1 Preliminary Remarks.

Game theory is a mathematical framework for analyzing conflict and cooperation. Early work was motivated by gambling and recreational games such as chess, hence the “game” in game theory. But it quickly became clear that the framework had much broader application. Today, game theory is used for mathematical modeling in a wide range of disciplines, including many of the social sciences, computer science, and evolutionary biology. In my notes, I draw examples mainly from economics.

These particular notes are an introduction to a formalism called a strategic form (also called a normal form). For the moment, think of a strategic form game as representing an atemporal interaction: each player (in the language of game theory) acts without knowing what the other players have done. An example is a single instance of the two-player game Rock-Paper-Scissors (probably already familiar to you, but discussed in the next section).

In companion notes, I develop an alternative formalism called an extensive form game. Extensive form games explicitly capture temporal considerations, such as the fact that in standard chess, players move in sequence, and each player knows the prior moves in the game. As I discuss in the notes on extensive form games, there is a natural way to give any extensive form game a strategic form representation. Moreover, the benchmark prediction for an extensive form game, namely that behavior will conform to a Nash equilibrium, is defined in terms of the strategic form representation.

There is a third formalism called a game in coalition form (also called characteristic function form). The coalition form abstracts away from the details of what players can do and focuses instead on what outcomes are physically possible for each coalition (group) of players. I do not (yet) have notes on games in coalition form.

The first attempt at a general treatment of games, von Neumann and Morgenstern (1944), analyzed both strategic form and coalition form games, but merged the analysis in a way that made it impossible to use game theory to study issues such as individual incentives to abide by agreements. The work of John Nash, in Nash (1951) and Nash (1953), persuaded game theorists to maintain a division between the analysis of strategic and coalition form games. For discussions of some of the

\footnote{\textcopyright 2018. This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License.}
history of game theory, and Nash’s contribution in particular, see Myerson (1999) and Nachbar and Weinstein (2016).

Nash (1951) called the study of strategic/extensive form games “non-cooperative” game theory and the study of coalition form games “cooperative game theory”. These names have stuck, but I prefer the terms strategic/extensive form game theory and coalition form game theory, because the non-cooperative/cooperative language sets up misleading expectations. In particular, it is not true that the study of games in strategic/extensive form assumes away cooperation.

A nice, short introduction to the study of strategic and extensive form games is Osborne (2008). A standard undergraduate text on game theory is Gibbons (1992). A standard graduate game theory text is Fudenberg and Tirole (1991); I also like Osborne and Rubinstein (1994). There are also good introductions to game theory in graduate microeconomic theory texts such as Kreps (1990), Mas-Colell, Whinston and Green (1995), and Jehle and Reny (2000). Finally, Luce and Raiffa (1957) is a classic and still valuable text on game theory, especially for discussions of interpretation and motivation.

I start my discussion of strategic form games with Rock-Paper-Scissors.

2 An example: Rock-Paper-Scissors.

The game Rock-Paper-Scissors (RPS) is represented in Figure 1 in what is called a game box. There are two players, 1 and 2. Each player has three strategies in the game: R (rock), P (paper), and S (scissors). Player 1 is represented by the rows while player 2 is represented by the columns. If player 1 chooses R and player 2 chooses P then this is represented as the pair, called a strategy profile, (R,P) and the result is that player 1 gets a payoff of -1 and player 2 gets a payoff of +1, represented as a payoff profile \((-1,1)\).

For interpretation, think of payoffs as encoding preferences over winning, losing, or tying, with the understanding that S beats P (because scissors cut paper), P beats R (because paper can wrap a rock . . . ), and R beats S (because a rock can smash scissors). If both choose the same, then they tie. The interpretation of payoffs is actually quite delicate and I discuss this issue at length in Section 3.3.

This game is called zero-sum because, for any strategy profile, the sum of payoffs is zero. In any zero-sum game, there is a number \(V\), called the value of the game,
with the property that player 1 can guarantee that she gets at least $V$ no matter what player 2 does and conversely player 2 can get $-V$ no matter what player 1 does. I provide a proof of this theorem in Section 4.5. In this particular game, $V = 0$ and both players can guarantee that they get 0 by randomizing evenly over the three strategies.

Note that randomization is necessary to guarantee a payoff of at least 0. In Season 4 Episode 16 of the Simpsons, Bart persistently plays Rock against Lisa, and Lisa plays Paper, and wins. Bart here doesn’t even seem to understand the game box, since he says, “Good old Rock. Nothing beats that.” I discuss the interpretation of randomization in Section 3.4.

3 Strategic Form Games.

3.1 The Strategic Form.

I restrict attention in the formalism to finite games: finite numbers of players and finite numbers of strategies. Some of the examples, however, involve games with infinite numbers of strategies.

A strategic form game is a tuple $(I, (S_i)_i, (u_i)_i)$.

- $I$ is the finite set, assumed not empty, of players with typical element $i$. The cardinality of $I$ is $N$; I sometimes refer to $N$-player games. To avoid triviality, assume $N \geq 2$ unless explicitly stated otherwise.

- $S_i$ is the set, assumed finite and not empty, of player $i$’s strategies, often called pure strategies to distinguish from the mixed strategies described below.

- $S = \prod_i S_i$ is the set of pure strategy profiles, with typical element $s = (s_1, \ldots, s_N) \in S$.

- $u_i$ is player $i$’s payoff function: $u_i : S \to \mathbb{R}$. I discuss the interpretation of payoffs in Section 3.3.

Games with two players and small strategy sets can be represented via a game box, as in Figure 1 for Rock-Paper-Scissors. In that example, $I = \{1, 2\}$, $S_1 = S_2 = \{R, P, S\}$, and payoffs are as given in the game box.

As already anticipated by the RPS example, we will be interested in randomization. For each $i$, let $\Sigma_i$ be the set of probabilities over $S_i$, also denoted $\Delta(S_i)$. An element $\sigma_i \in \Sigma_i$ is called a mixed strategy for player $i$. Under $\sigma_i$, the probability that $i$ plays pure strategy $s_i$ is $\sigma_i[s_i]$.

A pure strategy $s_i$ is equivalent to a degenerate mixed strategy, with $\sigma_i[s_i] = 1$ and $\sigma_i[\hat{s}_i] = 0$ for all $\hat{s}_i \neq s_i$. Abusing notation, I use the notation $s_i$ for both the pure strategy $s_i$ and for the equivalent mixed strategy.
Assuming that the cardinality of $S_i$ is at least 2, the strategy $\sigma_i$ is **fully mixed** iff $\sigma_i[s_i] > 0$ for every $s_i \in S_i$. $\sigma_i$ is **partly mixed** iff it is neither fully mixed nor pure (degenerate).

Since the strategy set is assumed finite, I can represent a $\sigma_i$ as a vector in either $\mathbb{R}^{|S_i|}$ or $\mathbb{R}^{|S_i|-1}$, where $|S_i|$ is the number of elements in $S_i$. For example, if $S_1$ has two elements, then I can represent $\sigma_1$ as either $(p, 1-p)$, with $p \in [0, 1]$ (probability $p$ on the first strategy, probability $1-p$ on the second), or just as $p \in [0, 1]$ (since the probabilities must sum to 1, one infers that the probability on the second strategy must be $1-p$). And a similar construction works for any finite strategy set. Note that under either representation, $\Sigma_i$ is compact and convex.

I discuss the interpretation of mixed strategies in Section 3.4. For the moment, however, suppose that players might actually randomize, perhaps by making use of coin flip or toss of a die. In this case the true set of strategies is actually $\Sigma_i$ rather than $S_i$.

$\Sigma = \prod_i \Sigma_i$ is the set of *mixed strategy profiles*, with typical element $\sigma = (\sigma_1, \ldots, \sigma_N)$. A mixed strategy profile $\sigma$ induces an independent probability distribution over $S$.

**Example 1.** Consider a two-player game in which $S_1 = \{T, B\}$ and $S_2 = \{L, R\}$. If $\sigma_1[T] = 1/4$ and $\sigma_2[L] = 1/3$ then the induced distribution over $S$ can be represented in a game box as in Figure 2.

