14

Multi-Stage Games

Recall that the normal for game had players choose their actions simultaneously, and ignored any real issues concerning sequential behavior. The extensive form games we have analyzed included the ability to analyze sequential behavior and conditional choices. In the extensive forms we have seen, players played sequentially, with actions and information unfolding, but the payoffs were delayed until the game reached a terminal node that was associated with the game ending.

In reality, dynamic play over time may be more complex than one game that unfolds over time. Instead, players can play one game that is followed by another, or maybe even several other games. Should we treat each game independently, or should we expect players to consider the sequence of different games as one "large" game?

For example, imagine the following twist on the familiar Prisoner's Dilemma. Suppose that first, two players play the familiar Prisoner's Dilemma with pure actions mum (M) and fink (F) with the payoff matrix, Player 2

		M	F
Player 1	M	$4, 4$	$-1, 5$
	F	$5, -1$	$1, 1$

stage 1

Now imagine that after this game is over, some time passes, and we are at a second point in time in which the same two players play the following *Revenge Game.* Each player can choose whether to join a gang (G) or to remain a "loner" (L). If they remain loners then they go their separate ways, regardless of the outcome of the first game, and obtain second period payoffs of 0 each. If both join gangs, then they will fight each other and suffer a loss, causing each to receive a payoff of -3. Finally, if player i joins a gang while player j remains a loner, then the loner suffers dearly and receives a payoff of -4, while the new gang member suffers much less and receives a payoff of -1. That is, the Revenge game is given by the following matrix,

Player 2

		L	G
Player 1	L	$0, 0$	$-4, -1$
	G	$-1, -4$	$-3, -3$

stage 2

From our previous analysis of static games of complete information, we know how to analyze these two games independently. In the first stage game, the Prisoner's dilemma, there is a unique Nash equilibrium, (F, F). In the second stage game, the revenge game, there are two pure strategy equilibria, (L, L) and (G, G), and a mixed strategy equilibrium where each player plays L with probability $\frac{1}{2}$ (verify this!).

Some natural questions follow: if players are rational and forward looking, should they not view these two games as one "grand" game? If they do, should we expect that their actions in the second stage game, namely the revenge game, will depend on the outcome of the first stage Prisoner's dilemma game that preceded it? However, if there is a unique Nash equilibrium in the fist stage, are the players not doomed to play it?

As we will soon see, by allowing players to condition future behavior on past outcomes, we may provide them with credible incentives to follow certain plays in earlier stage games that would not be possible if there was not a continuation game that followed. That is, we will soon see that in this grand game we described above, players may be able to avoid the bleak outcome of (F, F) in the first stage of interaction.

14.1 Set-up

To generalize from this simple example, a *multi-stage game* is a finite sequence of *stage-games,* each one being a game of complete but imperfect information (a simultaneous move game). These games are *played sequentially* by the *same players*, and the *total payoffs* from the sequence of games will be evaluated using the sequence of outcomes in the games that were played. We adopt the convention that each game is played in a *distinct period*, so that game 1 is played in period 1, game 2 in period 2, and so on. We will also assume that after each stage is completed, all the players observe the outcome of that stage, and that this information structure is common knowledge.[1]

A sequence of normal form games of complete information with the same players is not too hard to conceive. One example is the stylized game described above where first players play a Prisoner's dilemma, and then they play the Revenge game. Another can be several politicians who form a committee, and every well defined period they engage in strategic bargaining over legislative outcomes, the result of which affects their electorate, and in turn, affects their own utility. In each game, players have a set of actions they can choose, and the profiles of actions lead to payoffs for *that specific game*, which is then followed by another game, and another, until the sequence of games is over.

Once we realize that one game follows another, this implies that players can observe the outcomes of each game before another game is played. This observation

[1] It is conceivable to imagine games where some stages are played before the outcomes of previous stages are learned. In the extreme case of no information being revealed, this is no different than a simultaneous move game. If some, but not all the information is revealed, the analysis becomes very complicated, and is beyond the scope of this text.

is important because it allows players to *condition* their future actions on past outcomes. This is the idea at the center of multistage games: the ability to condition behavior may lead to a rich set of outcomes. In what follows, we will analyze the idea of conditional strategies, and the equilibria that can be supported using such strategies.

