6\}$

Player 1's partition $\mathcal{P}_1 = \{P_1^1 = \{1, 2, 3\}, P_1^2 = \{4, 5, 6\}\}$

Player 2's partition $\mathcal{P}_2 = \{P_2^1 = \{1\}, P_2^2 = \{2, 3, 4, 5\}, P_2^3 = \{6\}\}$.

Given the randomization and the information partitions, what would players do? Observing the signal, given what they can differentiate, they would use Bayes rule to infer the new probabilities for the events. If the die lands on 6, player 2 would know for sure that it is 6. Player 1 on the other hand cannot know for sure that it is 6. She can however know that the number is larger than 3. Applying the Bayes rule for the probability that the die has landed on 6, denoting with ω the true value on which die has landed, so the true state of the world, and with s the signal for the player,

$$P(\omega = 6 | s_1 = P_1^2 = \{4, 5, 6\}) = \frac{P(s_1 = P_1^2 = \{4, 5, 6\} | \omega = 6) P(\omega = 6)}{P(s_1 = P_1^2 = \{4, 5, 6\})} = \frac{1 * \frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

where the denominator comes from the fact that the total probability of the die landing on either 4,5 or 6 is $\frac{1}{2}$.

Try do the Bayesian update for when the die lands on 3 for both players.

With these basics in mind, let us define the correlated equilibrium.

Definition 2.13 (Correlated equilibrium)**.** *A correlated equilibrium of a strategic game G consists of*

1. *a finite probability space* (Ω, π) *where* Ω *is the state space,* π *probability measure over* Ω*,*

2. *for each player* $i \in \mathcal{I}$*, an information partition* $\mathcal{P}_i$ *of* Ω*,*

3. *for each player* $i \in \mathcal{I}$*, a function* $\sigma_i : \Omega \to A_i$*, with* $\sigma_i(\omega) = \sigma_i(\omega')$ *whenever we have* $\omega, \omega' \in P_i$ *for some* $P_i \in \mathcal{P}_i$*,*

   such that, for all $i \in \mathcal{I}$*, and any function* $\tau_i : \Omega \to A_i$*, whenever* $\omega \in P_i$ *and* $\omega' \in P_i$ *for some* $P_i \in \mathcal{P}_i$*, we have*

$$\sum_{\omega \in \Omega} \pi(\omega) u_i(\sigma_{-i}(\omega), \sigma_i(\omega)) \geq \sum_{\omega \in \Omega} \pi(\omega) u_i(\sigma_{-i}(\omega), \tau_i(\omega))$$

In words,

- State space, the probability measure over the state space, information partitions are all parts of the equilibrium.

- Further, for each player we have a function that assigns a probability to their actions in a given state of the world. **This function is the strategy**. A strategy assigns probabilities to actions conditional on the state of the world.

- For the states of the world between which the player cannot distinguish, she must play the same action: I cannot play one action if the die lands on 2 and another one if it lands on 3, if I cannot distinguish between these two events.

- Given these, in all states of the world, no player finds it optimal to deviate, and assign different probabilities to her actions, i.e. adopt a different strategy.

Let us give an example.

P2 / P1	L	R
T	6,6	2,7
B	7,2	0,0

In the game given by the first matrix, the PSNE give payoffs of (2,7), and (7,2). In mixed strategies, the NE yields payoffs $(4\frac{2}{3}, 4\frac{2}{3})$.

Consider now the following equilibrium: x, y, z are states of the world. We have $\Omega = \{x, y, z\}$. We have $\pi(x) = \pi(y) = \pi(z) = \frac{1}{3}$.

Let player 1's partition be $\mathcal{P}_1 = \{\{x\}, \{y, z\}\}$, and for player 2 let it be $\mathcal{P}_2 = \{\{x, y\}, \{z\}\}$.

This means: if the true state of the world is x, player 1 receives the signal, and she knows for sure that the state is x. Player 2 receives a signal as well, she can know that it is not z, but cannot distinguish if it is x or y. *(Can you reason similarly for y and z?)*

Consider now the following function σ_1:

- $\sigma_1(x) = B$
- $\sigma_1(y) = \sigma_1(z) = T$

So that Player 1 plays B whenever she observes x, and T whenever she receives a signal that the state is either y or z. **Strategy is a function that maps from the state of the world into actions.**

For Player 2, consider the following strategy:

- $\sigma_2(x) = \sigma_2(y) = L$

- $\sigma_2(z) = R$

Through the strategies σ_1 and σ_2, the states of the world translate into the outcomes of the game in the following manner:

P1 \ P2	L	R
T	y	z
B	x	-

P1 \ P2	L	R
T	1/3	1/3
B	1/3	-

Now that we have specified the state space, the information partitions, and the equilibrium strategies, we determine if it is an equilibrium by considering optimal behaviour given information.

Player 1's behaviour is optimal: in state x, player 1 knows that player 2 plays L and thus it is optimal for her to play B. No reason to deviate to T and get 2. In states y and z she assigns equal probabilities to player 2 using L and R, so that it is optimal for her to play T and obtain 4 in expectation. Playing B would lead to 3.5 in expectation. Symmetrically, player 2's behavior is optimal given player 1's, and hence we have a correlated equilibrium; the payoff profile is (5,5).

Note:

- Correlated equilibrium is a weaker notion than mixed strategy Nash equilibrium. As such, any MSNE is a correlated equilibrium, and can be constructed with a device.
- We considered cases where the signal tells the players to play a pure strategy, but it can equally signal them to play a mixed strategy.
- When constructing a correlated equilibrium, make sure to specify the state space, the probability measure over it, as well as the information partition for each

player.

Dynamic Games

3 Extensive Form Games

The strategic/normal form games we have seen so far does not allow for sequential structure of decision making: all actors decide on their strategies at once. In many games and real life situations, players perfectly or imperfectly observe each other's actions, and then carry out their strategies: think of games like chess, poker, and economic situations where one firm acts first, chooses an output, and the second firm can observe the output of its rival before choosing its output [5]

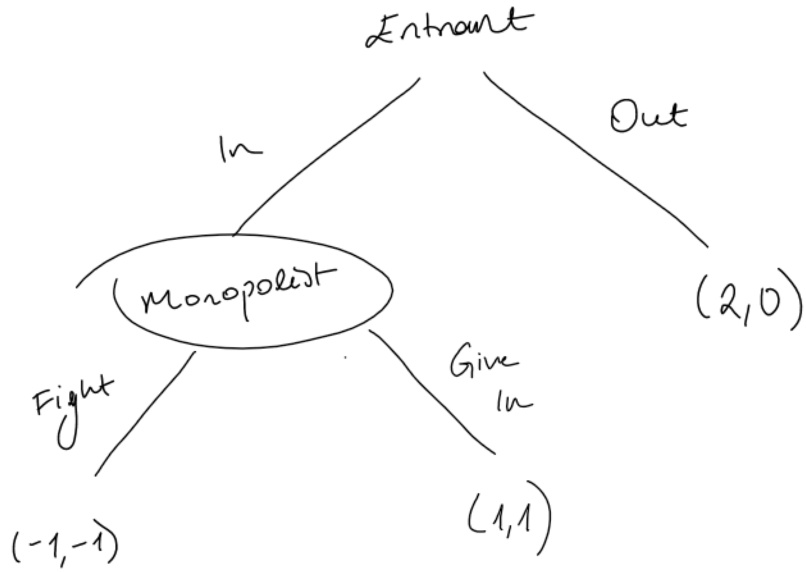
This class of games are called extensive form games. An extensive form game describes 1) entire set of players, 2) the order of moves, 3) which actions they are able to take each time they move, 4) what each player knows as they make decisions on their actions, 5) players' payoffs depending on the choices they make, and 6) probability distribution over any exogenous event.

Before we formalize this definition, let us give motivating examples.

Example 3.1 (Chain Store)**.** *A monopolist grocery store chain, Player 1, has a branch in a town. A local entrepreneur, Player 2, can enter the market by opening a new grocery store, or she can stay out. If she enters, the monopolist can react by accommodating the entrant, "giving in", or it can fight. If Player 2 decides not to enter, she gets 0,*

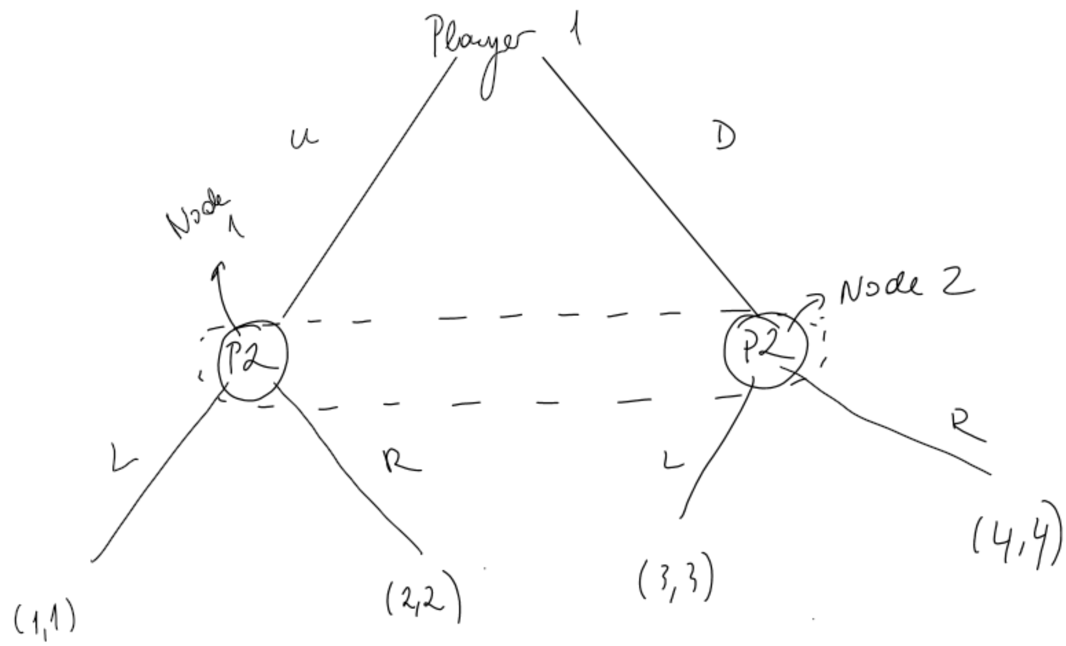
[5] This is the idea of Stackelberg equilibrium in a dupoly as opposed to Cournot equilibrium where both firms choose quantities at the same time.

and the monopolist gets 2. If she enters, and the monopolist fights, they each get -1, and if the monopolist gives in, they each get 1.



In this example, the monopolist can perfectly see what the entrepreneur has done before acting. This need not be the case. The following example depicts a situation where Player 2 cannot observe the action taken by Player 1.

Example 3.2 (Extensive Form of a Simultaneous Move Game)**.** *Player 1 gets to act first, and chooses between Up, Middle and Down. Player 2 then gets to play, and choose between Left and Right. But she cannot know what Player 1 has done before her. If Player 1 plays Up, the payoffs are (1,1) if Player 2 plays L and (2,2) if she plays R. If Player 1 plays D, then payoffs are (3,3) and (4,4) respectively for L and R.*

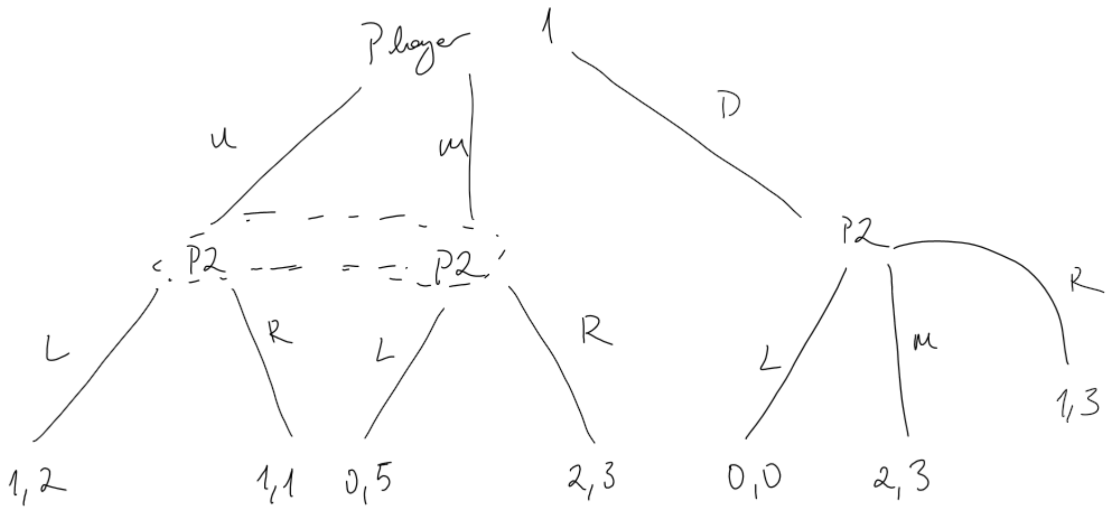


The dotted circle represents Player 2's **information set**: Player 2 cannot distinguish between the actions of Player 1, so node 1 and node 2 are in one information set.

In terms of information structure, we can allow for more complicated representations. In the following example, Player 2 can know if Player 1 moved Down, or not, but cannot distinguish between Up and Middle.

Example 3.3 (Extensive Form 3)**.**

Notice that Player 2 has to have the same actions available to her at two nodes if she is not able to distinguish between them. In the previous example, she cannot have the action M if Player 1 played Up, but not if Player 1 played Middle since this would allow P2 to infer the action of P1.

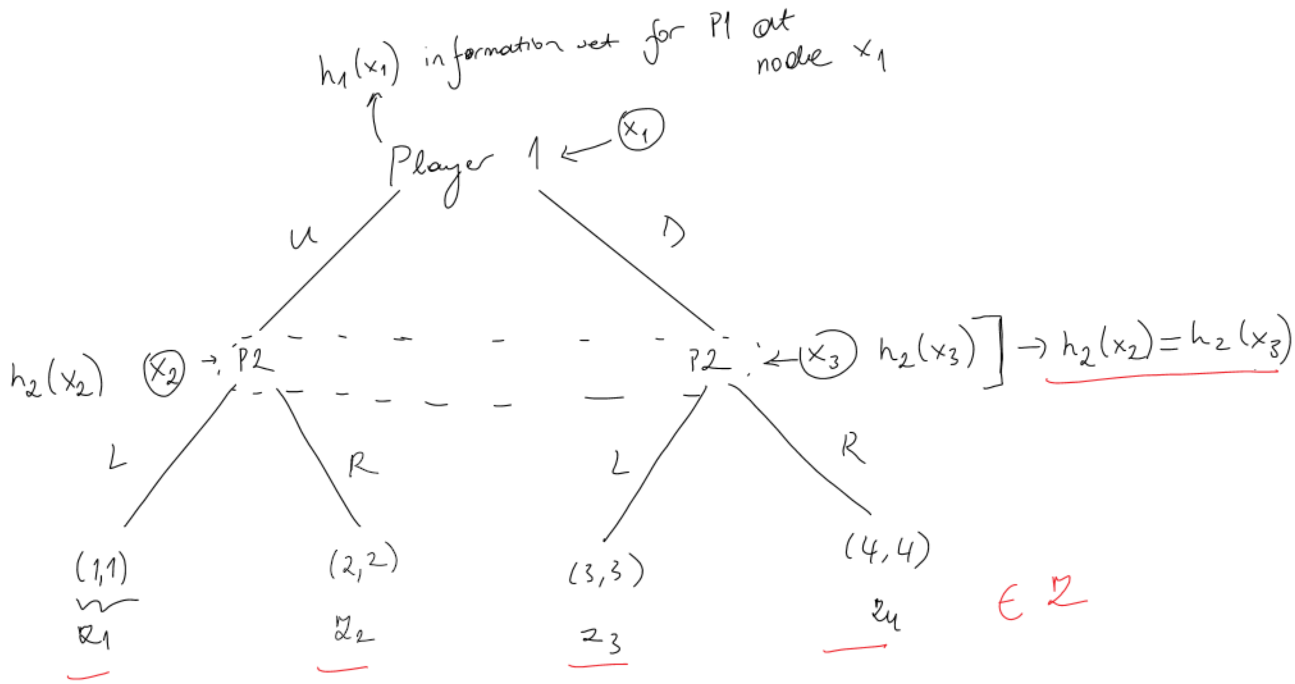


Let us now define the extensive form game formally.

Definition 3.1 (Extensive Form Games). *An extensive form game consists of*

1. *A finite set of N players, $i \in \mathcal{I}$, $\{1, 2, ...N\}$,*
2. *A finite game set of nodes $x \in X$, that form the game tree X, such that all nodes x have a single root,*
3. *A finite set of terminal nodes $z \in Z \in X$,*
4. *For each player i, information sets $H_i(x)$ for all nodes where i has the move, such that*
   - *where each $h_i(x) \in H_i$ defines a set of feasible actions $A_i(h_i(x))$, and consists of all the nodes between which player i cannot distinguish when they are at node x,*
   - *from each action $a_i(h_i)$ follows a unique node,*
   - *final nodes z are associated with singleton information sets for each player,*
5. *For each player i, payoff functions $u_i(z)$ that reflect the payoffs at a terminal node.*

Information Sets[6] Notice that the extensive form game, the game tree X is simply composed of the information sets of players. The information sets partition the game tree such that a node belongs at maximum to one information set. The interpretation of the information set $h_i(x)$ containing node x is that she is uncertain if she is at x or some other node $x' \in h_i(x)$.



