In class on March 1, we used an infinitely repeated prisoner’s dilemma to illustrate how the one-stage-deviation-principle can be a useful proof technique. This note checks and corrects some of the in-class calculations.

First, recall that the OSDP states that

A strategy profile (a combination of strategies, one for each player) is a subgame perfect Nash equilibrium if and only if no player has an incentive to deviate in a single stage or subgame and conform to his original strategy thereafter.

In order to show that the OSDP can be a useful proof technique, we will assume that two individuals play the prisoner’s dilemma repeatedly for infinitely many periods. The payoffs are shown below.

<table>
<thead>
<tr>
<th>Player 1</th>
<th>Deny</th>
<th>Confess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deny</td>
<td>1,1</td>
<td>-1, 2</td>
</tr>
<tr>
<td>Confess</td>
<td>2, -1</td>
<td>0,0</td>
</tr>
</tbody>
</table>

We are going to guess that, provided the players are sufficiently patient, so the discount factor \( \delta = \frac{1}{1 + r} \in (0,1) \) is large enough, the following strategy profile is a subgame perfect Nash equilibrium (SPNE) in the infinitely repeated game.

“Play DD unless any player Confessed in the past. In that case, play DC in even periods and CD in odd periods. If anybody deviated from the DC, CD, etc. sequence, play CC in all periods”

We can think of the proposed SPNE as a set of three “phases,” where

Phase I= “The Cooperation Phase” with (DD)
Phase II=”The Punishment Phase” with the less beneficial (2,-1), (-1,2), etc. or (-1,2), (2,-1) etc. payoff sequence
Phase III=”The Scaled-up Punishment Phase”

Intuitively, the players start by cooperating on DD, which maximizes their joint payoff. If somebody “cheats” they start the punishment phase, phase II. If somebody cheats again in the punishment phase, they scale up the punishment to the most mutually harmful outcome, which forces the minimum possible payoff (0) on each player. (Note that they cannot force -1 on any player forever since that player can always switch to confess and guarantee herself 0).

Proof of the SPNE claim:

(i) In a multi-stage game, every stage lasts until the stage game starts over. In our example, the stage game starts over every time period. In order to use the OSDP to prove the equilibrium, we must therefore prove that the players cannot benefit from deviating in any single period (=any single stage) and conforming to their original strategies in the remaining periods.

(ii) Due to observation (i) we “just” need to check that
(1) **Player 1 will not deviate in a phase I EVEN period.** If player 1 deviates in a phase I EVEN period and then conforms in phase II, she gets a payoff of 2 immediately and then there begins a sequence (since the next period is odd) of DC, DC, CD, DC, etc. This sequence pays her alternately 2 and (-1). On the other hand, if she conforms in phase I, the game remains in phase I and she just gets a payoff 1 per period. She has no single-stage deviation incentive in a phase I EVEN period if the deviation payoff is less than the payoff to conforming, so

\[
2 + \delta^2(-1) + \delta^3 2 + \delta^4(-1) + \delta^5 2 + \delta^6(-1) + \ldots \leq 1 + \delta + \delta^2 + \delta^3 + \ldots \quad \Leftrightarrow \quad 2 + \delta \sum_{t=0}^{\infty} \delta^{2t} (2 + \delta(-1)) \leq \sum_{t=0}^{\infty} \delta^t \quad \Leftrightarrow \\
2 + \delta (2 - \delta) \sum_{t=0}^{\infty} \delta^{2t} \leq \sum_{t=0}^{\infty} \delta^t \quad \Leftrightarrow \\
2 + \delta (2 - \delta) \sum_{t=0}^{\infty} (\delta^2)'^t \leq \sum_{t=0}^{\infty} \delta^t
\]

The reason we write it this way is that we can now use the property of geometric series that

if \( x \in (0,1) \), then \( \sum_{t=0}^{\infty} x^t = \frac{1}{1-x} \), we can therefore rewrite the preceding inequality as

\[
2 + \delta (2 - \delta) \frac{1}{1-\delta^2} \leq \frac{1}{1-\delta} \quad \Leftrightarrow \quad 2(1-\delta) + (2-\delta)\delta \frac{1-\delta}{1-\delta^2} \leq 1 \quad \Leftrightarrow \\
(2-\delta) \left(1 + \delta \frac{1-\delta}{1-\delta^2}\right) - \delta \leq 1
\]

It is hard to solve (2) for the minimum necessary \( \delta \) value since it is a fifth-order polynomial (after we multiply by \( 1-\delta^2 \)). However, we just have to prove that (2) holds as the discount factor \( \delta \to 1 \), so when we take the limit. We cannot take the limit right away since \( \lim_{\delta \to 1} \frac{1-\delta}{1-\delta^2} = 0/0 \). Nonetheless, we can use L’Hopital’s rule, which says that, when it appears that \( \lim_{x \to a} \frac{f(x)}{g(x)} = \frac{0}{0} \) or \( \lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\infty}{\infty} \), the limit is actually \( \lim_{x \to a} \frac{f'(x)}{g'(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)} \). In our case, therefore, \( \lim_{\delta \to 1} \frac{1-\delta}{1-\delta^2} = \lim_{\delta \to 1} \frac{-1}{-2\delta} = \frac{1}{2} \). The limit of the left hand side of (2) as \( \delta \to 1 \) is therefore

\[
\lim_{\delta \to 1} (2-\delta) \left(1 + \delta \frac{1-\delta}{1-\delta^2}\right) - \delta = \lim_{\delta \to 1} \left(1 + \frac{1}{2}\right) - 1 = \frac{3}{2} - 1 = \frac{1}{2} < 1.
\]

Thus, as \( \delta \) approaches one there is no deviation incentive.

