THE FOLK THEOREM IN REPEATED GAMES OF INCOMPLETE INFORMATION

Martin W. Cripps and Jonathan P. Thomas

No.439

WARWICK ECONOMIC RESEARCH PAPERS

THE UNIVERSITY OF WARWICK

DEPARTMENT OF ECONOMICS

THE FOLK THEOREM IN REPEATED GAMES OF INCOMPLETE INFORMATION

Martin W. Cripps and Jonathan P. Thomas
Department of Economics
University of Warwick
Coventry CV4 7AL
England

No.439

July 1995

This paper is circulated for discussion purposes only and its contents should be considered preliminary.

The Folk Theorem in Repeated Games of Incomplete Information

Martin W. Cripps and Jonathan P. Thomas*
University of Warwick,
Coventry CV4 7AL, UK.

March 1995
Revised June 1995

ABSTRACT: The paper analyzes the Nash equilibria of discounted repeated games with one-sided incomplete information. If the informed player is arbitrarily patient relative to the uninformed player, then the characterization is essentially the same as that in the undiscounted case. This implies that even small amounts of incomplete information can lead to a discontinuous change in the equilibrium payoff set. For the case of equal discount factors, however, a result akin to the folk theorem holds when a complete information game is perturbed by a small amount of incomplete information.

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

JEL CLASSIFICATION NUMBERS: C73, D83, L14.

* Our thanks are due to seminar participants at Erasmus and Carlos III Madrid and also to Marco Celentani for their comments and suggestions.

1. Introduction

The folk theorem for discounted repeated games of complete information asserts that any feasible strictly individually rational payoff vector can be sustained in equilibrium provided both players are sufficiently patient. A natural question which arises from the literature surrounding the folk theorem is whether a similar result holds when there is asymmetric information about the players' preferences over outcomes in the game. This is of interest in its own right; it has implications, for example, in situations where agents are repeatedly trading and valuations are uncertain. It is also important for the question of how robust complete information results are to *small* perturbations of the information structure. Suppose there is a small chance that the opponent's preferences might be different from their representation in some complete information game: does this small perturbation lead to a substantial change in the predictions of the model?

In this paper we consider discounted repeated games between two players when the stage-game payoff matrix of one of the players in unknown to the other player. As the players become patient, we shall investigate to what extent the (Nash) equilibria of the game can be characterized by a corresponding folk theorem result. We find that the answer to this question depends critically upon what is assumed about the relative rates of patience of the two players. If the players' discount factors are equal, then as they become very patient we can establish a continuity result for small perturbations of complete information games. If the informed player is sufficiently patient relative to the uninformed player, however, it turns out that the characterization is different, and such a continuity result no longer holds; we establish that all equilibria are approximately payoff equivalent to equilibria in which the informed player acts to reveal her information at the start of the game (see the end of the following subsection for a more complete summary of the main results).

1.1. An Example and Statement of Main Results

In this subsection, to motivate the main ideas of the paper, we shall analyse a simple example of an incomplete information game. While the general argument is much more complex, the example illustrates well the main issues and results. (The relationship between the current paper and the existing literature is deferred until the next section.) We consider the game in Figure 1 in which there are two players, 1 (row player) and 2 (column player). Player 1 has a stage-game payoff matrix given by either A_1 or A_2, while player 2 has the payoff matrix B. As in the usual Bayes-Nash approach, we imagine that nature selects A_κ, $\kappa = 1$ or 2, before the start of the game with probabilities $(1-p)$ and p respectively $(0 < p < 1)$ and only player 1 observes this choice; hence p is player 2's prior belief that player 1 is of "type 2". Each period $t = 0, 1, 2, \ldots$ the simultaneous move stage game (A_κ, B) is played out, and it is assumed that players can observe all previous actions. For the moment assume that there is a common discount factor δ, $0 < \delta < 1$, with the payoff at time t being weighted by $(1-\delta)\delta^t$. Apart from the incomplete information, the game is a standard discounted repeated game (see Section 2 for the formal description).

	L	R
U	3, 1	0, 0
D	0, 0	1, 3

(A_1, B)

	L	R
U	1, 1	1, 0
D	0, 0	0, 3

(A_2, B)

Figure 1 – Battle of the Sexes

This game is a perturbed version of Battle of the Sexes Game, in which type 2 is a "commitment type" of player 1, existing with probability p; playing the top row (U) is a dominant strategy in the repeated game for this type.[1] It will become clear that a crucial issue in characterizing the possible equilibria is the *speed of revelation of the informed player's information*. To start, consider first a *fully revealing Nash equilibrium* in which the types distinguish themselves in the first period (i.e. type 1 chooses D and type 2 chooses U). What conditions must such an equilibrium satisfy? After the first period,

[1]While the analysis below does not generally restrict the payoff matrices of the various types of player 1, such a commitment type simplifies the example in two important ways. First, on the equilibrium path, type 2 always chooses U, and consequently we need only concentrate on the behaviour of type 1. Secondly, punishing player 1 for deviating is simple since only type 1 needs to be punished.

$t = 0$, the game essentially becomes one of two possible complete information games, depending on which type is revealed. If type 2 is revealed, the equilibrium thereafter involves U being played each period. Clearly player 2 must respond so as to get at least his minmax payoff ($= 3/4$) from $t = 1$ onwards (i.e., the equilibrium must be individually rational for both players). On the other hand, if type 1 is revealed, in addition to the usual constraints on equilibria of a complete information game (essentially, again, individual rationality for both players), there is an *incentive compatibility* constraint which must be observed. Type 1 has the option at $t = 0$ of mimicking type 2. Consequently the incentive compatibility constraint requires that the payoff type 1 gets from revealing her type and playing out some complete information equilibrium of the game (A_1, B) must be at least as great as the payoff she would get from mimicking type 2. This latter payoff, discounted to period $t = 1$, is at least $9/4$, since in the type 2 equilibrium player 2 must play L (against U) at least three quarters of the time. Hence, in any fully revealing equilibrium, for δ near one, type 1 must receive at least approximately $9/4$. Since this argument is independent of p, it follows that even small perturbations of the repeated Battle of the Sexes game only have fully revealing equilibria in which type 1 does substantially better than her minmax payoff of $3/4$.

It is a result from the theory of *undiscounted* incomplete information games (of the type considered here) that *all* equilibria are payoff equivalent to revealing equilibria; that is, even for equilibria where revelation does not occur immediately, there are payoff- equivalent revealing equilibria (see Section 3 for a full discussion of these results). We shall show that this is not the case with symmetric discounting. Indeed for small perturbations of a complete information game there will be equilibria, not (in general) involving immediate revelation, which give overall payoffs close to any payoffs which are feasible and individually rational in the complete information game, as δ goes to one. Hence for symmetric discounting there is a continuity result for small perturbations of the information structure of complete information games.

To see how type 1 can receive payoffs substantially below $9/4$ in the example when δ is close to one, it is only necessary to *delay* revelation, so both types of player 1 pool on U for $t = 0, 1, \ldots, T-1$, and at $t = T$ type 1 reveals her type by playing D. If type 2 is revealed at time T, then an equilibrium of the complete information game between type

2 and player 2 is played out from $T+1$ which gives player 2 a payoff of 0.8 (discounted to $T+1$); this is possible for δ close enough to one. Notice that if type 1 was to mimic type 2, she would receive a payoff of 2.4. If type 1 is revealed, then an equilibrium of the repeated Battle of the Sexes game is played out, in which type 1 receives 2.5 and player 2 receives 1.5 — an equilibrium for δ near one.[2] At T player 2 plays R, and so incentive compatibility is satisfied. (A deviation by player 2 at T is not profitable for δ near one assuming minmax punishments are used.)

Hence we have constructed a revealing equilibrium starting at $t = T$, provided δ is close to one. In this equilibrium type 1 receives a payoff of $(1-\delta) + \delta(2.5)$, and player 2 receives $p \cdot \delta(0.8) + (1-p)((1-\delta)(3) + \delta(1.5))$. For δ close to one and p close to zero, these payoffs are respectively approximately 2.5 and 1.5. Next, for $t < T$, (U, R) is played out by both types. Choose T so that the weight in the discounted sum placed on the first T periods, $(1-\delta)\sum_{t=0}^{T-1}\delta^t$, is approximately 0.4, so player 2 gets a payoff of approximately $(0.4)(0) + (0.6)(1.5) = 0.9$ provided p is very small, and hence the threat of minmax punishments will prevent a deviation. Type 1 of player 1 gets approximately $(0.4)(0) + (0.6)(2.5) = 1.5$. So provided δ is close to one and p is close to zero, an equilibrium can be constructed in which type 1 gets considerably less than 9/4, her minimum payoff in a revealing equilibrium. Revelation still occurs, but not immediately.

Is it possible to drive type 1 down to her minmax payoff of 3/4 in such a pure strategy equilibrium? The answer is no. The worst equilibrium for type 1 would have a revealing equilibrium from T which gave her exactly 9/4 and which gave player 2 the highest feasible payoff consistent with this, 7/4.[3] The initial pooling phase can have a weight in the discounted sum of no more than 4/7 since player 2 then gets overall at most $(3/7)(7/4) = 3/4$ when p is very small (a larger p can only reduce the weight on the pooling phase); any extension of the pooling phase would imply that player 2's overall payoff drops below his minmax payoff. This gives a lower bound on type 1's payoff of $(3/7) \cdot (9/4) = 27/28 > 3/4$. It is easily checked that no other configuration can lead to a lower payoff.

To construct an equilibrium in which type 1 is held close to her minmax payoff it

[2] Because this is a revealing equilibrium from T onwards, type 1 must receive at least 9/4 as argued above.

[3] It is being assumed that δ is close to one so the actual period of revelation is insignificant for payoffs.

is necessary not only to delay revelation, but to make it *gradual*. The equilibrium will consist of the previously constructed delayed revelation equilibrium, *preceded* by an initial pooling phase of (U, R) which ends with a random move by type 1. More precisely, suppose that at $t = T'$ type 1 plays D with probability 0.5, while player 2 plays R. If type 1 chooses U, the above equilibrium is followed. If type 2 chooses D, a complete information equilibrium is played which gives type 1 the same payoff she would get from playing U, that is, approximately 1.5. Player 2 receives the maximum payoff consistent with this, approximately 2.5 (for δ close to one payoffs at T' are insignificant). Now payoffs discounted to T' are approximately 1.5 for type 1 and $(0.5)(0.9) + (0.5)(2.5) = 1.7$ for type 2. By having the initial phase of (U, R) sufficiently long, type 1's payoff can be reduced close to 0.75, her minmax payoff, without violating individual rationality for player 2 (provided the weight in the discounted sum on the first phase is appromately 0.5; the integer problem is the only constraint which prevents this being achieved exactly). To conclude, for all δ sufficiently close to one, and all p sufficiently close to zero, type 1 can be held arbitrarily close to her minmax payoff.

The example illustrates the need for a mixed-strategy equilibrium to attain certain payoffs — equivalently, revelation of player 1's information must be gradual; this is in contrast to the complete information case where only pure-strategy equilibria are needed. The example also illustrates the purpose served by the gradual information revelation: it permits player 2's overall payoff to be increased relative to the payoff player 2 receives along the path played by type 2. In turn, this permits actions to be taken along that path which are undesirable for type 2 and hence also for type 1. This relaxes the incentive compatibility constraint, which was the constraint which guaranteed type 1 a high payoff in the revealing equilibrium.

The characterization that type 1 can be driven down to her minmax payoff for small perturbations, fails, however, if player 1 is very patient *relative to player 2*. In the example it was argued that type 1's payoff could be held down by gradual revelation because the path followed by type 2 could then be made unattractive from type 1's point of view, thus relaxing the incentive compatibility constraint. If, however, player 1 is very patient relative to player 2, the part of the game in which type 2's path is unattractive becomes insignificant from the point of view of player 1; this implies that the incentive compatibility constraint will again require that type 1 gets at least approximately 9/4 (in this case there is a *discontinuity* with the complete information game as p goes to zero). It is the periods of learning which can be used to hold type 1's payoff down, but for a very patient player 1, the periods of learning are insignificant in the calculation of payoffs.

In what follows, we shall treat these two discounting cases in reverse order to the above discussion. Our first main result states that for arbitrary given initial beliefs, for a fixed value of player 2's discount factor, and for player 1's discount factor sufficiently close to one, the equilibrium payoffs to player 1 (for each of a finite number of types) must approximately satisfy the conditions of a revealing equilibrium. This implies continuity with the undiscounted case (holding prior beliefs constant): as the players' discount factors go to one, if player 1's discount factor goes to one sufficiently fast relative to that of player 2, then the limiting set of equilibrium payoffs for player 1 must satisfy the necessary conditions appropriate for the model with no discounting.

In the second part of the paper, the symmetric discounting case is analysed. Where the informed player has only two types, we establish a continuity result with complete information games as the probability of one of the types goes to zero: for each type of the informed player, and for any feasible strictly individually rational payoff vector in the game between this type and player 2, there is Nash equilibrium of the incomplete information game with these payoffs (to the type and to player 2) provided the players are sufficiently patient and *provided initial beliefs put sufficiently high probability on that type*. Since there is no such continuity result for undiscounted games as the size of the perturbation goes to zero, it can be concluded that the equilibrium characterization which exists for the undiscounted case is only the limit (as discount factors go to one, holding beliefs constant) of the discounted case if the limit is taken in a particular way,[4] and in particular it is *not* the limit of the discounted case if both players' discount factors are equal.

[4]This provides another contrast with complete information games where it can be shown that the projection of the set of equilibrium payoffs onto the more patient player's axis converges, as both discount factors go to one, to the folk theorem projection, so that in terms of the payoffs possible to the more patient player, the way the two discount factors go to one does not matter.

1.2. Relation to the Literature

Complete information discounted repeated games have been studied in great detail. By contrast, the situation where one or more players' preferences may be unknown to the opponent(s) has received relatively little attention in the discounted repeated games literature, despite the attention it has received in both the signalling games and the bargaining literature. Some recent results exist however. Kalai and Lehrer (1993) and Jordan (1995) have established that play must converge to Nash play of the true game.[5] Jordan (1995) has also proved the existence of equilibrium for this class of games. Perfect Bayesian equilibria of such games must have a Markov property (Bergin (1989)). McKelvey and Palfrey (1992, 1993) have studied belief stationary equilibria. Schmidt (1993b) has obtained a strong results in a finitely repeated bargaining game in which he also imposes a belief-stationarity restriction on the equilibrium concept.