All nodes $x_1, x_2, x_3, z_1, z_2, z_3, z_4$ are elements of the set that is the game tree X. $h_1(x_1) = x_1$ and is a singleton: Player 1 knows where she is. The information set at the first node is always singleton. Same goes for the information sets at the terminal nodes z_1, z_2, z_3, z_4: when the game ends, all of players know their payoffs. The information sets are correspondences of the nodes, and return nodes: $h_2(x_2) = h_2(x_3) = \{x_2, x_3\}$: when at x_2 or x_3, Player 2 has the same information set consisting of x_2 and x_3. Notice that, as in the definition, if an information set consists of several nodes, the feasible action set in all these nodes are the same: $A_2(h_2(x_2)) = A_2(h_2(x_3)) = \{L, R\}$. This is because the set of feasible actions are correspondences of information sets.

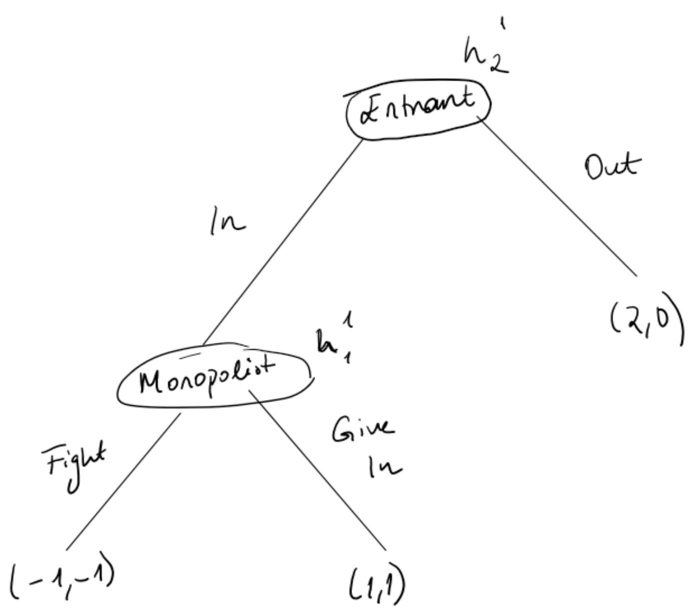
[6]Fudenberg ad Tirole, p80

4 Solution Concepts

4.1 Normal Form Representation and Nash Equilibrium

We can represent the extensive form games in the normal form. As we know, normal form games are those where players make an action plan once and for all. With the extensive form interpretation we allow player i to wait until it is her turn, and then decide on her action. With the normal form representation, each player makes a complete contingent plan of action at the beginning of the game: she sends a robot or a friend to play, telling them exact instructions on what to do if any of the information sets are reached.

For the Chain Store game from Example 3.1, we can write the normal form in the following matrix form



where the superscripts reflect the information sets in which the action is taken.

Since we can represent more complex games like this one in the strategic form,

Table 4.1: Chain Store, Normal Form

Monopolist \ Entrant	In^1	Out^1
Fight^2	-1,-1	$2^*, 0^*$
Give In^2	$1^*, 1^*$	$2^*, 0$

where players choose contingent plans, we can find its Nash equilibria. Looking at best responses, we see that if the monopolist fights, entrant decides to stay out, and if stays out, monopolist is indifferent between fighting and giving in: (Fight^2, Out^1) is a Nash equilibrium. Likewise, (Give In^2, In^1) is also a Nash equilibrium.

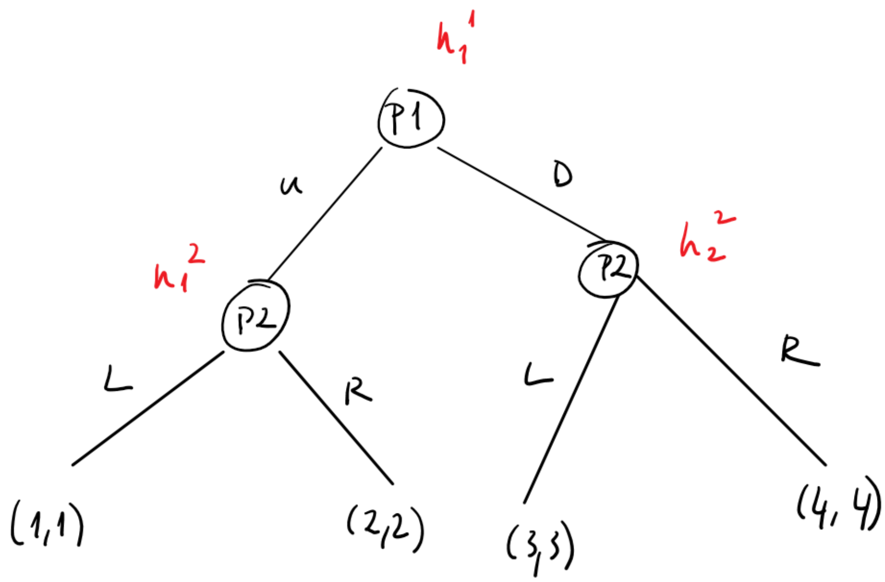
Now that we have a notion of strategies as contingent plan of actions, let us also define an important concept: behaviour strategies, and explain the relationship between them and mixed strategies.

Definition 4.1 (Behaviour strategies). *A behaviour strategy for player i , π_i is a map from information sets in H_i to probability distribution over feasible actions/moves: $\pi_i(h_i) \in \Delta(A(h_i))$.*

Behaviour strategies are probability distributions over the feasible actions at a given information set such that the probabilities assigned to each of the feasible actions sum up to 1 at all information sets. Take the game in Example 1.3. Player 2 has two information sets, so she would have two behaviour strategies: one if player 1 has played U or M, one if she played D. In the former case, a behaviour strategy defines the probability distribution between L and R, and in the latter, between L, M and R.

Mixed strategies on the other hand are probability distributions over all pure strategies in the contingent plan form.

Let us take the example that we have seen in the lecture to show the distinction (p8 of the slides).



For Player 2, we talk about a behaviour strategy at node h_1^2 and h_2^2, and we can denote them as $\pi_2(h_1^2)$ and $\pi_2(h_2^{)}2$. These can be of form $(\frac{1}{3}L, \frac{2}{3}R)$, or $(\frac{2}{5}L, \frac{3}{5}R)$ etc. Probability distribution is over L and R, the actions that are available at that given information set[7] and as you see sum up to 1. A mixed strategy on the other hand would be a probability distribution over the contingent plans as we defined them above, so $\sigma_2 \in \Delta(L^1L^2, L^1R^2, R^1R^2, R^1L^2)$. An example of a mixed strategy would be $\sigma_2 = (\sigma_2(L^1L^2), \sigma_2(L^1R^2), \sigma_2(R^1R^2), \sigma_2(R^1L^2)) = (\frac{1}{4}, \frac{1}{2}, 0, \frac{1}{4})$.

Theorem 1 (Kuhn's Theorem)**.** *Ever mixed strategy gives rise to a unique behaviour strategy.*

The converse is not true.

[7]The actions available are the same at both information sets in this case, but they need not be.

4.2 Extensive Form, Backward Induction and Subgame Perfect Nash Equilibrium

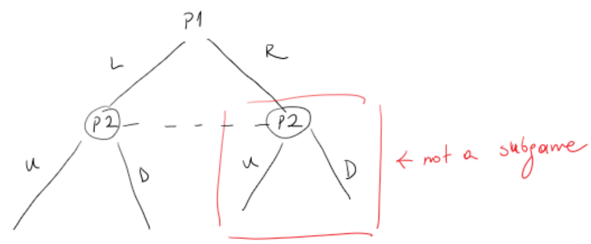
In the Chain Store game, Nash equilibria yield unreasonable results. (Fight^2, Out^1) is a Nash equilibrium, but we know that once the entrant enters, the monopolist no longer wishes to carry out its equilibrium strategy. Fighting is not a credible threat: if we reach Player 1's information set h_1^1, she will prefer to give in. Since Player 2 knows that, she would enter. From this perspective, the only reasonable outcome should be (Give In^2, In^1).

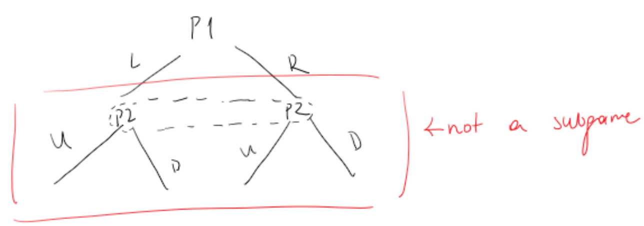
This process of thinking about the outcome of extensive games is very intuitive, and is formalized by the procedure of backward induction, which yields the **subgame perfect Nash equilibrium** of a game. Let us first define what a subgame is.

Definition 4.2 (Subgame)**.** *A subgame $G = (\mathcal{I}', X', Z', H(X'), A'(h))$ of an extensive-form game X such that*

1. *Begins at an information set that is singleton $h(x_k) = \{x_k\}$,*
2. *All nodes succeeding x_k are included in the subgame, and only these are included,*
3. *If $x' \in G$, and $x' \in h(x')$, $x'' \in h(x')$, then $x'' \in G$ such that we do not chop up any information sets.*

That a subgame starts at a singleton information set and includes all players means that **at the beginning of the subgame all previous actions are known to all players**. Not chopping up any information set means that we do not have any situations in the subgame that would not have arisen in the original game. Below are two examples that are not subgames: in the first one, Player 2 knows that R has been played: we chopped up an information set. That should not happen: we can't have outcomes that are impossible in the original game. In the second one, Player 2 does not know what has been played, and this can't happen in a subgame: we need to start from a singleton information node such that everyone knows what happened. In fact, the only subgame in this game is the original game itself.





Now we can define the subgame perfect equilibrium.

Definition 4.3 (Subgame Perfect Nash Equilibrium)**.** *A subgame perfect Nash equilibrium is a strategy profile where for each player, the strategy consists of actions taken at all information sets where the player has to move such that in all subgames, the Nash equilibrium is played.*

In other words, we require all players to carry out their Nash equilibrium strategies in every subgame of the entire game. In order to find this strategy, we use **backward induction**. In games of perfect information, the backward induction will give us the subgame perfect Nash equilibrium.

4.2.1 Backward Induction

1. Start at the end of the game tree, and identify the Nash equilibria for each of the final subgames (i.e. those that have no other subgames nested within them).

2. Find the Nash equilibria in these subgames. Replace the game by its Nash equilibrium strategy profiles and payoffs.

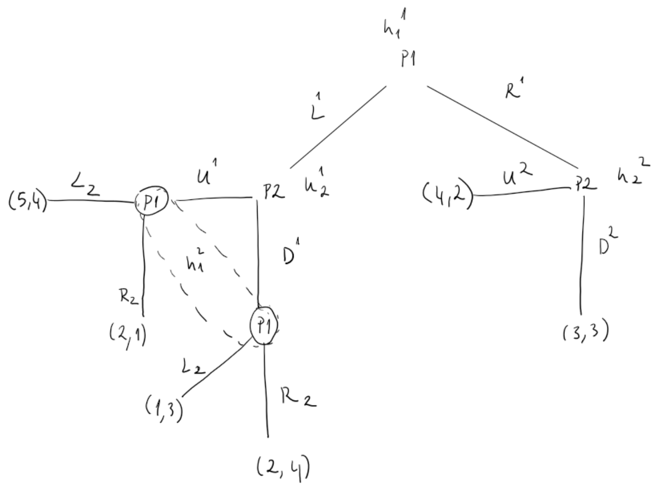
3. Repeat steps 1 and 2 for the reduced game. Continue the procedure until every move in the extensive game X is determined. This collection of moves at the various information sets of X constitutes a strategy profile that is SPNE. Notice that we also specify the actions at information sets that are not necessarily reached.

4. If multiple equilibria are never encountered in any step of this process, this profile of strategies is the unique SPNE. If multiple equilibria are encountered, the full set of SPNEs is identified by repeating the procedure for each possible equilibrium that could occur for the subgames in question.

Above, we applied backward induction to Chain Store Game to find its subgame perfect equilibrium (*Can you find its subgames?*). We can apply a similar reasoning to a more complicated game in extensive form. Following these steps described above.

Example 4.1. *Player 1 gets to move first. If she plays R^1, Player 2 gets to choose between U^2 and D^2, and game is over. Otherwise, a stage game is played where Player 2 can choose between U^1 and D^1, and Player 1 chooses between R^2 and L^2 [8].*

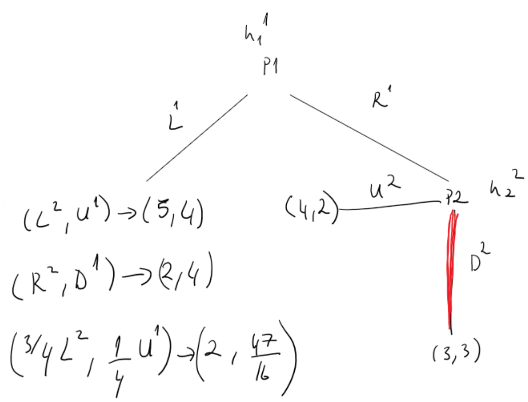
[8] Notice why this is a stage game. It seems like Player 2 acts first, and Player 1 follows. But neither Player 2 nor Player 1 can observe each other's actions before playing. This is exactly the definition of a simultaneous move game: it is not about the order of the moves, but that all players move without observing opponents' actions.



Let us apply the procedure as described above.

1. Identify the final subgames and their NE: consider the final subgame, where Player 1 plays L^1, and its MSNE are (L^2, U^1), (R^2, D^1), and $(\frac{3}{4}L^2, \frac{1}{4}U^1)$. Furthermore, we can see that, if Player 1 plays R^1, Player 2's best response is D^2.

2. Replace them in the extensive form:



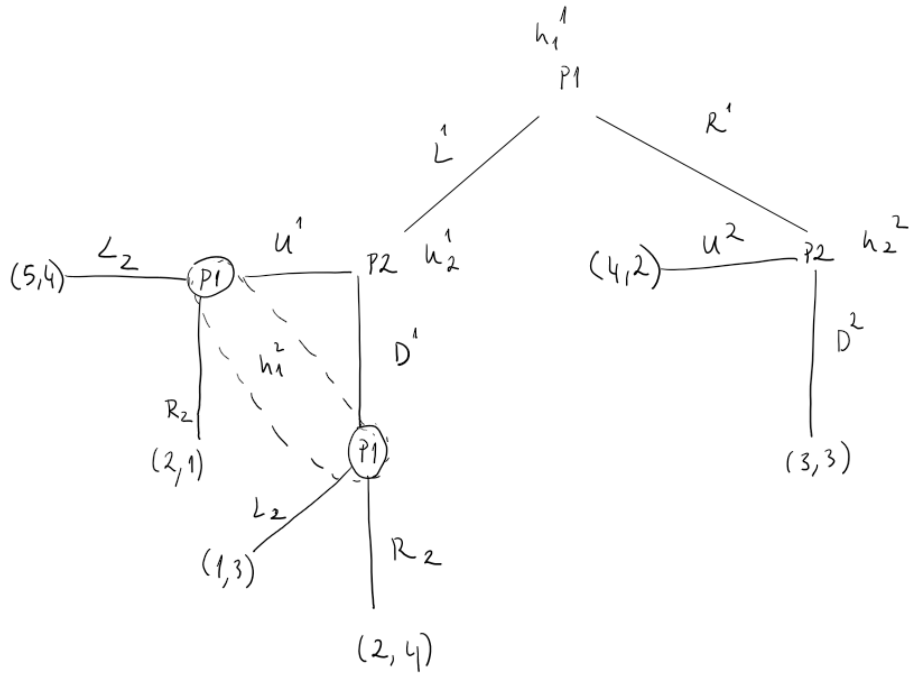
3. Now, we have one subgame, which is this reduced form extensive game as depicted above. We analyze the best responses for each equilibrium:

- Consider (L^2, U^1), it yields $(5,4)$. If Player 1 plays L^1, these are the payoffs. Playing R^1 yields $(3,3)$, since Player 2 would play D^2. Player 1 would choose L^1. The SPNE is $((L^1, L^2), (U^1, D^2))$. Notice that we are writing the strategies at the information sets that are not reached: we specify what Player 2 would have done if h_2^2 is reached.
- Consider (R^2, D^1), it yields $(2,4)$. Player 1 is better off playing R^1, since Player 2 would play D^2. The SPNE is $((R^1, R^2), (D^1, D^2))$. Again we specify what would have happened in the subgame that is not reached.
- Finally, consider the fully mixed strategy equilibrium. Again, Player 1 is better off playing R^1, since Player 2 would play D^2. The SPNE is $((R^1, 3/4L^2), (D^2, 1/4U^1))$.

4. The entire set of SPNE is
$\left\{\left((L^1, L^2), (U^1, D^2)\right), \left((R^1, R^2), (D^1, D^2)\right), \left((R^1, 3/4L^2), (D^2, 1/4U^1)\right)\right\}$.

5 Self-Confirming Equilibrium

So far, apart from dominance, we have studied solution concepts that are very demanding in terms of players' knowledge on the opponents' strategies. Think of subgame perfect Nash equilibrium: we require all players to know exactly what their opponents would do at all subgames, not only those that are reached, and respond optimally to these strategies. Consider the last game we have studied in the previous chapter.