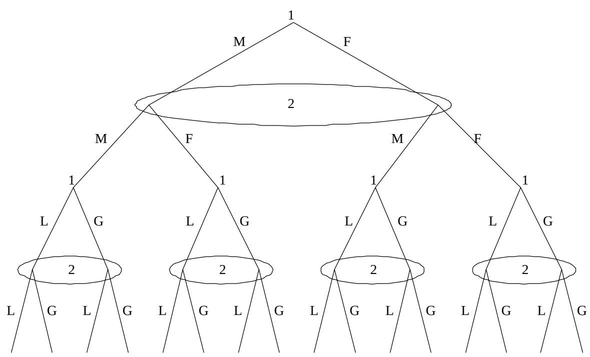
14.1.1 Conditional Strategies

Imagine that our players are playing some multistage game that consists of T stage games. If we treat each game independently, so players do not make any attempt to link them in any way, then strategies would be simple: each player treats every game independently, and commits to some action for each of the stage games in turn.

Players may, however, use strategies that create a *strategic link* between the stage games in ways that are no different than what children do in preschool: "if you play with the dolls this way now, I will play hide and seek with you later, and if you don't then I won't play with you later." More generally, players can use strategies of the form "if such-and-such happens in games $1, 2, ..., t-1$ then I will choose action a_{it} in game t...". Take the "Prisoner-Revenge" game above. Player 1, for example, can use the following strategy: "I will play F in the first game and I will play L in the second game only if player 2 played M in the first game. If player 2 played F in the first game then I will play G in the second."

A convenient way to introduce conditional strategies is by writing down the extensive form of the complete multistage game. As such, we concentrate on what the players can do at every stage of the game, and once a stage is competed, another stage begins. The important part is to correctly determine the information sets of each player. Since the outcomes of every stage are revealed before the next stage is played, then the number of information sets must be identified with the possible number of outcomes from previous stages.

Turning again to the Prisoner-Revenge example, the first stage game has four outcomes. This implies that in the second stage game, each player should have four information sets, each associated with one of the four outcomes of the first game. Note, however, that in each stage game the players do not know what their opponent is choosing, which means that we must have some information sets with more than one node in them. Indeed, ignoring payoffs, the extensive form will have the structure shown in Figure X.X:



As we can now clearly see from the figure, in the first stage the players choose between M and F without knowing what their opponent is choosing. In the second stage, they choose between L and G without knowing what their opponent is choosing, but each knows exactly how the first stage game concluded: either with MM, MF, FM or FF.

This implies that a pure strategy for this Prisoner-Revenge game will be a description of what each player will do in each of his information sets. We can therefore define a strategy for player i as a quintuple, $s_i = (s_i^1, s_i^{MM}, s_i^{MF}, s_i^{FM}, s_i^{FF})$, where s_i^1 is what player i will do in the first stage Prisoner's dilemma game, and s_i^{ab} is what player i will do in the second stage Revenge game, conditional on player 1 played a and player 2 played b in the first stage. Thus, the set of pure strategies for each player can be formally written as,

$$S_i = \left\{ (s_i^1, s_i^{MM}, s_i^{MF}, s_i^{FM}, s_i^{FF}) : s_i^1 \in \{M, F\} \text{ and } s_i^{MM}, s_i^{MF}, s_i^{FM}, s_i^{FF} \in \{L, G\} \right\}.$$

Notice that since each entry of the quintuple is binary, each player has 32 pure strategies, which means that if we try to draw a matrix it will be of size 32×32 and will have 1,024 pure strategy combinations! For this reason, it is most useful to consider the extensive form when deciding to analyze such a game.

More generally, in a multistage game that consists of T stage games a pure strategy of player i will be a list of conditional pure strategies of the following form: $S_i = \{s_i^1, s_i^2(h_1), ..., s_i^t(h_{t-1}), ..., s_i^T(h_{T-1})\}$ where h_{t-1} is the set of outcomes that occurred up to period t, not including period t, and $s_i^t(h_{t-1})$ is an action from the t^{th} stage game. With can conveniently think of h_{t-1} as the *history* of events that occurred up to period t. The different histories are associated with different information sets, each being associated with a distinct sequence of outcomes from the previous stage games combined.

