Prisoners' Dilemma & Repeated Games

Prisoners’ Dilemma

- It is probably the classic example of the theory of strategy and its implications for predicting the behavior of game players.
- The prisoners’ dilemma is a game in which each player has a ***dominant strategy***.
- But the equilibrium that arises when all players use their dominant strategies provides a ***worse outcome for every player*** than would arise if they all used their dominated strategies instead.
- The paradoxical nature of this equilibrium outcome leads to several more ***complex questions about the nature of the interactions***

Prisoners' Dilemma

- The "prisoners" can find another outcome that both prefer to the equilibrium outcome, but they find it difficult to bring about.

- Question – whether and how the players in a prisoners' dilemma can ***attain and sustain their mutually beneficial cooperative outcome***, overcoming their separate incentives to defect for individual gain?

- We first review the standard prisoners' dilemma game and then develop three categories of solutions.

- The first and most important method of solution consists of ***repetition of the standard one-shot game***. We then consider two other potential solutions that rely on ***penalty (or reward) schemes*** and on the ***role of leadership***.

The Basic Game (Review)

- Husband and wife suspected of murder. Each is interrogated separately and can choose to confess to the crime or to deny any involvement.

- The payoff matrix that they face –

		WIFE	
		Confess (Defect)	Deny (Cooperate)
HUSBAND	Confess (Defect)	10 yr, 10 yr	1 yr, 25 yr
	Deny (Cooperate)	25 yr, 1 yr	3 yr, 3 yr

- The numbers shown indicate years in jail; therefore low numbers are better for both players.

The Basic Game (Outcome)

- Both players here have a **dominant strategy**.
- Each does **better to confess**, regardless of what the other player does.
- The equilibrium outcome entails both players deciding to confess and **each getting 10 years in jail**.
- If they both had chosen to **deny** any involvement, however, they would have been **better off**, with only **3 years of jail** time to serve
- In any prisoners' dilemma (class of) game, there is always a ***cooperative strategy*** and a ***cheating*** **or** ***defecting strategy***

The Basic Game (Outcome)

- ***Deny is the cooperative strategy***; both players using that strategy yields the best outcome for the players.

- ***Confess is the cheating or defecting strategy***; when the players do not cooperate with one another, they choose to Confess in the ***hope of attaining individual gain at the rival's expense***.

- Thus, players in a prisoners' dilemma can always be labeled, according to their choice of strategy, as either ***defectors or cooperators***.

The Basic Game (Alternative Outcome)

- Although we speak of a cooperative strategy, the prisoners' dilemma game is non-cooperative – the players make their decisions and implement their choices ***individually***

- If the two players could discuss, choose, and play their strategies jointly (eg, the prisoners were in the same room and could give a joint answer to the question of whether they were both going to confess) there would be no difficulty about their achieving the outcome that both prefer

- The essence of the questions of **whether, when, and how** a prisoners' dilemma can be resolved is the ***difficulty of achieving a cooperative*** (jointly preferred) ***outcome through non-cooperative*** (individual) ***actions.***

Solution 1 (Repetition)

- Of all the mechanisms that can sustain cooperation in the prisoners' dilemma, the best known and the most natural is **repeated play of the game.** Repeated or ongoing relationships between players imply special characteristics for the games that they play against one another.

- In the prisoners' dilemma, this result manifests itself in the fact that each player ***fears that one instance of defecting will lead to a collapse of cooperation in the future***.

