# Cooperation in the Prisoner's Dilemma

**Pevchiv Daniil**[1, 2,*]

[1] Moscow Institute of Physics and Technology, Russia
[2] New Economic School, Russia

## Abstract

**This article investigates the iterative game "Prisoner's Dilemma" with a focus on a reformist approach presented by the state. We study the strategic setup of the game aimed at stimulating cooperation among players and enhancing social welfare. The research addresses issues of taxation, compensation to change players' incentives, and probabilistic verification of strategies. By combining these concepts into a single model, we obtain results that may be practically useful.**

**Keywords**: Prisoner's Dilemma

## 1 Introduction

The Prisoner's Dilemma is characterized by the asymmetry between the benefits of selfish actions of an individual and the benefits of cooperative strategies. Pursuing personal interests within society can potentially lead to a decrease in social welfare. Our research aims to study the static approach to influencing players' incentives.

Research on the conditions under which cooperation can be established has mainly been conducted in the second half of the twentieth century ([1], [2], [3]). The literature explores the impact and effectiveness of social incentives on cooperation ([4]). Social incentives include taxes for opportunism and compensations for those who play for the benefit of society. While there is some evidence that cooperation can be spontaneous, especially in small groups, efforts to develop cooperation in larger groups can be effective.

This study complements the literature on the structure and impact of social incentives on cooperation by presenting a model and static analysis. The model will include the probability of opportunistic behavior, a tax that all players are forced to pay at the beginning of each game to introduce an authority that will seek violators, as well as compensation for players who fall into a deception situation. The conclusion will briefly discuss possible applications of this model.

## 2 Model

### 2.1 Under what parameters do we still have a Prisoner's Dilemma?

Consider the classic Prisoner's Dilemma with the following payoff matrix.

| | Cooperation | Opportunism |
|---|---|---|
| Cooperation | $3, 3$ | $0, 5$ |
| Opportunism | $5, 0$ | $1, 1$ |

Since it is not always possible to identify violators in real life, we propose to introduce a probability $p \in [0, 1]$, which defines the chance of detecting a player's opportunistic strategy in each individual game. In this model, if we catch a player violating, their payoff in this game will be $0$, while the opponent's payoff will remain unchanged.

**Under what probability p do we still have a Prisoner's Dilemma?**

To do this, we need to check that for any move of the opponent, it is more profitable for us to deceive them:

«C»: $3 < p \cdot 0 + (1-p) \cdot 5,\ 1 - p > \frac{3}{5},\ p < \frac{2}{5}$.

«D»: $0 < p \cdot 0 + (1-p) \cdot 1,\ p < 1$.

**At $p < \frac{2}{5}$ we still have a Prisoner's Dilemma.**

Let's find the expected payoff for players from each of the moves. Let our opponent play the "cooperation" strategy with probability $a$ and "opportunism" with probability $1 - a$.

$\pi^C = 3a + 0 > \pi^D = p(5a + (1-a)) \cdot 0 + (1-p)(5a + (1-a)) \quad \forall a \in [0, 1]$

We want the expected payoff from the cooperative strategy to be greater.

From this, we get that

$$p > \frac{1+a}{1+4a} = 1 - \frac{3a}{4a+1} \geqslant 1 - \frac{3}{5} = \frac{2}{5}.$$

**At $p > \frac{2}{5}$ the cooperation strategy will dominate.**

Therefore, in this model, players will have no incentives to switch to the cooperation strategy, so let's modify it.

In our model, we consider the probability $p$ as analogous to law enforcement, which is necessary to stimulate players to choose the cooperation strategy. Just as taxes are collected to ensure the functioning of law enforcement, introducing a tax $t$ into our model is a natural step to ensure its operation.

Players pay a tax $t$ at the beginning of each turn. If a player plays the cooperation strategy "C" and their opponent chooses opportunism "D", the player will receive a compensation of $r + t$. It is also logical to assume that $p$ depends on $t$. The higher the $t$, the higher the $p$.

