ESSAYS ON REPEATED GAMES

Shivani Nayyar

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF ECONOMICS

Adviser: Thomas Palfrey

June 2009

Abstract

The theory of repeated games explores how mutual help and cooperation are sustained through repeated interaction, even when economic agents are completely self-interested beings. This thesis analyzes two models that involve repeated interaction in an environment where some information is private.

In the first chapter, we characterize the equilibrium set of the following game. Two players interact repeatedly over an infinite horizon and occasionally, one of the players has an opportunity to do a favor to the other player. The ability to do a favor is private information and only one of the players is in a position to do a favor at a time. The cost of doing a favor is less than the benefit to the receiver so that, always doing a favor is the socially optimal outcome. Intuitively, a player who develops the ability to do a favor in some period might have an incentive to reveal this information and do a favor if she has reason to expect future favors in return.

We show that the equilibrium set expands monotonically in the likelihood that someone is in a position to do a favor. It also expands with the discount factor. However, there are no fully efficient equilibria for any discount factor less than unity. We find sufficient conditions under which equilibria on the Pareto frontier of the equilibrium set are supported by efficient payoffs. We also provide a partial characterization of payoffs on the frontier in terms of the action profiles that support them.

In the second chapter, we use numerical methods to compute the equilibrium value set of the exchanging favors game. We use techniques from Judd, Yeltekin and Conklin (2003) to find inner and outer approximations of the equilibrium value set which, together, provide bounds on it. Any point contained in the inner approxima-

tion is certainly an equilibrium payoff. Any point not in the outer approximation is certainly not in the value set.

These inner and outer monotone approximations are found by looking for boundary points of the relevant sets and then connecting these to form convex sets. Working with eight boundary points gives us estimates that are coarse but still capture the comparative statics of the equilibrium set with respect to the discount factor and the other parameters. By increasing the number of boundary points from eight to twelve, we obtain very precise estimates of the equilibrium set. With this tightly approximated equilibrium set, the properties of its inner approximation provide good indications of the properties of the equilibrium set itself. We find a very specific shape of the equilibrium set and see that payoffs on the Pareto frontier of the equilibrium set are supported by current actions of full favors. This is true so long as there is room for full favors, that is, away from the two ends of the frontier.

The third chapter extends the concept of Quantal Response equilibrium, a statistical version of Nash equilibrium, to repeated games. We prove a limit Folk Theorem for a two person finite action repeated game with private information, the very specific additive kind introduced by the Quantal Response model. If the information is almost complete and the discount factor is high enough, we can construct Quantal Response Equilibria very close to any feasible individually rational payoffs. This is illustrated numerically for the repeated Prisoners' Dilemma game.

Acknowledgements

The completion of this thesis was not possible without the guidance of my advisers, Prof. Thomas Palfrey and Prof. Stephen Morris. I thank them for their time, their constant encouragement and their belief in me. I benefitted greatly from their knowledge and their intellect. Among other things, they have taught me how to shape my ideas and ask the right questions. I thank Prof. Dilip Abreu for many useful discussions and in helping me fit my research in the larger framework. As for filling in the details, I could not have done this without the brilliant help of Prof. Satoru Takahashi.

I thank my parents, Harish and Rita Nayyar, my earliest and most important teachers. They taught me the value of knowledge, diligence and doing things well. I was lucky to have the support of my wonderful brother, Mohit. He has been my personal step by step survival guide in graduate school just as he was in high school.

I thank the wonderful friends I made at Princeton: Divya Arora, Nandita Gawade, Chaitanya Guttikar, Satyen Kale and Sairam Subramanian. Among other things, they taught me to ask for more from life. My colleagues in the basement of Fisher Hall were an inspiring bunch to work with and some, especially Karen Kaiser and Wioletta Dziuda, will be friends for life.

I thank my partner in this journey, my husband, Samvit. Having reached this milestone before me, he has waited patiently for me. In the last year he has been, even more than usual, my biggest support and source of strength.

To my family

Contents

Chapter 1

Efficient Exchange of Favors with Private Information

1.1 Introduction

The theory of repeated games explains how mutual help and cooperation are sustained through repeated interaction, even when economic agents are completely self-interested beings. This paper examines how cooperation is sustained in the presence of private information and when there is a lack of immediate reciprocity.

Imagine a technology intensive industry where technological breakthroughs are hard to come by and the market is served almost entirely by two firms who use similar but not identical technologies. Suppose one of these firms makes a significant cost saving discovery which is more suitable to the other firm's technology. While communicating this discovery to the rival firm will be the socially efficient thing to do, will this firm consider sharing the information? With this setting in mind, we study a model where two players interact repeatedly over an infinite horizon and occasionally, one of the players has an opportunity to do a favor to the other player. The ability to do a favor is private information and only one of the players is in a position to do a favor at a time. The cost of doing a favor is less than the benefit to the receiver so that, always doing a favor is the socially optimal outcome. Intuitively, a player who develops the ability to do a favor in some period might have an incentive to reveal this information and do a favor if she has reason to expect future favors in return.

One can think of several other situations in economics and in political economy which can be modelled as people exchanging favors. For example, politicians might support legislation introduced by other politicians, expecting similar support in the future, when trying to pass their own legislations. Our model is a repeated game with privately informed and impatient players and we focus on the set of equilibrium outcomes for this game. It is instructive to think about the reasons why players may not be able to implement the socially optimal outcome that requires them to do favors whenever in a position to do so. The first is impatience. If a player values future payoffs less than current payoffs, she will attach a lower value to a promised favor to be delivered in the future. Moreover, the more favors a player has already accumulated, the less inclined she will be to collect yet another future favor due to its lower marginal value. The second factor is the frequency with which the opponent receives the chance to return a favor. If one player is in a position to do a favor much more frequently than her opponent, her incentive to do the favor is likely to be significantly lower than in the case where they are more equally matched. Given these constraints, we ask what are the most efficient outcomes that can be supported as equilibria in this game.

A two person model of *trading favors* was first studied by Mobius (2001). He analyzed one particular equilibrium where players provide favors on the basis of an intuitive counting mechanism. Players keep track of the difference between favors done by them and by their opponent, in particular, how many favors are *owed* by a player. Since the value of each additional favor owed to her gets progressively lower, a player provides a favor only if the favors she is owed is less than a finite number. The last favor provided is the first one for which the marginal value becomes negative.

Hauser and Hopenhayn (2008) characterize the Pareto frontier of equilibrium outcomes in the Mobius model. In Mobius' equilibrium, the rate of exchange of favors is always one. Hauser and Hopenhayn show that there are significant efficiency gains over this equilibrium by allowing the terms of trade to vary. In particular, on the Pareto frontier, each additional favor provided commands a higher price in terms of future favors promised. This induces players to provide more favors than in the case where the price is always one. Our model differs from that of Hauser and Hopenhayn (2008) in two respects. First, it is in discrete time as opposed to their model which is in continuous time. Second, their paper assumes a symmetric arrival rate of the opportunity to do a favor for both players. Our model allows for asymmetric opportunities of doing favors.

To characterize the equilibrium set of this infinitely repeated game, we follow Abreu, Pearce and Stacchetti (1990) and use the recursive approach to analyzing repeated games. This approach builds on the fact that perfect public equilibria can be viewed as combinations of current strategies and continuation values. More generally, restricting attention to public strategies, one can define payoffs that are *supported* by certain current actions and continuation payoffs. These continuation payoffs reflect the expected value of future favors and also act as state variables. At any point in time, the repeated game is characterized by these state variables. These determine the current actions and the impact of the observable public outcomes on future continuation values.

In the standard model from Abreu, Pearce and Stacchetti (1990), the set of current actions is discrete and finite. Our model differs in that the action space is continuous and we allow for favors ranging from no favor to a full favor (zero to one). Some results that are standard for the finite action but continuous signal model also extend to our model. In particular, we can establish the compactness (Lemma 1.1) and convexity (Lemma 1.2) of the equilibrium set.

In Section 3, we analyze the Pareto frontier of the set of equilibrium payoffs. Under certain intuitive parametric conditions, equilibrium values on the frontier are supported by efficient continuation payoffs (Lemma 1.4). In general, if no favors are observed, players might be punished with continuation values that are inefficient. We find that if with high probability, no one is in a position to do a favor, then inefficient punishments will not be used. Intuitively, if no favor is done, the likelihood that the opponent cheated is low and the value from such severe punishments is low.

Hauser and Hopenhayn (2008) find in the continuous time model that the frontier of the equilibrium set is supported by payoffs also on the frontier or, is self-generating. They argue that this is because of the negligible information content, in a continuous time model, of the event that no favor is done. Our result shares the property that inefficient continuation values are not used when the information content of no favors is low but is weaker than their result. We can show that for payoffs on the frontier, the continuation values must be drawn from the outer boundary of the equilibrium set, a set that includes the Pareto frontier but is potentially larger.

We also want to understand what strategies support equilibrium payoffs on the Pareto frontier and on the entire outer boundary of the equilibrium set. We find that boundary points that are not in the interior of the Pareto frontier are supported by one player doing zero favors (Proposition 1.1). Such payoff pairs are reached when one player has done many more favors than their opponent. The interpretation is that when one player has collected many favors for the future, they wait for some favors to be returned before doing any more.

We also find that some payoff pairs in the interior of the Pareto frontier are supported by players doing full favors. This is the case if there is a region on the frontier where the slope lies between two ratios: the ratio of the benefit of a favor to the cost of a favor and its reciprocal. These are also the slopes of the two arms of the set of feasible long run payoffs for this game which is similar to that of the repeated Prisoner's Dilemma.

We find cooperation among the players supported in equilibrium even when they receive opportunities to do favors with unequal probabilities. Our model does impose some restrictions on how much asymmetry is allowed in the ability of the two players to do favors. Assuming without loss of generality that q, the probability that player 2 will be in a position to do a favor in a period is greater than p, the probability that player 1 will be in a position to do a favor, we have to assume that p/q is greater than c/b. The larger the gap between the cost and benefit of a favor, the larger the asymmetry allowed.

In Section 4, we do comparative statics for the equilibrium set. As long as the asymmetry in p and q is within the required bounds, increasing either probability increases the amount of cooperation that can be supported in equilibrium. Theorem 1.1 shows that holding one player's probability of being in a position to do a favor constant, the higher the probability of the other player, the larger the equilibrium set (weakly).

Theorem 1.2 shows us that in equilibrium, the potential for cooperation is enhanced with more patient players. Specifically, the equilibrium set expands monotonically in the discount factor, in terms of weak set inclusion. Again, this is a standard result for the Abreu, Pearce and Stacchetti (1990) model. However, the proof here is different due to the structural differences of the model.

In a similar vein to this paper, Athey and Bagwell (2001) apply techniques from the theory of imperfect monitoring in repeated games (Abreu, Pearce and Stacchetti (1990), Fudenberg, Levine, Maskin (1994)) to a repeated game with private information. They study optimal collusion between two firms with privately observed cost shocks. They find a discount factor strictly less than 1 with which first-best payoffs are achieved as equilibria. In contrast, in the model in this paper, an exact Folk Theorem is not obtained. For any discount factor less than 1, there exist feasible individually rational payoffs that are not in the equilibrium set. This is shown in Theorem 1.3.

A potential source of this difference in findings is the difference in the shapes of the feasible sets. The Athey and Bagwell (2001) game has a feasible set that has a straight line Pareto frontier. In our game the set of feasible individually rational payoffs has a kinked Pareto boundary, similar to that of a repeated Prisoner's Dilemma. Azevedo and Moreira (2007) show that the generic repeated Prisoner's Dilemma has no efficient equilibrium for any discount factor. They also generalize this result to an *Anti-Folk Theorem* for a larger class of games.

Abdulkadiroglu and Bagwell (2005) analyze a repeated trust game with private information which has a similar flavor. In each period, either player might receive some income and this is private information. This player can then choose to exhibit trust by investing this income with the other player. The investment either succeeds or fails and this outcome is privately observed by the trustee. If the investment is successful, the trustee can choose to "reciprocate" and share the returns with the investor. They found that players are "willing to exhibit trust and thereby facilitate cooperative gains only if such behavior is regarded as a favor that must be reciprocated either immediately or in the future."

1.2 Model

There are two players represented by $i = 1, 2$. At each time period $t = 0, 1, 2, ..., \infty$, one of the players might get a chance to do the other a favor. The cost of doing a favor is c. The benefit to the recipient of the favor is b. $b > c$ and the socially efficient outcome is that a favor is always provided. There is a common discount factor δ and players seek to maximize the present discounted values of their utilities. We assume that the players are risk-neutral.

At any time period t, the state space is $\Omega = \{F_1, F_2, F_\phi\}$. In state F_1, player 1 is in a position to do player 2 a favor; in state F_2, player 2 is in a position do player 1 a favor and in state F_ϕ, neither player is in a position to do a favor. The ability to do a favor is private information. In other words, player i's information set is $\{(F_i), (F_{i \neq j}, F_\phi)\}$ and a player who is not in a position to do a favor does not know whether or not her opponent is in a position to do a favor. The state at time t depends on the realization of the random variable S^t, $S^t = s^t \in \Omega$. S^t is independently and identically distributed over time and in any period, takes the value F_1 with probability p, F_2 with probability q and F_ϕ with probability $r = 1 - p - q$. Throughout this paper (chapter), we assume that $r > 0$ so that there is always a positive probability that neither player is in a position to do a favor.

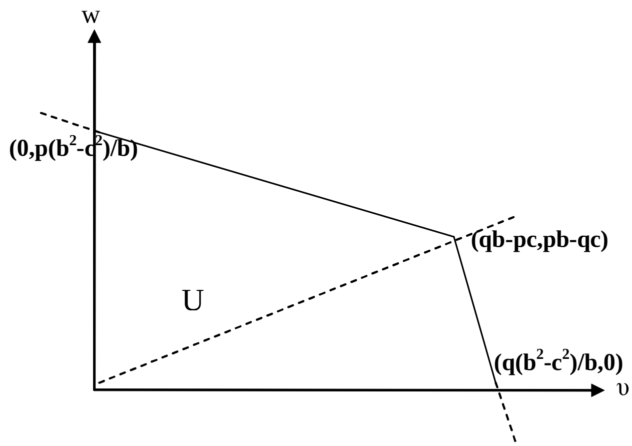
Without loss of generality, we assume that $q > p$, so player 2 is in a position to do a favor with a higher probability in any period. We assume that $p/q > c/b$. Under this assumption the set of long run feasible payoffs of the game looks like the feasible set of the repeated Prisoner's Dilemma (Figure 1). By making this assumption, we are restricting the amount of asymmetry allowed or how much larger q can be than p. The larger the gap between b and c, the more the asymmetry allowed.

Favors are perfectly divisible and players can provide full or partial favors, $x_i \in [0, 1]$, $i = 1, 2$. Figure 1 shows U, the set of feasible long run payoffs of the repeated game. If players do a favor each time they get a chance, the long-run payoffs are $(qb - pc, pb - qc)$. Since $p/q > c/b$, $(qb - pc, pb - qc) > (0, 0)$. If player 2 always does a favor while player 1 never does one, the payoffs are $(qb, -qc)$. If player 1 always does a favor while player 2 never does one, the payoffs are $(-pc, pb)$. We restrict attention to non-negative payoffs.

In the stage game, a strategy for player i is a decision whether to and how much of a favor to provide, if she finds herself in a position to do so. Let X^t be a variable that records, for period t, both the identity of a player who does a favor and how much of a favor is done. $X^t = (X^t(1), X^t(2))$ where $X^t(1) \in \{1, 2\}$ records the identity and $X^t(2) \in (0, 1]$ records the quantity. Let $X^t(1), X^t(2) = 0$ when no favor is done.

In the repeated game, at time t, player i observes her own private history (when she was in a position to do a favor) and a public history of when and what size of favors were provided in the past. At time t, let $h^t = \{X^0, X^1, ..., X^{t-1}\}$ denote a public history. H^t is the set of all possible public histories at time t. We restrict

Figure 1: Set of feasible long run payoffs (q > p)



attention to sequential equilibria in which players condition only on public histories and their current type but not on their private history of types. Such strategies are called public strategies and such sequential equilibria are called perfect public equilibria (Fudenberg, Levine and Maskin (1994)). Let $\sigma_i^t : H^t \times \Omega \to [0,1]$ denote a public strategy for player i at time t and let σ_i denote a sequence of such strategies for $t = 0, 1, ...\infty$, such that,

$$\sigma_i^t(h^t, s^t) = \left\{ \begin{array}{c} x_i^t(h^t) \in [0,1] \text{ if } s^t = F_i \\ 0 \text{ if } s^t \neq F_i \end{array} \right\}.$$

From Abreu, Pearce and Stacchetti (1990), it is known that perfect public equilibria can be expressed recursively. More generally, if we restrict attention to public strategies, any payoff pair (υ, w) can be *factorized* into current and continuation values, $(v_c, w_c) : X^t \to U$, that depend only on current public outcomes. For our purposes, it will be convenient to adopt the following notation. If, at time t, $X^t(1) = 1$, let the continuation values be denoted by $(v_1(x_1), w_1(x_1))$, where $x_1 = X^t(2) > 0$; if $X^t(1) = 2$, let the continuation values be denoted by $(v_2(x_2), w_2(x_2))$, where $x_2 = X^t(2) > 0$; and finally, if $X^t(1), X^t(2) = 0$, let the continuation values be denoted by (v_ϕ, w_ϕ). Also, define $v_1(0) = v_2(0) = v_\phi$ and $w_1(0) = w_2(0) = w_\phi$.

We define an operator B that maps sets of payoff pairs to sets of payoff pairs. For $W \subset R^2$, $B(W)$ is the set of payoffs that can be supported by continuation payoffs in W. More precisely, $B(W) = \{(v, w)$: there exist $(x_1, x_2) \in [0, 1]^2$ and

$$(v_1(x_1), w_1(x_1)), (v_2(x_2), w_2(x_2)), (v_\phi, w_\phi) \in W$$

such that

$$v = p(-(1-\delta)cx_1 + \delta v_1(x_1)) + q((1-\delta)bx_2 + \delta v_2(x_2)) + r\delta v_\phi \tag{1.1}$$

$$w = p((1-\delta)bx_1 + \delta w_1(x_1)) + q(-(1-\delta)cx_2 + \delta w_2(x_2)) + r\delta w_\phi \tag{1.2}$$

$$-c(1-\delta)x_1 + \delta v_1(x_1) \geq \delta v_\phi \tag{1.3}$$

$$-c(1-\delta)x_2 + \delta w_2(x_2) \geq \delta w_\phi \; \}. \tag{1.4}$$

For $(v, w) \in B(W)$, we say that (v, w) are *implemented* by x_1,x_2 and $(v_1(x_1), w_1(x_1)), (v_2(x_2), w_2(x_2)), (v_\phi, w_\phi) \in W$. Note that $(v_1(x_1), w_1(x_1))$, $(v_2(x_2), w_2(x_2))$ and (v_ϕ, w_ϕ) may not be distinct. If (v, w) is implemented by $x_1 = 0$, then $(v_1(0), w_1(0)) = (v_\phi, w_\phi)$. Similarly, if (v, w) is implemented by $x_2 = 0$, then $(v_2(0), w_2(0)) = (v_\phi, w_\phi)$.

The operator $B(.)$ holds a unique place in the repeated games literature. It has several well known properties. It maps convex sets to convex sets and is monotonic, $B(W) \subseteq W$ for any $W \subset R^2$. Let $E(\delta)$ be the set of perfect public equilibrium payoffs, given δ. Following the literature, if $W \subseteq B(W)$, we say that W is a self generating set. It is well known that for any self generating set W, $W \subseteq E(\delta)$, and $E(\delta)$ is the largest bounded self generating set.

1.3 The Equilibrium Set and its Pareto Frontier

The set of equilibrium payoffs $E(\delta)$ depends on p and q. Lemma 1.1 shows that $E(\delta)$ is compact. To establish compactness of $E(\delta)$, we cannot directly apply the standard result from Abreu, Pearce and Stacchetti (1990) (Lemma 1). In the standard model, the set of actions is finite. Once compactness is established for equilibria associated with an action pair, the finite union over all action pairs yields the compactness of $E(\delta)$. In our model, the set of actions is continuous. Compactness can still be established directly as shown below.

Let ClW denote the closure of $W \subset R^2$. We show that $ClE(\delta)$ is a self-generating set which implies that $ClE(\delta) \subseteq E(\delta)$ and that $E(\delta)$ is closed. Since $E(\delta) \subset U$, it is bounded. A closed and bounded set in R^2 is compact.

Lemma 1.1 *$E(\delta)$ is a compact set.*

Proof. We will show that $B(ClE(\delta))$ is a closed set which implies $ClB(ClE(\delta)) \subseteq B(ClE(\delta))$. For any $W \subset R^2$, consider a converging sequence $(\upsilon^n, w^n) \in B(W)$ supported by $(x_1^n, x_2^n) \to (x_1, x_2), (\upsilon_1^n, w_1^n) \to (\upsilon_1, w_1), (\upsilon_2^n, w_2^n) \to (\upsilon_2, w_2), (\upsilon_\phi^n, w_\phi^n) \to$

(υ_ϕ, w_ϕ). Since $(\upsilon_1, w_1), (\upsilon_2, w_2), (\upsilon_\phi, w_\phi) \in ClW$, (υ, w), the limit of (υ^n, w^n), is in $B(ClW)$. So $B(ClW)$ is a closed set.

Now we have $ClE(\delta) \subseteq ClB(E(\delta)) \subseteq ClB(ClE(\delta)) \subseteq B(ClE(\delta))$ where the third set inclusion follows from the monotonicity of $B(.)$. This implies that $ClE(\delta)$ is self generating and hence $ClE(\delta) \subseteq E(\delta)$, $E(\delta)$ is closed and hence compact. ■

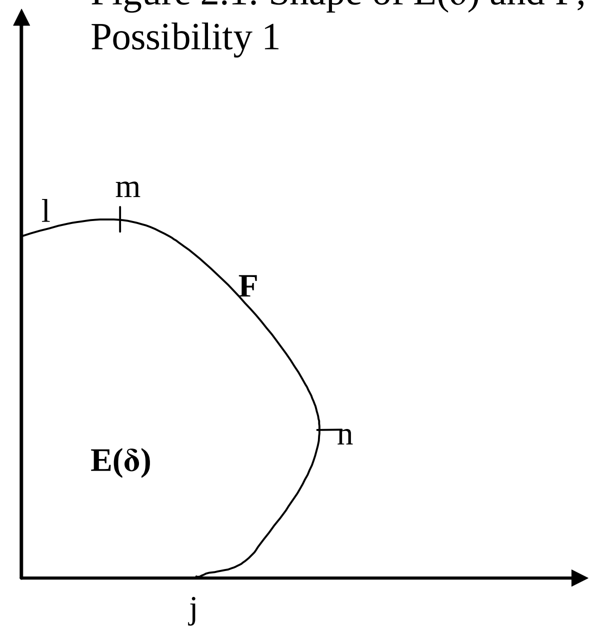
In Lemma 1.2, we see that $E(\delta)$ is convex. Let CoW denote the convex hull of $W \subset R^2$. We show that $CoE(\delta) \subseteq E(\delta)$ which implies that $E(\delta)$ is convex.

Lemma 1.2 *$E(\delta)$ is a convex set.*

Proof. We will show that for any $W \subset R^2$, $CoW \subseteq B(CoW)$. Consider some $\lambda \in [0,1]$. If (υ, w) is implemented by x_1,$x_2 \in [0,1]$ and $(\upsilon_1, w_1), (\upsilon_2, w_2), (\upsilon_\phi, w_\phi) \in W$; and $(\widetilde{\upsilon}, \widetilde{w})$ by $\widetilde{x_1}$,$\widetilde{x_2} \in [0,1]$ and $(\widetilde{\upsilon_1}, \widetilde{w_1}), (\widetilde{\upsilon_2}, \widetilde{w_2}), (\widetilde{\upsilon_\phi}, \widetilde{w_\phi}) \in W$, then $(\lambda\upsilon + (1-\lambda)\widetilde{\upsilon}, \lambda w + (1-\lambda)\widetilde{w})$ is implemented by $\lambda x_1 + (1-\lambda)\widetilde{x_1}$, $\lambda x_2 + (1-\lambda)\widetilde{x_2} \in [0,1]$ and $(\lambda\upsilon_1 + (1-\lambda)\widetilde{\upsilon_1}, \lambda w_1 + (1-\lambda)\widetilde{w_1}), (\lambda\upsilon_2 + (1-\lambda)\widetilde{\upsilon_2}, \lambda w_2 + (1-\lambda)\widetilde{w_2})$, $(\lambda\upsilon_\phi + (1-\lambda)\widetilde{\upsilon_\phi}, \lambda w_\phi + (1-\lambda)\widetilde{w_\phi}) \in CoW$. To see this, we check:

$$p(-c(1-\delta)(\lambda x_1 + (1-\lambda)\widetilde{x_1}) + \delta(\lambda\upsilon_1 + (1-\lambda)\widetilde{\upsilon_1})) + q(b(1-\delta)(\lambda x_2 + (1-\lambda)\widetilde{x_2})$$
$$+\delta(\lambda\upsilon_2 + (1-\lambda)\widetilde{\upsilon_2})) + r\delta(\lambda\upsilon_\phi + (1-\lambda)\widetilde{\upsilon_\phi}) = \lambda\upsilon + (1-\lambda)\widetilde{\upsilon}$$
$$p((1-\delta)b(\lambda x_1 + (1-\lambda)\widetilde{x_1}) + \delta(\lambda w_1 + (1-\lambda)\widetilde{w_1})) + q(-c(1-\delta)(\lambda x_2 + (1-\lambda)\widetilde{x_2})$$
$$+\delta(\lambda w_2 + (1-\lambda)\widetilde{w_2})) + r\delta(\lambda w_\phi + (1-\lambda)\widetilde{w_\phi}) = \lambda w + (1-\lambda)\widetilde{w}.$$

Figure 2.1: Shape of E(δ) and F, Possibility 1



We also check the incentive constraints (1.3) and (1.4):

$$-c(1-\delta)(\lambda x_1 + (1-\lambda)\widetilde{x_1} + \delta(\lambda v_1 + (1-\lambda)\widetilde{v_1})$$

$$= \lambda(-c(1-\delta)x_1 + \delta v_1) + (1-\lambda)(-c(1-\delta)\widetilde{x_1} + \delta\widetilde{v_1})$$

$$\geq \lambda\delta v_\phi + (1-\lambda)\delta\widetilde{v_\phi}$$

$$-c(1-\delta)(\lambda x_2 + (1-\lambda)\widetilde{x_2} + \delta(\lambda w_2 + (1-\lambda)\widetilde{w_2})$$

$$= \lambda(-c(1-\delta)x_2 + \delta w_2) + (1-\lambda)(-c(1-\delta)\widetilde{x_2} + \delta\widetilde{w_2})$$

$$\geq \lambda\delta w_\phi + (1-\lambda)\delta\widetilde{w_\phi}.$$

Now we have $CoE(\delta) \subseteq B(CoE(\delta))$ and from self-generation, $CoE(\delta) \subseteq E(\delta)$. This implies that $E(\delta)$ is a convex set. ■

We now consider E^B, the boundary of the equilibrium set. In Figure 2.1, which shows a possible shape of the equilibrium set, E^B is given by the segment (l, j). E^B is partly defined by the following optimization problem. Problem 1: $g(\upsilon) = \underset{(x_1,x_2,(\upsilon_1,w_1),(\upsilon_2,w_2),(\upsilon_\phi,w_\phi))}{Max} \{w : (\upsilon, w) \in E(\delta)\}$. This is stated in an expanded form below.