(2) **Player 1 will not deviate in a phase I ODD period.** If player 1 deviates in a phase I ODD period and then conforms in phase II, she gets a payoff 2 immediately and then there begins a sequence (since the next period is even) of DC, CD, DC, etc. This sequence pays her alternately (-1) and (2). If she conforms in phase I, the game remains in phase I and she just gets a payoff 1 per period. She has no single-stage deviation incentive in a phase I ODD period if
\[ 2 + \delta(-1) + \delta^2(2) + \delta^3(-1) + \delta^4(2) + \ldots \leq 1 + \delta + \delta^2 + \delta^3 + \ldots \Leftrightarrow \]
\[ 2 + \delta \sum_{i=0}^{\infty} \delta^{2i} (-1 + \delta(2)) \leq \sum_{i=0}^{\infty} \delta^i. \quad (3) \]

Inequality (3) follows from inequality (1) since \(-1 + 2\delta < 2 - \delta \Leftrightarrow \delta < 1\), which we assume.

(3) **Player 1 will not deviate in a phase II EVEN period.** In a phase II EVEN period, player 1 is supposed to play D and player 2 C, so the strategy profile is DC. If player 1 deviates, she will change it to CC and her payoff will be 0 instead of (-1). After the deviation, phase III begins with CC in each period, so she gets 0. On the other hand, if she conforms in the phase II EVEN period, the game remains in phase II and she gets the payoffs from the DC, CD, DC, etc. sequence. She has no single-stage deviation incentive in a phase II EVEN period if

\[ 0 + \delta0 + \delta^20 + \delta^30 + \ldots \leq -1 + \delta2 + \delta^2(-1) + \delta^32 + \delta^4(-1) + \ldots \Leftrightarrow \]
\[ 0 \leq 1 + \delta \sum_{i=0}^{\infty} \delta^{2i} (2 + \delta(-1)) \Leftrightarrow 1 \leq \delta(2 - \delta) \sum_{i=0}^{\infty} \delta^{2i} \Leftrightarrow 1 \leq \delta(2 - \delta) \sum_{i=0}^{\infty} (\delta^2)^i \Leftrightarrow \]
\[ 1 \leq \frac{\delta(2 - \delta)}{1 - \delta^2} \quad (4) \]

As \( \delta \to 1 \), the right hand side becomes large and the inequality holds (or we can solve it explicitly by rewriting it as \( 1 - \delta^2 \leq 2\delta - \delta^2 \Leftrightarrow \delta \geq 0.5 \)).

(4) **Player 1 will not deviate in a phase II ODD period.** In a phase II ODD period, player 1 is supposed to play C and player 2 D, so the strategy profile is CD. If player 1 deviates, she will change it to DD and her payoff will be 1 instead of 2. After the deviation, phase III begins with CC in each period, so she gets 0. On the other hand, if she conforms in the phase II ODD period, the game remains in phase II and she gets the payoffs from the CD, DC, CD, etc. sequence. She has no single-stage deviation incentive in a phase II ODD period if

\[ 1 + \delta0 + \delta^20 + \delta^30 + \ldots \leq 2 + \delta(-1) + \delta^2(2) + \delta^3(-1) + \delta^4(2) + \ldots \Leftrightarrow \]
\[ -1 \leq \delta \sum_{i=0}^{\infty} \delta^{2i} (-1 + \delta2) = \delta(2\delta - 1) \sum_{i=0}^{\infty} (\delta^2)^i = \delta \frac{2\delta - 1}{1 - \delta^2} \quad (5) \]

which must at least hold for \( \delta \geq 0.5 \), since then the right hand side of (5) is positive.

(5) **Player 1 will not deviate in a phase III period.** In a phase III period, player 1 is supposed to play C and player 2 C, so the strategy profile is CC. If player 1 deviates, she will change it to DC and her payoff will be (-1) instead of 0. After the deviation, phase III continues with CC in each period, so she gets 0. On the other hand, if she conforms in phase III period, the strategy sequence remains CC and the payoff is still 0 in each period. She has no single-stage deviation incentive in a phase III period if
\[-1 + \delta \sum_{r=0}^{\infty} \delta' 0 \leq 0 + \delta \sum_{r=0}^{\infty} \delta' 0 \iff -1 \leq 0, \quad (6)\]

which always holds.

(6) We actually also need to disprove all of player 2’s deviation incentives. However, the deviation incentives for player 2 in phase I and II EVEN periods are identical to the deviation incentives of player 1 in phase I and II ODD periods and vice versa. In stage III, the players have the same deviation incentives. The proof then follows from the preceding.

Finally, note that, in order to complete the proof, we used the OSDP and only ruled out single-stage deviations. For example, inequalities (1) and (3) checked that player 1 did not want to deviate in just a single phase I period taking as given that she would conform to the phase II behavior her SPNE strategy specifies. Similarly, inequalities (4)-(5) checked that she would not deviate in phase II and then conform to the phase III behavior her SPNE strategy specifies. Admittedly, combining (1)-(5) rules out double phase I-II deviations. But we still have not ruled out deviating, say, 5 times in a row in stage III. The OSDP shows that we do not have to check for such multi-stage deviations to prove our SPNE: if it is not an SPNE, we will find a problem just by checking the single-stage deviation gains. If we do not find a single-stage gain, it is an SPNE.