

Reputation in Perturbed Repeated Games

Martin W. Cripps
University of Warwick

Klaus M. Schmidt
University of Bonn

Jonathan P. Thomas
University of Warwick

March 1993

ABSTRACT: The paper analyzes reputation effects in general perturbed repeated games with discounting. If there is some positive prior probability that one of the players is committed to play the same (pure or mixed) action in every period, then this provides a lower bound for her equilibrium payoff in all Nash equilibria. This bound is tight and independent of what other types have positive probability. It is generally lower than Fudenberg and Levine's (1989, 1992) bound for games with a long-run player facing a sequence of short-run opponents. The bound cannot be improved by considering types playing history dependent commitment strategies.

KEYWORDS: Reputation, Folk Theorem, repeated games, incomplete information.

JEL CLASSIFICATION NUMBERS: C73, D83, L14.

1. Introduction

This paper analyzes reputation effects in general perturbed repeated games with discounting. If there is some positive prior probability that one of the players is committed to play the same action in every period, then this provides a lower bound for her equilibrium payoff in all Nash equilibria. Considering all (pure and mixed) actions of the stage game, we derive the highest such lower bound. This bound is tight and independent of what other types have positive probability. We also show that this bound cannot be improved by considering types playing more complicated, history dependent, commitment strategies. Our bound is described and explained intuitively in the second half of this introduction.

The idea that a small amount of incomplete information in a repeated game can lead to predictions quite different from those in the corresponding complete information game has received much attention since the analysis of Kreps, Milgrom, Roberts and Wilson (1982). They showed that if the finitely repeated prisoner's dilemma is perturbed such that there is a small probability that one of the players is a type who always plays the tit-for-tat strategy then there will be cooperation in all but a bounded number of periods in any sequential equilibrium. Unfortunately, this result does depend critically on the precise form of perturbation. The Folk Theorem of Fudenberg and Maskin (1986) demonstrates that any individually rational and feasible payoff vector can be sustained as an equilibrium outcome if the modeler chooses the "right" incomplete information.

Recently, however, such "reputation" arguments have been shown to lead to sharp predictions in a more robust fashion: Aumann and Sorin (1989) allow for the existence of many different types which play like automata with bounded recall. They study repeated common interest games and show that if there exists a type for each player who always plays the cooperative action, then, no matter what other automata are around, the payoffs of all pure strategy Nash equilibria will be close to the cooperative outcome.

Fudenberg and Levine (1989,92) impose no restriction at all on what types may have positive probability, and prove a remarkably strong result for games in which a long-lived player faces a sequence of short-lived opponents. Consider the strategy to which the long-lived player would most like to publicly commit herself, i.e., the strategy that maximizes her payoff subject to the constraint that the short-lived players will choose a best response

against it. If there exists a “commitment type” who always plays this strategy and if the long-run player is sufficiently patient, then she will get at least her “commitment payoff” in any Nash equilibrium (pure or mixed) of the game. Again, this result is independent of what other types may have positive probability.

Schmidt (1993a) generalizes this result to a class of games with two long-lived players, referred to as “games of conflicting interests”. These are defined as having stage games in which the pure strategy to which, say, player one would most like to publicly commit herself holds the other player down to his minmax payoff. If there is some prior probability that player one is a type committed to this strategy, then, as player one becomes more patient, there is a lower bound on her Nash equilibrium payoff converging to the commitment payoff. Moreover it is shown that for any stage game not of this class the result does not hold, that is, there are equilibria which give player one a payoff bounded below her commitment payoff no matter how patient she is.

In the current paper we examine two questions arising from this literature. First, is there a lower bound (higher than the minmax payoff) in repeated games that are not of conflicting interests? We derive such a bound for general stage games. This bound is tight in the sense that for any given prior probability of the commitment type equilibria exist with an equilibrium payoff arbitrarily close to the bound. In games of conflicting interests the bound is equal to the commitment payoff. We also generalize Schmidt’s result by allowing for commitment types who play the same mixed strategy in each period. Thus, our results can be applied to games with moral hazard and unobservable strategies.

Secondly, we ask whether types committed to playing more complicated strategies, such as tit-for-tat, might provide a better lower bound than a simple commitment type. We allow for types playing any pure strategy that can be implemented by a finite automaton. It is shown that our lower bound cannot be improved. Hence, if a player could create some uncertainty about her type in the mind of her opponent, she could do no better than make her opponent believe that she might be a particular commitment type playing the same action each period.

In the following we describe in more detail the bound that we derive, and give some intuition for it. Our argument builds on a basic statistical result which has been estab-

lished by Fudenberg and Levine (1989, 92). If, say, player one “mimics” a commitment type by always taking the commitment action, then in equilibrium player two can expect the commitment action not to be played with more than a given probability only a bounded number of times. This number does not depend upon the particular Nash equilibrium which is being played, nor on the discount factors of the two players. It is simply implied by the fact that player two uses Bayes’ rule to update his beliefs about player one’s type. Now suppose (as Fudenberg and Levine do) that the stage game is played repeatedly by a patient player one against a sequence of short-lived players two. In each period the respective short-lived opponent will play a best response against the expected action of player one. Hence, if player one always plays the commitment action, the number of times that a short-lived player two does not play a best response must be bounded, because players two eventually become convinced that the commitment action will be played in each period. Thus, if player one is sufficiently patient she can guarantee herself a payoff very close to the commitment payoff by simply mimicking this strategy. This gives Fudenberg and Levine’s lower bound for any Nash equilibrium.

This argument does not work, however, when player two is long-lived and cares about future payoffs. The reason is that even if the opponent is convinced that the commitment action is very likely to be played next period, he will not want to play a best response to it, if by so doing his future utility is sufficiently adversely affected. The problem hinges on the distinction between *on*- and *off*-equilibrium path behaviour. It is possible to convince a long-lived opponent that *on* the equilibrium path the commitment action will continue to be played with high probability. But it is not possible for player one to signal what would be done *off* the equilibrium path. Thus an equilibrium may specify that the opponent does not play a best response to the commitment action in every period which is sustained by the possibility that player one may severely punish her opponent in the future should he ever deviate. Only if the opponent is short-lived and does not care about the future is such an equilibrium behaviour impossible.

Nevertheless, the above argument furnishes a lower bound on Nash equilibrium payoffs, though one which is generally lower than what could be attained if public commitment to the commitment action were possible. Suppose again that player one adopts the mimicking strategy. Once her opponent is convinced that the commitment action will

be pursued in the future with very high probability, then he certainly must be playing a response which gives him at least his minmax payoff against the commitment action. Because her opponent is not necessarily playing a best response as in the short-lived case, the average payoff received by player one can be lower than the commitment payoff.¹ To find a lower bound we have to consider that individually rational response by player two which gives player one the lowest payoff. Since it is always an option to mimic the commitment type, the payoff just described must be a lower bound on player one's payoff in any Nash equilibrium provided that she is patient enough to wait for the learning to take place. More precisely, we show that for any prior probability of the commitment type, and for any given discount factor of player two, the first player's equilibrium payoffs are bounded below by an amount converging to this lower bound as player one's discount factor converges to unity. This result is very robust: it is independent of the existence of types of player one other than the commitment type, and is also robust to the existence of small amounts of incomplete information about player two.

The same lower bound has been established for the case where both players evaluate payoffs according to long run averages by Cripps and Thomas (1992), using methods based on Hart's (1985) characterization of Nash equilibria of repeated games with one-sided incomplete information. In comparison very little is known about the discounted case and the case where both players are incompletely informed. The method we use here relies upon the martingale convergence arguments first established by Fudenberg and Levine (1992), in contrast to the standard martingale techniques used in the zero discounted case.

The rest of the paper is organized as follows. In Section 2 the model is described. Section 3 establishes some preliminary results on Bayesian learning in a Nash equilibrium. In Section 4 we derive the lower bound described above and discuss its robustness. There we also consider the case of two-sided uncertainty, and we discuss the role of the relative patience of the two players. Section 5 gives a few examples to illustrate our results. Finally, in Section 6 we ask whether our bound can be improved. First we show that if the commitment type plays the commitment action on and off the equilibrium path, and if this type has a non-negligible probability, then a higher bound can be established.

¹Only in games of conflicting interests must he play a best response.

However, this higher bound converges to the original bound as the perturbation becomes small, i.e., as the probability of the commitment type goes to 0. Second, we allow for more complicated commitment strategies, namely any pure strategy that can be implemented by a finite automaton. We show that the existence of types playing such strategies does not improve our bound. Section 7 concludes and outlines directions for future research.

2. The Model

The description of the model follows closely Fudenberg and Levine (1989,1992) and Schmidt (1993a). There are two players called “one” (she) and “two” (he). In every period each player selects an “action” a_i out of a finite action space A_i , $i \in \{1, 2\}$. The stage game g may have an arbitrary extensive form, so an action should be thought of as a contingent plan of moves for the stage game. A mixed action is denoted by $\alpha_i \in \mathcal{A}_i$. The payoff function of player i in the stage game is given by $g_i(a_1, a_2)$, and, in an abuse of notation, $g_i(\alpha_1, \alpha_2)$ denotes the expected payoff if the mixed action profile (α_1, α_2) is being played.

The T -fold repetition of g is denoted by G^T , where T may be finite or infinite. Our results are stated for the infinite horizon case, but all of them carry over to finitely repeated games if T is large enough. The overall payoff for player i in the repeated game from period t onwards (and including period t) is given by

$$V_i^t(\delta_i) = \sum_{\tau=t}^{\infty} \delta_i^{\tau-t} g_i(a_1^\tau, a_2^\tau), \quad (1)$$

where δ_i , $0 \leq \delta_i < 1$, denotes the respective discount factor. We follow the convention to normalize payoffs so that stage game and repeated game payoffs can be expressed on the same scale. The normalized or average payoff is

$$v_i(\delta_i) = (1 - \delta_i)V_i^1 = (1 - \delta_i) \sum_{\tau=1}^{\infty} \delta_i^{\tau-1} g_i(a_1^\tau, a_2^\tau). \quad (2)$$

The reference to δ_i will be omitted if there is no ambiguity.

Throughout the paper we restrict attention to the case where both players observe the realized action profile (a_1^t, a_2^t) after each period. Of course, if a mixed action has been played, only the outcome of the randomization process can be observed. It would

be possible to extend the analysis to the more general case where players observe a signal about each others' strategies after each period. In our case this signal is just the realized action profile. But the more general formulation also includes games of moral hazard, where the observed signal contains only statistical information about the actions that have been taken. Furthermore, even if the signal is deterministic, it need not reveal how a player would have played at an unreached information set of the stage game.² This more general case has been dealt with by Fudenberg and Levine (1992) and we refer the reader to their work for more details. At the expense of a considerable increase of notation we could have stated all our results in this more general formulation, but since the generalization follows exactly the lines of Fudenberg and Levine we chose not to do so.

Now consider a perturbation of this game such that there is some incomplete information about the payoff function of player i , $i \in \{1, 2\}$. In period 0 the “type” ω_i of player i is drawn by nature out of the countable set $\Omega_i = (\omega_i^0, \omega_i^1, \dots)$ according to the probability measure μ_i . Let ω_i^0 denote the “normal” type of player i whose payoff function is as described in the unperturbed game, i.e.

$$g_i(a_1, a_2 \mid \omega_i^0) = g_i(a_1, a_2). \quad (3)$$

We will omit the reference to ω_i^0 in the payoff functions of the normal type of each player. The other types will sometimes be called “irrational types”. They may have arbitrary payoff functions $g_i(a_1^t, a_2^t, \omega_i, t)$ which may be non-stationary. We also allow for the possibility that some of them are not expected utility maximizers but automata that follow arbitrary strategies.

