JOURNAL OF ECONOMIC THEORY **69**, 387–410 (1996)
ARTICLE NO. 0060

Reputation in Perturbed Repeated Games*

MARTIN W. CRIPPS

Department of Economics, Warwick University, Coventry CV4 7AL, United Kingdom

KLAUS M. SCHMIDT

Wirtschaftspolitische Abteilung, Bonn University, Adenauerallee 24, 53113 Bonn, Germany

AND

JONATHAN P. THOMAS†

Department of Economics, Warwick University, Coventry CV4 7AL, United Kingdom

Received July 2, 1993; revised March 13, 1995

The paper analyzes reputation effects in perturbed repeated games with discounting. If there is some positive prior probability that one of the players is committed to play the same (pure) action in every period, then this provides a lower bound for her equilibrium playoff 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 bound for games with a long-run player facing a sequence of short-run opponents. The bound cannot be improved by considering types playing finitely complicated history-dependent commitment strategies. *Journal of Economic Literature* Classification Numbers: C73, D83, L14.

1. INTRODUCTION

This paper analyzes reputation effects in perturbed repeated normal form games with discounting. If one of the players is sufficiently more patient than her opponent and if there is some positive prior probability that she

* The second and the third author are grateful to the Economics Department at MIT and the Econometric Institute, Erasmus University, respectively, for their hospitality. We thank Nabil Al-Najjar, Drew Fudenberg, Georg Nöldeke, two referees, and an associate editor of this journal for helpful comments and suggestions. The paper also benefitted from comments by seminar participants at Bonn University, CentER, LSE, the European Summer Symposium in Economic Theory in Gerzensee (Switzerland), and the Tokyo Center of Economic Research International Summer Conference in Tateshina (Japan). Financial support for this project by Deutsche Forschungsgemeinschaft, SFB 303 at Bonn University, is gratefully acknowledged.

† To whom correspondence should be addressed.

0022-0531/96 $18.00

is a "commitment type" who always plays the same action in every period, then this provides a lower bound for her equilibrium payoff in all Nash equilibria. The reason is that the normal type of this player can mimic the commitment type and build up a reputation for always playing the commitment strategy. Considering all pure actions of the stage game, we derive the highest such lower bound. This bound is tight in the sense that there always exists an equilibrium of the perturbed game involving only one additional type of player 1, the commitment type, such that player 1's equilibrium payoff is arbitrarily close to this bound. Furthermore, the bound is independent of which other types have positive probability. We also show that this bound cannot be improved by considering types playing finitely complicated, history-dependent, commitment strategies.

Most of the literature on reputation effects deals with the case where a long-run player (player 1) faces a sequence of short-run opponents. Fudenberg and Levine [6, 7] have shown that if there is some incomplete information such that a Stackelberg type, who always plays the Stackelberg strategy of the stage game, has strictly positive probability, then the long-run player can build up a reputation for always playing the Stackelberg strategy and achieve at least her Stackelberg payoff in any Nash equilibrium of the game.[1]

The two long-run player case is considerably more complicated. If player 2 is long-lived and cares about future payoffs, then he will not play a best response to the expected action of player 1 in any given period, if, by doing so, he expects to be punished in later periods and his future utility is sufficiently adversely effected. This drives a wedge between the highest payoff player 1 can obtain by publicly committing herself and the lowest payoff she can guarantee herself in all Nash equilibria by building up a reputation. Note first, that if player 2 is long-lived it is no longer optimal to commit the Stackelberg strategy of the stage game. In general, player 1 could obtain an even higher payoff by committing herself to a history-dependent strategy; for example, the tit-for-tat strategy in the repeated prisoner's dilemma game.[2] On the other hand, it becomes more difficult to build up a reputation. 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. However, it is not possible for player 1 to signal what would be done off the equilibrium path. Thus an equilibrium

[1] This result requires that the stage game is either a simultaneous move game, or, in a sequential move game, that the short-run players always observe whether or not the Stackelberg strategy has been played.

[2] Here the Stackelberg strategy in the stage game is to defect which guarantees player 1 at most his minmax payoff. Note that the optimal response of a patient player 2 to the tit-for-tat strategy in the repeated game is to cooperate which is *not* a best response in the stage game.

may specify that the opponent does not play a best response to the strategy player 1 is mimicking, and the equilibrium is sustained by the possibility that player 1 may severely punish player 2 in the future should he ever deviate.

The first paper which attempts to extend Fudenberg and Levine [6, 7] to the two long-run player case is Schmidt [10]. He identifies the only case where the wedge discussed above does not exist. If the game is of "conflicting interests" in the sense that the strategy to which player 1 would most like to commit herself minmaxes player 2, and if player 1 mimics a type who follows this strategy, then player 2 must eventually play a best response. Otherwise, he would get less than his minmax payoff which is impossible in Nash equilibrium. Hence, Fudenberg and Levine's [6] lower bound obtains and coincides with the Stackelberg payoff. Schmidt also shows that if the game is not of conflicting interests the result does not hold, that is, there are equilibria which give player 1 a payoff bounded below her Stackelberg payoff no matter how patient she is.

In this paper, we examine two main questions arising from this literature. First, is there a non-trivial lower bound on Nash equilibrium payoffs in repeated games that are not of conflicting interests?[3] The answer is yes. In Section 3, we derive such a bound for general stage games using "simple" commitment types, i.e., types which play the same action in every period. Of course, this bound must be lower than the Stackelberg payoff, but it is generally higher than player 1's minmax payoff and may still be very useful in restricting the set of equilibrium payoffs. Furthermore, the bound is tight in the sense that for any given prior probability of the commitment type, a game can be constructed with equilibria giving player 1 a payoff arbitrarily close to the bound. This result generalizes Schmidt [10] and Fudenberg and Levine [6, 7] and shows that reputation effects have implications in general two long-run player games.

Note, however, that mimicking a "simple" commitment type is not necessarily optimal. The second main question (addressed in Section 4), is whether types committed to 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.) Perhaps surprisingly, the answer is negative. 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.

[3] While there are some interesting examples for games of conflicting interests (e.g., the "chain store game" or the "game of chicken"), it clearly is a very narrow class of games.

The intuition for our lower bound is roughly as follows. Suppose that player 1 mimics a simple commitment type. 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 1 can be lower than the bound derived by Fudenberg and Levine [6, 7]. To find a lower bound we have to consider that *individually rational* response by player 2 which gives player 1 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 1's payoff in any Nash equilibrium provided that she is patient enough to wait for the reputation building to take place. In games of conflicting interests this bound is equal to the Stackelberg payoff.[4]

This result is very robust: it is independent of the existence of types of player 1 other than the commitment type, and is also robust to the existence of small amounts of incomplete information about player 2. Note, however, that it requires that player 1 is sufficiently patient as compared to player 2; i.e., one of the players has to have sufficiently more at stake than the other for the reputation effect to work to her advantage. This condition is very natural and discussed extensively in Section 3.

