

A Reason for Unreason: Returns-Based Beliefs in Game  
Theory

Chander Velu, Sriya Iyer and Jonathan R. Gair

September 2010

CWPE 1058

# A Reason for Unreason: Returns-Based Beliefs in Game Theory

Chander Velu

*Judge Business School, University of Cambridge,  
Trumpington Street, Cambridge, CB2 1AG, UK<sup>\*†</sup>*

Sriya Iyer

*Corresponding author: Faculty of Economics, University of Cambridge,  
Sidgwick Avenue, Cambridge CB3 9DD, UK<sup>‡</sup>*

Jonathan R. Gair

*Institute of Astronomy, Madingley Road, CB3 0HA Cambridge, UK<sup>§</sup>*

(Dated: August 5, 2010)

## Abstract

Players cooperate in experiments more than game theory would predict. We introduce the ‘returns-based beliefs’ approach: the expected returns of a particular strategy in proportion to total expected returns of all strategies. Using a decision analytic solution concept, Luce’s (1959) probabilistic choice model, and ‘hyperpriors’ for ambiguity in players’ cooperability, our approach explains empirical observations in various classes of games including the Prisoner’s and Traveler’s Dilemmas. Testing the closeness of fit of our model on Selten and Chmura (2008) data for completely mixed  $2 \times 2$  games shows that with loss aversion, returns-based beliefs explain the data better than other equilibrium concepts.

Keywords: Rationality, Subjective Probabilities, Returns-Based Beliefs.

---

\* For helpful comments we would like to thank David Banks, Kaushik Basu, James Bono, Adam Brandenberger, Guilherme Carmona, Vincent Crawford, Philip Faulkner, Alberto Feduzi, Daniel Friedman, Jayant Ganguli, Vincent Mak, N. Panchapakesan, Amnon Rappoport, David Wolpert and Jun Xue.

<sup>†</sup>Electronic address: [c.velu@jbs.cam.ac.uk](mailto:c.velu@jbs.cam.ac.uk)

<sup>‡</sup>Electronic address: [sriya.iyer@econ.cam.ac.uk](mailto:sriya.iyer@econ.cam.ac.uk)

<sup>§</sup>Electronic address: [jgair@ast.cam.ac.uk](mailto:jgair@ast.cam.ac.uk)

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). One of the most widely analyzed games in economics, mathematics, biology and other sciences is the Prisoner's Dilemma, a two-by-two noncooperative game. The Prisoner's Dilemma lies at the heart of important concepts in game theory such as the 'Nash equilibrium' (Nash 1951). Empirical tests demonstrate that in the real world people are often more cooperative than that predicted by the outcome of the Prisoner's Dilemma game and other variations such as the Traveler's Dilemma game. In this paper we provide a reason for this unreason<sup>1</sup>: an explanation for why cooperative strategies might be played in the one-shot Prisoner's Dilemma and the Traveler's Dilemma games. We propose an alternative equilibrium concept to the Nash by which people might form their beliefs to play their strategies which we call '*returns-based beliefs*'. We show that this might explain better cooperation in the Prisoner's Dilemma and Traveler's dilemma games. We also test the closeness of fit of our returns-based beliefs model with hyperpriors, and by incorporating loss aversion, using data from Selten and Chmura (2008) for completely mixed  $2 \times 2$  games. We show that the returns based belief model is able to explain the data better than the other stationary concepts as outlined in Selten and Chmura (2008).

In Nash equilibrium, each player chooses the action that maximizes their returns subject to the opponent's choice and no player can gain by changing their strategy unilaterally. By contrast, we use a decision analytic approach where individuals form subjective probabilities over the actions of the individual's opponent and choose a mixed strategy profile over the actions based on the relative returns<sup>2</sup>. We use the probabilistic choice model developed axiomatically by Luce (1959). The returns based beliefs approach is both more sympathetic

---

<sup>1</sup> We attribute this expression to Herbert Simon (1987) who wrote, 'Sometimes the term rational (or logical) is applied to decision making that is consciously analytic, the term nonrational to decision making that is intuitive and judgmental, and the term irrational to decision making and behaviour that responds to the emotions or that deviates from action chosen "rationally". We will be concerned then, with the nonrational and the irrational components of managerial decision making and behaviour. Our task, you might say, is to discover the reason that underlies unreason.' (Simon, 1987, p.57)

<sup>2</sup> Kadane and Larkey (1982) in their seminal paper propose a decision analytic solution concept over the game theoretic solution concept. Rios, Rios and Banks (pp. 845-849, 2009) provide a recent discussion of the difference between game theoretic and decision analytic solution concepts. Also, see Bono and Wolpert (2009).

to and more consistent with the empirical testing of games. Returns-based beliefs bring squarely into the picture the emphasis on the relative attractiveness of individual actions in choosing the optimal strategy. Our approach treats *beliefs* rather than *strategies* as the primary concept (p.139, Binmore 2009; Nau and McCardle 1990). In doing so, we assume that players' subjective beliefs are in equilibrium with reference to each other. Therefore, consistent with a decision theoretic perspective, players adopt strategies on the basis of their respective subjective beliefs.

Our model has similarities with the Quantal Response Equilibrium (QRE) model proposed by McKelvey and Palfrey (1995) and the Boundedly Rational Nash Equilibrium (BRNE) model of Chen, Friedman and Thisse (1997). The QRE can be viewed as an application of stochastic choice theory to strategic games or as a generalization of the Nash equilibrium that allows noisy optimization (Haile, Hortacsu and Kosenok 2008). The QRE and its close variant the BRNE assumes that the decision maker might take a suboptimal action, and that the probability of doing so increases with the expected payoff of the action. We follow the BRNE in that decision makers are not expected utility maximizers in the conventional sense but exhibit a tendency toward utility maximization that is not optimal (Chen, Friedman and Thisse 1997). Each decision maker follows the random choice interpretation of discrete choice theory under which each player selects a mixed strategy in order to achieve the best possible outcome given the uncertainty regarding the other player's choice. In the BRNE and QRE models, the decision maker plays strategies proportional to the expected payoff with some error. The error the decision maker makes could be interpreted as either unmodeled costs of information processing (McKelvey and Palfrey 1995) or as unmodeled determinants of utility from any particular strategy (Chen, Friedman and Thisse 1997). We build on the QRE and BRNE models to develop our approach. Our model has some mathematical similarities to the QRE and BRNE. In particular, the 'returns-based beliefs' is the expected returns of a particular strategy, in proportion to the total expected returns of all strategies. However, the returns based beliefs approach differs significantly from the QRE and the BRNE because our approach does not assume unmodeled costs of information processing or unmodeled determinants of utility from any particular strategy. We also do not assume that the decision maker is making errors or mistakes. Rather in the returns based beliefs model, individuals need to form subjective beliefs about each others possible plays. Since there is strategic uncertainty or ambiguity about these beliefs we invoke

the concept of hyperpriors in the formation of such beliefs. In this respect our model differs both conceptually and mathematically from both the QRE and the BRNE. Our model based upon returns-based beliefs is also different to other alternative models that provide an explanation for non-equilibrium outcomes such as the level-k models (Costa-Gomes and Crawford 2006) and cognitive hierarchy (CH) model (Camerer, Ho and Chong 2004). The level-k and CH models assume that decision makers are heterogenous in their level of sophistication with respect to their strategic thinking. The returns-based beliefs model differs from the level-k and the CH models in that the former does not assume heterogeneity in the levels of sophistication in thinking by decision makers.

We first show how the returns based belief model without ambiguity is able to provide qualitative results that are consistent with empirical findings from the Prisoner's Dilemma and the Traveler's Dilemma. We then show that the returns based belief model with hyperpriors to incorporate ambiguity is able to better explain the empirical data from completely mixed  $2 \times 2$  games from Selten and Chmura (2008) once loss aversion is included compared to other stationary concepts. Thus, by putting forward the concept of returns-based beliefs we contribute to the extensive literature which tries to reconcile game theoretic predictions and empirical experimental results (Holt and Roth 2004). The next section develops the RBB model with the one-shot Prisoner's Dilemma game and the Traveler's Dilemma game as the context. Section II applies the RBB model to Selten and Chmura's (2008) completely mixed  $2 \times 2$  games. Section III concludes.

## I. RETURNS-BASED BELIEFS

Game theory models systematically human behavior when strategic interactions exist. In conventional game theory, the solution concept such as 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. 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' (Harsanyi

1982, 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 judgment, the likelihood that the outcome will be obtained (DeGroot 1975, p. 4; Savage 1954). 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. This approach calls for a decision theoretic approach whereby players form subjective beliefs about the opponent in choosing an action (Roth and Schoumaker 1983). However, as pointed by Myerson (1991, p.114), 'A fundamental difficulty may make the decision-theoretic approach impossible to implement, however. To assess his subjective probability distribution over other players' strategies, player  $i$  may feel that he should try to imagine himself in their situations. When he does so, he may realize that the other players cannot determine their optimal strategies until they have assessed their subjective probability distributions over  $i$ 's possible strategies. Thus, player  $i$  may realize that he cannot predict his opponents' behavior until he understands what an intelligent person would expect him rationally to do, which is, of course, the problem that he started with.' To resolve the issue, psychological motives could be used to understand behavior in game theoretic settings. Economists have suggested examining the process of cognitive reasoning (Rubinstein 2007). Research in psychology and economics shows that people have a bias towards wanting to cooperate (Farrell and Rabin 1996; Andreoni and Miller 1993). We follow this line of reasoning to propose a reasonable way of forming such subjective probabilities.

We argue that the willingness to cooperate might be influenced by past experience, generating 'strategic uncertainty' regarding the conjecture about the choice of the other player. We define strategic uncertainty as uncertainty concerning the actions and beliefs (and beliefs about the beliefs) of others (Brandenberger 1996). Researchers have argued that strategic uncertainty can arise even when all possible actions and returns are completely specified and are common knowledge (Ellsberg 1959; 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 players<sup>3</sup> (Brandenberger and Dekel 1987). This is succinctly summarized by Rabin 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). Following this line of reasoning, a player who knows that the non-Nash equilibrium belief is held by the opponent could form a subjective assessment of the opponent’s play by taking this belief into account (Basu 1990). For example, in the Prisoner’s Dilemma a player might play the cooperate strategy if they expect the other player to play the cooperate strategy. Therefore, player 1 has a profitable deviation by playing cooperate with a positive probability when player 2 plays cooperate with a positive probability<sup>4</sup>.

In order to illustrate our model 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 Luce’s (1959) probabilistic choice model. In this approach, players form subjective beliefs and then act based upon the expected returns, given these beliefs, of a particular strategy, in proportion to the total expected returns of all strategies. We call this the returns-based beliefs model.

### **A. The model**

In this section we develop the basic elements of a one-shot game and then introduce discrete choice theory based on subjective probabilities. We use below a particular variation

---

<sup>3</sup> We 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, p.47) points out, although it makes little difference to the mathematics, conceptually this distinction between randomization and subjective beliefs to explain the mixed strategies is a pertinent one. This interpretation is also sympathetic to Harsanyi’s (1973) purification interpretation of mixed strategy where mixing represents uncertainty in a player’s mind about how other players will choose their strategies, rather than deliberate randomization (Morris 2008).