- If the ***value of future cooperation*** is large and exceeds what can be gained in the short term by defecting, then the ***long-term individual interests*** of the players can automatically and tacitly keep them from defecting, ***without the need for any additional punishments or enforcement by third parties***

Solution 1 (Repetition)

- We consider the meal-pricing dilemma faced by the two restaurants, Xavier's Tapas and Yvonne's Bistro

- For simplicity, assume that only two choices of price are available: the jointly best (collusive) price of \$26 or the Nash equilibrium price of \$20

- The payoffs (profits measured in hundreds of dollars per month) for each restaurant can be calculated by using the quantity (demand) functions are calculated as earlier –

$$\pi_x = (P_x - 8)Q_x = (P_x - 8)(44 - 2P_x + P_y)$$
$$= (26 - 8)(44 - 52 + 26) = 18 \times 18 = 324$$

Solution 1 (Repetition)

		YVONNE'S BISTRO	
		20 (Defect)	26 (Cooperate)
XAVIER'S TAPAS	20 (Defect)	288, 288	360, 216
	26 (Cooperate)	216, 360	324, 324

- As in any prisoners' dilemma, each store has a dominant strategy to defect and price its meals at \$20, although both stores would prefer the outcome in which each cooperates and charges the higher price of \$26 per meal

Solution 1 (Repetition)

- Let's start by supposing that the two restaurants are initially in the cooperative mode, each charging the higher price of $26.

- If one restaurant—say, Xavier's—deviates from this pricing strategy, it can increase its profit from 324 to 360 (from $32,400 to $36,000) for one month.

- But then ***cooperation has dissolved*** and Xavier's rival, Yvonne's, will see no reason to cooperate from then on.

- Once cooperation has broken down, presumably permanently, the profit for Xavier's is 288 ($28,800) each month instead of the 324 ($32,400) it would have been if Xavier's had never defected in the first place.

Solution 1 (Repetition)

- By gaining 36 (360-326) in one month of defecting, Xavier's gives up 36 (324-288) each month thereafter by destroying cooperation.

- Even if the relationship lasts as little as three months, it seems that defecting is not in Xavier's best interest.

- A similar argument can be made for Yvonne's.

- Thus, if the two restaurants ***competed on a regular basis*** for at least three months, it seems that ***we might see cooperative behavior*** and high prices rather than the defecting behavior and low prices predicted by theory for the one-shot game

Finite Repetition

- The solution of the dilemma is not actually that simple.
- ***What if the relationship did last exactly three months?***
- Then strategic restaurants would want to analyze the full three-month game and choose their optimal pricing strategies.
- Each would use ***rollback*** to determine what price to charge each month.
- Starting their analyses with the third month, they would realize that, at that point, ***there was no future relationship to consider***.

Finite Repetition

- Each restaurant would find that it had a dominant strategy to defect.

- Given that, there is effectively no future to consider in the second month either.

- Each player knows that there will be mutual defecting in the third month, and therefore both will defect in the second month; defecting is the dominant strategy in month 2 also.

- Then the same argument applies to the first month as well. Knowing that both will defect in months 2 and 3 anyway, ***there is no future value of cooperation in the first month***. Both players defect right from the start, and **the dilemma persists**.

Finite Repetition

- This result is very general.
- As long as the relationship between the two players in a prisoners' dilemma game lasts a ***fixed and known length of time***, the dominant-strategy equilibrium with defecting should prevail in the last period of play.
- When the players arrive at the end of the game, there is never any value to continued cooperation, and so they defect.
- Then ***rollback predicts mutual defecting all the way back to the very first play***.

Infinite Repetition

- What would happen if the relationship did not have a predetermined length?

- ***What if the two restaurants expected to continue competing with one another indefinitely***?

- In repeated games of any kind, the sequential nature of the relationship means that players can adopt ***strategies that depend on behavior in preceding plays of the games***.

- Such strategies are known as **contingent strategies**

Infinite Repetition

- Most contingent strategies are **trigger strategies.**
- A player using **a trigger strategy** plays cooperatively as long as her rival(s) do so, but any defection on their part "**triggers**" a period of **punishment**, of specified length, in which she plays non-cooperatively in response.
- Two of the best-known trigger strategies are the grim strategy (also known as grim trigger) and tit-for-tat.
- The **grim strategy** entails cooperating with your rival until such time as she defects from cooperation; once a defection has occurred, you punish your rival (by choosing the Defect strategy) on every play ***for the rest of the game***

Infinite Repetition

- **Tit-for-tat (TFT)** is not so harshly unforgiving as the grim strategy and is famous for its ability to solve the prisoners' dilemma ***without requiring permanent punishment***.