**Under what parameters $p, t, r$ do we still have a Prisoner's Dilemma?**

«C»: $3 - t < (p \cdot 0 + (1 - p) \cdot 5) - t$, the tax does not affect and $p < \frac{2}{5}$.

«D»: $pr + (1 - p) \cdot (0 - t) < (p \cdot 0 + (1 - p) \cdot 1) - t$,

$p(r + t + 1) < 1, r + t < \frac{1-p}{p}, p \in [0, \frac{2}{5})$.

Let's find the expected payoff for players from each of the moves. Let our opponent play the "cooperation" strategy with probability $a$ and "opportunism" with probability $1 - a$.

$\pi^C = a \cdot (3 - t) + (1 - a) \cdot (p \cdot (0 + r) + (1 - p) \cdot (0 - t)) = 3a + pr + pt - t - apt - apr$

$\pi^D = a \cdot (p \cdot (0 - t) + (1 - p) \cdot (5 - t)) + (1 - a) \cdot (p \cdot (0 - t) + (1 - p) \cdot (1 - t)) = 4a - 4pa + 1 - p - t$

We want the expected payoff from the cooperative strategy to be greater.

From this, we get that $4pa + p - 1 - a + pr + pt - apt - apr > 0 \quad \forall a \in [0, 1], p < \frac{2}{5}, r + t < \frac{1-p}{p}$

Therefore,

$$p > \frac{a + 1}{4a + 1 + (1 - a)(r + t)}.$$

Therefore,

$$\frac{a + 1}{4a + 1 + (1 - a)(r + t)} < \frac{2}{5}$$

Considering that we solve this inequality for any $a \in [0, 1]$, we get that $r + t > \frac{3}{2}$.

Therefore,

$$\frac{3}{2} < r + t < \frac{1 - p}{p}.$$

$\frac{3}{2} < \frac{1-p}{p}$ is true for $p \in (0, \frac{2}{5})$.

**Thus, if $r + t > \frac{3}{2}, p < \frac{2}{5}$, we still remain within the framework of the Prisoner's Dilemma, but the cooperation strategy dominates the opportunism strategy.**

### 2.2 Example of model application

Consider the following game: there are $N$ companies participating in a repeated Prisoner's Dilemma. Each company has two strategies: switch to biodegradable materials or continue using plastic. In this situation, the state acts as a law enforcement agency, aiming to persuade companies to abandon the use of plastic in their products.

Let's consider a game between two companies:

| | Bio materials | Plastic |
|---|---|---|
| Bio materials | $3, 3$ | $1, 5$ |
| Plastic | $5, 1$ | $1, 1$ |

Why are such coefficients realistic? Each company benefits more from using plastic, as it reduces costs and thus increases profits. Suppose that if only one company uses plastic, the state cannot definitively determine which company violated the rules and therefore must conduct an investigation. With a probability $p \in [0, 1]$, it can identify the violator and punish them with a fine (for example, seize all profits). However, if there are two violators, the state will clearly establish that both companies have violated the rules and punish both.

For such an authority to function in the system, we will impose a tax $t$ on companies at the beginning of each period. If one company switches to biodegradable materials while its competitor does not, it will receive a subsidy of $r + t$. But this will only be paid if law enforcement can find the violator.

### 2.3 Solving the problem in general form

Consider the classic Prisoner's Dilemma with the following payoff matrix:

| | Cooperation | Opportunism |
|---|---|---|
| Cooperation | $\beta, \beta$ | $\sigma, \alpha$ |
| Opportunism | $\alpha, \sigma$ | $\gamma, \gamma$ |

The distribution of payoffs should be such that: $\alpha > \beta > \gamma > \sigma$.

Under what parameters do we still have a Prisoner's Dilemma?

«C»: $\beta - t < (p \cdot 0 + (1-p) \cdot \alpha) - t, p < 1 - \frac{\beta}{\alpha}$.