The results of Kalai and Lehrer and Jordan on convergence to Nash play are informative about the long-run behaviour of an equilibrium, but to be able to say anything about the overall payoffs from the beginning of the game—what we are interested in here—it is necesssary to know something about how fast convergence takes place relative to the rate of discounting of payoffs and also, possibly, what happens in the shorter run.[6] By exploiting a fundamental result due to Fudenberg and Levine (1992) on the speed of learning, the case where the informed player is arbitrarily patient relative to the uninformed player can be completely solved purely on the basis of "long-run" considerations. As was illustrated in the above example, however, a more detailed consideration of the shorter run is needed for the symmetric discounting case.

A closely related literature is the recent "reputation" literature on games with a small amount of incomplete information of a particular type. Here, complete information games are perturbed with a small possibility that one or more players might be irrational or "commitment" types which are committed to playing a particular *strategy*.[7]

[5] Their approach also covers a broader learning environment than the standard one.

[6] In this respect the undiscounted case is much more straightforward since long-run behaviour and payoffs can be characterised using martingale arguments, and the speed of convergence to the long-run does not affect average payoffs.

[7] It is possible to motivate fixed repeated game strategies as dominant strategies for particular preferences, and even mixed strategy commitment types can be motivated in this fashion (Fudenberg and Levine (1992)).

The question of whether such perturbations can lead to substantial changes in the set of equilibrium payoffs as compared to the corresponding complete information game — in the form of sharper or more realistic predictions—has received much attention since Kreps, Milgrom, Roberts and Wilson's (1982) analysis of the finitely repeated prisoner's dilemma. A contrasting result was established by Fudenberg and Maskin (1986), who demonstrated that any strictly individually rational and feasible payoff vector can be sustained as an equilibrium outcome in a finitely repeated game if the modeler chooses the "right" incomplete information. Recent results, which allow for more general perturbations, include those of Aumann and Sorin (1989), Schmidt (1993a), Cripps, Schmidt and Thomas (1995), Celentani et al. (1994) and Aoyagi (1993). Our results have implications for small perturbations of the information structure of a complete information game with types which are *rational* in the sense maximizing expected discounted stage-game payoffs. They show that small perturbations of a complete information game can lead to large changes—in the form of a lower bound for the informed player—in the equilibrium payoff set provided the informed player is sufficiently patient relative to the uninformed player. This result is reminiscent of recent results in the reputation literature, where a relatively patient informed player is a common assumption (indeed, the characterizations of Schmidt (1993a) and Cripps, Schmidt and Thomas (1995) are implied by our more general results when the perturbation introduces a type with a payoff matrix that makes a particular action dominant in the repeated game). The application of our results allows reputations for being a different normal type (i.e., for having a different stage-game payoff matrix) to be analysed. For equally patient players, however, we establish a continuity result with the complete information game.

In contrast to the discounted literature, *undiscounted* repeated games of incomplete information have been studied in some depth, especially in the zero-sum case and where payoffs are time averaged (see Mertens, Sorin and Zamir (1993)). In the non-zero-sum undiscounted payoffs case Hart (1985) in a seminal paper has given a complete characterization of Nash equilibria for a broad class of two-person games. Hart studied games where one player is informed of both her own and the other player's stage-game payoff function, with the other player uninformed. While this characterization is extremely complex, it has been shown by Shalev (1994) that a simpler characterization obtains for the case of "known own payoffs", where the uninformed player is aware of his or her own payoff function (this is a special case of the general model studied by Hart). This assumption of known own payoffs is the one most often made in economic applications, and it is the case we study here. Shalev's characterization is described in detail in Section 3, but amounts to the proposition that all equilibria are payoff equivalent to fully revealing equilibria, that is equilibria in which all types take actions at the start of the game which reveal their types. Such equilibria can be characterized in a relatively straightforward fashion. As discussed above, we obtain the same characterization as Shalev, in the limit as discounting goes to zero, but only if the limit is taken so that the informed player is arbitrarily patient relative to the uninformed player.

2. The Model

The infinitely repeated game $\Gamma(\mathbf{p}, \delta_1, \delta_2)$ is defined as follows. There are two players called "1" (she) and "2" (he). At the start of the game, player 1's "type" k is drawn from a finite set K (where K also denotes the number of elements) according to the probability distribution $\mathbf{p} = (p_k)_{k \in K} \in \Delta^K$ (the unit simplex of R^K), and informed to player 1. Hence p_k will denote the prior probability of type k. We shall assume that each type has strictly positive probability: $p_k > 0$ for all k. In every period $t = 0, 1, 2, \ldots$, player 1 selects an "action" i^t out of a finite action space I, where I has at least two elements, while player 2 simultaneously chooses an action j^t from the finite set J. Payoffs at stage t to type k of player 1 and to player 2 are respectively$A_k(i^t, j^t)$ and $B(i^t, j^t)$. Player i discounts payoffs with discount factor δ_i, and we normalize payoffs so that stage-game and repeated-game payoffs can be expressed on the same scale. The payoff to type k of player 1 is

$$\tilde{a}_k = (1 - \delta_1) \sum_{t=0}^{\infty} \delta_1^t A_k(i^t, j^t) , \tag{1}$$

and that to player 2 is

$$\tilde{b} = (1 - \delta_2) \sum_{t=0}^{\infty} \delta_2^t B(i^t, j^t) . \tag{2}$$

Throughout the paper we restrict attention to the case where both players observe the realized action profile (i^t, j^t) after each period. Let $H^t = (I \times J)^{t+1}$ be the set of all possible histories h^t up to and including period t. A mixed (behavioral) strategy for type k of player 1 is a sequence of maps $\sigma_k = (\sigma_k^0, \sigma_k^1, \cdots)$, $\sigma_k^t : H^{t-1} \to \Delta^I$. We define $\sigma = (\sigma_k)_{k\in K}$. Likewise, a mixed strategy for player 2 is a sequence of maps $\tau = (\tau^0, \tau^1, \cdots)$, $\tau^t : H^{t-1} \rightarrow \Delta^J$. The prior probability distribution $\mathbf{p}$, together with a pair of strategies (σ, τ), will induce a probability distribution over infinite histories and hence over discounted payoffs. We use $E_{\mathbf{p},\sigma,\tau}$ to denote expectations with respect to this distribution, and abbreviate to E where there is no ambiguity. Players are assumed to maximise expected payoffs, and a Nash equilibrium is defined as a pair of strategies (σ, τ) such that for each k

$$E_{\mathbf{p},\sigma,\tau}\left[\tilde{a}_k \mid k\right] \geq E_{\mathbf{p},\sigma',\tau}\left[\tilde{a}_k \mid k\right] \quad \text{for all } \sigma', \tag{3}$$

(where the conditional expectation given type k uses only the probability distribution induced by σ_k and τ) and

$$E_{\mathbf{p},\sigma,\tau}[\tilde{b}] \geq E_{\mathbf{p},\sigma,\tau'}[\tilde{b}] \quad \text{for all } \tau'. \tag{4}$$

Finally we shall need the following. Let

$$\hat{a}_k = \min_{f\in\Delta^I} \max_{g\in\Delta^J} A_k(f,g) \tag{5}$$

be type k's minmax payoff, where we use the notational abuse that $A_k(f,g)$ is the expected value of $A_k(i,j)$ when mixed actions f and g are followed. Likewise

$$\hat{b} = \min_{g\in\Delta^J} \max_{f\in\Delta^I} B(f,g). \tag{6}$$

3. A Relatively Patient Informed Player

We start by considering the case where the discount factor of player 2 is taken as fixed, and we let the discount factor of player 1, the informed player, go to one. This case corresponds closely to the undiscounted case; necessary conditions which must be satisfied by player 1's payoffs in the undiscounted case must also be (approximately) satisfied in the discounted case as $\delta_1 \rightarrow 1$. These necessary conditions can be interpreted as requiring payoff equivalence to some fully revealing equilibrium. Before we consider the discounted case, we briefly review the existing results in the undiscounted case.

3.1 Review of the Undiscounted Case[8]

Using the limit of the means criterion, with an arbitrary Banach limit, Hart (1985) gave a complete characterization for the general class of games with one-sided incomplete information, which includes the possibility that the uninformed player is unaware of his own payoff function. This characterization is based on bi-martingales. For the case we are interested in, namely where each player is aware of their own payoffs but one of the players does not know the payoffs of the other player,[9] a much simpler characterization has been found, using Hart's results, by Shalev (1994), as discussed in the Introduction.[10] Denote this game by $\Gamma(\mathrm{p}, 1, 1)$. We shall show that essentially the same characterization as that of Shalev can be obtained for the discounted case provided the informed player is arbitrarily patient relative to the uninformed player.

We define first individual rationality in this setting. For $\mathrm{q} \in \Delta^K$, let $a(\mathrm{q})$ be player 1's minmax payoff in the one-shot game with payoffs given by $\sum_{k\in K} q_k A_k(i,j)$. A vector payoff $\mathrm{x} = (x_k)_{k\in K}$ is said to be *individually rational* for player 1 if

$$\mathrm{q}\cdot\mathrm{x} \geq a(\mathrm{q}) \quad \text{for all } \mathrm{q} \in \Delta^K. \tag{7}$$

This condition is, by Blackwell's (1956) *approachability theorem*, necessary and sufficient for player 2 to have a strategy $\hat{\tau}$ against which no type $k \in K$ can achieve a limiting average payoff greater than x_k by any choice of strategy, that is, $E_{\mathrm{p},\sigma,\hat{\tau}}[a_k \mid k] \leq x_k$ for all σ and k. For player 2 the definition of individual rationality is the usual one from complete information repeated games: a payoff y for player 2 is *individually rational* if

$$y \geq \hat{b}. \tag{8}$$

Let $\pi = (\pi^{ij})_{i,j} \in \Delta^{I\times J}$ be a joint distribution over $I \times J$ (i.e. a correlated strategy). This will generate a vector of payoffs for player 1 and a payoff for player 2 of

$$A_k(\pi) = \sum_{i\in I, j\in J} \pi^{ij} A_k(i,j)\ , \tag{9}$$

$$B(\pi) = \sum_{i\in I, j\in J} \pi^{ij} B(i,j). \tag{10}$$

[8]See Forges (1992) for an excellent survey of the literature.

[9]That is, in the general case the informed player knows both her own and the opponent's payoff matrix, while the opponent knows only the probability distribution over the possible types of the informed player; here the uninformed player's payoff matrix does not vary with the informed player's type.

[10]For a direct proof of these results which does not use the bi-martingale approach, see Koren (1988).

Let $\Pi = (\Delta^{IJ})^K$ be the set of all correlated strategy profiles for each type, $(\pi_k)_{k \in K}$. Then

Definition 1 *Define $\Pi_0 \subset \Pi$ to be the subset of profiles satisfying conditions* (i) *(individual rationality): $(A_k(\pi_k))_{k \in K}$ is individually rational for player 1, and $B(\pi_k)$ is individually rational for player 2 for each $k \in K$, and* (ii) *(incentive compatibility): $A_k(\pi_k) \geq A_k(\pi_{k'})$ for all $k, k' \in K$.*

We can state

Result 1 (Shalev) *Payoffs* $(\mathbf{a}, b)$ *are Nash equilibrium payoffs of* $\Gamma(\mathbf{p}, 1, 1)$ *if and only if there exists a profile of correlated strategies $(\pi_k)_{k \in K} \in \Pi_0$ such that $A_k(\pi_k) = a_k$ for all $k \in K$ and $\sum_{k \in K} p_k B(\pi_k) = b$.*

The basic idea behind the result is as follows. Every Nash equilibrium is payoff equivalent to a *completely revealing joint plan*, which involves each type k of player 1 revealing its type at the beginning of the game and then playing out a deterministic equilibrium which has relative frequencies over action profiles corresponding to the joint distribution π_k. The revelation is achieved by constructing deterministic strategies for each type which differ during a finite number of periods at the start of the game. Since the payoffs are time-averaged, this revelation phase does not affect average payoffs.

It is straightforward to show that if payoffs satisfy the conditions of the result, then a completely revealing joint plan can be constructed which is an equilibrium and which delivers these payoffs. Condition (i) of the definition of Π_0 says that following the equilibrium gives each type of player 1 and also player 2 individually rational payoffs. Since all strategies are deterministic, deviations by player 2 are punished by minmaxing him; this makes deviations unprofitable. Deviations by player 1 from one of the k possible equilibrium strategies (i.e. playing an action inconsistent with *any* type dependent strategy) are punished by player 2 playing his Blackwell strategy corresponding to the equilibrium payoff vector $\mathbf{a}$. This ensures that no type can benefit from such a deviation. Condition (ii) is an incentive compatibility requirement; no type prefers to mimic another type's strategy. Consequently type k can benefit neither by mimicking one of the other types, nor by playing some other strategy, and so we have a Nash equilibrium. The converse, namely that any equilibrium payoffs satisfy the conditions of the result, is considerably more difficult to establish.

3.2 The Discounted Case

In this subsection we shall show that the characterization for the undiscounted case also holds for the discounted case provided player 1 is sufficiently patient relative to player 2. We start by establishing two preliminary results. First, in Lemma 1, we show that if player 2's equilibrium strategy gives him less than $\hat{b}$ when he plays against k, then he must expect that the probability distribution over outcomes generated by the repeated play of type k's strategy differs from the one generated by the "expected" equilibrium strategy of player 1, where the expectation is taken over all possible types using player 2's beliefs. Furthermore, because player 2 discounts future payoffs, there must be a significant difference between these distributions in the not too distant future. The second result (Result 2) is based on a fundamental proposition of Fudenberg and Levine (1992), which establishes that if player 1 follows type k's strategy, then player 2 cannot continue to believe that the true probability distribution over outcomes is significantly different from the one generated by type k's strategy. Taken together, these results imply that if player 1 always plays according to type k's strategy, then player 2 cannot continue to respond with a strategy which gives him less than $\hat{b}$ against this strategy. Eventually he will learn that his opponent plays type k's strategy, and he will choose a response which gives him at least his minmax payoff.