<sup>4</sup> Conventionally, 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 which includes the cooperate strategy. Therefore, a mixed strategy can exist if one player experiences uncertainty with respect to his conjecture about the choice of the other player.

of the one-shot game formulation of McKelvey and Palfrey (1995). Let  $(N, S, \pi)$  be a finite game.  $N$  denotes the set of players. Each player  $i \in N$  has a set of pure strategies,  $S^i$  with elements  $s^i$ . The set of strategy profiles of all players other than  $i$  is  $S^{-i} = \otimes_{j \neq i} S^j$  with elements,  $s^{-i}$ . The benefit a player  $i$  derives from playing a strategy profile  $s^i \in S^i$  is  $\pi_{s^i}^i$ ;  $\pi^i = \{\pi_{s^i}^i\}_{s^i \in S^i}$ . Each player knows who is in the game,  $N$ , and the strategy sets available to each other,  $S^i \forall i \in N$ . However, each player is ignorant of the subjective probability distribution held by the other player. We discuss below our approach to how players form a reasonable subjective probability belief structure.

In the following discussion, we assume that each decision maker follows the random choice interpretation of discrete choice theory under which each player selects a mixed strategy in order to achieve the best possible outcome given the uncertainty regarding the other player's choice. 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, p. 115). Kadane and Larkey (1982, p. 115) have expressed the implications of this line of thought very neatly as follows: *'If I think my opponent will choose strategy  $i$  ( $i = 1, \dots, I$ ) 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  $j$ .....the opponent's utilities are important only in that they affect my views  $\{p_i\}$  of what my opponent may do....'* Although any possible distribution of probabilities could be a possibility based upon the subjective method of forming them, we propose a reasonable subjective probability belief that the players might use when they do not know each other or their respective histories. We call this *'returns-based beliefs'*.

We believe it is reasonable to assume that the decision maker would assign probabilities based on the expected returns from playing the different strategies. Similarly, we 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 of 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  $j$  of player  $i$  can be expressed as a ratio of the expected payoff for that decision over the total expected payoff for all decisions:  $\Pi_j^i / \sum_k \Pi_k^i$ , where  $\Pi_j^i$  is the expected returns associated with decision  $j$ . 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 of choosing  $x$  over  $y$ ,  $P(x, y) = v(x) / [v(x) + v(y)]$  where  $v(x)$  and  $v(y)$  are the scale values of choosing  $x$  and  $y$  respectively (Abelson and Bradley 1954). Our model has similarities with the Quantal Response Equilibrium model (QRE) proposed by McKelvey and Palfrey (1995) and the Bounded Rationality Nash Equilibrium (BRNE) model of Chen, Friedman and Thisse (1997) in that all strategies with a positive payoffs are played with positive probabilities in proportion to their expected payoffs. However, we differ from the QRE and the BRNE models in an important respect in that our model does not require the specification of an error distribution. On the other hand, we invoke the concept of *hyperpriors* whereby players form beliefs over their beliefs in order to take into account the ambiguity related to the cooperative stance of the opponent. We discuss these points further below in Section II on comparing the returns based belief model to other non-equilibrium models.

In a game between multiple players, we assume that each player has a probability distribution over the choices available. This probability distribution,  $P^i$ , over the elements  $S^{-i}$  is defined such that  $P^i(s^{-i})$  is the probability associated with  $s^{-i} \in S^{-i}$ . We operationalize our model by assuming an expected return framework, so that the expected payoff of the  $j$ th pure strategy of player  $i$ , given  $P^i$ , is as follows:

$$\Pi_j^i(P^i) = \sum_{s^{-i} \in S^{-i}} P^i(s^{-i}) \pi_j^i(s^{-i}) \quad (1)$$

where  $\pi_j^i(s^{-i})$  is player  $i$ 's payoff from choosing a pure strategy  $j$  when the other players choose  $s^{-i}$  and  $P^i(s^{-i})$  is the belief held by player  $i$  about the probability the other players will choose  $s^{-i}$ . The decision probabilities over pure strategies for player  $i$  in turn follow the

specification outlined above which is proportional to the expected returns as follows:

$$p_j^i = \frac{\Pi_j^i(P^i)}{\sum_{k=1}^m \Pi_k^i(P^i)} \quad (2)$$

This model admits a Nash-like equilibria in belief formation such that the belief probabilities match the decision probabilities for all players. This equilibrium in beliefs can be found by iterating between the expected payoff in equation (1) and the decision probabilities in equation (2)<sup>5</sup>.

The returns based belief model can be used to find equilibria for games between any number of players. However, for clarity, the games discussed in the rest of this paper will be games played between 2 players only. We will use  $u_{ij}^1$  and  $u_{ij}^2$  to denote the payoffs to player 1 and 2, respectively, if player 1 plays move  $i$  and player 2 plays move  $j$ . We will denote a mixed strategy of player 1 by  $\mathbf{p} = \{p_i\}$ , where  $p_i$  is the probability that player 1 chooses  $s^i \in S^1$ , and denote a mixed strategy of player 2 by  $\mathbf{q} = \{q_j\}$ , where  $q_j$  is the probability that player 2 chooses  $s^j \in S^2$ . We will use a superscript ‘ $b$ ’ to denote belief probabilities, i.e.,  $\mathbf{q}^b$  is the mixed strategy that player 1 believes player 2 will play and vice versa for  $\mathbf{p}^b$ . We will denote the equilibrium solution in which  $\mathbf{p} = \mathbf{p}^b$  and  $\mathbf{q} = \mathbf{q}^b$  by a superscript ‘rbb’, i.e.,  $\mathbf{p}^{\text{rbb}}$ ,  $\mathbf{q}^{\text{rbb}}$ .

We assume that the way players act on their beliefs is based on the Luce (1959) expected returns approach. However, there will remain some ambiguity as to the cooperative stance of the other player and hence their disposition to form beliefs in such a manner. We capture this ambiguity via the concept of hyperpriors that are formed over the beliefs developed via the expected returns approach as outlined above. Following this interpretation, the belief probability will arise as an integral over player 1’s belief distribution,  $P(\mathbf{q})$ , for player 2

$$q_i^b = \int q_i P(\mathbf{q}) d\mathbf{q} = q_i^{\text{rbb}} \int \left[ \frac{q_i}{q_i^{\text{rbb}}} P(\mathbf{q}) \right] d\mathbf{q} = q_i^{\text{rbb}} X_i \quad (3)$$

where the last equality defines  $X_i$  from the belief distribution<sup>6</sup>. Writing the moments of the belief probability distribution in this way is an attempt to split the belief probability into a game dependent part, the equilibrium solution  $q_i^{\text{rbb}}$ , and a player dependent part,  $X_i$ . We

<sup>5</sup> Similar to the discussion in Aumann and Brandenburger (1995) on the epistemic conditions to get a Nash equilibrium in a two player game, mutual knowledge between the players that they are cooperative, i.e., play a Luce type strategy, would be sufficient to get a returns based belief equilibrium.

<sup>6</sup> Note that  $\sum X_j q_j^{\text{rbb}} = 1$  since  $\sum q_j = 1$ .

expect players who believe their opponent to be cooperative to play close to the equilibrium solution, and therefore  $X_i \approx 1$ . The Luce (1959) expected returns model provides a mapping between beliefs and actions, which allows the determination of these  $X_i$  parameters from experimental data. This will be explored further in Section II.

In standard statistical language, the strategy  $\mathbf{q}^b$  is player 1's prior belief on the pure strategy that player 2 will play. The individual components of  $\mathbf{q}$ , which characterize the prior distribution, are hyperparameters. Player 1 may not be certain about the mixed strategy that player 2 will adopt and so he may have a distribution over the possible mixed strategies of player 2 and hence over the hyperparameters. Such a distribution is termed a hyperprior and is denoted by  $P(\mathbf{q})$  in equation (3). In a theoretical approach, one could prescribe a hyperprior and derive the corresponding  $X_i$ 's from equation (3). However, the consequences of the hyperprior are entirely encoded in the mixed belief strategy that the player forms, and hence in the  $X_i$ 's. Thus, when presented with experimental data it is more practical to work with the measured  $q_i^b$ 's or  $X_i$ 's that can be determined from the observed strategies through the Luce (1959) rule.

We first examine the returns based belief model without ambiguity to provide qualitative results that are consistent with empirical results of the Prisoner's Dilemma and the Traveler's Dilemma. In particular, due to the limitation of the data for the Prisoner's Dilemma and Traveler's Dilemma games we will only consider the returns-based-belief equilibrium solution,  $(\mathbf{p}^{\text{rbb}}, \mathbf{q}^{\text{rbb}})$ , in which the belief and true probabilities coincide, i.e.,  $\mathbf{p}^b = \mathbf{p}$  and  $\mathbf{q}^b = \mathbf{q}$ . We then go on to examine the empirical evidence from the Selten and Chmura (2008) data in section II B with the returns based belief model incorporating ambiguity using hyperpriors as described here<sup>7</sup>.

---

<sup>7</sup> We note that deviations from the RBB equilibrium solution could also arise if a player makes an error in computation when evaluating the values associated with each move, and these values are ultimately used to determine their choice probabilities. In this case, we can use  $Y_i$  to represent the fractional error that player 1 makes when assessing the value of move  $i$ , so that he assigns a value  $Y_i v_i^{\text{rbb}}$ , where  $v_i^{\text{rbb}}$  is the value of that move in the RBB equilibrium solution. As a result of these errors, player 1 will play a strategy with the belief probabilities

$$q_i^b = \frac{Y_i v_i^{\text{rbb}}}{\sum_{j=1}^N Y_j v_j^{\text{rbb}}} = \frac{Y_i q_i^{\text{rbb}}}{\sum_{j=1}^N Y_j q_j^{\text{rbb}}} = X_i q_i^{\text{rbb}} \quad (4)$$

where  $\sum X_i q_j^{\text{rbb}} = 1$ . Mathematically, this interpretation is therefore identical to the ambiguity in beliefs characterised by the hyperprior.

To find the returns-based-belief equilibrium solution, the idea is that player 1 computes the expected payoff,  $\Pi^i(s_i, \mathbf{q}) = \sum u_{ij}^1 q_j$  of each pure action  $s_i \in S^1$ , given a mixed action  $\mathbf{q}$  of the other player. Player 1 then plays the mixed action,  $\mathbf{p} = \{p_i\}$ , where  $\{p_i\}$  is the probability of playing move  $s^i \in S^1$ , and

$$p_i = \frac{\sum_j u_{ij}^1 q_j}{\sum_i \sum_j u_{ij}^1 q_j} \quad (5)$$

if player 1 knew that the opponent played  $\mathbf{q}$ . This defines a mapping from player 2's mixed action to player's 1 mixed actions. Similarly player 2 would play a mixed action,  $\mathbf{q} = \{q_i\}$ , where  $\{q_i\}$  is the probability of playing move  $s^i \in S^2$  and

$$q_i = \frac{\sum_j u_{ji}^2 p_j}{\sum_i \sum_j u_{ji}^2 p_j} \quad (6)$$

if player 2 knew that player 1 played  $\mathbf{p}$ . The combination of these mappings defines a mapping from player 1/2 mixed strategies to player 2/1 mixed strategies and we assume the players then play a fixed point of this mapping. Results in Chen et al. (1997) can be used to show that there exists such a fixed point equilibrium in the case that the pay-offs to both players from all moves are non-negative<sup>8</sup>. Hence the equilibrium can be thought of as a set of consistency conditions whereby a player's choice probabilities are correct in the sense of reflecting what the player must do when the choice probabilities are best replies to the strategies of the other player who forms such subjective choice probabilities in proportion to the expected returns of the strategies. A more general set up of the returns based belief approach is described in Appendix I. We illustrate the existence of such a fixed point equilibrium through the numerical example below.

## B. The Prisoner's Dilemma

We consider the Prisoner's Dilemma game 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

---

<sup>8</sup> Chen et al. (1997) show that a fixed point equilibrium exists if probabilities are assigned in proportion to  $u^\mu$ , where  $u$  is the expected utility of the move and  $\mu$  is a 'rationality' parameter which varies between 0 and  $\infty$ . This relies on the assumption that the utilities,  $u$ , are positive. Although the basis of our model is very different, our model relates to the Chen et al. (1997) model when  $\mu = 1$  and the utility is taken to be the expected return from the move. This is the case without a hyperprior. Therefore, the Chen et al. (1997) proof will carry through to our model in the case where the pay-offs are non-negative. We discuss more general scenarios in Appendix I.