<table>
<thead>
<tr>
<th></th>
<th>$L$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>1/12</td>
<td>2/12</td>
</tr>
<tr>
<td>$B$</td>
<td>3/12</td>
<td>6/12</td>
</tr>
</tbody>
</table>

Figure 2: An independent distribution over strategy profiles.

Abusing notation, let $u_i(\sigma)$ be the expected payoff under this independent distribution; that is,

$$u_i(\sigma) = \mathbb{E}_\sigma[u_i(s)] = \sum_{s \in S} u_i(s) \times \sigma_1[s_1] \times \cdots \times \sigma_N[s_N].$$

Finally, I frequently use the following notation. A strategy profile $s = (s_1, \ldots, s_N)$ can also be represented as $s = (s_i, s_{-i})$, where $s_{-i} \in \prod_{j \neq i} S_j$ is a profile of pure strategies for players other than $i$. Similarly, $\sigma = (\sigma_i, \sigma_{-i})$ is alternative notation for $\sigma = (\sigma_1, \ldots, \sigma_N)$, where $\sigma_{-i} \in \prod_{j \neq i} \Sigma_j$ is a profile of mixed strategies for players other than $i$.

### 3.2 Correlation.

The notation so far builds in an assumption that any randomization is independent. To see what is at issue, consider the following game.
Example 2. The game box for one version of Battle of the Sexes is in Figure 3. The players would like to coordinate on either \((A, A)\) or \((B, B)\), but they disagree about which of these is better. □

If players were actually to engage in the Battle of the Sexes game depicted in Figure 3, then one plausible outcome is that they would toss a coin prior to play and then execute \((A, A)\) if the coin lands heads and \((B, B)\) if the coin lands tails. This induces a correlated distribution over strategy profiles, which I represent in Figure 4.

\[
\begin{array}{cc}
A & B \\
A & 1/2 & 0 \\
B & 0 & 1/2 \\
\end{array}
\]

Figure 4: A correlated distribution over strategy profiles for Battle of the Sexes.

Note that under this correlated distribution, each player plays \(A\) half the time. If players were instead to play \(A\) half the time independently, then the distribution over strategy profiles would be as in Figure 5.

\[
\begin{array}{cc}
A & B \\
A & 1/4 & 1/4 \\
B & 1/4 & 1/4 \\
\end{array}
\]

Figure 5: An independent distribution strategy profiles for Battle of the Sexes.

The space of all probability distributions over \(S\) is \(\Sigma^c\), also denoted \(\Delta(S)\), with generic element denoted \(\sigma^c\). Abusing notation (again), I let \(u_i(\sigma^c) = \sum_{s \in S} u_i(s)\sigma^c[s]\). For strategies for players other than \(i\), the notation is \(\sigma^c_{-i} \in \Sigma_{-i} = \Delta(S_{-i})\). The notation \((\sigma_i, \sigma^c_{-i})\) denotes the distribution over \(S\) for which the probability of \(s = (s_i, s_{-i})\) is \(\sigma_i[s_i]\sigma^c_{-i}[s_{-i}]\). Any element of \(\Sigma\) induces an element of \(\Sigma^c\): an independent distribution over \(S\) is a special form of correlated distribution over \(S\).

Remark 1. A mixed strategy profile \(\sigma \in \Sigma\) is not an element of \(\Sigma^c\), hence \(\Sigma \not\subseteq \Sigma^c\). Rather, \(\sigma\) induces a distribution over \(S\), and this induced distribution is an element of \(\Sigma^c\).

In the Battle of the Sexes of Figure 3, to take one example, it takes three numbers to represent an arbitrary element \(\sigma^c \in \Sigma^c\) (three rather than four because the four
numbers have to add up to 1). In contrast, each mixed strategy $\sigma_i$ can be described by a single number (for example, the probability that $i$ plays $A$), hence $\sigma$ can be represented as a pair of numbers, which induces a distribution over $S$. Thus, in this example, the set of independent strategy distributions is two dimensional while the set of all strategy distributions is three dimensional. More generally, the set of independent strategy distributions is a lower dimensional subset of the set of all strategy distributions (namely $\Sigma^c$). □

3.3 Interpreting Payoffs.

In most applications in economics, payoffs are intended to encode choice by decision makers, and this brings with it some important subtleties.

To make the issues more transparent, I introduce some additional structure. Let $Y$ be a set of possible prizes for the game and let $\gamma(s) = (s, y)$ be the outcome (recording both the strategy profile and the resulting prize) when the strategy profile is $s$. Finally, let $X = S \times Y$ be the set of outcomes.\(^2\)

As in standard decision theory, I assume that players have preferences over $X$ and that, for each $i$, these preferences have a utility representation, say $v_i$. I include $s$ in the description of the outcome $(s, y)$ because it is possible that players care not only about the prizes but also about the way the game was played. The payoff function $u_i$ is thus a composite function: $u_i(s) = v_i(\gamma(s))$.

Example 3. Recall Rock-Paper-Scissors from Section 2. In the game as usually played, $Y = \{(\text{win, lose}), (\text{lose, win}), (\text{tie, tie})\}$, \(^3\) where, for example, (win, lose) means that player 1 gets “win” (whatever that might mean) while player 2 gets “lose”. $\gamma$ is given by, for example $\gamma(R, P) = ((R, P), (\text{lose, win}))$. The game box implicitly assumes that preferences over outcomes depend only on the prize, and not on $s$ directly. The game box assumes further that for player 1, the utility representation for preferences assigns a utility of 1 to (win, lose), -1 to (lose, win), and 0 to (tie, tie), with an analogous assignment for player 2. □

With this foundation, I can now make a number of remarks about the interpretation of payoffs in decision theoretic terms.

1. In many economic applications, the prize is a profile of monetary values (profits, for example). It is common practice in such applications to assume that $u_i(s)$ equals the prize to $i$: if $s$ gives $i$ profits of $1$ billion, then $u_i(s) = 1$ billion. This assumption is substantive.

\(^2\)I can accommodate the possibility that $s$ induces a non-degenerate distribution over $Y$, but I will not pursue this complication.
(a) The assumption rules out phenomena such as altruism or envy. In contrast, the general strategic form formalism allows, for example, for \( u_i(s) \) to be the sum of the individual prizes (a form of perfect altruism).

(b) The assumption rules out risk aversion. If a player is risk averse, then his payoff will be a concave function of his profit.

2. It can be difficult to “test” game theory predictions such as Nash equilibrium in a lab. The experimenter can control \( \gamma \), but the \( v_i \), and hence the \( u_i \), are in the heads of the subjects and not directly observable. In particular, it is not a violation of game theory to find that players are altruistic or spiteful. This flexibility of the game theory formalism is a feature, not a bug: the goal is to have a formalism that can be used to model essentially any strategic interaction.

3. If \( v_i \) represents choice over (pure) outcomes as in standard decision theory, then payoff maximization is built into the payoffs; it is not a separate assumption. But this logic does not carry over to lotteries. An assumption that players maximize expected payoffs is substantive.

4. While it may be reasonable to assume that there is common knowledge of \( \gamma \) (everyone knows the correct \( \gamma \), everyone knows that everyone knows the correct \( \gamma \), and so on), there may not be even mutual knowledge of the \( v_i \) and hence of the \( u_i \) (players may not know the utility functions of the other players). Games in which there is not common knowledge of the \( u_i \) are called games of incomplete information. I discuss approaches to modeling such environments later in the course.

The interpretation of payoffs in terms of decision theory is not the only one possible. For example, in some applications of game theory to evolutionary biology, the “strategies” might be alleles (alternative versions of a gene) and payoffs might be the expected number of offspring.

3.4 Interpreting Randomization.

There are three main interpretations of randomization in games. These interpretations are not mutually exclusive.

1. **Objective Randomization.** Each player has access to a randomization device. In this case, the true strategy set for player \( i \) is \( \Sigma_i \). \( S_i \) is just a concise way to communicate \( \Sigma_i \).