The entire set of SPNE was
$\Big\{\Big(L^1L^2, U^1D^2\Big), \Big(R^1R^2, D^1D^2\Big), \Big(R^1 3/4L^2, D^2 1/4U^1\Big)\Big\}.$

Take (R^1, R^2, D^1, D^2). The subgame at h_2^1 is not even reached, but players have strategies for the case it is reached, and they *exactly* know what their opponents would be playing. Indeed, this is the definition of Nash equilibrium. Recall from our definition 2.5 in the chapter on static games: in a Nash equilibrium, 1) players know exactly each others' strategies, and 2) given these correct beliefs, strategies are payoff maximizing.

To the extent that we can interpret equilibria in games as outcomes of a learning process, we may think of equilibria that are not Nash in the sense that players do not hold correct beliefs on opponents' plays. Indeed, if some paths are never played and observed, how can players exactly know what their opponents would be doing on these off-equilibrium paths? It can only happen if players always held correct beliefs, not as an outcome of a learning process where players revise their beliefs using their observations of previous plays, and change their beliefs if the latter are contradicted. If in an equilibrium, some information sets are never reached, some players may hold wrong beliefs about what their opponents would be doing in these information sets. Since these information sets are never reached, the player with the wrong beliefs never get to be contradicted and revise their beliefs. The equilibrium is *self-confirming* in that sense.

This is the motivating idea of *Self-Confirming Equilibrium* (SCE) by Fudenberg and Levine: developing a notion of equilibrium that may arise if players are not experimenting enough. It is an equilibrium in the sense that this is where the play of the game converges to, but it is not Nash since players are allowed to hold wrong beliefs for the play of their opponents at the information sets that are never reached.

Before we define SCE formally, we need to introduce some notation, some of which we already know from extensive games.

Notation

- Game tree X, with $x \in X$ denoting nodes, $z \in Z \subset X$ terminal nodes
- Information sets $h \in H$, $H_i \subset H$ information sets of player i when she gets to play, $H_{-i} = H \setminus H_i$ information sets where the opponents get to play
- Set of feasible actions at an information set $A(h_i)$
- Pure strategy for player i, s_i, is a map from information sets in H_i to actions satisfying $s_i(h_i) \in A(h_i)$, S_i set of all such strategies
- s, s_{-i} strategy profile, and strategies of opponents. Mixed strategy profile $\sigma \in \times_i^I \Delta(S_i)$
- $u_i : Z \to \mathbb{R}$, payoff functions. Payoffs are functions of the terminal nodes.

- $Z(s_i)$ subset of terminal nodes that are reachable when s_i is played. $H(s_i)$ set of all information sets that can be reached when s_i is being played. For mixed strategies, let $H(\sigma_i) = \cup_{s_i \in supp(\sigma_i)} H(s_i)$ so that information sets reachable if a mixed strategy by player i is simply the union of the information sets that are reached when pure strategies s_i are played that are given positive probability of playing.

- Given the mixed strategies of all players, the information sets that are reached with positive probability under σ: $\bar{H}(\sigma)$. Notice that, if σ_{-i} is completely mixed we have $\bar{H}(s_i, \sigma_{-i}) = H(s_i)$: every information set is potentially reachable given that s_i has positive probability. **This part is crucial to understand.**

- A behaviour strategy is denoted π_i, is a map from information sets to feasible actions: $\pi_i(h_i) \in \Delta(A(h_i))$.

- By Kuhn's theorem, each mixed strategy σ induces a unique behaviour strategy denoted $\hat{\pi}_i(\cdot|\sigma_i)$. This is the probability distribution over actions that are feasible at h_i, which is induced by opponents' strategy σ_{-i}. If they play σ_{-i}, my information set h_i is reached.

Recall the discussion in class. We use the notation of both mixed and behaviour strategies. We are looking for mixed strategy profiles to be Nash or self-confirming equilibria. But, in self-confirming equilibrium, we are interested in what players observe as a function of the information sets that are reached, we use the notation of behaviour strategies. Mixed strategies are probability distributions over all pure strategies at all information sets, whereas behaviour strategies are specific to information sets, allowing us to focus only on the information sets that are reached with positive probability.

Now let us adopt a slightly different definition of Nash equilibrium to apply it to extensive form games that is used in the paper. In Fudenberg and Levine (1993), a Nash equilibrium is defined as following.

Definition 5.1 (Nash equilibrium)**.** *A Nash equilibrium is a mixed strategy profile* σ *such that for each* $s_i \in supp(\sigma_i)$ *there exists beliefs* μ_i *such that*

1. s_i *maximizes* $u_i(\cdot, \mu_i)$, *and*
2. $\mu_i\big[\big\{\pi_{-i}|\pi_j(h_j) = \hat{\pi}_j(h_j|\sigma_j)\big\}\big] = 1$ *for all* $h_j \in H_{-i}$

In words, each player optimizes given her beliefs, and her beliefs are a point mass over the true distribution. True distribution here is the behaviour strategies of all the opponents.

In self-confirming equilibrium, we adopt a weaker requirement for the beliefs. We expect the players to be right only for the beliefs they have on the opponents' behaviour strategies on the information sets that are reached with positive probability. Formally,

Definition 5.2 ((Unitary) Nash equilibrium)**.** *A Nash equilibrium is a mixed strategy profile* σ *such that for each* $s_i \in supp(\sigma_i)$ *there exists beliefs* μ_i *such that*

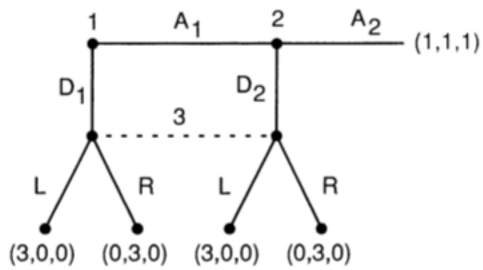
1. s_i *maximizes* $u_i(\cdot, \mu_i)$, *and*
2. $\mu_i\big[\big\{\pi_{-i}|\pi_j(h_j) = \hat{\pi}_j(h_j|\sigma_j)\big\}\big] = 1$ *for all* $j \neq i$ *and* $h_j \in \bar{H}(s_i, \sigma_{-i})$.

Notice what has changed in the second requirement. For each player i, the beliefs are a mass point on the true distribution of the behaviour strategies of the opponents, if the information sets where these strategies are played out are played with positive probability *given what that player does*: If my opponent gets to play only because of my move, then in equilibrium, I must be right about what she will do. Otherwise, one may have wrong beliefs. I am however allowed to be wrong about what she would have done if she never gets to play. Likewise, if a player acts before me and never again, such that whether or not they act is not a function of my strategies, I do not need to hold correct beliefs on their play.

Very important: A player holding correct beliefs means that that player's beliefs correspond exactly to the probability distribution the opponents are adopting at a given information set.

What we defined above is a **unitary self-confirming equilibrium**. We can also allow for players in the same position (like player 1s, player 2s etc.) to have different beliefs on the plays of the opponents. An equilibrium with such set of beliefs is called a **heterogeneous self-confirming equilibrium.**

Let us give examples to the definitions we have studied. Take the horse game from the class.



Example 5.1.

Player 1 A1			
	P2		
3*P3		A_2	D_2
	L	1,1,1	3,0,0
	R	1,1,1	0,3,0

Player 1 D1			
	P2		
3*P3		A_2	D_2
	L	3,0,0	3,0,0
	R	0,3,0	0,3,0

y In this game, there are four PSNE:

$$PSNE = \{(D_1, A_2, L), (D_1, D_2, L), (D_1, D_2, R), (A_1, D_2, R)\}$$

The equilibria where P1 plays D_1 embody what P2 would have played if P1 had played A_1. P1 and P3 are sure about what P2 would have done in the other case.

With self-confirming equilibrium we relax this constraint. Consider the following beliefs: P1 beliefs about P3 is that P3 plays R, while P2 beliefs are such that P3 plays

L for sure. Then, for P1 A_1 is a BR, and for P2 A_2 is a BR, and $\sigma = (A_1, A_2, \cdot)$ is a self confirming equilibrium.

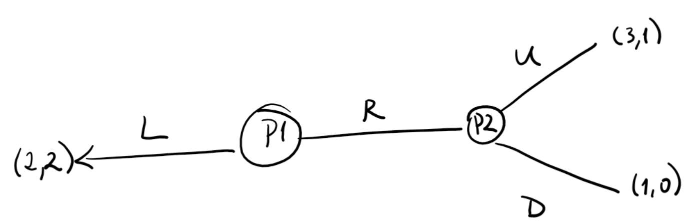
Notice that our proposed self-confirming equilibrium cannot be a Nash equilibrium, since for Nash equilibria beliefs need to be correct off-path. In this case, P1 and P2 have to have the same beliefs over P3's strategy: both need to believe L or both R. If both beliefs were P3 playing L, then for P1 D_1 is a profitable deviation, while if both beliefs are P3 playing R, D_2 is a profitable deviation.

To conclude, our self-confirming equilibrium is:

$$\left\{\sigma = (A_1, A_2, L), \mu_1 = (\Pr(A_2) = 1, \Pr(R) = 1) = 1, \mu_2 = (\Pr(L) = 1) = 1\right\}$$

Note that we specify both the strategies for each player, as well as the beliefs of players 1 and 2 on the act of the opponents. How about the beliefs of Player 3? We do not have to specify, or alternatively we can assign any belief to them. This is because Player 3 never gets to play, so that her beliefs are irrelevant to the play.

How about a **heterogeneous self-confirming equilibrium**? Let us take another example from class.



Example 5.2 (Heterogeneous self-confirming equilibrium).

The SPNE here is (R, U). There is an additional NE that is (L, D), which is not SPNE. There are no mixed NE since as soon as Player 1 mixes, Player 2 can no longer be indifferent: U always yields a higher payoff. We can however construct the following heterogeneous self-confirming equilibrium: There are two types of Player 1: types A and B. Types A believe that if they were to play R, Player 2 would choose D. A best response according to these beliefs is to play L. Player 1 of type B believes correctly that if they play R, Player 2 plays U with probability 1. Indeed, this is the strategy of Player 2. A heterogeneous self-confirming equilibrium is then given by

$$\left\{\sigma_1^A = L, \sigma_1^B = R, \sigma_2 = U, \mu_1^A = (\Pr(U) = 0) = 1, \mu_1^B = (\Pr(U) = 1) = 1\right\}$$

6 Repeated Games

We now turn our attention to games that are repeated over time, either for a fixed duration or infinitely. Players in these settings need to take into account not only the payoff in a given period, but also how their behaviour may affect the play in the future, and by that, their future payoffs. As we will see, repetition and long term interactions change the strategic nature of games substantially.

We will study two types of repeated interactions. First, we will focus on games that are played between one long-run and one short-run players. Then we will focus on settings where all players are long-run players, and study the Folk Theorem.

6.1 Long vs Short Run Players

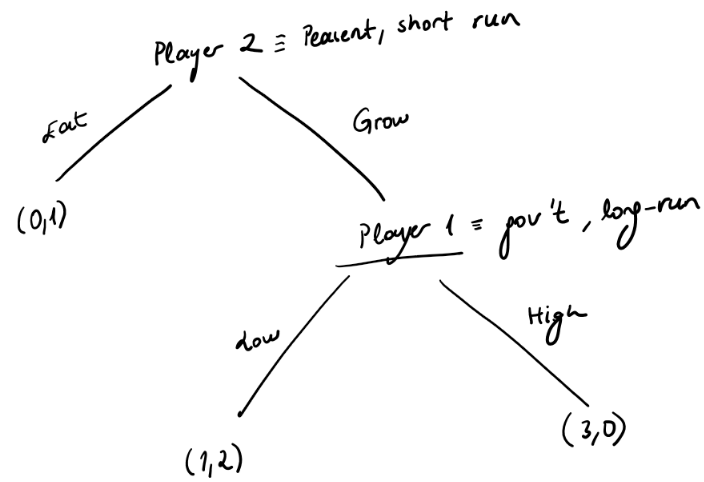
In this part, we study interactions between a long-run and many short-run players. The long-run players are always the same, they play the same game for a long duration. They therefore need to internalize the effect of their actions on the future course of the game. The short run players on the other hand play the game with the long run player once and then are replaced by other short run players. They only care about the static payoffs.

This class of games are applied in political economy and macroeconomics. You can think of a government, a central banker, a firm as a long run player, and citizens, consumers as the short run players. This is obviously not because the citizens actually disappear after one period. We model the interaction this way because a government, a central bank or any institution for that matter internalize the fact that their actions will have an effect on how citizens respond to them in the long run. From the viewpoint of a regular citizen, this is not the case: the government or the central bank will not take my individual saving decision into account when making choices in the next period. I would therefore act myopically, maximizing my payoff today, not considering how the government may react. In that sense, the short run player is like a representative agent for many long-lived agents.

For the remaining of this section, we will refer to the long run player as Player 1, and short run player as Player 2, and use these terms interchangeably.

Let us give a motivating example that we saw in the lectures. We will make use of this example throughout this part of the notes.

Example 6.1 (Peasant Dictator Game)**.** *Player 2, peasant acts first and can either grow crops, or eat. If she eats, she gets 1, and the dictator gets 0. If she grows crops, then the dictator gets to tax the peasant. The dictator can either employ a high tax rate or a low one.*



The only subgame perfect Nash equilibrium of this game is $\{Eat, High\}$. This is unsatisfactory, since grow and low tax would yield a higher payoff for both players. But there is a time inconsistency problem. Even though the dictator would prefer to tax low so that the peasant grows the crops, once the time comes to tax, it would have no incentive to carry out this strategy, she would want to impose high taxes.

What happens if these games are repeated over time? Can we then sustain an equilibrium where the peasant grows, and the government imposes low taxes? Indeed, as we will show, the repetition changes the strategic nature of the game. Player 1 can achieve outcomes that are not possible in a stage game through **reputation**. In stage games, it is possible that one of the players commit to a certain action in order to achieve a better outcome. In repeated games, reputation can act as a substitute to commitment.

In what follows, we will try to understand which equilibria are sustainable in such a setting, and if reputation can yield the same outcomes as commitment.The roadmap is: 1) introduce notation related to long vs short-run player games, 2) understand that playing the static NE repeatedly is an equilibrium, 3) explore what payoffs are possible in a static setting to then finally 4) derive the best dynamic equilibrium payoffs.

6.1.1 Notation

- Action for player i at time t is denoted a_t^i. a_t is the actions of all players at time t.
- History of the game at time t is the actions played by all players up to that point including time t play: $h_t = (a_1, a_2, \cdots, a_t)$. We denote the null history as h_0, this is before anything has been played.
- We define behaviour strategy for player i as $\alpha_t^i = \sigma(h_{t-1}^i)$, so they are the mixed strategies played at a given history. A player conditions her actions, her behaviour strategies, on the history of the play that she has observed so far, not on what the opponents play at t, therefore h_{t-1}.
- $u^i(a_t)$ denotes the player i's payoff from the actions at time t.
- If Player i is the long-run player, we assume that they maximize the **average discounted utility** given by $(1-\delta)\sum_{t=1}^{T} \delta^{t-1} u^i(a_t)$.[9]
- δ is the discount factor. For short run players, we have $\delta = 0$.

Notice the similarities between history and information sets in the extensive form game notation. In a sense, each period of a play is a subgame of the entire game.

6.1.2 Nash equilibrium and Subgame Perfection in Repeated Games

Proposition 6.1 (Stage game NE)**.** *Stage game Nash equilibrium is also a Nash equilibrium in the repeated game.*

[9] $\sum_{t=1}^{T} \delta^{t-1} u^i(a_t)$ is the discounted present value, multiplying it by $1-\delta$ gives us the average discounted value. The weights $(1-\delta)\sum_{t=1}^{T} \delta^{t-1})$ sum up to 1, which is why we call this the average discounted payoff.

This means that playing the Nash equilibrium of the stage game in every period will still be an equilibrium in the repeated game, since there are no profitable deviations.

Proposition 6.2 (SPNE in repeated games)**.** *If a strategy profile is Nash, then it is also subgame perfect.*

As we mentioned, a subgame in this setting is basically a time period. Playing the stage NE in every period means playing the NE in every subgame, which is the definition of subgame perfection. This does not mean that the unique SPNE is playing the static Nash equilibrium in every period, but it means that it is a SPNE.

6.1.3 Static Game Benchmarks

Remember, we are interested in the best dynamic equilibrium that we can sustain. For that, we first study what can be achieved in a static game both as a Nash equilibrium, and using **commitment**. Let us define these benchmarks. Remember that they apply not to repeated games, but to the games that are played once. There are three static benchmarks: 1) Nash equilibrium, 2) pure Stackelberg/precommitment, 3) mixed Stackelberg/precommitment, and 4) minmax.

Definition 6.1 (Stackelberg/commitment)**.** *In a Stackelberg equilibrium, the long run player gets to commit to an action. When this action is in pure strategies, we talk of pure precommitment or pure Stackelberg. If it is a mixed strategy, then we talk of mixed precommitmnet/Stackelberg.*

In a Stackelberg equilibrium, the leader either moves first, and lets the opponent observe her action, and then move, or she can also use an action that clearly shows the intent of playing an action. A classical example is when the Spanish commander Hernán Cortes ordered the ships to be burnt when they arrived to Mexico, making it clear that the Spanish forces will not retreat. Incidentally, another such precommitment example is when Ummayad forces invaded Spain and Portugal about 800 years previously to the conquest of Mexico, and the commander Tāriq ibn Ziyād ordered the ships to be burnt. You see that in these examples, a player acts in a way that it makes one of the previously available actions impossible. Before, the forces could fight or retreat, and now the latter option is no longer available.

