POLITECNICO DI TORINO

Department of Management and Production Engineering

Master Degree in Engineering and Management

Thesis

Neuroeconomics: how neuroscience can impact game theory

Supervisor:

prof. Luigi Buzzacchi

Candidate:

Francesca Vallomy

Academic Year 2019/2020

Index

INTRODUCTION 4

CHAPTER I - TRADITIONAL GAME THEORY 6

1.1 AREAS OF APPLICATION OF GAME THEORY 7
1.2 GAME DEFINITION 9
1.2.1 INFORMATION 11
1.2.2 TIME 12
1.2.3 REPRESENTATION 13
1.2.4 CONSTANT AND VARIABLE SUM 14
1.2.5 COOPERATION 15
1.2.6 EQUILIBRIUM 16
1.2.7 DOMINANCE CRITERIA 17
1.3 PRECURSORS OF GAME THEORY 19
1.4 THEORY OF GAMES AND ECONOMIC BEHAVIOR 22
1.4.1 MINIMAX THEOREM 22
1.5 NASH EQUILIBRIUM 26
1.5.1 DEFINITION OF BEST RESPONSE 27
1.5.2 DEFINITION OF NASH EQUILIBRIUM 27
1.5.3 PRISONER'S DILEMMA 28
1.5.4 REPEATED GAME 31
1.5.4.1 Strategies in a repeated game 33
1.5.4.1.1 Tit-for-tat strategy 33
1.5.4.1.2 Grim strategy 33
1.5.4.2 Repeated Prisoner's Dilemma 34
1.5.4.3 Reputation in Repeated Games 35
1.5.5 PARETO OPTIMALITY 36
1.5.6 LIMITATIONS OF NASH EQUILIBRIUM 38
1.6 THEORY OF RATIONAL CHOICE 39
1.7 GAMES ACCORDING TO RATIONAL THEORY 44
1.7.1 ULTIMATUM GAME 44
1.7.2 TRUST GAME 46

CHAPTER II - LIMITATIONS OF TRADITIONAL GAME THEORY AND BEHAVIOURAL GAME THEORY 47

2.1 DRAWBACKS OF TRADITIONAL GAME THEORY 48
2.1.1 NASH EQUILIBRIUM INEFFICIENCY 48
2.1.2 GAMES WITH MULTIPLE NASH EQUILIBRIA 53
2.2 DRAWBACKS OF THE THEORY OF RATIONAL CHOICE 55
2.2.1 INTERNAL LIMITATIONS OF THE THEORY OF RATIONAL CHOICE 58
2.2.2 GENERAL LIMITATIONS OF THE THEORY OF RATIONAL CHOICE 59

2.3 BEHAVIOURAL GAME THEORY **61**
2.3.1 THEORY OF SOCIAL UTILITY 64
2.3.1.1 Theories of fairness payoff-driven 66
2.3.1.2 Theories of fairness intention-driven 72
2.3.2 LEARNING THEORY 78
2.3.3 FURTHER CONSIDERATIONS ON BEHAVIOURAL GAME THEORY 83
2.3.4 MENTAL REPRESENTATIONS 84
2.3.4.1 Framing 85
2.3.4.2 Editing 86
2.3.5 CONSIDERATIONS ABOUT BEHAVIOURAL GAME THEORY 89

CHAPTER III - NEUROSCIENCE'S APPLICATION TO GAME THEORY 91

3.1 HOW NEUROSCIENCE CAN IMPACT GAME THEORY **92**
3.1.1 COGNITIVE NEUROSCIENCE 93
3.1.2 NEUROECONOMICS 94
3.2 BRAIN ANATOMY **95**
3.2.1 BRAIN AREAS IMPORTANT IN ECONOMIC EXPERIMENTS 98
3.2.1.1 Cerebral cortex 99
3.2.1.2 Striatum 101
3.3 NEUROIMAGING TECHNIQUES **102**
3.3.1 TECHNIQUES BASED ON ELECTROMAGNETISM 103
3.3.1.1 Electroencephalography (EEG) 103
3.3.1.2 Magnetoencephalography (MEG) 104
3.3.2 TECHNIQUES BASED ON HEMODYNAMIC REACTIONS 105
3.3.2.1 Positron emission tomography (PET) 105
3.3.2.2 Functional magnetic resonance imaging (fMRI) 105
3.3.3 GOAL OF NEUROIMAGING TECHNIQUES 106
3.4 NEUROSCIENCE EXPERIMENTAL GAMES **107**
3.4.1 ULTIMATUM GAME IN NEUROSCIENCE 108
3.4.2 ALTRUISTIC PUNISHMENT IN NEUROSCIENCE 111
3.4.3 TRUST GAME IN NEUROSCIENCE 114
3.4.4 REPEATED TRUST GAME IN NEUROSCIENCE 116
3.4.5 ITERATED PRISONER'S DILEMMA IN NEUROSCIENCE 120
3.5 REWARD SYSTEM **124**
3.6 FURTHER CONSIDERATIONS ABOUT ECONOMIC EXPERIMENTS **125**

CONCLUSIONS 127

THANKSGIVING 129

BIBLIOGRAPHY 130

SITOGRAPHY 132

Introduction

The object of study of this thesis is the analysis of the divergences between the predictions of players' decisions and their subsequent actions provided by traditional game theory models and behavioral game theory ones. The hypothesis that this thesis seeks to evaluate is that social behaviors predicted by behavioral game theory utility functions are more convergent than traditional game theory models to the real economic decision-making. To prove this, some experiments in which neuroscience techniques are applied to the games, whose theoretical models have already been introduced in the standard game theory, are investigated.

The reason why this argument has been chosen for the completion of this thesis relies on the fact that game theory has been increasingly studied in recent years and this is due not only to its importance in mathematics and economics, but especially to its multidisciplinary nature which entails its development also in fields of research different from the applied mathematics from which its initial development began. Therefore, the deepening of its application in another emerging sector such as the neuroeconomics one turns out to be interesting and challenging since it represents one of the first steps of a research on which the experts of the sector still have much to be studied.

As regard the structure of the chapters, the first chapter is dedicated to a general introduction of traditional game theory. A definition of the game and of all the elements that characterize it, is introduced and then is followed by a brief historical excursus on the origin of the traditional game theory from its precursors to the “*Theory of Games and Economic Behavior”* by John von Neumann and Oskar Morgenstern which corresponds by convention to the effective emergence of the standard game theory. The last part of the chapter is focused on other important elements of game theory, such as the Nash equilibrium and the concept of pareto optimality and on the assumptions underlying the traditional utility functions according to the theory of rational choice. Moreover, an explanation, in a theoretical point of view, of some well-known games is provided. All these parts represent an important basis for the subsequent development of the thesis in the following chapters. The first part of the second chapter analyses the limitations of standard game theory models and of the theory of rational choice on which its main assumptions are based. The heart of the thesis is represented by the second part of the second chapter in which behavioral game theory is finally introduced. In this part, some utility functions models incorporating new variables which takes into consideration players’ emotional and psychological aspects are reported and analyzed in order to draw some initial considerations about divergences between the outcomes predicted by the first traditional models and the more recent behavioral models. Lastly, in the third chapter, after a brief explanation about the cognitive neuroscience domain, the brain anatomy and the most used neuroimaging techniques, neuroscience experiments are investigated to better understand the social behavioral game theory models.

CHAPTER I

Traditional Game Theory

1.1 Areas of application of game theory

The fields of game theory application are several and can range from war games to microeconomic and macroeconomic areas and again from stock exchange to politics. The purpose of this thesis is not to analyze game theory applications in the real world, but to analyze in general terms individual decision-making in the context of strategic interaction[1] with other individuals. However, in order to introduce the foundations of this theory, it is useful to start with a general overview of its fields of application. One of the sectors in which game theory finds its first practical application is certainly the military one. An evident example of this fact is the institution of the Rand Corporation by US in 1948. During the Cold War, the goal of the American Corporation was to apply the concepts coming out from the study of game theory to the Cold War. In fact, the conflict between the Soviet Union and the United States is comparable to a game in which the two players, represented by the two federal States, try to pursue the right strategic decision in order to get the world economic dominance. In an economical point of view, game theory is especially used to understand oligopoly and collusion phenomena. Game theory applications concern both the microeconomic field and the macroeconomic one. As regard microeconomics, the interaction between two firms in order to fix a price for a good, between a firm and a consumer to establish the sale conditions of a product or between a creditor and a debtor for the granting

[1] Game theory deals with strategic interactions among individuals and analyses different typologies of interactions such as repeated or not repeated interactions and interactions with prefect or imperfect information. Individual behaviours turn out to be interdependent, that is to say able to influence each other.

of a loan are common examples of game theory cases. On the other hand, although to a lesser degree, there are macroeconomic situations referable to game theory. An example is provided by the strong economies' decisions in the Eurozone whether to help weak economies giving up to part of their resources and the latter' decisions whether to renounce to their sovereignty and adopt the economic measures imposed by the first ones. Another macroeconomic case are the negotiations about commercial duties between USA and China. More specifically, another economical field that highlights the importance of game theory is the stock exchange where the simultaneous decisions taken by speculators represent a competitive game. Even in a political point of view, game theory has been developed leading to the definition in 1954 of the Shapley-Shubik[2] index. This index is used in mathematical terms to determine the probability of a successful political alliance. An important clarification which is useful to note at this point, is the distinction among cooperative and non-cooperative games. The distinction among different types of games based on cooperation will be analyzed in more detail in the subchapter *1.2.4*.

Finally yet importantly, board games have surely played an important role in the evolution of game theory. However, in this type of games such as the poker or the chess, the outcomes do not depend on probabilities, but are mostly connected to general optimal strategies. Moreover, as it will be analyzed in the next chapters

[2] All the game theory applications described prior the Shapley-Shubik index are classified as non-cooperative games since they don't assume binding agreements between the parties. On the contrary, the political index just described belongs to cooperative games since the goal of the players is to form a coalition.

of the thesis, in more recent years, game theory has been the object of study also of fields such as psychology and neuroscience. These fields have allowed a deeper understanding of the brain functioning in decision-making.

1.2 Game definition

Before to look deeper for the evolution of traditional game theory as a recognized sector of mathematics and economics, it is necessary to provide a brief definition of what a game is. A game can be defined as every situation of interaction between two or more individuals in which strategic behavior is a relevant factor in the decisional process. The most important elements to be identified in every type of game are:

- Players
- Actions
- Strategies
- Payoffs

Players are obviously the participants of the game, which interact to each other during the decisional process. The actions are the set of the possible moves available to the players. In a game composed by $i=1, \ldots, N \in I$ players where $N \geq 2$, each player can dispose of a set $A^i = \{a^i_1, a^i_2, \ldots, a^i_{ni}\}$ of n_i possible actions. An action represents a move taken by a player at a certain point of the game. The strategies represent instead the set of all the potential plans of action. Hence, the concept of strategy must not be confused with the concept of action. If each player i chooses a strategy $a^i_k \in A^i$, a strategy profile $\boldsymbol{a} = (a^1_{i1}, a^2_{i2}, \ldots, a^N_{iN})$ is defined. Each strategy profile identifies eligible actions for all the possible decisional situations that the player can face during the game. More precisely, a strategy profile is modeled in mathematical form by a vector of strategies for all players, as shown just before.

Strategies can be distinguished into pure and mixed. A pure strategy defines in a clear way which particular choice the player will make in any situation he might face during the game. A mixed strategy is a distribution of probabilities on the set of pure strategies available to the player. If the set of pure strategies is composed by n elements, a mixed strategy can be mathematically represented by a vector $\boldsymbol{a} = (a_1, a_2, \ldots, a_N)$ with $a_i \geq 0$ and $\sum_{i=1}^{N}(a_i) = 1$. Stated another, each pure strategy can be seen as a particular case of mixed strategy, which assigns a probability of 1 to that pure strategy.

Finally, payoffs represent the set of the outcomes of the game for each players' decision. Stated in a more technical way, payoffs describe the utility derived by economic agents from the occurrence of a certain combination of strategies. Therefore, the payoff of each player i is defined by a utility function which associates to each strategy profile a real number.

$$U_i: \boldsymbol{a} \rightarrow \mathrm{IR}$$

The description of the results that emerge in a certain type of game is called *solution* of a game.

Games are classified based on different features. Generally, the most important ones, which are taken into account in order to identify a game, are:

- Information
- Time
- Representation
- Sum
- Cooperation
- Equilibrium
- Dominance Criteria

1.2.1 Information

An important feature to be considered in game classification is the type of information provided to players. There are games with perfect or imperfect information and games with complete or incomplete information.

In a perfect information game, each player knows exactly what happened during the development of the game and its own conditions and other player's ones at the time of the decision. Some examples of perfect information game are board games such as the chess or the old Chinese game *Go*. In these games, players can observe cards and stones laying on the board having a complete picture of previous competitors' moves. In games of imperfect information, there is at least one player who does not know the moves of other players. Therefore, players are not able to valuate precisely the payoffs got by their competitors because of the asymmetry of the information.

Perfect information must not be confused with complete information. This last feature in a game means that every player has complete and perfect knowledge in every moment about the elements that define the game, such as the number of players, the set of strategies and the payoffs achievable as a result of a certain move. Game with complete information represents a theoretical and unrealistic situation. In fact, in everyday reality no subject usually has all the information necessary to make a decision. On the other hand, in games with incomplete information not all players have clear information about all the elements that characterize the game.

1.2.2 Time

An important feature, which is relevant to take into account when classifying a game is the temporary variable. A first classification of games, when considering the time, is into static and dynamic games. In a static game, players make decisions simultaneously. Simultaneous games are generally games with imperfect information in which players take actions without knowing previous or simultaneous actions of the other players. In this type of game, actions and strategies are equivalent.

Instead, in a dynamic game, players make decisions in a sequential way and hence there is a significant difference between actions and strategies. In fact, in this case a strategy consists of a set of actions that a player plans to choose as a response to all possible combinations of actions chosen by its opponents. In general, sequential games are games with perfect information.

Considering for simplicity that the hypothesis of complete information holds, dynamic games with complete information are classified into not repeated games and repeated games. Not repeated games are those games in which players make decisions only after having observed other players' moves. Instead, repeated games involve the repetition over time of a one-shot game. Moreover, repeated games can be repeated a finite number of times or infinitely.

The attention given to time in the representation of a game is due to the fact that the payoff of the same game can be different if the game is one-shot or it is repeated. The Prisoner's Dilemma analyzed thereafter is a clear example of this fact.

1.2.3 Representation

Game representation is a tool used for analysis and formal description of problems in game theory. The main types of game representations are:

- normal form
- extensive form

The normal representation or strategic representation of the game is based on a matrix on which the possible strategies of the players and the payoffs associated with each strategy combination are reported. Player 1's strategies are represented on the matrix rows, while player 2's strategies are represented on the matrix columns. Each cell of the matrix represents a meeting point between the row strategy and the column strategy played by the two platers. Therefore, it represents the payoff which results from those particular strategies. Usually, the normal form is used to represent simultaneous games.

In the extensive form, the players' strategies are represented through a decision tree; hence, it is best suited than the normal form to represent sequential games and strategic decisions in time sequence. The nodes of the decision tree represent the state of the game while the arcs the possible strategies of the player.

1.2.4 Constant and Variable Sum

Games can also be divided in constant sum games and in variable sum games. Constant sum games are by definition those games in which the sum of the payoff of all players always corresponds to a constant outcome. Stated another, the sum of players' gain coincides to the sum of players' loss. Zero sum games are nothing but a particular case of constant sum games in which the algebraic sum of players' payoffs is equal to the constant sum of zero. An example of constant sum game is the poker, which is a competitive game where final payoff remains constant even if the distribution changes.

Variable sum games are instead those games in which the sum of all players' payoffs can differ depending on the strategy pursued by players. This type of games can be cooperative or competitive.

1.2.5 Cooperation

Cooperative and non-cooperative games represent the most important distinction in game theory. In the economic field, the large majority of traditional game theory studies concern non-cooperative games and only in recent years, with the emergence of the neuroeconomics field, the role played by cooperation in game theory has been more deeply analyzed.

Non-cooperative games are all those games in which players cannot create alliances neither communicate with each other independently on the fact that they have common goals to reach or not. The most well-known example of non-cooperative game is the Prisoner's Dilemma that will be analyzed more in detail in the subchapter *1.5.3*. Non-cooperative games can be solved using the so-called maximum strategy according to which players adopt an individualistic rational behavior with the purpose to maximize their own benefit. Moreover, the backward induction[3] is a method usually applied in traditional game theory to solve a finite[4] non cooperative game. The basic idea of this methodology is that the

---

[3] According to the backward induction procedure, it is possible to predict a player's decision based on the actions taken by its successors. Therefore, using this method, the analysis of the game starts from the last actions and move forward to the first one.

[4] A finite game is a game with a finite number of players and strategies.

rationality of players allows to predict their behavior, so that it is possible to identify the choice of a player based on the choices of his successors. Thus, it is possible to begin the analysis of the finite game from the end to the first moves. John Nash, one of the most important figures in traditional game theory, analyzed non-cooperative games leading to the formalization of the famous Nash Equilibrium, which will be analyzed later.

On the other hand, cooperative games are those games in which the agreements that players can establish between each other can be considered binding. More specifically, cooperative games can be divided in Non-Transferable Utility games and Transferable Utility games in which players can transfer part of their utility to their partners. TU-games are a particular case of NTU-games. In TU-games, utility functions of the players must be equivalent to allow the transferability of the utility from a player to another. The most important concept related to cooperative games in contrast to non-cooperative ones is that in this case, there is not a unique strategic outcome, but different possible solutions can be achieved depending on the grade of collaboration of players.

1.2.6 Equilibrium

One of the most meaningful concepts in non-cooperative games is represented by the equilibrium one. To clarify this concept in a simple way, it can be said that the point of equilibrium models a sort of stationary state from which no player has incentive to unilaterally deviate. A strategy profile $a^* = (a^*_1, a^*_2, \ldots, a^*_i, \ldots,$

a^*_N) is in equilibrium if no player can increase its payoff by choosing a different strategy from a^*_i when all other players of the game choose the strategy profile a^*_{-i}.

The equilibrium concept has been deeply analyzed in game theory by John Nash. For this reason, the equilibrium is usually referred to as Nash equilibrium. In subchapter *1.5* this concept will be dealt more in-depth.

1.2.7 Dominance criteria

In game theory, a strategy is dominant if the payoff it leads to is always higher than the one guaranteed by all other alternatives, whatever the strategy chosen by other players. An example of dominance criterion is provided by the matrix represented below.

Player 2

Player 1		***Choice x***	***Choice z***
	Choice y	**A** (0,1)	**B** (1,2)
	Choice w	**C** (1,2)	**D** (2,3)

Figure 1.1 Matrix of payoffs

If player 1 chooses strategy ***y***, he obtains a payoff between 0 and 1 while if he chooses strategy ***w***, he will obtain a payoff between 1 and 2. Therefore, it is clear that for player 1 strategy ***w*** is the best one, regardless of other players' preferences. In the same way, for player 2 choice ***z*** is the best one among the alternatives available. Consequently, the optimal equilibrium is given by the strategy C since it represents the meeting point of the two dominant strategies of the players.

