

Returns-Based Beliefs and The Prisoner's Dilemma

Chander Velu* Sriya Iyer†

November 18, 2008

Abstract

Returns-based beliefs provides an explanation for the anomaly between the theory and empirics for the one-shot and finitely-repeated Prisoner's Dilemma games. Even in a fully specified game, there is strategic uncertainty as players attempt to coordinate their actions. Therefore players form subjective probabilities of the actions of their opponents. We provide a new method termed the 'returns-based beliefs' approach of forming subjective probabilities that is based upon the expected returns of a particular strategy, in proportion to the total expected returns of all strategies. This method can be applied even in the absence of knowledge of the players' respective histories.

JEL classification: C72, D43

Keywords: Prisoner's Dilemma, Rationality, Subjective Probabilities, Returns-Based Beliefs

*Judge Business School, University of Cambridge. c.velu@jbs.cam.ac.uk

†Corresponding author: Faculty of Economics, University of Cambridge. Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD. Tel: 44 1223 335257. Fax: 44 1223 335475. Sriya.Iyer@econ.cam.ac.uk.

1 Introduction

Economists have highlighted a number of game-theoretic contradictions and paradoxes in which individual decision-making in real-world situations is at odds with what is predicted by game theory (Goeree and Holt 2001; Luce and Raiffa 1957; Selten 1978; Rosenthal 1981; Reny 1993; Binmore 1987; Bicchieri 1989; Petit and Sugden 2001). One of the most widely analysed games in economics is the Prisoner's Dilemma, a two-by-two noncooperative game. Some empirical tests of this game demonstrate that in the real world people are often more cooperative than that predicted by the outcome of this game in theory. The Prisoner's Dilemma also lies at the heart of important concepts in game theory such as the 'Nash equilibrium' (Nash 1951). In this paper we provide an explanation for why cooperative strategies might be played in a one-shot and finitely repeated Prisoner's Dilemma game. We propose an alternative method by which people might form their beliefs to play their strategies. We call this '*returns-based beliefs*' which is the expected returns of a particular strategy, in proportion to the total expected returns of all strategies. We argue that this belief structure, which is based upon subjective probabilities, might explain the rationale behind the adoption of cooperative strategies in a one-shot and finitely repeated Prisoner's Dilemma game.

The payoff to the Prisoner's Dilemma game is as illustrated in Figure 1 where two agents have to decide whether to cooperate or to defect. Let us call the agents player 1 and player 2 respectively. If both cooperate they both get a payoff of 4. However, both player 1 and player 2 could be better off by playing Defect if the other player continues to play Cooperate. If player 1 chooses to cooperate and player 2 chooses to defect then player 2 gains with a payoff of 6 while player 1 loses with a payoff of only 0.5 and vice versa if player 2 cooperates but player 1 defects. Therefore, both players would reason that they are better off defecting and as a result end up obtaining a payoff of 1 each. This is

clearly less than the Pareto optimum of obtaining a payoff of 4 each by both cooperating. However, because the players are unable to communicate beforehand, game theory predicts that they will both play the Defect strategy, thereby making them both worse off. This is the only Nash equilibrium of the Prisoner's Dilemma as there is no incentive for any of the players to change their strategies. Yet, empirical testing of the Prisoner's Dilemma in laboratory experiments has shown that people are prone to play the cooperative strategies far more often than game theory alone might lead us to predict in both one-shot games and in finitely repeated games. These studies have also demonstrated that as the benefits from cooperation increase, players are more likely to cooperate, and that as the loss from not cooperating increases, the likelihood of cooperation increases as well (Sally 1995).

For example, theoretically it has been shown that the strategy where both players cooperate is possible to obtain in a repeated Prisoner's Dilemma game (Wilson 1986, Shubik 1970, Kreps et al 1982, Axelrod 1980a, 1980b, 1984). Equally, scholars have emphasised that cooperation in a one-shot Prisoner Dilemma game is an important finding of experimental research that needs to be further understood (Field 2001, Janssen 2008). Some economists have proposed various explanations for these experimental findings: for example, altruistic punishment among genetically unrelated people when the gains from reputation are small or absent (Fehr and Gächter, 2002); the ability to recognize untrustworthy opponents (Janssen 2008); and the incorporation of notions of fairness into game theory through which people help others that help them and hurt others that hurt them (Rabin 1993). Moreover, in the case of the finitely repeated Prisoner's Dilemma game, the leading theoretical explanations for cooperative behaviour here are reputation building and altruism (Kreps et al 1982, Andreoni and Miller 1993, Cooper et al. 1996). To all of these explanations, we add an alternative: we propose in this paper a new explanation for cooperation in a one-shot and a finitely repeated Prisoner's Dilemma game that is based upon subjective probabilities and returns based beliefs. Our explanation for the level of

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	(4, 4)	(0.5, 6)
	Defect	(6, 0.5)	(1, 1)

Figure 1: Payoffs and Dominated Claims

cooperation through the evolution of the game in the repeated Prisoner's Dilemma is not based on reputation or altruism, but directly on the consequences of strategic interactions as a result of returns based belief formation.