To solidify this idea of conditional strategies consider a game with n firms, that are choosing prices in a sequence of markets. In stage 1 they choose prices in market 1, and this market clears, and this is followed by $T-1$ more market games with their own demand structures. If each firm selects it's price p_i^t in period (stage game) t, then a pure strategy for firm i is a list of prices $p_i^1, p_i^2(h_1), ..., p_i^t(h_{t-1}), ..., p_i^T(h_{T-1})$ where $p_i^1 \geq 0$ and $p_i^t(h_{t-1}) \geq 0$ for all $t \in \{2, ..., T\}$, and each history h_t is the sequence of previously chosen prices: $h_1 = (p_1^1, p_2^1, ..., p_n^1)$, $h_2 = ((p_1^1, p_2^1, ..., p_n^1), (p_1^2, p_2^2, ..., p_n^2))$, and generally,

$$h_t = ((p_1^1, p_2^1, ..., p_n^1), (p_1^2, p_2^2, ..., p_n^2), ..., (p_1^t, p_2^t, ..., p_n^t)).$$

Notice that for this example a pure strategy for player i is a list of continuous functions, each from the set of continuously many histories to a selected price.

Just as we have defined pure strategies for multistage games, we can define behavioral strategies in which every player can mix between his or her pure actions in every information set. The formal extension is straightforward: in a multistage game that consists of T stage games a behavioral strategy of player i will be a list of conditional randomizations of the following form: $\Sigma_i = \{\sigma_i^1, \sigma_i^2(h_1), ..., \sigma_i^t(h_{t-1}), ..., \sigma_i^T(h_{T-1})\}$ where h_{t-1} is the set of outcomes that occurred up to period t, not including period t, and $\sigma_i^t(h_{t-1})$ is a randomization over actions from the t^{th} stage game.

The definition of strategies for the multistage game followed a rather natural structure of defining what people know when they have to choose actions, and using the extensive form to represent the derived structure of conditional strategies. Since we generally treat a game as a triplet of players, strategies and payoffs, we need to construct well defined payoffs for the multistage game.

14.1.2 Payoffs

What does it mean to evaluate the *total payoffs* from a sequence of outcomes in each of the sequentially played stage-games? This is where we turn to a well defined notion from economic analysis (and standard cost–benefit analysis) of *present value*.

In particular, we will *translate* any sequence of payments from the games played to a *single value* that can be evaluated at the beginning of time, before the first game is played. To do this, we take the payoffs derived from each of the individual games, and add then up to derive a total payoff from the anticipated outcomes. However, it may naive to evaluate payoffs from each of the games in the same way. In particular, we will follow the convention that payoffs that are further away in the future are "worth" less that payoffs obtained earlier in the sequence of play.

It is easy to justify this assumption using simple terms from financial markets. For example, if one game is played today and yields a payment of 11, and another is played in a year and it too will yield a payment of 11, then in *today's value* this second payment will be worth less than 11. If, say, there is a 10% interest rate at which we can borrow "utils" (in monetary terms) then we can borrow 10 today and repay 11 in a year, which is the future payoff from the game. Thus, the 11 next year is exactly worth 10 today. An alternative way to say this is that there is a *discount factor* of $\delta = \frac{1}{1.1}$ and the future payment of 11 in the next period (next year) is worth $\delta \cdot 11 = 10$ today.

Another common way to justify this notion of discounting, or *impatience*, is that today's game is played now, and is certain to occur, but there is some uncertainty about tomorrow's game actually occurring. For example, if two players are playing a game in period one, there may be some probability $\delta < 1$ that tomorrow's game will indeed be played, however, with probability $1 - \delta$ it will not. Then, if player i expects to get some utility $\widehat{u}_i$ from tomorrows game if it is played, and a utility of zero if it is not, then his expected utility is $\delta \cdot \widehat{u}_i$.