There are two types of dominant strategies: strictly dominant strategy and weakly dominant strategy. A strategy $a_i \in A_i$ is strictly dominant if $\forall \; a'_i \neq a_i$ and $\forall \; \boldsymbol{a}_{-i} \in A_{-i}$, the inequality $U_i\, (a_i, \boldsymbol{a}_{-i}) > U_i(a'_i, \boldsymbol{a}_{-i})$ holds $\forall \; i \neq j$. If every player of the game pursues a strictly dominant strategy, the combination of these strategies is said game solution with strictly dominant strategies. This type of solution is the most possible robust solution which can be found in game theory since it does require the minimum number of assumptions about the behavior of the players. In fact, it is sufficient to assume that each player acts in the best way from its point of view and no assumption is necessary about the behavior of other players. In contrast to the strictly dominant strategy according to which the utility function of the dominant strategy is higher than the utility function of the other strategies in all circumstances, the utility function of the weakly dominant strategy is higher than the utility function of the other strategies only in some circumstances. In other circumstances the utility function of the weakly dominant strategy could be equivalent to the utility function of the weakly dominated strategies. In mathematical terms, a strategy $a_i \in A_i$ is weakly dominant if $\forall \; a'_i \neq a_i \in A_i$, $\forall \; i \neq j$ the following inequality holds:

$U_i\ (a_i,\ \boldsymbol{a}_{-i}) \geq U_i\ (a'_i,\ \boldsymbol{a}_{-i})$ and $\exists\ \boldsymbol{a}''_{-i} \in A_{-i}\ \ \forall\ i \neq j$ such that $U_i\ (a_i,\ \boldsymbol{a}''_{-i}) > U_i\ (a'_i,\ \boldsymbol{a}''_{-i})$

The interesting fact to focus on is that the dominance criteria can be adopted only under the assumption that the players are perfect rational and have common knowledge of the game. According to the rationality assumption, each player acts in order to maximize its payoff, ordering the possible outcomes from best to worst. The assumption of common knowledge states that each player has the same level of understanding about the rules of the game and the payoffs resulting from a particular action. In fact, without such assumptions, players could be influenced in their decisions by some other factors, that is what happens in the real economic world and that will be analyzed more in details in the next chapters. Another point that is worth dwelling on is that there might be particular cases where it is impossible to apply the concept of dominance.

1.3 Precursors of game theory

The origin of game theory has usually made to coincide with the publication of John Von Neumann and Oskar Morgenstern book "*Theory of games and economic behavior"* in 1944. However, the beginnings of the theory go back to the XVI century when Machiavelli in his book "*Il Principe*" described what could be defined as the court games, which took place in particular in the Pope's Renaissance court. The idea of Niccolò Machiavelli was that the courtiers' actions were strategic actions of players and what happened in the court was the outcome of their interactive decisions. A century later, Blaise Pascal and Pierre de Fermat exchanged a correspondence of six letters about the calculation of probabilities in a gambling problem requested by De Méré, laying the foundations for probability. So, with these written records, the pillars of game theory were established.

Another important precursor of game theory was the mathematician Émile Borel, the father of zero-sum games. By definition, zero sum games are those games in which the total gain of one player corresponds exactly to the total loss of other players. Hence, zero sum games are classified as competitive games. The important role played in the context of decision-making by competition will be investigated further in the development of the thesis. The number of players in zero sum games can go from a minimum of two players to a maximum of infinite players and they play in a context of perfect information. In fact, in the course of its drafting, this thesis will prove that the presence or the absence of competition in game theory is a relevant factor in the analysis of the outcome of a game.

The so-called Matching Pennies game is one of the most famous examples of zero- sum game. The procedure of this game is very simple, but it is important to understand the basis of game theory and its future development. Player Matcher and player Nonmatcher have a penny each and have to decide whether to turn the side of the coin on head or tail. If the game end up with the pennies having the same figure, the winner is player Matcher, who will get +1 as showed in the figure below in case of *(Heads, Heads)* or *(Tails, Tails)*. Otherwise, if the outcome is *(Heads, Tails)* or vice versa, the winner is the player Nonmatcher.

		Nonmatcher	
		Heads	***Tails***
Matcher	***Heads***	(+1, -1)	(-1, +1)
	Tails	(-1, +1)	(+1, -1)

Figure 1.2 Matching Pennies

Some other well-known examples of zero-sum games are Poker and gambling in which the sum of the payoffs at the end of players' interaction is zero. For this reason, constant sum games are classified as competitive games.

Moreover, Borel is remembered for its contribution to the measure theory and probability theory, which can be considered in part as the basis of game theory. Not for nothing, he was the first highbrow who, during the 'twenties of the last century, coined the expression "théorie des jeux" referred to social decision-making discipline. The first theorem which has represented an important step in the development of the traditional game theory is the Zermelo-Kuhn Theorem that takes the name from its inventors. This theorem states that a finite game with perfect information has a Nash equilibrium in pure strategies.

1.4 Theory of Games and Economic Behavior

In 1944, the publication of the book "*Theory of Games and Economic Behavior*" by the mathematician John von Neumann and the economist Oskar Morgenstern marked the passage of game theory from a simple subject of study to a recognized interdisciplinary field connecting in particular the mathematical and the economical fields. The title of the book makes a direct reference to the behavior of individuals in decision-making context. This type of behavior is completely dominated by rationality and no trace of human emotion influences players in their choices. The main source of reference of this important book is the article "*On the Theory of Parlor Games*" published by von Neumann in 1928 in which the author proves the minimax theory, the most important theorem at the basis of traditional game theory.

1.4.1 Minimax theorem

In 1926, von Neumann and Morgenstern theorized the so-called minimax theorem. The minimax or maximin theorem states that each finite[5] constant-sum game has at least one minimax or maximin equilibrium point in pure or mixed strategies. A recursive algorithm models the theorem allowing the players to choose at each step of the game the most rational strategy to pursue, aware that they could not expect a better payoff by choosing another strategy. The goal of

[5] In a finite game, the sets of strategies *Ai* and *Bj* are compact because finite and hence without accumulation points. In fact, the minimax theorem is a duality theorem which can be applied to linear programming problems on condition that the sets are convex and compact.

the algorithm is to minimize the maximum loss the players can face while at the same time maximizing the minimum benefit achievable.

To better understand the concept behind this theorem, it is useful to provide an explanation of the theorem in mathematical terms. First of all, it is necessary to define the requirements that the payers must satisfy. These requirements concern the rationality of both players and the fact that the choice of the strategy by both the participants is driven exclusively by their own personal individual benefit.

Given a finite constant sum game and two players 1 and 2, player 1 has a set of *m* strategies available and player 2 has a set of *n* strategies available. Given the strategy pursued by player 1, its choice is identified by *i* and in the same way the choice of player 2 is identified by *j*. So, there will be a list of numbers $a_{i,j}$ with $i=1,\ldots,m$ and $j=1,\ldots,n$ represented in the matrix below. These numbers represent the payoff of player 1.

	B_1	**B_2**	…	**B_n**
A_1	$a_{1,1}$	$a_{1,2}$	…	$a_{1,n}$
A_2	$a_{2,1}$	$a_{2,2}$	…	$a_{2,n}$
…	…	…	…	…
A_m	$a_{m,1}$	$a_{m,2}$	…	$a_{m,n}$

Figure 1.3 Matrix of a game *mxn*

The goal is to identify the optimal strategy for player 1. If player 1 chooses the strategy A_i, $i=1,\ldots,m,$ then player 2, in order to counteract player 1, will choose the strategy B_j, $j=1,\ldots n,$ such that the payoff $a_{i,j}$ of player 1 will be the minimum one. So, among the strategies Ai, the payoff $\alpha_i = min_j\ a_{i,j}$ of player 1 must be considered.

	B₁	**B₂**	**…**	**Bₙ**	α_i
A₁	$a_{1,1}$	$a_{1,2}$	…	$a_{1,n}$	α_1
A₂	$a_{2,1}$	$a_{2,2}$	…	$a_{2,n}$	α_2
…	…	…	…	…	…
Aₘ	$a_{m,1}$	$a_{m,2}$	…	$a_{m,n}$	α_m

Figure 1.4 Matrix of a game *mxn* with payoff of player 1

Therefore, if player 1 chooses the strategy *Ai*, its gain cannot exceed *αi*. Consequently, player 1 in order to increase its benefit, will choose, among the set of available values *αi*, the maximum one. That is to say the value $\alpha = max_i\ \alpha_i$. The maximin value of the game can also be formulated in mathematical terms as follow:

$$\alpha = max_i\ \alpha_i = max_i\ min_j\ a_{i,\ j}$$

The same reasoning applies for player 2 leading to the minimax value $\beta_j = max_i$ $a_{i,j}$, which can also be formulated as follow.

$$\beta = min_j \beta j = min_j max_i \beta_{i,j}$$

The table below gives a clearer representation of how the values of the game are obtained:

	B_1	**B_2**		**B_n**	α_i
A_1	$a_{1,1}$	$a_{1,2}$	…	$a_{1,n}$	$\alpha 1$
A_2	$a_{2,1}$	$a_{2,2}$	…	$a_{2,n}$	$\alpha 2$
…	…	…	…	…	…
A_m	$a_{m,1}$	$a_{m,2}$	…	$a_{m,n}$	αm
β_j	β_1	β_2	…	β_n	

Figure 1.5 Matrix of a game mxn with payoffs of player 1 and 2

1.5 Nash Equilibrium

One of the most important figures in game theory is with no doubt John Nash. Nash was born in West Virginia in 1928. He graduated in mathematics and developed a PhD thesis about non-cooperative games in the University of Princeton where there were professors such as Einstein and Von Neumann. He worked also as a consultant in the Rand Corporation during the period of the Cold War applying the game theory strategies to the war decisions. He has surely been one of the most brilliant mathematicians of the XX century. In fact, during his PhD studies he developed the basic mathematical principles of game theory. Thanks to the important results obtained by these studies, he was awarded the Nobel Prize for Economics in 1994. The most important concept developed concerning non-cooperative games and which made a change in game theory is the Nash Equilibrium, which was introduced by Nash in his paper "*Non-Cooperative Games*" . The Nash Equilibrium is considered the very first attempt to link the mathematical field and economical field because of the mathematical demonstration which proves how rational agents pursue economical strategies. The theorem of Nash Equilibrium is a generalization of the minimax theory of Von Neumann, described in subchapter *1.4.1,* to the broader case of variable sum games.

1.5.1 Definition of best response

Before defining the Nash Equilibrium, it is necessary to take a step back and provide a definition of the fundamental concept of best response.

Given the strategy profile $\boldsymbol{a}_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots a_n)$ of all players different from player i and the strategy $a^*_i \in A_i$ available to player *i-th,* the strategy a^*_i is defined as the best response to strategy $\boldsymbol{a}_{-i}$ if it does not exist any strategy $a_i \in A_i$ which provide a higher payoff to the player: $U_i\ (a^*_i, \boldsymbol{a}_{-i}) \geq U_i\ (a_i, \boldsymbol{a}_{-i}),\ \forall\ a_i \in A_i$.

1.5.2 Definition of Nash Equilibrium

Clarified this concept, it is now possible to proceed with the definition in mathematical terms of the Nash equilibrium. In mathematical terms:

$\boldsymbol{a}^* = (a^*_1, \ldots, a^*_i, \ldots, a^*_n)$ is a Nash Equilibrium if for each player $i \in I$, $U_i\ (a^*_i, \boldsymbol{a}_{-i}) \geq U_i\ (a_i, \boldsymbol{a}_{-i})$

As already anticipated in the subchapter *1.2.5* about equilibrium conditions, a strategy profile is in equilibrium if no player has the interest to deviate unilaterally from this point, that is to say no player can increase its payoff by choosing a strategy different from the Nash Equilibrium a^*_i when all other players choose the strategy profile $\boldsymbol{a}^*_{-i}$.

With respect to the concept of dominant strategy, the Nash equilibrium is a broader concept and consequently requires less assumptions. In fact, both the solutions in strictly dominant strategies and in weakly dominant strategies are Nash Equilibrium. For example, the Prisoner's Dilemma presents a solution in strictly dominant strategies and this solution is a Nash Equilibrium. The famous game Battle of Sex presents multiple solutions in weakly dominant strategies and these solutions are both Nash Equilibria.

As previously mentioned in subchapter *1.2.6* about dominance criteria, there might be situations in which it is not possible to identify a dominant strategy. However, even if in the game there are not strictly nor weakly dominant strategies, it is possible sometimes to identify a set of strategies, one for each player, that lead to the point of equilibrium. Anyway, this type of equilibrium guarantees only a state from which it is no convenient for the player to deviate unilaterally from. Therefore, there is the possibility that the final solution of the game is not the best one in absolute terms and indeed is worse than another solution for all players. The Prisoner's Dilemma described in the following subchapter is a clear example of this fact.

1.5.3 Prisoner's Dilemma

The most well-known example in game theory to better understand the Nash Equilibrium is the Prisoner's Dilemma. It is a simultaneous game with complete information proposed by Albert Tucker in the fifties of the twentieth century. The success of this game is due to the fact that it was developed in the middle of the Cold War and it was applied to analyze diplomatic-military cases. As it can be deducted from the title of the game, the players in this case are two criminals set in the context of decision-making. It is a non-cooperative game in which both the prisoners are closed in two separate cells without the possibility to agree in advance and communicate with each other becoming aware of the strategy adopted by the other. They are both accused of having committed a crime and are simultaneously asked by the investigators whether they want to cooperate with the other prisoner or defect. The following matrix describes the payoff resulting from the match of the choices of the two prisoners.

Prisoner 2

		Cooperate	***Defect***
Prisoner 1	***Cooperate***	(-1,-1)	(-3,0)
	Defect	(0,-3)	(-2,-2)

Figure 1.6 Prisoner's Dilemma Matrix

The three possible outcomes resulting from the game are the following.

- If only one of the prisoners accuses the other, the prisoner who decided to cooperate and that has been accused is sentenced to three years, while the one who has accused can avoid the sentence.
- If both the prisoners accuse the other, they are both sentenced to two years
- If both the prisoners cooperate with each other, they are both sentenced to only one year.

Being a non-cooperative game, it is more convenient for each prisoner to *Defect* because in this way its payoff is higher, independently on other player's decision. In fact, if player 1 decides to *Defect*, for player 2 it is better to *Defect* as well in order to avoid spending three years in prison. If player 1 decides instead to *Cooperate*, for player 2 is even more convenient to *Defect* in order to avoid jail. Obviously, considering rationality as common knowledge[6], the same rational reasoning is made by player 1 when making its decisions. Hence, the Nash equilibrium is given by the point *(Defect, Defect)* which represents a solution in strictly dominant strategies since both players want to avoid the worst possible condition which implies to spend three years in jail by cooperating with the defector. In fact, the strategy *Defect* is a dominant strategy because guarantees always a

[6] According to the mathematician Robert Aumann, in the *"Interactive epistemology I: Knowledge"(1999)*, it is the assumption of rationality as common knowledge for players which lead to the paradox of choosing the pair of strategies (*Defect, Defect*) even if the payoff resulting from the strategies (*Cooperate, Cooperate*) is higher.

higher payoff than strategy *Cooperate,* whatever the strategy chosen by the other player. However, the best payoff of the game is not given by the point of equilibrium, but is provided by the strategy *(Cooperate, Cooperate)* which corresponds to the maximum payoff and which represents the so-called Pareto optimality. So, the dilemma is given by the fact that mutual cooperation would provide the highest payoffs to both players, but is not the Nash equilibrium; hence, according to the normative approach adopted by traditional game theory, the cooperative outcome is irrational and it is not chosen by players. At this point, an important question is to analyze whether the cooperation is point of equilibrium in the finite repeated Prisoner's Dilemma game.

1.5.4 Repeated game

Before starting with the illustration of the repeated Prisoner's Dilemma game, it is useful to provide a brief explanation about what a finite repeated game is. First of all, it is important to specify that a game is not repeated if it is played only once by the same players. If a game G is repeated more than once by the same players, it becomes the so-called *stage game* of the repeated game G^*. The main properties of a repeated game G^* are the following:

- At each repetition of the repeated game *G*,* the set of strategies and the preference relationships[7] available to players are the same of the stage game G;
- Players' payoffs of the repeated game are given by summing the payoffs obtained in each repetition;
- A player's strategy in a repeated game is not simply a list of alternatives to play in each repetition of the game. In fact, the action of an agent at iteration N could have been affected by what has occurred until the iteration N-1, and hence could be contingent[8] to the *history of the game.*

Considering for simplicity a game with two players i and j, the *history* of a repeated game is defined as the set $H = A_i \times A_j$, where A_i and A_j are the sets of possible strategies respectively for player i and j. If $A_i = \{s_1, s_2\}$ and $A_j = \{t_1, t_2\}$, then s_1, s_2, t_1, t_2 are the possible histories of the game at each repetition. For example, it could be $h_1 = \{s_2, t_1\}$ in the first game repetition and $h_2 = \{s_2, t_2\}$ in the second one. It is assumed that the history of the game is known by all player.

[7] The concept of preference relationship will be explained more in details in the subchapter about the theory of rational choice.

[8] Contingent strategies are strategies that indicate in each stage of the game a choice that depends on the previous history.

1.5.4.1 Strategies in a repeated game

The strategies which are applied to a repeated game in which the number of repetitions N is not known in advance by players are contingent strategies. Most of the contingent strategies adopted are called *trigger strategies* and the most well-known trigger strategies are the tit-for-tat strategy and the grim strategy.

1.5.4.1.1 Tit-for-tat strategy

Tit-for-tat strategy has been developed in the 80's by the mathematical psychologist Anatol Rapoport to solve the repeated Prisoner's Dilemma. According to this strategy, the first player starts the game playing the *Cooperate* strategy and then he plays the strategy previously adopted by the other player. So, if player 2 cooperates in all the steps of the game, player 1 will cooperate as well for the entire duration of the game. On the contrary, if player 2 decides to defect at a certain point of the game, in the next iteration player 1 will defect as well.

1.5.4.1.2 Grim strategy

As in the tit-for-tat strategy, in the grim strategy the game starts with cooperation. The only difference in this strategy with respect to the first one is that after the first defection, in all next steps of the game is always played the defect strategy.

1.5.4.2 Repeated Prisoner's Dilemma

The initial hypothesis of the repeated Prisoner's Dilemma is that all players know the number of repetitions of the game. At the N-th iteration of the game after the history *h,* the payoffs are respectively $x(h)= x_1 + x_2 +\ldots+x_n$ for Prisoner 1 and $y(h)=y_1 + y_2 +\ldots+y_N$ for Prisoner 2. This means that at iteration n+1 the stage Prisoner's Dilemma game is strategically identical to the one at iteration n=1, since the constants are simply added to the players' payoffs, as described in the matrix below.