«D»: $p \cdot (\sigma + r) + (1-p) \cdot (\sigma - t) < (p \cdot 0 + (1-p) \cdot \gamma) - t$,

$r + t < \frac{\gamma - \sigma}{p} - \gamma$.

Let's find the expected payoff for players from each of the moves. Let our opponent play the "cooperation" strategy with probability $a$ and "opportunism" with probability $1 - a$.

$\pi^C = a \cdot (\beta - t) + (1-a) \cdot (p \cdot (\sigma + r) + (1-p) \cdot (\sigma - t)) = a\beta + pr + pt + \sigma - t - apt - apr - a\sigma$

$\pi^D = a \cdot (p \cdot (0 - t) + (1-p) \cdot (\alpha - t)) + (1-a) \cdot (p \cdot (0 - t) + (1-p) \cdot (\gamma - t)) = a\alpha - ap\alpha + \gamma - p\gamma - t - a\gamma + ap\gamma$

We want $\pi^C > \pi^D$.

Therefore, $p > \dfrac{a(\alpha - \gamma - \beta) + \gamma - \sigma}{\gamma + r + t + a(\alpha - r - t - \gamma)}$,

$\dfrac{a(\alpha - \gamma - \beta) + \gamma - \sigma}{\gamma + r + t + a(\alpha - r - t - \gamma)} < 1 - \dfrac{\beta}{\alpha} \quad \forall a \in [0, 1], r + t < \frac{\gamma - \sigma}{p} - \gamma$.

Let's normalize the coefficients in the payoff matrix. Let $\beta = 1, \gamma = 0$.

Then the last inequality can be rewritten as:

$\dfrac{a(\alpha - 1) - \sigma}{r + t + a(\alpha - r - t)} < 1 - \dfrac{1}{\alpha} \quad \forall a \in [0, 1], r + t < \frac{-\sigma}{p}$.

Solving this inequality for given $\sigma, \alpha$, we will obtain an estimate for $r + t$.

You can view the solutions depending on the parameters at this link.

## 3 Experiment

One of the key results of our research is that the conditions under which the expected value from the cooperative strategy exceeds that from the opportunistic strategy appear quite simple. Returning to paragraph 2-3, we established that $p < \frac{2}{5}$. This limitation seems natural and realistic.

As for the limitations on tax and subsidy, they depend on their sum: we concluded that $r + t > \frac{3}{2}$. In this regard, an important question arises: how will people behave in real life?

Hypothesis: changes in the parameters $r$ and $t$ will affect people's behavior differently. In particular, an increase of one conditional unit of one parameter may elicit a stronger reaction than a change in the other. We plan to conduct an experiment to test this hypothesis.

### 3.1 Experiment Description

**Our hypothesis:** the subsidy will have a greater impact on individuals. The agent understands that in case of an unfavorable outcome, the state will pay them insurance. Thus, they won't be afraid to make a cooperative move. Experiment description:

1. First game: classic prisoner's dilemma. This allows players to recall the rules, and we can identify how they behave: opportunistically or cooperatively. This ensures there are no illusions that they changed their behavior due to our parameters.

2. Second game: classic prisoner's dilemma with probabilistic checking of opportunistic strategy. If a player plays opportunistically, we can find the violator with probability p (known to all players) and automatically take away their entire winnings (the other player's winnings remain unchanged).

3. Third game: we add taxes to this model, which we collect from players at the beginning of each round (to support law enforcement), and subsidies, which we give to players who were "deceived" in the game.

**Instructions for experiment participants.** It is implemented in the game itself (generally very similar to the classic prisoner's dilemma). Here's an example from one of our games: in this study, you will be randomly paired. Each of you simultaneously and anonymously chooses one of two strategies: cooperate or deceive. Based on your choice and your opponent's choice, you will receive corresponding payouts, which you can see in the payoff matrix. Note that if you choose an opportunistic strategy, there is a probability p=0.4 that it will be detected and your winnings in this round will be zero. This means the risk of cheating can lead to losing your potential winnings in this round. Tax and subsidy: At the beginning of each round, we will collect tax from each player. In the first four rounds, the tax will be 0, and in the next four rounds — 2. If you played cooperatively and your opponent cheated, and we were able to detect it, you will receive a subsidy. In the first four rounds, the subsidy will be 2, and in the last four rounds — 0. In each cell, the amount on the left is your payout, and on the right is for the other participant.