More generally, consider a multi-stage game in which there are T stage-games played in each of the periods $1, 2, ..., T$. Let $\delta \in [0, 1]$ represent the discount factor that applies to measure future payoffs in comparable, present terms. If we value the future exactly as much as we value the present, then $\delta = 1$. If we value the future less then the present, then $\delta < 1$. Let u_i^t represent the payoff for player i from playing the t^{th} stage-game (the game played in period t). We denote by u_i the *total payoff* of player i from playing a sequence of games in a multi-stage game, and define it as,

$$\begin{aligned} u_i &= u_i^1 + u_i^2\delta + u_i^3\delta^2 + \cdots + u_i^T\delta^{T-1} \\ &= \sum_{t=1}^{T} u_i^t\delta^{t-1}. \end{aligned}$$

which is the discounted sum of payoffs that the player expects to get in the sequence of games. The payoffs one period away are discounted once, two periods away are discounted twice (hence the δ^2) and so on. For example, if in the Prisoner-Revenge game we have the sequence of play be (F, M) followed by (L, G), then given a discount factor δ for each player, the payoffs for players 1 and 2 are $u_1 = 5 + \delta(-4)$ and $u_2 = -1 + \delta(-1)$. Similarly, the outcome (F, F) in the first game followed by (G, L) in the second game will yield utilities of $u_1 = -1 + \delta(-1)$ and $u_2 = -1 + \delta(-4))$ for players 1 and 2 respectively. We can now complete the extensive form of the Prisoner-Revenge game to demonstrate all the possible outcomes, and the total sum of discounted payoffs, as shown in figure X.X:

It is worth noting that a higher discount factor δ means that the players are more patient, and will care more about future payoffs.

14.2 Subgame Perfect Equilibria

Since multi-stage games are dynamic in nature, and the past play is revealed over time, it is natural to turn to subgame perfect equilibrium as a solution concept. In particular, rational players should play sequentially rational strategies, which justifies the concept of SPE. The question is then, how do we find SPE for such a game? The following result will offer a natural first step:

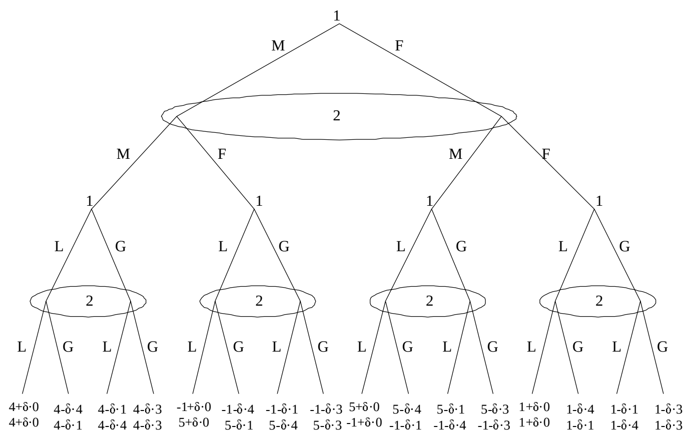

1
M
F
2
M
F
M
F
1
1
1
1
L
G
L
G
L
G
L
G
2
2
2
2
L
G
L
G
L
G
L
G
L
G
L
G
L
G
L
G
4+δ·0 4+δ·0
4-δ·4 4-δ·1
4-δ·1 4-δ·4
4-δ·3 4-δ·3
-1+δ·0 5+δ·0
-1-δ·4 5-δ·1
-1-δ·1 5-δ·4
-1-δ·3 5-δ·3
5+δ·0 -1+δ·0
5-δ·4 -1-δ·1
5-δ·1 -1-δ·4
5-δ·3 -1-δ·3
1+δ·0 1+δ·0
1-δ·4 1-δ·1
1-δ·1 1-δ·4
1-δ·3 1-δ·3

FIGURE 14.1.

Proposition 10 *Consider a multistage game with T stages. If $\sigma^{1*}, \sigma^{2*}, ..., \sigma^{T*}$ is a sequence of Nash equilibria strategy profiles for the independent stage games, then there exists a subgame perfect equilibrium in a multistage game in which the equilibrium path coincides with the path generated by $\sigma^{1*}, \sigma^{2*}, ..., \sigma^{T*}$.*