	Cooperate	***Defect***
Cooperate	$-1+x_h$; $-1+y_h$	$-3+x_h$; $0+y_h$
Defect	$0+y_h$; $-3+x_h$	$-2+x_h$; $-2+y_h$

Figure 1.7 Repeated Prisoner's Dilemma Matrix

In this type of game, as in the Prisoner's Dilemma not repeated, the cooperative outcome would be the most beneficial to all players, but it is inconsistent with the strategies of the rational individuals. In fact, in a Prisoner's Dilemma game *G** composed by N=10 repetitions, even if both players have understood the reciprocal advantages of choosing to cooperate, the tenth iteration is equivalent in a strategical point of view to a not repeated Prisoner's Dilemma since there is no deterrent of vengeance in the next round. Hence, in the last iteration the dominant strategy is to *Defect*, whatever the history until that point in the game. If in the last iteration of the game, the best response of all players is to *Defect,* also in the penultimate iteration there is no deterrent to defection. Applying this kind of reasoning iteratively and assuming rationality as common knowledge, the dominant strategy in each iteration of G^* is represented by defect. Cooperation does not represent a rational solution in the finitely repeated version of the Prisoner's Dilemma neither. The Prisoner's Dilemma repeated a finite number of times has a unique Nash Equilibrium represented by the point *(Defect, Defect),* which is played by players at each repetition of the game.

1.5.4.3 Reputation in Repeated Games

In the repeated games, an important concept to be taken into consideration when analysing players' strategies is the reputation. Reputation in game theory is the estimation of a player's credibility by other players. For example, a player's decision of not maintaining agreements previously arranged with other players affects its reputation and its subsequent future strategic interactions with others.

In a finite repeated game with N iterations and in which the number of repetitions N is known by all players, such as the Prisoner's Dilemma previously presented, at the N-th iteration, perfect rational players have the convenience to defect and not to cooperate since after the N-th iteration there will not be another iteration whose outcome could be adversely affected by their reputation. Hence, the last round of the game is similar to a one-shot game situation. By applying the backward induction method, as illustrated in the finite repeated Prisoner's Dilemma, players will defect also in the previous iterations. Therefore, in finite repeated game in which N is known by players, reputation is not a factor which affect the strategic interaction among players.

Instead, in a game repeated an indefinite number of times or in a game repeated a finite number of times in which players do not know in advance the number of repetitions, reputation is an important factor affecting strategic interaction among players since, in this case, the time horizon of the game is unknown to anyone. Deviant behaviours from cooperative agreements generate a short-term advantage but affect the reputation of the player and the credibility of his future threats and promises, thus causing long-term disadvantages. Hence, future strategic countermoves of other players are affected by player's reputation.

The concept of reputation building will be come into play also in *Chapter III* when analysing other repeated game with the use of neuroimaging techniques.

1.5.5 Pareto optimality

A solution is said Pareto optimal if and only if it does not exist any possible way to increase the payoff of a player without decreasing the payoff of another player.

In mathematical terms:

$$\nexists\, a_{i'} \in A_i \mid U_i(a_{i'}, a^*_{-i}) \geq U_i(a^*_i, a^*_{-i}) \land U_j(a_j, a^*_{-i}) \geq U_j(a_j, a^{*\prime}_{-i}) \quad \forall j \land \forall a_j \in A_j$$

As evidenced in the previous subchapter, Nash Equilibrium may not coincide with Pareto optimality. There may be other strategies' combinations which lead to improve the payoff of some players without decreasing anyone's payoff, or even, as in the case of the Prisoner's Dilemma with the strategy *(Cooperate, Cooperate)* , to increase the payoff of all players. Briefly, the best payoff for all players may not coincide with the point of equilibrium. In mathematical terms: $U(a^\circ_1, a^\circ_2, \ldots, a^\circ_i, \ldots, a^\circ_n) > U(a^*_1, a^*_2, \ldots, a^*_i, \ldots, a^*_n)$ $\forall$ player i, where $a^* = (a^*_1, a^*_2, \ldots, a^*_i, \ldots, a^*_n)$ is the Nash Equilibrium and $a^\circ = (a^\circ_1, a^\circ_2, \ldots, a^\circ_i, \ldots, a^\circ_n)$ is the strategy profile of Pareto optimality.

In the example of the Prisoner's Dilemma described before, all the strategies' combinations different from the Nash Equilibrium are points of Pareto optimality. In fact, considering the combinations of strategies *(Cooperate, Cooperate)*, *(Cooperate, Defect)* and *(Defect, Defect)*, it is not possible to find another combination of strategies which implies an increase in the payoff of one of the players without a decrease in the payoff of the other player. For example, considering the combination *(Cooperate, Defect)* to which corresponds the payoff (-3,0) for *Prisoner 1* and *Prisoner 2* respectively, moving to the other combinations of strategies implies an increase of the payoff for *Prisoner 1* but at the same time a decrease of the payoff for *Prisoner 2*.

1.5.6 Limitations of Nash Equilibrium

In conclusion, Nash Equilibrium has surely represented an important step forward in game theory, but it still does not coincides exactly with outcomes of the empirical world since it does not take into account some variables that exist in the real economic world, such as for example the propensity to risk of players or the passing of time in making decisions.

However, even in the cases in which the Equilibrium and Pareto optimality do not match, it is possible to reach a situation in which everyone gets the best result, on condition that the assumptions of perfect rationality and selfishness are not complied with. When among the players there is a relationship of trust such as to lead to a form of cooperation among players, it is more likely that the conditions are improved with respect to the point of equilibrium. The establishment of a form of cooperation among players means that all players act in order not to get the best payoff for themselves, but in order to get the best result for all. In this way, they indirectly obtain the best possible payoff also for themselves. This means that all players can enhance their conditions moving jointly[9] away from the point of equilibrium. However, since collective rationality often contrasts with individual rationality, a binding agreement is necessary in order to establish

[9] It is important to underline that in order to improve their conditions with respect to the point of equilibrium, players have to move away from the point of equilibrium not unilaterally, but jointly. In fact, if only some players move away from the point of equilibrium, players who did not move from it could take advantage of a higher payoff.

a form of cooperation among the players, as it has been already stated in the subchapter *1.2.4* about cooperative games.

1.6 Theory of Rational Choice

Subjects such as philosophy, sociology and economics have tried to explain during the years, with their own technical instruments, the concept of rationality. One of the most relevant theories developed by the economic research in this field and which can be considered as the paradigm on which traditional game theory put its foundations is represented by the theory of rational choice.

All traditional game theory developed prior the emergence of neuroeconomics relies on the importance of the rationality assumption in decision-making. According to this concept, players of game theory can be identified in the figure of a sort of *homo oeconomicus*[10] who is rational[11] and interested only in maximizing its own utility function. The selfishness which characterizes the *homo oeconomicus* makes it not subject to emotional conditioning or, stated more formally, its utility function is not affected by the utility functions of other economic subjects.

[10] The concept of *homo oeconomicus* has been introduced by John Stuart Mill in his essay "*Principles of Political Economy*" (1836) and belongs to the field of microeconomics of individual decisions.

[11] Briefly, a rational individual is one who is always able to order its preferences. These preferences satisfy the set of von Neumann–Morgenstern rationality axioms.

It is important to specify that the underlying rationality model refers to an *instrumental rationality*[12] able to process and use available information to achieve specific purposes. Traditional game theory is enclosed in general terms in the broader field of neoclassical economics. In contrast with the earlier theories of classical economics, according to which the price of a good is given by its cost of production, neoclassical economics assumes that the most important factor for determining a product's price is the utility to consumers. Therefore, purchasing decisions made by consumers are based on the evaluation of the level of benefit derived from the good. This idea is based on the rational behavior theory[13] which states that economic agents act rationally when making decisions. However, what is exactly meant by rationality is an important aspect to linger. In fact, the rational behavior theory does not necessarily imply that consumer preferences are based on monetary benefit derived from a good or a service. The utility derived from a particular good, with respect to another could depend for example on emotions and in general on non-monetary factors.

The theory of rational choice has been developed starting from the fifties of the last century and is mainly based on an axiomatic method according to which the decision-making process is a rational deductive process. The actions undertaken by the agents are the direct result of the rational choices which are based on a

[12] According to the instrumental rationality, players choose the right means in order to achieve their self-interest goals. This type of rationality has been defined by Max Weber with the term Zweckrationalität.

[13] The Theory of Rational Choice has been exemplified through the Theory of Rational Consumer Behaviour.

rational system of preferences. An important clarification to specify is that individuals' choices can be pursued in three different conditions depending on the information context:

- certainty
- uncertainty
- risk

Choices under conditions of certainty refer to situations where the consequences of an action can be determined a priori. When the consequences of an action cannot be identified a priori, choices are made in a context of uncertainty or risk. The difference between uncertainty and risk conditions is that under uncertainty conditions, the economic agent cannot associate a probability with the occurrence of future events. Instead, in case of risk conditions, each event is associated with a certain probability that the future event will occur.

The rationality assumption turns out to be fundamental because is translated into coherence of decisions taken by economic agents. According to this system of preferences, there is a binary system which establishes the relationship of preferences between pairs of actions. By definition, given the outcome of a player's choice $a \in A \subseteq \mathrm{IR}^2_+$ where A represents the set of possible actions of the player, a relationship of preference is a binary relationship over the set of outcomes A allowing comparison in terms of preferability between $\forall\, a, a' \in A$. It is possible to identify three main types of relationships of preference.

- $a > a'$ represents a strict relationship of preference since a is strictly better than a'.
- $a \succsim a'$ represents a weak relationship of preference since a is at least as much preferred as a'.
- $a \sim a'$ represents the relationship of indifference.

In some way, it is the preference system that brings the rationality, consistency and coherence in the criteria of choices. The axiom that establishes this preferential order of possible decisions providing a property of coherence to the decision-making behavior is called axiom of revealed preferences[14]. Another important aspect to underline when dealing with this theory is that the preferences are not measured qualitatively but conveys only ordinal information. According to the theory of rational choice, the economic agent adopts perfectly consistent criteria of choice in relation to a set of rules defined a priori called axioms of rationality[15] whose meaning is highlighted by the axiom of revealed preferences.

The axioms of rationality include:

---

[14] The theory of revealed preferences has been proposed by the economist Paul Anthony Samuelson in 1938. Briefly, it states that under constant consumer behaviour conditions, preferences of economic agents do not change.

[15] It is necessary to point out that it would be more appropriate to enclose the axioms of rationality within decision theory than game theory. The main difference between game theory and decision theory is that the primary goal of the first theory is to find the point of equilibrium while the aim of the decision theory is to find the optimal choice of an economic agent.

- completeness axiom
- transitivity axiom
- continuity axiom
- independence axiom

According to the completeness axiom, each player is able to order its preferences that in mathematical terms can be translated as $\forall$ *a, a'* $\in$ *A* the relationship *a* > *a'* or *a'* >~ *a* hold. The axiom of transitivity states that among three possible alternatives, if a player prefers *a* to *a'* and *a'* to *a''*, it must be true that alternative *a* is preferred to *a''*. In mathematical terms, $\forall$ *a, a', a''* $\in$ A if *a*>~ *a'* and *a'* >~ *a''* then the relationship *a* >~ *a''* must be true. When these conditions of completeness and transitivity are verified, a relationship of preference is said to be rational. The axiom of continuity states that considering two alternatives *a* and *a'*, if *a* is preferred to *a',* tiny deviations on these alternatives do not change the relationship of preference. In particular, if a player prefers alternative *a* to *a'*, the any alternative sufficiently close to *a* is preferred to any other alternative sufficiently close to *a'*. Finally, the axiom of independence states that if an individual prefers alternative *a* with respect to alternative *a'*, the presence of another alternative *a''* does not change the previous relationship of preference. To these four properties, the axiom of non-satiation, which states that it is always preferable for a consumer to consume more, must be added.

However, as it will be analyzed in the following chapter, the use of the theory of rational choice for predicting economic decisions has been the subject of several disputes in recent years with the progress of behavioral economics research.

1.7 Games according to rational theory

In addition to the Prisoner's Dilemma, there are other *experimental* economic games which are worthwhile to describe in order to underline the importance of the perfect rationality assumption in the context of decision-making enclosed in traditional game theory and how a change of this assumption can lead to a change of the outcome of the game. The games which are explained hereafter are the Ultimatum game and the Trust game. In the majority of cases, the decisions observed by players of these games in case of experimental situations are in contradiction with the economic theory of rational choice because these decisions are not driven by factors different from perfect rationality.

1.7.1 Ultimatum Game

The two players which take part in the so-called ultimatum game are the proposer and the responder. The proposer receives an amount of money X without having done anything to deserve them. He has the possibility to keep part of this sum of money on condition that he gives part of that amount of money to the other player and after that this latter accepts the offer of the proposer. The proposer, when deciding which is the optimal split of money to offer to the responder, has to take into account and guess which is the minimum amount of money the other player is willing to accept. In fact, if the responder rejects the split offered, then both players get zero. In a rational point of view, considering a total sum of money X=10 $, the minimum amount of money accepted by the responder is one euro since it is more profitable for him than zero euro. So, according to the theory of rational choice, proposer must offer one of the ten euro received and he must keep the remaining sum of money for himself. In fact, in the perfect rationality context on which traditional game theory is based, even if this type of split proposed by the first player is clearly not fair toward the second player, the responder prefers to receive a small amount of money than receiving nothing. However, as it will be analyzed more in details in the next chapters, in the real economic world, in most cases the responder refuses sum of money lower than half of the amount to be shared.

Another version of the Ultimatum game is represented by the Dictator game in which the responder has not the possibility to reject the amount of money offered by the proposer. The name given to this type of game suggests the imposition of the decision by the proposer who plays the role of a dictator toward the responder who is excluded from the possibility of choosing. The split offered by the dictator to the responder represents a form of altruism by the first player and it has been demonstrated by Camerer in his paper "*Behavioral game theory: experiments in* *strategic interactions*” that in most cases dictators offer a split of 25% of their total fixed amount.

1.7.2 Trust Game

The first version of the trust game was proposed by *Berg, Dighaut, & McCabe,* in 1995*,* where it was presented as an investment game. In this version of the game, the first player, that is to say the investor, is provided a sum of money of 10 $ and he has the possibility to choose between keeping the entire amount of money for himself or send a portion of the money to the second player, that is to say the trustee. If the investor chooses to pursue this second alternative, the trustee can then decide whether to keep the tripled amount invested by the investor totally for himself or reciprocate the trust given by the first player by sending back to him a part of the tripled invested amount. According to standard game theory, perfect rational players do not trust their partners and do not reciprocate trust toward their partners. Therefore, according to traditional game theory, the Nash equilibrium is given by the investor's strategy of investing zero and, consequently, by the second player's strategy of reciprocating zero.

In *Chapter III,* the trust game is taken up and analyzed in an empirical point of view. These game experiments will probe that, in some circumstances, trusting moves by investor and reciprocation of trust by trustee, may occur.

CHAPTER II

Limitations of Traditional Game Theory

and

Behavioural Game Theory

2.1 Drawbacks of Traditional Game Theory

In traditional game theory, the main goal of Nash equilibrium is to respond to the problem of determining, on the basis of rationality principles, the solution to any decision-making game. In fact, if players' strategies are in equilibrium, everyone gets from the game the maximum payoff which he can obtain given the rational choices pursued by other players. Considering that, as demonstrated by Nash, every finite game presents at least one point of equilibrium, such concept seems to offer a general solution to the main problem examined in game theory.

However, there are two main limitations related to the Nash equilibrium. The first limitation has already been mentioned in *Chapter I* and concerns the fact that the point of Nash equilibrium is not always efficient and may not coincides with the one of Pareto optimality. The second limitation relates to the fact that some games present more than one Nash equilibrium. This typology of games puts into evidence the drawbacks of game theory in determining which point of equilibrium will be chosen by the players.

2.1.1 Nash equilibrium inefficiency

As already shown in *Chapter I*, the famous Prisoner's Dilemma rightly makes the point on the inefficiency of Nash equilibrium. In fact, the individualistic and selfish choices of the players lead to a strictly Pareto inefficient equilibrium since the payoff provided by the strategy *(Cooperate, Cooperate)* would have been more beneficial for both players. Therefore, the Nash equilibrium inefficiency refers to all those decision-making situations characterized by a point of equilibrium in which each strategy is the best response to other players' strategies, but the social utility globally got by all players is not optimal.

It is possible to measure quantitatively the equilibrium inefficiency. First of all, it is necessary to define a *social function*, that is to say an objective function defined on the outcomes of the game. Through this function, it is possible to compare different equilibrium points between them or evaluating equilibrium solutions with respect to other solutions not in equilibrium. When defining equilibrium inefficiency, five parameters shall be taken into account:

- how players' utilities are formulated
- the social function adopted
- the concept of solution[16] adopted
- with respect to which solution inefficiency can be assessed
- how solutions can be compared

Players' utilities can be formulated in a cost or in a profit point of view. Obviously, if players' utility is expressed as a cost, the goal is to minimize the social function, while in the other case their aim is to maximize the social function. As regard the social function, one form of function which is commonly adopted is the utilitarian one. According

[16] The solution concept in game theory is used to describe those strategies that should be pursued by decision makers as a result of their rationality assumptions. Stated another way, the term solution concept identifies that choice which, according to rational absolute criteria, is considered an acceptable choice by all players.

to the utilitarian form, the social welfare function[17] *SW* is defined as the sum of individuals' utilities.

$$SW(u_i) = \sum_{i=1}^{n} u_i$$

where u_i is the utility function of a particular economic agent $i \in I$. One of the main properties of the utilitarian function[18] is that the social function increases or remains constant with the increase of a player *i*'s individual utility, assuming that utilities of other players and other conditions are constant. In fact, $\frac{\partial SW}{\partial u_i} \geq 0 \ \forall \ i \in I$. Moreover, individual utility functions have cardinal measurability and can hence be compared among them. The solution concept with respect to which the inefficiency is assessed may be represented for example by Nash equilibriums. The efficiency of a solution can be calculated as the ratio of the value of the social function in the considered solution and the value of the social function in the optimal solution. Considering *SW* as the social welfare function, *S* as the solution concept considered and *OPT* as the value of the solution that maximizes *SW* and considering the utility of players as a cost, it is possible to define in mathematical terms:

$$PoA = max_{s \in S} \frac{SW(s)}{SW(OPT)}$$

[17] Player's individual preferences are mathematically represented by the utility function of the player. The collective preferences are instead represented by social welfare function which is dependent on the utility functions of single players.

[18] It is important to specify that there are also social welfare functions not linear. For example, the Rawlsian SW, the Cobb-Douglas or Bernoulli-Nash SW and the isoelastic SW are not linear. The scope of this thesis is not to deepen SW functions, but it is important to specify that welfare economics is a broad discipline. Among the major experts of welfare economics, there is Amartya Kumar Sen, an Indian economist, philosopher and Nobel Prize in Economics award winner in 1998.

$$PoS = min_{s \in S} \frac{SW(s)}{SW(OPT)}$$

PoA is given by the ratio between the worst solution with respect to *SW* and the optimal solution. *PoA* stands for *Price of Anarchy* and measures how much the value of the social function can worsen due to the lack of cooperation among players. It is an upper bound since it represents the maximum inefficiency of the equilibrium solution. *PoS* is given by the ratio between the best solution with respect to *SW* and the optimal solution. *PoS* stands for *Price of Stability* and measures the minimum worsening of the value of the social function required to choose a solution $s \in S$. It is a lower bound since it represents the minimum inefficiency required so that Nash equilibrium property, whereby players do not deviate from the point of equilibrium, is verified. If the concept of solution adopted is the set of Nash equilibriums of the game, then *PoA* measures the maximum inefficiency that players' selfish behaviour can trigger on overall players' behaviour. If players' utilities are considered as costs, then the main goal is to minimize the social function and $PoA \geq 1$. If $PoA \approx 1$, it means that the cost of all Nash equilibriums is close to the optimal one and therefore the players' selfish behaviours do not cause great damages.