Why is it impossible for player 1 to improve the lower bound by mimicking a more (but finitely) complicated commitment type? Recall that history-dependent commitment strategies can considerably improve the Stackelberg payoff of player 1 in two long-run player games. The problem hinges again on the difference between *on*- and *off*-equilibrium path behaviour. If player 1 can publicly commit herself, she commits to a strategy which determines her actions on *and* off the equilibrium path. In particular, she can credibly threaten to punish player 2 if he deviates from a certain path of behaviour. On the other hand, in Nash equilibrium player 1 can build up a reputation only for what she is going to do *on* the equilibrium path, since information sets *off* the path are never reached. We show that it is always possible to construct a Nash equilibrium giving a payoff to player 1 arbitrarily close to our bound, no matter how sophisticated the strategy of the commitment type player 1 could mimic is.

This argument suggests that if player 1's *off*-equilibrium path strategy were observed occasionally, then she might be able to build up a reputation for *off*-equilibrium path behaviour. Two recent papers (written after

[4] The same lower bound has been established for the case where both players evaluate payoffs according to long-run averages by Cripps and Thomas [4]. In comparison, very little is known about the discounted case and the case where both players are incompletely informed.

the first version of this paper had been submitted) build on this observation. Celentani *et al.* [2] consider games in which each player observes only an imperfect signal about the action taken by his opponent. If each signal has strictly positive probability given any action profile, then every information set is reached with strictly positive probability and there is no "off the equilibrium path." Hence, if player 1 mimics a commitment type whose strategy depends only on the history in the last finite number of periods, player 2 will eventually learn what her strategy is. Similarly, Aoyagi [1] considers trembling hand perfect equilibria and requires that with some fixed and strictly positive probabilities each player must take each action in every period. Both papers show that in these circumstances the wedge between the Stackelberg payoff and lower bound on player 1's equilibrium payoff disappears if player 1 is patient enough as compared to player 2. We discuss the relation of our work to these papers in more detail in Section 5.

2. The Model

The description of the model follows closely Fudenberg and Levine [6, 7] and Schmidt [10]. There are two players called "1" (she) and "2" (he). In every period they play a normal form game g; i.e., each player selects an "action" a_i out of a finite action space A_i, $i \in \{1, 2\}$.[5] A mixed action is denoted by $\alpha_i \in \mathscr{A}_i$. 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. We follow the convention of normalizing payoffs so that stage game and repeated game payoffs can be expressed on the same scale. The normalized or average payoff of player i from period t onwards (and including period t) is

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

and the average payoff from period 1 for a given δ_i is $v_i(\delta_i) \equiv v_i^1(\delta_i)$, where δ_i, $0 \leqslant \delta_i \leqslant 1$, denotes the respective discount factor. The reference to δ_i will be omitted if there is no ambiguity.

Consider a perturbation of this game such that there is some incomplete information about the payoff function of player 1. In period 0 the "type" ω

[5] All our results carry over to extensive form games as long as all actions are perfectly observed by both players at the end of each period.

of player 1 is drawn by nature out of the set Ω according to the probability measure μ. Let ω^0 denote the "normal" type of player 1 whose payoff function is as described in the unperturbed game, i.e., $g_1(a_1, a_2 \mid \omega^0) = g_1(a_1, a_2)$. We will omit the reference to ω^0 in the payoff function of the normal type. The other types will sometimes be called "irrational types." They may have arbitrary payoff functions which may be non-stationary.[6]

We are particulary interested in what impact the presence of "commitment types" has on the set of equilibrium outcomes. A commitment type is an irrational type for whom it is a dominant strategy in the repeated game to follow a particular strategy $\hat{s}_1$. In Section 3, we derive a lower bound on player 1's (normal type) equilibrium payoff using "simple" commitment types; types, that is, who always play the same action $\hat{a}_1$ in every period. Such a type is denoted by $\omega(\hat{a}_1)$. In Section 4, we consider more general commitment types who may be committed to play a more complicated, history-dependent, strategy.

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 $s_1 \in S_1$ for player 1 in the perturbed repeated game is a sequence of maps $s_1^t\colon \Omega_1 \times H^{t-1} \to A_1$, while a pure strategy $s_2 \in S_2$ of player 2 is a sequence of maps $s_2^t\colon H^{t-1} \to A_1$. Correspondingly, $\sigma_i = (\sigma_i^1, \sigma_i^2, \ldots) \in \Sigma_i$ denotes a mixed (behavioral) strategy of player i. Let minmax $g_i = \min_{\alpha_{-i} \in \mathscr{A}_{-i}} \max_{\alpha_i \in \mathscr{A}_i} g_i(\alpha_1, \alpha_2)$ be player i's minmax payoff. A Nash equilibrium requires that both players choose a strategy to maximize their expected average payoff taking the strategy of the opponent as given, where expectations are taken with respect to the joint distribution over types and infinite histories induces by μ and (σ_1, σ_2).

3. The Lower Bound on Equilibrium Payoffs

In this section, we will establish a lower bound for the equilibrium payoff of the normal type of player 1 which generalizes the results of Fudenberg and Levine [6, 7] and Schmidt [10]. As in these papers, our lower bound is based on the fact that player 1 always has the option to mimic a simple

[6] Note that there are two restrictions imposed here: First, we only consider the case of one-sided incomplete information. The extension of our results to the case of two-sided uncertainty is straightforward and exactly the same as in Schmidt [11]. In particular, all our results hold up to a small ε if the perturbation is small, i.e., if the probability that player 2 is of the normal type is close enough to 1. Second, we do not consider the case where player 1 could mimic an irrational type who is committed to a mixed strategy. Both generalizations can be found in the working paper version of this paper [3]. The latter one builds on Fudenberg and Levine [7] and is not trivial. However, considering this case requires much more notation and significantly increases the length of the proofs.

commitment type who plays the same action in every period. However, as discussed in the introduction, there are two important differences to the above mentioned papers: First, we do not restrict attention to "games of conflicting interests" or games where a long-run player faces a sequence of short-run players. While in these special cases player 2 must eventually play a best response to the commitment strategy, this is not the case in general two long-run player games. Second, in general two long-run player games player 1 might do better by mimicking a history-dependent commitment strategy. In Section 4, we demonstrate that our lower bound is tight and cannot be improved upon by considering more complicated commitment types.