proof: Fix a sequence of Nash equilibria for each stage game, and let σ^{t*} denote the profile of strategies that form the Nash equilibrium for the stage game t, where σ_i^{t*} is player i's strategy in the profile σ^{t*}. Consider the following *unconditional* strategies for each player i: in each stage $t \in \{1, 2, ..., T\}$ play σ_i^{t*} *regardless* of the sequence of plays before stage t (that is, for any information set in stage t choose σ_i^{t*}). It should be clear that for each of the "last" subgames that start at stage T this is a Nash equilibrium. Now consider all the subgames that start at stage $T-1$. Since the actions in stage T do not depend on the outcomes of stage $T-1$, the payoffs from play in this stage are equal to the payoffs of the stage game at time $T-1$, plus a constant that has two components: the fixed payoff resulting from the previous games and the constant payoff from the game at stage T. This implies that σ_{T-1}^* followed by σ_T^* in any subgame at stage T is a Nash equilibrium in any subgame starting at time $T-1$. Using backward induction, this argument implies that at any stage t, current play does not affect future play, and this inductive argument in turn implies that in any subgame these strategies constitute a Nash equilibrium. ■

At a first glance this proof may be somewhat confusing, but it follows from a very intuitive argument. To see this recall the Prisoner-Revenge game described above, and consider the sequence of stage game Nash equilibria, (F, F) followed by (L, L). Now consider the following multi-stage strategies s_i for players $i \in \{1, 2\}$:

$$s_i = (s_i^1, s_i^{MM}, s_i^{MF}, s_i^{FM}, s_i^{FF}) = (F, L, L, L, L) \text{ for } i \in \{1, 2\}.$$

That is, each player plays F in the first stage game, and then plays L in the second stage-game *regardless of what was played in the first stage.* Now observe that with this pair of strategies the players are clearly playing a Nash equilibrium in each of the four subgames in the second stage. This follows because the first period payoffs are set in stone and do not depend on the second period choices, and in each of these second period subgames both players playing L is a Nash equilibrium. Then moving back to the whole game (the first stage subgame), the payoffs in the second period are a constant and are $(-\delta 3, -\delta 3)$, so each player has a dominant strategy which is to play F, and thus both playing F is a Nash equilibrium. This argument immediately implies that there is another (pure strategy) SPE of the multi stage game in which the players play

$$s_i = (s_i^1, s_i^{MM}, s_i^{MF}, s_i^{FM}, s_i^{FF}) = (F, G, G, G, G) \text{ for } i \in \{1, 2\}.$$

The logic is identical since both players playing G is a Nash equilibrium of the revenge game.[2]

The proposition should therefore not be too surprising. It basically says that if we look at a sequence of plays that is a Nash equilibrium in each game *independently*, then players should be content playing these stage-game strategies in each stage game. By considering such unconditional strategies, we are removing any *strategic link* between stages, and thus each stage can be treated as if it were an independent game.

The interesting question that remains to be answered is whether we are ignoring behavior that can be supported as a SPE if we actually use a strategic linkage between the games. In other words, if we allow players to condition their future play on past play in a sequentially rational manner, will we be able to support play in early stages that is not a Nash equilibrium in the stage games?

The answer is *yes*. This of course will depend on the specific game in consideration, and to see this consider the Prisoner-Revenge game given above. The trick will be to *condition the behavior in the second stage* on the actions taken in the first stage. This is of course possible because there are multiple Nash equilibria in the second stage-game, and sequential rationality implies that players *must play* one of these equilibria in the second stage.[3]

[2] There is also a mixed strategy Nash equilibrium in the revenge game, so there is a third SPE that follows from the proposition: F followed by the mixtures determined by the mixed strategy Nash equilibrium in the revenge game, unconditional on the outcome of the fist stage-game.

[3] This is general to any finite stage game of length T. In the last stage T players *must play a Nash equilibrium* of that stage game since there is no future that depends on the actions taken in stage T. We will revisit this point shortly when we discuss repeated games.

It seems appealing to try and support a SPE in which the players will play (M, M) in the first period. This would be interesting because as we know, in the stand-alone Prisoner's dilemma, playing (M, M) cannot be supported as an equilibrium. However, given the appended Revenge game, we can use conditional strategies.