Below, we shall heavily exploit the fact that player 2 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 2 could continue to play a strategy which gives him less than his minmax payoff against type k's strategy if he believes that there is a high enough probability that player 1 is playing according to some other strategy. To be more precise: It may be the case that the equilibrium strategy of player 2 yields strictly less than $\hat{b}$ against type k, as long as it yields at least $\hat{b}$ in expectation against all types of player 1, where the expectation is taken according to the beliefs of player 2. However, the following lemma says that in this case the strategy of type k and the expected strategy of player 1 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 2 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 2, paired with first the repeated play of type k's strategy and secondly with the expected equilibrium strategy of player 1, are arbitrarily close to each other. Then the distribution over payoffs for player 2 would be almost the same in both cases. Thus, if he gets strictly less than $\hat{b}$ against k, he must also get less than $\hat{b}$ against the expected equilibrium strategy of player 1, a contradiction.

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

$$y^N = \left((i^{t+1}, j^{t+1}), \ldots, (i^{t+N}, j^{t+N})\right) . \tag{11}$$

For given equilibrium strategies (σ, τ) we let $\mathrm{q}^N(\cdot \mid h^t)$ be the distribution over these outcomes and likewise $\mathrm{q}^N(\cdot \mid h^t, k)$ the distribution conditional additionally upon player 1's true type being k (defined for h^t having positive probability conditional on type k). We define for any two distributions q^N and $\hat{\mathrm{q}}^N$,

$$\| \mathrm{q}^N - \hat{\mathrm{q}}^N \| := \max_{y^N} \left| \mathrm{q}^N(y^N) - \hat{\mathrm{q}}^N(y^N) \right| . \tag{12}$$

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

$$\tilde{a}_k^{t+1} = (1 - \delta_1) \sum_{\tau=t+1}^{\infty} \delta_1^{\tau-t-1} A_k(i^\tau, j^\tau) , \tag{13}$$

and that for player 2 as

$$\tilde{b}^{t+1} = (1 - \delta_2) \sum_{\tau=t+1}^{\infty} \delta_2^{\tau-t-1} B(i^\tau, j^\tau) . \tag{14}$$

Lemma 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 type k. Suppose that conditional upon player 1 being type k the expected continuation payoff for player 2 is*

$$E\left[\tilde{b}^{t+1} \mid h^t, k\right] \;\leq\; \hat{b} - \epsilon\,. \tag{15}$$

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

$$\| \, \mathrm{q}^N(\cdot \mid h^t) - \mathrm{q}^N(\cdot \mid h^t, k) \, \| \; > \; \eta\,. \tag{16}$$

Proof: See Appendix.

The next result shows that if player 1 follows the strategy of type k, then there can be only a finite number of periods in which the probability distribution over outcomes predicted by player 2 differs significantly from the true distribution. Eventually, player 2 will predict future play (almost) correctly.

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

Result 2 (Fudenberg and Levine) *Given integers N and n, with $N > 0$ and $0 \leq n < N$, and for every $\xi > 0$, $\psi > 0$ and a type k of player 1 with $p_k > 0$, there is an m depending only on N, ξ, ψ, and p_k such that for any (σ, τ) and any h^t consistent with (σ, τ), the probability, conditional on player 1's true type being k, that there are more than m periods $t \in T(n, N)$ with*

$$\| \, \mathrm{q}^N(\cdot \mid h^t) - \mathrm{q}^N(\cdot \mid h^t, k) \, \| \; > \; \psi \tag{17}$$

is less than ξ.

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

We are now able to establish that if player 1 follows type k's strategy, then player 2 must eventually respond so as to get at least his minmax payoff against type k's strategy. Consider a history of any Nash equilibrium in which player 1 is type k. By Lemma 1 we know that, on the one hand, player 2 can reply with a strategy which yields less than $\hat{b}$ against type k's strategy only if he believes that the probability distribution over outcomes generated by player 1's expected equilibrium strategy differs significantly from the one generated by type k's strategy over the next N periods. On the other hand, by Result 2 we know that there are at most m periods in which player 2 may believe that there is a significant difference if the true distribution is the one generated by type k's strategy. Thus, there can be at most $m \cdot N$ periods, $m \cdot N < \infty$, in which player 2 expects player 1's strategy to be significantly different from type k's strategy, and consequently if a sufficiently high discount factor (i.e. δ_1 as oppposed to δ_2) is used to evaluate player 2's payoffs, these $m \cdot N$ periods will be insignificant and player 2 must get approximately his minmax payoff against type k. [11]

First, for a fixed equilibrium, we define the average frequencies over action profiles conditional on type k when the discount factor is δ as follows:

$$\pi_k^{ij}(\delta) = (1-\delta) E\left[\sum_{t=0}^{\infty} \delta^t 1(i,j,t) \middle| k\right], \tag{18}$$

for each i and j, where $1(i,j,t)$ is the indicator function for the action profile (i,j) occurring at date t. It is easy to check that the equilibrium payoffs are $E[\tilde{a}_k \mid k] = A_k(\pi_k(\delta_1))$ for each k and $E\left[\tilde{b}\right] = \sum_{k \in K} p_k B(\pi_k(\delta_2))$.

Lemma 2 *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 average frequencies over action profiles for each $k \in K$ in any Nash equilibrium, calculated using discount factor δ_1, $\pi_k(\delta_1)$, satisfy*

$$B(\pi_k(\delta_1)) \geq \hat{b} - \phi\ . \tag{19}$$

[11]The formal argument is a bit more involved. Since type k's strategy may be mixed we can only say that there will be more than $m \cdot N$ such periods with probability less than ξ.

Proof: See Appendix.

Given the result of Lemma 2, we are now in a position to establish the main result of this section, namely that Shalev's equilibrium characterization holds approximately as a necessary condition provided that player 1 is sufficiently patient relative to player 2. Recall that this implies that the equilibrium is approximately payoff equivalent to a revealing equilibrium. This theorem is a characterization of the equilibrium payoffs of player 1 only; since different discount factors are being used, the usual feasibility constraint on the average payoff profile across both players does not apply here. First we need to define the set of payoff vectors which player 1 can receive in equilibrium in the undiscounted case (i.e., the projection of the equilibrium payoff set onto the space of player 1's payoffs). Recall that Π_0 is the set of all correlated strategy profiles which satisfy individual rationality and incentive compatibility. We define

$$\mathbf{A}^* = \{(A_1(\pi_1), A_2(\pi_2), \ldots, A_K(\pi_K)) : \pi \in \Pi_0\} \ . \tag{20}$$

Theorem 1 *Let δ_2, $0 < \delta_2 < 1$, and $\mathbf{p} \gg \mathbf{0}$ be fixed. Then for any $\epsilon > 0$ there exists a $\underline{\delta}_1 < 1$ such that for all $1 > \delta_1 > \underline{\delta}_1$, if payoffs $\mathbf{a}$ are Nash equilibrum payoffs to player 1 in $\Gamma(\mathbf{p}, \delta_1, \delta_2)$, then $\mathbf{a}$ lies approximately in $\mathbf{A}^*$ in the sense that*

$$min_{\mathbf{x}\in\mathbf{A}^*} \parallel \mathbf{a} - \mathbf{x} \parallel < \epsilon \ . \tag{21}$$

Proof: See Appendix.

Theorem 1 developed necessary conditions which equilibrium payoffs must satisfy asymptotically. In the undiscounted model, the necessary conditions were also *sufficient* (see Result 1). A similar result can be established with discounting provided the inequalities in the conditions of Definition 1 are assumed to hold strictly. We say that a payoff vector $\mathbf{a}$ is *strictly individually rational* for player 1 if there exists some individually rational point $\mathbf{x}$ with $a_k > x_k$ for all k.

We can state

Theorem 2 *Suppose that $(\pi_k)_{k\in K} \in \Pi_0$ satisfies* (i) *(strict individual rationality): $(A_k(\pi_k))_{k\in K}$ is strictly individually rational for player 1, and $B(\pi_k)$*

is strictly individually rational for player 2 for each $k \in K$, and (ii) *(strict incentive compatibility): $A_k(\pi_k) > A_k(\pi_{k'})$ for all $k, k' \in K$. Then for any $\epsilon > 0$ there exists a $\underline{\delta}$ such that whenever $1 > \delta_1$, $\delta_2 > \underline{\delta}$, there exists a Nash equilibrium of $\Gamma(\mathbf{p}, \delta_1, \delta_2)$ with payoffs $(\mathbf{a}, b)$ satisfying $|A_k(\pi_k) - a_k| < \epsilon$ for all $k \in K$ and $|\sum_{k \in K} p_k B(\pi_k) - b| < \epsilon$.*

Proof: Omitted.

The proof is straightforward and follows closely the argument given below Result 1 for constructing a completely revealing joint plan, with each type k revealing itself during the first few periods and thereafter playing approximately according to π_k. One complication which arises is the punishment of player 1. See below in Section 4 for a discussion of Blackwell punishment strategies with discounting.

3.3 Applications to the Reputation Literature

The characterization given in Theorem 1 is independent of the value of $\mathbf{p}$; in particular it holds for even small perturbations of a complete information game. Suppose that p_1 is close to one, and that there are only two types of player 1. Type 1 is a "normal type"[12], with arbitrary preferences A_1, and type 2 a type committed to an action i^* in that her payoffs from the row corresponding to i^* are all equal and strictly greater than any other payoff. Then Theorem 1 has the following implications. Consider Π_0 in this case (see Definition 1): by condition (i) that $A_2(\pi_2)$ is individually rational, π_2 must specify all play occurs in the row corresponding to i^*, that is $\sum_{j \in J} \pi_2^{i^* j} = 1$. Also by condition (i), $B(\pi_2)$ must be individually rational for player 2. Consequently π_2 must specify a probability distribution over the i^* row which gives player 2 at least $\hat{b}$. By condition (ii), incentive compatibility, π_1 must give type 1 at least what she can get from following π_2, so it can be concluded that type 1, the "normal type", must get from any $\pi \in \Pi_0$ at least what she would get in the stage game from playing i^* when player 2 responds with the least favorable (mixed) response to i^* (from type 1's point of view) which gives

[12] We shall use "normal type" to refer to a type such as we have been studying so far, that is, with general stage-game preferences A_k and discount factor δ_1.

player 2 at least his minmax payoff. This is a lower bound on type 1's payoffs in $\mathbf{A}^*$ (see (20)). Applying Theorem 1, for a fixed δ_2 and for any desired degree of approximation, equilibrium payoffs for player 1 must be approximately equal to or greater than this bound for all δ_1 sufficiently close to one. For conflicting interest games,[13] this gives the result of Schmidt (1993) that the Stackelberg payoff can asymptotically be guaranteed by player 1. For general games, this is exactly the bound established in Cripps et al. (1994).

The recent reputation literature typically deals with perturbations which allow a large class, or even an unrestricted class, of types, including irrational types, whereas our results only deal with a finite number of normal types. It is however straightforward to check that the necessary conditions on a Nash equilibrium developed in Theorem 1 remain as necessary conditions even when arbitrary other types of player 1 are present. In that case, the payoff vector $\mathbf{a}$ needs to be interpreted as the payoffs to a particular subset of types of player 1 consisting only of normal types. In view of this observation, the application to commitment types discussed above gives exactly the results (in terms of lower bounds) of the work cited.

An interesting question in the reputation literature arises as to whether building a reputation for being a (different) normal type can provide a patient player with a lower bound on equilibrium payoffs which exceeds that furnished by types of the kind considered above which are committed to playing a constant action. In other words, if the relatively patient player could create some uncertainty about her preferences in the mind of her opponent, would it be better to make the opponent think that she might be a constant action type or a type with more sophisticated preferences? We can answer this question by relating our results to those of Israeli (1989, quoted in Forges, 1992) who analysed the best normal type from the reputation point of view in the undiscounted model (see also Shalev (1994)). The answer turns out to be surprisingly simple: the best lower bound on type 1's payoffs can be obtained by assuming that type 2 has diametrically opposed preferences to player 2: $A_2 = -B$. Moreover this lower bound cannot be improved upon by having additional types with different preferences (this statement is true within the model studied here, that is, with a finite number of normal types). If we define $\underline{a}_1 := \min_{\mathbf{a} \in \mathbf{A}^*}\{a_1\}$ to

[13]Defined as games in which the action to which player 1 would most like to commit also minmaxes player 2. In calculating the payoff from committing to an action to which there are multiple best responses, it is assumed that the least favorable one from player 1's point of view is chosen.

be the lower bound on type 1's payoffs in the undiscounted game, then if $A_2 = -B$,

$$\underline{a}_1 = \max_{f \in \Delta^I} \min_{g \in \Delta^J(f)} A_1(f, g) \ , \tag{22}$$

where

$$\Delta^J(f) \ := \ \{g \in \Delta^J \mid B(f, g) \geq \hat{b}\} \ . \tag{23}$$

This lower bound is independent of whether there are other types ($k \geq 3$) and cannot be bettered. We can relate Israeli's result to the discounted model by using Theorem 1: if $A_2 = -b$, then for a given δ_2 and a given $\epsilon > 0$, there exists a $\underline{\delta}_1$ such that for all $1 > \delta_1 > \underline{\delta}_1$, a lower bound on type 1's equilibrium payoff is given by $\underline{a}_1 - \epsilon$ (where $\underline{a}_1$ is as defined by (22)). Since, as noted above, Theorem 1 derives necessary conditions for equilibrium, this lower bound is valid no matter which other types, normal or otherwise, are present; it depends only on the existence of a positive probability type with preferences which are the opposite of those of player 2 (though in this general case we do not know whether a better bound is available from a type other than a normal one). Since this is the best possible bound that can be derived from a "normal" type, it must be at least as good as that derived from a type committed to playing the same action each period, as there is a normal type which behaves in this fashion. Indeed, the expression given in (22) can be interpreted as the best that type 1 could get in the complete information stage game between herself and player 2 by committing to some mixed strategy f given that player 2 responds by choosing g in the least favorable way subject only to (player 2's) individual rationality. This is a generalization of the description of the bound given above for fixed action commitment types, in the sense that it allows player 1 to choose a *mixed* strategy, and it is shown in Cripps and Thomas (1995) that in games with no discounting this is indeed the bound associated with the best mixed strategy commitment type. Consequently the $-B$ type is equivalent, in terms of its value to type 1, to a type which is committed to the best mixed strategy.