Game theory models systematically human behavior when strategic interactions exist. In conventional game theory, the solution concept such as a Nash equilibrium is critical in forming the basis for the prior distribution of beliefs that players hold. In determining the outcome of the game these prior beliefs held by the players are fulfilled in equilibrium. However, a player's actions are determined by her beliefs about other players which may depend upon their real-life contexts such as custom or history (Aumann and Dreze 2008). Game theory is also a normative theory that describes how people ought to behave rather than a descriptive theory about how people actually behave (Kadane and Larkey 1982, 1983). For example, Harsanyi (1982) contended that normative game theory was not as helpful as 'an empirically supported psychological theory making probabilistic predictions about the strategies people are likely to use, . . . given the nature of the game and given

their own psychological makeup' (p.122). This psychological makeup might be conditioned by the past experience of individuals' beliefs about an opponent's play. This is termed the '*subjective*' or personal interpretation of probability. Subjective probability is the probability that a person assigns to a possible outcome, or some process based on his own judgement, the likelihood that the outcome will be obtained (DeGroot 1975, pp. 4).

In a similar vein, Herbert Simon (1957) distinguished between subjective and objective rationality. Subjective rationality is behavior that is rational given the perceptual and evaluational premises of the subject. Objective rationality is behavior that is rational as viewed by the experimenter. Human behavior is likely to be subjectively rational but not necessarily objectively rational. The implication is that the experiences of the individual might feed into the so-called perceptive and evaluational premises of the individual and influence thereby the subjective probabilities, which then influences the strategies chosen. An implication of the subjective probability approach is that the chosen strategy might not be consistent with the equilibrium predictions of an objectively rational outcome (Roth and Schoumaker 1983b). The players' experience is an important determinant of the player's expectations which might lead to outcomes that might not be the Nash equilibrium prediction (Roth and Schoumaker 1983a).

Subjective probabilities might have important consequences for the outcome of the Prisoner's Dilemma game and for its empirical testing. We suggest an alternative basis upon which beliefs in game theory might be formed - returns-based beliefs - and we present the corresponding numerical results. Returns-based beliefs are important when forming subjective probabilities because we can show that they can be applied even when the players are not knowledgeable about their respective histories. We show that as long as the temptation to defect is not large, or the benefits to cooperate are significant, then players are likely to cooperate. Our results correspond very closely to empirical studies of the Prisoner's Dilemma. Section 2 revisits the Prisoner's Dilemma and provides an

explanation for the reconciliation of the empirical findings and the theoretical predictions of the Prisoner's Dilemma using returns based beliefs. Section 3 concludes.

2 Cooperative behavior in a Prisoner's Dilemma

This section provides an explanation for cooperative behavior in the Prisoner's Dilemma. We show that a plausible explanation for the empirical evidence conducted on the Prisoner's Dilemma game may be based upon agents playing mixed strategies. We discuss a manner in which agents form beliefs that is based upon their expected returns. We discuss how this forms the subjective probabilities, and how the numerical results obtained correspond closely with previous empirical evidence on this issue.

2.1 The formation of beliefs

In the Prisoner's Dilemma game, the elimination of dominated strategies is used to derive the Nash equilibrium of (Defect, Defect). However, clearly because the empirical evidence is at odds with the theoretical outcome of a Nash equilibrium, there must be an alternative manner in which agents are choosing their strategies. Other economists have argued that perhaps we need to examine psychological motives more to understand behavior in the Prisoner's Dilemma game and have suggested examining the process of cognitive reasoning (Rubinstein 2006 and 2007). In this section, we propose an explanation that is based on players wanting to cooperate on their strategies, as such cooperation provides better returns than non-cooperation.

In the case of the Prisoner's Dilemma, ideally, the agents would like to cooperate by coordinating their actions on the joint claims that will maximize their returns, which is (4,4). Intrinsically, each player knows the benefits of cooperation and hence he or she may actually play the cooperative strategy with positive probabilities, that coordinates

with the other player. For example, research has shown that human beings are prone to cooperative behavior based on reciprocity (Axelrod 1984, Axelrod and Dion 1988). Therefore, the history of human interactions are likely to influence the general disposition of players to want to cooperate¹. We need to factor this cooperative bias into our decision making framework in order to predict how players should behave in a competitive situation where cooperation is possible and could produce better outcomes² (Friedman 1996).

In this paper we argue that the willingness to cooperate might be influenced by past experience, generating ‘*strategic uncertainty*’. We define strategic uncertainty as uncertainty concerning the actions and beliefs (and beliefs about the beliefs) of others (Brandenberger 1996, Morris and Shin 2002). Researchers have argued that strategic uncertainty can arise even when all possible actions and returns are completely specified and are common knowledge (Van Huyck et al 1990). The rational decision-maker has to form beliefs about the strategy that the other decision maker will use as a result of strategic uncertainty. As a consequence, players form their beliefs about the probabilities that other players play in order to determine in turn their best-response strategy. Hence, the best response strategy of one player is likely to be based upon the mixed strategy of the other player. The mixture is because of the uncertainty regarding the conjecture about the choice by the other players³ (Brandenburger and Dekel 1989; Brandenberger 2007). This is succinctly sum-

¹Farell and Rabin (1996) have argued that even with communication there would be ‘cheap talk’ and hence the communication is not credible. However, Sally (1995) shows that communication does result in an increase in cooperation among players.

