Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/55/20231018

Research on the Iterated Prisoner Dilemma Based on Combinations of Two-side Strategies

Jie Zhang[1,a,*]

[1]*International Department, The Affiliated High School of SCNU, Guangzhou, 510665, China*
a. zhangj.joseph2022@gdhfi.com
**corresponding author*

Abstract: The endless prisoner's dilemma game is described as a "primer" using the perspectives of the bilateral parties. Current research focuses on which player responds most effectively, which is a kind of self-protect behavior. Content analysis and mathematical justification were undertaken for the prisoner's dilemma iteration to understand better how to improve the game in a partially competitive and cooperative collaborative setting. This aims to apply the two-side strategies to the real world, especially the political events in which the whole human outcome should be maximized. Participants, who were mostly chosen from game theory specialists from an experiment, provided the decision parameters and rules, in which they have to decide among certain personal strategy choices. The comparison's findings show that two-sided cooperation produces the best results. Still, two-sided inaction results in the poorest, and single cooperation and single noncooperation result from personal sacrifice and personal selfishness, which is a psychological phenomenon.

Keywords: infinite prisoner dilemma, two-side strategies, content analysis

1. Introduction

1.1. Research Background

1.1.1. The Description of the Game

Prisoners A and B are charged with a crime and detained separately, and each can give a confession(cooperation) or deny(noncooperation). Neither prisoner knows what the other will choose to do [1]. The payoff outcomes in each situation are listed in Table 1 [1].

Table 1: Game utility outcomes.

		B	
		Do cooperation	Cannot Cooperation
A	Do Cooperation	-2,-2	-20,0
	Cannot Cooperation	0,-20	-10,-10

DOI: 10.54254/2754-1169/55/20231018

1.1.2. Iterated and Infinite Prisoner Dilemma

The author can even purposefully switch from a single prisoner's dilemma to an iterated prisoner's dilemmaby continually interacting with the same people, a recent trend that has been applied to some sectors [2]. Participants in an iterated prisoner's dilemma, are able to learn about their counterpart's behavioral inclinations, based on the history dependent reciprocity, which distinguishes it from the original concept of a prisoner dilemma.

Particularly, the essay focuses on the infinite prisoner dilemma, a kind of iterated prisoner dilemma with infinite iteration. The purpose of this is to use the characteristic of infinity to make the calculation more representative and comparative.

1.1.3. Future Discount

Usually, future discounts are calculated into future payoffs in infinite games, which means the future payoffs are lower than the current ones because people prefer to be satisfied now [3]. This factor leads to a drop in incentives to cooperate because the sum of infinite payoffs achieved by cooperation may be even lower than those gained by defecting in the initial round. The future discount usually represents a constant between 0 and 1 (use δ to represent it). Thus, the payoff in the nth round is the payoff of the first round times δ^n-1 in a certain circumstance.

1.1.4. Strategies Chosen

In an iterated prisoner dilemma, the participants can gradually encourage collaboration, penalize defection, and treat cheating as noncooperation. Some common strategies have been listed and proved by previous researchers. The majority of strategies that occur in iterated prisoner dilemma are Tit for Tat strategy (TFT), always defection strategy (AD), and grim trigger strategy (GT), but some strategies that are not that commonly chosen will still be studied in the essay, like always cooperation (AC), Win-Stay-Lose-Shift (WSLS), Suspicious Tit for Tat (STFT) [4]. Besides, the Limited Punishment strategy (LP) is also included in the study. Even if there is an endless variety of tactics, this essay still takes some representative choices to figure out the best outcome to show a new perspective of the strategies chosen.

1.1.5. Pros and Cons of Strategies

Players employ the "tit for tat" tactic in the game to punish the noncooperation behavior or reward the cooperation behavior of the other player, which is a method of self-protection. It works out well because of the simplicity of strategy rules, and it is tricky to find the best response except for always cooperation, facing other’s tit-for-tat strategy.

In addition, the Win-Stay-Lose-Shift strategy (which is also called perfect TFT) is advantageous because it can achieve cooperation even if two players have witnessed the dual-noncooperation (fighting), which can be applied in reality well, for instance, countries relieved from World War Two and set up the United Nations.

Similarly, the suspicious Tit for Tat strategy frequently occurs when two oppositions are doubtful of each other, such as the Cold War. It starts with defection in the first round and then repeats what the other did in the last round, so it can avoid one-sided cooperation, which brings the worst outcome for the player, but it damages the cooperation, on the other hand.