4. A "Folk Theorem" Result for Equally Patient Players

In this section we consider games $\Gamma(\mathrm{p}, \delta)$ where the players are equally patient, $\delta = \delta_1 = \delta_2$, and there are only two types of player 1, $\#K = 2$. We denote this class of games by $\Gamma(\mathrm{p}, \delta)$, so $\Gamma(\mathrm{p}, \delta) := \Gamma(\mathrm{p}, \delta, \delta)$. We show, in a sense to be made more precise, that the (Nash) Folk Theorem can be extended to the repeated games $\Gamma(\mathrm{p}, \delta)$ when p_1 is large. [14] In the repeated game of complete information played between type 1 and player 2 the Folk Theorem asserts that, given any profile of feasible and strictly individually rational payoffs (a_1, b), there is a Nash equilibrium where the players receive these payoffs if the players are sufficiently patient. We will extend this result in the following way. Again let (a_1, b) be any profile of feasible and strictly individually rational payoffs for the complete information game played by type 1 and player 2. Then Theorem 3 shows that there exists $\underline{\delta}, \underline{p}_1 < 1$ such that the pair (a_1, b) are the players' equilibrium payoffs in $\Gamma(\mathrm{p}, \delta)$ if $\delta > \underline{\delta}$ and $p_1 > \underline{p}_1$. Thus introducing a small amount of uncertainty about the type of player 1 does not reduce the set of equilibrium payoffs in any significant way when both of the players are sufficiently patient.

To prove our result we will follow the usual folk theorem strategy and construct an equilibrium where the payoffs (a_1, b) are received. We begin this section by discussing the notion of punishments and individual rationality in the games $\Gamma(\mathrm{p}, \delta)$. Then Lemma 3 characterizes the long-run behaviour of the players at the folk theorem equilibrium by describing an equilibrium that is played out when all learning is finished. Lemma 4 describes a finite sequence of actions that holds type 1 to her minmax level, $\hat{a}_1$, or below whilst at the same time rewarding type 2. This is used in Lemma 5 which shows how a given equilibrium of $\Gamma(\mathrm{p}, \delta)$ can be preceded by repetitions of this finite sequence to get an equilibrium of the game where type 1's payoff is reduced and player 2's payoff is also altered. Finally in Theorem 3 repeated use of Lemma 5 together with the long-run behaviour described in Lemma 3 are combined to prove the result. This argument is a generalization of the construction presented for the Battle of the Sexes example in the Introduction.

It is helpful at this stage to recall the Folk Theorem of Fudenberg and Maskin

[14] In the previous section, in contrast, the characterization was valid for all values of p.

(1991) for the repeated game of complete information, $\Gamma_1(\delta)$, played between type 1 and player 2. Define the set of feasible and (uniformly for a given ϵ) strictly individually rational payoffs for the complete information game between type k and player 2: $G_k(\epsilon) := \{(A_k(\pi), B(\pi)) | A_k(\pi) \geq \hat{a}_k + \epsilon, B(\pi) \geq \hat{b} + \epsilon, \pi \in \Delta^{IJ}\}$, $k \in K$. The Folk Theorem for $\Gamma_1(\delta)$ combined with Fudenberg and Maskin (1991), Lemma 2, implies that provided δ is sufficiently close to unity, then for any payoffs $(a_1, b) \in G_1(\epsilon)$, there exists a Nash equilibrium for $\Gamma_1(\delta)$ where the players play a deterministic sequence of actions, receive the payoffs (a_1, b), and their continuation payoffs after any finite history are within $\epsilon/2$ of (a_1, b). [15]

Result 3 (Fudenberg and Maskin) *Let $\epsilon > 0$ be given. There is a $\underline{\delta} < 1$ such that if $\delta > \underline{\delta}$ and $(a_1, b) \in G_1(\epsilon)$ then there exists an equilibrium of $\Gamma_1(\delta)$ where the sequence of actions $\{(i^t, j^t)\}_{t=0}^{\infty}$ is played and:*
$a_1 = (1-\delta)\sum_{t=0}^{\infty} \delta^t A_1(i^t, j^t)$, $b = (1-\delta)\sum_{t=0}^{\infty} \delta^t B(i^t, j^t)$ *with*

$$\begin{aligned} \left| (1-\delta)\textstyle\sum_{t=s}^{\infty} \delta^{t-s} A_1(i^t, j^t) - a_1 \right| &\leq \epsilon/2 \qquad \forall s, \\ \left| (1-\delta)\textstyle\sum_{t=s}^{\infty} \delta^{t-s} B(i^t, j^t) - b \right| &\leq \epsilon/2 \qquad \forall s. \end{aligned}$$

There are two points to stress concerning Result 3. First for the result to have any significance it is necessary for there to exist a pair $(a_1, b) \in G_1(\epsilon)$ for some $\epsilon > 0$, that is, there must be a profile of strictly individually rational payoffs. Second the result works because each player can punish deviations by holding their opponent to his/her minmax level. These two features will appear in a different guise in the repeated game of incomplete information.

In the repeated games with incomplete information studied here the punishment strategies for player 2 are more complex because the deviations can be made by either of player 1's types; the punishment must *simultaneously* punish both possible types. In a repeated game without discounting the problem of simultaneously punishing the two types is solved by Blackwell's approachability result (Blackwell (1956)), briefly discussed in the previous section. Let $\mathrm{x} := (x_k)_{k \in K}$ be a vector of payoffs for the types of player 1. The set of payoffs $\{\mathrm{y} | \mathrm{y} \leq \mathrm{x}\}$ is said to be *approachable* by player 2 if he has a strategy,

[15]While their result is established for subgame-perfect equilibria, we only need the result for Nash equilibria, the proof of which is straightforward.

τ, that guarantees type k gets no more than x_k whatever strategy, σ, player 1 uses. Thus the set $\{\mathbf{y}|\mathbf{y} \leq \mathbf{x}\}$ is approachable if $\mathbf{x}$ is a vector of feasible punishment payoffs for player 2 to impose on the types of player 1. Therefore, the vector $\mathbf{x} = (x_k)_{k\in K}$ is *individually rational (IR)* if the set $\{\mathbf{y}|\mathbf{y} \leq \mathbf{x}\}$ is approachable, and the convex set of all such IR payoffs $(x_k)_{k\in K}$ is characterised by (7). The above definition of individual rationality applies to player 1's undiscounted payoffs, however; here we are interested in *discounted* repeated games. In discounted games it is in general impossible to impose the severe punishment described by the approachable sets $\{\mathbf{y}|\mathbf{y} \leq \mathbf{x}\}$, but as the players become more patient player 2 is able to approximate these punishments arbitrarily closely. First define the notion of ϵ-IR payoffs.

Definition 2 *Let $\epsilon > 0$ be given. The vector $\mathbf{x} = (x_k)_{k\in K}$ is* ϵ-individually rational (ϵ-IR) *if the set $\{ \mathbf{y} \mid \mathbf{y} + \epsilon\mathbf{1} \leq \mathbf{x} \}$ is approachable.* [16]

By Zamir (1992) there is a lower threshold on the discounting, δ_ϵ, so that if $\delta > \delta_\epsilon$ then player 2 can hold the types of player 1 down any ϵ-IR payoff in $\Gamma(\mathbf{p}, \delta)$. [17] Note that without loss of generality we can assume $\delta_\epsilon > \underline{\delta}$ so Result 3 also applies for $\delta > \delta_\epsilon$.

The second element of Result 3 is the existence of strictly individually rational payoffs. We will assume (in (A.1)) that we can find strictly individually rational payoffs for the repeated game of incomplete information $\Gamma(\mathbf{p}, \delta)$.

(A.1) *There exists $\breve{\pi}_k \in \Delta^{IJ}$ for all $k \in K$ and $\epsilon > 0$ such that $(A_k(\breve{\pi}_k))_{k\in K}$ is ϵ-IR and $B(\breve{\pi}_k) > \hat{b}$ for all $k \in K$.*

[16] The notation $\mathbf{1}$ is used to denote a vector of 1's of unspecified dimension; the dimension should be clear from the context.

[17] Let Cav $a(\mathbf{p})$ be the (pointwise) smallest concave function $g(\mathbf{p})$ satisfying $g(\mathbf{p}) \geq a(\mathbf{p})$ where $a(\mathbf{p})$ is defined in (7). Then by Zamir (1992 p.126) Cav $a(\mathbf{p})$ is the value for the zero-sum repeated game of incomplete information with no discounting that is played when player 2's payoffs are $(-A_k(i,j))_{k\in K}$. (Player 2 can achieve his value by using a strategy to approach a given set of payoffs.) Now consider the zero-sum discounted repeated game of incomplete information with the same payoffs. By Zamir (1992 p.119) the value function for this game, $v_\delta(\mathbf{p})$, exists and since $\sharp K = 2$, by Zamir (1992 p.125) this satisfies $0 \leq v_\delta(\mathbf{p}) - \text{Cav } a(\mathbf{p}) \leq M\sqrt{\{(1-\delta)/(1+\delta)\}}$. Thus as $\delta \to 1$ the punishments that can be imposed in the discounted game converge uniformly to the punishments that can be imposed in the undiscounted game. Suppose that player 2 could not hold the types of player 1 to some ϵ-IR payoffs as $\delta \to 0$, then this would contradict the uniform convergence of the two value functions.

Here strict individual rationality is defined by a strict inequality and approachability rather than in relation to the players' minmax levels. As in the complete information case there are always weakly individually rational payoffs, that is, there exists $\breve{\pi}_k$ and an individually rational vector $(\breve{\omega}_k)_{k\in K}$ so that: $A_k(\breve{\pi}_k) \geq \breve{\omega}_k$, $B(\breve{\pi}_k) \geq \hat{b}$, for all $k \in K$. [18] Assumption A.1 implies that the game of complete information played between each type k and player 2 has strictly individually rational payoffs ($G_k(\epsilon) \neq \emptyset$ for some $\epsilon > 0$) and thus it cannot be the case, for example, that one of player 1's types plays a zero-sum game with player 2. [19] It is, nevertheless, a natural extension of the assumption made in the complete information case.

Using A.1 we can now describe a particular Nash equilibrium which we refer to as the *terminal equilibrium*. The terminal equilibrium will serve to describe the players' long-run behaviour. That is, the terminal equilibrium will be the equilibrium that is eventually played out if type 1 mimics type 2 indefinitely in the "Folk Theorem" construction below. Of course, because the players are discounting, the payoffs in the terminal equilibrium may have very little impact on the players' expected payoffs at the start of the game. The terminal equilibrium is revealing and so in general the incentive compatibility conditions discussed in the introduction will bind most tightly at it. In fact we choose the payoffs at the equilibrium so type 1 receives a payoff $\bar{a}_1(\epsilon)$ which is her largest individually rational payoff in $G_1(\epsilon)$, that is, $\bar{a}_1(\epsilon) := \max\{\ a_1 \mid (a_1, b) \in G_1(\epsilon)\ \}$. All the other players get strictly individually rational payoffs. Below we will also use M to denote an upper bound on the absolute magnitude of the players' payoffs, that is, $M := \max_{(i,j)} \max\{\ \max_k |A_k(i,j)|\ ,\ |B(i,j)|\ \}$.

Lemma 3 *Assume A.1 and let $\epsilon > 0$ sufficiently small be given. If $\delta > \max\{\ \delta_\epsilon\ ,\ 2M(2M+\epsilon)^{-1}\ \}$, then $\Gamma(\mathrm{p}, \delta)$ has a* terminal equilibrium *with the payoffs $(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta})$ satisfying $\bar{\alpha}_1 = \bar{a}_1(5\epsilon)$, $\bar{\alpha}_2 > \hat{a}_2 + 5\epsilon$ and $\bar{\beta} \geq \hat{b} + 5\epsilon$.*

Proof: See Appendix.

[18] Let $\hat{g}$ be a strategy that ensures player 2 receives his minmax payoff, $B(i, \hat{g}) \geq \hat{b}$ for $i = 1, 2, ..., I$. By playing $\hat{g}$ in every period he can hold type k to the payoff $\max_i A_k(i, \hat{g})$, so the set $\{\mathrm{y}|\mathrm{y} \leq (\max_i A_k(i, \hat{g}))_{k\in K}\}$ is approachable. If $\breve{\pi}_k$ is the correlated strategy generated by player 2 playing $\hat{g}$ and type k playing a best response then we will have $A_k(\breve{\pi}_k) \geq \max_i A_k(i, \hat{g})$ and $B(\breve{\pi}_k) \geq \hat{b}$ for $k \in K$.

[19] It also rules out commitment types, as in the example in the introduction; extending the argument below to the commitment type case is straightforward.

The assumption A.1 allows us to calculate a particularly simple equilibrium for $\Gamma(\mathbf{p},\delta)$ with desirable properties which we will use later. However, the assumption is by no means necessary for the existence of an equilibrium of $\Gamma(\mathbf{p},\delta)$ as shown by the work of Jordan (1995).