Problem 1:

$$g(\upsilon) = \underset{\{x_1,x_2,(\upsilon_1,w_1),(\upsilon_2,w_2),(\upsilon_\phi,w_\phi)\}}{Max} p((1-\delta)bx_1 + \delta w_1) + q(-(1-\delta)cx_2 + \delta w_2) + r\delta w_\phi$$

s.t.

$$p(-(1-\delta)cx_1 + \delta\upsilon_1) + q((1-\delta)bx_2 + \delta\upsilon_2) + r\delta\upsilon_\phi = \upsilon$$

$$-c(1-\delta)x_1 + \delta\upsilon_1 \geq \delta\upsilon_\phi \tag{1.5}$$

$$-c(1-\delta)x_2 + \delta w_2 \geq \delta w_\phi \tag{1.6}$$

$$x_1, x_2 \in [0,1], (\upsilon_1, w_1), (\upsilon_2, w_2), (\upsilon_\phi, w_\phi) \in E(\delta).$$

Also consider Problem 2: $h(w) = \underset{(x_1,x_2,(\upsilon_1,w_1),(\upsilon_2,w_2),(\upsilon_\phi,w_\phi))}{Max} \{\upsilon : (\upsilon, w) \in E(\delta)\}$.

Lemma 1.3 *The value functions in Problems 1 and 2, $g(\upsilon)$ and $h(w)$ are concave and continuous.*

Proof. This follows from Lemma 1.1 and Lemma 1.2. ■

We can define E^B, the boundary of the equilibrium set. To avoid cluttering, we suppress the dependence on p, q and δ.

Definition 1.1 $E^B = \{(\upsilon, w) \in E(\delta)$ *such that* $w = g(\upsilon)\} \cup \{(\upsilon, w) \in E(\delta)$ *such that* $\upsilon = h(w)\}$.

We now define the Pareto frontier of $E(\delta)$, the "downward sloping" part of E^B. Call it F. Let $m = (\upsilon_m, w_m)$ and $n = (\upsilon_n, w_n)$ denote the endpoints of F. These are shown in Figure 2.1, which shows one possible shape of the equilibrium set $E(\delta)$.

Definition 1.2 $F = \{(\upsilon, w) \in E(\delta)$ *such that* $w = g(\upsilon)\} \cap \{(\upsilon, w) \in E(\delta)$ *such that* $\upsilon = h(w)\}$.

The next Lemma shows that under certain parametric conditions, the Pareto frontier F is supported by "efficient" payoffs or payoffs on E^B. This is to be compared to a result in Hauser and Hopenhayn (2008) who find in a similar model that efficient equilibria are always supported by efficient continuous values.

In general, supporting (υ, w) on the frontier might require inefficient continuation values. For example, the punishment payoffs (υ_ϕ, w_ϕ) might have to be inefficient in order to give incentives to the players to do favors when they can. For $(\upsilon, w) \in F$, it is easy to check that the continuation values (υ_1, w_1) and (υ_2, w_2) are efficient. We find that for high values of r, punishments payoffs are also efficient. Recall that $r = 1 - p - q$ is the probability that neither player was in a position to do a favor. Intuitively, if r is high, the public signal of no favor is not a good indicator that the opponent shirked from doing a favor. In this case, a severe inefficient punishment is just not worthwhile.

These conditions are sufficient conditions. It remains to check whether they are necessary. In a similar continuous time model, Hauser and Hopenhayn (2008) find that the Pareto frontier is self generating and efficient equilibria are always supported by efficient continuous values. The reason is again that in a continuous time model, (a signal of) no favor has very low informational content.

Lemma 1.4 *$(\upsilon, w) \in F$ is implemented by $(\upsilon_1, w_1), (\upsilon_2, w_2) \in E^B$. If $p, q < \frac{1}{(1+\frac{b}{c})}$, $(\upsilon, w) \in F$ is implemented by $(\upsilon_\phi, w_\phi) \in E^B$.*

Proof. Suppose $(\upsilon_1, w_1) \notin E^B$. Then there exists $(\upsilon_1 + \varepsilon, w_1 + \varepsilon)$ in the equilibrium set. Replacing the original continuation values with these makes both players better off. This contradicts the fact that $(\upsilon, w) \in F$. Similarly, it can be argued that $(\upsilon_2, w_2) \in E^B$.

Suppose $(\upsilon_\phi, w_\phi) \notin E^B$. Suppose $x_1, x_2 > 0$. Then we can add ε to both υ_ϕ and w_ϕ, still be within $E(\delta)$ and make both players better off thus contradicting $(\upsilon, w) \in F$. To make sure that the incentive constraints still hold we reduce x_1 and x_2 as shown below:

$$\begin{aligned} -c(1-\delta)(x_1 - \frac{\delta\varepsilon}{(1-\delta)c}) + \delta\upsilon_1 &\geq \delta(\upsilon_\phi + \varepsilon) \\ -c(1-\delta)(x_2 - \frac{\delta\varepsilon}{(1-\delta)c}) + \delta w_1 &\geq \delta(w_\phi + \varepsilon) \end{aligned}$$

The expected payoffs for player 1 is now:

$$\begin{aligned} \upsilon' &= p(-c(1-\delta)(x_1 - \frac{\delta\varepsilon}{(1-\delta)c}) + \delta\upsilon_1) + q(b(1-\delta)(x_2 - \frac{\delta\varepsilon}{(1-\delta)c}) + \delta\upsilon_2) \\ &\quad + r\delta(\upsilon_\phi + \varepsilon) \\ \upsilon' - \upsilon &= -q\frac{b}{c}\delta\varepsilon + (p+r)\delta\varepsilon > 0 \Longleftrightarrow q < \frac{1}{(1+\frac{b}{c})} \end{aligned}$$

And for player 2:

$$w' = p(b(1-\delta)(x_1 - \frac{\delta\varepsilon}{(1-\delta)c}) + \delta w_1) + q(-c(1-\delta)(x_2 - \frac{\delta\varepsilon}{(1-\delta)c}) + \delta w_2)$$
$$+ r\delta(w_\phi + \varepsilon)$$
$$w' - w = -p\frac{b}{c}\delta\varepsilon + (q+r)\delta\varepsilon > 0 \Longleftrightarrow p < \frac{1}{(1+\frac{b}{c})}$$

If $x_1 = 0$, then by definition, $(v_1, w_1) = (v_\phi, w_\phi)$. Replacing (v_ϕ, w_ϕ) with $(v_\phi + \varepsilon, w_\phi) \in E(\delta)$ makes one player better off without making the other worse off, contradicting $(v, w) \in F$. Similarly for when $x_2 = 0$.

QED ■

Now we try to further characterize the Pareto frontier of the equilibrium set. Proposition 1.1 shows what action profiles support the end points of F. In Figure 2.1, these would be points m and n. Point m must be supported by $x_2 = 0$ and point n by $x_1 = 0$. Proposition 1.1 also states that all points between l and m are supported by $x_2 = 0$ and points between n and j are supported by $x_1 = 0$. The interpretation is that in the region from l to m, player 2 has collected many favors and stops doing them and the same is true for player 1 in the region from n to j. In between points m and n we expect players to do non-zero favors. (However, note that we have not proved that in this region they must do non-zero favors.)

When one player has done many more favors than the other, they are in a good position and stop doing any more. Only when the opponent returns some of these favors and the players are back in the region (m, n) does this player start doing favors again.

More precisely, we can show that if $(v, w) \in F$ is supported by $x_1 > 0$ and $x_2 > 0$, then the slope at (v, w) lies between $\frac{-b}{c}$ and $\frac{-c}{b}$. At points between l and m and between n and k, the slope is outside these bounds.

In Lemma 1.5, we show that for any v on the frontier supported by $x_1 > 0$ and $x_2 > 0$, the left derivative of $g(.)$, $g'(v^-)$ and the left derivative of $h(.)$, $h'(w^-)$ is weakly smaller than $\frac{-c}{b}$. For points in the interior of F, $h'(w^-)$ is the same as the right derivative of $g(.)$, $g'(v^+)$. This formulation is required as $g(.)$ may not be differentiable, that is, there might be kinks in the frontier. The proof (in the appendix) is based on the following idea. Suppose the slope at a point supported by $x_1 > 0$ and $x_2 > 0$ on the frontier was less than $\frac{-b}{c}$. Consider reducing x_1 by a small amount ε. This means that player one does a smaller favor. Now note that player 2's loss relative to player 1's gain will be $-\frac{b}{c}$. If the slope is lower than $-\frac{b}{c}$, we can lower x_1 and support a pair of payoffs outside the frontier. This would be a contradiction.

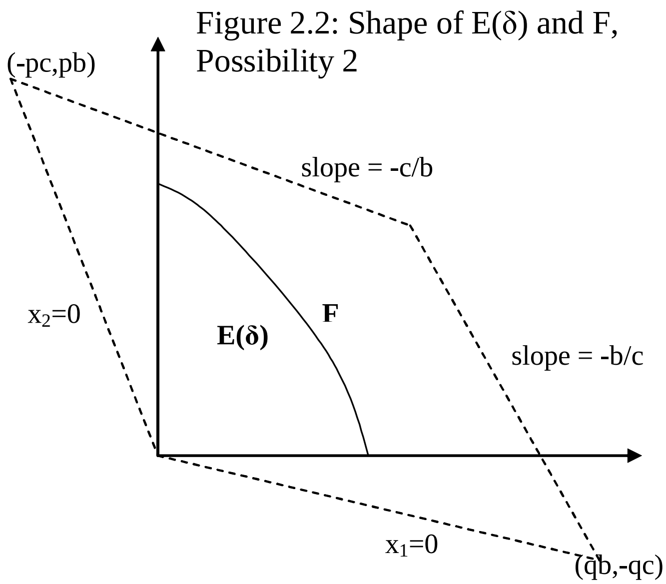
Lemma 1.5 *If* $(v, w) \in F$ *and* $g'(v^-) > \frac{-c}{b}$, *then* $x_2 = 0$. *If* $(v, w) \in F$ *and* $h'(w^-) > \frac{-c}{b}$, *then* $x_1 = 0$.

Proof. In the Appendix. ■

Lemma 1.5 allows us to state the following.

Proposition 1.1 *The endpoints of* F, $m = (v_m, w_m)$ *and* $n = (v_n, w_n)$ *must be supported by* $x_2 = 0$ *and* $x_1 = 0$ *respectively. Moreover,* $(v, w) \in E^B$ *such that* $v < v_m$ *must be supported by* $x_2 = 0$ *and* $(v, w) \in E^B$ *such that* $w < w_m$ *must be supported by* $x_1 = 0$.

Proof. This follows from Lemma 1.5. ■



As mentioned above, we expect payoffs in between m and n, in the interior of F, to be supported by non-zero current actions. That is, neither player has collected so many favors that they wait for some to be returned before doing any more. We define below a region $F^* \subseteq F$ where the slope $g'(.)$ lies between $\frac{-b}{c}$ and $\frac{-c}{b}$. This is a subset of F and is to be thought of as points lying close to the middle of F. Figure 2.2 shows one other possible shape for F and $E(\delta)$. We can see how at the ends of F, the slope might be outside these bounds.

Definition 1.3 $F^* = \{(v, w) \in F$ *such that* $\frac{-b}{c} \leq g'(v^+) \leq g'(v^-) \leq \frac{-c}{b}\}$.

Proposition 1.2 below shows that for payoffs in F supported by continuation payoffs in F^*, players must do full favors. Roughly this shows that payoffs away from the extremes of F are supported by full favors.

Proposition 1.2 *(1) If* $(\upsilon, w) \in F$ *is implemented by* $(\upsilon_1, w_1) \in F^*$ *then* $x_1 = 1$. *(2) If* $(\upsilon, w) \in F$ *is implemented by* $(\upsilon_2, w_2) \in F^*$ *then* $x_2 = 1$.

Proof. (1) Suppose $(\upsilon, w) \in F$ is implemented by $(\upsilon_1, w_1) \in F^*$ and $x_1 < 1$. Then for a small ε, we can make the following adjustments in x_1 and υ_1 without violating the incentive constraint (1.5): $-(1-\delta)c(x_1+\varepsilon) + \delta(\upsilon_1 + \frac{(1-\delta)c\varepsilon}{\delta}) \geq \delta\upsilon_\phi$ and make player 2 better off without making player 1 worse off.

The players' expected payoffs are:

$$(1-\delta)(-pc(x_1+\varepsilon) + qbx_2) + \delta(p(\upsilon_1 + \frac{(1-\delta)c\varepsilon}{\delta}) + q\upsilon_2 + r\upsilon_\phi) = \upsilon$$

$$(1-\delta)(pb(x_1+\varepsilon) - qcx_2) + \delta(pg(\upsilon_1 + \frac{(1-\delta)c\varepsilon}{\delta}) + qw_2 + rw_\phi) =$$

$$w + (1-\delta)pb\varepsilon - p\delta(g(\upsilon_1) - g(\upsilon_1 + \frac{(1-\delta)c\varepsilon}{\delta}))$$

From the definition of F^* and Lemma 1.5, at $\upsilon_1 : g'(\upsilon_1^+) \geq \frac{-b}{c}$, $p\frac{g(\upsilon_1)-g(\upsilon_1+\frac{(1-\delta)c\varepsilon}{\delta})}{\frac{(1-\delta)c\varepsilon}{\delta}} \leq p\frac{b}{c}$ or $(1-\delta)pb\varepsilon - p\delta(g(\upsilon_1) - g(\upsilon_1 + \frac{(1-\delta)c\varepsilon}{\delta})) \geq 0$ which means that player 2 is better off.

The proof for (2) is analogous.

QED ■

To summarize, it is not completely clear what action profiles support payoffs in the interior of F. There might be points in F that do involve partial or even zero favors. However, if there is a region in F such where the slope lies between $\frac{-b}{c}$ and $\frac{-c}{b}$ then we do expect to see some efficient payoffs that are supported by full favors.

1.4 Comparative Statics

In this section, we examine how the equilibrium set $E(\delta)$ changes with changes in the parameters p, q and δ. We find that the amount of cooperation that can be sustained in equilibrium is enhanced by more patient players and also by increasing either p or q.

For the next theorem, we denote the equilibrium set be denoted by $E_{p,q}$ (for given δ), hence making the dependence on p and q explicit. Theorem 1.1 shows that, holding q constant, the equilibrium set expands (weakly) in p. Similarly, holding p constant, the equilibrium set expands (weakly) in q. However, we have to make sure that as q expands, the condition $\frac{p}{q} > \frac{c}{b}$ is still satisfied.

Theorem 1.1 *Given δ, for $1 > p' > p$, $E_{p,q} \subseteq E_{p',q}$. Given δ, for $\frac{pb}{c} > q' > q, E_{p,q} \subseteq E_{p,q'}$.*

Proof. First, we show that for any convex $W \subseteq R^2$, $B_p(W) \subseteq B_{p'}(W)$ where $B(W)$ is the operator defined above and q is held constant. Consider $(\upsilon, w) \in B_p(W)$ supported by $x_1, x_2 \in [0,1]$ and (υ_1, w_1), (υ_2, w_2) and $(\upsilon_\phi, w_\phi) \in W$. Let $\upsilon_1' = \frac{p}{p'}\upsilon_1 + (1 - \frac{p}{p'})\upsilon_\phi$ and let $w_1' = \frac{p}{p'}w_1 + (1 - \frac{p}{p'})w_\phi$. (υ, w) can be supported by p', $(x_1\frac{p}{p'}, x_2)$ and (υ_1', w_1), (υ_2', w_2) and $(\upsilon_\phi, w_\phi) \in W$:

$$\begin{aligned}(1-\delta)(-p'cx_1\frac{p}{p'} + qbx_2) + \delta(p'(\frac{p}{p'}\upsilon_1 + (1-\frac{p}{p'})\upsilon_\phi) + q\upsilon_2 + (1-p'-q)\upsilon_\phi) &= \upsilon \\ (1-\delta)(p'bx_1\frac{p}{p'} - qbx_2) + \delta(p'(\frac{p}{p'}w_1 + (1-\frac{p}{p'})w_\phi) + qw_2 + (1-p'-q)w_\phi) &= w\end{aligned}$$

We also check that the player 1's incentive constraint is satisfied (player 2's is unchanged):

$$\begin{aligned} -(1-\delta)cx_1\frac{p}{p'}+\frac{p}{p'}\delta\upsilon_1+(1-\frac{p}{p'})\delta\upsilon_\phi &\geq \delta\upsilon_\phi \\ \text{or } -(1-\delta)cx_1\frac{p}{p'}+\frac{p}{p'}\delta\upsilon_1 &\geq \frac{p}{p'}\delta\upsilon_\phi \text{ from (1.3)} \end{aligned}$$

We have shown $E_p \subseteq B_p(E_p) \subseteq B_{p'}(E_p)$. Hence E_p is a self-generating set under p' and $E_p \subseteq E_{p'}$. The proof for $q' > q$ is analogous.

QED ■

Theorem 1.2 shows that $E(\delta)$ expands with δ.

Theorem 1.2 *Given p, q, for $1 > \delta' > \delta$, $E(\delta) \subseteq E(\delta')$.*

Proof. First, we show that for any convex $W \subseteq R^2$ that includes the point $(0,0)$, $B_\delta(W) \subseteq B_{\delta'}(W)$ where $B_\delta(W)$ is the operator defines above with respect to δ and p, q are held constant. Consider $(\upsilon, w) \in B_\delta(W)$ supported by $x_1, x_2 \in [0,1]$ and (υ_1, w_1), (υ_2, w_2) and $(\upsilon_\phi, w_\phi) \in W$. (υ', w') (see below) can be supported by $\delta', (0,0)$ and $(\frac{\delta}{\delta'}\upsilon_1, \frac{\delta}{\delta'}w_1)$, $(\frac{\delta}{\delta'}\upsilon_2, \frac{\delta}{\delta'}w_2)$ and $(\frac{\delta}{\delta'}\upsilon_\phi, \frac{\delta}{\delta'}w_\phi) \in W$. Be definition, for $(x_1, x_2) = (0,0)$, $(\upsilon_1, w_1) = (\upsilon_\phi, w_\phi)$ and it is easy to check that the incentive constraints (1.3) and (1.4) are satisfied.

$$\begin{aligned} \upsilon' &= \delta'(p\frac{\delta}{\delta'}\upsilon_1+q\frac{\delta}{\delta'}\upsilon_2+r\frac{\delta}{\delta'}\upsilon_\phi) \\ &= \upsilon-(1-\delta)(-pcx_1+qbx_2) \\ w' &= \delta'(p\frac{\delta}{\delta'}w_1+q\frac{\delta}{\delta'}w_2+r\frac{\delta}{\delta'}w_\phi) \\ &= w-(1-\delta)(pbx_1-qcx_2) \end{aligned}$$

(υ'', w'') below can be supported by $\delta', (\frac{\delta}{\delta'}x_1, \frac{\delta}{\delta'}x_2)$ and $(\frac{\delta}{\delta'}\upsilon_1, \frac{\delta}{\delta'}w_1)$, $(\frac{\delta}{\delta'}\upsilon_2, \frac{\delta}{\delta'}w_2)$ and $(\frac{\delta}{\delta'}\upsilon_\phi, \frac{\delta}{\delta'}w_\phi) \in W$. First we check that the incentive constraints are satisfied:

$$\begin{aligned}
\frac{\delta}{\delta'}(\upsilon_1 - \upsilon_\phi) &\geq \frac{(1-\delta)}{\delta}(\frac{\delta}{\delta'}x_1) \text{ (from (1.3))} \geq \frac{(1-\delta')}{\delta'}(\frac{\delta}{\delta'}x_1) \\
\frac{\delta}{\delta'}(w_2 - w_\phi) &\geq \frac{(1-\delta)}{\delta}(\frac{\delta}{\delta'}x_2) \text{ (from (1.4))} \geq \frac{(1-\delta')}{\delta'}(\frac{\delta}{\delta'}x_2)
\end{aligned}$$

$$\begin{aligned}
\upsilon'' &= (1-\delta')(-pc\frac{\delta}{\delta'}x_1 + qb\frac{\delta}{\delta'}x_2) + \delta'(p\frac{\delta}{\delta'}\upsilon_1 + +q\frac{\delta}{\delta'}\upsilon_2 + r\frac{\delta}{\delta'}\upsilon_\phi) \\
&= \upsilon + ((1-\delta')\frac{\delta}{\delta'} - (1-\delta))(-pcx_1 + qbx_2) \\
&= \upsilon + \frac{(1-\delta')}{\delta'}(-pcx_1 + qbx_2) \\
w'' &= (1-\delta')(pb\frac{\delta}{\delta'}x_1 - qc\frac{\delta}{\delta'}x_2) + \delta'(p\frac{\delta}{\delta'}w_1 + +q\frac{\delta}{\delta'}w_2 + r\frac{\delta}{\delta'}w_\phi) \\
&= w + ((1-\delta')\frac{\delta}{\delta'} - (1-\delta))(pbx_1 - qcx_2) \\
&= w + \frac{(1-\delta')}{\delta'}(pbx_1 - qcx_2)
\end{aligned}$$

Now $(\upsilon, w) = (1-\alpha)(\upsilon', w') + \alpha(\upsilon'', w'')$ where $\alpha = \frac{(1-\delta)\delta'}{1-\delta\delta'}$. Since W is convex, we have shown that $(\upsilon, w) \in B_{\delta'}(W)$. Since $(0,0) \in E(\delta)$ and $E(\delta)$ is convex, we now have that $E(\delta) \subseteq B_\delta(E(\delta)) \subseteq B_{\delta'}(E(\delta))$. Hence $E(\delta)$ is a self-generating set under δ' and $E(\delta) \subseteq E(\delta')$. ■

Having shown the monotonicity of the equilibrium set in the discount factor, the question arises whether in the limit as $\delta \to 1$, a Folk Theorem is obtained. Before tackling that question, we first establish that it is not possible to get an *exact* Folk Theorem. There are no equilibria that lie on the frontier of the feasible set U for any $\delta < 1$. This is established in Theorem 1.3 below. For the proof, we use the fact that equilibrium payoffs on the frontier must be supported by current payoffs and continuation values also on the frontier. More specifically, in supporting an equilibrium payoff profile on the line segment $[L, M]$ (or $[M, N]$), continuation values for public histories with positive probabilities must lie on $[L, M]$ (or $[M, N]$). This follows from the fact that the ex-ante average payoff is a linear combination of current payoffs and continuation values.

Lemma 1.6 shows that any equilibrium payoff profile (v, w) on the line segment $[L, M]$ (or on $[M, N]$) (see Figure 3) must be part of a self-generating interval.

Lemma 1.6 *An equilibrium payoff profile (v, w) on the line segment $[L, M]$ (or on $[M, N]$) belongs to a self-generating interval of equilibrium values, $[(v_a, w_a), (v_b, w_b)]$.*

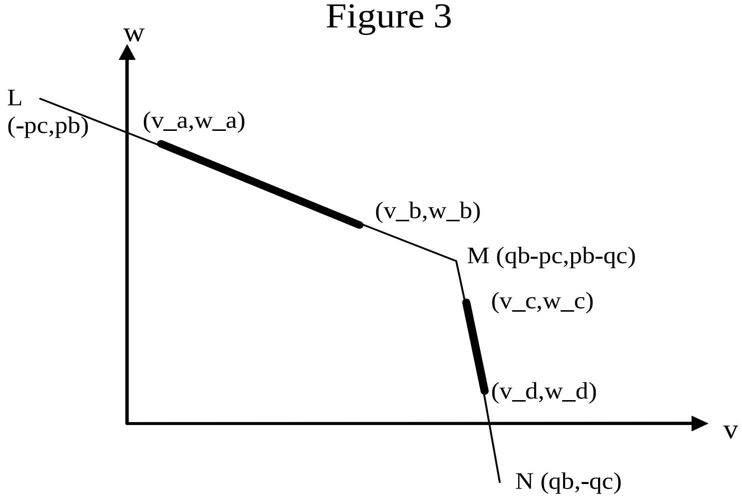
Proof. Consider an equilibrium payoff profile $(v, w) \in E(\delta) \cap [L, M] = B(E(\delta)) \cap [L, M]$. Since any payoff profile in $[L, M]$ must be supported by continuation values also on $[L, M]$, for any W, $B(W) \cap [L, M] \subset B(W \cap [L, M])$. Therefore $B(E(\delta)) \cap [L, M] \subset B(E(\delta) \cap [L, M])$ and $E(\delta) \cap [L, M]$ is a self-generating set. Let $E(\delta) \cap [L, M] = [(v_a, w_a), (v_b, w_b)]$. ■

Theorem 1.3 *There does not exist a $\delta < 1$ such that any point on the Pareto frontier of the feasible set is obtained as a point in $E(\delta)$.*

Proof. Consider supporting some equilibrium on $[L, M]$. From Lemma 1.6, it must be part of a self generating interval $[(v_a, w_a), (v_b, w_b)]$. Also, the current payoffs must lie on $[L, M]$. This implies that $x_1 = 1$.

We show that there is no such self-generating interval. Note that the equation of $[L, M]$ is $v = \frac{-b}{c} w + \frac{p(b^2 - c^2)}{b}$. Suppose there were such an interval $[(v_a, w_a), (v_b, w_b)]$ as

Figure 3



shown in the figure above. That is, suppose (v_b, w_b) can be implemented with x_1, x_2 and continuation values that satisfy $v_a \leq v_2 \leq v_\phi \leq v_1 \leq v_b$. From the incentive constraints, we have $v_\phi \leq v_1 - \frac{c(1-\delta)}{\delta}x_1$ and $v_2 \leq v_\phi - \frac{b(1-\delta)}{\delta}x_2$ (or $w_2 \geq w_\phi + \frac{c(1-\delta)}{\delta}x_2$). Let $v_\phi = v_1 - \frac{c(1-\delta)}{\delta}x_1$ and $v_2 = v_\phi - \frac{b(1-\delta)}{\delta}x_2$. Then $v = (1-\delta)(qbx_2 - pc) + \delta(pv_1 + qv_2 + rv_\phi) = \delta v_1 - c(1-\delta) < v_b$ for $\delta < 1$. This is a contradiction.