The limited punishment strategy is used for punishment for a certain period to the other player’s noncooperation, which enhances the punishment to a certain amount (which is decided by the period of time), but it also brings with a disadvantage that the other can estimate the punishment period, and

DOI: 10.54254/2754-1169/55/20231018

cheat for benefits when the action turns from punishment to cooperation, which means the other earns for the one-off utilities in every punishment period.

The grim trigger strategy is a special type of limited punishment that punishes forever. This kind of strategy is used for frightening, but it usually brings a fatal result to both sides. For example, a nuclear weapon that can destroy the world can be seen as a grim trigger. If the side uses the weapon, there would be no cooperation from then on, which can also be seen as a game over and the bankruptcy of utility accumulation to both sides. However, nuclear weapon exists to avoid unfair or insane behaviors that happen, just like grim trigger contributes to cooperation.

The always defection and always cooperation strategies are similar in the form of permanent actions, which seems like unwiseness because the other players easily use this kind of strategy. However, the study of these strategies is not to research the insane groups. Instead, the strategies can be treated as a situation that both players with exterior powers achieve.

1.2. Research Framework

Recent common research is divided into two groups: new types of strategies in infinite prisoner's dilemma with different parameters of games and an evaluation of the ability to implement strategies based on experimental data regarding the frequency of strategies exists. Generally, the research focuses on one side's behavior toward different environments. However, this essay inspects the two-sided game. Both players can use the given strategy to compete with each other, and both contribute equally to outcomes. The originality of the study is not to maximize one player's benefit. Instead, it is to achieve the best situation for the whole game.

Except for the angle of view, the study is conducted with pure calculation instead of real experiments because the proof of theories is more important and economical. Therefore, the study is based on some parameters and variants that are the most well-known and recommended in iterated prisoner dilemma and finding the relationships between parameters and how they result in the best outcome for the game.

2. Method

2.1. Mathematical Calculations

Instead of experiments and real subjects playing the games, the study is based on pure mathematical calculations. The outcome of each situation of each round is given in Table 1. A future discount value is represented with the symbol: 'δ' ($0<\delta<1$). 'U' represents the sum of infinite payoffs of 'player A' (account that the action and the payoff outcomes are symmetric for both players; the calculation of a single player is representative of the whole game).

Thus, among the given set of strategies, players A and B choose one of them, and they are matched with each other. The real actions in each round will be figured out, and the total payoffs will be calculated and compared to find the best outcome with the relationships between parameters.

2.2. Content Analysis

To display the single outcome of every round and provide the materials for infinite payoffs calculation, each player's two strategies are matched and played by content analysis. This is based on the hypothesis that both players decide to use regular strategies in the same round. However, in real life, people may change their strategies or start to play a certain strategy irregularly, which makes the situation extremely complicated. The research simplifies the decisions and can be brought to every single round of real-life games, but not for the long-term strategies chosen because the environment

DOI: 10.54254/2754-1169/55/20231018

is influenced by macro or micro factors. Overall, the content analysis results in specific action performances in each round.

3. Results

3.1. The Permutation and Combination of Two-side Strategies

There are 49 combinations of the two players' strategies (seen in Table 2).

Table 2: Combinations of two-side strategies.

Player A/Player B	TFT	WSLS	STFT	GT	LP	AD	AC
TFT	$U^*_{1.1}$	$U^*_{1.2}$	$U^*_{1.3}$	$U^*_{1.4}$	$U^*_{1.5}$	$U^*_{1.6}$	$U^*_{1.7}$
WSLS	$U^*_{2.1}$	$U^*_{2.2}$	$U^*_{2.3}$	$U^*_{2.4}$	$U^*_{2.5}$	$U^*_{2.6}$	$U^*_{2.7}$
STFT	$U^*_{3.1}$	$U^*_{3.2}$	$U^*_{3.3}$	$U^*_{3.4}$	$U^*_{3.5}$	$U^*_{3.6}$	$U^*_{3.7}$
GT	$U^*_{4.1}$	$U^*_{4.2}$	$U^*_{4.3}$	$U^*_{4.4}$	$U^*_{4.5}$	$U^*_{4.6}$	$U^*_{4.7}$
LP	$U^*_{5.1}$	$U^*_{5.2}$	$U^*_{5.3}$	$U^*_{5.4}$	$U^*_{5.5}$	$U^*_{5.6}$	$U^*_{5.7}$
AD	$U^*_{6.1}$	$U^*_{6.2}$	$U^*_{6.3}$	$U^*_{6.4}$	$U^*_{6.5}$	$U^*_{6.6}$	$U^*_{6.7}$
AC	$U^*_{7.1}$	$U^*_{7.2}$	$U^*_{7.3}$	$U^*_{7.4}$	$U^*_{7.5}$	$U^*_{7.6}$	$U^*_{7.7}$

3.2. The Action Performance of Each Combination

'C' represents confession(noncooperation), and 'D' represents denying(cooperation), and here is the result:

If both players choose Tit for Tat, the performance would be like they cooperate forever because they both start with cooperation. They repeat cooperation following what the other player does in the last round, which achieves a positive reciprocity: (D, D) (D, D) ...