²To add to this line of reasoning, psychologists have argued that cooperation may be prompted by altruism, by the desire to conform to social norms, or by adhering to the dictates of one’s conscience (Dawes 1980). In addition, economists have shown that ‘people’s natural tendency to cooperate’ is an important trait that subjects bring to experimental situations from the outside (Andreoni and Miller 1993, p. 571).

³We are not assuming that the opponent is using a randomized strategy. The mixture merely reflects the representation of player 1’s belief about player 2. As Wilson (1986, pp.47) points out, although

marised by Rabin (1993) that ‘In psychological games, there can be a difference between interpreting mixed strategies literally as purposeful mixing by a player versus interpreting them as uncertainty by other players’ (Rabin 1993, p.1286).

The issue of mixed strategies is also relevant when one considers the concept of the elimination of dominated strategies. The concept of elimination of dominated strategies assumes that a strategy that is dominated will never be played. However, empirical studies show that dominated co-operative strategies might be played with positive probabilities (Cooper et al 1990). The empirical findings show that playing dominated cooperative strategies with positive probabilities could change the results in a significant way. For example, forming beliefs about the strategies that the other decision maker might use is a subjective assessment based on the previous experiences of the player. Hence, research has shown that it is possible for the player to expect an opponent to play all possible strategies, including strictly dominated strategies, with positive probabilities. For example, Roth and Schoumaker (1983b) in their seminal study showed that more than one outcome that is not the Nash equilibrium outcome, could be considered to be consistent with perfect rationality when the ‘outcome depends on subjective expectations of the players which are not determined by the data of the game’ (p. 1338). If the experience of the individual has been to get a share of a bargain that is not a Nash equilibrium outcome, then it is reasonable to assume that the player will continue to expect the non-Nash equilibrium outcome in the next game (Roth and Schoumaker 1983a, 1983b). Therefore, a player who knows that the non-Nash equilibrium belief is held by the opponent could be deemed to be rational when forming a subjective assessment of the opponent’s play by taking this belief into account (Basu 1990). This might be construed as an error by the player if one were to look at this game purely from the perspective of objective rationality. In this

it makes little difference to the mathematics, conceptually this distinction between randomization and subjective beliefs to explain the mixed strategies is a pertinent one.

case, rationality of perception which might exist does not necessarily correspond to the objective probabilities that would result if the dominated strategies were to be deleted (as in the case of the Prisoner's Dilemma). Correspondingly, we argue that 'subjective' rationality coexists with the rationality of perception driven by the previous experiences of the individuals. Given such a perception, we need to re-examine the process of deletion of dominated strategies.

The deletion of dominated strategies assumes that the probability of playing these strategies is zero and hence they are not expected to be played. As discussed earlier, when subjective probabilities are held, dominated strategies could be played with positive probabilities. Therefore, could one delete dominated strategies when such subjective probabilities are held? We attempt to answer this by examining the game in Figure 1. The figure shows the payoff for the various strategies of both players. In the case of elimination of dominant strategies, player 1 would not play cooperate as player 2 could trump him by playing Defect. Therefore, player 1 would be better off playing Defect and so will player 2 in that case and we reach the Nash equilibrium of (Defect, Defect) by eliminating the co-operate strategies which is dominated. However, if we assume that player 1 holds the belief that player 2 is likely to cooperate even with a very small probability because of some historical experience of player 2 (for example, the need to want to coordinate the strategy with player 1 and earn the same payoff) then it would be rational for player 1 to also choose the strategy cooperate with a positive probability. Now it is easy to see that it is no longer the case that strategy Cooperate is dominated. This is because if player 2 were to play both Cooperate and Defect with some positive probability, player 1 could do better by mixing between Cooperate and Defect than to play Defect. This result is shown using the example of the numerical model in the next section. The common practice is to eliminate strictly dominated strategies from a normal form game in formulating equilibria in mixed strategies (Kohlberg and Mertens 1986). However, the practice of eliminating

strategies (Cooperate, Cooperate) via a process of deletion of dominated strategies is not appropriate, because playing a mixed strategy between Cooperate and Defect could yield a better payoff than playing the Nash equilibrium strategy of Defect when the other player plays all strategies with some positive probability⁴.

In order to illustrate our results we need some plausible set of assumptions about how agents form beliefs with respect to the probability of the opponent's strategy. We suggest that one way in which agents might do so is by basing their decisions on '*returns-based beliefs*'. In this method, player 1 plays strategies based on the probabilities of the proportion of returns over the total returns for each possible claim by player 2. The next section describes the approach and results associated with this method of belief formation.