Suppose there is a type $\hat{\omega} = \omega(\hat{a}_1)$ with $\mu(\hat{\omega}) > 0$ who is committed to always play $\hat{a}_1$. We are going to show that if player 1 chooses to follow this strategy (which must happen with a probability at least $\mu(\hat{\omega}) > 0$ in any Nash equilibrium), then player 2 cannot continue to respond with a strategy which gives him less than his minmax payoff against $\hat{a}_1$. To put this more formally, let

$$\mathcal{M}(\hat{a}_1) = \{\alpha_2 \in \mathcal{A}_2 \mid g_2(\hat{a}_1, \alpha_2) \geqslant \text{minmax } g_2\} \tag{2}$$

be the set of all strategies of player 2 which yield at least minmax g_2 against $\hat{a}_1$. Suppose that player 2 picks a strategy out of $\mathcal{M}(\hat{a}_1)$ which is worst for player 1. Then the best strategy to which player 1 would like to commit herself is

$$a_1^* \in \arg\max_{a_1 \in A_1} \min_{\alpha_2 \in \mathcal{M}(a_1)} g_1(a_1, \alpha_2). \tag{3}$$

Define the corresponding payoff of player 1,

$$g_1^* = \min_{\alpha_2 \in \mathcal{M}(a_1^*)} g_1(a_1^*, \alpha_2), \tag{4}$$

to be her "minimal commitment payoff."[7] Now we can state the main result of this section:

THEOREM 1. *Let ω^* be a commitment type who always plays a_1^*, and let $\mu(\omega^0) > 0$ and $\mu(\omega^*) > 0$ 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 1 in any Nash equilibrium is at least*

$$E[v_1(\delta_1) \mid \omega^0] > g_1^* - \phi. \tag{5}$$

[7] Since we restrict attention to pure commitment strategies, it is possible that $g_1^* <$ minmax g_1. In such cases we adopt the convention of setting $g_1^* =$ minmax g_1. It should also be mentioned that although we concentrate here on the best strategy, Eq. (4) defines a bound for an arbitrary strategy if a_1^* is replaced by any a_1.

Proof. See Appendix A.

The proof of the Theorem is complicated and relegated to the Appendix, but we want to give some intuition for it. The argument proceeds in three steps: Suppose that player 1 sticks to always playing a_1^*. First we show that if there is a period t where player 2's equilibrium strategy yields a continuation payoff against the repeated play of a_1^* which is smaller than minmax g_2, then player 2 must expect player 1 to deviate from a_1^* with strictly positive probability in the not too distant future, otherwise player 2 would be anticipating a payoff less than minmax g_2, impossible in Nash equilibrium.

The second step uses a proposition on statistical inference which is due to Fudenberg and Levine [6]. It says that if player 2 always observes player 1 playing a_1^*, then he cannot continue to believe that it is "unlikely" that player 1 is going to play a_1^* in the future. Taken together, these results imply that if player 1 sticks to playing a_1^*, then player 2 cannot continue to respond with a strategy which gives him less than minmax g_2. Eventually, after some finite number of periods, he will learn that his opponent plays a_1^* and he will choose a response which gives him at least his minmax payoff against the commitment strategy. If player 1 is sufficiently patient, then any finite number of periods does not count very much in her overall payoff, and she can guarantee herself almost the minimal commitment payoff g_1^*.[8]

The lower bound on player 1's equilibrium payoff has two important properties: First, it holds for *all* Nash equilibria. Note that our proof does not exploit all of the rationality requirements imposed by the Nash concept. In particular, our result does not depend upon whether or not players correctly predict the behaviour of their opponent off the equilibrium path. Thus our result carries over to the more general notion of all self-confirming equilibria.[9] Second, our lower bound is independent of what other

[8] The formal argument is considerably more involved and complicated by the fact that players 1 and 2 use different discount factors. See Appendix A for details. It should be noted that the proof is *essentially* different from the proof used by Schmidt [11]. Schmidt's argument requires that player 1 mimicks a commitment type which minmaxes player 2, while this is not required for the argument developed here. Hence, the lower bound which can be obtained for games not of conflicting interests using Schmidt's proof (see Schmidt [11, p. 344]) is generally *strictly* weaker than our lower bound. This can easily be seen from the example of the Battle of the Sexes given below, where mimicking a type who minmaxes player 2 would guarantee player 1 less than her minmax payoff.

[9] See Fudenberg and Levine [8] for a definition and characterization of self-confirming equilibria. In our context, the set of self-confirming equilibrium outcomes and the set of Nash equilibrium outcomes coincides, so we stick to the more familiar Nash concept.

types of player 1 may exist with positive probability as long as the commitment type is present. Furthermore, it can be shown that our lower bound carries over (up to an arbitrarily small ε) to the case of two-sided uncertainty if the probability of the normal type of player 2 is sufficiently close to 1.[10] Thus, as in Fudenberg and Levine [6] and Schmidt [10], the 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 Theorem 1 the discount factor of player 2 is fixed while δ_1 has to be chosen close enough to 1; i.e., player 1 has to be sufficiently patient as compared to player 2. In particular, our lower bound does not hold if both players are equally patient and if their common discount factor goes to 1.[11] To see why this is the case suppose the discount factor of player 2 increases so that he cares more about future payoffs. Then, he may continue for a longer period of time to play a strategy which gives him less than minmax g_2 against type ω^*, even if he expects player 1 to play a_1^* in the following periods with a very high probability. Thus, player 1 has to wait for a longer period of time until she can be sure that player 2 will respond to a_1^* with a strategy which gives him at least minmax g_2. In order to get the same lower bound player 1 has to be sufficiently more patient.[12]

[10] See Cripps, *et al.* [3, Theorem 2].

[11] A counterexample can be found in Cripps and Thomas [5], who construct an equilibrium of a common interest game with one-sided incomplete information and a common discount factor δ for both players which violates our lower bound for any $\delta < 1$. In Cripps and Thomas [4], the same lower bound as in current paper is derived for the case where both players use the limit of the means criterion for averaging payoffs. Hence only if the discount factors converge to unity in a particular way—with δ_1 converging faster than δ_2—is there continuity between the discounted and undiscounted cases. See also the discussion following the example at the end of this section.

[12] In economic applications it is typically difficult to justify why players face different discount rates. Therefore we would like to offer a different interpretation: Suppose player 1 is a "big" player who plays the same game in each period against L "small" opponents, denoted by $2_1, ..., 2_L$. Each of the small players 2_j is long-lived and has the same discount factor as the "big" player 1. Divide each period in L subperiods and suppose that player 1 plays sequentially against a small player 2_j in subperiod j, $j = 1, ..., L$. From the perspective of player 1, increasing the number of small players is equivalent to increasing her discount factor. To see this note that along the equilibrium path all small players share the same expectation about player 1's equilibrium strategy. Thus, the maximum number of times where player 1 may receive less than g_1^* remains constant as the number of small players increases. Hence, the costs for building up a reputation remain roughly the same (they slightly increase because they may have to be incurred earlier, so they are discounted less), while the benefits multiply with the number of small players. Thus, we can say that player 1 can exploit reputation effects to increase her equilibrium payoff if she is sufficiently "big" as compared to her opponents. We are grateful to Nabil Al-Najjar for suggesting this interpretation.