In this class we do not discuss how the long run players get to commit, but rather focus on which actions they would commit to. Before we apply precommitment to Peasant-Dictator game, let me write the best responses of Player 2.

Define $\alpha^1(L)$ the probability of low taxation. This will be useful later on.

$$BR^2(\sigma^1(L)) = \begin{cases} Eat & \text{if } \alpha^1(L) < \frac{1}{2} \\ \Delta(Eat, Grow) & \text{if } \alpha^1(L) = \frac{1}{2} \\ Grow & \text{if } \alpha^1(L) > \frac{1}{2} \end{cases} \tag{4}$$

In the Peasant-Dictator game, **the pure precommitment/Stackelberg equilibrium** is that the dictator commits to low taxation, and the peasant grows. This equilibrium yields payoff of 1 to the dictator, the long run player.

How about the **mixed precommitment payoff**? First, find the probability of low taxation that will make the peasant indifferent between growing and eating: $1 = 2p + 0(1-p) \implies p = 1/2$. If the probability is lower, the peasant always eats, if it is higher, they always grow. If the dictator commits to imposing low taxation with this probability, the peasant is indifferent. We assume that when indifferent, the short run player acts in the way that leader would prefer. In this case, we assume that the peasant grows if she is indifferent. How would the dictator randomize? She would choose the lowest probability of low taxation that still makes the peasant grow the crops, which is 1/2. What is the payoff to the dictator? She gets 2 in expectation: $\frac{1}{2}1 + \frac{1}{2}3 = 2$.

As a general notation for payoffs with pure and mixed commitment/Stackelberg, we can write

$$u^1_{ps} = max_{a^1} \quad max_{\alpha^2 \in BR^2(\alpha^1)} \quad u^1(\alpha^1, \alpha^2)$$

$$u^1_{ms} = max_{\alpha^1} \quad max_{\alpha^2 \in BR^2(\alpha^1)} \quad u^1(\alpha^1, \alpha^2)$$

where ms is mixed Stackelberg, and ps the pure Stackelberg. These expressions show that in commitment, Player 1 chooses the action a^1 or the mixing probability α^1 in such a way that, given that Player 2 best responds, Player 1's payoff is maximized.

Definition 6.2 (Minmax)**.** *A player's minmax, or reservation utility is defined as*

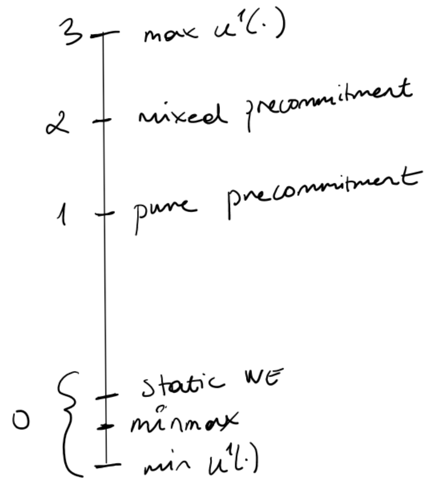
$$\underline{u}_i = \min_{\alpha_{-i}}[\max_{\alpha_i} g_i(\alpha_i, \alpha_{-i})]$$

that is, the lowest payoff that player i's opponents can hold him to by any choice of α_{-i}*, provided that player i correctly foresees* α_{-i} *and plays a best response to it. This is what we found above.*

For the minmax, we do not ask the player is opponents to behave optimally. We merely ask, what is the worst they can do to Player i, given that the latter will react optimally. Note that player i's payoff is at least $\underline{v}_i$ in any static equilibrium and in any NE of the repeated game. This is why we can also refer to $\underline{u}_i$ as the reservation utility. We thus know that no equilibrium of the repeated game can give any player a payoff lower than this amount.

In the Peasant-Dictator game, the minmax for Player 1 is 0. The worst the peasant can do is to Eat always such that the dictator gets 0. This is also the lowest payoff in the payoff space, but this does not need always to be the case.

These three are the static benchmarks we need. The other benchmark is the static Nash equilibrium, which in this case is (Eat, High), yielding 0 to the long run player. We can represent the payoff space of the dictator on the real line:



6.1.4 Best Dynamic Equilibrium

Now we turn to our analysis of which payoffs are possible in repeated games. We analyze two settings: 1) when the time horizon is finite such that the game is played for T periods, and this is known to all players, and 2) infinite horizon.

Finite Horizon

When the time horizon is finite, we solve the game by backward induction. The last subgame is the last period where the game is played, T. Since the game ends after this period, this is no different than a stage game. The dictator has no incentive to tax low if the peasant is growing the crops, so the peasant eats. The static Nash equilibrium is the only Nash equilibrium of this subgame.

Now think of what happens in $T-1$. Players know that, whatever they do now, Nash equilibrium will be played in the next period. What they do in $T-1$ will not change that. As the peasant, I know that in the next period the dictator will impose high taxes. The dictator cannot convince me otherwise by imposing low takes today, at $T-1$. So the dictator would impose high taxes, and I would not grow any crops. So again, static Nash equilibrium is played.

We can see that the same reasoning will apply to $T-2$, $T-3$, and so on until time period 1. **If the game is finitely repeated, the only subgame perfect Nash equilibrium is the repetition of the static Nash equilibrium in every period**. We have the following proposition.

Proposition 6.3 (SPNE of the finitely repeated game)**.** *Consider a repeated game G^T for $T < \infty$. Suppose that the stage game G has a unique pure strategy equilibrium $a^\star$. Then G^T has a unique SPNE. In this unique SPNE, $a^t = a^\star$ for each $t = 0, 1, \ldots, T$ regardless of history.*

Infinite Horizon

When time horizon is infinite, the reasoning above no longer applies. We cannot do backward induction since there is no T. At every period, there is at least some chance that the game will go on in the next period [10]. Then, in every period, players know that their actions today may have consequences in the future. This mechanism allows players to play actions they would not in a setting where there is no future. For that reason, how patient the Player 1 is, how high δ is, that is, how much she cares about the future consequences of current actions will be crucial to determine if we can sustain a different equilibrium than the static Nash equilibrium of a game.

Before we study how we construct equilibria in a more general setting, let us show one equilibrium we can construct for Peasant-Dictator game when the time horizon is infinite.

Consider the following equilibrium. Play {Low, Grow} if it is the first time the game is being played or if this profile has always been played. If at any point something else happens, then revert to the static Nash equilibrium. The candidate equilibrium strategies can be written as

[10] Note that, infinite horizon does not necessarily mean that the game will never end, but rather that there is no definite time period that it ends.

$$a_t^{Dictator} = \begin{cases} Low & \text{if t = 1 or } h_{t-1} = \{(Low, Grow), (Low, Grow), ..., (Low, Grow)\} \\ High & \text{otherwise} \end{cases}$$

$$a_t^{Peasant} = \begin{cases} Grow & \text{if t = 1 or } h_{t-1} = \{(Low, Grow), (Low, Grow), ..., (Low, Grow)\} \\ Eat & \text{otherwise} \end{cases}$$

We claim that for δ high enough, this is a subgame perfect Nash equilibrium with average payoff 1 for the long run player.

Proof:

Step 1: Consider the subgames after someone deviated. We claim that playing {Eat, High} forever is subgame perfect. Indeed, this is the Nash equilibrium of the stage game, so playing it in every period is by definition subgame perfect.

Step 2: Is $(Low, Grow)$ an equilibrium in the sense that are there no profitable deviations?

For the short-run player, this is a static problem. If low is being played, grow is the best response: $BR_t^2(Low) = Grow$ for all t.

For the long run player, we need to weigh the profit from deviating versus the future loss resulting from deviation and reversion to static Nash equilibrium.

Average discounted utility from playing Low is: $u(Low|BR^p(Low) = Grow) = (1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}1 = 1$.

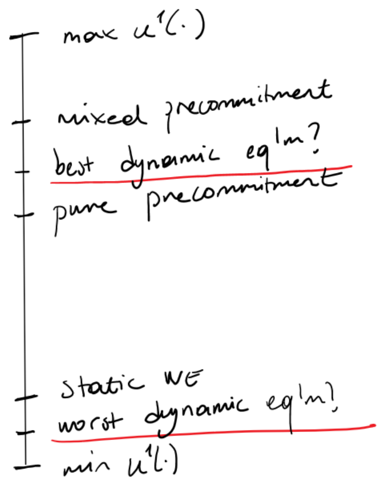
If the dictator sets high taxes at any t, she gets 3 in that period since the peasant is growing, and then 0 for all future periods since we then revert to the static Nash equilibrium. The average discounted payoffs from deviation is $(1-\delta)\sum_{t=1}^{\infty}\delta^0 3+\delta^{t-1}0 = 3(1-\delta)$.

For the equilibrium to be sustainable, i.e. deviation not to be profitable, we need

$$1 \geq 3(1-\delta) \implies \delta \geq \frac{2}{3}$$

We can conclude that, for $\delta \geq \frac{2}{3}$, our candidate equilibrium is indeed an equilibrium.

Note that, this is one of the many, many equilibria we can construct. We constructed one that yields a higher average payoff than the static Nash equilibrium. That is great. But is 1 actually the best we can do? Is it the **best dynamic equilibrium payoff?** We want to understand where the best dynamic equilibrium lies in this payoff space:



Proposition 6.4 (Best dynamic equilibrium payoff)**.** *The best dynamic equilibrium payoff a long run player can achieve is given by*

$$\bar{v}^1 = max_{\alpha^2 \in BR^2(\alpha^1)} \quad min_{a^1 | \alpha^1(a^1) > 0} \quad u^1(a^1, \alpha^2)$$

What this means is that the best equilibrium payoff is obtained by 1) finding BRs of the short run player to different behaviour strategies of long run player α^1, given by $BR^2(\alpha^1)$ for all α^1, 2) taking the minimum of the payoffs when strategy profile $(\alpha^1, BR^2(\alpha^1))$ is played given by $min_{a^1|\alpha^1(a^1)>0} \quad u^1(a^1, \alpha^2)$ for each pair, 3) among these minima, finding the maximum. That maximum gives us the best dynamic equilibrium payoff.

Before we show why, let me show how we do that so you have an idea of what this process looks like. Take the Peasant-Dictator game.

First, we look at different values of α^1, the probability distribution over the action space of Player 1. We then look at the best responses of Player 1: $\alpha^2 \in BR^2(\alpha^1)$. Define $\alpha^1 = P^1(Low)$, the probability with which the dictator imposes low taxes.

α^1	$BR^2(\alpha^1)$
1	Grow
$\in (1/2, 1)$	Grow
$1/2$	$\in \Delta\{Grow, Eat\}$
$\in (0, 1/2)$	Eat
0	Eat

Now, for each pair $(\alpha^1, BR^2(\alpha^1))$, we write the payoffs that are attainable, which is the support of the payoffs, and then the minimum of each support, so worst in support or $min_{a^1|\alpha^1(a^1)>0} \quad u^1(a^1, \alpha^2)$.

α^1	$BR^2(\alpha^1)$	Support	Worst in Support
1	Grow	1	1
$\in (1/2, 1)$	Grow	[1,3]	1
$1/2$	$\in \Delta\{Grow, Eat\}$	[0,3] & [0,2]	< 1
$\in (0, 1/2)$	Eat	0	0
0	Eat	0	0

where, the support in the third row comes from the fact that Player 1 mixes with probability 1/2, Player 2 grows with some probability q. The support of the strategies of Player 1 is $Low, High$. For Low, expected payoff is q, whereas for High, it is $3q$. For any value of q, worst in support is utility from playing Low. As q changes between 0 and 1, the worst in the support is somewhere between 0 and 1.

Now we look at the pair $(\alpha^1, BR^2(\alpha^1))$ with the highest worst in support, which are the first two. 1 is indeed the best dynamic equilibrium payoff.

We take one of them, for instance the first row. This corresponds to $(Low, Grow)$. This was the equilibrium candidate that we have shown was indeed an equilibrium previously.

In the section below, I show step by step why we just did what we did to find the best dynamic equilibrium payoff, how we characterize the best dynamic equilibrium payoffs.

General Deterministic Case

Characterization of equilibrium payoffs, proof to Proposition 6.4

Developed by Fudenberg, Kreps and Maskin, Repeated Games with Long-run and Short-run Players, ReStud, 1990

We first determine the bounds where best dynamic equilibrium payoff must lie. Let $\bar{v}^1$ be the best dynamic equilibrium payoff.

Upper bound of $\bar{v}^1$

Proposition 6.5. $\bar{v}^1$ *is smaller than the payoff obtained in mixed precommitment.*

Proof:

In best dynamic equilibrium, following a history, LRP plays a behaviour strategy SRP plays BR to it Short run player gets $u^2(h_t) = max_{\alpha^2} u^2(\sigma^1(h_t), \alpha^2)$, so when she plays her best response against the strategy of Player 1, given history h_t.

Therefore, Player 1 is getting at most the utility $u^1(h_t) \leq max\ u^2(\alpha^1(h_t), \alpha^2)$ given that $\alpha^2 \in BR^2(\alpha^1(h_t))$. So that α^2 must be a best response to what Player 1 is playing. The RHS is the definition of mixed precommitment: Player 1 acts first, and commits to a strategy such that Player 2 responds optimally. Since Player 1 acts first, she chooses the α^1 that, together with Player 2's best response will yield the best payoff. We just showed that the best dynamic payoff must be smaller than this number.

We just showed that the upper bound of best dynamic equilibrium payoff is the mixed precommitment.

Lower bound of $\bar{v}^1$

Clearly, no equilibrium strategy can yield a lower payoff than the minmax. This is the lower bound. In this class, for simplicity, we further **assume** $\underline{v}^1 = n$ so that the worst dynamic equilibrium is simply the repeated Nash equilibrium, and that is the minmax. Now that we have our bounds, let us determine the value of best dynamic equilibrium payoff.

Call $\underline{v}^1$ the worst dynamic payoff for Player 1, and n the Nash equilibrium payoff. Denote $w^1(a^1)$ the payoff the long run player gets in the second period onward **depending on what they do today, in the first period**[11]. v^1 how much she gets in equilibrium. Remember, players condition their actions to what they observed up to one period before, so on h_{t-1}. That is why Player 1's continuation payoff depends on her actions a period before. The short run player reacts in the next period.

We want to find an equilibrium in behaviour strategies $\alpha = (\alpha^1, \alpha^2)$ that yields the highest possible payoff for Player 1.

[11]Recall the discussion in class. $w^1(a^1)$ is an object that encompasses all future payoffs starting from next period. Because we work with averages, it is an average discounted payoff of all future payoffs.

1. Fix $\alpha = (\alpha^1, \alpha^2)$ that is an equilibrium from first period onward.
2. First, for the short run player, Player 2, we only need to check that $BR^2(\alpha^1) = \alpha^2$, since they only care about the present.
3. For Player 1 to stick to equilibrium, we need

$$v^1 \geq (1-\delta)u^1(a^1, \alpha^2) + \delta w^1(a^1) \tag{5}$$

where a^1 is any possible action Player 1 gets, either the equilibrium action α^1, or a different action when they deviate.

This is the incentive constraint 1.

4. Then, for any action employed with positive probability, we must have

$$v^1 = (1-\delta)u^1(a^1, \alpha^2) + \delta w^1(a^1),\ \alpha^1(a^1) > 0 \tag{6}$$

since all actions employed with positive probability in equilibrium must yield the same payoff. This is the IC 2. [12]

5. These are the ICs for period 1. In infinite horizon, the game is the same for period 2, and for any other period.
6. The continuation payoff $w^1(a^1)$ should lie between the payoff of best and worst dynamic equilibrium, since we need an equilibrium in every subgame, i.e. in every period. So I cannot do any better or worse in the following subgames than what is possible in an equilibrium.

$$\begin{aligned} \underline{v}^1 &\leq w^1(a^1) \leq \bar{v}^1 \\ n &\leq w^1(a^1) \leq \bar{v}^1 \end{aligned} \tag{7}$$

where we use the simplifying assumption $n = \underline{v}^1$.

[12]In the Peasant Dictator game, if in equilibrium the dictator is employing both Low and High, then these must have the same discounted payoff. If high taxation yields higher utility today, then it yields a lower utility in the future than taxing low, in a way that it yields the same utility as playing Low today. Then if Low yields a lower payoff today, its continuation payoff must be higher.

7. The Equations 5, 6 and 7 are the conditions that any w^1 has to satisfy. We ask, how big can $w^1(a^1)$ and $\bar{v}^1$ be such that the ICs are satisfied.

8. Look at IC2. It shows that all actions we employ with positive probability give us the same average utility. Some of these actions yield a higher payoff **today** than the others. Then, these actions must have lower continuation payoffs starting from tomorrow: higher the $u(a^1, \alpha^2)$, lower the $w^1(a^1)$ and vice versa. So $w^1(a^1)$ is highest, when $u(a^1, \alpha^2)$ has its lowest value.

9. Highest we can do for $w^1(a^1)$ is $\bar{v}^1$ from Equation 7. Plugging it into IC 6, we get

$$\bar{v}^1 = (1 - \delta)u^1(a^1, \alpha^2) + \delta\bar{v}^1 \tag{8}$$

and $u^1(a^1, \alpha^2)$ is the lowest payoff you get by an action you play with positive probability.