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

FIG. 1: Payoffs for the Prisoner’s Dilemma game.

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 regardless of what the other player does. Playing defect is the dominant strategy for both players and is the only Nash equilibrium of the Prisoner’s Dilemma as there is no incentive for any of the players to change their strategies. This is clearly less than the Pareto optimum of obtaining a payoff of 4 each by both cooperating. Yet, empirical testing of the one-shot Prisoner’s Dilemma in laboratory experiments has shown that people are prone to play the cooperative strategies far more often than the Nash equilibrium might lead us to predict. 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).

Scholars have emphasized that cooperation in a one-shot Prisoner Dilemma game is an important finding of experimental research that needs to be further understood (Janssen 2008). 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). To all of these explanations, we add an alternative: a new explanation for cooperation in a one-shot Prisoner’s Dilemma game that is based upon subjective probabilities and returns based beliefs.

In the case of the Prisoner’s Dilemma, ideally, the agents would like to cooperate by coordinating their actions on the joint strategies 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, which coordinates with the other player. For example, research has shown that human beings are prone to cooperative behavior based on reciprocity (Axelrod 1984; Farrell and Rabin 1996). Therefore, the history of human interactions is likely to influence the general disposition of players to want to cooperate<sup>9</sup>. 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 outcomes<sup>10</sup>.

In the Prisoner's Dilemma,  $S^i = \{C, D\}$  and  $N = \{1, 2\}$ . We could begin with any arbitrary beliefs and through a process of iteration each player updates their beliefs using the expected payoff in equation (1) and the decision probabilities in equation (2). 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. We discuss the case with hyperpriors on the beliefs later in this paper.

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' (Sally 1995, p.75). 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

---

<sup>9</sup> Farrell 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.

<sup>10</sup> To 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 (Farrell and Rabin 1996). 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).

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, which for 2 players reduce to

$$\text{Loss} = \frac{C(2) - D(0)}{C(2)} \quad (7)$$

$$\text{Temptation} = \frac{D(1) - C(2)}{C(2)} \quad (8)$$

where  $C(x)/D(x)$  are the payoffs to cooperators/defectors when there are  $x$  cooperators. For the game in Figure 1, the value of the indices are 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 to increase from 4 to 5.8, the loss index increases from 75.0% to 82.8% which in turn increases the probabilities of cooperation from 0.39 to 0.47. 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. This result is consistent with empirical evidence from the Prisoner's Dilemma game.

However, the question that remains is whether a one unit increase in the returns to the  $(C, C)$  outcome has a stronger or weaker effect on the willingness to cooperate compared to a one unit decrease in the  $(D, C)$  payoff. Would a policymaker trying to increase the incidence of cooperation be better off trying to increase the returns from cooperation or decrease the temptation to defect from unanimous cooperation? If the  $(C, C)$  payoff in the PD game of Figure 1 is increased by one unit to 5, we find the probability of cooperation increases from 0.39 to 0.44. However, if instead the  $(D, C)$  payoff to the defector is decreased by one unit to 5, the probability of cooperation increases by a lesser amount from 0.39 to 0.42. In both cases, the temptation index decreases (from 50% to 20%/25%) but in the first case the loss index also increases (from 75% to 80%). This indicates that policy makers who are interested in encouraging cooperation are better suited to focus their policy on increasing the benefits from cooperation rather than on decreasing the benefits from non-cooperation from a similar level of effort<sup>11</sup>.

Our approach is different to previous studies that show the plausibility of cooperation in the one-shot Prisoner's Dilemma game discussed earlier. In this paper we propose an

---

<sup>11</sup> While we have illustrated this result using the particular game shown in Figure 1, this result carries over to any PD-type game. A proof is given in Appendix II.

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.

### C. The Prisoner's Dilemma with Loss Aversion

The argument we have made above for a player to, rationally, choose to play the dominated cooperate strategy in the Prisoner's Dilemma is that the (C,C) solution has a better payoff than the natural Nash solution of (D,D). The idea is that a player is willing to cooperate based on the belief that the other is willing to cooperate in the hope that this tendency will lead to a better payoff overall. Two general features that we would like a model of this decision process to include would be 1) if the (C,C) strategy does not pay better than the (D,D) strategy, then there should be zero probability that the player would choose C; 2) a player who is averse to risk should be less inclined to play C, since there is a risk that his opponent will defect, leading to a worse payoff for himself. The model that we have adopted, in which the players play mixed strategies according to the Luce decision rule, does indeed exhibit both of these features, if the PD payoff matrix is transformed to include loss aversion (see, for example, Selten and Chmura (2008) and Brunner et al. (2010)).

Player 1 can compute his *guaranteed payoff* by finding the max-min of the elements in his payoff matrix,  $u^1$ , i.e., for each possible move,  $i$ , he computes his minimal payoff,  $\min_j u_{ij}^1$ , over possible strategies of his opponent, and then finds the maximum,  $\max_i(\min_j u_{ij}^1)$ . If the maximum occurs for  $i = I$ , for which the minimum is at  $j = J$ , then by playing move  $I$  the player could guarantee a payoff at least as big as  $u_{IJ}^1$ . The matrix can then be transformed by subtracting  $u_{IJ}^1$  from each element of the matrix. The transformed payoff matrices represent the payoffs relative to the guarantee. The RBB approach naturally places unfavorable weight on negative payoffs, so this transformation already encodes loss aversion. However, the standard approach is to multiply negative entries in the transformed matrix by a multiplier,  $\lambda$ . This multiplier is a measure of the degree of loss aversion of the player.

In Appendix III we analyze the PD game including loss aversion and show that it has



precisely the features we wanted (as noted in (1) and (2) above). A solution with some cooperation exists if there is an incentive to mutual cooperation, i.e., the payoff from defect-defect is worse than that from cooperate-cooperate. As that incentive goes away, so does the tendency to cooperate. In addition, we see that the more loss averse a player is, the less inclined they will be to cooperate and sufficiently loss-averse players will always choose the Nash equilibrium. These nice features illustrate why the Luce choice rule provides a good model for cooperative behavior.

#### D. The Traveler's Dilemma

Another game that has attracted the interest of economists and others is Basu's (1994) Traveler's Dilemma game that is based on a story about two travelers who holiday on a tropical island and then return having purchased identical antiques. Whilst returning, the airline that they have flown back on damages their antiques irreparably, but promises them adequate compensation, requesting them to make claims for that compensation independently. The airline manager, who is unaware of the true cost of the antiques, makes the following proposition to the two travelers: Each traveler has to write down the cost of the antique  $s^i$  ( $i = 1, 2$ ), which can take a value between 80 and 200<sup>12</sup>.

If traveler 1 and traveler 2 write down the same number ( $s^1 = s^2$ ), then the manager assumes that they are telling the truth and both travelers are paid the sum of money written down. If traveler 1 writes down a number larger than traveler 2 ( $s^1 > s^2$ ), then it is assumed that traveler 1 is lying relative to traveler 2. In this case, the airline manager regards  $s^2$  as the cost of the antique and pays traveler 1 the sum of  $(s^2 - r)$ , while traveler 2 gets the sum of  $(s^2 + r)$  where  $r$  is the reward/penalty. Traveler 1 thus receives a penalty for inflating the cost of the antique, while traveler 2 is suitably rewarded for his honesty. The Traveler's Dilemma game thus involves choosing the amount to claim,  $(s^1, s^2)$ , to maximize the travelers' respective payoffs. It is well known that the unique Nash equilibrium of this game is for both players to play (80, 80). However, studies of the Traveler's Dilemma have asserted that when the dilemma is tested in empirical laboratory situations, when the

---

<sup>12</sup> In the original paper (Basu 1994), the cost of the antique can take a value between 2 and 100. However, to avoid negative payoffs we have altered the payoff to conform with the examples in empirical studies (Capra et al 1999, Goeree and Holt 2001).

penalty is small, the Nash outcome does not obtain and when the penalty is large, the outcome is very near to the Nash outcome of (80, 80) (Capra et al 1999, Goeree and Holt 2001). One of the most striking features of these studies is that the anomalous result for the low penalty case does not disappear even when subjects play the game repeatedly and so have the benefit of learning from past experience.

We show that a plausible explanation for the empirical evidence conducted on this game may be based upon agents playing mixed strategies through a succession of profitable deviation from the Nash equilibrium strategy of playing 80 based on the returns based beliefs model outlined earlier. We can follow the same iteration process between players 1 and 2 whereby each player updates their beliefs using the expected payoff in equation (1) and the decision probabilities in equation (2) until the probabilities converge. The results are shown in Figure 2. Players 1 and 2 are most likely to play the claim strategy that maximizes this expected value. Figure 2 shows the expected values for different penalties such as 5, 30, 40 and 60. The expected value is maximized at claim 190 for penalty value 5, 138 for penalty value 30, 114 for penalty value 40, and 80 for penalty value 60. The claim strategy decreases from close to 200 for small penalty values to fall sharply to the Nash equilibrium value at about penalty level 55, and remains at this level thereafter for higher penalty values. Table I shows the claims for the game based on returns-based beliefs (Column 1), game one in Capra et al (Column 2) and the average for game 8-10 in Capra et al (Column 3) which incorporates the ability of players to learn. As shown in Figure 2 this feature of the data conforms closely to the empirical observation of the Traveler's Dilemma in studies by Capra et al (1999) and Goeree and Holt (2001) which show the '*intuitive inverse relationship between average claims and the parameter that determines the relative cost of having the higher claim*' (Capra et al 1999, p.680).

Our explanation for the reconciliation between the theoretical Nash equilibrium outcome and the empirical findings are different to the ones provided by Capra et al (1999) in several ways. First, the Capra model has a time element whereby there is learning taking place over time between the actual play and the predicted claims based on the model. Therefore, their model has an error parameter and a learning parameter to allow for this adjustment to take place. Since our model does not assume learning over time, these parameters are not relevant for our analysis. Second, the Capra model assumes that the decision probabilities