   An interesting case of this is reported in Sontag and Drew (1998). Military submarines occasionally implement hard turns in order to detect possible trailing submarines; such maneuvers are called “clearing the baffles.” In order to be as effective as possible, it is important that these turns be unpredictable.
Sontag and Drew (1998) reported that a captain of the USS Lapon used dice in order to randomize. Curiously, in Clancy (1984), a classic military techno-thriller, much of the plot turns on a CIA analyst correctly predicting when and how a (fictional) top Russian submarine commander would clear the baffles of his submarine.

2. *Empirical Randomization.* From the perspective of an observer (say an experimentalist), \( \sigma_i[s_i] \) is the frequency with which \( s_i \) is played.

The observer could, for example, be seeing data from a cross section of play by different players (think of a Rock-Paper-Scissors tournament with many simultaneous matchings). \( \sigma_i[s_i] \) is the fraction of players in role \( i \) of the game who play \( s_i \). Nash discussed this interpretation explicitly in his thesis, Nash (1950b). Alternatively, the observer could be seeing data from a time series: the same players play the same game over and over, and \( \sigma_i[s_i] \) is the frequency with which \( s_i \) is played over time.

3. *Subjective Randomization.* From the perspective of player \( j \neq i \), \( \sigma_i[s_i] \) is the probability that \( j \) assigns to player \( i \) playing \( s_i \).

Consider again the cross sectional interpretation of randomization, in which many instances of the game are played by different players. If players are matched randomly and anonymously to play the game then, from the perspective of an individual player, the opponents are drawn randomly, and hence opposing play can be “as if” random even if the player knows that individual opponents are playing pure strategies.

An important variant of the cross sectional interpretation of randomization is the following idea, due to Harsanyi (1973). As discussed in Section 3.3, players may know the \( \gamma \) of the game (giving prizes), but not the \( v_i \) of the other players (giving preferences over prizes), and hence may not know the \( u_i \) (giving payoffs) of the other players. Suppose that player \( j \) assigns a probability distribution over possible \( u_i \), and for each \( u_i \), player \( j \) forecasts play of some pure strategy \( s_i \). Then, even though player \( j \) thinks that player \( i \) will play a pure strategy, player \( i \)’s play is effectively random in the mind of player \( j \) because the distribution over \( u_i \) induces a distribution over \( s_i \).

A distinct idea is that in many games it can be important not to be predictable. This was the case in Rock-Paper-Scissors, for example (Section 2).

A particular concern is that if the game is played repeatedly then one’s behavior should not follow some easily detected, and exploited, pattern, such as \( R, P, S, R, P, S, R, \ldots \). A player can avoid predictability by literally randomizing each period. An alternative, pursued in Hu (2014), is the idea that even if a pure strategy exhibits a pattern that can be detected and exploited in principle, it may be impossible to do so in practice if the pattern is sufficiently complicated.
A subtlety with the subjective interpretation is that if there are three or more players, then two players might have different subjective beliefs about what a third player might do. This is assumed away in our notation, where \( \sigma_i \) does not depend on anything having to do with the other players.

4 Nash Equilibrium.

4.1 The Best Response Correspondence.

Given a profile of opposing (mixed) strategies \( \sigma_{-i} \in \Sigma_{-i} \), let \( BR_i(\sigma_{-i}) \) be the set of mixed strategies for player \( i \) that maximize player \( i \)'s expected payoff; formally,

\[
BR_i(\sigma_{-i}) = \{ \sigma_i \in \Sigma_i : \forall \hat{\sigma}_i \in \Sigma_i, u_i(\sigma_i, \sigma_{-i}) \geq u_i(\hat{\sigma}_i, \sigma_{-i}) \}.
\]

An element of \( BR_i(\sigma_{-i}) \) is called a best response to \( \sigma_{-i} \).

Given a profile of strategies \( \sigma \in \Sigma \), let \( BR(\sigma) \) be the set of mixed strategy profiles \( \hat{\sigma} \) such that, for each \( i \), \( \hat{\sigma}_i \) is a best response to \( \sigma_{-i} \). Formally,

\[
BR(\sigma) = \{ \hat{\sigma} \in \Sigma : \forall i \; \hat{\sigma}_i \in BR_i(\sigma_{-i}) \}.
\]

\( BR \) is a correspondence on \( \Sigma \). Since \( \Sigma_i \) is compact for each \( i \), \( \Sigma \) is compact. For each \( i \), expected payoffs are continuous, which implies that for any \( \sigma \in \Sigma \) and any \( i \), \( BR_i(\sigma_{-i}) \) is not empty. Thus, \( BR \) is a non-empty-valued correspondence on \( \Sigma \).

4.2 Nash Equilibrium.

The single most important solution concept for games in strategic form is Nash equilibrium (NE). Nash himself called it an “equilibrium point.” A NE is a (mixed) strategy profile \( \sigma^* \) such that for each \( i \), \( \sigma^*_i \) is a best response to \( \sigma^*_{-i} \), hence \( \sigma^* \in BR(\sigma^*) \); a NE is a fixed point of \( BR \).

**Definition 1.** Fix a game. A strategy profile \( \sigma^* \in \Sigma \) is a Nash equilibrium (NE) iff \( \sigma^* \) is a fixed point of \( BR \).

In separate notes (Game Theory Basics III), I survey some motivating stories for NE. It is easy to be led astray by motivating stories, however, and confuse what NE “ought” to be with what NE, as a formalism, actually is. For the moment, therefore, I focus narrowly on the formalism.

A NE is pure if all the strategies are pure. It is fully mixed if every \( \sigma_i \) is fully mixed. If a NE is neither pure nor fully mixed, then it is partly mixed. Strictly speaking, a pure NE is a special case of a mixed NE; in practice, however, I may sometimes (sloppily) write mixed NE when what I really mean is fully or partly mixed.

The following result was first established in Nash (1950a).
**Theorem 1.** Every (finite) strategic form game has at least one NE.

**Proof.** For each $i$, $\Sigma_i$ is compact and convex and hence $\Sigma$ is compact and convex. BR is a non-empty valued correspondence on $\Sigma$ and it is easy to show that it has a closed, convex graph. By the Kakutani fixed point theorem, BR has a fixed point $\sigma \in \text{BR}(\sigma)$.

**Remark 2.** Non-finite games may not have NE. A trivial example is the game in which you name a number $\alpha \in [0,1)$, and I pay you a billion dollars with probability $\alpha$ (and nothing with probability $1-\alpha$). Assuming that you prefer more (expected) money to less, you don’t have a best response and hence there is no NE.

The difficulty in the above example was that the strategy set was not compact. A related example has $\alpha \in [0,1]$, which is now compact, but with payoff of zero if you choose $\alpha = 1$. Here the problem is that payoff function is not continuous.

The most straightforward extension of Theorem 1 is to games in which each player’s strategy set is a compact subset of some measurable topological space and the payoff functions are continuous. In this case, the set of mixed strategies is weak-* compact. One can take finite approximations to the strategy sets, apply Theorem 1 to get a NE in each finite approximation, appeal to compactness to get a convergent subsequence of these mixed strategy profiles, and then argue, via continuity of payoffs, that the limit must be an equilibrium in the original game.

This approach will not work if the utility function is not continuous and, unfortunately, games with discontinuous utility functions are fairly common in economics. A well known example is the Bertrand duopoly game, discussed in Example 11 in Section 4.4. For a survey of the literature on existence of NE in general and also on existence of NE with important properties (e.g., pure strategy NE or NE with monotonicity properties), see Reny (2008).

**Remark 3.** A related point is the following. The Brouwer fixed point theorem states that if $D \subseteq \mathbb{R}^N$ is compact and convex and $f : D \to D$ is continuous then $f$ has a fixed point: there is an $x \in D$ such that $f(x) = x$. Brouwer is the underlying basis for the Kakutani fixed point theorem, which was used in the proof Theorem 1, and one can also use Brouwer more directly to prove NE existence (as, indeed, Nash himself subsequently did for finite games in Nash (1951)). Thus, Brouwer implies NE.