This gives us immediately $\bar{v}^1 = u^1(a^1, \alpha^2)$, where, $u^1(a^1, \alpha^2)$ is the smallest possible payoff we get, given α^2, in a given period. Formally we write it as $min_{a^1|\alpha^1(a^1)>0}u^1(a^1, \alpha^2)$. This is called **worst in support**: among actions that are played with positive probability, so given α^1, the one that yields the worst payoff is the best dynamic equilibrium payoff. Now all the equalities are satisfied: IC2 and IC3.

How about IC1? That the best dynamic equilibrium payoff should be larger or equal to playing any other action? So far, we worked with a fixed α. Now, remark that, Player 1 *chooses* α^1 to maximize the best dynamic payoff. So, the average discounted payoff of the best dynamic equilibrium is the static payoff of the action that yields the worst static payoff among those Player 1 chooses with positive probability. But, Player 1 chooses which actions she plays with positive probability: she chooses α^1.

11. Which α^1s are we allowed to choose? We need to ensure that $\alpha^2 \in BR^2(\alpha^1)$. Then the problem of the Player 1 is

$$\bar{v}^1 = max_{\alpha^2 \in BR^2(\alpha^1)} \quad min_{a^1|\alpha^1(a^1)>0} \quad u^1(a^1, \alpha^2)$$

which is what we wanted to show.

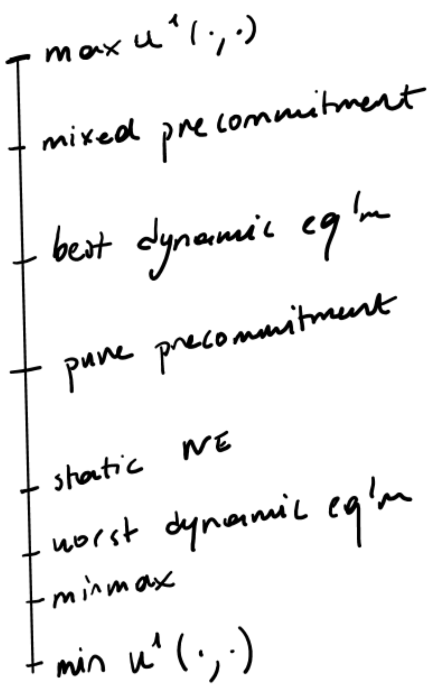
We have argued above that $\bar{v}^1$ lies below the payoff from mixed precommitent. We now argue that it is above pure precommitment.

Proposition 6.6 (Best dynamic payoff, pure and mixed Stackelberg)**.** *We have that*

$$u^1_{ps} \leq \bar{v}^1 \leq u^2_{ms}$$

This follows from the definition of u^1_{ps}, u^2_{ms} and $\bar{v}^1$. The first two we have defined when discussing the static benchmarks.

The payoff space can then be depicted as follows.



Deviations

Now we showed how to find the equilibrium profile that yields highest payoff. We also need to specify what happens if Player 1 deviates. This is true not just when we want to sustain the best dynamic equilibrium, but any equilibrium. That is, what will the continuation payoff to Player 1 be if she deviates to a^1? What is $w^1(a^1)$?

Using the ICs,

$$w^1(a^1) = \frac{\bar{v}^1 - (1-\delta)u^1(a^1, \alpha^2)}{\delta} \geq n^1$$

So that the continuation payoff is $\bar{v}^1$ as δ converges to 1. If it is small, we have that the continuation payoff is simply the static Nash payoff.

Remark that this equality should hold for all a^1, and it will give us a threshold for δ. By simply plugging in the action that yields highest payoff today given Player 2's strategy, so the deviation to the action with highest static payoff, and replacing $w^1(a^1)$ and $\bar{v}^1$, you can find the critical δ. This is indeed what we did while checking the (Low, Grow) equilibrium for Peasant Dictator game.

a^1 is then High. Player 2 first plays Grow, and then in the next period, upon observing High, reverts to Eat. $w^1(a^1)$ is then 0. $\bar{v}^1$ is 1. Plugging them into the above equation,

$$0 = \frac{1-(1-\delta)3}{\delta} \to \delta = \frac{2}{3}$$

as we had showed earlier as well.

Summary

When asked to find the best dynamic equilibrium/best subgame perfect dynamic equilibrium for Player 1, we

1. Construct a table with all pairs of $(\alpha^1, BR^2(\alpha^1))$
2. For each pair, find the support of payoffs. This is the set of payoffs that are reachable under the strategy profile $(\alpha^1, BR^2(\alpha^1))$
3. For each support, find the minimum. Then, find the maximum among the minima.
4. There may be multiple. Choose one $(\alpha^{1*}, BR^2(\alpha^{1*}))$ and construct the following equilibrium: Players play $(\alpha^{1*}, BR^2(\alpha^{1*}))$ if $t = 1$ or if $(\alpha^{1*}, BR^2(\alpha^{1*}))$ has always been played. Revert to the static Nash equilibrium otherwise.
5. Notice that Player 2 would never deviate since $\alpha^2 = BR^2(\alpha^{1*}))$
6. Player 1 would not deviate as long as average discounted equilibrium payoff is larger than the average payoff from deviating in one period, and then obtaining the static Nash equilibrium payoff. Set up the condition. Use the deviation to an action that gives the highest payoff today. That is the a^1 that requires the highest level of patience. This will give us a cutoff for discount factor $\delta(a^1)$, above which our proposed equilibrium will be sustained.

6.2 Long Run Players, Folk Theorem

The second class of repeated games we study concerns repeated games where all of the players are long run players. We ask again which Nash equilibria other than the static Nash equilibria can be sustained if the game is repeated. This question is in the core of a series of theorems known as the Nash Folk Theorems. We will see that once the game is **infinitely** repeated, set of Nash equilibria contain further equilibria than the stage game Nash equilibrium.

The idea, similarly to the games with short and long run players, is that the deviation from the equilibrium will result in a punishment phase. Notice that, in the stage games, a Nash equilibrium is defined as a strategy profile where no one has a profitable deviation. In repeated games, we are able to sustain equilibria that are not Nash within the stage game: where there are profitable deviations **in the stage game**. This is because even if there is a strategy that yields a higher payoff within that period, that strategy is associated with a lower continuation payoff, which is the punishment. Punishments can take many forms and can be infinitely complicated. In this class, we focus on the so-called **grim strategies**: any deviation causes the players to carry out a punitive action **forever.**

Before we formally express the Folk Theorem, let us define two notions.

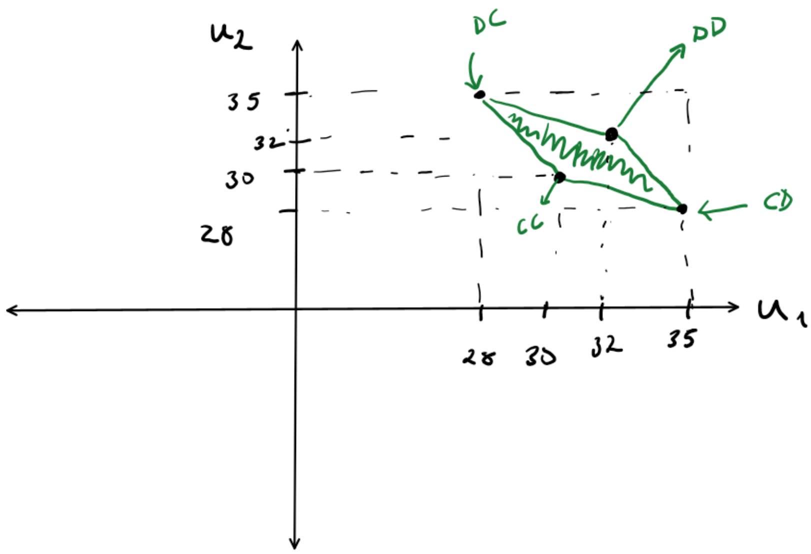
Definition 6.3 (Social feasibility)**.** *A payoff profile is socially feasible if it can be obtained as a result of a strategy profile that is constructed by some convex combination of the strategy profiles of the stage game. In other words, the set of socially feasible payoffs is given by the convex hull of the payoffs associated with all possible strategy profiles.*

Definition 6.4 (Individual rationality)**.** *A payoff profile is called individually rational if, for all players i, the payoff is at least as high as the minmax.*

To illustrate, let us take the example from the lecture slides.

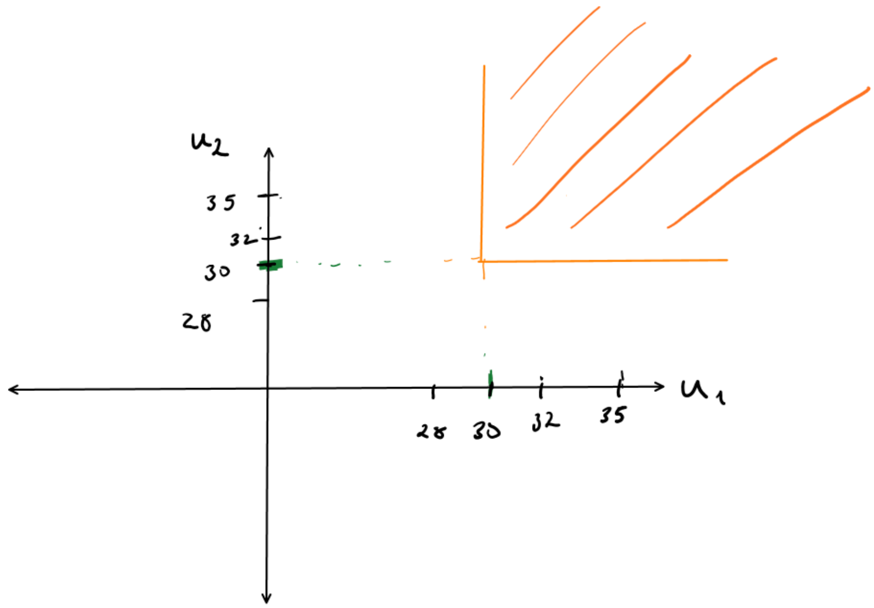
P1 \ P2	D	C
D	32,32	28,35
C	35,28	30,30

The set of socially feasible payoffs can be illustrated as below

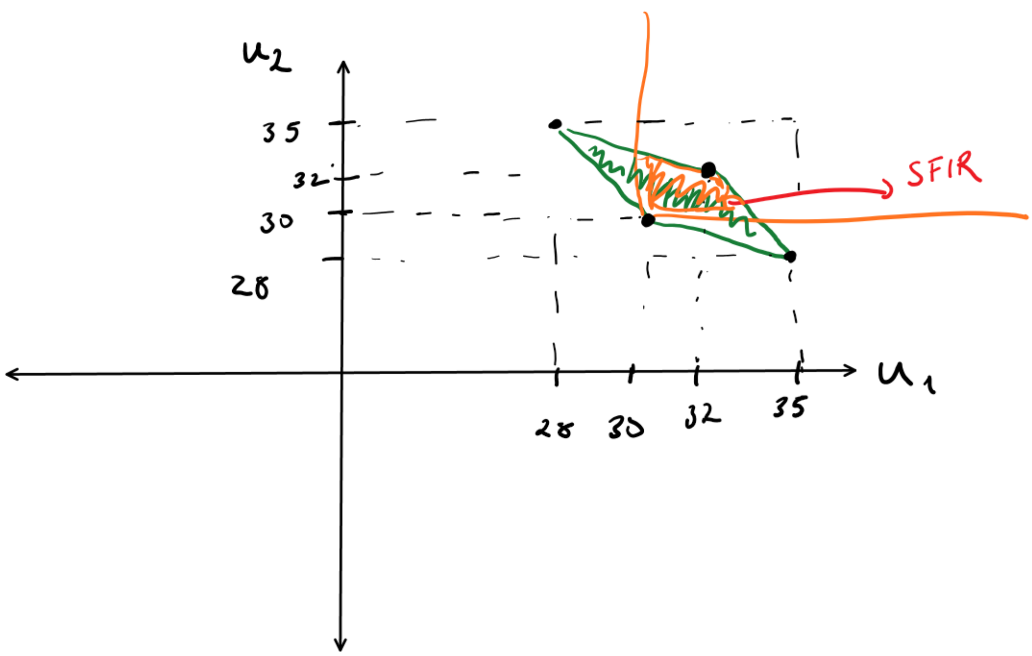


We first mark the payoffs given by any of the four strategy profiles of this game, and take a convex combination. These payoff profiles are all reachable under some public randomization.

Let us now illustrate the individually rational payoffs. For that, we need to find the minmax for each player. For Player 1, what is the worst Player 2 can do to her given that P1 will react optimally? Player 2 can either play D, to which the best response is C, resulting in 35 for P1, or they can play C, to which again the best response is C (it is the strictly dominant strategy), yielding 30. Then, worst P2 can do is to play C, so that the minmax for Player 1 is 30. The same is true for the other player due to symmetry. We can show the region of individually rational payoffs as below:



The region of **socially feasible, individually rational payoffs** is simply the intersection of these two sets.



Let us state a version of the folk theorem.

Theorem 2 (Folk Theorem)**.** *For every feasible payoff vector v with $v_i > \underline{v}_i \; \forall i$, there* *exists a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$, there is a Nash equilibrium of $G^\infty(\delta)$ with payoffs $v = (v_1, \ldots, v_I)$.*

In other words, any payoff profile that is in the socially feasible, individually rational payoff space can be sustained as a Nash equilibrium of the repeated game for discount factors that is higher than some $\underline{\delta} < 1$.

Just as in the long run short run players, we need to define what happens if anyone deviates, and ensure that none of the players do not have an incentive to deviate. The condition for each player not to deviate will allow us to determine $\underline{\delta}$, the lowest discount factor for which the equilibrium is sustainable. An example of such a game is tacit collusion, which you are expected to do in Exercise 7 of Problem Set 2.

To test your understanding of the concepts we invoked in this section, you can do the following exercise from a past problem set.

Exercise: The Folk Theorem

For each of the following simultaneous move games, find the static Nash equilibria, and give an accurate sketch of the socially feasible individually rational region.

	L	**R**
U	4,3	0,7
D	5,0	1,2

	L	**R**
U	6,6	5,0
D	0,5	0,0

Decision Theory

In this part of the class, we deviate from game theory where we study strategic interactions, to study how individuals make decisions when they face uncertainty related to the future. We will introduce the some results that pose a challenge to the standard economic theory, a critical analysis of the answers behavioural economics provides. Finally, we will introduce dynamic programming, which is an optimization solution to dynamic problems.

7 Decision under Uncertainty: Equity Premium Puzzle, Habit Formation

An important question in macroeconomics is how people allocate their savings across assets of different risk levels and returns. Equity premium puzzle, which we will introduce now, relates to empirical findings of how people do their portfolio allocation, which, according to the theoretical models we use, imply an unreasonably large risk aversion. Before we move on to exploring more in detail this phenomenon, we define some concepts such as utility derived from wealth, and risk aversion.

Portfolio allocation relates to **wealth**, which is a different concept than consumption. While it may be natural to think that people derive utility from consumption of goods or services, we need to link wealth to consumption so to be able to have a theory of why people care about wealth, and how much. We define wealth as the **present**

value of consumption: $W = \frac{c}{1-\delta}$. The **utility of wealth** is then the **present value of utility of consumption**:

$$U(W) = \frac{u(c)}{1-\delta} = \frac{u((1-\delta)W)}{1-\delta}$$

13

As mentioned above, we want to understand people's allocation of their wealth across different assets, i.e how they do portfolio allocation. How much risk are people willing to take given a certain amount of return? Moreover, how does their attitude towards risk change with the wealth they possess? **Relative risk aversion** measures attitudes towards lotteries that are proportional to wealth. The **coefficient of relative risk aversion (CRRA) of wealth** is given by

$$\rho_W = -\frac{U''(W)W}{U'(W)} = -\frac{(1-\delta)^2 u''(c)W}{(1-\delta)u''(c)} = \frac{u''(c)c}{u'(c)} = \rho_c$$

so that it is the same as the **coefficient of relative risk aversion of consumption**.

We now turn to a model of portfolio allocation model.

Simple Portfolio Choice Model

The agent has an initial wealth W. There are two assets in the economy: a risk-free asset of government bonds that has return r_b, and a risky asset of stocks. The stocks have a return given by $r_s = \bar{r}_s + \sigma y$, where y is a random variable with mean 0 and variance 1: $E[y] = 0$, $Var(y) = E[y^2] = 1$.

The **equity premium**, λ is defined as the difference between the mean return of the risk-free and the risky asset: $\lambda = \bar{r}_s - r_b$.

[13] For this part of the lectures, we assume that the individuals maximize present value, given by $\sum_{t=1}^{\infty} \delta^{t-1} u_t$, as is commonly assumed in macroeconomics or finance.

The agent's problem is to choose the allocation of her wealth among the risk free and risky assets. Denote the share of the risky asset in the portfolio with α. The wealth of the agents after investment choice and after the return is realized is given by $W + W(\alpha r_s + (1-\alpha)r_b) = W + W(r_b + \alpha(\lambda + \sigma y))$.

