UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
INSTITUTO TÉRCIO PACITTI DE APLICAÇÕES E PESQUISAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

THIAGO MENDES DE MELO

ON THE COLLUSION CONDITIONS IN THE INDEFINITELY REPEATED PRISONERS' DILEMMA UNDER DIFFERENT DECISION CRITERIA – A COMPUTATIONAL APPROACH

Rio de Janeiro
2016

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
INSTITUTO DE MATEMÁTICA
INSTITUTO TÉRCIO PACITTI DE APLICAÇÕES E PESQUISAS COMPUTACIONAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

THIAGO MENDES DE MELO

ON THE COLLUSION CONDITIONS IN THE INDEFINITELY REPEATED PRISONERS' DILEMMA UNDER DIFFERENT DECISION CRITERIA – A COMPUTATIONAL APPROACH

Dissertation submitted in partial fulfillment of the requirements for the degree of Master (Computer Science) in Informatics Graduate Program, Universidade Federal do Rio de Janeiro (PPGI-UFRJ).

Advisor: Eber Assis Schmitz, Ph.D

Rio de Janeiro
2016

M528 Melo, Thiago Mendes de

On the collusion conditions in the indefinitely repeated Prisoners' Dilemma under different decision criteria – a computational approach / Thiago Mendes de Melo. – 2016.
66 f.: il.

Dissertation (Master in Computer Science) – Universidade Federal do Rio de Janeiro, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Programa de Pós-Graduação em Informática, Rio de Janeiro, 2016.

Advisor: Eber Assis Schmitz, Ph.D.
1. Game Theory. 2. Iterated Prisoners' Dilemma. 3. Decision Theory. 4. Collusion. – Dissertation. I. Schmitz, Ph.D, Eber Assis (Advisor). II. Universidade Federal do Rio de Janeiro, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Programa de Pós-Graduação em Informática. III. Título

CDD

THIAGO MENDES DE MELO

On the collusion conditions in the indefinitely repeated Prisoners' Dilemma under different decision criteria – a computational approach

Dissertation submitted in partial fulfillment of the requirements for the degree of Master (Computer Science) in Informatics Graduate Program, Universidade Federal do Rio de Janeiro (PPGI-UFRJ).

Rio de Janeiro, _____ of ______________, _________.

____

Prof. Eber Assis Schmitz, Ph.D (Advisor)

____

Prof. Antonio Juarez Alencar, D.Phil

____

Prof. Márcio de Oliveira Barros, D.Sc

____

Prof. Denis Silva da Silveira, D.Sc

AGRADECIMENTOS

Gostaria de agradecer muito ao meu orientador pelo seu incentivo e seu otimismo, que foram de extrema importância durante todo meu percurso no mestrado. Nada do que está escrito nestas páginas teria sido possível se não fosse pela sua orientação.

Agradeço muito a meus familiares e amigos cujo apoio foi imprescindível para que eu pudesse chegar até aqui. Seu incentivo é o que me ajuda e inspira para superar os desafios de qualquer jornada.

RESUMO

MELO, Thiago Mendes de. **On the collusion conditions in the indefinitely repeated Prisoners' Dilemma under different decision criteria – a computational approach**. 2016. 64 f. Dissertação (Mestrado em Informática) - PPGI, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

A competição entre empresas é uma parte essencial da operação dos mercados, proporcionando inovação, produtividade e crescimento, que geram benefícios, criação de riqueza e redução da pobreza. Competição, entretanto, tende a reduzir os lucros totais das empresas. Como consequência, empresas tentam chegar a acordos através de um comportamento chamado conluio. Embora um conluio possa aumentar os lucros de empresas rivais, algumas delas podem explorá-lo em benefício próprio, ao custo das demais empresas que obedecem o conluio. Para desencorajar esse tipo de atitude, empresas em um conluio frequentemente fazem ameaças de retaliação contra quaisquer desvios praticados pelas empresas participantes.

O estudo das condições em que conluios podem ser alcançados usando o Dilema dos Prisioneiros Iterado (IPD) é de grande importância, não apenas para as empresas rivais, mas também para as agências reguladoras. As condições para formação de conluios encontradas na literatura consideram apenas um critério de racionalidade, que consiste em escolher a consequência que traz a maior utilidade esperada dos ganhos. Entretanto, condições para formação de conluio dependem fortemente das racionalidades dos tomadores de decisão, que não estão limitadas à maximização da utilidade esperada.

Esta dissertação tem dois objetivos: (1) propor um novo modelo para determinar se a formação de um conluio é possível no IPD considerando diferentes critérios de decisão além da utilidade esperada para mostrar, usando um conjunto de exemplos, que as condições para formação de um conluio dependem fortemente dos critérios de decisão utilizados pelos agentes; (2) mostrar que a aplicação de tais critérios requer, na maior parte das vezes, cálculos muito elaborados, tornando-os quase inacessíveis para o raciocínio humano apenas. Para utilizá-las com sucesso é necessário auxílio de ferramentas computacionais. Como subproduto desta pesquisa, uma ajuda computacional baseada na linguagem R foi criada para auxiliar no processo de seleção das melhores ações a serem tomadas de acordo com cada critério.

Palavras-chave: teoria dos jogos. dilema dos prisioneiros iterado. teoria da decisão. conluio

ABSTRACT

MELO, Thiago Mendes de. **On the collusion conditions in the indefinitely repeated Prisoners' Dilemma under different decision criteria – a computational approach**. 2016. 64 f. Dissertation (Master in Computer Science) - PPGI, Instituto de Matemática, Instituto Tércio Pacitti de Aplicações e Pesquisas Computacionais, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

Competition among firms is an essential part of the operation of markets, providing innovation, productivity and growth, all of which generate benefits, wealth creation and reduction of poverty. Competition, on the other hand, tends to minimize firms' total profits. As a consequence, firms try to engage in agreements in a behaviour called collusion. Although collusion can increase competing firms profitability, it can also be exploited by one of the competitors to its own advantage at the expense of the other collusion obeying companies. To discourage this kind of behaviour, colluding firms often pose threats to retaliate deviations by participant companies.

The study of the conditions on which collusive behaviour can be achieved using the Iterated Prisoners' Dilemma, or IPD, turns out to be of great importance, not only for the competing firms but also for regulating agencies. The conditions for collusion formation found in the literature consider only one rationality criterion, which consists in choosing the consequence that brings the largest expected utility of the outcomes. However, conditions for collusion are highly dependant on decision makers' rationalities, which are not limited to expected utility maximization.

This dissertation has two objectives: (1) propose a new model for determining whether collusion is possible in the IPD considering decision criteria other than just expected utility to show, by using a set of examples, that the collusion conditions are highly dependent on the rational decision criteria used by the agents; (2) to show that the application of these criteria require, most of the time, rather elaborate computations, making them almost inaccessible to human reasoning alone. Their successful application requires support from computational tools. As a byproduct of this research, a computational aid, based on the R language, was created to help the process of selecting the best actions according to each criteria.

Keywords: game theory. iterated prisoners' dilemma. decision theory. collusion

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

PD	Prisoners' Dilemma
IPD	Iterated Prisoners' Dilemma
FSD	First-order Stochastic Dominance
SSD	Second-order Stochastic Dominance
RSD	Risk-seeking Stochastic Dominance
CPT	Cumulative Prospect Theory

CONTENTS

1 INTRODUCTION

Competition among firms is an essential part of the operation of markets, providing innovation, productivity and growth, all of which generate benefits such as wealth creation and reduction of poverty. Economic models of oligopoly competition, such as Cournot-Nash and Bertrand models [24], predict that, under circumstances of perfect competition, firms total profit in equilibrium is lower than the monopoly profits. This is due to the fact that maximization of profits of the competing firms is achieved by the reduction of the competitor's profit[6]. As a consequence, firms try to engage in agreements, such that all competitors could increase their profits, usually at the consumer's expense. This type of behaviour is called collusion. Since this kind of activity is harmful for consumers, regulation agencies provide many rules intended to prevent price fixing and cartel formation (such as antitrust laws and competition policy).

The oligopoly competition models, however, do not take into account the (frequently) collusive behaviour that may arise among firms in the market [6]. We can cite examples of collusion that are still found in some markets, such as in the oil industry with OPEC – an example of explicit cartel – or in the danish ready-mixed concrete market – an implicit cartel [1].

Although collusion can increase competing firms profitability, it can also be exploited by one of the competitors to its own advantage at the expense of the other collusion obeying companies. This is due to the fact that one firm may feel tempted to lower its prices in order to obtain competitive advantage over its rivals, who will be setting collusive, and, therefore, higher prices. To discourage this kind of behaviour, colluding firms often pose threats to retaliate deviations by participant companies. Examples of threats are starting a price war or flooding the market, both of which will reduce the profits of the deviant company [6].

Collusive behaviour can be achieved provided that the threats are both credible and sufficiently severe and that the value of the future payoffs is not discounted too much [27]. This is due to the fact that firms in a market compete on a daily basis and during indefinitely large periods of time. Under these circumstances, deviation would lead to harmful financial consequences, while colluding could be beneficial for them. The study of the conditions on which collusive behaviour can be achieved (and maintained) turns out to be of great importance, not only for the competing firms but also for regulating agencies trying to impede the occurrence of this type of situation. The theoretical conditions for collusion can be studied using tools supplied by game theory and decision theory (such as dominance, Nash equilibrium and expected utility [22, 21, 16]).

Game theory is the study of mathematical models of interactions that aims to understand how rational agents behave in situations in which their decisions are interdependent [22, 20]. In economics, it plays a big role in helping to understand competitive strategy, which includes situations of competition and collusion.

The Prisoners' Dilemma (PD) is probably the most studied case in Game Theory. This game and its variations, mainly the Iterated Prisoners' Dilemma or IPD, can be used to model collusive situations like the ones previously described, such as cartel formation, R&D and advertisement investment. The application of game theory concepts to these models makes it possible to determine the conditions that make collusion possible (or not). The IPD will be the model for all collusion situations presented in this study.

A very simple example can help understanding the situations where firms do make collusion agreements. Imagine that two supermarkets compete in the same neighbourhood, they share the local market equally and have equivalent market power. To attract more customers, either may offer discounts in its products, which will earn it more profit by increasing its market share as long as it is the only one to do so. If the two use the same strategy, however, they'll be cutting down their profit margin while keeping the same market share, which will result in loss to both, thus making the cartel a reasonable option. To discourage deviations from the cartel, each supermarket threatens to start a permanent price war if discounts are ever offered. Because the game is repeated indefinitely, the payoffs of the firms can be described as probability density functions (PDF). But how does one know if a collusion is possible in this situation?