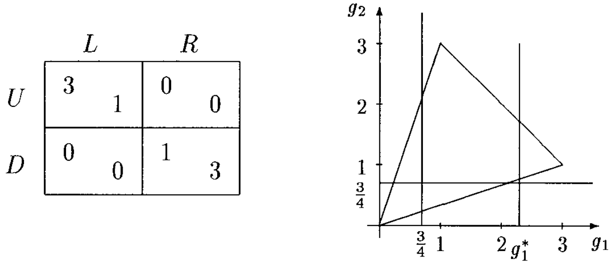

FIG. 1. The Battle of the Sexes

How useful is the lower bound? Recall that our argument also works for mixed commitment strategies. Thus, $g_1^* \geqslant \text{minmax } g_1$ because minmax g_1 is derived from minimizing player 1's payoff over all $\alpha_2 \in \mathscr{A}_2$, while g_1^* is obtained from minimizing over the smaller set $\mathscr{M}(\alpha_1^*)$, i.e., the set of all strategies which give player 2 at least his minmax payoff. Hence, our lower bound can be used to restrict the set of Nash equilibrium outcomes as compared to the prediction of the Folk Theorem.[13] Our lower bound is most powerful if the game is of "conflicting interests" in the sense of Schmidt [10] in which case it coincides with the Stackelberg payoff. If the game is not of conflicting interests our lower bound can still be very useful. To illustrate this consider the Battle of the Sexes game depicted in Fig. 1.[14]

Player 1's minmax strategy is to play U with probability 3/4 and D with probability 1/4 which holds player 2 down to a payoff of 3/4. Thus, player 2 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 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 1 for whom it is a dominant strategy to always play U. In order to get at least his minmax payoff against this type, player 2 must play L with probability at least 3/4. Thus, if player 1 mimics this type and if she is patient enough as compared to player 2, then she will get a payoff of at least 9/4 in any Nash equilibrium of the repeated game. This considerably reduces the set of Nash

[13] For a more detailed discussion of the Folk Theorem for games with asymmetric information (Fudenberg and Maskin [9]), and in what sense the lower bound is in contrast to its prediction, see Schmidt [11, p. 341].

[14] Note, that the graph illustrates only the stage-game payoffs g_i. If the players have different discount factors it is inappropriate to represent the normalized payoffs of the repeated game on the same scale as the stage-game payoffs. See Schmidt [11, footnote 8].

equilibrium payoffs for player 1 as compared to the prediction of the Folk Theorem.

The Battle of the Sexes game is completely symmetric. We mentioned already that our results carry over to the case of two-sided uncertainty provided the perturbation is small. Thus, if there is a type of player 2 who always plays R, then, if player 2 is patient enough as compared to player 1, he can guarantee himself at least 9/4, too. This indicates again that our lower bound cannot apply if both players have the same discount factors, in the two-sided case with equal discount factors both players would be able to guarantee themselves 9/4. However, the payoff vector (9/4, 9/4) is not in the feasible set. Thus, there are only two possibilities. First, a Nash equilibrium does not exist. However, we know that our result holds for finitely repeated games as well if T is large enough; with T finite, we have a finite game, so existence is not a problem. Hence, the only remaining possibility is that our lower bound does not hold if there is two-sided uncertainty and both parties have the same discount factor. Only if player 1 is sufficiently more patient than player 2 does the reputation effect work to her advantage, and vice versa.

4. The Value of History-Dependent Commitment Strategies

Up to this point we have focused on commitment types that played a constant action in every period. The question arises whether player 1 can guarantee herself a better lower bound than g_1^* (as defined in Theorem 1) 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. 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 1 could choose a single type (to be given positive prior probability by player 2), then *she could do no better than choose a commitment type which plays the same action in every period.* As a by-product of this analysis the tightness of our bound will be established: for every game there exists a perturbation involving the commitment type with 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, with each state $x \in X$ being associated with an action $a_1 \in A_1$, and there is a transition rule which associates with each state and action of player 2 (x, a_2) a new state for the automaton. The automaton starts in some initial state x_0, with the action associated with

x_0 being played in period 1, and so on. The pure strategy implemented by this automaton will be denoted $\hat{s}_1$.[15]

The following theorem shows that it is not possible to improve on g_1^* by mimicking any finite automaton which plays a strategy $\hat{s}_1$. That is, even if very complicated commitment types exist, it is always possible to construct, in a perturbation involving only that additional type, a Nash equilibrium in which player 1, normal type, gets at most $g_1^* + \varepsilon$, where $\varepsilon > 0$ may be arbitrarily small.[16] To avoid detailed considerations of some awkward boundary cases we assume that feasible strictly individually rational stage-game payoffs exist which give player 1 more than g_1^*:[17]

Assumption 1. Let F be the convex hull of the unperturbed stage-game feasible payoff set. Then there exists $(v_1, v_2) \in F$ with $v_1 > g_1^*$ and $v_2 >$ minmax g_2.

Then we have:

THEOREM 2. *Fix any pure strategy $\hat{s}_1$ that can be implemented by a finite automaton. For any $\varepsilon > 0$ 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 1 gets an average equilibrium payoff which is smaller than $g_1^* + \varepsilon$.*

Proof. See Appendix B.

Note that Theorem 2 implies that our bound is tight: If we consider the automaton which always plays a_1^*, then the theorem tells us that there is an equilibrium of a perturbed game in which $\omega(a_1^*)$ has positive probability such that the equilibrium payoff for player 1 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 $\omega(\hat{s}_1)$ has positive probability and a Nash equilibrium of this perturbed game in which the normal type of player 1 gets less than $g_1^* + \varepsilon$. Consider the most simple such perturbation in which player 1 is either the normal type (with probability $1-\mu$) or the

[15] The 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.

[16] It is unclear whether this result extends to infinite automata. See the discussion at the end of this section.

[17] The other cases are dealt with briefly in [3].

commitment type (with probability μ). 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* (*player* 1 *builds a reputation for following the automaton strategy*): Suppose there exists a pure strategy $\hat{s}_2$ of player 2 such that $(\hat{s}_1, \hat{s}_2)$ yields a continuation payoff (after any period t) strictly higher than minmax g_2 to player 2 and strictly in between minmax g_1 and $g_1^* + \varepsilon$ to player 1. In this case the following strategies constitute a Nash equilibrium if both players are patient enough and if μ is small enough: The normal and the commitment type of player 1 both play $\hat{s}_1$ and player 2 responds with $\hat{s}_2$. Any deviation is minmaxed by the other player.[18] By construction, the normal type of player 1 gets less than $g_1^* + \varepsilon$ in this equilibrium.

— *A separating equilibrium* (*player* 1 *distinguishes herself from the automaton*): Suppose no such strategy for player 2 exists. In this case any strategy combination $(\hat{s}_1, \hat{s}_2)$ will at some point either hold player 1 weakly below her minmax payoff, or give player 2 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 1, 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 deviation from $\hat{s}_1$ and is rewarded in the continuation equilibrium with a payoff strictly above her minmax payoff but strictly below $g_1^* + \varepsilon$. Player 2 also gets strictly more than minmax g_2 in this continuation equilibrium. He follows $\hat{s}_2^*$ up to the period where the normal type of player 1 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 1 gets less than $g_1^* + \varepsilon$.