The utility from the final wealth is given by $U(W+W(r_b+\alpha(\lambda+\sigma y)))$. The agent's problem can be solved by deriving this expression with respect to α:

$$\frac{\partial U(W + W(r_b + \alpha(\lambda + \sigma y)))}{\partial \alpha} = 0$$

$$U'(W + Wr_b + W\alpha\lambda + W\alpha\sigma y))(W(\lambda + \sigma y)) = 0$$

Let us define $W' = W + Wr_b + W\alpha\lambda$. This is the expected value of the final wealth after investment, since $E[\sigma y] = 0$. Then we can write the above equation as

$$U'(W' + W\alpha\sigma y))(W(\lambda + \sigma y)) = 0$$

We take the first-order Taylor approximation of the FOC. Remember that the approximation of a function of form $U'(W' + \varepsilon)$ is $U'(W') + \varepsilon U'(W'')$. Taking $\varepsilon = W\alpha\sigma y$, we can approximate the above equation as

$$U'(W')(W(\lambda + \sigma y)) + W\alpha\sigma y U''(W')(W(\lambda + \sigma y)) = 0$$

$$U'(W')(W(\lambda + \sigma y)) + W^2\alpha\sigma^2 y^2 U''(W') = 0$$

Now, let us take expectations of this expression, and recall $\mathbb{E}[y] = 0, \mathbb{E}[y^2] = 1$

$$U'(W')W\lambda + W^2\alpha\sigma^2 U''(W') = 0$$

Dividing by W and rearranging:

$$\rho = \frac{U''(W')W}{U'(W')} = \frac{\lambda}{\alpha\sigma^2}$$

This is the FOC. λ and σ^2 are exogenously given, and ρ is also taken as given, it is a preference parameter. The agent chooses α such that the FOC holds in expectation.

7.1 Equity Premium Puzzle

In 1981, Robert Shiller gathered 100 years of data on stock and bonds returns, and consumption data. Mehra and Prescott have taken the simple portfolio choice model to data in their 1985 paper Equity Premium: A Puzzle. They use data on annual returns of stocks and US treasury bills between 1889–1978. The mean return of bonds is $r_b = 1.9\%$, whereas for stocks, it is 7.5%. Recall that in our model, the equity premium is the difference between these two returns, so $\lambda = 5.6\%$.

The standard deviation of stocks is $\sigma = 0.181$. Plugging these numbers in the FOC we just derived, we find

$$\rho = \frac{0.056}{\alpha 0.181^2} = 1.709\alpha^{-1}$$

That is, larger values of α imply a lower relative risk aversion. α can be at most 1, implying a CRRA of 1.709. Of course, α is not 1: people do not on average invest all their wealth in private equity. Can we deduce α from the data?

To do so, let us assume that the consumption is some constant fraction of final wealth, given by $c = \phi W_1$. Recall that we defined final wealth, which we now denote as W_1 as $W + W(r_b + \alpha(\lambda + \sigma y)$. Let us define s^2 as the standard deviation of consumption, which is $\frac{Var(c)}{E(c)^2}$. Since consumption is a constant fraction of final wealth, the two have the same standard deviation, so that $s^2 = \frac{Var(W_1)}{E(W_1)^2}$.

$$s^2 = \frac{Var(W_1)}{E(W_1)^2} = \frac{\alpha^2\sigma^2 W^2}{W'^2} \approx \alpha^2\sigma^2$$

where, recall that we had defined $W' = E(W_1)$, the expected value of the final wealth, and $Var(W_1) = Var(W + W(r_b + \alpha(\lambda + \sigma y)) = \alpha^2\sigma^2 W^2$ [14]. Finally, we use the fact that the wealth does not change very much from one period to the other, which is why the ratio of wealth today to expected wealth tomorrow is approximately 1.

Using this identity, we can deduce α from the data since we can observe the standard deviation and the variance of stock returns. α^{-1} then turns out to be 5.17, implying that $\rho = 8.84$. This coefficient implies an unreasonably high risk aversion, which is the **equity premium puzzle**. In other words, there seems to be an **excess return on risky assets**.

Why is that a puzzle, and how do we know that CRRA cannot be that large? Two considerations give us an indication of the magnitude of this coefficient.

First, in standard theory the coefficient of relative risk aversion determines our willingness to save. Assume that the utility is given by $u(c) = \frac{c^{1-\rho}}{1-\rho}$ so that we have a constant relative risk aversion coefficient, ρ. Suppose further that the economy is growing at a constant rate γ, such that $c_t = \gamma c_{t-1}$. With this utility functional form, we have that $u'(c) = c^{-\rho}$, so that the FOC for the consumption-saving decision is

$$\frac{u'(c_t)}{u'(c_{t+1})} = \frac{c_t^{-\rho}}{c_{t+1}^{-\rho}} = \frac{c_t^{-\rho}}{\gamma^{-\rho} c_t^{-\rho}} = \delta(1+r) \rightarrow \gamma^\rho = \delta(1+r)$$

The average annual growth rate of the US economy in Shiller's data is 1.8%. Plugging this number into γ, even with δ as high as possible, a CRRA of 8.84 implies that the interest rate should be 17%, which is much larger than what it actually is. If

[14] Recall that $Var(x + a) = Var(x)$ if a is constant. Here $W + W(r_b + \alpha(\lambda)$ is the constant term.

CRRA were really as high as 8.84, people would want to borrow so much more in order to increase consumption today.

A less known part of the puzzle is what this level of CRRA would imply about how stock market responds to news. It has been empirically shown that the stock market goes up with good news, and down when there is bad news. Boldrin and Levine (2001) show that with a coefficient of relative risk aversion above 1, stock prices should fall after good news. High levels of risk aversion imply that the stock prices dramatically increases after bad news hit. This is due to the general equilibrium effect of interest rates response dominating, which happens with high levels of CRRA.

7.2 Habit Formation

Equity premium puzzle, or the excess return on risky assets is a result of observed low volatility of consumption growth. With such low volatility of consumption growth, high returns on risky assets can be supported only if one assumes that even smallest consumption fluctuations are very painful to consumers. In other words, one must assume that consumers are extraordinarily risk averse [15].

Habit formation has been proposed to resolve the equity premium puzzle. Habit formation models break the time separability of standard preferences, allowing the utility of consumption in a given period to depend on past consumption. Past consumption then acts almost as a "subsistence level of consumption". Small drops in consumption generates a much larger drop in habit-adjusted consumption. In other words, people really dislike small divergences from their average consumption level: they are much more risk averse on the short term, than on the long term.

In the model we studied above, there is a direct link between relative risk aversion and intertemporal elasticity of substitution. Specifically, the former is the inverse of the

[15] Stephanie Schmitt-Grohé and Martín Uribe, New Palgrave Dictionary of Economics, Palgrave Macmillan, May 2008.

latter. Habit persistence on the other hand creates a wedge between these two. With the same data, Constantinides (1990) show that the estimated CRRA with the habit formation model is 2.2. Excess return on risky assets can be generated even with low levels of relative risk aversion if we allow for time inseparability with habit formation models.

An example of a model of habit formation is in Question 5 of PS3.

8 Decision under Uncertainty: Behavioural Critique

In this part, we study some experimental and real life results that seem to contradict the way in which the economists model human behaviour. They mostly concern expected utility representation of preferences, and discounting of the future.

8.1 Allais Paradox

Expected utility theorem is a central and pervasively used result in economics. The result, stated below is due to continuity of preferences and independence axiom [16].

Expected Utility Theorem

Suppose that the rational preference relation $\succsim$ on the space of lotteries L satisfies the continuity and independence axioms. Then, $\succsim$ admits a utility representation of the expected utility form. That is, we can assign a number u_n to each outcome $i = 1, \dots, N$ in such a manner that for any two lotteries $L = (p_1, \dots, p_n)$ and $L' = (p'_1, \dots, p'_n)$ we have:

$$U : \mathcal{L} \to \qquad U(L) = u_i p_i + \cdots + u_n p_n$$

$$L \succsim L' \quad \leftrightarrow \quad \sum_{i=1}^{n} u_i p_i \geq \sum_{i=1}^{n} u_i p'_i$$

[16] For a reminder, you can check out MWG Chapter 6.

where $u = (u_i, \ldots, u_n)$ are called Bernoulli utilities of outcomes.

The axioms on which the expected utility is based are not very strong assumptions. Furthermore, expected utility theorem allows us to work with utility functions that are analytically very easy. For these reasons, its use in both micro and macroeconomic theory is pervasive. A central question is then whether people actually act in the way that expected utility theorem would predict.

The **Allais paradox** constitutes one of the oldest and most famous challenge to expected utility theorem. Consider the following decision. You can choose between two lotteries, one paying 1 billion dollars for sure, and the other one where you have 10% chance of getting 5 billion dollars, 1% chance of getting nothing. What would you choose? People typically take the first lottery, getting a billion dollars for sure.

Now consider another decision between two lotteries. In the first lottery, you have 10% chance of getting 5 billion dollars, and 90% of getting nothing. In the second one, you have 11% chance of getting 1 billion or nothing. What would you go for? Most people would choose the first, as the additional 4 billion dollars seem like a reasonable enough return to an additional 1% risk of getting nothing.

There is nothing wrong with any of the decisions per se. The problem arises if people choose the first lottery in the first problem, and the first one also in the second problem. Choosing the first lottery in the first problem implies by expected utility theorem implies $u(1) > .1u(5) + .89u(1) + .01u(0) \rightarrow u(5) < 1.1u(1) - .1u(0)$. Choosing the first lottery in the second decision problem on the other hand implies $.1u(5) + .9u(0) \geq 0.11u(1) + 0.89u(0) \rightarrow u(5) > 1.1u(1) + -.1u(0)$, a contradiction[17].

The Allais paradox sparked a debate on whether expected utility representation of preferences actually capture how people make decisions. Prospect theory suggests two answers to this puzzle. The first one is that people tend to exaggerate low probabilities,

[17]You are asked to do a similar exercise in PS3 Question 3

especially when they are linked to losses. That is, faced with a rare event event like the stock market crashing, a terror attack, or .01 chance of getting nothing in the first decision problem above, people treat these events as if they have a higher probability: let's say, instead of .01, they assign .1 chance to getting nothing. Secondly, prospect theory suggests that people do not care so much about their overall well-being, that is, the expected utility, but about gains and losses from a reference point.

8.2 Subjective Uncertainty

Ellsberg Paradox

There are two urns each containing 100 balls. It is known that urn A contains 50 red and 50 black, but urn B contains an unknown mix of red and black balls.

The following bets are offered to a participant:

Bet 1A: get 1 if red is drawn from urn A, 0 otherwise

Bet 2A: get 1 if black is drawn from urn A, 0 otherwise

Bet 1B: get 1 if red is drawn from urn B, 0 otherwise

Bet 2B: get 1 if black is drawn from urn B, 0 otherwise

Typically, participants were seen to be indifferent between bet 1A and bet 2A, consistent with expected utility theory, but were seen to strictly prefer Bet 1A to Bet 1B *and* Bet 2A to 2B. This seems like a surprising result: you should either think that urn B has more red balls, or more black balls. If you think the first, then you would choose 1B over 1A, 2A to 2B. Otherwise you would choose 2B to 2A, and 1A to 1B. But you should not think simultaneously that the urn has both more red and more black balls. This result is generally interpreted to be a consequence of **ambiguity aversion**; people intrinsically dislike situations where they cannot attach probabilities to outcomes, in this case favouring the bet in which they know the probability and utility outcome. In a sense, we can think of this aversion as believing that Nature plays against you: if you bet on urn B having more red balls, it will actually have more black balls and vice versa.

What should people do in such cases? If they dislike unknown probabilities, because they think that Nature will attach a higher probability to their least preferred event after they act, they can randomize their actions, turning the subjective probabilities into objective probabilities. In a way, this is similar to playing a game where the opponent may take advantage of your strategies, such as Matching Pennies or Penalty Shots. Randomization prevents the opponents from taking advantage of a player's strategies, and in this case, Nature acts as an opponent [18].

8.3 Risk Aversion in the Laboratory

In laboratory experiments, we often observe what appears to be risk averse behaviour over small amount of money. Recall how we defined wealth at the beginning of this chapter. It is the discounted sum of all consumption a person does in their lifetime. The stakes at laboratory experiments are minuscule compared to this amount. Yet, people seem to avoid taking risks with such small amounts. This finding contradicts utility functions with diminishing returns to scale. Such functions imply that people dislike uncertainty, that they are risk averse. But for small stakes, expected utility maximisers with concave utility functions should be almost risk neutral [19].

Extrapolating their behaviour to higher stakes show us how unreasonably risk averse people would need to be for them to behave the way they do in the laboratory. Suppose a risk averse person who turns down a 50-50 bet between losing 100 dollars and gaining 105 dollars. Then, we know that from an initial wealth level of \$340,000 that person will turn down a 50-50 bet of losing \$4,000 and gaining \$635,670. You are asked to

[18]That is in people's minds of course. But since we are interested in how people make decisions, that is what we care about the.

[19]Rabin, Matthew. "Risk Aversion and Expected-Utility Theory: A Calibration Theorem." Econometrica 68, no. 5 (2000): 1281–92.

derive the CRRA with data from a laboratory experiment in Q1 of PS3. The estimated CRRA is unreasonably high. It seems that people do not take their own lifetime wealth as a reference point in the laboratory, but something else.

The main point here is that people seem to act in a different way in laboratory than in real life. Since most of the results in prospect theory, which may seem as an alternative to the expected utility theorem's paradoxes, are due to experimental results, how much can we rely on these?

8.4 Present Bias

Suppose someone asks you whether you would like to have 175 euros now, or 192 euros in 4 weeks. What would you choose? In Keren and Roelsofsma's experimental study, 82% of the participants chose to have the money now.

How about if you are asked whether you would like to have 175 euros in 26 weeks, or 192 euros in 30 weeks. Now, only 37% of the participants chose to have the money in 26 weeks.

We usually model preferences over time with exponential discounting. Agents maximize $\sum_{t=1}^{\infty} \delta^{t-1} u(c_t)$. By this representation, the way someone weighs the utility of today versus four weeks and how they weigh the utility in 26 weeks versus in 30 weeks should be exactly the same. This experiment on the other hand shows that this is not the case. This phenomenon is called the **present bias**: the inclination to prefer a smaller present reward to a larger later reward, but reversing this preference when both rewards are equally delayed.

Hyperbolic Discounting

Quasi-hyperbolic discounting, developed by Richard Strotz in Review of Economic Studies in 1955, is an alternative way of modelling preferences that may deliver results that are consistent with present bias. The utility function has the form $u(c_1) +$

$\theta \sum_{t=2}^{\infty} \delta^{t-1} u(c_t)$, where θ is the weight attached with future utilities. $\theta < 1$, where 1 is the weight of present utility, indicates present bias.

Uncertainty of the Future

An alternative explanation relate to uncertainty over future. Fernandez-Villaverde and Mukherji (2003) point out that we know our present needs much better than our future needs. Future is uncertain. Most people in the experiment may have needed the money right away, or they knew exactly what to buy with it, as opposed to their situation within a month. On the other hand, whether they will need the money in 26 weeks more urgently or in 30 is has the same level of uncertainty. When asked whether they want 175 euros in 26 weeks vs 192 in 30 weeks, people choose to have more money later, since in expectation, both time periods have the same marginal utility of consumption. Indeed, when the subjects were asked to make the same decision with uncertain rewards, a very different result arises. Keren and Roelsofsma conducted the same experiment when there was only a 50 % chance of getting the money. The results are in the table below.

fraction Making Choice With Uncertain Reward

Scenario	Choices	Probability of reward	
Scenario		1.0 (60)	0.5 (100)
1	\$175 now	0.82	0.39
	\$192 in 4 weeks	0.18	0.61
2	\$175 in 26 weeks	0.37	0.33
	\$192 in 30 weeks	0.63	0.67

We see that when the chance of reward is only 50%, people behave treat the present and future rewards in the same way as they do with a certain future reward. This indicates that the result has less to do with present bias, and more with the uncertainty of the future.

Self-Commitment

A further aspect to the decision problem in this experiment is self-commitment. Suppose now that you have the options of getting 175 euros in 26 weeks, 192 euros in 30 weeks, or decide in 26 weeks whether you want it at that moment, or in 4 weeks. If the reason people chose to get the reward in the present is uncertainty, having the option to change decisions should be preferred: in 26 weeks, you will have a better knowledge of your preferences and needs, so the option to reevaluate your choice should be valuable to you. A rational utility maximizer would always value the flexibility. Some people would however choose to have the 192 euros in 30 weeks, and not take the option of changing the choice. Why is that? We may know that when the time comes, we will take the money immediately, even though we would prefer to have more money in 30 weeks now. Our current self and future self may have different preferences, often referred as time inconsistency. People try to limit their options, an act of self-commitment, which is not in line with standard economic modelling of human behaviour.

An example of self-commitment is shown in the paper of DellaVigna and Malmendier, *Paying Not to Go to the Gym*. They find that people subscribe to annual memberships, which, per visit cost them $17. Per visit fee on the other hand is $10 dollars. On average, these users forgo savings of $600 during their membership. This may be an example of people making a commitment to go to the gym.

8.5 Procrastination: Irrationality or Learning?

[20]

Prominent on Akerlof's list of "behavioral" phenomena is procrastination. This is something we are all familiar with – but what is it exactly? When I quizzed a behavioral economist for examples he came up with the following list.