One can also look at the collusion formation matter as a decision problem for the competing companies, as a choice between the best of two the alternatives, either to collude or not. A generic decision problem D can be modeled as

$$D : A \times B \to C \tag{1.1}$$

in which A is the set of possible actions the decision maker may take, B is the set of possible states of the world, and C is the set of possible consequences [5].

The decision maker is supposed to choose an action in A such that, depending on the state of the world B will lead to the desired outcome in C. If the state of the world is known at the time of decision, then the decision problem is trivial: choose the a_i that provides the best consequence. But when the decision must be made ex-ante, the problem is classified as a decision under uncertainty or decision under risk depending the type of the knowledge available about the future states.

The decision maker is defined to be a rational person when he/she makes choices that are always consistent with his/her preferences among the possible outcomes[5]. This implies in the existence of a preference order relation in the set C. But, how does one define the preference order of the outcomes when they are uncertain? Or, formally, how does one define a preference order when each consequence is a random variable? Decision Theory provides several criteria (methods or rules for basing decisions on) that can be used to choose among different alternatives [9], both under uncertainty and under risk. So how does one define if a collusive engagement is possible under each different criteria?

The conditions for collusion formation found in the literature consider only one rationality criterion, which consists in choosing the consequence that brings the largest expected utility of the PDF of the consequences [27, 6, 34]. Nothing exceptional to this fact, since von Neumann proved that, under some conditions, the criterion of maximum expected utility is the optimal choice for a rational agent [22]. On the other hand, the criterion can be difficult (or impossible) to use, because it relies on the knowledge of the preference curve of all agents, which is rarely the case in real world situations [16]. Also, the probability distribution of the payoffs can be quite skewed, which may also make this metric of limited help to the decision maker whose decisions take into account a risk perspective [28].

The first objective of this dissertation is to propose a new model for determining whether collusion is possible in the IPD. A model considering decision criteria (methods of choosing among alternatives) other than expected utility. We show, through a set of examples, that the collusion conditions are highly dependent on the decision criteria used by the agents. Although these conditions are dependent on very personal views of the world, provided that the agent uses them in a consistent manner, they must also be considered as rational. As a result, in the same situation, one criterion would lead to collusion and another to defection. The second objective is to show that the application of these criteria requires, most of the time, rather elaborate computations, making them almost inaccessible to human reasoning alone. Their successful application requires support from computational tools. As a byproduct of this research, a computational framework, based on the R language, was created for helping the process of selecting the best actions according to each criteria.

This dissertation is organized as follows: section 2 describes the conceptual framework, presenting an overview of Game Theory concepts used in the analysis presented here, the formal definition of an IPD and the decision methods utilized in this study; section 3 presents an example of an IPD situation, shows the solutions based on the different decision criteria presented and shows the computational approach used to achieve the results presented in this study; section 4 compares our study to its related works; finally, section 5 presents the conclusions and final remarks.

2 CONCEPTUAL FRAMEWORK

2.1 GAME THEORY

2.1.1 Fundamentals of Game Theory

Game Theory can be shortly defined as the study of mathematical models designed for analysing situations in which rational individuals make decisions that influence one another's welfare, in other words, they're mutually interdependent. These situations are called *games*, the individuals are called *players* and the main goal of game theory is to understand players' behaviour in the games [20, 27].

The complete definition of a game requires the knowledge of: who the players are, what are the actions available for each player, the payoffs for each possible combination of choices made by the players and, sometimes, the chronological order of the actions taken. Thus, a strategic game model contains the basic elements as follows:

Definition 1 (Players) $N = \{P_1, P_2, \ldots, P_n\}$ *where* N *is a finite set of players. The players in a game are decision-maker agents whose behaviors dictate their choices. In order to establish interaction between strategies, or mutual interdependence, at least two players are required to compete in a game.*

Definition 2 (Strategies) $S_i = \{s_1, s_2, \ldots, s_n\}$ *is the finite set of pure strategies for player* i*. Let* $S_i' = \prod(S_i)$ *be the finite set of mixed strategies for players except* i*. If* $\prod(X)$ *be the set of all probability functions over a set* X*, then* S_i' *is the set* *of mixed strategies for player i. A strategy is an entire description of how a player could behave in response to what can be observed about its opponents' actions during a game. A game is considered symmetric when the set of pure strategies is identical for all players, and the payoff to playing a given strategy depends only on the strategies being played, not on who plays them.*

Definition 3 (Strategy profiles) *Let* $S_p = \{S_1 \times S_2 \times \cdots \times S_n\}$ *be she set of strategy profiles. The set* S_p *of strategy profiles is the Cartesian product of the* $S_i^{\prime}s$*, consisting of every possible game combination.*

Definition 4 (Utility function) *Then* $(u_i : S_p \rightarrow \mathbb{R})$ *is the utility function for player i. An utility function represents a player's preference relation over each strategy profile. Given that this study involves monetary rewards as outcomes, e.g. profits, a payoff function is suitable to evaluate them. In general, rational players are induced and stimulated to prefer higher payoffs, and naturally choose a strategy profile that promotes the maximization of their utilities.*

A basic assumption of Game Theory is that *players* are rational, meaning that they are supposed to act in their own self-interest with the objective of maximizing the expected utility of their own payoffs and not caring about the payoffs of others [27]. This is the rationality principle described by maximization of utility [22].

Determining the best strategy in each situation, however, depends on the context or essence of the game, which may involve cooperation or competition. When cooperation exists, players can assume binding and enforceable commitments with one another, so that real intentions of each individual are flagged and formalized. By contrast, in a non-cooperative game, players need to conceal their strategies from each other, making them individualistic by nature and thus resembling the real world competition environment between companies [27].

2.1.2 Static Games and the Prisoners' Dilemma

In a static game, decisions are made in isolation, that is, without knowledge of others' decisions. Since players do not have information about which strategy has been adopted by their opponents, these are, inherently, games of imperfect information [27]. In addition, these games are predominantly represented in normal (matrix) form, since the amount of information available to players is constant within a game (no information is obtained during its course), and timing of decisions has no effect on players' choices [27]. A common interpretation of a static game is that it models a unique event, also called stage game or one-off game [27].

The widely discussed static game called "The Prisoners' Dilemma" describes a game with prison sentence rewards (thus the reason for its name) [26]. The dilemma's description is given as follows: two members of a criminal gang are arrested and placed under solitary confinement, unable to communicate with each other. The police doesn't have enough evidence to convict them on the main charge and thus the officer makes the same offer to each prisoner, saying that if he/she accuses his/her partner then he/she will be set free while the accused will spend three years in jail on the main charge. However, if both accuse each other, the two will be sentenced to two years in prison. On the other hand, if both stay silent, they will be be sentenced to one year in jail on a minimum charge. So the prisoners need to choose between accusing their partner or staying silent.

This game can be described by its players, their possible actions (accuse, stay silent) and the payoffs for each situation, and can be represented in the form of a table such as in Table 2.1. This is called a *normal form* representation [27].

		Prisoner 2	
		Staying Silent	Accusing
Prisoner 1	Staying Silent	-1, -1	-3, 0
	Accusing	0, -3	-2, -2

Table 2.1: The Prisoners' Dilemma in Normal Form.

The game table shows in each cell the number of years in jail (negative value of the sentences being imposed) with the first number being outcome for the first prisoner followed by the outcome for the second. For example, if both testify against each other, the outcome will be two years in prison (-2) for each one of them, while unilateral betrayal gives the accuser a payoff of liberty (0) whilst the accused has to spend three years (-3) in prison.

Assuming that both prisoners are rational, each one will reason as follows: "my partner has only two choices, if he chooses to give me up to the police (accusing), I will be better off turning him in as well (accusing); on the other hand, if he chooses to keep quiet (staying silent), I will still be better off by turning him in (accusing). So, my best decision, independently of what he does, is to accuse him". The same chain of reasoning occurs to the second prisoner and so the predicted behavior to the dilemma is both prisoners accusing each other and thus spending two years in jail each.

For a game to be a Prisoners' Dilemma it has to be configured such as shown in Table 2.2, in which *Collusion* and *Defection* are the possible strategies and the payoffs T for *temptation*, R for *reward*, P for *punishment* and S for *sucker*. The relation between these payoffs must be $T > R > P > S$, although it is also acceptable if $T \geq R > P > S$ or yet $T > R > P \geq S$ [14].

		Player 2	
		Collusion	Defection
Player 1	Collusion	R, R	S, T
	Defection	T, S	P, P

Table 2.2: Symmetric game of the Prisoners' Dilemma class.

2.1.3 Solution of a Game

A solution to a game is a way of predicting how it will be played and giving its result through a description of the strategies adopted by all players [27]. It is a systematic description of the outcomes that may emerge in a game [25]. There are some different ways of finding solutions to games.

As observed in the prisoners' reasoning in section 2.1.2, since *accusing* always gives a better payoff than *staying silent* despite the other player's strategy, the former is said to be a strictly dominant strategy whilst the latter is said to be strictly dominated.

In certain types of games, however, the strict dominance concept may not predict a unique solution when strictly dominated strategies are non-existent, thus iterated elimination cannot be employed. On the other hand, every game with a finite number of players that may choose from among a finite number of strategies has at least one Nash Equilibrium [21]. Players are in Nash equilibrium if they are choosing their best responses in relation to each other players' choices and no player has incentive to change his/her strategy. In other words, no player can be made better-off by changing his/her strategy unilaterally [27].

In a Prisoners' Dilemma Game, the Nash Equilibrium is that none of the players (given that both are rational) will collude with each other, which makes this situation not Pareto efficient, because both could have better payoffs if they both changed strategies and colluded. But for that to happen, they would have to choose a strategy that is dominated and thus risk having serious losses in their payoffs if the other player defects. That's why it is called a dilemma.

2.1.4 Dynamic Games and the Iterated Prisoners' Dilemma

Games can be dynamic if at least one player is able to observe the actions of others and thus condition his/her strategy accordingly. A game can be dynamic if decisions are not simultaneous and some player(s) know others' choices before having to move, or if a static game is repeated a number of times and players can learn from their previous interactions. Being able to conditioning optimal actions on what other players have done in the past is an essential feature of all dynamic games [27].

Another aspect of dynamic games is that players can pose threats or make promises as a form of trying to prevent undesired behaviour from other players. When considering this kind of maneuver, it's fundamental that these threats and promises are credible or believable, meaning that it's in the player's interest to carry it out at the appropriate time. Therefore, credibility is another important concept, since rational players will only believe in credible statements [27].

Situations like the one in the Prisoners' Dilemma, in which the equilibrium is not Pareto efficient can be avoided if games are repeated a sufficient number of times. If a Prisoners' Dilemma class game is repeated a number of times, it is called *Iterated Prisoners' Dilemma* or IPD.