Let $H^t = (A_1 \times A_2)^t$ be the set of all possible histories h^t up to and including period t . A pure strategy $s_i \in S_i$ for player i in the perturbed repeated game is a sequence of maps $s_i^t : \Omega_i \times H^{t-1} \rightarrow A_i$. Correspondingly, $\sigma_i = (\sigma_i^1, \sigma_i^2, \dots) \in \Sigma_i$ denotes a mixed (behavioral) strategy of player i .

We are particularly interested in what impact the presence of “commitment types” has on the set of equilibrium outcomes. A commitment type is an irrational type who

²For example, if a buyer chooses not to buy, then she does not learn what quality the seller would have provided had she bought.

wants to play the same (possibly mixed) action $\hat{\alpha}_i$ in every period.³ Such a type is denoted by $\omega(\hat{\alpha}_i)$.

Finally we introduce some definitions which will be used frequently. Let

$$\underline{g}_i = \min_{a_i \in A_i} \min_{a_{-i} \in A_{-i}} g_i(a_1, a_2) \quad (4)$$

be the *worst payoff* player i , normal type, can get in the stage game,

$$\bar{g}_i = \max_{a_i \in A_i} \max_{a_{-i} \in A_{-i}} g_i(a_1, a_2) \quad (5)$$

be the *best payoff* for player i , and

$$\min \max g_i = \min_{\alpha_{-i} \in A_{-i}} \max_{\alpha_i \in A_i} g_i(a_1, a_2) \quad (6)$$

be her minmax payoff.

3. Bayesian Inference and Equilibrium Payoffs

For notational simplicity we consider only the case of one-sided uncertainty in this section, i.e., player two is the normal type with probability 1. The following results are statements on how player two predicts future play and hold for any type of player two. Two-sided incomplete information is reintroduced in Section 4.

Suppose there is a type $\hat{\omega}_1 = \omega_1(\hat{\alpha}_1)$ who is committed to always play the (possibly mixed) action $\hat{\alpha}_1$.⁴ We want to show that if player one chooses to follow this strategy (which happens with a probability of at least $\mu_1(\hat{\omega}_1) > 0$ in any Nash equilibrium), then player two cannot continue to respond with a strategy which gives him less than his

³Note, that a commitment type who strictly prefers to play the same mixed action in every period cannot be an expected utility maximizer. Fudenberg und Levine (1992) have chosen a different way to formalize the idea that a player may be able to maintain a reputation for always playing a mixed action. They assume a continuum of types with non-stationary preferences. Each type prefers a particular sequence of actions over all others. A type that is committed to a mixed strategy can be generated through integration over a subset of such types. We refer the reader to Fudenberg and Levine (1992, p. 566) for technical details. Their formulation has the advantage that all types are expected utility maximizers, but it is considerably more complicated. The way in which a mixed strategy commitment type is modelled does not affect the results.

⁴Type $\hat{\omega}_1$ may be an automaton that always plays $\hat{\alpha}_1$, or it may be a type with a payoff function such that it is a dominant strategy in the repeated game to always play $\hat{\alpha}_1$. The crucial point is that $\hat{\omega}_1$ plays $\hat{\alpha}_1$ on the equilibrium path of any Nash equilibrium.

minmax payoff against $\hat{\alpha}_1$. This in turn implies a lower bound for the equilibrium payoff of the normal type of player one which will be characterized in more detail in Section 4.

In this section we establish two preliminary results. First, in Proposition 1, we show that if player two's equilibrium strategy gives him less than $\min \max g_2$ when he plays against $\hat{\omega}_1$, then he must expect that the probability distribution over outcomes generated by the repeated play of $\hat{\alpha}_1$ differs from the one generated by the "expected" equilibrium strategy of player one, where the expectation is taken over all possible types weighted with the probability measure $\mu_1(\omega_1)$. Furthermore, because player two discounts future payoffs, there must be a significant difference between these distributions in the not too distant future. The second result (Proposition 2) uses a theorem by Fudenberg and Levine (1992) to establish that if player one sticks to always playing $\hat{\alpha}_1$, then player two cannot continue to believe that the true probability distribution over outcomes is significantly different from the one generated by $\hat{\alpha}_1$. Taken together, these results imply that if player one always plays according to $\hat{\alpha}_1$, then player two cannot continue to respond with a strategy which gives him less than $\min \max g_2$. Eventually, he will learn that his opponent plays $\hat{\alpha}_1$, and he will choose a response which gives him at least his minmax payoff.

In Proposition 1 we heavily exploit the fact that player two can guarantee himself at least his minmax payoff in every Nash equilibrium. The problem is, however, that the minmax payoff is a lower bound not for the actual but only for the expected equilibrium payoff. Thus, player two could continue to play a strategy which gives him less than his minmax payoff against $\hat{\alpha}_1$ if he believes that there is a high enough probability that player one will eventually play according to some other strategy. To be more precise: It may be that the equilibrium strategy of player two yields strictly less than $\min \max g_2$ against type $\hat{\omega}_1$, as long as it yields at least $\min \max g_2$ in expectation against all types of player one, where the expectation is taken according to the beliefs of player two. However, the following proposition says that in this case the strategy of type $\hat{\omega}_1$ and the expected strategy of player one must lead to significantly different probability distributions over outcomes in the not too distant future.

The intuition for this result is simple: Given that player two discounts future payoffs, everything that happens after some finite period N is insignificant for today's expected payoff. Suppose the probability distributions over nodes in the game tree up to period N

generated by the equilibrium strategy of player two, paired with first the repeated play of $\hat{\alpha}_1$ and secondly with the expected equilibrium strategy of player one, are arbitrarily close to each other. Then the distribution over payoffs for player two would be almost the same in both cases. Thus, if he gets strictly less than $\min \max g_2$ against $\hat{\omega}_1$, he must also get less than $\min \max g_2$ against the expected equilibrium strategy of player one, a contradiction.

To express this formally consider after any history h^t the set of possible outcomes over the next N periods, that is $(A_1 \times A_2)^N$ with typical element

$$y^N = \left((a_1^{t+1}, a_2^{t+1}), \dots, (a_1^{t+N}, a_2^{t+N}) \right) , \quad (7)$$

where a_i^τ is the action played at period τ by player i . For given equilibrium strategies (σ_1, σ_2) we let $p^N(\cdot | h^t)$ be the distribution over these outcomes and likewise $p^N(\cdot | h^t, \hat{\omega}_1)$ the distribution conditional additionally upon player one's true type being $\hat{\omega}_1$ (defined for h^t having positive probability under the type conditioning). We define for any two distributions p^N and \hat{p}^N ,

$$\| p^N - \hat{p}^N \| \equiv \max_{y^N} \left| p^N(y^N) - \hat{p}^N(y^N) \right| . \quad (8)$$

Finally, define the continuation payoff for player i , normal type, discounted to period $t+1$ as:

$$v_i^{t+1}(\delta_i) = (1 - \delta_i) \sum_{\tau=t+1}^{\infty} \delta_i^{\tau-t-1} g_i(a_1^\tau, a_2^\tau) , \quad (9)$$

and its non-normalized counterpart as $V_i^{t+1}(\delta_i) = \frac{v_i^{t+1}}{1-\delta_i}$.

Proposition 1 *Let $\hat{\alpha}_1$ be any (mixed) action and suppose that there is a type $\hat{\omega}_1$ with $\mu_1(\hat{\omega}_1) > 0$ for whom it is a dominant strategy in the repeated game always to play $\hat{\alpha}_1$. Let $\delta_2 < 1$ and $\epsilon > 0$ be given and consider any Nash equilibrium and any history h^t which has positive probability in this equilibrium conditional upon $\hat{\omega}_1$. Suppose that conditional upon player one being $\hat{\omega}_1$ the expected continuation payoff for player two is*

$$E \left[v_2^{t+1} | h^t, \hat{\omega}_1 \right] \leq \min \max g_2 - \epsilon . \quad (10)$$

Then there exists a finite integer N and a number $\eta > 0$, both depending only on δ_2 and ϵ , such that

$$\| p^N(\cdot | h^t) - p^N(\cdot | h_t, \hat{\omega}_1) \| > \eta . \quad (11)$$

Proof: To simplify notation let $p^N = p^N(\cdot | h^t)$ and $\hat{p}^N = p^N(\cdot | h_t, \hat{\omega}_1)$. Choose N to be the smallest integer such that

$$\delta_2^N \frac{\bar{g}_2 - \underline{g}_2}{1 - \delta_2} < \frac{\epsilon}{2(1 - \delta_2)} . \quad (12)$$

Next, define $\tilde{V}_2^{t+1}(y^N)$ to be the payoff to the normal type of player two over the next N periods discounted to period $t + 1$, that is

$$\tilde{V}_2^{t+1}(y^N) = \sum_{\tau=t+1}^{t+N} \delta_2^{\tau-t-1} g_2(a_1^\tau, a_2^\tau) . \quad (13)$$

For probability distribution p^N its expectation is $E_{p^N} [\tilde{V}_2^{t+1}(y^N)] = \sum_{y^N} p^N \tilde{V}_2^{t+1}(y^N)$, and since this is a continuous function of p^N there exists an $\eta > 0$ such that $\| p^N - \hat{p}^N \| \leq \eta$ implies that $|E_{p^N} [\tilde{V}_2^{t+1}(y^N)] - E_{\hat{p}^N} [\tilde{V}_2^{t+1}(y^N)]| \leq \frac{\epsilon}{2(1-\delta_2)}$. Hence assume to the contrary of the proposition that $\| p^N - \hat{p}^N \| \leq \eta$; then

$$|E_{p^N} [\tilde{V}_2^{t+1}(y^N)] - E_{\hat{p}^N} [\tilde{V}_2^{t+1}(y^N)]| \leq \frac{\epsilon}{2(1 - \delta_2)} . \quad (14)$$

Hence

$$\begin{aligned} & E [V_2^{t+1} | h^t] - E [V_2^{t+1} | h^t, \hat{\omega}_1] \\ & < E_{p^N} [\tilde{V}_2^{t+1}(y^N)] + \frac{\delta_2^N \bar{g}_2}{1 - \delta_2} - E_{\hat{p}^N} [\tilde{V}_2^{t+1}(y^N)] - \frac{\delta_2^N \underline{g}_2}{1 - \delta_2} \\ & < |E_{p^N} [\tilde{V}_2^{t+1}(y^N)] - E_{\hat{p}^N} [\tilde{V}_2^{t+1}(y^N)]| + \frac{\delta_2^N (\bar{g}_2 - \underline{g}_2)}{1 - \delta_2} \\ & < \frac{\epsilon}{1 - \delta_2} , \end{aligned} \quad (15)$$

so that

$$E [v_2^{t+1} | h^t] < E [v_2^{t+1} | h^t, \hat{\omega}_1] + \epsilon . \quad (16)$$

But using (??) this implies

$$E [v_2^{t+1} | h^t] < \min \max g_2 \quad (17)$$

which is impossible.