[20] Excerpt from *Is Behavioral Economics Doomed?*, David K. Levine, 2012

Paying taxes the day before the deadline

Christmas shopping on Christmas eve

Buying party supplies for something like a New Years Eve party or a 4th of July party at the last minute

Buying Halloween costumes at the last minute

Delaying the purchase of concert tickets

Waiting to buy plane tickets for Thanksgiving

Here is the thing: none of these is the least irrational. In each case an unpleasant task is delayed until the deadline. But if the task is unpleasant and we are impatient – as economists assume we are – then the best thing to do is to wait until the deadline. In the folk story:

The king had a favorite horse that he loved very much. It was a beautiful and very smart stallion, and the king had taught it all kinds of tricks. The king would ride the horse almost every day, and frequently parade it and show off its tricks to his guards.

A prisoner who was scheduled to be executed soon saw the king with his horse through his cell window and decided to send the king a message. The message said, "Your Royal Highness, if you will spare my life, and let me spend an hour each day with your favorite horse for a year, I will teach your horse to sing."

The king was amused by the offer and granted the request. So, each day the prisoner would be taken from his cell to the horse's paddock, and he would sing to the horse "La-la-la-la" and would feed the horse sugar and carrots and oats, and the horse would neigh. And, all the guards would laugh at him for being so foolish.

One day, one of the guards, who had become somewhat friendly with the prisoner, asked him, "Why do you do such a foolish thing every day singing to the horse, and letting everyone laugh at you? You know you can't teach a horse to sing. The year will pass, the horse will not sing, and the king will execute you."

The prisoner replied, "A year is a long time. Anything can happen. In a year the king may die. Or I may die. Or the horse may die. Or... The horse may learn to sing."

The focus on procrastination is behavioral economics at its worst. Here a phenomenon that for the most part is rational and sensible is promoted to a glaring contradiction of standard theory that requires an elaborate psychological explanation. It is true in some of the examples above that there might be a cost of delaying: tickets might sell out before the deadline and so forth. However that simply introduces a trade-off between buying early and closer to the deadline, and different rational people with different degrees of patience, and who value the tickets differently may well choose to behave differently, some procrastinating and some not.

We previously discussed the DellaVigna and Malmendier (2006) study of health club memberships. They provide evidence that people pay extra to self-commit to exercising. They also discuss procrastination: the fact that people after they stop attending delay canceling their memberships. Unlike the example above there is no issue of delaying an unpleasant task until a deadline. So: is this the irrational procrastination Akerlof is concerned about? DellaVigna and Malmendier's data shows that people typically procrastinate for an average of 2.3 months before canceling their self-renewing membership. The average amount lost is nearly $70 against canceling at the first moment that attendance stops.

Leaving aside the fact that it may take a while after last attending to make the final decision to quit the club, we are all familiar with this kind of procrastination. Why cancel today when we could cancel tomorrow instead? Or given the monthly nature of the charge, why not wait until next month. One behavioral interpretation of procrastination is that people are naïve in the sense that they do not understand that they are procrastinators. That is, they put off until tomorrow, believing they will act tomorrow, and do not understand that tomorrow they will face the same problem and put off again. There may indeed be some people that behave this way. But if we grant that people who put off cancellation are making a mistake, there are several kinds of untrue beliefs they might hold. One is that they falsely believe that they are not procrastinators. DellaVigna and Malmendier assert that canceling a membership is a simple inexpensive procedure. Supposing this to be true, it might be that people falsely believe that it will be a time consuming hassle. Foolishly they think canceling will involve endless telephone menus, employees who vanish in back rooms for long periods of time, and all the other things we are familiar with whenever we try to cancel an automatic credit card charge.

The question to raise about the "naïve" interpretation is this. Which is more likely: that people are misinformed about something they have observed every day for their entire lives (whether or not they are procrastinators) or something that they have observed infrequently and for which the data indicates costs may be high (canceling)? Learning theory suggests the latter – people are more likely to make mistakes about things they know little about. Behavioral economics argue the former is more likely.

9 Dynamic Programming

Dynamic programming is a solution to dynamic optimization problems. A problem is dynamic if the actions taken in the current period have consequences for the future. In the games we have studied in the first chapter, as well as the consumer and firm problems we studied in Microeconomics I, people were faced with static choices: either the game was played once, or people/firms could adjust at every period, essentially facing static optimization problems.

In most real life situations, problems are of dynamic nature. For consumers, consumption decisions today will influence have much resources is available in the future periods tomorrow. For firms, hiring and capital acquisition decisions are associated with adjustment costs in the next period. There is an irreversibility of actions, either making actions in the next period impossible or costly. Furthermore, the future is uncertain. When making hiring decisions today, a firm does not know with certainty the level of demand in the future. With infinite time horizon, the problem appears to be very complex: we need to find the optimal action for each period, keeping in mind how actions today will change the actions available at every consecutive period over an infinite horizon, as well as the different states of the world. Dynamic programming provides an efficient solution to such problems.

Setting

Formally, we have

- a finite action space, $a \in A$,
- a finite state space, $y \in \mathcal{Y}$,
- time invariant transition probability between states $\pi(y'|y, a)$ such that the probability of next period's state depends on the state in and actions taken current period,
- per period utility given by $u(a, y)$ such that it depends on the actions taken and the current state
- finite histories h that includes all information on the state until current time period: $h_t = \{y_1, ..., y_t\}$, and history space H
- strategies σ that map from history space to action space, so that actions depend on the history, i.e realized states in current and past periods

Current state y_t is known at the beginning of the period.

Markov strategies are a special class of strategies where strategies only depend on the current state, and all past states are irrelevant to the strategy. We will not prove here that a Markov strategy exists, but we take it as given and focus on this case. The relevant state is then the current state at each period, and we consider strategies that are functions of that state, $\sigma(y)$.

A value function at time t is the average discounted value given state y,

$$v(y_t, \sigma(y_t)) = (1-\delta)\sum_{t=1}^{\infty}\sum_{y_t \in \mathcal{Y}} \delta^{t-1} u(\sigma(y_t), y_t)\pi(y_t|y_{t-1}), \sigma(y_{t-1})) \tag{9}$$

The **dynamic programming problem** at any time period and at any state y is to obtain **the value function**, i.e maximizing the above equation by choosing $\sigma \in \Sigma$. **The value function** is given by

$$V(y_t) = max_{\sigma\in\Sigma}\ (1-\delta)\sum_{t=1}^{\infty}\sum_{y_t \in \mathcal{Y}} \delta^{t-1} u(\sigma(y_t), y_t)\pi(y_t|y_{t-1}, \sigma(y_{t-1})) = max_{\sigma\in\Sigma}\ v(y_t, \sigma) \tag{10}$$

such that it is the value associated with the strategy that maximizes the average discounted payoff.

We can write this expression by separating the current utility and the discounted average utility of all future periods:

$$V(y_t) = max_{\sigma\in\Sigma}\ (1-\delta)u(\sigma(y), y) + \delta(1-\delta)\sum_{t=1}^{\infty}\sum_{y_{t+1} \in \mathcal{Y}} \delta^{t-1} u(\sigma(y_{t+1}), y_{t+1}))\pi(y_{t+1}|y_t, \sigma(y_t))$$

21

[21]Notice that we denote **the** value function with capital V, and **a** value function with small v.

Notice that the second part of the RHS of the above equation is $\delta v(y_{t+1})$, as it is defined in Equation 9. We can then write the problem as

$$V(y) = max_{\sigma \in \Sigma} \ (1-\delta)u(\sigma(h_1), y(h_1)) + \delta v(y')\pi(y'|y, \sigma(y))$$

This is called the Bellman equation. The decision problem is the same at every period given a particular state. If I face the same state now or in ten periods does not change the nature of the decision problem. This stationarity of the problem allows us to omit time subscripts that we have been using so far. **The value function**, that is, the solution to this optimization problem is the fixed point of the Bellman equation[22]: the value associated with the optimal strategy of a given state in the current period is the same value in the subsequent periods given that state. We are looking for the V such that the value functions on the RHS and LHS in the Bellman equation are the same.

To gain intuition, let us consider what $V(y)$. Remember that the state space is finite, and can be represented as a vector. V is then also a vector, with each element $V(y)$ corresponding to the maximized value of a given state.

Take the job search model in the lecture slides. Our state space is given by $\mathcal{Y} = \{unemployed, bad\ job,\ good\ job\}$. The value function is a vector of three elements consisting of the maximized value for each state.

$$V = \begin{bmatrix} V(unemployed) \\ V(bad\ job) \\ V(good\ job) \end{bmatrix}$$

Action space when state is bad job is to quit the job, becoming unemployed, or to

[22]The existence of the fixed point of the Bellman equation is omitted from these notes, you can check the lecture slides for further information.

continue having a bad job, It is an empty set when it is unemployed or having a good job. This is because in this problem, an unemployed person gets exogenously matched with a good job with probability a, with a bad job with a probability c, and the state of good job is absorbing. A person with a bad job has an exogenous probability of moving up to a good job of b. Per period payoff of being unemployed is 0, of having a bad job is 1, and of having a good job is d.

We can then write.

$$\begin{bmatrix} V(u) \\ V(b) \\ V(g) \end{bmatrix} = \begin{bmatrix} (1-\delta)0 + \delta(aV(g) + cV(b) + (1-a-c)V(u)) \\ max\{(1-\delta)1 + \delta(bV(g) + (1-b)V(b)), (1-\delta)1 + \delta V(u)\} \\ d \end{bmatrix} \tag{11}$$

That the value function V is the fixed point to the Bellman equation tells us that the $V(u), V(b), V(g)$ on the RHS are the same as the $V(u), V(b), V(g)$ on the LHS. We can then solve the optimization problem of the agent with the bad job (she is the only one facing a decision here) in the following way:

1. Plug in d for $V(g)$ in the value of being unemployed and having a bad job:

$$V(b) = max\{(1-\delta)1 + \delta(bd + (1-b)V(b)), (1-\delta)1 + \delta V(u)\}$$

$$V(u) = \delta(ad + cV(b) + (1-a-c)V(u))$$

2. When quitting the job is optimal, it means that

$$V(b) = (1-\delta) + \delta V(u)$$

3. Substitute this expression into $V(u)$

$$V(u) = \delta(ad + c((1-\delta) + \delta V(u)) + (1-a-c)V(u))$$
$$V(u) = \frac{ad + c(1-\delta)}{1 - c\delta^2 - \delta(1-a-c)}$$ [23]

4. Now recall we are in the case where quitting is optimal so that $\delta V(u) \geq +\delta(bd + (1-b)V(b))$ Replacing $V(b) = (1-\delta) + \delta V(u)$

$$V(u) \geq (bd + (1-b)((1-\delta) + \delta V(u))$$

$$V(u) \geq \frac{bd + (1-b)((1-\delta)}{1-\delta(1-b)}$$

Using our expression for $V(u)$ that we derived,

$$\frac{ad + c(1-\delta)}{1 - c\delta^2 - \delta(1-a-c)} \geq \frac{bd + (1-b)((1-\delta)}{1-\delta(1-b)}$$

so that when the parameters satisfy this condition, an agent with a bad job finds it optimal to quit, receiving 1 today and becoming unemployed in the next period. The stationarity of the problem means that this is true for any time period where the agent has a bad job. Furthermore, when making decisions today, the agent knows this value function, and knows that he would quit a bad job. When she decides to quit the job and become unemployed, she does so knowing the value of being unemployed $V(u)$, which embeds also the case of being matched with a bad job $V(b)$, in which case the agent would again quit the job.

[23] We can do that precisely because V is the fixed point of the Bellman equation such that $V(u)$ on both RHS and LHS are the same.

Games with Incomplete Information

10 Bayesian Games

Up until now, payoffs for players depended on the strategy profile, and their position (whether they are Player 1, 2, and so on). In Bayesian games, we encounter a different setting of information, that is, incomplete information over payoffs of opponents. Naturally, not knowing the utility an opponent derives from an outcome creates an uncertainty over their strategies. In order to model how players act given this type of uncertainty, we think that there are different games that can be played depending on the payoffs of players. Payoffs are private information to the players, which determine their **types**. In general, types encompass all private information that is relevant for the game. While players do not know which type of opponents they face, they know the probability distribution over types, and can form expectations. Such games are called **Bayesian games**.

A Bayesian game starts with Nature, who acts first and assigns a type to the players. Everyone knows their own type. The types may or may not be correlated, perfectly or imperfectly among the players. That is, my type may or may not be informative about the types of my opponents. With own type known, and expectation over the types of opponents computed, the players play the game.

A classical example for this setting is principle-agent problems. Nature goes first and assigns a level of productivity or a labour cost to an employee, the agent. This information is private information for the employee. The employer, principal, on the other hand would know the distribution from which Nature drew: she would know that 10% of employees have high marginal productivity, whereas 90% has low productivity.

Another example is the Cournot game where firms decide on quantity. In the version we studied in the first part of the class, the profit function of both firms were common knowledge: $\pi_i = q_i(p - c_i)$. Suppose now that there are different types of firms that differ in their marginal cost of production, such that c_i is private information. It can be thought as a random variable that is drawn from a cost distribution $f(c)$: with 50% it is $c = 1$, and 50% it is $c = 3$. Nature acts first, assigns a cost to each of the firms with equal probability. Then, a firm does not know exactly the payoff/profit function of its rival, but knows the distribution from which the rival's (and own) costs are drawn. Notice that it in this case, because the draws are iid, the type of one firm is not informative about the rival's type.

Let us now formally define a Bayesian game.

Definition 10.1 (Bayesian games)**.** *Formally, a Bayesian game consists of the following elements:*

1. *A stage game with a finite set of N players, $i \in \mathcal{I}$, $\{1, 2, ...N\}$,*
2. *For each player i a finite set of types $\theta_i \in \Theta_i$,*
3. *A prior probability distribution of types for each players, denoted $p(\theta)$, known to all players, where $p(\theta_{-i}|\theta_i)$ denotes the conditional probability distribution over opponents' types given one's own type, computed using the Bayes' law,*
4. *For each player i, a finite (pure) action space A_i, which is a list that consists of the actions available to them and gives rise to set of action profiles A,*
5. *For each player i, a payoff function $u^i(a, \theta)$, which represents the utility that Player i derives from the outcome of the game a given all players' types θ.*

11 Solution Concepts

11.1 Bayesian Equilibrium

Definition 11.1. *A Bayesian equilibrium or Bayes-Nash equilibrium of game is a Nash equilibrium of a Bayesian game where strategies are* ***maps from types to actions:*** $s_i : \Theta_i \to A_i$*. That is, given own type* θ_i*, strategies of all types of opponents', and the probability distribution of opponents' types.* $s_i(\theta_i)$ *is payoff maximizing:*

$$s_i(\theta_i) = \; max_{s_i} \sum_{\theta_{-i}} p(\theta_{-i}|\theta_i) u_i(s_i, s_{-i}(\theta_{-i}), \theta)$$

where θ *is the vector of types of all players.*

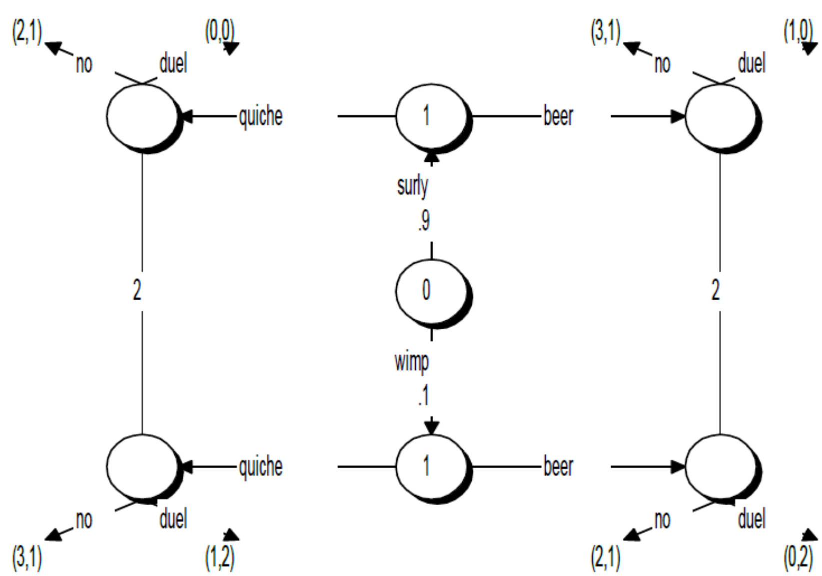
Each player's strategy consists of choosing a (pure or mixed) action for each of her types, such that they maximize payoffs given the strategies of opponents and the probability distribution over different types of opponents.

We differentiate between **separating, semi-separating and pooling equilibria**. Separating or semi-separating equilibria are those where different types take different actions such that, if there is a player who observes the action taken before they move, they can infer information on which type they are faced with **in equilibrium** (this information is perfect in separating equilibria). Pooling equilibria are those where strategies are the same for all types for a player $s_i(\theta) = s_i(\theta_l) \; \forall l \neq k$.

We say that Bayesian equilibrium is the Nash equilibrium of a Bayes game. Recall what how we defined a Nash equilibrium: 1) there are no profitable deviations given the strategies of opponents, and 2) players have perfect knowledge over the strategies of opponents.

Recall also that in Nash equilibrium, we could use players' indifference in order to support some equilibria through some unreasonable strategies on the off-path information sets. A similar problem is encountered with the Bayesian equilibrium of Bayes games, which is why we introduce sequential rationality. Before we delve into sequential rationality, let us give the classic example of Beer-or-Quiche game, which has politically not aged very well.