These sort of dilemmas are more common than they appear to be. In modern economy, for example, IPD situations can be found in cartel formation, advertising and marketing expenditures; and new product or technology development. Not to mention applications in other areas, such as public relations and animal behaviour [32, 7].

If a Prisoners' Dilemma class game is played only once, the Nash equilibrium for both players is to defect, which is not the Pareto efficient solution. However, if this situation is repeated, then it could be more profitable if both players colluded. This can be achieved by establishing a punishment strategy of permanently changing the strategy to defection if one of the players tries to defect, thus discouraging selfish behaviour. This is called a *trigger strategy* because it is activated by an action performed by either of the players, and the punishment strategy is known as *grim* because it changes permanently the players' strategies [27].

Let us consider the case where the game is played an infinite number of times and payoffs are discounted in the future, which makes future values less interesting than the current ones. This can be expressed by discounting future values by a factor discount rate $0.0 < \delta < 1.0$. This means that, some value V at time k will be worth $V\delta^k$.

The payoffs obtained by the player in an IPD can then be expressed as:

Collusion Payoff $= R + R\delta + R\delta^2 + ... = \sum_{i=0}^{\infty} R\delta^i = \frac{R}{1-\delta}$

Defection Payoff$= T + P\delta + P\delta^2 + ... = T + \sum_{i=1}^{\infty} P\delta^i = T + \frac{P\delta}{1-\delta}$

Expected value solution says that collusive behaviour is possible if

$$\frac{R}{1-\delta} \geq T + \frac{P\delta}{1-\delta} \tag{2.1}$$

with $\delta = 1/(1+t)$, where t the interest rate.

Alternatively, the IPD can be repeated an indefinite number of times by adding a stochastic component to its structure, such as a probability p (which is beyond the control of either player) that the game will finish by the end of each round. This way, the end of the game is no longer known by the players and the discount rate used in equation 2.1 is now evaluated as $\delta = (1 - p)/(1 + t)$ (for a derivation of this formula see Appendix A).

2.2 DECISION THEORY

2.2.1 Fundamentals of Decision Theory

Decision Theory can be defined as a set of knowledge, concepts and analytical techniques developed to describe the rational process of decision making in order to help understand it and help decision makers make the best choice when faced with multiple possible alternatives [19].

Decision making is a process consisting of the following steps: problem specification, developing the alternatives, describing the consequences, relating alternatives to consequences and, finally, making the decision [9].

The decision situations we consider are cases in which a decision maker has to choose between a list of mutually exclusive options. In other words, from among the alternatives, one and only one choice can be made. Each of these choices might have one or more possible consequences that are beyond the control of the decision maker, which, again, are mutually exclusive.

Decision theory can be distinguished between decision under uncertainty and decision under risk. For example, when a gamble consists of rolling a fair die, the probability of it landing on any face is known to be $p = 1/6$, making this a problem of decision under risk, but when the gamble consists of rolling an unbalanced die, the probabilities associated to it landing on each face are no longer known, they're uncertain and thus this is a problem of decision under uncertainty [16]. Risk can be measured and is defined by a pair of values $(x, p(x))$, while uncertainty is immeasurable, because the probabilities of each event are not known [12].

There are lots of different methods for evaluating risky assets. Some are very popular, such as mean-variance analysis, semi-variance analysis and value-at-risk [16]. However, these methods can be too subjective to analyse the situations in this study (requiring that the decision-maker establishes a threshold value associated with the risk) and therefore are not very suited for this purpose. Considering this, the chosen methods that will be used in the analysis in this study are presented in the following subsections.

2.2.2 Decision making under uncertainty

When facing a situation in which the outcome of the game is uncertain, meaning that there are no probabilities associated to what might happen, the principle of indifference (principle of insufficient reason) [11] states that all events are considered equally probable because there is no evidence that one of them has higher chance of occurring. In these situations, a player may choose among a number of different criteria to decide his/her course of action. They are:

2.2.2.1 Maximin

Here, the decision rule is to consider the worst consequence of each possible course of action and choose the one that has the least worst consequence. It is the most pessimistic decision to be used; therefore the chosen strategy is the one that maximizes the outcome for the worst-case situation [5].

2.2.2.2 Maximax

The very opposite of Maximin, this is the most optimistic rule to look at, as the decision maker considers the best possible outcome for each course of action and chooses one that gives the maximum outcome [5].

2.2.2.3 Minimax regret

"Regret" (also referred as "opportunity loss") is defined as the difference between what we actually get and the best position that we could have gotten if a different course of action had been taken. The minimum regret is evaluated by listing the maximum amount of regret for each possible strategy and selecting the one that corresponds to the minimum [29, 5].

2.2.3 Decision making under risk

When the probabilities associated to each possible outcome in a game can be measured, then we have decision under risk. In this study, only the following criteria will be covered:

2.2.3.1 Expected value

Consider a lottery $L = \{p_1x_1, p_2x_2, \ldots, p_nx_n\}$ where p_i is the probability of payoff x_i be obtained. The expected value $E[X]$ of this lottery is the weighted mean of the possible payoffs (calculated as in equation 2.2) and determines the average payoff a player would get if he/she played the lottery an infinite number of times [18].

$$E[X] = \sum_{i=1}^{n} p_i x_i \tag{2.2}$$

2.2.3.2 Expected utility

This criterion is defined by expected utility theory [22] and takes into consideration decision maker's risk attitude [4], which follows a utility function. A decision maker's risk attitude can be determined by the minimum value they would trade a risky asset for. This minimum value is called the *certainty equivalent* (ce) of that asset. If a lottery L has an expected payoff of $E[X]$, then the decision maker is risk-averse if his certainty equivalent is $ce < E[X]$, risk-neutral if $ce = E[X]$ and risk-seeking if $ce > E[X]$. Expected utility theory gives the optimal decision if the decision maker follows certain axioms [22].

The way to calculate the expected utility payoff is very similar to calculating the expected value, but considering the utility function instead of the absolute payoff. It is given by $E[X] = \sum_{i=1}^{\infty} U(x_i)p_i$ for discrete distributions and by $E[X] = \int_{-\infty}^{\infty} U(x)dF(x)$ for continuous distributions, where $F(x)$ is the cumulative distribution function (CDF) of the distribution.

2.2.3.3 Stochastic dominance

Roughly speaking, one lottery A is said to dominate another lottery B if A gives at least as good result as B for each probability p according to a utility function $U(x)$. For example, if there is a lottery L and \$1 is added to each of its payoffs, then the new lottery dominates the old one.

Stochastic dominance decision rules are, thus, used to determine whether a lottery dominates another. They consist of two phases: one *objective phase*, in which lotteries are evaluated and dominated ones are deemed inefficient and eliminated; and one *subjective phase*, in which lotteries that are not dominated are evaluated by the decision maker [16]. They only require knowledge of the distributions' CDFs. For second-order or higher stochastic dominance, the decision maker's risk attitude is also required because each rule is meant for a determined risk attitude.

First-order stochastic dominance (FSD) describes that, for every non-decreasing utility function (i.e. any utility function $U(x)$ with $U'(x) \geq 0$ and a strict inequality at some x), lottery A dominates lottery B if, given their cumulative distribution functions $F_A(x)$ and $F_B(x)$,

$$F_A(x) \leq F_B(x) \tag{2.3}$$

for all x with a strict inequality at some x. This means that, for every payoff x, A has at least as high probability of giving at least x than lottery B [16]. The probability of getting a value higher than x in B is given by $1 - F_B(x)$, and the probability of getting higher than x in A is given by $1 - F_A(x)$. Thus, graphically, F_A would have to be below F_B for the condition $(1 - F_A(x)) > (1 - F_B(x))$ to be satisfied and for FSD to exist.

Second-order stochastic dominance (SSD) describes that, for every non-decreasing concave (implying risk-aversion) [16] utility function (i.e. any utility function U with

$U'(x) \geq 0$ and $U"(x) \leq 0$ with strict inequality in at least one x), lottery A dominates lottery B if, given their cumulative distribution functions $F_A(x)$ and $F_B(x)$,

$$\int_{-\infty}^{x} [F_B(t) - F_A(t)]dt \geq 0 \tag{2.4}$$

for all x with a strict inequality at some x. This means that the area below F_B must be greater or at least equal to that below F_A for every value of x from the lower values to the higher ones [16].

Risk-seeking stochastic dominance (RSD) is similar to SSD, but with convex utility functions, or $U"(x) \geq 0$ (implying risk-seeking behaviour) [16] and is defined by

$$\int_{x}^{\infty} [F_B(t) - F_A(t)]dt \geq 0 \tag{2.5}$$

meaning that the area below F_B must be greater or at least equal to that below F_A for every value of x from the higher values of the distribution to the lower ones [16].

2.2.3.4 Prospect theory criterion

Based on Prospect Theory [10], this criterion also considers risk profiles, but argues that expected utility theory's results do not reflect on real life decision situations because it ignores two factors: individuals tend to overweight outcomes that are certain in relation to outcomes which are only probable (certainty effect) and generally overweight small probabilities in relation to large ones (probability weighting). These two elements help explain things such as gambling with small probabilities of winning and insurance [10].

Prospect Theory argues that people are generally risk-averse for gains and risk-seeking for losses. A latter version of the theory, called Cumulative Prospect Theory (CPT) [33] made it possible to calculate the certainty equivalent of multiple-outcome lotteries. For more information on how to find the CPT value of a asset see Appendix B.

3 COLLUSION CONDITIONS IN THE IPD

In this section, an example will be described which will be the object of the analysis throughout the following subsections. First, game theory concepts will be used to model the situation in the form of a game, which will help in the latter analysis. Afterwards, decision theory concepts will be used to show the collusion conditions, first under uncertainty and then under risk. Results found will be compared to show how each decision criteria influences in the collusion conditions in the IPD.

The figures shown throughout this section were produced using the R package that was created for this study (see section 3.4 for more information). Mostly, they are graphs comparing the values obtained with defection to those obtained with a collusion (according to each decision criteria) intended to illustrate which strategy is preferred in a given scenario (with different probabilities of the game ending and different interest rates).

3.1 EXAMPLE

Since the intent of this study is to analyse situations as close to reality as possible, the following example contains outcomes for both gains and losses because, in real life, most of the investment decision problems have probability distributions that include chances of gains as well as chances of losses [17].