Finally, suppose that $(\hat{s}_1, \hat{s}_2^*)$ gives player 2 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 actions period by period since this guarantees player 2 his minmax payoff each period. But then, by the definition of g_1^*, there exists an $\hat{s}_2^*$ such that $(\hat{s}_1, \hat{s}_2^*)$ does not give player 1

[18] This equilibrium is not a sequential equilibrium since *not* punishing a deviation by player 2 will give the normal type of player 1 full reputation of being the automaton and this may be highly desirable.

more than g_1^*. Again, we can construct a separating equilibrium, but now the normal type of player 1 is rewarded with $g_1^* + \varepsilon$ if she reveals herself.

Theorem 2 establishes the existence of a Nash equilibrium with the desired payoff for any given commitment strategy $\hat{s}_1$ which can be implemented by a finite automaton. Note that the lower bound on the discount factor, $\bar{\delta}$, may depend on the number of states of the automaton. However, if we consider uniformly bounded automata, it follows immediately that for any number of states, m, there exists a $\bar{\delta}$ such that Theorem 2 holds for *all* automata with less than or equal to m states.[19]

The reason why we restricted attention to finite automata is as follows. The arguments given above require that continuation payoffs after any period t are bounded away from minmax g_1; otherwise the threat of mutual minmaxing may not be sufficient to deter a deviation. A second major complication with an infinite automaton is that the above division into pooling and separating equilibria fails in the sense that there may not exist critical discount factors above which an equilibrium may break down at some higher discount factor.

5. Summary and Conclusions

We have shown that if a player has the possibility of acquiring 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 [6, 27] 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.

Theorem 2 shows that it is not possible with finite automata to find a better lower bound for all *Nash* equilibria. We noted, however, that the pooling equilibrium constructed in the proof is not sequential. An important question for future research is whether a better lower bound can be obtained if the equilibrium notion is refined. Celentani *et al.* [2] give an example for a perturbed repeated game where there exists a *perfect Bayesian* equilibrium giving player 1 considerably less than her Stackelberg payoff. Thus, there is no hope that requiring perfection is sufficient to

[19] This is true since for a fixed number of states there is only a finite number of automata. Hence, we can simply take $\bar{\delta}$ to be the maximum critical discount factor over this finite set.

obtain the Stackelberg payoff as a lower bound in general games. However, it may still be possible to improve the bound we obtained in Theorem 1.[20]

An alternative strategy is to introduce trembles or imperfect observability so that every information set is reached with strictly positive probability. Aoyagi [1] and Celentani *et al.* [2] have shown that in this case player 1 will obtain at least his Stackelberg payoff if he is sufficiently patient. Note, however, that the order of limits is crucial here. References [1] and [2] fix the probabilities with which each information set must be reached and then make the limit as player 1's discount factor goes to 1. If the order of limits is reversed, such that for any given discount factor of player 1 the probabilities of trembles go to 0, then we are back to the case considered in this paper where only the weaker lower bound obtains.

APPENDIX A: Proof of Theorem 1

The basic idea of the proof is to show that if the normal type of player 1 mimics the commitment strategy a_1^*, then she will get close to g_1^* and hence her equilibrium payoff must be at least this amount. We proceed in three steps: Lemma 1, Lemma 2 and the final proof of the theorem. For a verbal description of these steps and the basic intuition see the discussion following the theorem. The strategy "always play a_1^*" is denoted by s_1^*.

Lemma 1. *Let $\mu^* = \mu(\omega^*) > 0$, $\delta_2 < 1$ and $\varepsilon > 0$ be given and consider any Nash equilibrium $(\hat{\sigma}_1, \hat{\sigma}_2)$ and any history $\hat{h}^t$ which has positive probability in this equilibrium conditional upon ω^*. Let $\hat{\sigma}_2(\hat{h}^t)$ be the equilibrium strategy of player 2 after this history, and let $\hat{s}_2(\hat{h}^t)$ be any pure strategy in the support of $\hat{\sigma}_2(\hat{h}^t)$. Suppose that the strategy profile $(s_1^*, \hat{s}_2(\hat{h}^t))$ yields a continuation payoff for player 2 of*

$$E(v_2^{t+1} \mid \hat{h}^t, \omega^*) \leqslant \text{minmax } g_2 - \varepsilon. \tag{6}$$

Then for any δ_2, $0 < \delta_2 < 1$, there exists a finite integer N and a positive number $\eta \in (0, 1]$, both depending only on δ_2 and ε, such that in at least one of the periods $t+1, t+2, ..., t+N$ the probability that player 1 does not play

[20] Cripps and Thomas [5] consider the set of all perfect Bayesian equilibrium payoffs in a perturbed common interest game. In this game, perfection reduces the set of Nash equilibria, but only to a limited extent. Schmidt [12] characterizes the set of all sequential equilibria satisfying a weak Markov property in a finitely repeated bargaining game. He exploits the special structure of the set of possible types to derive lower bounds for the equilibrium payoffs of both players. Surprisingly, his result holds independently of the relative patience of the two players.

a_1^* *(given that she always played* a_1^* *before and that player 2 played according to* $\hat{s}_2(\hat{h}^t)$*) must be at least* η.

Proof. Consider any history $\hat{h}^t$ up to and including period t that has positive probability given $(s_1^*, \hat{\sigma}_2)$. If the continuation strategy of player 2 after this history, $\hat{\sigma}_2(\hat{h}^t)$, is mixed, it will be convenient to think about it not as a behavioral strategy but rather as a probability distribution over pure strategies $\hat{s}_2(\hat{h}^t)$. Consider any such $\hat{s}_2(\hat{h}^t)$ in the support of $\hat{\sigma}_2(\hat{h}^t)$ and let $\pi^\tau = \text{Prob}(s_1^\tau = a_1^* \mid \hat{h}^{\tau-1})$. In words: π^τ is the probability player 2 attaches to the event that player 1 will play a_1^* in period τ, given that player 1 always played a_1^* before and that player 2 himself followed his (pure) strategy $\hat{s}_2(\hat{h}^t)$. Note that the π^τ are well defined numbers for all $\tau = t+1, t+2, ...$ (they are not random given that we specified a strategy profile).

Let $\underline{g}_i = \min_{a_i \in A_i} \min_{a_{-i} \in A_{-i}} g_i(a_1, a_2)$ be the *worst payoff* player i (normal type) can get in the stage game, and $\bar{g}_i = \max_{a_i \in A_i} \max_{a_{-i} \in A_{-i}} g_i(a_1, a_2)$ be the *best payoff* for player i. Define N to be the smallest integer such that

$$N > \frac{\ln \varepsilon - \ln(\bar{g}_2 - \underline{g}_2)}{\ln \delta_2}, \tag{7}$$

and let

$$\eta = \frac{\varepsilon - \delta_2^N(\bar{g}_2 - \underline{g}_2)}{N \cdot \bar{g}_2} > 0. \tag{8}$$