Since we would like to support the "good behavior" (M, M) as something that the players would like to reward, they can choose the "good equilibrium" of playing L in the second stage if M was played by both in the first stage. Then, if the good behavior was not followed, the players can impose a cost on themselves, or *punish themselves,* by switching to the "bad equilibrium" in the second stage. This means that each player will play the following strategy in the multi-stage game:

- Play M in stage 1. In stage 2 play L if (M, M) was played in stage 1, and play G if anything but (M, M) was played in stage 1. Using the previous notation:

$$s_i = (s_i^1, s_i^{MM}, s_i^{MF}, s_i^{FM}, s_i^{FF}) = (M, L, G, G, G) \text{ for } i \in \{1, 2\}$$

We now need to see if this pair of strategies can be a subgame perfect equilibrium. It is easy to see that in each of the subgames beginning in stage 2, the players are playing a Nash equilibrium (either both play L following (M, M), or they both play G following other histories of play). To check that this is a SPE we need to check that players would not want to deviate from M in the first stage of the game. In other words, is M a best response in period 1, given what each player believes about his opponent, and given the continuation games? Consider player 1 and observe that,

$$u_1(M, s_2) = 4 + 0 \cdot \delta \quad \text{and} \quad u_1(F, s_2) = 5 + (-3) \cdot \delta$$

which implies that M is a best response if and only if $4 \geq 5 - 3\delta$, or,

$$\delta \geq \frac{1}{3}.$$

This example is illuminating because it sheds light on two elements that are crucial to support behavior in the first stage (early periods in general) that is *not* a Nash equilibrium of this first stage. These elements are,

1. There must be at least two distinct equilibria in the second stage: a "stick" and a "carrot"

2. The discount factor has to be large enough

What is the meaning of these two elements? the first is necessary to let us use "reward and punishment" strategies that help support first-period play that is not a Nash equilibrium in the stand alone first stage game – each player would like to deviate from it if there were no future to consider. That is, since the play we are trying to support includes a profile of actions that *is not a Nash equilibrium* of the first stage, it must be true that in the first stage some players would benefit in the "short run" from deviating from the proposed path of play. To keep them on the path of play we must guarantee that such deviations will be met with credible "punishments", and these punishments take the form of moving from an equilibrium where the players get high payoffs to on where they get low payoffs. Thus, we use *long term losses* to deter the players from pursuing short term gains.[4]

This, however, is where the second element comes in: it must be that these long term losses are enough to deter the players from deviating. For this to be the case the players must value the payoffs that they will receive in the second period, which can only happen if they do not discount the future too much, or in other words, if they are *patient enough*. Thus, the *effective punishment* from deviating will depend first on the difference in payoffs of the two equilibria, the "reward" and "punishment", and second on the discount factor. A more figurative way to think about this is that the good and bad continuation equilibria are the "carrot" and the stick" respectively. Then, a smaller discount factor makes the carrot smaller, and the stick less painful, which in turn makes deterrence harder.

To see another example, let's try to support the path of play (F, M) followed by (L, L). To do this we use the crime-and-punishment strategies for each player as follows: Player 1 plays F in stage 1, and player 2 plays M. In stage 2 both players

[4] This implies that these multiple equilibria must have some structure on payoffs. In particular, we need either two equilibria that are Pareto ranked (one is better for everyone), so that we can punish any deviator with the same equilibrium, or we can have more equilibria that would allow for "targeted" punishments to the deviators alone. If the multiple equilibria are like in the battle of the sexes where they are not Pareto ranked then we can't punish both deviators and this will impose limits on the "crime and punishment" type of strategies.

play L if (F, M) was played in stage 1, while they play G if anything but (F, M) was played in stage 1.[5] Clearly, player 1 would not want to deviate from F in the first period since he's playing a dominant strategy in that period, followed by the best possible second stage equilibrium. For player 2 things are different. We have,

$$u_2(M, s_1) = -1 + 0 \cdot \delta \text{ and } u_2(F, s_1) = 1 - 3 \cdot \delta \ ,$$

which implies that M is a best response for player 2 in stage 1 if and only if $-1 \geq 1 - 3\delta$, or in other words,

$$\delta \geq \frac{2}{3} \ .$$

As we can see, to support this outcome we need a higher discount factor compared with supporting (M, M) in the first period. Compared to the earlier path of (M, M) followed by (L, L), here the future loss of 3 is enough to deter player 2's deviation only when the future is not discounted as much because the short term gains from deviating are equal to 2, whereas before they were equal only to 1. L10

[5] Again, using the previous notation this is $s_1 = (F, G, G, L, G)$ and $s_2 = (M, G, G, L, G)$.