Similarly, the other combinations would act like Table 3.

Table 3: Action performances of combinations.

A/B	TFT	WSLS	STFT	GT	LP	AD	AC
TFT	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(C,D).....	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(C,C)(C,C).....	(D,D)(D,D).....
WSLS	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(C,D).....	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(C,C)(D,C).....	(D,D)(D,D).....
STFT	(C,D)(D,C).....	(C,D)(D,C)......	(C,C)(C,C).....	(C,D)(D,C)(C,C).....	(C,D)(D,C)(C,C).....*	(C,C)(C,C).....	(C,D)(D,D).....
GT	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(C,D)(C,C).....	(D,D)(D,D).....	(D,D)(D,D)....	(D,C)(C,C)(C,C).....	(D,D)(D,D).....
LP	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(C,D)(C,C).....*	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(C,C)(C,C).....*	(D,D)(D,D).....
AD	(C,D)(C,C)(C,C).....	(C,D)(C,C)(C,C).....	(C,C)(C,C)...	(C,D)(C,C)(C,C).....	(C,D)(C,C)(C,C).....*	(C,C)(C,C).....	(C,D)(C,D).....
AC	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(D,D)(D,D)...	(D,D)(D,D).....	(D,D)(D,D).....	(D,C)(D,C).....	(D,D)(D,D).....

*The symbol means the last action performance is repeated for k periods

3.3. Calculations of PlayerA's Total Payoffs in Each Combination

Here is the result of the total payoffs of player A, seen in Table 4.

The number of combinations is too large, so only two calculation examples are given to show the process.

$U_{1.1}$= - 2 - 2*δ - 2*δ^2 - ... - 2*δ^n = $\frac{2}{\delta-1}$

DOI: 10.54254/2754-1169/55/20231018

$U_{1.3}$= - 20 - 20*δ^2 - ... - 20*δ^2n = $\frac{20}{δ^2-1}$

(n means infinity, and the calculation uses the geometric progression summing formula)

Table 4: Total payoffs of PlayerA in each combination.

A/B	TFT	WSLS	STFT	GT	LP	AD	AC
TFT	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{20}{\delta^2-1}$	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{20-10\delta}{\delta-1}$	$\frac{2}{\delta-1}$
WSLS	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{20}{\delta^2-1}$	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{20+10\delta}{\delta^2-1}$	$\frac{2}{\delta-1}$
STFT	$\frac{20\delta}{\delta^2-1}$	$\frac{20\delta}{\delta^2-1}$	$\frac{10}{\delta-1}$	$\frac{20-30\delta^2}{\delta-1}$	$\frac{-10\delta^{(k+2)}-10\delta^2+20\delta}{\delta-1}$	$\frac{10}{\delta-1}$	$\frac{2\delta}{\delta-1}$
GT	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{10\delta^2-20\delta+20}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{10\delta}{\delta-1}$	$\frac{2}{\delta-1}$
LP	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{-10\delta^{(k+2)}-10\delta^2-20\delta+10}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{-10\delta^{(k+1)}-10\delta+20}{\delta-1}$	$\frac{2}{\delta-1}$
AD	$\frac{10\delta}{\delta-1}$	$\frac{10\delta}{\delta-1}$	$\frac{10}{\delta-1}$	$\frac{10\delta}{\delta-1}$	$\frac{10\delta^{(k+1)}+10\delta}{\delta-1}$	$\frac{10}{\delta-1}$	0
AC	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{-18\delta+20}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{2}{\delta-1}$	$\frac{20}{\delta-1}$	$\frac{2}{\delta-1}$

In conclusion, 25 payoffs are ($\frac{2}{δ-1}$). 4 payoffs are ($\frac{10}{δ-1}$), which means player A's pain times 5 because the payoffs are negative. 4 payoffs are ($\frac{10δ}{δ-1}$). One payoff is ($\frac{20}{δ-1}$). One payoff is ($\frac{2δ}{δ-1}$). Among those payoffs, $\frac{2δ}{δ-1} < \frac{2}{δ-1} < \frac{10δ}{δ-1} < \frac{10}{δ-1} < \frac{20}{δ-1} < 0$.