2.2 Returns-based beliefs and the Prisoner's Dilemma

In the following discussion, we describe a possible subjective probability formation based on returns based beliefs. First, we discuss the concept of subjective probabilities and its implications for the Nash equilibrium in order to provide an understanding of why this method of forming beliefs might be reasonable. We assume that players are expected utility maximizers. The traditional approach in game theory when mixed strategies are used is for a player to choose probabilities (over their own strategies) in such a manner as to make the other player indifferent between the different strategies. The implication of

⁴Conventionally, any mixed strategy will have a support in pure strategies. However, the pure strategy (Cooperate) will get eliminated by the deletion of dominated strategies which suggests that the play of a mixed strategy based upon a support in pure strategies would not be apposite in this context. However, because of our argument that invokes subjective probabilities, all the strategies are played with positive probabilities. As a result we argue that the deletion of dominated strategies is no longer appropriate. Therefore, a mixed strategy can exist if one player experiences uncertainty with respect to his conjecture about the choice of the other player.

this is that each player's equilibrium strategy depends only on the other players' payoff and not their own in order to make the other player indifferent between the different strategies (Amaldoss and Jain 2001). However, this approach of making the opponent indifferent would not be the case when non-equilibrium strategies are chosen via the use of subjective probabilities. When such non-equilibrium outcomes are chosen, each player maximizes their own expected values based on their conjecture of what the opponent is likely to do. Therefore, the probabilities are chosen to maximize their own expected values rather than to make the opponent indifferent to the different strategies. As discussed earlier, driven by the desire to want to cooperate, for example in the Prisoner's Dilemma game, there is uncertainty regarding the conjecture about the choice of the other player. Hence, the player holds an opinion based on the subjective probability with respect to all of the unknown contingencies affecting his payoffs. In particular the player is assumed to have *'an opinion about the major contingency faced, namely what the opposing player is likely to do'* (Kadane and Larkey 1982, pp. 115). Kadane and Larkey (1982, pp. 115) have expressed the implications of this line of thought very neatly as follows: *'If I think my opponent will choose strategy i ($i = 1, . . . I$) with probability p_i , I will choose any strategy j maximizing $\sum_{i=1}^I p_i u_{ij}$, where u_{ij} , is the utility to me of the situation in which my opponent has chosen i and I have chosen jthe opponent's utilities are important only in that they affect my views $\{p_i\}$ of what my opponent may do....'*

Therefore, it follows that if player 2 is not expected to play the Nash equilibrium strategy than it might be optimal for player 1 also not to play the Nash equilibrium strategy as doing so would give player 1 a better payoff. This implies that the Nash equilibrium is a special case when each player is assumed to believe that the other is sure to play the Nash equilibrium strategy. Let us revisit the concept of objective and subjective probabilities to help clarify the context of this discussion: the Nash equilibrium solution concept assumes rationality from the perspective of an external observer. However, at the level of the

individual player, assumptions about the opponent's beliefs may be conditioned by past experience and hence be different from the priors held by the rational external observer. In situations of strategic interaction such as in a Prisoner's Dilemma game, the players might hold subjective probabilities that are different from the objective probabilities demanded by the Nash equilibrium solution concept. When these probabilities are the same we get the special case of the Nash equilibrium. However, there is no compelling reason a priori for these probabilities to definitely be the same. Although any possible distribution of probabilities could be a possibility based upon the subjective method of forming them, we shall try to propose a reasonable subjective probability belief that the players might use when they do not know each other or their respective histories. We shall call this *'returns-based beliefs'*, which we describe in more detail below.

We posit that the players have a desire to want to cooperate based on the premise that historical experience tells them that this might provide a better return. However, the rational decision maker has to form beliefs about the opponent's play due to the strategic uncertainty about what the opponent is likely to play. As discussed before, the decision maker is trying to maximize expected returns based on these beliefs. Therefore, it is reasonable to assume that the decision maker would assign probabilities based on the expected returns from playing the different strategies. Similarly, it is reasonable to assume that the opponent also assigns probabilities based on the opponent's expected returns given the probabilities of the focal decision maker. Following this line of reasoning, our analysis is based on a model for which the decision probabilities are proportional to the expected returns. We assume that agents form beliefs based upon the expected returns for a particular strategy over the total expected returns of all strategies, assuming the opponent plays all possible strategies. Our proposed approach has both theoretical and empirical justification. First, for the theoretical justification we defer to Luce (1959) who showed using probability axioms that if the ratio of probabilities associated with any two

decisions is independent of the payoff of any other decisions, then the choice probabilities for decision i can be expressed as a ratio of the expected payoff for that decision over the total expected payoff for all decisions: $\frac{\pi_i^e}{\sum_j \pi_j^e}$ where π_i^e is the expected returns associated with decision i . Second, this method of arriving at decision probabilities has been supported by empirical work which provides empirical justification for our approach. In particular, empirical research for paired comparison data supports the Luce (1959) method of arriving at decision probabilities such that the probability for choosing x over y , $P(x, y) = \frac{v(x)}{[v(x)+v(y)]}$ where $v(x)$ and $v(y)$ are the scale values of choosing x and y respectively (Abelson and Bradley 1954, Bradley and Terry 1952). We operationalize our model as follows. In this model, each player chooses among $j, m = 2$ possible strategies and the expected payoffs are given by the summation below:

$$\pi_i^e(j) = \sum_{m=1}^n \pi_i(j, m) p_i(m), \quad j, m = Cooperate, Defect \quad (1)$$

where $\pi_i(j, m)$ is player i 's payoff from choosing a claim equal to j when the other player claims m and $p_i(m)$ is the belief probabilities held by player i about player j playing strategy m . The decision probabilities in turn follow the specification outlined above which is proportional to the expected returns as follows,