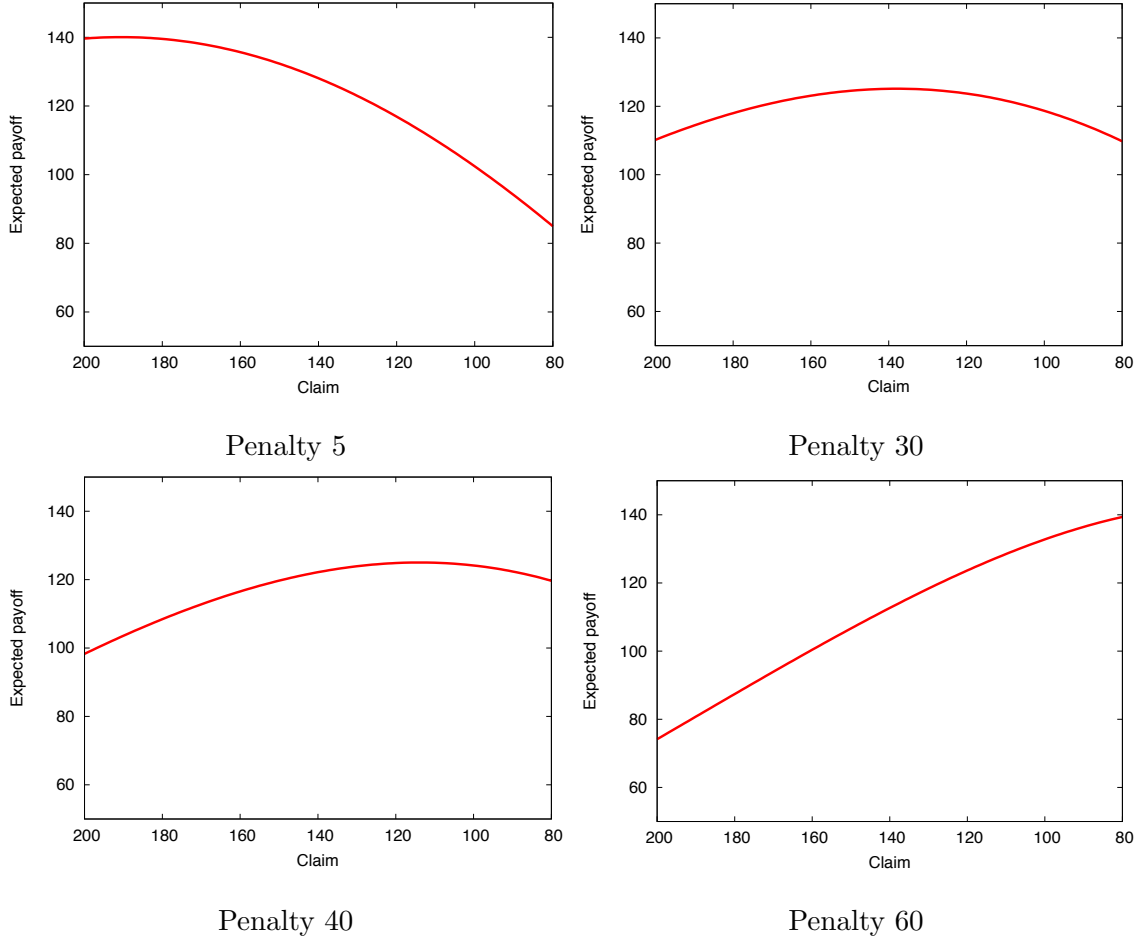


FIG. 2: Expected return versus claim in the Traveler’s Dilemma game, for penalties of 5, 30, 40 and 60, as labelled.

are proportional to an exponential function of the expected payoff<sup>13</sup>.

In the next section we will test our model further using empirical data from Selten and Chmura (2008) for  $2 \times 2$  fully mixed games. We discuss the model with loss-aversion and the concept of hyperpriors.

<sup>13</sup> Decision probabilities that are exponential functions of expected payoffs implies that the choice probabilities are unaffected by adding a constant to all expected payoffs (Capra et al 1999, pp. 683). The equivalent of the Capra model for decision probabilities without a time element and error parameter is  $p_j^i = \exp(\pi_j^i) / \sum_{k=1}^m \exp(\pi_k^i)$ .

Penalty	Maximum Claim	Capra et al.	Capra et al.
	Returns Based Beliefs (Column 1)	Game 1	Average for Game 8-10
5	190	178	195
10	180	180	186
20	160	160	119
25	149	150	138
50	86	130	85
80	80	120	81

TABLE I: Penalty and maximum claim in the Traveler’s Dilemma game

## II. EXPERIMENTAL $2 \times 2$ GAMES

In order to test the applicability of the RBB model in an empirical setting we will look at data from Selten and Chmura (2008). They presented empirical results for 12 games, each played between two players and with two possible moves for each player. The games are asymmetric, games 1-6 are constant sum games, but games 7-12 are non-constant sum. The payoffs for all twelve games are shown in Figure 3. The constant sum and non-constant sum games form pairs (e.g. Game 1 with Game 7, Game 2 with Game 8 and so on). As noted by Selten and Chmura (2008, p. 950), ‘the non-constant sum game in a pair is derived from the constant sum game in the pair by adding the same constant to player 1’s payoff in the column for R and player 2’s payoff in the row for U’.

### A. Loss aversion

Before analyzing the data from the Selten games, we first transform the pay-off matrices to include loss-aversion (consistent with the Impulse Balance Equilibrium in Selten and Chmura 2008, p. 947). It was demonstrated by Brunner et. al (2010) that the explanatory power of the various models compared by Selten and Chmura (2008) improves significantly once loss aversion is included before computing the equilibrium solutions. We include loss aversion by subtracting the guaranteed payoff from each entry in the matrix, as described in Section I C and then multiply negative entries in the transformed matrix by a multiplier,  $\lambda$ . The conventional approach is to take  $\lambda = 2$  (Selten and Chmura 2008), but we will show

<b>Game 1</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(10,8)</td><td style="border: 1px solid black; padding: 2px;">(0,18)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(9,9)</td><td style="border: 1px solid black; padding: 2px;">(10,8)</td></tr> </table>	(10,8)	(0,18)	(9,9)	(10,8)	<b>Game 7</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(10,12)</td><td style="border: 1px solid black; padding: 2px;">(4,22)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(9,9)</td><td style="border: 1px solid black; padding: 2px;">(14,8)</td></tr> </table>	(10,12)	(4,22)	(9,9)	(14,8)
(10,8)	(0,18)										
(9,9)	(10,8)										
(10,12)	(4,22)										
(9,9)	(14,8)										
<b>Game 2</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(9,4)</td><td style="border: 1px solid black; padding: 2px;">(0,13)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(6,7)</td><td style="border: 1px solid black; padding: 2px;">(8,5)</td></tr> </table>	(9,4)	(0,13)	(6,7)	(8,5)	<b>Game 8</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(9,7)</td><td style="border: 1px solid black; padding: 2px;">(3,16)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(6,7)</td><td style="border: 1px solid black; padding: 2px;">(11,5)</td></tr> </table>	(9,7)	(3,16)	(6,7)	(11,5)
(9,4)	(0,13)										
(6,7)	(8,5)										
(9,7)	(3,16)										
(6,7)	(11,5)										
<b>Game 3</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(8,6)</td><td style="border: 1px solid black; padding: 2px;">(0,14)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(7,7)</td><td style="border: 1px solid black; padding: 2px;">(10,4)</td></tr> </table>	(8,6)	(0,14)	(7,7)	(10,4)	<b>Game 9</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(8,9)</td><td style="border: 1px solid black; padding: 2px;">(3,17)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(7,7)</td><td style="border: 1px solid black; padding: 2px;">(13,4)</td></tr> </table>	(8,9)	(3,17)	(7,7)	(13,4)
(8,6)	(0,14)										
(7,7)	(10,4)										
(8,9)	(3,17)										
(7,7)	(13,4)										
<b>Game 4</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(7,4)</td><td style="border: 1px solid black; padding: 2px;">(0,11)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(5,6)</td><td style="border: 1px solid black; padding: 2px;">(9,2)</td></tr> </table>	(7,4)	(0,11)	(5,6)	(9,2)	<b>Game 10</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(7,6)</td><td style="border: 1px solid black; padding: 2px;">(2,13)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(5,6)</td><td style="border: 1px solid black; padding: 2px;">(11,2)</td></tr> </table>	(7,6)	(2,13)	(5,6)	(11,2)
(7,4)	(0,11)										
(5,6)	(9,2)										
(7,6)	(2,13)										
(5,6)	(11,2)										
<b>Game 5</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(7,2)</td><td style="border: 1px solid black; padding: 2px;">(0,9)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(4,5)</td><td style="border: 1px solid black; padding: 2px;">(8,1)</td></tr> </table>	(7,2)	(0,9)	(4,5)	(8,1)	<b>Game 11</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(7,4)</td><td style="border: 1px solid black; padding: 2px;">(2,11)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(4,5)</td><td style="border: 1px solid black; padding: 2px;">(10,1)</td></tr> </table>	(7,4)	(2,11)	(4,5)	(10,1)
(7,2)	(0,9)										
(4,5)	(8,1)										
(7,4)	(2,11)										
(4,5)	(10,1)										
<b>Game 6</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(7,1)</td><td style="border: 1px solid black; padding: 2px;">(1,7)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(3,5)</td><td style="border: 1px solid black; padding: 2px;">(8,0)</td></tr> </table>	(7,1)	(1,7)	(3,5)	(8,0)	<b>Game 12</b>	<table style="border-collapse: collapse; margin: auto;"> <tr><td style="border: 1px solid black; padding: 2px;">(7,3)</td><td style="border: 1px solid black; padding: 2px;">(3,9)</td></tr> <tr><td style="border: 1px solid black; padding: 2px;">(3,5)</td><td style="border: 1px solid black; padding: 2px;">(10,0)</td></tr> </table>	(7,3)	(3,9)	(3,5)	(10,0)
(7,1)	(1,7)										
(3,5)	(8,0)										
(7,3)	(3,9)										
(3,5)	(10,0)										

FIG. 3: Pay-offs for the 12 games in Selten and Chmura (2008). In each cell an entry  $(x, y)$  indicates a payoff of  $x$  for the row player and  $y$  for the column player.

results for both  $\lambda = 1$  and  $\lambda = 2$  in the following.

After carrying out this transformation, some of the initially positive pay-offs become negative. Our previous results no longer guarantee the existence of the RBB equilibrium solution under those circumstances. However, it is possible to prove for  $2 \times 2$ , constant-sum games, that an RBB equilibrium exists for any choice of the loss multiplier used when transforming the matrices to include loss-aversion. This proof is given in Appendix IV. In the non-constant sum case it is more difficult to prove generic existence of the RBB solutions, although such solutions do exist for all the games considered by Selten and Chmura (2008). Two of the games, numbers 8 and 10, have a singular behavior that arises from the equality

of pay-offs for some moves. The derivation of the RBB solution is more subtle in those cases and is described in Appendix V.

## B. Ambiguity in beliefs

In this section we will explain the empirical observations of the completely mixed  $2 \times 2$  games using the RBB model. The RBB model incorporates ambiguity using a hyperprior on the belief probability, as described in equation (3). The experimental data can be used to determine the properties of this hyperprior, which we characterize in terms of the  $X_i$  parameters. These are determined from the belief probabilities held by the players, which we must compute from the observational data. Under the Luce (1959) rule, the probability of move  $i$  for player 1 is just  $p_i \propto u_{ij}^1 q_j$ , where  $q_j$  is the player's belief about player 2's strategy. In the RBB equilibrium,  $q_j$  is determined from  $p_i$  via a similar equation. However, if player 1 is observed to play the strategy  $p_i^{\text{obs}}$ , then this equation can be inverted to give  $q_j^{\text{b}} \propto (u^1)_{jk}^{-1} p_k^{\text{obs}}$ , which is the belief that player 1 must have had about the strategy of player 2 in order to lead them to play as they did. For any results observed in an experiment, the belief probabilities  $q_j^{\text{b}}$  that would lead to the observed behavior,  $p_i^{\text{obs}}$  can be computed in this way. If the person is playing a pure RBB strategy, then  $q_j^{\text{b}} = q_j^{\text{rbb}}$ , but the ambiguity characterized by the hyperprior would lead to some deviation.

For two-move games, there is only one probability, which we denote  $q = q_1$ , as  $q_2 = 1 - q$ . In Figure 4 we show  $q^{\text{b}}$  versus  $q^{\text{rbb}}$  and  $p^{\text{b}}$  versus  $p^{\text{rbb}}$  for all of the data in the Selten and Chmura games. The two plots correspond to transformed games with  $\lambda = 1$  and  $\lambda = 2$  respectively. We see that there is a tight correlation between the beliefs that players hold and the beliefs predicted in the RBB equilibrium, which reinforces the idea that the RBB is a good model for describing choice behavior in games. However, there are some deviations as we expected.