Q.E.D.

The next result shows that if player one always plays $\hat{\alpha}_1$, then there can be only a finite number of periods in which the probability distribution over outcomes predicted by player two differs significantly from the true distribution. Eventually, player two will predict future play (almost) correctly.

Given integers N and n , with $N > 0$ and $0 \leq n < N$, define the set $T(n, N) = \{n, n + N, n + 2N, \dots\}$. Suppose that at the end of each of the periods $t \in T(n, N)$ player two makes predictions about the course of play over the following N periods. The Lemma says that if $\hat{\omega}_1$ is the true type of player one then, no matter how small $\mu_1(\hat{\omega}_1)$ and what strategies the other types of player one are supposed to play, in almost all periods player two will make predictions which are very close to the true predictions given player one's type. The proposition is a straightforward adaptation of the main theorem of Fudenberg and Levine (1992) which is stated for the case $N = 1$.

Proposition 2 (Fudenberg and Levine) *Given integers N and n , with $N > 0$ and $0 \leq n < N$, and for every $\xi > 0$, $\Delta > 0$ and a type of player one $\hat{\omega}_1$ with $\mu_1(\hat{\omega}_1) > 0$, there is a K depending only on N , ξ , Δ , and $\mu_1(\hat{\omega}_1)$ such that for any (σ_1, σ_2) , the probability, conditional on player one's true type being $\hat{\omega}_1$, that there are more than K periods $t \in T(n, N)$ with*

$$\| p^N(\cdot | h^t) - p^N(\cdot | h^t, \hat{\omega}_1) \| > \Delta \quad (18)$$

is less than ξ .

Proof: See Fudenberg and Levine (1992), Theorem 4.1.

Proposition 2 is closely related to Kalai and Lehrer's (1991) result that players will learn to play a Nash equilibrium if their priors about the other player's strategy contain a "grain of truth". In our model the grain of truth is that player two believes that, with probability $\mu_1(\hat{\omega}_1) > 0$, player one is a type who always plays $\hat{\alpha}_1$. If $\hat{\alpha}_1$ is played in every period, player two will eventually learn to predict future play correctly.

4. The Lower Bound on Equilibrium Payoffs

Consider a history of any Nash equilibrium in which player one plays $\hat{\alpha}_1$ in every period. By Proposition 1 we know that, on the one hand, player two can reply with a strategy which yields less than $\min \max g_2$ against $\hat{\alpha}_1$ only if he believes that the probability distribution over outcomes generated by player one's expected equilibrium strategy differs significantly from the one generated by $\hat{\alpha}_1$ over the next N periods. On the other hand, by Proposition 2 we know that there are at most K periods in which player two may believe that there is a significant difference if the true distribution is the one generated by $\hat{\alpha}_1$. Thus, there can be at most $K \cdot N$ periods, $K \cdot N < \infty$, in which player two chooses not to play a strategy which gives him less than $\min \max g_2$ against $\hat{\alpha}_1$.⁵

From the perspective of player one, normal type, this implies that if she sticks to always playing $\hat{\alpha}_1$, then player two must eventually restrict himself to choose strategies which yield at least $\min \max g_2$ against $\hat{\alpha}_1$. To put this more formally, let $\hat{\alpha}_1 \in \mathcal{A}_1$ be a given stage game action for player one and consider the set

$$\mathcal{M}(\hat{\alpha}_1) = \{ \alpha_2 \in \mathcal{A}_2 \mid g_2(\hat{\alpha}_1, \alpha_2) \geq \min \max g_2 \} , \quad (19)$$

i.e., the set of all mixed actions of player two which give him at least his minmax payoff against $\hat{\alpha}_1$. We now define player one's minimal commitment payoff of $\hat{\alpha}_1$ as

$$g_1^*(\hat{\alpha}_1) = \min_{\alpha_2 \in \mathcal{M}(\hat{\alpha}_1)} g_1(\hat{\alpha}_1, \alpha_2) . \quad (20)$$

Furthermore, let the worst payoff to player one if she plays $\hat{\alpha}_1$ and player two minimizes against it be denoted by

$$\underline{g}_1(\hat{\alpha}_1) = \min_{\alpha_2 \in \mathcal{A}_2} g_1(\hat{\alpha}_1, \alpha_2) . \quad (21)$$

Theorem 1 *Let $\hat{\omega}_1$ be a commitment type who always plays $\hat{\alpha}_1$, and let $\mu_1(\omega_1^0) > 0$ and $\mu_1(\hat{\omega}_1) > 0$ be given. Then, given $\delta_2 < 1$ and for any $\phi > 0$, there exists*

⁵The formal argument is a bit more involved. Since $\hat{\alpha}_1$ may be mixed we can only say that there will be more than $K \cdot N$ such periods with probability less than ξ . Furthermore, the argument is complicated by the possibility that players one and two may have different discount factors. See the proof of Theorem 1.

a $\underline{\delta}_1 < 1$ such that whenever $\underline{\delta}_1 < \delta_1 < 1$ the expected payoff of the normal type of player one in any Nash equilibrium is at least

$$E \left[v_1(\delta_1) \mid \omega_1^0 \right] > g_1^*(\hat{\alpha}_1) - \phi . \quad (22)$$

Proof: The basic idea is to show that if the normal type of player one mimics the commitment strategy $\hat{\alpha}_1$, then she will get close to $g_1^*(\hat{\alpha}_1)$, and hence her equilibrium payoff must be at least this amount. To simplify notation let $\hat{g}_1^* = g_1^*(\hat{\alpha}_1)$ and $\underline{\hat{g}}_1 = \underline{g}_1(\hat{\alpha}_1)$.

Fix $\phi > 0$. By the definition of the minimal commitment payoff, if $g_2(\hat{\alpha}_1, \alpha_2) \geq \min \max g_2$, then $g_1(\hat{\alpha}_1, \alpha_2) \geq \hat{g}_1^*$. We can choose $\epsilon > 0$ such that

$$g_2(\hat{\alpha}_1, \alpha_2) \geq \min \max g_2 - \epsilon \quad \text{implies} \quad g_1(\hat{\alpha}_1, \alpha_2) \geq \hat{g}_1^* - \frac{\phi}{3} . \quad (23)$$

Given ϵ there is an N and an η as in Proposition 1 such that (??) holds whenever (??) holds. Set $\Delta = \eta$ in Proposition 2, take any integer n , $0 \leq n < N$, and set $\xi = \frac{\phi}{3N(\hat{g}_1^* - \underline{\hat{g}}_1)}$. By Proposition 2 there is a finite K such that the probability that inequality (??) holds more than K times in $T(n, N)$ is less than ξ , so the probability that inequality (??) holds more than K times in $T(n, N)$ must also be less than ξ . Hence, considering all values for n , $0 \leq n < N$, we have that the probability that the events “ $E \left[v_2^{t+1} \mid h^t, \hat{\omega}_1 \right] \leq \min \max g_2 - \epsilon$ ” occur more than NK times given $\hat{\omega}_1$ is smaller than $N\xi$. [In other words, if player one follows the commitment strategy, the continuation payoff of player two, computed under the assumption that player one is following the commitment strategy, is unlikely to be much below the minmax payoff more than a certain number of times.]

Next, let $\hat{\alpha}_1(a_1)$ be the probability with which action a_1 is played under $\hat{\alpha}_1$, and let $\sigma_2^{t+1}(a_2 \mid h^t)$ be the random variable representing the probability with which player two plays action a_2 in period $t + 1$ after history h^t . Then

$$E \left[v_2^{t+1} \mid h^t, \hat{\omega}_1 \right] = (1 - \delta_2) E \left[\delta_2^{\tau-t-1} g_2 \left(\hat{\alpha}_1, \sigma_2^\tau(h^{\tau-1}) \right) \mid h^t, \hat{\omega}_1 \right] , \quad (24)$$

i.e., in period τ after history $h^{\tau-1}$ player two's expected payoff is generated by player one using $\hat{\alpha}_1$ and player two using $\sigma_2^\tau(h^{\tau-1})$. Player two's future expected payoff $E(v_2^{t+1} \mid h^t, \hat{\omega}_1)$ is therefore a convex combination of terms $g_2(\hat{\alpha}_1, \sigma_2^\tau)$.

Consider the convex set of payoffs (g_1, g_2) that are consistent with player one playing $\hat{\alpha}_1$ and denote this set by $F(\hat{\alpha}_1)$. Note that there cannot be a point in $F(\hat{\alpha}_1)$ with $g_1 < \hat{g}_1^*$

and $g_2 > \min \max g_2$, because \hat{g}_1^* is defined to be the smallest payoff player one can get if she plays $\hat{\alpha}_1$ and player two gets at least his minmax payoff. Therefore, it must be the case that if

$$E \left[v_2^{t+1} \mid h^t, \hat{\omega}_1 \right] \geq \min \max g_2 - \epsilon , \quad (25)$$

then

$$E \left[v_1^{t+1}(\delta_2) \mid h^t, \hat{\omega}_1 \right] \geq \hat{g}_1^* - \frac{\phi}{3} \quad (26)$$

where $v_1^{t+1}(\delta_2)$ signifies the discounted sum of player one's payoffs, calculated using δ_2 to discount instead of δ_1 . [The left hand side of the inequality (??) is just the expected value of $g_1(a_1, a_2)$ using the probability distribution just referred to; i.e. it is the right hand side of equation (??) with g_2 replaced by g_1 .] Consequently (??) can only fail more than NK times with probability less than $N\xi = \frac{\phi}{3(\hat{g}_1^* - \hat{g}_1)}$ (conditional on $\hat{\omega}_1$). Next,

$$\begin{aligned} E \left[v_1^{t+1}(\delta_2) \mid h^t, \hat{\omega}_1 \right] & \quad (27) \\ & = E \left[(1 - \delta_2)g_1(a_1^{t+1}, a_2^{t+1}) + \delta_2 v_1^{t+2}(\delta_2) \mid h^t, \hat{\omega}_1 \right] , \end{aligned}$$

so

$$(1 - \delta_2)E \left[g_1(a_1^{t+1}, a_2^{t+1}) \mid h^t, \hat{\omega}_1 \right] = E \left[v_1^{t+1}(\delta_2) - \delta_2 v_1^{t+2}(\delta_2) \mid h^t, \hat{\omega}_1 \right] . \quad (28)$$

Consequently the payoff to the normal type of player one if she follows the commitment strategy is⁶

$$\begin{aligned} E [v_1 \mid \hat{\omega}_1] & = (1 - \delta_1) \sum_{t=1}^{\infty} \delta_1^{t-1} E \left[g_1(a_1^t, a_2^t) \mid \hat{\omega}_1 \right] \quad (29) \\ & = E \left[\sum_{t=1}^{\infty} \delta_1^{t-1} \frac{1 - \delta_1}{1 - \delta_2} E \left[v_1^t(\delta_2) - \delta_2 v_1^{t+1}(\delta_2) \mid h^{t-1}, \hat{\omega}_1 \right] \mid \hat{\omega}_1 \right] \\ & = \frac{1 - \delta_1}{1 - \delta_2} \left\{ E \left[v_1^1(\delta_2) \mid \hat{\omega}_1 \right] + E \left[\sum_{t=1}^{\infty} E \left[\delta_1^{t-1} (\delta_1 - \delta_2) v_1^{t+1}(\delta_2) \mid h^{t-1}, \hat{\omega}_1 \right] \mid \hat{\omega}_1 \right] \right\} . \end{aligned}$$

Using the result on the number of times (??) holds, for $\delta_1 > \delta_2$ the random variable

$$\begin{aligned} & \sum_{t=1}^{\infty} E \left[\delta_1^{t-1} (\delta_1 - \delta_2) v_1^{t+1}(\delta_2) \mid h^{t-1}, \hat{\omega}_1 \right] \\ & \geq \left\{ \frac{\delta_1 - \delta_2}{1 - \delta_1} (\hat{g}_1^* - \frac{\phi}{3}) - (\delta_1 - \delta_2) (\hat{g}_1^* - \hat{g}_1) NK \right\} \quad (30) \end{aligned}$$

⁶Because (??) is a discounted sum, the partial sums which are always random variables converge everywhere and hence the infinite sum is measurable.

with probability at least $(1 - N\xi)$ conditional on $\hat{\omega}_1$, where we are using the fact that in the event that (??) fails no more than NK times, subtracting $(\hat{g}_1^* - \hat{g}_1)$ NK times undiscounted yields a payoff lower than the minimum possible. The random variable is at least $\frac{\delta_1 - \delta_2}{1 - \delta_1} \hat{g}_1$ otherwise.