Suppose that, after some period t, $E(v_2^{t+1} \mid \hat{h}^t, \omega^*) \leqslant \text{minmax } g_2 - \varepsilon$ and $\pi^\tau > 1 - \eta$ for all $\tau \in \{t+1, t+2, ..., t+N\}$. With probability $\prod_{\tau=t+1}^{t+N} \pi^\tau$ player 1 will play a_1^* in all of the next N periods, in which case player 2's payoff is at most $\text{minmax } g_2 - \varepsilon + \delta_2^N(\bar{g}_2 - \underline{g}_2)$. With probability $\sum_{i=0}^{N-1} (\prod_{\tau=t+1}^{t+i} \pi^\tau)(1 - \pi^{t+i+1})$ player 1 will not play a_1^* in at least one of the next N periods. In this case the best possible payoff for player 2 is $\bar{g}_2$. Thus we have

$$E(v_2^{t+1} \mid \hat{h}^t) \leqslant \prod_{\tau=t+1}^{t+N} \pi^\tau[\text{minmax } g_2 - \varepsilon + \delta_2^N(\bar{g}_2 - \underline{g}_2)] + \sum_{i=0}^{N-1} \left(\prod_{\tau=t+1}^{t+i} \pi^\tau \right) (1 - \pi^{t+i+1})\, \bar{g}_2. \tag{9}$$

Using $\pi^\tau \leqslant 1$ and $1 - \pi^\tau < \eta$, we get

$$E(v_2^{t+1} \mid \hat{h}^t) < \text{minmax } g_2 - \varepsilon + \delta_2^N(\bar{g}_2 - \underline{g}_2) + N\eta\bar{g}_2. \tag{10}$$

Note that η has been chosen such that $\delta_2^N(\bar{g}_2 - \underline{g}_2) + N\eta\bar{g}_2 = \varepsilon$. Therefore $E(v_2^{t+1} \mid \hat{h}^t) < \text{minmax } g_2$, a contradiction to the assumption that $\hat{s}_2(\hat{h}^t)$ is in the support of the equilibrium strategy for player 2. Q.E.D.

The following lemma on statistical inference is the second step of the proof. It is due to Fudenberg and Levine [6, Prop. 1]:

LEMMA 2. *Let $0 \leqslant \bar{\pi} < 1$. Let H^* be the set of all histories in which player 1 always plays a_1^*. Consider any strategy profile (σ_1, σ_2) which satisfies* $\text{Prob}(h \in H^* \mid \omega^*) = 1$. *Then*

$$\text{Prob}\left[n(\pi^t \leqslant \bar{\pi}) > \frac{\ln \mu^*}{\ln \bar{\pi}} \,\middle|\, h \in H^* \right] = 0, \tag{11}$$

where π^t denotes the probability attached by player 2 to the event that a_1^* is going to be played in period t, and $n(\pi^t \leqslant \bar{\pi})$ is the number (possibly infinite) of random variables π^t for which $\pi^t \leqslant \bar{\pi}$.

We are now ready to prove Theorem 1. To simplify notation let $\underline{g}_1^* = \min_{\alpha_2 \in A_2} g_1(a_1^*, \alpha_2)$. Fix $\phi > 0$. By the definition of the minimal commitment payoff, if $g_2(a_1^*, \alpha_2) \geqslant \text{minmax } g_2$, then $g_1(a_1^*, \alpha_2) \geqslant \underline{g}_1^*$. We can choose $\varepsilon > 0$ such that

$$g_2(a_1^*, \alpha_2) \geqslant \text{minmax } g_2 - \varepsilon \qquad \text{implies} \quad g_1(a_1^*, \alpha_2) \geqslant \underline{g}_1^* - \frac{\phi}{2}. \tag{12}$$

Given ε, by Lemma 1 there is an N and an η such that the probability that player 1 does not play a_1^* in at least one of the next N periods is at least η whenever (6) holds. Set $\bar{\pi} = 1 - \eta$ in Lemma 2. By Lemma 2, if player 1 follows s_1^* there are at most $K = \ln \mu^* / \ln(1 - \eta)$ periods in which the probability that player 1 does not play a_1^* is bigger than η. We conclude that there can be at most NK periods in which the events "$E(v_2^{t+1} \mid h^t, \omega^*) \leqslant \text{minmax } g_2 - \varepsilon$" occur.

Next,

$$E[v_2^{t+1} \mid h^t, \omega^*] = (1 - \delta_2)\, E\left[\sum_{\tau = t+1}^{\infty} \delta_2^{\tau - t - 1} g_2(a_1^*, \sigma_2^\tau(h^{\tau - 1})) \,\middle|\, h^t, \omega^* \right]. \tag{13}$$

Player 2's future expected payoff $E[v_2^{t+1} \mid h^t, \omega^*]$ is therefore a convex combination of terms $g_2(a_1^*, \sigma_2^\tau)$.

Consider the convex set of payoffs (g_1, g_2) that are consistent with player 1 playing a_1^*, i.e. $\{g(a_1^*, \alpha_2) \mid \alpha_2 \in \mathscr{A}_2\}$, and denote this set by $F(a_1^*)$. Note that by (12) there cannot be a point in $F(a_1^*)$ with $g_1 < g_1^* - (\phi/2)$, and $g_2 > \text{minmax } g_2 - \varepsilon$. Therefore, it must be the case that if $E[v_2^{t+1} \mid h^t, \omega^*] \geqslant \text{minmax } g_2 - \varepsilon$, then

$$E[v_1^{t+1}(\delta_2) \mid h^t, \omega^*] \geqslant g_1^* - \frac{\phi}{2}. \tag{14}$$

Recall that $v_1^{t+1}(\delta_2)$ signifies the discounted sum of player 1's payoffs, calculated using δ_2 to discount instead of δ_1. [The left hand side of the inequality (14) is just the right hand side of equation (13) with g_2 replaced by g_1.] Consequently (14) cannot fail more than NK times (conditional on ω^*). Next,

$$E[v_1^{t+1}(\delta_2) \mid \omega^*] = E[(1-\delta_2)\, g_1(a_1^*, a_2^{t+1}) + \delta_2 v_1^{t+2}(\delta_2) \mid \omega^*], \tag{15}$$

so

$$(1-\delta_2)\, E[g_1(a_1^*, a_2^{t+1}) \mid \omega^*] = E[v_1^{t+1}(\delta_2) - \delta_2 v_1^{t+2}(\delta_2) \mid \omega^*]. \tag{16}$$

Hence, the payoff to the normal type of player 1 if she follows the commitment strategy is

$$\begin{aligned}
E[v_1 \mid \omega^*] &= (1-\delta_1) \sum_{t=1}^{\infty} \delta_1^{t-1} E[g_1(a_1^*, a_2^t) \mid \omega^*] \\
&= E\left[\sum_{t=1}^{\infty} \delta_1^{t-1} \frac{1-\delta_1}{1-\delta_2} [v_1^t(\delta_2) - \delta_2 v_1^{t+1}(\delta_2)] \mid \omega^* \right] \\
&= \frac{1-\delta_1}{1-\delta_2} \Bigg\{ E[v_1^1(\delta_2) \mid \omega^*] \\
&\quad + E\left[\sum_{t=1}^{\infty} E[\delta_1^{t-1}(\delta_1 - \delta_2)\, v_1^{t+1}(\delta_2) \mid h^t, \omega^*] \mid \omega^* \right] \Bigg\}.
\end{aligned} \tag{17}$$