$$D_i(j) = \frac{\pi_i^e(j)}{\sum_{m=1}^n \pi_i^e(m)} \quad (2)$$

In our model we assume a Nash-like equilibrium in belief formation such that the belief probabilities matches the decision probabilities for both players 1 and 2 respectively. This symmetry in probabilities is achieved by iterating between the expected payoff in equation (1) and the decision probabilities in equation (2). To begin the analysis, let us assume that

player 1 believes that player 2 plays each of the strategies with equal probabilities i.e. 0.5. We then multiply 0.5 with the rewards of player 1 (as per Figure 1) to get the expected returns as shown in Figure 2. For example, the number 2.00 (row one and column one in Figure 2) is obtained by multiplying 0.5 by 4 (the first number in parentheses in row one and column one in Figure 1). The second last column labelled ‘Total’ shows the total rewards for a particular claim for player 1 for all possible claims by player 2. For example, 2.25 is the sum of all the rewards (2.00+0.25) for the cooperate strategy by player 1 when player 2 plays cooperate and defect with equal probabilities.

We would now need to calculate the probabilities that player 1 will play the various claim strategies. As discussed above, player 1 is concerned about his opponent’s returns only to the extent that he wants to maximize his own returns subject to the opponent’s play. Therefore, it might be reasonable to assume that player 1 assigns probabilities to each of the claim strategies proportional to the expected returns of playing that strategy. The last column in Figure 2 depicts the probabilities that player 1 would play a particular claim in response to the various claim strategies for player 2. This is the same as the conjecture that player 2 has about player 1’s probabilities. This is derived by dividing the reward player 1 gets for a particular claim strategy by the total rewards for all possible claims of player 1. For example, the number in the last column and first row of Figure 2, 0.392 is obtained by dividing 2.25 (the number in the first row of the column labelled ‘Total’ in Figure 2) by the total of 6.05 (the number in the far right row in the column labelled ‘Total’ in Figure 2).

However, so far player 1 and player 2 have different beliefs about each other’s probabilities of playing a claim. Player 2’s expected returns can be calculated by applying the probabilities that player 1 will play each of the strategies as calculated from Figure 2 above. In a similar way, player 1’s revised probabilities for each of the claim strategies can be calculated based on the returns based method described for player 1 above. We now

		Player 2		Total	Prob.
		Cooperate	Defect		
Player 1	Cooperate	2.00	0.25	2.25	0.392
	Defect	3.00	0.5	3.80	0.608
				6.05	1.00

Figure 2: Player 1's Probabilities

revise the expected returns in Figure 2 with the new probabilities (as compared to the equal probabilities that we started out with). This process provides updated probabilities for player 2 for each claim strategy, shown by the revised numbers for the last row of Figure 2. This process can be repeated until the probabilities for players 1 and 2 converge. Conducting this iterative process shows that these probabilities do actually converge after about three to four iterations. Since the players are symmetric, it is not unreasonable to assume without any further information about history or preferences that they would have the same subjective beliefs about each other.

The probabilities converge for both players to 0.387 for the cooperate strategy and 0.613 for the defect strategy. Therefore, the 'returns based beliefs' show that the players will play the cooperate strategies with positive probabilities which is in line with empirical evidence. In addition the empirical evidence shows that as the benefits from cooperation increases the players are more likely to cooperate. Sally (1995) showed using a metanalysis of over 100 studies that *'The one major consistency with rational self interest is that the temptation to defect decreases the level of cooperation'* (p.75, Sally 1995). Another

way of looking at this is that as the opportunity to increase one's reward by defecting from unanimous cooperation decreases, then the likelihood of cooperation decreases. In addition, the analysis shows that as the loss from not cooperating increases, the likelihood of cooperation increases too. Sally (1995) proposed a method of calculating the Temptation Index and Loss Index as follows:

$$Loss = (C(x) - D(0))/C(x) \quad (3)$$

$$Temptation = (D(x - 1) - C(x))/C(x) \quad (4)$$

where

$n = 2$ (number of players)

$x =$ number of operators ($0 < x < n$)

$C(x) =$ payoff for each of the x cooperators

$D(x) =$ payoff to each of the $(n - x)$ defectors

Based on our example in Figure 1, the value of the indices are as follows,

$Loss = (4 - 1)/4 = 3/4$ (75%) and $Temptation = (6 - 4)/4 = 1/2$ (50%). As one allows the benefit from cooperation from the (Cooperate, Cooperate) payoff increase from 4 to 5.8, the loss index increases from 75.0% to 82.8% which in turn increases the probabilities of co-operation from 0.39 to 0.47 as shown in Figure 3. In addition, as the payoff from (Defect, Cooperate) is increased from 6 to 7.8, the temptation index increases from 50.0% to 85.0% which in turn decreases the probability of cooperation from 0.39 to 0.35 as shown in Figure 4. This result is consistent with empirical evidence from the Prisoner's Dilemma games.