As described above, deviations will arise because the player is ambiguous as to the cooperative stance of the other player and therefore the likelihood that their opponent will play the RBB solution. In order to incorporate this ambiguity we examine the RBB model with hyperpriors. As mentioned earlier, the belief distribution,  $P(\mathbf{q})$ , is a *hyperprior* as it is a distribution for the hyperparameters,  $\{q_j\}$ , that characterize player 1's prior belief,  $\mathbf{q}$ , for the action of player 2. A particular choice of hyperprior will lead to particular  $\{q_i^{\text{b}}\}$ 's, but

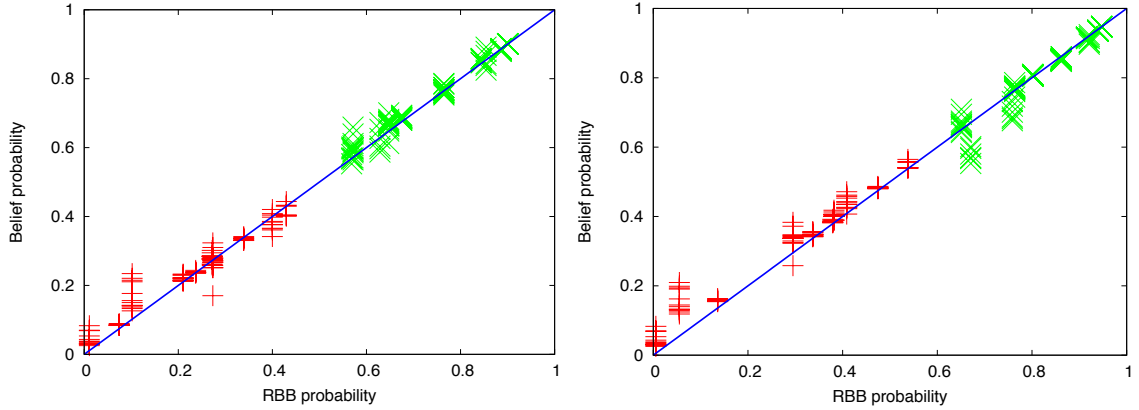


FIG. 4: Belief probability (y axis) versus RBB equilibrium probability (x axis) for all data in Selten and Chmura. The red pluses are values for  $p$ , the probability for the row player, while the green crosses are for  $q$ , the probability of the column player. The line indicates  $q^b = q^{\text{rbb}}$  to guide the eye. The left panel uses a multiplier  $\lambda = 1$  to transform the game, while the right panel uses  $\lambda = 2$ .

the observational consequence of the hyperprior is entirely encoded in the  $X_i$  parameters. These represent the fractional change from the RBB equilibrium solution. The deviation could also be characterized by the values of the observed belief probabilities,  $q_i^b$ , directly, but using the  $X_i$  parameters allows us to factorize the belief into a game-dependent part,  $\mathbf{q}^{\text{rbb}}$ , and a part that is a property of the players in the game,  $\{X_i\}$ . This is not the only way in which such a factorization could be accomplished, but it provides a good description of the observed data and it has a simple interpretation as a fractional modification of the belief probability. The  $\{X_i\}$ 's can be computed for any observed data, but to test whether they are indeed game independent for a particular set of players would require new experiments<sup>14</sup>.

In a given data set, the probability constraint  $\sum q_i^b = \sum q_i^{\text{rbb}} = 1$  means that the  $X_i$ 's are not all independent. Alternatively, we can define  $r_i = X_i/X_N$  for  $i = 1, \dots, N - 1$  and then the  $r_i$ 's can independently take any values in  $[0, \infty)$ . In the case of the Selten and Chmura data, each game has only two moves and therefore one independent probability and one  $r_i$ ,  $r_1 = X_1/X_2$ . This quantity can be computed from the observed probability  $q^b$  and the RBB

<sup>14</sup> This is a common problem to any analysis of the belief distribution, since it is reasonable to suppose that a player might formulate his beliefs about his opponent in a way that depends on the properties of the game.

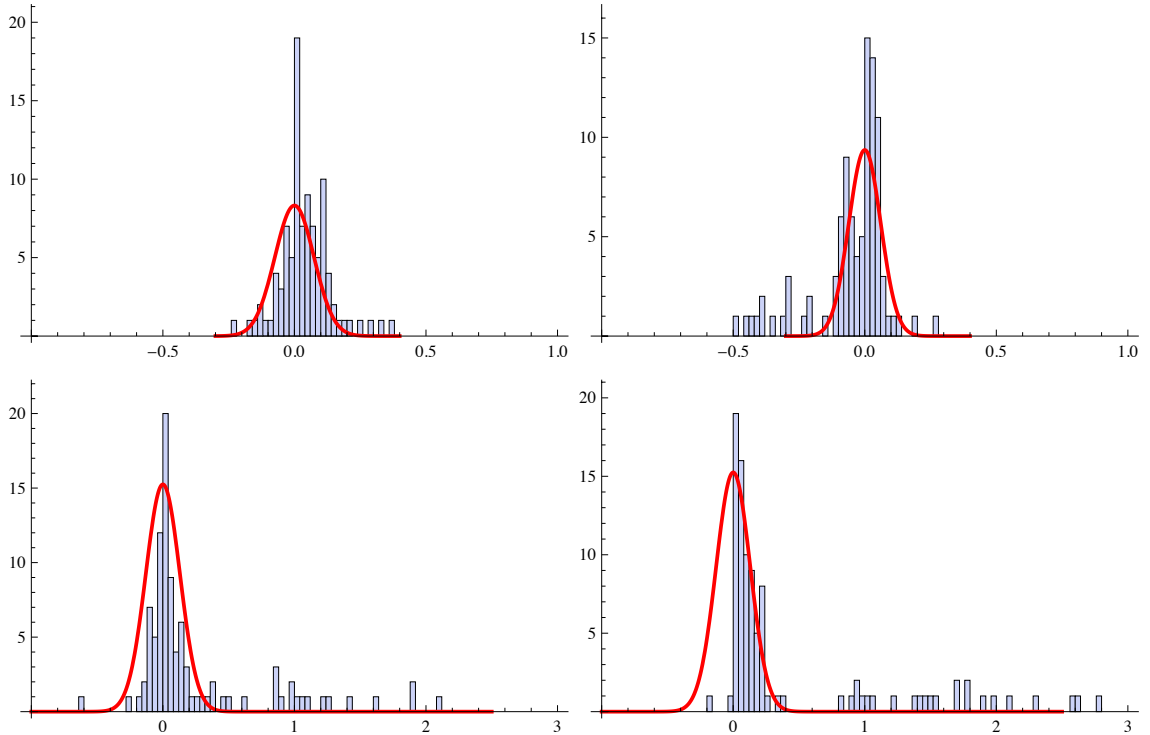


FIG. 5: Distribution of the parameter  $r_1$  for player 1's belief probability,  $q^b$ , (upper panels) and player 2's belief probability,  $p^b$ , (lower panels). The solid lines are the best-fit zero-mean log-normal distributions to the data. The lefthand panels use a multiplier  $\lambda = 1$  to transform the games, while the righthand panels use  $\lambda = 2$ .

equilibrium probability  $q^{\text{rbb}}$  as

$$r_1 = \frac{q^b(1 - q^{\text{rbb}})}{q^{\text{rbb}}(1 - q^b)}. \quad (9)$$

The  $r_1$  parameter encodes information about the players playing the game. We want to characterize the distribution of values this parameter can take using the experimental data. The two moves for player 1 are treated asymmetrically in  $r_1$ , as we have chosen move 1 to be in the numerator. As the moves could be ordered arbitrarily, using  $r_2 = X_2/X_1 = 1/r_1$  is equally valid. If  $r_1$  is an intrinsic property of people, we would therefore expect its distribution to be inversion symmetric, i.e., the distribution of  $1/r_1$  should resemble that of  $r_1$ . A simple distribution that has this property is a log-normal distribution with zero mean. In Figure 5 we show the distribution of  $r_1$  for the Selten and Chmura data, along with the best-fit zero-mean log-normal distributions. The left and right hand panels are for loss aversion multipliers  $\lambda = 1$  and  $\lambda = 2$  respectively.

The fits in Figure 5 reproduce the data quite well, which lends support to the picture we



have developed, namely that the players are formulating their beliefs according to the RBB equilibrium strategy, but with fractional modifications that arise from the player’s beliefs about the cooperability of their opponent. Characterized in this way, the beliefs seem to be consistent with being drawn from a zero-mean log-normal distribution. There are a few things to note from the figure. Firstly, we have treated player 1 and player 2 independently by analyzing the  $q^b$  and  $p^b$  results separately. The games considered do not treat player 1 and player 2 equivalently in terms of the payoffs. Nonetheless, the widths of the best-fit distributions are quite similar in both cases, which supports the notion that the  $X_i$ ’s are an intrinsic characteristic of the players rather than the game. Secondly, we note that there are some outliers in the distribution for  $p^b$ . These all come from the separate replicates of game 1. Game 1 has a very small value for  $p^{\text{rbb}}$  — the row player is unlikely to play the upper move due to the potential penalty involved. This has two consequences — (i) the statistical errors in measurements of this number are proportionally larger; (ii) statistical errors are magnified, resulting in larger errors in the  $X_i$ ’s. For this reason, we treated these points as outliers when deriving the best-fit distributions. Finally, we see that the fits seem to be somewhat better for  $\lambda = 1$  than for  $\lambda = 2$ . The choice of this parameter is to some extent arbitrary and  $\lambda = 2$  is chosen by convention, but this result might indicate that this set of players were not particularly loss averse.

### C. Comparison to other equilibrium concepts

In Selten and Chmura (2008), the observations described here were compared to the predictions of several equilibrium concepts — the Nash equilibrium, the QRE, the Action Sampling equilibrium, the Payoff Sampling equilibrium and the Impulse Balance equilibrium. In the original analysis, it was found that the impulse balance equilibrium was a much better predictor of behavior. However, the impulse balance approach naturally incorporates loss-aversion and it was subsequently demonstrated by Brunner et al. (2010) that if the analysis was repeated on the transformed game including loss aversion with  $\lambda = 2$ , the equilibrium concepts other than Nash are able to explain the data as well as the impulse balance results. To compare the RBB equilibrium predictions to those of these other concepts, it is natural to use the same deviation parameter,  $r_i$ , introduced above. The alternative equilibrium concepts are not formulated in terms of beliefs about a player’s opponent so in this context

the  $X_i$  parameters are most readily interpreted as the fractional errors made by the player in the computation of the equilibrium solution. The RBB solution represents an equilibrium in beliefs and therefore we compare the belief probabilities to the RBB solution, rather than the move probabilities. This approach is natural in the RBB model because of the assumptions made about the players' decision process, but the other equilibrium concepts are not formulated in terms of beliefs about the opponent <sup>15</sup>. For the other models, we must therefore compare the move probabilities for each player to the predictions of the equilibrium concept by defining

$$r_1 = \frac{q^{\text{obs}}(1 - q^{\text{pred}})}{q^{\text{pred}}(1 - q^{\text{obs}})} \quad (10)$$

where  $q^{\text{obs}}$  is the observed probability that player 2 plays move left and  $q^{\text{pred}}$  is the corresponding probability predicted by the particular equilibrium concept under consideration.

In Figure 6 we show the distribution of the logarithm of this parameter,  $r_1$ , for each of the equilibrium concepts listed above. The bottom right panel shows the corresponding distribution for the RBB model for comparison. We used a loss aversion multiplier  $\lambda = 2$  for consistency with Brunner et al. (2010). The shape of these distributions is broadly consistent with a normal distribution in all cases so the treatment of the deviations from the model described here could be applied to any of the other equilibrium concepts. The width of the distribution reflects the amount of departure from the equilibrium solution seen in the data. We see that the Nash prediction is generally worse than the other predictions, which is consistent with previous results in Selten and Chmura (2008) and Brunner et al. (2010). It is also clear that the RBB solution is generally better than the other concepts, showing a much narrower distribution in the deviation parameter.