The converse is also true: given a compact convex set $D \subseteq \mathbb{R}^N$ and a continuous function $f : D \times D$, one can construct a game in which the strategy set for player 1 is $D$ and, if a NE exists, it must be that $s_1 = f(s_1)$. One such construction can be found here.

**4.3 NE and randomization.**

The following fact says that a best response gives positive probability to a pure strategy only if that pure strategy is, in its own right, also a best response.
Theorem 2. For any finite game and for any $\sigma_{-i}$, if $\sigma_i \in BR_i(\sigma_{-i})$ and $\sigma_i(s_i) > 0$ then $s_i \in BR_i(\sigma_{-i})$.

Proof. Since the set of pure strategies is assumed finite, one of these strategies, call it $s_i^*$, has highest expected payoff (against $\sigma_{-i}$) among all pure strategies. Let the expected payoff of $s_i^*$ be $c_i^*$. For any mixed strategy $\sigma_i$, the expected payoff is the convex sum of the expected payoffs to $i$'s pure strategies. Therefore, $c_i^*$ is the highest possible payoff to any mixed strategy for $i$, which implies that, in particular, $s_i^*$ is a best response, as is any other pure strategy that has an expected payoff of $c_i^*$.

I now argue by contraposition. Suppose $s_i \notin BR_i(\sigma_{-i})$, hence $u_i(s_i, \sigma_{-i}) < c_i^*$. If $\sigma_i[s_i] > 0$, then $u_i(\sigma_i, \sigma_{-i}) < c_i^*$, hence $\sigma_i \notin BR_i(\sigma_{-i})$, as was to be shown. □

Theorem 2 implies that a NE mixed $\sigma_i$ will give positive probability to two different pure strategies only if those pure strategies each earn the same expected payoff as the mixed strategy. That is, in a NE, a player is indifferent between all of the pure strategies that he plays with positive probability. This provides a way to compute mixed strategy NE, at least in principle.

Example 4 (Finding mixed NE in 2 x 2 games). Consider a general 2 x 2 game given by the game box in Figure 6. Let $q = \sigma_2(L)$. Then player 1 is indifferent between $T$ and $B$ iff

$$qa_1 + (1-q)b_1 = qc_1 + (1-q)d_1,$$

or

$$q = \frac{d_1 - b_1}{(a_1 - c_1) + (d_1 - b_1)},$$

provided, of course, that $(a_1 - c_1) + (d_1 - b_1) \neq 0$ and that this fraction is in $[0, 1]$.

Similarly, if $p = \sigma_1(T)$, then player 2 is indifferent between $L$ and $R$ if

$$pa_2 + (1-p)c_2 = pb_2 + (1-p)d_2,$$

or

$$p = \frac{d_2 - c_2}{(a_2 - b_2) + (d_2 - c_2)},$$

again, provided that $(a_1 - c_1) + (d_1 - b_1) \neq 0$ and that this fraction is in $[0, 1]$.

Note that the probabilities for player 1 are found by making the other player, player 2, indifferent. A common error is to do the calculations correctly, but then to write the NE incorrectly, by flipping the player roles. □
Example 5 (NE in Battle of the Sexes). Battle of the Sexes (Example 2) has two pure strategy NE: \((A, B)\) and \((B, A)\). There is also a mixed strategy NE with

\[
\sigma_1(A) = \frac{8 - 0}{(10 - 0) + (8 - 0)} = \frac{4}{9}
\]

and

\[
\sigma_2(A) = \frac{10 - 0}{(10 - 0) + (8 - 0)} = \frac{5}{9}
\]

which I can write as \(((4/9, 5/9), (5/9, 4/9))\); where \((4/9, 5/9)\), for example, means the mixed strategy in which player 1 plays \(A\) with probability \(4/9\) and \(B\) with probability \(5/9\). □

Example 6 (NE in Rock-Paper-Scissors). In Rock-Scissor-Paper (Section 2), the NE is unique and in it each player randomizes \(1/3\) each across all three strategies. That is, the unique mixed strategy NE profile is

\(((1/3, 1/3, 1/3), (1/3, 1/3, 1/3))\).

It is easy to verify that if either player randomizes \(1/3\) across all strategies, then the other payer is indifferent across all three strategies. □

Example 7. Although every pure strategy that gets positive probability in a NE is a best response, the converse is not true: in a NE, there may be best responses that are not given positive probability. A trivial example where this occurs is a game where the payoffs are constant, independent of the strategy profile. Then players are always indifferent and every strategy profile is a NE.

4.4 More NE examples.

Example 8. Figure 7 provides the game box for a game called Matching Pennies, which is simpler than Rock-Paper-Scissors but in the same spirit. This game has

\[
\begin{array}{c|cc}
H & T \\
\hline
H & 1, -1 & -1, 1 \\
T & -1, 1 & 1, -1 \\
\end{array}
\]

Figure 7: Matching Pennies.

one equilibrium and in this equilibrium both players randomize 50:50 between the two pure strategies. □

Example 9. Figure 8 provides the game box for one of the best known games in Game Theory, the Prisoner’s Dilemma (PD) (the PD is actually a set of similar games; this is one example). PD is a stylized depiction of friction between joint incentives (the sum of payoffs is maximized by \((C, C)\)) and individual incentives (each player has incentive to play \(D\)).
It is easy to verify that $D$ is a best response no matter what the opponent does. The unique NE is thus $(D, D)$.

A PD-like game that shows up fairly frequently in popular culture goes something like this. A police officer tells a suspect, “We have your partner, and he’s already confessed. It will go easier for you if you confess as well.” Although there is usually not enough detail to be sure, the implication is that it is (individually) payoff maximizing not to confess if the partner also does not confess; it is only if the partner confesses that it is payoff maximizing to confess. If that is the case, the game is not a PD but something like the game in Figure 9, where I have written payoffs in a way that makes them easier to interpret as prison sentences (which are bad).

\[
\begin{array}{cc}
\text{Silent} & \text{Confess} \\
0, 0 & -4, -2 \\
-2, -4 & -3, -3 \\
\end{array}
\]

Figure 9: A non-PD game.

This game has three NE: (Silent, Silent), (Confess, Confess), and a mixed NE in which the suspect stays silent with probability $1/3$. In contrast, in the (true) PD, each player has incentive to play $D$ regardless of what he thinks the other player will do. □

Example 10 (Cournot Duopoly). The earliest application of game theoretic reasoning that I know of to an economic problem is the quantity-setting duopoly model of Augustin Cournot, in Cournot (1838).

There are two players (firms). For each firm $S_i = \mathbb{R}_+$. Note that this is not a finite game; one can work with a finite version of the game but analysis of the continuum game is easier in many respects. For the pure strategy profile $(q_1, q_2)$ ($q_i$ for “quantity”), the payoff to player $i$ is,

\[(1 - q_1 - q_2)q_i - k,\]

where $k > 0$ is a small number (in particular, $k < 1/9$), provided $q_i > 0$ and $q_1 + q_2 \leq 1$. If $q_i = 0$ then the payoff to firm $i$ is 0. If $q_i > 0$ and $q_1 + q_2 > 1$ then the payoff to firm $i$ is $-k$.

The idea here is that both firms produce a quantity $q_i$ of a good, best thought of as perishable, and then the market clears somehow, so that each receives a price
1 − q_1 − q_2, provided the firms don’t flood the market (q_1 + q_2 > 1); if they flood the market, then the price is zero. I have assumed that the firm’s payoff is its profit. There is a fixed cost (k, for ease of notation taken to be the same for both firms) but no marginal costs.

Ignore k for a moment. It is easy to verify that for any probability distribution over q_2, provided E[q_2] < 1, the best response for firm 1 is the pure strategy,

\[
\frac{1 - E[q_2]}{2}.
\]

If E[q_2] ≥ 1, then the market is flooded, in which case any q_1 generates the same profit, namely 0. Because of this, if k = 0, then there are an infinite number of “nuisance” NE in which both firms flood the market by producing at least 1 in expectation. This is the technical reason for including k > 0 (which seems realistic in any event): for k > 0, if E[q_2] ≥ 1, then the best response is to shut down (q_1 = 0). Adding a small marginal cost of production would have a similar effect.