Using this in (??) and taking the limit as $\delta_1 \rightarrow 1$ yields

$$\lim_{\delta_1 \rightarrow 1} E[v_1 | \hat{\omega}_1] \geq (1 - N\xi) \left(\hat{g}_1^* - \frac{\phi}{3} \right) + N\xi \hat{g}_1 ; \quad (31)$$

hence, since $N\xi = \frac{\phi}{3(\hat{g}_1^* - \hat{g}_1)}$, we get

$$\begin{aligned} \lim_{\delta_1 \rightarrow 1} E[v_1 | \hat{\omega}_1] &\geq \hat{g}_1^* - \frac{\phi}{3} - \frac{\phi}{3(\hat{g}_1^* - \hat{g}_1)} \left(\hat{g}_1^* - \hat{g}_1 - \frac{\phi}{3} \right) \\ &= \hat{g}_1^* - \frac{2\phi}{3} + \frac{\phi^2}{9(\hat{g}_1^* - \hat{g}_1)} \\ &> \hat{g}_1^* - \frac{2\phi}{3} . \end{aligned} \quad (32)$$

Choosing $\underline{\delta}_1$ such that the left hand side of (??) is within $\frac{\phi}{3}$ of its limit, we have for $\delta_1 \geq \underline{\delta}_1$

$$E[v_1 | \hat{\omega}_1] \geq \hat{g}_1^* - \phi . \quad (33)$$

Consequently, by mimicking type $\hat{\omega}_1$, player one is guaranteed a payoff of $\hat{g}_1^* - \phi$. The result follows. *Q.E.D.*

Let us now come back to the case of two-sided uncertainty. Propositions 1 and 2 are statements on how player two predicts future play and are independent of player two's actual type. If there are several types of player two, then player two will make his predictions conditional on his true type. Except for this the statements of the propositions are the same.

However, the lower bound for player one's equilibrium payoff is affected by the possibility of other types. By Theorem 1 we know that if player one faces the normal type of player two, then she will get at least $g_1^*(\hat{\alpha}_1) - \phi$ in any Nash equilibrium. This happens with probability $\mu_2(\omega_2^0)$. If she plays against some other type, which happens with probability $1 - \mu_2(\omega_2^0)$, the worst that can happen to her if she sticks to $\hat{\alpha}_1$ is that she

gets $\underline{g}_1(\hat{\alpha}_1)$ in every period. Thus, we can give the following lower bound for her expected payoff:⁷

Corollary 1 *Let $\mu_i(\omega_i^0) > 0$, $i \in \{1, 2\}$, and $\mu_1(\hat{\omega}_1) > 0$ be given. Then, given $\delta_2 < 1$, and for any $\phi > 0$ there exists a $\underline{\delta}_1 < 1$ such that whenever $\underline{\delta}_1 < \delta_1 < 1$ the expected payoff of the normal type of player one in any Nash equilibrium is at least*

$$E \left[v_1(\delta_1) \mid \omega_1^0 \right] > \left(1 - \mu_2(\omega_2^0) \right) \underline{g}_1(\hat{\alpha}_1) + \mu_2(\omega_2^0) (g_1^*(\hat{\alpha}_1) - \phi) . \quad (34)$$

Up to this point we have taken the strategy $\hat{\alpha}_1$ which could be mimicked by the normal type of player one as given. Let us now ask: What is the strategy which maximizes the lower bound of player one's equilibrium payoff? We will pose this question for “small” perturbations of the informational structure, i.e. with probability $1 - \epsilon$ each player believes that the other player is normal. But there is a small probability $\epsilon > 0$ that the opponent has a different payoff function compared to the unperturbed game. Note that we do not restrict the set of possible types. For small perturbations of the informational structure the following theorem gives the best lower bound for player one's equilibrium payoff.

Theorem 2 *For any given $\delta_2 < 1$ and $\psi > 0$ there exists an $\epsilon > 0$, such that for all ϵ -perturbations of G that satisfy $\mu_i(\omega_i^0) \geq 1 - \epsilon$ and $\mu_1(\omega_1^*) > 0$ the following holds: There exists a $\underline{\delta}_1 < 1$, such that if $\underline{\delta}_1 < \delta_1 < 1$, then the expected payoff of the normal type of player one in any Nash equilibrium of the ϵ -perturbed game is at least*

$$E \left[v_1(\delta_1) \mid \omega_1^0 \right] > g_1^* - \psi , \quad (35)$$

where

$$g_1^* = \sup_{\alpha_1 \in \mathcal{A}_1} \min_{\alpha_2 \in \mathcal{M}(\alpha_1)} g_1(\alpha_1, \alpha_2) \quad (36)$$

is the best lower bound from mimicking a commitment type ω_1^* who always plays the same action in every period.

⁷We can also establish the following trivially: Suppose player two has two “normal types”. Then compute the lower bound given each type of player two and take the weighted average according to their prior probabilities. This average is the lower bound for player one's expected equilibrium payoff. The argument extends to any finite number of types.

Proof: Fix $\psi > 0$ and let α_1^* be an action such that

$$g_1^*(\alpha_1^*) \equiv \min_{\alpha_2 \in \mathcal{M}(\alpha_1^*)} g_1(\alpha_1^*, \alpha_2) \geq g_1^* - \frac{\psi}{2}. \quad (37)$$

Let $\omega_1^* = \omega_1(\alpha_1^*)$ and suppose that $\mu_1(\omega_1^*) > 0$. By Corollary 1 we know that for any $\delta_2 < 1$ and $\phi > 0$ there exists a $\underline{\delta}_1 < 1$ such that if $\underline{\delta}_1 < \delta_1 < 1$ player one, normal type, can guarantee himself at least

$$E(v_1(\delta_1) \mid \omega_1^0) > (1 - \mu_2(\omega_2^0)) \underline{g}_1(\alpha_1^*) + \mu_2(\omega_2^0) (g_1^*(\alpha_1^*) - \phi). \quad (38)$$

Suppose that $\underline{g}_1(\alpha_1^*) < g_1^*$, otherwise the lower bound is trivial. Choose $\phi > 0$ and $\epsilon > 0$ such that

$$\frac{\psi}{2} > \epsilon (g_1^* - \underline{g}_1(\alpha_1^*)) + (1 - \epsilon)\phi. \quad (39)$$

Using $\mu_2(\omega_2^0) > 1 - \epsilon$ we get

$$\begin{aligned} E(v_1(\delta_1) \mid \omega_1^0) &> \epsilon \underline{g}_1(\alpha_1^*) + (1 - \epsilon) (g_1^*(\alpha_1^*) - \phi) \\ &> \epsilon \underline{g}_1(\alpha_1^*) + (1 - \epsilon) (g_1^* - \phi) - \frac{\psi}{2} \\ &= g_1^* - \epsilon (g_1^* - \underline{g}_1(\alpha_1^*)) - (1 - \epsilon)\phi - \frac{\psi}{2} \\ &> g_1^* - \psi. \end{aligned} \quad (40)$$

Q.E.D.

This lower bound on equilibrium payoffs has several important properties:

- It holds for all Nash equilibria, no matter whether they are in pure or in mixed strategies, whether they are perfect Bayesian or whether they satisfy any other refinement. Even more generally, it holds for all self-confirming equilibria. A self-confirming equilibrium is a weaker notion than Nash equilibrium because it does not require that a player correctly predicts the other player's play off the equilibrium path. However, in our context the set of self-confirming equilibrium outcomes and the set of Nash equilibrium outcomes coincides,⁸ so we can stick to the more familiar Nash concept.

⁸See Fudenberg and Levine (1993) for a definition and characterization of self-confirming equilibria.

- The results do not depend on the assumption that the game is infinitely repeated. All our arguments go through for a finite horizon if there are enough repetitions.
- The lower bound offered above is independent of what other types of player one and two may exist with positive probability. The earlier literature on repeated games with incomplete information assumed that there is only one “crazy” type that may have positive probability. We find it implausible to a priori restrict the set of possible types. Our approach rather follows Fudenberg and Levine (1989) in allowing for arbitrarily many different types with arbitrary payoff functions. Thus, our result does not hinge delicately on the specific informational assumptions imposed by the modeler but is robust against further perturbations of the informational structure.

It is important to note that in Theorems 1 and 2 the discount factor of player two is fixed while δ_1 has to be chosen close enough to 1. Put differently, player one has to be patient enough as compared to player two. To see why this is the case suppose the discount factor of player two increases. Since he is more patient, he cares more about future payoffs. Thus, he may continue for a longer period of time to play a strategy which gives him less than $\min \max g_2$ against type $\hat{\omega}_1$, even if he expects an equilibrium strategy of player one which is very close to always playing $\hat{\alpha}_1$. That is, the number N in Proposition 1 increases. Thus, if player one wants to mimic $\hat{\omega}_1$, she has to wait for a longer period until she can be sure that player two will respond with a strategy which gives player two at least $\min \max g_2$. To get the same lower bound for player one’s average equilibrium payoff she has to be sufficiently more patient.

In Theorem 1 we do not give an explicit formula for $\underline{\delta}_1$ as a function of δ_2 , so we cannot use this result to characterize how much more patient player one has to be than player two in order to obtain the lower bound. However, in an earlier draft of this paper we considered the special case where $\hat{\alpha}_1$ is restricted to be a pure strategy. In this case we can compute $\underline{\delta}_1(\delta_2)$ and show that

$$\lim_{\delta_2 \rightarrow 1} \frac{1 - \underline{\delta}_1(\delta_2)}{1 - \delta_2} = 0, \tag{41}$$

i.e., in the limit as player two’s discount factor approaches 1, player one is infinitely more patient than player two.

In economic applications it is typically difficult to justify why two players face different discount rates. This is why we would like to offer a different interpretation:⁹ Suppose player one is a “big” player who plays the same game against L “small” opponents, denoted by $2_1, \dots, 2_L$. Each of the small players 2_j is long-lived and has the same discount factor as the “big” player one. Divide each period in L subperiods and suppose that player one plays sequentially against the small player 2_j in subperiod j , $j = 1, \dots, L$. Suppose that each of the small players can observe the play not only in his own game but also in all other games player one is playing against the other players 2_k , $k \neq j$. It is easy to see that increasing the number of small players is equivalent to increasing the discount factor of player one.¹⁰ Thus, we can say that player one can exploit reputation effects to increase his equilibrium payoff if he is sufficiently big as compared to his opponents.