Similarly it can be shown that there are no equilibria on $[M, N]$. ■

Theorem 1.3 is similar to the impossibility result of Azevedo and Moreira (2007). They show that for almost every game with imperfect monitoring, there are feasible strictly individually rational payoffs that are not public perfect equilibria. Specifically, they show this for an imperfect monitoring version of the Prisoner's Dilemma game or the partnership game. This is to be contrasted with the Athey and Bagwell (2001) who found a δ strictly less than 1 with which first-best payoffs are achieved as equilibria.

We now try to use the framework of Fudenberg, Levine and Maskin (1994) to check if a Folk theorem (inexact) is obtained in our game. We are able to show that a Folk theorem holds for a discretized version of the game. In particular, we allow, instead of a continuum of favors $x_i \in [0,1]$, partial favors in discrete amounts: $x_i \in \{0, \varepsilon_i, 2\varepsilon_i, ..., k_i\varepsilon_i, ...m_i\varepsilon_i\}$ where $\varepsilon_i > 0$ and $m_i\varepsilon_i = 1$. In this modified model, Proposition 1.3 shows that in the limit as δ goes to 1, $E(\delta)$ approaches the feasible set U. In the terminology of Fudenberg, Levine and Maskin (1994), it can be shown that all stage game strategy profiles have pairwise full rank.

The argument here applies to a modification of our original model. It shows that for arbitrarily small ε_i, for any $W \subset U$, there exists a $\underline{\delta} < 1$ such that for $\delta \geq \underline{\delta}$, U belongs to $E(\delta)$. We would like to show that this is also true in the limit as ε_i goes to 0 (continuous actions). We do not have a formal argument for that yet. One of the issues that might arise is that as ε_i goes to 0, $\underline{\delta}$ might go to 1 . We have not ruled this out.

In the modified model, for $k = 0, 1, 2, ..., m_i$, let $a_i(k) = k\varepsilon_i$ denote the strategy: do $k\varepsilon_i$ favor if in a position to do a favor, none if not. To be able to compare directly with Fudenberg, Levine and Maskin (1994), let $y = \{0, \varepsilon_1, 2\varepsilon_1, ..., 1_1, \varepsilon_2, 2\varepsilon_2, ..., 1_2\}$ be the set of possible public outcomes where 1_i stands for player i doing a full favor. Let $\pi(y \mid a_1, a_2)$ be the probability of observing y given strategy profile (a_1, a_2). It can be shown that the rank of matrix $\Pi(a_1, a_2) = \pi(y \mid ., a_{j\neq i})$ is 1 less than full rank. Then from Theorem 6.2 in Fudenberg, Levine and Maskin (1994), we get the Folk Theorem.

Proposition 1.3 *In the discretized version of the game, for any* $\varepsilon_i > 0$*, for any* $W \subset U$*, there exists* $\underline{\delta} < 1$*, such that for all* $\delta \geq \underline{\delta}$*,* $W \subseteq E(\delta)$*.*

Proof. Let $\pi(y \mid a_1, a_2)$ be the probability of observing y given strategy profile (a_1, a_2). Given strategy profile (a_1, a_2), $\Pi_i(a_1, a_2)$ is the matrix $\pi(y \mid ., a_{j \neq i})$ and $\Pi(a_1, a_2) = \begin{bmatrix} \Pi_1(a_1, a_2) \\ \Pi_2(a_1, a_2) \end{bmatrix}$. For any profile of strategies (a_1, a_2), the matrix $\Pi(a_1, a_2)$ has 1 less than full rank. That is, all stage game strategy profiles have pairwise full rank and from Theorem 6.2 in Fudenberg, Levine and Maskin (1994), we get the Folk Theorem. ■

As an example, consider the strategy profile: $(a_1, a_2) = (2\varepsilon_1, 2\varepsilon_2)$. The matrix $\Pi(a_1, a_2)$ is shown below. Row $k\varepsilon_i$ denotes the strategy: Do $k\varepsilon_i$ favor if in a position to do a favor, none if not; while column $k\varepsilon_i$ denotes a public outcome.

$$\Pi(a_1, a_2) = \begin{array}{c|ccccccccccc} & 0 & \varepsilon_1 & 2\varepsilon_1 & ... & ... & 1 & \varepsilon_2 & 2\varepsilon_2 & ... & ... & 1 \\ 0 & 1-p & 0 & 0 & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 \\ \varepsilon_1 & 1-2p & p & 0 & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 \\ 2\varepsilon_1 & 1-2p & 0 & p & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 \\ ... & 1-2p & 0 & 0 & ... & ... & 0 & 0 & q & 0 & 0 & 0 \\ 1 & 1-2p & 0 & 0 & 0 & 0 & p & 0 & q & 0 & 0 & 0 \\ 0 & 1-p & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \varepsilon_2 & 1-2p & 0 & p & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 \\ 2\varepsilon_2 & 1-2p & 0 & p & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 \\ ... & 1-2p & 0 & p & 0 & 0 & 0 & 0 & 0 & ... & ... & 0 \\ 1 & 1-2p & 0 & p & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q \end{array}$$

To see that the rank of matrix $\Pi(a_1, a_2)$ is 1 less than full rank, note that the rows corresponding to $a_1 = 2\varepsilon_1$ and $a_2 = 2\varepsilon_2$ are identical.

1.5 Conclusion

This paper studied cooperation in an infinitely repeated game with private information. We provide a discrete time version of the Hauser and Hopenhayn (2008) model. We also generalize their model to the asymmetric case where both players may not get the opportunity to do favors with the same probability. Moreover, we answer questions about the equilibrium set not addressed by them. These include showing the monotonicity of the equilibrium set in the discount factor, the monotonicity of the equilibrium set in the likelihood that somebody will be in a position to do a favor and that an exact Folk Theorem does not hold in this model.

A natural extension of the present model is to an n-player game. However, this would raise a number of interesting issues. In a two person model, the meaning of *exchanging favors* is very clear. With more than two players, when in a position to do a favor, it is not so clear whom a player will provide a favor to. This will require careful modelling with respect to the values of favors from different opponents and the cost of doing favors to different opponents. If we assume that the benefit and cost are identical for all players, we will still have to incorporate in the strategies some rules on how favors are done. For example, a player might do one favor for each opponent before doing any second favors. With an appropriate generalization to the n-player case, it is reasonable to still expect the comparative statics results for the equilibrium set that we see in the current model. It is harder to say what the equilibrium strategies for the Pareto frontier will look like.

An n-person extension suggests a possible application to the recent economic innovation of microcredit or credit to poor rural households who were earlier deemed un-creditworthy. A lot of excitement has been generated by loan repayment rates of close to 95 percent (Morduch (1999)) in institutions like the Grameen Bank of Bangladesh. Several studies[1] have attributed the high repayment rate to the novel "group-lending" or *joint liability* feature found in these lending contracts.

As Besley and Coate (1995) explain, joint liability leads to interdependence between the buyers, especially at the time of repayment of the loan. The resultant "repayment game" can also be analyzed in the setup of my model. If one group member has a failed investment and another member has a high enough return from her investment that she is in a position to repay the former's loan along with her own, then the second group member is in a position to do a favor. If borrowers take repeated loans from the bank, they can trade favors and repay each others' loans.

The model here is general enough to accommodate several economic applications. For one, it shares the adverse selection feature of a rich literature in *risk-sharing*. Atkeson and Lucas (1992) study an exchange economy where, each period, a certain endowment has to be allocated among a large number of consumers who receive privately observed taste shocks that affect their marginal utility of consumption. If reporting a high value of the shock leads to a higher current allocation, then all consumers have an incentive to report a higher shock. They characterize the incentive compatible efficient equilibria in this environment. Thomas and Worrall (1990) study a model where a risk neutral lender offers insurance to a risk-averse borrower who receives i.i.d. income shocks. The income shocks are private information and not observable by the lender. The borrower has an incentive to report low income in each period and they characterize the Pareto frontier of equilibrium outcomes in this set-up[2].

[1] See Stiglitz (1990), Varian (1990) and Besley and Coate (1995).
[2] Also see Townsend(1982) and Green (1987) and Hertel (2004).

This connection suggests a possible application of the model. Think of a two person model with two possible income levels: high and low. In the presence of risk aversion, the high income person can stabilize the low income person's consumption and the benefit to the latter would be greater than the cost to the former. In a repeated game, there is a possibility to exchange favors.

In the model here, as in Hauser and Hopenhayn (2008), opportunities to do favors are independent across time periods. A possible direction for future research is a model where the opportunities are correlated rather than independent across time periods. For example, a player's ability to do a favor might be negatively correlated to her ability in the past. This set up would likely aid greater cooperation. Allowing for asymmetric probabilities as in this paper is a first step in that direction. If current probabilities are to depend on what transpired in the previous period, we must first relax the assumption of symmetric probabilities.

1.6 Appendix - Proof of Lemma 1.5

Proof. Part 1: We show that if $x_1 > 0$, then $h'(w^-) \leq \frac{-c}{b}$.

Suppose $h'(w^-) = \lim_{\varepsilon \to 0} \frac{h(w)-h(w-\varepsilon)}{\varepsilon} > \frac{-c}{b}$. Then there exists ε^* such that for all $\varepsilon \leq \varepsilon^*$ and $(h(w), w)$ and $(h(w-\varepsilon), w-\varepsilon)$ on the frontier,

$$h(w) - h(w-\varepsilon) > \frac{-c}{b}\varepsilon \tag{1.7}$$

Let $\varepsilon 1 = \frac{\varepsilon^*}{(1-\delta)pb}$. If $x_1 - \varepsilon 1 \geq 0$, let $\widehat{x_1} = x_1 - \varepsilon 1$. If we replace x_1 by $\widehat{x_1}$, the players' payoffs will be the following. Player 1:

$p(-(1-\delta)c(x_1 - \frac{\varepsilon^*}{(1-\delta)pb}) + \delta \upsilon_1) + q((1-\delta)bx_2 + \delta \upsilon_2) + r\delta \upsilon_\phi = h(w) + \frac{c}{b}\varepsilon^*$

Player 2:

$p((1-\delta)b(x_1 - \frac{\varepsilon^*}{(1-\delta)pb}) + \delta g(\upsilon_1)) + q(-(1-\delta)cx_2 + \delta g(\upsilon_2)) + r\delta w_\phi = w - \varepsilon^*.$

Note that replacing x_1 by $\widehat{x_1}$ does not violate the incentive constraints (1.5) and (1.6). Therefore, we have a new point on the frontier $(h(w) + \frac{c}{b}\varepsilon^*, w - \varepsilon^*)$ where, compared with $(h(w-\varepsilon^*), w-\varepsilon^*)$ player 2 is equally well off and player 1 is better off since $h(w) - h(w-\varepsilon) > \frac{-c}{b}\varepsilon$ from (1.7). Therefore $(h(w+\varepsilon^*), w+\varepsilon^*)$ is not on the frontier. This is a contradiction.

If $x_1 - \frac{\varepsilon^*}{(1-\delta)pb} < 0$, let $\widehat{x_1} = 0$. Note that $x_1 - \frac{\varepsilon^*}{(1-\delta)pb} < 0$ implies $(1-\delta)pbx_1 < \varepsilon^*$. Let $(1-\delta)pbx_1 = \widetilde{\varepsilon} < \varepsilon^*$. With $\widehat{x_1} = 0$, player 1's payoff:

$p(\delta\upsilon_1) + q((1-\delta)bx_2 + \delta\upsilon_2) + r\delta\upsilon_\phi = h(w) + \frac{c}{b}\widetilde{\varepsilon}$

Player 2's payoff:

$p\delta g(\upsilon_1) + q(-(1-\delta)cx_2 + \delta g(\upsilon_2)) + r\delta w_\phi = g(\upsilon) - bp(1-\delta)x_1 = w - \widetilde{\varepsilon}$

Note that replacing x_1 by $\widehat{x_1}$ does not violate the incentive constraints (1.5) and (1.6). Therefore, we have a new point on the frontier $(h(w) + \frac{c}{b}\widetilde{\varepsilon}, w - \widetilde{\varepsilon})$ where, compared with $(h(w-\widetilde{\varepsilon}), w-\widetilde{\varepsilon})$ player 2 is equally well off and player 1 is better off since$h(w) + \frac{c}{b}\widetilde{\varepsilon} > h(w-\widetilde{\varepsilon})$ from (1.7), which is a contradiction.

Part 2: We show that if $x_2 > 0$, then $g'(\upsilon^-) \leq \frac{-c}{b}$ or $f'(g(\upsilon)^+) \geq \frac{-b}{c}$ where $f = g^{-1}$.

Suppose $f'(g(\upsilon)^+) < \frac{-b}{c}$. Then there exists ε^* such that for all $\varepsilon \leq \varepsilon^*$ and $(\upsilon, g(\upsilon))$ and $(\upsilon+\varepsilon, g(\upsilon+\varepsilon))$ on the frontier,

$$\frac{f(g(\upsilon)+\varepsilon) - f(g(\upsilon))}{\varepsilon} < \frac{-b}{c}$$

or

$$f(g(v)+\varepsilon) < v - \frac{b\varepsilon}{c} \tag{1.8}$$

Let $\varepsilon 1 = \frac{\varepsilon^*}{(1-\delta)qc}$. If $x_2 - \varepsilon 1 \geq 0$, let $\widehat{x_2} = x_2 - \varepsilon 1$. If we replace x_2 by $\widehat{x_2}$, the players' payoffs will be the following. Player 1:

$p(-(1-\delta)c(x_1+\delta v_1) + q((1-\delta)b(x_2 - \frac{\varepsilon}{(1-\delta)qc}) + \delta v_2) + r\delta v_\phi = v - \frac{b\varepsilon^*}{c}$

Player 2:

$p((1-\delta)bx_1 + \delta g(v_1)) + q(-(1-\delta)c(x_2 - \frac{\varepsilon}{(1-\delta)qc}) + \delta g(v_2)) + r\delta w_\phi = g(v) + \varepsilon^*$.

Note that replacing x_2 by $\widehat{x_2}$ does not violate the incentive constraints (1.5) and (1.6). Therefore, we have a new point on the frontier $(v - \frac{b\varepsilon^*}{c}, g(v) + \varepsilon^*)$ where, compared with $(f(g(v)+\varepsilon^*), g(v)+\varepsilon^*)$ player 2 is equally well off and player 1 is better off since $v - \frac{b\varepsilon^*}{c} > f(g(v)+\varepsilon^*)$ from (1.8). Therefore $(f(g(v)+\varepsilon^*), g(v)+\varepsilon^*)$ cannot be on the frontier. This is a contradiction.

If $x_2 - \varepsilon 1 < 0$, let $\widehat{x_2} = 0$. Note that $x_2 - \varepsilon 1 < 0$ implies $(1-\delta)qcx_2 < \varepsilon^*$. Let $(1-\delta)qcx_2 = \widetilde{\varepsilon} < \varepsilon^*$. With $\widehat{x_2} = 0$, player 1's payoff:

$p(-(1-\delta)cx_1 + \delta v_1) + q(\delta v_2) + r\delta v_\phi = v - q(1-\delta)bx_2 = v - \frac{b}{c}\widetilde{\varepsilon}$.

Players 2's payoff:

$p((1-\delta)bx_1 + \delta g(v_1)) + q(\delta g(v_2)) + r\delta w_\phi = g(v) + q(1-\delta)cx_2 = g(v) + \widetilde{\varepsilon}$.

Note that replacing x_2 by $\widehat{x_2}$ does not violate the incentive constraints (1.5) and (1.6). Therefore, we have a new point on the frontier $(v - \frac{b\widetilde{\varepsilon}}{c}, g(v) + \widetilde{\varepsilon})$ where, compared with $(f(g(v)+\widetilde{\varepsilon}), g(v)+\widetilde{\varepsilon})$ player 2 is equally well off and player 1 is better off since $v - \frac{b\widetilde{\varepsilon}}{c} > f(g(v)+\widetilde{\varepsilon})$ from (1.8). Therefore $(f(g(v)+\widetilde{\varepsilon}), g(v)+\widetilde{\varepsilon})$ cannot be on the frontier, which is a contradiction.

QED ■

Chapter 2

Efficient Exchange of Favors with Private Information: Computing the Equilibrium Set

2.1 Introduction

We use numerical methods to compute the equilibrium value set of the exchanging favors repeated game from Chapter 1. We use techniques from Judd, Yeltekin and Conklin (2003) that focus on finding inner and outer approximations of the equilibrium value set which, together, provide bounds on it. Any point contained in the inner approximation is certainly an equilibrium payoff. Any point not in the outer approximation is certainly not in the value set.

The numerical methods of Judd, Yeltekin and Conklin (2003) use techniques suggested by the recursive analysis of repeated games (the details of the algorithm are presented in Section 3). In Abreu, Pearce and Stacchetti (APS) (1990), finding the equilibrium set involves finding the largest bounded fixed point of a monotone set valued operator. The key properties of the APS operator are that it maps convex sets to convex sets and is monotone. Judd, Yeltekin and Conklin (2003) use two set valued operators which inherit these properties: inner and outer monotone approximations of the APS operator. Moreover, the largest bounded fixed point of the inner monotone approximation yields an inner approximation to the equilibrium set while the largest bounded fixed point of the outer monotone approximations yields an outer approximation.

Roughly, inner and outer monotone approximations of the APS operator are found by first looking for boundary points of the relevant sets and then connecting these. The computations here are done with two different degrees of precision.

First, we find the required approximate sets using a set of eight boundary points. Despite being coarse, these estimates are able to get a handle on the comparative statics of the equilibrium set. We find that the (inner and outer approximations of the) equilibrium set expands as the discount factor gets larger (Section (4.1.1)). In the model we analyze, players are occasionally in a position to do a favor to an opponent. In the case of symmetric probabilities of doing favors, the larger the probability, the larger the equilibrium set suggested by our estimates (Section (4.1.2)). Holding one player's probability constant and expanding the other's also results in a larger equilibrium set.

Next we make the approximations more precise by using twelve instead of eight boundary points. By increasing the number of boundary points to connect in each iteration, we obtain smoother sets that result in smoother inner and outer approximations. This small adjustment is good enough to greatly improve the precision of the estimates. We find that the inner and outer approximations of the equilibrium set are very close to each other and almost coincide (Section (4.2)). These superior estimates suggest a very specific shape of the equilibrium payoff set. The Pareto frontier of the set extends from the x-axis to the y-axis.

With this tightly approximated equilibrium set, the properties of its inner approximation provide good indications of the properties of the equilibrium set itself. We see that payoffs on the Pareto frontier of the equilibrium set are supported by current actions of full favors. This is true so long as there is room for full favors that is, away from the two ends of the frontier.

Judd, Yeltekin and Conklin (2003) develop techniques for repeated games with perfect monitoring. Our application of their algorithm to the current game shows that these can be extended to games with imperfect monitoring. The algorithm also requires a discrete action space. For this, we alter the basic model slightly. Besides no favor and a full favor (one), players can do a range of favors in between and these go up in small discrete amounts.

The rest of the paper is organized as follows. Section 2 reviews the model from *Efficient Exchange of Favors with Private Information* (2009). Section 3 presents the details of the algorithm used to obtain the results in Section 4. Section 5 concludes.

2.2 Model

There are two players represented by $i = 1, 2$. At each time period $t = 0, 1, 2, ..., \infty$, one of the players might get a chance to do the other a favor. The cost of doing a favor is c. The benefit to the recipient of the favor is b. $b > c$ and the socially efficient outcome is that a favor is always provided. There is a common discount factor δ and players seek to maximize the present discounted values of their utilities. We assume that the players are risk-neutral.

At any time period t, the state space is $\Omega = \{F_1, F_2, F_\phi\}$. In state F_1, player 1 is in a position to do player 2 a favor; in state F_2, player 2 is in a position do player 1 a favor and in state F_ϕ, neither player is in a position to do a favor. The ability to do a favor is private information. In other words, player $i's$ information set is

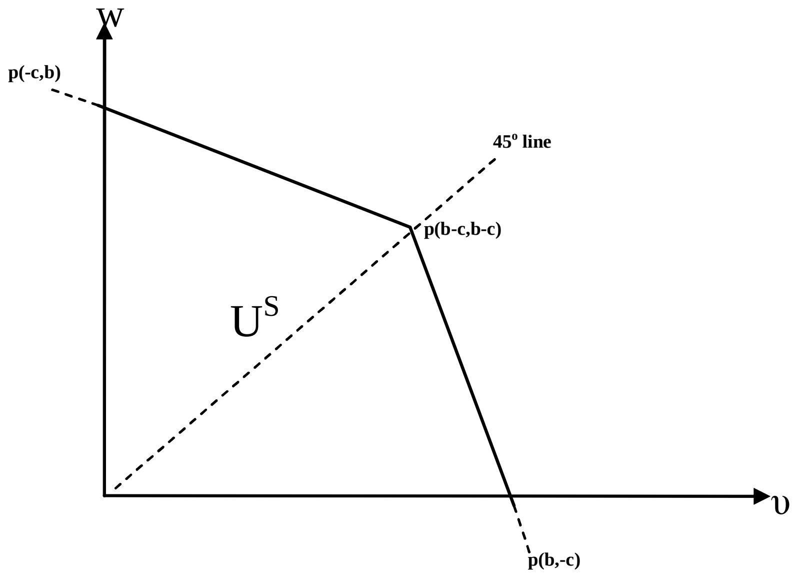
$\{(F_i), (F_{i \neq j}, F_\phi)\}$ and a player who is not in a position to do a favor does not know whether or not her opponent is in a position to do a favor. The state at time t depends on the realization of the random variable S^t, $S^t = s^t \in \Omega$. S^t is independently and identically distributed over time and in any period, takes the value F_1 with probability p, F_2 with probability q and F_ϕ with probability $r = 1-p-q$. Throughout this paper, we assume that $r > 0$ so that there is always a positive probability that neither player is in a position to do a favor.

Without loss of generality, we assume that $q > p$, so player 2 is in a position to do a favor with a higher probability in any period. We assume that $p/q > c/b$. Under this assumption the set of long run feasible payoffs of the game looks like the feasible set of the repeated Prisoner's Dilemma. By making this assumption, we are restricting the amount of asymmetry allowed or how much larger q can be than p. The larger the gap between b and c, the more the asymmetry allowed.

Favors are perfectly divisible and players can provide full or partial favors, $x_i \in [0, 1]$, $i = 1, 2$. Figure 1 shows U^S, the set of feasible long run payoffs for the game in the symmetric case when players 1 and 2 receive the chance to do a favor with equal probabilities. Let this probability be p. Note that $0 < p < \frac{1}{2}$ and $r = 1 - 2p > 0$. The feasible set looks like that of the standard Prisoner's Dilemma, scaled according to p. For example, if players do a favor each time they get a chance, the long-run payoffs are $(p(b - c), p(b - c))$. If player 1 always does a favor while player 2 never does none, the payoffs are $(-pc, pb)$. This is a discrete time version of the Hauser and Hopenhayn (2008).

Figure 2 gives the set of feasible payoffs for the general case and we call it U. If players do a favor each time they get a chance, the long-run payoffs are $(qb - pc, pb -$

Figure 1: Feasible long run payoffs in the symmetric case.

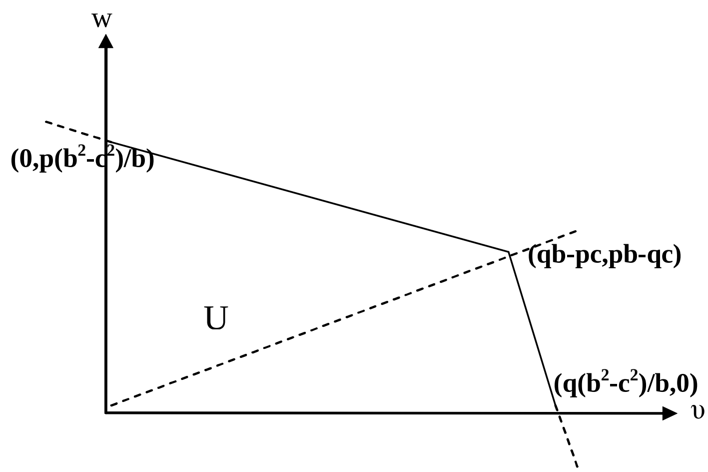


qc). If player 2 always does a favor while player 1 never does one, the payoffs are $(qb, -qc)$. If player 1 always does a favor while player 2 never does one, the payoffs are $(-pc, pb)$. We restrict attention to non-negative payoffs.

In the stage game, a strategy for player i is a decision whether to and how much of a favor to provide, if she finds herself in a position to do so. Let X^t be a variable that records, for period t, both the identity of a player who does a favor and how much of a favor is done. $X^t = (X^t(1), X^t(2))$ where $X^t(1) \in \{1, 2\}$ records the identity and $X^t(2) \in (0, 1]$ records the quantity. Let $X^t(1), X^t(2) = 0$ when no favor is done.

In the repeated game, at time t, player i observes her own private history (when she was in a position to do a favor) and a public history of when and what size of favors were provided in the past. At time t, let $h^t = \{X^0, X^1, ..., X^{t-1}\}$ denote a

Figure 2: Feasible long run payoffs in the asymmetric case (q > p).



public history. H^t is the set of all possible public histories at time t. We restrict attention to sequential equilibria in which players condition only on public histories and their current type but not on their private history of types. Such strategies are called public strategies and such sequential equilibria are called perfect public equilibria (Fudenberg, Levine and Maskin (1994)). Let $\sigma_i^t : H^t \times \Omega \to [0,1]$ denote a public strategy for player i at time t and let σ_i denote a sequence of such strategies for $t = 0, 1, ...\infty$, such that,

$$\sigma_i^t(h^t, s^t) = \left\{ \begin{array}{c} x_i^t(h^t) \in [0,1] \text{ if } s^t = F_i \\ 0 \text{ if } s^t \neq F_i \end{array} \right\}.$$

From Abreu, Pearce and Stacchetti (1990), it is known that perfect public equilibria can be expressed recursively. More generally, if we restrict attention to public strategies, any payoff pair (υ, w) can be *factorized* into current and continuation values, $(\upsilon_c, w_c) : X^t \rightarrow U$, that depend only on current public outcomes. For our purposes, it will be convenient to adopt the following notation. If, at time t, $X^t(1) = 1$, let the continuation values be denoted by $(\upsilon_1(x_1), w_1(x_1))$, where $x_1 = X^t(2) > 0$; if $X^t(1) = 2$, let the continuation values be denoted by $(\upsilon_2(x_2), w_2(x_2))$, where $x_2 = X^t(2) > 0$; and finally, if $X^t(1), X^t(2) = 0$, let the continuation values be denoted by (υ_ϕ, w_ϕ). Also, define $\upsilon_1(0) = \upsilon_2(0) = \upsilon_\phi$ and $w_1(0) = w_2(0) = w_\phi$.