As an aside, this illustrates an issue that sometimes arises in economic modeling: sometimes one can simplify the analysis, at least in some respects, by making the model more complicated.\(^3\)

With k ∈ (0, 1/9), the NE is unique and in it both firms produce 1/3. One can find this by noting that, in pure strategy equilibrium, q_1 = (1 − q_2)/2 and q_2 = (1 − q_1)/2; this system of equations has a unique symmetric solution with q_1 = q_2 = q = (1 − q)/2, which implies q = 1/3.

Total output in NE is thus 2/3 (and price is 1/3). The NE of the Cournot quantity game is intermediate between monopoly (where quantity is 1/2) and efficiency (in the sense of maximizing total surplus; since there is zero marginal cost, efficiency calls for total output of 1). □

Example 11 (Bertrand Duopoly). Writing in the 1880s, about 50 years after Cournot, Joseph Bertrand objected to Cournot’s quantity-setting model of duopoly on the grounds that quantity competition is unrealistic. In Bertrand’s model, firms set prices rather than quantities.

Assume (to make things simple) that there are no production costs, either marginal or fixed. I allow p_i, the price charged by firm i, to be any number in \(\mathbb{R}_+\), so, once again, this is not a finite game.

• p_i ≥ 1 then firm i gets a payoff of 0, regardless of what the other firm does. Assume henceforth that p_i ≤ 1 for both firms.

• If p_1 ∈ (0, p_2) then firm 1 receives a payoff of (1 − p_1)p_1. In words, firm 1 gets the entire demand at p_1, which is 1 − p_1. Revenue, and hence profit, is (1 − p_1)p_1. Implicit here is the idea that the good of the two firms are

\(^3\)If E[q_2] is not defined, then player 1’s best response does not exist; by definition, this cannot happen in NE, so I ignore this case.
homogeneous and hence perfect substitutes: consumers buy whichever good is cheaper; they are not interested in the name of the seller.

- If $p_1 = p_2 = p$, both firms receive a payoff of $(1 - p)p/2$. In words, the firms split the market at price $p$.
- If $p_1 > p_2$, then firm 1 gets a payoff of 0.

Payoffs for firm 2 are defined symmetrically. It is not hard to see that firm 1’s best response is as follows.

- $p_2 > 1/2$ then firm 1’s best response is $p_1 = 1/2$.
- If $p_2 \in (0, 1/2]$ then firm 1’s best response set is empty. Intuitively, firm 1 wants to undercut $p_2$ by some small amount, say $\varepsilon > 0$, but $\varepsilon > 0$ can be made arbitrarily small. Technically, the existence failure for best response arises because of a discontinuity in the payoff function. One can address the existence failure by requiring that prices lie on a finite grid (be expressed in pennies, for example), but this introduces other issues.
- If $p_2 = 0$, then any $p_1 \in \mathbb{R}_+$ is a best response.

Notwithstanding the fact that best response isn’t defined over much of the domain, there is a unique pure NE (and in fact the unique NE period), namely the price profile $(0, 0)$.

The Bertrand model thus predicts the competitive outcome with only two firms. At first this may seem paradoxical, but remember that I am assuming that the goods are homogeneous. The bleak prediction of the Bertrand game (bleak from the viewpoint of the firms; consumers do well and the outcome is efficient) suggests that firms have strong incentive to differentiate their products, and also strong incentive to collude (which may be possible if the game is repeated, as I discuss later in the course).

Example 12. Francis Edgeworth, writing in the 1890s, about ten years after Bertrand, argued that firms often face capacity constraints, and these constraints can affect equilibrium analysis. Continuing with the Bertrand duopoly of Example 11, suppose that firms can produce up to 1/2 at zero cost but then face infinite cost to produce any additional output. This would model a situation in which, for example, each firm had stockpiled a perishable good; this period, a firm can sell up to the amount stockpiled, but no more, at zero opportunity cost (zero since any amount not sold melts away in the night).

Suppose that sales of firm 1 (taking into consideration the capacity constraint
and ignoring the trivial case with \( p_1 > 1 \) are given by,

\[
q_1(p_1, p_2) = \begin{cases} 
0 & \text{if } p_1 > p_2 \text{ and } p_1 > 1/2, \\
1/2 - p_1 & \text{if } p_1 > p_2 \text{ and } p_1 \leq 1/2, \\
(1 - p_1)/2 & \text{if } p_1 = p_2 \leq 1, \\
1 - p_1 & \text{if } p_1 < p_2 \text{ and } p_1 > 1/2, \\
1/2 & \text{if } p_1 < p_2 \text{ and } p_1 \leq 1/2.
\end{cases}
\]

The sales function for firm 2 is analogous.

For interpretation, suppose that there are a continuum of consumers, with a consumer whose “name” is \( \ell \in [0, 1] \) wanting to buy at most one unit of the good at a price of at most \( v_\ell = 1 - \ell \). Then, integrating across consumers, market demand facing a monopolist would be \( 1 - p \), as specified.

With this interpretation, consider the case \( p_1 > p_2 \) and \( p_1 \leq 1/2 \). Then \( p_2 < 1/2 \) and demand for firm 2 is \( 1 - p_2 > 1/2 \). But because of firm 2’s capacity constraint, it can only sell \( 1/2 \), leaving \( (1 - p_2) - 1/2 > 0 \) potentially for firm 1, provided firm 1’s price is not too high. Suppose that the customers who succeed at buying from firm 2 are the customers with the highest valuations (they are the customers willing to wait on line, for example). Then the left over demand facing firm 1, called residual demand, is \( 1/2 - p_1 \).

Note that this specification of residual demand is a substantive assumption about how demand to the low priced firm gets rationed. We could instead have assumed that customers are assigned to firm 2 randomly. That would have resulted in a different model.

In contrast to the Bertrand model, it is no longer an equilibrium for both firms to charge \( p_1 = p_2 = 0 \). In particular, if \( p_2 = 0 \), then firm 1 is better off setting \( p_1 = 1/4 \), which is the monopoly price for residual demand. In fact, there is no pure strategy equilibrium here. There is, however, a symmetric mixed strategy equilibrium, which can be found as follows.

Note first that if the maximum profit from residual demand is \( 1/16 = (1/4 \times 1/4) \). A firm can also get a profit of \( 1/16 \) by being the low priced firm at a price of \( 1/8 \) (and selling \( 1/2 \)). The best response for firm 1 to pure strategies of firm 2 is then,

\[
\begin{array}{ll}
1/2 & \text{if } p_2 > 1/2, \\
\emptyset & \text{if } p_2 \in (1/8, 1/2], \\
1/4 & \text{if } p_2 \leq 1/8.
\end{array}
\]

The best response for firm 2 to pure strategies of player 1 is analogous.

Intuitively, the mixed strategy equilibrium will involve a continuous distribution, and not put positive probability on any particular price, because of the price undercutting argument that is central to the Bertrand game. So, I look for a symmetric equilibrium in which firms randomize over an interval of prices. The obvious interval to look at is \( [1/8, 1/4] \).
For it to be an equilibrium for firm 1 to randomize over the interval \([1/8, 1/4]\), firm 1 has to get the same profit from any price in this interval.\(^4\) The profit at either \(p_1 = 1/8\) (where the firm gets capacity constrained sales of 1/2) or \(p_1 = 1/4\) (where the firm sells 1/2-1/4=1/4 to residual demand) is 1/16. Thus for any other \(p_1 \in (1/8, 1/4)\) it must be that, letting \(F_2(p_1)\) be the probability that player 2 charges a price less than or equal to \(p_1\),

\[
F_2(p_1)p_1(1/2 - p_1) + (1 - F_2(p_1))p_1(1/2) = 1/16,
\]

where the first term is profit when \(p_2 < p_1\) and the second term is profit when \(p_2 > p_1\); because the distribution is continuous, the probability that \(p_2 = p_1\) is zero and so I can ignore this case.