### 3.2 Experiment Results

The game implementation, results, and experiment analysis can be found here: github.com Let's list the percentage of cooperative moves in all rounds of the three games (detailed analysis can be found in files uploaded to GitHub):

In the classic prisoner's dilemma: Percentage of cooperative moves by round: Round 1: 41.7%; Round 2: 25.0%; Round 3: 25.0%; Round 4: 16.7%.

In prisoner's dilemma with probabilistic checking: Percentage of cooperative moves by round: Round 1: 75.0%; Round 2: 75.0%; Round 3: 75.0%; Round 4: 66.7%.

In prisoner's dilemma with probabilistic checking, taxes and subsidies: Percentage of cooperative moves by round: Round 1: 83.3%; Round 2: 83.3%; Round 3: 83.3%; Round 4: 66.7%; Round 5: 50.0%; Round 6: 33.3%; Round 7: 25.0%; Round 8: 50.0%.

In the last four rounds, the percentage of cooperative moves decreased. This can be explained by the fact that we stopped paying subsidies to players, meaning we stopped insuring players in case someone wanted to deceive them.

*Conclusion: the result is visible. In the first game, the percentage of cooperative moves was low (and decreased over rounds). In the second game, the percentage of cooperative moves was high and remained stable throughout the game (only dropping in the last round when one player changed strategy). In the third game, the percentage of cooperative moves was highest in rounds where we paid subsidies to players, confirming our hypothesis. In the final rounds, when we stopped paying subsidies, the percentage of cooperative moves decreased. Thus, we confirmed our hypothesis that player subsidies encourage cooperation. We also demonstrated the effectiveness of our model in practice.*

## 4 Conclusion

In our research, we propose ways to resolve the Prisoner's Dilemma. The first step is to check players for the use of opportunistic strategy. With a set probability $p$, we can identify that a player resorts to opportunism. Upon detection of a violation, the player receives nothing in this turn. If $p > 2/5$, then it is more profitable for each player to apply the cooperative strategy, allowing us to avoid the Prisoner's Dilemma. In our model, we interpret the probability $p$ as analogous to law enforcement actions. Similarly, as taxes are collected to ensure the functioning of law enforcement, introducing a tax $t$ into our model became a natural step to ensure its operation. Players pay a tax $t$ at the beginning of each turn. If we fail to assist a player and they are deceived, we pay compensation $r$. In the third chapter of the article, we define the limitations for the parameters $p, r, t$, under which it becomes more profitable for players to use the cooperative strategy.

We will conclude this article with a discussion of how to further develop the ideas presented in this work.

What can be added to the model? Some consumers are willing to pay extra for recyclable products that minimize harm to nature. This allows us to include demand in the model and calculate profits based on this demand.

It is also worth considering the model from an evolutionary perspective, although it may seem that this will not lead to interesting results. In our model, players choosing cooperative strategies are provided with additional incentives to cooperate.

It may be interesting to study a model in which initially all companies use plastic. Then a law enforcement agency is introduced to the market, which starts with a small probability of punishing dishonest companies. It will be interesting to see how quickly the system transitions to more environmentally friendly companies depending on the number of dishonest players at the initial stage.

## References


[1] Axelrod Robert, The Evolution of Cooperation [CrossRef]

[2] Hardin Garrett, The Tragedy of the Commons[CrossRef]

[3] Ernest Fehr and Simon Gachter, Cooperation and Punishment in Public Goods Experiments [CrossRef]

[4] Van Lange, Rockenbach, and Yamagishi, Reward and Punishment in Social Dilemmas[CrossRef]