We define an operator B that maps sets of payoff pairs to sets of payoff pairs. For $W \subset R^2$, $B(W)$ is the set of payoffs that can be supported by continuation payoffs in W. More precisely, $B(W) = \{(\upsilon, w) :$ there exist $(x_1, x_2) \in [0, 1]^2$ and

$$(\upsilon_1(x_1), w_1(x_1)), (\upsilon_2(x_2), w_2(x_2)), (\upsilon_\phi, w_\phi) \in W$$

such that

$$\upsilon = p(-(1-\delta)cx_1 + \delta\upsilon_1(x_1)) + q((1-\delta)bx_2 + \delta\upsilon_2(x_2)) + r\delta\upsilon_\phi$$
$$w = p((1-\delta)bx_1 + \delta w_1(x_1)) + q(-(1-\delta)cx_2 + \delta w_2(x_2)) + r\delta w_\phi$$
$$-c(1-\delta)x_1 + \delta\upsilon_1(x_1) \geq \delta\upsilon_\phi$$
$$-c(1-\delta)x_2 + \delta w_2(x_2) \geq \delta w_\phi \ \}.$$

For $(\upsilon, w) \in B(W)$, we say that (υ, w) are *implemented* by x_1,x_2 and $(\upsilon_1(x_1), w_1(x_1)), (\upsilon_2(x_2), w_2(x_2)), (\upsilon_\phi, w_\phi) \in W$. Note that $(\upsilon_1(x_1), w_1(x_1))$,

$(\upsilon_2(x_2), w_2(x_2))$ and (υ_ϕ, w_ϕ) may not be distinct. If (υ, w) is implemented by $x_1 = 0$, then $(\upsilon_1(0), w_1(0)) = (\upsilon_\phi, w_\phi)$. Similarly, if (υ, w) is implemented by $x_2 = 0$, then $(\upsilon_2(0), w_2(0)) = (\upsilon_\phi, w_\phi)$.

The operator $B(.)$ holds a unique place in the repeated games literature. It has several well known properties. It maps convex sets to convex sets and is monotonic, $B(W) \subseteq W$ for any $W \subset R^2$. Let $E(\delta)$ be the set of perfect public equilibrium payoffs, given δ. Following the literature, if $W \subseteq B(W)$, we say that W is a self generating set. It is well known that for any self generating set W, $W \subseteq E(\delta)$, and $E(\delta)$ is the largest bounded self generating set.

2.3 Methodology

We use iterative procedures from Judd, Yeltekin and Conklin (2003) to find inner and outer approximations for the equilibrium value set, $E(\delta)$.

For a convex set $V \subset R^2$, inner approximations are convex hulls of points on the boundary of V. For $W \subseteq U$, we use Algorithm 1 from Judd, Yeltekin and Conklin (2003) to find points on the boundary of the set $B(W)$. Convex hulls of boundary points yield inner approximations to $B(.)$. These inner approximations are convex and monotonic sets and we call them inner monotone approximations to $B(W)$.

The method to find the boundary points (Algorithm 1 in Judd, Yeltekin and Conklin (2003)) is the following. We fix a set of subgradients, call it H, and locate boundary points $z = (\upsilon, w)$ of $B(W)$ where the subgradient of $B(W)$ at z is in H. The input for Algorithm 1 is a set of vertices Z, such that $Co(Z)$=W. For each subgradient $h_l \in H$, we find the action pair and continuation values in W that maximize a weighted sum of the players' current and future discounted payoffs, while satisfying the incentive compatibility constraints. The weights are given by h_l. Each $h_l \in H$ gives a point z_l, the maximized payoff pair, on the boundary of $B(W)$.

The convex hull of all such $z_l's$ gives an inner approximation to $B(W)$. We then look for a fixed point in inner approximations as this gives us a self generating set which is an inner approximation to the value set, $E(\delta)$ (see Theorem 7 and Proposition 8 in Judd, Yeltekin and Conklin (2003)). We define a function that keeps track of the distance between consecutive iterations and perform the iterations until the distance is less than some number ε. This is our stopping rule. Precisely, the distance function sums the distances between pairs of points where each pair corresponds to the same subgradient, $h_{l.}$

Inner approximations are contained in the value set but might be significantly smaller. Outer approximations for any convex set $W \subset R^2$ are polytopes defined by supporting hyperplanes of W. The boundary points for $B(.)$ generated in Algorithm 1 can also be used to find outer approximations to $B(.)$. This involves constructing hyperplanes (in our case, lines) through these boundary points. If z_l is a point on the boundary of W and h_l is the corresponding subgradient, then the required line is $z.h_l = z_l.h_l$. The outer approximation is the intersection of the half spaces defined by these hyperplanes. This too maps convex sets to convex sets and preserves monotonicity.

For the results in the next section, the value of b is 2 and that of c is .5. Partial favors can be done in discrete amounts, $x_i = \{0, .1, .2, ..., .9, 1\}$. Each simulation starts by finding $B(U)$ where U is the set of feasible individually rational payoffs. The stopping rule used is $\varepsilon = 10^{-6}$. In Section 4.1, we use a set of eight subgradients, $H^8 = \{(-1,-1), (0,-1), (1,-1), (1,0), (1,1), (0,1), (-1,1), (-1,0)\}$, in that order (counter-clockwise). In Section 4.2, we use a larger set of twelve subgradients, $H^{12} = \{(-1,-1),(0,-1),(1,-1),(1,0),(1,.33),(1,.66),(1,1),(.66,1),(.33,1),(0,1),(-1,1),(-1,0)\}$.

2.4 Results

2.4.1 Eight Subgradients

As mentioned above, we look for the largest fixed point of inner monotone approximations of $B(.)$. Figure 4.1 shows a few iterations of the inner monotone approximation operator on the way to finding the fixed point. Here, $p = q = .35$ and $\delta = .9$. For each iteration, we form a convex set by connecting the eight points described above. We can see in Figure 4.1 that these iterative sets are nested. We also find that the distance between consecutive iterations decreases monotonically with each subsequent iteration. We use a cutoff rule so that when the distance is small enough, we consider the last iteration to be the fixed point. With the stopping rule $\varepsilon = 10^{-6}$, a fixed point is achieved at the 34th iteration with these parameters. Note that the outermost set in Figure 4.1 is the feasible set, U^S (or U, in general).

For the first iteration, we pick eight points on the boundary of the feasible set U. These are the vertices of U (four points) and the midpoints of its four arms. These can be seen as dots on the boundary of the outermost set in Figure 4.1. Each subsequent iteration potentially yields eight points. However, the subgradients $(1,-1),(1,0)$ always yield the same point. This point is the same point for each iteration and lies on the x-axis. Also, $(0,1),(-1,1)$ always yield the same point. This point is always the same point on the y-axis. For each iteration we end up with six boundary points. The subgradient $(-1,-1)$ always yields the point $(0,0)$.

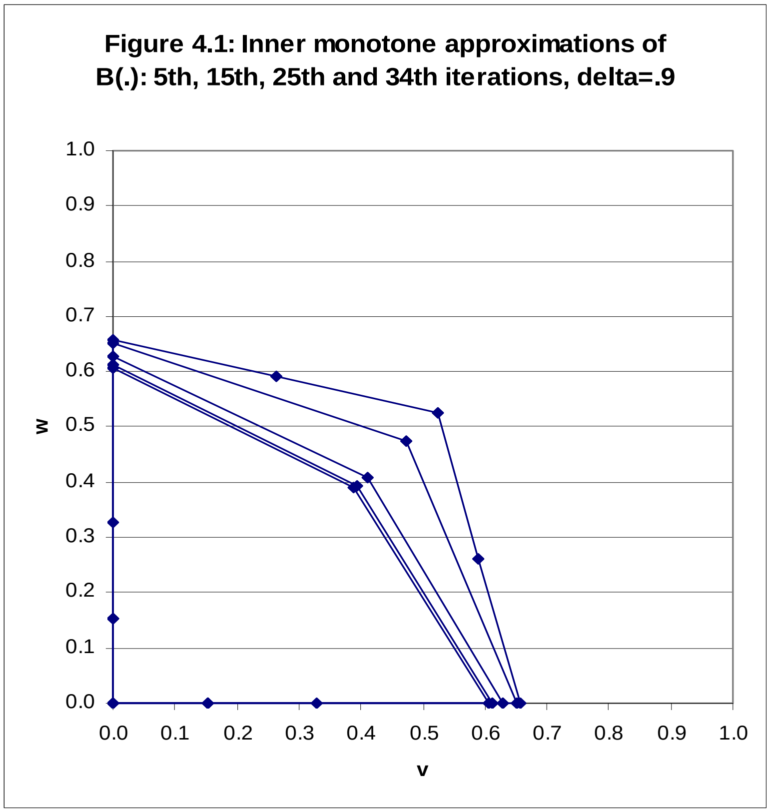


Table 2.1 below shows the boundary points from the last iteration (the fixed point) for the parameters used above, $\delta = .9$ and $p = q = .35$. It also shows the values of x_1, x_2 and the continuation values $(v_1, w_1), (v_2, w_2)$ and (v_ϕ, w_ϕ) that support the each of these boundary points. The first row gives the different values of h_l, the subgradients used. For example, the symmetric boundary point corresponding to subgradient $(1, 1)$, $(.3889, .3889)$ is supported by $(x_1, x_2) = (1, 1)$ and continuation values $(v_1, w_1) = (.4448, .2897), (v_2, w_2) = (.2897, .4448)$ and $(v_\phi, w_\phi) = (.3892, .3892)$.

If we look at the points corresponding to the subgradients $(1, 0)$, $(1, 1)$ and $(0, 1)$, this suggests that points on the Pareto frontier, except the end points, are supported by full favors. At the boundary point corresponding to $(1, 1)$, the expected payoff is

h_l	$(-1,-1)$	$(0,-1)$	$(1,-1)$	$(1,0)$	$(1,1)$	$(0,1)$	$(-1,1)$	$(-1,0)$
v	0	.1530	.6060	.6060	.3889	0	0	0
w	0	0	0	0	.3889	.6060	.6060	.1530
x_1	0	0	0	0	1	1	1	.9
x_2	0	.9	1	1	1	0	0	0
v_1	0	.1	.6064	.6064	4448	.0556	.0556	0.5
w_1	0	0	0	0	.2897	.5754	.5754	.1
v_2	0	.1	.5754	.5754	.2897	0	0	0
w_2	0	.05	.0556	.0556	4448	.6064	.6064	.1
v_ϕ	0	.1	.6064	.6064	.3892	0	0	0
w_ϕ	0	0	0	0	.3892	.6064	.6064	.1

Table 2.1: Eight point inner approximation of symmetric $E(\delta)$

$(.3889, .3889)$. If player 1 is in a position to do a favor, she does one and the resulting continuation values are $(.4448, .2897)$. If player 2 is in a position to do a favor, she does one and the resulting continuation values are $(.2897, .4448)$. If neither players does a favor, the continuation values are $(.3892, .3892)$.

The subgradient $(1, 0)$ yields a boundary point on the x-axis: $(.6060, 0)$. Here, player 1 does not do a favor. Since players 2's continuation value here is 0 and cannot be lowered further, there is no room to compensate player 1 for an additional favor at this point. However, if player 2 gets a chance to do a favor, she does and the resulting continuation values are $(.5754, .0556)$. Similarly, at the boundary point corresponding to subgradient $(0, 1)$, player 2 does a zero favor while player 1 does a full favor.

Note that the above discussion is for the inner approximation of $E(\delta)$. If we had a much tighter and better approximation (as in Section 4.2), these assertions could be made for $E(\delta)$ itself.

Table 2.2 shows the values of x_1, x_2 and the continuation values (v_1, w_1), (v_2, w_2) and (v_ϕ, w_ϕ) that support the boundary points for the last iteration (the fixed point) in the case where $\delta = .9$, $p = .35$ and $q = .5$. The first row gives the different values of h_l.

h_l	$(-1,-1)$	$(0,-1)$	$(1,-1)$	$(1,0)$	$(1,1)$	$(0,1)$	$(-1,1)$	$(-1,0)$
v	0	.19	.9042	.9042	.7532	0	0	0
w	0	0	0	0	.3783	.6412	.6412	.1530
x_1	0	0	0	0	1	1	1	.9
x_2	0	.9	1	1	1	0	0	0
v_1	0	.1	.9046	.9046	.7536	.0556	.0556	0.5
w_1	0	0	0	0	.3787	.6220	.6220	.1
v_2	0	.1	.8824	.8824	.7536	0	0	0
w_2	0	.05	.0556	.0556	.3787	.6414	.6414	.1
v_ϕ	0	.1	.9046	.9046	.6980	0	0	0
w_ϕ	0	0	0	0	.3231	.6414	.6414	.1

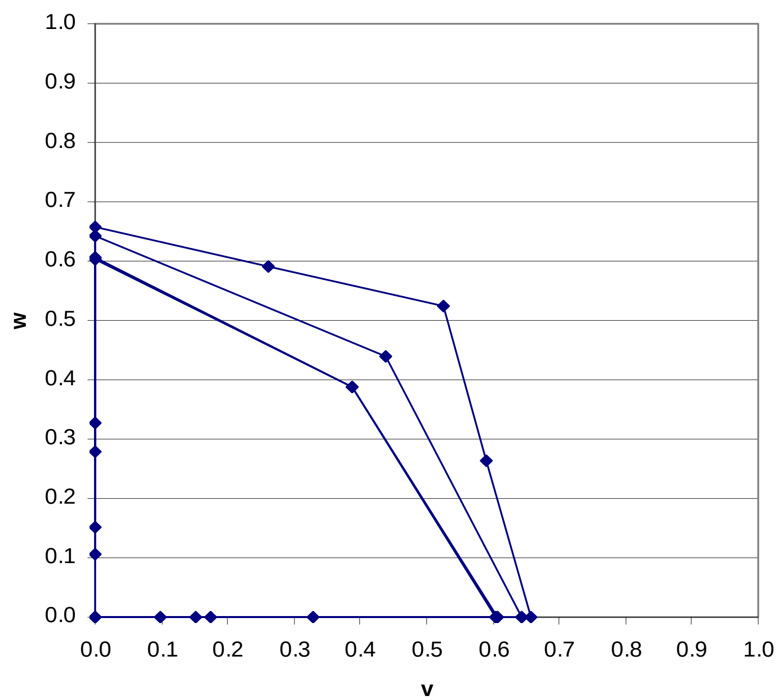
Table 2.2: Eight point inner approximation of asymmetric $E(\delta)$

This table suggests that as in the symmetric case, points on the Pareto frontier of the inner approximation of $E(\delta)$, except the end points, are supported by full favors. Consider the boundary point corresponding to the subgradient $(1, 0) : (.9042, 0)$. Player 2 does a full favor here. resulting in the continuation values $(.8824, .0556)$ while player 1 does none. This is because there is no room to compensate player 1 for any more favors. Similarly, at the boundary point corresponding to the subgradient $(0, 1) : (0, .6412)$, player 1 does a full favor while player 2 does none.

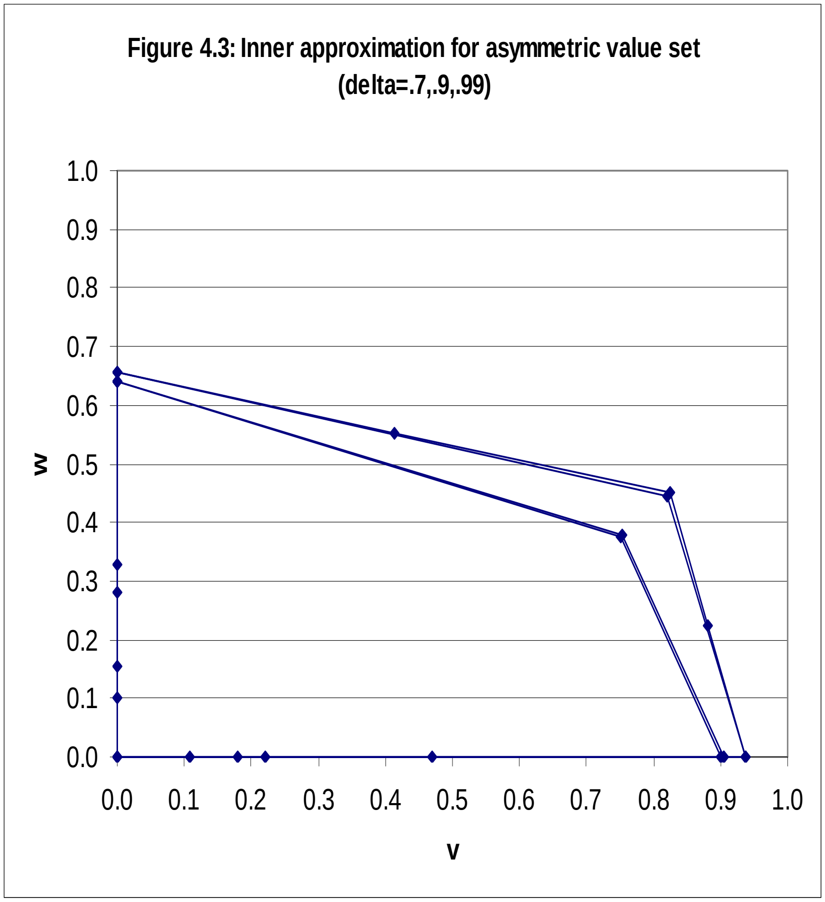
Again, what we are describing here is the inner approximation of the equilibrium set which is contained in it but probably a strict subset with the coarseness of the approximation here.

Inner Monotone Approximations - Comparative Statics with respect to δ : For the symmetric case $p = q = .35$, Figure 4.2 also shows the inner approximations of $E(\delta)$ for different values of the discount factor, $\delta = .7, .9$ and $.99$. These are fixed points of the inner monotone approximations operator (see Figure 4.1). We see monotonicity in δ for the inner approximations. This illustrates Theorem 1.2 in Chapter 1 and suggests strict monotonicity of the equilibrium set in the discount factor.

Figure 4.2: Inner approximations for symmetric value set (delta=.7,.9,.99)



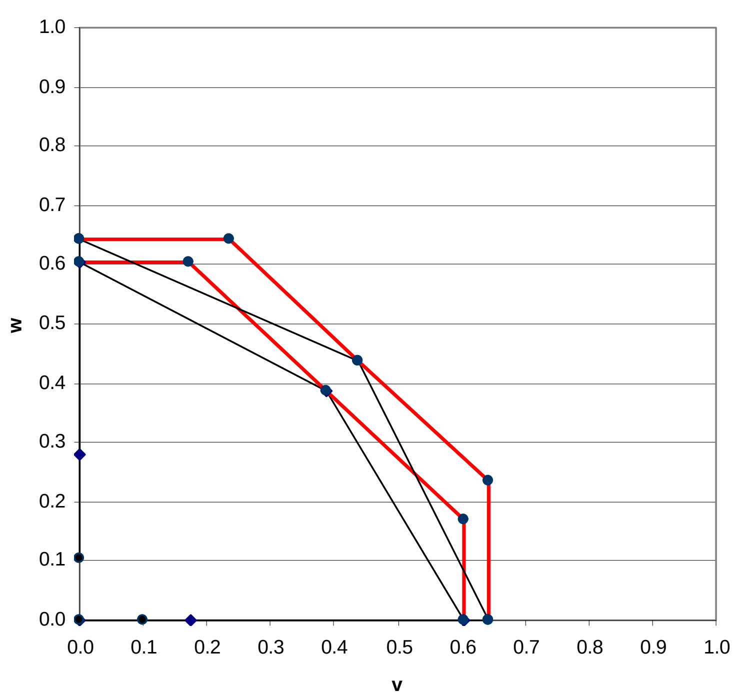
In Figure 4.2, the outermost set is the feasible set, U^S (see Figure 1). In Figure 4.3 below, we observe the same monotonicity in the asymmetric model where $p = .35$ and $q = .5$. Again, the outermost set is the feasible set, U.



Since outer approximations use the same points as inner approximations, monotonicity in δ is also seen in outer approximations (see Figures 4.4 and 4.5).

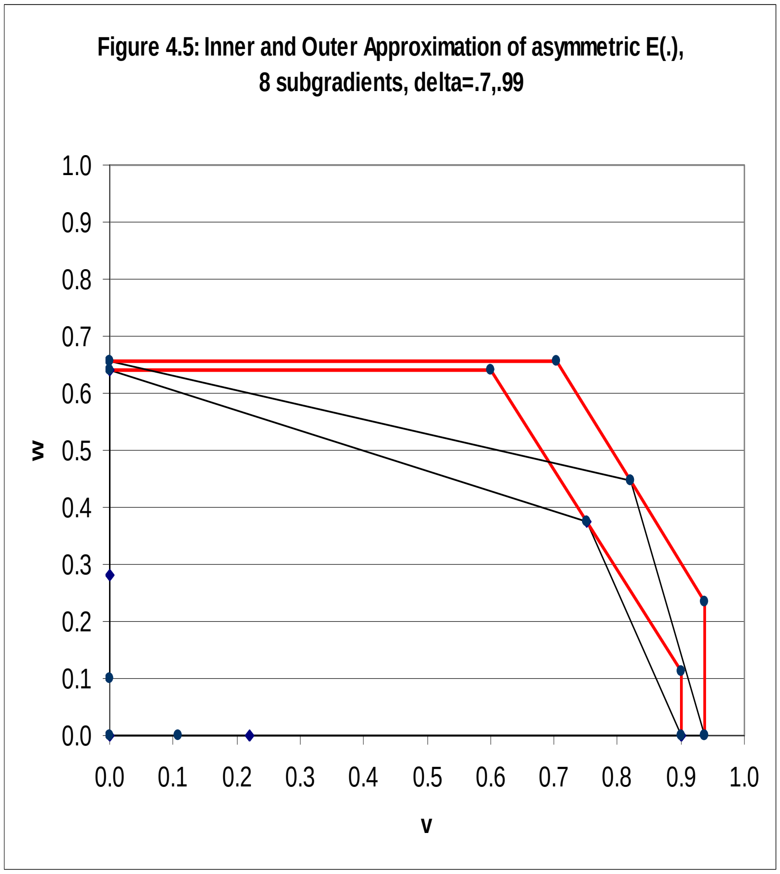
Outer Monotone Approximations Figure 4.4 shows the inner and outer approximations (thick, red lines) for the value set for the case $p = q = .35$ and $\delta = .7, \delta = .99$. The outer approximation was found by finding the tangent hyperplanes through the boundary points that are perpendicular to the subgradients in H^8 and the intersection of the corresponding half spaces.

Figure 4.4: Outer and Inner approximations of symmetric E(.), 8 subgradients, delta = .7,.99



We do see monotonicity in δ in the outer approximations just as we did for the inner approximations. The result would be sharper if any points on the Pareto frontier of $E(\delta = .99)$ were ruled out from the Pareto frontier of $E(\delta = .7)$. This is probably not the case due to the coarseness of these estimates. We do expect to see this with the much tighter approximations obtained with twelve subgradients.

Figure 4.5 shows the inner and outer approximations (thick, red lines) for the value set for the case $p = .35, q = .5$ and $\delta = .7$, $\delta = .99$.



The Pareto frontier of the equilibrium value set must lie within the bounds provided by the inner and the outer approximations. Our estimates suggest a very specific shape of the equilibrium set. In particular, the Pareto Frontier extends from the x-axis to y-axis or in other words, the outer boundary of the set does not bend backwards.

Inner Monotone Approximations - Comparative Statics with respect to p, q **:** Figure 4.6 shows inner approximations of $E(\delta)$ for the symmetric case for three different values of $p(= q)$. These are $p = .25, .35$ and $.45$ respectively. The discount factor is $\delta = .9$ in all three cases. We see monotonicity in p. This confirms Theorem 1.1 from Chapter 1.