For a more direct comparison to Selten and Chmura (2008) and Brunner et al. (2010) we can compare goodness of fits for these various equilibrium concepts. We first compute the squared error in the equilibrium prediction for each data point, i.e.,  $(p_X - p_{\text{obs}})^2 + (q_X - q_{\text{obs}})^2$ , where  $(p_X, q_X)$  are the equilibrium model predictions for “up” and “left”. For the RBB model we compare the belief probabilities to the equilibrium predictions, as above, while for the other concepts we compare the move probabilities to the predictions. For consistency with

---

<sup>15</sup> Note that it is possible to reformulate some of the other equilibrium concepts, for instance the QRE and Nash solutions, in terms of a hyperprior on the belief probabilities of one player about the other. Such models are worth further investigation.

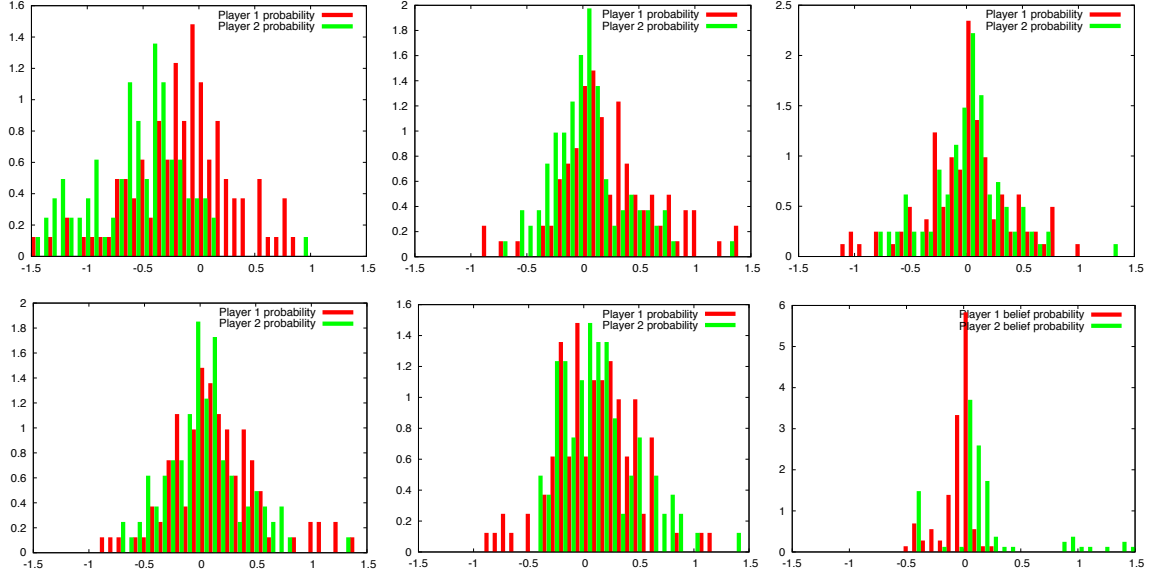


FIG. 6: Distribution of the error parameter,  $\log(r_1)$ , for the Selten and Chmura (2008) data when compared to the predictions of various equilibrium concepts. The top row of panels show, from left to right, the Nash equilibrium, the QRE and the Action Sampling equilibrium. The bottom row of panels show, from left to right, the Payoff Sampling equilibrium, the Impulse Balance equilibrium and, for comparison, the RBB equilibrium described in this paper. We note that in each bin of the histogram the bars for player 1 and player 2 have been deliberately offset from one another for clarity, but refer to the same range of values in the bin.

Brunner et al. (2010) we compute the QRE with  $\lambda = 1.05$ , the action sampling equilibrium with  $n = 12$  and the payoff-sampling equilibrium with  $n = 6$ . In Table II we show p-values comparing these mean squared errors between the models using the Wilcoxon signed-rank matched-pairs test. Where a significant difference was found, we also did a one-sided test on the same data to see which model predicted significantly smaller errors. The results of these one-sided tests are also shown in the table. For these tests we use only data from Games 1-6, 7, 9, 11 and 12 due to the issue in computing the RBB solution in Games 8 and 10. These p-values strikingly reinforce the conclusion we drew from the error distributions — the RBB solution is significantly better at explaining the data than the other equilibrium concepts. We used the Wilcoxon test to allow direct comparison to Selten and Chmura (2008) and Brunner et al. (2010). Brunner et al. (2010) point out correctly that the Wilcoxon test is a parametric test, and the underlying distributional assumptions may not apply to the Selten and Chmura data. Brunner et al. (2010) also presented results based on the non-parametric

	Payoff sampling	QRE	Impulse balance	RBB	Nash
Action	n.s.	n.s.	n.s.	0.1% (g: 0.1%)	0.1% (l: 0.1%)
Sampling	n.s.	10% (l: 5%)	5% (l: 5%)	0.1% (g: 0.1%)	0.1% (l: 0.1%)
( $n = 12$ )	n.s.	n.s.	n.s.	5% (g: 2%)	1% (l: 0.5%)
Payoff		n.s.	n.s.	0.1% (g: 0.1%)	0.1% (l: 0.1%)
Sampling		0.1% (l: 0.1%)	0.1% (l: 0.1%)	0.1% (g: 0.1%)	0.1% (l: 0.1%)
( $n = 6$ )		0.5% (g: 0.5%)	2% (g: 1%)	5% (g: 5%)	2% (l: 1%)
QRE			10% (l: 5%)	0.1% (g: 0.1%)	0.1% (l: 0.1%)
( $\lambda = 1.05$ )			0.2% (l: 0.1%)	0.1% (g: 0.1%)	0.1% (l: 0.1%)
			10% (g: 5%)	10% (g: 5%)	2% (l: 1%)
Impulse				0.1% (g: 0.1%)	0.1% (l: 0.1%)
Balance				0.1% (g: 0.1%)	0.1% (l: 0.1%)
				n.s. (g: 10%)	1% (l: 0.5%)
RBB					0.1% (l: 0.1%)
					0.1% (l: 0.1%)
					0.2% (l: 0.1%)

TABLE II: P-values in favour of row concepts, using two-tailed matched-pairs Wilcoxon signed rank test (rounded to the next higher level among 0.1%, 0.2%, 0.5%, 1%, 2%, 5% and 10%). Where results are significant for the two-tailed test, we quote the significance of the one-sided test in brackets, with “l” indicating the mean of the error distribution for the row value is significantly lower than that of the the column value, and “g” indicating it is significantly greater. In each entry of the table, the first row uses data from Games 1-6, 7, 9, 11 and 12; the middle row uses data from Games 1-6 only; the last row uses data from Games 7, 9, 11 and 12 only.

Kolmogorov-Smirnov<sup>16</sup> and robust rank-order tests. We also performed those statistical tests and the results were consistent with the conclusions from the Wilcoxon test.

In summary, we have demonstrated in this section how the hyperprior concept provides a natural way to characterize departures from the equilibrium solution observed in exper-

<sup>16</sup> We note that the Kolmogorov-Smirnov test is also not entirely appropriate here, as it is not a matched-pairs test. However, it is non-parametric and does tell us whether the distributions of mean-squared errors are significantly different between the equilibrium concepts.

iments. Our framework attempts to split up the deviation into pieces that are player dependent (the  $X_i$ 's) and pieces that are game dependent. We have found that the player-dependent pieces are well fit by a log-normal distribution and that this kind of model can explain experimental data very closely. This type of analysis can be readily adapted to other equilibrium concepts and games. For data in Selten and Chmura (2008), the RBB model appears to explain the data significantly better than other equilibrium concepts when the comparison is done in terms of belief probabilities, which is natural in the RBB framework.

### III. CONCLUSION

Game theorists have long been interested in the empirical testing of game-theoretic anomalies. A sizeable experimental literature shows that there is a dissonance between the theory and empirics of games in a strategic setting where defection is predicted by theory to be the optimal outcome over cooperation. Players in experimental settings cooperate far more than the theory would predict. For example, 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.

In this paper we provide a reason for this unreason and explain this anomaly by providing a new method termed the 'returns-based beliefs' approach in which players form subjective beliefs and then act based upon the expected returns, given these beliefs, of a particular strategy, in proportion to the total expected returns of all strategies. Our approach treats beliefs rather than strategies as the primary concept. In so doing, we assume that the players' beliefs are in equilibrium. Our approach combines a decision analytic solution concept where individuals form subjective probabilities over the actions of the individual's opponent and then choose a mixed strategy profile over the actions, using the probabilistic choice model developed axiomatically by Luce (1959), that is based on the relative returns of each strategy given these subjective probabilities. We hypothesize that when agents form subjective beliefs about the strategies of the other players then there might be opportunities for profitable deviation from the Nash equilibrium strategies when the other player is expected to deviate from playing the Nash strategy. We incorporate the concept of hyperpriors to account for the ambiguity of the cooperative stance of the players.

We test our model on various classes of games: for example, we show how our approach provides a closer description of empirical observations in both the Prisoner's Dilemma and the Traveler's Dilemma games. Our approach explains 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. We also show that this method can be applied to the Traveler's Dilemma; we show that the strategies that players will choose are inversely related to the size of the reward or penalty, and accord very closely with the empirical findings. This provides some initial support that our model delivers predictions in games commonly used in game theory, well beyond the Prisoner's Dilemma.

Finally, we test the closeness of fit of our model using data from Selten and Chmura (2008) for completely mixed  $2 \times 2$  games. In particular, we show that the returns-based belief model is able to explain the empirical data from completely mixed  $2 \times 2$  games once loss aversion is included, with small deviations about the equilibrium solution. We show that the RBB belief equilibrium is generally better at explaining the data than the other stationary concepts discussed in Selten and Chmura (2008), showing a much narrower distribution in the deviation parameter between the actual and the predicted values and a significantly smaller mean-square error.

We argue that returns-based beliefs are useful in game theory because they allow us to reconcile game theoretic predictions and empirical experimental results. We believe that if returns-based belief formation is a possible explanation for the experimental observations in these games, then this may amplify the possibility of testing other observable anomalies in game theory. The sensitivity of the equilibrium to changes in payoff is an important aspect of our model that accords more with reality than the conventional Nash equilibrium. Moreover, even if returns-based beliefs do provide predictions in some situations, there may be other avenues for modeling non-equilibrium behavior, which when combined with our model may provide additional insights. Therefore, we hope that this study provides a helpful prelude for better understanding the impulses which govern human behavior.