(C,C) Payoff	Loss Index (%)	Cooperative Probability
	(Column 1)	(Column 2)
4.0	75.0	0.39
4.2	76.2	0.4
4.4	77.3	0.41
4.6	78.3	0.42
4.8	79.2	0.43
5.0	80.0	0.44
5.2	80.8	0.45
5.4	81.5	0.45
5.6	82.1	0.46
5.8	82.8	0.47

Figure 3: Loss Index and Cooperative Probability

(D,C) Payoff	Temptation Index (%)	Cooperative Probability
	(Column 1)	(Column 2)
6.0	50.0	0.39
6.2	55.0	0.38
6.4	60.0	0.37
6.6	65.0	0.37
6.8	70.0	0.36
7.0	75.0	0.36
7.2	80.0	0.35
7.4	85.0	0.35

Figure 4: Temptation Index and Cooperative Probability

Our approach is different to previous studies that show the plausibility of cooperation in the one-shot Prisoner's Dilemma game. For example, Fehr and Gächter (2002) show that people experience negative emotions towards defectors and are therefore willing to punish them. This act is called 'altruistic punishment'. Although such punishment is costly and yields no material benefits to the punisher it acts as a deterrent to defection which can then explain cooperation in one shot games among strangers. Jansenn (2008) shows that the option available to players to withdraw from playing and the ability to recognize untrustworthy opponents could create conditions where players cooperate in one-shot Prisoner's Dilemma games. This is because the ability of players to withdraw from the game and not play an untrustworthy opponent and the consequent social welfare preference of not wanting to be seen as contributing to the level of untrustworthiness in society results in cooperation in one-shot Prisoner's Dilemma games. On the other hand, Rabin (1993) showed that incorporating fairness into game theory could result in cooperative outcomes more than conventional game theory would predict. The idea is that people like to help those who are seen to be helping them and hurt those who are seen to be hurting them. Therefore, Rabin (1993) shows that so long as the material gains from defecting are not too large each player would want to help the other by cooperating which results in the (Cooperate, Cooperate) outcome (p. 1288, Rabin 1993).

In this paper we propose an alternative explanation for the (Cooperate, Cooperate) outcome in the Prisoner's Dilemma that is based on subjective probabilities and returns based beliefs formation. Our approach enables us to provide a plausible explanation for two anomalies between the theoretical predictions of one-shot Prisoner's Dilemma games and the empirical evidence: (1) the inverse relationship between cooperation and the temptation to defect and (2) the positive relationship between cooperation and the gains from cooperation.

The next section extends the application of the returns-based belief approach to the

finitely repeated Prisoner's Dilemma.

2.3 Finitely Repeated Game

The discussion so far has been on one-shot Prisoner's Dilemma games. However, cooperation is also evident from empirical studies of finitely repeated Prisoner's Dilemma games (Andreoni and Miller 1993, Cooper et al. 1996, Dal Bo 2005). Economic theory predicts (Defect, Defect) as the equilibrium play for all rounds of the game in a finitely repeated Prisoner's Dilemma. However, empirical evidence shows a high level of cooperation in early rounds of the game with the rate of cooperation declining towards the final rounds of the finitely repeated Prisoner's Dilemma game (Andreoni and Miller 1993, Dal Bo 2005). In addition to the declining rate of cooperation, Cooper et al. (1996) found that the rate of cooperation was higher in the finitely repeated Prisoner's Dilemma game than the one-shot Prisoner's Dilemma game. To date, the leading theories of cooperation are based on either reputation building or altruism. Krepps et al. (1982) show that cooperation is possible in the finitely repeated Prisoner's Dilemma due to the presence of incomplete information about the true payoffs of the rival. In such a setting a small belief that an opponent will cooperate is sufficient for cooperative play as players would like to build their reputation for cooperation. A leading alternative explanation of cooperation is based on players being altruistic. In the altruism theory, one can think of the players receiving an additional payoff from being cooperative. So for example, Cooper et al. (1996) argue that the reputation model is inconsistent with cooperation in the one-shot Prisoner Dilemma game as there is no opportunity to build reputation. In addition, these authors show that altruism alone cannot explain the cooperation rates over time in a finitely repeated Prisoner's Dilemma game, as the observed cooperation rate is higher than that predicted by the theory with altruistic players.

We provide an alternative explanation for the observed cooperation in finitely repeated Prisoner's Dilemma games. As before, let us assume that the players are playing the game as shown in Figure 1. However, we now extend our analysis to the case where the players repeatedly play the game 5 times. Our approach to calculating the level of cooperation involves a five step procedure:

(1) First calculate the level of cooperation and defection based on the returns based approach discussed in the previous section for the last game (Game 5). In this case, and as before, the probability of playing the cooperate strategy and the defect strategy would be 0.387 and 0.613 respectively.

(2) Calculate the value of the game. This is the expected value given the probabilities for cooperation and defection respectively. The value of the game is, $2.52 = 0.387*(0.387*(4)+0.613*(0.5))+0.613*(0.387*(6)+0.613*(1))$.

(3) The value of Game 5 is then added to each component of the returns to each player (each column and each row) in the game as shown in Figure 1 to work out the payoff matrix for Game 4.

(4) The cooperation and defect rates are calculated for Game 4 based on returns based beliefs as discussed in the previous section, and the value of the game is calculated as per the method in point (2) above.

(5) The value of Game 4 is then added to the payoff matrix as per point (3) above to get the payoff of Game 3. The same procedure is repeated to calculate the cooperative and defect strategies for Game 3, Game 2 and Game 1 respectively.