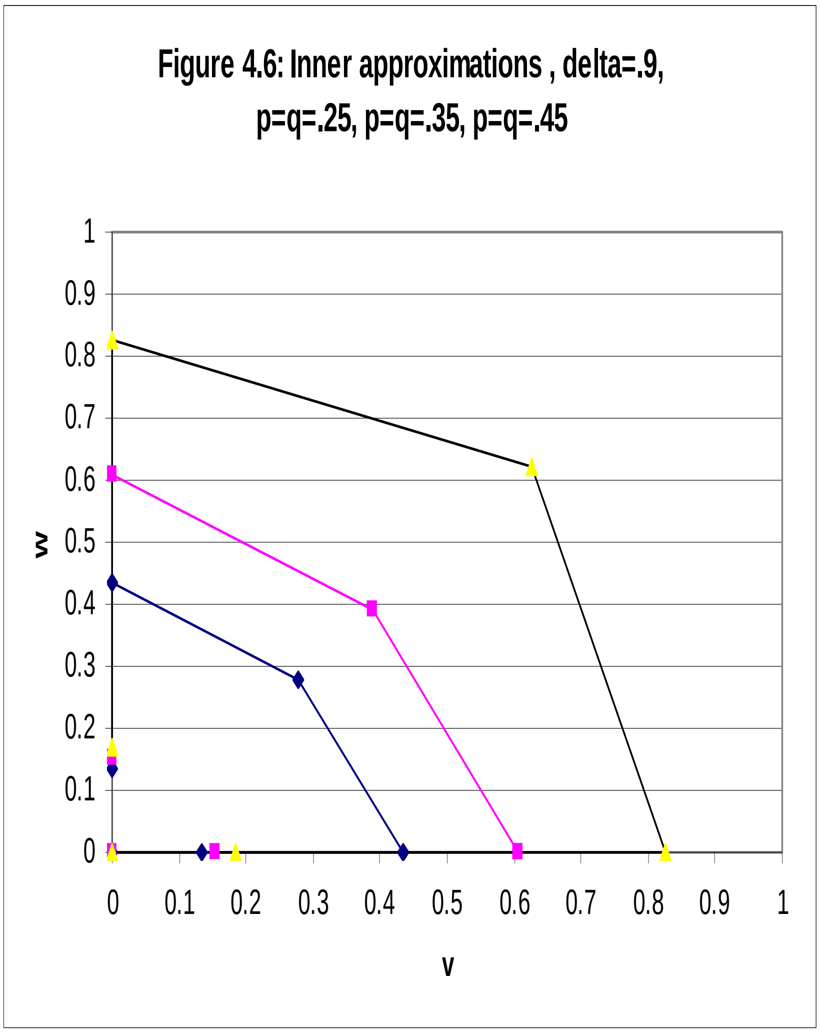
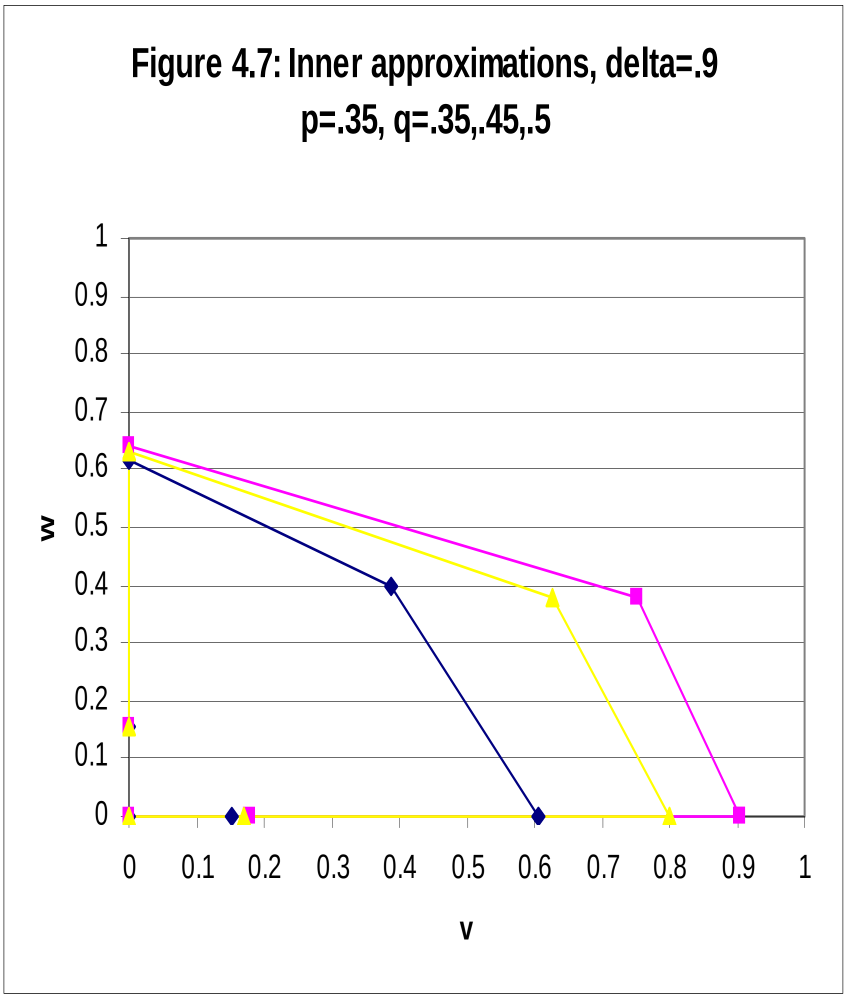


Figure 4.7 shows inner approximations of $E(\delta)$ for $p = .35, q = .35$; $p = .35, q = .45$ and $p = .35, q = .5$ respectively. The discount factor is $\delta = .9$ in all three cases. This shows how the equilibrium set becomes larger as we increase q while holding p constant. Again, this expansion of the inner approximations suggests a stronger version of Theorem 1.1 (strict rather than weak monotonicity).



2.4.2 Twelve Subgradients

In this section, we expand the set H to include twelve subgradients, $H^{12} = \{(-1,-1), (0,-1), (1,-1), (1,0), (1,.33), (1,.66), (1,1), (.66,1), (.33,1), (0,1), (-1,1), (-1,0)\}$. Each iteration now yields twelve points, potentially. (However, for each iteration we end up with ten boundary points. The subgradients $(1,-1), (1,0)$ always yield the same point. Also, $(0,1), (-1,1)$ always yield the same point.) We have expanded the number of subgradients in a way that we are able to get more points on the Pareto frontier of each iteration.

Figure 4.8: Inner monotone approximations of B(.): 12 subgradients, 5th, 15th, 25th, and 34th iterations, delta=.9

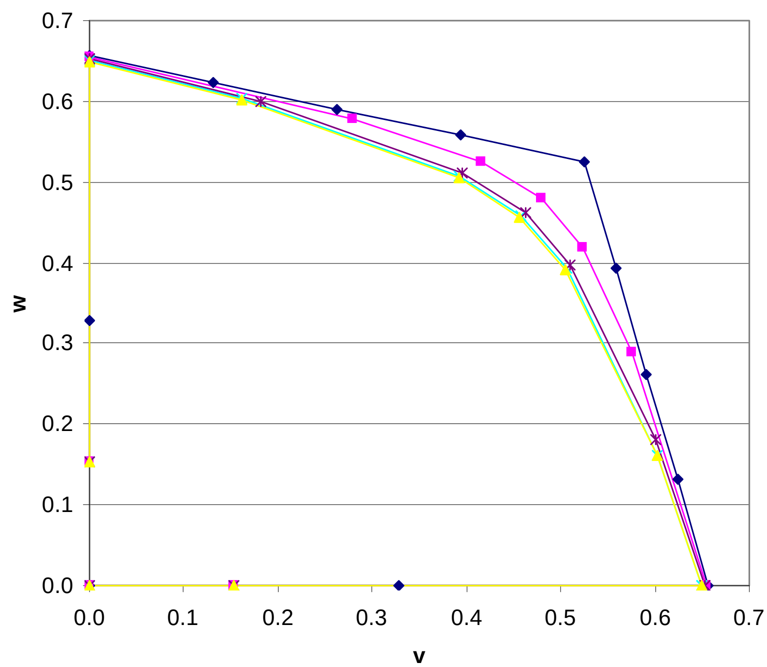


Figure 4.8 shows some iterations of the inner monotone approximation operator for the symmetric case: $p = q = .35$ and $\delta = .9$. With the stopping rule, $\varepsilon = 10^{-6}$, a fixed point is achieved at the 34th iteration.

Figure 4.9 shows, for the parameters $p = q = .35$ and $\delta = .9$, the inner (fixed point from figure 4.8 above) and outer approximations of $E(\delta)$. These provide a very tight bound for $E(\delta)$. By increasing the number of subgradients to twelve, the precision of our exercise has increased greatly. We can also confirm what we found with eight subgradients: the Pareto Frontier extends from the x-axis to y-axis.

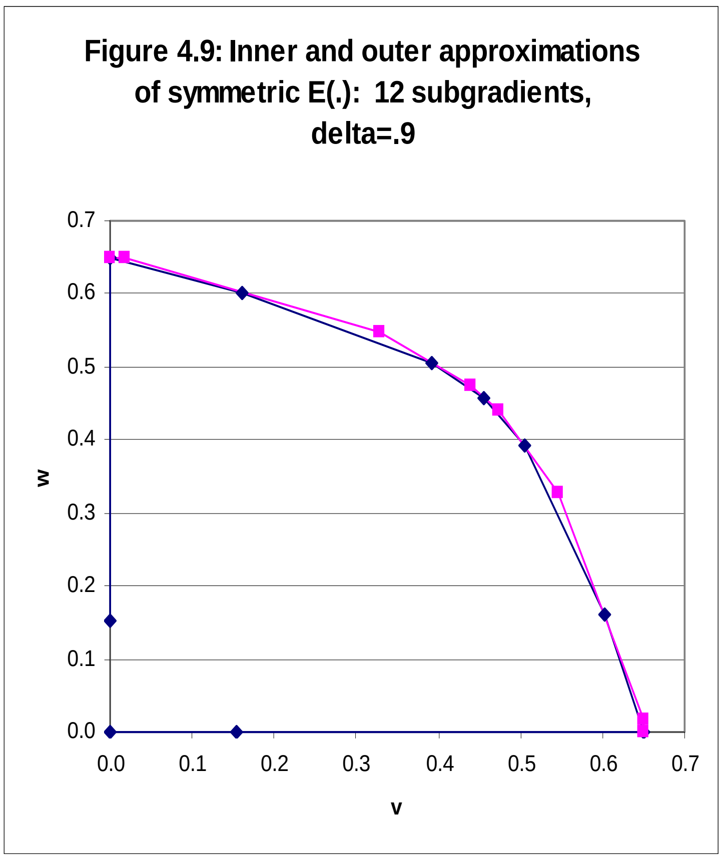


Table 2.3 shows how the boundary points of the inner approximation of $E(\delta)$ are supported. It also shows the values of x_1, x_2 and the continuation values (v_1, w_1),

(v_2, w_2) and (v_ϕ, w_ϕ) that support the each of these boundary points. The first row gives the different values of h_l, the subgradients used. The columns for the subgradients $(-1,-1)$, $(0,-1)$ and $(-1,0)$ have not been shown. These columns are identical to those in Table 2.1, the case of eight subgradients.

Since we have such a tight approximation of $E(\delta)$, the boundary points of the inner approximation of $E(\delta)$ can be thought of as boundary points of $E(\delta)$ itself. For instance, the boundary point corresponding to the subgradient $(1,1)$, $(.4564, .4568)$ is supported by $(v_1, w_1) = (.5053, .3921)$, $(v_2, w_2) = (.3915, .3915)$ and $(v_\phi, w_\phi) = (.4497, .4501)$.

h_l	$(1,-1)$	$(1,0)$	$(1,.33)$	$(1,.66)$	$(1,1)$	$(.66,1)$	$(.33,1)$	$(0,1)$	$(-1,1)$
v	.6494	.6494	.6022	.5056	.4564	.3918	.1613	0	0
w	0	0	.1613	.3915	.4568	.5054	.6022	.6494	.6494
x_1	0	0	1	1	1	1	1	1	1
x_2	1	1	1	1	1	1	1	0	0
v_1	.6495	.6495	.6495	.5541	.5053	.4568	.2235	.0556	.0556
w_1	0		0	.2759	.3921	.4566	.5762	.6332	.6332
v_2	.6332	.6332	.5762	.4568	.3915	.2766	0	0	0
w_2	.0556	.0556	.2235	.4566	.5057	.5539	.6495	.6495	6495
v_ϕ	.6495	6495	.5995	.4986	.4497	.4012	.1680	0	0
w_ϕ	0		.1680	.4010	.4501	.4984	.5995	.6495	6495

Table 2.3: Twelve point inner approximation of $E(\delta)$

If we look at the values of (x_1, x_2) supporting the boundary points, we see that points on the Pareto frontier, except the end points, are supported by full favors. For example, at the boundary point corresponding to the subgradient $(1,1)$, if player 1 is in a position to do a favor, she does a full favor and the resulting continuation values are $(.5053, .3921)$. If player 2 is in a position to do a favor, she does a full favor and the resulting continuation values are $(.3915, .5057)$. If neither players does a favor, the continuation values are $(.4497, .4501)$.

Corresponding to the subgradient $(.66, 1)$, the expected payoffs on the frontier are $(.3918, .5054)$. Corresponding to the subgradient $(1, .66)$, the expected payoffs on the frontier are $(.5059, .3915)$. At both these points, both players do full favors if in a position to do one. This is also true for points corresponding to $(.33, 1)$ and $(1, .33)$ so for all points on the Pareto frontier except the endpoints.

Extending the approximation with twelve subgradients to the asymmetric case is part of future work.

2.5 Conclusion

This paper estimates numerically the equilibrium value set of a repeated exchanging favors game with private information. We adapt techniques from Judd, Yeltekin and Conklin (2003) to evaluate equilibria in repeated games of perfect information. These techniques utilize the usual recursive analysis of repeated games.

We get a lot more precision when we expand the set of subgradients from eight to twelve. Essentially, by looking for more boundary points with each iteration, we find smoother inner and outer approximations. Future work will involve extending the twelve point approximation to the asymmetric model. We expect to see much tighter approximations just as in the symmetric case and better insights into players' behavior on the Pareto frontier of the equilibrium set.

Chapter 3

Quantal Response Equilibrium in Repeated Games

3.1 Introduction

Quantal Response Equilibrium (McKelvey and Palfrey, 1995) is a statistical version of Nash Equilibrium with probabilistic best response functions in which better strategies are more likely to be played than worse strategies. This paper extends the Quantal Response Equilibrium (QRE) concept to repeated games and proves a limit Folk Theorem.

An attractive property of QRE is that systematic deviations from Nash equilibria are predicted without introducing systematic errors. In the basic setup, the payoff from each action that a player can take is subject to random error which can be interpreted as error in calculating payoffs or as privately known payoff disturbances. By imposing a certain structure on the errors (the marginal distribution of the error associated with any action has an extreme value distribution), we can focus on Logit Quantal Response Equilibria.

Logit QRE are a class of QRE parameterized with a parameter λ that is inversely related to the level of error. Given other players' probabilistic strategies, players evaluate their payoffs from alternative actions. Based on these payoffs, the logistic quantal response function prescribes probabilities for playing different actions and QRE is defined as a fixed point of the quantal response function.

Logit Quantal Response Equilibria in extensive form games (McKelvey and Palfrey, 1998) are defined by using the agent model of how an extensive form game is played. At each information set, the log probability of choosing an action is proportional to its continuation payoff where λ is the proportionality factor. Finding the set of Agent QRE involves solving a system of equations, one for each player, for each information set, for each possible action.

It is possible to extend the concept of Agent QRE to infinite horizon games and to infinitely repeated games in particular. We can solve for Agent QRE in such games by restricting attention to strategies that are measurable with respect to a finite number of states. Battaglini and Palfrey (2007) use QRE to study an infinite horizon policy game with endogenous status quo outcomes. This dynamic bargaining problem has proved hard to solve theoretically once non linear utility functions are introduced. They solve the game numerically by finding the limit of the Markov Logit QRE as the error term λ goes to infinity. In this paper, we do not focus on finding the limit equilibrium of a game. Rather, we show that for all finite (two player) repeated games any individually rational and feasible payoff profile can be supported by a limit QRE.

In general, in the limit as $\lambda \to \infty$, players choose perfect best responses. This paper shows that for large enough λ and a discount factor close to one, repeated game QRE payoffs can get arbitrarily close to any feasible individually rational payoff of the static game. This is shown in two steps. At first we assume that the minmax strategies are pure strategies. It is then possible to show that the following strategies constitute a QRE. If in any period, a player deviates from the prescribed action, this triggers a punishment phase in which players play the minmax strategies against one another with probabilities close to one. The punishment phase is long enough that for high λ, deviation is unlikely in the first place, that is, the prescribed action is played with probability close to one.

Next we allow the minmax strategies to be mixed strategies. The punishment phase still requires players to play the minmax strategies. The complication is that mixed strategies are not observable. To make sure that players play the strategies in the support of the mixed minmax strategies in the right proportions, there is a punishment linked to each action in the support of the minmax strategy. The punishment linked to the least desirable action for the player is zero. At the end of the punishment phase, the cumulative punishment from all the other actions is meted out. This is done in a way that ensures that players play their minmax strategies in the correct way.

The rest of the paper is organized as follows. Section 2 defines QRE in infinitely repeated games by extending the concept of Agent QRE. In Section 3, we prove a limit Folk Theorem for a repeated Prisoner's Dilemma Game. This result is illustrated with computations that find equilibria to support different individually rational and feasible payoffs for high values of λ. Section 4 proves the Folk Theorem for a general two person finite action repeated game and Section 5 concludes.

3.2 Model

3.2.1 The stage game G

The stage game G is a finite normal form game. For each player $i = 1, 2, ..., L$, the set S_i represents the finite set of actions in the game. Let $|S_i|$ denote the number of stage game strategies available to player i. The set of action profiles is $S = \prod_{i=1}^{L} S_i$. S_{-i} is the set of action profiles of all players except i, $S_{-i} = \prod_{j \neq i} S_j$.

The payoff vector is $u = (u_1, u_2, ..., u_L)$ where the function $u_i : S \to \mathbb{R}$ gives player $i's$ payoff. Let $\triangle_i$ be the set of probability measures on S_i. Also assume that there is a public randomization device available to players.

3.2.2 The repeated game $G^{\infty}(\delta)$

The stage game is played repeatedly at $t = 1, 2, ...\infty$. Let $h^t = \{s^\tau\}_{\tau=1...t}$, $s^\tau \in S$ denote the history of play up to period t. If $h^{t-1} = \{s^\tau\}_{\tau=1...t-1}$, then we can write $h^t = \{h^{t-1}, s^t\}$. H_t denotes the set of all t period histories.

Player $i's$ strategy $p_i = \{p_i^1, p_i^2, \ldots, p_i^t, \ldots\}$, $p_i^t : H^t \to \triangle_i$, specifies a probability distribution over actions for any possible history h_t. A strategy profile is given by $p = (p_1, ..., p_L)$. Let $p_{s_i}^{h_t}$ denote the probability of playing s_i, following a history of h_t in accordance with strategy profile p.

The common discount factor is δ. Let $W^{h_t} = \{W_1^{h_t}, \ldots, W_L^{h_t}\}$ be the vector of expected continuation payoffs after history h_t.

$$W^{h_t} = (1-\delta)\sum_{\tau=t}^{\infty} \delta^{\tau-t} u(s^\tau \mid p, h^t)$$

Given strategy profile p, let $U_{s_i}^{h_t}(p)$ be the payoff to player i from playing s_i, after history h_t. Then

$$U_{s_i}^{h_t}(p) = \sum_{s_{-i} \in S_{-i}} p_{s_{-i}}^{h_t}[(1-\delta)u_i(s_i, s_{-i}) + \delta W_i^{\{h_t,(s_i,s_{-i})\}}]$$

Definition 3.1 *A Logit(λ) Quantal Response Function is given by*

$$p_{s_i}^{h_t}(p,\lambda) = \frac{e^{\lambda U_{s_i}^{h_t}(p)}}{\sum\limits_{s_j \in S_i} e^{\lambda U_{s_j}^{h_t}(p)}}$$

Definition 3.2 *A Logit(λ) Quantal Response Equilibrium is a fixed point of the Logit Quantal Response function or, a strategy profile p that satisfies*

$$p_{s_i}^{h_t} = \frac{e^{\lambda U_{s_i}^{h_t}(p)}}{\sum\limits_{s_j \in S_i} e^{\lambda U_{s_j}^{h_t}(p)}}$$

for all $h^t \in H^t$, $i = 1, ..., L$, $s_i \in S_i$.

3.3 Repeated Prisoners' Dilemma

3.3.1 Folk Theorem

For the standard Prisoner's Dilemma (Table 3.1), infinitely repeated, we construct a Quantal Response equilibrium that supports expected payoffs very close to $u(s_1^*, s_2^*) > 0$. We assume that a public randomizing device is available and (s_1^*, s_2^*) can be correlated strategies.

	C	D
C	a,a	-b,a+t
D	a+t,-b	0,0

Table 3.1: Prisoner's Dilemma

We partition the set of possible histories into the following two states. If both players have always played (s_1^*, s_2^*) then they are in the cooperative state, State C

which has a continuation value V_i^C. If either of them has ever deviated, then they are in the defection state, State D which has a continuation value V_i^D.

In State C, the probabilities of the players playing s_1^*, s_2^* are $p_{s_1^*}^C$ and $p_{s_2^*}^C$ respectively. In State D, the probabilities of playing C are $p_1^D$and p_2^D. We construct a Quantal Response equilibrium for the repeated game such that $p_{s_1^*}^C \approx 1, p_{s_2^*}^C \approx 1, p_1^D \approx 0$ and $p_2^D \approx$0. Let $p = (p_{s_i^*}^C, \{p_{s_i}^C\}_{s_i \neq s_i^*}, p_i^D)_{i=1,2}$. Note that for $i = 1, 2,\ j \neq i$,

$$\begin{aligned} V_i^C &= p_{s_i^*}^C p_{s_j^*}^C [(1-\delta)u_i(s_i^*, s_j^*) + \delta V_i^C] + \sum_{s_j \neq s_j^*} p_{s_i^*}^C p_{s_j}^C [(1-\delta)u_i(s_i^*, s_j) + \delta V_i^D] \\ &\quad + \sum_{s_i \neq s_i^*} p_{s_i}^C p_{s_j^*}^C [(1-\delta)u_i(s_i, s_j^*) + \delta V_i^D] + \sum_{\substack{s_i \neq s_i^*, \\ s_j \neq s_j^*}} p_{s_i}^C p_{s_j}^C [(1-\delta)u_i(s_i, s_j) + \delta V_i^D] \end{aligned}$$

and

$$\lim_{p_{s_i^*}^C, p_{s_j^*}^C \to 1} V_i^C = u_i(s_1^*, s_2^*) \tag{3.1}$$

Also,

$$V_i^D = (1-\delta)[p_i^D p_j^D a - p_i^D (1-p_j^D) b + p_j^D (1-p_i^D)(a+t)] + \delta V_i^D$$

$$\lim_{p_i^D, p_j^D \to 0} V_i^D = 0 \tag{3.2}$$

In state C, the payoff from playing s_i^* is $U_{s_i^*}^C$ and the payoff from $s_i \neq s_i^*$ is $U_{s_i}^C$:

$$U_{s_i^*}^C = p_{s_j^*}^C[(1-\delta)u_i(s_i^*, s_j^*) + \delta V_i^C] + \sum_{s_j \neq s_j^*} p_{s_j}^C[(1-\delta)u_i(s_i^*, s_j) + \delta V_i^D]$$

$$U_{s_i}^C = p_{s_j^*}^C[(1-\delta)u_i(s_i, s_j^*)] + \sum_{s_j \neq s_j^*} p_{s_j}^C[(1-\delta)u_i(s_i, s_j) + \delta V_i^D$$

In state D, U_i^{DC} is the payoff from playing C and U_i^{DD} is the payoff from playing D:

$$U_i^{DC} = p_j^D(1-\delta)a + (1-p_j^D)(1-\delta)(-b) + \delta V_i^D$$

$$U_i^{DD} = p_j^D(1-\delta)(a+t) + \delta V_i^D$$

These are the Logit Quantal Response functions for $i = 1, 2,\ j \neq i$ in States D and C respectively:

$$\begin{aligned} p_i^D(\lambda, p) &= \frac{1}{1 + \exp \lambda(U_i^{DD} - U_i^{DC})} \\ &= \frac{1}{1 + \exp \lambda(p_j^D(1-\delta)t + (1-p_j^D)(1-\delta)b)} \end{aligned} \tag{3.3}$$

$$\begin{aligned} p_{s_i^*}^C(\lambda, p) &= \frac{1}{1 + \sum_{s_i \neq s_i^*} \exp \lambda(U_{s_i}^C - U_{s_i^*}^C)} \\ &= [1 + \sum_{s_i \neq s_i^*} \exp \lambda(p_{s_j^*}^C(1-\delta)(u_i(s_i, s_j^*) - u_i(s_i^*, s_j^*)) \\ &+ \sum_{s_j \neq s_j^*} p_{s_j}^C(1-\delta)(u_i(s_i, s_j) - u_i(s_i^*, s_j)) + p_{s_j^*}^C\delta(V_i^D - V_i^C))]^{-1} \end{aligned} \tag{3.4}$$

Proposition 3.1 *For $\varepsilon > 0$ and any feasible individually rational payoff profile of the Prisoner's Dilemma, $u(s_1^*, s_2^*)$, there exist $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$, there exists $\underline{\lambda}(\delta)$ with the following property. For $\delta \geq \underline{\delta}$ and $\lambda \geq \underline{\lambda}(\delta)$, there exists a Quantal Response equilibrium in the repeated game in which $V_i^C > u(s_1^*, s_2^*) - \varepsilon$.*

Proof. From (3.1), $\lim_{p^C_{s_i^*}, p^C_{s_j^*} \to 1} V_i^C = u_i(s_1^*, s_2^*)$ and we can find small enough ε^C such that if $p^C_{s_1^*} = 1 - \varepsilon^C$, then

$$V_i^C > u(s_1^*, s_2^*) - \varepsilon \tag{3.5}$$

We show that for Player i, there exist $\underline{\varepsilon}^C$ and $\underline{\varepsilon}^D$ which ensure (3.5) and there exists $\underline{\delta} < 1$ such that if $\delta > \underline{\delta}$ and we assume that $p^C_{s_j^*} \geq 1 - \underline{\varepsilon}^C$ in the current period and $p^C_{s_1^*}, p^C_{s_2^*} \geq 1 - \underline{\varepsilon}^C$ and $p_1^D, p_2^D \leq \underline{\varepsilon}^D$ in all future periods, then $\lim_{\lambda \to \infty} p^C_{s_i^*}(\lambda, p) = 1$. For $s_i \in S_i$, $s_i \neq s_i^*$:

$$\begin{aligned}
& \lim_{p^C_{s_i^*} \to 1, p_i^D \to 0} U^C_{s_i} - U^C_{s_i^*} \\
= & \lim_{p^C_{s_i^*} \to 1, p_i^D \to 0} p^C_{s_j^*}(1-\delta)[u_i(s_i, s_j^*) - u_i(s_i^*, s_j^*)] + \\
& \sum_{s_j \neq s_j^*} p^C_{s_j}(1-\delta)[u_i(s_i, s_j) - u_i(s_i^*, s_j)] + p^C_{s_j^*}\delta(V^i_D - V^i_C) \\
= & \lim_{p^C_{s_i^*} \to 1, p_i^D \to 0} p^C_{s_j^*}(1-\delta)[u_i(s_i, s_j^*) - u_i(s_i^*, s_j^*)] - p^C_{s_j^*}\delta u_i(s_i^*, s_j^*) \\
& \text{(from (3.1) and (3.2))} \\
< & \; 0 \qquad (3.6) \\
\text{when } \delta > & \; \frac{u_i(s_i, s_j^*) - u_i(s_i^*, s_j^*)}{u_i(s_i, s_j^*)}.
\end{aligned}$$

Let $\underline{\delta} = \max_{s_i} \frac{u_i(s_i,s_j^*)-u_i(s_i^*,s_j^*)}{u_i(s_i,s_j^*)} = \frac{a+t-u_i(s_i^*,s_j^*)}{a+t}$. For $(s_i^*, s_j^*) = (C, C)$, $\underline{\delta} = \frac{t}{a+t}$. Consider the Logit Quantal Response function in State C for $i = 1, 2, j \neq i$ (see (3.4)):

$$p^C_{s_i^*}(\lambda, p) = \frac{1}{1 + \sum\limits_{s_i \neq s_i^*} \exp \lambda[U^C_{s_i} - U^C_{s_i^*}]}$$

From (3.6), for $\delta \geq \underline{\delta}$, $\lim_{\lambda\to\infty} p^C_{s_i^*}(\lambda, p) = 1$ and we can find a $\underline{\lambda}^C_{s_i^*}$ such that for $\delta \geq \underline{\delta}$, $\lambda \geq \underline{\lambda}^C_{s_i^*}$, $p^C_{s_i^*}(\lambda, p) \geq 1 - \underline{\varepsilon}^C$. For $\{p^C_{s_i}\}_{s_i \neq s_i^*}$, we can find $\underline{\lambda}^C_{s_i}$ such that for $\lambda \geq \underline{\lambda}^C_{s_i}$, $p^C_{s_i}(\lambda, p) \leq \underline{\varepsilon}^C$. Let $\underline{\lambda}^C = \max(\underline{\lambda}^C_{s_1^*}, \underline{\lambda}^C_{s_2^*}, \{\underline{\lambda}^C_{s_1}\}_{s1 \neq s_1^*}, \{\underline{\lambda}^C_{s_2}\}_{s2 \neq s_2^*})$.