How useful is our lower bound? If we are interested in games where the informational structure is only slightly perturbed, i.e., games where $\mu_i(\omega_i^0)$ is close to 1, then the lower bound offered in (??) is in general higher than player one’s minmax payoff because $\min \max g_2$ is derived from minimizing player one’s payoff over all $\alpha_2 \in \mathcal{A}_2$, while g_1^* is obtained from minimizing over the smaller set $\mathcal{M}(\alpha_1^*)$, i.e. the set of all strategies which give player two at least his minmax payoff. Thus, our lower bound can be used to restrict the set of Nash equilibrium outcomes as compared to the prediction of the Folk theorem. This will be illustrated with some examples in Section 5.

The second question that arises is whether player one can guarantee himself more than the lower bound given in (??) by mimicking any other type who does not follow a constant but a more complicated, possibly history dependent strategy. In section 6 we will show that if we are looking for a lower bound for all Nash equilibria and if we do not impose any restriction on what other types may have positive probabilities, then the lower bound derived in (??) is indeed the best we can get even if a large class of other commitment types using arbitrarily complex history dependent strategies are considered.

⁹We are grateful to Nabil Al Najjar for suggesting this interpretation.

¹⁰The argument still works (but is slightly more complicated) if player one plays against all L opponents at the same time.

5. Examples

We now consider some examples to illustrate the power and the limitations of our results and to relate them to the literature. Let us start with the well known “Battle of the Sexes” game depicted in Figure 1.

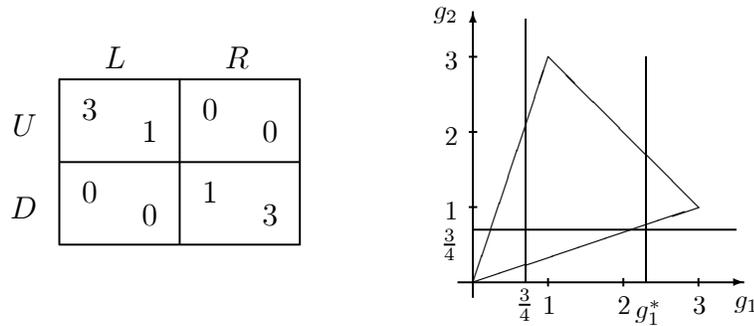


Figure 1 – The Battle of the Sexes

Player one’s minmax strategy is to play U with probability $\frac{3}{4}$ and D with probability $\frac{1}{4}$ which holds player two down to a payoff of $\frac{3}{4}$. Thus, player two must get at least that much in expected terms in any Nash equilibrium. Furthermore, the Folk theorem predicts that any feasible payoff vector that gives each player at least $\frac{3}{4}$ can be sustained as an equilibrium outcome.

Now suppose that the informational structure is perturbed such that with positive probability there exists a type of player one for whom it is a dominant strategy to always play T . In order to get at least his minmax payoff against this type, player two must play L at least with probability $\frac{3}{4}$. Thus, if player one mimics this type and if she is patient enough as compared to player two, then she will get at least a payoff of $\frac{9}{4}$ in any Nash equilibrium of the repeated game. This considerably reduces the set of Nash equilibrium payoffs as compared to the prediction of the Folk theorem.¹¹

The Battle of the Sexes game is completely symmetric. Thus, if there is a type of player two who always plays R , then, if player two is patient enough as compared to

¹¹For a more detailed discussion of the Folk theorem, and in which sense our lower bound is in contrast to its prediction, see Schmidt (1993a, ??f).

player one, he can also guarantee himself at least $\frac{9}{4}$. Note, however, that the payoff vector $(\frac{9}{4}, \frac{9}{4})$ is not in the feasible set. That is, it is impossible that in equilibrium both players build up a reputation for always playing their most preferred commitment strategy. This shows, that the outcome must depend on the relative patience of the two players. If player one is sufficiently more patient than player two, or, if she has sufficiently more at stake (e.g. because she plays against many player twos simultaneously), then she can use the reputation effect to her advantage and vice versa.

However, even if player one is arbitrarily more patient than player two, i.e. if $\frac{1-\delta_1}{1-\delta_2}$ is arbitrarily close to 0, we are not back to the case of Fudenberg and Levine where a long-run player faces a sequence of short-lived opponents. In the example of the Battle of the Sexes Fudenberg and Levine's result implies that if player one sticks to always playing U , then her short-lived opponents will eventually play a best response, i.e. they will play L with probability 1. This gives player one her Stackelberg payoff of 3. In contrast, if player two is long-lived it is not true that player two has to play a best response against U even if he believes that U is certain to be played in every period along the equilibrium path. As discussed in the introduction, the reason is that player two may be afraid that he will be punished off the equilibrium path if he deviates now and plays a best response against U . This is why we can establish only a smaller lower bound, $\frac{9}{4}$, which is derived from the fact that even if player two does not play a best response he must get at least his minmax payoff.

Our lower bound is most powerful if the game is of “conflicting interests” in the sense of Schmidt (1993a), i.e., if the strategy to which player one would most like to commit herself holds the other player down to his minmax payoff. In this case player two must eventually play a best response against player one's commitment strategy in order to get at least $\min \max g_2$. This is why for games of conflicting interests we get the same lower bound as Fudenberg and Levine. Examples of such games include the Chain-Store game and the Game of Chicken. We generalize Schmidt's results in two respects. First of all, our theory, which contains his results as a special case, applies to all games, not only to games of conflicting interests.¹² Secondly, Schmidt (1993a) considered only pure strategy

¹²Schmidt (1993a, p.20) suggests a different and much weaker generalization of his results. Suppose that the game is not of conflicting interests but that there is a type of player one who is committed to hold player two down to his minmax payoff. Eventually player two must play a best response against this

commitment types while we also allow for mixed strategies.

The next example shows that it may be worthwhile for a player to mimic a mixed strategy. Consider the stage game depicted in Figure 2. Suppose player one mimics a type who plays U in every period. Eventually player two must respond R with probability 1 in order to get at least his minmax payoff of 0. This gives player one a payoff of 1.

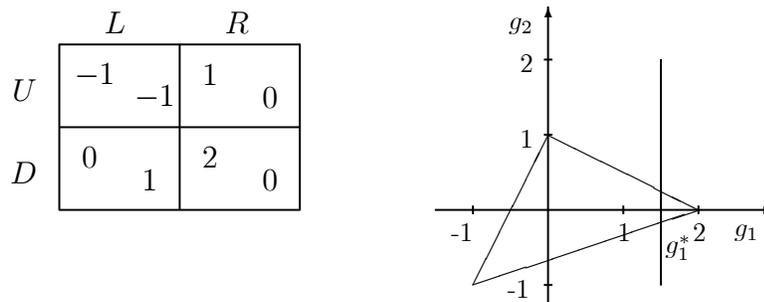


Figure 2 – Commitment to a Mixed Strategy.

However, player one can guarantee herself more if she can mimic mixed strategy types. Suppose there is a type who plays U with probability $\frac{1}{2} + \epsilon$ and D with probability $\frac{1}{2} - \epsilon$ in every period. Again, in order to get at least his minmax payoff player two must respond to this strategy by playing R with probability 1. But this gives player one an expected payoff of $\frac{3}{2} - \epsilon$.

Let us finally give an example where our lower bound is of no use. Consider the Prisoners' Dilemma game of Figure 3. No matter which strategy is mimicked by player one, player two can always guarantee himself his minmax payoff by playing D , so player one cannot guarantee herself more than 1, which is his minmax payoff. Thus, our lower bound does not restrict the set of Nash equilibrium payoffs as compared to the prediction of the Folk theorem.

strategy. Thus, mimicking this commitment type guarantees player one a payoff which is lower than his Stackelberg payoff but which may still be higher than his minmax payoff. Note that we allow for this kind of commitment types as well, so our best lower bound must be at least as high as the one of Schmidt. Note further, that in many games Schmidt's lower bound is of no use because it is strictly lower than the minmax payoff. For example, in the Battle of the Sexes game Schmidt's commitment type would play U with probability $\frac{3}{4}$ and D with probability $\frac{1}{4}$. The best response of player two which is worst for player one is to always play R which gives player one a payoff of $\frac{1}{4}$ (less than her minmax payoff), so the lower bound is trivially satisfied.

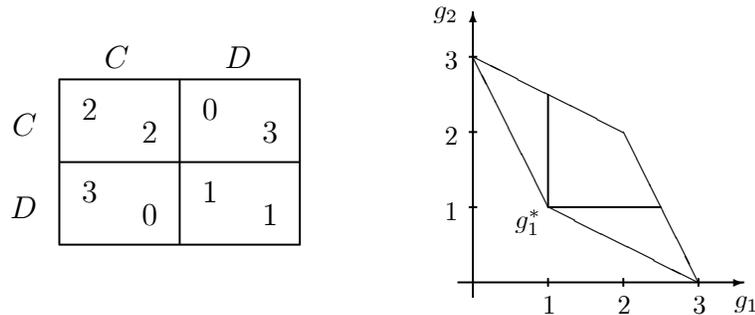


Figure 3 – The Prisoner’s Dilemma.

This is in interesting contrast to the approach of Watson (1993) who shows that if the infinitely repeated Prisoners’ Dilemma is perturbed only with strategies of bounded recall, including the “tit-for-tat” strategy, then the payoffs associated with the set of pure strategy equilibria is close to the cooperative, efficient payoff vector. Aumann and Sorin (1989) have a similar result for common interest games. The difference between Watson’s approach and ours is not that his allows for the possibility of mimicking a “tit-for-tat” type (i.e. a type that uses a history dependent commitment strategy) and our approach does not. We will consider such kind of commitment strategies in the next section and show that they do not improve our lower bound. The crucial difference is that Watson restricts attention to pure strategy equilibria, while we consider the set of *all* Nash equilibria.

6. Improvements of the Lower Bound

In this section we ask whether it is possible to improve the lower bound derived in Section 4. First we show that the lower bound can be improved if we consider perturbations of the informational structure in which a particular commitment type has a non-negligible probability (Proposition 3). Then we come back to small perturbations of the informational structure and consider the possibility that player one may mimic not a constant strategy but a more complicated history dependent strategy. In Theorem 3 we show that the lower bound of Section 4 is tight even if we allow for commitment types playing arbitrarily complicated history dependent strategies, more precisely, any pure strategy which

can be implemented by a finite automaton.

6.1. Commitment Types with Non-negligible Probabilities

Theorems 1 and 2 are based on considerations relating to behaviour on the equilibrium path: Provided the commitment strategy is followed on the equilibrium path by some type of player 1, then the bound follows, and it does not matter what such a player might do should player two not adhere to his equilibrium strategy. In general though, off equilibrium path considerations lead to a tighter bound if there is a type following $\hat{\alpha}_1$ whatever happens. Intuitively, the punishment that can be imposed on player two is less severe since with a certain probability player one will play $\hat{\alpha}_1$ rather than a minmax strategy. This implies that once player two becomes convinced that $\hat{\alpha}_1$ will be played on the equilibrium path in the future, his continuation payoff must be generally greater than the minmax payoff. This rules out certain long-run responses to $\hat{\alpha}_1$ and hence raises the payoff from playing $\hat{\alpha}_1$.