- Playing TFT involves cooperating on the first play and then choosing, in each future period, the action chosen by your rival in the preceding period of play. Thus, when playing TFT, you cooperate with your rival if she cooperated during the most recent play of the game and defect (as punishment) if your rival defected.

- The punishment phase lasts only as long as your rival continues to defect; you will ***return to cooperation*** one period after she chooses to do so.

Infinite Repetition

- **Grim Strategy:** If Xavier defects in period 1, Yvonne defects for all the periods to come – Xavier looses 36 (324-288) in each period starting from period 2 while he makes profit of 36 (360-324) in the first period by defecting

		YVONNE'S BISTRO	
		20 (Defect)	26 (Cooperate)
XAVIER'S TAPAS	20 (Defect)	288, 288	360, 216
	26 (Cooperate)	216, 360	324, 324

- **TFT:** Xavier defects first period. Yvonne retaliates in the 2nd period. But Xavier's could get back to cooperation in the 3rd period which would be reciprocated by Yvonne.

Infinite Repetition

		YVONNE'S BISTRO	
		20 (Defect)	26 (Cooperate)
XAVIER'S TAPAS	20 (Defect)	288, 288	360, 216
	26 (Cooperate)	216, 360	324, 324

- It is important to realize here, however, that Xavier's extra 36 from defecting is gained in the first month. Its losses 108 (216 instead of 324) are ceded in the 2nd month. Yvonne gets 360 in 2nd month

- Therefore, the relative importance of the two depends on the relative importance of the present versus the future.

Rate of return

- Generally, money (or profit) that is earned today is better than money that is earned later because, even if you do not need (or want) the money until later, you can invest it now and earn a return on it until you need it.

- We use the symbol r to denote this rate of return.

- Thus, one dollar invested generates r dollars of interest and/or dividends and/or capital gains

- Xavier loses 108 and gains 36 by defecting once. ***Is defecting for only one month worth it***?

Present Value

- We need to calculate the present value (PV) of the 108 loss in 2nd month to make it comparable to the 36 gain the 1st month.

- PV $= (\frac{108}{1+r})$ and hence, for any value of r, we can now determine the exact number of dollars that, earned this month, would be worth 108 next month

- So, for Xavier, defecting is wothwhile if $36 > (\frac{108}{1+r})$ or, $r>2$. Thus, Xavier's should choose to defect once against a rival playing TFT only if the monthly total rate of return exceeds 200% which is quite impossible!
- Therefore, it is better for Xavier's to continue cooperating than to try a single instance of defecting when Yvonne's is playing TFT.

Infinite Repetition

- **What about the possibility of defecting once and then continuing to defect forever?**

- This second option of Xavier's gains the restaurant 36 in the first month but loses it 36 in every month thereafter into the future if the rival restaurant plays TFT.

- To determine whether such a strategy is in Xavier's best interest again depends on the present value of the losses incurred.

- But this time the losses are incurred over an **infinite horizon** of future months of competition.

Infinite Repetition

- Mathematically, the future loss can be written as the sum of an infinite number of terms –

$$\frac{36}{1+r}+\frac{36}{(1+r)^2}+\frac{36}{(1+r)^3}+\frac{36}{(1+r)^4}+\ldots\ldots$$

- The ratio $\left(\frac{1}{1+r}\right)$ is generally called the discount factor and denoted by δ and as $r>0$ we have $\delta<1$

- The series will boil down to $\left(\frac{36}{r}\right)$

- Then Xavier defects forever only if $36 > 36/r$, or, $r>1$

Infinite Repetition

- Defecting forever is beneficial in this particular game only if the monthly rate of return exceeds 100%, another unlikely event.