Now assume that $p^D_j \in [0, \underline{\varepsilon}^D]$. Consider the Logit Quantal Response function in State D for $i = 1, 2, j \neq i$ (see (3.3)):

$$p^D_i(\lambda, p) = \frac{1}{1 + \exp \lambda[p^D_j(1-\delta)t + (1-p^D_j)(1-\delta)b]}$$

Since $p^D_j(1-\delta)t + (1-p^D_j)(1-\delta)b > 0$, $\lim_{\lambda\to\infty} p^D_i(\lambda, p^D_j) = 0$ and we can find a $\underline{\lambda}^D_i$ such that for $\lambda \geq \underline{\lambda}^D_i$, $p^D_i(\lambda, p^D_j) \leq \underline{\varepsilon}^D$. Define $\underline{\lambda}(\delta) = \max(\underline{\lambda}^C, \underline{\lambda}^D_1, \underline{\lambda}^D_2)$.

Let $\delta \geq \underline{\delta}$, $\lambda \geq \underline{\lambda}$. The Quantal Response Function is then a continuous function mapping a compact and convex set: $[1-\underline{\varepsilon}^C, 1]^2 \times [0, \underline{\varepsilon}^C]^{K^0} \times [0, \underline{\varepsilon}^D]^2$ into itself where $K^0 = |S_1| + |S_2|$ if (s_1^*, s_2^*) is a correlated action profile and $K^0 = |S_1| + |S_2| - 2$ if (s_1^*, s_2^*) is a pure action profile.

From the Brouwer fixed point theorem, there exists a Quantal Response Equilibrium. As required, for $\lambda \geq \underline{\lambda}$, we have a Quantal Response Equilibrium $p^C_{s_1^*}(\lambda)$, $p^C_{s_2^*}(\lambda) \geq 1 - \underline{\varepsilon}^C$ and $p^D_1(\lambda)$, $p^D_1(\lambda) \leq \underline{\varepsilon}^D$. From (3.5), $V^C_i > u(s_1^*, s_2^*) - \varepsilon$.

QED ■

3.3.2 Computations

We compute exact values for Quantal Response Equilibria (QRE) that support $u(s_1^*, s_2^*)$. More precisely, we find equilibrium strategies that involve playing (s_1^*, s_2^*) with probability close to 1 in the State C and defecting with almost certainty in the state D.

We look for $(p_{s_1^*}^C,\ p_{s_2^*}^C,\ p_1^D,\ p_2^D)$ that are fixed points of the Quantal Response function, for a range of values of λ. We set some initial values for these probabilities: $p0 = (p0_{s_1^*}^C,\ p0_{s_2^*}^C,\ p0_1^D,\ p0_2^D)$. These imply values for the continuation values V_i^C, V_i^D and for the payoff functions $U_{s_i^*}^C, U_{s_i}^C, U_i^{DC}, U_i^{DC}$. Through the Quantal Response function, $U_{s_i^*}^C, U_{s_i}^C, U_i^{DC}$ and U_i^{DC} imply values for $(p_{s_1^*}^C,\ p_{s_2^*}^C,\ p_1^D,\ p_2^D)$: call these $(\widehat{p}_{s_1^*}^C,\ \widehat{p}_{s_2^*}^C,\ \widehat{p}_1^D,\ \widehat{p}_2^D)$. We find $p0$ to minimize the difference between $p0$ and $(\widehat{p}_{s_1^*}^C,\ \widehat{p}_{s_2^*}^C,\ \widehat{p}_1^D,\ \widehat{p}_2^D)$. If the minimized distance is 0, we have a fixed point.

Since the QRE for $\lambda = 0$ must be $(.5, .5, .5, .5)$, we set $p0 = (.5, .5, .5, .5)$. Then for each successive value of λ from 0 to 100 (or more), we set $p0$ to the fixed point for the previous value of λ. In the following simulations, $a = 1, t = 1, b = 1$ and $t/(a + t) = .5$. The specific Prisoner's Dilemma game used is shown in Table 3.2.

	C	D
C	1,1	-1,2
D	2,-1	0,0

Table 3.2: Prisoner's Dilemma

We first examine the case where $\delta = .9$ and $(s_1^*, s_2^*) = (C, C)$. According to Proposition 3.1, we should be able to find Quantal Response Equilibria to support the cooperative outcome so long as $\delta > t/(a + t)$.

For $(s_1^*, s_2^*) = (C, C)$ let $p_{s_1^*}^C = p_1^C$ and $p_{s_2^*}^C = p_2^C$. The results are shown in Figures 1 and 2 (to be read together). Figure 1 plots p_1^C, player 1's probability of playing C in the cooperative state C, as a function of λ. p_2^C, player 2's probability of playing C in the cooperative state C is identical to p_1^C as we would expect given the symmetry of the game, and is not shown here.

As shown in Figure 1, we find two branches for $p_1^C (= p_2^C)$. The solid, blue series begins at .5 for $\lambda = 0$ and converges to 0. For values of λ larger than 10, there is a branch staring close to 1 and quickly reaching 1[1]. This is the dashed, pink colored series. These probabilities comprise QRE in the repeated Prisoner's Dilemma game together with the values of $p_1^D (= p_2^D)$ shown in Figure 2.

Figure 2 plots p_1^D, player 1's probability of playing C in the defection state D, against different values of λ. The solid, blue series begins at .5 for $\lambda = 0$ and converges to 0. The dashed, pink series corresponds to the dashed, pink series in Figure 1 and represents equilibrium probabilities for $\lambda \geq 10$. It happens to coincide with the solid, blue series (for $\lambda \geq 10$) and is hard to see on its own in the figure.

[1] I first ran the program for $\lambda = 0 : 100$. This gave the solid, blue branch. Then I checked to see if (1,1,0,0) was a fixed point for $\lambda = 100$. It was and then I ran the program for $\lambda = 100$ to 0. I found that I got fixed points (minimized distance is 0) for values of $\lambda \geq 10$. I followed a similar procedure for other values of δ and $u(s_1^*, s_2^*)$.

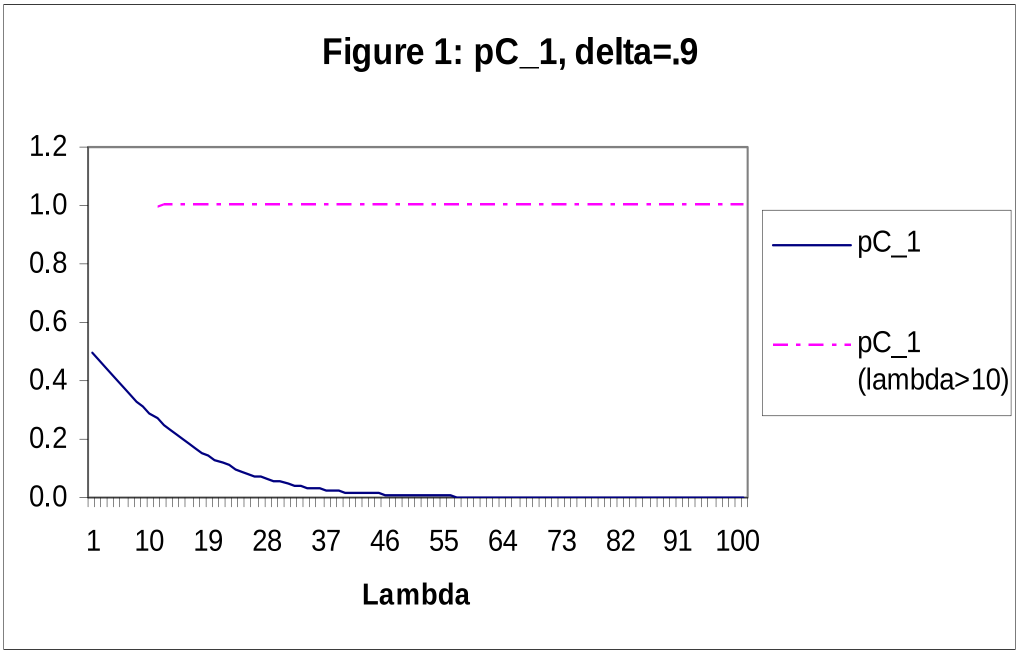


Figure 1: pC_1, delta=.9
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1
10
19
28
37
46
55
64
73
82
91
100
Lambda
pC_1
pC_1 (lambda>10)

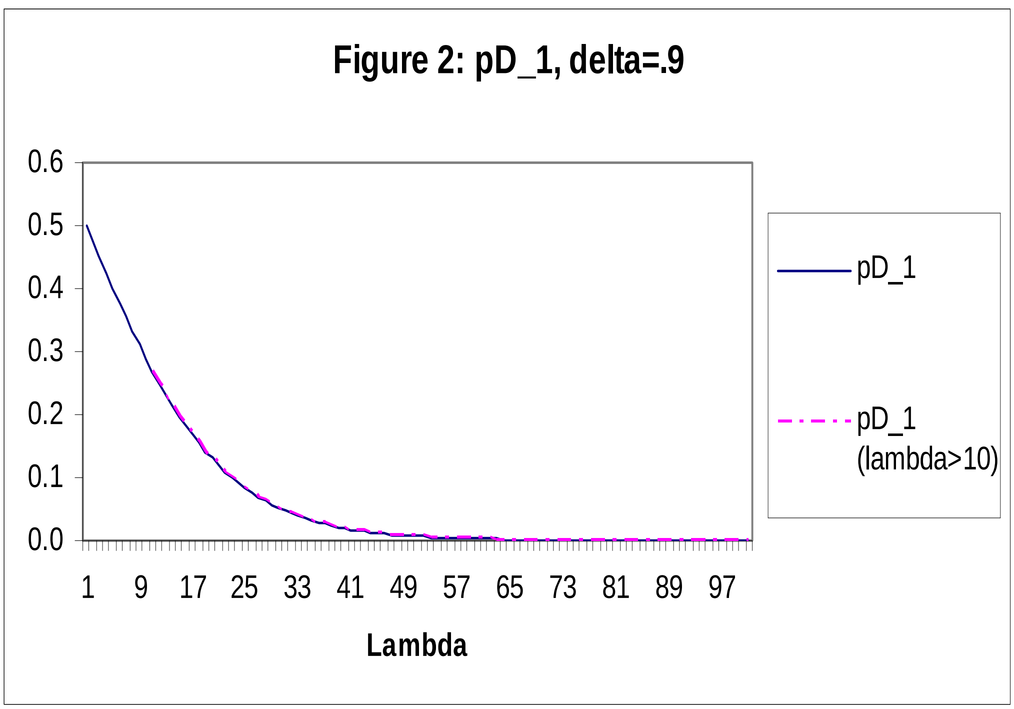
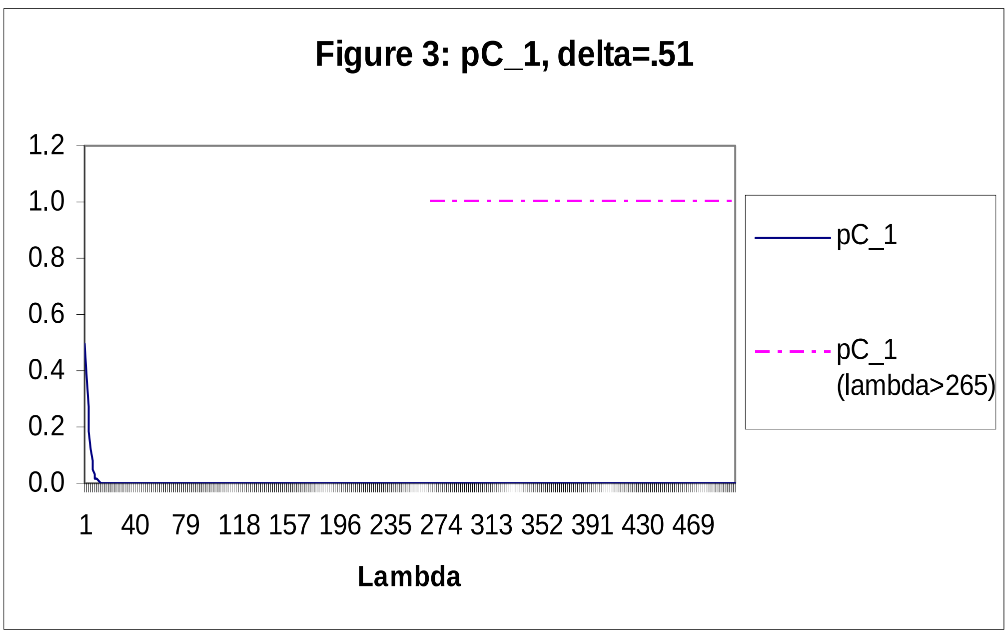
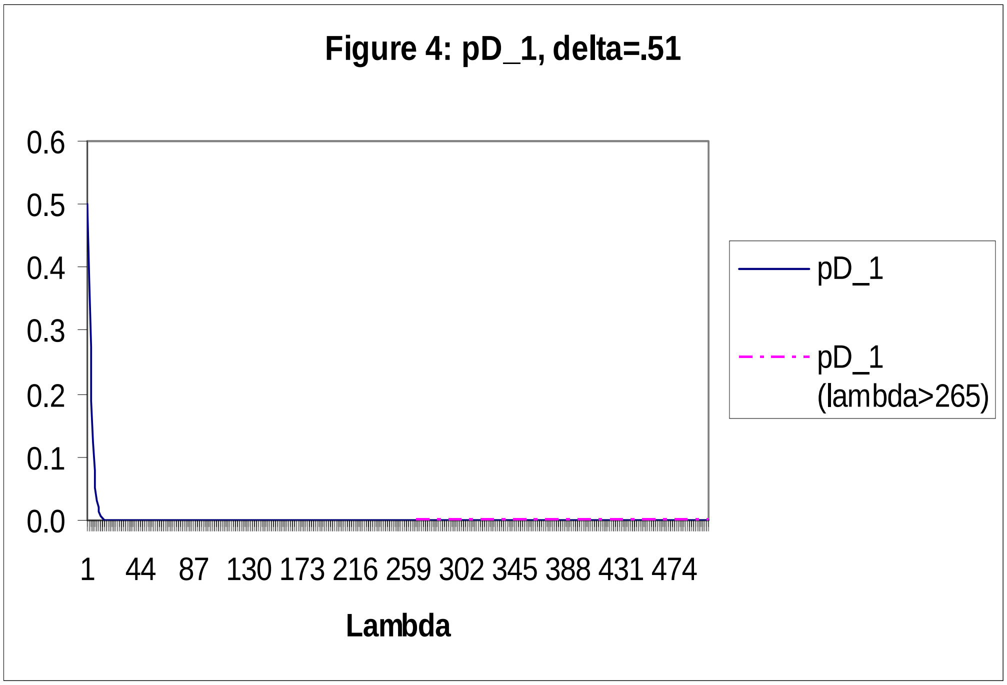


Figure 2: pD_1, delta=.9
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1
9
17
25
33
41
49
57
65
73
81
89
97
Lambda
pD_1
pD_1 (lambda>10)

As noted above, we should be able to find Quantal Response Equilibria to support the cooperative outcome so long as $\delta > t/(a+t) = .5$ here. We check the case where this condition is just met, $\delta = .51$. Figures 3 and 4, which should be read together, show equilibria that support the cooperative outcome (C, C) when $\delta = .51$. Here too we use the notation: for $(s_1^*, s_2^*) = (C, C)$ let $p_{s_1^*}^C = p_1^C$ and $p_{s_2^*}^C = p_2^C$.

In Figure 3, there are two branches for p_1^C $(= p_2^C)$. The solid, blue series begins at .5 for $\lambda = 0$ and converges to 0. For values of $\lambda \geq 265$, there is a branch of p_1^C at 1. This is the dashed, pink colored series. In Figure 4, we see p_1^D $(= p_2^D)$. The solid, blue series begins at .5 and converges to 0. The dashed, pink series begins at $\lambda \geq 265$, is exactly at 0 and is not clearly visible in the figure.





We compute Quantal Response Equilibria that support $u(s_1^*, s_2^*) = (.3, .8)$. These are shown in Figures 5 and 6 below, which should be read together. For values of $\lambda < 43$, we see a unique equilibrium. This branch starts at $(.5, .5, .5, .5)$ converges to $(0, 0, 0, 0)$. For values of $\lambda \geq 43$, there is a second equilibrium. As λ increases, this quickly converges to supporting $u(s_1^*, s_2^*) = (.3, .8)$ with probability one.

In Figures 5 and 6, the dashed lines represent player 1 and the solid lines player 2. The plots sometimes coincide and are difficult to distinguish. In figure 5, the solid, blue and dashed, pink series begin at $\lambda \geq 43$, represent the two players and support $u(s_1^*, s_2^*) = (.3, .8)$. The solid, red and dashed, dark blue series represent the other branch. In figure 6, there are four series converging to 0. Two of these begin at $\lambda = 43$. These represent the equilibrium that supports $u(s_1^*, s_2^*) = (.3, .8)$ and correspond to the solid, blue and dashed, pink lines in Figure 5.

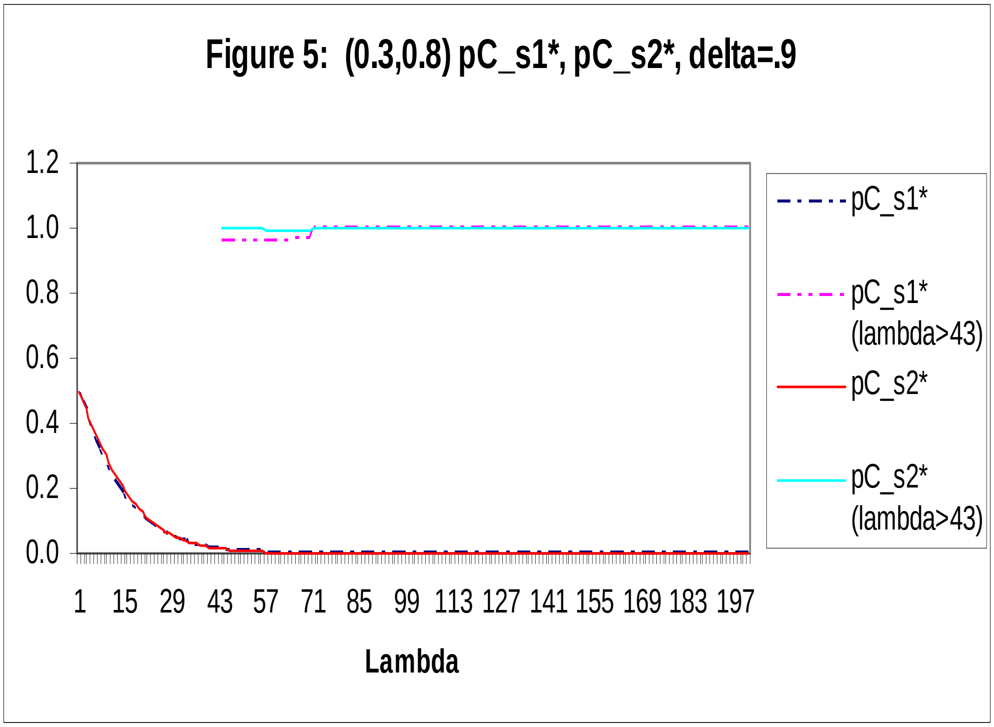


Figure 5: (0.3,0.8) pC_s1*, pC_s2*, delta=.9
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1 15 29 43 57 71 85 99 113 127 141 155 169 183 197
Lambda
pC_s1*
pC_s1* (lambda>43)
pC_s2*
pC_s2* (lambda>43)

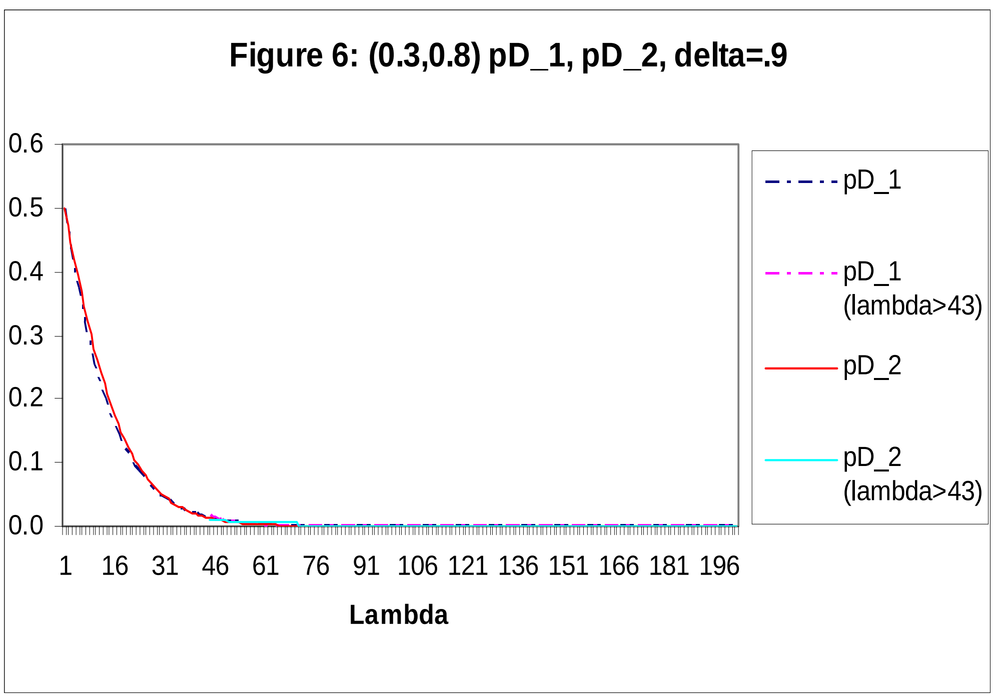
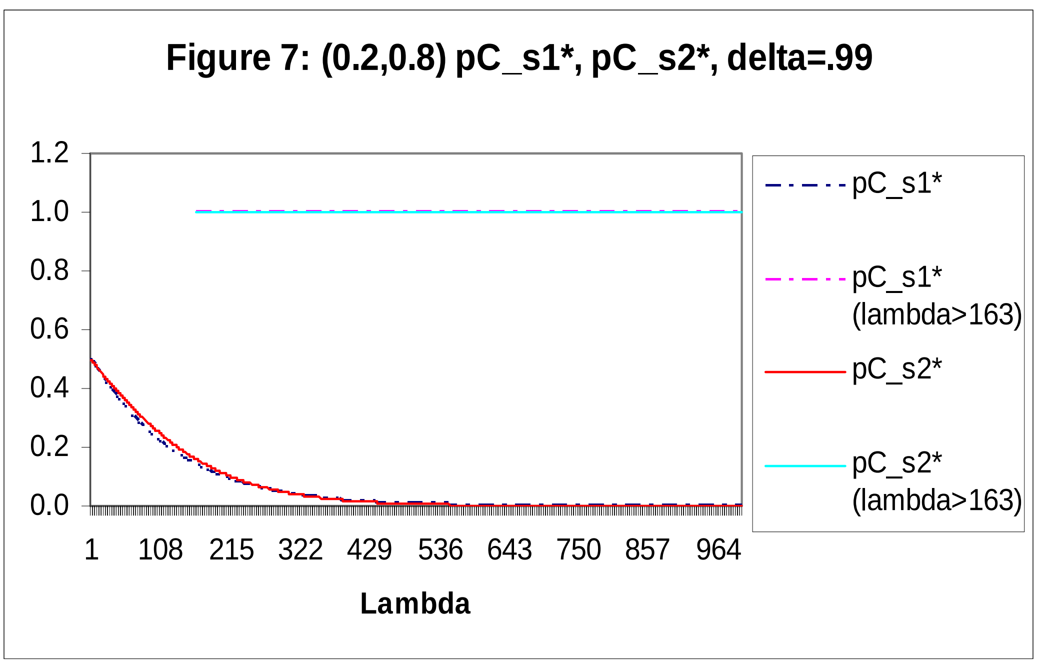
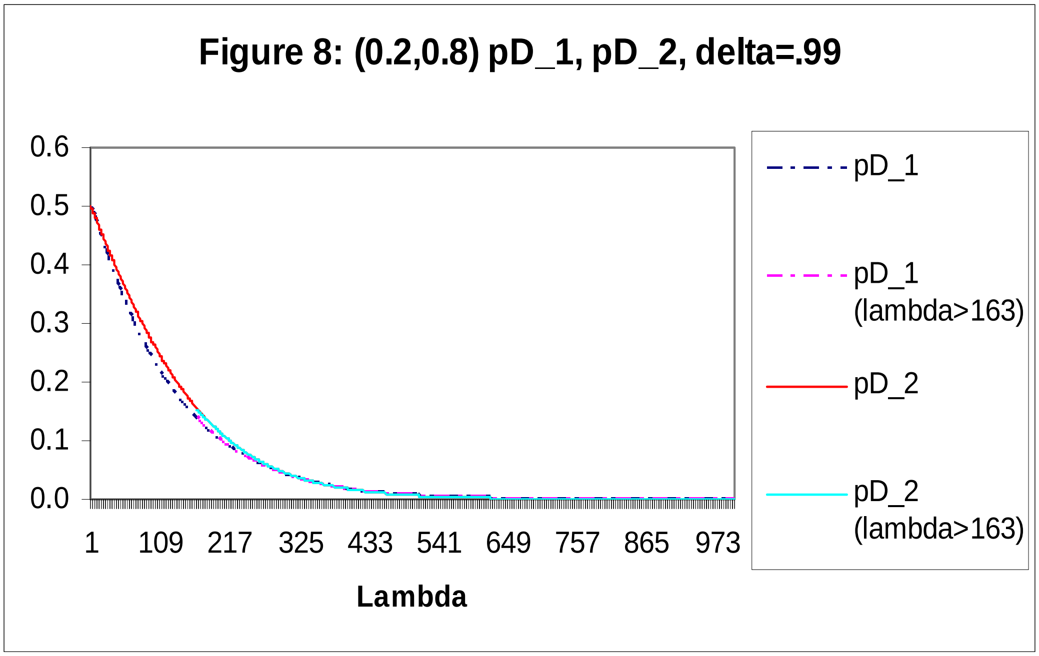


Figure 6: (0.3,0.8) pD_1, pD_2, delta=.9
0.6
0.5
0.4
0.3
0.2
0.1
0.0
1 16 31 46 61 76 91 106 121 136 151 166 181 196
Lambda
pD_1
pD_1 (lambda>43)
pD_2
pD_2 (lambda>43)

We find Quantal Response Equilibria that support: $u(s_1^*, s_2^*) = (.2, .8)$. These are shown in Figures 7 and 8 below, which should be read together. Again, for low values of λ, we see a unique equilibrium, which starts at $(.5, .5, .5, .5)$ and converges to $(0, 0, 0, 0)$. For higher values of λ, we see a second branch that supports $u(s_1^*, s_2^*) = (.2, .8)$.