We consider the case where the commitment strategy is a pure strategy $\hat{\alpha}_1$. Let $\hat{\omega}_1$ be a type for whom it is a dominant strategy to always play this action in the repeated game and let $\hat{\mu}_1 \equiv \mu_1(\hat{\omega}_1)$. We define player two's best response payoff as

$$\bar{g}_2(\hat{\alpha}_1) = \max_{a_2 \in A_2} g_2(\hat{\alpha}_1, a_2) . \quad (42)$$

Next define

$$\begin{aligned} \tilde{B}(\hat{\alpha}_1, \delta_2, \hat{\mu}_1) = & \quad (43) \\ \{ \alpha_2 \in A_2 \mid g_2(\hat{\alpha}_1, \alpha_2) \geq \hat{\mu}_1 \bar{g}_2(\hat{\alpha}_1) + (1 - \hat{\mu}_1) \left((1 - \delta_2) \underline{g}_2 + \delta_2 \min \max g_2 \right) \} \end{aligned}$$

and

$$\tilde{g}_1(\hat{\alpha}_1, \delta_2, \hat{\mu}_1) = \min_{\alpha_2 \in \tilde{B}(\hat{\alpha}_1, \delta_2, \hat{\mu}_1)} g_1(\hat{\alpha}_1, \alpha_2) . \quad (44)$$

Note that $\tilde{g}_1(\hat{\alpha}_1, \delta_2, \hat{\mu}_1) \geq g_1^*(\hat{\alpha}_1)$ because the minimum is taken over a smaller set.

Proposition 3 *Let $\mu_i(\omega_i^0) > 0$, $i \in \{1, 2\}$ and $\hat{\mu}_1 \equiv \mu_1(\hat{\omega}_1) > 0$ be given, where $\hat{\omega}_1$ is a type playing $\hat{\alpha}_1$ after any history. Then, given $\delta_2 < 1$ and for any $\phi > 0$, there exists a $\underline{\delta}_1 < 1$ such that whenever $\underline{\delta}_1 < \delta_1 < 1$ the expected payoff of the normal type of player one in any Nash equilibrium is at least*

$$\left(1 - \mu_2(\omega_2^0)\right) \underline{g}_1(\hat{\alpha}_1) + \mu_2(\omega_2^0) (\tilde{g}_1(\hat{\alpha}_1, \delta_2, \hat{\mu}_1) - \phi) . \quad (45)$$

Proof: Consider, after any history h_t which occurs with positive probability when player one plays \hat{a}_1 every period, the least continuation payoff that the normal type of player two can receive in equilibrium. He always has the option of doing the following: play a best response against \hat{a}_1 so long as \hat{a}_1 has been played in the past, and as soon as \hat{a}_1 is not played, play a strategy which guarantees $\min \max g_2$ thereafter. This strategy yields player two a payoff at least equal to

$$\mu_1(\hat{\omega}_1)\hat{g}_2(\hat{a}_1) + (1 - \mu_1(\hat{\omega}_1)) \left((1 - \delta_2)\underline{g}_2 + \delta_2 \min \max g_2 \right) \quad (46)$$

since the worst that can happen is that he is not facing type $\hat{\omega}_1$ and player one does not play \hat{a}_1 in period t and proceeds to minmax him thereafter, while against type $\hat{\omega}_1$ his payoff is $\hat{g}_2(\hat{a}_1)$; player two's subjective beliefs after h^t attach probability at least $\hat{\mu}_1$ to type $\hat{\omega}_1$ since this probability cannot fall as long as \hat{a}_1 has been followed. The proofs of Proposition 1 and Theorem 1 can be repeated with (??) replacing $\min \max g_2$, and the result follows. *Q.E.D.*

For example in the Battle of the Sexes game with $\mu_2(\omega_2^0) = 1$, δ_2 close to 1, and $\mu_1(\hat{\omega}_1) = \frac{1}{2}$, we get a lower bound approaching $\frac{21}{8}$. As $\mu_1(\hat{\omega}_1)$ approaches 1, the expression in (??) approaches $\hat{g}_2(\hat{a}_1)$, the best response payoff of player two, and consequently $\tilde{g}_1(a_1, \delta_2, \hat{\mu}_1)$ approaches the Fudenberg-Levine commitment payoff.

Note that Proposition 3 is useful only if one of the players assigns a considerable probability to the event that his opponent is committed to one particular strategy. In most applications, however, we are interested in small perturbations, where there may be many different types which have a very small prior probability. We will consider this case in the rest of this section.

6.2. History Dependent Commitment Strategies

Up to now we focused on commitment types that played a constant action in every period. The question arises whether player one can guarantee herself a higher lower bound than g_1^* (as defined in Theorem 2) if she mimics a more complicated commitment type, i.e. a type who follows a history dependent strategy. In this section we consider pure history dependent strategies that can be implemented by a finite automaton, and we define g_1^* now as the bound from mimicking the best *pure* strategy commitment type. We shall see

that the answer to the above question is negative: commitment types cannot be improved upon by such automata. In other words, if the normal type of player one could choose a type (to be given positive prior probability by player two), then she could do no better than choose a commitment type who plays the same action in every period. As a by-product of this analysis the tightness of our bound will be established: there are Nash equilibria arbitrarily close to the bound.

We start by defining what we mean by a finite automaton. This is a machine (formally a “Moore” machine) which has a finite number M of states represented by the set X , and to each state $x \in X$ is associated an action $a_1 \in A_1$, and there is a transition rule which associates with each state and action of player two (x, a_2) a new state for the automaton. At period one the automaton starts in some initial state x_0 . The pure strategy implemented by this automaton will be denoted \hat{s}_1 .¹³

The following theorem shows that it is not possible to improve on g_1^* by mimicking any finite automaton who plays a strategy \hat{s}_1 . That is, even if very complicated commitment types exist, it is always possible to construct a Nash equilibrium in which player one, normal type, gets at most $g_1^* + \epsilon$, where $\epsilon > 0$ may be arbitrarily small.¹⁴

Theorem 3 *Fix any pure strategy \hat{s}_1 that can be implemented by a finite automaton. For any $\epsilon > 0$ there exists a perturbation of G in which $\omega(\hat{s}_1)$ has strictly positive probability, there exists a $\bar{\delta}$ with $0 < \bar{\delta} < 1$, and for any $\bar{\delta} < \delta_1, \delta_2 < 1$ there exists a Nash equilibrium of this perturbed game such that the normal type of player one gets an average equilibrium payoff which is smaller than $g_1^* + \epsilon$.*

Proof: See Appendix.

Note that Theorem 3 implies that our bound is tight: If we consider the automaton

¹³The automaton does not specify what action will be taken after histories in which the automaton has not followed its own strategy. For Nash equilibria however this is not important and arbitrary actions can be specified after such histories.

¹⁴We conjecture that this result extends to infinite automata. However, we have been able to show this only for the case of no discounting. To establish our result for history dependent *mixed* strategies would be hopelessly complicated because we would have to construct a mixed strategy equilibrium for a game with incomplete information.

which always plays a_1^* , then the theorem tells us that there is an equilibrium of a perturbed game in which $\omega_1(a_1^*)$ has positive probability such that the equilibrium payoff for player one is arbitrarily close to g_1^* .

The proof of the theorem is complicated and relegated to the appendix. In the rest of this section we try to give some intuition for it. We have to find a perturbation of G in which $\omega_1(\hat{s}_1)$ has positive probability and a Nash equilibrium of this perturbed game in which the normal type of player one gets less than $g_1^* + \epsilon$. Consider the most simple such perturbation in which player one is either the normal type (with probability $1 - \mu$) or the commitment type (with probability μ), while there is no uncertainty about the type of player two. Depending on the strategy \hat{s}_1 we can construct at least one of two types of Nash equilibria with the desired property:

- *A pooling equilibrium:* Suppose there exists a pure strategy \hat{s}_2 of player two such that (\hat{s}_1, \hat{s}_2) yields a continuation payoff (after any period t) strictly higher than $\min \max g_2$ to player two and strictly in between $\min \max g_1$ and $g_1^* + \epsilon$ to player one. In this case the following strategies constitute a Nash equilibrium if both players are patient enough: The normal and the commitment type of player one both play \hat{s}_1 and player two responds with \hat{s}_2 . Any deviation is minmaxed by the other player. By construction, the normal type of player one gets less than $g_1^* + \epsilon$ in this equilibrium.
- *A separating equilibrium:* Suppose no such strategy for player two exists. In this case any strategy \hat{s}_2 of player two will at some point either hold player one weakly below her minmax payoff, or give player two weakly less than his minmax payoff (or both). Consider a strategy \hat{s}_2^* which is a best response against \hat{s}_1 . If we are in the former subcase we can construct an equilibrium as follows: Player one, normal type, follows \hat{s}_1 up to the point where her expected continuation payoff against \hat{s}_2^* no longer exceeds her minmax payoff. At this point she reveals herself by deviating from \hat{s}_1 and is rewarded in the continuation equilibrium with a payoff strictly above her minmax payoff but strictly below $g_1^* + \epsilon$. Player two also gets strictly more than $\min \max g_2$ in this continuation equilibrium. He follows \hat{s}_2^* up to the period where the normal type of player one is supposed to reveal herself. If he observes a deviation from \hat{s}_1 at this point, he plays the above mentioned continuation equilibrium. If

\hat{s}_1 was played in this period he continues with \hat{s}_2^* . Note that he is playing a best response in this subgame. Off the equilibrium path deviations are minmaxed. Again, it is easy to see that if players are sufficiently patient these strategies constitute a Nash equilibrium in which player one gets less than $g_1^* + \epsilon$.

Finally, suppose that (\hat{s}_1, \hat{s}_2^*) gives player two weakly less than her minmax payoff. In this case, because \hat{s}_2^* is a best response against \hat{s}_1 , \hat{s}_2^* is simply any short-run best response to the automaton's action period by period since this guarantees player two his minmax payoff each period. But then, by the definition of g_1^* , (\hat{s}_1, \hat{s}_2^*) cannot give player one more than g_1^* . Again, we can construct a separating equilibrium, but now the normal type of player one is rewarded with $g_1^* + \epsilon$ if she reveals herself.

One technical twist in these arguments is that continuation payoffs after any period t have to be bounded away from $\min \max g_i$; otherwise the threat of mutual minmaxing is not sufficient to deter a deviation. This is why we have to restrict attention to finite automata.

7. Conclusions

If a player has the possibility to acquire a reputation for being a commitment type who always takes the same action in every period, and if she is sufficiently patient as compared to her opponent, then this yields a lower bound for her payoff in any Nash equilibrium. This bound is robust against further perturbations of the informational structure. In games that are not of “conflicting interests” our bound is weaker than the one derived by Fudenberg and Levine (1989, 92) for games in which a long-lived player faces a sequence of short-lived opponents. However, our bound is tight and still useful to reduce the set of equilibrium payoffs predicted by the Folk theorem.