Solving for \(F_2\) yields

\[
F_2(p_1) = \frac{1}{2p_1} - \frac{1}{16p_1^2}.
\]

The symmetric equilibrium then has \(F_1 = F_2\), where \(F_2\) is given by the above expression. The associated density is

\[
f_2(p_1) = \frac{1}{8p_1^3} - \frac{1}{2p_1^2}
\]

and is graphed in Figure 12. This makes some intuitive sense. If we imagine this equilibrium being repeated, then most of the time prices will be close to 1/8, reflecting the undercutting logic of the Bertrand game, but from time to time one of the prices will be close to 1/4.\(^4\)

\(^4\)In a continuum game like this, this requirement can be relaxed somewhat, for measure theoretic reasons, but this is a technicality as far as this example is concerned.

Figure 10: Density for the NE in the Bertrand-Edgeworth Game example.
Edgeworth argued that in a situation such as this, since there is no pure strategy equilibrium, prices would just bounce around. Nash equilibrium makes a more specific prediction: the distribution of prices is generated by the mixed strategy equilibrium.

There is an intuition, first explored formally in Kreps and Scheinkman (1983), that the qualitative predictions of Cournot’s quantity model can be recovered in a more realistic, multi-period setting in which firms make output choices in period 1 and then sell that output in period 2. The Edgeworth example that I just worked through is an example of what the period 2 analysis can look like. If the capacity constraints had been, instead, 1/3 each (the equilibrium quantities in the Cournot game) then you can verify that the equilibrium of the Bertrand-Edgeworth game would, in fact, have been for both firms to charge a price of 1/3. □

4.5 The MinMax Theorem.

Recall from Section 2 that a two-player game is zero-sum iff \( u_2 = -u_1 \). A game is constant-sum iff there is a \( c \in \mathbb{R} \) such that \( u_1 + u_2 = c \).

Theorem 4 below, called the MinMax Theorem (or, more frequently, the Min-iMax Theorem), was one of the first important general results in game theory. It was established for zero-sum games in von Neumann (1928). Although it is rare to come across zero-sum, or constant-sum, games in economic applications, the MinMax Theorem is useful because it is sometimes possible to provide a quick proof of an otherwise difficult result by constructing an auxiliary constant-sum game and then applying the MinMax Theorem to this new game; an example of this is given by the proof of Theorem 6 in Section 4.8.

The MinMax Theorem is an almost immediate corollary of the following result.

Theorem 3. In a finite constant-sum game, if \( \sigma^* = (\sigma_1^*, \sigma_2^*) \) is a NE then,

\[
    u_1(\sigma^*) = \max_{\sigma_1} \min_{\sigma_2} u_1(\sigma) = \min_{\sigma_1} \max_{\sigma_2} u_1(\sigma).
\]

Proof. Suppose that \( u_1 + u_2 = c \). Let \( V = u_1(\sigma^*) \). Since \( \sigma^* \) is a NE, and since, for player 2, maximizing \( u_2 = c - u_1 \) is equivalent to minimizing \( u_1 \),

\[
    V = \max_{\sigma_1} u_1(\sigma_1, \sigma_2^*) = \min_{\sigma_2} u_1(\sigma_1^*, \sigma_2).
\]

Then,

\[
    \min_{\sigma_2} \max_{\sigma_1} u_1(\sigma) \leq \max_{\sigma_1} u_1(\sigma_1, \sigma_2^*) = \min_{\sigma_2} u_1(\sigma_1^*, \sigma_2) \leq \max_{\sigma_1} \min_{\sigma_2} u_1(\sigma) \tag{2}
\]

On the other hand, for any \( \tilde{\sigma}_2 \),

\[
    u_1(\sigma_1^*, \tilde{\sigma}_2) \leq \max_{\sigma_1} u_1(\sigma_1, \tilde{\sigma}_2),
\]

\footnote{Yes, von Neumann (1928) is in German; no, I haven’t read it in the original.}
hence,
\[
\min_{\sigma_2} u_1(\sigma_1^*, \sigma_2) \leq \max_{\sigma_1} u_1(\sigma_1, \tilde{\sigma}_2).
\]
Since the latter holds for every \(\tilde{\sigma}_2\),
\[
\min_{\sigma_2} u_1(\sigma_1^*, \sigma_2) \leq \min_{\sigma_1} \max_{\sigma_2} u_1(\sigma_1, \sigma_2).
\]
Combining with expressions (1) and (2),
\[
V = \min_{\sigma_2} u_1(\sigma_1^*, \sigma_2) = \min_{\sigma_1} \max_{\sigma_2} u_1(\sigma).
\]
By an analogous argument,
\[
V = \max_{\sigma_1} u_1(\sigma_1, \sigma_2^*) = \max_{\sigma_1} \min_{\sigma_2} u_1(\sigma).
\]
And the result follows. ■

**Theorem 4** (MinMax Theorem). *In any finite constant-sum game, there is a number \(V\) such that*
\[
V = \max_{\sigma_1} \min_{\sigma_2} u_1(\sigma_1, \sigma_2) = \min_{\sigma_1} \max_{\sigma_2} u_1(\sigma).
\]

*Proof. By Theorem 1, there is NE, \(\sigma^*\). Set \(V = u_1(\sigma^*)\) and apply Theorem 3. ■*

**Remark 4.** The MinMax Theorem can be proved using a separation argument, rather than a fixed point argument (the proof above implicitly uses a fixed point argument, since it invokes Theorem 1, which uses a fixed point argument). □

By way of interpretation, one can think of \(\max_{\sigma_1} \min_{\sigma_2} u_1(\sigma)\) as the payoff that player 1 can guarantee for herself, no matter what player 2 does. Theorem 3 implies that, in a constant-sum game, player 1 can guarantee an expected payoff equal to her NE payoff. This close connection between NE and what players can guarantee for themselves does not generalize to non-constant-sum games, as the following example illustrates.

**Example 13.** Consider the game in Figure 11, which originally appeared in Aumann and Maschler (1972). This game is not constant sum. In this game, there is a

\[
\begin{array}{cc}
L & R \\
T & 1,0 & 0,1 \\
B & 0,3 & 1,0 \\
\end{array}
\]

**Figure 11: The Aumann-Maschler Game.**

unique NE, with \(\sigma_1^* = (3/4, 1/4)\) and \(\sigma_2^* = (1/2, 1/2)\) and the equilibrium payoffs are \(u_1(\sigma^*) = 1/2\) and \(u_2(\sigma^*) = 3/4.\)
But in this game, $\sigma_1^*$ does not guarantee player 1 a payoff of 1/2. In fact, if player 1 plays $\sigma_1^*$ then player 2 can hold player 1’s payoff to 1/4 by playing $\sigma_2 = (0, 1)$ (i.e., playing $R$ for certain); it is not payoff maximizing for player 2 to do this, but it is feasible. Player 1 can guarantee a payoff of 1/2, but doing so requires randomizing $\sigma_1 = (1/2, 1/2)$, which is not an strategy that appears in any Nash equilibrium of this game. □

4.6 Never a Best Response and Rationalizability.

**Definition 2.** Fix a game. A pure strategy $s_i$ is never a best response (NBR) iff there does not exist a profile of opposing mixed strategies $\sigma_{-i}$ such that $s_i \in \text{BR}_i(\sigma_{-i})$.

If $s_i$ is NBR then it cannot be part of any NE. Therefore, it can be, in effect, deleted from the game when searching for NE. This motivates the following procedure.

Let $S^1_i \subseteq S_i$ denote the strategy set for player $i$ after deleting all strategies that are never a best response. $S^1_i$ is non-empty since, as noted in Section 4.1, the best response correspondence is not empty-valued. The $S^1_i$ form a new game, and for that game we can ask whether any strategies are NBR. Deleting those strategies, we get $S^2_i \subseteq S^1_i \subseteq S_i$, which again is not empty. And so on.