Using the procedure outlined above we are able to calculate the cooperate and defect probabilities for each round of the finitely repeated Prisoner's Dilemma game and the results are shown in Table 1.

The results are striking indeed. One of the most significant features of the results

Probabilities	Game 1	Game 2	Game 3	Game 4	Game 5
Cooperate	0.493	0.485	0.471	0.443	0.387
Defect	0.507	0.515	0.529	0.557	0.613

Table 1: Probabilities of Cooperation and Defect from Finitely Repeated Prisoner’s Dilemma

shown in Table 1 is that they are entirely consistent with the observations from previous empirical studies of the finitely repeated Prisoner’s Dilemma game (Cooper 1996). In particular, the results show that

- (1) Cooperation is possible in the finitely repeated Prisoner’s Dilemma game.
- (2) Higher levels of cooperation are evident in early parts of the game with the level of cooperation declining as one approaches the final game.
- (3) The level of cooperation in early parts of the finitely repeated game is higher than in the one shot game.

Significantly, our explanation for the level of cooperation as the repeated Prisoner’s Dilemma game evolves is not based on either reputation or altruism arguments, but emerges directly as a consequence of the strategic interactions resulting from returns based belief formation.

3 Conclusion

The empirical testing of the one-shot Prisoner’s Dilemma game in laboratory experiments has shown that when the temptation to defect is low, or the benefits to cooperation are high, players do not play the Nash equilibrium outcome as suggested by the game (Sally 1995). In this paper, we show that even when the game is fully specified, there is strategic

uncertainty when players try to coordinate their actions, based on an understanding that both cooperating could yield a better outcome. We argue that this uncertainty induces players to play mixed strategies. Although the common practice is to eliminate strictly dominated strategies from normal form games in formulating equilibria in mixed strategies, empirical studies have shown that dominated cooperative strategies might be played with positive probabilities (Cooper et al 1990). Based on such findings, we hypothesize that when agents form subjective probabilities about the strategies of the other players, strategies that were previously eliminated via deletion of dominated strategies cannot be eliminated any longer. This is because the use of mixed strategies might result in better returns than the Nash equilibrium strategy of playing defect. We show that using mixed strategies (including the dominated strategies) allows us to provide a plausible explanation for the empirical evidence discussed in previous studies of the Prisoner's Dilemma. We show that as long as the temptation to defect is not large or the benefit of cooperation is large, then players are likely to play the cooperative strategy. Our approach enables us to provide a plausible explanation for two anomalies between the theoretical predictions of one-shot Prisoner's Dilemma games and the empirical evidence: first, the inverse relationship between cooperation and the temptation to defect; and second, the positive relationship between cooperation and the gains from cooperation. In addition, in the finitely repeated Prisoner's Dilemma game we are able to provide a possible explanation for the observation that (1) higher levels of cooperation are evident in early parts of the game with the level of cooperation declining as one approaches the final game and (2) the level of cooperation in early parts of the finitely repeated game is higher than in the one shot game. Our explanation for the level of cooperation through the evolution of the repeated Prisoner's Dilemma game is not based on other leading theories of cooperation such as reputation or altruism, but on the consequences of strategic interactions as a result of returns-based belief formation. Our proposed new method of returns-based beliefs forming

the subjective probabilities is based upon the expected returns of a particular strategy, in proportion to the total expected returns of all strategies. Moreover, returns-based beliefs are important because we can show that they can be applied even when the players do not know their respective histories. We believe that if returns-based belief formation is in fact a possible explanation for the experimental observations of the Prisoner's Dilemma game, then this might amplify the possibility of testing other observable anomalies in game theory.

References

Abelson, R.M., R.A. Bradley. 1954. A2 X A2 factorial with paired comparisons. *Biometrics*.10(4) 487-502.

Amaldoss, W., S. Jain. 2001. David vs. Goliath: An analysis of asymmetric mixed-strategy games and experimental evidence. *Management Science*. 48(8) 972-991.

Andreoni, J. A., J. H. Miller. 1993. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The Economic Journal*. 103(418). 570-585.

Aumann, R. J., J.H. Dreze. 2008. Rational expectations in games. *American Economic Review*. 98(1), 72-86.

Axelrod, R. 1980a. Effective choice in the Prisoner's Dilemma. *The Journal of Conflict Resolution*. 24(1), 3-25.

Axelrod, R. 1980b. More effective choice in the Prisoner's Dilemma. *The Journal of Conflict Resolution*. 24(3), 379-403.

Axelrod, R. 1984. *The Evolution of Cooperation*. Basic Books. New York. NY.

Axelrod, R., D. Dion 1988. The further evolution of cooperation. *Science*. 242(4884). 1385-1390.

Basu, K. 1990. On the non-existence of a rationality definition for extensive games. *International Journal of Game Theory*. 19 33-44.

Bicchieri, C. 1989. Self-refuting theories of strategic interaction: A paradox of common knowledge. *Erkenntnis*. 30(1-2) 69-85.

Binmore, K. G. 1987. Modelling rational players, Part 1. *Economics and Philosophy*. October 3(2).

Bradley, R. A., M. E. Terry. 1952. Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* 39 (3/4) 324-345.