Our last theorem shows that it is not possible to find a better lower bound for all Nash equilibria. However, it is an important question of future research whether a better lower bound can be obtained if the equilibrium notion is refined and/or if the set of possible types is restricted. Very few first steps in this direction have been made. We mentioned the work of Aumann and Sorin (1989) and Watson (1993) who restrict attention to pure strategy

Nash equilibria. Cripps and Thomas (1993) consider the set of perfect equilibrium payoffs in a perturbed common interest game. In this game perfection reduces the set of Nash equilibrium outcomes, but very limitedly so. Unfortunately, even in this extremely simple game the necessary computations turn out to be very complicated. Schmidt (1993b) characterizes the set of all sequential equilibria satisfying a weak Markov property in a finitely repeated bargaining game with incomplete information about the seller's cost. He exploits the special structure of the set of possible types to derive a sharp prediction of the equilibrium payoffs for both players. Surprisingly, his results hold independent of the relative patience of the two players. More work in this direction needs to be done.

Appendix

Proof of Theorem 3: We will use the following perturbation of G in which player one has two possible types. With probability $1 - \mu$ she is the normal type, with probability μ she is a finite automaton that follows the pure strategy \hat{s}_1 . There is no uncertainty about the type of player two.

The automaton that implements \hat{s}_1 has a finite number of states. There must exist at least one subset of states which can be reached starting from x_0 such that once reached, the automaton must stay within this subset, and such that from each state within this set any other state can be reached. Denote such a subset by Z . Notice that there must be a sequence of actions by player two $(a_2^1, a_2^2, \dots, a_2^{\hat{t}})$ that will “steer” the automaton into a state $\hat{x} \in Z$ at time $\hat{t} + 1$ within $\hat{t} \leq M - 1$ periods. Let $\hat{h}^{\hat{t}}$ be the corresponding history: $\hat{h}^{\hat{t}} = ((\hat{s}_1(h^0), a_2^1), (\hat{s}_1(\hat{s}_1(h^0), a_2^1), a_2^2), \dots)$; this will prove very useful below. If $x_0 \in Z$ then $\hat{t} = 0$ is possible, but it will be convenient if we define in this case $\hat{t} = 1$ and choose some arbitrary action for player two a_2^1 so that $\hat{h}^{\hat{t}}$ is not empty.

For our purposes the automaton can be characterized in terms of the set of long run average payoffs which can be attained starting in a state $x \in Z$ given that player one is following the automaton strategy and that player two is following an arbitrary pure strategy \hat{s}_2 (this clearly does not depend upon which state in Z is the initial state). Define

$$P = \left\{ \lim_{T \rightarrow \infty} \sum_{t=\hat{t}}^T \frac{(g_1(a_1^t, a_2^t), g_2(a_1^t, a_2^t))}{(T - \hat{t})} \middle| \begin{array}{l} \text{player one follows } \hat{s}_1, \text{ player two follows any } s_2 \\ \text{such that the limit exists} \end{array} \right\}.$$

Let $V(\delta_1, \delta_2) = \{(v_1(\delta_1), v_2(\delta_2))\}$ for some $0 \leq \delta_1, \delta_2 < 1$.¹⁵ In order to avoid tedious notation we assume that players can use correlated strategies. Thus, V is always convex. The set of feasible payoffs in the stage game is a subset of V for all δ_1, δ_2 . Thus, clearly $P \subseteq V$.

Let g_1^* be the highest bound from a pure strategy commitment type (as defined in Theorem 2 but where the maximum is taken over pure strategies only). Suppose that some $\epsilon > 0$ is given. Our aim is to show that with the given automaton we can find critical values $1 > \bar{\mu} > 0, 0 < \bar{\delta} < 1$ such that whenever $0 < \mu < \bar{\mu}$ and $1 > \delta_1, \delta_2 > \bar{\delta}$, an equilibrium can be found which gives player one less than $g_1^* + \epsilon$. Hence there will be a positive value for μ such that, for fixed $\delta_2 > \bar{\delta}$, no matter how patient player one becomes relative to player two, our previous bound cannot be improved upon.

Without loss of generality we impose the following assumption:

Assumption 1 *For all $0 \leq \delta_1, \delta_2 < 1$ there exist $(v_1(\delta_1), v_2(\delta_2))$ such that $v_1(\delta_1) > g_1^*$ and $v_2(\delta_2) > \min \max g_2$.*

¹⁵Even though the set V always depends on δ_1, δ_2 the reference to discount factors will sometimes be omitted if there is no ambiguity.

Note that Assumption 1 implies that there exists a strictly individually rational payoff vector in the stage game. If this assumption is not satisfied for some (δ_1, δ_2) a simple argument shows that player one cannot get more than $g_1^* + \epsilon$ in any Nash equilibrium: Two cases have to be distinguished:

- (i) For some $0 \leq \delta_2 < 1$ there does not exist a $v_2(\delta_2) > \min \max g_2$. In this case player two has to play a best response to player one's action in every period in order to get at least $\min \max g_2$. But, by the definition of the lower bound, this cannot give player one a payoff higher than g_1^* .
- (ii) For some $0 \leq (\delta_1, \delta_2) < 1$ there exists a $(v_1(\delta_1), v_2(\delta_2))$ with $v_2(\delta_2) > \min \max g_2$, but there does not exist a point $(\tilde{v}_1(\delta_1), \tilde{v}_2(\delta_2))$ with $\tilde{v}_1(\delta_1) > g_1^*$ and $\tilde{v}_2(\delta_2) > \min \max g_2$. Denote the set of such (δ_1, δ_2) by Δ . Note that there must exist a point (g_1, g_2) in the set of feasible stage game payoffs with $g_2 > \min \max g_2$. Fix such a point. Recall that $(g_1, g_2) \in V(\delta_1, \delta_2)$ for all $0 \leq \delta_1, \delta_2 < 1$, and that $V(\delta_1, \delta_2)$ is convex. Consider the line L from (g_1, g_2) through $(g_1^*, \min \max g_2)$. Now suppose that for some $(\delta_1, \delta_2) \in \Delta$ there is a point $(\hat{v}_1(\delta_1), \hat{v}_2(\delta_2))$ such that $\hat{v}_1(\delta_1) > g_1^* + \epsilon$ and this point lies strictly above line L . Then there must also exist a point $(\hat{\hat{v}}_1(\delta_1), \hat{\hat{v}}_2(\delta_2))$ with $\hat{\hat{v}}_1(\delta_1) > g_1^*$ and $\hat{\hat{v}}_2(\delta_2) > \min \max g_2$, a contradiction. Thus, we conclude that for all $(\delta_1, \delta_2) \in \Delta$ and for all $(v_1(\delta_1), v_2(\delta_2)) \in V(\delta_1, \delta_2)$, $v_1(\delta_1) > g_1^* + \epsilon$ implies $v_2(\delta_2) < \min \max g_2 - \alpha$ for some $\alpha > 0$.

If player two faces the automaton type of player one, the best that can happen to him is that he gets \bar{g}_2 in every period. Thus we have

$$\begin{aligned} v_2(\delta_2) &\leq \mu \bar{g}_2 + (1 - \mu) v_2(\delta_2 \mid \omega_1^0) \\ &< \eta + v_2(\delta_2 \mid \omega_1^0) \end{aligned} \tag{47}$$

By choosing the perturbation to be small, we can make $\eta = \mu \bar{g}_2$ arbitrarily small. Choose $\eta = \alpha$. Thus, player one can get more than $g_1^* + \epsilon$ only if player two gets an expected payoff strictly smaller than his minmax payoff, a contradiction.

The following lemma describes some properties of the limiting set P which we shall need.

Lemma 1 (i) P is a convex set; (ii) For any $\eta > 0$, any initial state $x \in Z$, and any point $(\hat{v}_1, \hat{v}_2) \in P$, there exists $\tilde{\delta} < 1$ such that for $\delta_1, \delta_2 > \tilde{\delta}$ there exists a strategy for player two against the automaton, such that discounted continuation payoffs are always within η of (\hat{v}_1, \hat{v}_2) ; (iii) There is a point $(\tilde{v}_1, \tilde{v}_2) \in P$ with $\tilde{v}_1 \leq g_1^*$ and $\tilde{v}_2 \geq \min \max g_2$.

Convexity follows straightforwardly from the ability to switch between states together with zero discounting. The second property says that any point in P can be approximated by a point in the discounted set as discounting goes to zero. The third property follows

because player two can play a strategy against each action of player one which mimics the response played in the definition of the lower bound.¹⁶

The concept of a best response against the automaton will also prove important below. If on the equilibrium path at any stage of the game player two becomes convinced that he faces the automaton type in state x , then his continuation strategy must be a best response against the automaton strategy, and there always exists such a best response which, after at most $M - 1$ periods, leads to an outcome path which cycles with periodicity equal to some integer K where $K \leq M$ (this follows immediately from dynamic programming). We refer in this case to *a best response starting from x with cycle K* .

Depending upon the set P there are three cases to consider.

Case 1 *There exists $(\hat{v}_1, \hat{v}_2) \in P$ with $\min \max g_1 < \hat{v}_1 < g_1^* + \epsilon/2, \hat{v}_2 > \min \max g_2$.*

In this case it is possible to construct a pooling equilibrium in which player one follows the automaton strategy, player two plays a strategy corresponding to (\hat{v}_1, \hat{v}_2) , and player one is minmaxed should she reveal herself. Consider the following strategies for the normal types:

Equilibrium path. Player one follows the automaton strategy. Player two plays according to $\hat{h}^{\hat{t}}$ and from time $\hat{t} + 1$ onwards plays as in Lemma 2(ii) so that payoffs are within η of (\hat{v}_1, \hat{v}_2) , where η is defined below.

Off equilibrium path. Any deviation is minmaxed by the other player.

“Off equilibrium path” means a history which has unconditional probability zero. In this case, both types of player one follow the same pure strategy, and so any deviation from this is punished by player two, who plays to minmax *the normal type* of player one. Likewise the normal type of player one minmaxes player two should he deviate. Any deviation by player two will result in an expected continuation payoff of at most $(1 - \mu) \min \max g_2 + \mu \bar{g}_2$ since with probability $(1 - \mu)$ he will be facing the normal type who will play a minmax strategy. Choose $0 < \delta'_2 < 1, \eta' > 0$ and $\bar{\mu} > 0$ to satisfy

$$(1 - \delta'_2) \bar{g}_2 + \delta'_2 ((1 - \bar{\mu}) \min \max g_2 + \bar{\mu} \bar{g}_2) \leq (1 - (\delta'_2)^{M-1}) \underline{g}_2 + (\delta'_2)^{M-1} (\hat{v}_2 - \eta')$$

where the L.H.S. is an upper bound on the deviation payoff and the R.H.S. is a lower bound on the payoff from holding to the equilibrium strategy, given that deviation might occur as early as the first period, \underline{g}_2 may be received along $\hat{h}^{\hat{t}}$, and $\hat{t} \leq M - 1$. Likewise a δ''_1 and an η'' can be found such that a corresponding inequality holds for player one (who is minmaxed with probability one after deviation). Choose $\eta = \min\{\eta', \eta''\}$ and letting $\tilde{\delta}$ be as in Lemma 2(ii), set $\bar{\delta} = \max\{\tilde{\delta}, \delta'_2, \delta''_1\}$. Hence for $\bar{\delta} < \delta_1, \delta_2 < 1, 0 < \mu < \bar{\mu}$ the above strategies are feasible and constitute a Nash equilibrium.

¹⁶The formal proof of this lemma is purely technical and available from the authors upon request.