## References

- Abelson, R M, and RA Bradley. 1954. "A2 X A2 factorial with paired comparisons." *Biometrics*, 10(4): 487-502.
- Andreoni, J A, and JH Miller. 1993. "Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence." *Economic Journal*, 103(418): 570-585.
- Aumann, Robert, and Adam Brandenburger. 1995. "Epistemic Conditions for Nash Equilibrium." *Econometrica*, 63(5). 1161-1180.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*, Basic Books.
- Basu, Kaushik. 1990. "On the non-existence of a rationality definition for extensive games." *International Journal of Game Theory*, 19: 33-44.
- Basu, Kaushik. 1994. "The Traveler's Dilemma: Paradoxes of Rationality in Game Theory". *American Economic Review (Papers and Proceedings)*, 84(2) 391-395.
- Binmore, Kenneth. 2009. *Rational Decisions*. Princeton, NJ: Princeton University Press.
- Bono, James, and David Wolpert. 2009. "How to use decision theory to choose among mechanisms", Paper presented at *From Game Theory to Game Engineering Workshop*, Oxford Man Institute, 17 and 18 September.
- Brandenberger, Adam. 1996. "Strategic and structural uncertainty in games." In *Wise Choices: Decisions, Games and Negotiations*, eds RJ Zeckhauser, RL Keeney, JK Sebenius. Boston: Harvard Business School Press.
- Brandenburger, Adam, and Eddie Dekel. 1987. "Rationalizability and correlated equilibrium." *Econometrica*, 55: 1391-1402.
- Brunner, C., Colin H. Camerer and Jacob K. Goeree. 2010. "A correction and re-examination of 'Stationary concepts for experimental 2 games'." *American Economic Review*, forthcoming.
- Camerer, Colin F, T-H Ho and J.K. Chong. 2004. "A cognitive hierarchy model of games." *Quarterly Journal of Economics*, 119(3): 861-898.
- Capra, M C., J K. Goeree., R. Gomez., C. H. Holt. 1999 "Anomalous behavior in a traveler's dilemma?" *American Economic Review*, 89(3): 678-690.
- Chen, Hsiao-Chi, W. James Friedman, Jacques-Francois Thisse. 1997. "Bounded rational Nash equilibrium: A probabilistic choice approach" *Games and Economic Behavior*, 18: 32-54.

- Costa-Gomes, M A and Vincent P Crawford. 2006. "Cognition and behavior in two person guessing games: An experimental study." *American Economic Review*, 96(5): 1737-1768.
- DeGroot. M.H. 1975. *Probability and Statistics*. Reading, Mass: Addison-Wesley.
- Ellsberg, Daniel. 1959. "Recurrent objections to the minimax strategy: Rejoinder", *The Review of Economics and Statistics*, 41(1): 42-43.
- Farrell, J and Matthew Rabin. 1996. "Cheap talk." *Journal of Economic Perspectives*, 10(3): 103-118.
- Fehr, Ernst, and S Gächter. 2002. "Altruistic punishment in humans." *Nature*, 415: 137-140.
- Goeree, Jacob K, and Charles A. Holt. 2001 "Ten little treasures of game theory and ten intuitive contradictions." *American Economic Review*, 91(5): 1402-1422.
- 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.
- Haile, P A, A Hortacsu, G Kosenok. 2008. "On the empirical content of quantal response equilibrium." *American Economic Review*, 98(1): 180-200.
- Harsanyi, J C. 1973. "Games with randomly distributed payoffs: a new rationale for mixed strategy equilibrium points." *International Journal of Game Theory*, 2: 1-23
- Harsanyi, J C. 1982. "Subjective probability and the theory of games: Comments on Kadane and Larkey's paper." *Management Science*, 28(2):120-124.
- Holt, Charles A and Al E. Roth. 2004. "The Nash Equilibrium: A Perspective." *Proceedings of the National Academy of Sciences*, 23 March, 101(12): 3999-4002.
- Kadane, J B, P D Larky. 1982. "Subjective probability and the theory of games." *Management Science*, 28(2):113-120.
- Luce, D R. 1959. *Individual Choice Behavior*. New. New York: Wiley.
- Luce, D R and H. Raiffa. 1957. *Games and Decisions*. New York: Wiley.
- McKelvey, R D and RP Palfrey. 1995. "Quantal response equilibria for normal form games." *Games and Economic Behavior*, 10(1): 6-38.
- Morris, Stephen. 2008 "Purification". *The New Plagrove Dictionary of Economics. Second Edition*. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.
- Myerson, R. 1991. *Game Theory: Analysis of Conflict*. Harvard University Press, MA



- Nash, John F. 1951. “Non-cooperative games.” *Annals of Mathematics*, 54(2): 286-95.
- Nau, R.F. and K.F. McCardle. 1990. “Coherent Behavior in Noncooperative Games.” *Journal of Economic Theory*, 50: 424-444.
- Rabin, Matthew. 1993. “Incorporating Fairness into Game Theory and Economics.” *American Economic Review*, 83(5): 1281-1302.
- Rios Insua, D., J. Rios and David Banks. 2009. “Adversarial risk analysis.” *Journal of the American Statistical Association*, 104 (486): 841–854.
- Rosenthal, R W. 1981. “Games of perfect information, predatory pricing and the chain-store paradox.” *Journal of Economic Theory*, 25(1): 92-100.
- Roth, Al and F. Schoumaker. 1983. “Subjective probability and the theory of games.” *Management Science*, 29(11): 1337-1340.
- Rubinstein, Ariel. 2007. “Instinctive and cognitive reasoning: Response times study.” *Economic Journal*, 117:1243-1259.
- Savage, L.J. 1954. *The Foundations of Statistics*. New York: Wiley.
- Sally, D. 1995. “Conversation and cooperation in Social Dilemmas.” *Rationality and Society*, 7(1): 58-92.
- Selten, Reinhart. 1978. “The chain-store paradox.” *Theory and Decision*, 9(2): 127-59.
- Selten, Reinhart and Thorsten, Chmura. 2008. Stationary concepts for experimental 2x2-Games. *American Economic Review*. 98(3). 938-966.
- Simon, Herbert A. 1987. “Making Management Decisions: the Role of Intuition and Emotion.” *Academy of Management Executive*, February: 57-64.
- Van Huyck J B, RC Battalio, RO Beil. 1990. “Tacit coordination games, strategic uncertainty, and coordination failure.” *American Economic Review*, 80(1): 234-248.

## APPENDIX A: APPENDIX I: RETURNS BASED BELIEF EQUILIBRIUM

Consider a game played between two players, and use  $u_{ij}^1, u_{ij}^2$  to denote the payoffs to player 1 and 2 respectively when player 1 plays move 1 and player 2 plays move 2. Player 1 plays an RBB strategy if, when player 2 plays a strategy  $\mathbf{q} = \{q_j\}$ , player 1 adopts a strategy  $\mathbf{p} = \{p_i\}$  in which

$$p_i = \frac{u_{ij}^1 q_j}{\sum_k u_{kj}^1 q_j} \quad (\text{A1})$$

and similarly for player 2. In an RBB equilibrium we assume both player 1 and player 2 are playing the RBB strategy and therefore

$$p_i = \frac{1}{\sum_l u_{lj}^1 u_{kj}^2 p_k} u_{ij}^1 u_{kj}^2 p_k \quad (\text{A2})$$

An RBB equilibrium is a fixed point of this mapping. If we denote the first fraction as  $1/\lambda$  we see that, in matrix notation,  $\mathbf{u}^1 \mathbf{u}^{2T} \mathbf{p} = \lambda \mathbf{p}$ , i.e., the RBB equilibrium is given by eigenstates of the matrix  $\mathbf{u}^1 \mathbf{u}^{2T}$ .

In principle, this matrix will have  $m$  eigenstates, but we must impose the requirement that  $\mathbf{p}$  and  $\mathbf{q}$  are *probability vectors*<sup>17</sup>. Therefore we require  $\lambda \geq 0$  and that both  $\mathbf{p}$  and  $\mathbf{u}^{2T} \mathbf{p}$  are probability vectors. If the pay-offs are non-negative to both players from all moves, then the mapping

$$\mathbf{p} \rightarrow \frac{\mathbf{u}^1 \mathbf{u}^{2T} \mathbf{p}}{\mathbf{e}^T \mathbf{u}^1 \mathbf{u}^{2T} \mathbf{p}} \quad (\text{A3})$$

with  $\mathbf{e} = (1, 1, \dots, 1)$  defines a mapping from probability vectors to probability vectors and so must have a fixed point, from the fixed point theorem. This is the essential basis of the proof given in Chen et al. (1997) in a different setting that we mentioned earlier.

If some of the pay-offs are negative, then there is no guarantee that one of the eigenvectors will lie in the probability vector quadrant. This arises, for example, when transforming the pay-off matrices to account for loss aversion in the analysis of the Selten and Chmura (2008) games described in Section II. In that context it is possible to prove that an RBB equilibrium exists for any constant sum game (see Appendix IV) and for the particular non-constant-sum games given in Selten and Chmura (2008), RBB equilibria are found to exist when using the usual loss aversion multipliers.

In a more general context, if no RBB equilibrium exists as defined by the above definition, an RBB solution can be found by zeroing any negative probabilities encountered. In other words, when player 2 plays a strategy  $q_j$ , we define

$$v_i = \begin{cases} u_{ij}^1 q_j & \text{if } u_{ij}^1 q_j \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (\text{A4})$$

and then say that player 1 plays an RBB strategy if he adopts a strategy  $p_i$  in which

$$p_i = \frac{v_i}{\sum_k v_k}. \quad (\text{A5})$$

---

<sup>17</sup> We define a *probability vector* to be an  $m$ -tuple,  $\{p_1, \dots, p_m\}$  with  $p_i \geq 0 \forall i$  and  $\sum p_i = 1$ .

Under this definition, the fixed point theorem will guarantee the existence of an RBB equilibrium in all games. We will refer to this algorithm as the *probability-zeroing algorithm* elsewhere.

The concept of an RBB equilibrium can be extended naturally to games between more than two players and the existence of an equilibrium when pay-offs are positive or when using the zeroing algorithm carries over again from the fixed point theorem. However, the interpretation of the RBB equilibrium as an eigenstate of a matrix is lost when there are more than two players.

## APPENDIX B: APPENDIX II: TEMPTATION VERSUS PENALTY IN THE PD

We consider a general PD game with payoffs

$$\begin{array}{rcc}
 & \text{Cooperate} & \text{Defect} \\
 \text{Cooperate} & (C, C) & (A, B) \\
 \text{Defect} & (B, A) & (D, D)
 \end{array} \tag{B1}$$

The probability of cooperation in the RBB equilibrium solution for this game is

$$q_C = \frac{Cq_C + A(1 - q_C)}{(C + B - A - D)q_C + A + D} \tag{B2}$$

which may be rewritten in the form

$$f(q_C, A, B, C, D) = (C + B - A - D)q_C^2 + (2A + D - C)q_C - A = 0 \tag{B3}$$

We want to compare the effect on the cooperative probability,  $q_C$ , of a one unit increase in the cooperative payoff,  $C$ , versus the effect of a one unit decrease in the mixed payoff,  $B$ . This can be achieved by computing the ratio

$$-\frac{\partial q_C / \partial C}{\partial q_C / \partial B} \equiv -\frac{\partial f / \partial C}{\partial f / \partial B} = \frac{(1 - q_C)}{q_C} \tag{B4}$$

where the minus sign is introduced since we compare an increase in  $C$  to a decrease in  $B$ , and the second expression follows from differentiation of equation (B3). We see that the effect of an increase in  $C$  is greater than the effect of a comparable decrease in  $B$ , provided  $q_C < 0.5$ .

Writing  $X = C + B - A - D$  and  $Y = A + B$ , the solution of equation (B3) can be found to be

$$q_C = \frac{1}{2} - \frac{1}{2X} \left( Y - \sqrt{Y^2 - X[B - C + D - A]} \right) \quad (\text{B5})$$

To be a PD game, we require  $A < D < C < B$ , so  $X > 0$  and  $B - C + D - A > 0$ , which means the term in the square root is less than  $Y^2$ , so the term in round brackets is positive and we deduce that  $q_C < 1/2$ . Thus, the effect on the probability of cooperation of a unit increase in  $C$  is always greater than the effect of a unit decrease in  $B$ .

### APPENDIX C: APPENDIX III: PD WITH LOSS AVERSION

We set up the Prisoner's Dilemma such that the payoff from cooperation is  $X$ , from defection is  $Y$ , the reward from defecting while the opponent cooperates is  $R$  and the penalty for cooperating when the opponent defects is  $P$ . The auxiliary conditions are that  $R > 0$ ,  $P > 0$  and  $X > Y$ . The guaranteed payoff is  $Y$ , so we include loss aversion by subtracting this value from the entries in the payoff matrix and multiplying any negative entries by a multiplier,  $\lambda$ . The payoff matrix before and after transformation is shown below, denoting cooperate by "C" and defect by "D". Note that we treat the players symmetrically, i.e., with the same loss aversion  $\lambda$ .