Another important aspect concerning inefficiency is that the increase of the number of players' strategies starting from a game with efficient equilibriums points may lead to unstable solutions and inefficient equilibriums. To clarify this concept, a simple example is useful to be explained. In the game presented in the matrix below, the outcome provided by the strategies *(c,a)* is clearly the only point of equilibrium and it is efficient.

	a	***b***
c	(2,2)	(0,0)
d	(0,0)	(-1,-1)

Figure 2.1 Matrix of the game

However, if the strategies *e* and *f,* with their respective payoffs, are added to the game, the solution changes. The game is represented by the matrix below.

	a	***b***	***e***
c	(2,2)	(0,0)	(0,3)
d	(0,0)	(-1,-1)	(0,1)
f	(3,0)	(1,0)	(1,1)

Figure 2.2 Matrix of the game with added strategies

In this case, *(f,e)* is the only equilibrium but is inefficient.

2.1.2 Games with multiple Nash equilibria

The other main drawback of game theory concerning the presence in a game of multiple Nash equilibria is well-highlighted by a particular class of games called coordination games. Coordination games are by definition those games in which there are pure strategies Nash equilibria when players choose the same strategies or corresponding ones. These games are referred to as games of *coordination* since they, unlike other games, do not express a real conflict between the parties. In game theory, coordination problems arise when players have identical preferences over strategic combinations, but there are two or more Nash equilibrium solutions. The consequence is that economic agents may fail to achieve equilibrium. In fact, even if each player chooses an action associated with an equilibrium, the selected equilibrium may not be the same chosen by all the other players. One of the most famous two-player coordination game is the so-called Battle of the Sexes, represented in the matrix below. In this game, the players are depicted by an engaged couple which has to make a decision between two alternatives. They have to decide simultaneously whether to go to the theater or to a football match. The attendance of the theater is the preferred alternative by player 1, while the football match is the preferred one by player 2. In any case, both of them prefer to spend their time together. It should be noted that in this game there is not a preferable solution that both players would choose if they could talk to each other, as was *(Cooperate, Cooperate)* in the case of the Prisoner's Dilemma. The BoS game presents two Nash equilibria in pure strategies: *(Theater, Theater)* and *(Football Match, Football Match)* and there is not a dominant strategy for any of the player. In fact, if player 1 chooses to go to the theater, the possible outcome he can obtain is 2 > 0, in the event that also the other player chooses the theater alternative, or 0 < 1 if the other player chooses the football match. The BoS game presents also a Nash equilibrium in mixed strategies in which each player chooses his preferred alternative with a higher probability of approximately 60% with respect to the other alternative. However, this equilibrium is not efficient.

Player 2

Player 1		***Theater***	***Football match***
	Theater	(2,1)	(0,0)
	Football match	(0,0)	(1,2)

Figure 2.3 Matrix of the Battle of the Sexes game

In a non-cooperative point of view, there is not a solution to this game. Instead, in a cooperative perspective, the possible solutions of the game are represented by the payoffs (2,1) and (1,2). Hence, one of the two players need to give up to part of its benefit. In order to identify the solution, it is possible to opt for a von Neumann random choice with probability (p,1-p). Moreover, in case of a repeated game, the von Neumann choice would provide the frequency with which players have to go to the theater or to the football match.

The paradox of the BoS game is that the assumption of rationality as common knowledge does not allow the players to choose the same strategy and the game is likely to end in a missed appointment.

In summary, coordination games face three different but interconnected problems:

- the alignment of the choices of individual players so as to avoid conflicting decisions;
- the identification, between the different available alternatives, of the Pareto optimal solution;
- the reduction of the risk to select, because of the strategic uncertainty, suboptimal equilibria.

2.2 Drawbacks of the Theory of Rational Choice

Limitations of the theory of rational choice concern intrinsic limitations of the model as well as external factors which affect the rational formation of the economic agents' choices. Traditional game theory is based on perfect rationality assumptions of the economic agent who is well represented by the figure of the *homo oeconomicus*. However, the concept of homo oeconomicus has been largely criticised in the last years for its oversimplification. In fact, the paradigm of the neoclassical homo oeconomicus appears incomplete in a context of uncertainty and of strategic interdependence, such as the context of game theory, which requires economic agents to be able to make decisions based on other players' choices. Economists have advanced traditional game theory by extending the individual rational choice theory to the interindividual choices, typical of the game theory framework, to overcome this incompleteness of the homo oeconomicus figure. In this way, the initial instrumental rationality, mentioned in *Chapter I,* has been substituted, with the development of the behavioural game theory, by the *cognitive rationality* which is more suitable for a game in which also the beliefs evolved during the game by players towards their opponents are important to be taken into account.

As regard the criticisms made to the rational choice theory over the years, it is possible to mention the American economist Herbert Simon who has demonstrated, through computer simulations, that is impossible for human beings to achieve the optimal solution of a decision-making problem. According to his experiments' results, economic agents are rational only to a limited extent. According to this principle of *bounded rationality*, economic agents are conditioned in their choices by the possibility to consider only some of the possible alternatives and the inability to know all the possible consequences of a particular choice. Simon states that decision makers have to take into account not only the outcomes of other players' choices but also the procedure followed by others. This concept, when applied to game theory, puts into evidence that the presence of more potential players' procedures to make a rational choice may lead to different outcomes, assuming that conditions of the game are the same, and it would be more complex also for players to prefigure their opponents' choices. More recently, the fathers of the modern behavioural economics, Daniel Kahneman and Amos Tversky, have introduced in their paper *"Prospect theory: Decision Making Under Risk"*[19] the so-called "framing effect" phenomenon. The framing effect depicts the set of all those inevitable external factors that can affect the economic decisions of an individual. According to the framing effect, players' choices can change based on the way a problem is explicated. Moreover, it underlines that in the majority of cases, economic agents are more likely to take risks when payoffs are described as losses than gains. Also, some supporters of the methodological individualism, on which the concept of *homo oeconomicus* is based, have criticized this economic model of perfect rationality in favour of limited and imperfect rationality ones. For example, Jon Elster, in its theorization of the precommitment strategy[20], highlights aspects concerning the constraints of human rationality by introducing the concept of imperfect rational individual. According

[19] The paper *"Prospect theory: Decision Making Under Risk"* develops the prospect theory in contrast with the expected utility theory of von Neumann and Morgenstern. According to this alternative decisional theory, it is important to describe empirically how individuals really behave when facing a decision considering also the uncertainty factor which may play an important role in decision-making.

[20] According to the theory of precommitment, in some circumstances an economic agent limits voluntarily the number of possible alternatives to choose. Such a theory refers to the canticle XII of *"The Odyssey"* where Ulysses orders to his sailors to tie him to the ship's mast in order not to heed to the sirens' call. In fact, such canticle highlights the human weakness of will and the man's necessity to pre-commit himself to achieve perfect rationality.

to another important economist, Friedrich von Hayek, economic agents are influenced in their decision-making process by established social rules of conduct[21] even if, in the majority of cases, they are rarely aware of these rules.

2.2.1 Internal limitations of the Theory of Rational Choice

As already stated in *Chapter 1*, the theory of rational choice can be exemplified by the analysis of economic behaviour and more precisely by the theory of consumer behaviour[22]. The reason of this exemplification is that in the critique of the rationality axioms defined to explain the economic behaviour of the agents in a deductive way, the internal drawbacks of the theory of rational choice implicitly emerge. The completeness axiom is violated in all those circumstances in which the economic agents are not able to define a fixed preferential order among the available alternatives. As regards the transitivity axiom, an example is useful to be provided in order to understand its groundlessness. Considering the alternatives *a*, *a'* and *a''*, an economic agent prefers slightly *a* to *a'*. However, this difference is not so clear as to be perceived in a definite way. Consequently, the agent claims to be indifferent between *a* and *a'*. The same happens to the alternatives *a'* and *a''*. Anyway, it could happen that in the comparison of alternatives *a* and *a''*, the agent could prefer largely *a''* over *a* and he could affirm this

[21] In subchapter *2.3.4.2*, it will be showed, dealing with the topic of focal points, how economic agents' decisions can be affected by social knowledge.

[22] According to the consumer theory, if the consumer complies with the rationality axioms, the definite choice he makes must correspond to the optimal one.

preference. Therefore, the agent is indifferent between *a* and *a'* and between *a'* and *a''*, but it is not true that he is indifferent between *a* and *a''*. So, in this example the transitivity axiom does not hold. As regard the non-satiety axiom, the concept of "more is better" is not always true. This axiom hides a clear vision of the human being considered as selfish. His own goal would be exclusively to increase without limits his own satisfaction totally disregarding others' benefit. On the contrary, there can be situations in which the benefit perceived by the consumer decreases with the increase of consumption and sympathy toward others could makes feel the economic agent's benefit as dependent on that of others.

2.2.2 General limitations of the Theory of Rational Choice

Among the more general aspects that put into evidence the limitations of the theory of rational choice, there are the so-called concept of *by-products* and *akrasia.* The concept of by-products refers to all those outcomes which are the result of actions pursued for purposes other than to act in an intentional and rational way. Other situations in which the decisions and the actions of the economic agents fail to meet the criteria of the classical economic model of rationality are those which can be enclosed in the concept of akrasia. Akrasia defines the weakness of will of human being which lead to the inability of the agent to act based on what she considers the most useful choice. Moreover, during the decision-making process, some non-rational mechanisms of formation of the preferences could come into play affecting the choices of individuals. In fact, the theory of rational choice does not explain how preferences are formed and assumes these preferences as external variables from the theory and fixed over time. On the contrary, preferences may change over time, including the time of the decision-making process. In addition, even if the process starting from the ordering of the preferences and leading to the choice of a particular action is a rational process, such preferences may not have been formed by a merely rational process. Hence, the foundations of the theory of rational choice crumble. Among the non-rational mechanisms of preferences' formation there are the adaptive preferences. This typology of preferences is developed during a change of thought of an agent. In particular, the judgment changes from positive to negative. Since it is not possible for the economic agent to pursue its primary preference choice, he tries to convince himself that this is not his real primary preference. On the other hand, counter adaptive preferences represent the other side of the coin of adaptive preferences. In this case, the judgment changes from negative to positive. They are created when the economic agent prefers something different from what he can really get. Other types of preferences formed in an irrational way are the preferences formed through learning. They concern all those cases in which the economic agent changes his mind about her preferences' order after having experienced the outcomes obtained by such preferences.

2.3 Behavioural Game Theory

Behavioural game theory, and behavioural economics in more general terms, should not be considered as a completely new economic discipline which contrasts with the underlying assumption of rationality supposed by neoclassical theory. In fact, economists such as Adam Smith made already reference in their works to the role of empiric behaviour in economic decisions. For example, in his book *"The Theory of Moral Sentiments"*, Smith analysed the psychological factors at the basis of human behaviour, thus laying the foundations for the development of the subsequent behavioural economic theory. Then, towards the end of the XIX century, the emergence of the neoclassical economics contributed to identify the economic decision-maker in the figure of the rational homo oeconomicus, thus separating the psychological field from the economic one. From the middle of the last century there was then a new rapprochement of the psychology field to the economy one and the paper "*Prospect theory: Decision Making Under Risk*" published in 1979 by Tversky and Kahneman[23] has surely marked an important step in the development of behavioural economics. The branch of psychology that has largely affected behavioural game theory is the cognitive decision theory whose aim is to understand agents' cognitive processes underlying real-world decision making. In this regard, the instrumental rationality, on which traditional game theory is based, has been replaced, with the

[23] Amos Tversky and Daniel Kahneman, besides being considered the founders of the behavioural economics, have discovered important and innovative methods of cerebral functions visual representation that will be analysed in *Chapter III*.

development of the behavioural game theory, with the *cognitive rationality*[24], which is more suitable for a game in which also the players' beliefs towards their opponents are important to be considered. Since 1980s, behavioural economics has been one of the major areas of development in the economic field. This was in partly due to the onset dissatisfaction of the economists with respect to traditional economic models and partly because of the multidisciplinarity[25] of behavioural economics.

The step forward made by behavioural game theory with respect to traditional game theory concerns the construction of economic models capable of predicting players' behaviours better, that is to say closer to real economic world, than standard game theory models do. However, unlike what could be thought, behavioural game theory does not pretend to disavow all the traditional game theory previously analysed. In fact, the traditional economic models revolving around the existence of a perfect rational economic agent have represented in most cases the starting point for the more advanced behavioural models. In short, behavioural economics adds elements concerning individuals' behaviour such as misleading influences in the interpretation of information, interdependence of preferences, emotions, learning in repeated games and limited rationality of players

[24] Cognitive rationality guarantees the logical coherence of the whole set of rational criteria choices on which the economics agent's beliefs rely on.

[25] Behavioural economics include elements from economics, psychology and sociology and combine all these elements to achieve a more precise understanding of individuals' economic behaviour.

to the traditional economic theory. An important difference between traditional game theory and behavioural game theory is that the first one is classified as a normative theory, while the second one as a positive theory[26]. Traditional game theory starts from the assumption that players rational, that is to say capable to order their choice alternatives in a scale of preferences, and clever, that is to say able to make logical and complex reasoning, in order to explain why players pursue certain strategies when they are in situations of strategic interdependence and in order to determine which mathematical equilibria may result from the interaction among the economic subjects. Therefore, standard game theory draws implications, that is to say logical consequences, from its assumptions[27]. Behavioural game theory, instead, tries also to explain, through modelization of new utility functions, agents' behaviours which are not in line with behaviours predicted by tradition game theory. In all sectors in which it is applied, the normative approach establishes norms aimed to subject to rules of conduct which can be more or less binding. In this sense, traditional game theory does not stand as a predictive tool for human behaviour, but rather as a suggestion for how people should behave.

[26] Positive theory is described by economists as "*what is*" while normative theory as "*what ought to be*".

[27] Standard game theory models are based on micro foundation; this means that in these games the equilibrium points are obtained from the processes of optimization of the economic agents that make economic choices.

Behavioural game theory can be split into different macro categories:

- theory of social utility
- learning theory in repeated games
- mental representation of the games

2.3.1 Theory of Social Utility

In the context of behavioural game theory, the traditional utility function used in classical game theory dependent only on players' own payoffs turns out to be incomplete. In behavioural game theory, the utility function includes also psychological and emotional factors in its formula. The branch of behavioural game theory is concerned with understanding why some decision makers prefer to lose part of their wealth to reward those who have helped them or punish those who have hurt them is the theory of social preferences. In this regard, different theories of fairness whose purpose is to define utility functions mirroring these possible social behaviours have been developed in recent years. According to these theories of fairness[28], players take into consideration in their decision-making if the game is played equally and fairly by other players and, in particular, if payoffs are distributed approximately equally among economic agents. In general, the

[28] In this context, fairness must be referred to as a particular typology of social preference as altruism, reciprocity, inequity aversion. Social preferences are studied in behavioural economics and psychology to underline the fact that economic agents in their decision-making process take into account not only their own material payoffs, but also other individuals' payoffs and their intentions.

same player will be altruistic with regard to players who have helped her and vengeful towards players who have hurt her. All these "emotional" behaviours have important economic repercussion[29]. Before proceeding with the analysis of the most important types of social utility functions, some initial definitions must be provided. The set of the economic decision-makers is identified as usual by $i = 1,...,N \in I$ and the set of material resources allocations by the vector $X = (x_1, x_2,\dots, x_N)$. Therefore, in mathematical terms an individual i is said to have social preferences if, for any x_i, the utility of i is affected by x_j. Based on these definitions, a first step to formulate the utility function for player i, taking into consideration also its social preferences, is to include in its formula not only x_i but also x_j with $j \neq i$. For example, considering a two-player game and their respective consumptions x_1 and x_2, utility function for player 1 can be written mathematically in the following way.

$$u_1 (x_1, x_2) = \nu (x_1) + \alpha \cdot \nu (x_2)$$

The α coefficient represents what Adam Smith defined as the sympathy[30] coefficient that player 2 exercises toward player 1. If $\alpha > 0$, player 1 benefits from an

[29] As an example of economic consequence of players' social goals, it is sufficient to think of a consumer that, considering unfair the price of a producer, does not buy the good even if the material payoff provided to him would be beneficial. The positive side of the coin is instead represented by a consumer who buy the product whose material payoff is lower than the effective price only because of altruistic actions made by the producer in his regard.

[30] The principle of sympathy proposed by Adam Smith is in net contrast with that of self-interest typical of traditional game theory. In this regard, in "*The Theory of Moral Sentiments"* Smith wrote: *"How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it.*

increase of player 2 consumption. On the contrary, if $\alpha < 0$, the higher is the second player's consumption the lower is the benefit obtained by the first player. In this regard, Edgeworth in his book *"Mathematical Psychics"* wrote: *"We must modify the utilitarian integral by multiplying each pleasure, except the pleasures of the agent himself, by a fraction – a factor doubtless diminishing with what may be called the social distance between the individual agent and those of whose pleasures he takes account"*. Anyway, what is important about this sympathy coefficient is that the utility of each player in a game may be affected positively or negatively by other players' consumption.

An important step forward in behavioural game theory with respect to the traditional one is that in the Prisoner's Dilemma, in some circumstances, the cooperation of the players is an outcome of the game. Some examples about Prisoner's Dilemma that will be reported in this subchapter prove that the cooperative outcome can be an equilibrium. In fact, experimental evidence shows that players cooperate both in the not repeated Prisoner's Dilemma and in the repeated Prisoner's Dilemma, in particular in the first iterations of the finite repeated game.

2.3.1.1 Theories of fairness payoff-driven

The theories of fairness classified as payoff-driven define social utility functions dependent not only on the payoff of the player they are referred to but also on other players' payoffs.

A lot of studies focused on social utility function in the context of altruistic social preference. In general, a player is considered altruistic if his utility function increases not only with his own consumption but also with other players' consumptions. In mathematical terms, this means that the first partial derivative of u (x_1, x_2,…, x_N) is strictly positive with respect to (x_1, x_2,…, x_N).

Charness and Rabin (2000) have introduced a particular form of altruism denominated *quasi maximin preferences*. According to the criterion of quasi maximin preferences, every player's benefit increases not only with his own material payoff, but also with the fair share of welfare he contributes to other players. The starting point to get this social utility function is given by defining a disinterested social welfare function.

$$SW(x_1, x_2,..., x_N) = \delta \cdot min\{x_1,..., x_N\} + (1-\delta) \cdot (x_1+...+x_N)$$

where $\delta \in (0,1)$ is the weight of the quasi maximin criterion. This weight is adopted by players to help the player in the worst social conditions, represented by the first term, versus to help the social welfare of all players in its whole, represented by the second term. For the sake of simplicity, the model does not consider δ as player-specific, but identical for all players. The utility function for a certain player i is defined as follows.

$$u_i(x_1, x_2,..., x_N) = (1-\gamma) \cdot x_i + \gamma \cdot [\delta \cdot min\{x_1,..., x_N\} + (1-\delta) \cdot (x_1+...+x_N)]$$

The utility function is dependent on player *i*'s own payoff and on the social welfare function. The constant $\gamma \in (0,1)$ indicates the sensitivity of player *i* toward its own payoff versus the social interest. This model is used to explain examples of giving and kindness behaviours in the dictator game as well as voluntary contributions in the public goods game. However, one of its limitation is that it cannot be applied to explain punishment behaviours toward unjustified self-interested behaviour of other players.