In Figures 7 and 8, the dashed lines correspond to player 1 and the solid ones to player 2. The series sometimes coincide and are difficult to distinguish. In figure 7, the solid, blue and dashed, pink series begin at $\lambda \geq 163$, represent the two players and support $u(s_1^*, s_2^*) = (.2, .8)$. The solid, red and dashed, dark blue series represent the other branch. In figure 8, there are four series converging to 0. Two of these begin at $\lambda = 163$. These represent the equilibrium that supports $u(s_1^*, s_2^*) = (.2, .8)$ and correspond to the solid, blue and dashed, pink lines in Figure 7.





3.4 A limit Folk Theorem for a two player game.

We will construct Quantal Response equilibria to support expected payoffs very close to any feasible and individually rational $(v_1, v_2) = u(a_1, a_2)$ where (a_1, a_2) can be correlated strategies.

3.4.1 Minimax strategies are pure strategies

Let $m_j \in \arg\min_{s_j} \max_{s_i} u_i(s_i, s_j)$ be a minimax strategy against Player i. Let $\underline{v}_i = \max_{s_i} u_i(s_i, m_j)$. $\underline{v}_i$ is the maximum payoff player i can achieve when her opponent is playing their minimax strategies against her. Normalize $(\underline{v}_1, \underline{v}_2) = 0$. Let $b_i = \max_{s} u_i(s_i, s_j)$. First note that, following Fudenberg and Maskin(1986), given $(v_1, v_2) >$

0, we can find $\underline{\delta}$ and $\underline{n}$ such that the following conditions hold:

$$\upsilon_i > b_i(1-\underline{\delta}) + \underline{\delta}(1-\underline{\delta}^{\underline{n}})u_i(m_i, m_j) + \underline{\delta}^{\underline{n}+1}\upsilon_i \tag{3.7}$$

$$(1-\underline{\delta}^{\underline{n}})u_i(m_i, m_j) + \underline{\delta}^{\underline{n}}\upsilon_i > 0 \tag{3.8}$$

To see this, note that for $\underline{\delta}$ close enough to 1, $\upsilon_i > b_i(1-\underline{\delta})$. If (3.7) does not hold for $\underline{n} = 1$, since for large enough $\underline{\delta}$, $(1-\underline{\delta}^{\underline{n}})u_i(m_i, m_j) + \underline{\delta}_i^{\underline{n}}\upsilon_i$ is decreasing in $\underline{n}$, we can raise $\underline{n}$ until it does. Also, by picking $\underline{\delta}$ large enough, we can ensure that (3.8) is satisfied with this $\underline{n}$.

Condition (3.7) says that player i prefers υ_i to getting her best possible payoff once, $u_i(m_i, m_j)$ for $\underline{n}$ periods and υ_i forever after that. Later, in the proof of Theorem 3.1, we will use a perturbed version of (3.7) as an incentive compatibility condition (C2) that shows that under certain conditions, players do not gain from deviating from a strategy where (a_1, a_2) is played with probability close to 1. (3.8) will be useful in establishing a condition (C3) that shows that players are willing to punish, with probability close to 1, a player who deviates from (a_1, a_2).

States: For $(\upsilon_1, \upsilon_2) = u(a_1, a_2) > 0$, $\delta > \underline{\delta}$ and the corresponding $n(\delta)$, we partition the set of histories into $n+1$ states. The initial state is 0.

In State 0: If Player $i (i = 1, 2)$ plays a_i, the system stays in State 0. If either player plays $s_i \neq a_i$, the system moves to State B_n. Player i plays $s_i \in S_i$ with probability $p^0_{s_i}$.

In State $B_\varpi(\varpi = n, n-1, n-2, ..., 2)$: If Player i plays m_i move to $B_{\varpi-1}$. If either player deviates then move to B_n. Player i plays $s_i \in S_i$ with probability $p^{\varpi}_{s_i}$.

In State B_1: If Player i plays m_i move to State 0. If either player deviates then move to B_n. Player i plays $s_i \in S_i$ with probability $p^1_{s_i}$.

Continuation payoffs: Given a strategy profile p, let W_i^{ϖ} be the continuation payoff for player i in state ϖ, where $\varpi = n, n-1, n-2, ..., 2, 1, 0$.

$$\begin{aligned} W_i^0 &= p^0_{a_i} p^0_{a_j}[(1-\delta)\upsilon_i + \delta W_i^0] + (1-\delta)[p^0_{a_i} \sum_{s_j \neq a_j} p^0_{s_j} u_i(a_i, s_j) + p^0_{a_j} \sum_{s_i \neq a_i} p^0_{s_i} u_i(s_i, a_j) \\ &+ \sum_{\substack{s_i \neq a_i \\ s_j \neq a_j}} p^0_{s_i} p^0_{s_j} u_i(s_i, s_j)] + (1 - p^0_{a_i} p^0_{a_j})\delta W_i^n \end{aligned} \tag{3.9}$$

Note that:

$$\lim_{p^0_{a_i} \to 1, p^0_{a_j} \to 1} W_i^0 = \upsilon_i. \tag{3.10}$$

$$\begin{aligned} W_i^{\varpi} &= p^{\varpi}_{m_i} p^{\varpi}_{m_j}[(1-\delta)u_i(m_i, m_j) + \delta W_i^{\varpi - 1}] \qquad (3.11) \\ &+ (1-\delta)[p^{\varpi}_{m_i} \sum_{s_j \neq m_j} p^{\varpi}_{s_j} u_i(m_i, s_j) + p^{\varpi}_{m_j} \sum_{s_i \neq m_i} p^{\varpi}_{s_i} u_i(s_i, m_j) \\ &+ \sum_{\substack{s_i \neq m_i \\ s_j \neq m_j}} p^{\varpi}_{s_i} p^{\varpi}_{s_j} u_i(s_i, s_j)] + (1 - p^{\varpi}_{m_i} p^{\varpi}_{m_j})\delta W_i^n \\ \text{where } \varpi &= n, n-1, n-2, ..., 2, 1, 0. \qquad (3.12) \end{aligned}$$

Note that:

$$\begin{aligned}
\lim W_i^{\varpi} &= (1-\delta)u_i(m_i, m_j) + \delta \lim W_i^{\varpi-1} \\
&= (1-\delta^2)u_i(m_i, m_j) + \delta^2 \lim W_i^{\varpi-1} \\
&\quad \ldots\ldots\ldots\ldots \\
&\quad \ldots\ldots\ldots\ldots \\
&= (1-\delta^{\varpi})u_i(m_i, m_j) + \delta^{\varpi} \lim W_i^0 \\
&= (1-\delta^{\varpi})u_i(m_i, m_j) + \delta^{\varpi} \upsilon_i \text{ (from 3.10).}
\end{aligned} \tag{3.13}$$

where the limits are over $p_{a_i}^0, p_{a_j}^0, p_{m_i}^{\varpi}, p_{m_j}^{\varpi} \longrightarrow 1$.

Now we show that for $\delta > \underline{\delta}$ and $\varepsilon > 0$, we can find $\underline{\varepsilon}^0$ and $\underline{\varepsilon}^{\varpi}$ ($\varpi = n, n-1, n-2, ..., 1$) such that for the set of strategies which satisfy $p_{a_i}^0 = 1 - \varepsilon^0 > 1 - \underline{\varepsilon}^0$ and $p_{m_i}^{\varpi} = 1 - \varepsilon^{\varpi} > 1 - \underline{\varepsilon}^{\varpi}$, the following four conditions (C0-C3) are satisfied.

In the following conditions, the limits are over $p_{a_1}^0, p_{a_2}^0, p_{m_1}^{\varpi}, p_{m_2}^{\varpi} \longrightarrow 1$. Each of the following inequalities is strictly satisfied in the limit as $p_{a_1}^0, p_{a_2}^0, p_{m_1}^{\varpi}, p_{m_2}^{\varpi} \longrightarrow 1$. Therefore each is satisfied for small enough ε^0 and ε^{ϖ} and *all* are *simultaneously* satisfied by the minimum, $\underline{\varepsilon}^0$ and the minimum, $\underline{\varepsilon}^{\varpi}$. (Sometimes we do not need to take all these limits. See specific discussion below.)

C0. For $\varepsilon > 0$, $W_i^0 > \upsilon_i - \varepsilon$. (From 3.10, for any $\varepsilon > 0$, we can find small enough ε^0 such that $W_i^0 > \upsilon_i - \varepsilon$.)

C1. For $\varpi = n, n-1, n-2, ..., 2:$

a. $W_i^{\varpi} < W_i^{\varpi-1}$. From (3.13) since $\lim W_i^{\varpi} = (1-\delta^{\varpi})u_i(m_i, m_j) + \delta^{\varpi}\upsilon_i < (1-\delta^{\varpi-1})u_i(m_i, m_j) + \delta^{\varpi-1}\upsilon_i = \lim W_i^{\varpi-1}$. In particular, this implies that $W_i^n < W_i^{\varpi}$ for $\varpi = n-1, n-2, ..., 1$.

b. $W_i^1 < W_i^0$. Since $\lim W_i^1 = (1-\delta)u_i(m_i, m_j) + \delta\upsilon_i < \upsilon_i = \lim W_i^0$ from (3.10), (3.13) and because $\max\limits_{s_i \in S_i} u_i(s_i, m_j) = \underline{\upsilon}_i = 0$ by assumption.

C2. For all $s_i \in S_i$,

$$(1-\delta)[p_{a_j}^0(u_i(s_i, a_j) + \sum_{s_j \neq a_j} p_{s_j}^0 u_i(s_i, s_j)] + \delta p_{a_j}^0 W_i^n < (1-\delta)[\upsilon_i + \sum_{s_j \neq a_j} p_{s_j}^0 u_i(a_i, s_j)] + \delta p_{a_j}^0 W_i^0$$

or

$$(1-\delta)[p_{a_j}^0(u_i(s_i, a_j) - \upsilon_i) + \sum_{s_j \neq a_j} p_{s_j}^0 u_i(s_i, s_j) - \sum_{s_j \neq a_j} p_{s_j}^0 u_i(a_i, s_j)] + \delta p_{a_j}^0 (W_i^n - W_i^0) < 0 \quad (3.14)$$

(3.14) can be thought of as a perturbed version of Condition (3.7). It will be used in the proof of Theorem 3.1 to show that when $p_{a_j}^0$ $(j \neq i)$ in the current period and $p_{a_i}^0$, $p_{a_j}^0$, $p_{m_i}^{\varpi}$, $p_{m_j}^{\varpi}$ in the future periods are close enough to 1, $\lim\limits_{\lambda \to \infty} p_{a_i}^0 = 1$. Note that we only need $p_{a_j}^0$ $(j \neq i)$ in the current period and $p_{a_i}^0$, $p_{a_j}^0$, $p_{m_i}^{\varpi}$, $p_{m_j}^{\varpi}$ in the future periods to be close enough to 1. (3.14) is true for small ε^0 and ε^ϖ since:

$$\lim(1-\delta)[p^0_{a_j}(g_i(s_i,a_j)-\upsilon_i)+\sum_{s_j\neq a_j}p^0_{s_j}u_i(s_i,s_j)-\sum_{s_j\neq a_j}p^0_{s_j}u_i(a_i,s_j)]$$

$$+\delta p^0_{a_j}(W^n_i-W^0_i)$$

$$=(1-\delta)[u_i(s_i,a_j)-\upsilon_i]+\delta(\lim W^n_i-\lim W^0_i)$$

$$=(1-\delta)u_i(s_i,a_j)+\delta[(1-\delta^n)u_i(m_j,m_i)+\delta^n\upsilon_i]-\upsilon_i \quad \text{(from (3.10)}$$

and (3.13).)

< 0 From (3.7)

C3. For $\varpi = n, n-1, n-2, ..., 1$, for all s_i,

$$\begin{aligned}&(1-\delta)[p^\varpi_{m_j}u_i(s_i,m_j)+\sum_{s_j\neq m_j}p^\varpi_{s_j}u_i(s_i,s_j)]+p^\varpi_{m_j}\delta W^n_i\\ < \quad &(1-\delta)[p^\varpi_{m_j}u_i(m_i,m_j)+\sum_{s_j\neq m_j}p^\varpi_{s_j}u_i(m_i,s_j)]+p^\varpi_{m_j}\delta W^{\varpi-1}_i\end{aligned}$$

or

$$\begin{aligned}&(1-\delta)[p^\varpi_{m_j}(u_i(s_i,m_j)-u_i(s_i,m_j))+\sum_{s_j\neq m_j}p^\varpi_{s_j}u_i(m_i,s_j)\\ &-\sum_{s_j\neq m_j}p^\varpi_{s_j}u_i(m_i,s_j)]+p^\varpi_{m_j}\delta(W^n_i-W^{\varpi-1}_i)<0\end{aligned} \tag{3.15}$$

(3.15) will be used later to show that when $p^\varpi_{m_j}$ $(j \neq i)$ in the current period and $p^0_{a_i}$, $p^0_{a_j}$, $p^\varpi_{m_i}$, $p^\varpi_{m_j}$ in the future period are close enough to 1, $\lim_{\lambda\to\infty} p^\varpi_{m_i} = 1$. Again, note that we only need p_{m_j} $(j \neq i)$ in the current period and $p^0_{a_i}$, $p^0_{a_j}$, $p^\varpi_{m_i}$, $p^\varpi_{m_j}$ in the future period to be close enough to 1 for C3 to hold.

(3.15) is true for small for small ε^0 and ε^{ϖ} since:

$$\lim(1-\delta)[p^{\varpi}_{m_j}(u_i(s_i,m_j)-u_i(m_i,m_j))+\sum_{s_j\neq m_j}p^{\varpi}_{s_j}u_i(m_i,s_j)-\sum_{s_j\neq m_j}p^{\varpi}_{s_j}u_i(m_i,s_j)]$$

$$+p^{\varpi}_{m_j}\delta(W_i^n-W_i^{\varpi-1})$$

$$=(1-\delta)[u_i(s_i,m_j)-u_i(m_i,m_j)]+\delta(\lim W_i^n-\lim W_i^{\varpi-1})$$

$$=(1-\delta)u_i(s_i,m_j)+\delta\lim W_i^n-[(1-\delta)u_i(m_i,m_j)+\delta\lim W_i^{\varpi-1}]\text{ (from (3.13).)}$$

$$\leq(1-\delta)u_i(s_i,m_j)+\delta\lim W_i^n-[(1-\delta)u_i(m_i,m_j)+\delta\lim W_i^{n-1}]\text{ (From }\mathbf{C1.}\text{)}$$

$$=(1-\delta)u_i(s_i,m_j)+(\delta-1)[(1-\delta^n)u_i(m_i,m_j)+\delta^n\upsilon_i]\quad\text{(from (3.13).)}$$

$$\leq(\delta-1)[(1-\delta^n)u_i(m_i,m_j)+\delta^n\upsilon_i]\text{ (since }\max u_i(s_i,m_j)=\underline{\upsilon}_i=0.\text{)}$$

$$<0\quad\text{(since }(1-\delta^n)u_i(m_i,m_j)+\delta^n\upsilon_i>0\text{ from (3.8).)}$$

Theorem 3.1 *For $\varepsilon>0$ and any feasible individually rational $(\upsilon_1,\upsilon_2)=u(a_1,a_2)$, there exists $\underline{\delta}<1$ such that for all $\delta\geq\underline{\delta}$, there exists $\underline{\lambda}(\delta)>0$ with the following property. For $\delta\geq\underline{\delta}$ and $\lambda\geq\underline{\lambda}(\delta)$, there exists a Logit Quantal Response equilibrium of the infinitely repeated game $G^{\infty}(\delta)$ in which $|W_i^0-u_i(a_1,a_2)|<\varepsilon$.*

Proof. Use (3.7) and (3.8) to find $\underline{\delta}$. For $\delta>\underline{\delta}$, we are looking for a strategy profile with $p's$ that solve a system of equations: one equation for each action for each state for each player. In the analysis below, we assume that when deciding on their current action, players take as given that in all future periods, the strategy profile satisfies $p^0_{a_1},p^0_{a_2}>1-\underline{\varepsilon}^0$ and $p^{\varpi}_{m_1},p^{\varpi}_{m_2}>1-\underline{\varepsilon}^{\varpi}$. If in addition to this the opponent's current period strategy satisfies these conditions, we know that C2 and C3 are satisfied.

State 0: Let $p_i^0 = (p_{a_i}^0, \{p_{s_i}^0\}_{s_i \neq a_i})$. Assume $p_{a_j}^0 > 1 - \underline{\varepsilon}^0$. Consider the Logit QRE function:

$$p_{a_i}^0(p, \lambda) = \frac{1}{1 + \sum\limits_{s_i \in S_i} e^{\lambda(U_{s_i}^0(p_j^0) - U_{a_i}^0(p_j^0))}}$$

where $U_{a_i}^0(p_j^0)$ is player 1's payoff from $s_i = a_i$ and $U_{s_i}^0(p_j^0)$ from $s_i \neq a_i$.

$$U_{a_i}^0(p_j^0) = p_{a_j}^0((1-\delta)\upsilon_i + \delta W_i^0) + (1-\delta)\sum_{s_j \neq aj} p_{sj}^0 u_i(a_i, s_j) + (1 - p_{a_j})\delta W_i^n$$

$$U_{s_i}^0(p_j^0) = p_{a_j}^0(1-\delta)u_i(s_i, a_j) + (1-\delta)\sum_{sj \neq a_j} p_{s_j}^0 u_i(s_i, s_j) + \delta W_i^n$$

$$\begin{aligned} U_{s_i}^0(p_j^0) - U_{a_i}^0(p_j^0) &= (1-\delta)[p_{a_j}^0(u_i(s_i, a_j) - \upsilon_i) \\ &\quad + \sum_{s_j \neq a_j} p_{s_j}^0 u_i(s_i, s_j) - \sum_{sj \neq a_j} p_{sj}^0 u_i(a_i, s_j)] + \delta p_{a_j}(W_i^n - W_i^0) \\ &< 0. \text{ (From C2, since } p_{a_j}^0 > 1 - \underline{\varepsilon}^0.) \end{aligned}$$

This implies that $\lim\limits_{\lambda \to \infty} p_{a_i}^0(p, \lambda) = 1$. Since $p_{a_i}^0(p, \lambda)$ increases continuously in λ, there exists a $\underline{\lambda}_i^0$ such that $p_{a_i}^0(p, \lambda) \geq 1 - \underline{\varepsilon}^0$ for $\lambda \geq \underline{\lambda}_i^0$. Also for $p_{s_i}^0(s_i \neq a_i)$, there exists $\underline{\lambda}_{s_i}^0$ such that $p_{s_i}^0(p, \lambda) \leq \underline{\varepsilon}^0$.

Let $\underline{\lambda}^0 = \max(\underline{\lambda}_1^0, \{\underline{\lambda}_{s_1}^0\}_{s_1 \neq a_1}, \underline{\lambda}_2^0, \{\underline{\lambda}_{s_2}^0\}_{s_2 \neq a_2})$.

State B_ϖ: Let $p_i^\varpi = (\{p_{m_i}^\varpi\}, \{p_{s_i}^\varpi\}_{s_i \neq m_i})$. Assume $p_{m_j}^\varpi > 1 - \underline{\varepsilon}^\varpi$. Consider the Logit QRE function:

$$p_{m_i}^\varpi(p, \lambda) = \frac{1}{1 + \sum\limits_{s_i \in S_i} e^{\lambda(U_{s_i}^\varpi(p_j^\varpi) - U_{m_i}^\varpi(p_j^\varpi))}}$$

where $U^{\varpi}_{m_i}(p^{\varpi}_j)$ is player 1's payoff from $s_i = m_i$ and $U^{\varpi}_{s_i}(p^{\varpi}_j)$ from $s_i \neq m_i$.

$$\begin{aligned} U^{\varpi}_{m_i}(p^{\varpi}_j) &= p^{\varpi}_{mj}((1-\delta)u_i(m_i, m_j) + \delta W_i^{\varpi-1}) + (1-\delta)\sum_{s_j \neq m_j} p^{\varpi}_{s_j} u_i(m_i, s_j) \\ &\quad +(1 - p_{m_j})\delta W_i^n \\ U^{\varpi}_{s_i}(p^{\varpi}_j) &= p^{\varpi}_{m_j}(1-\delta)u_i(s_i, m_j) + (1-\delta)\sum_{s_j \neq m_j} p^{\varpi}_{s_j} u_i(s_i, s_j) + \delta W_i^n \end{aligned}$$

$$\begin{aligned} U^{\varpi}_{s_i}(p^{\varpi}_j) - U^{\varpi}_{m_i}(p^{\varpi}_j) &= (1-\delta)[p^{\varpi}_{m_j}(u_i(s_i, m_j) - u_i(m_i, m_j)) + \sum_{sj \neq m_j} p^{\varpi}_{s_j} u_i(s_i, s_j) - \\ &\quad \sum_{s_j \neq m_j} p^{\varpi}_{s_j} u_i(m_i, s_2)] + p^{\varpi}_{m_j}\delta(W_i^n - W_i^{\varpi-1}) \\ &< 0 \text{ (From C3 since } p^{\varpi}_{m_j} > 1 - \underline{\varepsilon}^{\varpi}. \text{)} \end{aligned}$$

This implies that $\underset{\lambda \to \infty}{lim} p^{\varpi}_{m_i}(p, \lambda) = 1$. Since $p^{\varpi}_{mi}(p, \lambda)$ increases continuously in λ, there exists a $\underline{\lambda}^{\varpi}_i$ such that $p^{\varpi}_{m_i}(p, \lambda) \geq 1 - \underline{\varepsilon}^{\varpi}$ for $\lambda \geq \underline{\lambda}^{\varpi}_i$. Similarly, for $p^{\varpi}_{s_i}(s_i \neq m_i)$, there exists a $\underline{\lambda}^{\varpi}_{s_i}$ such that $p^{\varpi}_{s_i}(p, \lambda) \leq \underline{\varepsilon}^{\varpi}$ for $\lambda \geq \underline{\lambda}^{\varpi}_{s_i}$.

Let $\underline{\lambda}^{\varpi} = \max(\underline{\lambda}^{\varpi}_1, \{\underline{\lambda}^{\varpi}_{s_1}\}_{s_1 \neq m_1}, \underline{\lambda}^{\varpi}_2, \{\underline{\lambda}^{\varpi}_{s_2}\}_{s_2 \neq m_2})$.

Now, let $\underline{\lambda}(\delta) = \max(\underline{\lambda}^0, \underline{\lambda}^n, \underline{\lambda}^{n-1}, ...\underline{\lambda}^1)$. Let $K^0 = |S_i| + |S_j| - 2$ if (a_1, a_2) is a pure action pair and $K^0 = |S_i| + |S_j|$ if (a_1, a_2) is a correlated action pair. Let $K = |S_i| + |S_j| - 2$.

For $\delta > \underline{\delta}$ and $\lambda \geq \underline{\lambda}(\delta)$, the Quantal response function is a continuous function from the closed and convex set $[1 - \underline{\varepsilon}^0, 1]^2 \times [0, \underline{\varepsilon}^0]^{K^0} \times [1 - \underline{\varepsilon}^n, 1]^2 \times [0, \underline{\varepsilon}^n]^K \times ... \times [1 - \underline{\varepsilon}^1, 1]^2 \times [0, \underline{\varepsilon}^1]^K$ into itself.

From the Brouwer Fixed Point Theorem, there exists a Quantal Response Equilibrium with the required properties $p^0_{a_i} \geq 1 - \underline{\varepsilon}^0$, $p^{\varpi}_{m_i} \geq 1 - \underline{\varepsilon}^{\varpi}$. From C0, $W^0_i \geq u_i(a_1, a_2) - \varepsilon$.

QED ■

3.4.2 Minimax strategies are (unobservable) mixed strategies

Now let M_i be Player i's minimax strategy against player j and let $|M_i| = m_i$. The strategies in the support of M_i are $\{a_i(k)\}_{k=1}^{m_i}$ where $q_i(k) > 0$ is the probability with which $a_i(k)$ is played in M_i.

To implement $(\upsilon_1, \upsilon_2) = u(a_1.a_2)$, we could try using the same repeated game strategies as in Section 3.4.1. In the punishment phase(s) ($B^{\varpi} = B^n, .., B^1$) both players would have to play their minimax strategies and any deviations from this would entail beginning the punishment phase all over again (B^n). But now the minimax strategies are mixed strategies. There is no way to know if a player is mixing or randomizing with the correct probabilities as specified by M_i.

Let $u_i(s_i, M_j) \equiv \sum_{k=1}^{m_j} q_j(k) u_i(s_i, a_j(k))$, the payoff to i from s_i given that j is playing according to m_j. Also, $u_j(M_i, s_j) \equiv \sum_{k=1}^{m_i} q_i(k) u_j(a_i(k), s_j)$ and $u_i(M_i, M_j) = \sum_{k=1}^{m_i} \sum_{k'=1}^{m_j} q_i(k) q_j(k') u_i(a_i(k), a_j(k'))$. Let $\underline{\upsilon}_i = \max_{s_i} u_i(s_i, M_j)$, the maximum payoff player i can achieve when her opponent is playing their minimax strategies against her. Normalize $(\underline{\upsilon}_1, \underline{\upsilon}_2) = 0$.