Case 2 *There exists $(\hat{v}_1, \hat{v}_2) \in P$ with $\min \max g_1 \geq \hat{v}_1$ and $\hat{v}_2 > \min \max g_2$, but no $(v_1, v_2) \in P$ with $v_1 > \min \max g_1, v_2 > \min \max g_2$.*

In this case a separating equilibrium will be constructed. Consider a best response against the automaton starting in a state $x \in Z$ which leads to a cycle of no more than M periods. For high discount factors the payoffs generated by this must approximately lie in P - see the corollary below - and hence in Case 2 offer player one no more than her minmax payoff. This means that player one can be induced to reveal her type provided she is rewarded by a strictly individually rational payoff: if she mimics the automaton at this point she would convince player two that he is playing against the automaton; player two would therefore play a best-response which would give player one a lower payoff than she would get by revealing her type.

Lemma 2 *There exists $1 > \hat{\delta}_2 > 0$ such that starting in any $x \in Z$, and for $1 > \delta_2 > \hat{\delta}_2$, there exists a best response against the automaton strategy with cycle no more than M which coincides with a best response in the zero discounting case, that is, which yields a payoff to player two of $\max_{(v_1, v_2) \in P} v_2$.*

Proof of Lemma 2: Consider a best response with cycle $K \leq M$ starting from $x \in Z$, and let the payoffs to player two in the cycle be $g_2^1, g_2^2, \dots, g_2^K$. Discounted payoffs starting from the first point in the cycle are

$$\frac{g_2^1 + \delta_2 g_2^2 + \dots + \delta_2^{K-1} g_2^K}{\sum_{i=1}^K \delta_2^{i-1}}$$

which converges to the average payoff $(1/K) \sum_{n=1}^K g_2^n$ as $\delta_2 \rightarrow 1$. Since there are only a finite number of such cycles, above a critical discount factor $\bar{\delta}_2$ a best response cycle in the non discounting case must also be a best response cycle in the discounting case; it thus corresponds to a point in P such that v_2 is maximised. Q.E.D.

Corollary 2 *As $\delta_2 \rightarrow 1$, all continuation payoffs of player two from a best-response against the automaton strategy at any $x \in Z$ converge to $\arg \max_{(v_1, v_2) \in P} v_2$.*

Choose $(\hat{v}_1, \hat{v}_2) \in V(\delta_1, \delta_2)$, $0 \leq \delta_1, \delta_2 < 1$, such that $\min \max g_1 < \hat{v}_1 < g_1^* + \epsilon/2$, $\hat{v}_2 > \min \max g_2$ (this is always possible by Case 2, Assumption 1, and the convexity of V). From the corollary and by definition of Case 2 there exists a $0 < \tilde{\delta} < 1$ such that for $1 > \delta_1, \delta_2 > \tilde{\delta}$ there are best response payoffs (v_1^*, v_2^*) (depending on δ_1, δ_2) starting from \hat{h}^t satisfying $v_1^* < \hat{v}_1 - \alpha$, for some $\alpha > 0$.

Consider the following strategies:

Equilibrium path. Both players follow \hat{h}^t for the first $\hat{t} - 1$ periods. At time \hat{t} player one normal type reveals his type by playing some $a_1^{\hat{t}} \neq \hat{s}_1(\hat{h}^{\hat{t}-1})$. Thereafter both play a Nash

equilibrium of the complete information game between the normal types with payoffs (\hat{v}_1, \hat{v}_2) . If $a_1^{\hat{t}} = \hat{s}_1(\hat{h}^{\hat{t}-1})$ then player two plays a best response against the automaton strategy as in Lemma 2.

Off equilibrium path. Deviations are minmaxed.

The idea is to reward player one for revealing her type with a payoff \hat{v}_1 and if she mimics the automaton she receives a payoff v_1^* less than (or equal) her minmax payoff.

These strategies are feasible, i.e., payoffs (\hat{v}_1, \hat{v}_2) can be sustained as an equilibrium outcome in the complete information game for any discount factors (δ_1, δ_2) , and additionally with μ close to zero it will pay neither player to deviate - see the argument for Case 1 for player one; for player two the expected continuation payoffs after period \hat{t} , compared to the minmax threat if player one is the normal type, will prevent deviation during $\hat{h}^{\hat{t}}$ for μ small enough, and thereafter if player one follows the automaton strategy player two cannot gain by deviating because he is by assumption playing a best response, and if player one is revealed to be the normal type then the continuation game by assumption is in equilibrium so deviations are not profitable. Player one has a new deviation possibility however: the option of mimicking the automaton at time \hat{t} and thereafter; however, if $\delta_1, \delta_2 > \tilde{\delta}$, she will suffer a loss of continuation payoff of at least $(\hat{v}_1 - v_1^*) > 0$, which for δ_1 sufficiently close to one will make this deviation unprofitable (likewise for deviations after first mimicking the automaton).

Finally, for δ_1 near one, the overall payoff to player one from adhering to the above strategy is within $\frac{\epsilon}{2}$ of the continuation payoff after \hat{t} , and by Case 2 $\hat{v}_1 + \frac{\epsilon}{2} < g_1^* + \epsilon$. Hence there exist $\tilde{\delta}$ and $\bar{\mu}$ such that for $\tilde{\delta} < \delta_1, \delta_2 < 1$, $0 < \mu < \bar{\mu}$ the above strategies are feasible and constitute a Nash equilibrium which gives player one a payoff less than $g_1^* + \epsilon$.

Provided that P contains points strictly above $\min \max g_2$ then P must fall into Case 1 or Case 2. This follows because otherwise there would be points $(\bar{v}_1, \bar{v}_2) \in P$ with $\bar{v}_1 \geq g_1^* + \epsilon/2$, $\bar{v}_2 > \min \max g_2$. But then from Lemma 1 (iii) there is a point $(\tilde{v}_1, \tilde{v}_2) \in P$ satisfying $\tilde{v}_1 \leq g_1^*$, $\tilde{v}_2 \geq \min \max g_2$. Hence there exists a convex combination of (\bar{v}_1, \bar{v}_2) and $(\tilde{v}_1, \tilde{v}_2)$, belonging to P by Lemma 1(i), and satisfying Case 1. Finally there is:

Case 3 *There does not exist $(v_1, v_2) \in P$ such that $v_2 > \min \max g_2$.*

Again it will be demonstrated that a separating equilibrium can be constructed.

Lemma 3 *For given δ_1, δ_2 , after the history $\hat{h}^{\hat{t}}$ there is a best response by player two against the automaton which implies that, for all $t > \hat{t} + M - 1$, a_1^t is a minmax strategy against player two and a_2^t is such that $g_2(a_1^t, a_2^t) = \min \max g_2$.*

Proof: From $\hat{t} + 1$ onwards there exists a best response against the automaton which leads, after at most $M - 1$ periods, to a cycle of length, say, $K \leq M$. Suppose player one does

not play a minmax strategy against player two every period in the cycle. Then, player two's best response must achieve a payoff greater than $\min \max g_2$. Since he can guarantee himself at least his minmax payoff in every period and sometimes gets strictly more, his discounted average payoff must be bigger than $\min \max g_2$ - contradicting the definition of case 3. *Q.E.D.*

It follows that, once the cycle starts, a best response strategy for player two is simply *any* short-run best response to the automaton action period by period as this guarantees player two his minmax payoff each period.

Note that by assumption 1 there exists $(\hat{v}_1, \hat{v}_2) \in V(\delta_1, \delta_2)$, such that $g_1^* < \hat{v}_1 < g_1^* + \epsilon$ and $\min \max g_2 < \hat{v}_2$ for all $0 \leq \delta_1, \delta_2 < 1$. Using this payoff vector a separating equilibrium can be constructed as follows:

Equilibrium path. Both players follow $\hat{h}^{\hat{t}}$ for the first $\hat{t} - 1$ periods. At time \hat{t} player one reveals his type by playing some $a_1^{\hat{t}} \neq \hat{s}_1(\hat{h}^{\hat{t}-1})$. Thereafter both play a Nash equilibrium of the complete information game between the normal types with payoffs (\hat{v}_1, \hat{v}_2) . If $a_1^{\hat{t}} = \hat{s}_1(\hat{h}^{\hat{t}-1})$ then thereafter player two plays a best response against the automaton which after no more than $M - 1$ periods specifies each period the short-run best response against the automaton action which minimises player one's payoff.

Off equilibrium path. Deviations are minmaxed.

As before, deviation from the above strategies is unprofitable when μ is close to zero and δ_1, δ_2 are close to one. Mimicking the automaton from period \hat{t} gives player one a payoff each period after $\hat{t} + M - 1$ no greater than g_1^* by definition of g_1^* , which must be inferior to revelation and receipt of a continuation payoff of \hat{v}_1 for δ_1 near one. Hence there exist $\bar{\delta}$ and $\bar{\mu}$ such that for $\bar{\delta} < \delta_1, \delta_2 < 1$, $0 < \mu < \bar{\mu}$ the above strategies are feasible and constitute a Nash equilibrium which gives player one a payoff less than $g_1^* + \epsilon$. *Q.E.D.*

References

- AUMANN, R. AND S. SORIN (1989): "Cooperation and Bounded Recall", *Games and Economic Behaviour*, 1, 5-39.
- CRIPPS, M. AND J. THOMAS (1992): "Reputation and Commitment in Two-Person Repeated Games", *Discussion Paper*, No. 9210, CentER, Tilburg University.
- CRIPPS, M. AND J. THOMAS (1993): "Reputation and Perfection in a Repeated Common Interest Games", *mimeo*, Warwick University
- FUDENBERG, D. (1993): "Explaining Cooperation and Commitment in Repeated Games", forthcoming in J.-J. Laffont (ed.), *Advances in Economic Theory, Sixth World Congress*, Cambridge: Cambridge University Press.
- FUDENBERG, D. AND D.K. LEVINE (1989): "Reputation and Equilibrium Selection in Games with a Patient Player", *Econometrica*, 57, 759 - 778.
- FUDENBERG, D. AND D.K. LEVINE (1992): "Maintaining a Reputation when Strategies are Imperfectly Observed", *Review of Economic Studies*, 59, 561-579.
- FUDENBERG, D. AND D.K. LEVINE (1993): "Self-Confirming Equilibrium", *Econometrica*, forthcoming.
- FUDENBERG, D. AND E. MASKIN (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information", *Econometrica*, 54, 533-554.
- HART, S. (1985): "Nonzero-Sum Two-Person Repeated Games with Incomplete Information" *Mathematics of Operations Research*, Vol. 10, 117-153.
- KALAI, E. AND E. LEHRER (1991): "Bayesian Learning and Nash equilibrium", *mimeo*, Northwestern University
- KREPS, D., P. MILGROM, J. ROBERTS AND R. WILSON (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma", *Journal of Economic Theory*, Vol. 27, 245-252.
- SCHMIDT, K. (1993a): "Reputation and Equilibrium Characterization in Repeated Games of Conflicting Interests", *Econometrica*, forthcoming.
- SCHMIDT, K. (1993b): "Commitment through Incomplete Information in a Simple Repeated Bargaining Game", *Journal of Economic Theory*, forthcoming.
- WATSON, J. (1993): "Cooperation in the Infinitely Repeated Prisoner's Dilemma with Perturbations", *Games and Economic Behaviour*, forthcoming.