In order to take into account also punishment behaviours in the social utility function, *Levine (1998)* developed a model including both altruism and spitefulness behaviours. The social utility function defined by Levine is formulated as follows.

$$u_i = x_i + \sum_{j \neq i} x_j \, \frac{(\alpha_i + \lambda \alpha_j)}{(1+\lambda)}$$

where $-1 < \alpha_i < 1$ and $0 \leq \lambda \leq 1$.

To understand the meaning of α_i, it is useful to consider a simplification of the formula above for $\lambda=0$.

$$u_i = x_i + \alpha_i \sum_{j \neq i} x_j$$

The value assumed by the constant α_i is a measure of the kindness of player *i* toward other players. For $\alpha_i > 0$, player *i* wants to help other players while for $\alpha_i < 0$ player *i* wants to hurt them. Instead, the constant λ explains why the same player *i* may behave kindly or unkindly depending on the different situations in which he finds himself. It estimates how players respond to other players' altruistic or spitefulness behaviours and is assumed to be constant for all players of the game. Considering $\lambda > 0$, a kind player with $\alpha_i > 0$ will be more altruistic toward players who behave kindly to him than toward spiteful players.

Other important theories have been developed around the concept of inequity aversion preference. These theories are "*A Theory of Fairness, Competition, and Cooperation (1999)"* proposed by Fehr and Schmidt and "*A Theory of Equity, Reciprocity and Competition (2000)"* by Bolton and Ockenfels. Inequity aversion models include in the traditional decision-making process aimed at maximizing the individual utility, also social preferences such as factors concerning the well-being of the group and the respect for the sense of equity. In short, the concept of inequity aversion preference indicates the player's will of distributing payoffs in a fair and equal way among all players. The inequity aversion's utility function proposed by the *Fehr and Schmidt (1999)* model for player *i* is the following.

$$u_i(x_1, \dots, x_N) = x_i - \frac{\alpha_i}{N-1}\sum_{j\neq i}\max\{x_j - x_i, 0\} - \frac{\beta_i}{n-1}\sum_{j\neq i}\max\{x_i - x_j 0\}$$

where $0 \le \beta_i \le 1$ and $\alpha_i \ge \beta_i \ge 0$.

The first term of the utility function is given as usual by the material payoff of player *i*. The second term and the third term represent the loss caused respectively by disadvantageous and advantageous inequity for player *i*.

It is important to put into evidence that $\frac{\partial u_i}{\partial x_j} \geq 0$ if and only if $x_{\mathrm{i}} \geq x_j$. This means that only if the payoff of j is equal or lower than the payoff of i, player *i* is willing to help him because only in this case player i's utility will increase with player j's payoff. To better understand the meaning of the constants α_{i} and β_i which measure respectively the sensitivities to disadvantageous and advantageous inequities for player *i,* it is useful to consider a two-player game. In this case, the above player i's utility function can be simplified in the following mathematical formula.

$$u_i(x_i, x_j) = x_i - \alpha_{\mathrm{i}} \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\}$$

Since α_{i}, $\beta_i \geq 0$, it is easy to note from the formula above that players are averse to payoff inequality. In fact, assuming x_i as the player i's fixed material payoff, the minus sign before the constants α_{i} and β_i ensures that the utility function of player *i* increases with the decrease of the difference between players' material payoffs x_i and x_j. Anyway, in case of inequality among material payoffs, each player prefers his payoff to be greater than others' payoffs. In mathematical terms this is translated as $\alpha_{\mathrm{i}} \geq \beta_i$. In fact, the difference $(x_j - x_i)$, which represents a disadvantageous inequity for player i, weighs more negatively on his utility than the difference $(x_i - x_j)$ which represents a more advantageous inequity for player *i*. Moreover, since $\beta_i \leq 1$ player *i* is not willing to give up to his benefit to reduce inequality. An important innovative aspect of this model is that players' sensitivity about payoff inequity is player-specific.

The *Bolton and Ockenfels (2000)* model is very similar to the *Fehr and Schmidt (1999)* one, in particular in the case of a two-player game. The social utility function proposed by the *Bolton and Ockenfels (2000)* model is represented as follows.

$$u_i = U_i(x_i, \sigma_i)$$

$$\sigma_i \quad \begin{cases} \dfrac{x_i}{\sum_{j=1}^{n} x_j} & if \sum_{j=1}^{n} x_j \neq 0 \\ \dfrac{1}{n} & if \sum_{j=1}^{n} x_j = 0 \end{cases}$$

The utility function is weakly increasing and concave in x_i for any given σ_i and is strictly concave in σ_i for any given x_i. The maximum value of the utility function is got at $\sigma_i = \frac{1}{n}$. In games with more than two players, an important difference between *Fehr and Schmidt* and *Bolton and Ockenfels* models is that in the first model a certain player *i*, when evaluating payoffs' inequity, takes into account his own payoff compared to the specific payoff of each of his opponents, while, in the second model, player *i* is interested in the difference between his own payoff and the average payoff of his opponents. Therefore, in the *Bolton and Ockenfels* model, player *i* could try to help an agent whose payoff is higher than player *i*'s one if the average payoff of his opponents is lower than player *i*'s one.

2.3.1.2 Theories of fairness intention-driven

In the previous subchapter, social utility functions defined for a certain player, take into account other players' payoffs, but do not consider the player's perceived intentions about others players' behaviours. To consider also this last important element in the social utility function, Matthew Rabin, a professor of behavioural economics and one of the first economists who studied the principle of reciprocity in economics, has developed an intention-driven theory of fairness. To clarify the idea of players' *reciprocity*, Rabin wrote in his paper *"Incorporating Fairness into Game Theory and Economics" (1993):* "*People do not seek to help other people uniformly, rather, they are willing to help others if they believe they will be generous to them.*" Therefore, it is important to distinguish the concept of reciprocity which concerns those players who cooperate only in view of future economic return, by altruism which concerns those players who cooperate unconditionally, that is to say in virtue of an intrinsic motivation and no matter the probability of a future economic return. The intention-based approach cannot be elaborated starting from the traditional game theory framework such as the previous payoff-driven theories. Because of players' interpretation of their opponents' behaviour, theories of fairness which are intention-driven must be enclosed in the field of psychological game theory[31].

[31] The so-called *psychological games* arose from the need to include reciprocity in non-cooperative games and enclose all those games in which players' payoffs depend not only on actions, but also on beliefs about actions. Psychological games analysed in *Rabin (1993)* model refer to psychological games previously introduced by *John Geanakoplos, David Pearce and Ennio Stacchetti (1989).* Contrary to *GPS* which analyse psychological games already

The *Rabin (1993)* model proposes, as already stated, an intention-driven theory of fairness. Considering for simplicity a two-player game, the sets of strategies of player 1 and player 2 are represented as usual by A_1 and A_2 and the material payoff of player i is represented by π_i: $A_1 \times A_2 \rightarrow$ IR. The subjective utility function of a particular player i is affected by three main levels:

- his own strategy a_i
- his belief about player $j \neq i$'s strategy b_j
- his belief about player j's belief about his strategy c_i

Therefore, strategy a_i corresponds to the level 0 of the iterative reasoning of player i, strategy b_j to the first level and strategy c_i to the second level and all these factors contribute to player i's social utility prediction.

At this point, it is possible to define the so-called *kindness function* $f_i(a_i, b_j)$ which measures how kind player i is toward player j, assuming that all players agree upon the meaning of fairness and kindness and apply them symmetrically. If player i thinks that player j will choose strategy b_j, it is possible to measure, through the kindness function, how kind has been player i by choosing strategy a_i. Given the highest possible payoff $\pi^h_j(b_j)$ and the lowest possible payoff $\pi^l_j(b_j)$ among Pareto-efficient points in the set of feasible payoffs for player j, the equitable or fair payoff for player j $\pi^e_j(b_j) = [\pi^h_j(b_j) + \pi^l_j(b_j)]/2$ and the worst player

including the emotional component, Rubin derives psychological games directly from traditional games.

j's possible payoff $\pi^{min}{}_j(b_j)$, it is possible to define the kindness function as follows.

$$f_i(a_i, b_j) = \frac{\pi_j(b_j, a_i) - \pi_j{}^e(b_j)}{\pi_j{}^h(b_j) - \pi_j{}^{min}(b_j)}$$

From the mathematical formula above, it is easy to note that if $f_j(a_i, b_j) = 0$, player i wants player j to get her equitable payoff. If $f_j(a_i, b_j) > 0$, player i is going to be kind toward player j since she wants player j to obtain a payoff higher than the equitable payoff and if $f_j(a_i, b_j) < 0$, player i wants player j to obtain less than her equitable payoff. Moreover, if the denominator $\pi_j^h(b_j)$ - $\pi_j^{min}(b_j)$ = 0, then player j would not be considered able to affect player i's payoff leading to $f_i(a_i, b_j) = 0$

It is also possible to define the function which represent player i's belief about player j's kindness, that is to say the perceived kindness, as follows.

$$f'_j(b_j, c_i) = \frac{\pi_i(c_i, b_j) - \pi_i{}^e(c_i)}{\pi_i{}^h(c_i) - \pi_i{}^{min}(c_i)}$$

As in kindness function, if the denominator $\pi_i{}^h(c_i) - \pi_i{}^{min}(c_i)$ then $f'_j(b_j, c_i)$ = 0. These kindness functions contribute to the definition of each player expected utility.

$$u_i(a_i, b_j, c_i) = \pi_i(a_i, b_j) + f'_j(b_j, c_i) \cdot [1 + f_i(a_i, b_j)]$$

where $\pi_i(a_i, b_j)$ is player's i material payoff when he chooses strategy a_i as a response to his belief about player's j strategy b_j. If player i thinks that player j

will not be kind toward him, that is to say $f'_j(\cdot) < 0$, then player i will choose a strategy such that $f_i(\cdot) < 0$. Similarly, if player i thinks that player j will be kind, he will choose an action a_i such that $f_i(\cdot) > 0$.

At this point, it is possible to define the fairness equilibrium. A pair of strategies $(a_1, a_2) \in (A_1, A_2)$ is a fairness equilibrium if for i=1,2 and $j \neq i$,

- $a \in \text{argmax}_{a \in Ai}\, u_i(a, b_j, c_i)$
- $c_i = b_i = a_i$

To better understand the concept of fairness equilibrium, some examples of psychological games are useful to be explained. Considering the BoS game represented in *figure 2.5* where X > 0; if player 1 thinks that:

- player 2 will choose the *Football match*
- player 2 thinks that player 1 will choose the *Theater*

then, player 1's thought about player 2's intention is that player 2 wants to cause him damage. Therefore, player 1 wants to return the damage to player 2 by choosing the *Theater,* even if this option means a lower material payoff for himself.

Player 2

Player 1		***Theater***	***Football match***
	Theater	(2X,X)	(0,0)
	Football match	(0,0)	(X,2X)

Figure 2.4 Matrix of BoS Rabin (1993)

If $c_1 = b_1 = a_1 =$ *Theater* and $c_2 = b_2 = a_2 =$ *Football match*, *Rabin (1993)* has demonstrated that the pair of strategies (*Theater, Football match*) is a fairness equilibrium for X < 1. In fact, given these values of c_1, b_1, a_1 and c_2, b_2, and a_2, player 2 chooses *Football match* and hence is unkind toward player 1 given his beliefs. Therefore, player 2's kindness function is f_2 = -1 and also f'_2 = -1 since player 1 thinks that player 2 wants to hurt him. As regard player 1, if he chooses the *Theater*, he is unkind toward player 2 given his beliefs and hence f_1= -1. In this case, the material payoff of player 1 is $\pi_1(a_1,b_2)=0$ and consequently his utility function is $u_1(a_1,b_1,c_1)=0+(-1)\cdot[1+(-1)]=0$. Instead, if player 1 chooses *Football match,* he is neither kind nor unkind and f_1 =0. In this case, $\pi_1(a_1,b_2) = X$ and $u_1(a_1,b_1,c_1)=X+(-1)\cdot[1+(0)]=X-1$. Hence, it is clear that for X < 1 it is better for player 1 to play the strategy *Theater* and (*Theater, Football match*) is a fairness equilibrium. It is important to underline that if the two players coordinate and are kind to each other, both (*Theater, Theater*) and (*Football match, Football match*) are fairness equilibria for $X \in IR$.

Beside this example about altruistic punishment which shows how players may give up part of their benefit to punish other players' unkind behaviours, there are also examples which demonstrate how players may renounce to their benefit to reward other player' altruistic behaviours, which can be enclosed in the concept of altruistic rewarding. In the Prisoner's Dilemma depicted in *figure 2.6*, if X is a value small enough such as X < ¼, then the pair of strategies (*Cooperate*, *Cooperate*) is a fairness equilibrium. In fact, for small values of X the benefit received by defecting instead of cooperating is not so high and players could agree to help each other by cooperating. However, this does not take away that the Nash equilibrium (*Defect*, *Defect*) is a fairness equilibrium as well.

		Player 2	
		Cooperate	***Defect***
Player 1	***Cooperate***	(4X,4X)	(0,6X)
	Defect	(6X,0)	(X,X)

Figure 2.5 Matrix of Prisoner's Dilemma Rabin (1993)

Therefore, thanks to the principle of reciprocity on which the *Rabin (1993)* model relies on, the cooperation outcome may be an equilibrium in the Prisoner's Dilemma. The principle of reciprocity, contrary to rationality and self-interest criteria on which traditional game theory is based, allows to reach the Pareto optimality explained in *Chapter I*.

Another important aspect to underline and which is well represented in the redesigned Prisoner's Dilemma in *figure 2.7* is that players when evaluating others' kindness take into consideration if the gesture of help is made voluntarily or not. In this new Prisoner's Dilemma version, player 2 has no choice but to cooperate and player 1, aware of this mandatory kindness, prefers not to cooperate with player 2 but to defect obtaining the highest possible payoff to the detriment of player 2. Therefore, when evaluating the effect of players' intentions in a game, it is necessary to consider both beliefs and actual possibilities of the player.

		Player 2
		Cooperate
Player 1	***Cooperate***	(4X,4X)
	Defect	(6X,0)

Figure 2.6 Matrix of Prisoner's Non-Dilemma Rabin (1993)

2.3.2 Learning theory

A lot of economic studies have deepened the theory proposed by Rabin to the case of sequential games. One of the most important is the theory of sequential reciprocity developed by Martin Dufwenberg and Georg Kirchsteiger starting from *Rabin (1993)* model.

As underlined by Rabin himself, one limitation of his model is that it does not take into account how players' beliefs could change in a sequential game, since he studied games only in normal form[32]. Hence, the main goal of the theory of sequential reciprocity is to analyse how players' beliefs about

[32] The normal form is used to analyse simultaneous games through a matrix representation, while the extensive form is used to represent sequential games through a decision tree.

others' strategies can change in more complex dynamic games where new subgames could come up. The model is applied to finite multi-stage games in which at each stage players have knowledge of other players' previous choices. As in the *Rabin (1993)* model, π_i is the *material payoff function* of player i and A_i is the set of player i's strategies. Besides the material payoff, a *reciprocity payoff* dependent on a certain player's beliefs about other players' beliefs and strategies, must be considered to evaluate the utility of the player. Another new set which must be considered in this model with respect to the previous *Rabin (1993)*, due to the repetition of the stage game, is given by the set of histories H which lead to subgames. Each strategy assigns a probability distribution on player i's possible alternatives at h. To represent beliefs, the model defines the set of player i's beliefs about player j's strategy as $B_{ij} = A_j$ and the set of player i's beliefs about player j's beliefs about player k's strategy as $C_{ijk} = B_{jk} = A_k$. To take into account how players' perceptions may change across different histories, the *Martin Dufwenberg and Georg Kirchsteiger* model indicates with $a_i(h)$ the updated strategy which define the same choices of a_i apart from the choices of history h which present in this case a probability equal to 1. In the same way, the updated beliefs $b_{ij}(h)$ and $c_{ijk}(h)$ are defined.

Before defining the sequential reciprocity equilibrium, other important definitions about the equitable payoff, the kindness function and the utility function are necessary to be provided. First of all, it must be clarified that the equitable payoff belongs to the set of efficient strategies which is defined below.

$$E_i = \{\, a_i \in A_i \mid \textit{there exist no } a'_i \in A_i \textit{ such that for all } h \in H,\ (a_j)_{j\neq 1} \in \prod_{j\neq i} A_j, \textit{ and } k \in N \textit{ it holds that } \pi_k(a'_i(h), (a_j(h))_{j\neq i}) \geq \pi_k(a_i(h), (a_j(h))_{j\neq i}), \textit{ with strict inequality for some } (h, (a_j)_{j\neq i}, k) \,\}$$

The efficiency set of player i contains all those strategy whose correspondent material payoffs are always equal or higher to the material payoff of another strategy in every history and for every other player's consequent choices. The following equitable payoff of player j is calculated as a sort of average between the maximum and the minimum efficient material payoff of j.

$$\pi_j^{ei}((b_{ij})_{j\neq i}) = \frac{1}{2} \cdot [\max\{\pi_j(a_i, (b_{ij})_{j\neq i}) \mid a_i \in A_i\} + \min\{\pi_j(ai, (b_{ij})_{j\neq i}) \mid a_i \in E_i\}]$$

The equitable payoff is used as a reference to measure the kindness function of player i toward player j. If the strategy adopted by player i corresponds to a material payoff for player j which is equal to the equitable payoff of player j, it means that the kindness function of player i is zero. If the material payoff of player i is higher than the equitable payoff of player j, then the kindness function is higher than zero proportionally and the same goes obviously for lower payoffs.

The kindness function of player i toward player $j \neq i$ at history $h \in H$ is defined as follows.

$$f_{ij}(a_i(h), (b_{ij}(h))_{j\neq i}) = \pi_j(a_i(h), (b_{ij}(h))_{j\neq i}) - \pi_j^{ei}((b_{ij}(h))_{j\neq i})$$

To introduce the concept of reciprocity, that is to say the fact that if player *i* helps player *j* then player *j* wants to return the favour while if player *i* has been unkind then player *j* wants in turn to hurt the first player, another function must be defined. The following function measures player's $j \neq i$ kindness toward player *i* according to the belief of player *i* at history *h*.

$$\lambda_{ijk}(b_{ij}(h), (c_{ijk}(h))_{k \neq j}) = \pi_i(b_{ij}(h), (c_{ijk}(h))_{k \neq j}) - \pi_k^{ej}((c_{ijk}(h))_{k \neq j})$$

As usual, the main goal of individuals in a game is to maximize the utility function. In this model, the utility function of player *i* at history $h \in H$ is formulated as follows.

$$u_i: A_i \times \prod_{j \neq i}(B_{ij} \times \prod_{k \neq j} C_{ijk}) \rightarrow \mathrm{IR}$$

$$u_i(a_i(h), b_{ij}(h), (c_{ijk}(h))_{k \neq j})_{j \neq i}) = \pi_i(a_i(h), (b_{ij}(h))_{j \neq i}) + \sum_{j \in N \setminus \{i\}}(Y_{ij} \cdot f_{ij}(a_i(h),$$

$$(b_{ij}(h))_{j \neq i}) \cdot \lambda_{ijk}(b_{ij}(h), (c_{ijk}(h))_{k \neq j})$$

The utility of player *i* depends clearly on player *i*'s material payoff and also on his reciprocity payoff with regard to any other player $j \neq i$ which is represented by the second term. Y_{ij} is a constant non-negative number which indicates the sensitivity of player *i*'s toward player *j*'s reciprocity. One of the main properties which can be deducted from the utility function is that the utility function increases if player *i* is kind toward player $j \neq i$ who is kind in his turn toward player *i* and the same goes if player *i* is unkind toward player $j \neq i$ who is unkind in his turn toward player *i*. Another important property is that the utility function increases if player *i* is kind toward player $j \neq i$ who is kind in his turn toward player $k \neq j$ and the same goes for unkindness behaviour. From the concept of utility, it is possible to define the sequential reciprocity equilibrium. The strategy profile $a^* = (a_i^*)_{i \in N}$ is a sequential reciprocity equilibrium if the following conditions are true.