Before reaching the terminal equilibrium type 2 and player 2 repeatedly play out a finite deterministic sequence of actions and most of the time type 1 also plays this sequence, but occasionally she will deviate from it with positive probability. The sequence is chosen so that on average type 1 gets a payoff close to or below her minmax payoff and type 2 receives a high payoff. Thus type 1 is forced to accept a very low payoff if she wants to acquire a reputation for being type 2. The finite sequence of actions is denoted $\{(\hat{i}^t,\hat{j}^t)\}_{t=0}^{T-1}$. The players' average payoffs over this finite sequence are denoted by $(\hat{A}_1,\hat{A}_2,\hat{B})$ where

$$(24) \qquad \hat{A}_k := \frac{1-\delta}{1-\delta^T}\sum_{s=0}^{T-1}\delta^s A_k(\hat{i}^t,\hat{j}^t), \qquad \hat{B} := \frac{1-\delta}{1-\delta^T}\sum_{s=0}^{T-1}\delta^s B(\hat{i}^t,\hat{j}^t).$$

The properties of this finite sequence of actions are described in the lemma below which shows that: (a) the sequence does hold type 1 close to or below her minmax payoff and (b) that the convex combination of $(\hat{A}_1,\hat{A}_2)$ and any ϵ-IR pair (α_1,α_2) is ϵ-IR for the two types provided type 1 receives at least $\hat{a}_1+\epsilon$. This second property guarantees that provided the terminal equilibrium is ϵ-IR then the terminal equilibrium preceded by any amount of repeated play of this sequence is also ϵ-IR, provided type 1 gets more than her minmax level. In the statement of the lemma below a parameter $\lambda \geq 1$ is defined. This parameter is used to define a linear upper bound $\hat{a}_1+\lambda\epsilon$ on the rate at which the lowest ϵ-IR payoff to type 1 increases as ϵ increases. That is, we choose λ to be a positive constant so that

$$\hat{a}_1+\epsilon\lambda \geq \inf\{\ A_1(\pi)\mid \pi\in\Delta^{IJ},\ (A_k(\pi))_{k\in K}\text{ is }\epsilon\text{−IR }\},$$

for all $\epsilon>0$ for which the set on the right is non-empty and if the set is empty for all $\epsilon>0$ define $\lambda=1$. We now state the following result.

Lemma 4 *Assume A.1 and let $\epsilon>0$ sufficiently small be given. There exists $\hat{\delta}<1$, a finite T and a sequence of profiles $\{(\hat{i}^t,\hat{j}^t)\}_{t=0}^{T-1}$ such that if $\delta>\hat{\delta}$, the players' average payoffs $(\hat{A}_k)_{k\in K}$, $\hat{B}$ over the sequence satisfy*

(a) $\hat{A}_1 < \hat{a}_1 + 5\epsilon\lambda + \epsilon/2$,

(b) *If* (α_1, α_2) *is* $(7\epsilon/2)$-IR *then* $(1-\rho)\hat{A}_1 + \rho\alpha_1 \geq \hat{a}_1 + (7\epsilon/2)$ *and* $\rho \in [0,1]$ *implies* $(1-\rho)(\hat{A}_1, \hat{A}_2) + \rho(\alpha_1, \alpha_2)$ *is* $(7\epsilon/2)$-IR.

Proof: See Appendix.

The next step in the argument (Lemma 5) shows that given an arbitrary equilibrium $(\alpha_1, \alpha_2, \beta)$ for the game $\Gamma(\mathrm{p}, \delta)$ there is an equilibrium for $\Gamma(1 - rp_2, rp_2, \delta)$ where one period of randomization and N repetitions of the sequence described in Lemma 4 are played before the game settles down to $(\alpha_1, \alpha_2, \beta)$. At this new equilibrium type 2 and player 2 coordinate their actions for the first NT periods to repeat N times the sequence of actions described in Lemma 4; they then play the equilibrium strategies to get the payoffs $(\alpha_1, \alpha_2, \beta)$. Type 1, on the other hand, randomizes in the first period; with probability $q = r(1-p_2)(1-rp_2)^{-1}$ she plays $\hat{i}_0$ and then continues to mimic type 2 for NT periods, whereas with probability $1-q$ she reveals her type by deviating from $\hat{i}_0$ and then plays an equilibrium of the complete information game $\Gamma_1(\delta)$. The players' payoffs from playing the finite sequence n times and then getting the equilibrium payoffs $(\alpha_1, \alpha_2, \beta)$ are denoted $(a_1(n), a_2(n), b(n))$ where

$$(25) \qquad a_k(n) := (1-\delta^{nT})\hat{A}_k + \delta^{nT}\alpha_k, \quad k = 1,2; \qquad b(n) := (1-\delta^{nT})\hat{B} + \delta^{nT}\beta.$$

How do the payoffs at the equilibrium of $\Gamma(1 - rp_2, rp_2, \delta)$ described in Lemma 5 differ from those in the original game $\Gamma(\mathrm{p}, \delta)$? Type 1 randomizes so at the equilibrium she is indifferent between mimicking type 2 and revealing her type; and thus her expected payoff in $\Gamma(1 - rp_2, rp_2, \delta)$ is $a_1(N)$. Given $\hat{A}_1$ is close to $\hat{a}_1$ it is clear that as n increases so $a_1(n)$ approaches $\hat{a}_1$. Thus in $\Gamma(1 - rp_2, rp_2, \delta)$, type 1's payoff is generally lower than it is in $\Gamma(\mathrm{p}, \delta)$. Player 2's payoff is a combination of what he expects to get against type 2 and what he expects to get when type 1 reveals her type. When type 1 does reveal her type she and player 2 will play out an equilibrium of the repeated game of complete information $\Gamma(\delta)$ with the payoffs $(a_1(N), b)$. (Type 1's payoffs are determined by her indifference between mimicking type 2 and revealing her type, but b is to some extent arbitrary.) However, r, the total probability that the sequence is played following the randomization, can take any value in $[0,1]$. By a judicious choice of r we can ensure that player 2's payoff in $\Gamma(1-rp_2, rp_2, \delta)$ will approach b for any feasible b. In fact provided $(a_1(N), b) \in G_1(2\epsilon, \epsilon)$, any payoff b can be chosen, where the set $G_1(2\epsilon, \epsilon)$ consists of those points that are at least ϵ in distance inside the boundary of $G_1(2\epsilon)$:

$$G_1(2\epsilon, \epsilon) := \{ (a_1, b) \mid \{(x,y) | \parallel (x,y)-(a_1,b) \parallel \leq \epsilon\} \subset G_1(2\epsilon) \}. \tag{26}$$

In the equilibrium of Lemma 5 an important consideration is the punishments imposed when player 1 deviates. Since both her types' equilibrium payoffs vary as the finite sequence is played out so too do the punishments for deviation. During the early periods of play type 1 faces heavy punishment for deviation (because she has a low equilibrium payoff) whereas in later periods she has light punishment if she deviates (because her expected payoffs from continuing with the equilibrium have grown). This varying punishment is necessary because, as the discussion of approachability above implies, it is usually impossible to heavily punish both types of player 1 simultaneously. The fact that it is possible to punish deviations appropriately is ensured by Lemma 4(b).

Lemma 5 *Assume A.1 let:* $\epsilon > 0$, $\delta > \delta_\epsilon^* := \max\{ \delta_\epsilon \ , \ \hat{\delta} \ , \ (1-\epsilon/(4M))^{1/T} \}$, *and an equilibrium* $(\alpha_1, \alpha_2, \beta)$ *of* $\Gamma(\mathrm{p}, \delta)$ *with* $\beta > \hat{b} + 2\epsilon$ *and* (α_1, α_2) (3ϵ)-IR *be given. If there is an* N *and* b *so that*

$$b(N) > \hat{b} + 2\epsilon, \qquad (a_1(N), b) \in G_1(2\epsilon, \epsilon), \tag{27}$$

then the game $\Gamma(1-rp_2, rp_2, \delta)$, $r \in [0,1]$, *has an equilibrium with the payoffs* $(\alpha_1', \alpha_2', \beta')$ *where* $\alpha_1' = a_1(N)$, $\alpha_2' \geq a_2(N)$, $|\beta' - b| \leq 2Mr$.

<u>Proof:</u> See Appendix.

By choosing r sufficiently close to zero in the above lemma we can find an equilibrium of $\Gamma(1-rp_2, rp_2, \delta)$ where the players get payoffs arbitrarily close to the point $(a_1(N), b) \in G_1(2\epsilon, \epsilon)$. Thus if the possible values of $(a_1(N), b)$ are close to all of the points in $G_1(\epsilon)$ we would already have proved our folk theorem result. This is not true in general for two separate reasons. The first is easiest to deal with: $a_1(N) \leq \alpha_1$ for all $N \geq 0$ so we must choose α_1 as large as possible for all possible points in $G_1(\epsilon)$ to be approximated by $(a_1(N), b)$. Lemma 3 provides the obvious choice of $(\alpha_1, \alpha_2, \beta)$ in Lemma 5, that is

$(\alpha_1, \alpha_2, \beta) = (\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta})$, because then $\alpha_1 = \bar{a}_1(\epsilon)$. The second reason arises because the lemma above does place a restriction on which points in $G_1(\epsilon)$ can be approximated; it requires that the constraint $b(N) > \hat{b} + 2\epsilon$ be satisfied. That is, player 2 must receive strictly more than his minmax payoff while the players play out the equilibrium after the randomization and in general this will restrict how low $a_1(N)$ can be made (see the introductory example). Theorem 3 addresses this issue by allowing periodic randomizations by type 1 that increase player 2's payoffs. Lemma 5 is used as an inductive step in the proof of Theorem 3 , so if when $(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta})$ is used as an initial equilibrium it is not possible for $a_1(N)$ to be close to $\hat{a}_1$, choose instead the intial randomization in Lemma 5 so that β' is large. Thus from the initial equilibrium $(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta})$ a second equilibrium $(\alpha_1', \alpha_2', \beta')$ has been found with $\alpha_1' < \alpha_1 = \bar{\alpha}_1$ and $\beta' > \hat{b} + 2\epsilon$. Now apply Lemma 5 a second time to the new equilibrium with the payoffs $(\alpha_1', \alpha_2', \beta')$ and determine if it is now possible to be close to all of the points in $G_1(\epsilon)$. And if not define a second equilibrium $(\alpha_1'', \alpha_2'', \beta'')$ with $\alpha_1'' < \alpha_1'$ and $\beta'' > \hat{b} + 2\epsilon$. The proof of the theorem shows that only a finite number of iterations of this argument are needed to ensure that any $(a_1, b) \in G_1(\epsilon)$ can be approximated. As only a finite number of applications of Lemma 5 are necessary there is a lower threshold $\underline{p}_1 < 1$, the lower threshold on the discounting in the theorem comes directly from Lemma 5.

Theorem 3 *Assume A.1 and let* $\nu > 0$ *sufficiently small be given. Then there exist* $\underline{\delta} < 1$, $\underline{p}_1 < 1$ *such that for all* $p > \underline{p}_1$, $\delta > \underline{\delta}$ *and* $(a_1^*, b^*) \in G_1(\nu)$ *the game* $\Gamma(\mathrm{p}, \delta)$ *has an equilibrium with the payoffs* $(\alpha_1, \alpha_2, \beta)$ *where* $|\alpha_1 - a_1^*| < \nu$ *and* $|\beta - b^*| < \nu$.

Proof: See Appendix.

5. Conclusion

This paper studies the Nash equilibrium payoffs that can arise in discounted repeated games of one-sided incomplete information. We have isolated two polar cases. In the first case the informed player is very patient relative to the uninformed player and the set of equilibrium payoffs is close to the set of equilibrium payoffs when there is no discounting. In the second case both players are equally patient and the set of payoffs approaches the set of equilibrium payoffs in the game of complete information as the common discount factor approaches unity and the probability of all other types shrinks. In the first case the time taken for play to converge to a stationary outcome is insignificant in the informed player's equilibrium payoff whereas in the second case it completely determines her equilibrium payoff; it is this that largely explains the difference between the two poles. Whether these results also apply for the perfect Bayesian equilibria is a subject of our current work and for a treatment of a special case see Cripps and Thomas (1995).

Appendix

We shall need the following definitions. Let

$$b_{min} = \min_{i \in I} \min_{j \in J} B(i,j) \tag{28}$$

be the *worst payoff* player 2 can get in the stage game,

$$b_{max} = \max_{i \in I} \max_{j \in J} B(i,j) \tag{29}$$

be the *best payoff* for player 2.

PROOF OF LEMMA 1:

To simplify notation let $\mathrm{q}^N = \mathrm{q}^N(\cdot \mid h^t)$ and $\hat{\mathrm{q}}^N = \mathrm{q}^N(\cdot \mid h^t, k)$. Choose N to be the smallest integer such that

$$\delta_2^N \frac{b_{max} - b_{min}}{1 - \delta_2} < \frac{\epsilon}{2(1-\delta_2)} . \tag{30}$$

Next, define $\tilde{V}_2^{t+1}(y^N)$ to be the payoff to player 2 over the next N periods discounted to period $t+1$, that is

$$\tilde{V}_2^{t+1}(y^N) = \sum_{\tau = t+1}^{t+N} \delta_2^{\tau - t - 1} B(i^\tau, j^\tau) . \tag{31}$$

For probability distribution q^N its expectation is $E_{\mathrm{q}^N}\left[\tilde{V}_2^{t+1}(y^N)\right] = \sum_{y^N} \mathrm{q}^N \tilde{V}_2^{t+1}(y^N)$, and since this is a continuous function of q^N, with compact domain, there exists an $\eta > 0$ such that $\| \mathrm{q}^N - \hat{\mathrm{q}}^N \| \leq \eta$ implies that

$$\left| E_{\mathrm{q}^N}\left[\tilde{V}_2^{t+1}(y^N)\right] - E_{\hat{\mathrm{q}}^N}\left[\tilde{V}_2^{t+1}(y^N)\right] \right| \leq \frac{\epsilon}{2(1-\delta_2)} . \tag{32}$$

Let η be as in the statement of the lemma, and assume to the contrary of the lemma $\| \mathrm{q}^N - \hat{\mathrm{q}}^N \| \leq \eta$; then by (32)

$$\begin{aligned} & E\left[b_2^{t+1} \mid h^t\right] - E\left[b_2^{t+1} \mid h^t, k\right] \\ & \quad < (1-\delta_2) E_{\mathrm{q}^N}\left[\tilde{V}_2^{t+1}(y^N)\right] + \delta_2^N b_{max} - (1-\delta_2) E_{\hat{\mathrm{q}}^N}\left[\tilde{V}_2^{t+1}(y^N)\right] - \delta_2^N b_{min} \\ & \quad < (1-\delta_2) \left| E_{\mathrm{q}^N}\left[\tilde{V}_2^{t+1}(y^N)\right] - E_{\hat{\mathrm{q}}^N}\left[\tilde{V}_2^{t+1}(y^N)\right] \right| + \delta_2^N (b_{max} - b_{min}) \\ & \quad < \epsilon . \end{aligned} \tag{33}$$

But using (15) this implies

$$E\left[\tilde{b}^{t+1} \mid h^t\right] < \hat{b} \tag{34}$$

which is impossible. *Q.E.D.*

PROOF OF LEMMA 2:

Fix an equilibrium and a type k and choose $\epsilon = \phi/3$ in Lemma 1; then there is an N and an η such that (16) holds whenever (15) holds. Set $\psi = \eta$ in Result 2, take any integer n, $0 \leq n < N$, and set $\xi = \frac{\phi}{3N(\hat{b}-b_{min})}$ (assuming that $\hat{b} > b_{min}$; the lemma is trivial otherwise). Then by Result 2 there is an m (finite) such that the probability that inequality (16) holds more than m times in $T(n,N)$ is less than ξ, so the probability that inequality (15) holds more than m 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, conditional upon type k, that the inequality

$$E\left[\tilde{b}^{t+1} \mid h^t, k\right] \leq \hat{b} - \phi/3 \tag{35}$$

holds more than Nm times is smaller than $N\xi = \frac{\phi}{3(\hat{b}-b_{min})}$.