Consider an hypothetical situation in which two supermarkets compete in the same neighbourhood. Both are equivalent in size, have the same market share and are interested in increasing their own profit. To that end, they can lower their prices in order to attract more customers, thus increasing their sales at the cost of their profit margin, which gives them a competitive advantage over their rival. However, this strategy is costly and can only be profitable for the supermarket if it is the only one to lower the prices, meaning that if both do so, they'll keep the same market share and will have their profit margin reduced, which will result in loss for both of them. The alternative is for both to form a cartel, thus keeping their market shares and their profit margins unchanged. The payoffs of the supermarkets' game are shown in Table 3.1.

		Supermarket 2	
		Cartel	Lower Prices
Supermarket 1	Cartel	3, 3	-2, 5
	Lower Prices	5, -2	-1, -1

Table 3.1: Supermarkets' game in normal form.

Imagine that this scenario is repeated every week and the supermarkets make their sales plan in the beginning of each week. Both also have some trouble with the tax inspection, meaning that there's a probability $p = 10\%$ that, at the end each the week, one of them may suffer an intervention and be forced to close, in which case the game ends. Payoffs are subject to an interest rate of $t = 5\%$.

To keep a non-cooperative collusion, supermarket 1 threatens supermarket 2 by stating that if 2 chooses to lower its prices, it will retaliate by lowering its own forever onwards, using a grim trigger strategy. This is a credible threat and, therefore, sufficient to keep a non-cooperative collusion. Let's also assume that supermarket 1 and supermarket 2 have the same rationality (thus using the same decision criteria) and both know that.

This situation can be seen as an IPD class game where the collusion strategy is the one of forming a cartel whilst the defection strategy is the one of lowering prices trying to get more customers.

3.2 COLLUSION CONDITIONS UNDER UNCERTAINTY

Now, the scenario previously described will be analysed to find the collusion conditions when using decision under uncertainty criteria. When we apply each of the uncertainty criteria to the one-off Prisoners' Dilemma, the solution is the same for all of them: the best strategy is to defect and lower the prices. This happens because the strategy of defecting strictly dominates the one of colluding and forming a cartel. But when considering a repeated version of the game, the value of the future influences in the decision and new payoff tables (Table 3.2 and Table 3.3, both portraying supermarket 1's point of view with $\delta = 1/(1+t)$) must be considered.

When deciding under uncertainty, the probability of the game lasting one round or lasting infinitely is considered to be the same, and thus analyzing these two extremes is enough to determine which is the best strategy according to each uncertainty criterion.

3.2.1 Maximax criterion

Since maximax is an optimistic approach, it assumes that the other player will be willing to collude and form a cartel (which is the best-case scenario). Therefore, this criterion chooses from the maximum between a series of collusions (if the game lasts indefinitely) or a single defection (if the game ends in one round), such as shown in Table 3.2 (where $max(A, B)$ means the maximum value between A and B).

		P2
		Collusion
P1	Collusion	$\max\left(\frac{R}{1-\delta}, R\right)$
	Defection	$\max\left(T + \frac{P\delta}{1-\delta}, T\right)$

Table 3.2: Maximax view of an IPD.

Note that the values of T, R, P and S can be all positive, all negative or mixed as long as they satisfy the conditions shown in the end of section 2.1.2.

When using maximax, collusion is possible if the maximum value obtained from collusion is greater than the maximum value obtained from defection. Applying this criterion to the example of the supermarkets' game, the maximum value gained from collusion would be $max(3/(1-\delta), 3)$ (63 if the game lasts indefinitely and 3 if it lasts one round) with, which would result in $max(63, 3) = 63$. The maximum value gained from defection would be $max(5 + (-1\delta/(1+\delta)), 5)$, which resolves to $max(5, -15) = 5$. In this case, the best option is clearly to collude, because in the best-case scenario, it gives a higher payoff.

3.2.2 Maximin criterion

Maximin, on the other hand, is a pessimistic approach. Therefore, it assumes that the other player is willing to defect and lower its prices (which is the worst-case scenario) and since the best response to a defection is also to defect, then the solution is that both players will defect and lower their prices, making the formation of a cartel impossible. Analytically, this is shown by the fact that $P > S$ (see section 2.1.2) and so $P + P\delta/(1-\delta) > S + P\delta/(1-\delta)$, thus defection rules out any possibility of collusion because colluding would always give one S in the first round.

In the supermarkets' game, for example, if maximin is used and supermarket

		P2
		Defection
P1	Collusion	$max\left(S+\frac{P\delta}{1-\delta}, S\right)$
	Defection	$max\left(\frac{P}{1-\delta}, P\right)$

Table 3.3: Maximin view of an IPD.

1 is sure that 2 will lower its prices, it needs to choose between lowering its prices too, getting -1 (P) in the first round and the value of P discounted in all the following rounds, or getting -2 (S) for sure and receive the punishment P discounted in all the following rounds. Since $P > S$, lowering the prices is the best choice.

3.2.3 Minimax regret criterion

Minimax regret is called an opportunistic approach because it chooses the least regrettable option, with regret being the difference between the most desirable outcome and the actual gain in a given scenario. Thus, minimax regret combines Table 3.2 and Table 3.3 and sees the regret associated to each situation. In this case, a cartel is only possible if

$$\begin{aligned} &max\left(\frac{R}{1-\delta}, R\right) > max(T+\frac{P\delta}{1-\delta}, T) \\ &\text{and} \\ &\left[max\left(\frac{R}{1-\delta}, R\right) - max\left(\frac{T-P\delta}{1-\delta}, T\right)\right] > (P-S) \end{aligned} \tag{3.1}$$

otherwise defection is the best option and the supermarkets will lower its prices (see Appendix C for more detailed explanation).

In the supermarkets' game, for example, according to the maximax the best strategy is to collude (gives a payoff of 63 against a maximum 5 from defecting) and according to the maximin the best strategy is to defect (gives -1 against -2 from colluding). The regret of not choosing collusion when the best-case scenario (maximax) occurs is $63 - 5 = 58$ and the regret of not choosing defection when the worst-case scenario (maximin) occurs is $-1 - (-2) = 1$. Since the regret of not choosing collusion is higher and the objective of this criterion is to minimize regret, collusion is the preferred strategy.

Although this is an interesting example on how decision criteria could influence collusion conditions in the IPD, these criteria assume all outcomes as equiprobable. Since they do not consider the probability associated to each event, they are not very suited for situations in which the probabilities are known to be different, which are the most common cases.

3.3 COLLUSION CONDITIONS UNDER RISK

Since the distribution of the payoffs is usually described by a probability function, decision under risk criteria can be used to determine which are the conditions for collusion in each situation. Figure 3.1 shows the PDFs of each strategy (collusion and defection) with the horizontal axis (value) being the possible payoffs and the vertical axis being the probability of getting that payoff. Thus, the collusion distribution given by $C_i = \left(\sum_{k=1}^{i} R\delta^{(i-1)}; p(1-p)^{(i-1)}\right)$ and defection distribution by $D_1 = (T, p)$ and $D_i = \left(\sum_{k=2}^{i} P\delta^{i-1}; p(1-p)^{(i-1)}\right)$.

3.3.1 Expected value criterion

The expected payoffs obtained with each strategy are shown in section 2.1.4 and the collusion conditions under expected value criterion are described in equation 2.1. However, now the probability p of the game ending must be accounted for and so the value of the discount rate is $\delta = (1-p)/(1+t)$. If the condition exposed

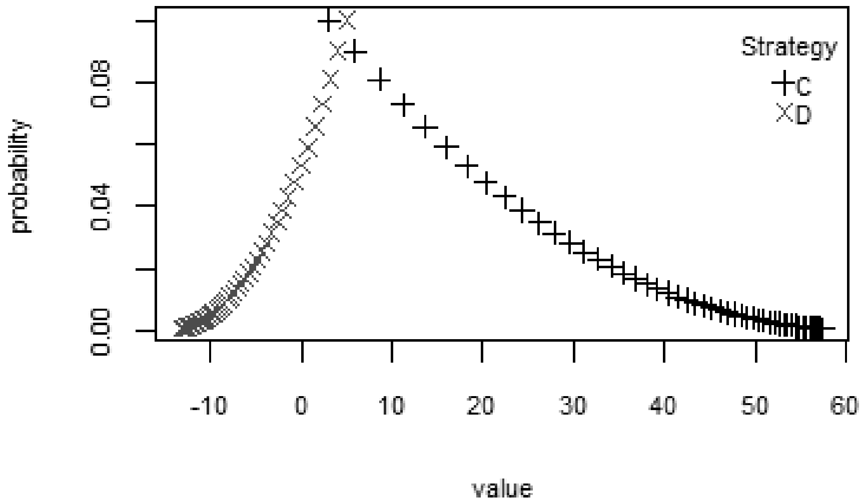

Figure 3.1: PDF of the payoffs for collusion(C) and defection(D) with T=5, R=3, P=-1, p=10% and t=5%.

in equation 2.1 is not satisfied, then the equilibrium is mutual defection and both supermarkets will lower their prices. In the case of the supermarkets' game, the condition would be

$$\frac{3}{1-\delta} > 5 + \frac{-1\delta}{1-\delta} \tag{3.2}$$

which resolves to $\delta > 1/3$, meaning that δ has to be greater than approximately 0.33 for a cartel to be the best strategy for both supermarkets according to the expected value criterion. Since $\delta = (1-0.1)/(1+0.05) \approx 0.86$, then collusion is possible and a cartel will be formed.

To illustrate how the criterion works, if we change the value of R to 2 and the value of P to 1, then δ has to be greater than 0.75 for collusion to happen. But if we change the values of p to 20% and t to 10%, for example, then $\delta \approx 0.72$, which is lower than 0.75 and so collusion is not possible and the supermarkets will lower their prices.

Nevertheless, people rarely consider only the mean value when making a decision because the probability distribution may be quite skewed and the mean may not have much significance (such as in the St. Petersburg Paradox [4], where the expected value is infinite). There are also matters such as the form of the distribution or the risk of loss that should also be considered.

3.3.2 Expected utility criterion

Utility theory states that individuals have particular utilities for different values, which gives each one a utility function. If there is complete information on preferences (i.e. the utility function $U(x)$ is known), then it is simple to calculate the maximum expected utility of a lottery.

If there is complete information and the utility function is known, then the expected utility of both strategies can be calculated such as shown under expected utility in section 2.2.3. $E[U(X_C)]$ being the expected utility of collusion and $E[U(X_D)]$ being that of defection, if the former is greater than the latter, then collusion is the best strategy. Thus, if supermarket 1 knows supermarket's 2 utility function $U(x)$, it could use the PDF of the distributions (Figure 3.1) to calculate the expected utility for each strategy and then identify which one supermarket 2 will choose.

One problem that arises when trying to identify the collusion conditions when both players intend to maximize expected utility is that they hardly ever have the same utility functions. In the oil cartel (OPEC), for example, each country has its own political positions and needs, which vary in terms of wealth, population, resources, etc, and thus, different utilities for the price of the oil barrel [35]. Therefore, it can happen that in the same situation the best strategy for one decision maker is not the best strategy for another.