Brandenberger, A. 1996. Strategic and structural uncertainty in games. R.J. Zeckhauser, R. L. Keeney, J. K. Sebenius, eds. *Wise Choices*. Harvard Business School Press, Boston.

Brandenburger, A. 2007. The power of paradox: some recent developments in interactive epistemology. *International Journal of Game Theory*. 465-492.

Brandenburger, A., E. Dekel. 1987. Rationalizability and correlated equilibrium. *Econometrica*. 55. 1391-1402.

Cooper, R.W., D.V.DeJong, R. Forsythe, T.W. Ross. 1990. Selection criteria in co-ordination games: some experimental results. *American Economic Review*. 80(1). 218-233.

Cooper, R.W., D.V.DeJong, R. Forsythe, T.W. Ross. 1996. Cooperation without reputation: Experimental evidence from Prisoner's Dilemma games. *Games and Economic Behavior*. 12. 187-218.

Dawes, R. M. 1980. Social dilemmas. *Annual Review of Psychology*. 31 169-93.

Dal Bo, P. 2005. Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *American Economic Review*. 95(5) 1591-1604.

DeGroot. M.H. 1975. *Probability and Statistics*. Addison-Wesley, Reading, Mass.

Farell, J., M. Rabin. 1996. Cheap talk. *Journal of Economic Perspectives*. 10(3) 103-118.

Fehr, E., Gächter, S., 2002. Altruistic punishment in humans. *Nature*. 415, 137–140.

Field, A.J., 2001. *Altruistically Included?* The University of Michigan Press, Ann Arbor.

Goeree, J. K., C. A. Holt. 2001. Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*. 91(5) 1402-1422.

Harsanyi, J. C. 1978. Bayesian decision theory and utilitarian ethics. *American Economic Review*. 68(2) May 223-228.

Harsanyi, J.C. 1982. Subjective probability and the theory of games: Comments on Kadane and Larkey's paper. *Management Science*. 28(2)120-124.

Janssen, M. 2008. Evolution of cooperation in a one-shot Prisoner's Dilemma based on recognition of trustworthy and untrustworthy agents. *Journal of Economic Behavior & Organization*. 65 458–471

Kadane, J. B., P.D. Larkey. 1982. Subjective probability and the theory of games. *Management Science*. 28(2) 113-120.

Kadane, J. B., P. D. Larkey. 1983. The confusion of is and ought in game theoretic contexts. *Management Science*. 29(12) 1365-1379.

Kolhberg, E., J-F. Mertens. 1986. On the strategic stability of equilibria. *Econometrica*. 54 1003-38.

Krepps, D.M., P. Milgrom, J. Roberts and R. Wilson. 1982. Rational cooperation in the finitely repeated Prisoners' Dilemma. *Journal of Economic Theory* 27 245-252.

Luce, D. R. 1959. *Individual Choice Behavior*. Wiley, New York.

Luce, D. R., H. Raiffa. 1957. *Games and Decisions*. Wiley, New York.

Morris, S., H.S.Shin, 2002. Measuring strategic uncertainty. Paper prepared for the Pompeu Fabra workshop on “Coordination, Incomplete Information, and Iterated Dominance: Theory and Empirics”, August 18-19.

Nash, J. F. 1951. Non-cooperative games. *Annals of Mathematics*. 54(2) 286-95.

Pettit, P., R. Sugden. 2001. The backward induction paradox. Y. Varoufakis, ed. *Game theory: Critical concepts in the social sciences*. Volume 4. Discontents. Routledge, London and New York, 103-16.

Rabin, M. 1993. Incorporating Fairness into Game Theory and Economics. *The American Economic Review*. 83(5) 1281-1302.

Reny, P. J. 1993. Common belief and the theory of games with perfect information. *Journal of Economic Theory*. 59(2) 257-74.

Rosenthal, R. W. 1981. Games of perfect information, predatory pricing and the chain-store paradox. *Journal of Economic Theory*. 25(1) 92-100.

Roth, A., F. Schoumaker. 1983a. Expectations and reputations in bargaining: An experimental study. *The American Economic Review*. 73(3) 362-372.

Roth, A., F. Schoumaker. 1983b. Subjective probability and the theory of games. *Management Science*. 1983b. 29(11) 1337-1340.

Rubinstein, A. 2007. Instinctive and cognitive reasoning: Response times study. *Economic Journal*. 117 1243-1259.

Rubinstein, A. 2006. Dilemmas of an economic theorist. *Econometrica*. 74 865-883.

Sally, D. 1995. Conversation and cooperation in Social Dilemmas. *Rationality and Society*. 7(1) 58-92.

Selten, R. 1978. The chain-store paradox. *Theory and Decision*. 9(2) 127-59.

Shubik, M. 1970. Game theory, behavior, and the paradox of the Prisoner’s Dilemma: Three solutions. *The Journal of Conflict Resolution*. 14(2) 181-193.

Simon, H. A. 1957. Comparison of game theory and learning theory. *Models of Man*. Wiley, New York, 274-279.

Van Huyck, J.B., R.C. Battalio, R.O. Beil. 1990. Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*. 0(1) 234-248.

Wilson, J. G. 1986. Subjective probability and the prisoner's dilemma. *Management Science*. 32(1), 45-55.