In addition, 2 payoffs are ($\frac{20}{δ^2-1}$). 2 payoffs are ($\frac{20δ}{δ^2-1}$). There are 10 special forms of payoffs: ($\frac{20-10δ}{δ-1}$), ($\frac{20+10δ}{δ^2-1}$), ($\frac{20-30δ^2}{δ-1}$), ($\frac{-10δ^{(k+2)}-10δ^2-20δ+20}{δ-1}$), ($\frac{-10δ^{(k+1)}-10δ+20}{δ-1}$), ($\frac{10δ^{(k+1)}+10δ}{δ-1}$), ($\frac{-18δ+20}{δ-1}$), ($\frac{10δ^2-20δ+20}{δ-1}$), ($\frac{-10δ^{(k+2)}-10δ^2+20δ}{δ-1}$), 0.

Obviously, the highest payoff is 0, which is the situation where Player A uses AD, and Player B uses AC, and the only circumstance applied to the strategies in the real world is the love of human beings. One side of the relationship claims, and the other lives forever.

3.4. Calculations of Sum of Players' Payoffs in Each Combination

Subsequently, the sum of the payoffs of the two players can be figured out in Table 5.

Table 5: The sum of the payoffs of the two players.

A/B	TFT	WSLS	STFT	GT	LP	AD	AC
TFT	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{20+20\delta}{\delta^2-1}$	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{20}{\delta-1}$	$\frac{4}{\delta-1}$
WSLS	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{20+20\delta}{\delta^2-1}$	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{20+20\delta+10\delta^2}{\delta^2-1}$	$\frac{4}{\delta-1}$
STFT	$\frac{20+20\delta}{\delta^2-1}$	$\frac{20+20\delta}{\delta^2-1}$	$\frac{20}{\delta-1}$	$\frac{40-20\delta-20\delta^2}{\delta-1}$	$\frac{-20\delta^{(k+2)}-20\delta^2+20}{\delta-1}$	$\frac{20}{\delta-1}$	$\frac{-16\delta+20}{\delta-1}$
GT	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{40-20\delta-20\delta^2}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{20\delta}{\delta-1}$	$\frac{4}{\delta-1}$
LP	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{-20\delta^{(k+2)}-20\delta^2+20}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{20}{\delta-1}$	$\frac{4}{\delta-1}$
AD	$\frac{20}{\delta-1}$	$\frac{20+20\delta+10\delta^2}{\delta^2-1}$	$\frac{20}{\delta-1}$	$\frac{20\delta}{\delta-1}$	$\frac{20}{\delta-1}$	$\frac{20}{\delta-1}$	$\frac{20}{\delta-1}$
AC	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{-16\delta+20}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{4}{\delta-1}$	$\frac{20}{\delta-1}$	$\frac{4}{\delta-1}$

Among those results, those with bigger numbers in numerators are worse at acting as a cooperation choice. $\frac{5}{7}$ of AD choices, whoever makes this strategy has the payoff of ($\frac{20}{\delta-1}$), which brings the highest pain to players. Overall, AD is harmful to cooperation and makes the utilities of the game worse.

4. Discussion

4.1. The Pros and Cons of Content Analysis Used

In this research, the content analysis focuses on a few strategies proved in the experiment, in which the ratio and rigorousness part comes from how the experiment was conducted in a way with real experimental subjects, but it also comes with a certain extent of randomness [4].

Anyway, the research lacks a full-sided analysis of the strategies chosen; for example, an experiment researching the effective choices based on a computer tournament with theories of several sides illustrates the result of two-sided cooperation, another experiment type [5].

Future research can be conducted in a similar or improved method, which discusses other effective choices in real life or the economic world [6].

4.2. The Pros and Cons of Mathematical Calculations Used

The mathematical rigorousness comes from pure mathematical calculations with fixed parameters and situations.

However, the numbers set are specific, lacking comprehensiveness and universality. Future research can use appropriate logarithms to make a set of data based on different parameters and one-round utility, for which human-force calculation is inefficient. In addition, the benefit comparison can be done with computer simulations or real experiments in which subjects evaluate their strategies, which are a kind of sub-proof to mathematical theories.

4.3. Inefficient Use of Final Data

The comparisons in this research do not fully use the data because of the inefficient human-force calculation, so future research can apply statistical tools to the research, and make an all-round conclusion based on the data.

4.4. The Influence of Historical Reciprocity

Additionally, behavior biases have an impact on players' judgments [7]. Some people and communities have cultivated psychological and behavioral biases over time, including a higher degree of interpersonal trust, an emphasis on the long term in interactions, and a propensity for either positive or negative reciprocity in cooperative or deviant behavior. The effects of reciprocity depend on historical context. These traits may evolve over time as a result of natural selection within a society or group selection among rival civilizations [8].