Some examples of different utility functions and their effects on the preferred strategies in the supermarkets' game are shown in Figure 3.2. To plot the graph, the expected value of each strategy (collusion and defection) were calculated and whichever one had the greatest value was considered the preferred strategy. The light areas represent the regions where collusion is possible (i.e. $E[U(X_C)] > E[U(X_D)]$) while the dark areas represent the regions where defection is the best strategy (i.e. $E[U(X_D)] > E[U(X_C)]$) according to each utility function.

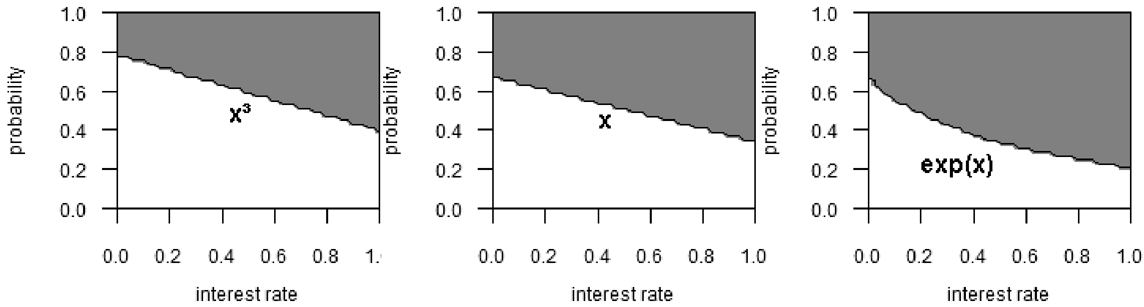

Figure 3.2: Collusion regions of the EU for the functions $g(x) = x^3$, $f(x) = x$ and $exp(x) = 1 - e^{-x}$ with $T = 5$, $R = 3$ and $T = -1$.

Utility functions with different concavities (see Figure 3.3) and thus different risk attitudes were used to show how variations in the utility can result in different collusion conditions for the same distributions. This shows that collusion may not arise if the decision problem falls in the domain where one of the players' best strategy is to defect even if the other's best strategy is to collude.

Table 3.4 gives an idea of how expected utility can affect the payoffs in the supermarkets' game. It is the same as Table 3.1 shown in section 3.1 but with the transformed values of T, R, P and S if supermarket 1 utility function were $exp(x) = 1 - e^{-x}$ and supermarket 2 utility function were $g(x) = x^3$.

To show how different utility functions can influence in the chosen strategy, suppose that in the supermarket's game the value of $T = 5$, $R = 1$ and $P = -1$.

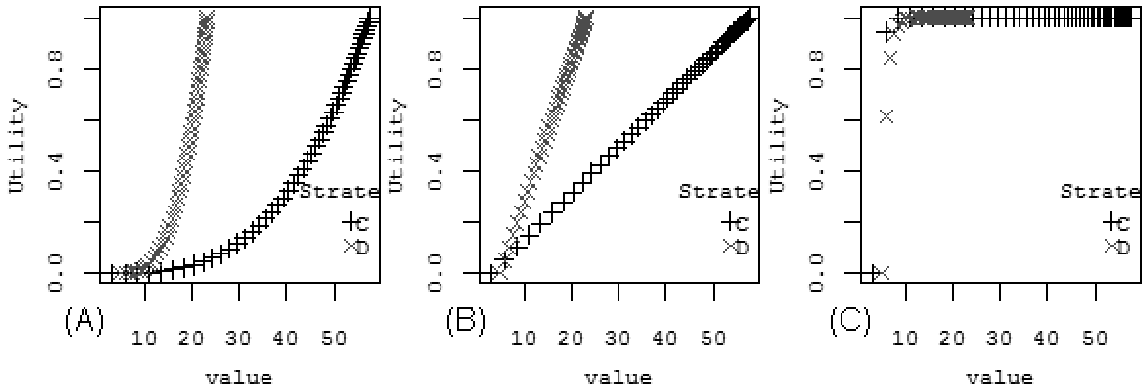

Figure 3.3: Utility functions' concavities of (A) x^3, (B) x and (C) $exp(x)$ for the payoffs collusion (C) and defection (D) in the supermarkets' game.

		Supermarket 2 $U_2(x) = x^3$	
		Collusion	Defection
Supermarket 1 $U_1(x) = 1 - e^{-x}$	Collusion	0.95, 27	-6.39, 125
	Defection	0.99, -8	-1.72, -1

Table 3.4: Supermarkets' game with transformed values according to expected utility functions $exp(x) = 1 - e^{-x}$ and $g(x) = x^3$.

Suppose also that the supermarkets know the tax inspection is tightening their surveillance and thus the probability of one of them being inspected and closed is $p = 35\%$. Interest rate is still $t = 5\%$. Imagine that supermarket 1 has a utility function $U(x) = 1 - e^{-x}$ and supermarket 2 has a utility function $U(x) = x$.

Finding the PDF for each strategy's distribution and calculating the expected utility one can find that, for supermarket 2, $E[U(X_C)] \approx 2.63$ and $E[U(X_D)] \approx 3.38$, making lowering prices the best strategy for it. As for supermarket 1, $E[U(X_C)] \approx 0.83$ and $E[U(X_D)] \approx -1.14$, making a cartel the best option. This situation shows that, under the same circumstances, it can happen that two decision makers choose different strategies because they have different utility functions.

To be able to use expected utility to solve a decision problem, the utility function of the decision maker must be known. If, however, there is not complete information and the utility function is not known, it can be really difficult to compare different lotteries [16]. There is also a possibility that people's decisions may violate expected utility theory axioms, such as the examples shown in Prospect Theory [10], which means that people don't always follow the rationality principle described by Expected Utility Theory.

3.3.3 Stochastic dominance criterion

Through the PDF (Table 3.1) of the distributions it is possible to find the CDF for each of them (Figure 3.4). Once again the horizontal axis (value) are the possible payoffs and the vertical axis being the cumulative probability of that payoff (i.e. the probability with which one can get a value less or equal that payoff).

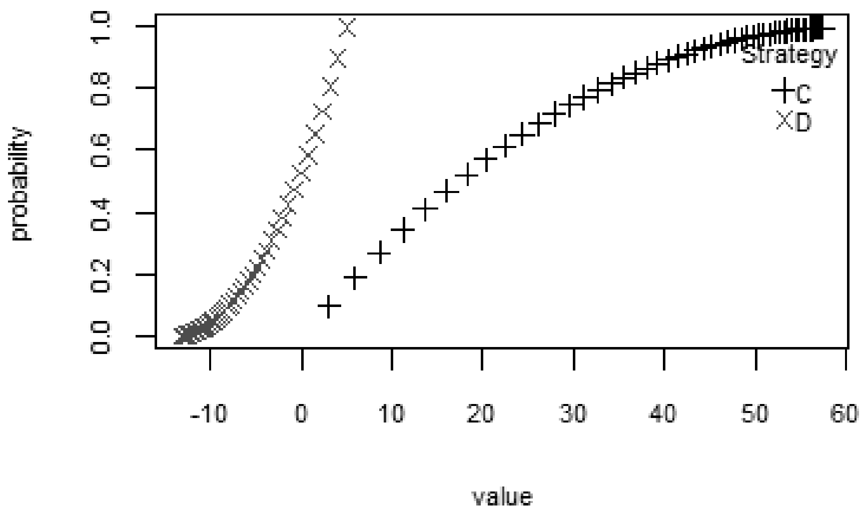

Figure 3.4: CDF of the payoff for collusion (C) and defection (D) with T=5, R=3, P=-1, p=10% and t=5%.

Using the criteria of stochastic dominance it is possible to determine which is the best strategy without knowledge of the utility function and knowing only the distribution. Applying the concepts shown in section 2.2.3, by examining the CDFs of both distributions (Figure 3.4) of the supermarkets' game, it is plain to see that collusion dominates defection by first-order dominance (because for any value X in both distributions, the probability of one supermarket getting at least X with collusion is always greater than with defection [16]) and thus, cartel formation is the best alternative.

The graphs that follow (Figures 3.5, 3.6 and 3.7) were plotted in a way that, if collusion dominates defection, then it is the preferred strategy and the area is marked with a 'C'. If it is the other way around, then the area is marked with a 'D'. If there is no dominance, then one cannot determine which one is the preferred strategy by objective means and thus the area is marked with a 'ND' for 'no dominance'.

Figure 3.5 shows the dominance in the supermarkets' game according to FSD. The areas enclosed between the origin and the curve are the regions where collusion dominates defection, whereas the outer areas represent no dominance between the two strategies. The regions where no dominance exists represent situations in which the decision maker cannot decide the best strategy by objective means, thus making it necessary to subjectively choose between his/her options.

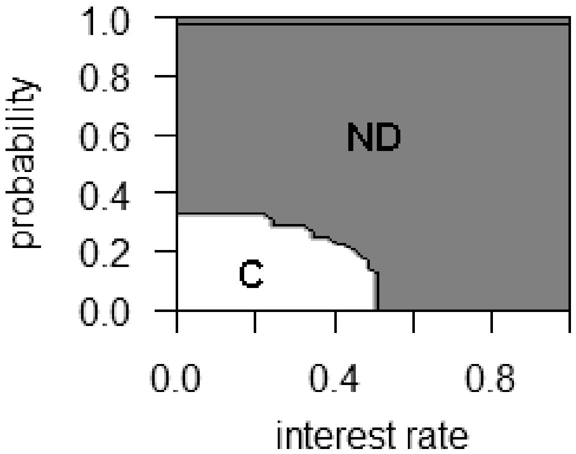

Figure 3.5: Dominance regions of collusion (C) and no dominance (ND) by FSD with T=5, R=3 and P=-1.

Although stochastic dominance can be used under incomplete information requiring only the risk-attitude of the decision maker, it allows only partial ordering, meaning that if there is no dominance then the different options can only be ranked by subjective means. Also, different risk profiles can render different results according to stochastic dominance decision rules.

For example, imagine that $T = 5$, $R = 2$ and $P = -1$ and supermarket 1 is risk-averse while supermarket 2 is risk-seeking. In this situation, 1 will have a collusion region such as that shown in Figure 3.6, while 2 will have one like Figure 3.7. Imagine now that $p = 40\%$ and $t = 5\%$. In this case, the decision problem will fall in the domain in which, for supermarket 1, there exists no dominance, while for supermarket 2 collusion dominates defection. So 2 will be willing to collude, but it doesn't know what 1 will do, making it impossible to predict whether there will be collusion or not.