	C	D		C	D
C	$X, X$	$Y - P, X + R$	C	$X - Y, X - Y$	$-\lambda P, X - Y + R$
D	$X + R, Y - P$	$Y, Y$	D	$X - Y + R, -\lambda P$	$0, 0$

Original payoff matrix

Payoff matrix including loss aversion

The RBB solution is given by eigenvalues,  $\mu$ , of the matrix

$$\begin{pmatrix} X - Y & -\lambda P \\ X - Y + R & 0 \end{pmatrix} \quad (\text{C1})$$

with the requirement that  $\mu > 0$  and the components of the eigenvector are strictly positive.

The eigenvalues are given by solutions of the equation

$$\mu((X - Y) - \mu) = \lambda P(X - Y + R) \quad (\text{C2})$$

This has positive solutions in the range  $0 \leq \mu \leq X - Y$  provided that  $X - Y \geq 0$  and

$$\left(\frac{X - Y}{2}\right)^2 \geq \lambda P(X - Y + R) \quad \Rightarrow \quad \lambda \leq \frac{(X - Y)^2}{4P(X - Y + R)} \quad (\text{C3})$$

and the corresponding eigenvector is parallel to  $(\lambda P, X - Y - \mu)$  so it is in the positive quadrant and it is thus a valid RBB solution.

If either condition is violated, there is no RBB solution, but the probability-zeroing algorithm will lead to the (D, D) Nash equilibrium.

**APPENDIX D: APPENDIX IV: TWO MOVE, CONSTANT-SUM GAMES**

In the first six of the Selten games, the form of the matrix is greatly simplified since there are only two moves for each player, and the games are constant sum. We can interchange the moves such that the guaranteed payoff is in the top left hand corner of the payoff matrix for player 1. If the constant sum is  $T$  then the payoff matrices take the form

$$u^1 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad u^2 = \begin{pmatrix} T - a & T - b \\ T - c & T - d \end{pmatrix} \quad (\text{D1})$$

with  $a < b$  and either  $c < a < d$  or  $d < a < c$  since  $a$  is the guaranteed payoff. If  $c < d$  then the two elements in the first column of player 2's payoff matrix are bigger than the corresponding elements in the second column (due to the fact it is a constant sum game), and so the second column is a dominated strategy. In the discussion of the Prisoner's Dilemma, we argued that there could be an incentive to play a dominated strategy if the players would mutually benefit from such cooperation. However, when  $c < d$ , the dominance in these games is of a more extreme form such that  $\min(T - a, T - c) > \max(T - b, T - d)$  so the worst payoff to player 2 from choosing the left hand column is better than the best payoff to player 2 from the right hand column. It is clear that in such circumstances the second column would never be played and so we can restrict to games with no such extremely dominated strategies<sup>18</sup>

Therefore we need  $d < a < c$ , and there are two cases: (i)  $c < b$ ; (ii)  $c > b$ . The

---

<sup>18</sup> Formally if the payoff to player  $i$  from strategy  $s^i \in S^i$  when his opponents play  $s^{-i} \in S^{-i}$  is  $u(s^i, s^{-i})$  then  $p^i \in S^i$  dominates  $q^i \in S^i$  if  $u(p^i, s^{-i}) \geq u(q^i, s^{-i}) \quad \forall s^{-i} \in S^{-i}$ , while  $p^i \in S^i$  has this extreme dominance over  $q^i \in S^i$  if  $\min \{u(p^i, s^{-i}); s^{-i} \in S^{-i}\} \geq \max \{u(q^i, s^{-i}); s^{-i} \in S^{-i}\}$ .

transformed payoffs are then

$$u^1 = \begin{pmatrix} 0 & b-a \\ c-a & \lambda_1(d-a) \end{pmatrix}, u^2 = \begin{pmatrix} c-a & \lambda_2(c-b) \\ 0 & c-d \end{pmatrix} \quad (\text{D2})$$

Case(i)

$$u^1 = \begin{pmatrix} 0 & b-a \\ c-a & \lambda_1(d-a) \end{pmatrix}, u^2 = \begin{pmatrix} b-a & 0 \\ \lambda_2(b-c) & b-d \end{pmatrix} \quad (\text{D3})$$

Case (ii)

where  $\lambda_1, \lambda_2$  are the loss-aversion multipliers described above, where we allow these to differ for the two players.

It can be shown that, irrespective of any further assumptions about the matrix elements, or the values of  $\lambda_1$  and  $\lambda_2$ , that these transformed games have an RBB equilibrium. The RBB equilibrium solution for player 1 is an eigenvector of the matrix  $u^1(u^2)^T$ . In these cases, this matrix becomes

$$\begin{aligned} \text{Case (i):} & \begin{pmatrix} \lambda_2(c-b)(b-a) & (b-a)(c-d) \\ (c-a)^2 + \lambda_1\lambda_2(d-a)(c-b) & \lambda_1(d-a)(c-d) \end{pmatrix} \\ \text{Case (ii):} & \begin{pmatrix} 0 & (b-a)(b-d) \\ (c-a)(b-a) & \lambda_1(d-a)(b-d) + \lambda_2(c-a)(b-c) \end{pmatrix} \end{aligned}$$

and we can see that in case (i) there is an RBB eigenstate

$$\begin{aligned} \text{Eigenvalue: } \mu_{(i)} &= \frac{1}{2} \sqrt{(\lambda_1(d-a)(c-d) + \lambda_2(b-a)(c-b))^2 + 4(b-a)(c-d)(c-a)^2} \\ &\quad - \frac{1}{2} (\lambda_1(a-d)(c-d) + \lambda_2(b-a)(b-c)), \end{aligned}$$

$$\begin{aligned} \text{Eigenvector: } \mathbf{p} &\propto \begin{pmatrix} (b-a)(c-d) \\ \mu_{(i)} - \lambda_2(b-a)(c-b) \end{pmatrix} \\ \mathbf{q} &\propto \begin{pmatrix} (b-a)(c-a)(c-d) \\ \mu_{(i)}(c-d) \end{pmatrix} \end{aligned}$$

and in case (i) there is an RBB eigenstate

$$\begin{aligned} \text{Eigenvalue: } \mu_{(ii)} &= \frac{1}{2} \sqrt{(\lambda_1(d-a)(b-d) + \lambda_2(c-a)(b-c))^2 + 4(b-a)^2(b-d)(c-a)} \\ &\quad - \frac{1}{2} (\lambda_1(a-d)(b-d) + \lambda_2(c-a)(c-b)), \\ \text{Eigenvector: } \mathbf{p} &\propto \begin{pmatrix} (b-a)(b-d) \\ \mu_{(ii)} \end{pmatrix} \\ \mathbf{q} &\propto \begin{pmatrix} (b-a)^2(b-d) - \lambda_2\mu_{(ii)}(c-b) \\ (b-d)\mu_{(ii)} \end{pmatrix} \end{aligned}$$

In the above  $\mathbf{p}$  represents the eigenvalue of  $u^1(u^2)^T$  which is the RBB probability for player 1, while  $\mathbf{q} = (u^2)^T \mathbf{p}$  is the corresponding RBB probability for player 2. To be a valid RBB solution, the eigenvalues must be positive and both  $\mathbf{p}$  and  $\mathbf{q}$  must lie in the positive quadrant. The preceding solutions satisfy these requirements for any choice of  $\lambda_1$  and  $\lambda_2$ . The only case for which this is not obvious in the above expressions is  $\mathbf{q}$  in case (ii), but it can be verified by expanding  $\mu_{(ii)}$  for  $\lambda_2 \rightarrow \infty$ .

## APPENDIX E: APPENDIX V: RBB SOLUTION FOR SELTEN GAMES 8 AND 10

Games 8 and 10 in the Selten and Chmura paper require more careful treatment in deriving the RBB solution, because the payoff matrices for player 2 contain multiple entries that are equal to the guaranteed payoff. This has the consequence that the transformed matrix includes a column of zeros, specifically the transformed matrices are

$$\begin{aligned} \text{Game 8: } u^1 &= \begin{pmatrix} 3 & -3\lambda_1 \\ 0 & 5 \end{pmatrix}, & u^2 &= \begin{pmatrix} 0 & 9 \\ 0 & -2\lambda_2 \end{pmatrix} \\ \text{Game 10: } u^1 &= \begin{pmatrix} 2 & -3\lambda_1 \\ 0 & 6 \end{pmatrix}, & u^2 &= \begin{pmatrix} 0 & 7 \\ 0 & -4\lambda_2 \end{pmatrix} \end{aligned}$$

This column of zeros means that the Kernel of the matrix is non-empty and so we will not be able to find eigensolutions. However, if we modify the payoffs by small amounts, of order  $\epsilon$ , there can be eigensolutions. In the RBB solution we have  $\mathbf{q}^{\text{rbb}} \propto (u^2)^T \mathbf{p}^{\text{rbb}}$  and  $\mathbf{p}^{\text{rbb}} \propto u^1 \mathbf{q}^{\text{rbb}}$ . The first of these equations applied to the perturbed matrix for Game 8 tells



us

$$\begin{pmatrix} q^{\text{rbb}} \\ 1 - q^{\text{rbb}} \end{pmatrix} \propto \begin{pmatrix} O(\epsilon) \\ (9 + 2\lambda_2)p^{\text{rbb}} - 2\lambda_2 + O(\epsilon) \end{pmatrix}. \quad (\text{E1})$$

This can give us a non-trivial solution in the limit that  $\epsilon \rightarrow 0$  provided that  $(9 + 2\lambda_2)p^{\text{rbb}} - 2\lambda_2 = O(\epsilon)$ . This defines an RBB solution  $p^{\text{rbb}}$ , from which the corresponding  $q^{\text{rbb}}$  can be computed using  $\mathbf{q}^{\text{rbb}} \propto (u^1)^{-1}\mathbf{p}^{\text{rbb}}$ . This solution is unique when the limit  $\epsilon \rightarrow 0$  is taken, independently of the way in which the matrix is initially perturbed. For the Selten game, these solutions are

$$\begin{aligned} \text{Game 8: } p^{\text{rbb}} &= \frac{2\lambda_2}{9 + 2\lambda_2}, & q^{\text{rbb}} &= \frac{6\lambda_2 + 27\lambda_1}{10\lambda_2 + 27\lambda_1 + 27} \\ \text{Game 10: } p^{\text{rbb}} &= \frac{4\lambda_2}{7 + 4\lambda_2}, & q^{\text{rbb}} &= \frac{24\lambda_2 + 21\lambda_1}{24\lambda_2 + 21\lambda_1 + 14} \end{aligned}$$

One important thing to note, however, is that while we can define a unique RBB equilibrium solution for this games, dealing with deviations from the solution arising from the hyperprior, which we discuss in the section IIB, is slightly more difficult. In general, if the observed strategy differs from the RBB solution, we can compute the belief probability which would have led to the observed strategy. For the Selten and Chmura (2008) Games 8 and 10, this can be done as normal for player 1, since the  $u^1$  matrix is non-singular. However, for player 2, using the perturbative technique described above, a deviation from the RBB strategy arises from the  $O(\epsilon)$  part of  $\mathbf{p}$ . Thus, the belief and RBB probabilities defined in this way coincide for any observed data. This is a consequence of the singularity in the transformed matrices, so in the analysis described in Sections IIB and IIC we will ignore the data for Games 8 and 10 because of this issue.