In actuality, they persuade crowds of people to "irrationally" choose solutions that are actually the best for everybody. In the game, players may change their attitudes and strategies as time passes, based on the other player's strategies chosen. This research only discusses a single strategy used in a certain period. Future research can combine different strategies and study how they change with the historical reciprocity.

4.5. Real-life Situations

In real life, certain circumstances seem like infinite prisoner dilemmas, including political events (nuclear wars), environmental pollution, and negative economic competition. In those cases, players try to escape from infinite prisoner dilemmas because the payoffs are usually negative, so achieving two-sided cooperation is essential.

In the real world, the third party can change the game with some methods. For instance, the government changes the incentives that different decision-makers confront. Many prisoners' issues can be resolved more cooperatively when cooperative behavior is enforced through reputation, norms, regulations, democratic or other group decision-making processes, and explicit social punishment for defections [9, 10].

5. Conclusion

In conclusion, the research makes a hypothesis of two-side strategies in an infinite prisoner dilemma game, including Tit for Tat strategy (TFT), always defection strategy (AD), and grim trigger strategy (GT), always cooperation (AC), Win-Stay-Lose-Shift (WSLS), Suspicious Tit for Tat (STFT) and Limited Punishment strategy (LP). With a future discount and certain parameters given by another paper, the research conducts content analysis to get a table of how players react in each round with two-side strategies. For the method part, content analysis and mathematical calculations are chosen to focus on the situation of a single round and make the comparison more rigorous. Mathematical calculations deal with each player's payoffs, and the comparisons indicate that two-sided cooperation achieves the best result for the game. In contrast, two-sided noncooperation leads to the worst case. The one-side noncooperation and one-side cooperation make one of the player's payoffs the highest, which seems to be a type of love and sacrifice in real life. In the discussion part, some of the research flaws are listed, and the future research path is enlightened, including the full use of data, other methods, the diversity of parameters, and statistical tools. In addition, some specific factors that may be important are mentioned, such as historical reciprocity and behavioral biases, which can be discussed in the future. Moreover, some real-life applications and examples are listed to show the comprehensive usefulness of the infinite prisoner dilemma game, including the game between nations, business, and personal careers.

The infinite prisoner dilemma can be discussed in different aspects of two-side or multiple-side strategies, which make the game more complex. Over the time, entrepreneurs take the advantage of reciprocity and biases, which leads to award of cooperation or the disadvantage of noncooperation.

References

[1] *Kuhn, Steven, "Prisoner's Dilemma", The Stanford Encyclopedia of Philosophy (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2019/entries/prisoner-dilemma/>.*

[2] *Industrial organization (see Friedman 1971, Green and Porter 1984 and Rotemberg and Saloner 1986), informal contracts (Klein and Leffler 1981), theory of the firm (Baker et al. 2002), public finance (Phelan and Stacchetti 2001) and macroeconomics (Rotemberg and Saloner 1986 and Rotemberg and Woodford 1990)*

[3] *Lyon, K. S. (1996). Why economists discount future benefits. Ecological Modelling, 92(2–3), 253–262.*

[4] *Dal Bó, P., & Fréchette, G. R. (2019). Strategy Choice in the Infinitely Repeated Prisoner's Dilemma. The American Economic Review, 109(11), 3929–3952.*

[5] *Axelrod, R. (1980). Effective Choice in the Prisoner's Dilemma. Journal of Conflict Resolution, 24(1), 3–25.*

[6] *Knoblauch, V. (1994). Computable strategies for repeated prisoner′s dilemma. Games and Economic Behavior, 7(3), 381–389. https://doi.org/10.1006/game.1994.1057*

[7] *Wikimedia Foundation. (2023a, August 12). Reciprocity (social psychology). Wikipedia. https://en.wikipedia.org/wiki/Reciprocity_(social_psychology)*

[8] *The cognitive problem of the behavioral decision theory through game theory: Biases and heuristics. (2021). Journal of Applied Business and Economics, 23(3). https://doi.org/10.33423/jabe.v23i3.4348*

[9] *Xiong, B. Q. (2021). How to help parents overcome the "prisoner's dilemma" of burden reduction Online Learning (10), 79. doi: CNKI: SUN: XOXZ. 0.2021-10-028*

[10] *Wang, P. P. (2022). Breaking through the Prisoner's Dilemma in Corporate Culture Human Resources (07), 65-67. doi: CNKI: SUN: RLZY.0.2022-07-018*