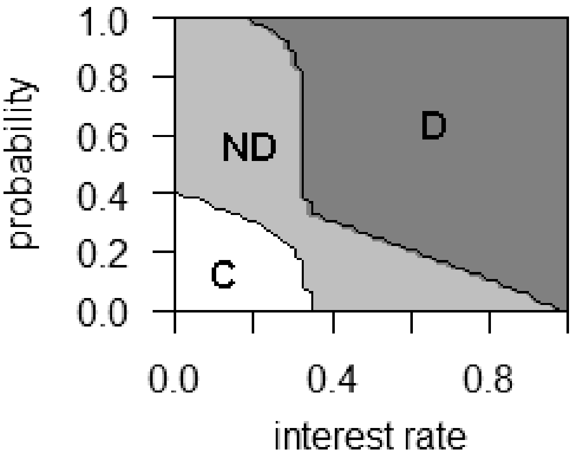

Figure 3.6: Dominant strategies' regions by SSD for the values of T=5, R=2 and P=-1. 'C' stands for collusion, 'D' for defection and 'ND' for no dominance.

This shows that variations in the risk profile can alter the collusion conditions and that it is only possible to predict if players will collude if the colluding strategy is the best option for both of them, otherwise subjective means will determine the outcome of the game and thus the solution won't depend on stochastic dominance decision rules alone.

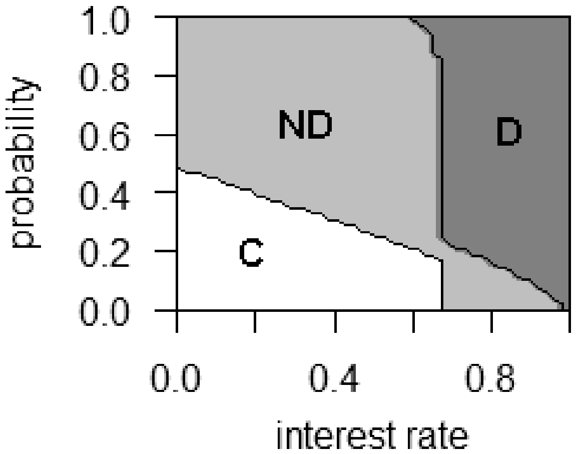

Figure 3.7: Collusion/defection regions by RSD for the values of T=5, R=2 and P=-1. 'C' stands for collusion, 'D' for defection and 'ND' for no dominance.

3.3.4 Cumulative prospect theory based criterion

The CPT method offers a function claimed to be a good approximation to most individual's preferences. Since calculating the value or the certainty equivalent of a lottery through CPT can be difficult because of the algorithm used and the many variables involved, a computational approach is of great help. After calculating the CPT value of each strategy, the one with the higher value is the preferred one.

Consider the example of the supermarkets' game where $T = 5$, $R = 1$, $P = -1$, $t = 5\%$ and $p = 40\%$. In this situation, using the expected value criterion, collusion is possible if $\delta > 0.67$, but in this case $\delta = 0.57$ and thus defection is the best decision. Using the CPT criterion, however, the value of cooperation is 2.27 against 2.17 of defection (these values can be calculated as shown in Appendice B), and so the best strategy is to collude. This happens because in the CPT values are transformed and probabilities are distorted, thus, even though there's a reasonably high probability of getting a payoff of 5 in the first round with defection followed by low probabilities of negative values, these low probabilities are over-weighted and, because of the loss aversion, collusion is slightly preferred.

Even though CPT describes a S-shaped utility function, it allows that each decision maker use his/her own parameters to describe his/her preference curve. Because of that, it may be that two decision makers using the CPT may have different opinions in relation to the same decision problem. An example of such case is shown in Figure 3.8, which pictures two different CPT value functions applied to the supermarkets' game. For each curve (A and B) the CPT value of each strategy was calculated. If the CPT value of cooperation is greater than that of defection, then cooperation is the preferred strategy and vice versa. Thus, the light area below the curve represents the region where collusion is possible (because colluding is the preferred strategy) while the dark area above represents the region where defection is the preferred strategy and thus collusion is not possible.

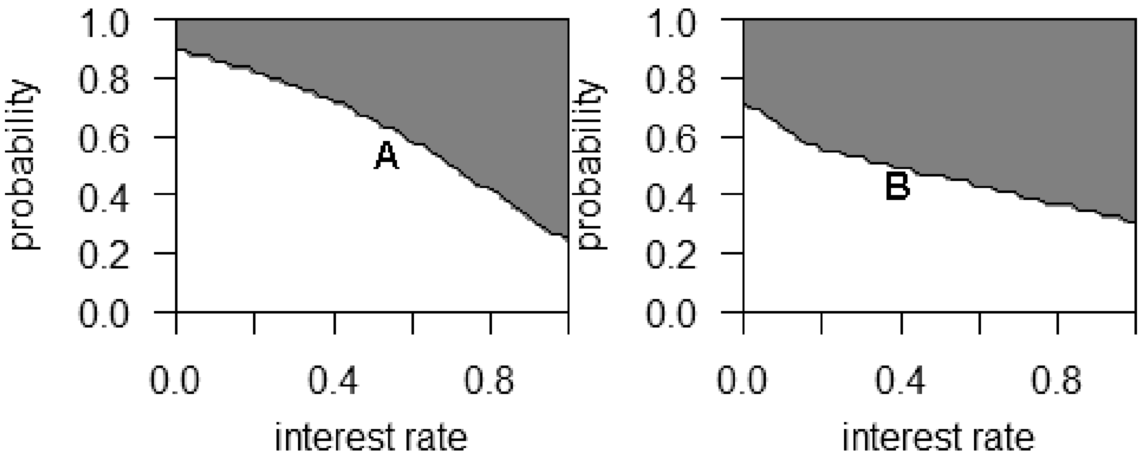

Figure 3.8: Collusion/defection regions by CPT criterion with T=5, R=3 and P=-1 for the function $A(\alpha = \beta = 1.5, \lambda = -2.25, \gamma = \delta = 0.5)$ and $B(\alpha = \beta = 0.1, \lambda = -2.25, \gamma = \delta = 1.0)$.

So, if supermarket 1 must decide on which strategy to use according to the CPT criterion, it will have to calculate the CPT value of each strategy and also have to know other details about supermarket 2 preferences (such as risk aversion, or probability distortion degree) in order to identify 2's preferred strategy, which could vary depending on the values of the parameters in its value function. This shows, once again, that the decision makers' must know each others' preferences and have the same CPT function in order to determine if collusion is possible or not.

Moreover, although the CPT function is said to be an approximation to most peoples' preferences, it is very restrictive. Also, it is arguable whether the function is really a good approximation because experiments used to determine it were made using either positive-only or negative-only lotteries and not mixed lotteries, which are the majority of real life problems [17].

3.4 COMPUTATIONAL APPROACH

In order to be able to make the right choice in every situation, the decision maker must be well informed. But even knowing the many decision criteria might not be enough, because it can be difficult to fit the game structure into each one of them and appropriately determine the best strategy to be used. On top of that, some of the decision methods involve reasonably complicated calculations. Thus, a computational approach is necessary to determine the best strategy.

The statistical computing language R was used to create a package (which can be obtained at `http://equipe.nce.ufrj.br/eber/IPD2DT\ipd2dt_1.0.tar.gz`) containing the algorithms that calculate the probability distributions of each strategy taking into account different interest rates and probabilities. The package also determines dominance between different distributions plots region graphs (such as those shown in section 3) that show the preferable strategy. Some of the package functions are described in Table E.1, found in appendix E. The functions extend the decision criteria to the domain of the IPD, making it so that the decision maker must only know the conditions of the game (T, R, P, S, t and p) in order to calculate the solutions.

This computational aid is intended to make it easier for the decision maker to determine when collusive behaviour is possible according to each criterion. It applies the decision criteria methods to the IPD scenarios so that both theories are put together and shows the results in a way that is simple to understand. In short, it helps the decision maker to analyse the possibilities and choose the best option according to the decision criterion of his/her preference.

4 RELATED WORKS

Kreps et al. [13] had pointed that incomplete information could lead to collusion throughout at least part of the finitely repeated prisoners' dilemma with mixed payoffs. The main difference between their study and ours is that, in theirs, the game is one of incomplete information (i.e. players are not certain about their opponent's rationality), whereas in ours, players have complete information on their opponent's rationality (making it a game of complete information). Another important difference is that, in their study, players know that the game is finite and when it will end, while in ours the game is indefinitely repeated and players do not know when it will end (there exists uncertainty about the future), which is an important factor that strongly influences decision making process.

In their study, Sabater-Grande and Georgantzis [28] draw attention to the fact that risk aversion is an important factor to be considered while solving an IPD. They show, through empirical experiments with positive-only outcomes, that risk aversion tends to lead to defection and that stochastic discount factors (such as the probability of the game ending) are more important than deterministic ones (such as interest rates) for determining whether collusion is possible or not. Their study differs from ours in the sense that, in our analysis, we consider the possibility of both types of discount factors (stochastic and deterministic) coexisting and that we account for situations in which there are positive and negative outcomes.

Van Assen and Snijders [34] show in their study how individuals who are not risk-neutral behave in IPDs with mixed outcomes. They show empirically that loss-aversion favours collusion these situations. This reinforces the fact that risk-orientation is an important factor to be considered in the IPD analysis. Their study differs from ours in the objectives, because theirs is mainly interested in showing the correlation between risk aversion and cooperation or loss aversion and collusion conditions in the IPD. Their approach also categorizes peoples' preferences as those of the CPT criterion, while our study considers other different rationalities. Another difference is that theirs i an empirical study while ours is a theoretical one.

Au et al. [2] show that risk-aversion can favour defection if collusion is the riskier strategy, while risk-seeking behaviour may favour collusion in the same situation, which suggests that the risk index associated to each strategy can also influence in the equilibrium of the PD. The main difference between their study and ours is that theirs consider only the one-off PD game and therefore does not take into consideration the risk associated to the IPD's payoffs' probability distribution, while in our study we consider the game in its iterated form and thus the risk associated to the distribution of the payoffs.

Ng et al. [23] also show, in their study, how players' risk orientation, game riskiness and expectation of cooperation affect collusion conditions in one-off PDs. Their study differs from ours in the fact that they analyse the one-off PD as opposed to ours, which analyses the iterated version of the PD, and also in the fact that theirs is an empirical study while ours is a theoretical one.

Guilfoos and Pape [8] show in their study how case-based decision theory can be used to predict if collusion is possible in the IPD. This study is closely related to ours in the sense that it observes the decision making process in the IPD through a rationality different from expected utility, using instead case-based decision theory in order to identify possibility of collusion. In our study we have the same objective (identifying collusion conditions in the IPD) but we consider other different rationalities, such as decision under uncertainty, stochastic dominance and the CPT, for instance.