- Thus, we would not expect Xavier's to defect against a cooperative rival when both are playing tit-for-tat

- We neither expect defection against a cooperative rival when both are playing grim strategy

- When both Yvonne and Xavier play TFT, the cooperative outcome is a Nash equilibrium of the game. Both playing TFT is a Nash equilibrium, and use of this contingent strategy solves the prisoners' dilemma for the two restaurants.

Games of Unknown Length

- It is possible that, in some repeated games, players might not know for certain exactly how long their interaction will continue.

- They may, however, have some idea of the probability that the game will continue for another period

- The present value of a loss next month is already worth only $\delta = (\frac{1}{1+r})$ times the amount earned.

- If in addition there is only a probability p (<1) that the relationship will actually continue to the next month, then next month's loss is worth only p times δ times the amount lost

Effective Rate of Return

- We call this **effective rate of return (*R*)** such that $\frac{1}{1+R} = p\delta$

- With a 5% actual rate of return on investments (r = 0.05, and so δ=1/1.05= 0.95) and a 50% chance that the game continues for an additional month (p =0.5), then R =1.1, or 110%.

- The ***high rates of return required to destroy cooperation*** (encourage defection) in these examples seem more realistic if we interpret them as ***effective rather than actual*** rates of return

- It becomes conceivable that defecting forever, or even once, might actually be to one's benefit if there is a large enough probability that the game will end in the near future

Games of Unknown Length

- Xavier's decision of permanent defecting is beneficial only when r exceeds 1, or 100%.

- If Xavier's faces the 5% actual rate of return and the 50% chance that the game will continue for an additional month, then the effective rate of return of 110% will exceed the critical value needed for it to continue defecting.

- Thus, the cooperative behavior sustained by the TFT strategy can break down if there is a sufficiently large chance that the repeated game might be over by the end of the next period of play—that is, by a sufficiently small value of p.

General Theory

- ***When is it worthwhile to defect against TFT-playing rivals?***

- We use a table with general payoffs that satisfy the standard structure of payoffs in the dilemma satisfying the relation **H > C > D > L** for the game to be a prisoners' dilemma, where –

1. C is the ***cooperative outcome***,
2. D is the payoff when ***both players defect from cooperation***,
3. H is the ***high payoff that goes to the defector*** when one player defects while the other cooperates, and
4. L is the ***low payoff that goes to the loser*** (the cooperator) in the same situation

General Theory

- A player defects once against a TFT-playing opponent only if $(H - C) > (C - L)/(1 + R)$ or $R > \frac{C-L}{H-C} - 1$
- a player defects forever against a TFT-playing opponent, or defects at all against a grim-playing opponent, only if $(H - C) > (C - D)/R$, or $R > \frac{C-D}{H-C}$

		COLUMN	
		Defect	Cooperate
ROW	Defect	D, D	H, L
	Cooperate	L, H	C, C

Solution: Penalties and Rewards

- One of the simplest ways to avert the prisoners' dilemma in the one-shot version of the game is to inflict some direct **penalty on the players when they defect.**

- When the ***payoffs have been altered to incorporate the cost of the penalty***, players may find that the dilemma has been resolved.

- If only one player defects, the game's outcome entails 1 year in jail for the defector and 25 years for the cooperator.
- The defector, though, getting out of jail early, might find the cooperator's friends waiting outside the jail.
- The physical harm caused by those friends might be equivalent to an additional 20 years in jail.

Solution: Penalties and Rewards

- With the additional 20 years in jail added to each player's sentence when one player confesses while the other denies, the game is completely different.