Next,

$$E\left[\tilde{b}^{t+1} \mid k\right] = E\left[(1-\delta_2)B(i^{t+1}, j^{t+1}) + \delta_2 \tilde{b}^{t+2} \mid k\right] , \tag{36}$$

so

$$(1-\delta_2)E\left[B(i^{t+1}, j^{t+1}) \mid k\right] = E\left[\tilde{b}^{t+1} - \delta_2 \tilde{b}^{t+2} \mid k\right] . \tag{37}$$

Hence, player 2's payoff against type k in the equilibrium, *calculated using player 1's discount factor*, is

$$\begin{aligned} B(\pi_k(\delta_1)) &= (1-\delta_1)\sum_{t=0}^{\infty} \delta_1^t E\left[B(i^t, j^t) \mid k\right] \\ &= \frac{1-\delta_1}{1-\delta_2}\sum_{t=0}^{\infty} \delta_1^t E\left[\tilde{b}^t - \delta_2 \tilde{b}^{t+1} \mid k\right] \\ &= \frac{1-\delta_1}{1-\delta_2}\left\{E\left[\tilde{b}^1 \mid k\right] + E\left[\sum_{t=0}^{\infty} E\left[\delta_1^t(\delta_1-\delta_2)\tilde{b}^{t+1}\middle| h^t, k\right]\middle| k\right]\right\} . \end{aligned} \tag{38}$$

Using the result on the number of times (35) holds, for $\delta_1 > \delta_2$ the random variable

$$\begin{aligned} &\sum_{t=0}^{\infty} E\left[\delta_1^t(\delta_1-\delta_2)\tilde{b}^{t+1}\middle| h^t, k\right] \\ &\geq \left\{\frac{\delta_1-\delta_2}{1-\delta_1}(\hat{b}-\frac{\phi}{3}) - (\delta_1-\delta_2)(\hat{b}-b_{min})Nm\right\} \end{aligned} \tag{39}$$

with probability at least $(1 - N\xi)$ conditional on k, where we are using the fact that in the event that (35) fails no more than Nm times, subtracting $(\hat{b} - b_{min})$ Nm times undiscounted yields a payoff lower than the minimum possible. The random variable is at least $\frac{\delta_1 - \delta_2}{1-\delta_1} b_{min}$ otherwise.

Using this in (38) gives a lower bound, say $\Phi(\delta_1, \delta_2)$, so that $B(\pi_k(\delta_1)) \geq \Phi(\delta_1, \delta_2)$, and notice that $\Phi(\delta_1, \delta_2)$ is independent of the particular equilibrium studied. Next, taking the limit as $\delta_1 \longrightarrow 1$ yields

$$\lim_{\delta_1 \to 1} \Phi(\delta_1, \delta_2) = (1 - N\xi)\left(\hat{b} - \frac{\phi}{3}\right) + N\xi b_{min} \; ; \tag{40}$$

hence, since $N\xi = \frac{\phi}{3(\hat{b} - b_{min})}$, we get

$$\begin{aligned} \lim_{\delta_1 \to 1} \Phi(\delta_1, \delta_2) &= \hat{b} - \frac{\phi}{3} - \frac{\phi}{3(\hat{b} - b_{min})}\left(\hat{b} - b_{min} - \frac{\phi}{3}\right) \\ &= \hat{b} - \frac{2\phi}{3} + \frac{\phi^2}{9(\hat{b} - b_{min})} \\ &> \hat{b} - \frac{2\phi}{3} \, . \end{aligned} \tag{41}$$

Choosing $\underline{\delta}_1^{(k)}$ such that $\Phi(\delta_1, \delta_2)$ is within $\frac{\phi}{3}$ of its limit ($\underline{\delta}_1^{(k)}$ depends only upon p_k, ϕ and δ_2), we have for $\delta_1 \geq \underline{\delta}_1^{(k)}$

$$B(\pi_k(\delta_1)) \geq \hat{b} - \phi \, . \tag{42}$$

Set $\underline{\delta}_1 = \max_{k \in K}\{\underline{\delta}_1^{(k)}\}$ and the result follows. *Q.E.D.*

<u>PROOF OF THEOREM 1:</u>

The strategy of the proof is to show that for player 1 sufficiently patient, the average frequencies over action profiles in a Nash equilibrium, *calculated using player 1's discount factor,* must approximately satisfy individual rationality and incentive compatibilty. We take δ_2 and $\mathbf{p}$ to be fixed throughout the proof. First consider condition (i) of Definition 1 of Π_0, individual rationality (for player 1). Let (σ, τ) be a Nash equilibrium pair of strategies for the game $\Gamma(\mathbf{p}, \delta_1, \delta_2)$, and suppose that the equilibrium payoff profile for player 1, $\mathbf{a} = (A_k(\pi_k(\delta_1)))_{k \in K}$, is not individually rational. Then by (7), there exists $\mathbf{q}^* \in \Delta^K$ such that $\mathbf{q}^* \cdot \mathbf{a} < a(\mathbf{q}^*)$. By the minimax theorem,

$$\mathbf{q}^* \cdot \mathbf{a} < \max_{f \in \Delta^I} \min_{g \in \Delta^J} \sum_k q_k^* A_k(f, g) \; , \tag{43}$$

so that if player 1 plays a mixed action f^* which attains the maximum in (43),

$$\mathbf{q}^* \cdot \mathbf{a} < \sum_k q_k^* A_k(f^*, g) \tag{44}$$

for all $g \in \Delta^J$. Denote by σ^* the repeated game strategy in which player 1 plays the mixed action f^* each period and independently of type k. Then

$$E_{\mathbf{p},\sigma^*,\tau}\left[(1-\delta_1)\sum_{t=0}^{\infty}\delta_1^t \sum_k q_k^* A_k(i^t, j^t)\right] \;>\; \mathbf{q}^* \cdot \mathbf{a}\ , \tag{45}$$

(Nb. k is not a random variable) so that

$$\sum_k q_k^* E_{\mathbf{p},\sigma^*,\tau}\left[\tilde{a}_k \mid k\right] \;=\; \sum_k q_k^* E_{\mathbf{p},\sigma^*,\tau}\left[(1-\delta_1)\sum_{t=0}^{\infty}\delta_1^t A_k(i^t, j^t) \mid k\right] \;>\; \mathbf{q}^* \cdot \mathbf{a}\ , \tag{46}$$

since given that σ^* does not vary with type, conditioning on k does not affect the distribution over histories. Because $\mathbf{q}^* \in \Delta^K$, it follows that

$$E_{\mathbf{p},\sigma^*,\tau}\left[\tilde{a}_k \mid k\right] \;>\; a_k \tag{47}$$

for at least one k, contradicting the definition of equilibrium. Hence individual rationality must be satisified for player 1 for any value of δ_1; that is, $\mathbf{a}$ satisfies (7). Next, condition (ii) of Definition 1 (incentive compatibility) must be satisfied for any δ_1, $0 < \delta_1 < 1$, since in any Nash equilibrium $A_k(\pi_k(\delta_1)) \geq A_k(\pi_k'(\delta_1))$ for all k, k' by the definition of equilibrium (recall that $A_k(\pi_k(\delta_1))$ is the equilibrium payoff of type k of player 1, and $A_k(\pi_k'(\delta_1))$ is the payoff type k would get from following the strategy of type k').

Finally, individual rationality for player 2 must be dealt with. Define

$$\hat{\Pi} := \{\pi \in \Pi \mid A_k(\pi_k) \geq A_k(\pi_{k'}) \ \text{ all } k, k' \text{and} (A_k(\pi_k))_{k \in K} \ \text{ is individually rational}\}\ , \tag{48}$$

and define the compact valued correspondence $\Psi : [0, \infty) \rightarrow\rightarrow \Pi$ by

$$\Psi(\phi) \;=\; \left\{\pi \mid B_k(\pi_k) \geq \hat{b} - \phi \ \text{ all } k \in K\right\}\ . \tag{49}$$

Since Ψ is clearly an upper hemi-continuous function of ϕ, it follows that the correspondence given by $\Psi \cap \hat{\Pi}$, which is non-empty (Shalev (1994)), is also upper hemi-continuous. Moreover, if the linear function $\mathbf{A}(\pi) := (A_1(\pi_1), A_2(\pi_2), \ldots, A_K(\pi_K))$ is defined on Π, the correspondence given by $\mathbf{A}[\Psi(\phi) \cap \hat{\Pi}]$ is an upper hemi-continuous function of ϕ, with value $\mathbf{A}^*$ at $\phi = 0$. Hence given ϵ, there is a $\bar{\phi} > 0$ such that for $0 \leq \phi < \bar{\phi}$, all payoffs in $\mathbf{A}[\Psi(\phi) \cap \hat{\Pi}]$ lie within ϵ of $\mathbf{A}^*$. Choose ϕ in Lemma 2 to be $\bar{\phi}$; the corresponding $\underline{\delta}_1$ is therefore as required for (21) to hold. *Q.E.D.*

PROOF OF LEMMA 3:

By (A.1) there exists $(\breve{\pi}_k)_{k\in K}$ and IR payoffs $(\breve{\omega}_k)_{k\in K}$ such that $A_k(\breve{\pi}_k) > \breve{\omega}_k$ and $B(\breve{\pi}_k) > \hat{b}$ for all k. Choose ϵ so small that: $A_k(\breve{\pi}_k) - \breve{\omega}_k > 5\epsilon$; $B(\breve{\pi}_k) - \hat{b} > 5\epsilon$ for $k \in K$; there exists $\bar{\pi}_1$ satisfying $(A_1(\bar{\pi}_1), B(\bar{\pi}_1)) \in G_1(\epsilon)$ and $A_1(\breve{\pi}_1) \leq A_1(\bar{\pi}_1) = \bar{a}_1(\epsilon)$. Let $\bar{\pi}_2 = \breve{\pi}_2$ if $A_2(\breve{\pi}_2) > A_2(\bar{\pi}_1)$ and let $\bar{\pi}_2 = \bar{\pi}_1$ otherwise.

The strategies in $\Gamma(\mathrm{p}, \delta)$ are as follows: In period zero player 2 plays $j = 1$ and type 2 plays a best response to $j = 1$ while type 1 plays a different action to signal her type. If type 1's signal is observed then the players play a deterministic sequence of action profiles that mimic $\bar{\pi}_1$. If type 2's signal is observed they mimic $\bar{\pi}_2$ in a similar fashion. Player 1 minmaxes 2 if he deviates and player 2 uses a punishment strategy to hold player 1 to at most $(\breve{\omega}_k)_{k\in K} + \epsilon\mathbf{1}$ if she deviates.

It is clear that type 2 does not gain by mimicking type 1 and nor does type 1 gain by mimicking type 2 (by definition of $\bar{a}_1(\epsilon)$). By construction $A_k(\bar{\pi}_k) - (\breve{\omega}_k + \epsilon) > 4\epsilon$ and $B(\bar{\pi}_k) - \hat{b} \geq \epsilon$ so the players lose at least $\delta\epsilon$ in total when they deviate, but any gains from deviation are bounded above by $2M(1-\delta)$ and $2M(1-\delta) < \delta\epsilon$ by assumption. *Q.E.D.*

PROOF OF LEMMA 4:

Blackwell's (1956) result implies the set $\{\mathrm{y} | \mathrm{y} \leq (\omega_k)_{k\in K}\}$ is approachable by player 2 if for any vector $\mathbf{z} \geq (\omega_k)_{k\in K}$ he has a stage-game mixed strategy $g_{\mathbf{z}}$ so that

$$(50) \qquad (\mathbf{z} - (\omega_1, \omega_2))(\ A_1(i, g_{\mathbf{z}}), A_2(i, g_{\mathbf{z}})\ ) \leq 0 \qquad \forall i \in I,$$

(see for example Zamir (1992)). This condition amounts to the requirement that the plane through (ω_1, ω_2) orthogonal to the vector $(\mathbf{z} - (\omega_1, \omega_2))$ separates the point $\mathbf{z}$ from the points $\{\ (A_1(i, g_{\mathbf{z}}), A_2(i, g_{\mathbf{z}}))\ |\ i \in I\ \}$. Now by (A.1) there are two cases.

(1) $A_1(\pi) > \hat{a}_1 \Leftrightarrow A_2(\pi) \leq \hat{a}_2$: In this case the set $\{\ \mathrm{y}\ | \mathrm{y} \leq (\hat{a}_1, \hat{a}_2)\ \}$ is approachable. (This follows by either choosing $g_{\mathbf{z}}$ to be the strategy that minmaxes type 1 or the strategy that minmaxes type 2, where appropriate, in (50).) Let $\hat{g}_1$ be a strategy that minmaxes

type 1 and let i' be one of type 2's best responses to $\hat{g}_1$ so: $A_1(i', \hat{g}_1) \leq \hat{a}_1$ and $A_2(i', \hat{g}_1) \geq \hat{a}_2$. Define $\hat{\pi}$ to be a rational approximation to the distribution over action profiles induced when player 1 uses i' and player 2 uses $\hat{g}_1$. Choose $\hat{\pi}$ so that $A_1(\hat{\pi}) \leq \hat{a}_1 + \epsilon$ and $A_2(\hat{\pi}) \geq \hat{a}_2 + 4\epsilon$ which is possible by (A.1) and for ϵ sufficiently small. Now choose an integer T and a finite sequence $\{(\hat{i}^t, \hat{j}^t)\}_{t=0}^{T-1}$ so that each profile (i, j) occurs $T\hat{\pi}_{ij}$ times in it. By continuity there exists a $\hat{\delta}$ so that $|\hat{A}_k - A_k(\hat{\pi})| < \epsilon/2$ for all $k \in K$ and $|\hat{B} - B(\hat{\pi})| < \epsilon/2$ for any $\delta > \hat{\delta}$. Then for any $\delta > \hat{\delta}$ it is true that $\hat{A}_1 < \hat{a}_1 + 3\epsilon/2$ so (a) in the lemma is certaintly true. Moreover, since $\hat{A}_2 > \hat{a}_2 + 7\epsilon/2$ and $\alpha_2 > \hat{a}_2 + 5\epsilon$ taking punishments $(\omega_1, \omega_2) = (\hat{a}_1, \hat{a}_2)$ implies (b) is also true.