- $a_i^*(h) \in \operatorname{argmax}_{ai \in Ai(h,a^*)} u_i(a_i, (b_{ij}(h), (c_{ijk}(h))_{k \neq j})_{j \neq i})$
- $b_{ij} = a_j^*$ *for every* $j \neq i$
- $c_{ijk} = a_k^*$ *for every* $j \neq i$ *and* $k \neq j$

for every player $i, j, k \in N$ and every history $h \in H$. According to the first condition, players pursue strategies in order to maximize their utility given their beliefs. The second and third conditions define the correctness of the initial beliefs.

An application of the concept of sequential reciprocity equilibrium to the sequential prisoner's dilemma game, represented through the decision tree of *figure 2.8*, has also been provided by Dufwenberg and Kirchsteiger. They analysed behaviours of both players based on reciprocity parameters Y_{ji} and Y_{ij}. As regard player 2, if the first player defects, then also player 2 will defect to render unkindness to the first player. Instead, if player 1 cooperates, then player 1 behaviours depends on the value of Y_{21}. In particular, if $Y_{21} > 1$, that is to say reciprocity of player 2 toward player 1 is high enough, then player 2 returns the kindness toward player 2 by cooperating as well. If $Y_{21} < \frac{1}{2}$, then player 2 defects since his reciprocity parameter is not so high. If $\frac{1}{2} < Y_{21} < 1$, that is to say for intermediate values of reciprocity, player 2's behaviour depends on randomized probabilities.

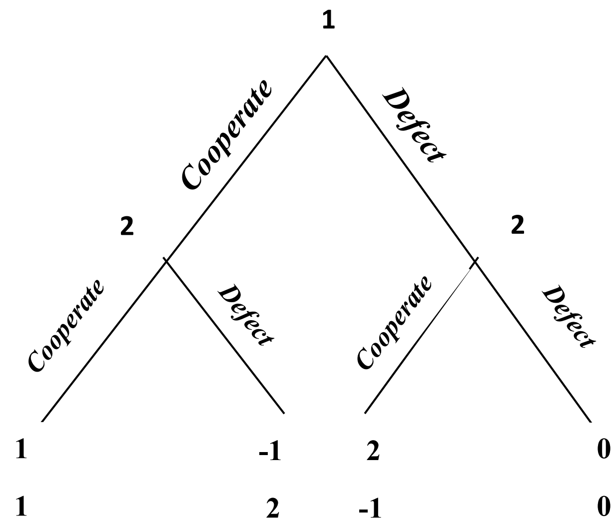

Figure 2.7 Sequential Prisoner's Dilemma decision tree

2.3.3 Further considerations on behavioural game theory

As underlined by the previous examples about the Prisoner's Dilemma, an important divergence between the outcomes of this game according to behavioural game theory and the traditional game theory is that in some circumstances the cooperative outcome is an equilibrium according to the economic models described in behavioural game theory. As already stated, experimental evidence is in line with the predictions of these behavioural game theory models and shows that players cooperate both in the not repeated Prisoner's Dilemma and in the repeated Prisoner's Dilemma.

Through the studies presented in this subchapter, behavioural economists have provided evidence that if traditional economic theory remains still firmly anchored to the principle of players' self-interest, it can only offer a partial and incomplete perspective of economic behaviours. Instead if traditional game theory incorporates in its original models the relational and reciprocal dimension typical of human being, a more complete representation of agents' economic behaviours can be provided.

2.3.4 Mental representations

Another important topic which must be considered when dealing with behavioural game theory is how players represent in their mind the game. Unlike the social utility functions and the learning theory which have been studied a lot in recent years, little research has been carried out with regard to players' mental representations[33]. This is mainly due to the fact that traditional game theory assumes that players have a common representation of the game[34]. However, observation of empirical experiments has put into evidence that sometimes different players mentally process games with the same form as if they were different or, on the contrary, process different games as if they were the same game. These examples can be enclosed in the concept of *framing* which has already been

[33] The concept of representation is meant as a player's internal cognitive representation of the elements of a game, on which decisional rules operate in order to produce a choice.

[34] In traditional game theory, games are represented through normal form by a matrix or extensive form by a decision-tree.

introduced in *subchapter 2.2*. According to framing, the same game described in two different ways can be played differently, or different games that can be represented in the same way can be played similarly. So, a frame is a knowledge structure which players use to construe a game and which may affect players' behaviour in a game. Moreover, in the so-called cases of editing, players may simplify games, for example by simplifying situations of decision-making under uncertainty or trying to simplify games with more than two players to make choices easier or, on the contrary, may enrich games by adding extra elements to the original game. Therefore, understanding how players represent mentally the games is crucial to apply empirically game theory.

2.3.4.1 Framing

As demonstrated in an experiment carried out by Colin Camerer, an expert in behavioural economics, some players exchange form of representation between the simultaneous ultimatum game and the Nash bargaining game. As already shown in *Chapter I*, in the ultimatum game a proposer offers a fraction x of the total quantity X to the responder who can accept or refuse the offer. If he accepts, players obtain respectively (X-x, x) while if he refuses, they obtain (0,0). In the Nash bargaining game, two players demand quantities x_1 and x_2. If the sum of the demand is lower than the total available amount, that is to say $x_1 + x_2 < X$, they obtain the requested quantities x_1 and x_2; otherwise, they both obtain 0. Usually, in the Nash bargaining game, players demand the 50% of the total amount

X in order to get the maximum benefit for themselves without damaging their opponent's benefit. This is mainly due to the fact that this game has not any asymmetry that allows to identify which player should get more. Instead, in the sequential ultimatum game, usually the proposers' offers hover around the 40% of the total amount and are accepted by responders. However, when the ultimatum game is played simultaneously and the responder have to indicate his minimum acceptable offer before knowing what will be effectively offered by the proposer, responders often indicate 50% of the total amount as the minimum acceptable offer. So, the interesting fact that emerge from this experiment is that in simultaneous ultimatum game responders act as if they were playing the Nash bargaining game. This happens because the asymmetry between the proposer and the responder which characterizes the sequential ultimatum game disappears when moving to a simultaneous ultimatum game.

2.3.4.2 Editing

Some examples of editing, in terms of enrichment of the starting game situation, are given by the focal points and by the fact that players may use time structure as a means of coordination.

The concept of focal point has been introduced by Thomas Shelling in his book "*The Strategy of Conflict*" in the context of coordination games. In this book, the author states that "*people can often concert their intentions or expectations with others if each knows that the other is trying to do the same*". According to

Shelling, the way in which strategies are labelled in coordination games can affect their probability to be chosen as focal. For example, it has been demonstrated that in a game in which two players can receive a benefit if they choose the same flower, 67% of players chooses the rose, playing a strategy that is commonly known and which requires explicit social cognition. Therefore, focality is considered as an enrichment of players' mental representation because ensures that not only experimental instructions given to players but also players' previous social knowledge affect the way in which the game is played.

As regard time structure, it is useful to remember that in the traditional game theory of von Neumann and Morgenstern, the information held by players about others' actions was considered more important than the timing of moves. The idea was that if a player had moved before and the second payer to make the move did not know first player's action, this was psychologically equivalent, from the point of view of the second player, to move simultaneously. However, some empirical experiments have proved clear evidence of the impact of this factor on players' decisions assuming information of all players as constant. In particular, *Cooper et al. (1993)* analysed a new version of the BoS game by comparing the different outcomes of this same game when played in a sequential or simultaneous way, as shown in *figure 2.4*. Considering the game played simultaneously, according to *Cooper et al. (1993),* players choose their preferred strategy *B* approximately 63% of the time. To compare these outcomes with the ones of the sequential case, they considered a game in which Player Row moves first and both players are aware of the sequentiality of these moves even if Player Column cannot know which action has been pursued before by Player Row. In traditional game theory, the decision trees describing this sequential game and the previous simultaneous game would be the same. According to the theory of von Neumann-Morgenstern, it is not necessary to mark temporally the knots of the tree because the fact that Player Row moved first should not matter to Player Column since he does not know which action was pursued by the first player. However, according to the experiments of *Cooper et al. (1993),* when both players know that Player Row moves first, Player Row chooses strategy *B,* that is to say his preferred strategy, 88% of the time; instead Player Column chooses his preferred strategy *B* only 30% of the time. This means that Player Column seeks to meet the preference of the player who moved first by choosing strategy *A* 70% of the time. One of the most well-known theories which explain why the first mover in a game can gain an advantage over the other player, is the virtual observability proposed by *Camerer, Knez and Weber (1996).* According to the virtual observability, individuals think deeper about events happened in the past than events not yet occurred. So, if the first mover expects the other player to think more carefully about his strategy, then he can take advantage of this tendency and this time structure is useful to create a sort of coordination among players.

		Player Column		***Choices Frequencies***	
		A	***B***	***Simultane-ous***	***Sequential***
Player Row	***A***	(0,0)	(2,6)	38%	12%
	B	(6,2)	(0,0)	62%	88%
Choices Frequen-cies	***Simultane-ous***	35%	65%		
	Sequential	70%	30%		

Figure 2.8 Matrix of BoS Cooper et al. (1993)

2.3.5 Considerations about behavioural game theory

As concern editing in terms of simplification, there are some experimental games which demonstrate that players, in games with more than two players, often simplify their representation by considering all their opponents as a single player. In some cases, this type of simplification can be useful, but in other cases can lead to systematic errors.

As underlined some years later by Andreoni and Samuelson in their paper *Building Rational Cooperation (2006),* the results obtained through these models, ‚enclosed in the field of behavioural game theory, broaden the notion of rationality used in traditional game theory, In fact, these models are capable of explaining cooperation behaviours without stating that this kind of behaviour is irrational. In order to do this, these models modify the structural hypothesis of rationality common knowledge and the notion of preference. The predictions made by behavioural game theory models came closest to the empirical results observed during the effective playing of the game by individuals than predictions made by traditional game theory models.

For example, thanks to the principle of reciprocity on which the *Rabin (1993)* model and the model rely on, according to these models the cooperation outcome may be an equilibrium in the Prisoner's Dilemma. Hence, the principle of reciprocity, contrary to rationality and self-interest criteria on which traditional game theory relies on allows to reach the Pareto optimality.

Through these studies, behavioural economists have provided evidence that if traditional economic theory remains still firmly anchored to the principle of players' self-interest, it can only offer a partial and incomplete perspective of economic behaviours. Instead if traditional game theory incorporates in its original models the relational and reciprocal dimension typical of human being, a more complete representation of economic behaviours can be provided.

CHAPTER III

Neuroscience's application to Game Theory

3.1 How neuroscience can impact game theory

In the previous *Chapter II,* the presentation of the studies sought by game theory experts about economic agents' behaviors models concerning altruism and spitefulness responses to their partners' actions or the players' will to play the game fairly and with the aim to contribute to an equal payoffs' distribution among all individuals, have put into evidence the limitations of the meaning of *perfect rationality* and *self-interest* assumed by traditional game theory. However, as underlined by some behavioral economists, even if players' behaviors predicted by the utility functions models of behavioral game theory are more convergent to empirical evidence than traditional game theory models, these behavioral game theory models are not sufficient to understand why such social strategies are pursued by economic agents. In order to really understand the cognitive mechanisms underlying certain players' behaviors, application of neuroimaging methods to the behavioral game theory sector turns out to be fundamental. In fact, neuroscience techniques allow to identify the activation of brain regions which are connected to the decision of an individual to pursue a particular social behavior or which are directly connected to the accomplishment of these social behaviors. Therefore, the application of the nascent neuroeconomics field[35] to game theory represents a step forward in the development of the game theory sector. For this reason, one of the main focus of the recent behavioral economics research is

[35] Neuroeconomics has been developed only in recent years. In fact, the first conference about neuroeconomics has been held only in 1997 at the Carnegie-Mellon University.

represented by finding the neurobiological roots at the basis of behaviors during economic games.

When considering the way in which neuroscience could inform game theory, it is necessary to take into consideration two approaches: the radical and the incremental approach. The radical neuroscientific approach takes a step back to the origin of the traditional economic theories and retraces its development considering the neuroscience influence. Instead, in an incremental perspective, neuroscience takes the existing traditional economic models as the starting point to which adding new variables and excluding those economic assumptions which lack of empirical veracity.

At this point, before starting with the application of neuroscience methods to experimental economic games, it is useful to take a step back and provide a brief explanation about what cognitive neuroscience and neuroeconomics are.

3.1.1 Cognitive neuroscience

Cognitive neuroscience is a branch of neuroscience whose scientific development has begun during the eighties of the last century thanks to the publication of the *Journal of Cognitive Neuroscience* and which analyses the cerebral functions connected to the human thought. Both cognitive psychologists such as G.A. Miller and S.M. Kosslyn and neuroscientists such as A. Damasio and J. Ledoux, who proposed new theories of the mind[36] with particular reference to the interweaving of emotional and cognitive factors, have contributed to the progress of cognitive neuroscience.

Neuroscience research has put into evidence the role of emotions as one of the primary sources from which individuals' actions and behaviors come out. Moreover, neuroscientific studies have underlined the agents' unawareness during decision-making and their incapacity to detect all typologies of mechanisms leading to their subsequent actions. Contrary to this new point of view, traditional economic theory is based on agents' rational and conscious processes: hence, it fails to grasp those mechanisms regulated by emotions that graft outside the level of individuals' consciousness. In this sense, the cognitive neuroscience domain can be useful to investigate what really happens in players' brain when economic games are submitted to them.

3.1.2 Neuroeconomics

As regard the neuroeconomics domain, the term "neuroeconomics" identifies an interdisciplinary research which combines cognitive neuroscience, economics and psychology and which contravenes and, at the same time, allows the progress of the neoclassical economics, according to which economic agents during

[36] By definition, theory of the mind, usually abbreviated with ToM, is the capability of an individual to ascribe mental states, such as intentions, desires emotions and beliefs to oneself and to other individuals.

decision-making are restricted by perfect and formalizable rationality constraints. Neuroeconomics is a very recent branch of behavioral economics_which has started to expand only in the nineties of the twentieth century when there has been some progress in the research about brain functioning and the relationship between agents' behavior and the corresponding brain activity. The innovation of the neuroeconomics approach lies in the use of cognitive neuroscience methods, which will be analyzed more in details in the following subchapter, in order to build economic models convergent as much as possible to empirical evidence. By adopting this new perspective, economic models should be able to explain real cognitive processes during decision-making in such a way that there is no longer the distinction between *homo oeconomicus* and *homo neurobiologicus*.

3.2 Brain anatomy

Before describing in more details those brain areas which are most relevant for the analysis of players' economic behaviors, it is useful to provide a brief and general introduction about the brain anatomy and its main functions. It is also important to specify that neuroanatomy studies the anatomy of the brain, while the scope of neuroscience, is to analyze the brain regions' functions.

The brain is the largest human organ and it is protected by the skull bones. Its main function is to coordinate all the other organs and systems' activities through the nervous system. It is composed by three parts: the cerebrum, composed by the diencephalon and the telencephalon, the brainstem and the cerebellum. The cerebrum is divided into two symmetrical hemispheres: the right and the left hemispheres. The cerebrum external surface, and more precisely the cerebral hemispheres, is covered by the cerebral cortex which is the peculiar gray matter, that is to say a nervous tissue at high content of neurons. The cerebral cortex represents approximately the 42% of the entire cerebral mass and its thickness varies between the two and the five millimeters. It includes approximately 16 billion neurons and 300 trillion synapses[37]. The cerebral cortex is characterized by a very particular macro-architecture: it consists of an alternation between deep grooves, called more properly sulcus, and fold or ridge, called gyrus. The cerebral cortex plays a pivotal role in the control of mental cognitive functions and it is the main neural information processing and integration center of the central nervous system. In particular, several parts of the cerebral cortex which are involved in behavioral game theory experiments are contained in the limbic system[38].

[37] Synapses are sites of functional contact between two neurons. These connecting points allow the transmission of information in the form of electrical signals.

[38] The limbic system is a group of structures of the telencephalon, which regulates motivated behaviours, that is to say decision-making, emotional processing and executive functions Among these structures the thalamus, the hypothalamus, the hippocampus, the amygdala, which plays a crucial role in emotions, are particularly important. The limbic system includes cerebral structures that have a key role in emotional reactions and behavioural responses and so, some of its brain structures are very interesting when analysing social decision-making.

By convention, the cerebral cortex of each brain hemisphere is divided into four major areas, called brain lobes:

- frontal lobe
- temporal lobe
- parietal lobe
- occipital lobe.

Besides the identification of the different areas of the cerebral cortex through the four lobes of each hemisphere, the cerebral cortex can also be divided in different cortical regions based on the type of function performed. These different cortical areas are:

- prefrontal cortex, related to emotions and problem-solving
- associative motor cortex, related to complex movements coordination
- primary motor cortex, related to voluntary movements
- associative sensitive cortex, related to processing of sensitive information
- primary somatosensory cortex, related to sensitive information recognition
- associative visual cortex, related to visual information processing
- visual cortex, related to recognition of simple visual stimuli
- Wernicke's area, related to language understanding
- associative auditory cortex, related to auditory information processing

- auditory cortex, related to recognition of sound quality
- inferior temporal cortex, memory-related processing
- Broca's area, related to speech production

The concept of function localization does not mean that a function is performed exclusively by a certain area since most functions are performed by neurons from different brain regions. What is important to underline is that certain areas have a closer relationship to certain functions than others. Thus, each area is designated primarily to perform a specific function.

3.2.1 Brain areas important in economic experiments

As it will be deducted from neuroscience experimental games described in the following subchapter, the brain areas which are most relevant when dealing with behavioral economics are the striatum and some areas of the cerebral cortex.

In the behavioral game theory domain, the cortical area which is more relevant in a functional point of view is obviously the prefrontal cortex, which include also the orbitofrontal and the dorsolateral cortex. In this subchapter, these brain areas, which will turn out to be important in economic games analyzed throught neuroimaging techniques support below, are described in an anatomical and functional point of view.