		WIFE	
		Confess	Deny
HUSBAND	Confess	10 yr, 10 yr	21 yr, 25 yr
	Deny	25 yr, 21 yr	3 yr, 3 yr

- No dominant strategy, two pure strategy NE – (C, C) and (D, D). Now each player finds that it is in his or her best interest to cooperate if the other is going to do so

Solution: Penalties and Rewards

- The game has changed from being a prisoners' dilemma to an assurance game

- Solving the new game requires selecting an equilibrium from the two that exist.
- One of them—the cooperative outcome—is clearly better than the other from the perspective of both players.
- Therefore, it may be easy to ***sustain it as a focal point*** if some convergence of expectations can be achieved

- The penalty in this scenario is inflicted on a defector ***only when his or her rival does not defect***

Solution: Penalties and Rewards

- Such discipline typically must be imposed by a third party with some power over the two players
- Think of organized crime gang that has a standing rule of never confessing to the police under penalty of extreme physical harm and hence, ***penalty imposed on anyone who defects***
- IEDS: (Deny, Deny)

		WIFE	
		Confess	Deny
HUSBAND	Confess	30 yr, 30 yr	21 yr, 25 yr
	Deny	25 yr, 21 yr	3 yr, 3 yr

Solution III: Leadership

- In PD game all the players stand to lose (and gain) the same amount from defecting (and cooperation).
- However, in actual strategic situations, **one player may be relatively "large" (a leader) and the other "small."**
- If the size of the payoffs is ***unequal enough***, so much of the harm from defecting may fall on the larger player that she acts cooperatively, even while knowing that the other will defect

- Saudi Arabia, for example, played such a role as the "swing producer" in for many years; to keep oil prices high, it cut back on its output when one of the smaller producers, such as Libya, expanded
- Leadership tends to be observed more often in games between nations than in games between firms or individual persons

Solution III: Leadership

- Two countries, same disease, total cost of population being affected is $3.2billion ($1.6bn each)
- Cost of vaccine research $2bn - simultaneous

		SOPORIA	
		Research	No Research
DORMINICA	Research	–2, –2	–2, 0
	No Research	0, –2	–1.6, –1.6

- Dominant strategy: (NR, NR)

Solution III: Leadership

- Population in two countries different

		SOPORIA	
		Research	No Research
DORMINICA	Research	–2, –2	–2, 0
	No Research	0, –2	–2.4, –0.8

- No Research is still the dominant strategy for Soporia.
- But Dorminica's best response is now Research
- Dorminica now stands to suffer such a large portion of the total cost of the disease that it finds it worthwhile to do the research on its own. This is true even though Dorminica knows well that Soporia is going to be a free rider and get a share of the full benefit of the research.

Evidence of PD

- Examples of the collapse of cooperation as players near the end of a repeated game are observed in numerous situations in the real world, as well as in the laboratory.

- The story of a long-distance bicycle (or foot) race is one such example. There may be a lot of cooperation for most of the race, as players take turns leading and letting others ride in their slipstreams; nevertheless, as the finish line looms, each participant will want to make a dash for the tape

- signs saying "no checks accepted" often appear in stores in college towns each spring near the end of the semester

The Kyoto Protocol

- One crucial example pertains to the international climate control agreement known as the Kyoto Protocol.

- Negotiated by the United Nations Framework Convention on Climate Change in 1997 as a tool for reducing greenhouse gas emissions

- The difficulty in achieving global reduction in greenhouse gas emissions comes in part from the prisoners' dilemma nature of the interaction.

The Kyoto Protocol

- Any individual country will have no incentive to reduce its own emissions, knowing that if it does so alone, it bears significant costs with little benefit to overall climate change.

- If others do reduce their emissions, the first country cannot be stopped from enjoying the benefits of the others' actions.

		THEM	
		Cut Emissions	Don't Cut
US	Cut Emissions	−1, −1	−20, 0
	Don't Cut	0, −20	−12, −12

The Kyoto Protocol - Status

- It went into effect in 2005 and its first phase expired in 2012.
- More than 170 countries signed on to the original treaty, although the United States was noticeably absent from the list.
- The protocol was extended at almost the last minute, in mid-December of 2012.
- It is now in place through 2020.

References

1. Games of Strategy (3rd Edition) by Avinash Dixit, Susan Skeath and David H. Riley Jr.; Viva-Norton [Chapter 11].