(2) There exists π^* such that $A_k(\pi^*) > \hat{a}_k$ for $k \in K$: Define π^0 so that it maximizes type 2's payoff given type 1 receives her minmax payoff: $\pi^0 \in \arg\max\{\ A_2(\pi) \mid A_1(\pi) = \hat{a}_1\ \}$. (Note $A_2(\pi^0) \geq \hat{a}_2$, because if π' represents the correlated strategy played when type 2 plays a best reponse to player 2's mixed strategy that minmaxes type 1 ($A_1(\pi') \leq \hat{a}_1$ and $A_2(\pi') \geq \hat{a}_2$), then the convex combination $\pi^\rho = \rho\pi' + (1-\rho)\pi^*$ can be chosen so that $A_1(\pi^\rho) = \hat{a}_1$ and $A_2(\pi^\rho) \geq \hat{a}_2$.) The point $(\hat{a}_1, A_2(\pi^0))$ is individually rational. This again follows by showing that the set $\{\ y \mid y \leq (\hat{a}_1, A_2(\pi^0))\ \}$ is approachable by taking g_z to either be the strategy that minmaxes type 1 or the strategy that minmaxes type 2 in (50) above, see the figures below. Choose $\hat{\pi}$ to be a rational approximation to π^0 so that there

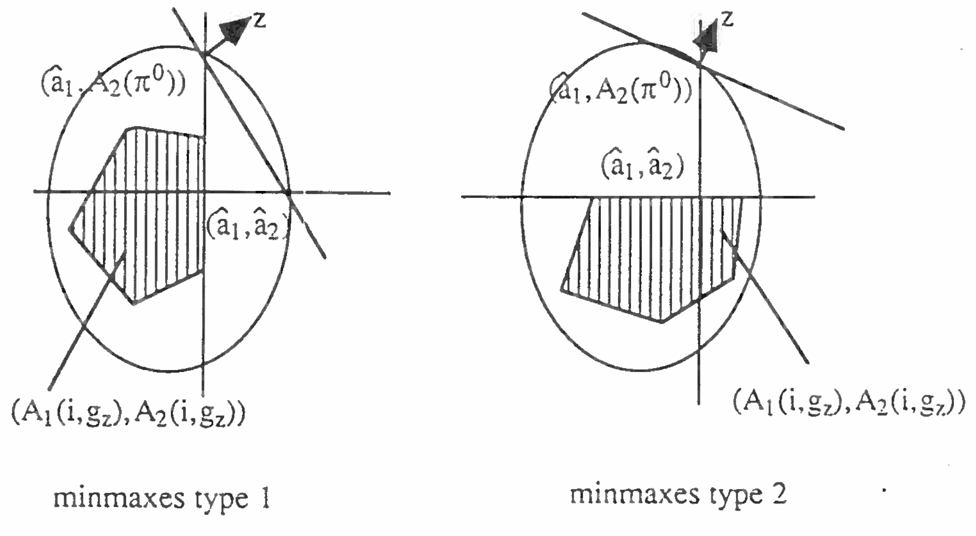

Figure 2 – $\{\ y \mid y \leq (\hat{a}_1, A_2(\pi^0))\ \}$ Approachable

exists an individually rational (ω_1, ω_2) satisfying

$$A_k(\hat{\pi}) > \omega_k + 4\epsilon, \qquad \forall k \in K; \qquad\qquad A_1(\hat{\pi}) < \hat{a}_1 + 5\lambda\epsilon.$$

Again mimic $\hat{\pi}$ by playing a finite deterministic sequence of actions; there exists an integer T and a finite sequence $\{(\hat{i}^t, \hat{j}^t)\}_{t=0}^{T-1}$ so that any profile (i,j) occurs $T\hat{\pi}_{ij}$ times in this sequence. By continuity, there exists a $\hat{\delta}$ such that for $\delta > \hat{\delta}$ the total payoffs over the sequence are close to those for $\hat{\pi}$, that is, $|\hat{A}_k - A_k(\hat{\pi})| < \epsilon/2$ for $\delta > \hat{\delta}$ and $k = 1, 2$. This implies statement (a) in the lemma and (b) follows from the convexity of the $(7\epsilon/2)$-IR set. *Q.E.D.*

<u>PROOF OF LEMMA 5:</u>

First define an infinite sequence $\{(\hat{i}^t, \hat{j}^t)\}_{t=0}^{\infty}$ by repeating the finite sequence of actions $\{(\hat{i}^s, \hat{j}^s)\}_{s=0}^{T-1}$ of Lemma 1, that is, for $t > T-1$ define $(\hat{i}^t, \hat{j}^t) := (\hat{i}^s, \hat{j}^s)$ where $s = t \bmod T$. We also use $\hat{h}^t := ((\hat{i}^s, \hat{j}^s))_{s=0}^{t-1}$ to define the history that arises if the players play out this infinite sequence for t periods. The players' equilibrium strategies in $\Gamma(1 - rp_2, rp_2, \delta)$ are now described.

Type 1

$t = 0$; play $\hat{i}^0$ with probability $q := r(1-p_2)(1-rp_2)^{-1}$ and $\underline{i}^0 \neq \hat{i}^0$ with probability $(1-q)$.

$0 < t < NT - 1$; if the history is $\hat{h}^t$ then play $\hat{i}^t$ in period t, if the history is $(\hat{h}^{t-1}, (\hat{i}^{t-1}, j))$ but $j \neq \hat{j}^{t-1}$ minmax player 2 thereafter.

If the history is $\hat{h}^{NT}$ then play the equilibrium strategy in $\Gamma(\mathbf{p}, \delta)$ to get the payoffs $(\alpha_1, \alpha_2, \beta)$.

If the history is $(\underline{i}^0, \hat{j}^0)$ then play pure strategies in the repeated game of complete information $\Gamma_1(\delta)$ to achieve the payoffs $\delta^{-1}(a_1(N), b) - \delta^{-1}(1 - \delta)(A_1(\underline{i}^0, \hat{j}^0), B(\underline{i}^0, \hat{j}^0))$ and minmax player 2 if he deviates. If the history is $(\underline{i}^0, j)$ but $j \neq \hat{j}^0$ then minmax player 2.

Type 2

$0 \leq t < NT - 1$; if the history is $\hat{h}^t$ then play $\hat{i}^t$ in period t. If the history is $(\hat{h}^{t-1}, (\hat{i}^{t-1}, j))$ and $j \neq \hat{j}^{t-1}$ minmax player 2 thereafter.

In period NT play the equilibrium of $\Gamma(\mathbf{p}, \delta)$ to get the payoffs $(\alpha_1, \alpha_2, \beta)$.

Player 2

> $0 \leq t < nT - 1$; if the history is $\hat{h}^t$ then play $\hat{j}^t$ in period t. If the history is $(\hat{h}^{t-1}, (i, \hat{j}^{t-1}))$ and $i \neq \hat{i}^{t-1}$ and $1 \leq t \leq NT$ play a strategy to ensure that player 1's types receive no more than $\underline{\omega}_k(t) + \epsilon$ for $k \in K$.
>
> If the history is $\hat{h}^{NT}$ then play the equilibrium strategy in $\Gamma(\mathbf{p}, \delta)$ to get the payoffs $(\alpha_1, \alpha_2, \beta)$.
>
> If the history is $(\underline{i}^0, \hat{j}^0)$ then play the pure strategy equilibrium in the repeated game of complete information $\Gamma_1(\delta)$ to achieve the payoffs $\delta^{-1}(a_1(N), b) - \delta^{-1}(1-\delta)(A_1(\underline{i}^0, \hat{j}^0), B(\underline{i}^0, \hat{j}^0))$. If player 1 deviates from this equilibrium or plays $i^0 \neq \hat{i}^0, \underline{i}^0$ then play a strategy to ensure that player 1's types receive no more than $\underline{\omega}_k(0) + \epsilon$ for $k \in K$.

(The payoffs from following these strategies certainly satisfy the conditions described in the Lemma. The punishment payoffs $(\underline{\omega}_k(t))_{k \in K}$ for $t = 0, 1, 2, \ldots$ will be defined below.)

Players' strategies are optimal in period $t \geq NT$: Type 1 reveals herself by playing the action $\underline{i}^0$ with probability $1 - r$ so player 2's priors are revised upwards if the history $\hat{h}^{NT}$ is played out. By Bayes' Rule his priors are revised from rp_2 to p_2 because $rp_2(rp_2 + (1 - rp_2)q)^{-1} = p_2$. If the history $\hat{h}^{NT}$ occurs, the players thus play the game $\Gamma(\mathbf{p}, \delta)$. By assertion $(\alpha_1, \alpha_2, \beta)$ are equilibrium payoffs for $\Gamma(\mathbf{p}, \delta)$ so the strategies described above certainly consitute an equilibrium given the history $\hat{h}^{NT}$.

Players' strategies are optimal in periods $t < NT$: We begin by considering player 1. In period zero type 1 plays a mixed strategy. Her strategy is optimal if she is indifferent between the actions $\hat{i}^0, \underline{i}^0$ and provided the payoff from $\hat{i}^0$ is no less than that from any other action $i \neq \hat{i}^0, \underline{i}^0$. By construction her expected payoff from the actions $\hat{i}^0, \underline{i}^0$ is $a_1(N)$ so she is indifferent. It is not obvious, however, that an equilibrium of $\Gamma_1(\delta)$ can be found with the payoffs $\delta^{-1}(a_1(N), b) - \delta^{-1}(1-\delta)(A_1(\underline{i}^0, \hat{j}^0), B(\underline{i}^0, \hat{j}^0))$ required by her strategy. To verify this notice that

$$\left| a_1(N) - \left(\frac{1}{\delta} a_1(N) - \frac{1-\delta}{\delta} A_1(\underline{i}^0, \hat{j}^0) \right) \right| = \frac{(1-\delta) \left| a_1(N) - A_1(\underline{i}^0, \hat{j}^0) \right|}{\delta} < \frac{(1-\delta) 2M}{\delta} < \epsilon.$$

(The last inequality follows from $\delta > 2M/(\epsilon + 2M)^{-1}$.) By assertion $(a_1(N), b) \in G_1(2\epsilon, \epsilon)$ so this establishes that Result 3 can be applied to these payoffs.

We now show that type 1 and type 2 do not benefit by deviating from the sequence $\{(\hat{i}^t, \hat{j}^t)\}$. To do this it is first necessary to define the punishments $(\underline{\omega}_k(t))_{k\in K}$. By assertion in the lemma (α_1, α_2) is (3ϵ)-IR and $a_1(N) \geq \hat{a}_1 + 3\epsilon$, now from Lemma 4(b) with $\rho = \delta^{NT}$ it follows that $(a_1(N), a_2(N))$ (defined in (25)) are (3ϵ)-IR. The set of (3ϵ)-IR payoffs is convex, so for any $0 \leq n \leq N$ the pair $(a_k(n))_{k\in K}$ is also (3ϵ)-IR. This means that for any $0 \leq n \leq N$ there exists an IR vector of payoffs, $(\omega_k(n))_{k\in K}$, such that

$$(a_k(n))_{k\in K} \geq (\omega_k(n))_{k\in K} + 3\epsilon\mathbf{1}. \tag{51}$$

$a_k(n)$ is type $k's$ equilibrium payoff if she expects to play n times through the finite sequence $\{(\hat{i}^t, \hat{j}^t)\}_{t=0}^{T-1}$, before reaching the equilibrium with payoffs $(\alpha_1, \alpha_2, \beta)$. In general, however, her expected payoff from abiding by the strategy above is $(1-\delta)\sum_{t=r}^{T-1} \delta^{t-r} A_k(\hat{i}^t, \hat{j}^t) + \delta^r a_k(n)$ where $r < T$. The difference between this and $a_k(n)$ satisfies

$$\begin{aligned}
\left|(1-\delta)\sum_{t=r}^{T-1} \delta^{t-r} A_k(\hat{i}^t, \hat{j}^t) + \delta^r a_k(n) - a_k(n)\right| &= \left|(1-\delta)\sum_{t=r}^{T-1} \delta^{t-r} A_k(\hat{i}^t, \hat{j}^t) - (1-\delta^r) a_k(n)\right| \\
&\leq (1-\delta^r)M + (1-\delta^r)M \\
&\leq 2M(1-\delta^T) \leq \epsilon/2 \qquad (52)
\end{aligned}$$

(The last line of (52) follows from the assumption made on δ.) Combining (51) and (52) implies that at for any $0 \leq t < NT-1$ where $(NT - 1 - t = nT + r)$ player 1's expected payoffs are at least $(5/2)\epsilon\mathbf{1}$ above some vector $(\omega_k(n))_{k\in K}$ of IR payoffs defined by (51). We will choose these to be the punishments $(\underline{\omega}_k(t))_{k\in K}$ described in the strategies above. Since (a) we have chosen δ small in the Lemma, (b) k's maximum payoff from deviation in period t is $(1-\delta)M + \delta(\underline{\omega}_k(t) + \epsilon)$ and (c) from above, her payoff from continuing is at least $\underline{\omega}_k(t) + (5/2)\epsilon$ type k never benefits by deviating from the sequence at time $t < NT - 1$.