Other studies use the IPD as model to develop new strategies for evolutionary players to learn and adapt to their adversaries strategies. some of these studies do this by considering players' reputations along with risk-orientation [15], others consider risk-orientation and use the standard deviation of the payoffs as income stream risk [36], and others consider expectation of cooperation as input for developing a new strategy [30]. Like ours, these studies show that factors such as other players reputations and risk-orientation are also something to be considered in competitive environments, but they differ from ours because their goal is to define an optimal strategy for a player to better react to any unknown environment in which it is put on, whereas our objective is to determine if collusive behaviour is possible in well defined scenarios in which players' rationalities are known.

Shigenaka et al. [31] also try to identify the possibility collusion formation in the IDP, but they add observation errors as another component to be considered in the situation. The main difference from their study and ours lies in the fact that they analyse the collusion conditions in the IPD with an additional variable that is the possibility of observation errors, which we do not account for in our study. On the other hand, they consider expected utility as the only rationality criteria while we consider other different rationalities.

5 FINAL REMARKS

5.1 CONCLUSION

Some of the most popular existent methods do not take the totality of the IPD's characteristics into account, such as the probability distribution of the payoffs and the fact that players have preferences in relation to the possible outcomes. However, these factors are relevant to the solution of the game and therefore must not be ignored.

As it was shown in previous sections, the collusion conditions in an IPD may vary according to the criterion used by each agent to determine the best strategy. This result contrasts with the recurrent solution of using the expected utility, a method rarely ever used to evaluate financial prospects and barely efficient for such purposes [16]. With these new perspectives comes a whole new set of possibilities of collusion conditions depending on each agent's decision criteria.

A major difference with expected utility is that there is only one scenario in which the decision maker is indifferent between two strategies (when they have the same utility) and it is not possible to determine whether there will be collusion or defection. Using criteria such as the FSD, for example, there can be whole regions where the decision maker is indifferent between collusion or defection and thus making it impossible to determine what decision he/she will make.

Under the circumstances shown in Figure 5.1, it is useful to know the possible combinations between strategies in order to try to make the best of each possible situation based on the agents' rationalities.

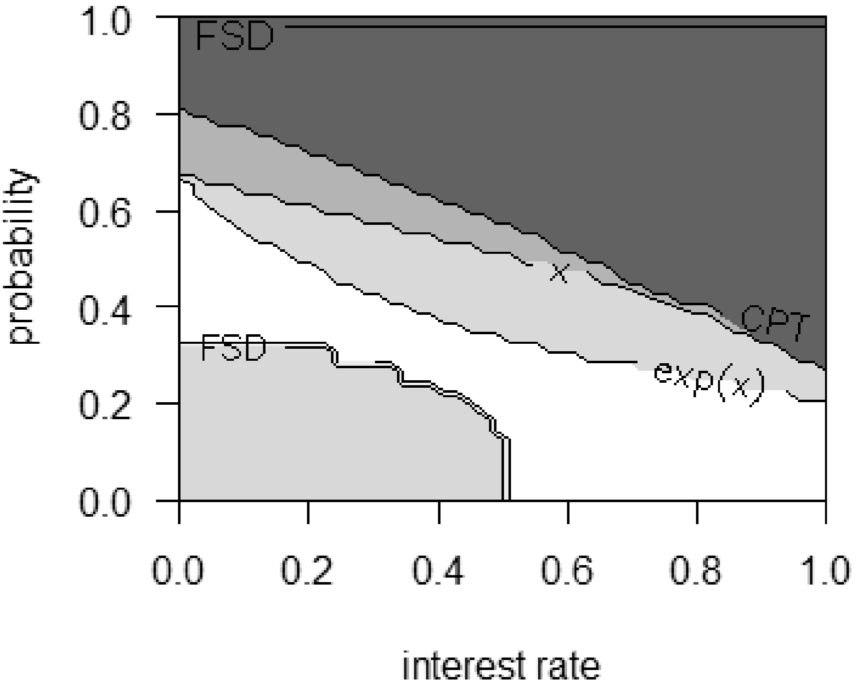

Figure 5.1: Different criteria regions for the values of T=5, R=3 and P=-1

To precisely determine the collusion conditions, it is important to know the decision criteria used by the decision makers. But the importance of these results is not limited only to the fact that the decision makers must know each other's rationalities to predict if there can be collusion or not. It can also be extended to situations between computational agents without human interference.

Finally, the question of whether there can be collusion or not in the IPD is far from settled because it can only be answered in particular cases in which the rationalities of the decision makers are known and are the same. The examples presented in this study intend to show these particular cases and, even so, there are situations in which the decision makers must choose subjectively.

Method	Preferred Strategy
Expected Value	Collude if $\frac{R}{1-\delta} > T + \frac{P\delta}{1-\delta}$ else defection
Maximax	Collude if $max(\frac{R}{1-\delta}, R) > max(T + \frac{P\delta}{1-\delta}, T)$ else defect
Maximin	Always Defect (see section 3.2.2)
	Collude if $\left[max\left(\frac{R}{1-\delta}, R\right) - max\left(\frac{T-P\delta}{1-\delta}, T\right)\right] > (P - S)$ else defect (see Appendix C)
Expected Utility	Collude if $E[U_C(x)] > E[U_D(x)]$
Stochastic Dominance	$CDF_C(x) \leq CDF_D(x)$ (FSD) $\int_{-\infty}^{x}[CDF_D(t) - CDF_C(t)]dt \geq 0$ (SSD) $\int_{x}^{\infty}[CDF_D(t) - CDF_C(t)]dt \geq 0$ (RSD)
Cumulative Prospect Theory	Calculated using the CPT calculation (Appendix B)

Table 5.1: Preferred strategy according to each decision criteria.

5.2 FUTURE WORKS

One limitation of this study which can be explored in future works is that it only considers situations in which both players have the same rationality, although it is quite common that the players have different rationalities, sometimes unknown to the other agents.

Another limitation is that it uses only one punishment strategy (grim trigger) ignoring less severe punishment strategies like the Tit-for-tat [3], for example. With different punishment strategies, one aspect that could be explored is the one in which players make mistakes in judgement, thinking that their opponents will defect when in fact they are inclined to collude and end up defecting.

Another possibility is to conduct and empirical study designed to show that (1) the decision makers rationalities can be really so diverse that, only by knowing the decision criteria used by the other player, the collusion conditions could be determined, and (2) it would be hard to achieve without the help of computational methods.

REFERENCES

[1] ALBAEK, S.; MØLLGAARD, P.; OVERGAARD, P. B. Government-Assisted Oligopoly Coordination? A Concrete Case. **The Journal of Industrial Economics**, Oxford, v.45, n.4, p.429–443, 1997.

[2] AU, W. T. et al. Risk and Prisoner's Dilemma: a reinterpretation of coombs' re-parameterization. **Journal of Behavioral Decision Making**, Chichester, v.25, n.5, p.476–490, 2012.

[3] AXELROD, R. M. The evolution of cooperation. **Science**, Washington, v.211, n.4489, p.1390–1396, 1981.

[4] BERNOULLI, D. Exposition of a new theory on the measurement of risk. **Econometrica: Journal of the Econometric Society**, Chicago, v.22, n.1, p.23–36, 1954.

[5] BINMORE, K. **Rational decisions**. Princeton: Princeton University Press, 2009.

[6] CABRAL, L. M. **Introduction to industrial organization**. Cambridge: MIT press, 2000.

[7] COLMAN, A. M. **Game theory and its applications in the social and biological sciences**. 2.ed. London: Routledge, 1999.

[8] GUILFOOS, T.; PAPE, A. D. Predicting human cooperation in the Prisoner's Dilemma using case-based decision theory. **Theory and Decision**, Dordrecht, v.80, n.1, p.1–32, 2016.

[9] HOLLOWAY, C. A. **Decision making under uncertainty**: models and choices. Englewood Cliffs: Prentice-Hall, 1979.

[10] KAHNEMAN, D.; TVERSKY, A. Prospect theory: an analysis of decision under risk. **Econometrica: Journal of the Econometric Society**, Chicago, v.47, n.2, p.263–292, 1979.

[11] KEYNES, J. M. **A treatise on probability**. London: Macmillan, 1921.

[12] KNIGHT, F. H. **Risk, uncertainty and profit**. Boston: Hart, Schaffner and Marx, 1921.

[13] KREPS, D. M. et al. Rational cooperation in the finitely repeated prisoners' dilemma. **Journal of Economic Theory**, San Diego, v.27, n.2, p.245–252, 1982.

[14] KUHN, S. Prisoner's Dilemma. In: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Stanford: Metaphysics Research Lab, 2009.

[15] LAM, K.-m.; LEUNG, H.-F. Existence of risk strategy equilibrium in games having no pure strategy Nash Equilibrium. **Agent computing and multi-agent systems: 10th Pacific Rim International Conference on Multi-Agents, PRIMA 2007, Bangkok, Thailand, November 21-23, 2007: revised papers**, Berlin, p.1–12, 2009.

[16] LEVY, H. **Stochastic dominance**: investment decision making under uncertainty. 2.ed. New York: Springer, 2006.

[17] LEVY, M.; LEVY, H. Prospect theory: much ado about nothing? **Management Science**, Providence, v.48, n.10, p.1334–1349, 2002.

[18] MARKOWITZ, H. Portfolio selection. **Journal of Finance**, New York, v.7, n.1, p.77–91, 1952.

[19] MCGRAW HILL ENCYCLOPEDIA OF SCIENCE & TECHNOLOGY. **10. ed.** New York: McGraw-Hill, 2007. (Mcgraw Hill Encyclopedia of Science and Technology, 5).

[20] MYERSON, R. B. **Game Theory**: analysis of conflict. Cambridge: Harvard University Press, 1991.

[21] NASH, J. Non-cooperative games. **Annals of mathematics**, Lawrenceville, v.54, n.2, p.286–295, 1951.

[22] NEUMANN, J. von; MORGENSTERN, O. **Theory of Games and Economic Behavior**. Princeton: Princeton University Press, 1944.

[23] NG, G. T. T.; AU, W. T. Expectation and cooperation in prisoner's dilemmas: the moderating role of game riskiness. **Psychonomic Bulletin & Review**, Austin, v.23, n.2, p.353–360, 2016.

[24] OSBORNE, M. J. **An introduction to game theory**. Oxford: Oxford University Press, 2003.

[25] OSBORNE, M. J.; RUBINSTEIN, A. **A Course in Game Theory**. Cambridge: MIT Press, 1994.