Since the game is finite, this process terminates after a finite number of rounds in the sense that there is a $t$ such that for all $i$, $S^{t+1}_i = S^t_i$. Let $S^R_i$ denote this terminal $S^t_i$ and let $S^R$ be the product of the $S^R_i$. The strategies $S^R_i$ are called the rationalizable strategies for player $i$; $S^R$ is the set of rationalizable (pure) strategy profiles. Informally, these are the strategies that make sense after introspective reasoning of the form, “$s_i$ would maximize my expected payoffs if my opponents were to play $\sigma_{-i}$, and $\sigma_{-i}$ would maximize my opponents’ payoffs (for each opponent, independently) if they thought . . . .” Rationalizability was first introduced into game theory in Bernheim (1984) and Pearce (1984).

**Theorem 5.** Fix a game. If $\sigma$ is a NE and $\sigma_i[s_i] > 0$ then $s_i \in S^R_i$.

**Proof.** Almost immediate by induction and Theorem 2. ■

**Remark 5.** The construction of $S^R$ uses maximal deletion of never a best response strategies at each round of deletion. In fact, this does not matter: it is not hard to show that if, whenever possible, one deletes at least one strategy for at least one player at each round, then eventually the deletion process terminates, and the terminal $S^t$ is $S^R$. □

**Example 14.** In the Prisoner’s Dilemma (Example 9, Figure 8), $C$ is NBR. $D$ is the unique rationalizable strategy. □
Example 15 (Rationalizability in the Cournot Duopoly). Recall the Cournot Duopoly (Example 10).

With \( k > 0 \), no pure strategy greater than \( 1/2 \) is a best response. Therefore, in the first round of deletion, delete all strategies in \( (1/2, \infty) \); we can actually delete more, as discussed below, but we can certainly delete this much. In the second round of deletion, one can delete all strategies in \( [0, 1/4) \). And so on. This generates a nested sequence of compact intervals. There is no termination in finite time, as for a finite game, but one can show that the intersection of all these intervals is the point \( \{1/3\} \): \( S_i^R = \{1/3\} \) for both firms. And, indeed, the NE is, as we already saw, \( (1/3, 1/3) \).

In fact, because of \( k > 0 \), firm 1 strictly prefers to shut down if \( \mathbb{E}[q_2] \) is too large; more precisely, at the first round of deletion, firm 1 prefers to shut down if \( q_2 > 1 - 2\sqrt{k} \). So we can also eliminate all \( q_1 \) in \( (0, \sqrt{k}) \), even at the first round. But note that this complication does not undermine the argument above; it only states that we could have eliminated more strategies at each round. □

A natural, but wrong, intuition is that any strategy in \( S_i^R \) gets positive probability in at least one NE.

Example 16. Consider the game in Figure 12. \( S_i^R = \{A, B, C\} \) for either player.

\[
\begin{array}{ccc}
A & B & C \\
A & 4, 4 & 1, 1 & 1, 1 \\
B & 1, 1 & 2, -2 & -2, 2 \\
C & 1, 1 & -2, 2 & 2, -2 \\
\end{array}
\]

Figure 12: Rationalizability and NE.

But the unique NE is \((A, A)\). □

4.7 Correlated Rationalizability.

Somewhat abusing notation, let \( \text{BR}_i(\sigma_{-i}) \) denote the set of strategies for player \( i \) that are best responses to the distribution \( \sigma_{-i} \in \Sigma_{-i} = \Delta(S_{-i}) \).

Definition 3. Fix a game. A pure strategy \( s_i \) is correlated never a best response (CNBR) iff there does not exist a distribution over opposing strategies \( \sigma_{-i} \) such that \( s_i \in \text{BR}_i(\sigma_{-i}) \).

Again, by Theorem 2, if \( s_i \) is never a best response, then so is any mixed strategy that gives \( s_i \) positive probability.

The set of CNBR strategies is a subset of the set of NBR strategies. In particular, any CNBR strategy is NBR (since a strategy that is not a best response to any distribution is, in particular, not a best response to any independent distribution). If \( N = 2 \) the CNBR and NBR sets are equal (trivially, since in this case there is
only one opponent). If \( N \geq 3 \), the CNBR set can be a proper subset of the NBR set. In particular, one can construct examples in which a strategy is NBR but not CNBR (see, for instance Fudenberg and Tirole (1991)). The basic intuition is that it is easier for a strategy to be a best response to something the larger the set of distributions over opposing strategies, and the set of correlated distributions is larger than the set of independent distributions.

If one iteratively deletes CNBR strategies, one gets, for each player \( i \), the set \( S_i^{CR} \) of correlated rationalizable strategies. Since, at each step of the iterative deletion, the set of CNBR is a subset of the set of NBR strategies, \( S_i^{R} \subseteq S_i^{CR} \) for all \( i \), with equality if \( N = 2 \). Let \( S^{CR} \) denote the set of correlated rationalizable profiles. The set of distributions over strategies generated by Nash equilibria is a subset of the set of distributions over \( S^{R} \), which in turn is a subset of the set of distributions over \( S^{CR} \).

4.8  Strict Dominance.

**Definition 4.** Fix a game.

- Mixed strategy \( \sigma_i \) strictly dominates mixed strategy \( \hat{\sigma}_i \) iff for any profile of opposing pure strategies \( s_{-i} \),
  \[ u_i(\sigma_i, s_{-i}) > u_i(\hat{\sigma}_i, s_{-i}). \]

- Mixed strategy \( \hat{\sigma}_i \) is strictly dominated iff there exists a mixed strategy that strictly dominates it.

If a mixed strategy \( \sigma_i \) strictly dominates every other strategy, pure or mixed, of player \( i \) then it is called strictly dominant. It is easy to show that a strictly dominant strategy must, in fact, be pure.

**Definition 5.** Fix a game. Pure strategy \( s_i \) is strictly dominant iff it strictly dominates every other strategy.

**Example 17.** Recall the Prisoner’s Dilemma of Example 9 and Figure 8. Then, for either player, \( C \) is strictly dominated by \( D \); \( D \) is strictly dominant. □

**Theorem 6.** Fix a game. For each player \( i \), a strategy is strictly dominated iff it is CNBR.

**Proof.** \( \Rightarrow \). If \( \hat{\sigma}_i \) strictly dominates \( \sigma_i \) then \( \hat{\sigma}_i \) has strictly higher payoff than \( \sigma_i \) against any distribution over opposing strategies (since expected payoffs a are a convex combination of payoffs against pure strategy profiles).

\( \Leftarrow \). Suppose that \( \hat{\sigma}_i \) is CNBR. Construct a zero-sum game in which the two players are \( i \) and \( -i \), the strategy set for \( i \) is \( S_i \) (unless \( \hat{\sigma}_i \) happens to be pure, in which case omit that), and the strategy set for \( -i \) is \( S_{-i} \). In this new game, a mixed
strategy for player $-i$ is a probability distribution over $S_{-i}$, which (if there are two or more opponents) is a correlated strategy $\sigma^c_{-i}$ in the original game. In the new game, the payoff to player $i$ is,

$$\tilde{u}_i(s) = u_i(s_i, s_{-i}) - u_i(\tilde{\sigma}_i, s_{-i}).$$

and the payoff to $-i$ is $\tilde{u}_{-i} = -\tilde{u}_i$.

Let

$$V = \min_{\sigma^c_{-i}} \max_{\sigma_i} \tilde{u}_i(\sigma_i, \sigma^c_{-i}).$$

Since $\sigma_i$ is CNBR, it follows that for every $\sigma^c_{-i}$ there is a $\sigma_i$ for which $\tilde{u}_i(\sigma_i, \sigma^c_{-i}) > 0$. Therefore $V > 0$. By Theorem 4,

$$V = \max_{\sigma_i} \min_{\sigma^c_{-i}} \tilde{u}_i(\sigma_i, \sigma^c_{-i}).$$

This implies that there exists a $\sigma^*_i$ (in fact, a NE strategy for player $i$ in the new game) such that for all $s_{-i} \in S_{-i}$, $\tilde{u}_i(\sigma^*_i, s_{-i})$, hence

$$u_i(\sigma^*_i, s_{-i}) > u_i(\tilde{\sigma}_i, s_{-i}),$$

so that $\sigma^*_i$ strictly dominates $\tilde{\sigma}_i$. ■

In the following examples, $N = 2$, so the CNBR and NBR strategy sets are equal.