Recall that (14) cannot fail more than NK times with probability one. Hence, for $\delta_1 > \delta_2$ and with probability one, the random variable

$$\begin{aligned}
&\sum_{t=1}^{\infty} E[\delta_1^{t-1}(\delta_1 - \delta_2)\, v_1^{t+1}(\delta_2) \mid h^t, \omega^*] \\
&\qquad \geqslant \frac{\delta_1 - \delta_2}{1-\delta_1} \left(g_1^* - \frac{\phi}{2} \right) - (\delta_1 - \delta_2)(g_1^* - \underline{g}_1^*)\, NK.
\end{aligned} \tag{18}$$

Using this in (17) and taking the limit as $\delta_1 \to 1$ yields $\lim_{\delta_1 \to 1} E[v_1 \mid \omega^*] \geqslant g_1^* - (\phi/2)$. Choosing $\underline{\delta}_1$ such that the left hand side of this expression is within $\phi/2$ of its limit, we have for $\delta_1 \geqslant \underline{\delta}_1$ that $E[v_1 \mid \omega^*] \geqslant g_1^* - \phi$. Consequently, by mimicking type ω^*, player 1 is guaranteed a payoff of $g_1^* - \phi$. The result follows.

APPENDIX B: PROOF OF THEOREM 2

We will use the following perturbation of G in which player 1 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 2.

The automaton that implements $\hat{s}_1$ has a finite number of states, M. 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 2 $(a_2^1, a_2^2, ..., a_2^{\hat{t}})$ that will "steer" the automaton into a state $\hat{x} \in Z$ at time $\hat{t}+1$ within $\hat{t} \leqslant M-1$ periods. Let $\hat{h}^{\hat{t}}$ be the corresponding history: $\hat{h}^{\hat{t}} = ((\hat{s}_1(h^0), a_2^1), (\hat{s}_1(\hat{s}_1(h^0), a_2^1), a_2^2), \ldots)$; 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 2 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 1 is following the automaton strategy and that player 2 is following an arbitrary pure strategy $\hat{s}_2$ (this clearly does not depend upon which state in Z is the initial state). Define

$$P = \left\{ \begin{array}{l} \lim\limits_{T\to\infty} \sum\limits_{t=\hat{t}}^{T} \dfrac{(g_1(a_1^t, a_2^t), g_2(a_1^t, a_2^t))}{(T-\hat{t})} \Bigg| \\ \quad \text{player 1 plays } \hat{s}_1\text{, player 2 plays any } s_2 \\ \quad \text{such that } \hat{h}^{\hat{t}} \text{ is followed and the limit exists} \end{array} \right\}.$$

Notice that P depends upon $\hat{s}_1$ and also on the set Z chosen when Z is not unique. Clearly $P \subseteq F$. The following lemma describes some properties of the limiting set P which we shall need.

LEMMA 3. (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 2 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 $\tilde{v}_1 \leqslant g_1^*$ and $\tilde{v}_2 \geqslant$ minmax 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 2 can play a strategy against each action of player 1 which mimics the response played in the definition of the lower bound.[21]

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 2 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 $K \leqslant M$ (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 minmax $g_1 < \hat{v}_1 < g_1^* + (\varepsilon/2)$, $\hat{v}_2 >$ minmax g_2.

In this case it is possible to construct a pooling equilibrium in which player 1 follows the automaton strategy, player 2 plays a strategy corresponding to $(\hat{v}_1, \hat{v}_2)$, and player 1 is minmaxed should should she reveal herself. Consider the following strategies for the normal types:

Equilibrium path. Player 1 follows the automaton strategy. Player 2 plays according to $\hat{h}^{\hat{t}}$ and from time $\hat{t}+1$ onwards plays as in Lemma 3(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 1 follow the same pure strategy, and so any deviation from this is punished by player 2, who plays to minmax *the normal type* of player 1. Likewise the normal type of player 1 minmaxes player 2 should he deviate. Any deviation by player 2 will result in an expected continuation payoff of at most $(1-\mu)$ minmax $g_2 + \mu \bar{g}_2$ 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

$$(1-\delta_2')\,\bar{g}_2 + \delta_2'((1-\bar{\mu}) \text{ minmax } g_2 + \bar{\mu}\bar{g}_2)$$
$$\leqslant (1-(\delta_2')^{M-1})\,\underline{g}_2 + (\delta_2')^{M-1}\,(\hat{v}_2 - \eta'), \tag{19}$$

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

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 $\hat{t} \leqslant M-1$. Likewise, a δ_1'' and an η'' can be found such that a corresponding inequality holds for player 1 (who is minmaxed with probability 1 after deviation). Choose $\eta = \min\{\eta', \eta''\}$ and letting $\tilde{\bar{\delta}}$ be as in Lemma 3(ii), set $\bar{\delta} = \max\{\tilde{\bar{\delta}}, \delta_2', \delta_1''\}$. Hence for $\bar{\delta} < \delta_1, \delta_2 < 1$, $0 < \mu < \bar{\mu}$ the above strategies are feasible and constitute a Nash equilibrium.

Case 2. There exists $(\hat{v}_1, \hat{v}_2) \in P$ with minmax $g_1 \geqslant \hat{v}_1$ and $\hat{v}_2 >$ minmax g_2, but no $(v_1, v_2) \in P$ with $v_1 >$ minmax g_1, $v_2 >$ minmax 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 no more than M periods. For high discount factors, the payoffs generated by this must approximately lie in P—see Corollary 1 below—and hence in Case 2 offer player 1 no more than her minmax payoff (since P is convex). This means that player 1 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 2 that he is playing against the automaton; player 2 would therefore play a best-response which would give player 1 a lower payoff than she would get by revealing her type.

LEMMA 4. *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* 2 *of* $\max_{(v_1, v_2) \in P} v_2$.

Proof. Consider a best response with cycle $K \leqslant M$ starting from $x \in Z$, and let the payoffs to player 2 in the cycle be $g_2^1, g_2^2, ..., g_2^K$. Discounted payoffs starting from the first point in the cycle are $\sum_{i=1}^{K} \delta_2^{i-1} g_2^i / \sum_{i=1}^{K} \delta_2^{i-1}$, which converge to the average payoff $(1/K) \sum_{n=1}^{K} g_2^n$ as $\delta_2 \to 1$. Since there are only a finite number of such cycles, above a critical discount factor $\hat{\delta}_2$ a best response cycle in the non-discounting case must also be a best response cycle in the discounting case. Note that a best response in the non-discounting case yields a payoff to player 2 of $\max_{(v_1, v_2) \in P} v_2$. Hence, if the discount factor is sufficiently close to 1, a best response in the discounting case corresponds to a point in P such that v_2 is maximized. Q.E.D.