[26] POUNDSTONE, W. **Prisoner's dilemma**. New York: Anchor Books, 1993.

[27] ROMP, G. **Game Theory** : introduction and applications. Oxford: Oxford University Press, 1997.

[28] SABATER-GRANDE, G.; GEORGANTZIS, N. Accounting for risk aversion in repeated prisoners' dilemma games: an experimental test. **Journal of Economic Behavior & Organization**, Amsterdam, v.48, n.1, p.37–50, 2002.

[29] SAVAGE, L. J. The theory of statistical decision. **Journal of the American Statistical Association**, New York, v.46, n.253, p.55–67, 1951.

[30] SHIBUSAWA, R.; OTSUKA, T.; SUGAWARA, T. Emergence of Cooperation in Complex Agent Networks Based on Expectation of Cooperation. In: INTERNATIONAL CONFERENCE ON AUTONOMOUS AGENTS & MULTIAGENT SYSTEMS, 2016, Singapore. **Proceedings...** [S.l.: s.n.], 2016. p.1333–1334.

[31] SHIGENAKA, F. et al. Achieving Sustainable Cooperation in Generalized Prisoner's Dilemma with Observation Errors. In: INTERNATIONAL CONFERENCE ON AUTONOMOUS AGENTS & MULTIAGENT SYSTEMS, 2016, Singapore. **Proceedings...** [S.l.: s.n.], 2016. p.1323–1324.

[32] SHY, O. **Industrial organization**: theory and applications. Cambridge: MIT press, 1995.

[33] TVERSKY, A.; KAHNEMAN, D. Advances in prospect theory: cumulative representation of uncertainty. **Journal of Risk and Uncertainty**, Boston, v.5, n.4, p.297–323, 1992.

[34] VAN ASSEN, M.; SNIJDERS, C. The effect of nonlinear utility on behaviour in repeated prisoner's dilemmas. **Rationality and Society**, Newbury, v.22, n.3, p.301–332, 2010.

[35] WEHREY, F. et al. **Saudi-Iranian Relations Since the Fall of Saddam**: rivalry, cooperation, and implications for us policy. Santa Monica: Rand, 2009.

[36] ZENG, W.; LI, M.; CHEN, F. Cooperation in the evolutionary iterated prisoner's dilemma game with risk attitude adaptation. **Applied Soft Computing**, Amsterdam, v.44, p.238–254, 2016.

APPENDIX A EXPECTED VALUE FORMULA

Consider $\delta = \frac{1}{1+t}$ and $q = (1-p)$, then the expected value of cooperation $E_c[X]$ is given by:

$$E_c[X] = \sum_{n=0}^{\infty}\left(pq^n \sum_{k=0}^{n} R\delta^k\right)$$

$\sum_{k=0}^{n} R\delta^k$ is the sum of the n terms of a geometric progression

with decreasing rate, thus it can be written as

$$= \sum_{n=0}^{\infty} pq^n \left[R\frac{1-\delta^{n+1}}{1-\delta}\right]$$

$$= \frac{pR}{1-\delta}\sum_{n=0}^{\infty} q^n \left[1-\delta^{n+1}\right]$$

$$= \frac{pR}{1-\delta}\sum_{n=0}^{\infty} q^n - \delta\sum_{n=0}^{\infty} q\delta^n$$

$$= \frac{pR}{1-\delta}\left(\frac{1}{1-q} - \frac{\delta}{1-q\delta}\right)$$

$$= \frac{pR}{1-\delta}\left(\frac{1-q\delta-\delta+q\delta}{(1-q)(1-q\delta)}\right)$$

$$= \frac{pR}{1-\delta}\frac{1-\delta}{p(1-q\delta)}$$

$$= \frac{R}{(1-q\delta)} = \frac{R}{(1-\frac{1-p}{1+t})}$$

APPENDIX B CUMULATIVE PROSPECT THEORY CALCULATION

Consider a lottery $L = \{p_1x_1, p_2x_2, \ldots, p_nx_n\}$. Cumulative Prospect Theory says that people have a utility for the outcomes that generally follows a power function, thus the value function for each possible outcome x_i in L is given by

$$v(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda\ x^{\beta} & \text{if } x < 0 \end{cases} \tag{B.1}$$

where, according to CPT, $\alpha = \beta = 0.88$ and the average λ is -2.25, meaning that the concavity of the preference curve is towards the horizontal axis. Also the probability weighting function for each probability p_i of each outcome x_i in L is

$$w^{+}(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{\frac{1}{\gamma}}}, w^{-}(p) = \frac{p^{\delta}}{(p^{\delta} + (1-p)^{\delta})^{\frac{1}{\delta}}} \tag{B.2}$$

with w^{+} being for positive values of x_i (gains) and γ equal to 0.61 while w^{-} is for negative values of x_i (losses) and δ equal to 0.69.

In order to calculate the CPT value of a lottery, first it's necessary to divide the distribution in two: $f^{+} = (x_n^{+}, \ldots, x_0^{+})$ and $f^{-} = (x_{-m}^{-}, \ldots, x_0^{-})$, then the weight π of each value x_i in the distribution is given by

$$\pi_n^{+} = w^{+}[p(x_n)] \text{ and } \pi_{-m}^{-} = w^{-}[p(x_{-m})] \tag{B.3}$$

$$\pi_i^{+} = w^{+}[p(x_i) + \cdots + p(x_n)]$$
$$-w^{+}[p(x_{i+1}) + \cdots + p(x_n)], 0 \leq i \leq n-1$$

$$\pi_i^{-} = w^{-}[p(x_{-m}) + \cdots + p(x_i)]$$
$$-w^{-}[p(x_{-m}) + \cdots + p(x_{i-1})], 1-m \leq i \leq 0$$

And so, the value of a prospect is given by

$$V(x) = \sum_{i=1}^{\infty} \pi_i v(x_i) \tag{B.4}$$

this is the CPT value from which it is possible to calculate the certainty equivalent corresponding to the prospect.

APPENDIX C COLLUSION CONDITIONS USING MINIMAX REGRET

To calculate the regret associated with each strategy, it is necessary to combine the maximax (Table 3.2) and maximin (Table 3.3) values in the IPD (see Table C.1). This way, it is possible to identify the best-case scenario in each situation (the maximum value in each row) and find the difference (regret) in relation to the other outcomes in the table.

	Maximax	Maximin
Collusion	$max\left(\frac{R}{1-\delta}, R\right)$	$max\left(S + \frac{P\delta}{1-\delta}, S\right)$
Defection	$max\left(T + \frac{P\delta}{1-\delta}, T\right)$	$max\left(\frac{P}{1-\delta}, P\right)$

Table C.1: Minimax regret table with $\delta = 1/(1+t)$

The conditions under which collusion is possible are: when the value of colluding is higher than that of defecting and enduring the punishment (i.e. collusion is the maximax preferred strategy). For that to happen, the condition

$$max(\frac{R}{1-\delta}, R) > max(T + \frac{P\delta}{1-\delta}, T) \tag{C.1}$$

must be satisfied; and if the regret associated to the collusion strategy is higher than that of the defection strategy (which is always equals to $P - S$, because when the other player defects, the strategy that maximizes the minimum value is to defect as well, getting P in the first round while colluding would give a payoff of S in the first round), thus the condition

$$\left[max\left(\frac{R}{1-\delta}, R\right) - max\left(\frac{T - P\delta}{1-\delta}, T\right)\right] > (P - S) \tag{C.2}$$

must also be satisfied.

In other words, the collusion strategy has to be both the best-case scenario for the maximax and the regret of not choosing the collusion has to be higher than the regret of not choosing the defection strategy, otherwise defection rules out.

APPENDIX D CUMULATIVE PROSPECT THEORY ALGORITHM

Algoritmo 1: Algorithm to compute the Cumulative Prospect Theory value of a lottery

Input: Lottery $L = x_1p_1, x_2p_2, \ldots, x_np_n$
Output: The certainty equivalent of L according to Cumulative prospect Theory

$L^+ \leftarrow$ positive values of x in L;
$L^- \leftarrow$ negative values of x in L;
$p_c \leftarrow 0.0$;
$v^+ \leftarrow 0.0$;
foreach x_i *in* L^+ **do**
  $p \leftarrow p_c + p_i$;
  $v^+ \leftarrow v^+ + x_i^\alpha \left(\frac{p^\gamma}{(p^\gamma+(1-p)^\gamma)^{\frac{1}{\gamma}}} - \frac{p_c^\gamma}{(p_c^\gamma+(1-p_c)^\gamma)^{\frac{1}{\gamma}}} \right)$;
  $p_c \leftarrow p$;
end foreach
$p_c \leftarrow 0.0$;
$v^- \leftarrow 0.0$;
foreach x_i *in* L^- **do**
  $p \leftarrow p_c + p_i$;
  $v^- \leftarrow v^- - \lambda x_i^\beta \left(\frac{p^\delta}{(p^\delta+(1-p)^\delta)^{\frac{1}{\delta}}} - \frac{p_c^\delta}{(p_c^\delta+(1-p_c)^\delta)^{\frac{1}{\delta}}} \right)$;
  $p_c \leftarrow p$;
end foreach
$V \leftarrow v^+ + v^-$;
if $V >= 0$ **then**
  $ce = V^{\frac{1}{\alpha}}$
end if
else
  $ce = (\frac{V}{\lambda})^{\frac{1}{\beta}}$
end if

APPENDIX E R FUNCTIONS DESCRIPTIONS

Function	Description
genCoopPdf	Generate the PDF of the gains obtained with collusion
genGrimPdf	Generate the PDF of the gains obtained with defection
runplotdf	Plots the distributions (PDF or CDF) obtained with collusion and defection
maximax	Determines which strategy gives the better payoff according to the Maximax criterion
minimaxRegret	Determines which strategy gives the better payoff according to the Minimax Regret criterion
evalEU	The expected utility value of each strategy according to a given utility function
fsd	Determines the dominance relation between two given distributions according to FSD
ssd	Determines the dominance relation between two given distributions according to SSD
ssd	Determines the dominance relation between two given distributions according to RSD
calcCPTValue	Evaluates the CPT value of a distribution
gridMaximax, gridMinimaxRegret, gridEU, gridFSD, gridSSD, gridRSD, gridCPT	Generates a grid showing the strategy which gives better payoff according to the given criterion for different values of p and t
runplotPrefCurveEU	Plots the utility curve of a given EU function
runplotMaximax, runplotMinimaxRegret, runplotEU, runplotFSD, runplotSSD, runplotRSD, runplotCPT	Plots the regions of collusion and defection for the given criterion for different values of p and t

Table E.1: Functions included in the R package and their respective descriptions.