Now also consider the equilibrium of $\Gamma_1(\delta)$ that is played out when $(\underline{i}^0, \hat{j}^0)$ happens in period zero. Since (a) we have chosen δ small in the Lemma, (b) $a_1(N) > \omega_1(N) + 5\epsilon/2$ and (c) by Result 3 the strategies giving the payoff $\delta^{-1}(a_1(N) - (1-\delta)A_1(\underline{i}^0, \hat{j}^0))$ can be chosen so that type 1's continuation payoff is never less than $\delta^{-1}(a_1(N) - (1-\delta)A_1(\underline{i}^0, \hat{j}^0)) - \epsilon/2$ it follows that type 1 never benefits by deviating from this equilibrium.

Type 2, however, may prefer to deviate from the equilibrium strategies above by mimicking type 1 and playing action $\underline{i}^0$ in period zero. If she does prefer this, there is an equilibrium where types 1 and 2 both play $\underline{i}^0$ in period zero and then play out the equilibrium of $\Gamma_1(\delta)$ described above whilst player 2 uses the above strategy. At this new equilibrium the payoffs to type 1 will not have changed, although type 2's payoff is now greater than $a_2(N)$. Player 2's payoff is not $(1-r)b + rb(N)$ but instead is b. So player 2's payoff β' certainly satsifies

$$| \beta' - b | \leq r | b(N) - b | \leq 2Mr.$$

Player 2 is minmaxed whenever he deviates and $b(N) \geq \hat{b} + 2\epsilon$ it is simple to show that it is never optimal for him to deviate. *Q.E.D.*

PROOF OF THEOREM 3:

Let $(a_1^*, b^*) \in G_1(\nu)$ and choose ϵ so that $\nu > 5\epsilon\lambda + \epsilon > 6\epsilon$, where λ is defined in Lemma 4. The first step in the proof is to show that given an equilibrium $(\alpha_1, \alpha_2, \beta)$ of $\Gamma(\mathbf{p}, \delta)$ that satisfies the conditions of Lemma 5, we can find a range $[\underline{\alpha}(\alpha_1, \beta), \alpha_1]$ such that for any (a_1^*, b^*) with $a_1^* \in [\underline{\alpha}(\alpha_1, \beta), \alpha_1]$ there is an equilibrium $(\alpha_1', \alpha_2', \beta')$ and a game $\Gamma(\mathbf{p}', \delta)$ so that $|\alpha_1' - a_1^*| < \nu$, $|\beta - b^*| < \nu$ provided $\delta > \underline{\delta} = \delta_\epsilon^*$ and $p_1' > \underline{p}_1$, where δ_ϵ^* is defined in Lemma 5. Define $\underline{N}$ to be the largest integer N such that $(a_1(N), b(N)) \in G_1(2\epsilon)$. Define $\underline{\alpha}(\alpha_1, \beta) = a_1(\underline{N})$, then $a_1(N) \in [\underline{\alpha}(\alpha_1, \beta), \alpha_1]$ for $0 \leq N \leq \underline{N}$. By (52) $|a_1(n) - a_1(n+1)| \leq \epsilon/2$ so for any $a_1^* \in [\underline{\alpha}(\alpha_1, \beta), \alpha_1]$ there exists $a_1(N)$ so that $|a_1(N) - a^*| \leq \epsilon/4$. By using Lemma 5 again it is possible to choose b so that $|b - b^*| < \epsilon$, but then

$$|\beta' - b^*| \leq |\beta' - b| + |b - b^*| \leq 2Mr + \epsilon \leq 2\epsilon,$$

by Lemma 5 and by choosing r so that $2Mr < \epsilon$. Since $2Mr < \epsilon$ implies $p_2' < \epsilon/(2M)$ we have that provided $p_1' > 1 - \epsilon/(2M)$ and $\delta > \underline{\delta}$ the game $\Gamma(\mathbf{p}', \delta)$ has an equilibrium with the payoffs $(\alpha_1', \alpha_2', \beta)$ so that $|\alpha_1' - a_1^*| < \epsilon/4$, $|\beta' - b^*| < 2\epsilon$. Now because our choice of ϵ we have the completed the first step.

Let $(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\beta})$, defined in Lemma 3, be the initial equilibrium of $\Gamma(\mathbf{p}, \delta)$. Then from above if $a_1^* \in [\underline{\alpha}(\bar{\alpha}_1, \bar{\beta}), \bar{a}_1(5\epsilon)]$ then for $(a_1^*, b^*) \in G_1(2\epsilon)$ the theorem above holds. So, if $\underline{\alpha}(\bar{\alpha}_1, \bar{\beta}) = \hat{a}_1 + 2\epsilon$ we are finished, because the theorem holds for all $(a_1^*, b^*) \in G_1(2\epsilon)$ with $a_1^* \in [\hat{a} + \nu, \bar{a}_1(5\epsilon)]$. If, however, $\underline{\alpha}(\bar{\alpha}_1, \bar{\beta}) > \hat{a}_1 + 2\epsilon$ then the line segment in (a_1, b)-space defined by the convex combination

$$(a_1(n), b(n)) = (1-\delta^{nT})(\hat{A}_1, \hat{B}) + \delta^{nT}(\bar{\alpha}_1, \bar{\beta})$$

intersects $b = \hat{b}+2\epsilon$ before it hits $a_1 = \hat{a}_1+2\epsilon$. In this case for all $\delta > \underline{\delta}$, $p_1' > (1/2)+(1/2)p_1$ we can choose an equilibrium $(\alpha_1', \alpha_2', \beta')$ of $\Gamma(\mathrm{p}', \delta)$ such that $|\alpha_1' - \underline{\alpha}(\bar{\alpha}_1, \bar{\beta})| < \epsilon/4$, $|\beta' - f(\underline{\alpha}(\bar{\alpha}_1, \bar{\beta}))| < 2\epsilon$ where

$$f(\alpha) := \frac{1}{2}(\hat{b}+2\epsilon) + \frac{1}{2}\max\{\ b \mid (\alpha, b) \in G_1(2\epsilon)\ \}. \tag{53}$$

This is achieved by choosing a payoff pair (a_1, b) so that $a_1 = a_1(\underline{N})$, b so that $|b - f(a_1)| \leq \epsilon$ and $r = 1/2$ in Lemma 5.

Given the equilibrium $(\alpha_1', \alpha_2', \beta')$ described in the previous paragraph we can repeat the above argument and find a new range $[\underline{\alpha}(\alpha_1', \beta'), \alpha_1']$ such that if $(a_1^*, b^*) \in G_1(2\epsilon)$ and $a_1^* \in [\underline{\alpha}(\alpha_1', \beta'), \bar{a}_1(5\epsilon)]$, then for $\delta > \underline{\delta}$ and $p_1'' > (1/2)+(1/4)+(1/4)p_1$ the game $\Gamma(\mathrm{p}'', \delta)$ has an equilibrium $(\alpha_1'', \alpha_2'', \beta'')$ such that: $|\alpha_1'' - a_1^*| < \nu$, and $|\beta'' - b^*| < \nu$.

If $\underline{\alpha}(\alpha_1', \beta') = \hat{a}+2\epsilon$ this is sufficient to prove the theorem; otherwise for all $\delta > \underline{\delta}$, $p_1'' > (1/2)+(1/4)+(1/4)p_1$ we can choose an equilibrium $(\alpha_1'', \alpha_2'', \beta'')$ of $\Gamma(\mathrm{p}'', \delta)$ such that $|\alpha_1'' - \underline{\alpha}(\alpha_1', \beta')| < \epsilon/4$, $|\beta'' - f(\underline{\alpha}(\alpha_1', \beta'))| < 2\epsilon$. And the entire argument above can be repeated once again. Provided it is possible to approximate any $(a_1^*, b^*) \in G_1(2\epsilon)$ in a finite number of iterations of this argument we have proved the theorem. To show that this is possible define the sequence $\{X_d\}$ as follows

$$\begin{aligned} X_0 &= \bar{a}_1(5\epsilon), \\ X_{d+1} &= (1-\rho_d)\hat{A}_1 + \rho_d X_d, \end{aligned}$$

where

$$\rho_d = \min\{\ \rho \mid (1-\rho)(\hat{A}_1, \hat{B}) + \rho(X_d, f(X_d))\ \},$$

and $f(.)$ is the function defined in (53). This sequence defines how the lower thresholds $\underline{\alpha}(.)$ evolve as the iterations are repeated. We need only consider the case where $\hat{b}+2\epsilon = (1-\rho)\hat{B} + \rho f(X_d)$, (as otherwise the result is trivial) but then it is possible to solving for ρ and to calculate X_d.

$$X_{d+1} - \hat{A}_1 = \frac{\hat{b}+2\epsilon-\hat{B}}{f(X_d)-\hat{B}}\left(X_d - \hat{A}_1\right) \leq \frac{\hat{b}+2\epsilon+M}{\hat{b}+5\epsilon+M}\left(X_d - \hat{A}_1\right). \tag{54}$$

The bound on the right follows from $\hat{B} > -M$ and $f(X_d) \geq \hat{b} + 5\epsilon$ (because we are only interested in approximating points $G_1(\nu)$). If we define $Q = (\hat{b} + 2\epsilon + M)/(\hat{b} + 5\epsilon + M)$ then $X_d - \hat{A}_1 \leq Q^d(\bar{\alpha}_1 - \hat{A}_1)$, where $Q < 1$. Since, by Lemma 4(a), $\hat{A}_1 < \hat{a}_1 + 5\epsilon\lambda + \epsilon/2$ and from above $\nu > 5\epsilon\lambda + \epsilon$ there is a finite number D so that $\hat{a}_1 + \nu - \hat{A}_1 \geq Q^D(\bar{\alpha}_1 - \hat{A}_1)$. Thus only a finite number of iterations of the above argument are needed for the intervals above to reach the point where $\underline{\alpha}(.) < \hat{a}_1 + \nu$. Thus there exists D finite so that for any $(a_1^*, b^*) \in G_1(\nu)$ any $\delta' > \underline{\delta}$ and $p_1' > \sum_{d=1}^{D}(\frac{1}{2})^d + p_1(\frac{1}{2})^D$ the game $\Gamma(p', \delta')$ has an equilibrium that satisfies the conditions of the theorem. *Q.E.D.*

References

AUMANN, R. J., AND S. SORIN (1989): "Cooperation and Bounded Recall," *Games and Economic Behavior*, 1, 5–39.

BERGIN, J. (1989): "A Characterization of Sequential Equilibrium Strategies in Infinitely Repeated Incomplete Information Games," *Journal of Economic Theory*, 47, 51–65.

BLACKWELL, D. (1956): "An Analog of the Minmax Theorem for Vector Payoffs," *Pacific Journal of Mathematics*, 65, 1–8.

CELENTANI, M., D. FUDENBERG, D. K. LEVINE, AND W. PESENDORFER (1994): "Maintaining a Reputation Against a Long-Lived Opponent," *Econometrica*, forthcoming.

CRIPPS, M. W., AND J. P. THOMAS (1992): "Reputation and Commitment in Two-Person Repeated Games without Discounting," *Econometrica*, forthcoming.

CRIPPS, M. W., AND J. P. THOMAS (1995): "*Reputation and Perfection in Repeated Common Interest Games*," mimeo.

CRIPPS, M. W., K. M. SCHMIDT, AND J. P. THOMAS (1995): "Reputation in Perturbed Repeated Games," *Journal of Economic Theory*, forthcoming.

FORGES, F. (1992): "Non-Zero-Sum Repeated Games of Incomplete Information," in *Handbook of Game Theory*, ed. by R. J. Aumann and S. Hart. Amsterdam: North Holland.

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 E. MASKIN (1986): "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," *Econometrica*, 54, 533–554.

FUDENBERG, D., AND E. MASKIN (1991): "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory*, 53, 428–438.

HART, S. (1985): "Nonzero-Sum Two-Person Repeated Games with Incomplete Information," *Mathematics of Operations Research*, 10, 117–153.

ISRAELI, E. (1989): "*Sowing Doubt Optimally in Two-Person Repeated Games*," M.Sc. Thesis (in Hebrew), Tel-Aviv University.

JORDAN, J. S. (1995): "Bayesian Learning in Repeated Games," *Games and Economic Behavior*, 9, 8–20.

KALAI, E., AND E. LEHRER (1993): "Rational Learning Leads to Nash Equilibrium," *Econometrica*, 61, 1019–1045.

KOREN, G. (1988): "*Two-Person Repeated Games with Incomplete Information and Observable Payoffs*," M.Sc. Thesis, Tel-Aviv University.

KREPS, D., P. MILGROM, J. ROBERTS AND R. WILSON (1982): "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 27, 245–252.

MCKELVEY, R. D., AND T. R. PALFREY (1992): "*Stationarity and Chaos in Infinitely Repeated Games of Incomplete Information*," California Institute of Technology, Social Science Working Paper No. 803.

MCKELVEY, R. D., AND T. R. PALFREY (1993): "*The Holdout Game: An Experimental Study of an Infinitely Repeated Game with Two-Sided Incomplete Information*," California Institute of Technology, Social Science Working Paper No. 804.

MERTENS,J-F., S. SORIN AND S. ZAMIR (1994): "*Repeated Games : Part A Background Material*," Universite Catholique De Louvain, C.O.R.E. Discussion Paper no.9420.

MYERSON, R. B. (1991): *Game Theory.* Cambridge: Harvard University Press.

SCHMIDT, K. M. (1993a): "Reputation and Equilibrium Characterization in Repeated Games of Conflicting Interests," *Econometrica*, 61, 325–352.

SCHMIDT, K.M. (1993b): "Commitment through Incomplete Information in a Simple Repeated Bargaining Game," *Journal of Economic Theory*, 60, 114–139.

SHALEV, J. (1994): "Nonzero-Sum Two-Person Repeated Games with Incomplete Information and Known-Own Payoffs," *Games and Economic Behavior*, 7, 246–259.

SORIN, S. (1992): "Repeated Games with Complete Information," in *Handbook of Game Theory*, ed. by R. J. Aumann and S. Hart. Amsterdam: North Holland.

VAN DAMME, E. E. C. (1987): *Stability and Perfection of Nash equilibria.* Berlin: Springer-Verlag.

ZAMIR, S. (1992): "Repeated Games of Incomplete Information: Zero Sum," in *Handbook of Game Theory* ed. by R. J. Aumann and S. Hart. Amsterdam: North Holland.