**Example 18.** Consider the game in Figure 13 (I’ve written payoffs only for player 1). $T$ is the best response if the probability of $L$ is more than 1/2. $M$ is the best response if the probability of $L$ is less than 1/2. If the probability of $L$ is exactly 1/2 the the set of best responses is the set of mixtures over $T$ and $M$. Hence $B$ is NBR.

$B$ is strictly dominated for player 1 by the mixed strategy $(1/2, 1/2, 0)$ (which gets a payoff of 5 no matter what player 2 does, whereas $B$ always gets 4). No strategy is strictly dominant.

Note that $B$ is not strictly dominated by any pure strategy. So it is important in this example that I allow for dominance by a mixed strategy. □
Consider the game in Figure 14. $T$ is the best response if the probability of $L$ is more than $3/5$ and $M$ is the best response if the probability of $L$ is less than $2/5$. If the probability of $L$ is in $(2/5, 3/5)$ then $B$ is the best response. If the probability of $L$ is $3/5$ then the set of best responses is the set of mixtures over $T$ and $B$. If the probability of $L$ is $2/5$ then the set of best responses is the set of mixtures over $M$ and $B$.

Note that $B$ is not a best response to any pure strategy for player 2. So it is important in this example that I allow for mixed strategies by player 2. $B$ is not strictly dominated for player 1. No strategy is strictly dominant. □

It follows from Theorem 6 that one can find the correlated rationalizable strategy sets, $S^{CR}_i$, by iteratively deleting strictly dominated strategies. Once again, this may help in narrowing down the search for Nash equilibria.

4.9 How many NE are there?

Fix players and strategy sets and consider all possible assignments of payoffs. With $N$ players and $|S|$ pure strategy profiles, payoffs can be represented as a point in $\mathbb{R}^{N|S|}$. Wilson (1971) proved that there is a precise sense in which the set of Nash equilibria is finite and odd (which implies existence, since 0 is not an odd number) for most payoffs in $\mathbb{R}^{N|S|}$. As we have already seen, Rock-Paper-Scissors has one equilibrium, while Battle of the Sexes has three. A practical implication of the finite and odd result, and the main reason why I am emphasizing this topic in the first place, is that if you are trying to find all the equilibria of the game, and so far you have found four, then you are probably missing at least one.

Although the set of NE in finite games is typically finite and odd, there are exceptions.

Example 20 (A game with two NE). A somewhat pathological example is the game in Figure 15. This game has exactly two Nash equilibria, $(A, A)$ and $(B, B)$. In

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1,1</td>
<td>0,0</td>
</tr>
<tr>
<td>B</td>
<td>0,0</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Figure 15: A game with two Nash equilibria.
particular, there are no mixed strategy NE: if either player puts any weight on $A$ than the other player wants to play $A$ for sure. □

The previous example seems somewhat artificial, and in fact it is extremely unusual in applications to encounter a game where the set of NE is finite and even. But it is quite common to encounter strategic form games that have an infinite number of NE; such games are common when the strategic form represents an extensive form game with a non-trivial temporal structure.

**Example 21 (An Entry Deterrence Game).** Consider the game in Figure 16. For motivation, player 1 is considering entering a new market (staying in or going out). Player 2 is an incumbent who, if there is entry, can either accommodate (reach some sort of status quo intermediate between competition and monopoly) or fight (start a price war).

There are two pure strategy NE, $(In, A)$ and $(O, F)$. In addition, there are a continuum of NE of the form, player 1 plays $O$ and player 2 accommodates, should entry nevertheless occur, with probability $q \leq 1/16$. More formally, every profile of the form $(O, (q, 1-q))$ is a NE for any $q \in [0, 1/16]$. Informally, the statement is that it is a NE for the entrant to stay out provided that the incumbent threatens to fight with high enough probability (accommodate with low enough probability).

It is standard in game theory to argue that the only reasonable NE of this game is the pure strategy NE $(In, A)$. One justification for this claim is that if player 1 puts any weight at all on $In$ then player 2’s best response is to play $A$ for sure, in which case player 1’s best response is likewise to play $In$ for sure. The formal version of this statement is to say that only $(In, A)$ is an admissible NE, because the other NE put positive weight on a weakly dominated strategy, namely $F$; I discuss weak dominance later in the course. For the moment, what I want to stress is that every one of this infinity of NE is a NE, whether we find it reasonable or not. □

Assuming that the number of equilibria is in fact finite and odd, how many equilibria are typical? In $2 \times 2$ games, the answer is 1 or 3. What about in general? A lower bound on how large the set of NE can possibly be in finite games is provided by $L \times L$ games (two players, each with $L$ strategies) in which, in the game box representation, payoffs are $(1, 1)$ along the diagonal and $(0, 0)$ elsewhere. In such games, there are $L$ pure strategy Nash equilibria, corresponding to play along the diagonal, and an additional $2^L - (L + 1)$ fully or partly mixed NE. The total number of NE is thus $2^L - 1$. This example is robust; payoffs can be perturbed slightly and there will still be $2^L - 1$ NE.
This is extremely bad news. First, it means that the maximum number of NE is growing exponentially in the size of the game. This establishes that the general problem of calculating all of the equilibria is computationally intractable. Second, it suggests that the problem of finding even one NE may, in general, be computationally intractable; for work on this, see Daskalakis, Goldberg and Papdimitriou (2008). Note that the issue is not whether algorithms exist for finding one equilibrium, or even all equilibria. For finite games, there exist many such algorithms. The problem is that the time taken by these algorithms to reach a solution can grow explosively in the size of the game. For results on the number of equilibria in finite games, see McLennan (2005).

4.10 The structure of NE.

As in Section 4.9, if players and strategy sets are fixed, then payoff functions can be represented as a vector in $\mathbb{R}^{|S|}$, giving payoffs for each player and each strategy profile. Let

$$\mathcal{N} : \mathbb{R}^{|S|} \to \mathcal{P}(\Sigma)$$

be the NE correspondence: for any specification of payoffs $u \in \mathbb{R}^{|S|}$, $\mathcal{N}(u)$ is the set of NE for the game defined by $u$. By Theorem 1, $\mathcal{N}$ is non-empty-valued.

The following result states that the limit of a sequence of NE is a NE (in particular, the set of NE for a fixed $u$ is closed) but that for some games there are NE for which there are no nearby NE in some nearby games.

**Theorem 7.** $\mathcal{N}$ is upper hemicontinuous but (for $|S| \geq 2$) not lower hemicontinuous.

**Proof.**

1. **Upper Hemicontinuity.** Since $\Sigma$ is compact, it suffices to prove that $\mathcal{N}$ has a closed graph, for which it suffices to prove that every point $(u, \sigma)$ in the complement of graph($\mathcal{N}$) is interior. Take any $(u, \sigma)$ in the complement of graph($\mathcal{N}$). Then there is an $i$ and a pure strategy $s_i$ for which

$$u_i(s_i, \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}) > 0.$$

By continuity of expected utility in both $u_i$ and $\sigma$, this inequality holds for all points within a sufficiently small open ball around $(u, \sigma)$, which completes the argument.

2. **Lower Hemicontinuity.** Consider the game in figure 17. For any $\varepsilon > 0$, the unique NE is $(A, L)$. But if $\varepsilon = 0$ then any mixed strategy profile is a NE, and in particular $(A, R)$ is a NE. Therefore, for any sequence of $\varepsilon > 0$ converging to zero, there is no sequence of NE converging to $(A, R)$. This example implies a failure of lower hemicontinuity in any game with $|S| \geq 2$. 

26
The set of NE need not be connected, let alone convex. But one can prove that for any finite game, the set of NE can be decomposed into a finite number of connected components; see Kohlberg and Mertens (1986). In the Battle of the Sexes (Example 5 in Section 4.3), there are three components, namely the three NE. In the entry deterrence game of Example 21 (Section 4.9), there are two components: the pure strategy (In, A) and the component of (O, (q, 1 – q)) for q ∈ [0, 1/16].

References


Luce, R. Duncan and Howard Raiffa. 1957. *Games and Decisions*. Wiley.