COROLLARY 1. *As* $\delta_1, \delta_2 \to 1$ *from below, all continuation payoffs 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 $(\hat{v}_1, \hat{v}_2) \in F$ such that minmax $g_1 < \hat{v}_1 < g_1^* + (\varepsilon/2)$, $\hat{v}_2 >$ minmax g_2 (this is always possible by Assumption 1 and the convexity of F). From the corollary and by definition of Case 2 there exists an $\alpha > 0$ satisfying $\alpha < \varepsilon$ and $\hat{v}_i - (\alpha/2) \geqslant$ minmax g_i for $i = 1, 2$, and 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 $\hat{x} \in Z$ satisfying $v_1^* < \hat{v}_1 - \alpha$.

Consider the following strategies:

Equilibrium path. Both players follow $\hat{h}^{\hat{t}}$ for the first $\hat{t} - 1$ periods. At time $\hat{t}$ player 1 normal type reveals her 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 within $\alpha/2$ of $(\hat{v}_1, \hat{v}_2)$. If $a_1^{\hat{t}} = \hat{s}_1(\hat{h}^{\hat{t}-1})$ then player 2 plays a best response against the automaton strategy as in Lemma 4.

Off equilibrium path. Deviations are minmaxed.

The idea is to reward player 1 for revealing her type with a payoff $\hat{v}_1$ and if she mimics the automaton she receives a payoff v_1^* close to her minmax payoff.

For δ_1, δ_2 sufficiently close to 1 these strategies are feasible (that is, payoffs within $\alpha/2$ of $(\hat{v}_1, \hat{v}_2)$ can be achieved which are an equilibrium in the complete information game), and additionally with μ close to zero it will pay neither to deviate - see the argument for Case 1 for player 1; for player 2 the expected continuation payoffs after period $\hat{t}$, compared to the minmax threat if player 1 is the normal type, will prevent deviation during $\hat{h}^{\hat{t}}$ for μ small enough, and thereafter if player 1 follows the automaton strategy player 2 cannot gain by deviating because he is by assumption playing a best response, and if player 1 is revealed to be the normal type then the continuation game by assumption is in equilibrium so deviations are not profitable. Player 1 has a new deviation possibility however: the option of mimicking the automaton at time $\hat{t}$ and thereafter, but, if $\delta_1, \delta_2 > \tilde{\delta}$, she will suffer a loss of continuation payoff of at least $(\hat{v}_1 - (\alpha/2) - v_1^*) > \alpha/2$, which for δ_1 sufficiently close to 1 will make this deviation unprofitable (likewise for deviations after first mimicking the automaton).

Finally, for δ_1 near 1, the overall payoff to player 1 from adhering to the above strategy will be within $\alpha/2$ of the continuation payoff after $\hat{t}$, and consequently no more than $\hat{v}_1 + \alpha < g_1^* + \varepsilon$. Hence there exists $\bar{\delta}$ 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 1 a payoff less than $g_1^* + \varepsilon$.

Provided that P contains points strictly above minmax 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 $\bar{v}_1 \geqslant g_1^* + (\varepsilon/2)$, $\bar{v}_2 >$ minmax g_2. However, then from Lemma 3(iii) there is a point $(\tilde{v}_1, \tilde{v}_2) \in P$ satisfying $\tilde{v}_1 \leqslant g_1^*$, $\tilde{v}_2 \geqslant$ minmax 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 3(i), and satisfying Case 1. Finally there is:

Case 3. There does exist $(v_1, v_2) \in P$ such that $v_2 >$ minmax g_2.

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

LEMMA 5. *In Case* 3, *for given* δ_1, δ_2, *after the history* $\hat{h}^{\hat{t}}$ *there is a best response by player* 2 *against the automaton which implies that, for all* $t > \hat{t} + M - 1$, a_1^t *is a minmax strategy against player* 2 *and* a_2^t *is such that* $g_2(a_1^t, a_2^t) =$ minmax 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, $K \leqslant M$. Suppose player 1 does not play a minmax strategy against player 2 every period in the cycle. Then, player 2's best response must achieve a payoff greater than minmax 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 minmax g_2 —contradicting the definition of Case 3. Q.E.D.

It follows that, once the cycle starts, a best response strategy for player 2 is simply *any* short-run best response to the automaton action period by period as this guarantees player 2 his minmax payoff each period. Note that by assumption 1 there exists $(\tilde{v}_1, \tilde{v}_2) \in F$, such that $g_1^* < \tilde{v}_1 < g_1^* + \varepsilon$ and minmax $g_2 < \tilde{v}_2$. Approximating 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 1 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 $(\hat{v}_1, \hat{v}_2)$ satisfying $g_1^* + \alpha < \hat{v}_1 < g_1^* + (\varepsilon/2)$, $\hat{v}_2 >$ minmax $g_2 + \alpha$, where α satisfies $0 < \alpha < \varepsilon/2$. If $a_1^{\hat{t}} = \hat{s}_1(\hat{h}^{\hat{t}-1})$ then thereafter player 2 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 minimizes player 1's payoff.

Off equilibrium path. Deviations are minmaxed.

An α can be found such that there is a complete information game equilibrium as specified above for all δ_1, δ_2 near enough to 1. As before, deviation from the above strategies is unprofitable when μ is close to zero and δ_1, δ_2 are close to 1. Mimicking the automaton from period $\hat{t}$ gives player 1 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 1. Hence there exists $\bar{\delta}$ 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 1 a payoff less than $g_1^*+\varepsilon$.

References

1. M. Aoyagi, Reputation and dynamic Stackelberg leadership with random public signals, mimeo, Univ. of Pittsburg, Nov. 1993.
2. M. Celentani, D. Fudenberg, D. Levine, and W. Pesendorfer, Maintaining a reputation against a patient opponent, mimeo, Harvard Univ., Dec. 1993.
3. M. Cripps, K. Schmidt, and J. Thomas, Reputation in perturbed repeated games, Discussion Paper No. A-410, SFB 303, Bonn Univ., June 1993.
4. M. Cripps and J. Thomas, Reputation and commitment in two-person repeated games without discounting, *Econometrica* **63** (1995), 1401–1420.
5. M. Cripps and J. Thomas, Reputation and perfection in repeated common interest games, mimeo, Warwick Univ., Feb. 1995.
6. D. Fudenberg and D. K. Levine, Reputation and equilibrium selection in games with a patient player, *Econometrica* **57** (1989), 759–778.
7. D. Fudenberg and D. K. Levine, Maintaining a reputation when strategies are imperfectly observed, *Rev. Econ. Stud.* **59** (1992), 561–579.
8. D. Fudenberg and D. K. Levine, Self-confirming equilibrium, *Econometrica* **61** (1993), 523–545.
9. D. Fudenberg and E. Maskin, The Folk Theorem in repeated games with discounting or with incomplete information, *Econometrica* **54** (1986), 533–554.
10. K. Schmidt, Reputation and equilibrium characterization in repeated games of conflicting interests, *Econometrica* **61** (1993), 325–351.
11. K. Schmidt, Commitment through incomplete information in a simple repeated bargaining game, *J. Econ. Theory* **60** (1993), 114–139.