3.2.1.1 Cerebral cortex

The prefrontal cortex (PFC) is an area of the cerebral cortex located on the anterior part of the frontal lobes. It is considered an associative polymodal area, since it receives cortical afferents from almost all other cortical areas, such as the thalamus[39] and from several subcortical structures, among which the most important is the limbic system. The prefrontal cortex plays an important role in the so-called executive functions, such as the anticipation, goal selection, planning of strategies, monitoring, attention, concentration and self-control of impulses and emotions. The prefrontal cortex is divided, in a functional point of view, into three structures: dorsolateral, medial and orbital. The dorsolateral prefrontal cortex (DLPFC) is responsible for the organization and planning of complex behaviors and of high-level cognitions. Hence, the DLPFC is involved in all those functions enclosed in the cognitive sphere. It is connected to the orbitofrontal cortex, the thalamus, the dorsal caudate nucleus[40] and other brain structures. The medial prefrontal cortex plays a role in both cognitive and emotional motivation. According to several cognitive neuroscience studies, the medial prefrontal cortex is involved in theory of mind processes, that is to say the subjects' capability to attribute mental states to oneself and others. This function of the medial prefrontal cortex

[39] The thalamus is a structure which composes together with the hypothalamus the diencephalon.

[40] The caudate nucleus is a component of the basal ganglia. The basal ganglia, such as the structures of the limbic system, are located on the subcortical area of the telencephalon. Among the most important components of the basal ganglia there are the dorsal striatum, composed by the caudate nucleus and the putamen and ventral striatum composed by the nucleus accumbens and the substantia nigra which produces dopamine (DA), a neurotransmitter which regulates motor and reward system in the striatum.

will be taken into consideration when analyzing the neuroimaging experiment about the trust game in the following subchapter. The function of the orbital prefrontal cortex is of regulating and inhibiting, if necessary, the elaboration of stimuli interfering with the current task, as well as the function of controlling impulses. According to recent studies, the prefrontal cortex is associated to decision-making processes. The prefrontal cortex performs all these functions in association with the thalamus and the basal ganglia, forming the so-called frontal-subcortical circuits.

Another part of the cerebral cortex associated to the limbic system and whose functions are relevant in social decisions is the insular cortex or insula. The insular cortex is located deep in the brain between the temporal and the frontal lobes and is divided into the anterior insula, more relevant for the scope of this thesis, and the smaller posterior insula. The insula is usually associated to negative emotions. In particular, it is related to pain and basic emotions such as anger, disgust, fear and sadness. Moreover, it is generally related to sensations of thirst, hunger and disgusting odor or taste.

Another part of the cerebral cortex included in the limbic system and which will turn out to be important in the neuroscience games experiments analyzed thereafter is the anterior cingulate cortex[41] (ACC). Both in the ultimatum game and in the iterated Prisoner's Dilemma experiments illustrated thereafter, the anterior cingulate cortex plays the role of a sort of "*mediator*" in the conflict between the

[41] The cingulate cortex is located in the medial area of the cerebral cortex and is divided into the anterior cingulate and the posterior cingulate cortex.

emotional and the cognitive spheres during economic agent's decision-making. As it can be deducted from its name, the anterior cingulate cortex is located in the frontal part of the cingulate cortex and is divided by experts into dorsal and ventral anterior cingulate cortex. The dorsal component, which is linked to the prefrontal cortex is associated principally to cognitive aspects, while the ventral component, which is linked to the amygdala, the nucleus accumbens, the hypothalamus, the hippocampus and the anterior insula is associated to emotional aspects. Moreover, ACC is associated to the response to pain, both related to physical sensations and negative social events.

3.2.1.2 Striatum

The striatum is a subcortical part of the telencephalon, which is located on the frontal lobes. Its name derives from its particular structure composed by alternate layers of grey matter and white matter. It is, as already stated, an important input component of the basal ganglia. The striatum receives input from many areas of the brain beyond the ganglia of the base, but sends the output only to other components of the basal ganglia. The striatum is composed by the caudate nucleus, which is part of the dorsal striatum, and the nucleus accumbens, which is part of the ventral striatum.

The caudate nucleus has a large and wide head which dwindle in a thin tail, forming a particular C shape. Obviously, each hemisphere contains one caudate nucleus, which is located close to the thalamus and deep in the brain. As regard the functions of the caudate nucleus, which is one of the most important component of the striatum when dealing with game theory experiments in the neuroscience domain, it is mainly associated to goal-directed action that is to say the capability of an individual to take a particular action, whose result in known by the decision-maker, based on the goal that he wants to achieve. The caudate nucleus can be defined as a sort of feedback processing because it is involved in decisional processes about actions to be taken based on past information available about past occurred events.

3.3 Neuroimaging techniques

The scope of this subchapter is to provide a short presentation about the neuroimaging techniques which are mainly used in neuroeconomics in order to identify which agents' brain areas present a major activation during social decision-making.

It is important to point out that scientific technologies are not only useful tools for the research progress in the field in which they were originally conceived, but also for the advancement of research in other fields. In this regard, neuroimaging techniques initially applied only to neuroscience domain are also more recently applied in behavioral economics. For the sake of simplicity, neuroimaging methods are usually distinguished into two categories: those methods that identify a certain brain region's activation by measuring neurons electromagnetic activity and those methods that identify the brain region of interest through variations in the level of blood.

3.3.1 Techniques based on electromagnetism

Among the most important techniques based on electromagnetism there are the electroencephalography and the magnetoencephalography.

3.3.1.1 Electroencephalography (EEG)

The electroencephalography technique is used to measure the electric activity of the brain. A number of electrodes which can vary from ten to twenty is put on the person's scalp along five lines, according to the standard *International 10-20 system*. The numbers ten and twenty of this system indicate the position on the scalp of the electrodes which must be put at a distance of 10% or 20% of the total distance between the anterior and the posterior part of the skull or between the right and the left part of the skull. The function of the electrodes is to detect the potential difference caused by the electric activity of the neurons present on the cerebral cortex and related to individuals' behavioral responses. The brain response to a certain stimulus, for example a cognitive or a motor stimulus, is called *event related potential*, abbreviated as ERP. EEG is the oldest brain imaging technique and has been invented in the early twentieth century by Hans Berger, from which the name Berger rhythm or alpha rhythm, which indicates the base frequency in EEG, derives. The temporal resolution of EEG is very accurate. In fact, through EEG, it is possible to identify neural activity in a time lapse of the order of few milliseconds. Unfortunately, the spatial resolution is not very advanced due to the so-called inverse problem[42].

3.3.1.2 Magnetoencephalography (MEG)

The magnetoencephalography is a technique complementary to the EEG since, contrary to the EEG which measures the electric activity of neurons flowing perpendicularly to the scalp, it measures the electric activity of neurons flowing parallelly to the scalp. Its name derives from the fact that the brain electric activity is measured through variations of the magnetic fields. To measure these magnetic changes, extremely sensitive magnetic sensors, made with superconducting circuits, are used. These sensors, indicated with the acronym SQUID (Superconducting quantum interference device), are able to measure even very small variations of magnetic field. MEG is a more recent technique than the EEG and it has been used the first time only in the second half of the twentieth century. As the EEG, the MEG method is very precise as regard the temporal resolution[43]; instead, the spatial resolution is not very accurate due, even in this case, to the inverse problem.

[42] EEG inverse problem consists in identifying the signal of the brain area (and the subsequent brain area of interest) starting from the final EEG data measured. The issue of the inverse problem is that there is not a unique solution to an inverse problem.

[43] MEG is able to detect the brain neural activity in only 10 milliseconds.

3.3.2 Techniques based on hemodynamic reactions

As regard the techniques based on hemodynamic reactions the most well-known ones and useful to describe for the scope of this thesis are the positron emission tomography and the functional magnetic resonance imaging.

3.3.2.1 Positron emission tomography (PET)

The Positron emission tomography (PET) technique is used to measure blood flow in the brain. In fact, blood flow is correlated to neural activity since the increase of neural activity in a certain brain region produces an increase of blood flow in the same region. When applying this technique, the first step to be taken is to inject modified molecules, called radio-nuclides, which emit positrons. The collision of the positrons with their antiparticles, that is to say the electrons which are already present, lead to a radiation to which the PET scanner is sensitive to.

3.3.2.2 Functional magnetic resonance imaging (fMRI)

The functional magnetic resonance imaging is the most recent and most frequently used neuroimaging technique. fMRI detects blood flow in the brain caused by variations in magnetic properties due to the oxygenation of the blood, the so-called *blood oxygen level dependent imaging*, abbreviated with the acronym BOLD. More specifically, through the BOLD effect, fMRI is able to reveal changes in hemoglobin's[44] magnetic properties caused by variations in the level of oxygen. In fact, the oxy-hemoglobin is diamagnetic while the deoxy-hemoglobin is paramagnetic. When a particular tissue of the brain is activated, the blood flow exceeds the oxygen consumption; hence the deoxy-hemoglobin replaces part of the oxy- hemoglobin leading to a resonance signal which is measured by MRI scanner. After that, MRI scanners represent tomograms[45] of the brain areas of interest. Contrary to EEG and MEG, fMRI and PET can only detect variations in the blood level with a precision of some hundreds of milliseconds in case of fMRI and of one minute in case of PET. However, as regard the spatial resolution, both PET and fMRI are more accurate than the two neuroimaging methods based on electromagnetism.

3.3.3 Goal of neuroimaging techniques

It is important to underline that the goal of the cognitive neuroscience applied to the game theory sector is not simply to localize which area of players' brain present a higher activity when having a particular interaction with their partner, as it could be thought through the analysis of the neuroimaging techniques. The goal of neuroscience is to localize the brain regions of interest correspondent to the economic agent's behavior in order to elucidate all those emotions which arise in

[44] Hemoglobin is a globular protein which transports oxygen.

[45] In radiology, the tomogram is the image obtained through tomography, that is to say a diagnostic technique whose main aim is to analyze individual planes of an organ.

players when they pursue social strategies predicted by social preferences models.

3.4 Neuroscience experimental games

In this subchapter some neuroimaging experiments about economic games are reported. In particular, there will be the presentation of the ultimatum game, an economic-exchange game about altruistic punishment, the trust game and the Prisoner's Dilemma, which have been already described in *Chapter I* according to the standard game theory point of view, in order to put into evidence the differences between players' empirical behaviors and players' behaviors predicted by traditional game theory. In the first two games presented, the brain structures associated to fairness social behavior, such as the refuse of an unfair offer by the responder in the ultimatum game and the punishment of unfair actions, is investigated. In the last games the focus is instead on the brain structures associated to cooperative actions, and specifically cooperation based on reciprocity. These game experiments, carried out with the support of neuroimaging techniques, represent an empirical demonstration of the utility functions' behaviors predictions that have been reported in *Chapter II* and which include also emotional components when evaluating decision-making of economic agents. The reason why experimenters apply to the games analysed thereafter neuroimaging techniques such as fMRI and PET is due to their non-invasive nature. Moreover, fMRI and

PET methods allow to create maps of functional connectivity of brain regions whose activation is temporally correlated, called "functional networks".

3.4.1 Ultimatum game in neuroscience

An important study which has been made in the neuroscience domain to better understand which responder's brain regions are activated when responder takes the decision to accept or refuse proposals in the ultimatum game is "*The Neural Basis of Economic Decision-Making in the Ultimatum Game"* by *Sanfey et al. (2003).* The neuroimaging technique used by the researchers in this experiment is the fMRI. The main results from this experiment are that during the ultimatum game there is, in case of unfair offers by the proposers, an activation of the responder's anterior insula, related principally to negative emotions and of the dorsolateral prefrontal cortex, related to cognition and rationality[46].

Before getting to the heart of the experiment, Sanfey and its team underline once again the problem of the divergences between the outcomes observed empirically and the outcomes predicted by traditional game theory when playing the ultimatum game. In fact, according to recent empirical observations carried out by behavioral economists, proposers offer in most cases a split of money of 50% of the total amount and responders reject offers which are lower than the 20% of the

---

[46] Rationality must be considered in the sense given by traditional game theory. In fact, when an unfair split is offered to the proposer, its negative emotional reaction which lead him to refuse the offer comes into conflict with its rationality to accept a little sum of money which is, in any case, higher than zero.

total amount approximately 50% of the time because they consider these offers unfair. The fact that these results occur also when the responder knows that there will not be a future game interaction with the proposer, is a further demonstration of the importance given by economic agents to behave according to fairness norms. According to the economic models of *Fehr and Schmidt (1999)* and *Bolton and Ockenfels (2000)* about inequity aversion, which suppose that players want to distribute payoffs equally and fairly among them, it would be reasonable that responders refuse unfair offers by proposers even if, by accepting the unfair offer, they could obtain an amount of money higher than zero.

The experiment consisted of 30 rounds of the ultimatum game in which in each round, the responder was offered randomly by a human or a computer proposer a split of 5$:5$, 7$:3$, 8$:2$ and 9$:1$ starting from the total amount of 10 $. Obviously, the split 5$:5$ was always accepted by responders as well as 7$:3$ in most cases. So, the attention of researchers directed toward responders' brain areas' activation for the unfair offers 8$:2$ and 9$:1$. A first evident observation concerned the fact that unfair offers from human beings were rejected more frequently than unfair offers from computer. This evidence was important to confirm that, when evaluating an offer, responders do not take into account only its fairness, that is to say that the split offered was fair, but also the proposer's intentionality, as underlined by the theory of fairness intention-driven of *Rabin et al. (1993)* which presents an higher weight of influence on responder's decision in case the proposer is represented by a human person than a machine. Scanning from the fMRI revealed that the responder's brain areas of major activation in case of unfair offer were three: the bilateral anterior insula, the dorsolateral prefrontal cortex (DLPFC) and the anterior cingulate cortex (ACC). According to other neuroimaging experiments, the anterior insula is related to negative emotions, first of all pain, but also basic emotions such as sadness, anger, distress, and even hunger and thirst sensations. As a confirmation of this, there is the fact that in this experiment the intensity of anterior insula increased when passing from 8$:2$ to 9$:1$ offer. Moreover, rejecting low and unfair offers is also correlated to responders' will to preserve their social reputation. Contrary to the anterior insula, the DLPFC is associated to the cognitive[47] sphere which in this case is represented by the responder's decision to accept even low offers, as predicted by traditional game theory. To demonstrate this, the experiment shows that the activity of DLPFC is higher than the one of the bilateral anterior insula in the case in which an unfair offer is accepted by the responder. This is due to the fact that the acceptance of unfair offer requires an effort from responder to put aside his negative emotions in order to rationally get the little reward. Moreover, another responder's cortical area whose activation increased in case of unfair offers was the anterior cingulated cortex. The researchers have defined the ACC as a sort of "mediator" in the conflict between the emotional and the cognitive spheres.

[47] In this context, the cognitive adjective replaces the rational one to put into evidence that the responder's willingness to refuse an unfair offer is not necessarily an irrational behavior, as believed by traditional game theory. Stated another, responder's rejection of an unfair offer in order to chastise proposer could have a background of rationality.

3.4.2 Altruistic Punishment in neuroscience

In the field of behavioural game theory, social preferences models have defined utility functions according to which players can punish partners' unfair behaviours or reward partners' cooperative behaviours. An example of altruistic punishment, that in game theory corresponds to players who give up to their own benefit to punish partner's unfair behaviors, has been already illustrated in *Chapter II* as a possible fairness equilibrium. The following experiment shows which players' brain areas are involved in altruistic punishment behaviors.

In their economic experiment contained in the paper "*The Neural Basis of Altruistic Punishment*", *De Quervain et al.* used PET neuroimaging technique to scan players' neural activity during an economic exchange game in which players decide to punish defectors present particular brain areas activation. Since altruistic punishment is based on the individual's intention to give up to part of its own benefit in order to punish the defector, the hypothesis underlying the experiment were that individuals get contentment when rightly punishing defectors. In fact, it was observed during the game that when a player did not respect conventional social norms of fairness and cooperation, his partner had in most cases the desire to punish his unkind behavior even at the cost of losing part of his benefit.

The economic game analyzed by experimenters was a two-player game. Before starting the game, a total amount of 10 monetary units (MUs) was provided to both players. In the first step, player 1 could decide whether to keep to himself the 10 MUs or give them to the other player. If the first player decided to trust his partner by providing to him the 10 MUs, player 2 would find himself with a total sum of 50 MUs[48] (the sum of the 10 MUs received at the beginning and the amount provided by player 1 which is quadrupled by the experimenter). In this case, the social norms of fairness and reciprocal cooperation implicitly laid down the game. At this point, player 2 had the possibility to send back half of the total 50 MUs to the first player or keep all the money to himself. In case player 2 decided not to reciprocate the trust of the first player acting selfishly, player 1 had the possibility to punish him, by evaluating his unfair behavior on a scale of twenty points.

According to the experiment, there are four possible conditions in which the first player can find himself during the game.

- Intentional and costly condition (IC)
- Intentional and free (IF)
- Intentional and symbolic (IS)
- Nonintentional and costly (NC)

In the IC condition, if player 1 decided to punish player 2 of one point, the total amount of money of the punished player was reduced by two MUs while the amount of the punisher was decreased by one MU. Hence, the punishment pursued by player 1 represented a cost for him. Instead, the condition IF did not have any cost for player 1, but led to a reduction of two MUs for player 2 for each

[48] The sum of the 10 MUs received at the beginning and the amount provided by player 1 which is quadrupled by the experimenter.

point assigned by the punisher. In the IS, the punishment had only a symbolic meaning and costs nothing to both players. Finally, in NC, the decision to keep the 50 MUs or send part of it back to the trustee was delegated to a device which made the choice randomly and the cost of punishment corresponded to the IC condition.

By comparing all these conditions, experimenters found that the caudate nucleus presented a higher activation in all those situations in which player 1 had the desire to punish the defector and could effectively punish him in monetary terms. Hence, the activation of the caudate nucleus was higher in IC and IF conditions in which punishment of defectors was effective with respect to the IS condition in which the punishment was only symbolic. Moreover, the higher the punishment intentionally inferred which led to higher satisfaction to the punisher, the higher the activation of this brain area. The association of the caudate nucleus to reward derived directly from goal directed-action, suggests that players derive a sort of reward from effective punishment. Another brain area which presented an increase of BOLD activation, even if with minor contribution with respect to the caudate nucleus, when there was a high desire of player to effectively punish defectors in IC and IF conditions was the thalamus, associated also to reward processing. Moreover, the importance the orbitofrontal cortex in the correlation between decision-making and separate cognitive processes was put into evidence by the activation of these regions in the IC condition, in which the player faces a trade-off between deriving satisfaction from effective punishment or avoiding the punishment cost.

3.4.3 Trust game in neuroscience

Besides the ultimatum game and the economic exchange game about altruistic punishment, another economic game which has been the object of several analyses in the nascent field of neuroeconomics is the trust game. Contrary to traditional game theory predictions, according to which the Nash equilibrium in the trust game is given by the investor's strategy of keeping the total amount of money for himself and the subsequent strategy of the second player to return nothing, empirical experiments of trust game have demonstrated that in most cases investors put their trust in trustees by investing a part of the total amount of money available. On his part, the trustees reciprocate in most cases the favor by returning to the investor an amount higher than the one got. Moreover, it is important to underline that the expression and the reciprocity of trust is very important when analyzing cooperative behavior because it is one of the major actors which usually affects cooperation among players.

The first neuroscientific experiment about trust game has been analyzed through the use of the fMRI technique by *McCabe et al.* in 2001. The aim of this experiment is to demonstrate the involvement of the prefrontal cortex, and more specifically of the medial prefrontal cortex, in the players' decision to cooperate due to their ability to attribute mental states to their partner which lead both of them to obtain a higher outcome by cooperating. The following two-person decision tree represents the trust game proposed by the experimentalists to the players.