To deal with unobservable mixed strategies, we modify the strategies from Section 3.4.1. The indexation of $\{a_i(k)\}_{k=1}^{m_i}$ depends on the QRE parameter λ and is such that for $k = 1, ..., m_i - 1$:

$$(1-\delta)u_i(a_i(k), M_j) - \frac{1}{\lambda}\log q_i(k) \leq (1-\delta)u_i(a_i(k+1), M_j) - \frac{1}{\lambda}\log q_i(k+1) \quad (3.16)$$

A low k represents a low payoff $u_i(a_i(k), M_j)$ relative to $\frac{1}{\lambda}q_i(k)$. As will be clear below, the equilibrium strategies involve an incentive scheme where a player is punished for playing a strategy that has a high payoff compared to how often it has to be played in M_i; more precisely, a high $(1-\delta)u_i(a_i(k), M_j)$ compared with $\frac{1}{\lambda}\log q_i(k)$. (Players will of course also be punished for playing *outside* the support of their minimax strategies).

Punishment: While in the punishment phase (B^{ϖ}) if a player plays an action *outside* the support of M_i, this can be detected and leads to restarting of the punishment phase. The strategies *in* the support of M_j each has an associated *punishment*. These punishments induce players to play according to their mixed minimax strategies.

We find punishments $y_i^{\varpi}(k)$, one for each $a_i(k)$ and each ϖ in the following way. Player i is punished for playing a strategy $a_i(k)$ which has a high one-period payoff $(1-\delta)u_i(a_i(k), M_j)$ compared to $\frac{1}{\lambda}\log q_i(k)$. Strategy $a_i(1)$ which has the lowest $(1-\delta)u_i(a_i(k), M_j) - \frac{1}{\lambda}\log q_i(k)$ is not punished, or $y_i^{\varpi}(1) = 0$. For $k = 2, ..., m_i$, the punishments satisfy the equation:

$$\begin{aligned} & (1-\delta)u_i(a_i(k+1), M_j) - \frac{1}{\lambda}\log q_i(k+1) - (1-\delta)u_i(a_i(k), M_j) \quad (3.17) \\ & +\frac{1}{\lambda}\log q_i(k) \\ = \; & \delta^{\varpi}(y_i^{\varpi}(k+1) - y_i^{\varpi}(k)) \end{aligned}$$

As shown below, in the limit, all punishments go to zero.

$$\begin{aligned}\lim_{\delta\to1,\ \lambda\to\infty} y_i^{\varpi}(k+1) &= \lim_{\delta\to1,\ \lambda\to\infty} y_i^{\varpi}(k) \\ &+\frac{(1-\delta)[g_i(a_i(k+1),M_j)-g_i(a_i(k),M_j)]-\frac{1}{\lambda}\log\frac{q_i(k+1)}{q_i(k)}}{\delta^{\varpi}} \\ &= \lim_{\lambda\to\infty} y_i^{\varpi}(k)-\frac{1}{\lambda}\log\frac{q_i(k+1)}{q_i(k)} = y_i^{\varpi}(k)\end{aligned} \tag{3.18}$$

and

$$\lim_{\delta\to1,\ \lambda\to\infty} y_i^{\varpi}(k)\to 0 \text{ for all } k \text{ (since } y_i^{\varpi}(1)=0) \tag{3.19}$$

Now consider $(\upsilon_1,\upsilon_2)>0$. As in 3.4.1, we find $\underline{\underline{\delta}}$ and $\underline{n}$ such that the following conditions hold:

$$\upsilon_i > b_i(1-\underline{\underline{\delta}})+\underline{\underline{\delta}}(1-\underline{\underline{\delta}}^{\underline{n}})u_i(M_i,M_j)+\underline{\underline{\delta}}^{\underline{n}+1}\upsilon_i \tag{3.20}$$

$$(1-\underline{\underline{\delta}}^{\underline{n}})u_i(M_i,M_j)+\underline{\underline{\delta}}^{\underline{n}}\upsilon_i>0 \tag{3.21}$$

For $\delta>\underline{\underline{\delta}}$ let $n(\delta)$ be the corresponding n satisfying (3.20) and (3.21). From (3.19), we can find $\widehat{\delta}$ and $\widehat{\lambda}$ such that for $\delta>\widehat{\delta}$ and $\lambda>\widehat{\lambda}$,

$$\upsilon_i - n(\delta)\max_{\varpi,k} y_i^{\varpi}(a_i(k))>0 \tag{3.22}$$

and

$$(1-\delta^n)u_i(M_i,M_j)+\delta^n(\upsilon_i-n(\delta)\max_{\varpi,k} y_i^{\varpi}(a_i(k)))>0 \tag{3.23}$$

Let $\underline{\delta}=\max(\widehat{\delta},\underline{\underline{\delta}})$.We refer to $n(\delta)$ as simply n in the following.

States: For $s_i \in \{a_i(k)\}_{k=1}^{m_i}$ we will use $s_i \in M_i$. State 0 and its strategies are identical to Part 1. In State $B_\varpi(\varpi = n, n-1, ..., 2)$: If Player i plays $s_i \in M_i$, the state in the next period is $B_{\varpi-1}$. Let $s_i^\varpi = k$ if strategy $a_i(k)$ is played by Player i in State ϖ. If player i plays $s_i \notin M_i$ the state switches to B^n.

In State B_1: If Player i plays $s_i \in M_i$, the state in the next period is State $C(z_i, z_j)$ where $z_i = \{s_i^\varpi\}_{\varpi=1}^n$ is the vector of the s_i^ϖ's for Player i. If Player i plays $s_i \notin M_i$ the state switches to B^n.

In State $C(z_i, z_j)$: Players play $(\widetilde{a_1}, \widetilde{a_2})$ such that $u_i(\widetilde{a_1}, \widetilde{a_2}) = \upsilon_i - \sum_{\varpi=1}^{n} y_i^\varpi(s_i^\varpi)$. If player i plays $s_i \neq \widetilde{a_i}$, the state switches to B^n. There are $m_i^n \times m_j^n$ such states as the vector z_i can take m_i^n values. The total number of states is $1 + n + (m_i m_j)^n$.

Continuation payoffs: Consider W_i^C, the continuation payoff for Player i in State C.

$$\lim_{p^C_{\widetilde{a_i}},\, p^C_{\widetilde{a_j}} \to 1} W_i^C = \upsilon_i - \sum_{\varpi=1}^{n} y_i^\varpi(s_i^\varpi) \; = \widetilde{\upsilon_i} > 0 \quad \text{(From (3.22))}$$

where $\widetilde{\upsilon_i} = \upsilon_i - \sum_{\varpi=1}^{n} y_i^\varpi(s_i^\omega)$. For $\varpi = n, ...1:$

$$\begin{aligned} \lim W_i^\varpi &= \lim(1-\delta^\varpi)u_i(M_i, M_j) + \delta^\varpi(\upsilon_i - \sum_{\varpi=1}^{n} y_i^\varpi(s_i^\varpi)) \qquad (3.24) \\ &= (1-\delta^\varpi)u_i(M_i, M_j) + \delta^\varpi \widetilde{\upsilon_i} > 0 \text{ (from (3.23))} \end{aligned}$$

as $p^\varpi_{a_i(k)} \to q_i(k)$ for $s_i \in M_i$, $p^\varpi_{a_j(k)} \to q_j(k)$ for $s_j \in M_j$, $p^\varpi_{s_i}$, $p^\varpi_{s_j} \to 0$ for all $s_i \notin M_i$, $s_j \notin M_j$ and $p^C_{\widetilde{a_i}}, p^C_{\widetilde{a_j}} \to 1$.

Now we show that for $\delta > \underline{\delta}$ and $\varepsilon > 0$, we can find $\underline{\varepsilon}^0, \underline{\varepsilon}^{\varpi}, \underline{\varepsilon}^{\varpi}(k), \underline{\varepsilon}^C$ such that $0 < (q_i(k) \pm \underline{\varepsilon}^{\varpi}(k)) < 1$ and for $p^0_{a_i} > 1 - \underline{\varepsilon}^0$, $\left| p^{\varpi}_{a_i(k)} - q_i(k) \right| < \underline{\varepsilon}^{\varpi}(k)$, $p^{\varpi}_{s_i} < \underline{\varepsilon}^{\varpi}$ $(s_i \notin M_i)$ and $p^C_{\widetilde{a}_i} > 1 - \underline{\varepsilon}^C$ the following conditions (C0-C4) hold. For $i = 1, 2$:

C0. $W^0_i > \upsilon_i - \varepsilon$. As in Section 3.4.1, from 3.10, for any $\varepsilon > 0$, we can find small enough ε^0 such that $W^0_i > \upsilon_i - \varepsilon$.

C1. $W^{\varpi}_i < W^{\varpi-1}_i$ $(\varpi = 1, ..n)$.From (3.24) since $W^{\varpi}_i \to (1 - \delta^{\varpi})u_i(M_i, M_j) + \delta^{\varpi}\widetilde{\upsilon}_i$, $W^{\varpi-1}_i \to (1 - \delta^{\varpi-1})u_i(M_i, M_j) + \delta^{\varpi-1}\widetilde{\upsilon}_i$ and $\widetilde{\upsilon}_i > u_i(M_i, M_j)$.

For the next condition note that, $U^0_{s_i}(s_i \in S_i)$ is Player i's payoff from s_i in state 0.

$$U^0_{a_i} = (1 - \delta)[p^{\varpi}_{a_j}\upsilon_i + \sum_{s_j \neq a_j} p^{\varpi}_{s_j} u_i(a_i, s_j)] + p^{\varpi}_{a_i}\delta W^0_i + (1 - p^{\varpi}_{a_i})\delta W^n_i$$

$$U^0_{s_i} = (1 - \delta)\sum_{s_j} p^{\varpi}_{s_j} u_i(s_i, s_j) + \delta W^n_i \text{ where } s_i \neq a_i$$

C2. $U^0_{a_i} - U^0_{s_i} > 0$ for all $s_i \neq a_i$. This is true because:

$$\begin{aligned}
\lim_{p^0_{a_j} \to 1} U^0_{a_i} - U^0_{s_i} &= (1 - \delta)[\upsilon_i - u_i(s_i, a_j)] + \delta(W^0_i - W^n_i) \\
&= \upsilon_i - [\delta(1 - \delta^n)u_i(M_i, M_j) + \delta^{n+1}\widetilde{\upsilon}_i] - (1 - \delta)u_i(s_i, a_j) \\
&\quad \text{(from (3.24))} \\
&> 0 \qquad \text{(from (3.20))}
\end{aligned}$$

For the next condition, let $p^{\varpi}_{M_j} = \sum_{k=1}^{m_j} p^{\varpi}_{a_j(k)}$. In State ϖ, for $k = 1, 2, ...m_i$:

$$\begin{aligned}
\lim U^{\varpi}_{a_i(k)} &= (1-\delta)\sum_{s_j \in S_j} p^{\varpi}_{s_j} u_i(a_i(k), s_j) + p^{\varpi}_{M_j}\delta W^{\varpi-1} + (1-p^{\varpi}_{M_j})\delta W^n \\
&= (1-\delta)u_i(a_i(k), M_j) + \delta W^{\varpi-1} \\
\text{as } p_{a_j(k)} &\to q_j(k) \text{ for } s_j \in M_j \text{ , } p_{s_j} \to 0 \text{ for all } s_j \notin M_j.
\end{aligned}$$

C3. $\lim_{p_{a_j(k)} \to q_j(k),\, p_{s_j} \to 0} U_{a_i(k+1)} - U_{a_i(k)} = \frac{1}{\lambda} \log \frac{q_i(k+1)}{q_i(k)}$ for $k = 1, 2, ...m_i - 1$. This is true because:

$$\lim U_{a_i(k+1)} - U_{a_i(k)} = (1-\delta)(u_i(a_i(k+1), M_j) - u_i(a_i(k), M_j)) + \delta^{\varpi}(y^{\varpi}_i(k) - y^{\varpi}_i(k+1))$$

as $p_{a_j(k)} \to q_j(k)$ for $s_j \in M_j$ and $p_{s_j} \to 0$ for $s_j \notin M_j$.

$$= \frac{1}{\lambda} \log \frac{q_i(k+1)}{q_i(k)} \qquad \text{From (3.17)}$$

For the next condition note that, in State ϖ :

$$U^{\varpi}_{M_i} = (1-\delta)\sum_{s_j} p^{\varpi}_{s_j} u_i(M_i, s_j) + p^{\varpi}_{M_j}\delta W^{\varpi-1} + (1-p^{\varpi}_{M_j})\delta W^n$$

For $s_i \notin M_i$:

$$U^{\varpi}_{s_i} = (1-\delta)\sum_{s_j} p^{\varpi}_{s_j} u_i(s_i, s_j) + \delta W^n$$

C4. $U^{\varpi}_{M_i} - U^{\varpi}_{s_i} > 0$ for $s_i \notin M_i$.

$$\lim U^{\varpi}_{M_i} - U^{\varpi}_{s_i} = (1-\delta)[u_i(M_i, M_j) - u_i(s_i, M_j)] + \delta(W^{\varpi-1} - W^n)$$

as $p^{\varpi}_{a_j(k)} \to q_j(k)$ for $s_j \in M_j$ and $p^{\varpi}_{s_j} \to 0$ for all $s_j \notin M_j$.

$$\begin{aligned} &\geq (1-\delta)[u_i(M_i,M_j) - u_i(s_i,M_j)] + \delta(W^{n-1} - W^n) \text{ from C3} \\ &= (1-\delta)u_i(M_i,M_j) + \delta(W^{n-1} - W^n) \ - (1-\delta)u_i(s_i,M_j) \\ &= (1-\delta)((1-\delta^n)u_i(M_i,M_j) + \delta^n \widetilde{\upsilon_i}) - (1-\delta)u_i(s_i,M_j) \\ &> 0 \qquad \text{from (3.23) and because } u_i(s_i,M_j) \leq 0. \end{aligned}$$

Theorem 3.2 *For* $\varepsilon > 0$ *and any feasible individually rational* $(\upsilon_1, \upsilon_2) = u(a_1, a_2)$, *there exists* $\underline{\delta} < 1$ *such that for all* $\delta \geq \underline{\delta}$, *there exists* $\underline{\lambda}(\delta) > 0$ *with the following property. For* $\delta \geq \underline{\delta}$ *and* $\lambda \geq \underline{\lambda}(\delta)$, *there exists a Logit Quantal Response equilibrium of the infinitely repeated game* $G^{\infty}(\delta)$ *in which* $|W_i^0 - u_i(a_1,a_2)| < \varepsilon$.

Proof. Use (3.20) and (3.21) to find $\underline{\underline{\delta}}$. For $\delta > \underline{\delta} = \max(\underline{\underline{\delta}}, \widehat{\delta})$, we are looking for a strategy profile with $p's$ that solve a system of equations: one equation for each action for each state for each player. In the analysis below, we assume that when deciding on their current action, players take as given that in all future periods, the strategy profile satisfies $p^0_{a_i} > 1 - \underline{\varepsilon}^0$ and $\left| p^{\varpi}_{a_i(k)} - q_i(k) \right| < \underline{\varepsilon}^{\varpi}(k)$ for $k = 1, ..., m_{i.}$ For $s_i \notin M_i$, let $p^{\varpi}_{s_i} < \underline{\varepsilon}^{\omega}$ and finally, $p^{C}_{\widetilde{a}_i} > 1 - \underline{\varepsilon}^C$. If in addition to this the opponent's current period strategy satisfies these conditions, we know that C2-C4 are satisfied.

State0: Let $p^0_i = (p^0_{a_i}, \{p^0_{s_i}\}_{s_i \neq a_i})$. Assume $p^0_{a_j} > 1 - \varepsilon^0$. From the Logit quantal response function:

$$p^0_{a_i}(p, \lambda) = \frac{1}{\sum\limits_{s_i \in S_i} e^{\lambda(U^0_{s_i} - U^0_{a_i})}}$$

From C2, $U^0_{s_i} - U^0_{a_i} < 0$ for all $s_i \neq a_i$. This implies that $\underset{\lambda\to\infty}{lim}\ p^0_{a_i}(p,\lambda) = 1$. We can find $\underline{\lambda_i}^0$ such that $p^0_{a_i}(p,\lambda) \geq 1 - \underline{\varepsilon}^0$ for $\lambda \geq \underline{\lambda}^0_i$. Also for $p^0_{s_i}(s_i \neq a_i)$, there exists $\underline{\lambda}^0_{s_i}$ such that $p^0_{s_i}(p,\lambda) \leq \underline{\varepsilon}^0$.

Let $\underline{\lambda}^0 = \max(\underline{\lambda}^0_1, \{\underline{\lambda}^0_{s_1}\}_{s_1\neq a_1}, \underline{\lambda}^0_2, \{\underline{\lambda}^0_{s_2}\}_{s_2\neq a_2})$.

State B$^\varpi$: Let $p^\varpi_i = (\{p^\varpi_{a_i(k)}\}^{m_i}_{k=1}, \{p^\varpi_{s_i}\}_{s_i\notin M_i})$. Assume that Player j is playing a stage game strategy very close to M_j i.e.$\left|p^\varpi_{a_i(k)} - q_i(k)\right| < \varepsilon^\varpi(k)$.

From the Logit quantal response function:

$$\frac{p^\varpi_{a_i(k+1)}}{p^\varpi_{a_i(k)}}(p,\lambda) = e^{\lambda(U_{a_i(k+1)} - U_{a_i(k)})}$$

From C3,

$$lim \frac{p^\varpi_{a_i(k+1)}}{p^\varpi_{a_i(k)}} = \frac{q_i(k+1)}{q_i(k)} \tag{3.25}$$

as $p_{a_j(k)} \to q_j(k)$ for $s_j \in M_j$ and $p_{s_j} \to 0$ for $s_j \notin M_j$. Also from the Logit quantal response function, for $s_i \notin M_i$:

$$\frac{p^\varpi_{s_i}}{p^\varpi_{M_i}} = e^{\lambda(U_{s_i} - U_{M_i})}$$

From C4, $U_{s_i} - U_{M_i} < 0$. Therefore for $p^\varpi_{s_i}(s_i \notin M_i)$, there exists a $\underline{\lambda}^\varpi_{s_i}$ such that $p^\varpi_{s_i}(p,\lambda) \leq \underline{\varepsilon}^\varpi$ for $\lambda \geq \underline{\lambda}^\varpi_{s_i}$. Combined with (3.25) this implies that we can find $\underline{\lambda}^\varpi_k$ such that $\left|p^\varpi_{a_i(k)} - q_i(k)\right| < \underline{\varepsilon}^\varpi(k)$ for $\lambda \geq \underline{\lambda}^\varpi_k$.

Let $\underline{\lambda}^\varpi = \max(\{\underline{\lambda}^\varpi_k\}^{m_i}_{k=1}, \{\underline{\lambda}^\varpi_{s_1}\}_{s_1\notin M_1}, \{\underline{\lambda}^\varpi_k\}^{m_i}_{k=1}, \{\underline{\lambda}^\varpi_{s_2}\}_{s_2\notin M_2})$.

State C: Since $W^C_i > 0$, the implementation of $(\widetilde{a_1}, \widetilde{a_2})$ is identical to the implementation of (a_i, a_j) in State 0 with a large enough $\underline{\lambda}^C$.

Let $\underline{\lambda}(\delta) = \max(\widehat{\lambda}, \underline{\lambda}^0, \underline{\lambda}^n, \underline{\lambda}^{n-1}, ...\underline{\lambda}^1, \underline{\lambda}^C)$. Now, let $K^0 = |S_i| + |S_j| - 2$ if (a_1, a_2) is a pure action pair and $K^0 = |S_i| + |S_j|$ if (a_1, a_2) is a correlated action pair. Let $K^C = |S_i| + |S_j| - 2$ if $(\widetilde{a_1}, \widetilde{a_2})$ is a pure action pair and $K^C = |S_i| + |S_j|$ if $(\widetilde{a_1}, \widetilde{a_2})$ is a correlated action pair.

For $\delta > \underline{\delta}$ and $\lambda \geq \underline{\lambda}(\delta)$, the Quantal response function is a continuous function from the closed and convex set $[1-\underline{\varepsilon}^0, 1]^2 \times [0, \underline{\varepsilon}^0]^{K^0} \times [q_1(1)-\underline{\varepsilon}^{\varpi}(1), q_1(1)+\underline{\varepsilon}^{\varpi}(1)] \times ... \times [q_2(m_2)-\underline{\varepsilon}^{\varpi}(m_2), q_2(m_2)+\underline{\varepsilon}^{\varpi}(m_2)] \times [0, \underline{\varepsilon}^{\varpi}]^{|S_i|+|S_j|-m_i-m_j} \times [1-\underline{\varepsilon}^C, 1]^2 \times [0, \underline{\varepsilon}^C]^{K^C}$ into itself.

From the Brouwer Fixed Point Theorem, there exists a Quantal Response Equilibrium with the required properties $p^0_{a_i} \geq 1 - \underline{\varepsilon}^0$, $\left|p^{\varpi}_{a_i(k)} - q_i(k)\right| < \underline{\varepsilon}^{\varpi}$ for $k = 1, ..., m_i$, $p^{\varpi}_{s_i} < \underline{\varepsilon}^{\varpi}$ for $s_i \notin m_i$ and $p^C_{\widetilde{a}_i} > 1 - \underline{\varepsilon}^C$. From C0, $W^0_i \geq u_i(a_1, a_2) - \varepsilon$.

QED ■

3.5 Conclusion

Folk Theorems are well established in the repeated games literature when there is perfect monitoring of opponents' actions or imperfect but public monitoring (Fudenberg, Levine and Maskin, 1994). Several papers look into which equilibria of these games are robust to private monitoring imperfections. For instance, Ely and Valimaki (2002) find equilibria for the repeated Prisoner's Dilemma that are robust to private monitoring and are able to support all feasible individually rational payoffs. Mailath and Morris (2002) find sufficient conditions under which imperfect public monitoring equilibria are robust to private monitoring.

A related literature studies robustness to incomplete information in dynamic and repeated games (see Chassang and Takahashi, 2009). The Quantal Response model introduces one specific kind of additive incomplete information. This paper shows that the perfect monitoring equilibrium strategies used to prove the Folk Theorem in Fudenberg and Maskin (1986) are robust to this kind of private information. If the information is almost complete and the discount factor is close enough to 1, there exist Quantal Response Equilibria sufficiently close to the equilibria in the original construction.

References

[1] Abdulkadiroglu, A. and K. Bagwell (2005), "Trust, Reciprocity and Favors in Cooperative Relationships", manuscript May 2005.

[2] Abreu, D. (1986), "Extremal Equilibria of Oligopolistic Supergames", *Journal of Economic Theory* 39, 191-225.

[3] Abreu, D. (1988), "On the Theory of Infinitely Repeated Games with Discounting", *Econometrica* 56(2), 383-396.

[4] Abreu, D., D. Pearce and E. Stacchetti (1990), "Towards a Theory of Discounted Repeated Games with Imperfect Monitoring", *Econometrica 58(5), September 1990, 1041-1063.*

[5] Atkeson, A. and R. Lucas (1992), "On Efficient Distribution with Private Information", *Review of Economic Studies* 59, 427-453.

[6] Athey, S. and K. Bagwell (2001), "Optimal Collusion with Private Information", *RAND Journal of Economics* 32, 428-65.

[7] Battaglini, M. (2005), "Long-Term Contracting with Markovian Consumers", *American Economic Review* 95, 637-658.

[8] Battaglini M. and Palfrey T., The Dynamics of Distributive Politics, mimeo, 2007.

[9] Besley, T. and S. Coate (1995), "Group Lending, Repayment Incentives and Social Collateral", *Journal of Development Economics* 46.

[10] Chassang S. and S. Takahashi (2009), Robustness to Incomplete Information in Repeated and Dynamic Games, mimeo, 2009.

[11] Ely J. and J. Valimaki (2002), A Robust Folk Theorem for the Prisoner's Dilemma, *Journal of Economic Theory* 102, 84-105..

[12] Fudenberg D. and E. Maskin (1986), The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, Econometrica,1986.

[13] Fudenberg, D., D. Levine and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information", *Econometrica* 62, 997-1039.

[14] Green, E. (1987), "Lending and the Smoothing of Uninsurable Income", in E. Prescott and N. Wallace, (Eds.), *Contractual Arrangements for Intertemporal Trade* (Minneapolis: University of Minnesota Press).

[15] C. Harris (1985), "Existence and characterization of Perfect Equilibrium in Games of Perfect Information", *Econometrica* 53(3), 613-628.

[16] Hauser C. and H. Hopenhayn, "Trading Favors: Optimal Exchange and Forgiveness", manuscript, May 2008.

[17] Hertel, J. (2004), "Efficient and Sustainable Risk Sharing with Adverse Selection", mimeo, Princeton University.

[18] Judd, K., Yeltekin S. and Conklin J. (2003), "Computing Supergame Equilibria", *Econometrica 71(4), July 2003, 1239-1254.*

[19] Kandori, M. (1992), The Use of Information in Repeated Games with Imperfect Monitoring, *Review of Economic Studies*, 581-593.

[20] Kocherlakota, N. (1996), "Implications of Efficient Risk Sharing without Commitment", *Review of Economic Studies* 63, 595-609.

[21] Mailath G.J. and S. Morris (2002), Repeated Games with Almost-Public Monitoring, *Journal of Economic Theory* 102, 189-228.

[22] Mailath, G.J. and Samuelson, L. (2006), "Repeated Games and Reputations", *Oxford University Press.*

[23] Mobius, M. (2001), "Trading Favors", manuscript May 2001.

[24] Morduch, J. (1999), "The Microfinance Promise", *Journal of Economic Literature* 37 (4), December 1999, 1569 - 1614

[25] Azevedo, E. and Moreira, H. (2007), "On the Impossibility of an Exact Imperfect Monitoring Folk Theorem", manuscript February 2007.

[26] McKelvey and Palfrey, Quantal Response Equilibria for Normal Form Games, Games and Economic Behavior, 1995.

[27] McKelvey and Palfrey, Quantal Response Equilibria for Extensive Form Games. Experimental Economics, 1998.

[28] Stiglitz, J.E. (1990), "Peer monitoring and credit markets", *World Bank Economic Review* 4, 351-366.

[29] Thomas, J. and T. Worrall (1990), "Income Fluctuations and Asymmetric Information: As Example of the Repeated Principal-Agent Problem", *Journal of Economic Theory* 51, 367-390.

[30] Varian, H.R. (1990), "Monitoring Agents with other agents", *Journal of Institutional and Theoretical Economics* 146, 153-174.