Example 11.1. *Consider the following Bayesian game. First, Nature moves and decides whether a man is surly with probability* 0.9*, or wimp with* $p = 0.1$*. The man, who knows his type, then walks into a bar where he makes an order for breakfast. If he is surly, he prefers to have a beer, and if he is a wimp, he prefers a quiche. Eating the most preferred option brings 3, and the other option brings 2 as payoff. There is a second man, Player 2, who observes what the first man eats, but not his type. This is depicted below as the information sets of Player 2. Once his information set, either beer or quiche is reached, Player 2 decides to duel Player 1 or not. Player 2 would obtain 1 if he does not duel. 2 from duelling a wimp, and 0 if he duels a surly type. Both types want to avoid duelling, it costs both of them 2.*



When faced with such a game, we consider the possible outcomes, and ask if they satisfy Nash equilibrium conditions. With two types and two actions available for each type, we can have four candidate equilibria in pure strategies: 2 pooling, one where each orders beer, one where each orders quiche and 2 separating: one where surly orders beer, the wimp the quiche and one where it is the other way around. In order to check if these can be sustained as equilibria, we need to see 1) what Player 2 would do in each case, when faced with an action, and 2) given Player 2's best responses, if players have any incentive to deviate.

Let us focus on a pooling equilibrium, and do an informal analysis in order to motivate the need for sequential rationality, which we will introduce later. Take the following equilibrium candidate[24]: both types order quiche. A Bayesian equilibrium in this game consists of a strategy profile where each player decides on a complete contingent plan for actions on all their information sets. In our example, Player 1s have one information set each, we propose that they both order Quiche. Player 2 has two information sets: one that is reached when Quiche is played, and one when Beer is played. Our proposal is play Not duel if Quiche is played, and Duel if Beer is played. Let us see if any player has a profitable deviation. If not, we have a Bayesian equilibrium.

At the information set after Quiche is played:

Best responses for Player 2: When the Player 2 observes quiche, he cannot infer new information from this action since all types are taking the same action in this equilibrium candidate. Player 2's belief that he is faced with a surly type, observing quiche is $\mu_2(s) = P(q|b) = \frac{P(q|s)P(s)}{P(q)} = \frac{1^*0.9}{1} = 0.9$. He prefers not to duel, since the expected utility of duelling is $E[D|\mu_2(s)] = 0\mu_2(s) + 2(1 - \mu_2(s)2 = 0.2$ versus the utility of not duelling, 1. Surly Player 1 obtains 2, and wimp Player 1 obtains 3.

At the information set after Beer is played, off-equilibrium path:

[24]Equilibrium candidate is simply a strategy profile, which we then analyze to decide if any player has incentive to deviate. Any strategy profile is an equilibrium candidate, and this is also the case for static or extensive form games of perfect information.

What is the best response of Player 2 if he observes that the Player 1 orders a beer? This is an action that does not happen in equilibrium. Remember, we want to sustain $\sigma_1(s) = \sigma_1(w) = Quiche$ as an equilibrium such that no player 1 wants to deviate. We can say that Player 2 duels when he sees beer. Is that a best response for Player 2? Yes, because remember that this is an action that never happens in equilibrium, so Player 2 is indifferent between duelling or not if that information set is reached. As before with extensive form games of complete information, Nash equilibria can be sustained through indifferences at information sets that are not reached in equilibrium.

Given that Player 2 will duel if beer is ordered, would any of the Player 1s deviate to beer? Player 1 of type wimp would not, he is having his favourite meal and not being fought. Surly Player 1 actually wants to have the beer. Would he deviate from quiche to beer, knowing that he will be duelled if he does? He would receive 1 with this deviation, less than what he receives from playing *Quiche*.

We see that if both Player 1s play *Quiche*, and Player 2 duels upon observing beer, Player 2 is best responding by not duelling when observing quiche, and no Player 1 would want to deviate. We just showed that the strategy profile $\{\sigma_1^s = \sigma_1^w = Quiche, \sigma_2(quiche) = Not\ duel, \sigma_2(beer) = Duel\}\}$ is **a Bayesian equilibrium.**

11.2 Sequentially Rational Equilibrium

Belief Formation

Before we move on to sequentially rational equilibrium, let us talk about how we define beliefs. In games of private or incomplete information, players need to form beliefs over the types of their opponents. A **belief** in Bayesian games is a **probability distribution over the types** conditional on prior probabilities, and actions. That means that the belief of any player would depend on their prior beliefs, that is, the probability distribution they have before observing any action, as well as the actions they observe. This process of updating beliefs upon observing some actions happens through Bayes law. The belief that the opponent is of type θ, given action a is given by $\mu(\theta|a) = \frac{P(a|\theta)P(\theta)}{P(a)}$ where $P(\theta)$ is the prior belief. In this class, we focus on games where the prior probability distribution is common knowledge such that all players share the same prior belief, which corresponds to the distribution from which Nature draws.

A **belief system** is an assignment of probabilities to every node in the game such that the sum of beliefs in any information set is 1.

Sequentially Rational Equilibrium

Recall the Chain Store game from the extensive form games. There were two equilibria, one of which we could not sustain as a subgame perfect equilibrium. The reason was that once we do arrive at a subgame, the strategy was no longer optimal. A similar problem arises extensive games of incomplete information. Why would Player 2 duel when they see beer being served? If actions are optimal given beliefs, some belief must be supporting the action of duelling. With sequential rationality, we want to have **consistency between actions and beliefs also along the off-equilibrium path**, and ensure that **players make optimal decisions to well-defined problems at all information sets**.

A **sequentially rational equilibrium**, which consists of both a strategy profile σ, and a belief system μ, we require consistency of beliefs and actions at all information sets, including those not reached in equilibrium. **Beliefs that are consistent with actions are derived through Bayesian updating.** In other words, when we talk about consistent beliefs, we mean that these beliefs are obtained by Bayesian updating using the prior probability distribution and actions of players. Consistency of beliefs imply that each belief should be updated according to the equilibrium strategies, the observed actions, and Bayes' rule on every path.

The problem we encountered in the above example was that when faced with actions that happen with zero probability, we cannot use Bayes law to update beliefs. In other words, the decision problem was not well-defined. In order to be still able to use the Bayes law so to derive beliefs that are consistent with actions, we perturb the strategy profile in a way that all actions receive positive probability. This perturbed strategy profile should converge to the equilibrium strategy profile at the limit. Beliefs are then constructed using this perturbed sequence. For instance, the equilibrium candidate we analyzed above had $(\sigma_1^s(Quiche) = 1, \sigma_1^w(Quiche) = 1)$, so that $\sigma_1^\theta(Beer) = 0$ for all θ. A perturbed strategy that converges to this profile would be $(1-\varepsilon^k, 1-\varepsilon^k)$ so that there is positive probability attached to both types playing beer. We can then use Bayesian updating to derive beliefs that are consistent with this perturbed strategy profile.

This procedure allows us to derive **consistent beliefs**.

Let us formally define sequentially rational equilibrium.

Definition 11.2. *A sequential equilibrium is a strategy profile σ and a belief μ_i for each player i such that beliefs are consistent and each player optimizes at each information set.*

Now let us solve Beer Quiche game, and find sequentially rational equilibria. We will also make distinctions between sequential rationality and Bayesian equilibrium. Let us first consider the two separating, then the two pooling equilibria.

Before starting to evaluate the equilibria, it is useful to 1) derive the best responses of the Player 2 depending on his beliefs, and 2) note down the most preferred outcomes for the Player 1s of different types.

Player 2 is indifferent between duelling and not when

$$u(D|\mu_2(s)) = u(ND|\mu_2(s))$$

$$0 * \mu_2(s) + 2(1 - \mu_2(s)) = 1 \rightarrow \mu_2(s) = \frac{1}{2}$$

We can write the best response correspondence as:

$$BR_2(\mu_2(s)) = \begin{cases} Duel \text{ if } \mu_2(s) < 1/2 \\ \Delta\{D, N\} \text{ if } \mu_2(s) = 1/2 \\ Not\ duel \text{ if } \mu_2(s) > 1/2 \end{cases}$$

Then, note that both types get a payoff of 2 from not duelling, regardless of what they eat, and if they also eat what they prefer the most, 1. That means that while their most preferred outcome is not having to duel, and having their most preferred option, the former is more important than the latter.

Separating Equilibria

1. Separating equilibrium 1:($\sigma_1^s = Beer, \sigma_1^w = Quiche$)

   - We start with Player 2. Upon seeing Beer, he knows that he is facing a surly type: $\mu_2(s|b) = \frac{p(beer|surly)p(surly)}{p(beer)} = \frac{1*0.9}{0.9*1+0*0.1} = 1$. We can see from the best response correspondence that he does not duel with these beliefs.
   - Upon seeing Quiche, he knows that he is facing a wimp type: $\mu_2(s|q) = \frac{p(quiche|surly)p(surly)}{p(quiche)} = \frac{0*0.9}{0.9*0+1*0.1)} = 0$. We can see from the best response correspondence that he duels.
   - Player 1 of surly type does not want to deviate to quiche, he is obtaining his best possible outcome, 3.
   - Player 1 of wimp type does want to deviate: he is having quiche and being duelled, getting 1. He would get 2 by choosing beer, leading the Player 2 to believe he is of type surly. We conclude that this cannot be an equilibrium.

Notice that it this case, we could use Bayes' law for both actions since they are both part of the equilibrium. Sequentially rational equilibrium corresponds to Bayesian equilibrium. This is always the case when all actions are played with positive probability in equilibrium.

2. Separating equilibrium 2:($\sigma_1^s = Quiche, \sigma_1^w = Beer$)

   - Upon seeing Quiche, Player 2 knows that he is facing a surly type: $\mu_2(s|q) = \frac{p(quiche|surly)p(surly)}{p(quiche)} = \frac{1*0.9}{0.9*1+0*0.1)} = 1$. We can see from the best response correspondence that he does not duel with these beliefs.
   - Upon seeing Beer, he knows that he is facing a wimp type: $\mu_2(s|b) = \frac{p(beer|surly)p(surly)}{p(beer)} = \frac{0*0.9}{0.9*0+1*0.1)} = 0$. We can see from the best response correspondence that he duels.
   - Player 1 of surly type does not want to deviate to beer. He does not like quiche all that much, but he is avoiding a duel.
   - Player 1 of wimp type does want to deviate: he is having beer and being duelled, that is the worst thing that can happen to him. He would get 3 by choosing quiche, leading the Player 2 to believe he is of type surly. We conclude that this cannot be an equilibrium.

3. Pooling equilibrium 1:($\sigma_1^s = \sigma_1^w = Beer$)

   - Upon seeing Beer, Player 2 has no new information: $\mu_2(s|b = \frac{p(beer|surly)p(surly)}{p(beer)} = \frac{1*0.9}{0.9*1+1*0.1)} = 0.9$. We can see from the best response correspondence that he does not duel with these beliefs.
   - Observing quiche is something that happens with 0 probability in equilibrium. We cannot apply Bayes law, and we need to construct a converging sequence to derive beliefs. Before we do so, we need to think of possible deviations by Player 1s, and how we can avoid them:

- Player 1 of surly type does not want to deviate to quiche, he is obtaining his best possible outcome, 3.
- Player 1 of wimp type may want to deviate to quiche if he were to be not duelled. Then, we need to construct the sequence in such a way that Player 2 believes he is facing a wimp with probability higher than 1/2, and duels. Then Player 1 of wimp type would not want to deviate.
- Take ($\sigma_1^s(Beer) = 1 - \varepsilon^{2k}, \sigma_1^w(Beer) = 1 - \varepsilon^k$), such that Player 1 of type wimp's probability of playing Quiche is ε^k), surly type's is ε^{2k}. It converges to (1,1), our equilibrium candidate.
- Observing quiche, beliefs are constructed as

$$\mu_2^k(s|quiche, \sigma^k) = \frac{p(quiche|surly)p(surly)}{p(quiche))} = \frac{\varepsilon^{2k}0.9}{0.9\varepsilon^{2k} + \varepsilon^k 0.1} = \frac{\varepsilon^k 0.9}{0.9\varepsilon^k + 0.1}$$

- Observing the equilibrium action on the other hand, we have

$$\mu_2^k(s|beer, \sigma^k) = \frac{p(beer|surly)p(beer)}{p(beer))} = \frac{(1-\varepsilon^{2k})0.9}{(1-\varepsilon^{2k})0.9 + 0.1(1-\varepsilon^k)}$$

- At the limit, we have $lim_{k\to\infty}\mu^k(quiche|\sigma^k) = \mu(s|quiche, \sigma) = 0$, and $lim_{k\to\infty}\mu^k(s|beer, \sigma^k) = \mu(s|beer, \sigma) = 0.9$.

- With this belief system, Player 2 believes that he sees a wimp when he observes the off-path action Quiche and duels, and assigns probability 0.9 to type being surly when he observes Beer. Notice that it would work with any sequence that leads Player 2 to believe the type is surly with probability less than 1/2.
- Player 1 of surly type does not want to deviate to quiche, he is obtaining his best possible outcome, 3.
- Player 1 of wimp type does not want to deviate: he is having beer avoiding a duel, getting 2. He would get 1 by choosing quiche, revealing his type. We conclude that this is an equilibrium, and we write it as

$$(\sigma, \mu) = \{\sigma_1^s = \sigma_1^w = b, \mu_2(s|b) = 0.9, \sigma_2(b) = ND, \mu_2(s|q) < 1/2, \sigma_2(q) = D\}$$

with $(\sigma, \mu) = lim_{k\to\infty}(\sigma^k, \mu^k)$ where $\sigma^k = (\sigma_1^s(Beer) = 1 - \varepsilon^{2k}, \sigma_1^w(Beer) = 1 - \varepsilon^k)$

4. Pooling equilibrium 2:$(\sigma_1^s = \sigma_1^w = Quiche)$

- Upon seeing Quiche, Player 2 has no new information: $\mu_2(s|q) = \frac{p(quiche|surly)p(surly)}{p(quiche)} = \frac{1*0.9}{0.9*1+1*0.1)} = 0.9$. We can see from the best response correspondence that he does not duel with these beliefs.
- Observing beer is something that happens with 0 probability in equilibrium. We cannot apply Bayes law, and we need to construct a converging sequence to derive beliefs. Before we do so, we need to think of possible deviations by Player 1s, and how we can avoid them:
  - Player 1 of wimp type does not want to deviate to quiche, he is obtaining his best possible outcome, 3.
  - Player 1 of surly type may want to deviate to quiche if he were to be not duelled. Then, we need to construct the sequence in such a way that Player 2 believes he is facing a wimp with probability higher than 1/2, and duels. Then Player 1 of surly type would not want to deviate.
  - Take $(\sigma_1^s(Quiche) = 1 - \varepsilon^{2k}, \sigma_1^w(Quiche) = 1 - \varepsilon^k)$, such that Player 1 of type wimp's probability of playing Quiche is ε^k, surly type's is ε^{2k}. It converges to (1,1), our equilibrium candidate.
  - Observing beer, beliefs are constructed as
  $$\mu^k(s|beer, \sigma^k) = \frac{p(beer|surly)p(surly)}{p(beer))} = \frac{\varepsilon^{2k}0.9}{0.9\varepsilon^{2k} + \varepsilon^k 0.1} = \frac{\varepsilon^k 0.9}{0.9\varepsilon^k + 0.1}$$
  - At the limit, we have $lim_{k\to\infty}\mu_2^k(s|beer, \sigma^k) = \mu_2(s|beer, \sigma) = 0$ and $lim_{k\to\infty}\mu^k(s|quiche, \sigma^k) = \mu(s|quiche, \sigma) = 0.9$.
  - With this belief, Player 2 believes that he sees a wimp when he observes the off-path action Beer and duels. Notice that it would work with any sequence that leads Player 2 to believe the type is surly with probability less than 1/2.

- Player 1 of wimp type does not want to deviate to quiche, he is obtaining his best possible outcome, 3.
- Player 1 of surly type does not want to deviate: he is having quiche and avoiding a duel, getting 2. He would get 1 by choosing beer, leading the Player 2 to think he is wimp and getting duelled. We conclude that this is an equilibrium, and we write it as

  $(\sigma, \mu) = \{\sigma_1^s = \sigma_1^w = q, \mu_2(s|q) = 0.9, \sigma_2(q) = ND, \mu_2(s|b) < 1/2, , \sigma_2(b) = D\}$

  with $(\sigma, \mu) = lim_{k\to\infty}(\sigma^k, \mu^k)$ where $\sigma^k = (\sigma_1^s(Quiche) = 1 - \varepsilon^{2k}, \sigma_1^w(Quiche) = 1 - \varepsilon^k)$

This game has two pure strategy pooling equilibria that are sequentially rational, and no pure strategy separating equilibrium.

Keep in mind

- Bayesian equilibrium is a weaker concept than sequentially rational equilibrium. If you can show that an equilibrium is not Bayesian, then you also show it is not a sequentially rational equilibrium. This is what we do in Question 2 of PS4, pooling equilibrium with low settlement.
- When constructing a sequence, think of which action you want to induce for the player who will respond to the perturbed strategies. Which type do you want them to believe they are facing? Looking at best responses here would help.
- After you know which type, give higher probability of deviation to that type. Remember $\varepsilon^k > \varepsilon^{2k}$
- When you use your converging sequence to derive beliefs, do not forget to do that both for on and off-path actions. That way, you are checking if your sequence is indeed converging to the equilibrium strategy, also inducing the equilibrium beliefs.