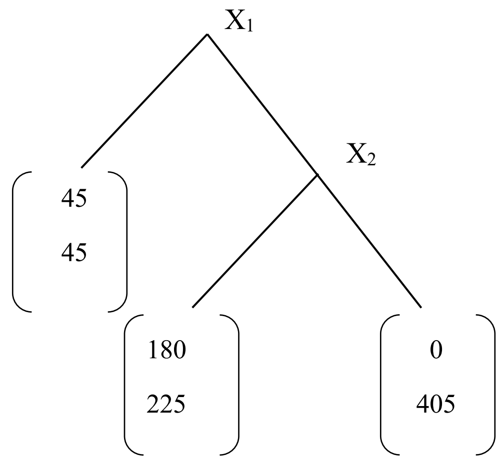

Figure 3.1 Decision tree of the trust game

The first decision maker could decide whether to distribute payoffs equally with his partner or to trust the other player by leaving to him the decision of the payoffs' distribution. In this second case, player 2 could decide whether to reciprocate the trust given by the first player or to take selfishly the advantage to gain the highest payoff. Experimenters observed that the first player decided to play the trusting move 50% of the time and the second player reciprocated its kindness by cooperating with him 75% of the time. The partner of the first or the second player could be identified by another human person or a computer. In this last case, the player was informed that the computer was already preprogrammed to play the game with the fixed probability previously mentioned respectively for the role of first player and second player. It has been noted that, in the case in which the first player decided to take the trusting move by shifting right and the second player decided to reciprocate its kindness leading to the cooperative outcome (180, 225), the prefrontal cortex of players was more active when their counterpart was represented by another person than a computer. This difference in the activation of the prefrontal cortex put once again into evidence that players' convergence to cooperation is higher in case of human partners. Instead, if there is not the intention to cooperate, the activation of the prefrontal cortex is low in case of both human and computer counterpart. Therefore, the conclusion about this experiment was that the correlation between players' ability to attribute mental states to oneself and to their partners and players' cooperative behaviors involve the activation of players' prefrontal cortex.

3.4.4 Repeated trust game in neuroscience

As already showed in *Chapter II* when dealing with the learning theory, when the trust game is played more than once, learning from early iterations can affect future players' behavior. An example of the association of the influence of reciprocity factor on other player's intention to trust to the corresponding brain region of interest activation, is provided by "*Getting to Know You: Reputation and Trust in a Two-Person Economic Exchange*" by *King-Casas et al. (2005)*, one of the most famous experiment studied in the neuroscientific domain about trust game.

In this experiment, experimenters decided to use the hyperscan[49] fMRI technique in order to measure neural activity of both players at the same time. The choice of this type of neuroimaging technique was due to the fact that, especially in sequential repeated games, drawing conclusions by analyzing neural activity of only one player at a time is incomplete and may lead to ignore some important neural processes underlying a certain player's behavior.

The trust game was played 10 consecutive times by a relative high number of players, N=48. In each iteration, the investor decided how much of the total sum of 20 $ to invest, that is to say which portion of the total 20 $ to give to the trustee. Then, the trustee could choose how much of the tripled portion invested previously by the first player to give back to the investor. Since the trust game was repeated, it was possible in this specific game experiment to observe trust behaviors of both trustee and investor, in contrast to the previous one-shot trust game experiment in which only the expression of trust by the investor could be observed. Moreover, it was possible to observe the divergences between standard game theory and this empirical game experiment about reputation building in a finite repeated game. An important discovery made in this experiment was, as already anticipated, that the player's reciprocity (by investor) has turned out to be the factor with the highest influence on the subsequent degree of trust expressed by the other player(trustee). Therefore, reciprocity could be considered as a social signal which predicted trust behaviors. In this specific experiment,

[49] Hyper scanning is a technology which allow to scan simultaneously brain activity of multiple subjects.

King Casas et al. defined the reciprocity factor as "*a fractional change in money sent across rounds by one player in response to a fractional change in money sent by their partner*". In particular, the reciprocity of the investor at iteration *j* is defined in mathematical terms as $r_j = (\Delta I_j - \Delta R_j)$, where ΔI_j represents the difference between the amount invested by the investor in the current round *j* and the amount invested in the previous iteration (*j* – 1) and ΔR_j is defined in mathematical terms as $\Delta R_j = \Delta R_{j-1} - \Delta R_{j-2}$ and measures the difference between the repayment made by the trustee toward the investor at iteration (*j* – 1) and at iteration (*j* – 2). For the sake of simplicity, researchers have distinguished three different types of reciprocity: the benevolent reciprocity, which indicate a bold investment by investor in response to lower trust degree manifested by trustee, the malevolent reciprocity, which represents the opposite situation of the benevolent reciprocity, that is to say the investor return an amount of investment lower than the one just received by the trustee, and the neutral reciprocity, according to which the sum invested by both investor and trustee does not change over the rounds. Researchers observed that benevolent reciprocity of investor was rewarded by trustee with a larger amount of investment in the subsequent round and in the same way, investor's malevolent reciprocity was punished by trustee with a lower amount of investment, as a demonstration of the fact that reciprocity by investor r_j affected effectively subsequent trustee trust behavior ΔR_j.

By comparing these types of behavioral reciprocities, a first observation which has come out was that the blood oxygenation level-dependent (BOLD) in the trustee's brain was higher in response to benevolent and malevolent reciprocities of the investor than by neutral reciprocity. Researchers hypothesized that this physiological effect could be associated to the trustee's sudden shock about the unexpected generosity or spitefulness of the investor. Moreover, considering only the benevolent and the malevolent reciprocities, the BOLD intensity differed only in the head of the caudate nucleus of the trustee, and was higher in case of response to benevolent reciprocity by investor than in case of malevolent reciprocity. In order to better understand these changes in the "intention to trust" by the trustee, experimentalists carried out cross-brain analyses and discovered that activation of the middle cingulate cortex in the investor's brain foresaw subsequent intention to trust of the trustee expressed by the activation of the caudate nucleus in a neural point of view. Another important observation was that, in the early iterations of the game, the intention to trust of the trustee occurred only after the investor benevolent reciprocity; instead, during the last iterations, the intention to trust could occur also before the investor benevolent behaviors and the brain structures of interest of the investor and the trustee presented a strongest correlation in this case, suggesting the importance of reputation building also in finite games[50]. This experiment shows how in the real economic world the actions pursued by players in the early stages of a finite game, whose number of iterations is known in advance by the players, can be relevant for the construction

[50] As specified in *Chapter I,* according to the standard theory the reputation is considered a factor of influence of players' strategic interaction only in games repeated an indefinite number of times, but not in finite repeated games due to the application of the backward induction from the last iteration N, which is a stage game, to the first iteration.

of their reputation and can hence affect the subsequent players' behaviors in the last iterations of the game, contrary to standard game theory's predictions. Therefore, in a repeated game both the signals related to reward magnitude and to response timing revealed by the fMRI come into play during social decision-making.

3.4.5 Iterated Prisoner's Dilemma in neuroscience

Another important experiment in the neuroscience domain which takes into consideration the importance of reciprocity behaviors in iterated games is the experiment made by *Rilling et al (2002)* about the Prisoner's Dilemma. In their paper "*A Neural Basis for Social Cooperation*", they reported the main discoveries emerged from two experiments carried out separately with the support of the fMRI to scan 36 women players' brain. The fMRI was applied only to scan some players, while other players were playing the game outside the scanner. The goal of this experiment was to identify which players' brain regions presented a higher activation when making sustained social cooperative behaviors based on reciprocal altruism and the game that lent itself very well to the pursuit of this research was undoubtedly the repeated Prisoner's Dilemma[51]. The players' neural system which turned out to be relevant during the game included the anteroventral striatum, the anterior cingulate cortex and the orbitofrontal cortex, all of which belong

[51] Iterated Prisoner's Dilemma has been the subject of studies of different sectors of research about social cooperation based on reciprocal altruism.

to the reward system[52] which is usually_related to individuals' will to refuse immediate reward in order to get higher long-term reward.

The matrix below represents the amount of money which players could be awarded based on their social interactions. Players scanned through fMRI were identified by Player 2 , while non-scanned players by Player 1.

Player 2

Player 1		*Cooperate*	*Defect*
	Cooperate	**$2**(2)	**$3**(0)
	Defect	**$0**(3)	**$1**(1)

Figure 3.2 Payoff matrix Prisoner's Dilemma

The main aim of the first experiment was to distinguish the neural areas of activation in case of cooperation and non-cooperation behaviors between players; in case of cooperation, the different areas activated in a social and non-social context were distinguished. The aim of the second experiment was instead to put into evidence the different responses given by the player when playing with a human or a computer partner. As in the neural experiments previously analyzed, the players were keener to cooperate with a human partner than a computer, even

[52] More details about the neural reward system will be provided in the *subchapter 3.5*.

if the probabilities of this last type of partner were derived from the strategies played in the first experiment by the human unconstrained partner. In both experiments mutual cooperation was played the highest number of times in case of a human partner. Moreover, it was observed that when cooperation was reached by players, the probability that even in the following iterations cooperation was played by both players was higher with respect to the probability of defection by one or both players.

Regarding the correspondent neural activity in player's brain, experimentalists analyzed both BOLD reaction in the player's brain to the outcome observed and during his decision-making. In fact, each stage game of the Iterated game consisted of 12 seconds in which players could decide whether to cooperate or defect followed by 9 seconds in which the outcome obtain was showed to the players.

As regard the BOLD reaction to the outcome, experimenters observed in player 2 that BOLD response to cooperative outcome was higher than BOLD response to the other outcomes. The major brain areas of activation for the (*Cooperate, Cooperate*) outcome were identified in the anteroventral striatum, the anterior cingulate cortex and the orbitofrontal cortex. All these brain areas are related to the reward processing, analyzed thereafter; hence, the activation of these areas was coherent with the player strategy to cooperate in order to obtain a long-term advantage to the detriment of the immediate gain that he would benefit by choosing to defect, betraying the trust of the partner. The difference between the anteroventral striatum and the orbitofrontal cortex is that the first brain structure is associated only to the cooperation with a human counterpart while the second can be associated to both a human or a computer counterpart. As regard the anteroventral striatum, since the ventral striatum is usually associated to reward, it has been observed in neuroscience experiments that the increase of this brain area activity is associated to the reward of the economic agent relatively to the reward provided to its partner and not to its own reward in absolute terms. To demonstrate that the benefit derived from choosing the cooperative outcome was principally due to the willingness to reciprocate the altruism expressed by the partner in the previous round and not simply for receiving a payoff of 2$, researchers verified what happened in the first experiment in case a sum of 2$ was provided to the same player in a non-social context that did not presuppose the establishment of a relationship of cooperation based on reciprocal altruism. In this case, the anteroventral striatum, the anterior cingulate cortex and the orbitofrontal cortex did not present the activation observed in case of mutual cooperation, proving the fact that the benefit received by player in a social context is not the same to receive an equal monetary amount in a non-social perspective.

As regard the BOLD reaction during the decision of a player to cooperate after having observed the other partner cooperation move, experimentalists observed the activation of the rostral anterior cingulate cortex. As already observed in the experiment about the ultimatum game, the anterior cingulate cortex is associated to the emotional-cognitive conflict and hence, it can be attributed once again to the conflict between receiving an immediate reward or delay a higher reward. Moreover, the higher the activation of the ventromedial prefrontal frontal cortex during players' decision to resist the immediate temptation of defecting in order to get a higher long-term reward was coherent with recent neural studies according to which ventromedial prefrontal frontal cortex is associated to long term rewards and punishment. Another brain structure interested in this process was the right post-central gyrus, located on the primary somatosensory cortex and involved in reciprocal cooperation. Overall, the activation of all these brain regions is associated to an increase in cooperation based on reciprocal altruism by players.

3.5 Reward System

It is important to underline that almost all the brain structures which were found to be relevant in the previous neuroimaging experiments belong to the so-called *neural reward system*. The reward system is a neural system which include several brain structures associated to economic agents' formation of preferences and decisions about performance of actions leading to reward. Stated another, the neural reward system includes all the brain structures which are involved in reward evaluation, when dealing with reward evaluation in case of one-shot games and of both reward evaluation and reinforcement learning in the repeated games. In an anatomical point of view, among these brain structures there are: the striatum, which is one of the core components of the reward system, the prefrontal cortex, the anterior cingulate cortex, the insular cortex and the thalamus. The important concept behind the belonging of these brain structures to the same neural reward system lies in the fact that, although decision-making or the observation of a certain outcome correspond to greatest activation of specific brain region, it is the entire neural network that connect each brain structure to other regions to drive players towards certain choices and the subsequent actions.

In particular, the importance of the reward system in behavioral game theory is due to the fact that this neural system drives players' behaviors toward long-term reward and ward off players from conflictual situations. In the economic games previously analyzed, the activation of the brain structures belonging to the reward system is coherent with players' cooperative behaviors in order to obtain a long-term advantage and also to effective punishment of defectors players in order to not incentivize unfair behaviors and to obtain a sort of reward.

3.6 Further considerations about economic experiments

The purpose of this last subchapter is to sum up and draw some conclusions about how the general functions of the brain structures analyzed in subchapters *3.2.1.1.* and *3.2.1.1.* can be translated into more specific functions when applied in the game theory context.

For example, the anterior cingulate cortex associated to the function of mediator in the emotional-cognitive conflict, comes in to play in game theory in those situations in which players have to decide whether is better to choose rationally the alternative leading to the highest benefit in monetary terms or to choose the less profitable alternative but coherent with fairness behavior. As regard the caudate nucleus, since the main functions of this brain structure concern goal-directed actions, this brain structure turned out to be important in those circumstances in which the player wanted to take a certain action to pursue a specific goal, such as the effective punishment of defector in the game about the altruistic punishment. In the same way, the prefrontal cortex associated to executive functions, comes into play in those situations in which players try to control their emotional impulses. Specifically, in the ultimatum game reported, that orbitofrontal cortex was associated to the responder's self-rational control of accepting an unfair offer leading to him a highest payoff in monetary terms than the one got by refusing it. In the same way, the medial prefrontal cortex associated to functions related to theory of mind processes, came out to important in context of reciprocal cooperation in which players' capability to attribute mental states to other players is essential to achieve social cooperation interactions.

Conclusions

The scope of this thesis was to analyze the divergences between the predictions made by mathematical models of traditional game theory and of the emergent behavioral game theory. To demonstrate the validity of the initial hypothesis, according to which social behaviors predicted by behavioral game theory get closer to what could actually be the decisions taken by the economic agents in a situation similar to the one proposed by the games, some experimental games, carried out with the support of recent neuroimaging techniques, have been analyzed.

The most important results which have been put into evidence by these experiments are that:

- Social behaviors of cooperation and fairness, including inequity aversion, predicted by behavioral game theory models present a background of empirical evidence according to the scientific techniques applied.
- There are some brain regions of interest, such as some areas of the cerebral cortex and the striatum, associated to social behaviors pursued by players.

Trying to understand what happens in our black box, that is to say the human brain, when decisions and actions are taken, is a step forward towards the comprehension of the mechanisms underlying these actions.

In almost all the experiments proposed, the brain areas were identified through the BOLD signal in fMRI. This method is better to predict input (decision) more than the effective output (action), because it provides only indirectly the images of the brain areas involved in decision-making process. Therefore, a more in-depth analysis would be to compare the results obtained in the same game experiment with the use of different neuroimaging techniques, such as the MEG and the EEG which, contrary to fMRI, record neuronal activity directly in the form of electromagnetic radiation. Another possible deepening could be to specify before the start of the experiment which are the social and cultural characteristics of the players tested in order to observe and evaluate the possible changes in the results. In general, since this field of research is still at the beginning of its development, many pieces still need to be added to achieve a more complete picture of the real motivations and factors driving economic agents in decision-making.

Thanksgiving

A conclusione della tesi, vorrei ringraziare tutte le persone che mi hanno accompagnato in questo percorso accademico, chi dall'inizio alla fine e chi magari solo per una parte di esso. Innanzitutto, ringrazio la mia famiglia per avermi sostenuto sia moralmente che economicamente in questo percorso e per aver investito in quello che sarà il mio futuro. Un ringraziamento particolare va anche al relatore della tesi, il prof. Luigi Buzzacchi, per avermi seguito attentamente nella stesura dell'elaborato e avermi incoraggiata a cercare di approfondire e migliorare sempre ogni sua parte. Infine, un ringraziamento speciale agli amici del Poli e non, che hanno creduto in me quando pensavo di non farcela e con cui ho condiviso preoccupazioni e piccoli traguardi raggiunti. Un grazie di cuore a tutti.

Bibliography

BRACALETTI S., La Teoria della scelta razionale. Applicazioni e problematiche, 2017

BOLTON G., OCKENFELS A., A Theory of Equity, Reciprocity and Competition, 2000

CAMERER C., Progress in Behavioral Game Theory, 1997

CAMERER C., Rappresentazione mentale dei giochi: Framing e editing, 1999

CAMERER C., Advances in Behavioral Economics, 2004

CAMERER C., Neuroeconomia, ovvero come le neuroscienze possono dare nuova forma all'economia, 2004

CAMPA R., Teoria dei giochi, neurofisiologia e scienze sociali, 2013

DELGADO M., Reward-Related Responses in the Human Striatum, 2007

DE QUERVAIN D., The Neural Basis of Altruistic Punishment, 2004

DUFWENBERG M., KIRCHSTEIGER G., A Theory of Sequential Reciprocity, 1998

FEHR E., SCHMIDT K., A Theory of Fairness, Competition, and Cooperation, 1999

FEHR E., SCHMIDT K., Theories of Fairness and Reciprocity - Evidence and Economic Applications, 2000

GARIBALDI P. BRUSCHI M., Introduzione alla Teoria dei giochi e ai modelli di oligopolio, 2014

GINTIS H., Behavioral Game Theory and Contemporary Economic Theory, 2005

GIRONDE S., La neuroeconomia. Come il cervello fa i nostri interessi, 2010

HAGEN E., HAMMERSTEIN P., Game theory and human evolution: A critique of some recent interpretations of experimental games, 2005

KAHNEMAN D., TVERSKY A., Prospect Theory: An Analysis of Decision under Risk, 1979

KENNING P., PLASSMANN H., Neuroeconomics: An overview from an economic perspective, 2005

KING-CASAS B., Getting to Know You: Reputation and Trust in a Two-Person Economic Exchange, 2005

LEE D., Game theory and neural basis of social decision making, 2008

LEVIN J., Fairness and Reciprocity, 2006

LIEBERMAN M., TABIBNIA G., Fairness and Cooperation Are Rewarding: Evidence from Social Cognitive Neuroscience, 2007

MCCABE K., A functional imaging study of cooperation in two-person reciprocal exchange, 2001

NASH J., John Nash racconta la teoria dei giochi, 2012

OSBORNE M., An Introduction to Game Theory, 2009

RABIN M., Incorporating Fairness into Game Theory and Economics, 1993

RABIN M., Social Preferences: Some Simple Tests and a New Model, 2000

RILLING J., A Neural Basis for Social Cooperation, 2002

SANFEY A., The Neural Basis of Economic Decision-Making in the Ultimatum Game, 2003

SANFEY A. Social Decision-Making: Insights from Game Theory and Neuroscience, 2007

SCHMIDT C., Neuroeconomia, 2013

Sitography

https://www.behavioraleconomics.com/

https://www.ilsole24ore.com/

http://www.treccani.it

https://www.okpedia.it